

**MATTHIAS GERDTS**

**Funktionalanalysis**

**Universität der Bundeswehr München  
Frühjahrssemester 2016**

ADRESSE DES AUTORS:

Matthias Gerdts

Institut für Mathematik und Rechneranwendung

Universität der Bundeswehr München

Werner-Heisenberg-Weg 39

85577 Neubiberg

E-Mail: [matthias.gerdts@unibw.de](mailto:matthias.gerdts@unibw.de)

WWW: [www.unibw.de/lrt1/gerdts](http://www.unibw.de/lrt1/gerdts)

Vorläufige Version: 27. Juni 2016

Copyright © 2016 by Matthias Gerdts

# Inhaltsverzeichnis

<b>1</b>	<b>Mehrdimensionale Integration</b>	<b>3</b>
1.1	Riemann-Integral . . . . .	3
1.2	Kurvenintegrale . . . . .	7
1.2.1	Parametrisierung einer Kurve nach der Bogenlänge . . . . .	11
1.2.2	Kurvenintegrale 1. Art (Integrale über die Bogenlänge) . . . . .	17
1.2.3	Kurvenintegrale 2. Art . . . . .	20
1.3	Das Lebesgue-Integral . . . . .	24
1.4	Flächen- und Volumenintegrale . . . . .	32
1.4.1	Flächen- und Volumenintegrale mit kartesischen Koordinaten . . . . .	34
1.4.2	Parametrisierung von Flächen und Volumen . . . . .	38
1.4.3	Flächen- und Volumenintegrale über parametrisierten Flächen und Volumen . . . . .	42
1.5	Oberflächenintegrale . . . . .	44
1.5.1	Flächenelemente . . . . .	44
1.5.2	Oberflächenintegrale 1. Art . . . . .	50
1.5.3	Oberflächenintegrale 2. Art . . . . .	53
1.6	Integralsätze . . . . .	56
1.6.1	Der Gauß'sche Integralsatz . . . . .	58
1.6.2	Der Integralsatz von Stokes . . . . .	60
<b>2</b>	<b>Räume, Metriken und Normen</b>	<b>61</b>
2.1	Metriken und metrische Räume . . . . .	61
2.2	Normen und normierte Vektorräume . . . . .	63
2.3	Banachräume . . . . .	74
<b>3</b>	<b>Approximation in Hilberträumen</b>	<b>80</b>
3.1	Bestapproximierende und Approximationsaufgabe . . . . .	80
3.2	Hilberträume und Projektionssatz . . . . .	83
3.3	Anwendung des Projektionssatzes auf lineare Ausgleichsprobleme . . . . .	89
3.4	Anwendung des Projektionssatzes bei der Fourieranalyse . . . . .	91
3.4.1	Komplexe Darstellung . . . . .	97

<b>4</b>	<b>Lineare Operatoren und lineare Differentialgleichungen</b>	<b>100</b>
4.1	Lineare Operatoren . . . . .	100
4.2	Lineare Funktionale und Dualräume . . . . .	107
4.3	Lineare Differentialgleichungen und Differentialoperatoren . . . . .	111
4.3.1	Bestimmung von Fundamentallösungen . . . . .	113
4.3.2	Bestimmung einer Partikulärlösung . . . . .	117
4.4	Die Laplace-Transformation . . . . .	121
<b>A</b>	<b>Anhang</b>	<b>126</b>
A.1	Diskrete Fourier-Transformation . . . . .	126
A.2	Tabellen zur Laplacetransformation . . . . .	131
A.3	Lineare Differentialgleichungssysteme 1. Ordnung . . . . .	134
	<b>Literaturverzeichnis</b>	<b>147</b>

# Vorwort

Diese Vorlesung vermittelt Grundlagen der Vektoranalysis und der Funktionalanalysis und richtet sich an Studierende im dritten Trimester im Bachelor-Studiengang Mathematical Engineering (ME).

Die Vorlesung baut auf den Vorlesungen Lineare Algebra und Analysis auf und hat zum Ziel, grundlegende funktionalanalytische Konzepte und Methoden zur mathematischen Beschreibung naturwissenschaftlich-technischer Prozesse zu vermitteln und zu eigenständigem logischen Denken anzuregen.

Die Vorlesung behandelt folgende Inhalte:

- Integration bei zwei und mehreren Variablen: Begriff des mehrdimensionalen Integrals, Berechnung auf Normalbereichen, Approximation durch Riemannsche Summen, Zylinder- und Kugelkoordinaten, Transformationsformel bei Koordinatentransformationen
- Kurven und Kurvenintegrale: Bogenlänge, Orientierung; Wegunabhängigkeit des skalaren Kurvenintegrals eines Vektorfeldes
- Flächen und Flächenintegrale: Flächenstücke, Orientierung, Flächeninhalt; Oberflächenintegral 1. Art (Integral eines skalaren Feldes über ein Flächenstück) und Flächenintegral 2. Art (Fluss eines Vektorfeldes durch ein Flächenstück), zusammengesetzte Flächen
- Die Integralsätze: Divergenz, Gaußscher Integralsatz; Gaußscher und Greenscher Integralsatz in der Ebene; Zirkulation, Wirbelstärke, Rotation, Stokesscher Integralsatz; Nabla-Kalkül
- Räume, Metriken und Normen
- Approximation
- Fréchet- und Gateaux-Ableitungen, schwache Ableitungen und Sobolew-Räume
- Nullstellen und Fixpunkte

**Allgemeine Tipps und Hinweise:**

- Mathematische Inhalte erschließen sich nicht von selbst und in den seltensten Fällen durch gelegentliches Überfliegen des Vorlesungsskriptes, sondern sie erfordern ein **aktives Mitarbeiten** und **viel Übung**.

Arbeiten Sie deshalb die Vorlesungen (und Übungen) nach!

Auch wenn die wöchentlichen Übungsaufgaben aus Kapazitätsgründen nicht eingesammelt und korrigiert werden können, so ist es zum Verständnis des Stoffes **sehr wichtig**, die Übungsaufgaben **regelmäßig** und **selbstständig** zu bearbeiten. Spätestens in der Prüfung, aber insbesondere auch im weiteren Verlauf des Studiums, ist eine gewisse Routine im Lösen von Aufgaben notwendig. Diese erlangt man nur durch eigenständiges Lösen von Aufgaben.

- Gehen Sie den Vorlesungsstoff und Übungsaufgaben mit Kommilitonen durch. Erklären und diskutieren Sie Definitionen und Sätze in eigenen Worten! Fragen Sie auch die Übungsleiter und den Leiter der Veranstaltung bei Unklarheiten.
- Es ist keine gute Idee, sich erst kurz vor der Prüfung mit dem Vorlesungsstoff zu beschäftigen. Dann ist es zu spät und es bleibt nichts hängen. Die Inhalte der Vorlesung tauchen an verschiedenen Stellen im Studium immer wieder auf und werden dann als bekannt vorausgesetzt.
- Wichtige Informationen zur Vorlesung, wie z.B. Ansprechpartner, Übungsaufgaben und Zusatzmaterial, finden sich auf der WWW-Seite

<http://www.unibw.de/lrt1/gerdts/funktionalanalysis>

Diese Seite wird laufend aktualisiert, so dass sie regelmäßig besucht werden sollte.

### Literatur:

- Göpfert, Riedrich, Tammer: Angewandte Funktionalanalysis, Motivationen und Methoden für Mathematiker und Wirtschaftswissenschaftler
- J. Wloka: Funktionalanalysis und Anwendungen
- H. Heuser: Funktionalanalysis
- Burg, K., Haf, H. and Wille, F.: Höhere Mathematik für Ingenieure. Band I: Analysis. Teubner, Stuttgart, 1985/2011.
- Meyberg, K. and Vachenauer, P. Höhere Mathematik I-II . Springer, Berlin-Heidelberg-New York, 2nd edition, 1993.
- Papula, L. Mathematik für Ingenieure und Naturwissenschaftler. Band 1-3, Vieweg+Teubner, Wiesbaden, 2008.

# Kapitel 1

## Mehrdimensionale Integration

In diesem Kapitel wird die Integration von Kurven und von Funktionen mehrerer Veränderlicher behandelt. Zunächst wiederholen wir einige Grundbegriffe der Integration und führen das Lebesgue-Integral ein.

### 1.1 Riemann-Integral

Anschaulich geht es bei der Integration um die Bestimmung des Flächeninhalts der von einer gegebenen Funktion mit der  $x$ -Achse eingeschlossenen Fläche. Die resultierende Gesamtfläche wird als das Integral über die gegebenen Funktion bezeichnet, wobei die Flächen unterhalb der  $x$ -Achse negativ in die Gesamtfläche eingehen und die Flächen oberhalb der  $x$ -Achse positiv eingehen, siehe Abbildung 1.1.

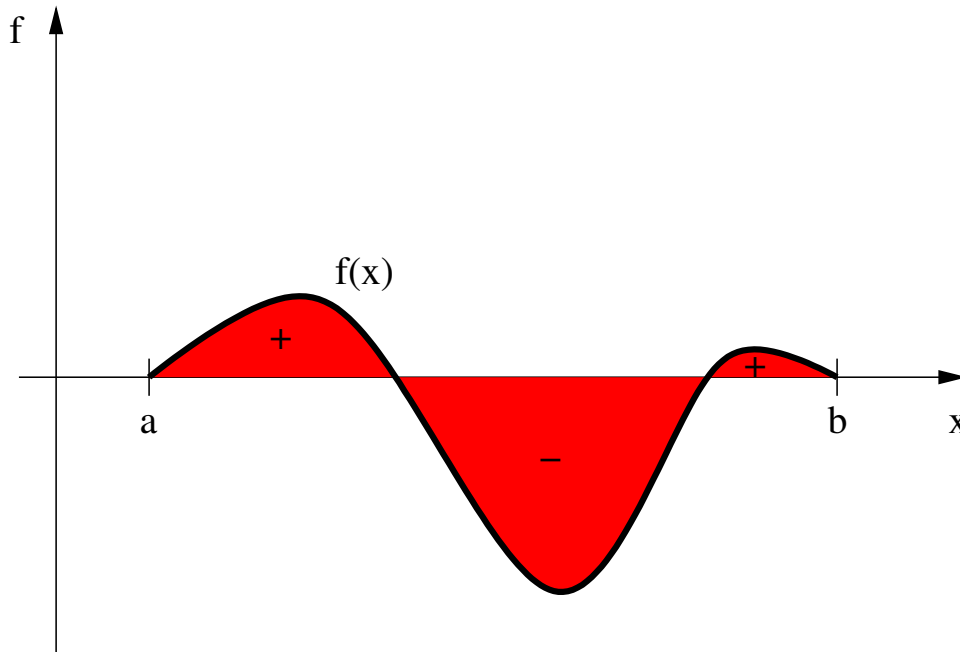
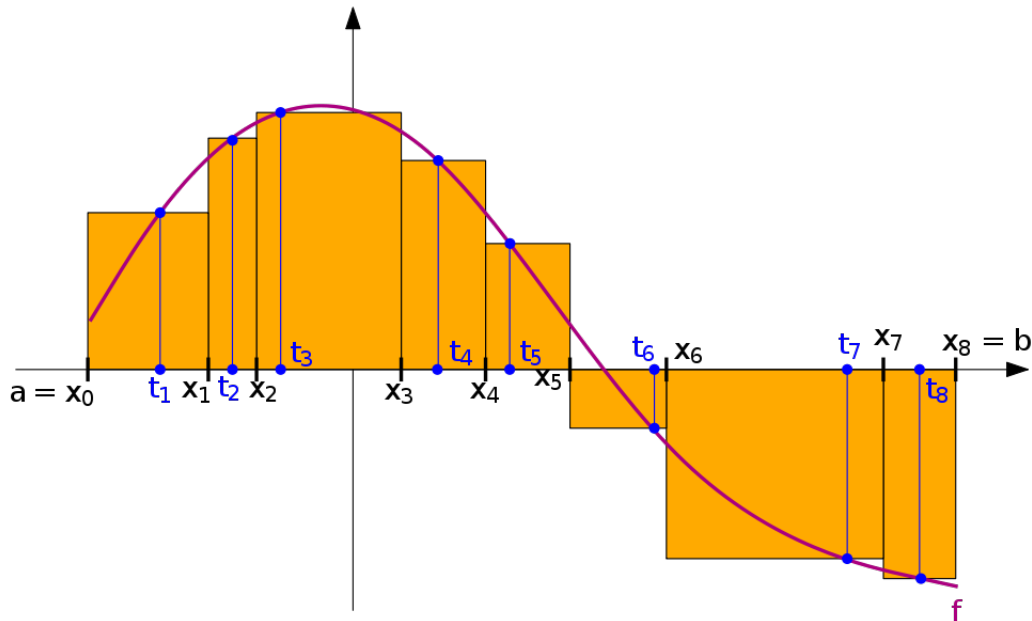


Abbildung 1.1: Integral einer Funktion  $f$ .

Das Riemann-Integral einer Funktion  $f$  auf dem Intervall  $[a, b]$  wird dabei durch einen Annäherungsprozess definiert bzw. motiviert, vgl. die folgende Abbildung (Quelle: Wikipedia).



Dazu wird das zu betrachtende Intervall  $[a, b]$  unterteilt in Zerlegungen der Form

$$Z = \{a = x_0 < x_1 < \dots < x_n = b\}.$$

Für beliebig gewählte Zwischenstellen  $t_i \in [x_{i-1}, x_i]$ ,  $i = 1, \dots, n$ , wird der Flächeninhalt approximiert durch die Riemann-Summe

$$S(f, Z) = \sum_{i=1}^n f(t_i)(x_i - x_{i-1}).$$

Die Funktion  $f$  heißt nun Riemann-integrierbar, wenn die Werte der Riemann-Summen für beliebig feine Zerlegungen und beliebige Zwischenstellen gegen einen festen Wert  $I$  konvergieren.  $I$  heißt dann **Riemann-Integral von  $f$  über  $[a, b]$**  und man schreibt hierfür

$$I = I[f] = \int_a^b f(x) dx.$$

Nicht alle Funktionen sind tatsächlich Riemann-integrierbar, aber man kann zeigen, dass z.B. die stetigen Funktionen auf einem kompakten Intervall Riemann-integrierbar sind.

### Rechenregeln:

- **Linearität des Integrals:** Für integrierbare Funktionen  $f$  und  $g$  und  $c \in \mathbb{R}$  gelten

$$\begin{aligned} \int_a^b f(x) + g(x) dx &= \int_a^b f(x) dx + \int_a^b g(x) dx, \\ \int_a^b cf(x) dx &= c \int_a^b f(x) dx. \end{aligned}$$



- **Additivität des Integrals:** Für  $a \leq c \leq b$  gilt

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

- **Monotonie des Integrals:** Für integrierbare Funktionen  $f$  und  $g$  mit  $f(x) \leq g(x)$  für alle  $x \in [a, b]$  gilt

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

Gilt sogar  $f(x) < g(x)$  für alle  $x \in [a, b]$ , so gilt

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

- Es gilt

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

- **Vertauschung der Integrationsgrenzen:**

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

- **Integration über Intervall der Länge 0:**

$$\int_a^a f(x)dx = 0.$$

Eine wichtige Rolle bei der Integration spielen Stammfunktionen.

### Definition 1.1.1 (Stammfunktion)

Die Funktion  $F : [a, b] \rightarrow \mathbb{R}$  heißt **Stammfunktion** von  $f : [a, b] \rightarrow \mathbb{R}$ , falls  $F'(x) = f(x)$  für alle  $x \in [a, b]$  gilt.

Gelingt es, eine Stammfunktion des Integranden zu bestimmen, so kann man Integrale leicht berechnen, wie der folgende Satz zeigt.

**Satz 1.1.2 (Hauptsatz der Differential- und Integralrechnung)**

Ist  $F$  eine Stammfunktion von  $f$  auf dem Intervall  $[a, b]$ , so gilt

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

Für die rechte Seite schreibt man häufig  $F(b) - F(a) = [F(x)]_a^b$ .

**Bemerkung 1.1.3**

Beachte, dass mit  $F$  auch jede Funktion  $F + c$  mit einer beliebigen Konstanten  $c \in \mathbb{R}$  eine Stammfunktion von  $f$  ist. Daher läßt man häufig die Intervallgrenzen weg und interpretiert eine Stammfunktion als **unbestimmtes Integral von  $f$** . Schreibweise:

$$F(x) = \int f(x)dx \quad \text{oder} \quad F(x) = c + \int f(x)dx.$$

■

Darüber hinaus gelten die bekannten Rechenregeln für Integrale:

**Satz 1.1.4**

Es gelten folgende Rechenregeln:

(a) **Partielle Integration:** Seien  $f$  und  $g$  stetig differenzierbar. Dann gilt

$$\int_a^b f'(x)g(x)dx = [f(x)g(x)]_a^b - \int_a^b f(x)g'(x)dx.$$

(b) **Substitutionsregel:** Seien  $f$  stetig und  $g$  stetig differenzierbar. Dann gilt

$$\int_a^b f(g(x))g'(x)dx = \int_{g(a)}^{g(b)} f(t)dt.$$

Schließlich betrachten wir noch uneigentliche Integrale.

**Definition 1.1.5 (Uneigentliches Integral)**

Unter einem **uneigentlichen Integral** verstehen wir die Integrale

$$\int_a^\infty f(x)dx, \quad \int_{-\infty}^\infty f(x)dx, \quad \int_{-\infty}^b f(x)dx,$$

welche formal definiert sind als

$$\begin{aligned}\int_a^\infty f(x)dx &= \lim_{b \rightarrow \infty} \int_a^b f(x)dx, \\ \int_{-\infty}^\infty f(x)dx &= \lim_{a \rightarrow -\infty, b \rightarrow \infty} \int_a^b f(x)dx, \\ \int_{-\infty}^b f(x)dx &= \lim_{a \rightarrow -\infty} \int_a^b f(x)dx.\end{aligned}$$

## 1.2 Kurvenintegrale

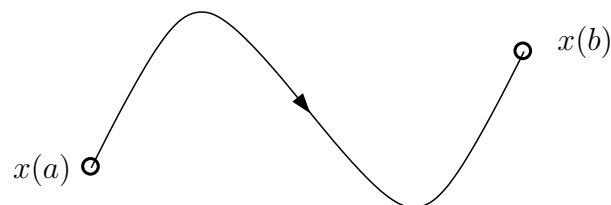
In diesem Abschnitt werden wir entlang von Kurven integrieren, um beispielsweise deren Länge (Bogenlänge) zu bestimmen. Solche Kurvenintegrale lassen sich auf eindimensionale (Riemann)-Integrale zurückführen.

Eine Kurve ist eine Abbildung  $x : [a, b] \rightarrow \mathbb{R}^d$  mit

$$x(t) = \begin{bmatrix} x_1(t) \\ \vdots \\ x_d(t) \end{bmatrix},$$

wobei  $t \in [a, b]$  den **Kurvenparameter** bezeichnet.

Im Fall  $d = 2$  handelt es sich um eine ebene Kurve, im Fall  $d = 3$  um eine räumliche Kurve. Durch die Parametrisierung der Kurve mit Kurvenparameter  $t$  wird die Kurve beginnend mit dem Anfangspunkt  $x(a)$  in Richtung des Endpunkts  $x(b)$  durchlaufen. Diese Durchlaufrichtung wird als **Orientierung der Kurve** bezeichnet, siehe nachfolgende Abbildung.



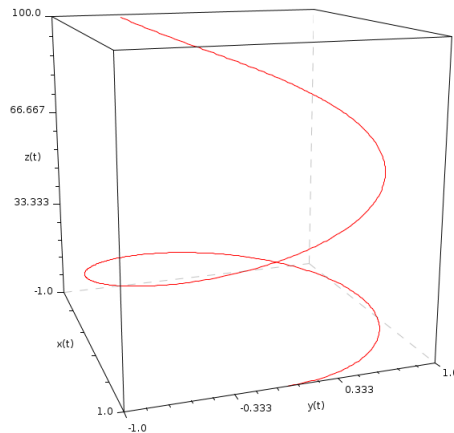
Darüber hinaus heißt eine Kurve **geschlossen**, wenn Anfangs- und Endpunkt übereinstimmen, also  $x(a) = x(b)$  gilt.

### Beispiel 1.2.1 (Kurve)

Die Funktion  $x : [0, 10] \rightarrow \mathbb{R}^3$  mit

$$x(t) = \begin{bmatrix} x_1(t) \\ x_2(t) \\ x_3(t) \end{bmatrix} := \begin{bmatrix} \cos(t) \\ \sin(t) \\ t^2 \end{bmatrix}$$

ist eine Kurve und beschreibt eine Spirale im  $\mathbb{R}^3$  mit Anfangspunkt  $[1, 0, 0]^\top$  und Endpunkt  $[\cos(10), \sin(10), 100]^\top$ .



Der Geschwindigkeitsvektor und der Beschleunigungsvektor lauten

$$x'(t) = \begin{bmatrix} -\sin(t) \\ \cos(t) \\ 2t \end{bmatrix}, \quad x''(t) = \begin{bmatrix} -\cos(t) \\ -\sin(t) \\ 2 \end{bmatrix}.$$

■

### Bemerkung 1.2.2

Das Beispiel zeigt, dass Geschwindigkeitsvektor  $x'(t)$  und Beschleunigungsvektor  $x''(t)$  im Allgemeinen nicht senkrecht aufeinander stehen, d.h. im Allgemeinen gilt  $\langle x'(t), x''(t) \rangle \neq 0$ .

■

Der Geschwindigkeitsvektor  $x'(t_0)$  zeigt in Richtung der Tangente an die Kurve  $x$  im Punkt  $x(t_0)$ . Dies kann man sich mit Hilfe von Abbildung 1.2 klarmachen. Der Richtungsvektor

$$\frac{x(t_0 + h) - x(t_0)}{h}$$

zeigt in Richtung der Sehne durch die Punkte  $x(t_0)$  und  $x(t_0+h)$ . Geht man zum Grenzwert  $h \rightarrow 0$  über, so folgt

$$x'(t_0) = \lim_{h \rightarrow 0} \frac{x(t_0 + h) - x(t_0)}{h}.$$

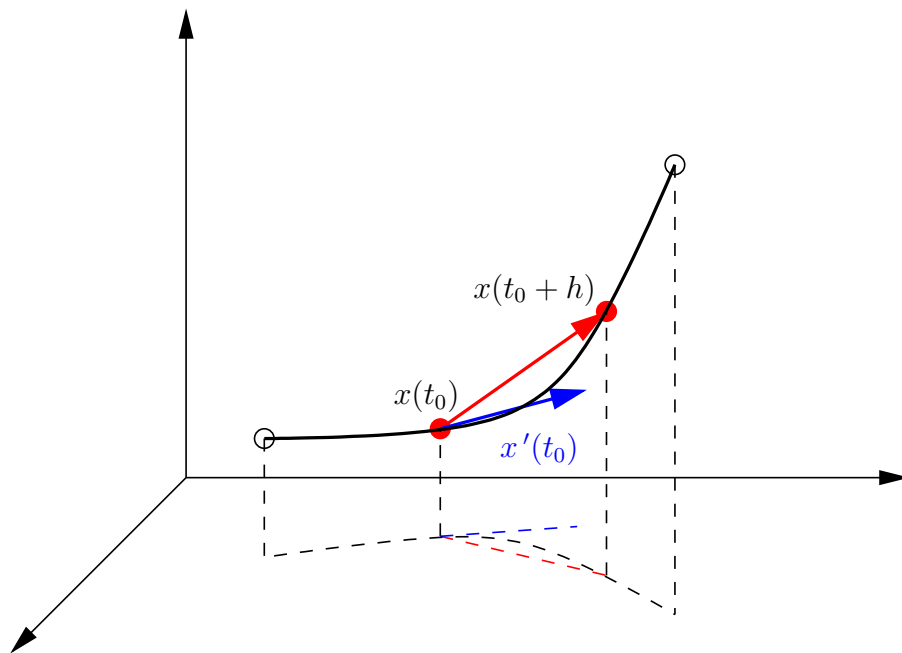


Abbildung 1.2: Kurve  $x$  (schwarz), Geschwindigkeitsvektor  $x'(t_0)$  (blau) und Richtungsvektor  $x(t_0 + h) - x(t_0)$  (rot).

### Definition 1.2.3 (Tangentenvektor)

(a) Der auf Länge 1 normierte Vektor

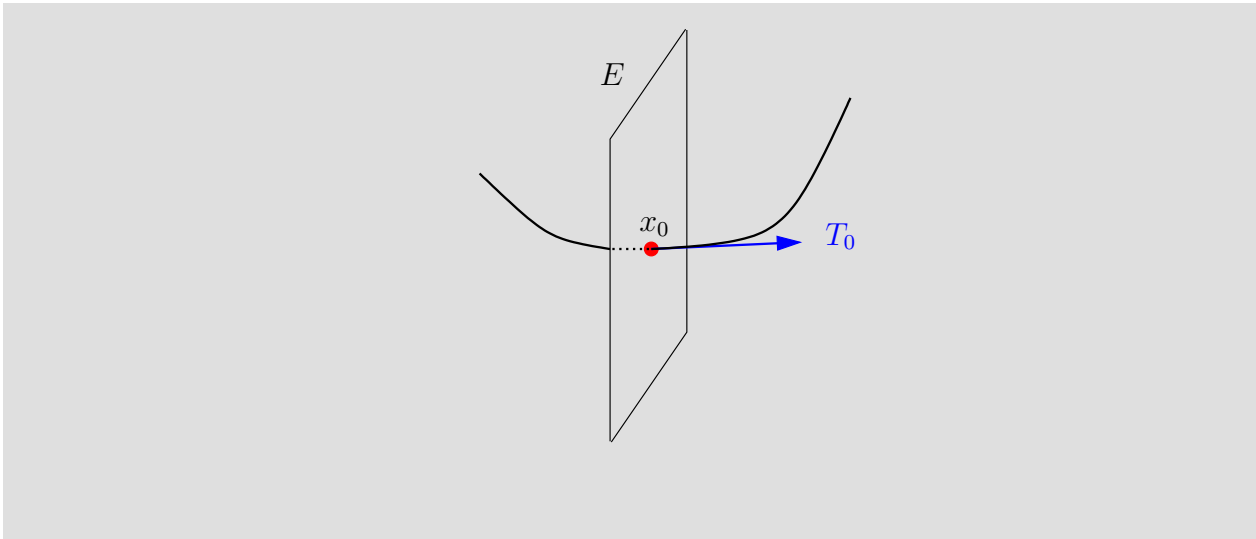
$$T(t) := \frac{x'(t)}{\|x'(t)\|}$$

heißt **Tangentenvektor** oder **Tangenteneinheitsvektor** der Kurve  $x$ .

(b) Die zum Tangentenvektor  $T_0 := T(t_0)$  im Punkt  $x_0 := x(t_0)$  orthogonale Ebene

$$E := \{x \in \mathbb{R}^d \mid \langle x - x_0, T_0 \rangle = 0\}$$

heißt **Normalebene** der Kurve  $x$  im Punkt  $x_0$ .



### Beispiel 1.2.4

(a) Die Spirale  $x : [0, 4\pi] \rightarrow \mathbb{R}^3$  mit

$$x(t) = \begin{bmatrix} x_1(t) \\ x_2(t) \\ x_3(t) \end{bmatrix} := \begin{bmatrix} \cos(t) \\ \sin(t) \\ t \end{bmatrix}$$

beschreibt die Bahn von Elektronen, die schief in ein homogenes (konstantes) Magnetfeld eingeschossen werden und sich auf einer solchen Schraubenbahn um die Feldrichtung ( $z$ -Achse) bewegen. Die Kurve hat den Tangentenvektor

$$T(t) = \frac{x'(t)}{\|x'(t)\|} = \frac{1}{\sqrt{2}} \begin{bmatrix} -\sin(t) \\ \cos(t) \\ 1 \end{bmatrix}.$$

Die Normalebene im Punkt  $x_0 = x(t_0)$  mit  $t_0 \in [0, 4\pi]$  lautet

$$E = \left\{ x = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \in \mathbb{R}^3 \mid x_3 = t_0 + (x_1 - \cos(t_0)) \sin(t_0) - (x_2 - \sin(t_0)) \cos(t_0) \right\}.$$

(b) Beschreibt die Funktion  $x(t)$  die Mittellinie eines Balkens, so liegt die Querschnittsfläche in der Normalebene der Kurve.

■

### Bemerkung 1.2.5

Wegen  $\|T(t)\|^2 = 1$  folgt für den Tangentenvektor an eine Kurve  $x$  durch Differenzieren

nach  $t$  die Beziehung

$$0 = \frac{d}{dt} \|T(t)\|^2 = \frac{d}{dt} (T(t)^\top T(t)) = \frac{d}{dt} \left( \sum_{i=1}^d T_i(t)^2 \right) = \sum_{i=1}^d 2T_i(t)T_i'(t) = 2\langle T(t), T'(t) \rangle.$$

Also stehen  $T$  und  $T'$  stets senkrecht aufeinander. ■

### 1.2.1 Parametrisierung einer Kurve nach der Bogenlänge

Eine Kurve  $x$  kann je nach Wahl bzw. Bedeutung des Kurvenparameters auf unterschiedliche Weise parametrisiert werden, d.h. sie kann unterschiedlich schnell durchlaufen werden. Eine natürliche Parametrisierung ergibt sich, wenn man die Länge der Kurve zur Parametrisierung verwendet, d.h. der Kurvenparameter  $t$  soll gleichzeitig der Länge der Kurve bezogen auf den Anfangspunkt entsprechen.

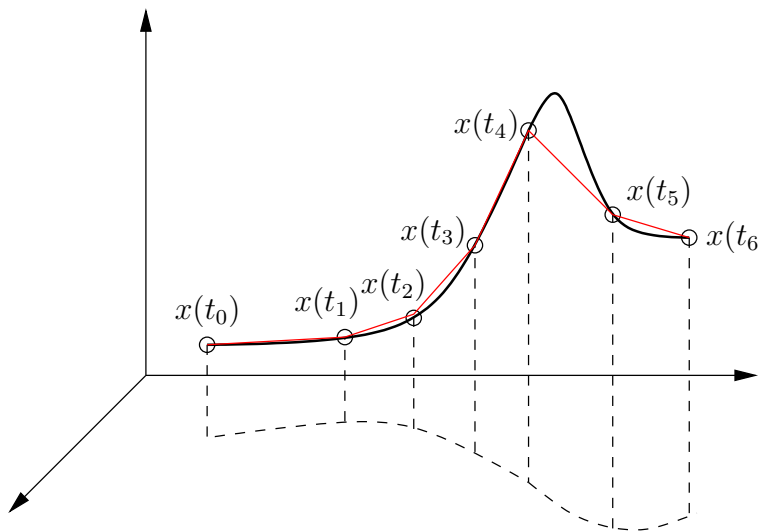
Wie können wir die Länge einer Kurve berechnen? Dazu stellen wir uns das Intervall  $[a, b]$  in äquidistante Teilabschnitte unterteilt vor, etwa

$$a = t_0 < t_1 < \dots < t_{N-1} < t_N = b \quad \text{mit } t_i = a + ih, \quad h = \frac{b-a}{N}, \quad N \in \mathbb{N}.$$

Dann approximieren wir die Kurve  $x$  durch einen **Polygonzug**  $p : [a, b] \rightarrow \mathbb{R}^d$ , der die Kurve  $x$  in jedem Intervall  $[t_i, t_{i+1}]$ ,  $i = 0, \dots, N-1$ , durch die Gerade

$$p(t) = x(t_i) + \frac{t - t_i}{t_{i+1} - t_i} (x(t_{i+1}) - x(t_i))$$

approximiert, vgl. Abbildung.



Die Länge des Polygonzugs  $p$  ist gegeben durch

$$L_N := \sum_{i=0}^{N-1} \|x(t_{i+1}) - x(t_i)\|.$$

Mit Hilfe des Mittelwertsatzes in Integralform ergibt sich daraus die Abschätzung

$$\begin{aligned}
 L_N &= \sum_{i=0}^{N-1} \|x(t_{i+1}) - x(t_i)\| \\
 &= \sum_{i=0}^{N-1} \left\| \int_0^1 x'(t_i + \tau(t_{i+1} - t_i))(t_{i+1} - t_i) d\tau \right\| \\
 &\leq \sum_{i=0}^{N-1} h \int_0^1 \|x'(t_i + \tau h)\| d\tau \\
 &= \sum_{i=0}^{N-1} \int_{t_i}^{t_{i+1}} \|x'(t)\| dt \\
 &= \int_a^b \|x'(t)\| dt =: L.
 \end{aligned}$$

Die Länge  $L$  kann als Grenzübergang der Werte  $L_N$  für  $N \rightarrow \infty$  (bzw.  $h \rightarrow 0$ ) angesehen werden und man schreibt

$$L = \int_K ds \quad \text{mit} \quad ds = \|x'(t)\| dt,$$

wobei der Index  $K$  im Integral  $\int_K$  andeutet, dass hier über eine Kurve integriert wird (Kurvenintegral).

### Definition 1.2.6 (Skalares Bogenelement, Bogenlänge)

Sei  $x : [a, b] \rightarrow \mathbb{R}^d$  eine stetig differenzierbare Kurve.

(a) Die Größe

$$ds = \|x'(t)\| dt$$

heißt **skalares Bogenelement**.

(b) Die **Bogenlänge**  $s(t)$  gibt die Länge der Kurve vom Anfangspunkt  $x(a)$  bis zum Punkt  $x(t)$  mit  $t \in [a, b]$  an und berechnet sich zu

$$s(t) = \int_a^t \|x'(\tau)\| d\tau.$$

Für  $t = b$  ergibt sich speziell die Länge der Kurve

$$L = s(b) = \int_a^b \|x'(\tau)\| d\tau = \int_K ds.$$



Das skalare Bogenelement  $ds$  beschreibt die Länge der Kurve zwischen zwei Kurvenpunkten  $x(t)$  und  $x(t+dt)$  für einen infinitesimal kleinen Schritt  $dt$ . Durch formales Umstellen erhält man  $\frac{ds}{dt} = \|x'(t)\|$ , also die Bahngeschwindigkeit.

### Beispiel 1.2.7

(a) Gegeben sei eine Funktion  $f : [a, b] \rightarrow \mathbb{R}$  gemäß  $t \mapsto f(t)$ . Wollen wir deren Länge bestimmen, so können wir die Kurve  $x : [a, b] \rightarrow \mathbb{R}^2$  mit

$$x(t) = \begin{bmatrix} t \\ f(t) \end{bmatrix}$$

betrachten. Diese beschreibt gerade den Graphen der Funktion  $f$ . Es gilt

$$x'(t) = \begin{bmatrix} 1 \\ f'(t) \end{bmatrix}, \quad ds = \|x'(t)\|dt = \sqrt{1 + f'(t)^2}dt.$$

Die Bogenlänge lautet damit

$$s(t) = \int_a^t \sqrt{1 + f'(\tau)^2} d\tau.$$

Speziell ergibt sich für  $f(t) := t$  und  $a = 0$  die Bogenlänge

$$s(t) = \int_0^t \sqrt{1 + 1} d\tau = \sqrt{2}t.$$

(b) Wir betrachten wieder eine Spirale  $x : [0, 2\pi] \rightarrow \mathbb{R}^3$  mit

$$x(t) = \begin{bmatrix} x_1(t) \\ x_2(t) \\ x_3(t) \end{bmatrix} := \begin{bmatrix} R \cos(t) \\ R \sin(t) \\ \alpha t \end{bmatrix}, \quad x'(t) = \begin{bmatrix} -R \sin(t) \\ R \cos(t) \\ \alpha \end{bmatrix},$$

wobei  $R$  und  $\alpha$  gegebene Konstanten sind. Es gilt

$$ds = \|x'(t)\|dt = \sqrt{(-R \sin(t))^2 + (R \cos(t))^2 + \alpha^2} dt = \sqrt{R^2 + \alpha^2} dt$$

Die Bogenlänge lautet damit

$$s(t) = \int_0^t \sqrt{R^2 + \alpha^2} d\tau = \sqrt{R^2 + \alpha^2} t.$$

Für eine Windung mit  $t = 2\pi$  ergibt sich speziell die Länge

$$L = s(2\pi) = \int_K ds = 2\pi \sqrt{R^2 + \alpha^2}.$$

■

Im Folgenden sind wir an differenzierbaren Kurven  $x(t)$  interessiert, deren Geschwindigkeitsvektor  $x'(t)$  nirgends verschwindet. Anschaulich bleibt man auf einer solchen Kurve also nicht stehen sondern bewegt sich stets vorwärts.

**Definition 1.2.8** (Reguläre Kurve)

Die Kurve  $x : [a, b] \rightarrow \mathbb{R}^d$  heißt **regulär**, wenn  $x$  stetig differenzierbar ist und  $\|x'(t)\| \neq 0$  für alle  $t \in [a, b]$  gilt.

Für reguläre Kurven  $x$  ist die Bogenlänge  $s(t)$  für alle  $t \in [a, b]$  streng monoton wachsend, da stets  $\|x'(t)\| > 0$  gilt. Folglich besitzt die Abbildung  $s : [a, b] \rightarrow [0, L]$  gemäß  $t \mapsto s(t)$  eine inverse Abbildung  $t : [0, L] \rightarrow [a, b]$  gemäß  $s \mapsto t(s)$ , die einer gegebenen Bogenlänge  $s$  eindeutig das zugehörige  $t$  zuordnet.

Damit können wir die Kurve  $x$  mit Hilfe der Bogenlänge  $s$  neu parametrisieren, indem wir für  $s \in [0, L]$  die Kurve  $y : [0, L] \rightarrow \mathbb{R}^d$  durch

$$y(s) := x(t(s))$$

definieren. Beachte, dass stets auch  $x(t) = y(s(t))$  gilt, so dass die Werte der Kurve dieselben bleiben. Durch diese Parametrisierung nach der Bogenlänge ändert sich jedoch die Durchlaufgeschwindigkeit durch die Kurve.

**Definition 1.2.9** (Parametrisierung nach der Bogenlänge)

Sei  $x : [a, b] \rightarrow \mathbb{R}^d$  eine stetig differenzierbare Kurve mit Bogenlänge  $L$  und  $\|x'(t)\| \neq 0$  für alle  $t \in [a, b]$ . Dann heißt die Kurve  $y : [0, L] \rightarrow \mathbb{R}^d$  mit

$$y(s) := x(t(s)), \quad s \in [0, L],$$

die **nach der Bogenlänge  $s$  parametrisierte Kurve**, wobei  $t : [0, L] \rightarrow [a, b]$  die Umkehrabbildung der Bogenlängenabbildung  $s : [a, b] \rightarrow [0, L]$  bezeichnet.

Man nennt die Parametrisierung nach der Bogenlänge auch die **natürliche Parametrisierung der Kurve  $x$** .

**Beispiel 1.2.10** (vgl. Beispiel 1.2.7)

Für die Spirale in Beispiel 1.2.7 haben wir die Bogenlänge

$$s(t) = \sqrt{R^2 + \alpha^2} t$$

berechnet. Die Umkehrabbildung lautet dann

$$t(s) = \frac{s}{\sqrt{R^2 + \alpha^2}},$$

womit sich die nach der Bogenlänge  $s$  parametrisierte Kurve wie folgt ergibt:

$$y(s) = x(t(s)) = \begin{bmatrix} R \cos\left(\frac{s}{\sqrt{R^2 + \alpha^2}}\right) \\ R \sin\left(\frac{s}{\sqrt{R^2 + \alpha^2}}\right) \\ \left(\alpha \frac{s}{\sqrt{R^2 + \alpha^2}}\right) \end{bmatrix}, \quad s \in [0, L].$$

■

Die Parametrisierung einer Kurve nach ihrer Bogenlänge hat einige Vorteile, da für die natürliche Parametrisierung Beziehungen gelten, die für beliebige Parametrisierungen im Allgemeinen nicht gelten. In Beispiel 1.2.1 (vgl. auch Bemerkung 1.2.1) haben wir gesehen, dass Geschwindigkeits- und Beschleunigungsvektor im Allgemeinen nicht senkrecht aufeinander stehen. Bei natürlicher Parametrisierung ist dies jedoch immer der Fall.

### Satz 1.2.11

Sei  $y : [0, L] \rightarrow \mathbb{R}^d$  eine nach ihrer Bogenlänge parametrisierte Kurve. Dann gelten folgende Eigenschaften:

(a)  $\|y'(s)\| = 1$  für alle  $s \in [0, L]$ .

(b) Erste und zweite Ableitung stehen senkrecht aufeinander:

$$\langle y'(s), y''(s) \rangle = 0 \quad \forall s \in [0, L].$$

### Beweis:

(a) Mit der Kettenregel folgt

$$y'(s) = x'(t(s))t'(s),$$

wobei die Ableitung  $t'(s)$  der Umkehrabbildung durch Ableiten der Identität  $s(t(s)) = s$  mit der Kettenregel nach  $s$  gegeben ist:

$$1 = s'(t)t'(s) \quad \implies \quad t'(s) = \frac{1}{s'(t)} = \frac{1}{\|x'(t)\|}.$$

Die letzte Beziehung  $s'(t) = \|x'(t)\|$  folgt aus der Definition der Bogenlänge durch Ableiten des Integrals nach der oberen Integrationsgrenze.

(b) Nach (a) gilt  $\|y'(s)\|^2 = 1$  für alle  $s \in [0, L]$ . Ableiten nach  $s$  liefert

$$\begin{aligned} 0 &= \frac{d}{ds} (\|y'(s)\|^2) = \frac{d}{ds} (y'(s)^\top y'(s)) = \frac{d}{ds} \left( \sum_{i=1}^n y_i'(s)^2 \right) \\ &= 2 \sum_{i=1}^n y_i'(s) y_i''(s) = 2 \langle y'(s), y''(s) \rangle. \end{aligned}$$

■

Häufig ist man an der Krümmung einer Kurve interessiert, da diese im physikalischen Kontext z.B. proportional zur Biegeenergie eines gebogenen Balkens ist.

### Definition 1.2.12 (Krümmung)

Die **Krümmung**  $\kappa$  einer Kurve ist definiert als die Richtungsänderung der Tangente bezogen auf die Bogenlänge, d.h.

$$\kappa(s) := \|T'(s)\| = \|y''(s)\|.$$

Der Wert  $\rho := \frac{1}{\kappa}$  heißt **Krümmungsradius**.

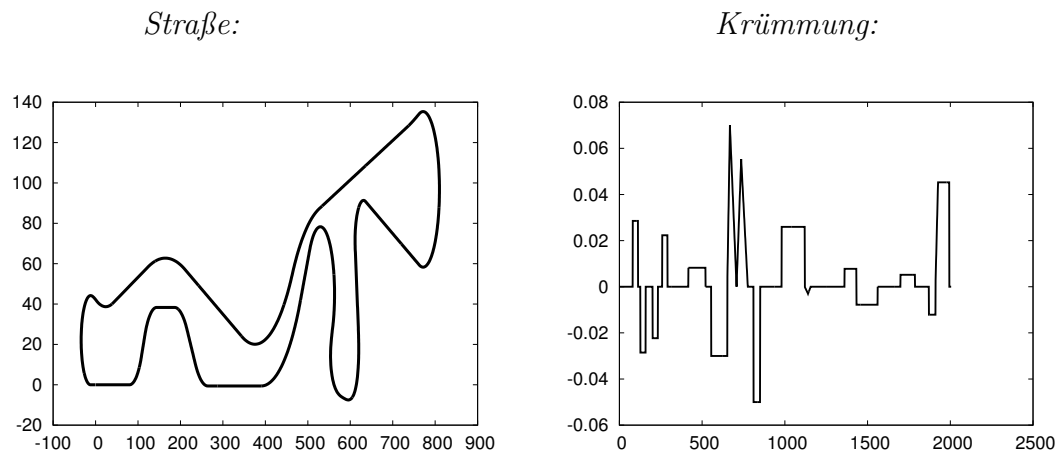
### Beispiel 1.2.13

- (a) Eine Gerade hat die Krümmung  $\kappa = 0$ .
- (b) Ein Kreis mit Radius  $r$  besitzt die Krümmung  $\kappa = 1/r$ .
- (c) Im Straßenbau werden Straßen als **Klothoiden** konstruiert. Klothoiden sind dadurch charakterisiert, dass ihre Krümmung stetig und stückweise linear von der Bogenlänge abhängt. Ebene Klothoiden in der  $(x, y)$ -Ebene sind durch Lösen des folgenden Differentialgleichungssystems charakterisiert:

$$x'(s) = \cos(\psi(s)), \quad y'(s) = \sin(\psi(s)), \quad \psi'(s) = \kappa(s),$$

wobei  $\kappa$  eine stetige und stückweise lineare Funktion ist. Straßen, die als Klothoiden konstruiert sind (etwa Autobahnausfahrten), sind komfortabel zu fahren, da die Fliehkraft sich stetig mit der Krümmung verändert.

Teststrecken oder Rennstrecken sind in der Regel nicht durch Klothoiden konstruiert, sondern erlauben Unstetigkeiten in der Krümmung, was insbesondere bei schneller Fahrt unangenehmer zu fahren ist. Die folgende Teststrecke ist an die BMW-Teststrecke in Aschheim angelehnt:



(d) Gegeben sei eine Funktion  $f : [a, b] \rightarrow \mathbb{R}$  gemäß  $t \mapsto f(t)$  und die dazugehörige Kurve  $x : [a, b] \rightarrow \mathbb{R}^2$  mit

$$x(t) = \begin{bmatrix} t \\ f(t) \end{bmatrix}.$$

Die Bogenlänge lautet damit

$$s(t) = \int_a^t \sqrt{1 + f'(\tau)^2} d\tau.$$

Wir parametrisieren die Kurve  $x$  nach der Bogenlänge  $s$  und erhalten

$$y : [0, L] \rightarrow \mathbb{R}^2, \quad s \mapsto y(s) = x(t(s)).$$

Insbesondere gilt dann (da  $x' \neq 0$ )

$$y'(s(t)) = \frac{x'(t)}{\|x'(t)\|} = \frac{1}{\sqrt{1 + f'(t)^2}} \begin{bmatrix} 1 \\ f'(t) \end{bmatrix},$$

vgl. auch den Beweis von Satz 1.2.11. Die Krümmung  $\|y''(s)\|$  mit  $s = s(t)$  ergibt sich durch Ableiten nach  $s = s(t)$  mit etwas Rechnung zu

$$\kappa(t) = \frac{f''(t)}{\sqrt{1 + f'(t)^2}^3}.$$

■

### 1.2.2 Kurvenintegrale 1. Art (Integrale über die Bogenlänge)

In diesem Abschnitt möchten wir Integrale über Funktionen entlang Kurven berechnen. Anwendungsgebiete solcher Kurvenintegrale lauten wie folgt:

#### Beispiel 1.2.14

Gegeben sei ein Draht mit Massendichte  $\rho$  (Masse pro Längeneinheit). Der Draht kann als Kurve  $K$  interpretiert werden.

(a) Die Masse des Drahtes bei inhomogener Massendichte ist gegeben durch das Kurvenintegral

$$m = \int_K \rho ds.$$

(b) Die Schwerpunktkoordinaten  $x_s = [x_s, y_s, z_s]^\top$  des Drahtes sind gegeben durch die Kurvenintegrale

$$x_s = \frac{1}{m} \int_K \rho x ds, \quad y_s = \frac{1}{m} \int_K \rho y ds, \quad z_s = \frac{1}{m} \int_K \rho z ds.$$

■

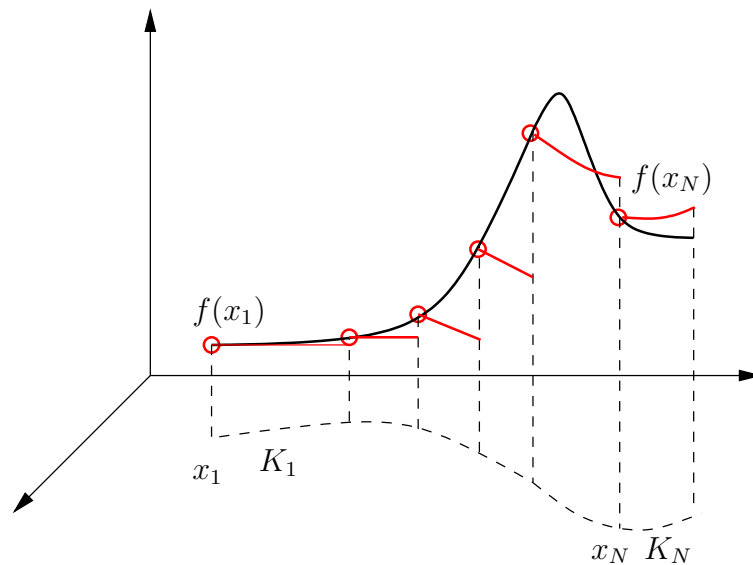


Abbildung 1.3: Konstruktionsprinzip des Kurvenintegrals 1. Art

### Definition 1.2.15 (Kurvenintegral 1. Art)

Gegeben sei eine Kurve  $x : [a, b] \rightarrow \mathbb{R}^d$  und eine Funktion  $f : \mathbb{R}^d \rightarrow \mathbb{R}$ , die mindestens entlang der Kurve definiert ist.

Für  $N \in \mathbb{N}$  sei die Kurve  $K := \{x(t) \in \mathbb{R}^d \mid t \in [a, b]\}$  in  $N$  Teilbögen  $K_1, \dots, K_N$  mit Bogenlängen  $ds_1, \dots, ds_N$  unterteilt und  $x_k$  ein beliebiger Punkt in  $K_k$  für  $k = 1, \dots, N$ .

Konvergieren die Werte

$$I_N := \sum_{k=1}^N f(x_k) ds_k$$

für  $N \rightarrow \infty$  unabhängig von der Wahl der Teilbögen  $K_1, \dots, K_N$  und der Punkte  $x_1, \dots, x_N$ , so heißt der Grenzwert das **Kurvenintegral 1. Art entlang der Kurve  $x$**  und der Wert wird mit

$$\int_K f(x) ds$$

bezeichnet, vgl. Abbildung 1.3.

Ist die Kurve  $x$  geschlossen und doppelpunktfrei (d.h. sie darf sich nicht überschneiden), so heißt das Kurvenintegral auch **Umlaufintegral** oder **Ringintegral** und wird mit

$$\oint_K f(x) ds$$

bezeichnet.

**Satz 1.2.16** (Berechnung des Kurvenintegrals 1. Art bei vorgegebener Parametrisierung)

Für die stetig differenzierbare Kurve  $x : [a, b] \rightarrow \mathbb{R}^d$  mit Parametrisierung  $t \mapsto x(t)$  und  $f : \mathbb{R}^d \rightarrow \mathbb{R}$  ist das Kurvenintegral 1. Art gegeben durch

$$\int_K f(x) ds = \int_a^b f(x(t)) \|x'(t)\| dt.$$

**Beweis:** Verwende das Bogenelement  $ds = \|x'(t)\| dt$ . ■

Mit Hilfe des Satzes kann das Kurvenintegral 1. Art also berechnet werden, indem man zunächst eine Parametrisierung der Kurve wählt, dann das Bogenelement ausrechnet und anschließend integriert.

**Beispiel 1.2.17**

(a) Für  $x : [0, \pi/2] \rightarrow \mathbb{R}^2$  und  $f : \mathbb{R}^2 \rightarrow \mathbb{R}$  gemäß

$$x(t) = \begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix} := \begin{bmatrix} r \cos(t) \\ r \sin(t) \end{bmatrix}, \quad f(x_1, x_2) := x_1 x_2$$

lautet das Kurvenintegral 1. Art wie folgt:

$$\begin{aligned}
 \int_K f(x_1, x_2) ds &= \int_K x_1 x_2 ds \\
 &= \int_0^{\pi/2} x_1(t) x_2(t) \sqrt{x_1'(t)^2 + x_2'(t)^2} dt \\
 &= \int_0^{\pi/2} r^2 \cos(t) \sin(t) \sqrt{(-r \sin(t))^2 + (r \cos(t))^2} dt \\
 &= r^3 \int_0^{\pi/2} \cos(t) \sin(t) dt \\
 &= r^3 \left[ \frac{\sin(t)^2}{2} \right]_0^{\pi/2} = \frac{r^3}{2}.
 \end{aligned}$$

(b) Ist die Kurve  $y : [0, L] \rightarrow \mathbb{R}^d$  nach der Bogenlänge  $s$  parametrisiert, so gilt wegen  $\|y'(s)\| = 1$  die Beziehung

$$\int_K f(y) ds = \int_0^L f(y(s)) ds.$$

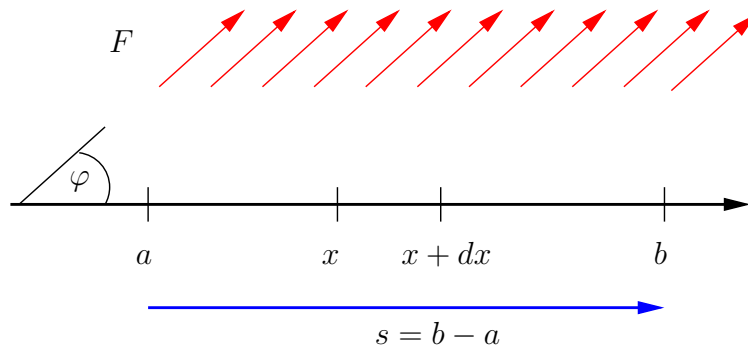
■

### 1.2.3 Kurvenintegrale 2. Art

#### Beispiel 1.2.18 (Arbeit im Kraftfeld)

Es soll die Arbeit berechnet werden, die von einem Kraftfeld (Vektorfeld) an einem Massenpunkt verrichtet wird, der sich entlang einer Kurve durch das Kraftfeld bewegt. Analoge Anwendungen ergeben sich bei der Betrachtung von Elektronen in einem Magnetfeld oder eines geladenen Drahtes in einem elektrischen Feld.

(a) Wir betrachten zunächst ein homogenes Kraftfeld (konstantes Vektorfeld) und die Bewegung des Massenpunkts entlang einer Geraden von  $a$  nach  $b$ .

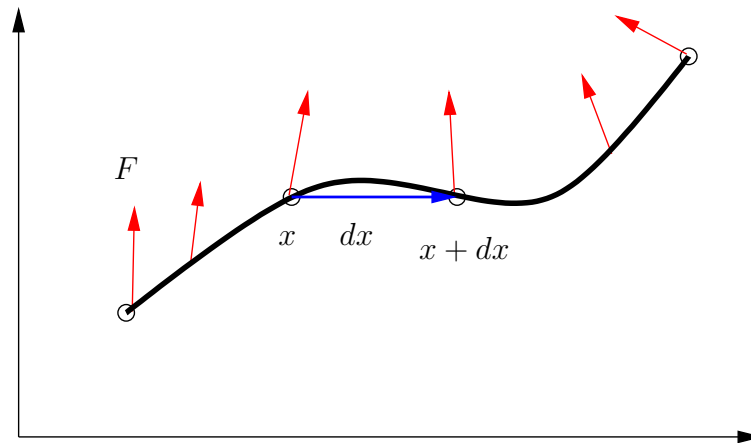


Die am Massenpunkt verrichtete Kraft lautet

$$W = \|F\| \cdot \|s\| \cdot \cos(\varphi) = \langle F, s \rangle.$$



(b) Jetzt betrachten wir ein inhomogenes (ortsabhängiges) Kraftfeld und die Bewegung des Massenpunkts entlang einer Kurve.



Die Kurve wird unterteilt in Kurvenabschnitte  $dx_k$ ,  $k = 1, \dots, N$ . Die Arbeit entlang eines infinitesimal kleinen Kurvenabschnitts lautet

$$dW_k = \langle F(\xi_k), dx_k \rangle,$$

wobei  $\xi_k$  ein beliebiger Punkt im  $k$ -ten Kurvenabschnitt ist. Durch Aufsummieren über alle Kurvenabschnitte erhalten wir eine Approximation der Gesamtarbeit:

$$W \approx \sum_{k=1}^N dW_k = \sum_{k=1}^N \langle F(\xi_k), dx_k \rangle,$$

die im Grenzübergang für  $N \rightarrow \infty$  gegen die geleistete Arbeit  $W$  konvergiert. ■

### Definition 1.2.19 (Kurvenintegral 2. Art)

Sei  $F : \mathbb{R}^d \rightarrow \mathbb{R}^d$  ein Vektorfeld und  $x : [a, b] \rightarrow \mathbb{R}^d$  eine Kurve. Konvergiert die in Beispiel 1.2.18, (b) konstruierte Summe

$$\sum_{k=1}^N \langle F(\xi_k), dx_k \rangle$$

für  $N \rightarrow \infty$  unabhängig von den gewählten Kurvenabschnitten und Zwischenpunkten  $\xi_k$ , so heißt der Grenzwert **Kurvenintegral 2. Art** und wird mit

$$\int_K \langle F(x), dx \rangle$$

bezeichnet.

Ist die Kurve geschlossen, so schreibt man

$$\oint_K \langle F(x), dx \rangle$$

und nennt ein solches Integral **Zirkulation des Vektorfeldes  $F$  entlang der Kurve**.

Im Unterschied zum Kurvenintegral 1. Art wird beim Kurvenintegral 2. Art nicht mit dem Bogenelement  $ds$  multipliziert, sondern man bildet das Skalarprodukt des vektoriiellen Kurvenlements  $dx = [dx_1, \dots, dx_d]^\top$ .

**Satz 1.2.20** (Berechnung des Kurvenintegrals 2. Art)

Für die stetig differenzierbare Kurve  $x : [a, b] \rightarrow \mathbb{R}^d$  mit Parametrisierung  $t \mapsto x(t)$  und dem Vektorfeld  $F : \mathbb{R}^d \rightarrow \mathbb{R}^d$  ist das Kurvenintegral 2. Art gegeben durch

$$\int_K \langle F(x), dx \rangle = \int_a^b \langle F(x(t)), x'(t) \rangle dt.$$

**Beweis:** Verwende das vektorielle Bogenelement

$$dx = \begin{bmatrix} dx_1 \\ \vdots \\ dx_d \end{bmatrix} = \begin{bmatrix} x'_1(t)dt \\ \vdots \\ x'_d(t)dt \end{bmatrix} = x'(t)dt.$$

■

**Bemerkung 1.2.21**

Durchläuft man die Kurve in umgekehrte Orientierung, also von  $b$  nach  $a$ , so ändert sich dadurch das Vorzeichen des Kurvenintegrals 2. Art, während das Kurvenintegral 1. Art sich dadurch nicht ändert.

■

**Beispiel 1.2.22**

Gegeben sei das Vektorfeld  $F : \mathbb{R}^2 \rightarrow \mathbb{R}^2$  mit

$$F(x_1, x_2) = \begin{bmatrix} 2x_1^2 + x_2 \\ 2x_2 \end{bmatrix}.$$

(a) Wir berechnen das Kurvenintegral 2. Art über die Kurve  $x : [0, 1] \rightarrow \mathbb{R}^2$  mit

$$x(t) = \begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix} := \begin{bmatrix} t \\ t^2 \end{bmatrix}.$$

Die Kurve beschreibt also einen Weg vom Punkt  $[0, 0]^\top$  zum Punkt  $[1, 1]^\top$ . Es gilt

$$\begin{aligned} \int_K \langle F(x), dx \rangle &= \int_0^1 \langle F(x(t)), x'(t) \rangle dt \\ &= \int_0^1 [2t^2 + t^2, 2t^2] \begin{bmatrix} 1 \\ 2t \end{bmatrix} dt \\ &= \int_0^1 3t^2 + 4t^3 dt \\ &= [t^3 + t^4]_0^1 = 2. \end{aligned}$$

(b) Nun berechnen wir Kurvenintegral 2. Art über einen anderen Weg vom Punkt  $[0, 0]^\top$  zum Punkt  $[1, 1]^\top$ , nämlich den direkten Weg, der durch die Kurve  $x : [0, 1] \rightarrow \mathbb{R}^2$  mit

$$x(t) = \begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix} := \begin{bmatrix} t \\ t \end{bmatrix}$$

gegeben ist. Es gilt

$$\begin{aligned} \int_K \langle F(x), dx \rangle &= \int_0^1 \langle F(x(t)), x'(t) \rangle dt \\ &= \int_0^1 [2t^2 + t, 2t] \begin{bmatrix} 1 \\ 1 \end{bmatrix} dt \\ &= \int_0^1 2t^2 + 3tdt \\ &= \left[ \frac{2}{3}t^3 + \frac{3}{2}t^2 \right]_0^1 = \frac{13}{6}. \end{aligned}$$

Das Beispiel zeigt, dass Kurvenintegrale 2. Art im Allgemeinen vom Weg abhängen und nicht nur von Anfangs- und Endpunkt der Kurve. ■

### Bemerkung 1.2.23 (Wegunabhängigkeit)

Falls das Vektorfeld  $F : \mathbb{R}^d \rightarrow \mathbb{R}^d$  ein **Potenzial**  $\phi : \mathbb{R}^d \rightarrow \mathbb{R}$  mit  $F = \nabla\phi$  besitzt, so gilt

$$\int_K \langle F(x), dx \rangle = \int_a^b \langle F(x(t)), x'(t) \rangle dt = \int_a^b \frac{d}{dt} \phi(x(t)) dt = \phi(x(b)) - \phi(x(a)).$$

In diesem Fall hängt das Kurvenintegral 2. Art also nur vom Anfangs- und Endwert der Kurve und nicht vom Verlauf dazwischen ab. Man nennt solche Vektorfelder auch **konservativ**. ■

### 1.3 Das Lebesgue-Integral

Das Riemann-Integral ist anschaulich definiert und es können nützliche Rechenregeln gezeigt werden, um Riemann-Integrale auszurechnen. Es ist allerdings technisch sehr aufwändig, den Begriff des Riemann-Integrals auf mehrdimensionale Bereiche  $\Omega \subset \mathbb{R}^d$  mit  $d > 1$  zu erweitern und damit dann Aussagen zu beweisen. Darüber hinaus fehlen dem Riemann-Integral einige wichtige mathematische Eigenschaften, etwa die Vertauschbarkeit von Integration und Grenzwertbildung (siehe Satz von Beppo Levi weiter unten). Es kann z.B. passieren, dass der Grenzwert von Riemann-integrierbaren Funktionen nicht mehr Riemann-integrierbar ist. Das Lebesgue-Integral beseitigt diese Unzulänglichkeiten und ermöglicht eine sehr allgemeine und in sich geschlossene Integrationstheorie, weshalb es in der mathematischen Analysis eine sehr wichtige Rolle spielt. Die Definition des Lebesgue-Integrals ist jedoch aufwändiger als die des Riemann-Integrals. Wir folgen der Darstellung in [BK11].

Zur Definition des Lebesgue-Integrals benötigt man zunächst ein sogenanntes Maß, welches bestimmten Teilmengen einer Menge  $\Omega \subset \mathbb{R}^d$  anschaulich ein Volumen (bzw. einen Inhalt) zuordnet. Wir beschränken uns in der Darstellung auf das Lebesgue-Maß, aber man kann Lebesgue-Integrale auch sehr viel allgemeiner über beliebigen Maßräumen definieren.

Das sogenannte **Lebesgue-Maß** ordnet einem  $d$ -dimensionalen (halboffenen) Intervall

$$[a, b) := [a_1, b_1) \times [a_2, b_2) \times \dots \times [a_d, b_d) \subset \mathbb{R}^d$$

mit  $a = (a_1, \dots, a_d)^\top$ ,  $b = (b_1, \dots, b_d)^\top$  und  $a < b$  sein Volumen

$$\mu([a, b)) = (b_1 - a_1) \cdot (b_2 - a_2) \cdots (b_d - a_d) = \prod_{i=1}^d (b_i - a_i)$$

zu, vgl. Abbildung 1.4.

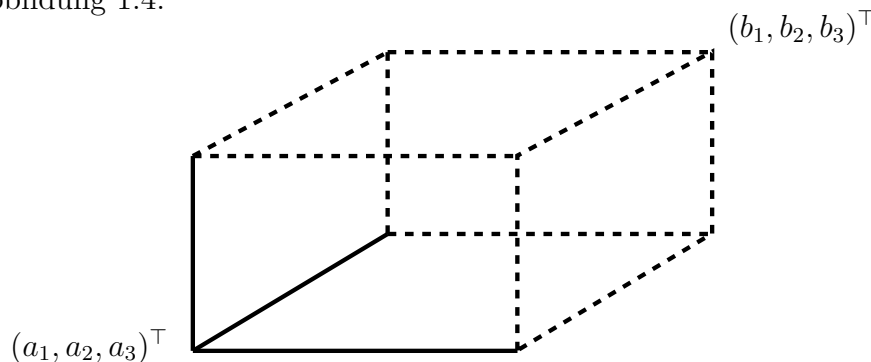


Abbildung 1.4: Halboffenes Intervall in  $\mathbb{R}^3$ .

Beachte, dass  $\mu$  den Wert 0 ergibt, falls  $a_j = b_j$  für ein  $j \in \{1, \dots, d\}$  ergibt. Ein solches degeneriertes Intervall besitzt also das Volumen 0 im  $\mathbb{R}^d$ . Man spricht dann auch von einer **Nullmenge**.

Mit diesem Maß kann man also einem  $d$ -dimensionalen Intervall dessen Volumen zuordnen. Leider ist nicht jede Teilmenge des  $\mathbb{R}^d$  ein  $d$ -dimensionales Intervall und man kann auch nicht jede Teilmenge des  $\mathbb{R}^d$  als Vereinigung von endlich vielen  $d$ -dimensionalen Intervallen darstellen. Wir benötigen also ein Konstrukt, welches erlaubt, sehr allgemeine Mengen aus den Intervallen zu bilden. Dies führt auf die sogenannte **(Borel-)  $\sigma$ -Algebra**. Darunter versteht man ein Mengensystem, welches zwar in seiner Allgemeinheit schwer vorstellbar ist, aber mathematisch sehr gute Eigenschaften besitzt.

**Definition 1.3.1** ( $\sigma$ -Algebra, Maße, Nullmengen)

Sei  $\Omega \subset \mathbb{R}^d$  eine nichtleere Menge.

(a) Jedes System  $\mathcal{A}$  von Teilmengen von  $\Omega$  mit den Eigenschaften

- (i)  $\Omega \in \mathcal{A}$ ;
- (ii) Ist  $A \in \mathcal{A}$ , so ist auch dessen komplementäre Menge  $A^c := \Omega \setminus A$  in  $\mathcal{A}$  enthalten;
- (iii) Für die abzählbar vielen Mengen  $A_1, A_2, \dots \in \mathcal{A}$  gilt, dass auch deren Vereinigung  $\bigcup_{i=1}^{\infty} A_i$  in  $\mathcal{A}$  enthalten ist;

heißt  **$\sigma$ -Algebra**. Das Paar  $(\Omega, \mathcal{A})$  heißt **messbarer Raum**. Die Elemente von  $\mathcal{A}$  sind die **messbaren Teilmengen** von  $\Omega$ .

(b) Eine Abbildung  $\mu : (\Omega, \mathcal{A}) \rightarrow [0, \infty)$  heißt **Maß**, falls gilt:

- (i)  $\mu(\emptyset) = 0$ ;
- (ii)  $\sigma$ -Additivität: Für abzählbar viele paarweise disjunkte messbare Mengen  $A_i \in \mathcal{A}$ ,  $i = 1, 2, \dots$ , gilt

$$\mu \left( \bigcup_{i=1}^{\infty} A_i \right) = \sum_{i=1}^{\infty} \mu(A_i).$$

Das Tripel  $(\Omega, \mathcal{A}, \mu)$  heißt **Maßraum**.

(c)  $A \subset \Omega$  heißt **Nullmenge**, falls  $A \in \mathcal{A}$  und  $\mu(A) = 0$  gilt.

Die Vorschriften in (a) geben an, wie man aus messbaren Mengen neue messbare Mengen erzeugen kann, nämlich durch Komplementbildung und Vereinigung von höchstens abzählbar vielen Mengen. Die Gesamtheit der dadurch erzeugbaren Mengen ist sehr reichhaltig.

**Bemerkung 1.3.2**

Interessant sind auch die Nullmengen, da diese sich von der leeren Menge bzgl. ihres Maßes nicht unterscheiden lassen. Später werden Eigenschaften häufig mit dem Zusatz **“fast überall (f.ü.)”** versehen, was bedeutet, dass die Eigenschaft mit Ausnahme auf einer Nullmenge gilt. Beispiele für Nullmengen in  $\mathbb{R}^2$  sind Geraden und einzelne Punkte. Im

$\mathbb{R}^3$  sind Ebenen Nullmengen. Allgemeiner sind Hyperebenen (also Ebenen der Dimension  $d - 1$ ) Nullmengen im  $\mathbb{R}^d$ . Des Weiteren ist die Menge der rationalen Zahlen  $\mathbb{Q}$  eine Nullmenge in der Menge der reellen Zahlen  $\mathbb{R}$ . ■

Es stellt sich die Frage nach einem Erzeugendensystem, mit dem man alle weiteren Mengen erzeugen kann. Hierbei heißt  $\mathcal{E} \subset \Omega$  ein **Erzeuger der  $\sigma$ -Algebra  $\mathcal{A}$** , falls  $\mathcal{A}$  die kleinste  $\sigma$ -Algebra in  $\Omega$  ist, die  $\mathcal{E}$  umfasst und man schreibt  $\sigma(\mathcal{E}) = \mathcal{A}$ . Die von den endlichen halboffenen Intervallen

$$[a, b) \quad \text{mit } a < b, a, b \in \mathbb{R}^d$$

auf  $\mathbb{R}^d$  erzeugte  $\sigma$ -Algebra nennt man **Borel- $\sigma$ -Algebra**, man bezeichnet sie mit  $\mathcal{B}$ . Für eine messbare Teilmenge  $\Omega \subset \mathbb{R}^d$  bildet dann  $\{A \subset \Omega \mid A \in \mathcal{B}\}$  eine  $\sigma$ -Algebra auf  $\Omega$ , die wir der Einfachheit halber ebenfalls mit  $\mathcal{B}$  bezeichnen. Abbildung 1.5 illustriert einige Elemente der Borel- $\sigma$ -Algebra für eine Menge  $\Omega$ .

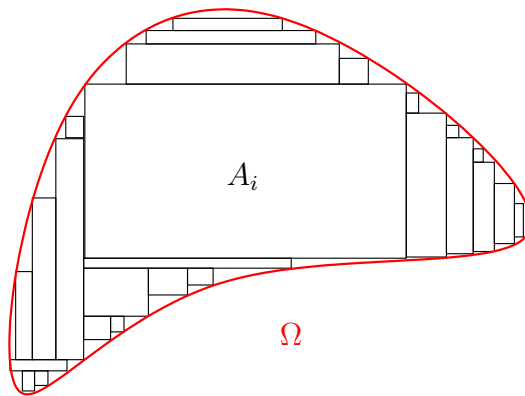


Abbildung 1.5: Illustration einiger Elemente  $A_i$  der Borel- $\sigma$ -Algebra für eine Menge  $\Omega \subset \mathbb{R}^2$ .

Man kann zeigen, dass  $(\mathbb{R}^d, \mathcal{B}, \mu)$  ein Maßraum ist, wobei  $\mu$  das oben definierte Lebesgue-Maß bezeichnet. Mithilfe des Lebesgue-Maßes können wir nun also jeder Borel-Menge aus  $\mathcal{B}$  ein Volumen zuordnen. Damit können wir auch Urbildern unter Abbildungen ein Volumen zuordnen, vgl. Abbildung 1.6.

### Definition 1.3.3 (messbare Abbildung)

Sei  $f : (\Omega, \mathcal{A}) \longrightarrow (\tilde{\Omega}, \tilde{\mathcal{A}})$  eine Abbildung zwischen zwei messbaren Räumen.  $f$  heißt **messbar**, falls  $f^{-1}(\tilde{A}) \in \mathcal{A}$  für alle  $\tilde{A} \in \tilde{\mathcal{A}}$ .

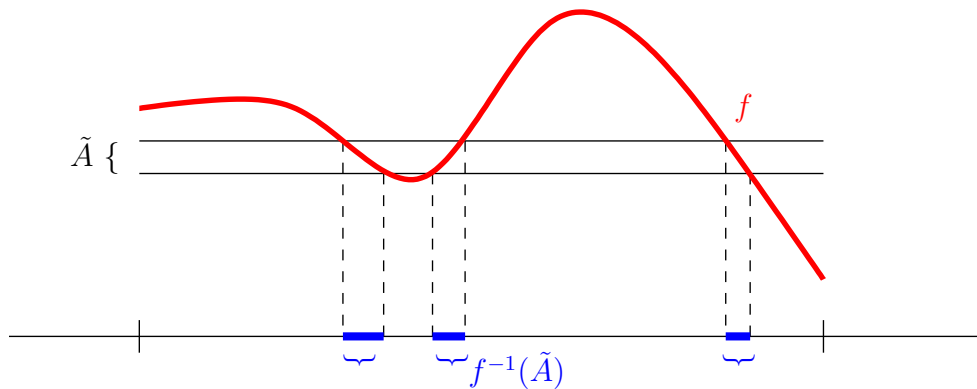


Abbildung 1.6: Messbarkeit einer Abbildung.

**Beispiel 1.3.4**

Beispiele für messbare Funktionen und Mengen sind:

(a) stetige Funktionen;

(b) charakteristische Funktion  $\chi_A : \mathbb{R}^d \rightarrow \mathbb{R}$  mit

$$\chi_A(x) = \begin{cases} 1, & \text{falls } x \in A, \\ 0, & \text{sonst;} \end{cases}$$

(c) Linearkombinationen  $\alpha f + \beta g$  mit  $\alpha, \beta \in \mathbb{R}$  von messbaren Funktionen  $f, g$ ;

(d) Produkt  $f \cdot g$  von messbaren Funktionen  $f$  und  $g$ ;

(e)  $\max\{f, g\}$ ,  $\min\{f, g\}$ ,  $|f|$ , falls  $f$  und  $g$  messbar sind;

(f) Die Mengen  $\{x \in \Omega \mid f(x) = g(x)\}$  und  $\{x \in \Omega \mid f(x) \leq g(x)\}$  sind messbar, falls  $f$  und  $g$  messbar sind.

■

Die folgende Eigenschaft wird uns in allgemeinerer Form noch bei der Berechnung von Integralen begegnen. Sei  $f : \mathbb{R}^d \rightarrow \mathbb{R}^d$  eine lineare Abbildung, d.h. man kann  $f$  darstellen als  $f(x) = Mx$  mit einer Matrix  $M \in \mathbb{R}^{d \times d}$ . Wendet man  $f$  auf eine messbare Menge  $B \subset \mathbb{R}^d$  an, so ist auch  $f(B)$  eine messbare Menge und es gilt

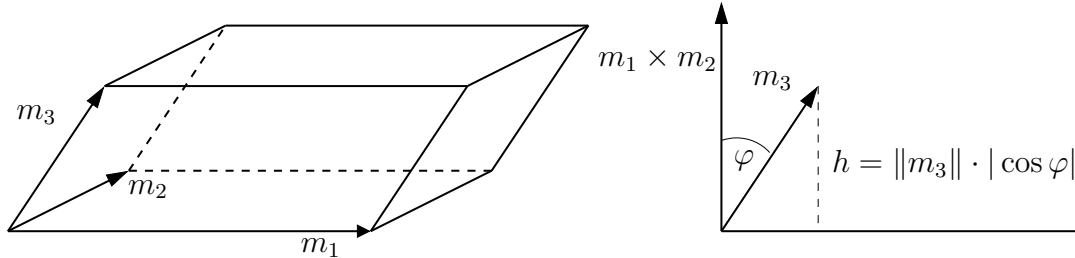
$$\mu(f(B)) = |\det(M)|\mu(B). \quad (1.1)$$

Ein formaler Beweis findet sich in [BK11, Satz III.4]. Wir illustrieren diese Eigenschaft an einem Beispiel.

**Beispiel 1.3.5 (Illustration)**

Wir illustrieren die Eigenschaft (1.1) im  $\mathbb{R}^3$  für den Einheitswürfel  $B$ , welcher durch

die Einheitsvektoren  $e_1 = (1, 0, 0)^\top$ ,  $e_2 = (0, 1, 0)^\top$  und  $e_3 = (0, 0, 1)^\top$  aufgespannt wird. Unter der Abbildung  $f$  werden die Einheitsvektoren auf die Vektoren  $m_i = Me_i$ ,  $i = 1, 2, 3$ , abgebildet, welche ein sogenanntes Parallelepiped (oder Spat) aufspannen (dies ist gerade die Menge  $f(B)$ ). Die Vektoren  $m_i$ ,  $i = 1, 2, 3$ , sind gerade die Spalten von  $M$ .



Das Volumen  $V = \mu(f(B))$  des Parallelepipeds berechnet sich zu

$$V = \|m_1 \times m_2\| \cdot h = \|m_1 \times m_2\| \cdot \|m_3\| \cdot |\cos \varphi| = |\langle m_1 \times m_2, m_3 \rangle| = |\det(M)|.$$

■

Das Lebesgue-Integral wird schrittweise zunächst für elementare Funktionen, dann für nichtnegative Funktionen und schließlich für allgemeine Funktionen eingeführt.

Unter einer elementaren Funktion wird eine messbare Funktion verstanden, die nur endlich viele Funktionswerte annimmt. Man kann sie sich als Treppenfunktion vorstellen, siehe Abbildung 1.7.

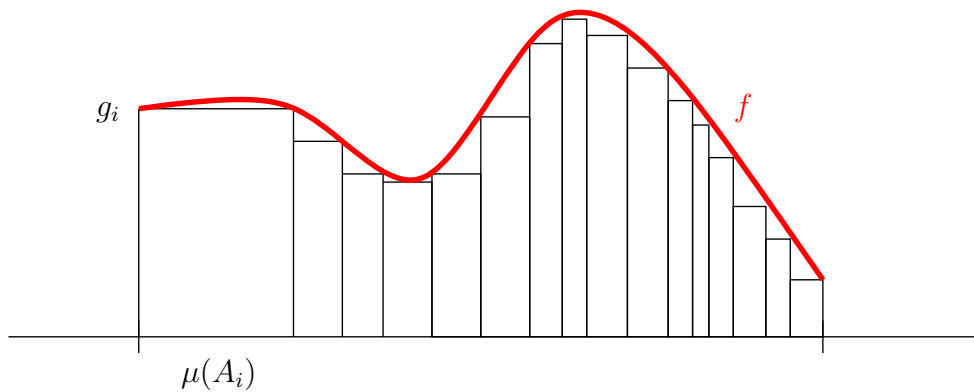


Abbildung 1.7: Elementare Funktion (Treppenfunktion) und Idee des Lebesgue-Integrals für  $f$ .

Ist  $g : \Omega \rightarrow \mathbb{R}$  eine solche elementare Funktion mit Werten in der Menge  $\{g_1, \dots, g_N\}$ , so kann man sie darstellen als

$$g(x) = \sum_{i=1}^N g_i \cdot \chi_{A_i}(x) \quad \text{mit } A_i := \{x \in \Omega \mid g(x) = g_i\}.$$

Für eine elementare Funktion  $g$  definiert man:



**Lebesgue-Integral einer elementaren Funktion:** Sei  $g : \Omega \rightarrow \mathbb{R}$  elementar mit Werten  $g_i, i = 1, \dots, N$ . Dann ist das Lebesgue-Integral definiert als

$$\int_{\Omega} g \, d\mu := \sum_{i=1}^N g_i \cdot \mu(A_i) \quad \text{mit } A_i = \{x \in \Omega \mid g(x) = g_i\},$$

wobei  $\mu$  das Lebesgue-Maß bezeichnet.

Mithilfe des Lebesgue-Integrals für elementare Funktionen definiert man nun das Lebesgue-Integral für messbare Funktionen  $f \geq 0$ , siehe Abbildung 1.7.

**Lebesgue-Integral für messbare, nichtnegative Funktionen:** Sei  $f \geq 0$  messbar. Dann ist das Lebesgue-Integral definiert als

$$\int_{\Omega} f \, d\mu := \sup \left\{ \int_{\Omega} g \, d\mu \mid \text{mit } g : \Omega \rightarrow \mathbb{R} \text{ ist elementar und } 0 \leq g \leq f \right\},$$

wobei  $\mu$  das Lebesgue-Maß bezeichnet.

Durch Zerlegung einer beliebigen messbaren Funktion  $f : \Omega \rightarrow \mathbb{R}$  in ihren Positiv- und Negativanteil gemäß

$$f = f^+ - f^- \quad \text{mit } f^+ = \max\{f, 0\}, \quad f^- = \max\{-f, 0\}$$

kann man das Lebesgue-Integral nun auch für  $f$  definieren.

### Definition 1.3.6 (Lebesgue-Integral)

Sei  $f : \Omega \rightarrow \mathbb{R}$  messbar und  $\mu$  das Lebesgue-Maß. Dann heißt  $f$  **(Lebesgue-)integrierbar**, falls  $\int_{\Omega} f^+ \, d\mu < \infty$  und  $\int_{\Omega} f^- \, d\mu < \infty$  gelten. Das **(Lebesgue-)Integral** ist definiert als

$$\int_{\Omega} f \, d\mu := \int_{\Omega} f^+ \, d\mu - \int_{\Omega} f^- \, d\mu.$$

Wegen  $|f| = f^+ + f^-$  ist  $f$  Lebesgue-integrierbar, genau dann, wenn  $\int_{\Omega} |f| \, d\mu < \infty$ .

Das Lebesgue-Integral besitzt die folgenden Eigenschaften, vgl. [BK11]:

- Es gilt  $|\int_{\Omega} f \, d\mu| \leq \int_{\Omega} |f| \, d\mu$ .
- **Monotonie:**  $\int_{\Omega} f \, d\mu \leq \int_{\Omega} g \, d\mu$  für Lebesgue-integrierbare Funktionen  $f$  und  $g$  mit  $f \leq g$ .

- **Linearität:**  $\int_{\Omega}(\alpha f + \beta g) d\mu = \alpha \int_{\Omega} f d\mu + \beta \int_{\Omega} g d\mu$  für  $\alpha, \beta \in \mathbb{R}$  und Lebesgue-integrierbare Funktionen  $f$  und  $g$ .
- Für messbare Funktionen  $f, g \geq 0$  gilt:

$$f \leq g \text{ fast überall} \implies \int_{\Omega} f d\mu \leq \int_{\Omega} g d\mu$$

$$f = g \text{ fast überall} \implies \int_{\Omega} f d\mu = \int_{\Omega} g d\mu$$

$$\int_{\Omega} f d\mu = 0 \iff f = 0 \text{ fast überall in } \Omega$$

$$\int_{\Omega} f d\mu < \infty \implies f < \infty \text{ fast überall in } \Omega$$

- **Satz von der monotonen Konvergenz (Satz von Beppo Levi):** Es gelte  $0 \leq f_1 \leq f_2 \leq \dots$  für messbare Funktionen  $f_i, i = 1, 2, \dots$ , und  $f := \sup_{i=1,2,\dots} f_i$ . Dann gilt

$$\int_{\Omega} f d\mu = \lim_{i \rightarrow \infty} \int_{\Omega} f_i d\mu.$$

(dies ist die Vertauschbarkeit von Integration und Grenzwertbildung)

- **Lemma von Fatou:** Es seien  $f, f_1, f_2, \dots \geq 0$  messbare Funktionen mit  $f \leq \liminf_i f_i$  fast überall. Dann gilt

$$\int_{\Omega} f d\mu \leq \liminf_{i \rightarrow \infty} \int_{\Omega} f_i d\mu.$$

- **Lebesgue'scher Konvergenzsatz:** Seien  $f_1, f_2, \dots$  messbare Funktionen, die fast überall (also bis auf eine Nullmenge) gegen die messbare Funktion  $f$  konvergieren. Des Weiteren gelte fast überall  $|f_i| \leq g$  für alle  $i \in \mathbb{N}$  für eine messbare Funktion  $g \geq 0$  mit  $\int_{\Omega} g d\mu < \infty$ . Dann sind die Funktionen  $f_1, f_2, \dots$  Lebesgue-integrierbar und es gelten die Beziehungen

$$\int_{\Omega} f_i d\mu \longrightarrow \int_{\Omega} f d\mu \quad \text{und} \quad \int_{\Omega} |f_i - f| d\mu \longrightarrow 0$$

für  $i \longrightarrow \infty$ .

### Bemerkung 1.3.7

*Verschiedene Schreibweisen für das Lebesgue-Integral sind*

$$\int_{\Omega} f d\mu \quad \text{oder} \quad \int_{\Omega} f(x) d\mu(x) \quad \text{oder} \quad \int_{\Omega} f(x) \mu(dx) \quad \text{oder} \quad \int_{\Omega} \mu(dx) f(x).$$

Wir werden häufig die klassische Schreibweise

$$\int_{\Omega} f(x) dx$$

verwenden. Hierbei droht eine Verwechslung mit dem Riemann-Integral, aber solange nichts anderes gesagt ist, meinen wir im Folgenden das Lebesgue-Integral bzw. gehen davon aus, dass Riemann- und Lebesgue-Integral beide existieren und gleich sind. ■

Es gibt Funktionen, die Lebesgue-integrierbar aber nicht Riemann-integrierbar sind.

**Beispiel 1.3.8 (Dirichlet-Funktion)**

Die Dirichlet-Funktion  $f : [0, 1] \rightarrow [0, 1]$  mit

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

ist Lebesgue-integrierbar mit Wert 0, aber nicht Riemann-integrierbar, da die Untersumme stets 0 und die Obersumme stets 1 ist. ■

Umgekehrt gibt es Riemann-integrierbare Funktionen, die nicht Lebesgue-integrierbar sind.

**Beispiel 1.3.9**

Die Funktion  $f : \mathbb{R} \rightarrow \mathbb{R}$  mit

$$f(x) = \frac{(-1)^n}{n} \quad \text{für } n-1 \leq x < n, \quad n \in \mathbb{N}$$

besitzt das Riemann-Integral

$$\int_0^{\infty} f(x) dx = \sum_{n=1}^{\infty} \frac{(-1)^n}{n} < \infty,$$

wobei die alternierende harmonische Reihe konvergiert, während das Lebesgue-Integral nicht existiert, weil

$$\int_0^{\infty} |f| d\mu = \sum_{n=1}^{\infty} \frac{1}{n}$$

die divergente harmonische Reihe ist. ■

**Bemerkung 1.3.10**

Ein Zusammenhang zwischen Riemann- und Lebesgue-Integral findet sich in Satz V.10 in [BK11]: Eine beschränkte Funktion  $f : [a, b] \rightarrow \mathbb{R}$  ist genau dann Riemann-integrierbar, wenn die Menge der Unstetigkeitsstellen von  $f$  eine Nullmenge bzgl. des Lebesgue-Maßes  $\mu$  ist. Dann gilt

$$\int_a^b f(x) dx = \int_{[a,b]} f \cdot \chi_C d\mu,$$

wobei  $C$  die Menge der Stetigkeitsstellen von  $f$  bezeichnet. Für eine stetige Funktion auf  $[a, b]$  gilt insbesondere  $\chi_C = 1$ . ■

Schließlich zitieren wir noch ein Resultat, welches man mithilfe von (1.1) zeigen kann, siehe [BK11, S. 107]. Dieses wird uns später bei der Berechnung von Integralen begegnen.

**Satz 1.3.11 (Substitutionsregel (Transformationsformel von Jacobi))**

Seien  $f : \Omega \rightarrow \mathbb{R}$  und  $\varphi : G \rightarrow \Omega$  Abbildungen mit  $G, \Omega \subset \mathbb{R}^d$  offen und  $f \geq 0$  messbar.  $\varphi$  sei eine bijektive Abbildung und  $\varphi$  und  $\varphi^{-1}$  seien stetig differenzierbar. Dann gilt

$$\int_{\Omega} f(x) \, dx = \int_G f(\varphi(\xi)) \cdot |\det(\varphi'(\xi))| \, d\xi.$$

$\det(\varphi'(\xi))$  heißt **Jacobische Funktionaldeterminante**.

## 1.4 Flächen- und Volumenintegrale

In Abschnitt 1.2 haben wir entlang Kurven integriert. In diesem Abschnitt gehen wir einen Schritt weiter und integrieren über Flächen oder Volumen und greifen auf das Lebesgue-Integral zurück, vgl. Abbildung 1.8.

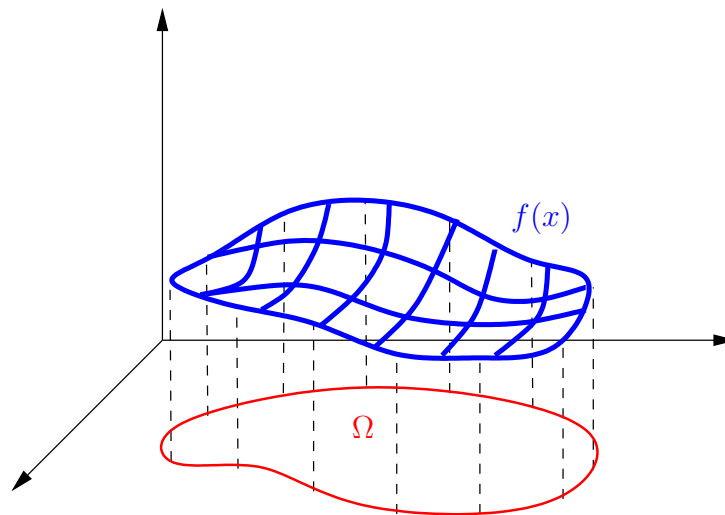


Abbildung 1.8: Illustration einer Funktion über einem Gebiet  $\Omega$ .

Gegeben sei eine Fläche oder ein Volumen  $\Omega \subset \mathbb{R}^d$  mit  $d = 2, 3$  und eine messbare Funktion  $f : \Omega \rightarrow \mathbb{R}$ . Gesucht ist das Integral

$$\int_{\Omega} f(x) \, d\mu \quad \text{bzw.} \quad \int_{\Omega} f(x) \, dx,$$

wobei  $\mu$  das Lebesgue-Maß bezeichnet. Für Flächenintegrale verwendet man anstatt  $x = [x_1, x_2]^\top$  häufig die (kartesischen) Koordinaten  $[x, y]^\top$  und schreibt

$$\iint_{\Omega} f(x, y) \, dx dy,$$

welches man als **Doppelintegral** bezeichnet. Entsprechend werden für Volumenintegrale anstatt  $x = [x_1, x_2, x_3]^\top$  die (kartesischen) Koordinaten  $[x, y, z]^\top$  verwendet, und man schreibt

$$\iiint_{\Omega} f(x, y, z) \, dx dy dz,$$

welches man als **Dreifachintegral** bezeichnet. Diese Schreibweise verbirgt eine Schwierigkeit, da suggeriert wird, dass zuerst über  $x$ , dann über  $y$  und schließlich über  $z$  integriert wird. Das stimmt nicht immer, wie der folgende Abschnitt zeigt.

Flächenintegrale werden z.B. zur Bestimmung des Schwerpunkts eines ebenen Bereiches  $\Omega$  benötigt (man kann sich dies als dünne Platte vorstellen) oder zur Berechnung des Flächeninhalts eines ebenen Bereiches  $\Omega$ .

#### Beispiel 1.4.1

Sei  $\Omega$  ein Gebiet im  $\mathbb{R}^2$ .

(a) Der Flächeninhalt  $A$  von  $\Omega$  ist gegeben durch

$$A = \int_{\Omega} d\mu .$$

(b) Der Schwerpunkt  $[x_s, y_s]^\top$  von  $\Omega$  ist gegeben durch

$$x_s = \frac{1}{A} \int_{\Omega} x \, d\mu, \quad y_s = \frac{1}{A} \int_{\Omega} y \, d\mu .$$

(c) Die axialen Flächenträgheitsmomente bzgl.  $x$ - und  $y$ -Achse sind gegeben durch

$$I_x = \int_{\Omega} y^2 \, d\mu, \quad I_y = \int_{\Omega} x^2 \, d\mu .$$

(d) Ist die Fläche ein Modell für eine dünne Platte mit Masse  $M$  mit Dichteverteilung  $\rho$  (Masse pro Flächeneinheit), so ergibt sich der Masseschwerpunkt  $[x_m, y_m]^\top$  durch

$$x_m = \frac{1}{M} \int_{\Omega} x \rho(x, y) \, d\mu, \quad y_m = \frac{1}{M} \int_{\Omega} y \rho(x, y) \, d\mu ,$$

wobei die Masse  $M$  gegeben ist durch

$$M = \int_{\Omega} \rho(x, y) \, d\mu .$$

■

Analoge Größen werden für 3-dimensionale Körper  $\Omega$  definiert.

### Beispiel 1.4.2

Sei  $\Omega$  ein Gebiet im  $\mathbb{R}^3$ .

(a) Das Volumen  $V$  von  $\Omega$  ist gegeben durch

$$V = \int_{\Omega} d\mu .$$

(b) Der Schwerpunkt  $[x_s, y_s, z_s]^{\top}$  von  $\Omega$  ist gegeben durch

$$x_s = \frac{1}{V} \int_{\Omega} x \, d\mu, \quad y_s = \frac{1}{V} \int_{\Omega} y \, d\mu, \quad z_s = \frac{1}{V} \int_{\Omega} z \, d\mu .$$

(c) Modelliert das Gebiet einen dreidimensionalen Körper mit Masse  $M$  mit Dichteverteilung  $\rho$  (Masse pro Volumeneinheit), so ergibt sich der Masseschwerpunkt  $[x_m, y_m, z_m]^{\top}$  durch

$$x_m = \frac{1}{M} \int_{\Omega} x \rho(x, y, z) \, d\mu, \quad y_m = \frac{1}{M} \int_{\Omega} y \rho(x, y, z) \, d\mu, \quad z_m = \frac{1}{M} \int_{\Omega} z \rho(x, y, z) \, d\mu ,$$

wobei die Masse  $M$  gegeben ist durch

$$M = \int_{\Omega} \rho(x, y, z) \, d\mu .$$

■

#### 1.4.1 Flächen- und Volumenintegrale mit kartesischen Koordinaten

Wir widmen uns nun der konkreten Berechnung des Flächenintegrals. Hierbei wird auch klar werden, warum man von einem Doppelintegral spricht.

### Beispiel 1.4.3 (Doppelintegral über rechteckigem Bereich ausgedrückt in kartesischen Koordinaten)

Wir möchten das Flächenintegral

$$I = \int_{\Omega} f(x, y) \, d\mu \quad \text{mit} \quad f(x, y) = xy^2$$

über dem ebenen Rechteck

$$\Omega = \{[x, y]^{\top} \in \mathbb{R}^2 \mid 1 \leq x \leq 3, 0 \leq y \leq 2\}$$

ausrechnen. Das Flächenintegral (Doppelintegral!) ergibt sich zu

$$I = \int_{y=0}^{y=2} \int_{x=1}^{x=3} xy^2 \, dx dy = \int_{y=0}^{y=2} \left[ \frac{1}{2} x^2 y^2 \right]_{x=1}^{x=3} dy = \int_{y=0}^{y=2} 4y^2 \, dy = \left[ \frac{4}{3} y^3 \right]_{y=0}^{y=2} = \frac{32}{3}.$$

Man könnte die Integrationsreihenfolge bzgl.  $x$  und  $y$  auch vertauschen und erhält in diesem Fall dasselbe Ergebnis:

$$I = \int_{x=1}^{x=3} \int_{y=0}^{y=2} xy^2 \, dy dx = \int_{x=1}^{x=3} \left[ \frac{1}{3} xy^3 \right]_{y=0}^{y=2} dx = \int_{x=1}^{x=3} \frac{8}{3} x \, dx = \left[ \frac{4}{3} x^2 \right]_{x=1}^{x=3} = \frac{32}{3}. \quad \blacksquare$$

Im vorangegangenen Beispiel war es egal, ob zunächst über  $x$  und dann über  $y$  integriert wird oder umgekehrt. Das folgende Beispiel zeigt, dass die Reihenfolge nicht einfach vertauscht werden darf, wenn die Intervallgrenzen von  $x$  bzw.  $y$  abhängen.

#### Beispiel 1.4.4 (Doppelintegral über komplizierterem Bereich ausgedrückt in kartesischen Koordinaten)

Wir möchten das Flächenintegral

$$I = \int_{\Omega} f(x, y) \, d\mu \quad \text{mit} \quad f(x, y) = y^2$$

über der Dreiecksfläche

$$\begin{aligned} \Omega &= \Omega_1 \cup \Omega_2, \\ \Omega_1 &= \{[x, y]^T \in \mathbb{R}^2 \mid 0 \leq x \leq 2, 0 \leq y \leq x\}, \\ \Omega_2 &= \{[x, y]^T \in \mathbb{R}^2 \mid 2 \leq x \leq 3, 0 \leq y \leq 6 - 2x\}, \end{aligned}$$

ausrechnen. Wegen der Additivität des Flächenintegrals gilt

$$\begin{aligned} \int_{\Omega} f(x, y) \, d\mu &= \int_{\Omega_1} f(x, y) \, d\mu + \int_{\Omega_2} f(x, y) \, d\mu \\ &= \int_{x=0}^{x=2} \int_{y=0}^{y=x} y^2 \, dy dx + \int_{x=2}^{x=3} \int_{y=0}^{y=6-2x} y^2 \, dy dx \\ &= \int_{x=0}^{x=2} \left[ \frac{1}{3} y^3 \right]_{y=0}^{y=x} dx + \int_{x=2}^{x=3} \left[ \frac{1}{3} y^3 \right]_{y=0}^{y=6-2x} dx \\ &= \int_{x=0}^{x=2} \frac{1}{3} x^3 \, dx + \int_{x=2}^{x=3} \frac{1}{3} (6 - 2x)^3 \, dx \\ &= \left[ \frac{1}{12} x^4 \right]_{x=0}^{x=2} + \left[ -\frac{1}{24} (6 - 2x)^4 \right]_{x=2}^{x=3} \\ &= \frac{4}{3} + \frac{2}{3} = 2. \end{aligned}$$

Beachte, dass die Intervallgrenzen für die Integrationsvariable  $y$  von  $x$  abhängen. Daher darf man hier nicht einfach die obige Integrationsreihenfolge vertauschen. Man kann die Fläche  $\Omega$  jedoch auch auf folgende Art darstellen (dies ist eine andere Parametrisierung von  $\Omega$ ):

$$\tilde{\Omega} = \{[x, y]^T \in \mathbb{R}^2 \mid 0 \leq y \leq 2, y \leq x \leq 3 - \frac{1}{2}y\} = \Omega.$$

Mit dieser Parametrisierung von  $\Omega$  folgt:

$$\begin{aligned} \int_{\tilde{\Omega}} f(x, y) \, d\mu &= \int_{y=0}^{y=2} \int_{x=y}^{x=3-\frac{1}{2}y} y^2 \, dx dy \\ &= \int_{y=0}^{y=2} [xy^2]_{x=y}^{x=3-\frac{1}{2}y} \, dy \\ &= \int_{y=0}^{y=2} y^2(3 - \frac{1}{2}y) - y^3 \, dy \\ &= \left[ y^3 - \frac{3}{8}y^4 \right]_{y=0}^{y=2} \\ &= 8 - 6 = 2. \end{aligned}$$

■

In Beispiel 1.4.4 traten ineinander verschachtelte Integrale auf. Man spricht hier auch von iterierten Integralen.

#### Definition 1.4.5 (Iteriertes Integral, Normalbereich)

Man spricht von einem **iterierten Integral**, wenn mehrere Integrale ineinander verschachtelt sind, wobei die Integrationsgrenzen der inneren Integrale von den Integrationsvariablen der äußeren Integrale abhängen, d.h. iterierte Integrale sind von der Form

$$\int_{x_1=a_1}^{x_1=b_1} \left[ \int_{x_2=a_2(x_1)}^{x_2=b_2(x_1)} \left[ \cdots \left[ \int_{x_d=a_d(x_1, \dots, x_{d-1})}^{x_d=b_d(x_1, \dots, x_{d-1})} f(x_1, \dots, x_d) dx_d \right] \cdots \right] dx_2 \right] dx_1,$$

wobei die Integrale nacheinander von innen nach außen zu berechnen sind. Bereiche der Form

$$\Omega = \left\{ x = [x_1, \dots, x_d]^T \in \mathbb{R}^d \mid \begin{array}{l} a_1 \leq x_1 \leq b_1, \ a_2(x_1) \leq x_2 \leq b_2(x_1), \ \dots, \\ a_d(x_1, \dots, x_{d-1}) \leq x_d \leq b_d(x_1, \dots, x_{d-1}) \end{array} \right\}$$

heißen **Normalbereiche**.

Der folgende Satz rechtfertigt die Vorgehensweise in den Beispielen 1.4.3-1.4.4.



**Satz 1.4.6** (Berechnung des Flächenintegrals in kartesischen Koordinaten über Normalbereiche)

Sei  $f : \mathbb{R}^2 \rightarrow \mathbb{R}$  stetig. Dann gelten die folgenden Aussagen:

(a) Für  $\Omega = \{[x, y]^\top \in \mathbb{R}^2 \mid a_1 \leq x \leq b_1, a_2 \leq y \leq b_2\}$  gilt

$$\int_{\Omega} f(x, y) \, d\mu = \int_{x=a_1}^{x=b_1} \int_{y=a_2}^{y=b_2} f(x, y) \, dy dx = \int_{y=a_2}^{y=b_2} \int_{x=a_1}^{x=b_1} f(x, y) \, dx dy.$$

(Satz von Fubini)

(b) Für  $\Omega = \{[x, y]^\top \in \mathbb{R}^2 \mid a_1 \leq x \leq b_1, a_2(x) \leq y \leq b_2(x)\}$  mit stetigen Funktionen  $a_2, b_2$  gilt

$$\int_{\Omega} f(x, y) \, d\mu = \int_{x=a_1}^{x=b_1} \int_{y=a_2(x)}^{y=b_2(x)} f(x, y) \, dy dx.$$

(c) Für  $\Omega = \{[x, y]^\top \in \mathbb{R}^2 \mid a_1(y) \leq x \leq b_1(y), a_2 \leq y \leq b_2\}$  mit stetigen Funktionen  $a_1, b_1$  gilt

$$\int_{\Omega} f(x, y) \, d\mu = \int_{y=a_2}^{y=b_2} \int_{x=a_1(y)}^{x=b_1(y)} f(x, y) \, dx dy.$$

Für Volumenintegrale (Dreifachintegrale) geht man vollkommen analog vor, weshalb wir nur ein Beispiel betrachten. Satz 1.4.6 gilt analog.

**Beispiel 1.4.7**

Das axiale Trägheitsmoment des Würfels

$$\Omega = \{[x, y, z]^\top \in \mathbb{R}^3 \mid -1 \leq x, y, z \leq 1\}$$

um die z-Achse ist gegeben durch

$$\begin{aligned} I &= \int_{-1}^1 \int_{-1}^1 \int_{-1}^1 x^2 + y^2 \, dx dy dz = \int_{-1}^1 \int_{-1}^1 \left[ \frac{1}{3} x^3 + y^2 x \right]_{-1}^1 dy dz \\ &= \int_{-1}^1 \int_{-1}^1 \frac{2}{3} + 2y^2 \, dy dz \\ &= \int_{-1}^1 \frac{4}{3} + \frac{4}{3} dz = \frac{16}{3}. \end{aligned}$$

■

### 1.4.2 Parametrisierung von Flächen und Volumen

Ähnlich wie bei der Parametrisierung von Kurven müssen wir die Flächen oder Volumina beschreiben, indem wir sie mit geeigneten Koordinaten **parametrisieren**. Bisher haben wir hierfür die “normalen” kartesischen (rechtwinkligen) Koordinaten verwendet. Anstatt der kartesischen Koordinaten werden häufig spezielle Koordinaten verwendet, etwa Polarkoordinaten, Kugelkoordinaten oder Zylinderkoordinaten. Dabei drückt man die üblichen kartesischen Koordinaten  $x = [x_1, x_2, \dots, x_d]^\top$  in  $\Omega$  durch andere, geeignetere Koordinaten  $\xi = [\xi_1, \xi_2, \dots, \xi_d]^\top$  aus einer Menge  $G \subset \mathbb{R}^d$  aus, d.h. man führt eine **Koordinatentransformation**

$$x(\xi) = \begin{bmatrix} x_1(\xi_1, \dots, \xi_d) \\ \vdots \\ x_d(\xi_1, \dots, \xi_d) \end{bmatrix}$$

durch. Diese Transformation kann als Abbildung  $x : G \rightarrow \Omega$  angesehen werden. Mit der Substitutionsregel in Satz 1.3.11 ergibt sich unter dieser Transformation die Beziehung

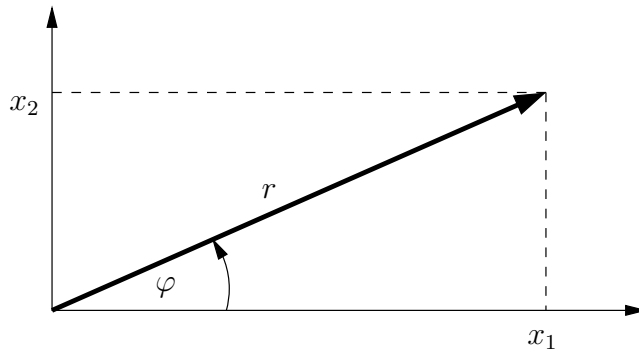
$$\int_{\Omega} f(x) dx = \int_G f(x(\xi)) \cdot |\det(x'(\xi))| d\xi.$$

Wir betrachten häufig verwendete Parametrisierungen.

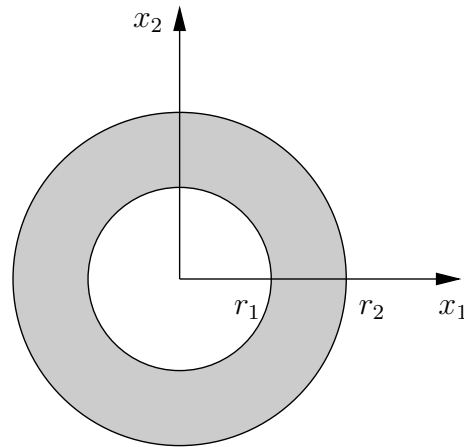
#### Polarkoordinaten

Polarkoordinaten haben wir schon im Zusammenhang mit komplexen Zahlen kennen gelernt. Jeder Punkt  $x = [x_1, x_2]^\top$  im  $\mathbb{R}^2$  ungleich dem Nullvektor kann durch dessen Abstand  $r$  zum Ursprung und den Polarwinkel  $\varphi \in [0, 2\pi)$  zwischen x-Achse und Ortsvektor eindeutig beschrieben werden:

$$\begin{aligned} x_1 &= x_1(r, \varphi) = r \cos(\varphi), \\ x_2 &= x_2(r, \varphi) = r \sin(\varphi). \end{aligned}$$



Mit Hilfe der Polarkoordinaten  $r$  und  $\varphi$  kann man Kreise oder auch Kreise mit Löchern sehr gut beschreiben, während dies mit kartesischen Koordinaten mitunter mehr Aufwand erfordert.

**Beispiel 1.4.8**

Der Ring  $M$  kann mit Polarkoordinaten  $r, \varphi$  sehr einfach dargestellt werden:

$$M = \{[r \cos(\varphi), r \sin(\varphi)]^\top \in \mathbb{R}^2 \mid r_1 \leq r \leq r_2, \varphi \in [0, 2\pi)\}.$$

Die Beschreibung mit Hilfe kartesischer Koordinaten  $x_1, x_2$  lautet:

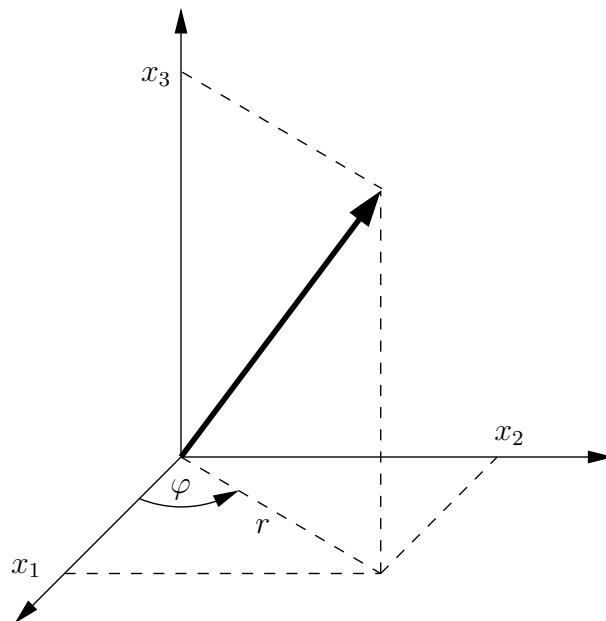
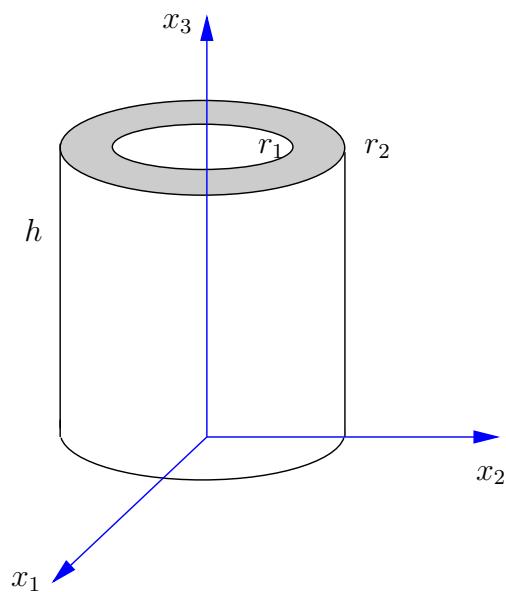
$$M = \{[x_1, x_2]^\top \in \mathbb{R}^2 \mid r_1^2 \leq x_1^2 + x_2^2 \leq r_2^2\}.$$

Schneidet man aus dem Ring ein Kreisstück heraus, so ist die Beschreibung mit Polarkoordinaten immer noch einfach möglich, indem  $\varphi$  eingeschränkt wird. Eine Beschreibung mit kartesischen Koordinaten ist dann jedoch aufwändiger. ■

**Zylinderkoordinaten**

Jeder Punkt  $x = [x_1, x_2, x_3]^\top$  im  $\mathbb{R}^3$  ungleich dem Nullvektor kann eindeutig durch **Zylinderkoordinaten**  $r, \varphi, z$  mit  $\varphi \in [0, 2\pi)$  und  $r \geq 0$  dargestellt werden, indem  $x_1$  und  $x_2$  durch Polarkoordinaten ausgedrückt werden:

$$\begin{aligned} x_1 &= x_1(r, \varphi, z) = r \cos(\varphi), \\ x_2 &= x_2(r, \varphi, z) = r \sin(\varphi), \\ x_3 &= x_3(r, \varphi, z) = z. \end{aligned}$$


**Beispiel 1.4.9**


Der zylindrische Ring  $M$  der Höhe  $h$  kann mit Zylinderkoordinaten  $r$ ,  $\varphi$ ,  $z$  sehr einfach dargestellt werden:

$$M = \{[r \cos(\varphi), r \sin(\varphi), z]^T \in \mathbb{R}^3 \mid r_1 \leq r \leq r_2, \varphi \in [0, 2\pi), 0 \leq z \leq h\}.$$

Die Beschreibung mit Hilfe kartesischer Koordinaten  $x_1$ ,  $x_2$ ,  $x_3$  lautet:

$$M = \{[x_1, x_2, x_3]^T \in \mathbb{R}^3 \mid r_1^2 \leq x_1^2 + x_2^2 \leq r_2^2, 0 \leq x_3 \leq h\}.$$

■

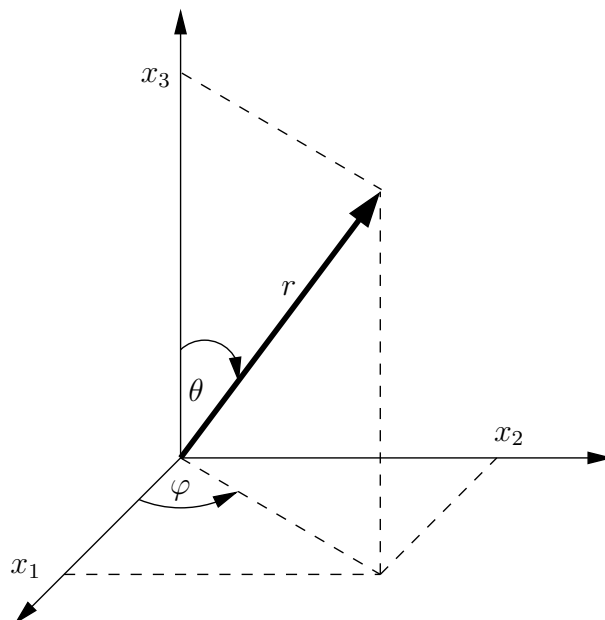
### Kugelkoordinaten

Zur eindeutigen Darstellung eines Punktes  $x = [x_1, x_2, x_3]^\top$  im  $\mathbb{R}^3$  ungleich dem Nullvektor können neben den Zylinderkoordinaten auch **Kugelkoordinaten**  $r, \varphi, \theta$  mit  $r \geq 0$ ,  $\varphi \in [0, 2\pi)$  und  $\theta \in [0, \pi]$  verwendet werden. Der Punkt  $x$  in Kugelkoordinaten lautet

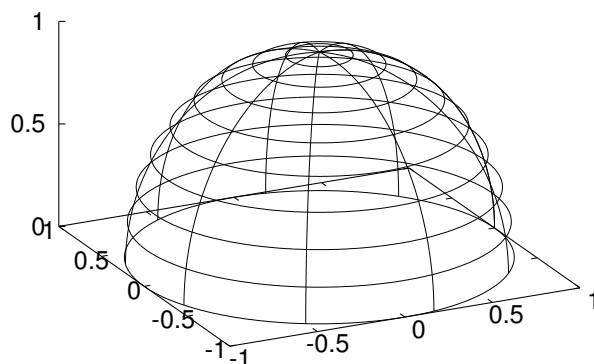
$$x_1 = x_1(r, \varphi, \theta) = r \cos(\varphi) \sin(\theta),$$

$$x_2 = x_2(r, \varphi, \theta) = r \sin(\varphi) \sin(\theta),$$

$$x_3 = x_3(r, \varphi, \theta) = r \cos(\theta).$$



#### Beispiel 1.4.10



Die Halbkugel  $M$  der Höhe 1 kann mit Kugelkoordinaten  $r, \varphi, \theta$  sehr einfach dargestellt werden:

$$M = \left\{ \begin{bmatrix} r \cos(\varphi) \sin(\theta) \\ r \sin(\varphi) \sin(\theta) \\ r \cos(\theta) \end{bmatrix} \in \mathbb{R}^3 \mid 0 \leq r \leq 1, \varphi \in [0, 2\pi), 0 \leq \theta \leq \frac{\pi}{2} \right\}.$$

Die Beschreibung mit Hilfe kartesischer Koordinaten  $x_1, x_2, x_3$  lautet:

$$M = \{[x_1, x_2, x_3]^\top \in \mathbb{R}^3 \mid x_1^2 + x_2^2 + x_3^2 \leq 1, x_3 \geq 0\}.$$

■

### 1.4.3 Flächen- und Volumenintegrale über parametrisierten Flächen und Volumen

Nach den Vorbereitungen zur Parametrisierung von Flächen und Volumen, berechnen wir nun Integrale über parametrisierten Flächen oder Volumen  $\Omega$  gemäß der Substitutionsregel

$$\int_{\Omega} f(x) \, dx = \int_G f(x(\xi)) \cdot |\det(x'(\xi))| \, d\xi,$$

d.h. wir integrieren eine Funktion  $f$  über der Fläche oder dem Volumen  $\Omega$ , die/das durch eine Menge  $G$  und eine Koordinatentransformation  $x = x(\xi)$  beschrieben wird.

#### Bemerkung 1.4.11

Häufig bezeichnet man solche Flächenintegrale auch mit

$$\int_G f \, d\Omega, \quad \iint_G f \, d\Omega \quad \text{oder} \quad \int_G f(x, y) \, dx dy.$$

Entsprechend werden Volumenintegrale mit

$$\int_G f \, d\Omega, \quad \iiint_G f \, d\Omega \quad \text{oder} \quad \int_G f(x, y, z) \, dx dy dz$$

bezeichnet. ■

#### Beispiel 1.4.12

Wir betrachten die Fläche

$$\Omega = \left\{ [x, y]^\top \in \mathbb{R}^2 \mid x = \xi_1 \cos(\xi_2), y = \xi_1 \sin(\xi_2), 0 \leq \xi_1 \leq R, 0 \leq \xi_2 \leq \frac{\pi}{2} \right\}$$

mit gegebenem  $R > 0$  ( $\Omega$  ist ein Viertelkreis mit Radius  $R$  um den Nullpunkt) und Parameterbereich

$$G = \left\{ [\xi_1, \xi_2]^\top \in \mathbb{R}^2 \mid 0 \leq \xi_1 \leq R, 0 \leq \xi_2 \leq \frac{\pi}{2} \right\}.$$

Die Jacobische Funktionaldeterminante lautet

$$D(\xi_1, \xi_2) = \det \begin{bmatrix} \cos(\xi_2) & -\xi_1 \sin(\xi_2) \\ \sin(\xi_2) & \xi_1 \cos(\xi_2) \end{bmatrix} = \xi_1 (\cos^2(\xi_2) + \sin^2(\xi_2)) = \xi_1.$$

Damit berechnet sich das Flächenintegral für die Funktion  $f(x, y) = xy$  zu

$$\begin{aligned} \int_{\Omega} f \, d\mu &= \int_G f \, d\Omega = \int_{\xi_1=0}^{\xi_1=R} \int_{\xi_2=0}^{\xi_2=\frac{\pi}{2}} f(x(\xi_1, \xi_2), y(\xi_1, \xi_2)) |D(\xi_1, \xi_2)| \, d\xi_2 d\xi_1 \\ &= \int_{\xi_1=0}^{\xi_1=R} \int_{\xi_2=0}^{\xi_2=\frac{\pi}{2}} \xi_1^2 \cos(\xi_2) \sin(\xi_2) |\xi_1| \, d\xi_2 d\xi_1 \\ &= \int_{\xi_1=0}^{\xi_1=R} \int_{\xi_2=0}^{\xi_2=\frac{\pi}{2}} \xi_1^3 \cos(\xi_2) \sin(\xi_2) \, d\xi_2 d\xi_1 \\ &= \int_{\xi_1=0}^{\xi_1=R} \left[ \frac{1}{2} \xi_1^3 \sin^2(\xi_2) \right]_{\xi_2=0}^{\xi_2=\frac{\pi}{2}} d\xi_1 \\ &= \int_{\xi_1=0}^{\xi_1=R} \frac{1}{2} \xi_1^3 \, d\xi_1 \\ &= \left[ \frac{1}{8} \xi_1^4 \right]_{\xi_1=0}^{\xi_1=R} = \frac{1}{8} R^4. \end{aligned}$$

■

Bei Volumenintegralen geht man analog vor.

### Beispiel 1.4.13

Wir berechnen das Volumen des Zylinders

$$Z = \left\{ \begin{bmatrix} \xi_1 \cos(\xi_2) \\ \xi_1 \sin(\xi_2) \\ \xi_3 \end{bmatrix} \mid 0 \leq \xi_1 \leq R, 0 \leq \xi_2 < 2\pi, 0 \leq \xi_3 \leq H \right\}$$

mit Radius  $R > 0$  und Höhe  $H > 0$ . Die Jacobische Funktionaldeterminante lautet

$$D(\xi_1, \xi_2, \xi_3) = \det \begin{bmatrix} \cos(\xi_2) & -\xi_1 \sin(\xi_2) & 0 \\ \sin(\xi_2) & \xi_1 \cos(\xi_2) & 0 \\ 0 & 0 & 1 \end{bmatrix} = \xi_1 \cos^2(\xi_2) + \xi_1 \sin^2(\xi_2) = \xi_1.$$

Das Volumen von  $Z$  ist gegeben durch das Volumenintegral

$$\begin{aligned} V &= \int_Z d\mu = \int_G dZ = \int_{\xi_1=0}^{\xi_1=R} \int_{\xi_2=0}^{\xi_2=2\pi} \int_{\xi_3=0}^{\xi_3=H} \xi_1 \, d\xi_3 d\xi_2 d\xi_1 \\ &= \int_{\xi_1=0}^{\xi_1=R} \int_{\xi_2=0}^{\xi_2=2\pi} H \xi_1 \, d\xi_2 d\xi_1 \\ &= \int_{\xi_1=0}^{\xi_1=R} 2\pi H \xi_1 \, d\xi_1 \\ &= \left[ 2\pi H \frac{1}{2} \xi_1^2 \right]_{\xi_1=0}^{\xi_1=R} = \pi H R^2. \end{aligned}$$



## 1.5 Oberflächenintegrale

In diesem Abschnitt betrachten wir Flächen im dreidimensionalen Raum, sogenannte Oberflächenintegrale. Oberflächenintegrale unterscheidet man in Oberflächenintegrale 1. Art, mit deren Hilfe man Oberflächeninhalte oder Schwerpunkte von Flächen im Raum (Schalen) berechnen kann, und Oberflächenintegrale 2. Art, mit deren Hilfe man z.B. ausrechnen kann, welche Flüssigkeitsmenge durch eine gegebene Fläche fließt. Oberflächenintegrale können wir nicht einfach dadurch berechnen, dass wir ein Volumenintegral verwenden, weil Oberflächen eine Nullmenge bzgl. des Lebesgue-Maßes im dreidimensionalen Raum darstellen.

### 1.5.1 Flächenelemente

Eine Kurve im Raum ist ein eindimensionales Objekt, welches durch den Kurvenparameter parametrisiert wird. Eine Fläche im Raum ist hingegen ein zweidimensionales Objekt und lässt sich durch zwei Parameter beschreiben. Wir verwenden im Folgenden  $u$  und  $v$  anstatt  $\xi_1$  und  $\xi_2$ .

#### Definition 1.5.1

Eine **Fläche  $S$  im Raum** lässt sich durch Flächenparameter  $u$  und  $v$  aus dem **Parametergebiet**  $G \subseteq \mathbb{R}^2$  und eine Vektorfunktion

$$x(u, v) = \begin{bmatrix} x_1(u, v) \\ x_2(u, v) \\ x_3(u, v) \end{bmatrix} = \begin{bmatrix} x(u, v) \\ y(u, v) \\ z(u, v) \end{bmatrix}$$

mit  $[u, v]^T \in G$  beschreiben.

#### Beispiel 1.5.2

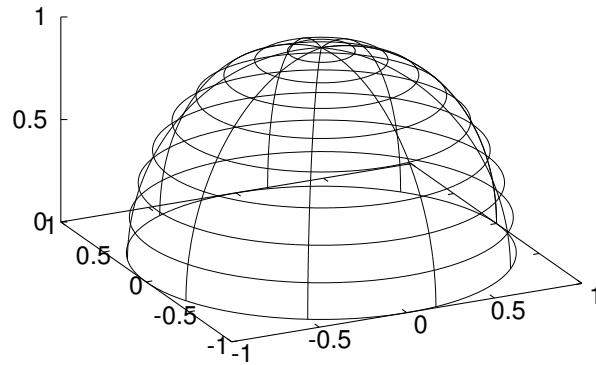
Die Halbkugeloberfläche mit Höhe 1 wird beschrieben durch

$$S = \left\{ \begin{bmatrix} \cos(u) \sin(v) \\ \sin(u) \sin(v) \\ \cos(v) \end{bmatrix} \in \mathbb{R}^3 \mid u \in [0, 2\pi), 0 \leq v \leq \frac{\pi}{2} \right\}.$$

Das Parametergebiet  $G$  lautet hierbei

$$G = \left\{ [u, v]^T \in \mathbb{R}^2 \mid u \in [0, 2\pi), 0 \leq v \leq \frac{\pi}{2} \right\}.$$



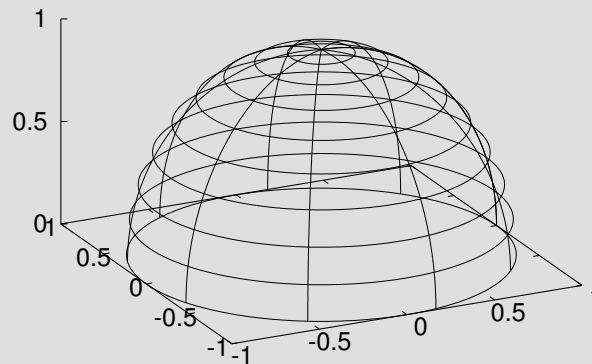


■

### Definition 1.5.3

Gegeben seien das Parametergebiet  $G$  und die Fläche  $S = \{x(u, v) \in \mathbb{R}^3 \mid [u, v]^T \in G\}$ .

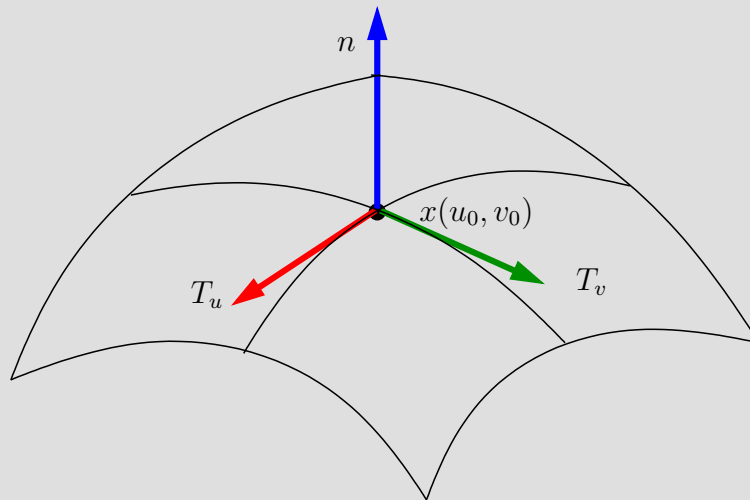
- (a) Wird einer der Parameter  $u$  oder  $v$  festgehalten, so erhält man spezielle Kurven  $x(c, v)$  bzw.  $x(u, c)$  im Raum innerhalb der Fläche  $S$ , wobei  $c$  eine Konstante sei. Diese Kurven heißen **Koordinatenlinien** oder **Parameterlinien**, vgl. Abbildung.



- (b) Die Vektoren

$$T_u(u_0, v_0) := \frac{\partial x}{\partial u}(u_0, v_0) \quad \text{und} \quad T_v(u_0, v_0) := \frac{\partial x}{\partial v}(u_0, v_0)$$

sind (nicht normierte) Tangentenvektoren an die Koordinatenlinien im Punkt  $x(u_0, v_0)$ , vgl. Abbildung.



(c) Der normierte Vektor

$$n(u_0, v_0) = \frac{T_u(u_0, v_0) \times T_v(u_0, v_0)}{\|T_u(u_0, v_0) \times T_v(u_0, v_0)\|}$$

heißt **Normalenvektor der Fläche  $S$  in  $(u_0, v_0)$** . Er steht senkrecht auf  $T_u(u_0, v_0)$  und  $T_v(u_0, v_0)$ . Das Vektorentupel  $(T_u, T_v, n)$  bildet ein Rechtssystem und heißt **begleitendes Dreibein der Fläche  $S$** .

(d) Die Parameterdarstellung der Fläche  $S$  heißt **regulär**, wenn die Abbildung  $x$  stetig differenzierbar ist und die Tangentenvektoren  $T_u$  und  $T_v$  linear unabhängig sind, also  $T_u(u, v) \times T_v(u, v) \neq 0$  für alle  $[u, v]^T \in G$  gilt.

#### Bemerkung 1.5.4

- Das Kreuzprodukt zweier Vektoren  $a, b \in \mathbb{R}^3$  ist definiert als  $a \times b = [a_2b_3 - a_3b_2, a_3b_1 - a_1b_3, a_1b_2 - a_2b_1]^T$ .
- Mit Hilfe der Tangentenvektoren bzw. des Normalenvektors kann die Tangentialebene  $T$  der Fläche  $S$  im Punkt  $x(u_0, v_0)$  in Parameterform dargestellt werden als

$$T = \{x \in \mathbb{R}^3 \mid x = x(u_0, v_0) + \lambda T_u(u_0, v_0) + \mu T_v(u_0, v_0), \lambda, \mu \in \mathbb{R}\}$$

bzw. in Normalenform als

$$T = \{x \in \mathbb{R}^3 \mid \langle n(u_0, v_0), x - x(u_0, v_0) \rangle = 0\}.$$

**Beispiel 1.5.5**

Kugeloberfläche der Kugel mit Radius 1 wird beschrieben durch

$$S = \left\{ \begin{bmatrix} \cos(u) \sin(v) \\ \sin(u) \sin(v) \\ \cos(v) \end{bmatrix} \in \mathbb{R}^3 \mid u \in [0, 2\pi), 0 \leq v \leq \pi \right\}.$$

Das Parametergebiet  $G$  lautet hierbei

$$G = \{[u, v]^T \in \mathbb{R}^2 \mid u \in [0, 2\pi), 0 \leq v \leq \pi\}.$$

Die Tangentenvektoren lauten

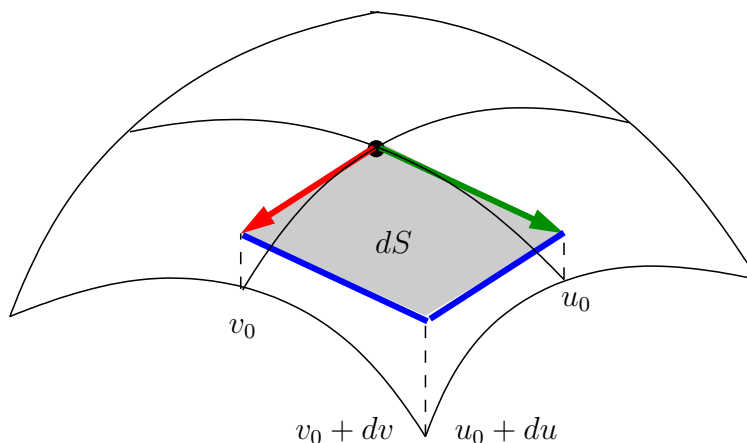
$$T_u(u, v) = \begin{bmatrix} -\sin(u) \sin(v) \\ \cos(u) \sin(v) \\ 0 \end{bmatrix}, \quad T_v(u, v) = \begin{bmatrix} \cos(u) \cos(v) \\ \sin(u) \cos(v) \\ -\sin(v) \end{bmatrix}$$

und der Normalenvektor lautet

$$n(u, v) = \frac{T_u(u, v) \times T_v(u, v)}{\|T_u(u, v) \times T_v(u, v)\|} = \begin{bmatrix} -\cos(u) \sin(v) \\ -\sin(u) \sin(v) \\ -\cos(v) \end{bmatrix}.$$

Die Fläche  $dS$ , die durch zwei infinitesimal benachbarte Koordinatenlinien zu den Koordinaten  $u_0$  und  $v_0$  bzw.  $u_0 + du$  und  $v_0 + dv$  begrenzt wird, ergibt sich ungefähr als Fläche des durch  $T_u(u_0, v_0)du$  und  $T_v(u_0, v_0)dv$  aufgespannten Parallelogramms:

$$dS = \|T_u(u_0, v_0)du \times T_v(u_0, v_0)dv\| = \left\| \frac{\partial x}{\partial u}(u_0, v_0) \times \frac{\partial x}{\partial v}(u_0, v_0) \right\| dudv.$$



**Definition 1.5.6** (skalares Flächenelement, vektorielles Flächenelement)

Die Größe

$$dS := \left\| \frac{\partial x}{\partial u} \times \frac{\partial x}{\partial v} \right\| dudv$$

heißt **skalares (Ober-)flächenelement**.

Die Größe

$$d\vec{S} := n dS = \left( \frac{\partial x}{\partial u} \times \frac{\partial x}{\partial v} \right) dudv$$

heißt **vektorielles (Ober-)flächenelement**.

**Beispiel 1.5.7**

Wir betrachten wieder die Kugeloberfläche der Kugel mit Radius 1:

$$S = \left\{ x(u, v) = \begin{bmatrix} \cos(u) \sin(v) \\ \sin(u) \sin(v) \\ \cos(v) \end{bmatrix} \in \mathbb{R}^3 \mid u \in [0, 2\pi), 0 \leq v \leq \pi \right\}.$$

Die Tangentenvektoren lauten

$$T_u(u, v) = \frac{\partial x}{\partial u}(u, v) = \begin{bmatrix} -\sin(u) \sin(v) \\ \cos(u) \sin(v) \\ 0 \end{bmatrix}, \quad T_v(u, v) = \frac{\partial x}{\partial v}(u, v) = \begin{bmatrix} \cos(u) \cos(v) \\ \sin(u) \cos(v) \\ -\sin(v) \end{bmatrix}.$$

Daraus folgt

$$\frac{\partial x}{\partial u}(u, v) \times \frac{\partial x}{\partial v}(u, v) = \begin{bmatrix} -\cos(u) \sin^2(v) \\ -\sin(u) \sin^2(v) \\ -\cos(v) \sin(v) \end{bmatrix}$$

und

$$\begin{aligned} \left\| \frac{\partial x}{\partial u}(u, v) \times \frac{\partial x}{\partial v}(u, v) \right\|^2 &= \cos^2(u) \sin^4(v) + \sin^2(u) \sin^4(v) + \cos^2(v) \sin^2(v) \\ &= \sin^4(v) (\cos^2(u) + \sin^2(u)) + \cos^2(v) \sin^2(v) \\ &= \sin^2(v) (\sin^2(v) + \cos^2(v)) \\ &= \sin^2(v). \end{aligned}$$

Das skalare Flächenelement lautet damit

$$dS = \left\| \frac{\partial x}{\partial u}(u, v) \times \frac{\partial x}{\partial v}(u, v) \right\| dudv = \sin(v) dudv.$$

**Beispiel 1.5.8**

Wir berechnen das Flächenelement der durch den Graphen einer Funktion  $f : \mathbb{R}^2 \rightarrow \mathbb{R}$  beschriebenen Fläche. In Parameterform besteht die Fläche aus den Vektoren

$$x(u, v) = \begin{bmatrix} u \\ v \\ f(u, v) \end{bmatrix}, \quad u, v \in \mathbb{R}.$$

Es gilt

$$\begin{aligned} T_u &= \frac{\partial x}{\partial u} = \begin{bmatrix} 1 \\ 0 \\ \frac{\partial f}{\partial u} \end{bmatrix}, \\ T_v &= \frac{\partial x}{\partial v} = \begin{bmatrix} 0 \\ 1 \\ \frac{\partial f}{\partial v} \end{bmatrix}, \\ \frac{\partial x}{\partial u} \times \frac{\partial x}{\partial v} &= \begin{bmatrix} -\frac{\partial f}{\partial u} \\ -\frac{\partial f}{\partial v} \\ 1 \end{bmatrix}. \end{aligned}$$

Damit berechnet sich das skalare Flächenelement zu

$$dS = \sqrt{1 + \left(\frac{\partial f}{\partial u}\right)^2 + \left(\frac{\partial f}{\partial v}\right)^2} \, dudv$$

und das vektorielle Flächenelement zu

$$d\vec{S} = \begin{bmatrix} -\frac{\partial f}{\partial u} \\ -\frac{\partial f}{\partial v} \\ 1 \end{bmatrix} \, dudv.$$

Im Spezialfall einer ebenen Fläche mit  $f(u, v) = \text{const}$  gilt  $dS = dudv$ . ■

**Beispiel 1.5.9**

Wir berechnen das Flächenelement einer ebenen Fläche, die durch die folgende Darstellung gegeben sei:

$$x(u, v) = \begin{bmatrix} x_1(u, v) \\ x_2(u, v) \\ 0 \end{bmatrix}, \quad u, v \in \mathbb{R}.$$

Es gilt

$$T_u = \frac{\partial x}{\partial u} = \begin{bmatrix} \frac{\partial x_1}{\partial u} \\ \frac{\partial x_2}{\partial u} \\ 0 \end{bmatrix},$$

$$T_v = \frac{\partial x}{\partial v} = \begin{bmatrix} \frac{\partial x_1}{\partial v} \\ \frac{\partial x_2}{\partial v} \\ 0 \end{bmatrix},$$

$$\frac{\partial x}{\partial u} \times \frac{\partial x}{\partial v} = \begin{bmatrix} 0 \\ 0 \\ \frac{\partial x_1}{\partial u} \frac{\partial x_2}{\partial v} - \frac{\partial x_2}{\partial u} \frac{\partial x_1}{\partial v} \end{bmatrix}.$$

Die dritte Komponente des Kreuzprodukts ist gerade die Determinante der Jacobimatrix der vektorwertigen Abbildung  $[u, v]^T \mapsto [x_1(u, v), x_2(u, v)]^T$ , d.h.

$$D = \det \begin{bmatrix} \frac{\partial x_1}{\partial u} & \frac{\partial x_1}{\partial v} \\ \frac{\partial x_2}{\partial u} & \frac{\partial x_2}{\partial v} \end{bmatrix} = \frac{\partial x_1}{\partial u} \frac{\partial x_2}{\partial v} - \frac{\partial x_2}{\partial u} \frac{\partial x_1}{\partial v}.$$

$D$  heißt **Jacobische Funktionaldeterminante**. Damit berechnet sich das skalare Flächenelement zu

$$dS = |D| du dv = \left| \det \begin{bmatrix} \frac{\partial x_1}{\partial u} & \frac{\partial x_1}{\partial v} \\ \frac{\partial x_2}{\partial u} & \frac{\partial x_2}{\partial v} \end{bmatrix} \right| du dv. \quad \blacksquare$$

### 1.5.2 Oberflächenintegrale 1. Art

Wir widmen uns der Konstruktion eines Oberflächenintegrals 1. Art und nutzen die anschauliche Konstruktion im Sinne eines Riemann-Integrals.

#### **Definition 1.5.10** (Oberflächenintegral 1. Art)

Gegeben sei eine Fläche  $S \subseteq \mathbb{R}^3$  im dreidimensionalen Raum und eine Funktion  $f : S \rightarrow \mathbb{R}$ , die auf der Fläche  $S$  definiert ist.

- (a) Die Fläche  $S$  sei in beliebige Teilflächen  $S_1, \dots, S_N$  mit Flächeninhalten  $dS_1, \dots, dS_N$  unterteilt.
- (b) Für beliebig gewählte Punkte  $x_k$  aus  $S_k$  oder dem Rand von  $S_k$ ,  $k = 1, \dots, N$ , sei

$$I_N := \sum_{k=1}^N f(x_k) dS_k.$$

Falls die Folge  $\{I_N\}_{N \in \mathbb{N}}$  unabhängig von der Wahl der Zerlegungen und der Wahl der Punkte  $x_k$  gegen den Wert  $I$  konvergiert, wobei der maximale Durchmesser der Teilflächen mit  $N \rightarrow \infty$  gegen Null streben soll, so heißt

$$I := \lim_{N \rightarrow \infty} I_N$$

das **Oberflächenintegral 1. Art von  $f$  über  $S$**  und man schreibt

$$I = \int_S f(x) \, dS \quad \text{oder} \quad \iint_S f \, dS.$$

Die Definition des Oberflächenintegrals 1. Art erfolgt vollkommen analog zur Definition eines Flächenintegrals. Der einzige Unterschied ist, dass man beim Oberflächenintegral 1. Art über Funktionen  $f$  integriert, die im dreidimensionalen Raum definiert sind und nicht nur in der Ebene wie beim Flächenintegral. Tatsächlich besteht der folgende Zusammenhang:

**Satz 1.5.11** (Berechnung des Oberflächenintegrals 1. Art bei gegebener Parameterdarstellung der Fläche)

Die Fläche  $S$  sei parametrisiert gemäß

$$S = \left\{ x(u, v) = \begin{bmatrix} x(u, v) \\ y(u, v) \\ z(u, v) \end{bmatrix} \in \mathbb{R}^3 \mid [u, v]^T \in G \right\}$$

mit Parameterbereich  $G \subseteq \mathbb{R}^2$  und stetig differenzierbaren Funktionen  $x, y, z$ . Dann kann das Oberflächenintegral 1. Art von  $f$  über  $S$  mit Hilfe eines Flächenintegrals über  $G$  berechnet werden gemäß

$$\begin{aligned} \int_S f(x) \, dS &= \int_G f(x(u, v)) \left\| \frac{\partial x}{\partial u} \times \frac{\partial x}{\partial v} \right\| \, dudv \\ &= \int_G f(x(u, v)) \sqrt{\left\| \frac{\partial x}{\partial u} \right\|^2 \left\| \frac{\partial x}{\partial v} \right\|^2 - \left\langle \frac{\partial x}{\partial u}, \frac{\partial x}{\partial v} \right\rangle^2} \, dudv. \end{aligned}$$

**Beweis:** Berechne das skalare Flächenelement und verwende die Darstellung

$$\|a \times b\| = \|a\| \cdot \|b\| \cdot |\sin \gamma| = \|a\| \cdot \|b\| \cdot \sqrt{1 - \cos^2 \gamma}$$

für das Kreuzprodukt:

$$\begin{aligned}
 dS &= \left\| \frac{\partial x}{\partial u} \times \frac{\partial x}{\partial v} \right\| dudv \\
 &= \left\| \frac{\partial x}{\partial u} \right\| \cdot \left\| \frac{\partial x}{\partial v} \right\| \cdot \sqrt{1 - \cos^2 \gamma} dudv \\
 &= \sqrt{\left\| \frac{\partial x}{\partial u} \right\|^2 \cdot \left\| \frac{\partial x}{\partial v} \right\|^2 - \left( \left\| \frac{\partial x}{\partial u} \right\| \cdot \left\| \frac{\partial x}{\partial v} \right\| \cdot \cos \gamma \right)^2} dudv \\
 &= \sqrt{\left\| \frac{\partial x}{\partial u} \right\|^2 \cdot \left\| \frac{\partial x}{\partial v} \right\|^2 - \left\langle \frac{\partial x}{\partial u}, \frac{\partial x}{\partial v} \right\rangle^2} dudv.
 \end{aligned}$$

■

Wird die Fläche  $S$  durch den Graphen einer Funktion  $g : \mathbb{R}^2 \rightarrow \mathbb{R}$  wie in Beispiel 1.5.8 beschrieben, d.h.

$$S = \left\{ \left[ \begin{array}{c} u \\ v \\ g(u, v) \end{array} \right] \in \mathbb{R}^3 \mid [u, v]^\top \in G \right\}, \quad G \subseteq \mathbb{R}^2,$$

so berechnet sich das skalare Flächenelement zu

$$dS = \sqrt{1 + \left( \frac{\partial g}{\partial u} \right)^2 + \left( \frac{\partial g}{\partial v} \right)^2} dudv.$$

In diesem Fall berechnet sich das Oberflächenintegral 1. Art also wie folgt:

**Satz 1.5.12** (Berechnung des Oberflächenintegrals 1. Art bei expliziter Darstellung der Fläche)

Die Fläche  $S$  sei parametrisiert gemäß

$$S = \left\{ x(u, v) = \left[ \begin{array}{c} u \\ v \\ g(u, v) \end{array} \right] \in \mathbb{R}^3 \mid [u, v]^\top \in G \right\}$$

mit Parameterbereich  $G \subseteq \mathbb{R}^2$  und stetig differenzierbarer Funktion  $g$ . Dann lautet das Oberflächenintegral 1. Art von  $f$  über  $S$

$$\int_S f(x) dS = \int_G f(u, v, g(u, v)) \sqrt{1 + \left( \frac{\partial g}{\partial u} \right)^2 + \left( \frac{\partial g}{\partial v} \right)^2} dudv.$$



**Beispiel 1.5.13**

Wir berechnen die Oberfläche der Kugel mit Radius  $R > 0$

$$K = \left\{ \left[ \begin{array}{c} R \cos(u) \sin(v) \\ R \sin(u) \sin(v) \\ R \cos(v) \end{array} \right] \mid 0 \leq u \leq 2\pi, 0 \leq v \leq \pi \right\}.$$

Wie in Beispiel 1.5.7 berechnen wir das skalare Flächenelement  $dS$  und erhalten

$$dS = R^2 \sin(v) \, du \, dv.$$

Alternativ kann man die Formel aus Satz 1.5.11 verwenden.

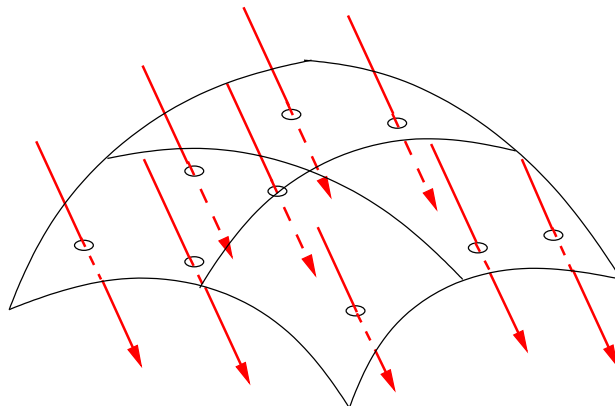
Damit ergibt sich der Flächeninhalt der Kugeloberfläche zu

$$\begin{aligned} I &= \int_S 1 \, dS = \int_{v=0}^{v=\pi} \int_{u=0}^{u=2\pi} R^2 \sin(v) \, du \, dv \\ &= \int_{v=0}^{v=\pi} 2\pi R^2 \sin(v) \, dv \\ &= [-2\pi R^2 \cos(v)]_0^\pi \\ &= 4\pi R^2. \end{aligned}$$

■

**1.5.3 Oberflächenintegrale 2. Art****Beispiel 1.5.14**

Strömung durch eine Fläche:



Welche Flüssigkeitsmenge fließt in einer infinitesimal kleinen Zeiteinheit  $dt$  durch die Fläche  $S$ ?

- (a) Dazu zerlegen wir die Fläche  $S$  wieder in Teilflächen  $S_k$ ,  $k = 1, \dots, N$ , mit Flächeninhalten  $dS_k$ ,  $k = 1, \dots, N$ .

- (b) Die Flüssigkeit legt in  $dt$  den Weg  $ds = \|v\|dt$  zurück, wobei  $v$  den Geschwindigkeitsvektor der Flüssigkeit bezeichnet.
- (c) Sei nun  $x_k \in S_k$  ein beliebiger Punkt und  $n(x_k)$  der Normalenvektor mit Länge 1 an die Fläche  $S$  im Punkt  $x_k$ . Bezeichnet  $\alpha_k$  den Winkel zwischen den Vektoren  $v(x_k)$  und  $n(x_k)$ , so ist

$$ds_k \cos(\alpha_k) = \|v(x_k)\| \cos(\alpha_k) dt$$

gerade der Anteil des zurückgelegten Wegs der Flüssigkeit in Richtung des Normalenvektors und in der Zeiteinheit  $dt$  strömt durch das Flächenstück  $S_k$  das Volumen

$$\begin{aligned} dV_k &= dS_k ds_k \cos(\alpha_k) \\ &= dS_k \|v(x_k)\| \cos(\alpha_k) dt \\ &= dS_k \|v(x_k)\| \|n(x_k)\| \cos(\alpha_k) dt \\ &= dS_k \langle v(x_k), n(x_k) \rangle dt. \end{aligned}$$

Anschaulich: Das maximale Volumen wird erreicht, wenn  $v(x_k)$  und  $n(x_k)$  parallel verlaufen. Stehen beide Vektoren senkrecht aufeinander, so fließt nichts durch das Flächenstück.

Mit Hilfe des vektoriellen Flächenelements  $d\vec{S}_k = n(x_k) dS_k$  gilt für den **Fluss**  $d\Phi := \frac{dV}{dt}$  die Beziehung

$$d\Phi_k = \langle v(x_k), d\vec{S}_k \rangle.$$

Der Fluss wird in der Einheit Flüssigkeitsmenge pro Zeiteinheit gemessen.

Den Gesamtfluss durch die Fläche  $S$  erhält man durch Summation über die Teilflächen:

$$\sum_{k=1}^N d\Phi_k = \sum_{k=1}^N \langle v(x_k), d\vec{S}_k \rangle.$$

- (d) Grenzübergang  $N \rightarrow \infty$ , wobei der Durchmesser der Flächenstücke gegen Null streben soll, liefert dann den Gesamtfluss

$$\Phi = \int_S \langle v, d\vec{S} \rangle.$$

■

### Definition 1.5.15 (Oberflächenintegral 2. Art)

Das in Beispiel 1.5.14 konstruierte Integral

$$\int_S \langle v, d\vec{S} \rangle$$

heißt **Oberflächenintegral 2. Art**, falls es unabhängig von der Wahl der Flächenelemente  $S_k$  und der Punkte  $x_k$  ist.

Analog zum Kurvenintegral 2. Art wird beim Oberflächenintegral 2. Art das Skalarprodukt zwischen Vektorfeld und vektorielltem Flächenelement gebildet. Das Vorzeichen des Oberflächenintegrals 2. Art hängt von der Orientierung des Normalenvektors  $n$  ab.

Wegen  $d\vec{S} = n dS$  besteht folgender Zusammenhang:

**Satz 1.5.16** (Zusammenhang zwischen Oberflächenintegralen 1. und 2. Art)

Das Oberflächenintegral 2. Art kann mit Hilfe des Oberflächenintegrals 1. Art berechnet werden gemäß

$$\int_S \langle v, d\vec{S} \rangle = \int_S \langle v, n \rangle dS.$$

Wegen

$$n = \frac{\frac{\partial x}{\partial u} \times \frac{\partial x}{\partial v}}{\left\| \frac{\partial x}{\partial u} \times \frac{\partial x}{\partial v} \right\|}$$

und

$$dS = \left\| \frac{\partial x}{\partial u} \times \frac{\partial x}{\partial v} \right\| dudv$$

folgt

$$d\vec{S} = n dS = \left( \frac{\partial x}{\partial u} \times \frac{\partial x}{\partial v} \right) dudv$$

für parametrisierte Flächen und es gilt

**Satz 1.5.17** (Berechnung des Oberflächenintegrals 2. Art)

Sei  $f : \mathbb{R}^3 \rightarrow \mathbb{R}^3$  ein Vektorfeld und

$$S = \left\{ x(u, v) = \begin{bmatrix} x(u, v) \\ y(u, v) \\ z(u, v) \end{bmatrix} \in \mathbb{R}^3 \mid [u, v]^T \in G \right\}$$

eine Fläche mit Parameterbereich  $G \subseteq \mathbb{R}^2$ . Dann lässt sich das Oberflächenintegral 2. Art mit Hilfe eines Flächenintegrals berechnen gemäß

$$\int_S \langle f, d\vec{S} \rangle = \int_G \left\langle f(x(u, v)), \frac{\partial x}{\partial u} \times \frac{\partial x}{\partial v} \right\rangle dudv.$$

**Beispiel 1.5.18**

Gegeben sei das Vektorfeld

$$f(x, y, z) = \begin{bmatrix} 2y \\ x \\ z \end{bmatrix}$$

und die Fläche

$$S = \left\{ x(u, v) = \begin{bmatrix} x(u, v) \\ y(u, v) \\ z(u, v) \end{bmatrix} = \begin{bmatrix} 2 \cos(u) \\ 2 \sin(u) \\ v \end{bmatrix} \in \mathbb{R}^3 \mid [u, v]^T \in G \right\}$$

mit Parameterbereich  $G = \{[u, v]^T \in \mathbb{R}^2 \mid 0 \leq u \leq \frac{\pi}{2}, 0 \leq v \leq 3\}$ . Die Fläche beschreibt die Oberfläche eines Viertelzylinders mit Radius 2 und Höhe 3.

Das vektorielle Flächenelement lautet

$$d\vec{S} = \left( \frac{\partial x}{\partial u} \times \frac{\partial x}{\partial v} \right) dudv = \left( \begin{bmatrix} -2 \sin(u) \\ 2 \cos(u) \\ 0 \end{bmatrix} \times \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \right) dudv = \begin{bmatrix} 2 \cos(u) \\ 2 \sin(u) \\ 0 \end{bmatrix} dudv.$$

Mit

$$\left\langle f(x(u, v)), \frac{\partial x}{\partial u} \times \frac{\partial x}{\partial v} \right\rangle = \left\langle \begin{bmatrix} 4 \sin(u) \\ 2 \cos(u) \\ v \end{bmatrix}, \begin{bmatrix} 2 \cos(u) \\ 2 \sin(u) \\ 0 \end{bmatrix} \right\rangle = 8 \sin(u) \cos(u) + 4 \cos(u) \sin(u)$$

berechnet sich das Oberflächenintegral 2. Art zu

$$\begin{aligned} & \int_S \langle f, d\vec{S} \rangle \\ &= \int_{v=0}^{v=3} \int_{u=0}^{u=\frac{\pi}{2}} 12 \sin(u) \cos(u) \, dv du = \int_{v=0}^{v=3} [6 \sin^2(u)]_{v=0}^{v=\frac{\pi}{2}} \, dv = \int_{v=0}^{v=3} 6 \, dv = 18. \end{aligned}$$

■

## 1.6 Integralsätze

Für gewöhnliche Integrale zweier Funktionen  $u, v : \mathbb{R} \rightarrow \mathbb{R}$  über dem Intervall  $[a, b]$  gelten der **Hauptsatz der Differential- und Integralrechnung**

$$u(b) - u(a) = \int_a^b u'(x) dx$$

und die **partielle Integrationsformel**

$$\int_a^b u'(x)v(x) dx = [u(x)v(x)]_a^b - \int_a^b u(x)v'(x) dx.$$

Gelten entsprechende Formeln auch für Flächen- oder Volumenintegrale?

Die folgenden Betrachtungen gelten für Flächen- und Volumenintegrale und wir bezeichnen mit  $\Omega \in \mathbb{R}^d$  mit  $d = 2$  oder  $d = 3$  ein zwei- bzw. dreidimensionales Gebiet, dessen **Rand** mit  $\Gamma = \partial\Omega$  bezeichnet wird. Die **äußere Normale** auf dem Rand, also derjenige Vektor, der in einem Randpunkt normal auf  $\Omega$  steht und nach außen zeigt, wird mit

$$n = \begin{bmatrix} n_1 \\ \vdots \\ n_d \end{bmatrix}, \quad \|n\| = 1,$$

bezeichnet. Mit  $\int_{\Omega} f \, d\Omega$  ( $= \int_{\Omega} f \, d\mu = \int_{\Omega} f \, dx$ ) wird das übliche Flächen- bzw. Volumenintegral bezeichnet.  $dS$  bezeichnet im Fall  $d = 2$  das skalare Bogenelement und im Fall  $d = 3$  das skalare Oberflächenelement und  $\vec{dS}$  das entsprechende vektorielle Element in Normalenrichtung.

**Satz 1.6.1 (Fundamentalformel)**

Sei  $u : \mathbb{R}^d \rightarrow \mathbb{R}$ ,  $d = 2$  oder  $d = 3$ , stetig partiell differenzierbar. Dann gilt für  $i = 1, \dots, d$

$$\int_{\Omega} \frac{\partial u}{\partial x_i} \, d\Omega = \int_{\Gamma} u \, n_i \, dS.$$

(Im Fall  $d = 1$  entspricht dies gerade dem Hauptsatz der Differential- und Integralrechnung!)

**Beweis:** An Stelle eines formalen Beweises illustrieren wir den Satz für den Spezialfall des quadratischen Gebiets  $\Omega = (0, 1) \times (0, 1)$  im  $\mathbb{R}^2$ . Der Rand des Gebiets setzt sich aus den 4 Kanten des Quadrats und den zugehörigen äußeren Normalenvektoren zusammen:

$$\begin{aligned} \Gamma_1 &= \{[x_1, 0]^T \in \mathbb{R}^2 \mid 0 \leq x_1 \leq 1\}, & n_1 &= \begin{bmatrix} 0 \\ -1 \end{bmatrix}, \\ \Gamma_2 &= \{[1, x_2]^T \in \mathbb{R}^2 \mid 0 \leq x_2 \leq 1\}, & n_2 &= \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \\ \Gamma_3 &= \{[x_1, 1]^T \in \mathbb{R}^2 \mid 0 \leq x_1 \leq 1\}, & n_3 &= \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \\ \Gamma_4 &= \{[0, x_2]^T \in \mathbb{R}^2 \mid 0 \leq x_2 \leq 1\}, & n_4 &= \begin{bmatrix} -1 \\ 0 \end{bmatrix}. \end{aligned}$$

Damit ergibt sich

$$\begin{aligned} \int_{\Omega} \frac{\partial u}{\partial x_1} d\Omega &= \int_{x_2=0}^{x_2=1} \int_{x_1=0}^{x_1=1} \frac{\partial u}{\partial x_1}(x_1, x_2) dx_1 dx_2 = \int_{x_2=0}^{x_2=1} u(1, x_2) - u(0, x_2) dx_2 \\ &= \int_{\Gamma_1} u \cdot 0 dS + \int_{\Gamma_2} u \cdot 1 dS + \int_{\Gamma_3} u \cdot 0 dS + \int_{\Gamma_4} u \cdot (-1) dS \\ &= \int_{\Gamma} u n_1 dS, \end{aligned}$$

wobei  $n_1$  jeweils die erste Komponente der äußeren Normalenvektoren bezeichnet. ■

### Beispiel 1.6.2

Wählt man in der Fundamentalformel speziell  $u(x_1, \dots, x_d) = x_i$  mit  $i \in \{1, \dots, d\}$ , so folgt

$$\int_{\Omega} d\Omega = \int_{\Gamma} x_i n_i dS.$$

Man kann den Flächeninhalt ( $d = 2$ ) bzw. das Volumen ( $d = 3$ ) also durch ein Kurven- bzw. Oberflächenintegral über dem Rand ausdrücken. Dies ist das Arbeitsprinzip eines Planimeters, vgl. <http://de.wikipedia.org/wiki/Planimeter>. ■

Setzt man  $u = vw$  in der Fundamentalformel, so erhält man die folgende Aussage:

### Satz 1.6.3 (Partielle Integration)

Für stetig partiell differenzierbare Funktionen  $v, w : \mathbb{R}^d \rightarrow \mathbb{R}$  mit  $d = 2$  oder  $d = 3$  gilt

$$\int_{\Omega} \frac{\partial v}{\partial x_i} w d\Omega = \int_{\Gamma} vw n_i dS - \int_{\Omega} v \frac{\partial w}{\partial x_i} d\Omega.$$

(Im Fall  $d = 1$  entspricht dies gerade der üblichen partiellen Integration!)

### 1.6.1 Der Gauß'sche Integralsatz

Wir betrachten die Strömung einer Flüssigkeit mit Geschwindigkeit  $u$ . Die Divergenz<sup>1</sup> der Geschwindigkeit ist ein Maß für das Auseinanderfließen der Strömung und

$$\int_{\Omega} \operatorname{div}(u) d\Omega$$

ist dann die pro Zeiteinheit in  $\Omega$  erzeugte/zugeflossene oder vernichtete/abgeflossene Flüssigkeitsmenge. Anschaulich ist klar, dass diese Flüssigkeitsmenge gerade die pro Zeiteinheit durch den Rand von  $\Omega$  herein- bzw. herausfließende Flüssigkeitsmenge sein muss,

<sup>1</sup>Die Divergenz einer Abbildung  $u : \mathbb{R}^d \rightarrow \mathbb{R}^d$  ist definiert als  $\operatorname{div}(u) = \frac{\partial u_1}{\partial x_1} + \dots + \frac{\partial u_d}{\partial x_d}$ .

welche durch das Oberflächenintegral

$$\int_{\Gamma} \langle u, n \rangle dS$$

gegeben ist. Dies ist die Aussage des Gauß'schen Integralsatzes:

**Satz 1.6.4 (Gauß'scher Integralsatz)**

Sei  $u : \Omega \subseteq \mathbb{R}^d \rightarrow \mathbb{R}^d$  ein differenzierbares Vektorfeld. Dann gilt

$$\int_{\Omega} \operatorname{div}(u) d\Omega = \int_{\Gamma} \langle u, n \rangle dS = \int_{\Gamma} \langle u, d\vec{S} \rangle.$$

**Beweis:** Die Fundamentalformel liefert

$$\int_{\Omega} \frac{\partial u_i}{\partial x_i} d\Omega = \int_{\Gamma} u_i n_i dS, \quad i = 1, \dots, d.$$

Summation über  $i = 1, \dots, d$  liefert

$$\int_{\Omega} \operatorname{div}(u) d\Omega = \int_{\Gamma} \langle u, n \rangle dS.$$

■

Mitunter findet man den Gauß'schen Integralsatz auch in einer anderen Form, wobei die Rotation<sup>2</sup> verwendet wird.

**Satz 1.6.5 (Gauß'scher Integralsatz, Rotationsvariante)**

Sei  $u : \Omega \subseteq \mathbb{R}^d \rightarrow \mathbb{R}^d$  mit  $d = 2, 3$  ein differenzierbares Vektorfeld. Dann gilt die Vektorgleichung

$$\int_{\Omega} \operatorname{rot}(u) d\Omega = \int_{\Gamma} n \times u dS.$$

**Beweis:** Nachrechnen mit Hilfe der Fundamentalformel. Im Fall  $d = 2$  wird dabei die Definition der Rotation  $\operatorname{rot} u = \frac{\partial u_2}{\partial x_1} - \frac{\partial u_1}{\partial x_2}$  verwendet und  $n \times u$  als  $u_2 n_1 - u_1 n_2$  definiert.

■

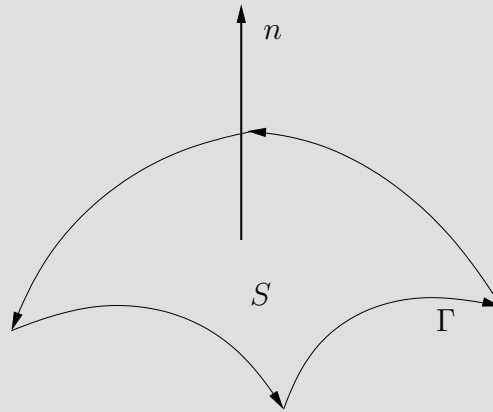
<sup>2</sup>Die Rotation einer Abbildung  $u : \mathbb{R}^d \rightarrow \mathbb{R}^d$  ist definiert als  $\operatorname{rot}(u) = \nabla \times u = \left[ \frac{\partial u_3}{\partial x_2} - \frac{\partial u_2}{\partial x_3}, \frac{\partial u_1}{\partial x_3} - \frac{\partial u_3}{\partial x_1}, \frac{\partial u_2}{\partial x_1} - \frac{\partial u_1}{\partial x_2} \right]^T$ .

### 1.6.2 Der Integralsatz von Stokes

Der Integralsatz von Stokes setzt Oberflächenintegrale in Relation zu Randintegralen. Anschaulich besagt der Satz, dass der Wirbelfluss eines Vektorfeldes  $u$  (Fluss von  $\text{rot}(u)$ ) durch die Fläche  $S$  gleich der Zirkulation von  $u$  entlang des Randes  $\Gamma$  der Fläche  $S$  ist.

#### Satz 1.6.6 (Integralsatz von Stokes)

Sei  $S \subseteq \mathbb{R}^3$  eine Fläche, die von einer geschlossenenen Kurve  $\Gamma$  berandet ist. Die Durchlaufrichtung durch die Kurve verlaufe dabei entgegen des Uhrzeigersinns um die Flächennormale  $n$  an  $S$ , vgl. Abbildung.



Sei  $u : \mathbb{R}^3 \rightarrow \mathbb{R}^3$  ein stetig partiell differenzierbares Vektorfeld. Dann gilt

$$\int_S \langle \text{rot}(u), d\vec{S} \rangle = \int_\Gamma \langle u, dx \rangle,$$

wobei  $dx$  das vektorielle Bogenelement bezeichnet.

**Beweis:** Wir beweisen den Satz für den Spezialfall einer ebenen Fläche in der  $x_1 - x_2$ -Ebene im  $\mathbb{R}^3$  mit Normalenvektor  $n = [0, 0, 1]^\top$ . Dann folgt wegen  $d\vec{S} = n dS$  die Beziehung

$$\int_S \langle \text{rot}(u), d\vec{S} \rangle = \int_S \left( \frac{\partial u_2}{\partial x_1} - \frac{\partial u_1}{\partial x_2} \right) dx_1 dx_2 = \int_\Gamma u_2 \tilde{n}_1 - u_1 \tilde{n}_2 ds,$$

wobei  $\tilde{n}_1$  und  $\tilde{n}_2$  die Komponenten der äußeren Normalen an die Fläche sind. Bezeichnet  $t$  den Tangenteneinheitsvektor an die Randkurve, so gilt  $t = [t_1, t_2, 0]^\top = [-\tilde{n}_2, \tilde{n}_1, 0]^\top$  und es folgt

$$\int_\Gamma u_2 \tilde{n}_1 - u_1 \tilde{n}_2 ds = \int_\Gamma u_2 t_2 + u_1 t_1 ds = \int_\Gamma \langle u, t \rangle ds.$$

Mit  $dx = x' dt = t \|x'\| dt = t ds$  folgt die Behauptung. ■



## Kapitel 2

### Räume, Metriken und Normen

In diesem Kapitel beschäftigen wir uns mit der Frage, wie wir Abstände und Längen in abstrakten Mengen bzw. Vektorräumen messen können. Dies führt auf Metriken bzw. Normen. Darüber hinaus führen wir vollständige normierte Räume – sogenannte Banachräume – ein, die die Basis in vielen Anwendungen der Funktionalanalysis sind.

#### 2.1 Metriken und metrische Räume

Sei  $X$  eine beliebige nichtleere Menge. Um den Abstand zwischen zwei Elementen  $x$  und  $y$  der Menge  $X$  bestimmen zu können, definieren wir eine **Metrik** (=Abstandsmaß).

##### Definition 2.1.1 (Metrik)

Eine **Metrik** ist eine Abbildung  $d : X \times X \rightarrow \mathbb{R}$  mit den folgenden Eigenschaften:

(i) **Definitheit:**  $d(x, y) = 0 \iff x = y$ ;

(ii) **Symmetrie:**  $d(x, y) = d(y, x)$  für alle  $x, y \in X$ ;

(iii) **Dreiecksungleichung:**  $d(x, y) \leq d(x, z) + d(z, y)$  für alle  $x, y, z \in X$ .

$(X, d)$  heißt **metrischer Raum**.

Wegen

$$0 = d(x, x) \leq d(x, y) + d(y, x) = 2d(x, y)$$

folgt  $d(x, y) \geq 0$  für alle  $x, y \in X$ .

Ein metrischer Raum ist ein sehr allgemeines Konstrukt, auf dem aber zumindest ein Abstands begriff existiert.

##### Beispiel 2.1.2 (Euklidischer Abstand im $\mathbb{R}^n$ )

Für  $X = \mathbb{R}^n$  definiert

$$d(x, y) = \sqrt{\sum_{i=1}^n (x_i - y_i)^2}$$

den euklidischen Abstand zwischen  $x = [x_1, \dots, x_n]^\top$  und  $y = [y_1, \dots, y_n]^\top$  in  $\mathbb{R}^n$ . ■

### Beispiel 2.1.3 (diskrete Metrik)

Die Abbildung

$$d(x, y) = \begin{cases} 0, & \text{falls } x = y, \\ 1, & \text{sonst} \end{cases}$$

erfüllt alle Eigenschaften einer Metrik auf einer beliebigen Menge  $X$  und heißt **diskrete Metrik**. ■

### Beispiel 2.1.4 (Geodäte auf Kugeloberfläche)

Die Menge  $X$  sei gegeben durch die Kugeloberfläche in  $\mathbb{R}^3$ . Auf dieser Menge definieren die Geodäten (kürzeste Verbindung zwischen 2 Punkten auf der Kugeloberfläche) eine Metrik. ■

### Beispiel 2.1.5 (Hausdorff-Metrik)

Sei  $X$  die Menge aller kompakten Teilmengen des  $\mathbb{R}^n$ . Dann definiert

$$d(A, B) = \max \{ \max \{ \text{dist}(a, B) \mid a \in A \}, \max \{ \text{dist}(b, A) \mid b \in B \} \}$$

die **Hausdorff-Metrik** auf  $X$  für die Mengen  $A, B \in X$ , wobei

$$\text{dist}(x, A) = \min \{ \delta(x, a) \mid a \in A \}$$

den Abstand des Punktes  $x$  zur Menge  $A$  unter Verwendung einer Metrik  $\delta$  auf dem  $\mathbb{R}^n$  angibt.  $\delta$  kann z.B. der euklidische Abstand sein. ■

Mithilfe des Abstandsbegriffs kann man bereits Konvergenz von Folgen definieren.

### Definition 2.1.6

In einem metrischen Raum  $(X, d)$  heißt die Folge  $\{x_n\}_{n \in \mathbb{N}} \subset X$  **konvergent** mit Grenzwert  $x \in X$ , wenn  $d(x_n, x) \rightarrow 0$  für  $n \rightarrow \infty$  gilt. Wir schreiben dann auch kurz  $x_n \rightarrow x$  oder  $\lim_{n \rightarrow \infty} x_n = x$ .

Interessant sind vor allem vollständige metrische Räume.

### Definition 2.1.7

Ein metrischer Raum  $(X, d)$  heißt **vollständig**, wenn jede Cauchy-Folge aus  $X$  einen Grenzwert in  $X$  besitzt. Dabei heißt eine Folge  $\{x_n\}_{n \in \mathbb{N}} \subset X$  **Cauchy-Folge**, wenn für jedes  $\varepsilon > 0$  ein  $N(\varepsilon) \in \mathbb{N}$  existiert, so dass  $d(x_n, x_m) < \varepsilon$  für alle  $n, m > N(\varepsilon)$  gilt.

In einem metrischen Raum ist jede konvergente Folge automatisch auch Cauchy-Folge. Die Umkehrung gilt nur in vollständigen Räumen.

Schließlich definieren wir noch Stetigkeit von Funktionen.

### Definition 2.1.8

Eine Abbildung  $f : X \rightarrow Y$  zwischen metrischen Räumen  $(X, d_X)$  und  $(Y, d_Y)$  heißt stetig im Punkt  $x \in X$ , wenn

$$\lim_{x_n \rightarrow x} f(x_n) = f(x)$$

gilt. Anders ausgedrückt:

$$\forall \epsilon > 0 \exists \delta > 0 \forall \tilde{x} \in X : d_X(\tilde{x}, x) < \delta \implies d_Y(f(\tilde{x}), f(x)) < \epsilon.$$

## 2.2 Normen und normierte Vektorräume

Wie wir gesehen haben, können metrische Räume sehr allgemein sein, wobei i.A. keine strukturellen Eigenschaften und Rechenregeln (Addition, skalare Multiplikation,...) vorliegen.

Im Folgenden arbeiten wir nicht mit einer beliebigen Menge  $X$ , sondern betrachten einen Vektorraum  $V$  (über  $\mathbb{K} = \mathbb{R}$  oder  $\mathbb{K} = \mathbb{C}$ ). In einem Vektorraum sind Operationen wie Addition und Multiplikation mit Skalaren erlaubt und es gelten die üblichen Rechenregeln (Assoziativgesetz, Kommutativgesetz, Distributivgesetz, Existenz von Null- und Einselement, Existenz eines additiv inversen Elements):

### Definition 2.2.1 (Vektorraum, linearer Raum)

Ein Vektorraum  $(V, +, \cdot)$  über dem Skalarenkörper  $\mathbb{K} = \mathbb{R}$  oder  $\mathbb{K} = \mathbb{C}$  besteht aus einer Menge  $V$ , auf der die Addition zweier Elemente und die Multiplikation mit einem Skalar aus  $\mathbb{K}$  mit den folgenden Eigenschaften definiert sind:

(a) **Abgeschlossenheit:** Mit  $x, y \in V$  ist auch  $x + y \in V$ . Mit  $x \in V$  und  $\lambda \in \mathbb{K}$  ist auch  $\lambda x \in V$ .

(b) **Kommutativgesetz:** Für alle  $x, y \in V$  gilt

$$x + y = y + x.$$

(c) **Assoziativgesetz:** Für alle  $x, y, z \in V$  gilt

$$x + (y + z) = (x + y) + z.$$

(d) **Existenz eines Nullelements:** Es existiert  $0 \in V$  (Nullelement), so dass für alle  $x \in V$  gilt:

$$x + 0 = x.$$

(e) **Existenz eines additiv inversen Elements:** Zu jedem  $x \in V$  gibt es genau ein Element  $y \in V$  mit

$$x + y = 0.$$

(f) **Neutralität der 1:** Für alle  $x \in V$  gilt

$$1 \cdot x = x.$$

(g) **Multiplikation mit Skalar:** Für alle  $x, y \in V$  und  $\lambda, \mu \in \mathbb{K}$  gelten

$$\lambda(\mu x) = (\lambda\mu)x,$$

$$(\lambda + \mu)x = \lambda x + \mu x,$$

$$\lambda(x + y) = \lambda x + \lambda y.$$

Man schreibt häufig  $V$  an Stelle von  $(V, +, \cdot)$ , wenn klar ist, welche Addition und welche Multiplikation gemeint ist. Die Elemente des Vektorraums  $V$  nennt man **Vektoren**.

Aus der linearen Algebra ist insbesondere der Vektorraum  $\mathbb{R}^n$  bekannt, aber es gibt auch sehr viel allgemeinere Vektorräume.

### Beispiel 2.2.2

(a)  $\mathbb{R}^n$  ist ein Vektorraum über  $\mathbb{K} = \mathbb{R}$  und  $\mathbb{C}^n$  ist ein Vektorraum über  $\mathbb{K} = \mathbb{C}$ .

(b) Die Menge aller stetigen Funktionen auf dem kompakten Intervall  $[a, b]$

$$C([a, b], \mathbb{R}) := \{f : [a, b] \longrightarrow \mathbb{R} \mid f \text{ ist stetig}\}$$

sowie die Menge aller  $k$ -mal stetig differenzierbaren Funktionen auf dem kompakten Intervall  $[a, b]$

$$C^k([a, b], \mathbb{R}) := \{f : [a, b] \longrightarrow \mathbb{R} \mid f \text{ ist } k\text{-mal stetig differenzierbar}\}$$

versehen mit der punktweisen Addition und Multiplikation sind für  $k \in \mathbb{N}$  Vektorräume. (Man könnte das Intervall  $[a, b]$  auch durch einen metrischen Raum ersetzen und die Menge aller stetigen Abbildungen analog definieren).

(c) Die Menge aller Polynome vom Höchstgrad  $n \in \mathbb{N}_0$

$$\mathcal{P}_n := \{a_0 + a_1x + \dots + a_nx^n \mid a_j \in \mathbb{R}, j = 0, \dots, n\}$$

definiert einen Vektorraum über  $\mathbb{R}$ .

(d) Die Gerade

$$\left\{ x(t) \in \mathbb{R}^2 \mid x(t) = t \begin{bmatrix} 1 \\ -1 \end{bmatrix}, t \in \mathbb{R} \right\}$$

definiert einen Vektorraum über  $\mathbb{R}$ .

(e) Die Gerade

$$\left\{ x(t) \in \mathbb{R}^2 \mid x(t) = \begin{bmatrix} 1 \\ 0 \end{bmatrix} + t \begin{bmatrix} 1 \\ -1 \end{bmatrix}, t \in \mathbb{R} \right\}$$

definiert **keinen** Vektorraum über  $\mathbb{R}$  (Warum nicht?).

(f) Die Folgenräume

$$\begin{aligned} \ell^p &= \left\{ \{x_n\}_{n \in \mathbb{N}} \mid x_n \in \mathbb{R}, \sum_{k=1}^{\infty} |x_k|^p < \infty \right\}, \quad p \in \mathbb{N}, \\ \ell^\infty &= \{ \{x_n\}_{n \in \mathbb{N}} \mid x_n \in \mathbb{R}, \{x_n\}_{n \in \mathbb{N}} \text{ ist beschränkt} \} \end{aligned}$$

sind Vektorräume. Für  $\ell^\infty$  prüft man die Eigenschaften (a)-(g) in Definition 2.2.1 leicht nach. Für  $\ell^p$ ,  $p \in \mathbb{N}$ , ist insbesondere die Abgeschlossenheit bzgl. der Addition in (a) nicht offensichtlich und muss nachgewiesen werden. Seien dazu zwei Folgen  $x = \{x_n\}_{n \in \mathbb{N}}$  und  $y = \{y_n\}_{n \in \mathbb{N}}$  in  $\ell^p$  gegeben. Dann gilt

$$\begin{aligned} \sum_{k=1}^{\infty} |x_k + y_k|^p &\leq \sum_{k=1}^{\infty} (|x_k| + |y_k|)^p \\ &\leq \sum_{k=1}^{\infty} (2 \max\{|x_k|, |y_k|\})^p \\ &\leq 2^p \sum_{k=1}^{\infty} \max\{|x_k|^p, |y_k|^p\} \\ &\leq 2^p \sum_{k=1}^{\infty} (|x_k|^p + |y_k|^p) \\ &= 2^p \left( \sum_{k=1}^{\infty} |x_k|^p + \sum_{k=1}^{\infty} |y_k|^p \right) < \infty. \end{aligned}$$

Also folgt  $x + y \in \ell^p$ .

■

Analog zu den Folgenräumen lassen sich die sehr wichtigen  $\mathcal{L}^p$ -Räume definieren (später werden wir diese noch leicht modifizieren und bekommen  $L^p$ -Räume). Diese sind von grundlegender Bedeutung in der Funktionalanalysis.

**Satz 2.2.3** ( $\mathcal{L}^p$ -Räume)

Sei  $\Omega \subset \mathbb{R}^d$  nichtleer. Für  $p \in \mathbb{N}$  und  $1 \leq p < \infty$  sind durch

$$\mathcal{L}^p(\Omega) := \left\{ f : \Omega \longrightarrow \mathbb{R} \mid f \text{ ist messbar und } \int_{\Omega} |f|^p d\mu < \infty \right\}$$

Vektorräume definiert. Ebenso ist

$$\mathcal{L}^\infty(\Omega) := \{ f : \Omega \longrightarrow \mathbb{R} \mid f \text{ ist messbar und fast überall gilt } |f| < \infty \}$$

ein Vektorraum.

**Beweis:** Für  $\mathcal{L}^\infty$  prüft man die Eigenschaften (a)-(g) in Definition 2.2.1 leicht nach. Für  $\mathcal{L}^p$ ,  $p \in \mathbb{N}$ , ist insbesondere die Abgeschlossenheit in (a) nicht offensichtlich und muss nachgewiesen werden.

Seien  $f, g \in \mathcal{L}^p(\Omega)$  mit  $1 \leq p < \infty$ . Für  $\lambda \in \mathbb{R}$  gilt

$$\int_{\Omega} |\lambda f|^p d\mu = |\lambda|^p \int_{\Omega} |f|^p d\mu < \infty \quad (1 \leq p < \infty).$$

Des Weiteren gilt

$$\begin{aligned} \int_{\Omega} |f + g|^p d\mu &\leq \int_{\Omega} (|f| + |g|)^p d\mu \\ &\leq \int_{\Omega} (2 \max\{|f|, |g|\})^p d\mu \\ &\leq 2^p \int_{\Omega} \max\{|f|^p, |g|^p\} d\mu \\ &\leq 2^p \int_{\Omega} |f|^p + |g|^p d\mu \\ &= 2^p \left( \int_{\Omega} |f|^p d\mu + \int_{\Omega} |g|^p d\mu \right) < \infty. \end{aligned}$$

■

Wir führen eine **Norm** auf dem Raum  $V$  ein, die wir im Folgenden mit  $\|\cdot\|_V$  bezeichnen werden. Anschaulich misst eine Norm die **Länge** eines Vektors.

**Definition 2.2.4** (Norm)

Eine **Norm** auf dem Vektorraum  $V$  über  $\mathbb{K} = \mathbb{R}$  oder  $\mathbb{K} = \mathbb{C}$  ist eine Abbildung  $\|\cdot\|_V : V \rightarrow \mathbb{R}$  mit den folgenden Eigenschaften:

- (i) **Definitheit:**  $\|x\|_V = 0$  genau dann, wenn  $x = 0$ ;
- (ii) **Translation:**  $\|\lambda x\|_V = |\lambda| \|x\|_V$  für alle  $x \in V$ ,  $\lambda \in \mathbb{K}$ ;
- (iii) **Dreiecksungleichung:**  $\|x + y\|_V \leq \|x\|_V + \|y\|_V$  für alle  $x, y \in V$ .

Ein mit einer Norm versehener Vektorraum  $V$  heißt **normierter Raum** und man schreibt hierfür auch  $(V, \|\cdot\|_V)$ .

**Bemerkung 2.2.5**

- (a) Jede Norm  $\|\cdot\|_V$  induziert eine Metrik  $d(x, y) := \|x - y\|_V$ . Somit ist jeder normierte Raum  $(V, \|\cdot\|_V)$  auch ein metrischer Raum und Begriffe wie Konvergenz, Cauchy-Folge und Stetigkeit von Abbildungen übertragen sich auf normierte Räume.
- (b) Aus der Dreiecksungleichung (iii) in Definition 2.2.4 folgt wegen  $\|x\|_V = \|x - y + y\|_V \leq \|x - y\|_V + \|y\|_V$  sofort die Dreiecksungleichung nach unten:

$$\left| \|x\|_V - \|y\|_V \right| \leq \|x - y\|_V \quad \forall x, y \in V.$$

■

**Beispiel 2.2.6** (Normen auf dem  $\mathbb{R}^n$ )

Auf dem reellen Vektorraum  $V = \mathbb{R}^n$  mit  $n \in \mathbb{N}$  sind durch

$$\begin{aligned} \|x\|_\infty &= \max_{k=1, \dots, n} |x_k|, && \text{(Maximumnorm),} \\ \|x\|_1 &= \sum_{k=1}^n |x_k|, && \text{(1-Norm, Manhattan-Norm),} \\ \|x\|_2 &= \sqrt{\sum_{k=1}^n |x_k|^2}, && \text{(2-Norm, euklidische Norm),} \\ \|x\|_p &= \left( \sum_{k=1}^n |x_k|^p \right)^{\frac{1}{p}}, \quad p \in \mathbb{N}, && \text{(p-Norm),} \end{aligned}$$

mit  $x = [x_1, \dots, x_n]^\top$  verschiedene Normen definiert. ■

Man nennt zwei Normen  $\|\cdot\|_a$  und  $\|\cdot\|_b$  **äquivalent**, falls es Konstanten  $c_1, c_2 > 0$  mit

$$c_1\|x\|_a \leq \|x\|_b \leq c_2\|x\|_a \quad \forall x \in V$$

gibt. Äquivalente Normen sind gewissermaßen austauschbar, weil konvergente Folgen bzgl. einer Norm auch konvergent bzgl. einer äquivalenten Norm sind. Bei nicht äquivalenten Normen gilt dies nicht. Tatsächlich sind alle Normen auf dem  $\mathbb{R}^n$  äquivalent (allgemeiner gilt dies sogar für alle endlichdimensionalen Vektorräume).

### Satz 2.2.7

Alle Normen auf dem  $\mathbb{R}^n$  sind äquivalent.

**Beweis:** Es genügt zu zeigen, dass irgendeine Norm  $\|\cdot\|$  auf dem  $\mathbb{R}^n$  äquivalent zur 1-Norm  $\|\cdot\|_1$  ist. Mit den Einheitsvektoren  $e_i$ ,  $i = 1, \dots, n$ , folgt

$$\|x\| = \left\| \sum_{i=1}^n x_i e_i \right\| \leq \sum_{i=1}^n |x_i| \cdot \|e_i\| \leq c_2 \|x\|_1$$

für eine Konstante  $c_2$  mit  $c_2 \geq \|e_i\| > 0$  für alle  $i = 1, \dots, n$ .

Wegen

$$|\|x\| - \|y\|| \leq \|x - y\| \leq c_2 \|x - y\|_1$$

ist die Abbildung  $f(x) = \|x\|$  stetig bzgl. der 1-Norm, denn aus  $\|x_k - x\|_1 \rightarrow 0$  folgt auch  $|f(x_k) - f(x)| \rightarrow 0$  für  $k \rightarrow \infty$ .

Die Einheitskugel  $S = \{x \in \mathbb{R}^n \mid \|x\|_1 = 1\}$  bzgl. der 1-Norm ist beschränkt und abgeschlossen, also kompakt. Nach dem Satz von Weierstraß<sup>1</sup> nimmt die stetige Funktion  $f$  auf  $S$  ihr Minimum  $c_1$  an, d.h. es gibt ein  $\hat{x} \in S$  mit  $c_1 = f(\hat{x}) = \|\hat{x}\| \geq 0$  und  $c_1 \leq f(x) = \|x\|$  für alle  $x \in S$ . Folglich gilt

$$c_1 \leq \frac{\|x\|}{\|x\|_1} \quad \forall x \in \mathbb{R}^n.$$

Tatsächlich ist  $c_1 > 0$ , denn ansonsten würde  $0 = \|\hat{x}\|$  und somit  $\hat{x} = 0$  gelten, so dass  $\hat{x}$  nicht in  $S$  wäre. Daraus folgt  $c_1\|x\|_1 \leq \|x\|$  für alle  $x \in \mathbb{R}^n$  mit  $c_1 > 0$ . ■

### Beispiel 2.2.8 (Matrixnormen)

<sup>1</sup>Der Satz von Weierstraß besagt gerade, dass jede stetige Funktion auf einer kompakten Menge ihr Maximum und Minimum annimmt.



Auf dem Raum aller reellen  $m \times n$ -Matrizen  $V = \mathbb{R}^{m \times n}$  mit  $m, n \in \mathbb{N}$  sind durch

$$\begin{aligned} \|A\|_\infty &= \max_{1 \leq i \leq m} \sum_{j=1}^n |a_{ij}|, && \text{(Zeilensummennorm),} \\ \|A\|_1 &= \max_{1 \leq j \leq n} \sum_{i=1}^m |a_{ij}|, && \text{(Spaltensummennorm),} \\ \|A\|_2 &= \sqrt{\rho(A^\top A)}, && \text{(Spektralnrm, } \rho = \text{betragsmäßig größter Eigenwert),} \\ \|A\|_F &= \sqrt{\sum_{\substack{i=1, \dots, m \\ j=1, \dots, n}} |a_{ij}|^2}, && \text{(Frobenius-Schur-Norm),} \end{aligned}$$

mit  $A = [a_{ij}] \in \mathbb{R}^{m \times n}$  verschiedene Normen definiert. Diese sind **äquivalent**. ■

### Beispiel 2.2.9 (Normen für stetige Funktionen)

Auf dem Vektorraum der stetigen Funktionen  $V = C([a, b], \mathbb{R})$  sind durch

$$\begin{aligned} \|x\|_\infty &= \max_{t \in [a, b]} |x(t)|, && \text{(Supremumsnorm),} \\ \|x\|_1 &= \int_a^b |x(t)| dt, && \text{(1-Norm),} \\ \|x\|_2 &= \sqrt{\int_a^b |x(t)|^2 dt}, && \text{(2-Norm),} \\ \|x\|_p &= \left( \int_a^b |x(t)|^p dt \right)^{\frac{1}{p}}, \quad p \in \mathbb{N}, && \text{(p-Norm),} \end{aligned}$$

verschiedene Normen definiert. Beachte, dass diese Normen i.A. **nicht äquivalent** sind, d.h. man findet i.A. keine positiven Konstanten  $c_1$  und  $c_2$ , so dass

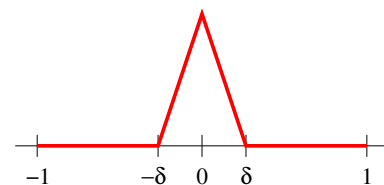
$$c_1 \|x\|_a \leq \|x\|_b \leq c_2 \|x\|_a$$

für alle  $x \in V$  gilt. Hierbei sind einseitige Abschätzungen aber durchaus möglich, denn alle  $p$ -Normen können auf einer kompakten Menge  $[a, b]$  wegen

$$\|x\|_p^p = \int_a^b |x(t)|^p dt \leq \int_a^b \max_{\tau \in [a, b]} |x(\tau)|^p dt \leq (b-a) \|x\|_\infty^p$$

nach oben durch die Supremumsnorm abgeschätzt werden. Umgekehrt ist eine solche Abschätzung nicht möglich. Um dies einzusehen, betrachten wir exemplarisch die 1-Norm und für  $\delta > 0$  die stetigen Funktionen

$$x_\delta(t) = \begin{cases} 1 + \frac{t}{\delta}, & \text{für } t \in [-\delta, 0), \\ 1 - \frac{t}{\delta}, & \text{für } t \in [0, \delta], \\ 0, & \text{sonst.} \end{cases}$$



Es gilt  $\|x_\delta\|_\infty = 1$  für alle  $\delta > 0$  und

$$\|x_\delta\|_1 = \int_{-1}^1 |x_\delta(t)| dt = 2 \int_0^\delta 1 - \frac{t}{\delta} dt = \delta.$$

Für  $\delta \rightarrow 0$  gilt also  $\|x\|_1 \rightarrow 0$  und somit kann es keine Konstante  $c_1 > 0$  mit  $c_1 \|x_\delta\|_\infty \leq \|x_\delta\|_1$  für alle  $\delta > 0$  geben. Damit ist eine Abschätzung der Form  $c_1 \|x\|_\infty \leq \|x\|_1$  mit einer Konstanten  $c_1 > 0$ , die für alle (!) stetigen Funktionen  $x \in C([a, b], \mathbb{R})$  gilt, nicht möglich. ■

### Beispiel 2.2.10 (Normen für stetig differenzierbare Funktionen)

Auf dem Vektorraum der stetig differenzierbaren Funktionen  $V = C^1([a, b], \mathbb{R})$  sind durch

$$\begin{aligned} \|x\|_{1,\infty} &= \max\{\|x\|_\infty, \|x'\|_\infty\}, \\ \|x\|_{1,p} &= \max\{\|x\|_p, \|x'\|_p\}, \quad p \in \mathbb{N}, \\ \|x\| &= |x(a)| + \|x'\|_p, \quad p \in \mathbb{N} \cup \{\infty\}, \end{aligned}$$

verschiedene Normen definiert. Auch diese Normen sind i.A. **nicht äquivalent**. ■

### Beispiel 2.2.11 (Normen für Folgen)

Die **Supremumsnorm**

$$\|x\|_{\ell^\infty} = \sup_{n \in \mathbb{N}} |x_n|$$

ist eine Norm auf dem Folgenraum  $\ell^\infty$ . Die Normeigenschaften kann man hierfür leicht nachrechnen.

Für die Folgenräume  $\ell^p$ ,  $p \in \mathbb{N}$ , sind durch

$$\|x\|_{\ell^p} = \left( \sum_{k=1}^{\infty} |x_k|^p \right)^{\frac{1}{p}}$$

Normen gegeben. Bis auf die Dreiecksungleichung in Definition 2.2.4 sind alle Normeigenschaften offensichtlich. Die Dreiecksungleichung folgt aber mit der Minkowski-Ungleichung in Satz 2.2.13 weiter unten. ■

Zum Beweis der Minkowski-Ungleichung, die wir zum Nachweis der Dreiecksungleichung benötigen, benötigen wir die Hölder-Ungleichung.

### Satz 2.2.12 (Hölder-Ungleichung für Folgen)

Für  $x \in \ell^p$  und  $y \in \ell^q$  mit  $1 \leq p \leq \infty$  und  $\frac{1}{p} + \frac{1}{q} = 1$  gilt

$$\|x \cdot y\|_{\ell^1} \leq \|x\|_{\ell^p} \cdot \|y\|_{\ell^q}.$$

(Hierbei ist  $q = \infty$  für  $p = 1$  und umgekehrt)

**Beweis:** Für  $p = 1$  und  $q = \infty$  erhält man

$$\|x \cdot y\|_{\ell^1} = \sum_{k=1}^{\infty} |x_k \cdot y_k| \leq \sum_{k=1}^{\infty} |x_k| \cdot \sup_{j \in \mathbb{N}} |y_j| = \|x\|_{\ell^1} \cdot \|y\|_{\ell^\infty}.$$

Für  $1 < p < \infty$  nutzen wir für Konstanten  $\alpha, \beta > 0$  die Konkavität des  $\log$  aus und erhalten die Abschätzung

$$\log\left(\frac{|x_k|}{\alpha} \cdot \frac{|y_k|}{\beta}\right) = \frac{1}{p} \log\left(\frac{|x_k|^p}{\alpha^p}\right) + \frac{1}{q} \log\left(\frac{|y_k|^q}{\beta^q}\right) \leq \log\left(\frac{1}{p} \frac{|x_k|^p}{\alpha^p} + \frac{1}{q} \frac{|y_k|^q}{\beta^q}\right)$$

und folglich

$$\frac{|x_k|}{\alpha} \cdot \frac{|y_k|}{\beta} \leq \left(\frac{1}{p} \frac{|x_k|^p}{\alpha^p} + \frac{1}{q} \frac{|y_k|^q}{\beta^q}\right).$$

Wähle nun  $\alpha = \|x\|_{\ell^p}$  und  $\beta = \|y\|_{\ell^q}$ . Falls  $\alpha = 0$  oder  $\beta = 0$ , so folgt  $x = 0$  bzw.  $y = 0$  und die Behauptung des Satzes ist richtig. Für  $\alpha, \beta > 0$  folgt durch Summation über  $k \in \mathbb{N}$  der obigen Ungleichung auf beiden Seiten

$$\sum_{k=1}^{\infty} \frac{|x_k|}{\alpha} \cdot \frac{|y_k|}{\beta} \leq \sum_{k=1}^{\infty} \left(\frac{1}{p} \frac{|x_k|^p}{\alpha^p} + \frac{1}{q} \frac{|y_k|^q}{\beta^q}\right) = \frac{1}{p} + \frac{1}{q} = 1$$

und folglich

$$\|x \cdot y\|_{\ell^1} = \left| \sum_{k=1}^{\infty} x_k \cdot y_k \right| \leq \sum_{k=1}^{\infty} |x_k| \cdot |y_k| \leq \alpha \cdot \beta = \|x\|_{\ell^p} \cdot \|y\|_{\ell^q}.$$

■

**Satz 2.2.13** (Minkowski-Ungleichung für Folgen)

Für  $x, y \in \ell^p$  mit  $1 \leq p < \infty$  gilt

$$\|x + y\|_{\ell^p} \leq \|x\|_{\ell^p} + \|y\|_{\ell^p}.$$

**Beweis:** Für  $p = 1$  erhalten wir

$$\|x + y\|_{\ell^1} = \sum_{k=1}^{\infty} |x_k + y_k| \leq \sum_{k=1}^{\infty} |x_k| + |y_k| = \sum_{k=1}^{\infty} |x_k| + \sum_{k=1}^{\infty} |y_k| = \|x\|_{\ell^1} + \|y\|_{\ell^1}.$$

Für  $p > 1$  nutzen wir die Hölder-Ungleichung aus.

$$\begin{aligned}
 \|x + y\|_{\ell^p}^p &= \sum_{k=1}^{\infty} |x_k + y_k|^p \\
 &= \sum_{k=1}^{\infty} |x_k + y_k| \cdot |x_k + y_k|^{p-1} \\
 &\leq \sum_{k=1}^{\infty} |x_k| \cdot |x_k + y_k|^{p-1} + \sum_{k=1}^{\infty} |y_k| \cdot |x_k + y_k|^{p-1} \\
 &\leq \left( \sum_{k=1}^{\infty} |x_k|^p \right)^{\frac{1}{p}} \cdot \left( \sum_{k=1}^{\infty} |x_k + y_k|^{(p-1)q} \right)^{\frac{1}{q}} \\
 &\quad + \left( \sum_{k=1}^{\infty} |y_k|^p \right)^{\frac{1}{p}} \cdot \left( \sum_{k=1}^{\infty} |x_k + y_k|^{(p-1)q} \right)^{\frac{1}{q}},
 \end{aligned}$$

wobei  $q = p/(p-1)$  und  $1/p + 1/q = 1$  gilt. Wegen  $q(p-1) = p$  und  $1 - 1/q = 1/p$  folgt nach Division durch die rechte Summe die Behauptung.

Für die Sonderfälle  $\sum_{k=1}^{\infty} |x_k + y_k|^p = 0$  oder  $\sum_{k=1}^{\infty} |x_k|^p = \infty$  bzw.  $\sum_{k=1}^{\infty} |y_k|^p = \infty$  ist die Behauptung trivialerweise erfüllt. ■

Die Hölder-Ungleichung und die Minkowski-Ungleichung lassen sich auch für Funktionen beweisen.

### Satz 2.2.14 (Hölder-Ungleichung für Funktionen)

Seien  $f, g : \Omega \rightarrow \mathbb{R}$  messbare Funktionen,  $p, q > 1$  reelle Zahlen mit  $\frac{1}{p} + \frac{1}{q} = 1$  und  $\int_{\Omega} |f|^p d\mu < \infty$  und  $\int_{\Omega} |g|^q d\mu < \infty$ .

Dann ist  $f \cdot g$  Lebesgue-integrierbar und es gilt die Ungleichung

$$\left| \int_{\Omega} f \cdot g d\mu \right| \leq \left( \int_{\Omega} |f|^p d\mu \right)^{1/p} \left( \int_{\Omega} |g|^q d\mu \right)^{1/q}.$$

**Beweis:** Der Beweis kann analog zum Beweis von Satz 2.2.12 geführt werden, indem Folgenglieder durch Funktionswerte und die Summen durch Integrale ersetzt werden ( $\rightsquigarrow$  Übung!) ■

### Satz 2.2.15 (Minkowski-Ungleichung für Funktionen)

Für messbare Funktionen  $f, g : \Omega \rightarrow \mathbb{R}$  gilt

$$\left( \int_{\Omega} |f + g|^p d\mu \right)^{\frac{1}{p}} \leq \left( \int_{\Omega} |f|^p d\mu \right)^{\frac{1}{p}} + \left( \int_{\Omega} |g|^p d\mu \right)^{\frac{1}{p}} \quad (1 \leq p < \infty).$$

**Beweis:** Der Beweis kann analog zum Beweis von Satz 2.2.13 geführt werden, indem Folgenglieder durch Funktionswerte und die Summen durch Integrale ersetzt werden ( $\rightsquigarrow$  Übung!) ■

Die Minkowski-Ungleichung für Folgen in Satz 2.2.13 wurde verwendet, um die Dreiecksungleichung für die Norm  $\|\cdot\|_{\ell^p}$  zu zeigen. Es liegt nun Nahe, dieselbe Vorgehensweise auch für Funktionen aus  $\mathcal{L}^p(\Omega)$  zu wählen. Hierbei tritt ein Problem auf. Definiert man analog zu den Folgenraumnormen die Größen

$$\|f\|_{\mathcal{L}^p(\Omega)} := \left( \int_{\Omega} |f|^p d\mu \right)^{\frac{1}{p}} \quad (1 \leq p < \infty) \quad (2.1)$$

und

$$\|f\|_{\mathcal{L}^\infty(\Omega)} := \inf\{c > 0 \mid |f| \leq c \text{ fast überall}\}, \quad (2.2)$$

so erfüllen diese die Eigenschaften  $\|\lambda f\|_{\mathcal{L}^p(\Omega)} = |\lambda| \cdot \|f\|_{\mathcal{L}^p(\Omega)}$  für alle  $\lambda \in \mathbb{R}$  und wegen der Minkowski-Ungleichung 2.2.15 auch die Dreiecksungleichung

$$\|f + g\|_{\mathcal{L}^p(\Omega)} \leq \|f\|_{\mathcal{L}^p(\Omega)} + \|g\|_{\mathcal{L}^p(\Omega)} \quad \forall f, g \in \mathcal{L}^p(\Omega)$$

mit  $1 \leq p \leq \infty$  (für  $p = \infty$  sieht man dies direkt).

Das Problem besteht nun darin, dass man aus  $\|f\|_{\mathcal{L}^p(\Omega)} = 0$  **nicht** folgern kann, dass  $f$  die Nullfunktion ist, da man Funktionen aus  $\mathcal{L}^p(\Omega)$  stets auf Nullmengen beliebig abändern kann, ohne die Größen  $\|\cdot\|_{\mathcal{L}^p(\Omega)}$  zu verändern. Daher sind die in (2.1) und (2.2) definierten Größen streng genommen keine Normen auf  $\mathcal{L}^p(\Omega)$ .

Allerdings kann man aus  $\|f\|_{\mathcal{L}^p(\Omega)} = 0$  folgern, dass  $f = 0$  fast überall gilt, d.h. bis auf eine Nullmenge gilt  $f = 0$ . Fasst man nun alle Funktionen, die sich nur auf einer Nullmenge unterscheiden, zu einer sogenannten **Äquivalenzklasse**

$$[f] := \{g : \Omega \rightarrow \mathbb{R} \mid g = f \text{ fast überall}\}$$

zusammen, so kann man für  $1 \leq p \leq \infty$  neue Vektorräume definieren, die aus diesen Äquivalenzklassen bestehen:

$$L^p(\Omega) := \{[f] \mid f \in \mathcal{L}^p(\Omega)\}.$$

Die Elemente dieser Räume sind also Klassen von Funktionen, die sich nur auf Nullmengen unterscheiden. Für diese Räume sind folglich alle Normeigenschaften erfüllt und es gilt:

**Satz 2.2.16**

Für  $1 \leq p \leq \infty$  sind durch  $\|\cdot\|_{L^p(\Omega)} := \|\cdot\|_{\mathcal{L}^p(\Omega)}$  Normen auf  $L^p(\Omega)$  definiert.

**Bemerkung 2.2.17**

Im alltäglichen Gebrauch unterscheidet man nicht zwischen  $f$  und der Äquivalenzklasse  $[f]$  und identifiziert  $\mathcal{L}^p(\Omega)$  und  $L^p(\Omega)$  miteinander. ■

**2.3 Banachräume**

Besonders wichtig sind normierte Räume, die vollständig sind.

Gemäß Definition 2.1.7 nennt man einen normierten Raum  $(V, \|\cdot\|_V)$  **vollständig**, wenn jede Cauchy-Folge aus  $V$  einen Grenzwert in  $V$  besitzt.

**Definition 2.3.1 (Banachraum)**

Ein vollständiger, normierter Vektorraum  $(V, \|\cdot\|_V)$  heißt **Banachraum**.

Wir betrachten wichtige Banachräume.

**Satz 2.3.2**

Der Raum  $\mathbb{R}^n$  versehen mit einer der Normen aus Beispiel 2.2.6 ist ein Banachraum.

**Beweis:** Dass  $\mathbb{R}^n$  vollständig ist, wird in der Analysis gezeigt. ■

**Satz 2.3.3**

Sei  $[a, b]$  kompakt. Der Vektorraum  $C([a, b], \mathbb{R})$  versehen mit der Norm  $\|\cdot\|_\infty$  ist vollständig, also ein Banachraum.

**Beweis:** Sei  $\{x_n\}_{n \in \mathbb{N}}$  eine Cauchy-Folge in  $C([a, b], \mathbb{R})$ . Wir werden die Existenz einer stetigen Funktion  $x \in C([a, b], \mathbb{R})$  zeigen, gegen die die Folge  $\{x_n\}_{n \in \mathbb{N}}$  konvergiert.

Da  $\{x_n\}_{n \in \mathbb{N}}$  eine Cauchy-Folge ist, gilt

$$\forall \varepsilon > 0 \exists N(\varepsilon) \in \mathbb{N} \forall n, m > N(\varepsilon) : \|x_n - x_m\|_\infty = \max_{t \in [a, b]} |x_n(t) - x_m(t)| < \varepsilon. \quad (2.3)$$

Damit existiert für jedes  $t \in [a, b]$  auf Grund der Vollständigkeit von  $\mathbb{R}$  der Grenzwert der Funktionswerte  $\{x_n(t)\}_{n \in \mathbb{N}}$ , den wir mit  $x(t)$  bezeichnen wollen. Auf diese Weise wird eine Abbildung  $x : [a, b] \rightarrow \mathbb{R}$  gemäß  $t \mapsto x(t) = \lim_{n \rightarrow \infty} x_n(t)$  definiert.

Wir zeigen, dass  $x$  stetig auf  $[a, b]$  ist. Dazu seien  $t_0 \in [a, b]$  und  $\varepsilon > 0$  beliebig. Wegen (2.3) gibt es zu  $\varepsilon/3$  ein  $N(\varepsilon/3)$  mit

$$|x_n(t) - x_m(t)| < \varepsilon/3 \quad \forall t \in [a, b], n, m > N(\varepsilon/3).$$

Für  $m \rightarrow \infty$  folgt daraus

$$|x_n(t) - x(t)| < \varepsilon/3 \quad \forall t \in [a, b], n > N(\varepsilon/3).$$

Hieraus folgt sofort

$$\|x_n - x\|_\infty = \max_{t \in [a, b]} |x_n(t) - x(t)| < \varepsilon/3 \quad \forall n > N(\varepsilon/3),$$

so dass  $\{x_n\}_{n \in \mathbb{N}}$  gegen  $x$  in der Norm  $\|\cdot\|_\infty$  konvergiert.

Da  $x_n$  stetig in  $t_0$  ist, gibt es ein  $\delta > 0$ , so dass

$$|x_n(t) - x_n(t_0)| < \varepsilon/3 \quad \forall t \in [a, b], |t - t_0| < \delta.$$

Damit folgt für alle  $t \in [a, b]$ ,  $|t - t_0| < \delta$  und  $n > N(\varepsilon/3)$  die Beziehung

$$|x(t) - x(t_0)| \leq |x(t) - x_n(t)| + |x_n(t_0) - x(t_0)| + |x_n(t) - x_n(t_0)| \leq \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3} = \varepsilon.$$

Damit ist die punktweise Stetigkeit von  $x$  in  $t_0$  gezeigt. Da  $t_0 \in [a, b]$  beliebig war, ist  $x$  punktweise stetig in  $[a, b]$ . ■

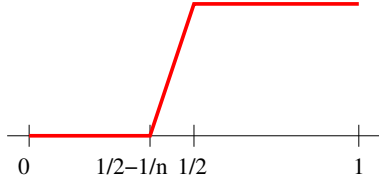
### Bemerkung 2.3.4

Der Vektorraum  $C^1([a, b], \mathbb{R})$  der auf der kompakten Menge  $[a, b]$  stetig differenzierbaren Funktionen versehen mit der Norm  $\|\cdot\|_{1, \infty}$  ist ebenfalls vollständig, also ein Banachraum. ■

Versieht man den Vektorraum der stetigen Funktionen mit einer anderen Norm, so muss der so entstandene normierte Raum nicht zwangsläufig ein Banachraum sein.

### Beispiel 2.3.5 (Beispiel eines nicht vollständigen Raums)

Der Vektorraum  $C([a, b], \mathbb{R})$  mit der Norm  $\|\cdot\|_p$  mit  $p \in \mathbb{N}$  ist **nicht vollständig**, also **kein Banachraum**. Betrachte z.B. für  $n \in \mathbb{N}$  die stetigen Funktionen

$$x_n(t) = \begin{cases} 0, & \text{für } 0 \leq t < \frac{1}{2} - \frac{1}{n}, \\ nt + 1 - \frac{n}{2}, & \text{für } \frac{1}{2} - \frac{1}{n} \leq t < \frac{1}{2}, \\ 1, & \text{für } \frac{1}{2} \leq t \leq 1. \end{cases}$$


Man rechnet leicht nach, dass  $\{x_n\}_{n \in \mathbb{N}}$  eine Cauchy-Folge bzgl. der Norm  $\|\cdot\|_p$ ,  $p \in \mathbb{N}$ , ist. Allerdings ist die Grenzfunktion

$$x = \lim_{n \rightarrow \infty} x_n = \begin{cases} 0, & \text{für } 0 \leq t < \frac{1}{2}, \\ 1, & \text{für } \frac{1}{2} \leq t \leq 1 \end{cases}$$

unstetig, also nicht in  $C([0, 1], \mathbb{R})$  enthalten. ■

### Bemerkung 2.3.6

Der Vektorraum  $C^1([a, b], \mathbb{R})$  versehen mit einer der Normen  $\|\cdot\|_p$ ,  $p \in \mathbb{N} \cup \{\infty\}$ , oder  $\|\cdot\|_{1,p}$ ,  $p \in \mathbb{N}$ , ist **nicht vollständig**, also kein Banachraum. Gegenbeispiele können analog zu Beispiel 2.3.5 konstruiert werden. ■

### Satz 2.3.7

Der Folgenraum  $\ell^\infty$  versehen mit der Supremumsnorm  $\|\cdot\|_{\ell^\infty}$  ist ein Banachraum. Für  $1 \leq p < \infty$  ist der Folgenraum  $\ell^p$  versehen mit der Norm  $\|\cdot\|_{\ell^p}$  ein Banachraum.

### Beweis:

- (a) Sei  $\{x_n\}_{n \in \mathbb{N}}$  eine Cauchyfolge in  $\ell^\infty$  (also eine Folgen von Folgen). Wir schreiben  $x_n = \{x_m^{(n)}\}_{m \in \mathbb{N}}$ . Für beliebiges  $\varepsilon > 0$  gibt es dann ein  $N(\varepsilon) \in \mathbb{N}$  mit

$$\|x_n - x_m\|_\infty = \sup_{k \in \mathbb{N}} |x_k^{(n)} - x_k^{(m)}| < \varepsilon \quad \forall n, m > N(\varepsilon).$$

Damit sind für festes  $k$  die Folgen  $\{x_k^{(n)}\}_{n \in \mathbb{N}}$  Cauchyfolgen in  $\mathbb{R}$  und wir konstruieren eine neue Folge  $z = \{z_k\}_{k \in \mathbb{N}}$  durch  $z_k := \lim_{n \rightarrow \infty} x_k^{(n)}$  für  $k \in \mathbb{N}$ . Aus

$$|x_k^{(n)} - x_k^{(m)}| < \varepsilon \quad \forall n, m > N(\varepsilon)$$

folgt für  $m \rightarrow \infty$  und alle  $k \in \mathbb{N}$  die Beziehung

$$|x_k^{(n)} - z_k| \leq \varepsilon \quad \forall n > N(\varepsilon)$$

bzw.

$$\|x_n - z\|_{\ell^\infty} = \sup_{k \in \mathbb{N}} |x_k^{(n)} - z_k| \leq \varepsilon \quad \forall n > N(\varepsilon)$$

Folglich konvergiert die Folge  $\{x_n\}_{n \in \mathbb{N}}$  gegen  $z$  in der Norm  $\|\cdot\|_{\ell^\infty}$ . Für das feste  $M = N(\varepsilon) + 1$  ist  $x_M \in \ell^\infty$  und somit beschränkt. Mit der Dreiecksungleichung folgt  $\|z\|_{\ell^\infty} \leq \|x_M\|_{\ell^\infty} + \|z - x_M\|_{\ell^\infty} \leq \|x_M\|_{\ell^\infty} + \varepsilon$ . Folglich ist  $z \in \ell^\infty$ , was die Vollständigkeit zeigt.

- (b) Wir gehen wie in (a) vor und definieren für eine Cauchyfolge  $\{x_n\}_{n \in \mathbb{N}}$  in  $\ell^p$  die Folge  $z = \{z_k\}_{k \in \mathbb{N}}$  durch  $z_k := \lim_{n \rightarrow \infty} x_k^{(n)}$  für  $k \in \mathbb{N}$ . Für beliebiges  $\varepsilon > 0$  gibt es dann ein  $N(\varepsilon) \in \mathbb{N}$  mit

$$\|x_n - x_m\|_{\ell^p} = \left( \sum_{k=1}^{\infty} |x_k^{(n)} - x_k^{(m)}|^p \right)^{\frac{1}{p}} < \varepsilon \quad \forall n, m > N(\varepsilon).$$



Insbesondere gilt dann für ein festes, aber beliebiges  $M$ :

$$\left( \sum_{k=1}^M |x_k^{(n)} - x_k^{(m)}| \right)^{\frac{1}{p}} < \varepsilon \quad \forall n, m > N(\varepsilon).$$

Übergang zum Grenzwert  $m \rightarrow \infty$  und anschließender Übergang  $M \rightarrow \infty$  impliziert

$$\|x_n - z\|_{\ell^p} = \left( \sum_{k=1}^{\infty} |x_k^{(n)} - z_k| \right)^{\frac{1}{p}} < \varepsilon \quad \forall n > N(\varepsilon).$$

Insbesondere ist  $z - x_{N(\varepsilon)+1} \in \ell^p$  und, da  $\ell^p$  ein Vektorraum ist, ist auch  $z = (z - x_{N(\varepsilon)+1}) + x_{N(\varepsilon)+1}$  in  $\ell^p$ , was die Vollständigkeit zeigt. ■

### Satz 2.3.8

Sei  $\Omega \subset \mathbb{R}^d$  nichtleer. Der Vektorraum  $L^p(\Omega)$  versehen mit der Norm  $\|\cdot\|_{L^p(\Omega)}$  ist für  $1 \leq p \leq \infty$  ein Banachraum.

**Beweis:** Der Nachweis der Vollständigkeit kann in [Wer95, S. 19] oder in [BK11, Satz VI.2 von Riesz-Fischer] nachgelesen werden. ■

### Bemerkung 2.3.9

Es gilt folgende Inklusionskette (für beschränktes Gebiet  $\Omega$ ):

$$\dots \subset C^k(\Omega, \mathbb{R}) \subset \dots \subset C^1(\Omega, \mathbb{R}) \subset C(\Omega, \mathbb{R}) \subset L^\infty(\Omega) \subset \dots \subset L^p(\Omega) \subset \dots \subset L^1(\Omega)$$

Es gibt zahlreiche weitere Funktionenräume, die insbesondere bei der Analyse von partiellen Differentialgleichungen und Problemen der optimalen Steuerung eine wichtige Rolle spielen, etwa Funktionen beschränkter Variation  $BV([a, b], \mathbb{R})$  oder absolutstetige Funktionen  $AC([a, b], \mathbb{R})$ .

Interessant sind aber vor allem die sogenannten **Sobolev-Räume**, die mithilfe von **schwachen (oder verallgemeinerte) Ableitungen** definiert werden. Um diese einzuführen, beschränken wir uns auf den eindimensionalen Fall und betrachten das Gebiet  $\Omega = (a, b) \subset \mathbb{R}$  und  $f \in C^1(\Omega, \mathbb{R})$ .

Sei nun  $\varphi$  eine sogenannte Testfunktion aus dem Raum  $C_0^\infty(\Omega, \mathbb{R})$  mit

$$C_0^\infty(\Omega, \mathbb{R}) := \{\varphi \in C^\infty(\Omega, \mathbb{R}) \mid \text{supp}(\varphi) \text{ ist kompakt und } \text{supp}(\varphi) \subset \Omega\},$$

wobei

$$\text{supp}(\varphi) := \overline{\{x \in \mathbb{R} \mid \varphi(x) \neq 0\}}$$

den **Support von**  $\varphi$  bezeichnet. Für jede Testfunktion  $\varphi \in C_0^\infty(\Omega, \mathbb{R})$  gilt dann  $\varphi(a) = \varphi(b) = 0$  und mithilfe der partiellen Integration folgt die Beziehung

$$\int_{\Omega} f'(x)\varphi(x)dx = - \int_{\Omega} f(x)\varphi'(x)dx.$$

Eine analoge Beziehung gilt mitunter auch noch, wenn  $f'$  nur fast überall existiert. Betrachte dazu die Funktion  $f(x) = |x|$  in  $\Omega = (-1, 1)$ . Die Ableitung  $f'$  existiert für  $x \neq 0$  und lautet

$$f'(x) = \begin{cases} 1, & \text{falls } x > 0, \\ -1, & \text{falls } x < 0. \end{cases}$$

Wir setzen  $g(x) = f'(x)$  für  $x \neq 0$  und definieren  $g(0) = 0$ , so dass  $f' = g$  fast überall gilt. Dann gilt

$$\begin{aligned} \int_{\Omega} g(x)\varphi(x)dx &= \int_{(-1,0)} (-\varphi(x))dx + \int_{(0,1)} \varphi(x)dx \\ &= [-x\varphi(x)]_{-1}^0 - \int_{(-1,0)} (-x\varphi'(x))dx + [x\varphi(x)]_0^1 - \int_{(0,1)} x\varphi'(x)dx \\ &= - \int_{\Omega} |x|\varphi'(x)dx \\ &= - \int_{\Omega} f(x)\varphi'(x)dx \end{aligned}$$

für jede Funktion  $\varphi \in C_0^\infty(\Omega, \mathbb{R})$ . Die Funktion  $g$  nennen wir **schwache (oder verallgemeinerte) Ableitung von  $f$** . Dies lässt sich verallgemeinern:

**Definition 2.3.10** (schwache Ableitung, 1d-Fall)

Sei  $\Omega \subset \mathbb{R}$  offen und nichtleer,  $f \in L^1(\Omega)$  und  $k \in \mathbb{N}$ . Die Funktion  $g \in L^1(\Omega)$  heißt  **$k$ -te schwache (oder verallgemeinerte) Ableitung von  $f$** , falls

$$\int_{\Omega} g\varphi \, d\mu = (-1)^k \int_{\Omega} f\varphi^{(k)} \, d\mu \quad \forall \varphi \in C_0^\infty(\Omega, \mathbb{R})$$

gilt.  $g$  wird auch mit  $f^{(k)}$  bezeichnet.

**Bemerkung 2.3.11**

Die Definition einer schwachen Ableitung kann auch auf Gebiete  $\Omega \subset \mathbb{R}^d$  mit  $d > 1$  erweitert werden. Die Definition der schwachen Ableitung erfolgt dann mithilfe von Multiindizes, vgl. [Wer95, Def. V.1.10], und gilt sinngemäß für die partiellen Ableitungen.



Mithilfe der schwachen Ableitungen lassen sich neue Räume definieren, die tatsächlich Banachräume sind. Eine Verallgemeinerung auf mehrdimensionale Gebiete ist wiederum möglich, siehe [Bur98].

**Satz 2.3.12 (Sobolev-Räume, 1d-Fall)**

Sei  $\Omega \subset \mathbb{R}$  offen und nichtleer,  $q \in \mathbb{N}$  und  $1 \leq p \leq \infty$ . Die **Sobolevräume**  $(W^{q,p}(\Omega), \|f\|_{W^{q,p}(\Omega)})$ , welche durch

$$W^{q,p}(\Omega) := \{f \in L^p(\Omega) \mid \|f^{(k)}\|_{L^p(\Omega)} < \infty, k = 0, 1, \dots, q\}$$

und

$$\|f\|_{W^{q,p}(\Omega)} := \max_{k=0,1,\dots,q} \{\|f^{(k)}\|_{L^p(\Omega)}\}$$

definiert sind, sind Banachräume. Speziell für  $p = 2$  bezeichnet man diese Räume auch mit  $H^q(\Omega)$ .

Hierbei sind die Ableitungen  $f^{(k)}$  mit  $k \geq 1$  im schwachen Sinne zu verstehen.

## Kapitel 3

### Approximation in Hilberträumen

In diesem Abschnitt untersuchen wir Approximationsprobleme in sogenannten Hilberträumen. Hierbei sind ein Vektorraum  $H$  (über  $\mathbb{K} = \mathbb{R}$  oder  $\mathbb{K} = \mathbb{C}$ ) und eine (konvexe) Teilmenge  $M \subset H$  sowie ein Vektor  $h \in H$  gegeben. Gesucht ist ein Vektor  $\hat{m}$  aus  $M$ , der  $h$  möglichst gut approximiert. Derartige Aufgabenstellungen treten in vielen Anwendungen auf. Konkret behandeln wir Ausgleichsprobleme und die trigonometrische Approximation. Besonders elegant lassen sich Ausgleichsprobleme in Hilberträumen mithilfe des Projektionssatzes behandeln.

#### 3.1 Bestapproximierende und Approximationsaufgabe

Um beurteilen zu können, wie gut ein Vektor  $m \in M$  den Vektor  $h \in H$  approximiert, greifen wir auf den Normbegriff aus Kapitel 2 zurück und betrachten folgende Approximationsaufgabe:

##### Problem 3.1.1 (Approximationsproblem)

Gegeben seien der normierte Raum  $(H, \|\cdot\|_H)$ , die Teilmenge  $M \subset H$  sowie  $h \in H$ . Finde  $\hat{m} \in M$  mit

$$\|h - \hat{m}\|_H \leq \|h - m\|_H \quad \forall m \in M. \quad (3.1)$$

Jedes  $\hat{m} \in M$  mit der Eigenschaft (3.1) heißt **Bestapproximierende an  $h$  in  $M$** .

Die Approximationsaufgabe lässt sich also auch so formulieren:

*Finde eine Bestapproximierende an  $h$  in  $M$ .*

##### Definition 3.1.2 (Projektion)

Die Bestapproximierende  $\hat{m}$  an  $h$  in  $M$  bezeichnet man auch als **Projektion von  $h$  auf  $M$**  und schreibt

$$\hat{m} = P_M(h).$$

Durch  $P_M : H \rightarrow M$ ,  $h \mapsto P_M(h)$ , wird eine Abbildung definiert, welche auch als **Projektion auf  $M$**  bezeichnet wird.

**Beispiel 3.1.3 (Lineares Ausgleichsproblem)**

Die Messung eines physikalischen Vorgangs hat an den Ortspunkten  $p_i = (p_{1,i}, p_{2,i})$ ,  $i = 1, \dots, 9$ , die folgenden Messwerte  $y_i$ ,  $i = 1, \dots, 9$ , ergeben:

Messpunkt $i =$	1	2	3	4	5	6	7	8	9
$p_{1,i}$	6	7	9	11	13	15	17	18	19
$p_{2,i}$	0	1	2	3	4	5	6	7	8
$y_i$	96	189	283	373	467	553	647	733	832

Es wird ein affin-linearer Zusammenhang zwischen den Messpunkten und den Messwerten vermutet und folgender Ansatz gewählt

$$y(p_1, p_2) = x_1 + x_2 p_1 + x_3 p_2,$$

wobei  $x = (x_1, x_2, x_3)^\top \in \mathbb{R}^3$  unbekannt ist. Ein naheliegender Versuch,  $x$  zu bestimmen, besteht darin, das lineare Gleichungssystem

$$y_i = y(p_{1,i}, p_{2,i}) = x_1 + x_2 p_{1,i} + x_3 p_{2,i}, \quad i = 1, \dots, 9,$$

zu lösen. In Matrixschreibweise lautet dies  $Ax = b$  mit

$$b = \begin{bmatrix} 96 \\ 189 \\ 283 \\ 373 \\ 467 \\ 553 \\ 647 \\ 733 \\ 832 \end{bmatrix}, \quad A = \begin{bmatrix} 1 & 6 & 0 \\ 1 & 7 & 1 \\ 1 & 9 & 2 \\ 1 & 11 & 3 \\ 1 & 13 & 4 \\ 1 & 15 & 5 \\ 1 & 17 & 6 \\ 1 & 18 & 7 \\ 1 & 19 & 8 \end{bmatrix}.$$

Offenbar ist dieses lineare Gleichungssystem überbestimmt, da die Anzahl der Gleichungen die Anzahl der Unbekannten übersteigt. Leider stellt sich bei näherer Betrachtung heraus, dass das Gleichungssystem **keine Lösung** besitzt. Der Versuch, den Vektor  $x$  über die Lösung des Gleichungssystems  $Ax = b$  zu bestimmen, schlägt also fehl.

Ein sinnvoller Zugang ist es daher, den Vektor  $x$  so zu wählen, dass der Fehler im Gleichungssystem  $\|Ax - b\|_2^2$  möglichst klein wird. Dies führt auf das **Minimierungsproblem**

$$\|Ax - b\|_2^2 \longrightarrow \min.$$

Dieses Problem ist auch als **lineares Ausgleichsproblem** oder **Least-Squares-Problem** bekannt, welches bereits auf Gauß zurück geht.



JOHANN CARL FRIEDRICH GAUSS

Geboren: 30.4.1777 in Braunschweig (Deutschland)

Gestorben: 23.2.1855 in Göttingen (Deutschland)

Wie passt dieses Ausgleichsproblem in unsere Aufgabenstellung aus Problem 3.1.1?

Wir setzen  $H = \mathbb{R}^9$  und versehen  $H$  mit der Euklidischen Norm  $\|\cdot\|_2$ , sowie  $h = b$  und  $M = \{m \in \mathbb{R}^9 \mid m = Ax, x \in \mathbb{R}^3\} \subset H$ . Dann gilt es,  $\hat{m} \in M$  zu finden mit  $\|h - \hat{m}\|_2 \leq \|h - m\|_2$  für alle  $m \in M$ . ■

### Beispiel 3.1.4 (Trigonometrische Approximation)

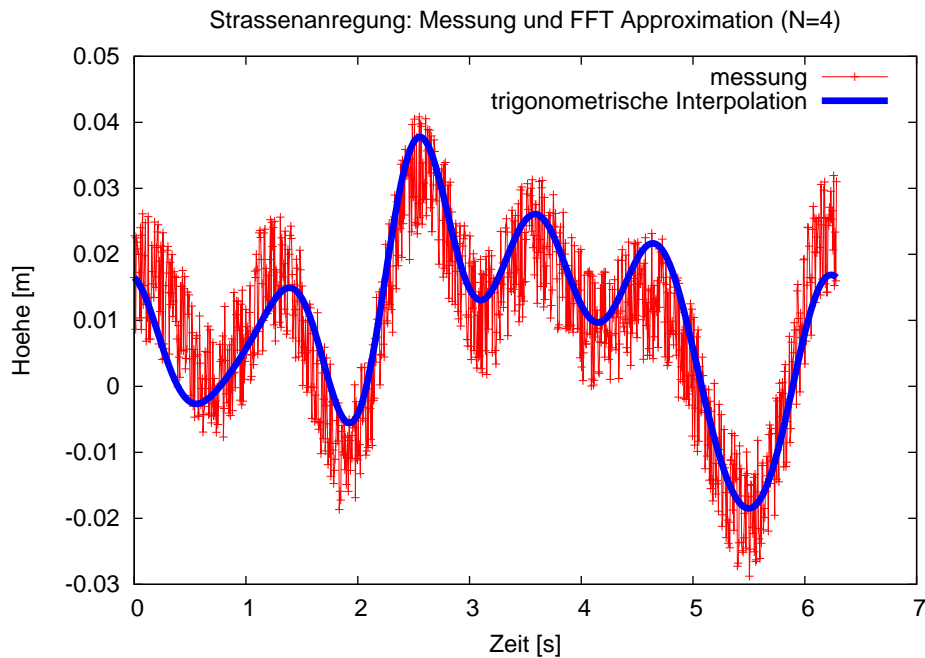
Gegeben sei eine  $T$ -periodische Funktion  $f \in L^2(\Omega)$  mit  $\Omega = [0, T]$ . Gesucht ist die Bestapproximierende an  $f$  in der Menge  $\mathcal{T}_n$  der sogenannten trigonometrischen Polynome vom Grad  $n$ , die definiert ist als

$$\mathcal{T}_n := \left\{ \frac{a_0}{2} + \sum_{k=1}^n a_k \cos(k\omega x) + b_k \sin(k\omega x) \mid \begin{array}{l} a_j \in \mathbb{R}, j = 0, \dots, n, \\ b_\ell \in \mathbb{R}, \ell = 1, \dots, n, \\ \omega = \frac{2\pi}{T} \end{array} \right\}.$$

Wir suchen also eine Funktion  $\hat{t}_n$  in  $\mathcal{T}_n$  mit

$$\|\hat{t}_n - f\|_{L^2(\Omega)} \leq \|t - f\|_{L^2(\Omega)} \quad \forall t \in \mathcal{T}_n.$$

Derartige Approximationen sind nützlich, um  $L^2$ -Funktionen, die sehr irregulär sein können, durch “angenehme” trigonometrische Polynome zu approximieren, siehe Abbildung.



■

### 3.2 Hilberträume und Projektionssatz

Während sich die Bestimmung einer Lösung von Problem 3.1.1 in einem allgemeinen normierten Raum als sehr schwierig herausstellt, erweisen sich Hilberträume als sehr geeignet für die Problemstellung und ermöglichen eine Charakterisierung der Lösung über den Projektionssatz. Hilberträume sind dabei spezielle Banachräume, auf denen ein Skalarprodukt definiert ist.

#### Definition 3.2.1 (Skalarprodukt, inneres Produkt)

Sei  $V$  ein Vektorraum über  $\mathbb{K} = \mathbb{R}$  oder  $\mathbb{K} = \mathbb{C}$ . Die Abbildung  $\langle \cdot, \cdot \rangle_V : V \times V \rightarrow \mathbb{K}$  heißt **inneres Produkt** oder **Skalarprodukt**, wenn die folgenden Bedingungen für alle  $x, y, z \in V$  und alle  $\lambda \in \mathbb{K}$  gelten:

$$(i) \quad \langle x, y \rangle_V = \overline{\langle y, x \rangle_V}$$

$$(ii) \quad \langle x + y, z \rangle_V = \langle x, z \rangle_V + \langle y, z \rangle_V$$

$$(iii) \quad \langle x, \lambda y \rangle_V = \lambda \langle x, y \rangle_V$$

$$(iv) \quad \langle x, x \rangle_V \geq 0 \text{ und } \langle x, x \rangle_V = 0, \text{ genau dann wenn } x = 0.$$

**Bemerkung 3.2.2**

- Jedes Skalarprodukt definiert eine Norm und eine Metrik auf  $V$  gemäß der Definition

$$\|x\|_V := \sqrt{\langle x, x \rangle_V}, \quad d(x, y) := \|x - y\|_V.$$

- Ein Skalarprodukt ermöglicht den Begriff der Orthogonalität. Zwei Vektoren  $x, y \in V$  heißen **orthogonal**, wenn  $\langle x, y \rangle_V = 0$  gilt.
- In einem Vektorraum  $V$  mit Skalarprodukt  $\langle \cdot, \cdot \rangle_V$  und induzierter Norm  $\|x\|_V = \sqrt{\langle x, x \rangle_V}$  gilt die **Parallelogrammgleichung**

$$\|x + y\|_V^2 + \|x - y\|_V^2 = 2\|x\|_V^2 + 2\|y\|_V^2.$$

Für orthogonale Vektoren  $x, y$  gilt der **Satz von Pythagoras**:

$$\|x + y\|_V^2 = \|x\|_V^2 + \|y\|_V^2.$$

■

**Definition 3.2.3** (Prä-Hilbertraum, Hilbertraum)

Ein mit einem Skalarprodukt versehener Vektorraum heißt **Prä-Hilbertraum** (oder **Innenproduktraum**). Ein vollständiger Prä-Hilbertraum heißt **Hilbertraum**.

**Beispiel 3.2.4** (Hilberträume und Prähilberträume)

- Der Raum  $\mathbb{R}^n$  mit dem üblichen Skalarprodukt  $\langle x, y \rangle_{\mathbb{R}^n} = x^\top y$  ist ein Hilbertraum.
- Der Raum  $\mathbb{C}^n$  mit dem Skalarprodukt  $\langle x, y \rangle_{\mathbb{C}^n} = x^* y$  mit  $x^* = \bar{x}^\top$  ist ein Hilbertraum.
- Der Folgenraum  $\ell^2$  versehen mit dem Skalarprodukt

$$\langle u, v \rangle_{\ell^2} := \sum_{k=1}^{\infty} u_k v_k$$

ist ein Hilbertraum.



(d) Der Raum  $L^2(\Omega)$  der auf dem Gebiet  $\Omega \subset \mathbb{R}^d$  quadratisch (Lebesgue-)integrierbaren Funktionen versehen mit dem Skalarprodukt

$$\langle u, v \rangle_{L^2(\Omega)} := \int_{\Omega} u(x)v(x)dx \quad (3.2)$$

ist ein Hilbertraum. In diesem Fall lautet die Hölder-Ungleichung

$$|\langle u, v \rangle_{L^2(\Omega)}| \leq \|u\|_{L^2(\Omega)} \cdot \|v\|_{L^2(\Omega)}$$

und wird **Cauchy-Schwarz-Ungleichung** genannt. Um zu zeigen, dass  $L^2(\Omega)$  tatsächlich ein Hilbertraum ist, muss lediglich nachgewiesen werden, dass (3.2) tatsächlich ein Skalarprodukt ist. Denn wegen

$$\langle u, u \rangle_{L^2(\Omega)} := \int_{\Omega} u(x)^2 dx = \|u\|_{L^2(\Omega)}^2$$

wissen wir bereits, dass  $L^2(\Omega)$  versehen mit der  $L^2$ -Norm vollständig ist.

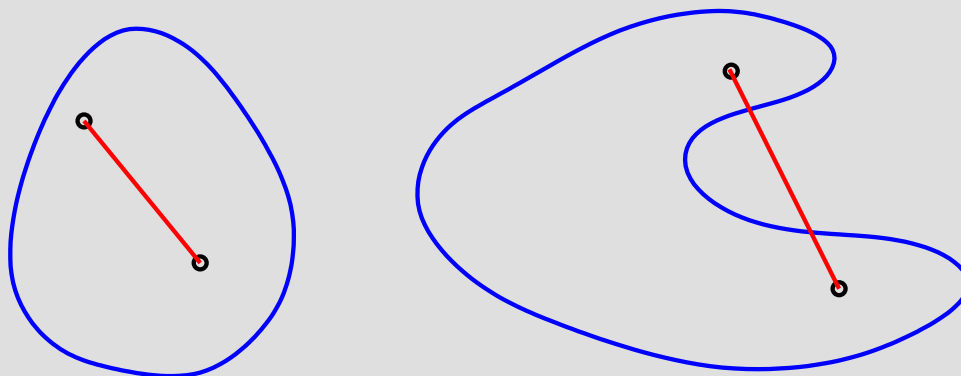
(e) Der Raum  $C([a, b], \mathbb{R})$  der auf  $\Omega = [a, b]$  stetigen Funktion versehen mit dem Skalarprodukt (3.2) ist ein Prä-Hilbertraum, aber kein Hilbertraum, da er bzgl. der durch das Skalarprodukt induzierten  $L^2$ -Norm nicht vollständig ist. ■

Wir wenden uns nun wieder der Aufgabe zu, eine Bestapproximierende  $\hat{m} \in M$  an  $h \in H$  zu bestimmen. Zunächst betrachten wir den Fall einer konvexen Menge  $M$ . Später werden wir dann den Spezialfall eines Unterraums daraus ableiten.

### Definition 3.2.5 (konvexe Menge)

Eine Menge  $M$  heißt **konvex**, falls mit je zwei Punkten aus  $M$  auch ihre gesamte Verbindungsstrecke in  $M$  liegt, d.h.

$$\forall m_1, m_2 \in M \quad \forall \lambda \in [0, 1] : \lambda m_1 + (1 - \lambda)m_2 \in M.$$



Konvexe (links) und nicht konvexe Menge (rechts)

Aus der geometrischen Anschauung ist klar, dass  $\hat{m}$  genau dann Bestapproximierende an  $h$  in  $M$  ist, wenn der Winkel zwischen  $h - \hat{m}$  und  $m - \hat{m}$  mindestens 90 Grad beträgt, vgl. Abbildung 3.1.

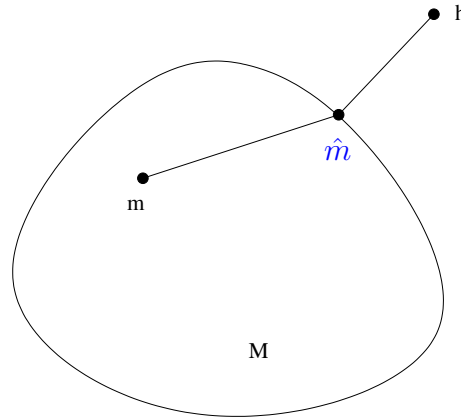


Abbildung 3.1: Bestapproximierende  $\hat{m}$  an  $h$  in  $M$ .

Dies besagt der folgende Satz.

**Satz 3.2.6 (Projektionssatz für konvexe Mengen)**

Seien  $H$  ein Hilbertraum über  $\mathbb{R}$  mit Skalarprodukt  $\langle \cdot, \cdot \rangle_H$ ,  $h \in H$  und  $M \subset H$  eine konvexe und abgeschlossene Teilmenge,  $M \neq \emptyset$ . Dann gelten:

- (i) Es existiert genau eine Bestapproximierende  $\hat{m}$  an  $h$  in  $M$ .
- (ii)  $\hat{m} \in M$  ist Bestapproximierende an  $h$  in  $M$  genau dann, wenn

$$\langle h - \hat{m}, m - \hat{m} \rangle_H \leq 0 \quad \forall m \in M. \quad (3.3)$$

**Beweis:** Wir zeigen zunächst (ii) und die Eindeutigkeitsaussage in (i).

$\Rightarrow$ : Seien  $m \in M$  und  $\lambda \in (0, 1]$  beliebig. Da  $M$  konvex ist, gilt  $w = (1 - \lambda)\hat{m} + \lambda m \in M$ . Da  $\hat{m}$  Bestapproximierende an  $h$  in  $M$  ist, folgt

$$\begin{aligned} \|h - \hat{m}\|_H^2 &\leq \|h - w\|_H^2 \\ &= \langle h - w, h - w \rangle_H \\ &= \langle h - (1 - \lambda)\hat{m} - \lambda m, h - (1 - \lambda)\hat{m} - \lambda m \rangle_H \\ &= \langle h - \hat{m} + \lambda(\hat{m} - m), h - \hat{m} + \lambda(\hat{m} - m) \rangle_H \\ &= \|h - \hat{m}\|_H^2 + 2\lambda \langle h - \hat{m}, \hat{m} - m \rangle_H + \lambda^2 \|\hat{m} - m\|_H^2. \end{aligned}$$

Division durch  $2\lambda > 0$  liefert

$$0 \leq \langle h - \hat{m}, \hat{m} - m \rangle_H + \frac{\lambda}{2} \|\hat{m} - m\|_H^2.$$

Für  $\lambda \rightarrow 0$  folgt daraus die Behauptung.

$\Leftarrow$ : Es gelte (3.3). Dann gilt für alle  $m \in M$  mit  $m \neq \hat{m}$ ,

$$\begin{aligned} \|h - m\|_H^2 &= \langle h - \hat{m} + \hat{m} - m, h - \hat{m} + \hat{m} - m \rangle_H \\ &= \|h - \hat{m}\|_H^2 + \langle h - \hat{m}, \hat{m} - m \rangle_H + \langle \hat{m} - m, h - \hat{m} \rangle_H + \|\hat{m} - m\|_H^2 \\ &\geq \|h - \hat{m}\|_H^2 + \|\hat{m} - m\|_H^2 \\ &> \|h - \hat{m}\|_H^2. \end{aligned}$$

Damit ist  $\hat{m}$  Bestapproximierende. Die Eindeutigkeit der Bestapproximierenden in (i) ist hiermit ebenfalls gezeigt, da obige Abschätzung auf einen Widerspruch führt, wenn man annimmt, dass es zwei unterschiedliche Bestapproximierende gibt.

Nun zeigen wir die Existenz einer Bestapproximierenden. Setze

$$\delta := \inf_{y \in M} \|h - y\|_H.$$

Dann existiert eine Folge  $\{y_k\}_{k \in \mathbb{N}} \subset M$  mit  $\|h - y_k\|_H \rightarrow \delta$  für  $k \rightarrow \infty$ . Mit der Parallelogrammgleichung folgt

$$\begin{aligned} \left\| \frac{y_n - y_m}{2} \right\|_H^2 &= \left\| \frac{(y_n - h) - (y_m - h)}{2} \right\|_H^2 \\ &= \underbrace{\frac{\|y_n - h\|_H^2 + \|y_m - h\|_H^2}{2}}_{\rightarrow \delta^2} - \underbrace{\left\| \frac{y_n + y_m}{2} - h \right\|_H^2}_{\geq \delta^2} \rightarrow 0. \end{aligned}$$

Beachte, dass  $(y_n + y_m)/2$  in  $M$  liegt. Damit ist  $\{y_k\}_{k \in \mathbb{N}}$  Cauchyfolge in  $H$  und, da  $H$  vollständig ist, existiert der Grenzwert  $\hat{y} = \lim_{k \rightarrow \infty} y_k$ . Da  $M$  abgeschlossen ist, gilt  $\hat{y} \in M$  und  $\delta = \|h - \hat{y}\|_H$ .  $\blacksquare$

Die Projektionsabbildung  $P_M : H \rightarrow M$ ,  $h \mapsto \hat{m} = P_M(h)$ , ist unter den Voraussetzungen des Satzes 3.2.6 Lipschitz-stetig.

### Satz 3.2.7

Sei  $H$  ein Hilbertraum und  $M \subset H$  eine abgeschlossene, konvexe, nichtleere Menge. Die Projektion  $P_M : H \rightarrow M$  ist Lipschitz-stetig mit Konstante 1, d.h. es gilt

$$\|P_M(h_1) - P_M(h_2)\|_H \leq \|h_1 - h_2\|_H \quad \forall h_1, h_2 \in H.$$

**Beweis:** Gemäß Projektionssatz existiert für jedes  $h \in H$  die eindeutig definierte Projektion  $P_M(h) \in M$  mit

$$\langle h - P_M(h), m - P_M(h) \rangle_H \leq 0 \quad \forall m \in M.$$

Für beliebige  $h_1, h_2 \in H$  folgt

$$\langle h_1 - P_M(h_1), P_M(h_2) - P_M(h_1) \rangle_H \leq 0, \quad \langle h_2 - P_M(h_2), P_M(h_1) - P_M(h_2) \rangle_H \leq 0$$

und somit

$$\begin{aligned} \langle P_M(h_1), P_M(h_1) - P_M(h_2) \rangle_H &\leq \langle h_1, P_M(h_1) - P_M(h_2) \rangle_H, \\ -\langle P_M(h_2), P_M(h_1) - P_M(h_2) \rangle_H &\leq -\langle h_2, P_M(h_1) - P_M(h_2) \rangle_H. \end{aligned}$$

Schließlich folgt mit der Cauchy-Schwarz-Ungleichung

$$\begin{aligned} \|P_M(h_1) - P_M(h_2)\|_H^2 &= \langle P_M(h_1) - P_M(h_2), P_M(h_1) - P_M(h_2) \rangle_H \\ &= \langle P_M(h_1), P_M(h_1) - P_M(h_2) \rangle_H - \langle P_M(h_2), P_M(h_1) - P_M(h_2) \rangle_H \\ &\leq \langle h_1, P_M(h_1) - P_M(h_2) \rangle_H - \langle h_2, P_M(h_1) - P_M(h_2) \rangle_H \\ &= \langle h_1 - h_2, P_M(h_1) - P_M(h_2) \rangle_H \\ &\leq \|h_1 - h_2\|_H \cdot \|P_M(h_1) - P_M(h_2)\|_H. \end{aligned}$$

■

Wir betrachten noch einen Spezialfall von Satz 3.2.6, der insbesondere für das lineare Ausgleichsproblem und die trigonometrische Approximation nützlich ist. Bei diesen Beispielen ist die Menge  $M$  ein abgeschlossener Untervektorraum von  $H$  (also auch konvex). Im Fall eines Untervektorraums ist aus der geometrischen Anschauung klar, dass  $\hat{m}$  genau dann Bestapproximierende an  $h$  in  $M$  ist, wenn  $h - \hat{m}$  senkrecht auf  $M$  steht, vgl. Abbildung 3.2.

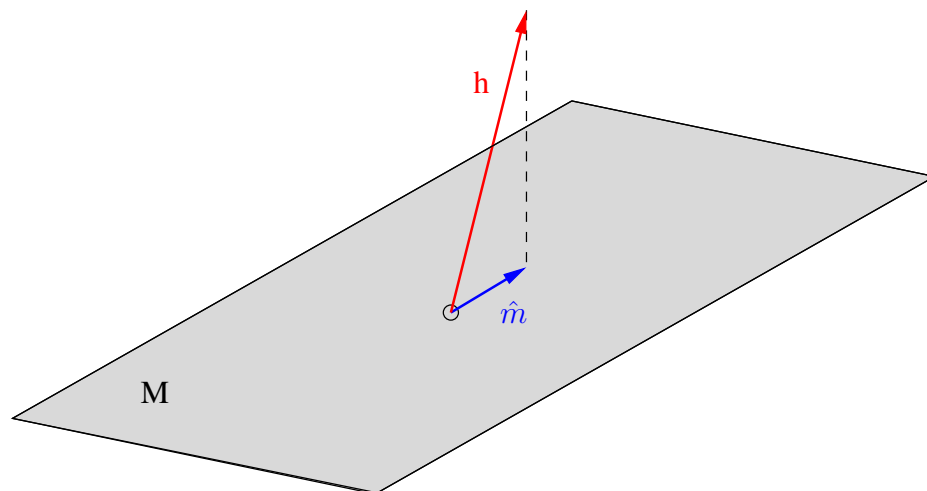
Dies besagt der folgende Satz, der für Hilberträume über  $\mathbb{R}$  direkt aus Satz 3.2.6 folgt, wenn man beachtet, dass in einem Untervektorraum mit  $m - \hat{m} \in M$  auch  $-(m - \hat{m}) \in M$  enthalten ist. Damit wird aus der Ungleichung (3.3) eine Gleichung.

### Satz 3.2.8 (Projektionssatz)

Sei  $H$  ein Hilbertraum mit Skalarprodukt  $\langle \cdot, \cdot \rangle_H$ . Sei  $h \in H$  und  $M \subset H$  ein abgeschlossener Untervektorraum von  $H$ . Dann gelten:

- (i) Es existiert genau eine Bestapproximierende  $\hat{m}$  an  $h$  in  $M$ .
- (ii)  $\hat{m} \in M$  ist Bestapproximierende an  $h$  in  $M$  genau dann, wenn

$$\langle h - \hat{m}, m \rangle_H = 0 \quad \forall m \in M. \quad (3.4)$$

Abbildung 3.2: Bestapproximierende  $\hat{m}$  an  $h$  in  $M$ .

### 3.3 Anwendung des Projektionssatzes auf lineare Ausgleichsprobleme

Wir betrachten das folgende Ausgleichsproblem, vgl. auch Beispiel 3.1.3.

#### Problem 3.3.1 (Lineares Ausgleichsproblem)

Gegeben seien eine Matrix  $A \in \mathbb{R}^{m \times n}$  und ein Vektor  $b \in \mathbb{R}^m$  mit  $m, n \in \mathbb{N}$ . Bestimme die Lösung  $x = [x_1, \dots, x_n]^\top \in \mathbb{R}^n$  des Minimierungsproblems

$$\text{Minimiere} \quad \frac{1}{2} \|Ax - b\|_2^2 \quad (3.5)$$

#### Bemerkung 3.3.2

- Im Fall  $m = n$  und  $A$  invertierbar, ist die Lösung eindeutig durch das lineare Gleichungssystem  $Ax = b$  bestimmt.
- Interessant ist insbesondere der Fall  $m > n$ , der in praktischen Anwendungen besonders häufig auftritt, zumeist ist  $m$  dort wesentlich grösser als  $n$ .
- Der Fall  $m < n$  kann ebenfalls auftreten. In diesem Fall ist das Gleichungssystem  $Ax = b$  unterbestimmt, es kann jedoch unlösbar sein, wenn  $\text{Rang}(A) \neq \text{Rang}(A|b)$ .

Wir formulieren das lineare Ausgleichsproblem in ein Approximationsproblem 3.1.1 um und definieren

$$\begin{aligned}(H, \|\cdot\|_H) &:= (\mathbb{R}^m, \|\cdot\|_2), \\ M &:= \{m \in \mathbb{R}^m \mid m = Ax, x \in \mathbb{R}^n\} = \text{Bild}(A), \\ h &:= b.\end{aligned}$$

Beachte, dass  $M$  ein abgeschlossener Untervektorraum von  $H$  mit Dimension  $\dim(M) = \text{Rang}(A)$  ist. Die Norm  $\|\cdot\|_2$  wird durch das Skalarprodukt  $\langle x, y \rangle_{\mathbb{R}^m} = x^\top y$  gemäß  $\|x\|_2 = \sqrt{\langle x, x \rangle_{\mathbb{R}^m}}$  induziert. Das lineare Ausgleichsproblem ist damit äquivalent zum folgenden Approximationsproblem:

Finde  $\hat{m} \in M$ , so dass

$$\|h - \hat{m}\|_2 \leq \|h - m\|_2 \quad \forall m \in M. \quad (3.6)$$

Wir suchen also die Bestapproximierende  $\hat{m}$  an  $h$  in  $M$ . Der Projektionssatz 3.2.8 besagt, dass  $\hat{m} = A\hat{x}$  genau dann Bestapproximierende an  $h = b$  in  $M$  ist, wenn die Bedingung

$$\begin{aligned}0 &= \langle h - \hat{m}, m \rangle_{\mathbb{R}^m} \\ &= \langle h - A\hat{x}, Ax \rangle_{\mathbb{R}^m} \\ &= (h - A\hat{x})^\top Ax \\ &= (A^\top b - A^\top A\hat{x})^\top x\end{aligned}$$

für alle  $m = Ax, x \in \mathbb{R}^n$ , gilt. Damit haben wir folgenden Satz bewiesen:

**Satz 3.3.3 (Gauß'sche Normalgleichungen)**

$\hat{x} \in \mathbb{R}^n$  löst das lineare Ausgleichsproblem genau dann, wenn die **Gauß'schen Normalgleichungen**

$$A^\top A\hat{x} = A^\top b$$

gelten.

Die Gauß'schen Normalgleichungen charakterisieren also die Lösung des linearen Ausgleichsproblems. Man kann noch zeigen, dass die Normalgleichungen (bzw. das lineare Ausgleichsproblem) eine eindeutige Lösung besitzen, falls  $\text{Rang}(A) = n$  gilt.

**Beispiel 3.3.4 (vgl. Beispiel 3.1.3)**

Wir betrachten wiederum Beispiel 3.1.3 und lösen das lineare Ausgleichsproblem (3.5) mit

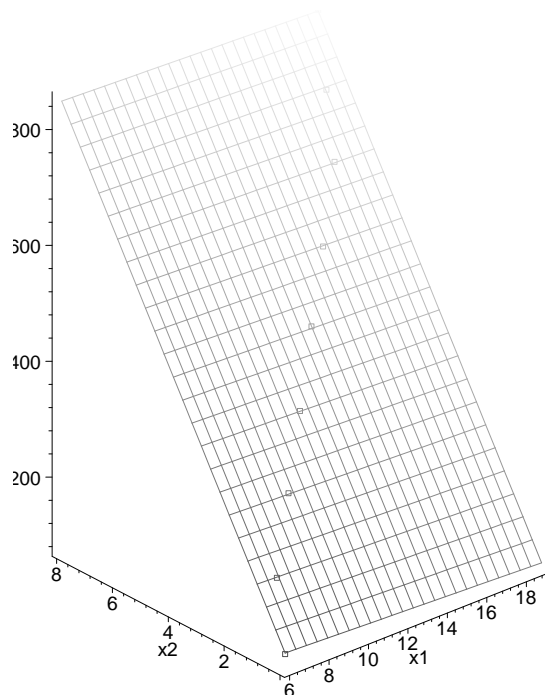
Hilfe der Normalgleichungen  $A^T Ax = A^T b$ . Es gilt

$$A^T A = \begin{pmatrix} 9 & 115 & 36 \\ 115 & 1655 & 565 \\ 36 & 565 & 204 \end{pmatrix}, \quad A^T b = \begin{pmatrix} 4173 \\ 62916 \\ 22176 \end{pmatrix}, \quad x = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}.$$

Die Lösung der Normalgleichungen lautet

$$x_1 = \frac{533}{5} = 106.6, \quad x_2 = -\frac{96}{65} \approx -1.476923077, \quad x_3 = \frac{6109}{65} \approx 93.98461538.$$

Grafisch ergibt sich folgendes:



### 3.4 Anwendung des Projektionssatzes bei der Fourieranalyse

Gegenstand dieses Kapitels ist die Approximation von periodischen Funktionen durch trigonometrische Funktionen. Periodische Funktionen treten in vielen Anwendungen auf, z.B. in der Signal- und Bildverarbeitung (jpeg/mpeg Kompression). Darüber hinaus gibt es auch periodische physikalische oder technische Vorgänge, etwa Schwingungen. Häufig liegen für derartige Prozesse lediglich Messwerte vor oder die beschreibenden Funktionen sind zu kompliziert und erlauben keine analytische Darstellung. Daher versucht man, diese Vorgänge durch periodische Funktionen zu approximieren und verwendet sogenannte trigonometrische Polynome, die aus einer Überlagerung von Sinus und Cosinus entstehen.

#### Beispiel 3.4.1

An der Kurbelwelle eines Einzylinder-Zweitakt-Motors wirkt eine Drehkraft  $f$ , die vom

Gasdruck und von der Massenträgheit herrührt. An den Stellen  $t_k = k \cdot \pi/8$ ,  $k = 0, 1, \dots, 15$ , wurden folgende Werte von  $f_k$  in [N] experimentell bestimmt:

$f_0$	$f_1$	$f_2$	$f_3$	$f_4$	$f_5$	$f_6$	$f_7$
-8250	-29430	-2286	5974	-8829	-25408	-22681	-28655
$f_8$	$f_9$	$f_{10}$	$f_{11}$	$f_{12}$	$f_{13}$	$f_{14}$	$f_{15}$
-8564	96560	45862	22092	-9025	-23514	-15127	12880

Zur Approximation der Drehkraft  $f$  wird die periodische Funktion

$$g(x) = a_0 + \sum_{j=1}^{\mu-1} (A_j \cos(jx) + B_j \sin(jx)) + a_\mu \cos(\mu x)$$

mit  $\mu = 8$  und den folgenden Koeffizienten verwendet:

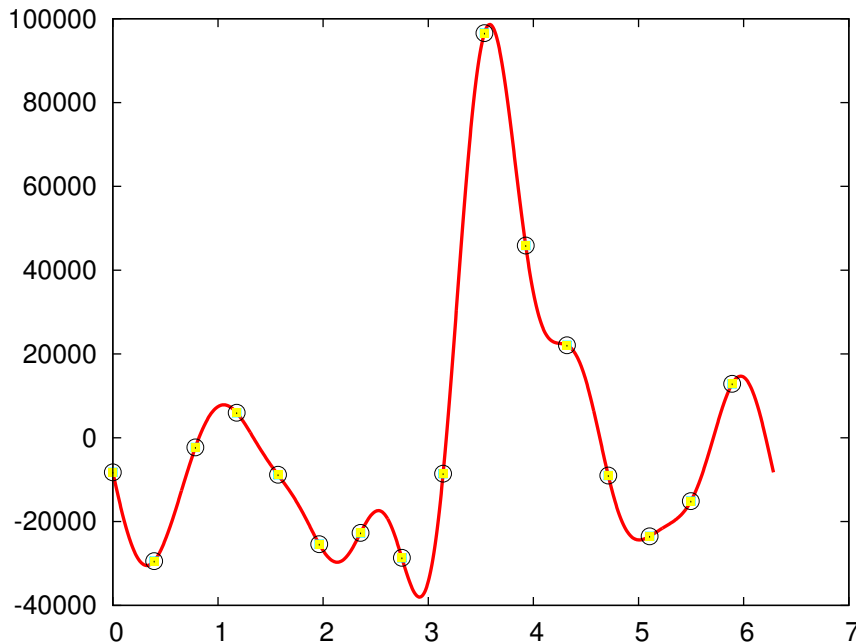
KOEFFIZIENTEN A:

0.999E+02 -0.140E+05 0.651E+04 0.123E+04 -0.505E+04 0.602E+04 -0.625E+04  
0.688E+04 -0.371E+04

KOEFFIZIENTEN B:

-0.150E+05 0.243E+05 -0.234E+05 0.740E+03 -0.135E+05 0.396E+04 -0.519E+04

Die resultierende Approximation der Drehkraft  $f$  sieht wie folgt aus:



■



**Definition 3.4.2** (periodische Funktion)

Die Funktion  $f : \mathbb{R} \rightarrow \mathbb{C}$  heißt **T-periodisch**, wenn es ein  $T > 0$  gibt mit

$$f(x + T) = f(x) \quad \text{für alle } x \in \mathbb{R}.$$

$T$  heißt **Periode** oder **Periodenlänge** von  $f$ .

Beachte, dass eine  $T$ -periodische Funktion  $f$  bereits vollständig durch die Werte im Intervall  $[0, T]$  festgelegt ist. Daher genügt es auch,  $f$  nur auf dem Intervall  $[0, T]$  zu betrachten. In den einführenden Mathematikveranstaltungen haben wir bereits die  $2\pi$ -periodischen Funktionen Sinus und Cosinus sowie die  $\pi$ -periodische Funktion Tangens kennen gelernt. Natürlich gibt es auch kompliziertere periodische Funktionen, wie z.B. die Bewegung eines Satelliten auf einer Umlaufbahn um die Erde oder die Bewegung der Erde auf einer Sonnenumlaufbahn.

Durch eine lineare Transformation kann man jede  $T$ -periodische Funktion auf eine  $2\pi$ -periodische Funktion transformieren (und umgekehrt).

**Hilfsatz 3.4.3**

Sei  $f$   $T$ -periodisch mit Periode  $T > 0$ . Dann ist die Funktion  $\tilde{f} : \mathbb{R} \rightarrow \mathbb{C}$  definiert durch

$$\tilde{f}(x) := f\left(\frac{T}{2\pi}x\right)$$

$2\pi$ -periodisch.

**Beweis:** Es gilt

$$\tilde{f}(x + 2\pi) = f\left(\frac{T}{2\pi}(x + 2\pi)\right) = f\left(\frac{T}{2\pi}x + T\right) = f\left(\frac{T}{2\pi}x\right) = \tilde{f}(x).$$

■

Wie in Beispiel 3.1.4 betrachten wir folgende Approximationsaufgabe:

**Problem 3.4.4** (Fourierapproximation)

Sei  $f \in L^2(\Omega)$  mit  $\Omega = [0, T]$  eine gegebene  $T$ -periodische Funktion. Gesucht ist die Bestapproximierende  $\hat{t}_n$  an  $f$  mit

$$\|\hat{t}_n - f\|_{L^2(\Omega)} \leq \|t - f\|_{L^2(\Omega)} \quad \forall t \in \mathcal{T}_n,$$

wobei  $\mathcal{T}_n$  für  $n \in \mathbb{N}_0$  die Menge aller trigonometrischen Polynome vom Grad  $n$  der Form

$$t(x) = \frac{a_0}{2} + \sum_{k=1}^n a_k \cos(k\omega x) + b_k \sin(k\omega x) \quad \text{mit} \quad \omega := \frac{2\pi}{T}$$

bezeichnet.

Trigonometrischen Polynome kann man interpretieren als Überlagerung von Sinus- und Cosinus-Schwingungen verschiedener Frequenzen  $k\omega$ , wobei die **Fourierkoeffizienten**  $a_0, \dots, a_n$  und  $b_1, \dots, b_n$  die Amplituden der Schwingungen angeben. Beachte, dass  $\cos(k\omega x)$  und  $\sin(k\omega x)$   $\frac{T}{k}$ -periodisch sind.

Wir berechnen nun die Fourierkoeffizienten mithilfe des Projektionssatzes. Beachte, dass  $\mathcal{T}_n$  ein abgeschlossener Untervektorraum von  $L^2(\Omega)$  ist. Der Projektionssatz 3.2.8 besagt, dass

$$\hat{t}_n(x) = \frac{\hat{a}_0}{2} + \sum_{k=1}^n \hat{a}_k \cos(k\omega x) + \hat{b}_k \sin(k\omega x) \quad \text{mit} \quad \omega := \frac{2\pi}{T}$$

genau dann Bestapproximierende an  $f$  ist, wenn

$$0 = \langle f - \hat{t}_n, t \rangle_{L^2(\Omega)}$$

für alle  $t \in \mathcal{T}_n$  gilt. Da sich jedes  $t \in \mathcal{T}_n$  als Linearkombination von den Funktionen  $c_\ell(x) = \cos(\ell\omega x)$ ,  $\ell = 0, \dots, n$ , und  $s_\ell(x) = \sin(\ell\omega x)$ ,  $\ell = 1, \dots, n$ , darstellen lässt, genügt es, diese Gleichung für eben diese Basisfunktionen zu betrachten: Für  $\ell = 0, \dots, n$  folgt

$$\begin{aligned} 0 &= \langle f - \hat{t}_n, c_\ell \rangle_{L^2(\Omega)} \\ &= \int_0^T \left( f(x) - \frac{\hat{a}_0}{2} - \sum_{k=1}^n \hat{a}_k \cos(k\omega x) + \hat{b}_k \sin(k\omega x) \right) \cos(\ell\omega x) dx \\ &= \int_0^T f(x) \cos(\ell\omega x) dx - \frac{\hat{a}_0}{2} \underbrace{\int_0^T \cos(\ell\omega x) dx}_{= \begin{cases} T, & \text{falls } \ell = 0, \\ 0, & \text{sonst} \end{cases}} \\ &\quad - \sum_{k=1}^n \hat{a}_k \underbrace{\int_0^T \cos(k\omega x) \cos(\ell\omega x) dx}_{= \begin{cases} T/2, & \text{falls } \ell = k, \\ 0, & \text{sonst} \end{cases}} - \sum_{k=1}^n \hat{b}_k \underbrace{\int_0^T \sin(k\omega x) \cos(\ell\omega x) dx}_{=0} \\ &= \int_0^T f(x) \cos(\ell\omega x) dx - \frac{T}{2} \hat{a}_\ell \end{aligned}$$

und für  $\ell = 1, \dots, n$  analog

$$\begin{aligned}
 0 &= \langle f - \hat{t}_n, s_\ell \rangle_{L^2(\Omega)} \\
 &= \int_0^T \left( f(x) - \frac{\hat{a}_0}{2} - \sum_{k=1}^n \hat{a}_k \cos(k\omega x) + \hat{b}_k \sin(k\omega x) \right) \sin(\ell\omega x) dx \\
 &= \int_0^T f(x) \sin(\ell\omega x) dx - \underbrace{\frac{\hat{a}_0}{2} \int_0^T \sin(\ell\omega x) dx}_{=0} \\
 &\quad - \sum_{k=1}^n \hat{a}_k \underbrace{\int_0^T \cos(k\omega x) \sin(\ell\omega x) dx}_{=0} - \sum_{k=1}^n \hat{b}_k \underbrace{\int_0^T \sin(k\omega x) \sin(\ell\omega x) dx}_{= \begin{cases} T/2, & \text{falls } \ell = k, \\ 0, & \text{sonst} \end{cases}} \\
 &= \int_0^T f(x) \sin(\ell\omega x) dx - \frac{T}{2} \hat{b}_\ell.
 \end{aligned}$$

Zusammenfassend haben wir gezeigt:

### Satz 3.4.5 (Fourier)

*Das trigonometrische Polynom*

$$\hat{t}_n(x) = \frac{\hat{a}_0}{2} + \sum_{k=1}^n \hat{a}_k \cos(k\omega x) + \hat{b}_k \sin(k\omega x) \quad \text{mit} \quad \omega = \frac{2\pi}{T}$$

*mit den Fourierkoeffizienten*

$$\begin{aligned}
 \hat{a}_0 &= \frac{2}{T} \int_0^T f(x) dx, \\
 \hat{a}_k &= \frac{2}{T} \int_0^T f(x) \cos(k\omega x) dx, \\
 \hat{b}_k &= \frac{2}{T} \int_0^T f(x) \sin(k\omega x) dx,
 \end{aligned}$$

$k = 1, \dots, n$ , ist die eindeutig bestimmte Lösung des Approximationsproblems 3.4.4.

### Beispiel 3.4.6

*Die  $2\pi$ -periodische Sägezahnfunktion*

$$f(x) = x - \pi, \quad x \in [0, 2\pi],$$

*soll im Intervall  $[0, 2\pi]$  in einer Fouriersumme entwickelt werden.*

Die Fourierkoeffizienten lauten:

$$\begin{aligned}
 a_0 &= \frac{1}{\pi} \int_0^{2\pi} (x - \pi) dx = \frac{1}{2\pi} [(x - \pi)^2]_0^{2\pi} = 0, \\
 a_k &= \frac{1}{\pi} \int_0^{2\pi} (x - \pi) \cos(kx) dx \\
 &= \frac{1}{\pi} \left( \left[ (x - \pi) \frac{\sin(kx)}{k} \right]_0^{2\pi} - \int_0^{2\pi} \frac{\sin(kx)}{k} dx \right) \\
 &= 0, \quad (k \in \mathbb{N}) \\
 b_k &= \frac{1}{\pi} \int_0^{2\pi} (x - \pi) \sin(kx) dx \\
 &= \frac{1}{\pi} \left( \left[ -(x - \pi) \frac{\cos(kx)}{k} \right]_0^{2\pi} + \int_0^{2\pi} \frac{\cos(kx)}{k} dx \right) \\
 &= \frac{1}{\pi} \left( -2 \frac{\pi}{k} \right) = -\frac{2}{k} \quad (k \in \mathbb{N}).
 \end{aligned}$$

Als Lösung erhält man somit

$$t_n(x) = -2 \sum_{k=1}^n \frac{\sin(kx)}{k}.$$

Abbildung 3.3 zeigt die Lösungen für  $n = 1, 2, 3, 4, 20$ .

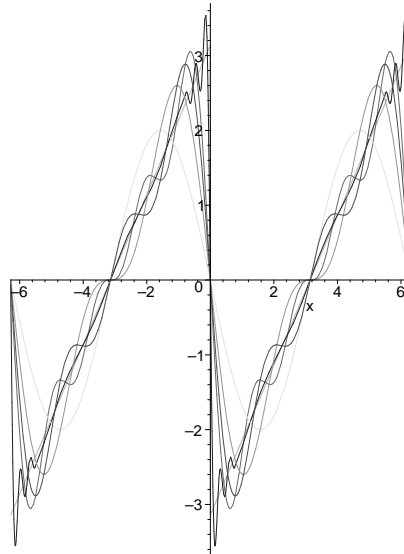


Abbildung 3.3: Approximation der Sägezahnfunktion durch trigonometrische Polynome vom Grad 1, 2, 3, 4 und 20.

Typisch bei der Approximation mit trigonometrischen Funktionen sind die Über- und Unterschwingen an Unstetigkeitsstellen von  $f$ , die sich auch durch Erhöhung des Polynomgrades nicht vermeiden lassen. Dieses Phänomen ist als **Gibbs-Phänomen** bekannt und

man beobachtet, dass die Höhe der Über- und Unterschwingung für hinreichend grosses  $n$  in etwa 18 % Prozent der Sprunghöhe beträgt. ■

### Bemerkung 3.4.7

Bei der Berechnung der Fourierkoeffizienten lassen sich Symmetrien ausnutzen:  
Ist  $f$  gerade, d.h. es gilt  $f(x) = f(-x)$  für alle  $x \in \mathbb{R}$ , so folgt  $b_k = 0$  für  $k \in \mathbb{N}$ .  
Ist  $f$  ungerade, d.h. es gilt  $f(x) = -f(-x)$  für alle  $x \in \mathbb{R}$ , so folgt  $a_k = 0$  für  $k \in \mathbb{N}_0$ . ■

### 3.4.1 Komplexe Darstellung

Die Darstellung der trigonometrischen Polynome ist auch als reelle Darstellung bekannt, auch wenn die Fourierkoeffizienten im Fall komplexwertiger Funktionen durchaus auch komplexe Werte annehmen können. Es gibt jedoch auch eine äquivalent komplexe Darstellung, die auf der komplexen Exponentialfunktion (**Euler-Formel**)

$$\exp(i\omega x) := \cos(\omega x) + i \sin(\omega x) \quad \forall x \in \mathbb{R} \quad (3.7)$$

mit  $\omega = 2\pi/T$  basiert, wobei hier und im Folgenden  $i$  für die komplexe Zahl  $i \in \mathbb{C}$  mit  $i^2 = -1$  steht. Mit Hilfe der Euler-Formel lassen sich Cosinus und Sinus für  $x \in \mathbb{R}$  ausdrücken als

$$\cos(\omega x) = \frac{1}{2} (\exp(i\omega x) + \exp(-i\omega x)), \quad \sin(\omega x) = \frac{1}{2i} (\exp(i\omega x) - \exp(-i\omega x)). \quad (3.8)$$

Mit Hilfe der Euler-Formel (3.7) folgt

$$\begin{aligned} & \frac{a_0}{2} + \sum_{k=1}^n (a_k \cos(k\omega x) + b_k \sin(k\omega x)) \\ &= \frac{a_0}{2} + \sum_{k=1}^n \left( \frac{a_k}{2} (\exp(ik\omega x) + \exp(-ik\omega x)) + \frac{b_k}{2i} (\exp(ik\omega x) - \exp(-ik\omega x)) \right) \\ &= \frac{a_0}{2} + \sum_{k=1}^n \left( \frac{1}{2} (a_k - ib_k) \exp(ik\omega x) + \frac{1}{2} (a_k + ib_k) \exp(-ik\omega x) \right) \\ &= \underbrace{\frac{a_0}{2}}_{=: \alpha_0} + \sum_{k=1}^n \underbrace{\frac{1}{2} (a_k - ib_k)}_{=: \alpha_k, k=1, \dots, n} \exp(ik\omega x) + \sum_{k=-n}^{-1} \underbrace{\frac{1}{2} (a_{-k} + ib_{-k})}_{=: \alpha_k, k=-n, \dots, -1} \exp(ik\omega x) \\ &= \sum_{k=-n}^n \alpha_k \exp(ik\omega x). \end{aligned}$$

### Definition 3.4.8 (komplexes trigonometrisches Polynom)

Die Funktion

$$t_n(x) = \sum_{k=-n}^n \alpha_k \exp(ik\omega x), \quad \alpha_k \in \mathbb{C}, \omega = \frac{2\pi}{T}, \quad (3.9)$$

heißt komplexes trigonometrisches Polynom vom Grad  $n$ .

**Satz 3.4.9** (Fourier, komplexe Darstellung)

Die Fourierkoeffizienten der Bestapproximierenden in (3.9) sind für  $k \in \mathbb{Z}$  gegeben durch

$$\alpha_k = \frac{1}{T} \int_0^T f(x) \exp(-ik\omega x) dx.$$

**Beweis:** Für  $k = 0$  gilt

$$\alpha_0 = \frac{a_0}{2} = \frac{1}{T} \int_0^T f(x) dx = \frac{1}{T} \int_0^T f(x) \exp(-i0\omega x) dx.$$

Für  $k = 1, 2, \dots$  gilt

$$\begin{aligned} \alpha_k &= \frac{1}{2}(a_k - ib_k) \\ &= \frac{1}{T} \int_0^T f(x) \cos(k\omega x) dx - i \frac{1}{T} \int_0^T f(x) \sin(k\omega x) dx \\ &= \frac{1}{T} \int_0^T f(x) \exp(-ik\omega x) dx. \end{aligned}$$

Für  $k = -1, -2, \dots$  gilt

$$\begin{aligned} \alpha_k &= \frac{1}{2}(a_{-k} + ib_{-k}) \\ &= \frac{1}{T} \int_0^T f(x) \cos(-k\omega x) dx + i \frac{1}{T} \int_0^T f(x) \sin(-k\omega x) dx \\ &= \frac{1}{T} \int_0^T f(x) \exp(-ik\omega x) dx. \end{aligned}$$

■

Liegt eine Fouriersumme in komplexer Darstellung vor, so kann man sie wie folgt in reelle Form umrechnen:

$$\begin{aligned} \sum_{k=-\infty}^{\infty} \alpha_k \exp(ik\omega x) &= \sum_{k=-\infty}^{\infty} \alpha_k \cos(k\omega x) + i \alpha_k \sin(k\omega x) \\ &= \alpha_0 + \sum_{k=1}^{\infty} (\alpha_k + \alpha_{-k}) \cos(k\omega x) + i(\alpha_k - \alpha_{-k}) \sin(k\omega x). \end{aligned}$$

Mit den Setzungen  $\alpha_0 = a_0/2$ ,  $a_k = \alpha_k + \alpha_{-k}$  und  $b_k = i(\alpha_k - \alpha_{-k})$  folgt die reelle Darstellung

$$\frac{a_0}{2} + \sum_{k=1}^{\infty} (a_k \cos(k\omega x) + b_k \sin(k\omega x)). \quad (3.10)$$

Falls  $\alpha_k = \overline{\alpha_{-k}}$  für  $k = 0, \dots, n$  gilt, so gilt  $a_k = 2\operatorname{Re}(\alpha_k)$  und  $b_k = -2\operatorname{Im}(\alpha_k)$  in (3.10) und sämtliche Koeffizienten in (3.10) sind reell. In diesem Fall spricht man von der reellen Fouriersumme und man kann komplett reell rechnen.

**Bemerkung 3.4.10 (diskrete Fouriertransformation)**

*Für eine gegebene Funktion  $f$  sind die Fourierkoeffizienten mitunter schwer zu berechnen, falls die auftretenden Integrale nicht analytisch berechenbar sind. Daher bietet es sich an, die Integrale durch ein geeignetes Verfahren zur Approximation von Integralen zu approximieren. Hierzu eignet sich auf Grund ihrer periodischen Struktur insbesondere die aus der Numerik bekannte zusammengesetzte Trapezregel. Dieser Ansatz führt auf die **diskrete Fouriertransformation** sowie die **schnelle Fouriertransformation (FFT)**. Beide eignen sich auch zur Approximation von Funktionen, deren Funktionswert nur an endlich vielen Stützpunkten (z.B. Messstellen) bekannt ist, siehe Anhang A.1. ■*

## Kapitel 4

# Lineare Operatoren und lineare Differentialgleichungen

Lineare Operatoren kennt man aus der linearen Algebra im Zusammenhang mit Matrizen. Jede Matrix  $A \in \mathbb{R}^{m \times n}$  definiert durch  $T(x) = Ax$  mit  $x \in \mathbb{R}^n$  eine lineare Abbildung  $T : \mathbb{R}^n \rightarrow \mathbb{R}^m$ . Umgekehrt lässt sich jede lineare Abbildung  $T : \mathbb{R}^n \rightarrow \mathbb{R}^m$  mithilfe einer Matrix  $A \in \mathbb{R}^{m \times n}$  darstellen, so dass  $T(x) = Ax$  für alle  $x \in \mathbb{R}^n$  gilt. Für Matrizen (und somit für lineare Abbildungen) lassen sich u.a. Begriffe wie Invertierbarkeit und Transponierung definieren. Wir verallgemeinern diese Konzepte auf allgemeine Vektorräume und Banachräume.

### 4.1 Lineare Operatoren

#### Definition 4.1.1 (Operator, linearer Operator)

Eine Abbildung  $T : X \rightarrow Y$  zwischen zwei Vektorräumen  $X$  und  $Y$  (über  $\mathbb{R}$ ) heißt **Operator**.

Ist  $T$  linear, d.h. gelten die Beziehungen

$$\begin{aligned}T(x_1 + x_2) &= T(x_1) + T(x_2), \\T(\lambda x) &= \lambda T(x)\end{aligned}$$

für alle  $x_1, x_2, x \in X$  und alle  $\lambda \in \mathbb{R}$ , so heißt  $T$  **linearer Operator**.

#### Beispiel 4.1.2 (Lineare Operatoren im $\mathbb{R}^n$ )

Sei  $A \in \mathbb{R}^{m \times n}$  eine Matrix. Dann ist durch

$$T(x) = Ax$$

eine lineare Abbildung  $T : \mathbb{R}^n \rightarrow \mathbb{R}^m$  gegeben, denn für  $x_1, x_2, x \in \mathbb{R}^n$  und  $\lambda \in \mathbb{R}$  gilt:

$$\begin{aligned}T(x_1 + x_2) &= A(x_1 + x_2) = Ax_1 + Ax_2 = T(x_1) + T(x_2), \\T(\lambda x) &= A(\lambda x) = \lambda Ax = \lambda T(x).\end{aligned}$$



Umgekehrt kann jede lineare Abbildung  $T : \mathbb{R}^n \rightarrow \mathbb{R}^m$  durch eine Matrix  $A \in \mathbb{R}^{m \times n}$  dargestellt werden: Um dies zu sehen, sei

$$x = \sum_{i=1}^n x_i e_i$$

ein beliebiger Vektor ( $e_i$  ist der  $i$ -te Einheitsvektor). Aus der Linearität von  $T$  folgt dann

$$T(x) = T\left(\sum_{i=1}^n x_i e_i\right) = \sum_{i=1}^n x_i T(e_i) = Ax$$

mit

$$A = \begin{bmatrix} T(e_1) & \cdots & T(e_n) \end{bmatrix}.$$

■

Wir betrachten nun normierte Vektorräume  $(X, \|\cdot\|_X)$  und  $(Y, \|\cdot\|_Y)$ . Wie üblich heißt der Operator  $T : X \rightarrow Y$  **stetig in**  $x_0 \in X$ , falls

$$\lim_{x \rightarrow x_0} T(x) = T(x_0)$$

gilt. Anders ausgedrückt: Für jedes  $\varepsilon > 0$  gibt es ein  $\delta > 0$ , so dass für  $x \in X$  die Implikation

$$\|x - x_0\|_X < \delta \quad \implies \quad \|T(x) - T(x_0)\|_Y < \varepsilon.$$

gilt. Die Menge aller stetigen linearen Operatoren von  $X$  nach  $Y$  wird mit  $\mathcal{L}(X, Y)$  bezeichnet und man kann jedem linearen Operator  $T : X \rightarrow Y$  durch

$$\|T\|_{\mathcal{L}(X, Y)} := \sup_{x \neq 0} \frac{\|T(x)\|_Y}{\|x\|_X} = \sup_{\|x\|_X \leq 1} \|T(x)\|_Y = \sup_{\|x\|_X = 1} \|T(x)\|_Y,$$

eine Norm zuordnen, die sogenannte **Operatornorm** (dass diese tatsächlich eine Norm ist, muss man mit den Normkriterien nachrechnen). Tatsächlich ist  $(\mathcal{L}(X, Y), \|\cdot\|_{\mathcal{L}(X, Y)})$  wieder ein Banachraum. Dazu genügt es, dass  $(Y, \|\cdot\|_Y)$  ein Banachraum ist, siehe [LS76, Satz 1, S. 100].

Wir haben in Beispiel 2.2.8 bereits einige Matrixnormen kennen gelernt. Tatsächlich sind diese Matrixnormen Operatornormen für die durch die Matrizen beschriebenen linearen Operatoren. Wir demonstrieren dies für die Zeilensummennorm  $\|\cdot\|_\infty$ .

### Beispiel 4.1.3 (Operatornorm und Matrixnorm)

Sei  $T \in \mathbb{R}^{m \times n}$  eine Matrix mit Einträgen  $t_{ij}$ ,  $i = 1, \dots, m$ ,  $j = 1, \dots, n$ . Diese definiert eine lineare Abbildung vom  $\mathbb{R}^n$  in den  $\mathbb{R}^m$  gemäß der Vorschrift  $x \mapsto Tx$ . Diese lineare Abbildung bezeichnen wir wieder mit  $T$  gemäß  $T : \mathbb{R}^n \rightarrow \mathbb{R}^m$ .

Hierbei seien die Räume  $\mathbb{R}^n$  und  $\mathbb{R}^m$  jeweils mit der Vektornorm  $\|\cdot\|_\infty$  versehen (damit sind es Banachräume). Wir rechnen die Operatornorm aus:

Es gilt

$$\begin{aligned} \|Tx\|_\infty &= \max_{1 \leq i \leq m} \left| \sum_{j=1}^n t_{ij} x_j \right| \\ &\leq \max_{1 \leq i \leq m} \sum_{j=1}^n |t_{ij}| \cdot |x_j| \\ &\leq \left( \max_{1 \leq j \leq n} |x_j| \right) \cdot \left( \max_{1 \leq i \leq m} \sum_{j=1}^n |t_{ij}| \right) \\ &= \|x\|_\infty \cdot \|T\|_\infty, \end{aligned}$$

wobei  $\|T\|_\infty$  die Matrixnorm (Zeilensummennorm)

$$\|T\|_\infty = \max_{1 \leq i \leq m} \sum_{j=1}^n |t_{ij}|$$

bezeichnet. Es gilt also  $\|T\|_{\mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)} \leq \|T\|_\infty$ .

Tatsächlich gilt sogar die Gleichheit. Um dies zu sehen, wähle  $i_0$  mit

$$\sum_{j=1}^n |t_{i_0,j}| = \max_{1 \leq i \leq m} \sum_{j=1}^n |t_{ij}|$$

und definiere  $\hat{x} \in \mathbb{R}^n$  mit

$$\hat{x}_j = \begin{cases} 0, & \text{falls } t_{i_0,j} = 0, \\ \frac{t_{i_0,j}}{|t_{i_0,j}|}, & \text{falls } t_{i_0,j} \neq 0. \end{cases}$$

Dann folgt  $\|\hat{x}\|_\infty = 1$  und

$$\begin{aligned} (T\hat{x})_{i_0} &= \sum_{j=1}^n t_{i_0,j} \hat{x}_j \\ &= \sum_{j=1}^n t_{i_0,j} t_{i_0,j} \frac{1}{|t_{i_0,j}|} \\ &= \sum_{j=1}^n |t_{i_0,j}| \cdot 1 \\ &= \left( \max_{1 \leq i \leq m} \sum_{j=1}^n |t_{ij}| \right) \cdot \|\hat{x}\|_\infty \\ &= \|T\|_\infty \cdot \|\hat{x}\|_\infty. \end{aligned}$$

Wegen  $(T\hat{x})_{i_0} = \|T\hat{x}\|_\infty$  folgt somit  $\|T\|_{\mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)} \geq \|T\|_\infty$  und damit haben wir insgesamt die Gleichheit gezeigt.

Analoge Beziehungen gelten auch für die Vektor- bzw. Matrixnormen  $\|\cdot\|_p$ ,  $p = 1, 2$ .

Damit sieht man auch sofort, dass lineare Abbildungen  $T : \mathbb{R}^n \rightarrow \mathbb{R}^m$  immer stetig sind, denn für  $\|T\|_\infty > 0$  (für  $\|T\|_\infty = 0$  ist die Stetigkeit klar), ein beliebiges  $\varepsilon > 0$ ,  $\delta < \varepsilon/\|T\|_\infty$ ,  $x_0 \in \mathbb{R}^n$  und  $\|x - x_0\| < \delta$  folgt

$$\|T(x) - T(x_0)\|_\infty = \|T(x - x_0)\|_\infty \leq \|T\|_\infty \cdot \|x - x_0\|_\infty < \varepsilon.$$

Damit ist  $T$  stetig bzgl. der Norm  $\|\cdot\|_\infty$ . Da im  $\mathbb{R}^n$  (bzw.  $\mathbb{R}^m$ ) alle Normen nach Satz 2.2.7 äquivalent sind, ist  $T$  bzgl. aller Normen stetig. ■

Wir haben im vorangegangenen Beispiel festgestellt, dass lineare Operatoren vom  $\mathbb{R}^n$  in den  $\mathbb{R}^m$  immer stetig sind. Dies gilt für allgemeinere Räume i.A. nicht mehr. Der folgende Satz liefert aber eine schöne Charakterisierung der Stetigkeit.

#### Satz 4.1.4

Sei  $T : X \rightarrow Y$  ein linearer Operator. Dann sind folgende Aussagen äquivalent:

- (a)  $T$  ist stetig.
- (b)  $T$  ist stetig in 0.
- (c)  $T$  ist beschränkt, d.h. es gibt eine Konstante  $C > 0$  mit  $\|T(x)\|_Y \leq C\|x\|_X$  für alle  $x \in X$ .

**Beweis:** Übungsaufgabe! ■

#### Beispiel 4.1.5 (Differentialoperator)

Sei  $A \in \mathbb{R}$  eine Zahl. Der Operator  $T : X \rightarrow Y$  mit  $X = (C^1([a, b], \mathbb{R}), \|\cdot\|_{1,\infty})$ ,  $Y = (C([a, b], \mathbb{R}), \|\cdot\|_\infty)$  sowie

$$T(x)(\cdot) = x'(\cdot) - Ax(\cdot)$$

ist ein stetiger linearer Operator. Hierbei deutet  $(\cdot)$  an, dass  $T(x)$ ,  $x$  und  $x'$  Funktionen sind. Der Funktionswert von  $T(x)$  an der Stelle  $t$  ist gegeben durch  $T(x)(t)$ .

Die Linearität ergibt sich wie folgt: Für  $x_1, x_2, x \in C^1([a, b], \mathbb{R})$  und  $\lambda \in \mathbb{R}$  gilt

$$\begin{aligned} T(x_1 + x_2) &= (x_1 + x_2)' - A(x_1 + x_2) = (x_1' - Ax_1) + (x_2' - Ax_2) = T(x_1) + T(x_2), \\ T(\lambda x) &= (\lambda x)' - A(\lambda x) = \lambda x' - \lambda Ax = \lambda T(x). \end{aligned}$$

Die Stetigkeit folgt gemäß Satz 4.1.4 aus der Beschränktheit:

$$\begin{aligned}
 \|T(x)\|_\infty &= \max_{t \in [a,b]} |T(x)(t)| \\
 &= \max_{t \in [a,b]} |x'(t) - Ax(t)| \\
 &\leq \max_{t \in [a,b]} (|x'(t)| + |A| \cdot |x(t)|) \\
 &\leq \max_{t \in [a,b]} |x'(t)| + |A| \max_{t \in [a,b]} |x(t)| \\
 &\leq C \max\{\|x'\|_\infty, \|x\|_\infty\} = C\|x\|_{1,\infty}
 \end{aligned}$$

mit  $C = 1 + |A|$ . ■

**Beispiel 4.1.6 (Lineares Kontrollsystem)**

Seien  $A, B \in \mathbb{R}$  Zahlen und  $\Omega = (a, b)$ . Der Operator  $T : Z \rightarrow Y$  mit  $Z = W^{1,2}(\Omega) \times L^2(\Omega)$ ,  $Y = L^2(\Omega)$  sowie

$$T(x, u)(\cdot) = x'(\cdot) - Ax(\cdot) - Bu(\cdot)$$

ist ein stetiger linearer Operator. Der Raum  $Z$  ist dabei versehen mit der Norm  $\|(x, u)\|_Z = \max\{\|x\|_{W^{1,2}(\Omega)}, \|u\|_{L^2(\Omega)}\}$  und der Raum  $Y$  mit der Norm  $\|\cdot\|_{L^2(\Omega)}$ .

Die Linearität rechnet man wie im vorigen Beispiel nach.

Die Stetigkeit folgt mit der Minkowski-Ungleichung (mit  $p = 2$ ) und Satz 4.1.4 aus der Beschränktheit:

$$\begin{aligned}
 \|T(x, u)\|_Y &= \|x'(\cdot) - Ax(\cdot) - Bu(\cdot)\|_{L^2(\Omega)} \\
 &\leq \|x'\|_{L^2(\Omega)} + |A|\|x\|_{L^2(\Omega)} + |B|\|u\|_{L^2(\Omega)} \\
 &\leq C \max\{\|x'\|_{L^2(\Omega)}, \|x\|_{L^2(\Omega)}, \|u\|_{L^2(\Omega)}\} \\
 &= C\|(x, u)\|_Z,
 \end{aligned}$$

wobei  $C = 1 + |A| + |B|$ . ■

**Beispiel 4.1.7**

Sei  $K : [a, b] \times [c, d] \rightarrow \mathbb{R}$  eine stetige Funktion und  $-\infty < a < b < \infty$  und  $-\infty < c < d < \infty$ . Dann ist durch

$$T(f)(\cdot) = \int_a^b K(t, \cdot) f(t) dt$$

ein stetiger linearer Operator  $T : (C([a, b], \mathbb{R}), \|\cdot\|_\infty) \rightarrow (C([c, d], \mathbb{R}), \|\cdot\|_\infty)$  definiert,

denn es gilt

$$\begin{aligned}
 \|T(f)\|_\infty &= \max_{c \leq s \leq d} \left| \int_a^b K(t, s) f(t) dt \right| \\
 &\leq \max_{c \leq s \leq d} \int_a^b |K(t, s)| \cdot |f(t)| dt \\
 &\leq \left( \max_{c \leq s \leq d} \int_a^b |K(t, s)| dt \right) \cdot \|f\|_\infty \\
 &\leq (b - a) \cdot \|K\|_\infty \cdot \|f\|_\infty \\
 &= C \|f\|_\infty,
 \end{aligned}$$

wobei die Konstante  $C = (b - a) \cdot \|K\|_\infty$  existiert, weil die stetige Funktion  $K$  auf dem kompakten Intervall  $[a, b] \times [c, d]$  beschränkt ist und dort nach dem Satz von Weierstraß ihr Minimum und Maximum annimmt. ■

Wie bei Matrizen kann man auch für lineare Operatoren  $T$  einen inversen Operator  $T^{-1}$  einführen. Falls dieser existiert, ist die eindeutige Lösung  $x$  der linearen Operatorgleichung  $T(x) = y$  formal durch  $x = T^{-1}(y)$  gegeben.

#### Definition 4.1.8 (Inverser Operator)

Sei  $T : X \rightarrow Y$  ein linearer Operator zwischen Vektorräumen  $X$  und  $Y$ .  $T^{-1} : Y \rightarrow X$  heißt **Inverser Operator von  $T$** , falls die Bedingungen

$$T^{-1}(T(x)) = x \quad \text{und} \quad T(T^{-1}(y)) = y$$

für alle  $x \in X$  und alle  $y \in Y$  gelten.

#### Satz 4.1.9

Der Inverse Operator  $T^{-1}$  eines linearen Operators ist ein linearer Operator.

**Beweis:** Für beliebige  $y_1, y_2 \in Y$  und  $x = T^{-1}(y_1 + y_2) - T^{-1}(y_1) - T^{-1}(y_2)$  gilt  $T(x) = 0$ . Daraus folgt

$$0 = T^{-1}(T(x)) = x = T^{-1}(y_1 + y_2) - T^{-1}(y_1) - T^{-1}(y_2)$$

also die additive Linearität

$$T^{-1}(y_1 + y_2) = T^{-1}(y_1) + T^{-1}(y_2).$$

Ebenso gilt für beliebiges  $\lambda \in \mathbb{R}$ ,  $y \in Y$  und  $x = T^{-1}(\lambda y) - \lambda T^{-1}(y)$  die Beziehung  $T(x) = 0$ , woraus wie oben  $T^{-1}(\lambda y) = \lambda T^{-1}(y)$  folgt. ■

**Bemerkung 4.1.10**

Beachte, dass der Inverse Operator eines stetigen linearen Operators i.A. **nicht** stetig ist. Ein Beispiel hierfür ist der lineare und beschränkte Operator  $T(x)(t) = \int_0^t x(\tau)d\tau$  auf den stetigen Funktionen mit Bildern in den stetigen Funktionen. Der Inverse Operator ist  $T^{-1}(y)(t) = \frac{d}{dt}y(t)$ . Dieser ist linear, aber nur für stetig differenzierbare Funktionen definiert. ■

Eine hinreichende Bedingung für die Existenz des inversen Operators liefert der folgende Satz. Dabei heißt ein Operator  $T : X \rightarrow Y$  surjektiv, wenn zu jedem  $y \in Y$  ein  $x \in X$  existiert mit  $y = T(x)$ .

**Satz 4.1.11**

Sei  $T : X \rightarrow Y$  ein linearer Operator zwischen normierten Vektorräumen  $(X, \|\cdot\|_X)$  und  $(Y, \|\cdot\|_Y)$ . Ist  $T$  surjektiv und gibt es eine Konstante  $C > 0$  mit

$$\|T(x)\|_Y \geq C\|x\|_X \quad \forall x \in X,$$

so existiert der inverse Operator  $T^{-1}$  von  $T$ , und er ist linear und beschränkt (also stetig).

**Beweis:** Gilt  $y = T(x_1) = T(x_2)$ , so folgt  $T(x_1 - x_2) = 0$  sowie

$$0 = \|T(x_1 - x_2)\|_Y \geq C\|x_1 - x_2\|_X.$$

Folglich ist  $x_1 = x_2$ , was die Injektivität von  $T$  zeigt. Jedem  $x \in X$  wird also ein eindeutiges  $y \in Y$  zugeordnet und umgekehrt. Daraus ergibt sich die Existenz von  $T^{-1}$  und somit auch dessen Linearität, weil  $T$  linear ist. Die Beschränktheit folgt mit

$$\|T^{-1}(y)\|_X \leq \frac{1}{C}\|T(T^{-1}(y))\|_Y = \frac{1}{C}\|y\|_Y.$$

Der folgende Satz betrachtet die Invertierbarkeit eines Operators unter Störungen. Ein Beweis findet sich in [LS76, Satz 3, S. 108].

**Satz 4.1.12**

Seien  $T, \Delta T : X \rightarrow Y$  lineare Operatoren zwischen normierten Vektorräumen  $(X, \|\cdot\|_X)$  und  $(Y, \|\cdot\|_Y)$ . Der inverse Operator  $T^{-1}$  existiere. Falls

$$\|\Delta T\|_{\mathcal{L}(X,Y)} \leq \frac{1}{\|T^{-1}\|_{\mathcal{L}(Y,X)}}$$

gilt, dann besitzt der Operator  $A = T + \Delta T$  einen inversen Operator  $A^{-1}$  mit

$$\|A^{-1} - T^{-1}\|_Y \leq \frac{\|\Delta T\|_{\mathcal{L}(X,Y)}}{1 - \|T^{-1}\|_{\mathcal{L}(Y,X)}\|\Delta T\|_{\mathcal{L}(X,Y)}} \|T^{-1}\|_{\mathcal{L}(Y,X)}^2.$$

## 4.2 Lineare Funktionale und Dualräume

Eine wichtige Teilklasse von (linearen) Operatoren stellen (lineare) Funktionale dar.

### Definition 4.2.1 (Funktional, lineares Funktional)

Eine Abbildung  $T : X \rightarrow \mathbb{R}$  von einem Vektorraum  $X$  (über  $\mathbb{R}$ ) in die reellen Zahlen heißt **Funktional**.

Ist  $T$  linear, d.h. gelten die Beziehungen

$$\begin{aligned} T(x_1 + x_2) &= T(x_1) + T(x_2), \\ T(\lambda x) &= \lambda T(x) \end{aligned}$$

für alle  $x_1, x_2, x \in X$  und alle  $\lambda \in \mathbb{R}$ , so heißt  $T$  **lineares Funktional**.

### Definition 4.2.2 (Dualraum)

Sei  $(X, \|\cdot\|_X)$  ein normierter Vektorraum. Die Menge aller stetigen linearen Funktionale  $\mathcal{L}(X, \mathbb{R})$  versehen mit der Norm

$$\|f\|_{X^*} := \|f\|_{\mathcal{L}(X, \mathbb{R})} = \sup_{\|x\|_X \leq 1} |f(x)|$$

heißt **Dualraum von  $X$**  und wird mit  $X^*$  bezeichnet.

### Bemerkung 4.2.3

$(X^*, \|\cdot\|_{X^*})$  ist ein Banachraum, siehe [Wer95, Korollar II.2.2, S. 58]. ■

### Beispiel 4.2.4 (Punktweise Auswertung einer stetigen Funktion)

Sei  $X = (C([a, b], \mathbb{R}), \|\cdot\|_\infty)$  der Raum der stetigen Funktionen auf dem Intervall  $[a, b]$  versehen mit der Supremumsnorm  $\|\cdot\|_\infty$  und  $t_0 \in [a, b]$  ein gegebener Punkt. Dann ist durch

$$T(f) = f(t_0)$$

ein stetiges lineares Funktional  $T : X \rightarrow \mathbb{R}$  definiert.

Zur Linearität: Für alle  $f, g \in X$  und  $\lambda \in \mathbb{R}$  gilt

$$\begin{aligned} T(f + g) &= (f + g)(t_0) = f(t_0) + g(t_0) = T(f) + T(g), \\ T(\lambda f) &= (\lambda f)(t_0) = \lambda f(t_0) = \lambda T(f). \end{aligned}$$

Zur Stetigkeit:

$$|T(f)| = |f(t_0)| \leq \max_{t \in [a, b]} |f(t)| = \|f\|_\infty.$$

■

**Beispiel 4.2.5 (Auswertung eines Integrals)**

Sei  $X = (L^1(\Omega), \|\cdot\|_{L^1(\Omega)})$  der Raum der (Lebesgue-)integrierbaren Funktionen auf  $\Omega$ .

Dann ist durch

$$T(f) = \int_\omega f d\mu$$

ein stetiges lineares Funktional  $T : X \rightarrow \mathbb{R}$  definiert.

Zur Linearität: Für alle  $f, g \in X$  und  $\lambda \in \mathbb{R}$  gilt

$$\begin{aligned} T(f + g) &= \int_\Omega f + g d\mu = \int_\Omega f d\mu + \int_\Omega g d\mu = T(f) + T(g), \\ T(\lambda f) &= \int_\Omega \lambda f d\mu = \lambda \int_\Omega f d\mu = \lambda T(f). \end{aligned}$$

Zur Stetigkeit:

$$|T(f)| = \left| \int_\Omega f d\mu \right| \leq \int_\Omega |f| d\mu = \|f\|_{L^1(\Omega)}.$$

■

Es stellt sich die Frage, wie der Dualraum eines normierten Vektorraums aussieht? Insbesondere ist man daran interessiert, die Elemente des Dualraums explizit darzustellen. Dies ist insbesondere für Hilberträume sehr angenehm.

**Satz 4.2.6 (Satz von Riesz)**

Sei  $X$  ein Hilbertraum mit Skalarprodukt  $\langle \cdot, \cdot \rangle_X$  und  $f \in X^*$ . Dann existiert ein eindeutiges Element  $y \in X$  mit

$$f(x) = \langle y, x \rangle_X \quad \forall x \in X$$

und  $\|f\|_{X^*} = \|y\|_X$ .



Der Satz von Riesz besagt, dass jedem stetigen linearen Funktional  $f$  auf  $X$  eindeutig ein Element  $y$  aus  $X$  zugeordnet werden kann. Umgekehrt definiert  $f(x) = \langle y, x \rangle_X$  für jedes  $y \in X$  ein stetiges lineares Funktional auf  $X$ . Man kann also die Elemente des Dualraums mit den Elementen des Hilbertraums identifizieren. Daher ist der Dualraum  $X^*$  eines Hilbertraums  $X$  wieder ein Hilbertraum.

**Beweis:** Für  $f \in X^*$  betrachte den Kern von  $f$ :

$$\text{Kern}(f) = \{x \in X \mid f(x) = 0\}.$$

Wegen der Stetigkeit von  $f$  ist  $\text{Kern}(f)$  ein abgeschlossener Unterraum von  $X$ . Im Fall  $f = 0$  wähle  $y = 0$ . Für  $f \neq 0$  wähle

$$w \in \text{Kern}(f)^\perp = \{w \in X \mid \langle w, x \rangle_X = 0 \forall x \in \text{Kern}(f)\}$$

mit  $f(w) = 1$ . Dann gilt  $x - f(x)w \in \text{Kern}(f)$  für jedes  $x \in X$  sowie

$$\langle w, x \rangle_X = \underbrace{\langle w, x - f(x)w \rangle_X}_{=0} + \langle w, f(x)w \rangle_X = f(x)\|w\|_X^2.$$

Mit  $y := \frac{1}{\|w\|_X^2}w$  folgt also  $f(x) = \langle y, x \rangle_X$ .

Die Eindeutigkeit sieht man wie folgt: Angenommen, es gibt ein  $\tilde{y} \in X$ ,  $\tilde{y} \neq y$  mit  $f(x) = \langle \tilde{y}, x \rangle_X$ . Dann folgt für alle  $x \in X$ ,

$$0 = f(x) - f(x) = \langle y, x \rangle_X - \langle \tilde{y}, x \rangle_X = \langle y - \tilde{y}, x \rangle_X.$$

Wählt man speziell  $x = y - \tilde{y}$ , so folgt hieraus  $0 = \|y - \tilde{y}\|_X$  und damit  $y = \tilde{y}$  im Widerspruch zur Annahme  $y \neq \tilde{y}$ .

Die Gleichheit der Norm zeigt man wie folgt: Mit der Cauchy-Schwarz-Ungleichung<sup>1</sup> folgt

$$\|f\|_{X^*} = \sup_{\|x\|_X \leq 1} |f(x)| = \sup_{\|x\|_X \leq 1} |\langle y, x \rangle_X| \leq \sup_{\|x\|_X \leq 1} \|y\|_X \cdot \|x\|_X = \|y\|_X,$$

und es gilt  $f(y/\|y\|_X) = \langle y, y/\|y\|_X \rangle_X = \|y\|_X$ , so dass  $\|f\|_{X^*} = \|y\|_X$ . ■

#### Beispiel 4.2.7

Die stetigen linearen Funktionale  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  lassen sich darstellen als

$$f(x) = c^\top x = \langle c, x \rangle_{\mathbb{R}^n}$$

mit einem Vektor  $c \in \mathbb{R}^n$ . ■

#### Beispiel 4.2.8

Die stetigen linearen Funktionale  $f : L^2(\Omega) \rightarrow \mathbb{R}$  lassen sich nach dem Satz von Riesz darstellen als

$$f(x) = \langle y, x \rangle_{L^2(\Omega)} = \int_{\Omega} y(t)^\top x(t) dt$$

<sup>1</sup>Die Cauchy-Schwarz-Ungleichung  $|\langle x, y \rangle_X| \leq \|x\|_X \cdot \|y\|_X$  gilt in Prä-Hilberträumen.

mit einer Funktion  $y \in L^2(\Omega)$ . ■

**Bemerkung 4.2.9**

Bei Banachräumen ist die Situation nicht mehr so angenehm, wie bei Hilberträumen. Der Dualraum eines Banachraums lässt sich i.A. nicht mehr mit sich selbst identifizieren. Einige Beispiele:

- (i) Für  $1 \leq p < \infty$  ist der Dualraum von  $L^p(\Omega)$  gegeben durch  $L^q(\Omega)$  mit  $1/p + 1/q = 1$ , d.h. jedes  $f \in L^p(\Omega)^*$  lässt sich darstellen als

$$f(x) = \int_{\Omega} y(t)x(t)dt, \quad \forall x \in L^p(\Omega)$$

mit einem  $y \in L^q(\Omega)$ . Für  $p = 1$  gilt die Konvention  $q = \infty$ .

- (ii) Der Dualraum von  $L^\infty(\Omega)$  besitzt keine schöne Struktur, insbesondere ist es **nicht** der Raum  $L^1(\Omega)$ .

- (iii) Der Dualraum der stetigen Funktionen ist der Raum der Funktionen beschränkter Variation, siehe [Lue69, Section 5.5, Theorem 1].

Für tiefere Details siehe [BK11, Wer95, Bur98, Alt02]. ■

Schließlich betrachten wir noch eine Verallgemeinerung der transponierten Matrix, nämlich den adjungierten Operator.

**Definition 4.2.10 (Adjungierter Operator)**

Sei  $T : X \rightarrow Y$  ein stetiger linearer Operator zwischen normierten Vektorräumen  $(X, \|\cdot\|_X)$  und  $(Y, \|\cdot\|_Y)$ . Der **adjungierte Operator**  $T^* : Y^* \rightarrow X^*$  ist ein stetiger linearer Operator, der definiert ist durch

$$T^*(y^*)(x) = y^*(T(x)) \quad \forall y^* \in Y^*, x \in X.$$

**Beispiel 4.2.11**

Für den stetigen linearen Operator  $T : \mathbb{R}^n \rightarrow \mathbb{R}^m$  gemäß  $T(x) = Ax$  mit einer Matrix  $A \in \mathbb{R}^{m \times n}$  ergibt sich:

$$y^*(T(x)) = y^*(Ax) = y^\top(Ax) = (A^\top y)^\top x = A^*(y^*)(x),$$

d.h. der adjungierte Operator von  $T$  (bzw.  $A$ ) ist der lineare Operator  $T^* : \mathbb{R}^m \rightarrow \mathbb{R}^n$  mit  $T^*(y) = A^\top y$ .

Hierbei wurde ausgenutzt, dass sich jedes stetige lineare Funktional  $y^* : \mathbb{R}^m \rightarrow \mathbb{R}$  schreiben lässt als  $y^*(x) = y^\top x$  mit einem Vektor  $y \in \mathbb{R}^m$ . Wir können ja den Hilbertraum  $\mathbb{R}^m$  nach dem Satz von Riesz mit sich selbst identifizieren. ■

**Beispiel 4.2.12 (Adjungierter Differentialoperator)**

Sei  $A \in \mathbb{R}^{n \times n}$  eine Matrix und  $\Omega = (0, 1)$ . Betrachte den stetigen linearen Operator  $T : X \rightarrow Y$  mit

$$T(x)(\cdot) = x'(\cdot) - Ax(\cdot)$$

und  $X = W_n^{1,2}(\Omega)$  und  $Y = L_n^2(\Omega)^2$

Wir rechnen den adjungierten Operator aus und nutzen aus, dass  $L_n^2(\Omega)$  ein Hilbertraum ist. Deshalb kann jedes  $y^* \in L_n^2(\Omega)^*$  dargestellt werden als

$$y^*(z) = \langle y, z \rangle_{L_n^2(\Omega)} = \int_{\Omega} y(t)^\top z(t) dt \quad \forall z \in L_n^2(\Omega)$$

mit eindeutigem  $y \in L_n^2(\Omega)$ . Damit erhalten wir mit partieller Integration (Achtung:  $y$  ist nur in  $L^2(\Omega)$  und darf nicht differenziert werden!)

$$\begin{aligned} y^*(T(x)) &= \int_{\Omega} y(t)^\top (x'(t) - Ax(t)) dt \\ &= - \left[ \left( - \int_t^1 y(s)^\top A ds \right) x(t) \right]_0^1 + \int_{\Omega} \left( y(t)^\top - \int_t^1 y(s)^\top A ds \right) x'(t) dt \\ &= \underbrace{\left( - \int_0^1 y(s)^\top A ds \right)}_{=\mu(0)^\top} x(0) + \int_{\Omega} \underbrace{\left( y(t)^\top - \int_t^1 y(s)^\top A ds \right)}_{=\mu'(t)^\top} x'(t) dt \\ &= \langle \mu(0), x(0) \rangle_{\mathbb{R}^n} + \langle \mu', x' \rangle_{L^2(\Omega)} \\ &= T^*(y^*)(x). \end{aligned}$$

### 4.3 Lineare Differentialgleichungen und Differentialoperatoren

Wir betrachten den stetigen linearen Differentialoperator  $T : C^n([a, b], \mathbb{R}) \rightarrow C([a, b], \mathbb{R})$  mit

$$T(x) = \frac{d^n}{dt^n} x + a_{n-1} \frac{d^{n-1}}{dt^{n-1}} x + \dots + a_1 \frac{d}{dt} x + a_0 x$$

mit reellen Zahlen  $a_j, j = 0, 1, \dots, n - 1$ , sowie die lineare Operatorgleichung

$$T(x) = f$$

<sup>2</sup>Der Index  $n$  gibt an, dass es sich hierbei um vektorwertige Funktionen mit  $n$  Komponenten handelt. Jede Komponente ist im  $W^{1,2}(\Omega)$  bzw. im  $L^2(\Omega)$ . Häufig nutzt man auch die Schreibweisen  $W^{1,2}(\Omega)^n$  bzw.  $L^2(\Omega)^n$ .

mit einer gegebenen Funktion  $f \in C([a, b], \mathbb{R})$ . Die Operatorgleichung stellt eine **lineare Differentialgleichung n-ter Ordnung mit konstanten Koeffizienten** dar, welche es zu lösen gilt:

$$T(x)(t) = x^{(n)}(t) + a_{n-1}x^{(n-1)}(t) + \dots + a_1x'(t) + a_0x(t) = f(t) \quad (4.1)$$

Man nennt diese lineare Differentialgleichung **homogen**, falls  $f \equiv 0$  ist, andernfalls **inhomogen**. Wir untersuchen zunächst den homogenen Fall mit  $f \equiv 0$ , also die Differentialgleichung

$$T(x)(t) = x^{(n)}(t) + a_{n-1}x^{(n-1)}(t) + \dots + a_1x'(t) + a_0x(t) = 0.$$

Aus der Linearität des Operators  $T$  folgt sofort das sogenannte

### Superpositionsprinzip:

Sind  $x_1$  und  $x_2$  Lösungen der homogenen linearen Differentialgleichung  $T(x) = 0$ , so ist auch jede Linearkombination  $c_1x_1 + c_2x_2$  für beliebige Konstanten  $c_1, c_2 \in \mathbb{R}$  eine Lösung mit  $T(c_1x_1 + c_2x_2) = c_1T(x_1) + c_2T(x_2) = 0$ .

Die Lösungen der homogenen linearen Differentialgleichung definieren wegen des Superpositionsprinzips einen Vektorraum und man kann zeigen, dass es **genau  $n$  linear unabhängige Lösungen** der homogenen linearen Differentialgleichungen gibt. Jede Menge  $\{x_1, \dots, x_n\}$  von  $n$  linear unabhängigen Lösungen der homogenen linearen Differentialgleichung nennt man **Fundamentalsystem**. Kennt man ein solches Fundamentalsystem, so lässt sich die allgemeine Lösung der homogenen Differentialgleichung schreiben als Linearkombination der Fundamentallösungen:

### Allgemeine Lösung von $T(x) = 0$ :

$$x(t) = \sum_{k=1}^n c_k x_k(t) \quad \text{mit } c_k \in \mathbb{R}, k = 1, \dots, n.$$

$\{x_k\}_{k=1, \dots, n}$ : linear unabhängige Fundamentallösungen mit  $T(x_k) = 0$ ,  $k = 1, \dots, n$

Die allgemeine Lösung der inhomogenen Differentialgleichung  $T(x) = f$  lässt sich dann wie folgt darstellen:

### Allgemeine Lösung von $T(x) = f$ (siehe (4.1)):

$$x(t) = x_p(t) + \sum_{k=1}^n c_k x_k(t) \quad \text{mit } c_k \in \mathbb{R}, k = 1, \dots, n.$$

$x_k$ : linear unabhängige Fundamentallösungen mit  $T(x_k) = 0$ ,  $k = 1, \dots, n$   
 $x_p$ : spezielle Lösung mit  $T(x_p) = f$  (genannt **Partikulärlösung**)

Zur Lösung der inhomogenen Differentialgleichung (4.1) benötigt man also eine Partikulärlösung  $x_p$  sowie ein Fundamentalsystem der zugehörigen homogenen Differentialgleichung.

### 4.3.1 Bestimmung von Fundamentallösungen

Zur Bestimmung von Fundamentallösungen wählen wir den Ansatz

$$x(t) = \exp(\lambda t) \quad (\lambda \text{ konstant}).$$

$x$  ist beliebig oft differenzierbar mit

$$x'(t) = \lambda \exp(\lambda t), x''(t) = \lambda^2 \exp(\lambda t), \dots, x^{(n)}(t) = \lambda^n \exp(\lambda t).$$

Einsetzen dieser Ableitungen in die Differentialgleichung liefert die Gleichung

$$\exp(\lambda t) (\lambda^n + a_{n-1}\lambda^{n-1} + \dots + a_1\lambda + a_0) = 0.$$

Wegen  $\exp(\lambda t) > 0$  ist diese Gleichung genau dann erfüllt, wenn  $\lambda$  eine Nullstelle des **charakteristischen Polynoms**

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

ist. Das charakteristische Polynom ist ein Polynom  $n$ -ten Grades und besitzt nach dem Hauptsatz der Algebra  $n$  Nullstellen, die reell oder komplex sein können. Im komplexen Fall treten Nullstellen immer als konjugiert komplexes Paar auf, d.h. mit  $\lambda$  ist auch  $\bar{\lambda}$  eine Nullstelle. Nullstellen können auch mehrfach auftreten. Wir diskutieren die verschiedenen Fälle.

#### Fall 1: Es treten nur einfache reelle Nullstellen auf

Bezeichnen  $\lambda_1, \dots, \lambda_n$  die  $n$  verschiedenen Nullstellen des charakteristischen Polynoms, so ergeben sich aus dem gewählten Ansatz die zugehörigen Lösungen

$$x_j(t) = \exp(\lambda_j t), \quad j = 1, 2, \dots, n,$$

welche linear unabhängig sind und ein Fundamentalsystem darstellen. Die allgemeine Lösung der homogenen Differentialgleichung lautet dann:

$$x(t) = c_1 x_1(t) + \dots + c_n x_n(t) = \sum_{j=1}^n c_j \exp(\lambda_j t)$$

mit beliebigen Koeffizienten  $c_1, \dots, c_n \in \mathbb{R}$ . die homogene Differentialgleichung.

**Beispiel 4.3.1**

Gegeben sei die lineare Differentialgleichung 2. Ordnung

$$x''(t) + 5x'(t) + 6x(t) = 0$$

mit Anfangswerten

$$x(0) = 2, \quad x'(0) = 3.$$

Der Ansatz  $x(t) = \exp(\lambda t)$  führt auf das charakteristische Polynom

$$\lambda^2 + 5\lambda + 6 = (\lambda + 2)(\lambda + 3) = 0 \quad \implies \quad \lambda_1 = -2, \lambda_2 = -3.$$

Daraus ergibt sich die allgemeine Lösung

$$x(t) = c_1 \exp(-2t) + c_2 \exp(-3t) \quad (c_1, c_2 \in \mathbb{R} \text{ konstant}).$$

Einsetzen der Anfangswerte:

$$c_1 + c_2 = 2, \quad -2c_1 - 3c_2 = 3 \quad \implies \quad c_1 = 9, c_2 = -7$$

Lösung des Anfangswertproblems:

$$x(t) = 9 \exp(-2t) - 7 \exp(-3t).$$

■

**Fall 2: Es treten einfache komplexe Nullstellen auf**

O.B.d.A. sei  $\lambda_1 = a + ib$  mit  $a, b \in \mathbb{R}$  eine komplexe Nullstelle des charakteristischen Polynoms. Dann ist auch  $\lambda_2 = \bar{\lambda}_1 = a - ib$  eine Nullstelle.

Die zugehörigen komplexen Lösungen lauten dann

$$\begin{aligned} x_1(t) &= \exp((a + ib)t) = \exp(at) \exp(ib t) \\ &= \exp(at) (\cos(bt) + i \sin(bt)), \\ x_2(t) &= \exp((a - ib)t) = \exp(at) \exp(-ib t) \\ &= \exp(at) (\cos(-bt) + i \sin(-bt)) \\ &= \exp(at) (\cos(bt) - i \sin(bt)). \end{aligned}$$

Wir sind jedoch nicht an komplexen Lösungen interessiert, sondern nur an reellen Lösungen. Diese können wir durch Linearkombinationen von  $x_1$  und  $x_2$  erhalten, denn es gilt:

$$\begin{aligned} \frac{1}{2} (x_1(t) + x_2(t)) &= \exp(at) \cos(bt), \\ \frac{1}{2i} (x_1(t) - x_2(t)) &= \exp(at) \sin(bt). \end{aligned}$$

Daraus ergeben sich die beiden reellen Lösungen

$$\exp(at) \cos(bt) \quad \text{und} \quad \exp(at) \sin(bt).$$

### Beispiel 4.3.2

Löse

$$x''(t) + x'(t) + x(t) = 0.$$

Das charakteristische Polynom und dessen Nullstellen lauten:

$$\lambda^2 + \lambda + 1 = 0, \quad \lambda_{1,2} = \underbrace{-\frac{1}{2}}_{=a} \pm i \underbrace{\frac{\sqrt{3}}{2}}_{=b}.$$

Lösung:  $(c_1, c_2 \in \mathbb{R})$

$$x(t) = c_1 \exp\left(-\frac{t}{2}\right) \cos\left(\frac{\sqrt{3}}{2}t\right) + c_2 \exp\left(-\frac{t}{2}\right) \sin\left(\frac{\sqrt{3}}{2}t\right).$$

Abbildung 4.1 zeigt eine spezielle Lösung. ■

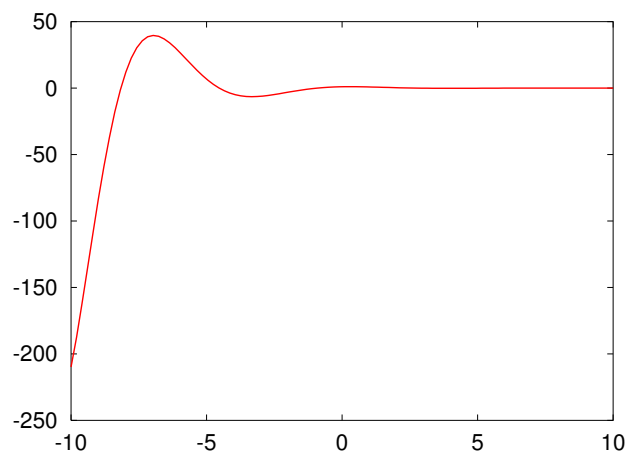


Abbildung 4.1: Lösung für  $c_1 = c_2 = 1$ .

### Fall 3: Es treten mehrfache reelle Nullstellen auf

Ist  $\lambda$  eine  $k$ -fache reelle Nullstelle des charakteristischen Polynoms, so sind die Funktionen

$$\exp(\lambda t), \quad t \exp(\lambda t), \quad \dots, \quad t^{k-1} \exp(\lambda t)$$

Lösungen der linearen homogenen Differentialgleichung.

**Fall 4: Es treten mehrfache komplexe Nullstellen auf**

Ist  $\lambda = a + ib$  (und damit auch  $\bar{\lambda} = a - ib$ ) eine  $k$ -fache komplexe Nullstelle des charakteristischen Polynoms, so sind die Funktionen

$$t^j \exp(at) \cos(bt), \quad t^j \exp(at) \sin(bt), \quad j = 0, \dots, k - 1.$$

Lösungen der linearen homogenen Differentialgleichung.

Wir deuten im folgenden Beispiel an, wie man auf die Lösungen in den Fällen 3 und 4 kommt. Dazu verwenden wir **d'Alemberts Methode**, die sich analog auch für homogene Differentialgleichungen mit variablen Koeffizienten anwenden läßt.

**Beispiel 4.3.3**

*Differentialgleichung:*

$$x''(t) + 4x'(t) + 4x(t) = 0. \tag{4.2}$$

*Charakteristisches Polynom:*

$$\lambda^2 + 4\lambda + 4 = (\lambda + 2)^2 = 0 \quad \Rightarrow \quad \lambda_1 = \lambda_2 = -2.$$

*Damit ist*

$$x_1(t) = c \exp(-2t) \tag{4.3}$$

*für alle Konstanten  $c$  Lösung.*

*Für ein Fundamentalsystem fehlt uns jedoch noch eine linear unabhängige zweite Lösung, die wir nun mit d'Alemberts Methode konstruieren.*

**d'Alemberts Idee:**

*Ersetze die Konstante  $c$  in (4.3) durch eine Funktion  $v(t)$  und versuche,  $v$  so zu bestimmen, dass  $v(t) \cdot x_1(t)$  Lösung von (4.2) ist.*

*Aus dem Ansatz von d'Alembert folgt*

$$\begin{aligned} x(t) &= v(t) \cdot x_1(t) = v(t) \exp(-2t), \\ x'(t) &= v'(t) \exp(-2t) - 2v(t) \exp(-2t) = \exp(-2t) (v'(t) - 2v(t)), \\ x''(t) &= -2 \exp(-2t) (v'(t) - 2v(t)) + \exp(-2t) (v''(t) - 2v'(t)) \\ &= \exp(-2t) (-4v'(t) + 4v(t) + v''(t)). \end{aligned}$$

*Einsetzen dieser Beziehungen in die Differentialgleichung (4.2) und Zusammenfassen von Termen liefert*

$$\exp(-2t) (v''(t) - 4v'(t) + 4v(t) + 4v'(t) - 8v(t) + 4v(t)) = 0$$



bzw.

$$v''(t) = 0.$$

Zweifache Integration liefert

$$v(t) = c_1 t + c_2,$$

wobei  $c_1$  und  $c_2$  reelle Integrationskonstanten sind. Einsetzen von  $v$  in (4.3) liefert die Lösungsdarstellung

$$x(t) = c_1 \underbrace{t \exp(-2t)}_{=:x_2(t)} + c_2 \underbrace{\exp(-2t)}_{x_1(t)}.$$

Hierin sind  $x_1$  und  $x_2$  linear unabhängige Lösungen und  $\{x_1, x_2\}$  ist somit ein Fundamentalsystem. ■

Mit Hilfe der Fälle 1-4 können wir nun homogene lineare Differentialgleichungen n-ter Ordnung vollständig lösen und ein Fundamentalsystem bestimmen.

### 4.3.2 Bestimmung einer Partikulärlösung

Nachdem wir homogene lineare Differentialgleichungen n-ter Ordnung mit konstanten Koeffizienten vollständig lösen können, benötigen wir nur noch eine Methode, um eine Partikulärlösung für die inhomogene Differentialgleichung zu bestimmen.

Häufig verwendet man hierfür spezielle Ansätze, die von der Form der Inhomogenität  $f$  abhängen. Die folgende Tabelle fasst geeignete Ansatzfunktionen  $x_p$  für einige Inhomogenitäten  $f$  zusammen. Hierin bezeichnet  $s$  die kleinste Zahl, so dass kein Term in  $x_p(t)$  Lösung der zugehörigen homogenen linearen Differentialgleichung ist:

$f(t)$	$x_p(t)$
$P_n(t) = \alpha_0 t^n + \alpha_1 t^{n-1} + \dots + \alpha_n$	$t^s (A_0 t^n + A_1 t^{n-1} + \dots + A_n)$
$P_n(t) \exp(\alpha t)$	$t^s (A_0 t^n + A_1 t^{n-1} + \dots + A_n) \exp(\alpha t)$
$P_n(t) \exp(\alpha t) \sin(\beta t)$	$t^s [(A_0 t^n + A_1 t^{n-1} + \dots + A_n) \exp(\alpha t) \cos(\beta t) + (B_0 t^n + B_1 t^{n-1} + \dots + B_n) \exp(\alpha t) \sin(\beta t)]$
$P_n(t) \exp(\alpha t) \cos(\beta t)$	$t^s [(A_0 t^n + A_1 t^{n-1} + \dots + A_n) \exp(\alpha t) \cos(\beta t) + (B_0 t^n + B_1 t^{n-1} + \dots + B_n) \exp(\alpha t) \sin(\beta t)]$

Die in den Ansätzen auftretenden Konstanten bestimmt man durch Einsetzen in die inhomogene Differentialgleichung.

**Beispiel 4.3.4** (Ansatz vom Typ der rechten Seite I)

Finde eine Partikulärlösung von

$$x''(t) - 3x'(t) - 4x(t) = 3 \exp(2t).$$

Ansatz vom Typ der rechten Seite:

$$x_p(t) = A \exp(2t), \quad A \text{ konstant.}$$

Einsetzen von  $x_p$  in die Differentialgleichung liefert

$$4A \exp(2t) - 6A \exp(2t) - 4A \exp(2t) = 3 \exp(2t) \quad \Leftrightarrow \quad -6A \exp(2t) = 3 \exp(2t)$$

und somit  $A = -\frac{1}{2}$ . Eine Partikulärlösung ist daher gegeben durch

$$x_p(t) = -\frac{1}{2} \exp(2t).$$

■

Das folgende Beispiel zeigt, dass der Faktor  $t^s$  in der obigen Tabelle benötigt wird, wenn sich heraus stellt, dass die Ansatzfunktion z.B. für  $s = 0$  gerade die Lösung der homogenen Gleichung ist.

**Beispiel 4.3.5**

Bestimme eine Partikulärlösung von

$$x''(t) + 4x(t) = 3 \cos(2t).$$

Ansatz:

$$x_p(t) = A \cos(2t) + B \sin(2t), \quad A, B \text{ konstant.}$$

Einsetzen von  $x_p$  liefert

$$\underbrace{(4A - 4A) \cos(2t) + (4B - 4B) \sin(2t)}_{=0} = 3 \cos(2t).$$

Es existiert keine Lösung!

**Begründung:**  $\{\sin(2t), \cos(2t)\}$  ist ein Fundamentalsystem der homogenen Gleichung. Daher liefert unser Ansatz gerade die allgemeine Lösung der homogenen Gleichung und kann daher keine Lösung der inhomogenen Gleichung sein.

Nächster Versuch:

$$x_p(t) = t (A \cos(2t) + B \sin(2t)), \quad A, B \text{ konstant}$$

(also  $s = 1$  in der obigen Tabelle). Einsetzen liefert

$$-4A \sin(2t) + 4B \cos(2t) = 3 \cos(2t) \quad \Rightarrow \quad A = 0, B = \frac{3}{4}.$$

Eine Partikulärlösung ist gegeben durch

$$x_p(t) = \frac{3}{4}t \sin(2t).$$

■

### Bemerkung 4.3.6

Ist die Inhomogenität eine Summe der angegebenen Funktionen, etwa  $f(t) = f_1(t) + f_2(t)$ , so kann eine Partikulärlösung durch Aufspaltung  $x_p(t) = x_{p,1}(t) + x_{p,2}(t)$  mit geeignetem Ansatz für  $x_{p,1}$  und  $x_{p,2}$  berechnet werden. ■

Die folgende Methode der **Variation der Konstanten** ist generell anwendbar und ist nicht auf bestimmte Ansatzfunktionen beschränkt.

### Beispiel 4.3.7 (Variation der Konstanten)

Bestimme eine Partikulärlösung von

$$x''(t) + 4x(t) = 2.$$

Die Lösung der homogenen Gleichung lautet

$$x_h(t) = c_1 \cos(2t) + c_2 \sin(2t), \quad c_1, c_2 \in \mathbb{R}.$$

#### Variation der Konstanten:

Ersetze  $c_1$  und  $c_2$  durch Funktionen  $u_1(t)$  und  $u_2(t)$  und bestimme  $u_1$  und  $u_2$ , so dass

$$x(t) = u_1(t) \cos(2t) + u_2(t) \sin(2t) \tag{4.4}$$

eine Lösung der inhomogenen Gleichung ist.

#### **Problem:**

Falls wir  $x$  einfach in die Differentialgleichung einsetzen, bekommen wir lediglich eine einzelne Gleichung für die beiden unbekanntenen Funktionen  $u_1$  und  $u_2$ . Dies erlaubt mehrdeutige Lösungen.

**Idee:** Formuliere eine zweite Gleichung, so dass  $u_1$  und  $u_2$  eindeutig festgelegt sind.

Es gilt

$$x'(t) = -2u_1(t) \sin(2t) + 2u_2(t) \cos(2t) + u_1'(t) \cos(2t) + u_2'(t) \sin(2t). \tag{4.5}$$

Wir fordern nun zusätzlich die Gleichung

$$u_1'(t) \cos(2t) + u_2'(t) \sin(2t) = 0 \quad (4.6)$$

aus folgenden Gründen: Erstens wird dadurch die Ableitung  $x'$  vereinfacht. Zweitens hängt  $x'$  nur von  $u_1$  und  $u_2$  ab und nicht von deren Ableitungen.

Mit (4.6) vereinfacht sich Gleichung (4.5) zu

$$x'(t) = -2u_1(t) \sin(2t) + 2u_2(t) \cos(2t). \quad (4.7)$$

Ableiten von (4.7) liefert

$$x''(t) = -4u_1(t) \cos(2t) - 4u_2(t) \sin(2t) - 2u_1'(t) \sin(2t) + 2u_2'(t) \cos(2t)$$

bzw.

$$x''(t) + 4 \underbrace{(u_1(t) \cos(2t) + u_2(t) \sin(2t))}_{=x(t) \text{ nach (4.4)}} = -2u_1'(t) \sin(2t) + 2u_2'(t) \cos(2t).$$

Damit ist die Differentialgleichung erfüllt, falls gilt

$$-2u_1'(t) \sin(2t) + 2u_2'(t) \cos(2t) = 2.$$

Beachte, dass  $u_1$  und  $u_2$  auch (4.6) erfüllen müssen. Deshalb müssen wir das lineare Differentialgleichungssystem erster Ordnung für  $u_1$  und  $u_2$  lösen:

$$\begin{aligned} u_1'(t) \cos(2t) + u_2'(t) \sin(2t) &= 0, \\ -2u_1'(t) \sin(2t) + 2u_2'(t) \cos(2t) &= 2. \end{aligned}$$

Die erste Gleichung kann nach  $u_2'$  aufgelöst werden:

$$u_2'(t) = -u_1'(t) \frac{\cos(2t)}{\sin(2t)}.$$

Einsetzen von  $u_2'$  in die zweite Gleichung liefert

$$u_1'(t) = -\sin(2t).$$

Daraus folgt

$$u_2'(t) = -u_1'(t) \frac{\cos(2t)}{\sin(2t)} = \cos(2t).$$

Integration von  $u_1'$  und  $u_2'$  liefert

$$u_1(t) = \frac{1}{2} \cos(2t) + c_1, \quad u_2(t) = \frac{1}{2} \sin(2t) + c_2.$$

Nach (4.4) haben wir die allgemeine Lösung

$$\begin{aligned} x(t) &= u_1(t) \cos(2t) + u_2(t) \sin(2t) \\ &= \frac{1}{2} \cos(2t)^2 + \frac{1}{2} \sin(2t)^2 + c_1 \cos(2t) + c_2 \sin(2t) \\ &= \frac{1}{2} + c_1 \cos(2t) + c_2 \sin(2t) \end{aligned}$$

und die Partikulärlösung

$$x_p(t) = \frac{1}{2}.$$

■

#### Bemerkung 4.3.8

Die hier vorgestellten Techniken und Methoden für inhomogene lineare Differentialgleichungen  $n$ -ter Ordnung mit konstanten Koeffizienten lassen sich auch auf Systeme von linearen Differentialgleichungen übertragen. Details finden sich im Anhang A.3. ■

## 4.4 Die Laplace-Transformation

Wählt man in Beispiel 4.1.7 die spezielle Funktion

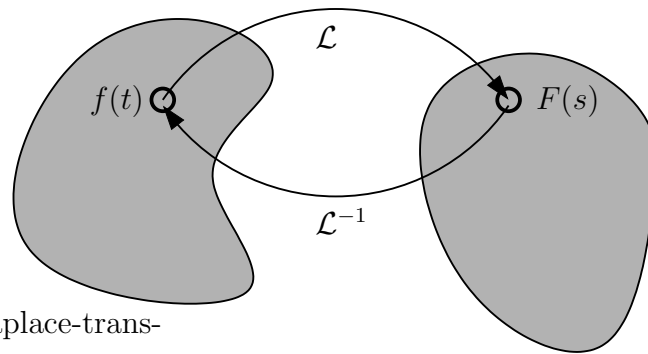
$$K(t, s) = \exp(-st)$$

sowie das unendliche Intervall  $[a, b] = [0, \infty)$ , so erhält man die **Laplace-Transformation**

$$\mathcal{L}\{f\}(s) := \int_0^{\infty} f(t) \exp(-st) dt.$$

Diese ordnet einer Funktion  $f : \mathbb{R}_+ \rightarrow \mathbb{R}$  die **Laplace-Transformierte**  $F(s) = \mathcal{L}\{f\}(s)$  zu. Dieser Operator ist linear, aber die Existenz der Laplace-Transformierten ist nicht für alle  $s$  gesichert. Daher bezeichnen wir  $f : \mathbb{R}_+ \rightarrow \mathbb{R}$  als **Laplace-transformierbar**, wenn  $F(s)$  für mindestens ein  $s \in \mathbb{R}$  existiert.

Die Laplace-Transformation ist eineindeutig (bijektiv), d.h. die Laplace-Transformierte  $F = \mathcal{L}\{f\}$  ist eindeutig durch  $f$  bestimmt und umgekehrt (Existenz vorausgesetzt). Insbesondere existiert die Inverse Laplace-Transformation  $f = \mathcal{L}^{-1}\{F\}$ , welche ebenfalls linear ist.



Menge der Laplace-transformierbaren Funktionen

Menge der Laplace-Transformierten

Abbildung 4.2: Laplace-Transformation und inverse Laplace-Transformation

Wir betrachten einige Laplace-Transformationen.

**Beispiel 4.4.1**

(i)  $f(t) = 1 \Rightarrow$

$$\mathcal{L}\{1\} = F(s) = \int_0^{\infty} \exp(-st) dt = \frac{1}{s}, \quad s > 0.$$

(ii)  $f(t) = t^n$  mit  $n \in \mathbb{N}$ :

Durch partielle Integration ergibt sich die Rekursionsvorschrift:

$$\begin{aligned} \mathcal{L}\{t^n\} &= \int_0^{+\infty} t^n \exp(-st) dt = \left[ \frac{1}{-s} t^n \exp(-st) \right]_0^{+\infty} + \frac{n}{s} \int_0^{\infty} t^{n-1} \exp(-st) dt \\ &= \frac{n}{s} \mathcal{L}\{t^{n-1}\} \text{ f\"ur } n > 0 \text{ und } s > 0. \end{aligned}$$

Mit (i) gilt dann

$$\begin{aligned} n = 1 & : \mathcal{L}\{t\} = \frac{1}{s} \mathcal{L}\{1\} = \frac{1}{s^2}, \quad s > 0 \\ n = 2 & : \mathcal{L}\{t^2\} = \frac{2}{s} \mathcal{L}\{t\} = \frac{2}{s^3} = \frac{2!}{s^3}, \quad s > 0 \\ n = 3 & : \mathcal{L}\{t^3\} = \frac{3}{s} \mathcal{L}\{t^2\} = \frac{6}{s^4} = \frac{3!}{s^4}, \quad s > 0 \\ & \vdots \\ n & : \mathcal{L}\{t^n\} = \frac{n!}{s^{n+1}} \quad \text{f\"ur } s > 0 \end{aligned}$$

(iii)  $f(t) = \exp(t) \Rightarrow$

$$\mathcal{L}\{\exp(t)\} = F(s) = \int_0^{\infty} \exp(t) \exp(-st) dt = \int_0^{\infty} \exp(-(s-1)t) dt = \frac{1}{s-1}, \quad s > 1.$$

$$f(t) = \exp(-t) \Rightarrow$$

$$\mathcal{L}\{\exp(-t)\} = F(s) = \int_0^{\infty} \exp(-t) \exp(-st) dt = \int_0^{\infty} \exp(-(s+1)t) dt = \frac{1}{s+1}, \quad s > -1.$$

(iv)  $f(t) = \sin(t) \Rightarrow$  Durch zweimalige partielle Integration ergibt sich

$$\mathcal{L}\{\sin(t)\} = F(s) = \int_0^{\infty} \exp(-st) \sin(t) dt = \dots = 1 - s^2 F(s).$$

Auflösen nach  $F(s)$ :

$$\mathcal{L}\{\sin(t)\} = F(s) = \frac{1}{s^2 + 1}, \quad s > 0.$$

(v)  $f(t) = \cos(t) \Rightarrow$  Durch partielle Integration ergibt sich

$$\mathcal{L}\{\cos(t)\} = F(s) = \int_0^{\infty} \exp(-st) \cos(t) dt = \dots = s \mathcal{L}\{\sin(t)\} = \frac{s}{s^2 + 1}$$

für  $s > 0$ .

■

Es gibt zahlreiche Rechenregeln für die Laplacetransformation. Wir betrachten hier nur eine, die es uns erlaubt, lineare Differentialgleichungen mit konstanten Koeffizienten zu lösen. Dabei geht man wie folgt vor:

- (i) Wende die Laplace-Transformation auf eine Differentialgleichung an.
- (ii) Löse die resultierende Gleichung für die Laplace-Transformierte  $Y = \mathcal{L}(y)$  der gesuchten Lösung  $y$  der Differentialgleichung.
- (iii) Führe die Rücktransformation durch und bestimme die zur Laplace-Transformierten  $Y$  gehörende Funktion  $y = \mathcal{L}^{-1}(Y)$ .

Der große Vorteil der Laplace-Transformation ist, dass die Differentialgleichung für  $y$  in Schritt (ii) in eine **algebraische Gleichung für  $Y$**  transformiert wird. Die resultierende algebraische Gleichung kann oft bzgl.  $Y$  gelöst werden. Die Hauptschwierigkeit besteht in der Rücktransformation

$$y = \mathcal{L}^{-1}\{Y\},$$

dem **inversen Problem der Laplace-Transformation**. Es gibt jedoch umfangreiche Tabellen von Funktionen und ihren Laplace-Transformationen mit deren Hilfe man die Rücktransformation in (iii) häufig durchführen kann, vgl. Tabelle A.2 im Anhang.

Wir möchten nun eine lineare Differentialgleichung mit konstanten Koeffizienten mithilfe der Laplacetransformation lösen und verdeutlichen die Vorgehensweise an einem konkreten Beispiel.

Gesucht ist die Lösung  $y$  des Anfangswertproblems

$$y''(t) + y(t) = f(t), \quad y(0) = 2, \quad y'(0) = 1,$$

mit  $f(t) = \sin(2t)$ .

Da die Laplacetransformation ein linearer Operator ist, können wir diese auf die Differentialgleichung anwenden und erhalten die Gleichung

$$\mathcal{L}(y'') + \mathcal{L}(y) = \mathcal{L}(f). \quad (4.8)$$

Die Laplace-Transformierte der Lösung  $y$  bezeichnen wir mit  $Y = \mathcal{L}(y)$  und die Laplace-Transformierte von  $f$  können wir in Tabelle A.2 nachschauen oder selber ausrechnen:

$$\mathcal{L}(f) = \frac{2}{s^2 + 4}.$$

Allerdings benötigen wir auch die Laplace-Transformierte von  $y''$ . Wie sieht die aus?

Mittels partieller Integration ergibt sich

$$\begin{aligned} \mathcal{L}\{y\} &= \int_0^{+\infty} \underbrace{y(t)}_u \underbrace{\exp(-st)}_{v'} dt = \left[ \frac{1}{-s} y(t) \exp(-st) \right]_0^{+\infty} - \int_0^{+\infty} y'(t) \frac{\exp(-st)}{-s} dt \\ &= \frac{y(0)}{s} + \frac{1}{s} \underbrace{\int_0^{+\infty} y'(t) \exp(-st) dt}_{=\mathcal{L}\{y'\}}, \quad s > 0. \end{aligned}$$

Damit folgt für  $s > 0$

$$\mathcal{L}\{y'\} = -y(0) + s\mathcal{L}\{y\}. \quad (4.9)$$

Rekursive Anwendung der Vorschrift (4.9) liefert dann Formeln für höhere Ableitungen:

$$\begin{aligned} \mathcal{L}\{y''\} = \mathcal{L}\{(y')'\} &= -y'(0) + s\mathcal{L}\{y'\} \\ &= -y'(0) + s(-y(0) + s\mathcal{L}\{y\}) \\ &= s^2\mathcal{L}\{y\} - (sy(0) + y'(0)), \\ \mathcal{L}\{y'''\} &= s^3\mathcal{L}\{y\} - (s^2y(0) + sy'(0) + y''(0)) \\ &\vdots \end{aligned}$$

Insgesamt:



$$\mathcal{L}\{y^{(n)}\} = s^n \mathcal{L}\{y\} - (s^{n-1}y(0) + \dots + sy^{(n-2)}(0) + y^{(n-1)}(0)).$$

Damit können wir nun die transformierte Differentialgleichung (4.8) formulieren:

$$s^2Y(s) - sy(0) - y'(0) + Y(s) = \frac{2}{s^2 + 4}.$$

Einsetzen der Anfangswerte  $y(0) = 2$  und  $y'(0) = 1$  und Auflösen nach  $Y(s)$  liefert

$$Y(s) = \frac{2s^3 + s^2 + 8s + 6}{(s^2 + 1)(s^2 + 4)}.$$

Partialbruchzerlegung mit dem Ansatz

$$Y(s) = \frac{as + b}{s^2 + 1} + \frac{cs + d}{s^2 + 4}$$

und Koeffizientenvergleich liefert  $a = 2, c = 0, b = 5/3, d = -2/3$  und

$$Y(s) = \frac{2s}{s^2 + 1} + \frac{5/3}{s^2 + 1} - \frac{2/3}{s^2 + 4}.$$

Nachschauen in der Tabelle A.2 im Anhang liefert die inverse Transformation

$$y(t) = 2 \cos(t) + \frac{5}{3} \sin(t) - \frac{1}{3} \sin(2t),$$

was die Lösung für das Anfangswertproblem ist.

#### **Bemerkung 4.4.2**

*Die obige Vorgehensweise lässt sich prinzipiell für inhomogene lineare Differentialgleichungen  $n$ -ter Ordnung mit konstanten Koeffizienten durchführen. Man kann die Laplace-Transformation analog auf Systeme von linearen Differentialgleichungen anwenden. Dies führt auf ein lineares Gleichungssystem für die Laplace-Transformierten, welches gelöst werden muss. Danach erfolgt dann die Rücktransformation wie üblich. ■*

# Kapitel A

## Anhang

### A.1 Diskrete Fourier-Transformation

Für eine gegebene Funktion  $f$  sind die Fourierkoeffizienten  $\hat{\alpha}_k$ ,  $k = 0, \pm 1, \dots, \pm n$ , in Satz 3.4.5 i.a. schwer zu berechnen, da die Integrale i.a. nicht analytisch berechenbar sind. Daher bietet es sich an, diese durch ein geeignetes Verfahren zur Approximation von Integralen zu approximieren. Hierzu eignet sich auf Grund ihrer periodischen Struktur insbesondere die zusammengesetzte Trapezregel, welche ein Verfahren zur Approximation von bestimmten Integralen der Form

$$I[f] := \int_a^b f(x) dx, \quad a, b \in \mathbb{R}, a < b,$$

ist. Dazu wird das Intervall  $[a, b]$  durch Einführung von äquidistanten Stützstellen

$$x_i = a + ih, \quad i = 0, \dots, n, \quad h = \frac{b - a}{n} \tag{A.1}$$

in Teilintervalle  $[x_i, x_{i+1}]$  zerlegt und auf jedem Teilintervall wird die Funktion  $f$  durch eine stetige und stückweise lineare Funktion approximiert, vgl. Abbildung A.1.

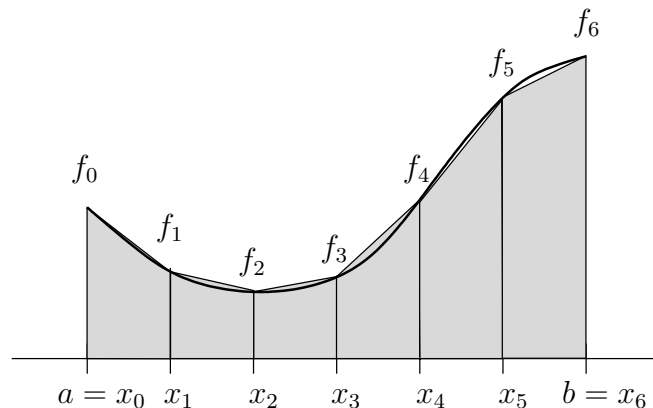


Abbildung A.1: Zusammengesetzte Trapezregel für  $n = 6$ .

Intervallweises Ausrechnen der Flächeninhalte (Trapeze) für die stückweise lineare Funk-

tion liefert die **zusammengesetzte Trapezregel** oder **iterierte Trapezregel**

$$T_n[f] := \frac{h}{2}f(a) + h \sum_{i=1}^{n-1} f(x_i) + \frac{h}{2}f(b).$$

Wir wenden nun die zusammengesetzte Trapezregel mit äquidistanten Stützstellen

$$x_k = \frac{2\pi k}{N}, \quad k = 0, \dots, N,$$

mit  $N \in \mathbb{N}$  an, um die Fourierkoeffizienten  $\hat{\alpha}$  für eine gegebene  $2\pi$ -periodische Funktion  $f$  zu approximieren.

Mit der Abkürzung  $g(x) = f(x)\overline{\exp(ikx)}$  erhalten wir die Approximationen

$$\begin{aligned} \hat{\alpha}_k &= \frac{1}{2\pi} \int_0^{2\pi} g(x) dx \\ &\approx \frac{1}{2\pi} \frac{2\pi}{N} \left( \frac{1}{2}g(x_0) + \sum_{j=1}^{N-1} g(x_j) + \frac{1}{2}g(x_N) \right) \\ &= \frac{1}{N} \sum_{j=0}^{N-1} g(x_j) \\ &= \frac{1}{N} \sum_{j=0}^{N-1} f(x_j) \overline{\exp(ikx_j)} \end{aligned}$$

für  $k = 0, \pm 1, \dots, \pm n$ . Es wird sich herausstellen, dass diese diskreten Fourierkoeffizienten eine analoge Rolle wie die kontinuierlichen Fourierkoeffizienten bei der Approximation von Funktionen auf dem Gitter  $\{x_0, \dots, x_N\}$  spielen. Um dies zu sehen, definieren wir

$$\begin{aligned} \omega(x) &= \exp(ix), \\ \omega_j^k &= \exp(ikx_j) = \omega(x_j)^k, \quad k = 0, \dots, n-1, \quad j = 0, \dots, N-1. \end{aligned}$$

Es gilt:

### Satz 1.1.1

Die Vektoren

$$\omega^k = \begin{pmatrix} \omega_0^k \\ \vdots \\ \omega_{N-1}^k \end{pmatrix}, \quad k = 0, \dots, n-1, \quad (\text{A.2})$$

bilden ein Orthonormalsystem in  $\mathbb{C}^N$  bzgl. des Skalarproduktes

$$\langle y, z \rangle := \frac{1}{N} z^* y.$$

**Beweis:** Für  $k \in \{0, \dots, n-1\}$  gilt

$$\langle \omega^k, \omega^k \rangle = \frac{1}{N} \sum_{j=0}^{N-1} \overline{\exp(ikx_j)} \exp(ikx_j) = \frac{1}{N} \sum_{j=0}^{N-1} 1 = 1.$$

Für  $k \neq \ell$  gilt

$$\begin{aligned} \langle \omega^\ell, \omega^k \rangle &= \frac{1}{N} \sum_{j=0}^{N-1} \overline{\exp(ikx_j)} \exp(i\ell x_j) \\ &= \frac{1}{N} \sum_{j=0}^{N-1} \exp(i(\ell - k)x_j) \\ &= \frac{1}{N} \sum_{j=0}^{N-1} \exp(i(\ell - k)2\pi/N)^j \\ &= \frac{1}{N} \cdot \frac{1 - \exp(i(\ell - k)2\pi/N)^N}{1 - \exp(i(\ell - k)2\pi/N)} \\ &= \frac{1}{N} \cdot \frac{1 - 1}{1 - \exp(i(\ell - k)2\pi/N)} \\ &= 0. \end{aligned}$$

■

Wir möchten nun gegebene Funktionswerte  $f_j$  an den Stützstellen  $x_j$ ,  $j = 0, \dots, N-1$ , bestmöglich approximieren. Wir erhalten:

### Satz 1.1.2

Seien  $f_j$ ,  $j = 0, \dots, N-1$ , gegebene komplexe Zahlen,  $n \leq N$  und

$$\hat{\alpha}_k = \frac{1}{N} \sum_{j=0}^{N-1} f_j \overline{\exp(ikx_j)} = \frac{1}{N} \sum_{j=0}^{N-1} f_j \overline{\omega_j^k}, \quad k = 0, \dots, n-1.$$

Dann ist

$$\hat{t} = \sum_{k=0}^{n-1} \hat{\alpha}_k \omega^k$$

die eindeutig bestimmte Bestapproximierende an  $f = (f_0, \dots, f_{N-1})^\top$  in dem durch  $\{\omega^k \mid k = 0, \dots, n-1\}$  aufgespannten Vektorraum bzgl. der euklidischen Norm, d.h. es gilt

$$\|f - \hat{t}\|_2 \leq \|f - t\|_2 \quad \forall t = \sum_{k=0}^{n-1} \alpha_k \omega^k.$$

Im Fall  $n = N$  gilt  $\|f - \hat{t}\|_2 = 0$ , d.h.

$$f_j = \sum_{k=0}^{n-1} \hat{\alpha}_k \omega_j^k, \quad j = 0, \dots, N-1.$$

**Beweis:** Für  $\ell = 0, \dots, n-1$  gilt

$$\langle f - \hat{t}, \omega^\ell \rangle = \langle f - \sum_{k=0}^{n-1} \hat{\alpha}_k \omega^k, \omega^\ell \rangle = \langle f, \omega^\ell \rangle - \sum_{k=0}^{n-1} \hat{\alpha}_k \langle \omega^k, \omega^\ell \rangle = \langle f, \omega^\ell \rangle - \hat{\alpha}_\ell = 0.$$

Nach dem Projektionssatz 3.2.8 ist  $\hat{t}$  also die eindeutig bestimmte Bestapproximierende an  $f$ . Im Fall  $n = N$  ist  $f \in \text{span}\{\omega^0, \dots, \omega^{n-1}\}$ , woraus die Behauptung folgt. ■

Wir betrachten nun den Fall  $N = n$  etwas genauer. Wegen

$$\sum_{k=0}^{n-1} \hat{\alpha}_k \exp(ikx_j) = \sum_{k=0}^{n-1} \hat{\alpha}_k \omega_j^k = f_j, \quad j = 0, \dots, N-1,$$

definiert

$$\hat{t}(x) = \sum_{k=0}^{n-1} \hat{\alpha}_k \exp(ikx) \tag{A.3}$$

ein interpolierendes trigonometrisches Polynom mit  $\hat{t}(x_j) = f_j$ ,  $j = 0, \dots, N-1$ . Definiere für  $k = 0, 1, \dots, n-1$

$$A_k = \frac{1}{n} \sum_{j=0}^{n-1} f_j \cos(kx_j), \quad B_k = \frac{1}{n} \sum_{j=0}^{n-1} f_j \sin(kx_j).$$

Damit gilt

$$\begin{aligned} A_0 &= \frac{1}{n} \sum_{j=0}^{n-1} f_j, & B_0 &= 0, \\ A_{n/2} &= \frac{1}{n} \sum_{j=0}^{n-1} (-1)^j f_j, & B_{n/2} &= 0 \quad (\text{falls } n \text{ gerade}), \\ \hat{\alpha}_k &= A_k - iB_k, & k &= 1, \dots, n-1, \\ \hat{\alpha}_{n-k} &= A_k + iB_k, & k &= 1, \dots, n-1. \end{aligned}$$

Weiter folgt für  $k = 1, \dots, n-1$

$$\begin{aligned} &\hat{\alpha}_k \exp(ikx) + \hat{\alpha}_{n-k} \exp(i(n-k)x) \\ &= (A_k - iB_k)(\cos(kx) + i \sin(kx)) + (A_k + iB_k)(\cos(kx) - i \sin(kx)) \\ &= 2(A_k \cos(kx) + B_k \sin(kx)). \end{aligned}$$

Zusammen mit dem interpolierenden trigonometrischen Polynom (A.3) erhalten wir daraus folgendes Resultat, welches den Begriff trigonometrische Interpolation rechtfertigt.

**Satz 1.1.3** (trigonometrische Interpolation)

Seien  $f_j$ ,  $j = 0, \dots, n-1$ , gegebene komplexe Zahlen und

$$\begin{aligned} a_0 &= \frac{1}{n} \sum_{j=0}^{n-1} f_j, \\ a_{n/2} &= \frac{1}{n} \sum_{j=0}^{n-1} (-1)^j f_j \quad \text{für } n \text{ gerade,} \\ a_k &= \frac{2}{n} \sum_{j=0}^{n-1} f_j \cos(kx_j), \quad k \in \mathbb{N}, k < \frac{n}{2}, \\ b_k &= \frac{2}{n} \sum_{j=0}^{n-1} f_j \sin(kx_j), \quad k \in \mathbb{N}, k < \frac{n}{2}. \end{aligned}$$

Setze für  $n = 2\ell + 1$  ( $n$  ungerade)

$$\hat{t}(x) = a_0 + \sum_{k=1}^{\ell} (a_k \cos(kx) + b_k \sin(kx))$$

bzw. für  $n = 2\ell$  ( $n$  gerade)

$$\hat{t}(x) = a_0 + \sum_{k=1}^{\ell-1} (a_k \cos(kx) + b_k \sin(kx)) + a_{\ell} \cos(\ell x).$$

Dann gilt  $\hat{t}(x_j) = f_j$  für  $j = 0, \dots, n-1$ .

## A.2 Tabellen zur Laplacetransformation

Nr.	Zeitfunktion $y(t)$	Transformierte $Y(s)$
1	1	$\frac{1}{s}$
2	$t^n$	$\frac{n!}{s^{n+1}}$
3	$\exp(-at)$	$\frac{1}{s+a}$
4	$\frac{1}{T} \exp\left(-\frac{t}{T}\right)$	$\frac{1}{1+sT}$
5	$\frac{1}{a} (1 - \exp(-at))$	$\frac{1}{s(s+a)}$
6	$1 - \exp(-t/T)$	$\frac{1}{s(1+sT)}$
7	$\frac{1}{b-a} (\exp(-at) - \exp(-bt))$	$\frac{1}{(s+a)(s+b)}$
8	$\frac{\exp(-t/T_1) - \exp(-t/T_2)}{T_1 - T_2}$	$\frac{1}{(1+sT_1)(1+sT_2)}$
9	$t \cdot \exp(-at)$	$\frac{1}{(s+a)^2}$
10	$\frac{1}{T^2} \cdot t \cdot \exp(-t/T)$	$\frac{1}{(1+sT)^2}$
11	$\frac{a}{b} + \frac{b-a}{b} \cdot \exp(-bt)$	$\frac{(s+a)}{s(s+b)}$
12	$1 + \frac{T-T_1}{T_1} \exp(-t/T_1)$	$\frac{(1+sT)}{s(1+sT_1)}$
13	$\frac{1}{a^2} (\exp(-at) - 1 + a \cdot t)$	$\frac{1}{s^2(s+a)}$

Tabelle A.1: Korrespondenzen zur Laplace-Transformation

Nr.	Zeitfunktion $y(t)$	Transformierte $Y(s)$
14	$T \cdot \exp(-t/T) + t - T$	$\frac{1}{s^2(1+sT)}$
15	$\frac{1}{a^2} [1 - (at+1) \exp(-at)]$	$\frac{1}{s(s+a)^2}$
16	$1 - \frac{T+t}{T} \exp(-t/T)$	$\frac{1}{s(1+sT)^2}$
17	$\frac{1}{ab} + \frac{1}{a-b} \left[ \frac{1}{a} \exp(-at) - \frac{1}{b} \exp(-bt) \right]$	$\frac{1}{s(s+a)(s+b)}$
18	$1 + \frac{T_1 \cdot \exp(-t/T_1) - T_2 \cdot \exp(-t/T_2)}{T_2 - T_1}$	$\frac{1}{s(1+sT_1)(1+sT_2)}$
19	$\frac{(c-b) \exp(-at) + (a-c) \exp(-bt) + (b-a) \exp(-ct)}{(a-b)(a-c)(c-b)}$	$\frac{1}{(s+a)(s+b)(s+c)}$
20	$[(T_1 - T_2)(T_1 - T_3)(T_2 - T_3)]^{-1} \cdot$ $[T_1(T_2 - T_3) \exp(-t/T_1) + T_2(T_3 - T_1) \exp(-t/T_2)$ $+ T_3(T_1 - T_2) \exp(-t/T_3)]$	$\frac{1}{(1+sT_1)(1+sT_2)(1+sT_3)}$
21	$\frac{1}{\omega} \cdot \sin(\omega t)$	$\frac{1}{s^2 + \omega^2}$
22	$\cos(\omega t)$	$\frac{s}{s^2 + \omega^2}$
23	$\frac{1}{2\omega} \cdot t \cdot \sin(\omega t)$	$\frac{s}{(s^2 + \omega^2)^2}$
24	$t \cdot \cos(\omega t)$	$\frac{s^2 - \omega^2}{(s^2 + \omega^2)^2}$
25	$\sinh(\omega t)$	$\frac{\omega}{s^2 - \omega^2}$

Tabelle A.2.: Fortsetzung



Nr.	Zeitfunktion $y(x)$	Transformierte $Y(s)$
26	$\cosh(\omega t)$	$\frac{s}{s^2 - \omega^2}$
27	$\frac{1}{\sqrt{1-D^2}\omega} \exp(-D\omega t) \sin(\sqrt{1-D^2} \omega t)$	$\frac{1}{s^2 + 2D\omega s + \omega^2}$
28	$\exp(-D\omega t) \sin(\sqrt{1-D^2} \omega t)$	$\frac{\sqrt{1-D^2} \omega}{s^2 + 2D\omega s + \omega^2}$
29	$\exp(-D\omega t) \cos(\sqrt{1-D^2} \omega t)$	$\frac{s + D\omega}{s^2 + 2D\omega s + \omega^2}$
30	$\exp(-\delta t) \sin(\omega_e t)$	$\frac{\omega_e}{(s + \delta)^2 + \omega_e^2}, \quad \omega_e = \sqrt{\omega - \delta^2}$ $= \sqrt{1 - D^2}\omega$
31	$\exp(-\delta t) \cos(\omega_e t)$	$\frac{s + \delta}{(s + \delta)^2 + \omega_e^2}, \quad \omega_e = \sqrt{\omega - \delta^2}$ $= \sqrt{1 - D^2}\omega$
32	$\exp(-\delta t) \sinh(\omega_e t)$	$\frac{\omega_e}{(s + \delta)^2 - \omega_e^2}, \quad \omega_e = \sqrt{\delta^2 - \omega^2}$
33	$\exp(-\delta t) \cosh(\omega_e t)$	$\frac{s + \delta}{(s + \delta)^2 - \omega_e^2}, \quad \omega_e = \sqrt{\delta^2 - \omega^2}$

Tabelle A.2.: Fortsetzung

### A.3 Lineare Differentialgleichungssysteme 1. Ordnung

In technischen Anwendungen treten häufig Systeme von gekoppelten Differentialgleichungen auf, die sich gegenseitig beeinflussen.

#### Definition 1.3.1 (Lineares Differentialgleichungssystem 1. Ordnung)

Ein lineares Differentialgleichungssystem 1. Ordnung ist von der Form

$$\begin{aligned}x_1'(t) &= a_{11}x_1(t) + a_{12}x_2(t) + \dots + a_{1n}x_n(t) + f_1(t) \\x_2'(t) &= a_{21}x_1(t) + a_{22}x_2(t) + \dots + a_{2n}x_n(t) + f_2(t) \\&\vdots \\x_n'(t) &= a_{n1}x_1(t) + a_{n2}x_2(t) + \dots + a_{nn}x_n(t) + f_n(t).\end{aligned}$$

In Matrixnotation lautet es

$$\vec{x}'(t) = A\vec{x}(t) + \vec{f}(t), \quad (\text{A.4})$$

wobei

$$A = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{bmatrix}, \quad \vec{x}(t) = \begin{bmatrix} x_1(t) \\ \vdots \\ x_n(t) \end{bmatrix}, \quad \vec{f}(t) = \begin{bmatrix} f_1(t) \\ \vdots \\ f_n(t) \end{bmatrix}.$$

Gilt  $\vec{f}(t) = 0$  für alle  $t$ , so heißt das Differentialgleichungssystem **homogen**, ansonsten **inhomogen**.

Jede zeitabhängige Vektorfunktion  $\vec{x}(t)$ , die das lineare Differentialgleichungssystem 1. Ordnung erfüllt, heißt **Lösung**.

#### Bemerkung 1.3.2

Es gibt eine enge Beziehung zwischen der linearen Differentialgleichung  $n$ -ter Ordnung in (4.1) mit konstanten Koeffizienten, d.h.

$$x^{(n)}(t) + a_{n-1}x^{(n-1)}(t) + \dots + a_1x'(t) + a_0x(t) = f(t)$$

und dem linearen Differentialgleichungssystem 1. Ordnung in (A.4).

Definiert man nämlich in (4.1) die Funktionen  $x_k$  für  $k = 1, \dots, n$  durch  $x_k(t) = x^{(k-1)}(t)$ , so erhält man das lineare Differentialgleichungssystem

$$\begin{aligned}x_1'(t) &= x_2(t), \\x_2'(t) &= x_3(t), \\&\vdots \\x_{n-1}'(t) &= x_n(t), \\x_n'(t) &= -a_0x_1(t) - a_1x_2(t) - \dots - a_{n-1}x_n(t) + f(t)\end{aligned}$$

bzw. in Matrixnotation

$$\underbrace{\begin{bmatrix} x_1'(t) \\ \vdots \\ x_{n-1}'(t) \\ x_n'(t) \end{bmatrix}}_{\vec{x}'(t)} = \underbrace{\begin{bmatrix} & & & 1 \\ & & \ddots & \\ & & & 1 \\ -a_0 & -a_1 & \dots & -a_{n-1} \end{bmatrix}}_A \underbrace{\begin{bmatrix} x_1(t) \\ \vdots \\ x_{n-1}(t) \\ x_n(t) \end{bmatrix}}_{\vec{x}(t)} + \underbrace{\begin{bmatrix} 0 \\ \vdots \\ 0 \\ f(t) \end{bmatrix}}_{\vec{f}(t)}.$$

Also kann jede lineare Differentialgleichung  $n$ -ter Ordnung (4.1) als lineares Differentialgleichungssystem erster Ordnung mit einer speziellen Matrix  $A$  und einer speziellen Inhomogenität  $\vec{f}$  geschrieben werden. ■

Analog zu den linearen Differentialgleichungen  $n$ -ter Ordnung gelten das Superpositionsprinzip und die allgemeine Lösungsdarstellung als Summe einer Partikulärlösung und eines Fundamentalsystems.

### Satz 1.3.3 (Superpositionsprinzip)

Sind  $\vec{x}$  und  $\vec{y}$  Lösungen des homogenen linearen Differentialgleichungssystems 1. Ordnung, so ist auch  $c_1\vec{x} + c_2\vec{y}$  für alle  $c_1, c_2 \in \mathbb{R}$  Lösung.

**Beweis:** Es gilt

$$\begin{aligned} (c_1\vec{x}(t) + c_2\vec{y}(t))' &= c_1\vec{x}'(t) + c_2\vec{y}'(t) \\ &= c_1A\vec{x}(t) + c_2A\vec{y}(t) \\ &= A(c_1\vec{x}(t) + c_2\vec{y}(t)). \end{aligned}$$

Die Lösungen des homogenen linearen Differentialgleichungssystems definieren also wieder einen Vektorraum. Der folgende Satz besagt, dass es genau  $n$  linear unabhängige Lösungen des homogenen linearen Differentialgleichungssystems gibt. Dabei heißen die  $m$  Vektorfunktionen  $\vec{x}_1, \dots, \vec{x}_m$  **linear unabhängig im Intervall**  $(a, b)$ , wenn aus der Gleichung

$$c_1\vec{x}_1(t) + \dots + c_m\vec{x}_m(t) = 0 \quad \forall t \in (a, b)$$

stets  $c_1 = \dots = c_m = 0$  folgt.

### Satz 1.3.4 (Allgemeine Lösung des homogenen linearen Differentialgleichungssystems 1. Ordnung)

Es gibt genau  $n$  linear unabhängige Lösungen  $\vec{x}_1, \dots, \vec{x}_n$  des homogenen linearen Differentialgleichungssystems

$$\vec{x}'(t) = A\vec{x}(t).$$

Die allgemeine Lösung lautet

$$\vec{x}(t) = \sum_{k=1}^n c_k \vec{x}_k(t),$$

wobei  $c_1, \dots, c_n$  beliebige reelle Zahlen sind.

### Definition 1.3.5 (Fundamentalsystem)

Jede Menge  $\{\vec{x}_1, \dots, \vec{x}_n\}$  von  $n$  linear unabhängigen Lösungen des homogenen linearen Differentialgleichungssystems heißt **Fundamentalsystem**.

Kennt man noch eine spezielle Lösung der inhomogenen Gleichung, so läßt sich die allgemeine Lösung wie folgt angeben:

### Satz 1.3.6 (Allgemeine Lösung des inhomogenen Differentialgleichungssystems 1. Ordnung)

Sei  $\vec{x}_p$  eine spezielle Lösung (genannt **Partikulärlösung**) der inhomogenen Differentialgleichung (A.4) und  $\{\vec{x}_1, \dots, \vec{x}_n\}$  ein Fundamentalsystem der zugehörigen homogenen Differentialgleichung.

Die allgemeine Lösung des inhomogenen Differentialgleichungssystems (A.4) lautet dann

$$\vec{x}(t) = \vec{x}_p(t) + \sum_{k=1}^n c_k \vec{x}_k(t),$$

wobei  $c_1, \dots, c_n$  beliebige reelle Zahlen sind.

Nach Satz 1.3.6 benötigt man zur Lösung von (A.4) also wieder eine Partikulärlösung  $\vec{x}_p$  sowie ein Fundamentalsystem der zugehörigen homogenen Differentialgleichung.

Wir widmen uns zunächst wieder dem homogenen Fall und berechnen hierfür Fundamentalsysteme. In Abschnitt A.3 werden wir dann Partikulärlösungen berechnen.

### Homogene lineare Differentialgleichungssysteme 1. Ordnung

Im Folgenden gelte  $\vec{f} \equiv 0$  in (A.4), d.h. wir betrachten das Differentialgleichungssystem

$$\vec{x}'(t) = A\vec{x}(t). \tag{A.5}$$

Zur Bestimmung einer Lösung wählen wir den Ansatz

$$\vec{x}(t) = \vec{v} \exp(\lambda t) \quad (\lambda \text{ und } \vec{v} \text{ konstant}).$$

$\vec{x}$  ist differenzierbar mit

$$\vec{x}'(t) = \lambda \vec{v} \exp(\lambda t).$$

Einsetzen dieser Ableitung in die Differentialgleichung liefert die Gleichung

$$\lambda \vec{v} \exp(\lambda t) = A \vec{v} \exp(\lambda t).$$

Wegen  $\exp(\lambda t) > 0$  ist diese Gleichung genau dann erfüllt, wenn  $\lambda$  das folgende **Eigenwertproblem** löst:

$$A \vec{v} = \lambda \vec{v}.$$

### Satz 1.3.7

Ist  $\lambda$  ein Eigenwert von  $A$  und  $\vec{v}$  ein zugehöriger Eigenvektor, so ist

$$\vec{x}(t) = \vec{v} \exp(\lambda t)$$

eine Lösung von (A.5).

Aus der Höheren Mathematik I wissen wir bereits, dass es genau  $n$  Eigenwerte zur Matrix  $A \in \mathbb{R}^{n \times n}$  gibt, wobei diese auch komplexwertig sein können und mehrfach auftreten können. Eigenvektoren zu verschiedenen Eigenwerten sind stets linear unabhängig, so dass die zu verschiedenen Eigenwerten gehörenden Lösungen stets linear unabhängig sind. Wie für lineare Differentialgleichungen  $n$ -ter Ordnung diskutieren wir wieder verschiedene Fälle.

### 1. Fall: Alle Eigenwerte sind reell und einfach

Für die  $n$  verschiedenen reellen Eigenwerte  $\lambda_1, \dots, \lambda_n$  und die zugehörigen linear unabhängigen Eigenvektoren  $\vec{v}_1, \dots, \vec{v}_n$  bilden die Funktionen

$$\vec{x}_1(t) = \vec{v}_1 \exp(\lambda_1 t), \dots, \vec{x}_n(t) = \vec{v}_n \exp(\lambda_n t)$$

ein Fundamentalsystem.

### Beispiel 1.3.8

Betrachte das homogene lineare Differentialgleichungssystem

$$\vec{x}'(t) = A \vec{x}(t) \quad \text{mit} \quad A = \begin{bmatrix} 1 & -1 & 1 \\ -1 & 1 & 1 \\ 1 & 1 & -1 \end{bmatrix}.$$

Die Eigenwerte von  $A$  berechnen sich zu  $\lambda_1 = 1$ ,  $\lambda_2 = 2$  und  $\lambda_3 = -2$ . Zugehörige Eigenvektoren lauten

$$\vec{v}_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \quad \vec{v}_2 = \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}, \quad \vec{v}_3 = \begin{bmatrix} 1 \\ 1 \\ -2 \end{bmatrix}.$$

Das Fundamentalsystem besteht aus den Lösungen

$$\vec{x}_1(t) = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \exp(t), \quad \vec{x}_2(t) = \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix} \exp(2t), \quad \vec{x}_3(t) = \begin{bmatrix} 1 \\ 1 \\ -2 \end{bmatrix} \exp(-2t).$$

Die allgemeine Lösung der Differentialgleichung lautet

$$\vec{x}(t) = c_1 \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \exp(t) + c_2 \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix} \exp(2t) + c_3 \begin{bmatrix} 1 \\ 1 \\ -2 \end{bmatrix} \exp(-2t).$$

■

## 2. Fall: Es treten einfache komplexe Eigenwerte auf

O.B.d.A. sei  $\lambda_1 = a + ib$  ein einfacher komplexer Eigenwert mit Eigenvektor  $\vec{v}_1$ . Dann ist auch der konjugiert komplexe Wert  $\lambda_2 = \bar{\lambda}_1 = a - ib$  ein einfacher Eigenwert und  $\vec{v}_1$  ist ein Eigenvektor zu  $\lambda_2$ .

Die zugehörigen komplexen Lösungen lauten dann

$$\begin{aligned} \vec{z}_1(t) &= \vec{v}_1 \exp((a + ib)t) \\ &= \vec{v}_1 \exp(at) (\cos(bt) + i \sin(bt)), \\ \vec{z}_2(t) &= \vec{v}_1 \exp((a - ib)t) \\ &= \vec{v}_1 \exp(at) (\cos(bt) - i \sin(bt)) \\ &= \vec{z}_1(t). \end{aligned}$$

Wir sind jedoch nicht an komplexen Lösungen interessiert, sondern nur an reellen Lösungen. Diese können wir durch Linearkombinationen von  $\vec{z}_1$  und  $\vec{z}_2$  erhalten, wobei zu beachten ist, dass auch die Eigenvektoren komplexe Vektoren sind. Es gilt

$$\begin{aligned} \vec{x}_1(t) &= \frac{1}{2} (\vec{z}_1(t) + \vec{z}_2(t)) = \operatorname{Re}(\vec{z}_1(t)), \\ \vec{x}_2(t) &= \frac{1}{2i} (\vec{z}_1(t) - \vec{z}_2(t)) = \operatorname{Im}(\vec{z}_1(t)). \end{aligned}$$

Daraus ergeben sich die beiden reellen Lösungen

$$\vec{x}_1(t) = \operatorname{Re}(\vec{z}_1(t)), \quad \vec{x}_2(t) = \operatorname{Im}(\vec{z}_1(t)).$$

### Beispiel 1.3.9

Betrachte das homogene lineare Differentialgleichungssystem

$$\vec{x}'(t) = A\vec{x}(t) \quad \text{mit} \quad A = \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix}.$$

Die Eigenwerte von  $A$  berechnen sich zu  $\lambda_1 = 1 + i$  und  $\lambda_2 = 1 - i$ . Zugehörige Eigenvektoren lauten

$$\vec{v}_1 = \begin{bmatrix} 1 \\ i \end{bmatrix}, \quad \vec{v}_2 = \begin{bmatrix} 1 \\ -i \end{bmatrix}.$$

Damit erhält man die komplexen Lösungen

$$\begin{aligned} \vec{z}_1(t) &= \begin{bmatrix} 1 \\ i \end{bmatrix} \exp(t)(\cos(t) + i \sin(t)) = \left( \begin{bmatrix} 1 \\ 0 \end{bmatrix} + i \begin{bmatrix} 0 \\ 1 \end{bmatrix} \right) \exp(t)(\cos(t) + i \sin(t)), \\ \vec{z}_2(t) &= \begin{bmatrix} 1 \\ -i \end{bmatrix} \exp(t)(\cos(t) - i \sin(t)) = \left( \begin{bmatrix} 1 \\ 0 \end{bmatrix} - i \begin{bmatrix} 0 \\ 1 \end{bmatrix} \right) \exp(t)(\cos(t) - i \sin(t)) \end{aligned}$$

und die reellen Lösungen

$$\begin{aligned} \vec{x}_1(t) &= \begin{bmatrix} 1 \\ 0 \end{bmatrix} \exp(t) \cos(t) - \begin{bmatrix} 0 \\ 1 \end{bmatrix} \exp(t) \sin(t), \\ \vec{x}_2(t) &= \begin{bmatrix} 1 \\ 0 \end{bmatrix} \exp(t) \sin(t) + \begin{bmatrix} 0 \\ 1 \end{bmatrix} \exp(t) \cos(t). \end{aligned}$$

Die allgemeine Lösung lautet damit

$$\vec{x}(t) = \exp(t) \begin{bmatrix} c_1 \cos(t) + c_2 \sin(t) \\ -c_1 \sin(t) + c_2 \cos(t) \end{bmatrix}.$$

■

### 3. Fall: Es treten mehrfache reelle Eigenwerte auf und algebraische und geometrische Vielfachheit sind gleich

Ist  $\lambda$  ein  $k$ -facher reeller Eigenwert und stimmen zusätzlich geometrische und algebraische Vielfachheit überein (wichtig!), so gibt es zu diesem Eigenwert  $k$  linear unabhängige Eigenvektoren  $\vec{v}_1, \dots, \vec{v}_k$ . Die Funktionen

$$\vec{x}_1(t) = \vec{v}_1 \exp(\lambda t), \dots, \vec{x}_k(t) = \vec{v}_k \exp(\lambda t)$$

sind dann linear unabhängige Lösungen des homogenen Differentialgleichungssystems.

**Beispiel 1.3.10**

Betrachte das homogene lineare Differentialgleichungssystem

$$\vec{x}'(t) = A\vec{x}(t) \quad \text{mit} \quad A = \begin{bmatrix} -\frac{1}{2} & 0 & \frac{1}{2} \\ 0 & -1 & 0 \\ \frac{1}{2} & 0 & -\frac{1}{2} \end{bmatrix}.$$

Die Eigenwerte von  $A$  lauten  $\lambda_1 = 0$  und  $\lambda_2 = -1$  (doppelter Eigenwert). Ein Eigenvektor zum Eigenwert  $\lambda_1 = 0$  lautet

$$\vec{v}_1 = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}.$$

Zum Eigenwert  $\lambda_2 = -1$  gehören die linear unabhängigen Eigenvektoren

$$\vec{v}_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \quad \vec{v}_3 = \begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix}.$$

Das Fundamentalsystem besteht aus den Lösungen

$$\vec{x}_1(t) = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \quad \vec{x}_2(t) = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \exp(-t), \quad \vec{x}_3(t) = \begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix} \exp(-t).$$

Die allgemeine Lösung der Differentialgleichung lautet

$$\vec{x}(t) = c_1 \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} + c_2 \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \exp(-t) + c_3 \begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix} \exp(-t).$$

■

**4. Fall: Es treten mehrfache komplexe Eigenwerte auf und algebraische und geometrische Vielfachheit sind gleich**

Dieser Fall kann analog zum 3. Fall behandelt werden, wobei die Lösungsdarstellungen aus dem 2. Fall zu Grunde liegen.

**5. Fall: Es treten mehrfache reelle oder komplexe Eigenwerte auf und algebraische und geometrische Vielfachheit sind nicht (!) gleich**



Ist  $\lambda$  ein  $k$ -facher reeller Eigenwert, wobei geometrische und algebraische Vielfachheit **nicht** übereinstimmen, so ist die Dimension des Eigenraums zum Eigenwert  $\lambda$  **kleiner als**  $k$ . Es gibt also nicht genügend linear unabhängige Eigenvektoren zum Eigenwert  $\lambda$ , um  $k$  linear unabhängige Lösungen zum Eigenwert  $\lambda$  angeben zu können.

**Beispiel 1.3.11** (vgl. HM I)

Die Matrix

$$A = \begin{bmatrix} \lambda & 1 & 0 \\ 0 & \lambda & 1 \\ 0 & 0 & \lambda \end{bmatrix} \in \mathbb{R}^{3 \times 3}$$

besitzt den Eigenwert  $\lambda$  mit algebraischer Vielfachheit 3 und geometrischer Vielfachheit 1. Der Eigenraum ist eindimensional und wird durch den Vektor  $[1, 0, 0]^\top$  aufgespannt. ■

In diesem Fall muss man die Eigenvektoren durch sogenannte **Hauptvektoren** ergänzen. Jeder Hauptvektor der Stufe 1 ist dabei ein Eigenvektor.

**Definition 1.3.12** (Hauptvektor der Stufe  $k$ )

Ein Vektor  $\vec{v}$  heißt **Hauptvektor der Stufe  $k$**  zum Eigenwert  $\lambda$  der Matrix  $A$ , wenn

$$(A - \lambda I)^k \vec{v} = \vec{0} \quad \text{und} \quad (A - \lambda I)^{k-1} \vec{v} \neq \vec{0}$$

gilt.

In der Höheren Mathematik I haben wir gelernt, dass jede Matrix eine Jordan-Zerlegung besitzt (Satz 4.1.19), woraus sich folgende Aussagen ableiten lassen:

**Satz 1.3.13**

(a) Zu jedem Eigenwert  $\lambda$  mit algebraischer Vielfachheit  $k$  gibt es  $k$  linear unabhängige Hauptvektoren.

(b) Hauptvektoren zu verschiedenen Eigenwerten sind linear unabhängig.

Nach (a) und (b) besitzt eine  $n \times n$ -Matrix damit stets  $n$  linear unabhängige Hauptvektoren.

Hat man einen Hauptvektor bestimmt, so ergibt sich damit die folgende Lösung des homogenen linearen Differentialgleichungssystems (verifiziere die Lösung durch Einsetzen in die Differentialgleichung).

**Satz 1.3.14**

Ist  $\vec{v}$  ein Hauptvektor der Stufe  $k$  zum Eigenwert  $\lambda$  der  $n \times n$ -Matrix  $A$ , so ist

$$\vec{x}(t) = \exp(\lambda t) \sum_{j=0}^{k-1} \frac{t^j}{j!} (A - \lambda I)^j \vec{v}$$

eine Lösung des homogenen linearen Differentialgleichungssystems (A.5).

**Beispiel 1.3.15**

Betrachte das homogene lineare Differentialgleichungssystem

$$\vec{x}'(t) = A\vec{x}(t) \quad \text{mit} \quad A = \begin{bmatrix} 1 & 0 & 0 & 0 \\ -1 & 2 & 0 & 0 \\ -1 & 1 & 2 & 0 \\ -1 & 1 & 0 & 2 \end{bmatrix}.$$

Die Eigenwerte der unteren Dreiecksmatrix  $A$  lassen sich leicht auf der Diagonalen ablesen:  $\lambda_1 = 1$  ist einfacher Eigenwert und  $\lambda_2 = 2$  ist dreifacher Eigenwert (die algebraische Vielfachheit ist also 3).

Zum Eigenwert  $\lambda_1 = 1$  gehört der Eigenvektor  $\vec{v}_1$  und zum Eigenwert  $\lambda_2$  gehören die linear unabhängigen Eigenvektoren  $\vec{v}_2$  und  $\vec{v}_3$ , wobei die Dimension des Eigenraums nur 2 ist (die geometrische Vielfachheit ist also 2):

$$\vec{v}_1 = \begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \quad \vec{v}_2 = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix}, \quad \vec{v}_3 = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}.$$

Für den Eigenwert  $\lambda_2$  müssen wir also einen Hauptvektor der Stufe 2 berechnen und dafür das lineare Gleichungssystem  $(A - \lambda_2 I)^2 \vec{v}_4 = \vec{0}$  lösen, also

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} v_{41} \\ v_{42} \\ v_{43} \\ v_{44} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}.$$

Es muss  $v_{41} = 0$  gelten. Zusätzlich sind wir nur an einem zu  $\vec{v}_2$  und  $\vec{v}_3$  linear unabhängigen Hauptvektor interessiert, so dass  $v_{42} \neq 0$  gelten muss. Wir können also etwa

$$\vec{v}_4 = \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix}$$

wählen. Damit stellen die folgenden vier linear unabhängigen Lösungen ein Fundamentalsystem dar, vgl. Satz 1.3.14:

$$\vec{x}_1(t) = \begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \end{bmatrix} \exp(t), \quad \vec{x}_2(t) = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix} \exp(2t), \quad \vec{x}_3(t) = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix} \exp(2t),$$

sowie

$$\begin{aligned} \vec{x}_4(t) &= \exp(2t) (\vec{v}_4 + t(A - \lambda_2 I)\vec{v}_4) \\ &= \exp(2t) \left( \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix} + t \begin{bmatrix} -1 & 0 & 0 & 0 \\ -1 & 0 & 0 & 0 \\ -1 & 1 & 0 & 0 \\ -1 & 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix} \right) = \begin{bmatrix} 0 \\ 1 \\ t \\ t \end{bmatrix} \exp(2t). \end{aligned}$$

■

Die Berechnung der Hauptvektoren ist mitunter nicht so einfach. Hat man jedoch die Hauptvektoren bestimmt, so lassen sich die Lösungen in Satz 1.3.14 einfacher angeben:

$$\begin{aligned} \vec{x}_1(t) &= \exp(\lambda t)\vec{v}_1, \\ \vec{x}_2(t) &= \exp(\lambda t) (\vec{v}_2 + t(A - \lambda I)\vec{v}_2) = \exp(\lambda t) (\vec{v}_2 + t\vec{v}_1), \\ &\vdots \\ \vec{x}_k(t) &= \exp(\lambda t) \left( \vec{v}_k + t\vec{v}_{k-1} + \dots + \frac{t^{k-1}}{(k-1)!}\vec{v}_1 \right). \end{aligned}$$

Wir betrachten noch den

### Spezialfall: Die geometrische Vielfachheit des Eigenwerts ist eins.

Für diesen Fall, in dem der Eigenraum also eindimensional ist, wurde in Bemerkung 4.1.20 in der Höheren Mathematik I angedeutet, wie Hauptvektoren einfach berechnet werden können. Dazu kann man beginnend mit einem Eigenvektor  $\vec{v}_1$  mit

$$(A - \lambda I)\vec{v}_1 = \vec{0}$$

die **Hauptvektorkette**

$$(A - \lambda I)\vec{v}_\ell = \vec{v}_{\ell-1}, \quad \ell = 2, \dots, k,$$

lösen. Hierzu muss in jedem Schritt ein lineares Gleichungssystem mit der nicht invertierbaren Matrix  $A - \lambda I$  gelöst werden. Der Lösungsvektor  $\vec{v}_\ell$  mit  $2 \leq \ell \leq k$  ist dann ein Hauptvektor der Stufe  $\ell$ .

**Beispiel 1.3.16**

Wir betrachten wiederum die Matrix

$$A = \begin{bmatrix} \lambda & 1 & 0 \\ 0 & \lambda & 1 \\ 0 & 0 & \lambda \end{bmatrix} \in \mathbb{R}^{3 \times 3}$$

mit dem 3-fachen Eigenwert  $\lambda$  und Eigenvektor  $\vec{v}_1 = [1, 0, 0]^\top$ . Zur Bestimmung eines Hauptvektors der Stufe 2 muss das folgende lineare Gleichungssystem gelöst werden:

$$\begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix} \vec{v}_2 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}.$$

Dieses Gleichungssystem besitzt z.B. die Lösung  $\vec{v}_2 = [0, 1, 0]^\top$ .

Zur Bestimmung eines Hauptvektors der Stufe 3 muss das folgende lineare Gleichungssystem gelöst werden:

$$\begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix} \vec{v}_3 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}.$$

Dieses Gleichungssystem besitzt z.B. die Lösung  $\vec{v}_3 = [0, 0, 1]^\top$ .

Die folgenden Lösungen des homogenen linearen Differentialgleichungssystems  $\vec{x}'(t) = A\vec{x}(t)$  stellen demnach ein Fundamentalsystem dar:

$$\vec{x}_1(t) = \exp(\lambda t) \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \quad \vec{x}_2(t) = \exp(\lambda t) \left( \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} + t \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \right) = \exp(\lambda t) \begin{bmatrix} t \\ 1 \\ 0 \end{bmatrix}$$

und

$$\vec{x}_3(t) = \exp(\lambda t) \left( \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} + t \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} + \frac{1}{2}t^2 \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \right) = \exp(\lambda t) \begin{bmatrix} \frac{1}{2}t^2 \\ t \\ 1 \end{bmatrix}$$

■

### Bestimmung einer Partikulärlösung für inhomogene lineare Differentialgleichungssysteme

Nachdem wir in Abschnitt A.3 die Lösung des homogenen linearen Differentialgleichungssystems (A.5) bestimmt haben, widmen wir uns nun der Berechnung einer Partikulärlösung für das inhomogene System (A.4). Dies geschieht wieder mittels **Variation der Konstanten**.

Seien linear unabhängige Lösungen  $\vec{x}_k$ ,  $k = 1, \dots, n$ , des homogenen linearen Differentialgleichungssystems gegeben. Für die zu bestimmende Partikulärlösung  $\vec{x}_p$  wählen wir den Ansatz

$$\vec{x}_p(t) = \sum_{k=1}^n u_k(t) \vec{x}_k(t)$$

mit noch zu bestimmenden Koeffizientenfunktionen  $u_k(t)$ ,  $k = 1, \dots, n$ . Einsetzen in die Differentialgleichung liefert

$$\begin{aligned} \vec{x}_p'(t) - A\vec{x}_p(t) - \vec{f}(t) &= \sum_{k=1}^n \left( u_k'(t) \vec{x}_k(t) + u_k(t) \underbrace{\vec{x}_k'(t)}_{=A\vec{x}_k(t)} \right) - \sum_{k=1}^n u_k(t) A\vec{x}_k(t) - \vec{f}(t) \\ &= \sum_{k=1}^n u_k'(t) \vec{x}_k(t) - \vec{f}(t). \end{aligned}$$

Damit die Differentialgleichung erfüllt ist, müssen die Koeffizientenfunktionen  $u_k(t)$ ,  $k = 1, \dots, n$ , also so gewählt werden, dass

$$\sum_{k=1}^n u_k'(t) \vec{x}_k(t) = \vec{f}(t)$$

gilt. Dies ist ein lineares Gleichungssystem für  $u_k'(t)$ ,  $k = 1, \dots, n$ , wenn man die Vektoren  $\vec{x}_k(t)$  als Spalten einer Matrix auffasst:

$$\left[ \vec{x}_1(t) \mid \cdots \mid \vec{x}_n(t) \right] \begin{bmatrix} u_1'(t) \\ \vdots \\ u_n'(t) \end{bmatrix} = \vec{f}(t). \quad (\text{A.6})$$

Damit ergibt sich folgender Algorithmus zur Bestimmung einer Partikulärlösung:

### Algorithmus 1.3.17 (Variation der Konstanten)

- (0) Gegeben: Fundamentalsystem  $\{\vec{x}_1, \dots, \vec{x}_n\}$  für (A.5).
- (1) Löse das lineare Gleichungssystem (A.6) für  $u_1'(t), \dots, u_n'(t)$ .
- (2) Integriere die Funktionen  $u_k'$  für  $k = 1, \dots, n$  (mit beliebigen Integrationskonstanten, etwa 0):

$$u_k(t) = \int u_k'(t) dt, \quad k = 1, \dots, n.$$

(3) Eine Partikulärlösung lautet

$$\vec{x}_p(t) = \sum_{k=1}^n u_k(t) \vec{x}_k(t).$$

### Beispiel 1.3.18

Löse

$$\vec{x}'(t) = A\vec{x}(t) + \vec{f}(t) \quad \text{mit} \quad A = \begin{bmatrix} -1 & 2 \\ 0 & -3 \end{bmatrix}, \quad \vec{f}(t) = \begin{bmatrix} 3t - 1 \\ -3t + 2 \end{bmatrix}.$$

Zunächst lösen wir das homogene System. Die Eigenwerte von  $A$  lauten  $\lambda_1 = -1$  und  $\lambda_2 = -3$  mit zugehörigen Eigenvektoren

$$\vec{v}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \quad \vec{v}_2 = \begin{bmatrix} 1 \\ -1 \end{bmatrix}.$$

Damit beschreiben die folgenden Lösungen ein Fundamentalsystem:

$$\vec{x}_1(t) = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \exp(-t), \quad \vec{x}_2(t) = \begin{bmatrix} 1 \\ -1 \end{bmatrix} \exp(-3t).$$

Die allgemeine Lösung des homogenen Systems lautet also

$$\vec{x}(t) = c_1 \begin{bmatrix} 1 \\ 0 \end{bmatrix} \exp(-t) + c_2 \begin{bmatrix} 1 \\ -1 \end{bmatrix} \exp(-3t)$$

mit beliebigen Konstanten  $c_1$  und  $c_2$ .

Zur Bestimmung einer Partikulärlösung lösen wir das Gleichungssystem

$$\begin{bmatrix} \vec{x}_1(t) & \vec{x}_2(t) \end{bmatrix} \begin{bmatrix} u_1'(t) \\ u_2'(t) \end{bmatrix} = \vec{f}(t)$$

bzw.

$$\begin{bmatrix} \exp(-t) & \exp(-3t) \\ 0 & -\exp(-3t) \end{bmatrix} \begin{bmatrix} u_1'(t) \\ u_2'(t) \end{bmatrix} = \begin{bmatrix} 3t - 1 \\ -3t + 2 \end{bmatrix}$$

nach  $u_1'$  und  $u_2'$  auf und erhalten

$$\begin{aligned} u_2'(t) &= -\exp(3t)(-3t + 2), \\ u_1'(t) &= \exp(t)(3t - 1 + \exp(-3t)\exp(3t)(-3t + 2)) = \exp(t). \end{aligned}$$

Integration mit Integrationskonstante 0 liefert

$$\begin{aligned}u_2(t) &= - \int \exp(3t)(-3t + 2)dt = (t - 1) \exp(3t), \\u_1(t) &= \exp(t).\end{aligned}$$

Damit ist durch

$$\begin{aligned}\vec{x}_p(t) &= u_1(t)\vec{x}_1(t) + u_2(t)\vec{x}_2(t) \\&= \exp(t) \begin{bmatrix} 1 \\ 0 \end{bmatrix} \exp(-t) + (t - 1) \exp(3t) \begin{bmatrix} 1 \\ -1 \end{bmatrix} \exp(-3t) \\&= \begin{bmatrix} 1 \\ 0 \end{bmatrix} + (t - 1) \begin{bmatrix} 1 \\ -1 \end{bmatrix} = \begin{bmatrix} t \\ 1 - t \end{bmatrix}\end{aligned}$$

eine Partikulärlösung gegeben. Nach Satz 1.3.6 lautet die allgemeine Lösung des inhomogenen Differentialgleichungssystems

$$\begin{bmatrix} t \\ 1 - t \end{bmatrix} + c_1 \begin{bmatrix} 1 \\ 0 \end{bmatrix} \exp(-t) + c_2 \begin{bmatrix} 1 \\ -1 \end{bmatrix} \exp(-3t)$$

mit beliebigen Konstanten  $c_1$  und  $c_2$ . ■

## Literaturverzeichnis

- [Alt02] Alt, H. W. *Lineare Funktionalanalysis*. Springer, Berlin-Heidelberg-New York, 4th edition, 2002.
- [BHW11] Burg, K., Haf, H. and Wille, F. *Höhere Mathematik für Ingenieure. Band I: Analysis*. Teubner, Stuttgart, 1985/2011.
- [BK11] Brokate, M. and Kersting, G. *Maß und Integral..* Basel: Birkhäuser, 2011. ISBN 978-3-7643-9972-6/pbk; 978-3-0346-0646-2/ebook.
- [BS89] Bronstein, I. and Semendjaev, K. A., editors. *Taschenbuch der Mathematik*. Verlag Harri Deutsch, Thun, 24th edition, 1989.
- [Bur98] Burenkov, V. I. *Sobolev Spaces on Domains*. Teubner-Texte zur Mathematik, Band 137. B.G. Teubner Stuttgart Leipzig, 1998.
- [LS76] Ljusternik, L. A. and Sobolew, W. I. *Elemente der Funktionalanalysis*. Verlag Harri Deutsch, Zürich-Frankfurt/Main-Thun, 1976.
- [Lue69] Luenberger, D. G. *Optimization by Vector Space Methods*. John Wiley & Sons, New York-London-Sydney-Toronto, 1969.
- [MV93] Meyberg, K. and Vachenauer, P. *Höhere Mathematik I-II*. Springer, Berlin-Heidelberg-New York, 2nd edition, 1993.
- [Pap08] Papula, L. *Mathematik für Ingenieure und Naturwissenschaftler. Band 1-3*. Vieweg+Teubner, Wiesbaden, 2008.
- [Wer95] Werner, D. *Funktionalanalysis*. Springer, Berlin-Heidelberg-New York, 1995.