# Mathematische Methoden

Johannes Berg Institut für Theoretische Physik Universität zu Köln

11. Juli 2015

## Inhaltsverzeichnis

In	Inhaltsverzeichnis					
1	Vektoren und Vektorräume					
	1.1	Verschiebungen in der Ebene	5			
	1.2	Raumpunkte und Verschiebungen	8			
	1.3	Axiome des Vektorraums	9			
	1.4	Basissysteme	12			
	1.5	Das Skalarprodukt und die Norm	14			
	1.6	Das Kreuzprodukt	16			
<b>2</b>	Lineare Algebra					
	2.1	Lineare Abbildungen	21			
	2.2	Änderung der Basis	28			
	2.3	Die Determinante	31			
	2.4	Lineare Gleichungen und lineare Algebra	36			
	2.5	Eigenvektoren und Eigenwerte	38			

3	Ana	alysis	41			
	3.1	Grenzwert	41			
	3.2	Differentiation	42			
	3.3	Integration	46			
	3.4	Mehrfachintegrale	50			
4	Vek	ctoranalysis	57			
	4.1	Integration von skalaren und vektoriellen Feldern	59			
	4.2	Linienintegrale	59			
	4.3	Das Oberflächenintegral	63			
	4.4	Das Volumenintegral	66			
	4.5	Differentialoperatoren	68			
	4.6	Die Rotation (engl. "curl")	72			
	4.7	Die Divergenz	74			
	4.8	Integraltheoreme	74			
5	Folgen und Reihen, Potenzreihen und komplexe Zahlen					
	5.1	Folgen	81			
	5.2	Reihen	82			
	5.3	Potenzreihen	83			
	5.4	Taylorreihe	85			
	5.5	Die komplexen Zahlen	88			
6	Dif	ferentialgleichungen	93			
	6.1	Differentialgleichungen in der Physik	93			
	6.2	Terminologie	95			
	6.3	Gewöhnliche DGL erster Ordnung	95			
	6.4	Gewöhnliche lineare DGL	99			
	6.5	Partielle DGL (PDG)	108			
7	Fourierreihen 115					
•	7.1	Fourierreihe der Rechteckfunktion	116			
			110			
	(.2)	Fourierkoeffizienten einer allgemeinen periodische Funktion	119			

## Überblick

"Let me end on a more cheerful note. The miracle of the appropriateness of the language of mathematics for the formulation of the laws of physics is a wonderful gift which we neither understand nor deserve. We should be grateful for it and hope that it will remain valid in future research and that it will extend, for better or for worse, to our pleasure, even though perhaps also to our bafflement, to wide branches of learning." (Eugene Wigner)

Diese Vorlesung gibt eine Einführung in mathematische Methoden, derer sich die Physik (und viele weitere Wissenschaften) bedient. Dabei nutzt die Physik mathematische Methoden durchaus auch als Handwerkszeug, also um Rechenmethoden für ein konkretes Problem zu finden. Wichtiger aber noch ist die Suche nach dem passenden mathematischen Formalismus um bestimmte physikalische Sachverhalte zu beschreiben. Ein Beispiel ist das zweite Newtonsche Gesetz  $\vec{F} = m\vec{a}$ . Implizit in diesem Formulierung ist die Annahme, dass die Kraft  $\vec{F}$  und die Beschleunigung  $\vec{a} \equiv d^2\vec{x}/dt^2$  Vektoren (im Ortsraum) sind<sup>1</sup>. Gemeinsam bilden sie eine Differentialgleichung  $\vec{F} = m\frac{d^2\vec{x}}{dt^2}$ , ihre Lösung  $\vec{x}(t)$  sagt die Bahnkurve voraus, die ein Teilchen unter dieser Kraft nimmt. Ohne Kenntnis von Differentialgleichungen und Vektoren lässt sich diese grundlegende Gleichung der Mechanik nicht verstehen.

Dabei treten oft dieselben mathematischen Konzepte in unterschiedlichen physikalischen Zusammenhängen auf. Hier ein kurzer Überblick verschiedener Teilgebiete der Physik und verwandter Gebiete, jeweils mit den mathematischen Methoden, die dort Verwendung finden.

- 1. Mechanik: Differentialgleichungen (DGL), Vektorrechung
- 2. Elektromagnetismus: DGL, Vektoranalysis
- 3. Quantenmechanik: DGL, lineare Algebra

<sup>&</sup>lt;sup>1</sup>Das Gleichheitszeichen  $\equiv$  bedeutet Gleichheit im Sinne einer Definition. In einer anderen Welt könnte die z.B. Kraft auch ungleich der Masse mal Beschleunigung sein, die Beschleunigung ist aber als zweite Ableitung des Ortsvektors *definiert*.

- 4. Relativitätstheorie und Kosmologie: Vektoren und Tensoren, Differentialgeometrie
- 5. Teilchenphysik und Quantenfeldtheorie: Gruppentheorie, *lineare Algebra*, Funktionentheorie, ...
- 6. Meteorologie: DGL, lineare Algebra
- 7. Geologie: Gruppentheorie, DGL, Statistik

Die kursiv gesetzten Themen bilden ein Grundgerüst mathematischer Methoden für die ersten Semester des Physikstudiums, und damit die Themen dieser Vorlesung. Dieses Skript basiert auf einer Vorlesungsmitschrift von Sebastian Breustedt und Nikolaus Rademacher, bei denen ich mich herzlich bedanke. Vielen Dank auch an Dennis Finck, Donald Köster und Steffen Karalus für engagiertes Korrekturlesen und Anmerkungen. Fehler (bestimmt noch viele) und Anregungen bitte direkt an berg@thp.uni-koeln.de. Die Lektüre dieses Skriptes ersetzt weder den Vorlesungsbesuch, noch das eingenständige Nacharbeiten des Stoffes anhand von Textbüchern, noch das selbstständige Bearbeiten der Übungen. Eins

## Vektoren und Vektorräume

Vektoren spielen eine zentrale Rolle bei der Beschreibung von physikalischen Vorgängen im dreidimensionalen Raum oder der vierdimensionalen Raumzeit. Wir beginnen mit einem einfachen Beispiel, der Verschiebung in der Ebene, bevor wir uns dem Begriff des Vektorraums auf allgemeinerem (und etwas abstrakterem) Niveau nähern.

## 1.1 Verschiebungen in der Ebene

Wir betrachten Verschiebungen in der Ebene.

Eine Verschiebung  $\vec{v}$  bildet jeden Punkt p der Ebene auf einen anderen Punkt p' ab. Abstände der Punkte untereinander und ihre Orientierung zueinander bleiben erhalten. Alle Verschiebungspfeile der Abbildung sind also einander äquivalent, und sollen als eine einzelne Verschiebung  $\vec{v}$  betrachtet werden.

Zwei Verschiebungen  $\vec{v}$  und  $\vec{w}$  lassen sich kombinieren ("verknüpfen"), indem man erst  $\vec{v}$  und dann  $\vec{w}$ ausführt. Das Ergebnis ist eine neue Verschiebung, die wir  $\vec{u}$  nennen. Wir schreiben

$$\vec{u} = \vec{w} + \vec{v} . \tag{1.1}$$

Diese Schreibweise nutzt aus, dass diese Verknüpfung zweier Verschiebungen Eigenschaften hat, die uns aus der Verknüpfung zweier Zahlen durch Addi-



tion geläufig sind. Wir bezeichnen sie daher auch als Addition zweier Verschiebungen.

Es gilt nämlich  $\vec{u} = \vec{v} + \vec{w}$ , also  $\vec{v} + \vec{w} = \vec{w} + \vec{v}$  und ebenso  $(\vec{v} + \vec{w}) + \vec{x} = \vec{v} + (\vec{w} + \vec{x})$ , die Hintereinanderausführung von Verschiebungen ist also kommutativ und assoziativ. Es existiert auch eine Verschiebung, die jeden Punkt auf sich selbst abbildet (das sogenannte Nullelement). Diese Eigenschaften hat die Hintereinanderausführung von Verschiebungen gemeinsam mit der Addition von Zahlen (das Nullelement ist dort die Zahl Null).



Diese Eigenschaften von Verschiebungen haben wir

übrigen nicht bewiesen. Diese folgen aus unserer Alltagserfahrung; streng genommen handelt es sich um experimentelle Ergebnisse, aus denen wir fordern die Verknüpfung von Verschiebungen in der Ebene möge kommutativ und assoziativ sein.

Zusätzlich zur Addition von zwei Verschiebungen lässt sich die Multiplikation einer Verschiebung  $\vec{v}$  mit einer Zahl  $\lambda$  als Streckung von  $\vec{v}$  um einen Faktor  $\lambda$  definieren. Dabei ist dann  $\vec{v} + \vec{v} = 2\vec{v}$ . Bei  $\lambda < 0$  ist die Streckung in Gegenrichtung gemeint. Es gelten  $\lambda \vec{v} + \nu \vec{v} = (\lambda + \nu)\vec{v}$  und  $\lambda(\nu \vec{v}) = (\lambda \nu)\vec{v}$  (wieder experimentelles Ergebnis, bzw. Forderung).

Durch Addition von Verschiebungen entsteht wieder eine Verschiebung (und nicht ein anderes Objekt), ebenso wie durch die Multiplikation einer Verschiebung mit einer Zahl wieder eine Verschiebung entsteht. Kurz: Verschiebungen kann man addieren (definiert durch Hintereinanderausführung) und mit Zahlen multiplizieren (definiert durch Streckung). Weitere Beispiele für physikalische und mathematische Objekte mit diesen Eigenschaften werden wir in Abschnitt 1.3 finden. Objekte mit diesen Eigenschaften werden wir dort als Vektoren bezeichnen.



So bildet die Menge der Verschiebungen mit den Operationen Hintereinanderausführung und Streckung einen sogenannten Vektorraum.

## 1.1.1 Basissysteme von Verschiebungen

Basissysteme erlauben, Verschiebungen durch Zahlen zu beschreiben, was konkrete Rechnungen oft stark vereinfacht.

Wir beginnen damit, zwei ausgezeichnete Richtungen in der Ebene ('rechts' und 'oben') und eine Längeneinheit ('1 m') zu wählen. Die Verschiebung  $\vec{e_1}$  verschiebe alle Punkte um eine Längeneinheit nach rechts,  $\vec{e_2}$  verschiebe alle Punkte um eine Längeneinheit nach oben. Alle Verschiebungen lassen sich dann durch Streckun-



gen von  $\vec{e_1}$  und  $\vec{e_2}$  und Hintereinanderausführung der resultierenden Verschiebungen generieren, z.B.  $\vec{v} = 2\vec{e_1} + \vec{e_2}$ . Ein Basissystem aus Verschiebungen der Länge 1, die senkrecht aufeinanderstehen (dazu später mehr) bezeichnet man als **kartesische Basis**. Eine Basis muss allerdings nicht aus solchen sogenannten **orthonormalen** Verschiebungen bestehen, in der Praxis hat eine solche Basis jedoch klare Vorteile (einen werden wir in Abschnitt 1.5.1 kennenlernen).

Jede Verschiebung  $\vec{v}$  in der Ebene ist damit durch 2 Zahlen charakterisiert,  $\vec{v} = v_1 \vec{e}_1 + v_2 \vec{e}_2$ . Diese Zahlen  $v_1$  und  $v_2$  heißen **Komponenten** von  $\vec{v}$  in der Basis  $\vec{e}_1, \vec{e}_2$ . Die Komponenten der Verschiebung  $\vec{v}$  lassen sich in Spaltenschreibweise zusammenfassen,

$$\mathbf{v} = \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} \;,$$

wobei  $\mathbf{v}$  aus den Komponenten der Verschiebung in der gewählten Basis besteht.  $\mathbf{v}$  wird auch als **Darstellung** der Verschiebung  $\vec{v}$  in der gewählten Basis bezeichnet. Gegeben die Basis, bezeichnet jedes  $\mathbf{v}$  eindeutig eine Verschiebung  $\vec{v}$ . Trotzdem handelt es sich um unterschiedliche Objekte,  $\vec{v}$  bezeichnet eine Verschiebung (ist also unabhängig von der Wahl einer Basis),  $\mathbf{v}$  die Komponenten von  $\vec{v}$  in einer bestimmten Basis<sup>1</sup>. Basisvektoren haben in ihrer Basis stets die Darstellung  $(1, 0, 0, \ldots), (0, 1, 0, \ldots), \ldots$  Die Aussage  $\mathbf{e}_1 = (1, 0, \ldots)$ ist korrekt, aber recht inhaltsleer, wenn nicht gesagt wird was  $\vec{e}_1$  ist (z.B. horizontale Verschiebung Richtung Osten um 1 m).

Für die Addition zweier Verschiebungen gilt nun

$$\vec{v} + \vec{w} = (v_1 \vec{e}_1 + v_2 \vec{e}_2) + (w_1 \vec{e}_1 + w_2 \vec{e}_2) = (v_1 + w_1)\vec{e}_1 + (v_2 + w_2)\vec{e}_2$$
(1.2)

(der zweite Schritt folgt aus der Assoziativität der Addition von Verschiebungen und der Forderung  $\lambda \vec{v} + \nu \vec{v} = (\lambda + \nu)\vec{v}$  an die Streckung). In Komponentenschreibweise

$$\mathbf{v} + \mathbf{w} = \begin{pmatrix} v_1 + w_1 \\ v_2 + w_2 \end{pmatrix} . \tag{1.3}$$

Dieses Ergebnis ist nicht zu unterschätzen! Wir haben nun eine Rechenregel, die die Addition (Hintereinanderausführung) von Verschiebungen durch die Addition von Zahlen beschreibt: Gegeben zwei Verschiebungen, addieren wir einfach ihre Komponenten und erhalten die Komponenten der Summe der Verschiebungen.

**Aufgabe:** Zeigen Sie analog, dass die Verschiebung  $\vec{w} = \lambda \vec{v}$  die Komponenten  $w_1 = \lambda v_1$ und  $w_2 = \lambda v_2$  hat.

<sup>&</sup>lt;sup>1</sup>In Textbüchern wird oft kein Unterschied zwischen diesen Objekten gemacht, der Sinn erschließt sich allerdings meist aus dem Kontext.

## 1.2 Raumpunkte und Verschiebungen

Es mag verwunderlich erscheinen, dass wir als Beispiele für Vektoren Verschiebungen genutzt haben, aber noch nicht über 'Ortsvektoren' gesprochen haben. Tatsächlich sind die Punkte einer Ebene oder des dreidimensionalen Raumes keine Vektoren, denn man kann Punkte im Raum nicht sinnvoll addieren. Mit der Wahl eines ausgezeichneten Punktes, des sogenannten **Ursprungs** können wir jedoch jeden Punkt im Raum durch eine Verschiebung beschreiben.

**Beispiel:** Die Universität befindet sich in Köln-Sülz am Punkt p. Anstatt der Beschreibung dieses Punktes im Raum kann man aber genauso gut angeben, welche Verschiebung  $\vec{r}$  uns vom Koordinatenursprung  $p_0$  (der in diesem Beispiel natürlich am Dom liegt) zur Universität bringt, wir schreiben  $p = p_0 + \vec{r}$ .

Nach Wahl eines Ursprungs lassen sich Punkte im Raum als Verschiebungen vom Ursprung, und damit als Vektoren beschreiben. Etwas präziser: jeder Punkt entspricht genau einer Verschiebung, und jede Verschiebung genau einem Punkt. Die Verschiebung  $\vec{r}$  wird als **Ortsvektor** bezeichnet. Sie beschreibt einen Raumpunkt allerdings nur bei gegebenem Ursprung.

Die wichtigste Anwendung des Konzept des Ortsvektors sind sogenannte **Bahnkurven**. Denken wir an eine Punktmasse, die sich durch den Raum bewegt. Zu jedem Zeitpunkt tbefindet sie sich an einem anderen Punkt p im Raum. Die Bahnkurve ist also durch eine Funktion der Zeit p(t), oder äquivalent bei Wahl eines Ursprungs durch den zeitabhängigen Ortsvektor  $\vec{r}(t)$  beschrieben. Wählen wir nun noch eine (feste, zeitlich unveränderte) Basis für die Ortsvektoren, beschreiben die Komponenten  $\mathbf{r}(t)$  die Bewegung der Punktmasse.



**Aufgabe:** Berechnen Sie die Komponenten einer dreidimensionalen spiralförmigen Bahnkurve um die z-Achse (Wendeltreppe, Korkenzieher). Vergleichen Sie Ihr Ergebnis mit

$$\mathbf{r}(t) = \begin{pmatrix} \cos(2\pi t)\\ \sin(2\pi t)\\ v_z t \end{pmatrix} . \tag{1.5}$$

### 1.3 Axiome des Vektorraums

Mengen, für deren Elemente eine Additionsoperation und eine Multiplikation mit Zahlen existiert treten in den unterschiedlichsten Zusammenhängen auf. Als konkretes Beispiel werden wir zwar weiter die Menge der Verschiebungen benutzen, es lassen sich allerdings auch für ganz andere Objekte die Operationen Addition und Multiplikation mit Zahlen definieren; sie bilden also auch Vektorräume. Daher ist eine formale Definition des Vektorraums sinnvoll; sie erlaubt eine einheitliche Beschreibung dieser Mengen.

**Info:** Wir benutzen Begriffe und Notation aus der Mengenlehre. Der Begriff der **Menge** wird nicht definiert. Ein Beispiel ist die Menge der natürlichen Zahlen, die wir als  $\mathbb{N} = \{1, 2, 3, \ldots\}$  bezeichnen. Ein weiteres Beispiel sind die Verschiebungen in der Ebene. Die **Untermenge** einer Menge M enthält Elemente dieser Menge, z.B. sind die natürlichen Zahlen eine Untermenge der reellen Zahlen  $\mathbb{R}$ , wir schreiben  $\mathbb{N} \subset \mathbb{R}$ .

Eine **Abbildung** zwischen zwei Mengen ist eine Zuordnung, die jedem Element einer Menge (der **Definitionsmenge**) jeweils ein Element einer zweiten Menge (**Zielmenge**) zuweist. Ist jedes Element der Zielmenge genau ein Abbild eines Elementes der Definitionsmenge, heißt die Abbildung eineindeutig oder **bijektiv**. Insbesondere existiert dann ein Inverses dieser Abbildung.

Als nächstes definieren wir das **Produkt**  $A \times B$  zweier Mengen A, B, als die Menge aller Paare der Elemente aus A und B. Für  $A = \{\text{Peter, Ida}\}$  und  $B = \{\text{Müller, Maier, Schmidt}\}$ ist  $A \times B$  gleich der Menge  $\{(\text{Peter, Müller}), (\text{Peter, Maier}), (\text{Peter, Schmidt}), (\text{Ida, Müller})$ (Ida, Maier), (Ida, Schmidt) $\}$ .

**Beispiel:** Ein Paar reeller Zahlen ist damit ein Element der Menge  $\mathbb{R} \times \mathbb{R}$ . Die Addition von zwei reellen Zahlen a, b ist eine Abbildung aus der Menge der Paare reeller Zahlen auf die Menge der reellen Zahlen, also  $\mathbb{R} \times \mathbb{R} \to \mathbb{R}$ ,  $(a, b) \mapsto a + b$ . Machen Sie sich klar, dass sich hier  $\mathbb{R} \times \mathbb{R}$  nicht auf die Operation der Multiplikation zweier Zahlen bezieht, sondern auf die Bildung der Paare reeller Zahlen (die Paare von Zahlen, die dann addiert werden).

**Beispiel:** Verschiebungen in der Ebene sind durch ein Zahlenpaar charakterisiert, kompakt schreiben wir  $\mathbf{v} \in \mathbb{R} \times \mathbb{R} \equiv \mathbb{R}^2$ . Als Beispiel für Abbildungen nutzen wir noch einmal die obige Raumkurve  $\mathbf{r}(t)$ . Formal betrachtet ist eine Raumkurve in zwei Dimensionen eine Abbildung aus den reellen Zahlen  $\mathbb{R}$  auf die Verschiebungen in der Ebene  $\mathbb{R}^2$ ,

$$\begin{aligned} \mathbf{r}(t) : \mathbb{R} &\to \mathbb{R}^2 \\ t &\mapsto \mathbf{r} = \mathbf{r}(t) = \begin{pmatrix} r_1(t) \\ r_2(t) \end{pmatrix} . \end{aligned} \tag{1.6}$$

Ausgerüstet mit diesem Werkzeug können wir jetzt den Begriff des Vektorraumes präzise fassen. Ein Vektorraum über den reellen Zahlen  $\mathbb{R}$  ist eine Menge V von Elementen (genannt Vektoren) und zwei Verknüpfungen. Die erste Verknüpfung ist die Addition von zwei Vektoren. Diese Verknüpfung bildet zwei Vektoren (ein Element von  $V \times V$ ) auf ein Element von V ab. Die zweite Verknüpfung ist die Multiplikation einer rellen Zahl mit einem Vektor, sie bildet also ein Element von  $\mathbb{R} \times V$  auf V ab.

Ein **Vektorraum** über den reellen Zahlen  $\mathbb{R}$  ist ein Tripel  $(V, +, \times)$  bestehend aus einer Menge V und zwei Verknüpfungen + und  $\times$ .

$$V \times V \longrightarrow V, (\vec{v}, \vec{w}) \longmapsto \vec{v} + \vec{w} \ (\vec{v}, \vec{w} \in V)$$

$$\mathbb{R} \times V \longrightarrow V, (\lambda, \vec{v}) \longmapsto \lambda \vec{v} (\lambda \in \mathbb{R}, \vec{v} \in V)$$

mit den folgenden Eigenschaften:

- 1. (V, +) bildet eine sogenannte Gruppe, d.h. 1.  $(\vec{v} + \vec{w}) + \vec{x} = \vec{v} + (\vec{w} + \vec{x}) \, \forall \vec{v}, \vec{w}, \vec{x} \in V$ (Assoziativität), 2. es existiert ein Nullelement  $\vec{e} : \vec{e} + \vec{v} = \vec{v} + \vec{e} = \vec{v} \, \forall \vec{v} \in V$ , 3. zu jedem  $\vec{v}$  existiert ein Inverses  $\vec{V}$ , so daß  $\vec{V} + \vec{v} = \vec{e} = \vec{v} + \vec{V}$
- 2.  $\vec{v} + \vec{w} = \vec{w} + \vec{v}$  (Kommutativität)
- 3.  $(\lambda + \mu)\vec{v} = \lambda\vec{v} + \mu\vec{v}$
- 4.  $\lambda(\mu \vec{v}) = (\lambda \mu) \vec{v}$
- 5.  $\lambda(\vec{v} + \vec{w}) = \lambda \vec{v} + \lambda \vec{w}$
- 6.  $1\vec{v} = \vec{v}$

Die Elemente der Menge V werden als Vektoren bezeichnet, verkürzend spricht man oft vom Vektorraum V und meint das Tripel  $(V, +, \times)$ . Schreibt man  $\vec{u} - \vec{v}$  für  $u + (-1\vec{v})$ , und  $\frac{\vec{u}}{\lambda}$  für  $(\frac{1}{\lambda})\vec{u}$ , so sind auch Subtraktion und Division definiert. Gelten diese Axiome auch für eine Untermenge  $U \subset V$ , spricht man von einem **Unterraum**  $(U, +, \times)$ . Dann generieren Addition und Multiplikation von Vektoren aus U nur Vektoren aus U. Zum Beispiel bilden Verschiebungen entlang einer Geraden bilden einen Unterraum der Verschiebungen in der Ebene. Begriffe wie "Länge", "Richtung" tauchen in diesen Axiomen nicht auf, sie lassen sich auch nicht für alle Vektorräume definieren.

**Aufgabe:** Prüfen Sie, ob die Verschiebungen in der Ebene, mit den Operationen der Hintereinanderausführung und der Streckung die Axiome des Vektorraumes erfüllen.

#### **Beispiel:**

- Verschiebungen (mit Hintereinanderausführung als Addition, Streckung als Multiplikation)
- Geschwindigkeiten; als Verschiebungen pro Zeit lassen sich Geschwindigkeiten genauso addieren und mit Zahlen multiplizieren wie Verschiebungen. Ein Massepunkt der sich mit konstanter Geschwindigkeit bewegt und pro Zeitintervall  $\tau$  um  $\vec{s}$  verschoben wird hat die Geschwindigkeit  $\vec{s}/\tau$ .
- *n*-Tupel von reellen Zahlen (1-Tupel: einzelne Zahl, 2-Tupel: ein Paar von Zahlen, ...) mit Addition definiert als Addition der einzelnen Zahlen, und analog der Multiplikation. z.B. 3-Tupel

$$\mathbf{x} = \begin{pmatrix} x \\ y \\ z \end{pmatrix}, \mathbf{x_1} + \mathbf{x_2} = \begin{pmatrix} x_1 + x_2 \\ y_1 + y_2 \\ z_1 + z_2 \end{pmatrix}.$$

Damit bilden auch die Vektorkomponenten einen Vektorraum.

• Funktionen f(x); durch Addition f(x) + g(x) oder Multiplikation  $\lambda f(x)$  für alle x entstehen neue Funktionen. Folgende Überlegung mag hilfreich sein: Approximieren wir eine Funktion f(x) durch diskrete Funktionswerte bei  $x_1, x_2, x_3, \ldots$  Diese Funktionswerte

$$\begin{pmatrix} f(x_1) \\ f(x_2) \\ f(x_3) \\ \dots \end{pmatrix}$$
(1.7)

bilden unter der Addition von Funktionen und der Multiplikation von Funktionen mit Zahlen einen Vektorraum (er ist gleich dem Raum der *n*-Tupel im obigen Beispiel). Legt man die diskreten Funktionswerte  $x_1, x_2, x_3, \ldots$  beliebig dicht, so approximiert man die Funktion f(x) beliebig gut. Daraus folgt auch, dass der von Funktionen gebildete Vektorraum unendlichdimensional ist.

• Und hier noch ein Beispiel für etwas, das kein Vektorraum ist: einzelne Kugeln Eis verschiedener Sorten. Die Menge { Vanille, Erdbeere, Schokolade } bildet *keinen* Vektorraum, weil keine (sinnvolle) Additionsregel für Eiskugeln existiert, bei der die Addition von zwei Kugeln wieder eine Eiskugel ergibt.

**Info:** Für den physikalischen Raum (ohne Wahl eines Ursprungs) ist ein sogenannter affiner Raum die adäquate Beschreibung. Ein **affiner Raum** (M, V, +) ist eine Menge M von Punkten, ein Vektorraum V und eine Verknüpfung +

$$\begin{split} M \times V \to M \\ (p, \vec{v}) \longmapsto p + \vec{v} \in M \end{split}$$

mit den Eigenschaften  $p + (\vec{u} + \vec{v}) = (p + \vec{u}) + \vec{v}$ 

und zu jedem Paar  $p, q \in M$  existiert genau ein Vektor  $\vec{v} \in V, p = q + \vec{v}$ .

Das heißt, Punkte  $p, q \in M$  des affinen Raumes können subtrahiert werden,  $p - q = \vec{v}$  (das Ergebnis ist die Verschiebung von q nach p). Punkte des affinen Raumes können aber nicht addiert werden, und bilden daher auch keinen Vektorraum.

#### 1.4 Basissysteme

Vektoren  $\vec{e}_1, \vec{e}_2, \vec{e}_3, \dots \vec{e}_n \in V$  nennt man ein **Erzeugendensystem**, wenn man jeden Vektor  $\vec{v} \in V$  schreiben kann als

$$\vec{v} = v_1 \vec{e}_1 + v_2 \vec{e}_2 + v_3 \vec{e}_3 + \dots + v_n \vec{e}_n v_1, \dots + v_n \in \mathbb{R}$$
(1.8)

Durch Multiplikation dieser Vektoren mit Zahlen  $v_1, v_2, \ldots, v_n$  und Addition der Vektoren lässt sich also jedes Element des Vektorraumes erzeugen. Man bezeichnet (1.8) als **Linearkombination** der Vektoren  $\vec{e_1}, \vec{e_2}, \vec{e_3}, \ldots, \vec{e_n} \in V$ .

Die Beispiele i)-iii) sind alle Erzeugendensysteme der Verschiebungen in der Ebene, iv) allerdings nicht.



Vektoren  $\vec{e}_1, \vec{e}_2 \dots \vec{e}_n$  heißen **linear unabhängig**, wenn gilt

sei 
$$v_1 \vec{e}_1 + v_2 \vec{e}_2 + v_3 \vec{e}_3 + \dots + v_n \vec{e}_n = 0$$
, dann gilt  $v_1 = v_2 = \dots = v_n = 0$ . (1.9)

Anders ausgedrückt kann aus einer Menge linear unabhängiger Vektoren kein Element als Linearkombination der anderen Vektoren geschrieben werden. Beispiele i) und ii) zeigen linear unabhängige Vektoren, nicht aber iii) oder iv). Vektoren  $\vec{e_1}, \vec{e_2}, \ldots, \vec{e_k}$  bilden eine sogenannte **Basis** von V, wenn sie sowohl ein Erzeugendensystem sind als auch linear unabhängig sind. Denn dann kann jeder Vektor  $\vec{v}$  auf eine einzige Art als eine Summe  $\vec{v} = v_1 \vec{e_1} + v_2 \vec{e_2} + \ldots v_k \vec{e_k}$  ausgedrückt werden: Erstens kann jeder Vektor als Linearkombination dieser Basisvektoren ausgedrückt werden, denn die Basisvektoren bilden ein Erzeugendensystem. Und zweitens, wäre die Darstellung nicht einzigartig dann hätten wir  $\vec{v} = \sum_{i=1}^{k} v_i \vec{e_i} = \sum_{i=1}^{k} v'_i \vec{e_i}$  mit  $v_i \neq v'_i$  für mindestens ein i. Damit hätten wir  $\sum_{i=1}^{k} (v_i - v'_i) \vec{e_i} = \vec{01}$  mit  $v_i - v'_i \neq 0$  für mindestens ein i, dann wären die Vektoren  $\vec{e_1}, \vec{e_2}, \ldots, \vec{e_k}$  aber (im Widerspruch zur Annahme) nicht linear unabhängig. Gegeben eine Basis  $\vec{e_1}, \vec{e_2}, \ldots, \vec{e_k}$  ist also jeder Vektor  $\vec{v}$  eineindeutig (bijektiv) durch das

Tupel  $\mathbf{v} = \begin{pmatrix} v_1 \\ v_2 \\ \cdots \\ v_k \end{pmatrix}$  bestimmt. Eineindeutig bedeutet, dass jedem  $\vec{v}$  genau ein  $\mathbf{v}$  zugeordnet

ist, und jedem **v** genau ein  $\vec{v}$ .

Die Zahl der Basisvektoren k heißt die Dimension des Vektorraumes. Interessanterweise ist sie unabhängig von der Wahl der Basis (was wir aber nicht beweisen werden).

Gegeben eine Basis  $\vec{e_1}, \vec{e_2}, \ldots$  ergeben sich aus den Axiomen des Vektorraumes Rechenregeln für die Komponenten eines Vektors. Damit läßt sich (zumindest in endlichdimensionalen Vektorräumen) mit Vektoren wie mit Zahlen rechnen:

$$\mathbf{x} = \mathbf{y} \text{ wenn (und nur wenn) } x_1 = y_1 \text{ und } x_2 = y_2$$
$$\mathbf{0} = \begin{pmatrix} 0\\0 \end{pmatrix}$$
$$\mathbf{x} + \mathbf{y} = \begin{pmatrix} x_1 + y_1\\x_2 + y_2 \end{pmatrix}$$
$$\lambda \mathbf{x} = \begin{pmatrix} \lambda x_1\\\lambda x_2 \end{pmatrix}$$

Der erste Punkt folgt daraus, dass die Darstellung eines Vektors  $\vec{x}$  in einer Basis als  $\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$  eindeutig ist.

Aufgabe: Leiten Sie den Rest der Rechenregeln aus den Axiomen des Vektorraumes her.

latex main.tex ;dvips main.dvi -o main.ps;ps2pdf main.ps

**Info:** Beim Rechnen mit Vektoren und ihren Komponenten treten häufig Summen auf, zum Beispiel  $v_1\vec{e_1} + v_2\vec{e_2} + v_3\vec{e_3}$ . Um unnötige Schreibarbeit zu vermeiden, nutzen wir das Summenzeichen  $\sum$ . Zum Beispiel ist 1 + 2 + 3 die Summe der natürlichen Zahlen *i* von 1 bis 3, was wir kompakt als  $\sum_{i=1}^{3} i$  schreiben. Statt *i* hätten wir natürlich auch eine andere Variable nutzen können. Analog ist  $v_1 \vec{e_1} + v_2 \vec{e_2} + v_3 \vec{e_3} = \sum_{i=1}^{3} v_i \vec{e_i}$ . Die Grenzen der Summe schreiben wir oft nicht mit, wenn sie aus dem Kontext folgen (z.B. in 3 Dimensionen wenn *i* von 1 bis 3 läuft). Doppelte Summen wie  $\sum_{i=1}^{3} \sum_{j=1}^{3} ij$  schreiben wir noch kompakter als  $\sum_{i,j=1}^{3} ij$ .

#### 1.5 Das Skalarprodukt und die Norm

**Beispiel:** Betrachten wir die Verschiebung  $\vec{v}$  eines Massepunktes gegen eine Kraft<sup>*a*</sup>  $\vec{F}$ . Die von der Kraft  $\vec{F} = F_2 \vec{e}_2$  am Massepunkt verrichtete Arbeit ist  $v_2 F_2$ , zeigte die Kraft in Richtung  $\vec{e}_1, \vec{F} = F \vec{e}_1$ , dann wäre die Arbeit  $v_1 F_1$ . Bei fester Kraft  $\vec{F}$  hängt die Arbeit von der Richtung der Verschiebung ab. Für  $\vec{F}, \vec{v}$  in belieber Richtung in drei Dimensionen ist die Arbeit in einer kartesischen Basis  $v_1 F_1 + v_2 F_2 + v_3 F_3$ . (1.10)  $\vec{F}$ 

Formal handelt es sich dabei um eine Abbildung von zwei Vektoren (Verschiebung, Kraft) auf eine reelle Zahl (Arbeit, eine sogenannte **skalare** Größe). Wir führen eine neue Verknüpfung von zwei Vektoren auf die reellen Zahlen ein, das sogenannte Skalarprodukt  $V \times V \to \mathbb{R}, (\vec{v}, vecw) \mapsto \vec{v} \cdot \vec{w}$ . In kartesischer Basis definieren wir das **Skalarprodukt** von zwei Vektoren im  $\mathbb{R}^n$ , **v** und **w** 

$$\vec{v} \cdot \vec{w} \equiv \mathbf{v} \cdot \mathbf{w} = v_1 w_1 + v_2 w_2 + v_3 w_3 + \dots + v_n w_n \tag{1.11}$$

Damit ist Arbeit definiert als das Skalarprodukt von Kraft und Verschiebung. Räume, in denen das obige Skalarprodukt definiert ist, heißen **euklidisch**. Aus dem Skalarprodukt eines Vektors mit sich selbst lässt sich die sogenannte **Norm**  $||\vec{v}||$  definieren, die Länge des Vektor,

$$||\vec{v}|| = \sqrt{\vec{v} \cdot \vec{v}} = \sqrt{v_1^2 + v_2^2 + v_3^2 + \dots} .$$
(1.12)

#### Geometrische Interpretation

Das Skalarprodukt hat eine geometrische Interpretation: Wählen wir die (orthonormalen) Basisvektoren  $\vec{e_1}, \vec{e_2}$  in der Ebene, die von  $\vec{v}, \vec{w}$  aufgespannt wird, und  $\vec{e_1}$  parallel zu  $\vec{w}$ , dann gilt  $\vec{v} \cdot \vec{w} = v_1 w_1 +$  $v_2 w_2 + v_3 w_3 + \ldots = v_1 w_1 = (||\vec{v}|| \cos(\phi))(||\vec{w}||) =$  $||\vec{v}|| ||\vec{w}|| \cos(\phi)$  Damit gibt das Skalarprodukt an, wie stark sich zwei Vektoren "ähnlich" sind, wie



stark sie überlappen:  $\vec{v} \cdot \vec{w}/(||\vec{v}|| ||\vec{w}||)$  ist 1, wenn  $\vec{v} = \vec{w}$ , es ist 0 wenn die beid en Vektoren senkrecht aufeinander stehen, und -1 wenn  $\vec{v} = -\vec{w}$ . In euklidischen Räumen gilt damit die Cauchy-Schwarzsche Ungleichung:

$$|ec{v}\cdotec{w}|\leq ||ec{v}||\,\,\|ec{w}\|$$
 .

### 1.5.1 Skalarprodukträume

Skalarprodukte treten in einem breiteren Kontext auf als Kraft und Arbeit. Mit Hilfe der Norm wird definiert was mit der Länge eines Vektors gemeint ist, mit dem Skalarprodukt wird definiert, wann zwei Vektoren orthogonal sind.

Ein Vektorraum V über  $\mathbb{R}$  heißt **Skalarproduktraum**, wenn für alle Vektoren  $\vec{v}, \vec{w}, \vec{w_1}, \vec{w_2} \in V$  ein Skalarprodukt  $\langle \vec{v}, \vec{w} \rangle$  definiert ist mit

- $\langle \vec{v}, \alpha \vec{w}_1 + \beta \vec{w}_2 \rangle = \alpha \langle \vec{v}, \vec{w}_1 \rangle + \beta \langle \vec{v}, \vec{w}_2 \rangle$
- $\langle \vec{v}, \vec{w} \rangle = \langle \vec{w}, \vec{v} \rangle$
- $\langle \vec{v}, \vec{v} \rangle \ge 0$  und wenn  $\langle \vec{v}, \vec{v} \rangle = 0 \curvearrowright \vec{v} = \vec{0}$

**Beispiel:** Wir definieren das Skalarprodukt zwischen zwei Funktionen f(x) und g(x) als  $\langle f, g \rangle = \int dx f(x) g(x)$ . Überzeugen Sie sich, dass diese Definition die obigen Bedingungen erfüllt.

Diese allgemeinen Eigenschaften helfen z.B. auch das Skalarprodukt von zwei Vektoren in einer nicht-kartesischen Basis  $\{\vec{e}_i\}$  auszurechnen, wenn die Skalarprodukte der Basisvektoren untereinander bekannt sind. Das Skalarprodukt von zwei Vektoren  $\vec{u} = \sum_i u_i \vec{e}_i$  und

 $\vec{v} = \sum_{j} v_j \vec{e}_j$  ist

$$\langle \sum_{i} u_{i} \vec{e_{i}}, \sum_{j} v_{j} \vec{e_{j}} \rangle \stackrel{i)}{=} \sum_{j} v_{j} \langle \sum_{i} u_{i} \vec{e_{i}}, \vec{e_{j}} \rangle$$

$$\stackrel{ii)}{=} \sum_{j} v_{j} \langle \vec{e_{j}}, \sum_{i} u_{i} \vec{e_{i}} \rangle$$

$$= \sum_{ij} u_{i} v_{j} \langle \vec{e_{j}}, \vec{e_{i}} \rangle \stackrel{ii)}{=} \sum_{ij} u_{i} v_{j} \langle \vec{e_{i}}, \vec{e_{j}} \rangle .$$

$$(1.13)$$

Sind also die Skalarprodukte zwischen allen Paaren von Basisvektoren  $\vec{e_i}, \vec{e_j}$  bekannt, kann mit diesem Ausdruck das Skalarprodukt zwischen  $\vec{u}$  und  $\vec{v}$  aus ihren Komponenten in der Basis  $\{\vec{e_i}\}$  berechnet werden. Dabei ist das Skalarprodukt zwischen diesen Vektoren unabhängig von der Wahl der Basis, denn Norm und Winkel ändern sich ja nicht unter durch eine Änderung der Basis. Allerdings ändert sich durch die Wahl der Basis ggf. der Zusammenhang (1.13) zwischen dem Skalarprodukt und den Komponenten in einer gegeben Basis. Wenn die Basis  $\{\vec{e_i}\}$  orthonormal ist, erhalten wir aus (1.13) mit  $\langle \vec{e_i}, \vec{e_j} \rangle =$  $\begin{cases} 1, & \text{wenn } i = j. \\ & \equiv \delta_{ij} & \text{wieder das euklidische Skalarprodukt (1.11), } \sum_{ij} u_i v_i \langle \vec{e_i}, \vec{e_j} \rangle = \end{cases}$ 

$$\begin{cases} 0, & \text{wenn } i \neq j \\ 0, & \text{wenn } i \neq j \end{cases} \equiv \delta_{ij} \text{ wieder das euklidische Skalarprodukt (1.11), } \sum_{ij} u_i v_j \langle \vec{e}_i, \vec{e}_j \rangle = \sum_i u_i v_i. \end{cases}$$

In einer orthonormalen Basis können mit Hilfe des Skalarproduktes auch leicht die Komponenten eines Vektors berechnet werden

$$\langle \vec{e}_i, \vec{v} \rangle = \langle \vec{e}_i, \sum_j v_j \vec{e}_j \rangle = \sum_j v_j \langle \vec{e}_i, \vec{e}_j \rangle = \sum_j v_j \delta_{ij} = v_i .$$
(1.14)

## 1.6 Das Kreuzprodukt

Ein weiteres Produkt von zwei Vektoren kann im dreidimensionalen Raum  $\mathbb{R}^3$  definiert werden, als Ergebnis dieser neuen Multiplikation von zwei Vektoren erhalten wir einen weiteren Vektor. Am leichtesten läßt sich das sogenannte Kreuzprodukt in einer kartesischen Basis definieren. Wir betrachten eine sogenanntes **rechtshändiges** kartesisches Basissystem im dreidimensionalen Raum, bei der die Basisvektoren  $\vec{e_1}, \vec{e_2}, \vec{e_3}$  orthonormal sind und wie Daumen, Zeigefinger und Mittelfinger Ihrer rechten Hand angeordnet sind. Das **Kreuzprodukt** zwischen  $\vec{v}, \vec{w} \in \mathbb{R}^3$  ist in einer solchen Basis definiert als

$$\mathbf{v} \times \mathbf{w} = \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix} \times \begin{pmatrix} w_1 \\ w_2 \\ w_3 \end{pmatrix} = \begin{pmatrix} v_2 w_3 - v_3 w_2 \\ v_3 w_1 - v_1 w_3 \\ v_1 w_2 - v_2 w_1 \end{pmatrix}$$
(1.15)

#### Geometrische Interpretation

1.  $\vec{v} \times \vec{w}$  ist ein Vektor, der senkrecht auf  $\vec{v}$  und  $\vec{w}$  steht.

$$\mathbf{v} \cdot (\mathbf{v} \times \mathbf{w}) = v_1 v_2 w_3 - v_1 v_3 w_2 + v_2 v_3 w_1 - v_2 v_1 w_3 + v_3 v_1 w_2 - v_3 v_2 w_1 = 0 , \qquad (1.16) \qquad \vec{w} \times \vec{v}$$

¥

 $\vec{v} \\ \phi$ 

 $\vec{w}$ 

und analog für  $\mathbf{w} \cdot (\mathbf{v} \times \mathbf{w})$ .

2.  $\|\vec{v}\times\vec{w}\|=\|v\|\|w\|\sin\phi,$  wobe<br/>i $\phi$ der Winkel zwischen  $\vec{v}$  und<br/>  $\vec{w}$ ist.

$$\begin{aligned} \|\vec{v} \times \vec{w}\|^{2} + (\vec{v} \cdot \vec{w})^{2} &= v_{2}^{2} w_{3}^{2} - 2v_{2} w_{3} v_{3} w_{2} + v_{3}^{2} w_{2}^{2} \qquad (1.17) \\ &+ v_{3}^{2} w_{1}^{2} - 2v_{3} w_{1} v_{1} w_{3} + v_{1}^{2} w_{3}^{2} \\ &+ v_{1}^{2} w_{2}^{2} - 2v_{1} w_{2} v_{2} w_{1} + v_{2}^{2} w_{1}^{2} \\ &+ (v_{1} w_{1} + v_{2} w_{2} + v_{3} w_{3})^{2} \\ &= (v_{1}^{2} + v_{2}^{2} + v_{3}^{2})(w_{1}^{2} + w_{2}^{2} + w_{3}^{2}) \\ &= \|\vec{v}\|^{2} \|\vec{w}\|^{2} \\ &\sim \|\vec{v} \times \vec{w}\|^{2} = \|\vec{v}\|^{2} \|\vec{w}\|^{2} - (\vec{v} \cdot \vec{w})^{2} = (1 - \cos^{2} \phi) \|\vec{v}\|^{2} \|\vec{w}\|^{2} \\ &= \sin^{2} \phi \|\vec{v}\|^{2} \|\vec{w}\|^{2} \end{aligned}$$

Damit ist das Kreuzprodukt von zwei Vektoren  $\vec{v}$  und  $\vec{w}$  ein Vektor, der senkrecht auf der von  $\vec{v}$  und  $\vec{w}$  aufgespannten Ebene steht. Seine Länge ist der Flächeninhalt des Parallelograms mit den Seiten  $\vec{v}$  und  $\vec{w}$ .

3. Diese geometrische Interpretation lässt allerdings noch die Richtung des Kreuzproduktes frei. Nach der Definition (1.15) folgt diese der "Rechte-Hand-Regel": Daumen =  $\vec{v}$ ; Zeige-finger =  $\vec{w}$ ; Mittelfinger =  $\vec{v} \times \vec{w}$ . Daraus folgt auch die wichtigste algebraische Eigenschaft des Kreuzproduktes, es ist **antikommutativ**,  $\vec{v} \times \vec{w} = -\vec{w} \times \vec{v}$ .

**Info:** Das Kreuzprodukt zweier Verschiebungen  $\vec{v} \times \vec{w}$  ist also selbst keine Verschiebung. Es hängt ausßerdem in zwei Punkten von der Wahl der Basis ab: in seiner Orientierung senkrecht zur von  $\vec{v}, \vec{w}$ , aufgespannten Oberfläche und in seiner Norm durch die Wahl der Längeneinheit. Es ist nur in drei Dimensionen definiert. Trotz dieser klaren Mängel wird ihnen das Kreuzprodukt an vielen Stellen begegenen. Wen diese Beschränkungen irritieren: eine allgemeine Definition findet sich im Rahmen des Cartan-Kalküls (Vorlesungen "Differentialgeometrie" oder "Geometrie in der Physik").

#### Das Kreuzprodukt und Flächen und Volumina

Das Kreuzprodukt  $\vec{v} \times \vec{w}$  ist nützlich um die Fläche des von  $\vec{v}$  und von  $\vec{w}$  aufgespannten Parallelogramms zu berechnen (durch die Norm von  $\vec{v} \times \vec{w}$ ). Eine weitere wichtige Anwendung aus der Geometrie betrifft den von drei Vektoren  $\vec{u}, \vec{v}, \vec{w}$ aufgespannten Körper, der sogenannte **Spat** oder das **Parallelepiped**. Sein Volumen ist die Grundfläche (das von  $\vec{v}, \vec{w}$ aufgespannte Parallelogramm), mal die Höhe des Parallelepipeds (senkrecht zur Grundfläche, Linie in blau). Da  $\vec{v} \times \vec{w}$ senkrecht zur Grundfläche steht ist das orientierte Volumen damit  $\vec{u} \cdot (\vec{v} \times \vec{w})$ , da das Skalarprodukt von  $\vec{u}$  und  $\vec{v} \times \vec{w}$ die Grundfläche mal die Komponente von  $\vec{u}$  senkrecht zur Grundfläche, angibt <sup>2</sup> Welches Paar der Vektoren  $\vec{u}, \vec{v}, \vec{w}$  man



zur Definition der Grundfläche nimmt kann dabei keine Rolle spielen, wir erwarten also  $\vec{u} \cdot (\vec{v} \times \vec{w}) = \vec{v} \cdot (\vec{w} \times \vec{u}) = \vec{w} \cdot (\vec{u} \times \vec{v})^3$ .

#### Das Kreuzprodukt und Rotationen

Die wichtigste Anwendung des Kreuzproduktes in der Mechanik liegt in Verbindung mit Rotationen. Betrachten wir eine sich drehende Scheibe und konstruieren einen Vektor  $\vec{\omega}$ , der senkrecht zur Scheibe steht und als Länge die Winkelgeschwindigkeit hat  $(2\pi/\text{Drehperiode})$ . Jeder Punkt auf der Scheibe bewegt sich nun mit einer anderen Geschwindigkeit  $\vec{v}$ . Legen wir den Ursprung in den Scheibenmittelpunkt, bewegt sich jeder Punkt auf der Scheibe senkrecht zu seinem Ortsvektor *und* senkrecht zu dem Winkelgeschwindigkeitsvektor  $\vec{\omega}$ , die ihrerseits senkrecht aufeinanderstehen. Betrag der Geschwindingkeit is  $||\vec{r}|||\vec{\omega}||$  und somit

$$\vec{v} = \vec{\omega} \times \vec{r} . \tag{1.18}$$

<sup>&</sup>lt;sup>2</sup>Das orientierte Volumen gibt über das Vorzeichen außer dem Volumen noch die Orientierung der Elemente des Spats an: das orientierte Volumen hat einen positiven Wert, wenn die 3 Elemente des Spats ein rechtshändiges System bilden, einen negativen Wert für ein linkshändiges System. Das Volumen des Spats ist dann einfach der Betrag des orientierten Volumens.

<sup>&</sup>lt;sup>3</sup>Bei diesen **zyklischen Vertauschungen** bleibt die Reihenfolge der drei Vektoren  $\vec{u}, \vec{v}, \vec{w}$  erhalten. Damit bleibt die Orientierung des Spats erhalten, z.B. bilden  $\vec{v}, \vec{w}, \vec{u}$  ein rechtshändiges System wenn  $\vec{u}, \vec{v}, \vec{w}$  ein rechtshändiges System bilden.

## Aufgabe:

Hier haben wir uns auf eine Scheibe beschränkt, so dass  $\vec{r}$  und  $\vec{\omega}$  senkrecht aufeinander stehen. Betrachten Sie einen rotierenden Zylinder, um zu zeigen, dass dieses Ergebnis allgemein gilt. (Hinweis: Nutzen Sie dass der Abstand eines Punktes von der Rotationsachse  $\sin(\phi)||\vec{r}||$  ist.) Zeigen Sie, dass dieses Ergebnis für beliebiger Körper bei Rotation mit fester Drehachse gilt.



Zwei

Lineare Algebra

Lineare Algebra beschäftigt sich mit linearen Abbildungen zwischen Vektorräumen. In der Mathematik entstand sie motiviert durch die Suche nach den Lösungen linearer Gleichungen. In der Physik ist die lineare Algebra das zentrale Werkzeug der Quantenmechanik. Wir führen die Darstellung einer linearen Abbildung durch eine Matrix ein und diskutieren Konzepte wie die Determinante einer Matrix und Eigenvektoren.

## 2.1 Lineare Abbildungen

Wir beginnen wieder mit einem einfachen Beispiel und betrachten unterschiedliche Möglichkeiten, aus Verschiebungsvektoren neue Verschiebungsvektoren zu generieren. Also Abbildungen  $L: V \to V', \vec{v} \mapsto \vec{v}'$ , die Vektoren  $\vec{v} \in V$  auf Vektoren  $\vec{v}' \in V'$  abbilden. Ein Beispiel ist (i) die Streckung aller Vektoren um einen Faktor. Denkbar sind auch (ii) Transformationen, die lediglich Komponenten eines Vektors entlang einer einzelnen Achse strecken, oder (iii) alle Vektoren um einen Winkel  $\theta$  drehen.



Diese drei Abbildung haben alle die wichtige Eigenschaft, Addition und Multiplikation im

Vektorraum V zu erhalten, d.h. zum Beispiel

$$\vec{v}_3 = \vec{v}_1 + \vec{v}_2 \Rightarrow \vec{v}'_3 = \vec{v}'_1 + \vec{v}'_2 ,$$
 (2.1)

wobei  $\vec{v}'_i = L(\vec{v}_i), i = 1, 2, 3$ . Abbildungen L mit dieser Eigenschaft werden wir als lineare Abbildungen bezeichnen, denn es folgt<sup>1</sup>

$$L(\vec{v}_1 + \vec{v}_2) = L(\vec{v}_1) + L(\vec{v}_2) .$$
(2.2)

Der Wunsch Addition und Multiplikation unter einer Transformation zu erhalten ist vom mathematischen Standpunkt aus verständlich. In der Physik treten lineare Abbildungen als lineare Näherung einer komplizierten nichtlinearen Funktion auf. Betrachten wir einen anisotropen<sup>2</sup> Kristall in einem homogenen elektrischen Feld. Wegen der ausgezeichneten Kristallrichtungen (z.B. Schichten in Graphit) fliesst Strom dann nicht genau in Richtung des Feldes (wie bei einem isotropen Medium). Ist das angelegte Feld jedoch klein, ist die Abbildung aus dem Raum der elektrischen Felder auf den Raum der Stromdichten näherungsweise linear.

#### Definition linearer Abbildungen

Eine Abbildung  $L: V \to W$  ist eine Vorschrift, die alle Elemente einer Menge V auf Elemente einer zweiten Menge W abbildet. Eine Abbildung  $L: V \to W, \vec{v} \mapsto \vec{w} = L(\vec{v})$ zwischen zwei Vektorräumen V und W ist **linear**, wenn gilt:

$$L(\alpha \vec{u} + \beta \vec{v}) = \alpha L(\vec{u}) + \beta L(\vec{v}) \quad \forall \vec{u}, \vec{v} \in V, \ \alpha, \beta \in \mathbb{R} .$$

$$(2.3)$$

Die Klammern werden oft weggelassen, und man schreibt  $L\vec{v}$ .

Beispiel: Weitere Beispiele für lineare Abbildungen sind

- Die Multiplikation von Zahlen mit einer Konstanten
  - Die Ableitung von Funktionen  $\frac{d}{dx}(f(x) + g(x)) = \frac{d}{dx}f(x) + \frac{d}{dx}g(x)$
  - Das Skalarprodukt von Vektoren  $\vec{v} \in V$  mit einem festen Vektor  $\vec{a}$  inV ist eine lineare Abbildung  $V \to \mathbb{R}, \vec{x} \mapsto \vec{a} \cdot \vec{x}$  mit  $\vec{a} \cdot (\alpha \vec{x}_1 + \beta \vec{x}_2) = \alpha \vec{a} \cdot \vec{x}_1 + \beta \vec{a} \cdot \vec{x}_2$ .
  - Der Zusammenhang zwischen einem Vektor  $\vec{v}$  in einem *n*-dimensionalen Vektorraum V und seinen *n* Komponenten **v** ist eine lineare Abbildung  $V \to \mathbb{R}^n$

Der **Kern** einer linearen Abbildung *L* sind alle Elemente von *V*, so dass  $L\vec{v} = \vec{0}$ , d.h.  $\ker(L) = \{\vec{v} \in \vec{V} | L\vec{v} = \vec{0}\}.$ 

 $\frac{1}{L(\vec{v}_1 + \vec{v}_2) = L(\vec{v}_3) = \vec{v}_3' = \vec{v}_1' + \vec{v}_2' = L(\vec{v}_1) + L(\vec{v}_2)}$ 

 $<sup>^2 \</sup>mathrm{Dies}$  bedeutet, dass Eigenschaften von der Richtung abhängen.

Das **Bild** einer linearen Abbildung  $L: V \longrightarrow W$  ist die Menge aller Elemente in W, die durch  $L\vec{v} = \vec{w}$ für ein  $\vec{v} \in V$  erreicht werden, d.h.  $\text{Im}(L) = \{\vec{w} \in W \mid \exists \vec{v} \in V : L\vec{v} = \vec{w}\}.$ 



Bild und Kern einer linearen Abbildung bilden wiederum Vektorräume, denn lineare Kombinationen von Elementen des Bildes/Kerns sind wieder Elemente des Bildes/Kerns. Die Dimension des Bildes heißt der **Rang** (engl. rank) der Abbildung.

## 2.1.1 Basisdarstellung einer linearen Abbildung

Wir betrachten eine lineare Abbildung  $L: V \longrightarrow W, \vec{v} \longmapsto \vec{w}$ . Die Dimension von V sei n und die Dimension von W sei m. Schreiben wir Vektoren in Komponenten einer Basis von V und einer Basis von W

$$\vec{v} = v_1 \vec{e}_1^v + \ldots + v_n \vec{e}_n^v$$

$$\vec{w} = w_1 \vec{e}_1^w + \ldots + w_m \vec{e}_m^w$$
(2.4)

ist die Kernfrage: wie hängen die Komponenten dieser Vektoren zusammen, wenn  $\vec{w} = L\vec{v}$ ?



'gelegt' vorstellen, mit komponentenweiser Multiplikation und Addition aller Terme. Die Spalten der Matrix **L** sind die Abbildungen der Basisvektoren. Betrachten wir die Basisvektoren als die Ecken eines sogenannten **Einheitswürfels**; graphisch lassen sich dann die Komponenten der Matrix **L** als Bild des Einheitswürfels unter der linearen Abbildung *L* verstehen,  $\mathbf{L}\left\{\begin{pmatrix}1\\0\end{pmatrix}, \begin{pmatrix}0\\1\end{pmatrix}\right\} = \left\{\begin{pmatrix}a\\b\end{pmatrix}, \begin{pmatrix}c\\d\end{pmatrix}\right\}$ .

Auch im allgemeinen Fall fragen wir zunächst, wie die Abbildung L auf die Basisvektoren  $\vec{e}_j^v, j = 1, \ldots, n$  wirkt. Da  $L\vec{e}_j^v \in W$  definieren wir  $L\vec{e}_j^v \equiv \sum_{i=1}^m L_{ij}\vec{e}_i^w$ . Aufgrund der Linearität von L steht damit schon die Wirkung von L auf einen beliebigen Vektor  $\vec{v} \in V$  fest:

$$\sum_{i}^{m} w_{i} \vec{e_{i}}^{w} = \vec{w} = L\vec{v} = L(\sum_{j}^{n} v_{j} \vec{e_{j}}^{v})$$

$$\stackrel{\text{Lin.}}{=} \sum_{j}^{n} v_{j} L(\vec{e_{j}}^{v}) = \sum_{j}^{n} v_{j} \left(\sum_{i}^{m} L_{ij} \vec{e_{i}}^{w}\right)$$

$$= \sum_{i}^{m} \left(\sum_{j}^{n} L_{ij} v_{j}\right) \vec{e_{i}}^{w} \qquad \curvearrowright \left[w_{i} = \sum_{j}^{n} L_{ij} v_{j}\right]. \qquad (2.7)$$

In Matrixform schreiben wir

$$\begin{pmatrix} w_1 \\ w_2 \\ \dots \\ w_m \end{pmatrix} = \begin{pmatrix} L_{11} \ L_{12} \ \dots \\ L_{21} \ L_{22} \ \dots \\ L_{31} \ \dots \ \dots \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \\ \dots \\ v_n \end{pmatrix} \equiv \begin{pmatrix} L_{11}v_1 + L_{12}v_2 + L_{13}v_3 + \dots \\ L_{21}v_1 + L_{22}v_2 + L_{23}v_3 + \dots \\ \dots \\ L_{m1}v_1 + L_{m2}v_2 + L_{m3}v_3 + \dots \end{pmatrix} , \qquad (2.8)$$

oder  $\mathbf{w} = \mathbf{L}\mathbf{v}^3$ .

So wie ein Spaltenvektor die Darstellung eines Vektors in einer gegebenen Basis ist, ist eine Matrix L mit Matrixelementen  $L_{ij}$  die Darstellung einer linearen Abbildung in einer Basis. Die Darstellung von  $L(\vec{v})$  ist dann das Matrixprodukt der Matrix L mit dem Spaltenvektor v. Das Matrixprodukt einer Matrix und einem Vektor ist durch (2.7) definiert.

**Info:** Eine drastische aber effektive Notation in diesem Kontext ist die sogenannte Einstein'sche Summenkonvention. Sie wird in der Elektrodynamik und vor allem in der Relativitätstheorie genutzt und spart dort viel Schreibarbeit. Statt  $w_i = \sum_j L_{ij}v_j$  einigt man sich auf die Regel, dass über paarweise auftretende Indizes summiert wird, also  $w_i = L_{ij}v_j$ .

#### Verkettung von linearen Abbildungen und Matrixmultiplikation

Abbildungen lassen sich hintereinander ausführen, zum Beispiel eine Rotation gefolgt von einer Streckung aller Vektoren um einen bestimmten Faktor. Wie hängt die Matrix der kombinierten Abbildung mit den Matrizen der beiden einzelnen Abbildungen zusammen? Wir betrachten zwei hintereinander ausgeführte (**verkettete**) lineare Abbildungen und zeigen, dass ihre Darstellungen in einer Basis der **Multiplikation** der entsprechenden Matrizen entspricht. Die erste Abbildung L bildet Elemente von des Raumes V auf Elemente des Raumes W ab, die zweite Abbildung K führt von W auf den Raum  $U; V \xrightarrow{L} W \xrightarrow{K} U$ . Diese Verkettung definiert eine Abbildung  $M \equiv K \circ L$  mit  $V \xrightarrow{M} U$ . (Mit dem Symbol  $\circ$ 

<sup>&</sup>lt;sup>3</sup>In dieser Schreibweise werden Matrizen konventionell mit Großbuchstaben bezeichnet. Auch die Notation mit Unterstrichen ist gebräuchlich und wird von mir an der Tafel verwendet;  $\underline{w} = \underline{L}\underline{v}$ .

beschreiben wir eine Verkettung zweier Abbildungen.) In Matrixschreibweise haben wir

$$w_{i} = \sum_{j} L_{ij} v_{j} , \quad u_{l} = \sum_{i} K_{li} w_{i} \quad \mathbf{w} = \begin{pmatrix} w_{1} \\ \cdots \\ w_{x} \end{pmatrix}, \mathbf{u} = \begin{pmatrix} u_{1} \\ \cdots \\ u_{m} \end{pmatrix}$$
$$u_{l} = \sum_{i} K_{li} \left( \sum_{j} L_{ij} v_{j} \right) = \sum_{j} \left( \sum_{i} K_{li} L_{ij} \right) v_{j}$$
$$\equiv \sum_{j} M_{lj} v_{j}$$
$$M_{lj} = \sum_{i} K_{li} L_{ij}$$

oder

$$\begin{pmatrix} M_{11} & M_{12} & M_{13} & \cdots \\ M_{21} & M_{22} & M_{23} & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix} = \begin{pmatrix} K_{11} & K_{12} & \cdots \\ K_{21} & K_{22} & \cdots \\ \vdots & \vdots & \ddots \end{pmatrix} \begin{pmatrix} L_{11} & L_{12} & \cdots \\ L_{21} & L_{22} & \cdots \\ \vdots & \vdots & \ddots \end{pmatrix}$$

Auch diese Multiplikation von zwei Matrizen kann man sich wieder merken als das Ablegen der Spalten der zweiten Matrix auf den Zeilen der ersten, komponentenweiser Multiplikation und Summation aller Terme.

Wichtigstes Ergebnis ist allerdings, dass wir die Verkettung von Abbildungen abgebildet haben auf die Multiplikation von Matrizen. Um das Ergebnis der Verkettung einer beliebigen Anzahl von Matrizen zu bestimmen, müssen lediglich Matrizen miteinander multipliziert werden.

Die Multiplikation von Matrizen hat einige Eigenschaften mit der Multiplikation von Zahlen gemeinsam, z.B. ist sie distributiv und assoziativ.

**Aufgabe:** Überprüfen Sie diese Behauptung, z.B. indem Sie die Komponenten der Matrix von zwei und drei verketteten Abbildungen betrachten.

Matrixmultiplikation ist allerdings im Allgemeinen nicht kommutativ, d.h.  $\mathbf{KL} \neq \mathbf{LK}$  (außer in speziellen Fällen). Diese mathematische Eigenschaft von Matrizen spielt eine zentrale Rolle in der Quantenphysik.

**Aufgabe:** Diese Eigenschaft ergibt sich aus der Verkettung von linearen Abbildungen. Betrachten Sie eine Rotation wie im Beispiel iii) in Abschnitt 2.1, verknüpft mit einer Streckung in *x*-Richtung wie im Beispiel ii). Sie erhalten unterschiedliche Abbildungen (und damit unterschiedliche Matrizen), je nachdem ob sie die Rotation oder die Streckung zuerst ausführen.

#### Matrixmultiplikation und das Euklidische Skalarprodukt

Auch für das Euklidische Skalarprodukt (1.11) lässt sich eine einfache Notation im Rahmen der Matrixmultiplikation einführen

$$\vec{v} \cdot \vec{w} = v_1 w_1 + v_2 w_2 + \ldots = \left(v_1 v_2 \ldots\right) \begin{pmatrix} w_1 \\ w_2 \\ \ldots \end{pmatrix} \equiv \mathbf{v}^T \mathbf{w} , \qquad (2.9)$$

wobei  $\mathbf{v}^T = (v_1 v_2 \dots)$  als **Spaltenvektor** definiert ist. Allgemeiner ist die **Transponierte** einer Matrix  $L^T$  durch die Vertauschung von Zeilen und Spalten definiert,  $L_{ij}^T = L_{ji}$ .

## Die Inverse einer linearen Abbildung

Gegeben eine lineare Abbildung  $L: V \to W, \vec{v} \mapsto \vec{w} = L\vec{v}$  ist  $K \equiv L^{-1}$  die Inverse von Lwenn gilt  $KL\vec{v} = \vec{v} \forall \vec{v} \in V$ .  $KL = \mathbb{1}$  ist also die **Identitätsabbildung**, die jeden Vektor auf sich selbst abbildet. Damit  $K: W \to V, \vec{w} \mapsto \vec{v} = K\vec{v}$  die Abbildung, die von jedem  $\vec{w} = L\vec{v}$  zurück auf  $\vec{v}$  führt. Nicht jede lineare Abbildung hat eine Inverse, denn nicht jede lineare Abbildung ist bijektiv. Die Abbildung, die jeden Vektor auf den Nullvektor abbildet, ist ein Extrembeispiel für eine nicht-invertierbare Abbildung:  $L\vec{v} = \vec{0} \forall \vec{v} \in V$ ). Insbesondere müssen die Dimensionen von V und W gleich sein,  $m = n^{4,5}$ .

**Info:** Um dies zu zeigen, beweisen wir, dass die Abbildungen von Basisvektoren von V eine Basis von W sind wenn L bijektiv ist. Erstens sind die Vektoren  $L\vec{e}_1, L\vec{e}_2, \ldots, L\vec{e}_n$  linear unabhängig: Für eine bijektive lineare Abbildung L gilt  $L\vec{0} = \vec{0}$ , denn wenn  $L\vec{0} = \vec{w} \neq \vec{0}$ gilt auch  $L\vec{0} = L(a\vec{0}) = aL\vec{0} = a\vec{w}$  was für  $a \neq 1$  zu einem Widerspruch führt. Diese Forderung

$$\vec{0} = L\vec{0} = a_1L\vec{e}_1 + a_2L\vec{e}_2 + \ldots + a_nL\vec{e}_n = L(a_1\vec{e}_1 + a_2\vec{e}_2 + \ldots + a_n\vec{e}_n)$$
(2.10)

führt wegen der Invertierbarkeit von L auf  $\vec{0} = a_1\vec{e}_1 + a_2\vec{e}_2 + \ldots + a_n\vec{e}_n$ . Da  $\vec{e}_1, \vec{e}_2, \ldots$  linear unabhängig sind gilt  $a_1 = a_2 = \ldots = 0$ , somit sind auch die Abbildungen  $L\vec{e}_1, L\vec{e}_2, \ldots, L\vec{e}_n$  linear unabhängig.

Zweitens sind  $L\vec{e}_1, L\vec{e}_2, \ldots, L\vec{e}_n$  ein Erzeugendensystem von W, denn da L bijektiv ist, gibt es für jedes  $\vec{w}$  ein  $\vec{v} : L\vec{v} = \vec{w}$ . Schreiben wir nun  $\vec{v} = a_1\vec{e}_1 + a_2\vec{e}_2 + \ldots + a_n\vec{e}_n$ , ist wegen der Linearität von  $L: \vec{w} = L\vec{v} = a_1L\vec{e}_1 + a_2L\vec{e}_2 + \ldots + a_nL\vec{e}_n$ .

<sup>&</sup>lt;sup>4</sup> Mathematiker bezeichnen Vektorräume mit gleicher endlicher Dimension als gleich, würden also V = W fordern. In der Physik wird gerne zwischen physikalisch unterschiedlichen Räumen unterschieden, z.B. zwischen Ortsvektoren und Geschwindigkeiten im dreidimensionalen Raum.

<sup>&</sup>lt;sup>5</sup>Wäre dies nicht so, dann könnte man aus dem zweidimensionalen Schatten einer Statue die dreidimensionale Statue rekonstruieren. Auch könnte man viele Komponenten der Darstellung eines Vektors eines hochdimensionalen Raumes umkehrbar durch weniger Komponenten ausdrücken. Die wiederholte Anwendung dieses Prinzips würde z.B. große Computerspeicher obsolet machen.

Ist K die Inverse von L, ist L auch die Inverse von K: Aus KL = 1 folgt LKL = L1 = Lund somit  $(LK)L\vec{v} = L\vec{v}$ ,  $\forall \vec{v}$ , und da für jedes  $\vec{w} \in W$  ein  $\vec{v} \in V$  existiert mit  $\vec{w} = L\vec{v}$ auch  $(LK)\vec{w} = \vec{w}$ ,  $\forall \vec{w}$ .

#### Die Inverse einer Matrix

Die Inverse einer linearen Abbildung ist ihrerseits eine lineare Abbildung<sup>6</sup>, ihre Darstellung (in denselben Basen wie L!) ist also eine quadratische Matrix mit

$$\begin{pmatrix} K_{11} & K_{12} & \dots \\ K_{21} & K_{22} & \dots \\ K_{31} & \dots & \dots \end{pmatrix} \begin{pmatrix} L_{11} & L_{12} & \dots \\ L_{21} & L_{22} & \dots \\ L_{31} & \dots & \dots \end{pmatrix} = \begin{pmatrix} 1 & 0 & \dots \\ 0 & 1 & \dots \\ 0 & 0 & 1 \\ \dots & \dots & \dots \end{pmatrix} \equiv \mathbb{1} .$$
 (2.11)

Die sogenannte Identitätsmatrix 1 ist die Darstellung der Identitätsabbildung und hat die Werte 1 auf der Diagonalen sonst nur Nullen,  $(1)_{ij} = \delta_{ij}$ . Diese Matrix bildet jeden Komponentenvektor auf sich selbst ab.

Für eine beliebige  $2 \times 2$  Matrix mit Komponenten a, b, c, d ist die Inverse

$$\frac{1}{ad-bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix} \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$
(2.12)

(nachrechnen!). Die Inverse existiert also nur, wenn  $ad - bc \neq 0$ , einem Term, dem wir in Abschnitt 2.3 wiederbegegnen werden. Inverse von größeren Matrizen lernen wir in Abschnitt 2.4 zu berechnen.

## 2.2 Änderung der Basis

Wir betrachten einen Vektor  $\vec{v}$  mit Komponenten  $\mathbf{v}$  in der Basis  $\vec{e_1}, \vec{e_2}, \ldots, \vec{e_n}$ . Was ändert sich, wenn wir statt dieses Basissystems eine andere Basis  $\{\vec{e_i'}\}$  nutzen? Was sind die Komponenten  $\mathbf{v'}$  von  $\vec{v}$  in der neuen Basis  $\{\vec{e_i'}\}$ ?

Kernpunkt ist, dass der Vektor  $\vec{v}$  selbst sich unter Basiswechsel nicht ändert.

$$\vec{v} = \sum_{j=1}^{n} v_j \vec{e}_j = \sum_{i=1}^{n} v'_i \vec{e}'_i$$
(2.13)

 $\overline{{}^{6}aK\vec{w}_{1} + bK\vec{w}_{2} = a\vec{v}_{1} + b\vec{v}_{2} = KL(a\vec{v}_{1} + b\vec{v}_{2}) = K(a\vec{w}_{1} + b\vec{w}_{2}) \ .$ 

Schreiben wir nun die alten Basisvektoren als Linearkombination der neuen,  $\vec{e_j} = \sum_{i=1}^n S_{ij}\vec{e_i}$ , erhalten wir

$$\vec{v} = \sum_{j=1}^{n} v_j \vec{e}_j = \sum_{ij=1}^{n} v_j S_{ij} \vec{e}'_i = \sum_{i=1}^{n} \left( \sum_{j=1}^{n} S_{ij} v_j \right) \vec{e}'_i \qquad (2.14)$$
$$\equiv \sum_{i=1}^{n} v'_i \vec{e}'_i \qquad \curvearrowright \left[ v'_i = \sum_{j=1}^{n} S_{ij} v_j \right] \qquad \mathbf{v}' = \mathbf{S} \mathbf{v} \ .$$

Damit gehen die Komponenten von  $\vec{v}$  in der neuen Basis aus einer Multiplikation einer Matrix mit dem Komponentenvektor in der alten Basis hervor.

**Info:** Was ist hier passiert? Die Abbildung  $V \xrightarrow{\text{alte Basis}} \mathbb{R}^n, \vec{v} \mapsto \mathbf{v} = \mathbf{b}_{\mathbf{e}}(\vec{v})$  ist eine bijektive lineare Abbildung von V auf den  $\mathbb{R}^n$ . Analog ist  $V \xrightarrow{\text{neue Basis}} \mathbb{R}^n$  bijektiv und linear. Somit ist auch die Abbildung von Komponenten in der alten Basis zu Komponenten in der neuen Basis  $\mathbb{R}^n \to \mathbb{R}^n, \mathbf{v} \mapsto \mathbf{b}_{\mathbf{e}'}(\mathbf{b}_{\mathbf{e}}^{-1}(\mathbf{v})) = \mathbf{v}'$  eine lineare invertierbare Abbildung, nämlich die Matrixmultiplikation mit  $\mathbf{S}$ .

**Beispiel:** Wir betrachten eine orthonormale Basis in zwei Dimensionen, die um  $45^{\circ}$ gedreht wird.

$$\vec{e}_{1} = \frac{1}{\sqrt{2}} (\vec{e}_{1}' - \vec{e}_{2}')$$

$$\vec{e}_{2} = \frac{1}{\sqrt{2}} (\vec{e}_{1}' + \vec{e}_{2}')$$

$$\vec{e}_{3} = \sum_{i} S_{ij} \vec{e}_{i}' \quad \land \qquad S_{11} = \frac{1}{\sqrt{2}} = S_{22} \qquad \land \qquad \mathbf{S} = \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix} / \sqrt{2}$$

$$S_{12} = \frac{1}{\sqrt{2}}, \quad S_{21} = -\frac{1}{\sqrt{2}}$$
mit  $\mathbf{v}' = \mathbf{S}\mathbf{v}$ . Damit ist für den horizon-talen Vektor nach rechts  $\vec{v} = \vec{e}_{1}$ 

$$\mathbf{v} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \mathbf{v}' = \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix} / \sqrt{2} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} \end{pmatrix}$$
(2.15)
und analog für einen Vektor nach oben oder einen Vektor entlang der Diagonalen

$$\mathbf{v} = \begin{pmatrix} 0\\1 \end{pmatrix} \qquad \mathbf{v}' = \begin{pmatrix} 1\\1 \end{pmatrix} \frac{1}{\sqrt{2}}$$
$$\mathbf{v} = \begin{pmatrix} 1\\1 \end{pmatrix} \qquad \mathbf{v}' = \begin{pmatrix} 2\\0 \end{pmatrix} \frac{1}{\sqrt{2}} = \begin{pmatrix} \sqrt{2}\\0 \end{pmatrix}$$

Bei letzten Beispiel lag der Vektor parallel zu einem Vektor des Basissystems  $\vec{e}'_i$ , damit hat er nur eine von Null verschiedene Komponente in der neuen Basis.

## Transformation von Matrizen

Analog zur Transformation von Vektorkomponenten lässt sich bestimmen wie die Einträge von Matrizen unter einem Basiswechsel transformieren.

$$\mathbf{w} = \mathbf{L}\mathbf{v}$$
 $\mathbf{w}' = \mathbf{S}\mathbf{w}$  $\mathbf{w} = \mathbf{S}^{-1}\mathbf{w}'$  $\mathbf{w}' = \mathbf{L}'\mathbf{v}'$  $\mathbf{v}' = \mathbf{S}\mathbf{v}$  $\mathbf{v} = \mathbf{S}^{-1}\mathbf{v}'$ (2.16)

$$\mathbf{S}^{-1}\mathbf{w}' = \mathbf{L}\mathbf{S}^{-1}\mathbf{v}'$$

$$\mathbf{w}' = \mathbf{S}\mathbf{L}\mathbf{S}^{-1}\mathbf{v}' \qquad \frown \qquad \mathbf{L}' = \mathbf{S}\mathbf{L}\mathbf{S}^{-1}$$
(2.17)

•

ergibt die Transformation von Komponenten einer Matrix unter Änderung der Basis. Basiswechsel sind sogenannte **passive** Transformationen, die Vektoren ändern sich nicht, lediglich ihre Komponenten (und die Basis). Im Gegensatz dazu ändern sich unter **aktiven** Transformationen die Vektoren selbst, die Basis bleibt allerdings im allgemeinen unverändert.

## 2.3 Die Determinante

Unter einer (aktiven) Transformation ändern sich auch die Volumina von geometrischen Objekten: in einer Dimension ändern sich die Längen, in zwei Dimensionen Flächeninhalte, etc., in drei Dimensionen die dreidimensionalen Volumina, etc<sup>7</sup>. Die Determinante einer Matrix gibt an wie sich Volumina unter der entsprechenden aktiven Transformation ändern.

Hier ein zweidimensionales Beispiel:



In diesem Beispiel, bei dem die x-Komponente um einen Faktor 2 gestreckt werden, ändert sich das zweidimensionale Volumen um einen Faktor 2.

Wir betrachen eine lineare Abbildung L von einem n-dimensionalen Vektorraum auf sich selbst. Die **Determinante** einer linearen Abbildung det(L) ist die vorzeichenbehaftete Volumen des Abbildung des n-dimensionalen Einheitswürfels unter der linearen Abbildung (aktive Transformation).

Vorzeichenbehaftet bedeutet, dass wenn sich die Orientierung der Basisvektoren unter L ändert, det(L) das Vorzeichen ändert. Die Volumenänderung ist unabhängig von der gewählten Basis. Gegeben eine Basis und damit eine Matrixdarstellung  $\mathbf{L}$  der Abbildung, bezeichnen wir ihre Determinante auch als Determinante der Matrix, det $(\mathbf{L}) \equiv \det(L)$ . Als erstes berechnen wir die Determinante von  $n \times n$  Matrizen unterschiedlicher Größe.

n = 1

 $\mathbf{L} = a$ , die einzige Komponente ändert sich also um den Faktor a, die Abbildung des Einheitswürfels (ein Vektor mit Länge a) hat das 'Volumen' a, also det(L) = a.

 $<sup>^7</sup>$ Unabhängig von der Dimension werden wir diese Größen als 'Volumen' im n-dimensionalen Raum bezeichnen.

n = 3

Aus Abschnitt 1.6 zum Kreuzprodukt kennen wir bereits das (vorzeichenbehaftete) Volumen eines Parallelepipeds. Die Bildpunkte der drei Basisvektoren  $\mathbf{u}, \mathbf{v}, \mathbf{w}$  spannen ein Parallelepiped mit vorzeichenbehaftetem Volumen  $\mathbf{u} \cdot \mathbf{v} \times \mathbf{w}$ , die Determinante einer  $3 \times 3$  Matrix ist also

$$\det \begin{pmatrix} u_1 & v_1 & w_1 \\ u_2 & v_2 & w_2 \\ u_3 & v_3 & w_3 \end{pmatrix} = u_1(v_2w_3 - v_3w_2) - u_2(v_1w_3 - v_3w_1) + u_3(v_1w_2 - v_2w_1) .$$
(2.18)

Verallgemeinerung auf n > 3

Eine Abbildung vom Raum der Matrizen auf die reellen Zahlen,  $\mathbb{R}^{n \times n} \to \mathbb{R}$  heißt Determinante, wenn und nur wenn gilt:

1. Normierung: 
$$\det \begin{pmatrix} 1 & 0 & 0 & \cdots \\ 0 & 1 & 0 & \cdots \\ 0 & 0 & 1 & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix} = 1$$

#### 2. Multilinearität:

$$\det \begin{pmatrix} a_{11} & \cdots & \lambda c_1 + \mu d_1 & \cdots \\ \vdots & \vdots & \lambda c_2 + \mu d_2 & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix}$$
$$= \lambda \det \begin{pmatrix} a_{11} & \cdots & c_1 & \cdots \\ \vdots & \vdots & c_2 & \cdots \\ \vdots & \vdots & c_3 & \ddots \end{pmatrix} + \mu \det \begin{pmatrix} a_{11} & \cdots & d_1 & \cdots \\ \vdots & \vdots & d_2 & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix}$$
$$3. \text{ alternierend:} \quad \det \begin{pmatrix} a_{11} & \cdots & a_1 & \cdots & a_1 & \cdots \\ a_{21} & \cdots & a_2 & \cdots & a_2 & \cdots \\ \vdots & \vdots & a_3 & \vdots & a_3 & \ddots \end{pmatrix} = 0$$

Diese Forderungen lassen sich aus geometrischen Überlegungen zu den Bildpunkten der Vektoren, die den Einheitswürfel aufspannen, herleiten: Die Normierung fordert, dass die Volumenänderung unter der Einheitsabbildung gleich eins ist. Zur Multilinearität betrachten wir die Spalten der

Matrix **L** als Bildvektoren der Basisvektoren  $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$  und  $\begin{pmatrix} 0 \\ 1 \end{pmatrix}$  (hier in



2D). Das graue Parallelogramm rechts (aufgespannt durch die Vektoren  $\mathbf{a}$  und  $\mathbf{b} + \mathbf{c}$  hat denselben Fächeninhalt wie die rote und grüne Fläche (aufgespannt durch  $\mathbf{a}$  und  $\mathbf{b}$ , bezw.  $\mathbf{a}$  und  $\mathbf{c}$ ) zusammen. Zur alternierende Eigenschaft: Sind die Bildvektoren von zwei unterschiedlichen Vektoren des Einheitswürfels parallel zueinander, ist der Flächeninhalt des von ihnen aufgespannten Parallelogramms null, damit ist das Volumen der Bildpunkte des Einheitswürfels null.

Die drei Forderungen an die Determinante sind also kompatibel mit geometrischen Vorstellungen von der Volumenänderung. Erstaunlich ist, dass sie die Determinante eindeutig festlegen (für den Beweis siehe zum Beispiel Kerner und v. Wahl). Vor allem genügt es daher, eine Rechenvorschrift zu finden, die diese drei Bedingungen erfüllt, um die Determinante zu finden.

### Rekursive Berechnung der Determinante (Laplace'sche Regel)

Schreiben wir eine Matrix  $\mathbf{A}$  bei der Zeile i und Spalte j<br/> entfernt wurden als  $\mathbf{A}_{ij}$ , dann ist

$$\det(\mathbf{A}) = \sum_{k=1}^{n} (-1)^{k+1} a_{k1} \det(\mathbf{A}_{k1})$$
(2.19)

Damit lässt sich die Berechung der Determinante einer  $n \times n$  Matrix reduzieren auf die Berechung der Determinante einer  $(n-1) \times (n-1)$  Matrix, u.s.w. Die Determinante einer  $1 \times 1$ -Matrix (a) ist dann a.

Beispiel: n = 2 $\mathbf{A} = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$   $\det(\mathbf{A}) = a_{11} \det(a_{22}) - a_{21} \det(a_{12})$   $k = 1 \qquad k = 2$  n = 3  $\mathbf{A} = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}$   $\det(\mathbf{A}) = a_{11} \det \begin{pmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{pmatrix} - a_{21} \det \begin{pmatrix} a_{12} & a_{13} \\ a_{32} & a_{33} \end{pmatrix}$   $+ a_{31} \det \begin{pmatrix} a_{12} & a_{13} \\ a_{22} & a_{23} \end{pmatrix}$ 

#### Berechnung der Determinante durch Permutationen (Leibniz'sche Regel)

Bei den Ergebnissen für die Determinante in n = 2 oder n = 3 Dimensionen fällt auf, dass bestimmte Kombinationen von Indexpaarungen auftreten. Die rekursive Berechnung der Determinante generiert alle **Permutationen** der Zahlen  $1, 2, 3, \ldots, n$ . Eine Permutation von Objekten ist ihre Anordnung in einer bestimmten Reihenfolge. Für die n = 3 Zahlen 1, 2, 3 gibt es 3! = 6 Permutationen (1, 2, 3), (1, 3, 2), (2, 1, 3), (2, 3, 1), (3, 1, 2), (3, 2, 1). Definiert man das Vorzeichen  $\sigma(P)$  einer Permutation P als 1 wenn die Zahl der paarweisen Vertauschungen die von  $1, 2, 3, \ldots, n$  zu der Permutation P führen gerade ist, und -, wenn sie ungerade ist, dann ist

$$\det(A) = \sum_{P} \sigma(P) A_{1P(1)} A_{2P(2)} \dots A_{nP(n)} , \qquad (2.20)$$

wobei wir mit P(n) das Element auf dem *n*-ten Platz der Permutation bezeichnet haben.

**Info:** Die Menge aller Permutationen von n Elementen bildet eine Gruppe (s. 1.3), denn Permutationen haben ein Inverses, ein Nullelement (1, 2, ..., n) und die Addition von Permutationen (Hintereinanderausführung) ist assoziativ. Diese Gruppe wird als  $S_n$  bezeichnet. Permutationen bilden jedoch keinen Vektorraum (warum nicht?).

#### Wichtige Eigenschaften der Determinante

- 1. Die Determinante einer Matrix, deren Einträge in einer Spalte alle Einträge null sind, ist null. Dies folgt aus geometrischen Betrachtungen, da alle Vektoren in Richtung des entsprechenden Einheitsvektors auf  $\vec{0}$  abgebildet werden, oder aus (2.19).
- 2. Sind zwei Spalten einer Matrix gleich, ist die Determinante ebenso null. Dies folgt wieder aus geometrischen Überlegungen, da die von den entsprechenden Einheitsvektoren aufgespannte Ebene auf eine einzelne Richtung abgebildet wird.
- 3. Sind einzelne Spalten einer Matrix linear abhängig, ist ihre Determinante null. Zum Beweis schreiben Sie eine Spalte als Linearkombination der anderen, wenden die Multilinearität an, sowie Eigenschaft 2.
- 4. Addiert man zu einer Spalte einer Matrix  $\mathbf{A}$  eine Linearkombination der übrigen Spalten, bleibt det( $\mathbf{A}$ ) unverändert. Folgt aus Multilinearität und Eigenschaft 3.
- 5. Vertauschung von Spalten einer Matrix führt zum Vorzeichenwechsel der Determinante. Halten wir die Notation einfach und schreiben die Determinante einer Matrix mit Spalte *i* **a**<sub>i</sub> und Spalte *j* **a**<sub>j</sub> als det(**a**<sub>i</sub>, **a**<sub>j</sub>) erhalten wir durch abwechselnde Anwendung von Eigenschaft 2 und Multilinearität det(**a**<sub>i</sub>, **a**<sub>j</sub>) = det(**a**<sub>i</sub>, **a**<sub>j</sub>) + det(**a**<sub>i</sub>, **a**<sub>i</sub>) = det(**a**<sub>i</sub>, **a**<sub>j</sub> + **a**<sub>i</sub>) = det(**a**<sub>i</sub>, **a**<sub>j</sub> + **a**<sub>i</sub>) - det(**a**<sub>j</sub> + **a**<sub>i</sub>, **a**<sub>j</sub> + **a**<sub>i</sub>) = det(-**a**<sub>j</sub>, **a**<sub>j</sub> + **a**<sub>i</sub>) = - det(**a**<sub>j</sub>,**a**<sub>j</sub> +**a**<sub>i</sub>) = - det(**a**<sub>j</sub>,**a**<sub>j</sub>) - det(**a**<sub>j</sub>,**a**<sub>j</sub>) = - det(**a**<sub>j</sub>,**a**<sub>j</sub>).
- 6.  $det(\mathbf{AB}) = det(\mathbf{A}) det(\mathbf{B})$ ; folgt aus der Volumenänderung unter einer Verkettung von linearen Transformationen.
- 7. Eine Matrix und ihre Transponiert haben dieselbe Determinante,  $\det(\mathbf{A}^T) = \det(\mathbf{A})$ ,  $(\mathbf{A}^T)_{ij} = A_{ji}$ . Folgt z.B. aus der Leibnitz'schen Regel. Obige Aussagen über Spalten von Matrizen gelten also auch analog über Zeilen.

Aufgabe: Beweisen Sie diese Eigenschaften.

## 2.4 Lineare Gleichungen und lineare Algebra

m lineare Gleichungen mit n Unbekannten

$$a_{11}x_1 + a_{12}x_2 + a_{13}x_3 + \dots + a_{1n}x_n = b_1$$
  

$$a_{21}x_1 + a_{22}x_2 + a_{23}x_3 + \dots + a_{2n}x_n = b_2$$
  

$$\vdots$$
  

$$a_{m1}x_1 + a_{m2}x_2 + a_{m3}x_3 \dots + a_{mn}x_n = b_m$$
  
(2.21)

(2.22)

können kompakt in Matrixform  $\mathbf{A}\mathbf{x} = \mathbf{b}$  geschrieben werden mit

$$\mathbf{A} = \begin{pmatrix} a_{11} & a_{12} & a_{13} & \cdots \\ a_{21} & a_{22} & a_{23} & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix} \qquad \mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \qquad \mathbf{b} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix}$$

**Beispiel:** Wann gibt es eine Lösung solcher Gleichungen ? Für n = m = 2 multiplizieren wir beide Gleichungen auf beiden Seiten

$$a_{11}x_1 + a_{12}x_2 = b_1 \qquad | \times a_{22}$$
  
$$a_{21}x_1 + a_{22}x_2 = b_2 \qquad | \times a_{12}$$

Subtraktion führt auf die explizite Lösung

$$(a_{11}a_{22} - a_{12}a_{21})x_1 = a_{22}b_1 - a_{12}b_2$$
$$x_1 = \frac{a_{22}b_1 - a_{12}b_2}{a_{11}a_{22} - a_{12}a_{21}}$$

Damit existiert für  $\mathbf{b} \neq 0$  nur dann eine Lösung der Gleichungen wenn die Determinante von  $\mathbf{A}$  nicht null ist. Und für  $\mathbf{b} = 0$  existiert eine **nicht-triviale** Lösung (bei der nicht alle  $x_i$  gleich null sind) nur wenn det $(\mathbf{A}) = 0$ . Die Lösbarkeit linearer Gleichungen ist also eng verknüpft mit Eigenschaften einer linearen Abbildung die durch  $\mathbf{A}$  dargestellt wird, konkret mit der Determinante det $(\mathbf{A})$ .

Der Zusammenhang zwischen den Determinanten von Matrizen und linearen Gleichungen lässt sich geometrisch verstehen. Die Lösungen einer einzelnen linearen Gleichung  $a_1x_1 + a_2x_2 + \ldots + a_nx_n = b$  in *n* Variablen beschreiben eine Ebene im *n*-dimensionalen Raum (genauer, eine sogenannte **Hyperebene** mit Dimension n-1). Diese Ebene ist definiert durch alle Vektoren **x**, deren Skalarprodukt **a** · **x** gleich *b* ist, also deren Projektion auf
den Vektor **a** senkrecht zu dieser Ebene gleich b ist. Ein System von m linearen Gleichungen beschreibt damit m Ebenen. Diese Ebenen können sich in einer  $n - 1, n - 2, n - 3, \ldots$ -dimensionalen Ebene schneiden (wenn einzelne Ebenen zusammenfallen), einer Linie, einem Punkt, oder in der leeren Menge (wenn die Ebenen alle parallel sind). Entsprechend kann ein System von linearen Gleichungen unendlich viele Lösungen haben, eine Lösung, oder keine Lösung.

Im Allgemeinen haben m Gleichungen in n Variablen einen n - m-dimensionalen Lösungsraum. Für m Gleichungen und n Unbekannte gibt es damit im Allgemeinen eine einzige Lösung  $\mathbf{x} = \mathbf{A}^{-1}\mathbf{b}$ . Wenn einzelne Gleichungen allerdings nicht **logisch unabhängig** von den anderen sind, also durch Multiplikation und Addition anderer Gleichungen gewonnen werden können, kann es vorkommen dass m Gleichungen in n Variablen entweder viele oder keine Lösungen haben. Dann ist  $\mathbf{A}$  auch nicht invertierbar. Zum Beispiel sind die Gleichungen  $3x_1 + x_2 = 6$  und  $6x_1 + 2x_2 = 12$  nicht logisch unabhängig, entsprechend gibt es viele Lösungen. Auf diese Weise lassen sich auch **inkonsistente** Gleichungen konstruieren, z.B.  $3x_1 + x_2 = 6$  und  $6x_1 + 2x_2 = 13$ . Gleichungen sind logisch unabhängig wenn keine der Zeilen von  $\mathbf{A}$  durch eine Linearkombination der anderen gewonnen werden kann, d.h. wenn die Zeilen von  $\mathbf{A}$  linear unabhängig sind. Sind die Zeilen von  $\mathbf{A}$  linear abhängig, ist det $(\mathbf{A}) = 0$ , die Determinante einer  $n \times n$  Matrix bestimmt ('determiniert') also ob ein System von n Gleichungen in n Unbekannten eine eindeutige Lösung hat. Analog ist  $\mathbf{A}$ invertierbar wenn und nur wenn ihre Determinante nicht null ist.

Ein einfacher Algorithmus<sup>8</sup> zur Lösung von linearen Gleichungen basiert auf der linearen Kombination von Gleichungen zur Elimination einzelner Variablen, wie in dem Beispiel oben. Bei der **Gauß'schen Elimination** schreibt man das das Gleichungsystem als sogenannte **vergrösserte Matrix** aus **A** und **b** nebeneinander gesetzt, z.B.

$$\begin{bmatrix} 1 & 3 & -2 & 5 \\ 3 & 5 & 6 & 7 \\ 2 & 4 & 3 & 8 \end{bmatrix}$$
(2.23)

und formt diese Matrix um. Jeder Umformungsschritt entspricht der Multiplikation von Gleichungen mit Zahlen und der Addition von Gleichungen. Konkret sind folgende elementaren Operationen erlaubt:

- Vertauschung von zwei Zeilen
- Multiplikation einer Zeile mit einer Zahl  $\neq 0$
- Addition eines Multiplen einer Zeile zu einer anderen Zeile

 $<sup>^{8}</sup>$ Ein Algorithmus ist eine eindeutige Handlungsvorschrift zur Lösung eines Problems und besteht typischerweise aus vielen Einzelschritten. Dabei bietet sich oft die Implementierung auf einem Computer an.

Ziel ist es durch diese Schritte die Matrix links zu einer Identitätsmatrix umzuformen, rechts steht dann der Lösungsvektor. In unserem Beispiel

$$\begin{bmatrix} 1 & 3 & -2 & | & 5 \\ 3 & 5 & 6 & | & 7 \\ 2 & 4 & 3 & | & 8 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 3 & -2 & | & 5 \\ 0 & -4 & 12 & | & -8 \\ 2 & 4 & 3 & | & 8 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 3 & -2 & | & 5 \\ 0 & -4 & 12 & | & -8 \\ 0 & -2 & 7 & | & -2 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 3 & -2 & | & 5 \\ 0 & 1 & -3 & | & 2 \\ 0 & -2 & 7 & | & -2 \end{bmatrix}$$
$$\rightarrow \begin{bmatrix} 1 & 3 & -2 & | & 5 \\ 0 & 1 & -3 & | & 2 \\ 0 & 0 & 1 & | & 2 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 3 & -2 & | & 5 \\ 0 & 1 & 0 & | & 8 \\ 0 & 0 & 1 & | & 2 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 3 & -2 & | & 5 \\ 0 & 1 & 0 & | & 8 \\ 0 & 0 & 1 & | & 2 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 3 & 0 & | & 9 \\ 0 & 1 & 0 & | & 8 \\ 0 & 0 & 1 & | & 2 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 0 & 0 & | & -15 \\ 0 & 1 & 0 & | & 8 \\ 0 & 0 & 1 & | & 2 \end{bmatrix} .$$
$$(2.24)$$

haben wir im ersten Schritt 3 mal der ersten Zeile von der zweiten abgezogen, um den ersten Koeffizienten auf Null zu setzen, dann analog für den ersten Koeffizienten der dritten Zeile. Nachdem damit die Koeffizienten der ersten Spalte bis auf den ersten auf Null gesetzt sind gehen wir zur zweiten Spalte über und so weiter, bis das Dreieck unter der Diagonalen nur die Einträge Null enthält. Dann gehen wir zum Dreieck über der Diagonalen und erhalten zum Schluss die Lösung  $x_1 = -15, x_2 = 8, x_3 = 2$ .

Mit dieser Methode lässt sich auch die Inverse einer Matrix **A** berechnen, nämlich durch Lösen der Gleichungen  $\mathbf{A}\mathbf{x}_1 = \begin{pmatrix} 1 \\ 0 \\ \dots \end{pmatrix}, \mathbf{A}\mathbf{x}_2 = \begin{pmatrix} 0 \\ 1 \\ \dots \end{pmatrix}, \dots,$  bzw. in Form einer vergrößerten

Matrix z.B.

$$\begin{bmatrix} 1 & 3 & -2 & | & 1 & 0 & 0 \\ 3 & 5 & 6 & | & 0 & 1 & 0 \\ 2 & 4 & 3 & | & 0 & 0 & 1 \end{bmatrix} .$$
(2.25)

Wird hier nach obigem Schema der linke Teil der vergrößerten Matrix zu einer Identitätsmatrix transformiert steht rechts dann die gesuchte Inverse.

## 2.5 Eigenvektoren und Eigenwerte

Eigenvektoren einer (aktiven) linearen Transformation sind Vektoren, die unter der Transformation ihre Richtung nicht ändern. Der transformierte Vektor ist also parallel zum ursprünglichen Vektor, er kann also nur um einen bestimmten Faktor gestreckt oder gestaucht werden. Dieser Faktor wird als Eigenwert bezeichnet.

Eigenvektoren werden sich als mächtiges Werkzeug bei der Beschreibung schwingender Systeme mit vielen Freiheitsgraden erweisen. Weitere wichtige Anwendungen liegen in der Quantenmechanik. Der Grund warum Eigenvektoren in vielen Gebieten der Physik und weit darüber hinaus eine Rolle spielen, ist dass sie die Wirkung einer Matrix (Darstellung einer linearen Transformation) effizient beschreiben: Kann man Eigenvektoren einer linearen Abbildung als Basis nutzen, ist die entsprechende Matrix diagonal (die Komponenten eines Vektors werden unter der Abbildung mit den Eigenwerten multipliziert).



Sei  $\mathbf{A} \in \mathbb{R}^{n \times n}$ . Ein von Null verschiedener Vektor  $\mathbf{v}$  heißt **Eigenvektor** von  $\mathbf{A}$ , wenn ein  $\lambda$  existiert mit  $\mathbf{A}\mathbf{v} = \lambda \mathbf{v}$ .  $\lambda$  heißt **Eigenwert** von  $\mathbf{v}$ .

Prinzipiell sind n unterschiedliche Eigenwerte möglich. (Nicht-parallele) Eigenvektoren mit demselben Eigenwerten heißen **degeneriert**. Eigenvektoren zum selben Eigenwert  $\lambda$  spannen einen Raum aus Eigenvektoren (mit Eigenwert  $\lambda$ ) auf.

## Die Bestimmung von Eigenvektoren

Wir schreiben  $\mathbf{Av} = \lambda \mathbf{v} = \lambda \mathbf{1} \mathbf{v}$  als  $(\mathbf{A} - \lambda \mathbf{1})\mathbf{v} = \mathbf{0}$ . In Abschnitt 2.1.1 hatten wir gezeigt, dass eine bijektive lineare Abbildung *L* Nullvektoren auf Nullvektoren abbildet,  $L\vec{0} = \vec{0}$ und  $L^{-1}\vec{0} = \vec{0}$ . Wenn also eine nichttriviale Lösung  $\mathbf{v} \neq \mathbf{0}$  von  $(\mathbf{A} - \lambda \mathbf{1})\mathbf{v} = \mathbf{0}$  existieren soll, darf  $(\mathbf{A} - \lambda \mathbf{1})$  also nicht invertierbar sein, muss also Determinante null haben. Für eine  $n \times n$ -Matrix können wir det $(\mathbf{A} - \lambda \mathbf{1}) = 0$  als polynomiale Gleichung n-ter Ordnung in  $\lambda$ auffassen und lösen. Diese Gleichung wird als **charakteristisches Polynom** bezeichnet. **Beispiel:** 

$$\mathbf{A} = \begin{pmatrix} 2 & -4 \\ -1 & -1 \end{pmatrix} \qquad 0 = \det(\mathbf{A} - \lambda \mathbf{I}) = \det\begin{pmatrix} 2 - \lambda & -4 \\ -1 & -1 - \lambda \end{pmatrix}$$
$$= -(2 - \lambda)(1 + \lambda) - 4 = \lambda^2 - \lambda - 6 = (\lambda + 2)(\lambda - 3)$$
mit Lösungen  $\lambda_1 = 3, \lambda_2 = -2$ 

Eigenvektoren lassen sich aus  $\mathbf{A}\mathbf{v} = \lambda \mathbf{v}$  (bei bekanntem  $\lambda$ ) berechnen.

$$(2-3)v_1 - 4v_2 = 0$$
  
-1v\_1 - (1+3)v\_2 = 0  $\Rightarrow v_1 = -4v_2$ 

Zu jedem Eigenwert  $\lambda$  (hier  $\lambda_1 = 3$ ) ergibt sich also eine Richtung (hier  $v_1 = -4v_2$ ), in der Eigenwektoren zu diesem Eigenwert liegen.  $v_1 = 4, v_2 = -1$  ist zum Beispiel ein Eigenvektor, seine und seine Vielfache ebenso (alle mit Eigenwert  $\lambda_1$ ):  $v_1 = 4, v_2 = -1$ spannt den eindimensionalen Eigenraum zum Eigenwert  $\lambda_1$  auf.

**Aufgabe:** Eigenvektoren gibt es natürlich nicht nur im  $\mathbb{R}^n$ . Wir betrachten den Vektorraum der Funktionen einer Veränderlichen, sowie die durch Ableitung definierte lineare Transformation  $f(x) \mapsto \frac{df}{dx}$ . Geben sie die Eigenvektoren und Eigenwerte an. Hinweis: Denken Sie an Exponentialfunktionen. Drei

Analysis

Analysis beinhaltet die Themen Grenzwerte, Differentiation und Integration. In der Physik spielen diese Themen eine zentrale Rolle bei der Beschreibung der Veränderung physikalischer Größen in Raum und Zeit: die Position eines Teilchens als Funktion der Zeit in der Mechanik, oder die räumliche und zeitliche Änderung elektromagnetischer Felder. Wir diskutieren die Differentiation von Funktionen einer und mehrerer Veränderlicher sowie Integrale über eine und mehrere Variablen.

Einiges dieser Themen ist Ihnen aus der Schule geläufig, wir beginnen daher nur mit einem kurzen Abriss der Themen Grenzwert, Differentiation und Integration von Funktionen einer Veränderlichen.

## 3.1 Grenzwert

Betrachten wir als Beispiel  $g(x) = (x - x_0)^2$ . Nähert sich x dem Wert  $x_0$ , nähert sich g(x) der Null an. Was bedeutet das "nähert sich der Null an"? Eine Abweichung von  $10^{-10}$  von Null? Oder  $10^{-100}$ ?

Der Sachverhalt lässt sich präziser formulieren: Fordern wir, dass der Abstand zwischen g(x) und 0 einen Wert  $\epsilon > 0$  nicht übersteige. Dabei können wir  $\epsilon$  so klein wählen wie wir wollen. Dann gibt es in unserem Beispiel einen Wert  $\delta$ , so dass sobald x näher als  $\delta$  an  $x_0$  liegt, diese Forderung erfüllt ist, d.h.  $|g(x)| < \epsilon$ .

Diese Denkweise wird zur Definition des Grenzwertes genutzt: Sei X eine Teilmenge von  $\mathbb{R}$  und  $x_0 \in \mathbb{R}$  ein Punkt, bei dem in jeder Umgebung unendlich viele Elemente von X liegen<sup>1</sup>. Gibt es für jedes  $\epsilon > 0$  ein  $\delta$ , so dass für alle  $x \in X$  mit  $0 < |x - x_0| < \delta$  gilt  $|g(x) - G| < \epsilon$ , dann bezeichnen wir G als **Grenzwert** von g(x) für x gegen  $x_0$ , oder

<sup>&</sup>lt;sup>1</sup>Hier fehlt noch der Begriff der Umgebung eines Punktes. Für reelle Zahlen ist das einfach, gegeben ein  $\epsilon > 0$  ist die  $\epsilon$ -Umgebung von  $x_0$  die Menge aller Punkte mit  $|x - x_0| < \epsilon$ .



Abbildung 3.1: Grenzwert von  $g(x) = x^2$  für  $x \to 0$ . Für G = 0 und  $\epsilon = 1/25$  existiert ein  $\delta$ , so dass für  $-\delta < x < \delta$  gilt  $|g(x) - G| < \epsilon$  (dunkelgraue Fläche (ein solches  $\delta$  lässt sich für jedes  $\epsilon < 0$  finden). Für G = 0.4 allerdings lässt sich für  $\epsilon = 1/25$  kein solches  $\delta$  finden (hellgraue Fäche).

 $\lim_{x\to x_0} g(x) = G$ . Der Unterschied zwischen dem Funktionswert g(x) und dem Grenzwert wird also beliebig klein, wenn man x genügend nahe bei  $x_0$  wählt. Dazu muss g(x) bei  $x_0$  nicht einmal definiert sein.

**Beispiel:** In unserem Beispiel  $g(x) = (x - x_0)^2$  ist  $g(x) - 0 < \epsilon$  für alle Werte von x mit  $|x - x_0| < \delta = \sqrt{\epsilon}$ .  $g(x) - G < \epsilon$  für alle Werte von x im Intervall  $x_0 - \delta < x < x_0 + \delta$  lässt sich nur mit G = 0 für alle Werte von  $\epsilon$  erfüllen. G = 0 ist daher als Grenzwert eindeutig bestimmt, siehe Abbildung.

Aus diesem Beispiel entnehmen wir, daß sich Grenzwerte nicht einfach "ausrechnen" lassen, wir haben zunächst G geraten und verifiziert. (Wir hätten aber auch nicht geraten, dass  $(x - x_0)^2$  nahe bei  $x_0$  andere Werte als solche nahe bei Null annimmt.)

## 3.2 Differentiation

## 3.2.1 Differentiation von Funktionen einer Veränderlichen

Die Ableitung f'(x) einer Funktion f(x) beschreibt die Steigung  $s = f'(x = x_0)$  von f(x) am Punkt  $x_0$ . Diese Formulierung verdeckt allerdings, wie nützlich das Konzept der Ableitung in der Praxis ist: Für eine hinreichend "glatte" Funktion schmiegt sich f(x)bei  $x_0$  an die Gerade durch den Punkt mit Koordinaten  $(x_0, f(x_0))$  mit Steigung s, f(x) kann also nahe  $x_0$  durch diese Gerade approximiert werden. Lokal können wir also die (im Allgemeinen) nichtlineare Funktion f(x) durch eine lineare Funktion approximieren



$$f(x) \approx f(x_0) + (x - x_0)f'(x_0)$$
, (3.1)

**Info:** Der Unterschied zwischen der Funktion f(x) und der beschriebenen Geraden ist wieder eine Funktion von x, die bei x = 0 verschwindet. Schreiben wir diesen Unterschied als xg(x) und approximieren g(x) wiederum durch eine Gerade, und führen dieses Prinzip iterativ fort, kommen wir auf eine Näherung für f(x), die im Prinzip beliebig genau sein kann. Diese Idee führt in Kapitel 5 auf die sogenannte Taylor-Reihe.

Wie können wir diese Ableitung (Steigung der Tangente zu f(x) an einem Punkt  $x_0$ ) bestimmen? Eine Gerade durch die Punkte mit Koordinaten  $(x_0, f(x_0))$  und  $(x_1, f(x_1))$ (genannt **Sekante**) hat Steigung  $S = \frac{f(x_1) - f(x_0)}{x_1 - x_0}$ . Je näher  $x_1$  bei  $x_0$  liegt, desto besser approximiert S die Steigung von f(x) bei  $x_0$ . Setzt man aber einfach  $x_1 = x_0$  ist allerdings der Nenner dieses Bruchs Null! Wir nutzen daher den Grenzwert:

Der Grenzwert  $f'(x) = \lim_{\delta x \to 0} \frac{f(x+\delta x) - f(x)}{\delta x}$  heißt (rechte) **Ableitung** einer Funktion f(x) $(f : \mathbb{R} \to \mathbb{R})$  am Punkt x (wenn der Grenzwert existiert). Analog definiert  $\lim_{\delta x \to 0} \frac{f(x-\delta x) - f(x)}{\delta x}$  die linke Ableitung, die sich von der rechten aber nur unterscheidet wenn f(x) bei  $x_0$  einen Knick hat.

In der Nähe von x lässt sich dann die Funktion durch eine Gerade  $f(x) \approx f(x_0) + (x - x_1)f'(x_0)$  nähern. Etwas präziser schreiben wir  $f(x) = f(x_0) + (x - x_0)f'(x_0) + \mathcal{O}((x - x_0)^2)$ , wobei der letzte Term für Terme der Ordnung  $(x - x_0)^2$  steht<sup>2</sup>.

<sup>&</sup>lt;sup>2</sup>  $\mathcal{O}(x^n)$  bedeutet "von der selben Ordnung wie  $x^n$  im Grenzfall  $x \to 0$ ", präziser  $f(x) = \mathcal{O}(g(x))$  wenn  $0 < \lim_{x \to 0} f(x)/g(x) < \infty$ . (Dieselbe Notation wird auch verwendet um die Asymptotik einer Funktion bei  $x \to \infty$  zu beschreiben, welche Asymptotik gemeint ist folgt aus dem Kontext.)

**Beispiel:** Betrachten wir die Funktion  $f(x) = x^2$ . Mit  $(x + \delta x)^2 = x^2 + 2x\delta x + \delta x^2$  ist

$$f'(x) = \lim_{\delta x \to 0} \frac{f(x+\delta x) - f(x)}{\delta x} = \lim_{\delta x \to 0} \frac{(x+\delta x)^2 - x^2}{\delta x} = \lim_{\delta x \to 0} (2x+\delta x) = 2x .$$
(3.2)

Die Ableitung höherer Polynome geht analog. Bei der Ableitung von Exponentialfunktionen  $f(x) = a^x$  erhalten wir

$$f'(x) = \lim_{\delta x \to 0} \frac{a^{x+\delta x} - a^x}{\delta x} = a^x \lim_{\delta x \to 0} \frac{a^{\delta x} - 1}{\delta x} .$$
(3.3)

Ableitung der Exponentialfunktion ist also die Exponentialfunktion selbst, mal eine Konstante. e = 2.7182818284590... ist definiert als die einzige positive Zahl für die diese Konstante den Wert eins hat;  $\lim_{h\to 0} \frac{e^h - 1}{h} = 1$ , damit ist die Exponentialfunktion  $e^x$  gleich ihrer Ableitung.

### Ableitung eines Vektors

Diese Definition der Ableitung funktioniert ohne Modifikation auch für die Ableitung eines Vektors, der von einer Veränderlichen abhängt. Beispiel ist eine Bahnkurve, die die Bewegung eines Massepunktes beschreibt, siehe Abschnitt 1.2. Mathematisch beschreiben wir diese Bahnkurve durch die (vektorwertige) Funktion  $\vec{r}(t)$ . Der Vektor  $\vec{r}(t+\delta t) - \vec{r}(t)$  (Differenz zweier Vektoren) gibt die Verschiebung des Massepunktes zwischen den Zeitpunkten t und  $t + \delta t$  an. Teilen wir durch  $\delta t$  erhalten wir im Grenzfall  $\delta t \to 0$  die Vektorgeschwindigkeit

$$\vec{v} = \frac{d\vec{r}}{dt} \equiv \lim_{\delta t \to 0} \frac{\vec{r}(t+\delta t) - \vec{r}(t)}{\delta t} .$$
(3.4)

Bei zeitlich konstanten Basisvektoren sind die Komponenten der Vektorgeschwindigkeit leicht zu bestimmen,

$$\vec{v} = \sum_{i=1}^{n} v_i \vec{e}_i = \frac{d}{dt} \sum_{i=1}^{n} r_i(t) \vec{e}_i = \sum_{i=1}^{n} \frac{dr_i}{dt} \vec{e}_i , \qquad (3.5)$$

und somit sind die Komponenten der Vektorgeschwindigkeit  $v_i = \frac{dr_i}{dt}$ . Ihre Einheiten sind (wegen des Faktors  $\delta t$ ) in (3.4) [Länge/Zeit], nicht wie bei dem Ortsvektor [Länge]. Trotzdem ist die Vektorgeschwindigkeit wiederum ein Vektor: Es handelt es sich um die Differenz zweier Ortsvektoren (Differenz zweier Vektoren ist ein Vektor), multipliziert mit  $1/\delta t$ , einer skalaren Größe. Würden wir also eine neue Basis nutzen, so daß sich die Komponenten von  $\vec{r}$  ändern, dann würden sich die Komponenten von  $\vec{v}$  auf die gleiche Art und Weise ändern wie die von  $\vec{r}$ . **Beispiel:** Wir kehren zurück zu der Bahnkurve aus Abschnitt 1.2, die die gleichförmige Bewegung eines Massepunktes entlang eines Halbkreises beschreibt

$$\mathbf{r}(t) = \begin{pmatrix} \cos(2\pi t) \\ \sin(2\pi t) \end{pmatrix} . \tag{3.6}$$

Die Geschwindigkeit des Massepunktes ist dann

$$\mathbf{v}(t) = \begin{pmatrix} -(2\pi)\sin(2\pi t)\\(2\pi)\cos(2\pi t) \end{pmatrix}$$
(3.7)

und die Beschleunigung  $\mathbf{a}(t) \equiv \frac{d\mathbf{v}}{dt}$ 

$$\mathbf{a}(t) = \begin{pmatrix} -(2\pi)^2 \cos(2\pi t) \\ -(2\pi)^2 \sin(2\pi t) \end{pmatrix} .$$
(3.8)

Damit stehen Geschwindigkeit und Ortsvektor bei der kreisförmigen Bewegung senkrecht zueinander, und die Beschleunigung ist parallel und in Gegenrichtung zum Ortsvektor.

Das 2te Newtonsche Gesetz,  $\vec{F} = m\vec{a}$  ist also nicht nur eine Verknüpfung der Größen Kraft, Masse und Beschleunigung, sondern enthält auch implizit die Forderung, dass die Kraft ein Vektor ist. Bei einer Basisänderung müssen sich die Komponenten der Kraft also genauso ändern wie die Komponenten eines Ortsvektors.

## 3.2.2 Die partielle Ableitung

Auch für Funktionen mehrere Veränderlicher können Ableitungen definiert werden. Betrachten wir zunächst eine Funktion f(x, y), die von zwei Variablen abhängt. Diese Funktion kann als gekrümmte Oberfläche über der x, y-Ebene visualisiert werden. Bei konstantem y = c hängt  $f(x, y = c) =: f_c(x)$ nur von x ab, so dass wir  $f'_c(x)$  definieren können. Eine zweite Ableitung kann gebildet werden, wenn statt y die Variable x = c konstant gehalten wird. Ist  $f(x_1, x_2, x_3, \ldots x_n)$  eine Funktion auf einem (offenen) Quader  $\subset \mathbb{R}^n$ , heißt der Grenzwert



$$\lim_{\delta x \to 0} \frac{f(x_1, x_2, \dots, x_i + \delta x_i, \dots) - f(x_1, x_2, \dots, x_i, \dots)}{\delta x_i} \equiv \frac{\partial f}{\partial x_i}$$
(3.9)

partielle Ableitung  $\frac{\partial f}{\partial x_i}$  von f nach  $x_i$  (wenn der Grenzwert existiert).

**Beispiel:** 

$$\frac{\partial}{\partial x_1}(x_1^3 + x_1x_2x_3 + x_1^2x_2) = 3x_1^2 + x_2x_3 + 2x_1x_2$$
$$\frac{\partial}{\partial x_2}(\sin(x_1x_2)) = x_1\cos(x_1x_2)$$

Ableitungen höherer Ordnung sind durch Hintereinander-Ausführung von  $\frac{\partial}{\partial x_i}$  und  $\frac{\partial}{\partial x_i}$ definiert. Dabei gilt die Schwarzsche Regel

$$\frac{\partial}{\partial x_i} \left( \frac{\partial}{\partial x_j} f \right) = \frac{\partial}{\partial x_j} \left( \frac{\partial}{\partial x_i} f \right) \equiv \frac{\partial^2}{\partial x_i \partial x_j} f$$
Beweis:  $\frac{\partial}{\partial x_i} \left( \frac{\partial f}{\partial x_j} \right) = \lim_{\delta x_i \to 0} \frac{1}{\delta x_i} \left( \frac{\partial f(x_1, \dots, x_i + \delta x_i, \dots x_j)}{\partial x_j} - \frac{\partial f(x_1 \dots x_i \dots x_j)}{\partial x_j} \right)$ 

$$= \lim_{\delta x_i \to 0} \lim_{\delta x_j \to 0} \frac{1}{\delta x_i \delta x_j} \left( f(x_i + \delta x_i, x_j + \delta x_j) - f(x_i + \delta x_i, x_j) - f(x_i, x_j + \delta x_j) + f(x_i, x_j) \right)$$
wenn die Grenzwerte vertauschen:
$$= \lim_{\delta x_i \to 0} \lim_{\delta x_i \to 0} \frac{1}{\delta x_i \delta x_i} \left( \dots \right) = \frac{\partial}{\partial x_i} \left( \frac{\partial f}{\partial x_i} \right)$$

$$= \lim_{\delta x_j \to 0} \lim_{\delta x_i \to 0} \frac{1}{\delta x_i \delta x_j} (\dots) = \frac{\partial}{\partial x_j} (\frac{\partial f}{\partial x_i})$$

Auch für partielle Ableitungen gibt es eine Kettenregel, die wir in einer Übung herleiten werden.

#### Integration 3.3

Die Idee der Zerlegung einer Größe in eine grosse Zahl (kleiner) Summanden tritt in den unterschiedlichsten Zusammenhängen auf, und führt stets auf eine Art von Integral. Der Ihnen bekannte Fall, das Problems die Fläche unter einer Kurve zu bestimmen, ist das prominenteste Beispiel. Dabei wird die Fläche unter einer Kurve in eine (grosse) Anzahl von Streifen zerlegt, die Fläche der (kleinen, daher fast rechteckigen) Streifen berechnet, und die Fläche der Teilstücke aufaddiert.

Betrachten wir z.B. eine Ladung Q, die sich auf einem Draht befindet. Gegeben die Ladungsdichte  $\rho(x)$  auf dem Draht pro Länge Draht (ein kleines Teilstück des Drahtes der Länge  $\Delta x$  hat also Ladung  $\rho(x)\Delta x$  ist die Gesamtladung ein Integral der Ladungsdichte über die Länge des Drahtes. Ist der Draht gerade, kommen wir auf ein Integral  $\int dx \rho(x)$ vom Typ "Fläche unter der Kurve". Die Idee des Zerlegens und Aufsummierens kleiner Terme funktioniert aber auch für einen gebogenen Draht, oder eine zweidimensionale Oberfläche oder ein dreidimensionales Volumen. Solche Probleme werden wir im nächsten Kapitel behandeln. Oberflächen- und Volumenintegrale führen auf Mehrfachintegrale, die wir hier definieren.

#### 3.3.1 Integrale von Funktionen einer Veränderlichen

Wir beginnen natürlich mit einem einfachen Fall, und greifen tatsächlich auf die "Fläche unter einer Kurve" zurück. Zur physikalischen Motivation können wir uns die Funktion f(x) als eine (lineare) Dichte in einer Dimension denken, die Masse oder Ladung eines kleinen Teilstücks der Länge  $\Delta x$  ist dann  $f(x)\Delta x$ , oder, geometisch gedacht, die Fläche eines schmalen Rechtecks mit Breite  $\Delta x$  und Höhe f(x). Klein bedeteute hier zunächst so klein, dass die Veränderung von f(x) im Intervall  $\Delta x$  vernachlässigt werden kann. Letztendlich werden wir den Grenzfall  $\Delta x \to 0$  betrachten.

Zur Berechnung der Fläche unter der Funktion f(x)zwischen  $x_a$  und  $x_b$  betrachten die Zerlegung des Intervalls  $[x_a, x_b]$  in eine grosse Anzahl N von Teilstücken;  $x_0 \equiv x_a, x_N \equiv x_b$  und  $x_1 < x_2 < x_3 \dots$ liegen zwischen  $x_a$  und  $x_b$ . Die Werte  $x_0, x_1, x_2, \dots$ werden auch als Stützstellen bezeichnet. Die Summe über alle Teilstücke der Länge  $x_1 - x_0$  (erstes Teilstück),  $x_2 - x_1$  (zweites Teilstück) ..., jeweils multipliziert mit  $f(x_0), f(x_1), \dots$  ergibt die Summe<sup>3</sup>

$$\sum_{i=0}^{N-1} (x_{i+1} - x_i) f(x_i) \equiv \sum_{i=0}^{N-1} \Delta x_i f(x_i) \qquad (3.10)$$



mit  $\Delta x_i \equiv x_{i+1} - x_i$ . Diese Summe wird als Riemann-Summe bezeichnet. Im nächsten Schritt machen wir die Längen der Teilstücke sehr klein; die Rechtecke der Abbildung approximieren die Fläche unter der Kurve also immer besser. Den Grenzfall bei dem die Längen der Teilstücke  $\Delta x_i$  gegen Null gehen, bezeichnen wir als das Riemannintegral

$$\int_{x_a}^{x_b} \mathrm{d}x \, f(x) = \lim_{\Delta x_i \to 0} \sum_{i=0}^{N-1} \Delta x_i f(x_i) \,. \tag{3.11}$$

In diesem Grenzfall geht natürlich auch die Zahl der Teilstücke gegen unendlich. Wie genau der Grenzfall  $\Delta x_i \to 0$  für alle  $\Delta x_i$  durchzuführen ist, ist damit noch nicht gesagt. Es stellt sich als ausreichend heraus, die Länge des längsten Teilstückes max<sub>i</sub>( $\Delta x_i$ ) gegen Null zu

<sup>&</sup>lt;sup>3</sup> Hier kann man fragen ob nicht  $f(x_{i+1})$  statt  $f(x_i)$  der korrekte Ausdruck wäre. Kernpunkt des Beweises der Existenz des sogenannten Riemann-Integrals ist, dass der Unterschied zwischen diesen beiden Fällen im Grenzfall  $N \to \infty$  verschwindet.

führen (siehe z.B. Berendt und Weimar). Funktionen, für die das Integral (3.11) existiert, heißen Riemann-integrierbar.

#### Das Integral und die Stammfunktion

Betrachen wir die Fläche unter der Kurve im Intervall [0, x] als Funktion von x, genannt F(x). Vergrößern wir nun das Intervall durch Hinzufügen eines (beliebig kleinen) Teilstücks  $\Delta x$ , vergrössert sich die Fläche um  $\Delta x f(x)$ , also  $F(x + \Delta x) = F(x) + \Delta x f(x) + \mathcal{O}(\Delta x^2)$  und somit

$$\frac{dF}{dx} = f(x) \ . \tag{3.12}$$

Aus dieser Gleichung ist die sogenannten **Stammfunktion** F(x) von f(x) nur bis auf eine additive Konstante bestimmt. Bei der Subtraktion von Stammfunktionen verschwindet diese Konstante allerdings und für allgemeine Integralgrenzen gilt dann

$$\int_{a}^{b} \mathrm{d}x \, f(x) = F(b) - F(a) \,. \tag{3.13}$$

Diese Aussage wird als Fundamentalsatz der Analysis bezeichnet. Integration und Differentiation sind also inverse Operationen; um das Integral einer Funktion f(x) zu bestimmen, suchen wir die Funktion F(x), deren Ableitung nach x gleich f(x) ist. Wir schreiben auch das sogannte **unbestimmte Integral** ohne die Integrationsgrenzen anzugeben als  $\int dx f(x) = F(x)$  (Gleichheit verstanden als "bis auf eine unbestimmte Konstante").

#### Techniken zur Berechnung von Integralen

Wichtigste Technik ist die Ableitungen vieler Funktionen zu kennen, und so die Stammfunktion zu raten und zu verifizieren: Eine allgemeine Methode um systematisch zu einer beliebigen Funktion die Stammfunktion anzugeben gibt es nicht. Es folgen einige nützliche Strategien.

#### Partielle Integration

Partielle Integration ist eine nützliche Technik, wenn der Integrand f(x) sich als Produkt zweier Funktionen schreiben lässt. Sei f(x) = u'(x)v(x). Aus der Produktregel der Ableitung erhalten wir  $\frac{d}{dx}(u(x)v(x)) = u'(x)v(x) + u(x)v'(x)$ . Umstellen der Terme und Integration auf beiden Seiten ergibt

$$\int_{a}^{b} \mathrm{d}x \, u'(x)v(x) = \int_{a}^{b} \mathrm{d}x \, \frac{d}{dx} \left( u(x)v(x) \right) - \int_{a}^{b} \mathrm{d}x \, u(x)v'(x) = \left[ u(x)v(x) \right]_{a}^{b} - \int_{a}^{b} \mathrm{d}x \, u(x)v'(x) \, .$$
(3.14)

Bei geschickter Wahl von u(x) und v(x) ist das resultierende Integral oft einfach zu bestimmen.

#### Integration durch Substitution

Dieses Integrationsprinzip folgt aus der Kettenregel der Differentialrechnung. Gegeben sei eine Funktion f(x) mit Stammfunktion F(x). Wir betrachten nun eine Funktion x(y)und definieren eine weitere Stammfunktion G(y) = F(x(y)) mit Ableitung g(y). Nach der Kettenregel ist  $\frac{dG}{dy} = \frac{dF}{dx}\frac{dx}{dy}$ , bezw.  $g(y) = f(x(y))\frac{dx}{dy}$ . Integration über x ergibt

$$\int_{y(a)}^{y(b)} \mathrm{d}y \, \frac{\mathrm{d}x}{\mathrm{d}y} f(x(y)) = \int_{y(a)}^{y(b)} \mathrm{d}y \, g(y) = [G(y)]_{y(a)}^{y(b)} = [F(x(y))]_{y(a)}^{y(b)} = [F(x)]_a^b = \int_a^b \mathrm{d}x \, f(x) \,.$$
(3.15)

Ist f(x) schwer zu integrieren, kann eine geschickt gewählte Substitution auf eine Funktion  $\frac{dx}{dy}f(x(y))$  führen, die leichter zu integrieren ist.

**Beispiel:** Gesucht sei das Integral der Funktion  $\sqrt{1-x^2}$  über das Intervall [0, 1]. Da  $\sqrt{1-\sin(t)^2} = \cos(t)$  erwarten wir, dass eine trigonometrische Variablensubstitution nützlich könnte. Wir führen eine neue Variable  $\sin(y) = x$ , bezw.  $y = \arcsin(x)$  ein und erhalten mit  $\frac{dx}{dy} = \cos(y)$ 

$$\int_0^1 \mathrm{d}x \,\sqrt{1-x^2} = \int_{\arccos(0)}^{\arcsin(1)} \mathrm{d}y \,\frac{\mathrm{d}x}{\mathrm{d}y} \sqrt{1-\sin^2(y)} = \int_0^{\pi/2} \mathrm{d}y \,\cos^2(y) \,, \tag{3.16}$$

das mit Hilfe partieller Integration berechnet werden kann. Oder mit der trigonometrischen Formel $\cos^2(y)=\frac{1+\cos(2y)}{2}.$ 

Die Variablensubstitution kann geometrisch interpretiert werden. Bei  $\int_a^b dx f(x) \approx \sum_i \Delta x f(x_i)$ unterteilen wir das Integrationintervall in Teilintervalle konstanter Länge  $\Delta x$  mit "Stützstellen" bei  $0, \Delta x, 2\Delta x, \ldots$  Dasselbe Intervall lässt sich mit neuen Stützstellen bei  $y(0), y(0) + \Delta y, y(0) + 2\Delta y, \ldots$  beschreiben, die Intervalllängen in beiden Fällen sind durch  $\Delta x \approx \Delta y \frac{dx}{dy}$ verknüpft. Eine Merkregel für die Variablensubstitution (3.15) ist  $dy = dx \frac{dy}{dx}$  (eine Aussage die hier nur unter dem Integral Sinn macht).

#### Differenzieren nach einem Parameter

Wenn der Integrand durch Ableitung nach einem Parameter aus einer Funktion hervorgeht, können wir das ausnutzen.

**Beispiel:** Gesucht sei das Integral  $\int_0^\infty dx \, x e^{-x}$ , das wir als  $\int_0^\infty dx \, x e^{-\alpha x}$  mit  $\alpha = 1$  schreiben.

$$\int_0^\infty dx \, x e^{-\alpha x} = \int_0^\infty dx \, \left(-\frac{d}{d\alpha}\right) e^{-\alpha x} = \left(-\frac{d}{d\alpha}\right) \int_0^\infty dx \, e^{-\alpha x}$$
$$= -\frac{d}{d\alpha} \left[-\frac{1}{\alpha} e^{-\alpha x}\right]_0^\infty = -\frac{d}{d\alpha} \alpha^{-1} = \frac{1}{\alpha^2}$$
(3.17)

Im zweiten Schritt haben wir Integration und Ableitung vertauscht. Für die endliche Riemannsumme (3.10) ist dies natürlich zulässig. Für das Riemannintegral ist es das mit Einschränkung auch (der Integrand muss stetig sein und seine Ableitung ebenso, siehe z.B. Berendt und Weimar).

## 3.4 Mehrfachintegrale

Ziel ist es, das Integral als Riemannsumme über eine Variable zu verallgemeinern. Anstatt (kleine) Intervalle einer Geraden (x-Achse) aufzusummieren, können wir dann kleine Flächenelemente oder Volumenelemente addieren. Damit können wir dann auch über Oberflächen und Volumen integrieren.

#### **Beispiel:**

In einem sehr beschränkten Sinne haben wir bereits ein Oberfächenintegral berechnet. Die Fläche unter der Kurve f(x) im Intervall  $[x_a, x_b]$  lässt sich nicht nur in schmale Rechecke mir Fläche  $\Delta x f(x)$  zerlegen, sondern auch in kleine Flächenelemente mit Fläche  $\Delta x \Delta y$ . Diese Flächenelemente haben ihre linke untere Ecke bei  $x_i, y_i$  und eine Fläche  $\Delta x \Delta y$ . Als Riemannsumme ist dann die Fläche unter der Kurve f(x)

$$\sum_{i=0,j=0}^{N_x-1,N_y(i)-1} (x_{i+1}-x_i)(y_{i+1}-y_i) = \sum_{\substack{i=0,j=0\\(3\ 18)}}^{N_x-1,N_y(i)-1} \Delta$$

die Summe der Flächen aller Elemente. Die Zahl der Flächenelemente entlang der y-Achse hängt hier von f(x) ab,  $N_y(i)\approx f(x_i)/\Delta y$  für konstante  $\Delta y_j$ .

Schreiben wir die Funktion  $f(x) = \int_0^{f(x)} dy$ , dann ist die Fläche unter der Kurve

$$\int_{x_a}^{x_b} \mathrm{d}x f(x) = \int_{x_a}^{x_b} \mathrm{d}x \int_0^{f(x)} \mathrm{d}y \equiv \int_{x_a}^{x_b} \int_0^{f(x)} \mathrm{d}x \mathrm{d}y \tag{3.19}$$

y

 $x_i \Delta y_j$  $x_a$  f(x)

x

 $\Delta x_i$ 

 $\Delta y_j$ 

 $x_b$ 

 $x_i, y_i$ 

Bei dieser Berechnung der Fläche wurden einzelne Flächenelemente aufsummiert. Das ist natürlich ein sehr eingeschränktes Problem. Auf einer 2D Oberfläche *S* könnte sich aber zum Beispiel elektrische Ladung mit einer Ladungsdichte  $\rho(x, y)$  pro kleiner Fläche befinden. Die Riemannsumme der Gesamtladung ist dann  $\sum_{i,j}^{N_x,N_y} \Delta x_i \Delta y_j \rho(x_i, y_j)$  und führt im Grenzfall beliebig kleiner Flächenelemente auf das Mehrfachintegral

$$\lim_{\Delta x_i \to 0, \Delta y_j \to 0} \sum_{i=0,j=0}^{N_x - 1, N_y - 1} \Delta x_i \Delta y_j \rho(x_i, y_j) = \int \int_S \mathrm{d}x \mathrm{d}y \rho(x, y) , \qquad (3.20)$$

(wenn der Grenzfall existiert)<sup>4</sup>. Dabei gilt

$$\int dx \left( \int dy \rho(x, y) \right) = \int dy \left( \int dx \rho(x, y) \right) \equiv \int \int dx dy \rho(x, y) .$$
(3.21)

<sup>&</sup>lt;sup>4</sup> Da klar ist wie viele Variablen hier integriert werden schreibt man auch  $\int_S dx dy \rho(x,y)$ , S bezeichnet die Fläche über die integriert wird.

Dieser **Satz von Fubini** (s. Berendt und Weimar) gilt für stetige Funktionen  $\rho(x, y)$  und besagt, dass es keine Rolle spielt, ob dieses Mehrfachintegral zunächst über die x-Variable oder über die y Variable ausgewertet wird. Für eine Riemannsumme wie (3.18) entsprechen diese beiden Möglichkeiten der Summenbildung über entweder horizontale oder über vertikale Teilstücke. Für Summen mit unendlich vielen Termen ist die Gleichheit nicht selbstverständlich.

**Beispiel:** Wir berechnen die Masse einer rechteckigen Platte S mit Begrenzung durch die Linien x = 0, x = 3, y = 0, y = 2 mit Dichte pro Fläche  $\rho(x, y) = xy$ . Die Masse eines kleinen rechteckigen Teilstücks mit Kangenlängen  $\Delta x, \Delta y$  ist  $\approx xy\Delta x\Delta y$ , die Gesamtmasse

$$M = \int \int_{S} dx dy(xy) = \int_{0}^{3} \int_{0}^{2} dx dy(xy) = \left(\int_{0}^{3} dxx\right) \left(\int_{0}^{2} dyy\right) = 9/2 \times 4/2 = 9(3.22)$$

Diese Vertauschbarkeit der Integrationen führt uns auf das Thema der Integrationsgrenzen, die bei Mehrfachintegralen komplizerter sein können als bei Integralen einer Veränderlichen. In (3.20) haben wir die Integrationsgrenzen nicht spezifiziert, und einfach ein Integral über die Fläche S angegeben. Die expliziten Integrationsgrenzen eines Mehrfachintegrals hängen davon ab, in welcher Reihenfolge wir die beiden Integrationen auswerten (das Ergebnis selbst hängt nicht von der Reihenfolge ab), also



ob wir die Flächenelemente in horizontale oder vertikale Streifen zusammensetzen. Werten wir im Beispiel oben zunächst das Integral (3.19) über y (bei konstantem x) aus, sind die Integrationsgrenzen einfach  $\int_{x_a}^{x_b} \int_0^{f(x)}$ . Werten wir zunächst das Integral über x (bei konstantem y) aus sind die Integrationsgrenzen bei diesem Beispiel komplizierter. Einige Oberflächen zerlegt man sinnvollerweise in Teilstücke, die separat ausgewertet werden (s. rechts).

#### Das Volumenintegral

Die gleiche Denkweise wie beim Integral über eine Oberfläche lässt sich auch auf Volumina anwenden. Das Volumen eines Quaders mit Kantenlängen  $\Delta x_i, \Delta y_j, \Delta z_k$  ist  $\Delta x_i \Delta y_j \Delta z_k$ . Gibt  $\rho(x, y, z)$  zum Beispiel eine Dichte pro kleinem Volumenelement an, dann ist die in in einem Volumen V enthaltene Masse durch die Riemannsumme

$$\sum_{i,j,k} \Delta x_i \Delta y_j \Delta z_k \rho(x_i, y_j, z_k)$$
(3.23)

approximiert, wobei die Summe über die Indizes i, j, k alle (kleinen) Volumenelemente addiert. Im Grenzfall unendlich kleiner Volumenelemente  $\lim_{\Delta x \to 0, \Delta y \to 0, \Delta z \to 0}$  erhalten wir

das Volumenintegral

$$\int_{V} \mathrm{d}x \mathrm{d}y \mathrm{d}z \rho(x, y, z) = \int \mathrm{d}x \int \mathrm{d}y \int \mathrm{d}z \rho(x, y, z) \ . \tag{3.24}$$

Wieder können die Integrationen in beliebiger Reihenfolge vorgenommen werden (unter milden Forderungen an den Integranden).

## 3.4.1 Substitution der Variablen in Mehrfachintegralen

Mehrfachintegrale sind also prinzipiell nicht schwieriger zu berechnen als Integrale von Funktionen einer Veränderlichen. Allerdings können die Integrationsgrenzen Schwierigkeiten machen. Geschickte Änderung der Integrationsvariablen kann hier helfen. Mit einer Variablensubstitution können Symmetrien eines Problems ausgenutz werden, z.B. wenn Integrationsgrenzen oder der Integrand eine Symmetrie wie die Kugelsymmetrie aufweist.



Jeder Punkt in der Ebene ist eindeutig durch ein Paar (x, y) und ein Paar  $(r, \phi), \phi \in [0, 2\pi]$  beschrieben (bis auf einen einzelnen Punkt, den Ursprung, dort ist  $\phi$  nicht eindeutig definiert), Die Abbildung von (x, y) auf  $(r, \phi)$  ist also invertierbar. Die Linien mit konstanter Variable r bilden Ringe, Linien mit konstantem  $\phi$  bilden radial nach außen laufende Geraden mit Steigung  $tan(\phi)$ . Eine kleine Änderung der Koordinaten (x, y) zu  $(x + \Delta x, y + \Delta y)$  definiert ein kleines Rechteck der Fläche  $\Delta x \Delta y$ . Eine kleine Änderung der Koordinaten  $r, \phi$  definiert eine Intersektion aus Ring und Kuchenstück, siehe Abbildung. Die Fläche dieses Objekts ist (für kleine Änderungen  $\Delta r, \Delta \phi$ , wenn dieses Objekt fast ein Recheck ist)  $\Delta r(r\Delta \phi)$ . Damit ist das Integral (3.25)

$$\int_{S} \mathrm{d}x \mathrm{d}y \rho(x, y) = \int_{0}^{1} \int_{0}^{2\pi} \mathrm{d}r \mathrm{d}\phi r \rho(r) = 2\pi \int_{0}^{1} \mathrm{d}r r \rho(r) \;. \tag{3.27}$$

Der zusätzliche Faktor r ergibt sich aus der Substitution der Variablen, man kann ihn als höherdimensionales Equivalent des Faktors  $\frac{dx}{dy}$  in (3.15) auffassen.

**Info:** Häufig auftretende Symmetrien sind die Radial- oder Kugelsymmetrie, oder eine zylindrische Symmetrie. Die dabei auftretenden Faktoren aus der Variablensubstitution lassen sich wie im Beispiel oben durch einfache geometrische Argumente bestimmen. Für generelle Variablensubstitutionen gibt es eine Verallgemeinerung der Substitutionsregel (3.15) auf Mehrfachintegrale. In n Dimensionen ist

$$\int_{V_x} \mathrm{d}x_1 \mathrm{d}x_2 \dots \mathrm{d}x_n f(\mathbf{x}) = \int_{V_y} \mathrm{d}y_1 \mathrm{d}y_2 \dots \mathrm{d}y_n \left| \det(\frac{\partial(x_1, x_2, \dots, x_n)}{\partial(y_1, y_2, \dots, y_n)}) \right| f(\mathbf{x}(\mathbf{y})) .$$
(3.28)

Dabei schreiben wir  $\mathbf{x}$  als Abkürzung für die Koordinaten  $x_1, x_2, \ldots, x_n$  und analog  $\mathbf{x}(\mathbf{y})$  als Abkürzung für  $x_1(y_1, \ldots, y_n), x_2(y_1, \ldots, y_n), \ldots$ 

Dieser **Transformationssatz** kann so verstanden werden: Betrachten wir das Netz, das von das von Linien konstanter Koordinaten gebildet wird. Für kartesische Koordinaten begrenzen diese Linien Quadrate (in 2 Dimensionen), Würfel (in 3D), bezw. *n*-dimensionale Hyperwürfel. In einem nichtkartesischen Koordinatensystem begrenzen die Linien konstanter Koordinaten andere geometrische Objekte, s. das obige Beispiel. Im Grenzfall bei dem die Änderung der Koordinaten von einer Linie zur nächsten klein ist, sind diese Objekte näherungsweise kleine Quader.

In kartesischen Koordinaten definieren  $\mathbf{x}$  und  $\mathbf{x} + \Delta \mathbf{x}$  einen (*n*-dimensionalen) Quader mit Kantenlängen  $\Delta x_1, \Delta x_2, \dots$  und Volumen  $\Delta x_1 \Delta x_2 \dots$  Die-



ser Quader ist auch durch die beiden diagonal genenüberliegenden Eckpunkte  $\vec{r}(\mathbf{x})$  und

 $\vec{r}(\mathbf{x} + \Delta \mathbf{x})$  definiert, seine Kanten sind  $\frac{\partial \vec{r}}{\partial x_i} \Delta x_i$ . In den neuen Koordinaten definieren  $\vec{r}(\mathbf{y})$  und  $\vec{r}(\mathbf{y} + \Delta \mathbf{y})$  auch einen kleinen Quader. Dessen Kanten sind durch  $\frac{\partial \vec{r}}{\partial y_j} \Delta y_j$  gegeben, mit Komponenten  $\frac{\partial x_i}{\partial y_j} \Delta y_j$ . Das Volumen des (hochdimensionalen) Spats, der durch diese Vektoren aufgespannt wird, ist die Determinante der sogenanten Jacobimatrix

$$\frac{\partial(x_1, x_2, \dots, x_n)}{\partial(y_1, y_2, \dots, y_n)} = \begin{pmatrix} \frac{\partial x_1}{\partial y_1} & \frac{\partial x_1}{\partial y_2} & \dots \\ \frac{\partial x_2}{\partial y_1} & \frac{\partial x_2}{\partial y_2} & \dots \\ \dots & \dots & \dots \end{pmatrix}$$
(3.29)

multipliziert mit  $\Delta y_1 \Delta y_2 \dots$  Diese Matrix beschreibt die lineare Abbildung vom Einheitsquader in den x-Koordinaten auf den neuen Einheitsquader in y-Koordinaten. Ihre Determinante in (3.28) beschreibt die Volumenänderung unter dieser Abbildung. Rechts ist  $\frac{\partial \vec{r}}{\partial y_j} \Delta y_j$  für das obige Beispiel (Polarkoordinaten in zwei Dimensionen) mit  $y_1 = r$  und  $y_2 = \phi$  gezeigt.

Beispiel: Wir werten die Jakobimatrix für die Transformation von kartesischen Koordinaten zu Polarkoordinaten aus. Hier erwarten wir aus den geometrischen Überlegungen oben, dass die Determinante der Jacobimatrix r ist. Mit  $x = r \cos(\phi)$  und  $y = r \sin(\phi)$ erhalten wir

$$\frac{\partial(x,y)}{\partial(r,\phi)} = \begin{pmatrix} \frac{\partial x}{\partial r} & \frac{\partial x}{\partial \phi} \\ \frac{\partial y}{\partial r} & \frac{\partial y}{\partial \phi} \end{pmatrix} = \begin{pmatrix} \cos(\phi) & -r\sin(\phi) \\ \sin(\phi) & r\cos(\phi) \end{pmatrix}$$
(3.30)

mit Determinante  $r(\cos^2(\phi) + \sin^2(\phi)) = r$ .

# Vektoranalysis

Wir führen Vektorfelder ein, die jedem Punkt im Raum einen Vektor zuordnen. Anwendungen sind Kraftfelder, wie z.B. das elektrische oder magnetische Feld. Wir diskutieren die Integration von solchen Feldern über Linien, Oberflächen und Volumina. Zur Beschreibung dieser Felder verallgemeinern wir das Konzept der Ableitung einer Funktion und führen eine Reihe von Differentialoperatoren ein.

Bisher haben wir uns mit den Elementen von Vektorräumen beschäftigt zur Beschreibung von Verschiebungen, Geschwindigkeiten, Beschleunigungen ( $\propto$  Kräften) eines Massepunktes. Mehrere Massepunkte könnten dann durch mehrere Vektoren beschrieben werden. Viele physikalische Situationen benötigen eine Erweiterung, die über den Vektorraum hinausgeht: ein Vektorfeld ordnet *jedem Punkt* im Raum einen Vektor zu.

## **Beispiel:**

- Zur Beschreibung eines strömenden Gases müssen wir berücksichtigen, dass Richtung und Geschwindigkeit der Gasteilchen vom Ort abhängen. Jedem Punkt im Raum ist also ein Vektor zugeordnet, der besagt in welche Richtung und mit welcher Geschwindigkeit Teilchen des Gases an diesem Punkt fließen. Die Geschwindigkeitsvektoren an benachbarten Punkten sind nicht unabhängig (wenn z.B. Gasteilchen erhalten bleiben).
- Das elektrische Feld  $\vec{E} = \vec{F}/q$  hat an unterschiedlichen Punkten im Raum unterschiedliche Richtungen und Stärken. Die Vektoren  $\vec{E}$  an unterschiedlichen Punkten sind nicht unabhängig, sie gehorchen den Maxwellgleichungen. Der Begriff des Vektorfeldes is zentral für den Elektromagnetismus.
- Analoges gilt für das magnetische Feld und für das Gravitationsfeld.

Sei M ein (affiner) Raum (s. 1.3), dann ist ein Vektorfeld  $\vec{v}(p)$  auf M eine Abbildung

$$\vec{v}: M \longrightarrow V, p \longmapsto \vec{v}(p)$$
, (4.1)

die jedem Punkt in M einen Vektor  $\vec{v}$  zuordnet.

Damit "sitzt" auf jedem Punkt des affinen Raumes ein Vektorraum. Die Wahl eines Vektors für jeden Raumpunkt definiert ein konkretes Vektorfeld. Ein Vektorfeld lässt sich in einer Basis darstellen, indem man für jeden Punkt im Raum eine Basis wählt.



Hier ist exemplarisch für alle Punkte des Raumes p an drei Punkten  $P_1, P_2, P_3$  eine kartesische Basis definiert. Zusätzlich haben diese Basisvektoren jeweils dieselbe Orientierung relativ zu einem kartesischen Koordinatensystem. Man spricht daher auch von einer **raumfesten Basis**. Aus nichtkartesischen Koordinaten kann man Basisysteme definieren, die nicht raumfest sind.

Mit einer Basis an jedem Punkt pkann nun der Vektor $\vec{v}$ an jedem Punkt dargestellt werden durch

$$\mathbf{v}(p) = \begin{pmatrix} v_1(p) \\ v_2(p) \end{pmatrix} \quad \text{für } \vec{v}(p) = v_1(p)\vec{e}_1 + v_2(p)\vec{e}_2 \tag{4.2}$$

und analog in höherdimensionalen Räumen.

Wählen wir für den affinen Raum M einen Ursprung  $p_0$  und schreiben  $p = p_0 + \vec{r}$ , dann ist das Vektorfeld in zwei Dimensionen charakterisiert durch

$$\mathbf{v}(\mathbf{r}) = \begin{pmatrix} v_1(\mathbf{r}) \\ v_2(\mathbf{r}) \end{pmatrix} = \begin{pmatrix} v_1(x,y) \\ v_2(x,y) \end{pmatrix} , \qquad (4.3)$$

in drei Dimensionen entsprechend

$$\mathbf{v}(\mathbf{r}) = \begin{pmatrix} v_1(x, y, z) \\ v_2(x, y, z) \\ v_3(x, y, z) \end{pmatrix}$$

In zwei Dimensionen ist ein Vektorfeld also durch zwei Funktionen von zwei Variablen charakterisiert, in drei Dimensionen durch drei Funktionen von drei Variablen. Analog zum Vektorfeld definieren wir ein **skalares Feld**, das jedem Punkt im Raum eine skalare Größe zuordnet, d.h. eine Abbildung

$$f: M \longrightarrow \mathbb{R}, p \longmapsto f(p) . \tag{4.4}$$

Ein Beispiel für ein skalares Feld ist die Dichte eines Gases, dessen Konzentration an unterschiedlichen Stellen unterschiedlich hoch sein mag, oder die Temperatur, die im Raum variiert.

## 4.1 Integration von skalaren und vektoriellen Feldern

Das Integral  $\int_{x_a}^{x_b} dx f(x)$  gibt z.B. das Gewicht eines Stabes in einer Dimension an, wenn dessen lineare Dichte f(x) ist.  $f(x)\Delta x$  ist dabei das Gewicht eines kleinen Teilstücks der Länge  $\Delta x$  des Stabes bei x. Wie könnte man nun sein Gewicht berechnen, wenn der Stab gebogen wäre? Die Zerlegung des Stabes in eine groß Zahl kleiner Teilstücke führt auf das sogenannte Linienintegral. Eine wichtige Anwendung dieses Integrals liegt in der Berechnung der Arbeit, die ein Kraftfeld entlang einer gegebenen Bahnkurve Bahnkurve verrichtet.

Die Verallgemeinerung auf Integrale von Feldern über gekrümmte Oberflächen und Volumina führt auf Flächen und Volumenintegrale. Sie spielen eine wichtige Rolle in der Elektrodynamik, wo z.B. elektrische Felder über Oberflächen oder Ladungsdichten über Volumina integriert werden.

## 4.2 Linienintegrale

Bleiben wir beim Beispiel des eindimensionalen gebogenen Drahtes mit räumlich veränderlicher Dichte. Der Draht bilde eine Kurve C im Raum. Die Dichte f pro Länge eines kurzen Teilstücks hängt nun vom Ort p ab, an dem sich das Teilstück befindet. Zerlegen wir die Kurve C in kleine Intervalle zwischen den Raumpunkten  $p_0, p_1, p_2, \ldots, p_N$  auf der Kurve C.  $\Delta \vec{r}_i \equiv p_{i+1} - p_i$  gibt die Verschiebung von  $p_i$  nach  $p_{i+1}$  an. Die Masse des *i*-ten Teilstücks ist dann die Länge des Teilstücks mal die Dichte,  $||\Delta \vec{r}_i|| f(p_i)$ . Die Gesamtmasse als Riemannsumme (vgl. die Riemannsumme (3.10)) ist damit

$$\sum_{i=0}^{N-1} ||(p_{i+1} - p_i)||f(p_i) = \sum_{i=0}^{N-1} ||\Delta \vec{r_i}||f(p_i)|.$$
(4.5)

 $q_0 = p_0 = p(t_0) = p(t_b) \bullet$ 

Zur Auswertung dieser Riemann-Summe im Grenzfall bei dem die einzelnen Schritte klein werden,  $||\Delta \vec{r_i}|| \rightarrow 0$ , definieren wir zunächst eine Parametrisierung der Kurve C. Eine Parametrisierung von C ist eine Funktion p(t), die jedem  $t \in [t_b, t_e]$  einen Punkt auf C zuordnet (b für Beginn, e für Ende). Eine **Parametrisierung** p(t) einer Kurve C als Punktmenge ist eine differenzierbare

 $q_1 = p_5 = p(t_5) = p(t_e)$ 

 $p_2 = p(t_2)$ 

 $p_1 = p(t_1)$ 

Abbildung mit  $p([t_b, t_e]) = C, p(t_b) = q_0$ (Anfangspunkt),  $p(t_e) = q_1$  (Endpunkt)

und  $\frac{dp}{dt} \neq 0$   $\forall$   $t \in [t_b, t_e]$ . Bei gegebenem Ursprung  $p_u$ , mit  $p = p_u + \vec{r}$  nutzen wir auch  $\vec{r}(t)$  als Parametrisierung der Kurve.

Eine Parametrisierung einer Raumkurve können sie durch Ameise visualisieren, die im Zeitintervall  $[t_b, t_e]$  die Raumkurve C (gebogener Draht) entlang krabbelt. Im kleinem Zeitintervall  $\Delta t_i = t_{i+1} - t_i$  legt sie die Strecke  $\approx \Delta t_i \frac{d\vec{r}(t_i)}{dt}$  zurück, dieses Teilstück trägt dann  $\approx \Delta t_i \frac{d\vec{r}(t_i)}{dt} f(p(t_i))$  zur Gesamtmasse des Drahtes bei. Dabei haben wir kleine Teilstücke durch Geraden approximiert (lineare Approximation),  $p(t_{i+1}) - p(t_i) = \vec{r}(t_{i+1}) - \vec{r}(t_i) = \Delta t_i \frac{d\vec{r}(t_i)}{dt} + \mathcal{O}(\Delta t_i^2)$ .

Im Grenzfall  $\Delta t_i \to 0 \,\forall i$  geht die Zahl dieser Teilstücke gegen unendlich und wir definieren das **Linienintegral** als

$$\int_{C} \mathrm{d}r f(\vec{r}) \equiv \lim_{||\Delta \vec{r_i}|| \to 0} \sum_{i=0}^{N-1} ||\Delta \vec{r_i}|| f(p_i) = \lim_{\Delta t_i \to 0} \sum_{i=0}^{N-1} \Delta t_i ||\frac{d\vec{r}(t_i)}{dt}|| f(p_i) = \int_{t_b}^{t_e} \mathrm{d}t ||\frac{d\vec{r}(t)}{dt}|| f(\vec{r}(t))$$
(4.6)

Bei der Notation  $\int_C \mathrm{d} r f(\vec{r})$  kann man sich dr als die Länge eines sehr kurzen Teilstücks denken.

Beispiel: Wir setzen f(p) = 1 um die Bogenlänge L des Halbkreises mit Radius 1 zu berechnen,  $L = \int_C dr$  (s. Beispiel in 1.2)  $\vec{r}(t) = \cos(2\pi t)\vec{e_1} + \sin(2\pi t)\vec{e_2}$   $\mathbf{r}(t) = \begin{pmatrix} \cos(2\pi t) \\ \sin(2\pi t) \end{pmatrix}$   $\int_c dr 1 = \int_0^{1/2} dt \left\| \left| \frac{d\vec{r}}{dt} \right| \right| 1$   $\frac{d\mathbf{r}}{dt} = 2\pi \begin{pmatrix} -\sin(2\pi t) \\ \cos(2\pi t) \end{pmatrix}$ ,  $\| \frac{d\mathbf{r}}{dt} \| = \sqrt{(2\pi \sin(2\pi t))^2 + (2\pi \cos(2\pi t))^2} = 2\pi$  $= \int_0^{1/2} dt \| \frac{d\mathbf{r}}{dt} \| = 2\pi \int_0^{1/2} dt = \pi$  (4.7)

Die Parametrisierung der Raumkurve C erlaubt also die Auswertung eines Linienintegrals als (Riemann-) Integral über eine Variable, den Parameter der Raumkurve. Eine wichtige Frage ist allerdings, ob die Definition des Linienintegrals (4.6) von der Wahl der Parametrisierung  $\vec{r}(t)$  abhängt. Wir betrachten dazu eine zweite Parametrisierung einer Kurve C,  $\vec{w}(t)$ .



Dann existiert eine differenzierbare Funktion t(u), die jedem u ein t zuordnet, so dass  $\vec{w}(u) = \vec{r}(t(u))$ . Jeder Schritt  $\Delta t > 0$  führt  $\vec{r}(t)$  weiter auf der Kurve C von  $q_0$  nach  $q_1$ , und ebenso jeder Schritt  $\Delta u$ . Daher ist dt/du > 0 entlang der Raumkurve und mit der Kettenregel  $\frac{d\vec{w}}{du} = \frac{d}{du}\vec{r}(t(u)) = \frac{d\vec{r}}{dt}\frac{dt}{du}$  erhalten wir

$$\int_{u_b}^{u_e} du \left\| \left| \frac{d\vec{w}}{du} \right| \right| f(\vec{w}(u)) = \int_{u_b}^{u_e} du \left\| \left| \frac{d\vec{r}}{dt} \right| \left| \frac{dt}{du} f(\vec{r}(t(u))) \right| = \int_{t_b}^{t_e} dt \left\| \left| \frac{d\vec{r}}{dt} \right| \right| f(\vec{r}(t))$$
(4.8)

Ein Wechsel der Parametrisierung läßt sich also als Variablensubstitution der Parametrisierungsvariablen interpretieren.

#### Das Linienintegral eines Vektorfeldes

Gesucht sei die Arbeit, die ein Kraftfeld  $\vec{F}(\vec{r})$  an einer Punktmasse verrichtet, die sich entlang einer Raumkurve C bewegt. Bei der Parametrisierung der Raumkurve können Sie wieder an eine Ameise denken, die im Zeitintervall  $\Delta t_i$  die Strecke  $\Delta \vec{r}_i \approx \Delta t_i \frac{d\vec{r}(t_i)}{dt}$  zurücklegt. Dabei verrichtet das Kraftfeld die Arbeit  $\approx \Delta \vec{r}_i \cdot \vec{F}(\vec{r}(t_i)) \approx \Delta t_i \frac{d\vec{r}(t_i)}{dt} \cdot \vec{F}(\vec{r}(t_i)).$ 



Wir definieren das Linienintegral eines Vektorfeldes

$$\int_{C} d\mathbf{r} \cdot \mathbf{F}(\mathbf{r}) \equiv \int_{t_{b}}^{t_{e}} dt \frac{d\mathbf{r}}{dt} \cdot \mathbf{F}(\mathbf{r}(t))$$

$$\overset{\text{kart. Basis}}{=} \int_{t_{b}}^{t_{e}} dt \frac{dx}{dt} F_{x}(\mathbf{r}(t)) + \int_{t_{b}}^{t_{e}} dt \frac{dy}{dt} F_{y}(\mathbf{r}(t)) + \int_{t_{b}}^{t_{e}} dt \frac{dz}{dt} F_{z}(\mathbf{r}(t))$$

$$(4.9)$$

Bei der Notation  $\int_C d\mathbf{r} \cdot \mathbf{F}(\mathbf{r})$  können sie sich  $d\mathbf{r}$  als kleinen Vektor entlang eines Teilstücks der Kurve C denken.

**Beispiel:** Wir betrachten ein Kraftfeld in 2 Dimensionen,  $\mathbf{F}(\mathbf{r}) = \begin{pmatrix} xy \\ -y^2 \end{pmatrix}$ . Ein Massepunkt bewegt sich entlang einer Parabel  $C : y = x^2$  vom Punkt (0,0) zu (1,1). Wir berechnen die am Massepunkt verrichtete Arbeit W mit der Parametrisierung  $x(t) = t, y(t) = t^2$  und damit  $\frac{d\mathbf{r}}{dt} = \begin{pmatrix} \frac{dx}{dt} \\ \frac{dy}{dt} \end{pmatrix} = \begin{pmatrix} 1 \\ 2t \end{pmatrix}$ . Wir erhalten  $W = \int_C d\mathbf{r} \cdot \mathbf{F}(\mathbf{r}) = \int_0^1 dt \frac{d\mathbf{r}}{dt} \cdot \mathbf{F}(\mathbf{r}(t)) = \int_0^1 dt \begin{pmatrix} 1 \\ 2t \end{pmatrix} \cdot \begin{pmatrix} t^3 \\ -t^4 \end{pmatrix}$   $= \int_0^1 dt(t^3 - 2t^5) = [\frac{1}{4}t^4 - \frac{2}{6}t^6]_0^1 = \frac{1}{4} - \frac{1}{3} = -\frac{1}{12}. \quad (4.10)$ 

Aus dem Spezialfall eines Feldes mit nur einer Komponente  $F_x(\vec{r}) = f(\vec{r})$  lässt sich das Linienintegral über eine ausgezeichnete Koordinate definieren

$$\int_{c} \mathrm{d}x f(\vec{r}) \equiv \int_{t_{b}}^{t_{e}} \mathrm{d}t \frac{\mathrm{d}x}{\mathrm{d}t} f(\vec{r}(t)) \ . \tag{4.11}$$

Hierbei werden bei der Zerlegung der Raumkurve C die Änderungen der x-Komponente entlang der Teilstücke aufsummiert.

Beispiel:

$$\int_{c} \mathrm{d}x = \int_{t_{b}}^{t_{e}} \mathrm{d}t \frac{\mathrm{d}x(t)}{\mathrm{d}t} = \int_{x_{b}}^{x_{e}} \mathrm{d}x = [x]_{x_{b}}^{x_{e}} = x_{e} - x_{b}$$
(4.12)

Im letzten Schritt wurde das Integral durch die Substitution x = x(t) ausgewertet.

Und schließlich lässt sich analog zum Linienintegral mit dem Skalarprodukt auch ein Integral mit Kreuzprodukt definieren.



$$\int_{c} \mathrm{d}\vec{r} \times \vec{F}(\vec{r}) \equiv \int_{t_{b}}^{t_{e}} \mathrm{d}t \frac{\mathrm{d}\vec{r}}{\mathrm{d}t} \times \vec{F}(\vec{r}(t)) \tag{4.13}$$

Dieses Integral tritt zum Beispiel bei der Berechnung des magnetischen Feldes eines stromdurchflossenen Drahtes auf.

## 4.3 Das Oberflächenintegral

Das Oberflächenintegral erlaubt Vektor- und Skalarfelder auf Oberflächen zu integrieren. Wir benötigen dazu das Konzept der Mehrfachintegrale aus Abschnitt 3.4. Ein Beispiel ist die Berechnung der Gesamtladung auf einer gekrümmten Oberfläche mit gegebener Dichte pro kleiner Fläche. Weitere wichtige Anwendungen in der Elektrostatik sind Oberflächenintegrale des elektrischen Feldes im Zusammenhang mit dem Gaußschen Integralsatz, den wir in Sektion 4.8 diskutieren werden.

Wir zerlegen wieder eine Oberfläche in kleine Teilflächen, über die wir dann summieren werden. Bei einer einzelnen kleinen Teilfläche nutzen wir, dass ihr Beitrag zum Integral lediglich von ihrem Flächeninhalt und von ihrer Orientierung im Raum abhängt: ist die Teilfläche hinreichend klein, dann ist der Integrand (z.B. eine Dichte oder ein Vektorfeld) konstant über die Teilfläche. Die genaue Form der Teilfläche (dreieckig, rund, elliptisch,  $\ldots$ ) ist dann irrelevant<sup>1</sup>. Zur Charakterisierung des Flächeninhalts und der Orientierung einer Teilfläche  $\Delta S$  definieren wir einen Vektor  $\Delta \vec{S}$ , der senkrecht zur Teilfläche  $\Delta S$  steht, und dessen Länge  $\|\Delta \vec{S}\|$  gleich ihrem Flächeninhalt ist. Dieser Vektor wird als vektorielles Oberflächelement oder Oberflächelement bezeichnet. Die Orientierung von  $\Delta \vec{S}$  wird von Fall zu Fall unter-



schiedlich gewählt. Bei geschlossenen Oberflächen wählt man die Richtung von  $\Delta \vec{S}$  meist so, dass die Oberflächenelemente nach aussen zeigen.

Die Summe über Nsolcher kleinen Flächenelemente führt uns im Grenzfall  $N \to \infty$ auf die Flächenintegrale

$$\int_{S} \mathrm{d}\vec{S} \cdot \vec{F}(\vec{r}) \equiv \lim_{N \to \infty} \sum_{\text{Teilflächen } i} \Delta \vec{S}_{i} \cdot \vec{F}(\vec{r}_{i}) \qquad \text{eines Vektorfeldes } \vec{F}(\vec{r}) \tag{4.14}$$

$$\int_{S} \mathrm{d}Sf(\vec{r}) \equiv \lim_{N \to \infty} \sum_{\text{Teilflächen } i} ||\Delta \vec{S}_{i}|| f(\vec{r}_{i}) \qquad \text{eines skalaren Feldes } f(\vec{r}) \qquad (4.15)$$

Zur konkreten Definition und Berechnung solcher Integrale benötigen wir wieder eine Parametrisierung. Zweidimensionale Oberflächen benötigen zwei Variablen (u, v) um einen

<sup>&</sup>lt;sup>1</sup>Vergleichen Sie mit dem Linienintegral, bei dem wir ein Teilstück zwischen  $p_i$  und  $p_{i+1}$  näherungsweise durch eine Verschiebung  $p_{i+1} = p_i + \Delta \vec{s_i}$  beschrieben haben.

Punkt  $p = p_0 + \vec{r}$  auf der Oberfläche eindeutig festzulegen. Eine **Parametrisierung** einer Oberfläche *S* (als Punktmenge) ist eine eineindeutige Abbildung von  $D \subset \mathbb{R}^2$  auf *X*,  $(u, v) \in D \mapsto \vec{r}(u, v) : p_0 + \vec{r}(u, v) \in S$ , die stetig differenzierbar ist und deren Inverses ebenso stetig differenzierbar ist.



Wir zerlegen nun D in kleine rechteckige Teilstücke mit Kantenlängen  $\Delta u_i, \Delta v_i$ , wie bei der Definition des Mehrfachintegrals in 3.4. Diese Zerlegung von D definiert auch eine Zerlegung der Oberfläche S in kleine Teilstücke. Näherungsweise (kleine  $\Delta u_i, \Delta v_i$ ) handelt es sich dabei um Parallelogramme mit Kanten  $\frac{\partial \vec{r}}{\partial u} \Delta u$  und  $\frac{\partial \vec{r}}{\partial v} \Delta v$ .  $u + \Delta u, v$ 

es sich dabei um Parallelogramme mit Kanten  $\frac{\partial \vec{r}}{\partial u} \Delta u$  und  $\frac{\partial \vec{r}}{\partial v} \Delta v$ . Zur Berechnung des Vektors des entsprechenden Flächenelementes  $\Delta \vec{S}$  nutzen wir das Kreuzprodukt

$$\Delta \vec{S} = \frac{\partial \vec{r}}{\partial u} \Delta u \times \frac{\partial \vec{r}}{\partial v} \Delta v$$
  
=  $\left(\frac{\partial \vec{r}}{\partial u} \times \frac{\partial \vec{r}}{\partial v}\right) \Delta u \quad \Delta v$  (4.16)  
 $\vec{r}(u,v)$  Punkt auf der Oberfläche

 $\vec{r}(u+\Delta u,v)-\vec{r}(u,v)$ 

 $u + \Delta u, v + \Delta v$ 

Gegeben eine Parametrisierung  $\vec{r}(u,v)$  erhalten wir damit für die Oberflächen<br/>integrale (4.14)

$$\int_{S} d\vec{S} \cdot \vec{F}(\vec{r}) \equiv \int_{D} du dv \left( \frac{\partial \vec{r}}{\partial u} \times \frac{\partial \vec{r}}{\partial v} \right) \cdot \vec{F}(\vec{r}(u,v))$$
$$\int_{S} dS f(\vec{r}) \equiv \int_{D} du dv \left\| \frac{\partial \vec{r}}{\partial u} \times \frac{\partial \vec{r}}{\partial v} \right\| f(\vec{r}(u,v))$$
(4.17)

**Beispiel:** Wir betrachten das Integral der Funktion  $f(\vec{r}) = 1$  über eine ebene quadratische Oberfläche in der z-Ebene.  $S = \{(x, y, z) \in \mathbb{R}^3 | 0 \le x \le 1, 0 \le y \le 1, z = 0\}$ . Die einfachste Parametrisierung dieser Oberfläche ist x(u, v) = u, y(u, v) = v, z(u, v) = 0, also  $\mathbf{r}(u, v) = \begin{pmatrix} u \\ v \\ 0 \end{pmatrix}$ .  $\frac{\partial \mathbf{r}}{\partial u} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \qquad \frac{\partial \mathbf{r}}{\partial v} = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 1 \end{pmatrix} \qquad \frac{\partial \mathbf{r}}{\partial u} \times \frac{\partial \mathbf{r}}{\partial v} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \qquad \frac{\partial \mathbf{r}}{\partial v} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \qquad \frac{\partial \mathbf{r}}{\partial v} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \qquad \frac{\partial \mathbf{r}}{\partial v} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \qquad \frac{\partial \mathbf{r}}{\partial v} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \qquad \frac{\partial \mathbf{r}}{\partial v} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \qquad \frac{\partial \mathbf{r}}{\partial v} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \qquad \frac{\partial \mathbf{r}}{\partial v} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \qquad \frac{\partial \mathbf{r}}{\partial v} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \qquad \frac{\partial \mathbf{r}}{\partial v} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \qquad \frac{\partial \mathbf{r}}{\partial v} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \qquad \frac{\partial \mathbf{r}}{\partial v} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \qquad \frac{\partial \mathbf{r}}{\partial v} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \qquad \frac{\partial \mathbf{r}}{\partial v} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \qquad \frac{\partial \mathbf{r}}{\partial v} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \qquad \frac{\partial \mathbf{r}}{\partial v} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \qquad \frac{\partial \mathbf{r}}{\partial v} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \qquad \frac{\partial \mathbf{r}}{\partial v} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \qquad \frac{\partial \mathbf{r}}{\partial v} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \qquad \frac{\partial \mathbf{r}}{\partial v} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \qquad \frac{\partial \mathbf{r}}{\partial v} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \qquad \frac{\partial \mathbf{r}}{\partial v} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \qquad \frac{\partial \mathbf{r}}{\partial v} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \qquad \frac{\partial \mathbf{r}}{\partial v} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \qquad \frac{\partial \mathbf{r}}{\partial v} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \qquad \frac{\partial \mathbf{r}}{\partial v} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \qquad \frac{\partial \mathbf{r}}{\partial v} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \qquad \frac{\partial \mathbf{r}}{\partial v} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \qquad \frac{\partial \mathbf{r}}{\partial v} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \qquad \frac{\partial \mathbf{r}}{\partial v} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \qquad \frac{\partial \mathbf{r}}{\partial v} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \qquad \frac{\partial \mathbf{r}}{\partial v} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \qquad \frac{\partial \mathbf{r}}{\partial v} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \qquad \frac{\partial \mathbf{r}}{\partial v} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \qquad \frac{\partial \mathbf{r}}{\partial v} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \qquad \frac{\partial \mathbf{r}}{\partial v} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \qquad \frac{\partial \mathbf{r}}{\partial v} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \qquad \frac{\partial \mathbf{r}}{\partial v} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \qquad \frac{\partial \mathbf{r}}{\partial v} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \qquad \frac{\partial \mathbf{r}}{\partial v} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \qquad \frac{\partial \mathbf{r}}{\partial v} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \qquad \frac{\partial \mathbf{r}}{\partial v} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \qquad \frac{\partial \mathbf{r}}{\partial v} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \qquad \frac{\partial \mathbf{r}}{\partial v} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \qquad \frac{\partial \mathbf{r}}{\partial v} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \qquad \frac{\partial \mathbf{r}}{\partial v} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \qquad \frac{\partial \mathbf{r}}{\partial v} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \qquad \frac{\partial \mathbf{r}}{\partial v} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \qquad \frac{\partial \mathbf{r}}{\partial v} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \qquad \frac{\partial \mathbf{r}}{\partial v} = \begin{pmatrix} 0 \\ 0 \\ 0$ 

**Beispiel:** Ein weniger einfaches Beispiel ist die Oberfläche einer Kugel mit Radius R. Zur Parametrisierung nutzen wir die **sphärischen Koordinaten** oder **Kugelkoodinaten** in 3D. Die beiden Winkelparameter  $\theta, \phi$  (statt u, v) sind so definiert, dass der Ortsvektor  $\vec{r}$  mit der z-Achse den **Polarwinkel**  $\theta$  einschließt, und seine Projektion auf die xy-Ebene mit der x-Achse den **Azimutwinkel**  $\phi$  einschließt.

$$\mathbf{r}(\theta,\phi) = \begin{pmatrix} x(\theta,\phi) \\ y(\theta,\phi) \\ z(\theta,\phi) \end{pmatrix} = \begin{pmatrix} R\sin\theta\cos\phi \\ R\sin\theta\sin\phi \\ R\cos\theta \end{pmatrix}$$
(4.18)

Die Werte des Polarwinkels  $\theta$  liegen zwischen 0 (Nordpol) und  $\pi$  (Südpol), die des Azimutwinkels  $\phi$  zwischen 0 und  $2\pi$ . Die Ableitung des Ortsvektors nach diesen Parametern ergibt

$$\frac{\partial \mathbf{r}}{\partial \theta} = R \begin{pmatrix} \cos \theta \cos \phi \\ \cos \theta \sin \phi \\ -\sin \theta \end{pmatrix} \qquad \frac{\partial \mathbf{r}}{\partial \phi} = R \begin{pmatrix} -\sin \theta \sin \phi \\ \sin \theta \cos \phi \\ 0 \end{pmatrix}$$
(4.19)

und damit das Kreuzprodukt

$$\frac{\partial \mathbf{r}}{\partial \theta} \times \frac{\partial \mathbf{r}}{\partial \phi} = \begin{pmatrix} R^2 \sin^2 \theta \cos \phi \\ R^2 \sin^2 \theta \sin \phi \\ R^2 \cos \theta \sin \theta (\cos^2 \phi + \sin^2 \phi) \end{pmatrix} = R^2 \sin \theta \begin{pmatrix} \sin \theta \cos \phi \\ \sin \theta \sin \phi \\ \cos \theta \end{pmatrix} .$$
(4.20)



## 4.4 Das Volumenintegral

Das Volumenintegral ermöglicht ein Skalarfeld  $\phi(\mathbf{r})$  oder ein Vektorfeld  $\mathbf{F}(\mathbf{r})$  über ein Volumin zu integrieren. Ein Beispiel ist das Integral einer Ladungsdichte über ein Volumen, oder der Schwerpunkt eines massiven Körpers. Zur Definition des Volumenintegrals gehen vor wie bei der Definition des Linien- und Oberflächenintegrals und unterteilen das Volumen in eine grosse Zahl kleiner Volumenelemente. Innerhalb eines Volumenelements ist der Integrand  $\phi(\mathbf{r})$  oder  $\mathbf{F}(\mathbf{r})$  näherungsweise konstant, so dass wir von jedem Volumenelement das Volumen  $\Delta V$  mit dem Integrand multiplizieren, und diese Beiträge aller Volumenelemente aufsummieren. Im Grenzfall bei dem die Dimensionen aller Volumenelemente gegen null gehen erhalten wir das Volumenintegral.

Die Parametrisierung eines Volumens erfordert drei Variablen (u, v, w), ein Raumpunkt ist also durch eine Funktion  $\vec{r}(u, v, w)$  festgelegt mit  $(u, v, w) \in D \subset \mathbb{R}^3$ . Die Kurven  $\vec{r}(u, v, w)$  bei denen jeweils alle Parameter bis auf einen festgehalten werden kann man sich (näherungsweise) als Netz aus lauter kleinen Parallelepipeden visualisieren. Diese Parallelepipede nutzen wir als Volumenelemente.

Das Volumen eines kleinen Parallelepipeds, das durch die Vektoren  $\frac{\partial \vec{r}}{\partial u}\Delta u$ ,  $\frac{\partial \vec{r}}{\partial v}\Delta v$  und

 $\frac{\partial \vec{r}}{\partial w} \Delta w$  aufgespannt wird ist

$$\Delta V = \left| \frac{\partial \mathbf{r}}{\partial u} \cdot \left( \frac{\partial \mathbf{r}}{\partial v} \times \frac{\partial \mathbf{r}}{\partial w} \right) \right| \Delta u \ \Delta v \ \Delta w \ . \tag{4.22}$$

Das Spatprodukt hat eine Orientierung (ist positiv für einen rechtshändigen Spat, negativ für einen linkshändigen, s. 1.6); daher ist das Volumen des von 3 Vektoren aufgespannten Parallelepipeds der Betrag des Spatproduktes.



 $\Delta V$  gibt das Volumen eines kleinen Volumen<br/>elementes an

Führen wir nun die Zahl der Volumenelemente gegen unendlich und die  $\Delta u, \Delta v, \Delta w$  für jedes einzelne Element gegen null, erhalten wir als Definition des Volumenintegrals

$$\int_{V} \mathrm{d}V\phi(\mathbf{r}) \equiv \int_{D} \mathrm{d}u \,\mathrm{d}v \,\mathrm{d}w \,\left| \frac{\partial \mathbf{r}}{\partial u} \cdot \left( \frac{\partial \mathbf{r}}{\partial v} \times \frac{\partial \mathbf{r}}{\partial w} \right) \right| \,\phi(\mathbf{r}(u,v,w)) \tag{4.23}$$

$$\int_{V} dV \mathbf{F}(\mathbf{r}) \equiv \int_{D} du \, dv \, dw \, \left| \frac{\partial \mathbf{r}}{\partial u} \cdot \left( \frac{\partial \mathbf{r}}{\partial v} \times \frac{\partial \mathbf{r}}{\partial w} \right) \right| \, \mathbf{F}(\mathbf{r}(u, v, w)) \tag{4.24}$$

Für kartesische Koordinaten ist das Spatprodukt eins,  $\int_V dV \phi(\mathbf{r}) = \int_V dx dy dz \phi(x, y, z)$ . Eine Änderung der Parametrisierung, z.B. von kartesischen Koordinaten zu Kugelkoordinaten, kann auch wieder als Variablensubstitution der Parametrisierungsvariablen verstanden werden, vgl. Transformationssatz in Abschnitt 3.4.1. **Beispiel:** Als Beispiel berechnen wir noch das Volumen einer Kugel mit Radius R. Zur Parametrisierung bieten sich wieder Kugelkoordinaten an

$$\mathbf{r}(r,\theta,\phi) = \begin{pmatrix} x(r,\theta,\phi) \\ y(r,\theta,\phi) \\ z(r,\theta,\phi) \end{pmatrix} = \begin{pmatrix} r\sin\theta\cos\phi \\ r\sin\theta\sin\phi \\ r\cos\theta \end{pmatrix}$$
(4.25)

mit einer **Radialkoordinate**  $r = \sqrt{x^2 + y^2 + z^2}$ , die den Abstand eines Punktes vom Ursprung im Kugelmittelpunkt angibt. Durch Ableitung nach der Radialkoordinate erhalten wir  $\frac{\partial \mathbf{r}}{\partial x} = \begin{pmatrix} \sin \theta \cos \phi \\ \sin \theta \sin \phi \end{pmatrix}$ , und zusammen mit dem Ergebnis aus dem Beispiel zum

halten wir  $\frac{\partial \mathbf{r}}{\partial r} = \left( \sin \theta \sin \phi \right)$ , und zusammen mit dem Ergebnis aus dem Beispiel zum

Flächenintegral

$$\frac{\partial \mathbf{r}}{\partial \theta} \times \frac{\partial \mathbf{r}}{\partial \phi} = r^2 \sin \theta \begin{pmatrix} \sin \theta \cos \phi \\ \sin \theta \sin \phi \\ \cos \theta \end{pmatrix}$$
(4.26)

erhalten wir das Spatprodukt

$$\frac{\partial \mathbf{r}}{\partial r} \cdot \left(\frac{\partial \mathbf{r}}{\partial \theta} \times \frac{\partial \mathbf{r}}{\partial \phi}\right) = r^2 \sin \theta \ . \tag{4.27}$$

Das Volumenelement in Kugelkoordinaten ist also  $r^2 \sin \theta \Delta r \Delta \theta \Delta \phi$ . Auch dieses Ergebnis lässt sich aus geometrischen Überlegungen herleiten; die Kanten des kleinen Quaders, der durch Koordinatenänderungen  $\Delta r, \Delta \theta$  und  $\Delta \phi$  generiert wird sind  $\Delta r, r\Delta \theta$  und  $r \sin \theta \Delta \phi$ . Das Volumen der Kugel ist dann

$$\int_{V} dV = \int_{D} dr \, d\theta \, d\phi \left| \frac{\partial \mathbf{r}}{\partial r} \cdot \left( \frac{\partial \mathbf{r}}{\partial \theta} \times \frac{\partial \mathbf{r}}{\partial \phi} \right) \right|$$
$$= \int_{D} dr \, d\theta \, d\phi r^{2} \sin \theta = \int_{0}^{R} dr r^{2} \int_{0}^{\pi} d\theta \sin \theta \int_{0}^{2\pi} d\phi = 4/3\pi R^{3} . (4.28)$$

## 4.5 Differentialoperatoren

Wie verändert sich  $\mathbf{v}$  eines Vektorfeldes  $\mathbf{v}(\mathbf{r})$  unter einer (kleinen) Änderung von  $\mathbf{r}$ ? Da Vektorfelder in drei Dimensionen durch drei Funktionen die von drei Veränderlichen abhängen beschrieben werden, ist das Konzept der Ableitung von Vektorfeldern mathematisch reicher als das der Ableitung einer einzelnen Funktion einer Variablen.

Wir werden drei sogenannte Differentialoperatoren auf Feldern diskutieren, die jeweils unterschiedliche Aspekte von Feldern charakterisieren; den Gradienten eines skalaren Feldes (ergibt ein Vektorfeld), die Rotation eines Vektorfeldes (ergibt ein Vektorfeld), und die Divergenz eines Vektorfeldes (ergibt ein skalares Feld).

## 4.5.1 Der Gradient

Für eine Funktion einer Veränderlichen hatten wir in Abschnitt 3.2.1 die Ableitung als lineare Näherung einer Funktion f(x) an einem Punkt verstanden,  $f(x + \Delta x) = f(x) + \frac{df}{dx}\Delta x + \mathcal{O}(Deltax^2)$ . Der Gradient verallgemeinert die Ableitung und gibt an, wie sich ein skalares Feld  $f(\mathbf{r})$  unter einer (kleinen) Änderung  $\Delta \mathbf{r}$  ändert.

Sei  $f(x_1, x_2, x_3, ...)$  eine Funktion  $\mathbb{R}^n \to \mathbb{R}, \mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ ... \end{pmatrix} \mapsto f(\mathbf{x}), \text{ und } x_1, x_2, ..., x_n \text{ kartesische Koordinaten des } \mathbb{R}^n$  des Punktes  $p_0 + \mathbf{x}$ , dann heißt der Spaltenvektor

$$\nabla f = \begin{pmatrix} \frac{\partial f}{\partial x_1} \\ \frac{\partial f}{\partial x_2} \\ \frac{\partial f}{\partial x_3} \\ \vdots \end{pmatrix}$$
(4.29)

der Gradient der Funktion  $f(x_1, x_2, x_3, ...) \equiv f(\mathbf{r})$  am Punkt  $p = p_0 + \mathbf{x}$ .

Betrachten wir die Funktion am Punkt p und an einem zweiten Punkt  $p + \Delta \mathbf{x}$ . Das Skalarprodukt von  $\nabla f$  mit einem Verschiebungsvektor  $\Delta \mathbf{x}$  gibt die Änderung von  $f(\mathbf{x})$  zu  $f(\mathbf{x} + \Delta \mathbf{x})$  in linearer Näherung

$$\Delta f \equiv f(\mathbf{x} + \Delta \mathbf{x}) - f(\mathbf{x}) = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i} \Delta x_i + \mathcal{O}(||\Delta \mathbf{x}||^2) = \nabla f \cdot \Delta \mathbf{x} + \mathcal{O}(||\Delta \mathbf{x}||^2) .$$
(4.30)

Der Gradient einer Funktion mehrerer Veränderlicher  $f(x_1, x_2, ...)$ , bzw.  $f(\mathbf{x})$  kann also als Verallgemeinerung der Ableitung einer Funktion einer Variablen betrachtet werden. Zur Visualisierung des Gradienten stellen wir eine Funktion f(p) durch Konturlinien dar, die wie Höhenlinien einer Landkarte Punkte mit gleichem Wert der Funktion f verbinden.



Dann steht der Vektor  $\nabla f$  senkrecht auf den Konturlinien, denn für ein  $\Delta \mathbf{x}$  entlang einer Kontour ist  $\nabla f \cdot \Delta \mathbf{x} = 0$ . Betrachten wir alle  $\Delta \mathbf{x}$  mit konstanter (kleiner) Norm. Da das Skalarprodukt zweier Vektoren konstanter Länge dann maximal ist, wenn beide Vektoren in dieselbe Richtung zeigen, zeigt  $\nabla f$  in die Richtung maximaler Veränderung von f (in linearer Näherung, also für kleine  $||\Delta \mathbf{x}||$ ).

Der Gradient zeigt damit in die Richtung der größten Veränderung von f.

Bei

spiel:  

$$f(x, y, z) = \sqrt{x^2 + y^2 + z^2}$$

$$\nabla f = \begin{pmatrix} \partial f / \partial x \\ \partial f / \partial y \\ \partial f / \partial z \end{pmatrix} = \frac{1}{2\sqrt{x^2 + y^2 + z^2}} \begin{pmatrix} 2x \\ 2y \\ 2z \end{pmatrix} = \frac{\vec{r}}{\|\vec{r}\|}$$

Hier bilden die Punkte mit gleichem Wert der Funktion f kugelförmige Schalen, entsprechend zeigt  $\nabla f$  radial nach außen. In zwei Dimensionen ist auch noch eine alternative Visualisierung möglich: Betrachtet man  $f(x, y) = \sqrt{x^2 + y^2}$  als Höhenprofil über der xy-Ebene erhält man einen Trichter mit Mittelpunkt beim Ursprung.

Optimierung unter Nebenbedingungen



Gesucht sei das (lokale) Maximum einer Funktion  $f : \mathbb{R}^n \to \mathbb{R}, \mathbf{x} \mapsto f(\mathbf{x})$  unter einer Nebenbedingung  $g(\mathbf{x}) = 0$ . An einem (lokalen) Maximum von  $f(\mathbf{x})$  unter der Randbedingung ist  $\nabla f$  orthogonal zur Oberfläche  $g(\mathbf{x}) = 0$  (sonst könnte man mit einem kleinen Schritt entlang der Oberfläche den Wert von f noch vergrößern). Damit ist  $\nabla f$  parallel zu  $\nabla g$ , d.h.  $\nabla f = \mu \nabla g$  oder  $\nabla (f - \mu g) = 0$ . Zusammen mit der Bedingung  $g(\mathbf{x}) = 0$  lässt sich diese Gleichung nach  $\mathbf{x}$  und  $\mu$  lösen.  $\mu$  wird als Lagrange-Multiplikator bezeichnet.

#### Gradientenfelder

Ein Vektorfeld  $\vec{v}(\vec{r})$  heißt **Gradientenfeld**, wenn eine Funktion  $f(\vec{r})$  existient, so dass

$$\vec{v}(\vec{r}) = \nabla f(\vec{r}) \tag{4.31}$$

Ein allgemeines Vektorfeld aus *d*-dimensionalen Vektoren wird durch *d* Funktionen beschrieben, ein Gradientenfeld ist also ein sehr spezielles Vektorfeld, denn es ist durch die eine Funktion  $f(\vec{r})$  vollständig beschrieben. Jedes Gradientenfeld weist eine Beziehung zwischen den partiellen Ableitungen seiner Komponenten auf: Sei  $\mathbf{v}(\mathbf{x}) = \nabla f$  ein Gradientenfeld. Aus der Schwarzschen Regel für partielle Ableitungen  $\frac{\partial^2 f}{\partial x_i \partial x_j} = \frac{\partial^2 f}{\partial x_j \partial x_i}$  gilt für die Komponenten  $v_i = \frac{\partial f}{\partial x_i}$  und  $v_j = \frac{\partial f}{\partial x_j}$  auch  $\frac{\partial v_j}{\partial x_i} = \frac{\partial v_i}{\partial x_j}$ . Ein Gradientenfeld im  $\mathbb{R}^3$  hat damit

$$\frac{\partial v_z}{\partial y} - \frac{\partial v_y}{\partial z} = 0$$

$$\frac{\partial v_x}{\partial z} - \frac{\partial v_z}{\partial x} = 0$$

$$\frac{\partial v_y}{\partial x} - \frac{\partial v_x}{\partial y} = 0 .$$
(4.32)

In einfach zusammenhängenden Räumen (grob ausgedrückt: ohne Löcher) ist diese Bedingung auch hinreichend für ein Gradientenfeld.

Info: Sei  $\mathbf{F}(\mathbf{x})$  ein Gradientenfeld, d.h.  $\mathbf{F}(\vec{r}) = \nabla f(\mathbf{x})$ 

$$\begin{split} \int_{c} d\mathbf{x} \cdot \vec{F}(\mathbf{x}) &= \int_{t_{b}}^{t_{e}} \mathrm{d}t \frac{d\mathbf{x}}{dt} \cdot \mathbf{F}(\mathbf{x}(t)) \qquad \frac{d}{dt} (f(\mathbf{x}(t))) = \nabla f \cdot \frac{d\mathbf{x}}{dt} \\ &= \int_{t_{b}}^{t_{e}} \mathrm{d}t \frac{d}{dt} (f(\mathbf{x}(t))) = [f(\mathbf{x}(t))]_{t_{b}}^{t_{e}} \end{split}$$

Damit ist das Linienintegral eines Gradientenfeldes wegunabhängig, es hängt nur von Anfangs- und Endpunkt ab. Das Gravitationsfeld ist ein Gradientenfeld, ebenso das elektrische Feld. Die von diesen Feldern verrichtete Arbeit um eine Punktmasse von Punkt  $q_0$  zu  $q_q$  zu bringen ist also unabhängig vom konkret gewählten Weg. Die skalare Funktion f wird in diesem Zusammenhang als Potential bezeichnet.

Ein Beispiel ist das elektrische Feld einer Punktladung am Ursprung,  $\mathbf{E}(\mathbf{r}) = \frac{q}{r^2} \frac{\mathbf{r}}{r} = \frac{q}{r^3} \mathbf{r}$ wobei  $r = ||\mathbf{r}||$ . Das skalare Potential  $\phi(\mathbf{r})$  ist (minus eins) mal die an einer zweiten Punktladung pro Ladungseinheit verrichtete Arbeit wenn man diese Ladung von einem unendlich weit entfernten Punkt zum Punkt  $\mathbf{r}$  bringt

$$\phi(\vec{r}) = -\int_{\infty \text{ nach } \mathbf{r}} \mathrm{d}\mathbf{r} E(\mathbf{r}) = q \int_{\mathbf{r} \text{ nach } \infty} \frac{\mathrm{d}\mathbf{r} \cdot \mathbf{r}}{r^3} .$$
(4.33)

Das dieses Integral unabhängig vom Weges ist, entlang dessen es ausgewertet wird, wählen wir als Integrationsweg einfach eine radiale Linie parametrisiert durch eine Parametrisierungsvariable t = r. Mit  $\frac{d}{dt}(\mathbf{r} \cdot \mathbf{r}) = 2\mathbf{r} \cdot \frac{d\mathbf{r}}{dt}$  und  $\frac{d}{dt}(\mathbf{r} \cdot \mathbf{r}) = \frac{d(r^2)}{dt} = 2r\frac{dr}{dt}$  erhalten wir

$$\phi(\mathbf{r}) = q \int_{r}^{\infty} dr \frac{r}{r^{3}} = q \int_{r}^{\infty} dr \frac{1}{r^{2}} = -\left[\frac{q}{r}\right]_{r}^{\infty} = q/r \tag{4.34}$$

## 4.6 Die Rotation (engl. "curl")

In 3D und kartesischen Koordinaten lässt sich für ein Vektorfeld  $\mathbf{v}(\mathbf{x})$  ein Feld von Spaltenvektoren definieren, die sogenannte **Rotation** von  $\mathbf{v}(\mathbf{x})$ , mit Komponenten

$$\nabla \times \mathbf{v} = \begin{pmatrix} \frac{\partial v_z}{\partial y} - \frac{\partial v_y}{\partial z} \\ \frac{\partial v_x}{\partial z} - \frac{\partial v_z}{\partial x} \\ \frac{\partial v_y}{\partial x} - \frac{\partial v_x}{\partial y} \end{pmatrix}$$
(4.35)

 $\nabla \times$  soll an das Kreuzprodukt  $\begin{pmatrix} \partial/\partial x \\ \partial/\partial y \\ \partial/\partial z \end{pmatrix} \times \begin{pmatrix} v_x \\ v_y \\ v_z \end{pmatrix}$  erinnern.
Aus (4.32) hat ein Gradientenfeld damit die Rotation Null,  $\nabla \times \nabla f = 0 \ \forall f(\mathbf{x})$ , oder kurz  $\nabla \times \nabla = 0$ .



Auch die Rotation eines Vektorfeldes lässt sich visualisieren.  $\mathbf{v}(\mathbf{x})$  sei das Geschwindigkeitsfeld einer strömenden Flüssigkeit.  $\nabla \times \mathbf{v}(\mathbf{x})$  beschreibt, wie ein kleiner Ball mit festem Mittelpunkt  $\mathbf{x}$  rotiert.  $\|\nabla \times \mathbf{v}\|$  ist proportional zur Winkelgeschwindigkeit, die Richtung von  $\nabla \times \mathbf{v}$  gibt die Drehachse an.



### 4.7 Die Divergenz

ist eine Abbildung von einem Vektorfeld  $\mathbf{v}(\mathbf{x})$  auf ein skalares Feld, definiert in kartesischen Koordinaten als

$$\nabla \cdot \mathbf{v} = \frac{\partial v_x}{\partial x} + \frac{\partial v_y}{\partial y} + \frac{\partial v_z}{\partial z}$$
(4.36)

(und analog in *n* Dimensionen). Die Notation  $\nabla \cdot$  soll an das Skalarprodukt erinnern:

$$\nabla \cdot \mathbf{v} = \begin{pmatrix} \partial/\partial x \\ \partial/\partial y \\ \partial/\partial z \end{pmatrix} \cdot \begin{pmatrix} v_x \\ v_y \\ v_z \end{pmatrix}$$
(4.37)



Auch die Divergenz hat eine intuitive Interpretation:  $\mathbf{v}(\mathbf{x})$  beschreibe die Bewegung einer Flüssigkeit oder eines Gases und sei gleich dem Geschwindigkeitsfeld. Zahl der Teilchen, die eine kleine Fläche pro Zeit und Flächeninhalt durchströmt ist dann  $\mathbf{v}(\mathbf{r}) \cdot \mathbf{S}$  (wobei  $\mathbf{S}$  ein Einheitsvektor senkrecht zur Oberfläche ist).  $\nabla \cdot \mathbf{v}$  gibt an, wieviele Teilchen netto aus einem kleinen Volumen hinausfließen (pro Volumen). Dieses Bild werden wir beim Gaußschen Integralsatz weiter ausbauen.

#### 4.8 Integraltheoreme

Der Fundamentalsatz der Analysis

$$\int_{a}^{b} \mathrm{d}x \,\left(\frac{d}{dx}F(x)\right) = F(b) - F(a) \tag{4.38}$$

verbindet die Werte von F(x) an den Rändern des Intervalls [a, b] mit dem Integral von  $\frac{dF}{dx}$  über das Intervall [a, b] (s. 3.3.1).

In höheren Dimensionen existieren Verallgemeinerungen dieses Satzes, die z.B. Integrale über ein Volumen einer Ableitung mit einem Integral über den Rand dieses Volumens (eine Oberfläche) verbinden. Wichtige Anwendungen dieser Integralsätze liegen in der Elektround Magnetostatik.

# 4.8.1 Der Gausssche Integralsatz

Wir betrachten ein Vektorfeld in 3 Dimensionen in kartesischer Basis und kartesischen Koordinaten

$$\mathbf{v}(\mathbf{r}) = \begin{pmatrix} v_x(x, y, z) \\ v_y(x, y, z) \\ v_z(x, y, z) \end{pmatrix}$$
(4.39)

 $\mathbf{v}(\mathbf{r})$  könnte zum Beispiel ein Strömungsfeld sein, wobei  $\mathbf{v} \cdot \Delta \mathbf{S}$  angibt, wie viele Teilchen pro Zeit durch ein kleines Flächenelement  $\Delta \mathbf{S}$  strömen.

Wir betrachten nun einen (kleinen) Quader mit Kantenlänge  $\Delta x, \Delta y, \Delta z$ .



Der Netto-Fluss in x-Richtung aus dem Quader heraus ist

$$(v_x|_2 - v_x|_1)\Delta y \ \Delta z \simeq \frac{\partial v_x}{\partial x}\Delta x \ \Delta y \ \Delta z$$

In y- und z-Richtung ist der Netto-Fluss:

$$\begin{array}{l} \frac{\partial v_y}{\partial y} \ \Delta x \ \Delta y \ \Delta z \end{array} \qquad \qquad \text{durch die hintere und vordere Wand} \\ \frac{\partial v_z}{\partial z} \ \Delta x \ \Delta y \ \Delta z \end{array} \qquad \qquad \qquad \text{durch obere und untere Wand} \end{array}$$

$$\left(\frac{\partial v_x}{\partial x} + \frac{\partial v_y}{\partial y} + \frac{\partial v_z}{\partial z}\right) \Delta x \ \Delta y \ \Delta z = \nabla \cdot \mathbf{v} \ \Delta x \ \Delta y \ \Delta z \tag{4.40}$$

ist der gesamte Netto-Fluss aus dem Quader heraus. Dieser Netto-Fluss lässt sich auch als Oberflächenintegral schreiben.



Der Netto Fluss in *x*-Richtung aus dem Quader hinaus ist also  $\Delta \vec{S}_1 \cdot \vec{v}|_1 + \Delta \vec{S}_2 \cdot \vec{v}|_2$ . Summiert man die Beiträge von den 6 Seiten des kleinen Quaders erhält man den Netto Fluss aus dem Quader hinaus  $\int d\mathbf{S} \cdot \mathbf{v}$  und damit

$$\nabla \cdot \mathbf{v} \,\Delta x \,\Delta y \,\Delta z = \int \mathrm{d}\mathbf{S} \cdot \mathbf{v}(\mathbf{r}) \qquad \text{für kleine Quader}$$
(4.41)

Für ein Volumen V, das sich in viele (kleine) Quader zerlegen lässt summieren wir dieses Ergebnis für einen Quader über die vielen Teilquader. Im Inneren des Volumens heben sich die Oberflächenintegrale auf der rechten Seite gegenseitig weg und so erhalten wir den **Gaussschen Integralsatz** 

$$\int_{V} \mathrm{d}V \nabla \cdot \mathbf{v}(\mathbf{r}) = \int_{\partial V} \mathrm{d}\mathbf{S} \cdot \mathbf{v}$$
(4.42)



**Beispiel:** Der Gausssche Integralsatz findet wichtige Anwendungen in der Elektrostatik. Für eine Ladungsverteilung mit radialer Symmetrie (Ladungsdichte hängt also nur vom Abstand r vom Ursprung ab) ist  $\rho(\mathbf{r}) = \rho(r)$ .

Die Maxwellgleichung  $\nabla \cdot \mathbf{E} = \rho(r)/\epsilon_0$  eingesetzt in das Volumenintegral einer Kugel über die Ladungsdichte (Ladung im Inneren der Kugel) ergibt

$$Q/\epsilon_0 = \int_V \mathrm{d} V \; \rho(r)/\epsilon_0 \stackrel{\text{Maxwell}}{=} \int_V dV \; \nabla \cdot \mathbf{E} = \int_{\partial V} \mathrm{d} \mathbf{S} \cdot \mathbf{E} = ||\mathbf{E}|| \int_{\partial V} \mathrm{d} S = ||\mathbf{E}|| \; 4\pi r^2 \; .$$

Der vorletzte Schritt folgt daher, dass das elektrische Feld einer kugelsymmetrischen Ladungsverteilung ebenso radialsymmetrisch ist. Auf der Kugeloberfläche ist  $||\mathbf{E}||$  also konstant, und  $\mathbf{E}$  zeigt radial nach innen oder aussen. Damit ist das elektrische Feld einer radialsymmetrischen Verteilung genau das Feld  $||\mathbf{E}|| = Q/(4\pi\epsilon_0 r^2)$ , das man erhalten würde, wenn die gesamte Ladung der Kugel im Ursprung konzentriert wäre!

Bei Ladungsverteilungen, die einer bestimmten Symmetrie unterliegen, lassen sich oft Volumina/Oberflächen finden, so dass sich das elektrische Feld wie in diesem Beispiel durch den Gaußschen Integralsatz leicht berechnen lässt. Man bezeichnet diese Oberflächen als Gaußsche Oberflächen.

### 4.8.2 Das Greensche Theorem in der Ebene

Wir beschränken uns auf 2 Dimensionen, um einen einfachsten Integralsatz herzuleiten, der Oberflächenintegrale mit Linienintegralen um den Rand der Oberfläche verbindet. Wir betrachten zwei Funktionen von x und y, P(x, y) und Q(x, y), mit kontinuierlichen partiellen Ableitungen. Diese Funktionen sind frei gewählt und ändern ihren Wert als Funktion von x, y. Auf einer besonders einfachen Fläche, einem Rechteck A, werden wir nun ein konkretes Oberflächenintegral über A mit einem konkreten Linienintegral über den Rand dieses Rechtecks  $\partial A$  vergleichen.

Oberflächenintegral:

$$\iint_{A} dx \, dy \frac{\partial Q}{\partial x} = \int_{a}^{b} dx \int_{c}^{d} dy \frac{\partial Q}{\partial x} = \int_{c}^{d} dy \left[Q(b, y) - Q(a, y)\right]$$
(4.43)
(4.43)
(4.43)
(4.43)
(4.43)
(4.43)
(4.43)
(4.43)
(5)
(4.43)
(4.43)
(4.43)
(4.43)
(4.43)
(4.43)
(4.43)
(4.43)
(4.43)
(4.43)
(4.43)
(4.43)
(4.43)
(4.43)
(4.43)
(4.43)
(4.43)
(4.43)
(4.43)
(4.43)
(4.43)
(4.43)
(4.43)
(4.43)
(4.43)
(4.43)
(4.43)
(4.43)
(4.43)
(4.43)
(4.43)
(4.43)
(4.43)
(4.43)
(4.43)
(4.43)
(4.43)
(4.43)
(4.43)
(4.43)
(4.43)
(4.43)
(4.43)
(4.43)
(4.43)
(4.43)
(4.43)
(4.43)
(4.43)
(4.43)
(4.43)
(4.43)
(4.43)
(4.43)
(4.43)
(4.43)
(4.43)
(4.43)
(4.43)
(4.43)
(4.43)
(4.43)
(4.43)
(4.43)
(4.43)
(4.43)
(4.43)
(4.43)
(4.43)
(4.43)
(4.43)
(4.43)
(4.43)
(4.43)
(4.43)
(4.43)
(4.43)
(4.43)
(4.43)
(4.43)
(4.43)
(4.43)
(4.43)
(4.43)
(4.43)
(4.43)
(4.43)
(4.43)
(4.43)
(4.43)
(4.43)
(4.43)
(4.43)
(4.43)
(4.43)
(4.43)
(4.43)
(4.43)
(4.43)
(4.43)
(4.43)
(4.43)
(4.43)
(4.43)
(4.43)
(4.43)
(4.43)
(4.43)
(4.43)
(4.43)
(4.43)
(4.43)
(4.43)
(4.43)
(4.43)
(4.43)
(4.43)
(4.43)
(4.43)
(4.43)
(4.43)
(4.43)
(4.43)
(4.43)
(4.43)
(4.43)
(4.43)
(4.43)
(4.43)
(4.43)
(4.43)
(4.43)
(4.43)
(4.43)
(4.43)
(4.43)
(4.43)
(4.43)
(4.43)
(4.43)
(4.43)
(4.43)
(4.43)
(4.43)
(4.43)
(4.43)
(4.43)
(4.43)
(4.43)
(4.43)
(4.43)
(4.43)
(4.43)
(4.43)
(4.43)
(4.43)
(4.43)
(4.43)
(4.43)
(4.43)
(4.43)
(4.43)
(4.43)
(4.43)
(4.43)
(4.43)
(4.43)
(4.43)
(4.43)
(4.43)
(4.43)
(4.43)
(4.43)
(4.43)
(4.43)
(4.43)
(4.43)
(4.43)
(4.43)
(4.43)
(4.43)
(4.43)
(4.43)
(4.43)
(4.43)
(4.43)
(4.43)
(4.43)
(4.43)
(4.43)
(4.43)
(4.43)
(4.43)
(4.43)
(4.43)
(4.43)
(4.43)
(4.43)
(4.43)
(4.43)
(4.43)
(4.43)
(4.43)
(4.43)
(4.43)
(4.43)
(4.43)
(4.43)
(4.43)
(4.43)
(4.43)
(4.43)
(4.43)
(4.43)
(4.43)
(4.43)
(4.43)
(4.43)
(4.43)
(4.43)
(4.43)
(4.43)
(4.43)
(4.43)
(4.43)
(4.43)
(4.43)
(4.43)
(4.43)
(4.43)
(4.43)
(4.43)
(4.43)
(4.43)
(4.43)
(4.43)
(4.43)
(4.43)
(4.43)
(4.43)
(4.43)
(4.43)
(4.43)
(4.43)
(4.43)
(4.43)
(4.43)
(4.43)
(4.43)
(4.43)
(4.43)
(4.43)
(4.43)
(4.43)
(4.43)
(4.43)
(4.43)
(4.43)
(4.43)
(4.43)
(4.43)
(4.43)
(4.43)
(4.43)
(4.43)
(4.43)
(4.43)
(4.43)
(4.4

$$\sim \iint_{A} \mathrm{d}x \, \mathrm{d}y \frac{\partial Q}{\partial x} = \oint_{\partial A} dy Q(x, y) \qquad \text{für das gewählte Rechteck}$$

$$(4.44)$$

$$\text{Analog}^{2}: -\iint_{A} dx \, \mathrm{d}y \frac{\partial P}{\partial y} = \oint_{\partial A} \mathrm{d}x P(x, y) \qquad (4.45)$$

$$\sim \iint_{A} \mathrm{d}x \, \mathrm{d}y \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y}\right) = \oint_{\partial A} (\mathrm{d}x P(x, y) + \mathrm{d}y Q(x, y))$$

$$(4.46)$$

Dieses Ergebnis gilt natürlich zunächst nur für rechteckige Flächen. Betrachten wir aber eine ebene Fläche A beliebiger Form (ohne Löcher), die in eine Menge (beliebig kleiner) Rechtecke zerlegt werden kann. Summieren wir (4.44) über all diese Rechtecke, so heben sich Beiträge der Linienintegrale *aus dem Inneren* von A gegenseitig weg. Die Summe von (1) über alle Rechtecke, in die A zerlegt wurde ergibt dann auf der linken Seite ein Oberflächenintegral, auf der rechten Seite ein Linienintegral über den Rand von A



$$\int_{A} \mathrm{d}x \, \mathrm{d}y \left( \frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y} \right) = \oint_{\partial A} (\mathrm{d}x \ P(x, y) + \mathrm{d}y \ Q(x, y)) \tag{4.47}$$

Dieses Ergebnis heißt Greensches Theorem in der Ebene.

**Beispiel:** Betrachten wir als konkretes Beispiel die Funktionen Q(x, y) = x und P(x, y) = 0.

$$\int_{A} dx \, dy \left( \frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y} \right) = \iint_{A} dx \, dy$$
$$\oint_{\partial A} (dx \, P(x, y) + dy \, Q(x, y)) = \oint_{\partial A} dy \, x$$

Dieses Ergebnis kann genutzt werden, um die Berechnung des Flächeninhaltes von A zu ersetzen durch durch die Berechnung eines Linienintegrals. Auf diesem Prinzip basieren alte mechanische Apparate zur Bestimmung von Flächeninhalten (http://de.wikipedia.org/wiki/Planimeter). Als Beispiel berechnen wir das entsprechende Linienintegral um den Viertelskreis mit Radius eins.

- Teilstück 1:  $x = 0 \curvearrowright$  kein Beitrag zum Integral
- Teilstück 2:  $\frac{dy}{dt} = 0 \curvearrowright$  kein Beitrag zum Integral



0

#### 4.8.3 Satz von Stokes

Betrachten wir wieder ein rechteckiges Flächenelement und legen die z-Achse senkrecht zu seiner Oberfläche. Wir betrachten ein Vektorfeld  $\mathbf{v}(\mathbf{r})$  und nutzen das Greensche Theorem mit  $Q(x, y) = v_y(x, y, z = 0)$  und  $P(x, y) = v_x(x, y, z = 0)$ . (z = 0 ist konstant auf dem Rechteck.)

$$\iint_{A} \mathrm{d}x \,\mathrm{d}y \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y}\right) = \iint_{A} \mathrm{d}x \,\mathrm{d}y \,\left(\frac{\partial v_{y}}{\partial x} - \frac{\partial v_{x}}{\partial y}\right) = \int_{A} \mathrm{d}\mathbf{S} \cdot \nabla \times \mathbf{v} \qquad (4.48)$$
  
z-Komponente von  $\nabla \times \mathbf{v}$ .

Für das Linienintegral über den Rand von A erhalten wir

$$\oint_{\partial A} (dx \ P(x, y) + dy \ Q(x, y)) = \oint_{\partial A} dx \ v_x(x, y, z) + dy \ v_y(x, y, z) + dz \ v_z(x, y, z)$$
$$= \oint_{\partial A} d\mathbf{r} \cdot \mathbf{v}(\mathbf{r}) \ . \tag{4.49}$$

Der letzte Schritt folgt aus dz/dt = 0 für jede Parametrisierung des Rechtecks. Aus dem Greenschen Theorem folgt für dieses Rechteck

$$\int_{A} \mathrm{d}\mathbf{S} \cdot (\nabla \times \mathbf{V}) = \oint_{\partial A} \mathrm{d}\mathbf{r} \cdot \mathbf{v}(\mathbf{r}) . \qquad (4.50)$$

Als beziehung zwischen Vektoren gilt dieser Ausdruck unabhängig von der Wahl der kartesischen Basis (auch wenn wir zur Herleitung eine konkrete Basis genutzt haben, in der das Rechteck in der *xy*-Ebene liegt).

Für eine Oberfläche, die man in (viele) kleine Rechtecke zerlegen kann, heben sich die Beiträge kleiner Rechtecke im Inneren von A gegenseitig weg. Aus der Summe von (4.50 über diese Rechtecke erhalten wir damit den **Satz von Stokes** (math.: Stokes-Kelvin-Theorem)

$$\int_{A} \mathrm{d}\mathbf{S} \cdot (\nabla \times \mathbf{v}) = \oint_{\partial A} \mathrm{d}\mathbf{r} \cdot \mathbf{v}(\mathbf{r}) \ . \tag{4}$$



Analog zum Gaussschen Integralsatz findet der Satz von Stokes Anwendung in der Magnetostatik.

Info: Mit dem Satz von Stokes können wir auch den Zusam-

menhang zwischen Vektorfeldern  $\mathbf{v}(\mathbf{r})$  mit Rotation  $\nabla \times \mathbf{v} = 0$ 

und Gradientenfeldern zeigen. Für ein Vektorfeld mit verschwindender Rotation sind nach (4.51) Linienintegrale um eine geschlossenen Kurve null. Damit sind auch Linienintegrale vom Weg unabhängig: Wir teilen eine geschlossene Kurve C in ein Teilstück  $C_1$  von Punkt  $p_1$  zu  $p_2$  und ein Teilstück  $C_2$  von  $p_2$  zu  $p_1$  und schreiben das Linienintegral von  $\mathbf{v}$  entlang einer Linie C als  $I_C$ . Aus  $I_{C_1} + I_{C_2} = I_C = 0$  folgt dann  $I_{C_1} = -I_{C_2}$ . Drehen wir nun die Integrationsrichtung in einem der Teilstücke um, erhalten wir aus zwei unterschiedlichen Integrationswegen von  $p_1$  nach  $p_2$  dasselbe Ergebnis, Linienintegrale sind also wegunabhängig. Mit diesem Ergebnis lässt sich auch  $f(\mathbf{r})$  konstruieren, so dass  $\mathbf{v} = \nabla f$ :  $f(\mathbf{r}) = \int_{C(p_0, p_0 + \mathbf{r})} d\mathbf{r} \cdot \mathbf{v}(\mathbf{r})$ . Dabei beschreibt  $C(p_0, p_0 + \mathbf{r})$  eine beliebige Kurve vom Ursprung zum Punkt  $p_0 + \mathbf{r}$ . Nach dem obigen Ergebnis hängt das so definierte skalare Feld nur vom Anfangs- und Endpunkt dieser Kurve ab.  $f(\mathbf{r})$  ist nur bis auf eine additive Konstante bestimmt, die durch die Wahl des Ursprungs festgelegt wird.

Randnotiz: Dass Vektorfelder mit Rotation null Gradientenfelder sind gilt für **einfach zusammenhängenden Räume** (vereinfacht: Räume ohne Löcher wie z.B.  $\mathbb{R}^3$ ; genauer Räume in denen für alle Punktpaare ein Weg existiert der sie verbindet, und jeder geschlossene Weg auf einen Punkt zusammengezogen werden kann). In nicht einfach zusammenhängenden Räumen kann nämlich noch im Inneren einer berandeten Oberfläche ein weiterer Rand existieren, auf dem das Linienintegral ggf. nicht verschwindet.

# Fünf

# Folgen und Reihen, Potenzreihen und komplexe Zahlen

Mit Hilfe des Grenzwerts einer unendlichen Folge von Zahlen lernen wir Summen mit einer unendlichen Zahl von Termen zu berechnen. Praktische Anwendungen liegen in der Entwicklung von Funktionen als Potenzenreihen, sogenannte Taylorreihen; weite Teile der theoretischen Physik fußen auf kunstvollen Reihenentwicklungen. Das zweites Thema dieses Kapitels ist die Erweiterung der reellen Zahlen auf die komplexen Zahlen.

#### 5.1 Folgen



Wir betrachten eine (unendliche) **Folge** von Zahlen  $B_1, B_2, B_3, \ldots$  Ein Beispiel ist die Folge  $1, 1/2, 1/3, \ldots$  mit  $B_n = 1/n$ . Die Abbildung zeigt, wie die Terme  $B_n$  für wachsendes n immer näher an Null heranrücken, ohne die Null für endliche n je zu erreichen. Dennoch rückt 1/n für wachsende n beliebig nahe an die Null heran. Dieses "beliebig nahe" lässt sich wieder, wie in 3.1 präziser fassen: egal wie klein man  $\epsilon > 0$  wählt, gibt es immer ein

N, so dass für alle n > N die Differenz zwischen  $B_n$  und einer Zahl B kleiner ist als  $\epsilon$ . (Hier ist B = 0.) Die Abbildung zeigt Beispiele für  $\epsilon = 1/5$  und 1/10. Existiert für eine Folge für jedes  $\epsilon > 0$  ein  $N \in \mathbb{N}$ , so dass für alle n > N gilt  $|B_n - B| < \epsilon$ , so nennen wir die Folge **konvergent** mit **Grenzwert** B und schreiben  $\lim_{n\to\infty} B_n = B$ . Diese Definition beschreibt das Verhalten der Folge  $B_n$  für unendliche n, kommt aber selbst eleganterweise völlig ohne den Begriff "unendlich" aus.

**Beispiel:** Wir zeigen, dass der Grenzwert  $\lim_{n\to\infty} 1/n$  gleich Null ist. Für ein gegebenes  $\epsilon > 0$  wählen wir  $N \in \mathbb{N}$  größer  $1/\epsilon$ . Also gilt für alle n > N dass  $|B_n - B| = |1/n - 0| = 1/n < 1/N < \epsilon$ , wobei die erste Ungleichung aus n > N, die zweite aus  $N > 1/\epsilon$  folgt. (Überzeugen Sie sich, dass dies für andere Werte von B als Null nicht gilt.)

## 5.2 Reihen

Wir betrachten Folgen, deren einzelne Terme durch Summen gebildet werden. Eine Folge von Zahlen, die durch die Teilsummen  $b_1, b_1 + b_2, b_1 + b_2 + b_3, \ldots$  gebildet wird,

$$B_n = \sum_{k=1}^n b_k av{5.1}$$

heißt **Reihe**. Ist die Folge dieser Teilsummen konvergent, sprechen wir von einer **konvergenten Reihe**. Damit läßt sich auch einer Summe mit unendlich vielen Termen ein Wert zuweisen, nämlich der Grenzwert der Folge  $B_1, B_2, B_3, \ldots$  Informell bezeichnet man  $b_1 + b_2 + b_3 + \ldots$  als Reihe und meint die Folge der Teilsummen (denn nur dann kann man Konvergenz definieren). Beispiel: Als Beispiel betrachten wir die sogenannte geometrische Reihe

$$\sum_{k=0}^{\infty} x^k = 1 + 0.1 + 0.01 + 0.001 + \dots = 1.1111\dots$$
 für  $x = 0.1$  (5.2)  
=  $1 + 2 + 4 + 8 + \dots$  für  $x = 2$ .

Ob die unendliche Reihe  $\sum_{k=0}^{\infty} x^k$  konvergent ist oder nicht hängt also von x ab. Die Summe  $B_n$  der ersten Terme bis  $x^n$  dieser Reihe lässt sich leicht berechnen

$$(1-x)B_n = (1-x)\sum_{k=0}^n x^k = \sum_{k=0}^n x^k - \sum_{k=0}^n x^{k+1}$$
(5.3)  
$$= (1+x+x^2+\ldots+x^n) - (x+x^2+\ldots+x^n+x^{n+1}) = 1-x^{n+1}$$
$$B_n = \frac{1-x^{n+1}}{1-x}$$
$$\land B = \lim_{n \to \infty} B_n = \frac{1}{1-x} \text{ für } |x| < 1$$

Das letzte Gleichheitszeichen der zweiten Zeile ergibt sich, weil sich alle Terme in  $\sum_{k=0}^{n} x^k$  und  $-\sum_{k=0}^{n} x^{k+1}$  gegenseitig wegheben, bis auf den ersten und den letzten Term. Für  $n \to \infty$  geht  $x^k$  gegen Null, wenn |x| < 1. Die geometrische Reihe konvergiert also für |x| < 1 und divergiert für |x| > 1.

#### 5.3 Potenzreihen

Reihen aus unterschiedlichen Potenzen einer Variablen heißen Potenzreihen.

$$f: \mathbb{R} \longrightarrow \mathbb{R}$$
$$f(x) = \sum_{k=0}^{n} a_k x^k = a_0 + a_1 x + a_2 x^2 + \dots + a_n x^n \qquad a_k \in \mathbb{R}$$

Die geometrische Reihe (5.2) ist ein Beispiel für eine Potenzreihe. Für endliches n ist f(x) ein Polynom n-ter Ordnung, z.B.  $f(x) = 3 + 7x + 19.5x^2 + \pi x^3$  mit  $a_0 = 3, a_1 = 19.5, \ldots$ Grund sich mit unendlichen Potenzreihen zu beschäftigen ist, dass sich eine Vielzahl von Funktionen durch unendliche Potenzreihen darstellen lässt. Bricht man die Potenzreihen nach einer endlichen Zahl von Termen ab erhält man oft eine nützliche Näherung.

#### Der Konvergenzradius

Hängen die Terme einer Reihe von einem Parameter x ab, kann es also sein, daß die Reihe für bestimmte Werte von x konvergiert, für andere aber nicht. Mit der geometrischen Reihe

(5.2) haben wir ein Beispiel kennengelernt.

Nehmen wir an, eine Potenzreihe in  $\sum_{k=0}^{\infty} a_k x^k$  konvergiere für einen bestimmten Wert von  $x = x_1 \neq 0$ .

 $\bigcap |a_k x_1^k| \le C \ \forall k$  wobei  $C \in \mathbb{R}$  eine endliche Konstante ist,  $denn \ a_k x_1^k \ \text{kann nicht beliebig mit } k \ \text{wachsen}$   $\sum_{k=0}^{\infty} a_k x^k = \sum_{k=0}^n a_k x_1^k \left(\frac{x}{x_1}\right)^k$   $mit \ |a_k x_1^k| \le C \ \text{und}$   $\left(\frac{x}{x_1}\right)^k \ \text{hat Betrag} < 1, \ \text{wenn } |x| < |x_1|$  (5.4)

 $\curvearrowright$ jeder Term in  $\sum a_k x^k$ ist betragsmäßig kleiner als die Terme in

$$\sum_{k=0}^{\infty} C \left| \frac{x}{x_1} \right|^k = C \sum_{k=0}^{\infty} q^k \qquad q = \left| \frac{x}{x_1} \right|$$

 $\sim$  für  $|x| < |x_1|$  konvergiert die Reihe  $\sum_k a_k x^k$ . D. h. wenn eine Reihe für einen Wert von x konvergiert, muss sie das für alle betragsmäßig kleineren Werte auch tun. Zu jeder Potenzreihe  $\sum_{k=0}^{\infty} a_k x^k$  gibt es also ein  $\rho \in [0, \infty]$  mit der Eigenschaft, dass die Reihe für  $|x| < \rho$  konvergiert, und für  $x > \rho$  divergiert.  $\rho$  heißt **Konvergenzradius**. Über das Verhalten der Potenzreihe bei  $x = \rho$  macht dieser Satz übrigens keine Aussage, dort gibt es auch kein allgemeine gültiges Verhalten: es gibt Reihen die am Konvergenzradius konvergieren, Reihen die divergieren, oder sogar an einem Punkt konvergiere, an einem anderen divergieren.

#### Ausblick (ohne Beweise)

- Der Konvergenzradius lässt sich durch einen Grenzwert  $\rho = \lim_{n \to \infty} \left| \frac{a_n}{a_{n+1}} \right|$  bestimmen, wenn ab einem bestimmten *n* alle  $a_n$  von Null verschieden sind, und wenn der Grenzwert existiert. (Für den Fall, dass er nicht existiert, s. Satz von Cauchy-Hadamard.)
- Eine Reihe heißt **absolut konvergent**, wenn auch die Reihe ihrer Absolutbeträge konvergiert.  $\sum_{k=0}^{\infty} b_k$  ist absolut konvergent, wenn  $\sum_{k=0}^{\infty} |b_k|$  konvergiert. Innerhalb ihres Konvergenzradius ist jede Reihe absolut konvergent.

• Absolut konvergente Reihen können Term für Term multipliziert werden

$$\left(\sum_{k=0}^{\infty} a_k\right) \left(\sum_{k'=0}^{\infty} b_{k'}\right) = \sum_{k \ k^1}^{\infty} a_k b_{k^1}$$

• eine Potenzreihe  $\sum_{k=0}^{\infty} a_k x^k$  heißt gleichmäßig konvergent in einem Bereich  $D \subset \mathbb{R}$  wenn

$$\forall x \in D \ \forall \epsilon > 0 \ \exists N \in \mathbb{N} : \ \forall n > N \ \left| \sum_{k=0}^{n} a_k x^k - \sum_{k=0}^{\infty} a_k x^k \right| < \epsilon$$

(Ein strengeres Kriterium als Konvergenz an jedem einzelnen Punkt x; eine Beziehung zwischen N und  $\epsilon$  muss jetzt statt an einem einzelnen Punkt im gesamten Definitionsbereich der Funktion gelten.) Innerhalb ihres Konvergenzradius konvergieren Potenzreihen gleichmäßig.

• Gleichmäßig konvergente Potenzreihen können Term für Term differenziert und integriert werden

$$\lim_{n \to \infty} \int dx \left( \sum_{k=0}^{\infty} a_k x^k \right) = \lim_{n \to \infty} \sum_{k=0}^n \left( \int dx \ a_k x^k \right)$$

und Limites (Grenzwerte) vertauschen bei gleichmäßiger Konvergenz der Reihe

$$\lim_{x \to A} \lim_{y \to B} f(x, y) = \lim_{y \to B} \lim_{x \to A} f(x, y)$$

Der Beweis dieser Aussagen ist ein wichtiger Teil von Vorlesungen in Analysis.

#### 5.4 Taylorreihe

Unser Ziel bei der Einführung von Potenzreihen war es gewesen, möglichst allgemeine Funktionen als Potenzreihen darzustellen (innerhalb des Konvergenzradius). Wir betrachten eine Reihe aus den Potenzen von x,

$$f(x) = \sum_{k=0}^{\infty} a_k x^k = a_0 + a_1 x + a_2 x^2 + a_3 x^3 + \dots$$
(5.5)

Dazu sind zunächst die Koeffizienten der Potenzreihe  $a_0, a_1, a_2$  zu bestimmen. Für k = 0 betrachten wir die linke und rechte Seite von (5.5) bei x = 0.

$$f(x) = \sum_{k=0}^{\infty} a_k x^k = a_0 + a_1 x + a_2 x^2 + \dots$$

$$k = 0: \qquad f(0) = a_0 + a_1 x|_{x=0} + a_2 x^2|_{x=0} + \dots = a_0 \qquad \qquad \bigcirc a_0 = f(0)$$

$$k = 1: \qquad f'(0) = a_1 + 2 a_2 x|_{x=0} + 3 a_3 x^2|_{x=0} + \dots = a_1 \qquad \qquad \bigcirc a_1 = f'(0)$$

$$k = 2: \qquad f^{(2)}(0) = 1 \times 2 a_2 + 3 2 a_3 x|_{x=0} + \dots = 2 a_2 \qquad \qquad \bigcirc a_2 = \frac{1}{2!} f^{(2)}(0)$$

$$k = 3: \qquad f^{(3)}(0) = 1 \times 2 \times 3 a_3 + 4 \times 3 \times 2 a_4 x|_{x=0} + \dots = 3! a_3 \qquad \bigcirc a_2 = \frac{1}{3!} f^{(3)}(0)$$
allgemeines k: 
$$\qquad \qquad \bigcirc a_k = \frac{1}{k!} f^{(n)}(0)$$

Analog können wir auch die Funktion f(x) als Potenzreihe in  $x - x_0$  schreiben, d.h.  $f(x) = \sum_{k=0}^{\infty} a_k (x - x_0)^k$ , mit neuen Koeffizienten  $a_k = \frac{1}{k!} f^{(n)}(x_0)$ . Die Potenzreihen in Potenzen von x und  $x - x_0$  sind natürlich eng miteinander verknüpft: Die Terme der Potenzreihe  $(x - x_0)^n$  lässt sich ausmultiplizieren und die Potenzreihe in Potenzen von x schreiben, wir erhalten dann wieder (5.5) mit den entsprechenden Koeffizienten.

Ist die Funktion  $f:(a,b) \Rightarrow \mathbb{R}$  beliebig oft differenzierbar und  $x_0 \in (a,b)$ , heißt

$$\sum_{k=0}^{\infty} \frac{f^{(k)}(x_0)}{k!} (x - x_0)^k$$

die **Taylorreihe** bei  $x_0$ , ihre Koeffizienten die **Taylorkoeffizienten**.  $R_n(x) = \sum_{k=n+1}^{\infty} \frac{f^{(n)}(x_0)}{k!} (x - x_0)^k$  heißt **Restglied** der Reihe (n-ter Ordnung).

**Beispiel:** Die Exponentialfunktion und trigonometrische Funktionen. Die Exponentialfunktion  $e^x$  hat eine besonders einfache Taylorreihe, da ihre erste Ableitung (sowie alle weiteren Ableitungen)  $e^x$  sind, damit ist die Taylorreihe u x = 0 gleich  $1 + x + \frac{1}{2!}x^2 + \frac{x^3}{3!} + \frac{x^4}{4!} + \ldots$  Für kleine Werte von x gibt die Taylorreihe oft bereits nach einer kleinen Zahl von Termen eine brauchbare Näherung;  $e^x = 1 + x + \mathcal{O}(x^2)$  (Taylorreihe bis zur ersten Ordnung in x),  $e^x = 1 + x + x^2/2 + \mathcal{O}(x^3)$  (Taylorreihe bis zur zweiten Ordnung in x, etc).



Die Abbildung zeigt die Exponentialfunktion und ihre Taylorentwicklung bei x = 0 bis zur n-ten Potenz von x für n = 0, 1, 2, 3, 4.

Als weiteres Beispiel noch die Taylorreihen für die trigonometrischen Funktionen bei  $x_0 = 0$ . Mit  $f(x) = \cos(x)$ ,  $\frac{df}{dx} = -\sin(x)$ ,  $\frac{d^2f}{dx^2} = -\cos(x)$  erhalten wir die Taylor-Reihe der Cosinus-Funktion  $1 - \frac{x^2}{2} + \frac{x^4}{4!} - \frac{x^6}{6!} + \dots$ , und analog die Taylor-Reihe der Sinus-Funktion  $x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + \dots$ 



Info: Wir haben nicht gefragt, welche Funktionen sich prinzipiell durch eine konvergente Taylorreihe (5.5) darstellen lassen. Funktionen die (lokal) gleich ihrer Taylorreihe sind

heißen **analytisch**. Die Frage, welche Funktionen analytisch sind, findet erst im Rahmen der Funktionentheorie (Theorie von Funktionen komplexer Veränderlicher) eine zufriedenstellende Antwort.

#### 5.5 Die komplexen Zahlen

Die Mathematik beruht auf den natürlichen Zahlen. Um bestimmte Gleichungen lösen zu können, wurden bisweilen neue Arten von Zahlen eingeführt, z.B. rationale und irrationale Zahlen oder negative Zahlen. Nur Gleichungen wie  $x^2 = -1$  blieben bislang ohne Lösung.

### Komplexe Zahlen und $\mathbb{R}^2$

Wie betrachten den Vektorraum  $\mathbb{R}^2 = \left\{ \begin{pmatrix} a \\ b \end{pmatrix} : a, b \in \mathbb{R} \right\}$  mit der Addition von Vektoren und Multiplikation mit reellen Zahlen definiert durch

$$\begin{pmatrix} a \\ b \end{pmatrix} + \begin{pmatrix} a' \\ b' \end{pmatrix} = \begin{pmatrix} a + a' \\ b + b' \end{pmatrix}$$

$$\lambda \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} \lambda a \\ \lambda b \end{pmatrix} \qquad \lambda \in \mathbb{R} .$$

$$(5.6)$$

Zusätzlich definieren wir noch eine Multiplikation von zwei Vektoren

$$\begin{pmatrix} a \\ b \end{pmatrix} \cdot \begin{pmatrix} a' \\ b' \end{pmatrix} = \begin{pmatrix} a a' - b b' \\ a b' + b a' \end{pmatrix} .$$
 (5.8)

Diese Multiplikation

ist kommutativ:

hat ein Einselement:

und erlaubt eine Lösung von  $x^2 = -1$ 

hat ein Inverses:

$$\begin{pmatrix} a'\\b' \end{pmatrix} \cdot \begin{pmatrix} a\\b \end{pmatrix} = \begin{pmatrix} a'a-b'b\\a'b+b'a \end{pmatrix} = \begin{pmatrix} aa'-bb'\\ab'+ba' \end{pmatrix} = \begin{pmatrix} a\\b \end{pmatrix} \cdot \begin{pmatrix} a'\\b' \end{pmatrix}$$

$$\begin{pmatrix} 1\\0 \end{pmatrix} \cdot \begin{pmatrix} a\\b \end{pmatrix} = \begin{pmatrix} a\\b \end{pmatrix}$$

$$\begin{pmatrix} a\\b \end{pmatrix}^{-1} \equiv \frac{1}{a^2+b^2} \begin{pmatrix} a\\-b \end{pmatrix} \qquad \begin{pmatrix} a\\b \end{pmatrix}^{-1} \cdot \begin{pmatrix} a\\b \end{pmatrix} = \begin{pmatrix} 1\\0 \end{pmatrix}$$

$$\begin{pmatrix} 0\\1 \end{pmatrix} \cdot \begin{pmatrix} 0\\1 \end{pmatrix} = -\begin{pmatrix} 1\\0 \end{pmatrix}$$

Wir schreiben  $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$  als 1 und  $\begin{pmatrix} 0 \\ 1 \end{pmatrix}$  als *i* mit  $i \cdot i = -1$ . a + ib entspricht dann  $\begin{pmatrix} a \\ b \end{pmatrix}$  und wird als **komplexe Zahl** bezeichnet. Jede polynome (algebraische) Gleichung *n*-ten Grades (mit komplexen Koeffizienten) hat *n* Lösungen in den komplexe Zahlen.



a wird als **Realteil**, b als **Imaginärteil** einer komplexen Zahl a + ib bezeichnet. Wir schreiben  $a = \Re(a + ib)$  und  $b = \Im(a + ib)$ . Es sind also zwei reelle Zahlen nötig um eine einzelne komplexe Zahl zu spezifizieren. Analog zur Zahlengeraden der reellen Zahlen, werden komplexe Zahlen graphisch in der Ebene dargestellt (**komplexe Ebene** oder **Gaußsche Ebene**); reelle Zahlen liegen dabei auf der x-Achse, rein imaginäre Zahlen auf der y-Achse.

Eine komplexe Zahl z = a + ib kann auch in Polarkoordinaten  $r = \sqrt{a^2 + b^2}$ ,  $\theta = \arctan \frac{b}{a}$  dargestellt werden.  $r = |z| \equiv \sqrt{a^2 + b^2}$  wird als **Betrag** einer komplexen Zahl z bezeichnet,  $\theta$  als ihr Argument.  $z^* = a - ib$  definiert die **komplex Konjugierte** von z = a + ib. Dann ist  $z z^* = (a + ib)(a - ib) = a^2 - i ab + i ba - (i)^2 b^2 = a^2 + b^2 = |z|^2$ 

Multiplikation von komplexen Zahlen erfolgt dann einfach durch Ausmultiplizieren

$$zz' = (a+ib)(a'+ib') = aa'+iab'+iba'+(i\cdot i)bb' = (aa'-bb')+i(ab'+ba'), \quad (5.9)$$

das Ergebnis stimmt mit (5.8) überein. Für die Division gilt nutzen wir  $\frac{z_1}{z_2} = \frac{z_1}{z_2} \frac{z_2^*}{z_3^*} = \frac{z_1 z_2^*}{|z_2|^2}$  (in Übereinstimmung mit dem inversen Element der Multiplikation).

**Die Euler-Formel** 

$$e^{ix} \equiv \sum_{k=0}^{\infty} \frac{(ix)^k}{k!} = 1 + ix - \frac{x^2}{2} + \dots$$
  
=  $\sum_{k=0,2,4,6,\dots} \frac{(ix)^k}{k!} + \sum_{k=1,3,5,7,\dots} \frac{(ix)^k}{k!}$   
=  $\sum_{k=0}^{\infty} (-1)^k \frac{x^{2k}}{(2k)!} + i \sum_{k=0}^{\infty} (-1)^k \frac{x^{2k+1}}{(2k+1)!}$   
=  $\cos x + i \sin x$ 

$$e^{ix} = \cos x + i \sin x \tag{5.10}$$

Die sogenannte **Eulerformel** verbindet trigonometrische Formeln mit der Exponentialfunktion. Eine komplexe Zahl z mit Betrag r und Argument  $\theta$  ist damit  $z = re^{i\theta}$ . Das Produkt zweier komplexer Zahlen  $z_1$  und  $z_2$  ist dann  $z_1z_2 = r_1r_2e^{i\theta_1}e^{i\theta_2} = r_1r_2e^{i(\theta_1+\theta_2)}$ , d.h. der Betrag des Produktes ist  $r_1r_2$ , sein Argument ist  $\theta_1 + \theta_2$ . Multiplikation komplexer Zahlen hängt also zusammen mit Drehungen in der komplexen Ebene.

Diese Eigenschaft komplexer Zahlen führt auf ihre wichtigste praktische Anwendung: Bei Schwingungen und Oszillationen erlaubt die Eulerformel (komplizierte) trigonometrische Berechnungen mit der (einfachen) Exponentialfunktion zu ersetzen. **Beispiel:** Eine einfache Anwendung der Eulerschen Formel sind die trigonometrischen Additionstheoreme

$$cos(x + y) + i sin(x + y) = e^{i(x+y)} = e^{ix} e^{iy} = (cos x + i sin x)(cos y + i sin y) (5.11) = cos x cos y - sin x sin y + i(sin x cos y + sin y cos x) .$$

Aus dem Realteil dieser Gleichung erhalten wir  $\cos(x + y) = \cos x \cos y - \sin x \sin y$ , aus dem Imaginärteil  $\sin(x + y) = \sin x \cos y + \sin y \cos x$ .

**Info:** Die komplexen Zahlen erscheinen zunächst abstrakt und schwer zugänglich. Das galt für die irrationalen Zahlen oder die negativen Zahlen aber sicher auch einmal. In der Mathematik spielen die komplexen Zahlen wegen der sogenannten algebraischen Abgeschlossenheit eine wichtige Rolle: Jede algebraische Gleichung von Grad größer Null hat eine Lösung in den komplexen Zahlen. (Dagegen gibt es Gleichungen, die keine reellen Lösungen haben, wie  $x^2 + 1 = 0$ .) In der Physik ist der Zusammenhang zwischen Trigonometrie und der Exponentialfunktion zentral bei der Beschreibung von Schwingungen und Oszillationen (s. Kapitel 6). Und schließlich wird die Quantenmechanik durch einen speziellen Vektorraum (Hilbertraum) über den komplexen Zahlen beschrieben.

Sechs

# Differentialgleichungen

Physik ist in weiten Teilen Naturbeschreibung mit Differentialgleichungen: viele Naturgesetze können mit Gleichungen formuliert werden, die eine Funktion sowie Ableitungen dieser Funktion enthalten. Wir lernen unterschiedliche Arten von Differentialgleichungen kennen und entwickeln verschiedene Lösungsstrategien.

# 6.1 Differentialgleichungen in der Physik

Einzelne Messungen können durch Zahlen beschrieben werden, die ihnen zugrundeliegenden Prozesse sind jedoch (meist) durch Funktionen beschrieben, z.B. eine Bahnkurve  $\mathbf{r}(t)$ oder ein Potential V(x, y, z).

Oft gehorchen diese Funktionen Differentialgleichungen

$$m \frac{d^2 \vec{r}}{dt^2} = \vec{F}(\vec{r}, t)$$
 Newton II  
$$\nabla \cdot \vec{E} = \rho(\vec{r})/\epsilon_0$$
 Maxwell I

Eine **Differentialgleichung** (DGL) ist eine Gleichung für eine unbekannte Funktion einer oder mehrerer Variablen, die die Funktion selbst und ihr Ableitungen verknüpft. Die klassische Mechanik, Elektrodynamik, Quantenmechanik basieren alle auf Differentialgleichungen. Bevor wir unterschiedliche Gleichungen klassifizieren und lösen lernen, diskutieren wir mehrere Beispiele. **Beispiel:** Das zweite Newtonsche Gesetz in einer Dimension, beschrieben durch die Variable y

$$F(y(t),t) = ma(t) = m\frac{d^2y(t)}{dt^2}$$

verbindet die Beschleunigung einer Punktmasse mit der Kraft die auf die Punktmasse wirkt. Letztere kann sowohl von der Zeit, als auch vom Ort abhängen. Wir betrachten mehrere Spezialfälle.

1. Konstante Kraft F = mg. Diese Fall beschreibt die Bewegung eines Teilchens in einem konstanten Gravitationsfeld

$$\frac{d^2y}{dt^2} = g \tag{6.1}$$

Diese DGL ist eine Gleichung für eine Funktion y(t) der Zeit, d.h. rechte und linke Seite sind Funktionen der Zeit (die rechte Seite ist in diesem Beispiel eine konstante Funktion). Die gesuchte Lösung, eingesetzt in die linke Seite der Differentialgleichung, muss für alle Zeiten t gleich der rechten Seite sein. Gesucht ist eine also Funktion y(t), deren zweite Ableitung nach t für alle Zeiten gleich g ist.

 $y(t) = \frac{1}{2}gt^2 + at + b$  löst diese Gleichung, da  $\frac{d^2}{dt^2}(gt^2/2) = g$  und  $\frac{d^2}{dt^2}(at + b) = 0$ . Die Koeffizienten a und b sind aus den Anfangsbedingungen  $y(t = 0) = y_0$ ,  $\frac{dy}{dt}(0) = v_0$  zu bestimmen: Einsetzen von t = 0 gibt  $a = y_0$  und  $b = v_0$ .

2. Zeitabhängige Kraft  $F = Am \cos(\Omega t)$ . Dieser Fall beschreibt ein Teilchen, das einer oszillierenden Kraft unterworfen ist

$$\frac{d^2y}{dt^2} = A\cos(\Omega t) \tag{6.2}$$

 $y(t) = -\frac{A}{\Omega^2}\cos(\Omega t) + at + b$  löst die Gleichung, da  $\frac{d^2}{dt^2}\cos(\Omega t) = -\Omega^2\cos(\Omega t)$  und  $\frac{d^2}{dt^2}(at+b) = 0.$ 

3. Ortsabhängige Kraft F(y) = -ky. Dieser Fall beschreibt eine Punktmasse an einer linearen Feder.

$$\frac{d^2y}{dt^2} = -\frac{k}{m} y(t) \equiv -\omega_0^2 y(t) \qquad , \ \omega_0^2 \equiv \frac{k}{m} \tag{6.3}$$

$$\frac{d^2y}{dt^2} + \omega_0^2 y(t) = 0 \tag{6.4}$$

(6.5)

Diese DGL unterscheidet sich qualitativ von den ersten beiden Beispielen. Gesucht ist eine Funktion y(t), deren zweite Ableitun**g**4nicht gleich einer bekannten Funktion ist, sondern proportional zur gesuchten Funktion selbst. Wir werden effektive Strategien zur Lösung solcher DGL kennenlernen, hier erraten wir die Lösung. Wir erwarten eine Oszillation der Masse an der Feder mit zeitlich konstanter Amplitude;  $y(t) = A \cos(\omega_0 t - \phi)$ . Einsetzen in die DGL zeigt, dass dieser Ansatz tatsächlich die DGL löst, da zweifaches Ableiten einer Kosinus oder Sinusfunktion wieder eine Kosinus/Sinusfunktion ist.

#### 6.2 Terminologie

Eine Gleichung für eine Funktion, die eine oder mehrere Ableitung enthält, heißt **Differentialgleichung**, die höchste Ordnung dieser Ableitungen heißt **Ordnung der DGL**. Die Variable, nach der abgeleitet wird, heißt **unabhängige Variable**, die Variable, die abgeleitet wird, heißt **abhängige Variable**. DGL, die nur Ableitungen nach einer einzelnen Variablen enthalten, heißen **gewöhnliche DGL**. DGL, die (partielle) Ableitungen nach mehreren Variablen enthalten, heißen **partielle DGL**.

### 6.3 Gewöhnliche DGL erster Ordnung

**Beispiel:** Radioaktiver Zerfall. Die Zahl der pro<br/> kleinem Zeitintervall zerfallenden Kerne ist proportional zur Zahl der vorhandenen Kerne <br/>  $\boldsymbol{y}$ 

$$\frac{dy}{dt} = -\lambda y(t) \tag{6.6}$$

Die Zahl radioaktiver Kerne nimmt exponentiell mit der Zeit ab: tatsächlich wird  $\frac{dy}{dx} + \lambda y(t) = 0$  gelöst durch  $y(t) = y_0 e^{-\lambda t}$ . Die Konstante  $y_0$  kann aus der Anfangsbedingung  $y(x = 0) = y_0 e^{-\lambda 0} = y_0$  bestimmt werden.

An dieser Stelle ist eine konzeptionelle Hürde zu überwinden: die DGL  $\frac{dy}{dx} = -\lambda y(t)$  ist auf der rechten Seite eine Funktion von y (hier die Funktion y selbst): Die Zahl der pro Zeiteinheit zerfallenden Kerne ist gegeben durch  $\lambda y$ , proportional zur Zahl der Kerne y. Die DGL setzt sie gleich der negativen Ableitung von y nach der Zeit. Finden wir allerdings die Lösung  $y(t) = Ae^{-\lambda t}$  und setzen sie in die DGL ein, steht auf der rechten Seite eine Funktion von t allein, nämlich  $-\lambda Ae^{-\lambda t}$ . Dennoch ist die DGL selbst eine Gleichung, die explizit y und t enthalten kann, erst ihre Lösung etabliert eine Beziehung  $y = y(t)^1$ .

<sup>&</sup>lt;sup>1</sup>Nicht anders ist es bei einer algebraischen Gleichung wie x - 3 = 4. Prinzipiell kann x jeden reellen Wert annehmen. Aber nur x = 7 löst die Gleichung. In dem obigen Beispiel kann die Zahl der Kerne y(t) zum Zeitpunkt t jeden Wert annehmen, aber nur  $y(t) = Ae^{-\lambda t}$  löst die DGL  $\frac{dy}{dt} = -\lambda y(t)$ .



Eine allgemeine DGL kann nach  $\frac{dy}{dx}$  aufgelöst werden (linke Seite), auf der rechten Seite stehen dann Terme, die von der abhängigen Variable y und vielleicht noch von der abhängigen Variable x abhängen. In dem Beispiel des radioaktiven Zerfalls ist die DGL (6.6) bereits nach  $\frac{dy}{dx}$  aufgelöst, auf der rechten Seite steht ein Term, der nur von y abhängt und als Funktion von y geschrieben werden kann  $\frac{dy}{dx} = f(y(x))$  mit  $f(y) = -\lambda y$ . Den letzten Ausdruck können wir auch als f(y(x)) = y(x) schreiben, f(y) = y betont, dass wir die rechte Seite dieser DGL kennen wenn wir y kennen (wenn 1000 Atome vorhanden sind, werden pro kleiner Zeiteinheit 1000 $\lambda$  Atome zerfallen, dazu müssen wir nicht die Uhrzeit (abhängige Variable) wissen). Die DGL erster Ordnung zum Beispiel

$$(\frac{dy}{dx})^3 = \sin(y(x)) + x^3$$
(6.7)

kann aufgelöst werden zu  $\frac{dy}{dx} = \sqrt[3]{\sin(y(x)) + x^3}$ . Die rechte Seite ist bekannt wenn y und x bekannt sind,  $\frac{dy}{dx} = f((y(x)), x)$  mit  $f(y, x) = \sqrt{\sin(y) + x^3}$ .

Allgemeine DGL erster Ordnung können also geschrieben werden als

$$\frac{dy}{dx} = f(y(x), x) \; ,$$

indem man die DGL nach der Ableitung erster Ordnung auflöst. Im Beispiel des radioaktiven Zerfalls (6.6) ist  $f(y, x) = -\lambda y$ . Die Lösung dieser DGL lässt sich als eine Kurve y(x)visualisieren, die überall die durch die Funktion f(y, x) vorgegebene Steigung f hat. Eine solche Kurve kann man geometrisch konstruieren, indem man die xy-Ebene zunächst mit kleinen Strichen füllt, die jeweils die Steigung f(y, x) haben. Gegeben eine Anfangsbedingung  $y(x) = y_x$  bei einem gegebenen x, können wir uns vorstellen den "Strichen zu folgen". Die entstehende Kurve ist überall tangential zu den kleinen Strichen, hat also überall die Steigung f(y(x), x) und löst damit die DGL. **Info:** Die graphische Darstellung legt auch eine einfache numerische Methode nahe, um DGL mit Hilfe eines Computers zu lösen. Ausgehend von y(0), das durch die Anfangsbedingung festgelegt sei, suchen wir nach Näherungen der Lösung auf einem diskreten Gitter  $x = 0, \Delta x, 2\Delta x, \ldots$  Dazu machen wir diskrete Schritte entlang des Gitters. Aus  $(y(\Delta x) - y(0))/\Delta x \approx dy/dx = f(y, x)$  am ersten Gitterpunkt folgt der Wert  $y(\Delta x) \approx y(0) + f(y(0), x)\Delta x$ . Dieser Schritt lässt sich nun wiederholen um iterativ  $y(2\Delta x), y(3\Delta x), \ldots$  zu bestimmen. Die Schrittweite  $\Delta x$  bestimmt die Genauigkeit des Ergebnisses. Mehr dazu in der Computerphysikvorlesung!

#### 6.3.1 Eindeutigkeit

Sei eine DGL  $\frac{dy}{dx} = f(y(x), x)$  auf einem Rechteck  $(a, b) \times (c, d)$  definiert, so dass ein endliches L > 0 existiert mit

$$|f(y,x) - f(\tilde{y},x)| \le (y - \tilde{y})L \qquad \forall x, y, \tilde{y} \text{ auf dem Rechteck}$$
(6.8)

so sind zwei Lösungen, die bei einem  $x_0 \in [a, b]$  denselben Wert haben, im gesamten Intervall [a, b] gleich (Satz von Picard-Lindelöf).

Die sogennante Lipschitz-Bedingung (6.8) ist in der Praxis fast immer erfüllt. Dann legt eine Anfangsbedinung die Lösung einer DGL erster Ordnung eindeutig fest. Für jede Gleichung, die vorgibt die Natur zu beschreiben, erwarten wir diese Eindeutigkeit: Gegeben die Anfangsbedingung und ein Naturgesetz wie das zweite Newtonsche Gesetz, ist der Zustand des Systems für alle weiteren Zeiten eindeutig festgelegt. Gäbe es mehrere Lösungen wäre diese Mechanik keine (vollständige) Beschreibung des Systems.

Ein Nebeneffekt ist Arbeitsersparnis: Schaffen wir es - egal wie - eine Lösung der DGL zu finden, die die Anfangsbedingung erfüllt, sind wir am Ziel; weitere Lösungen gibt es nicht.

#### 6.3.2 Gekoppelte Differentialgleichungen mehrerer Veränderlicher

Die graphische Konstruktion der Lösung einer DGL und die Eindeutigkeit der Lösung scheint zunächst auf DGL erster Ordnung beschränkt. Dieser Zugang lässt sich allerdings leicht auf DGL höherer Ordnung verallgemeinern. Dazu definieren wir DGL in mehreren Veränderlichen  $y_1(x), y_2(x), \ldots$ 

 $\begin{aligned} \frac{dy_1}{dx} &= f_1(y_1, y_2, y_3, x) & \text{gekoppelte DGL erster Ordnung} \\ \frac{dy_2}{dx} &= f_2(y_1, y_2, y_3, x) & \text{Vektorschreibweise } \frac{d\mathbf{y}}{dx} = \mathbf{f}(\mathbf{y}) \\ \frac{dy_3}{dx} &= f_3(y_1, y_2, y_3, x) \end{aligned}$ 

Man bezeichnet diese Differentialgleichungen als **gekoppelt**, wenn sie sich nicht als mehrere DGL in nur jeweils einer einzelnen Variablen schreiben lassen, d.h. wenn in  $f_1$  ausser  $y_1$  noch z.B.  $y_2$  etc. auftritt. Lösungen solcher Gleichungssysteme sind wieder eindeutig, wenn eine Verallgemeinerung der Lipschitz-Bedingung (6.8) erfüllt ist.

Differentialgleichungen höherer Ordung lassen sich als gekoppelte DGL erster Ordung schreiben.

**Beispiel:** Wir betrachten das zweite Newtonsche Gesetz als gewöhnliche DGL zweiter Ordung. Indem wir eine neue Variable  $v(t) = \frac{dy}{dt}$  einführen erhalten wir zwei gekoppelte DGL erster Ordnung

> $m \frac{dv}{dt} = F(y,t)$  zwei gekoppelte DGL  $\frac{dy}{dt} = v(t)$ .

### 6.3.3 Separation der Variablen

Eine Lösungsmethode existiert für einen Spezialfall von DGL erster Ordnung, nämlich sogenannte **separable** Gleichungen

$$\frac{dy}{dx} = h(y)g(x) , \qquad (6.9)$$

bei denen die rechte Seite ein Produkt von Funktionen von x und von y sind. Die gesuchte Lösung ist eine Funktion y(x), die (6.9) erfüllt. Mit dieser (noch unbekannten) Funktion y(x) können beide Seiten als Funktion von x verstanden werden. Umformung und Integration über ein Intervall  $[x_b, x_e]$  ergibt dann

$$\int_{x_b}^{x_e} \mathrm{d}x g(x) = \int_{x_b}^{x_e} \mathrm{d}x \frac{\mathrm{d}y}{\mathrm{d}x} h^{-1}(y(x)) = \int_{y(x_b)}^{y(x_e)} \mathrm{d}y h^{-1}(y).$$
(6.10)

Im letzten Schritt wurde durch Variablensubstitution integriert. Für gegebenes  $x_b$  ergibt die Auswertung des Integrals links eine Funktion von  $x_e$ , rechts eine Funktion von  $y_e$ , die nach  $y_e = y(x_e)$  aufgelöst die Lösung der DGL (6.9) ergibt.

**Beispiel:** Wir wenden die Methode der Separation der Variablen auf die Differentialgleichung (6.6) an, die radioaktiven Zerfall beschreibt. Anfangsbedingung sei  $y(t = 0) = y_0$ . Mit  $f(y,t) = -\lambda y = h(y)g(t)$  schreiben wir h(y) = y und  $g(t) = -\lambda$ . Aus (6.10) erhalten wir mit  $t_b = 0$ 

$$\int_{t_b}^{t_e} \mathrm{d}tg(t) = -\lambda \int_0^{t_e} \mathrm{d}t = -\lambda [t]_0^{t_e} = -\lambda t_e \tag{6.11}$$

$$= \int_{y(t_b)}^{y(t_b)} dy h^{-1}(y) = \int_{y_0}^{y(t_c)} dy \frac{1}{y} = [\ln(y)]_{y_0}^{y(t_e)} = \ln(y(t_e)) - \ln(y_0) = \ln\left(\frac{y(t_e)}{y_0}\right) .$$

Auflösen nach  $y(t_e)$  ergibt das erwartete Ergebnis  $y(t) = y_0 e^{-\lambda t}$ .

### 6.4 Gewöhnliche lineare DGL

Betrachten wir die gewöhnliche DGL *n*-ter Ordnung für y(x) mit einer reellen Veränderlichen x

$$a_n(x) \frac{d^n y}{dx^n} + a_{n-1}(x) \frac{d^{n-1} y}{dx^{n-1}} + \dots + a_1(x) \frac{dy}{dx} + a_0(x) y(x) = f(x)$$
(6.12)

Diese DGL heißt **lineare DGL** da y(x) und seine Ableitungen in jedem Term in erster Potenz auftreten. Viele fundamentale Gleichungen der Physik sind lineare DGL (Maxwell-Gleichungen der Elektrodynamik, Schrödingergleichung der Quantenmechanik). Die rechte Seite f(x) wird als **Inhomogenität** der DGL bezeichnet. Die Gleichung lässt sich in der Form

$$\left[a_n(x)\frac{d^n}{dx^n} + a_{n-1}(x)\frac{d^{n-1}}{dx^{n-1}} + \ldots + a_1(x)\frac{d}{dx} + a_0(x)\right]y(x) = f(x)$$
(6.13)

schreiben, wobei wir den Term in eckigen Klammern als eine Abbildung  $\mathcal{L} \equiv \left[a_n(x) \frac{d^n}{dx^n} + a_{n-1}(x) \frac{d^{n-1}}{dx^{n-1}} + \ldots + a_1(x) \frac{d}{dx} + a_0(x)\right]$ (einen Differentialoperator) von Funktionen auf Funktionen auffassen können, mit

$$y(x) \xrightarrow{\mathcal{L}} a_n(x) \frac{d^n y}{dx^n} + a_{n-1}(x) \frac{d^{n-1} y}{dx^{n-1}} + \dots + a_1(x) \frac{dy}{dx} + a_0(x) y(x) .$$
(6.14)

Der Differentialoperator  $\mathcal{L}$  einer linearen DGL ist eine lineare Abbildung (im Raum der Funktionen),  $\mathcal{L}(a y_1(x) + b y_2(x)) = a \mathcal{L} y_1(x) + b \mathcal{L} y_2(x).$ 

Dieser Zugang erlaubt die kompakte Schreibweise der DGL  $\mathcal{L} y = f$ . Die Lösung y folgt dann aus der Inversen des Operators  $\mathcal{L}$ : Gesucht ist y(t), so daß der Differentialoperator angewandt auf y(t) die Funktion f(t) ergibt.

Aus der Linearität von  $\mathcal{L}$  ergibt sich auch eine wichtige mathematische Eigenschaft, die alle homogenen (f(t) = 0) DGL teilen. Seien  $y_1(x)$  und  $y_2(x)$  Lösungen einer linearen, homogenen DGL, d.h.  $\mathcal{L} y_1 = 0$ ,  $\mathcal{L} y_2 = 0$ . Dann ist auch die Linearkombination  $a y_1 + b y_2$ Lösungen der DGL, denn

$$\mathcal{L}(a y_1 + b y_2) = a \mathcal{L} y_1 + b \mathcal{L} y_2 = 0.$$
(6.15)

Ebenso ist  $\lambda y_1$  eine Lösung. Die Lösungen einer linearen homogenen DGL bilden also einen Vektorraum, man kann sie addieren und mit Zahlen multiplizieren und erhält wieder eine Lösung der DGL.

**Beispiel:** Als Beispiel betrachten wir die lineare DGL  $\frac{d^2y}{dt^2} + \omega_0^2 y(t) = 0$ ,  $\mathcal{L} = \frac{d^2}{dt^2} + \omega_0^2$  ist der dazugehörige (lineare) Differentialoperator, mit ihm kann die Gleichung als  $\mathcal{L}y = 0$ geschrieben werden. Diese homogene lineare gewöhnliche DGL beschreibt zum Beispiel eine Punktmasse an einer linearen Feder,  $m\frac{d^2y}{dt^2} = -ky$ . Sowohl  $y(t) = \cos(\omega_0 t)$  als auch  $y(t) = \sin(\omega_0 t)$  lösen diese DGL. Auch  $y(t) = a\cos(\omega_0 t) + b\sin(\omega_0 t)$  ist eine Lösung für beliebige  $a, b, \text{denn } \frac{d^2y}{dx^2} = -a\omega_0^2\cos(\omega_0 t) - b\omega_0^2\sin(\omega_0 t) = -\omega_0^2y(t)$ . Diese Linearkombination von Sinus- und Kosinusfunktionen gleicher Kreisfrequenz wiederum eine oszillierende Funktion, da  $A\cos(\omega_0 t - \phi) = A\cos(\omega_0 t)\cos(-\phi) + A\sin(\omega_0 t)\sin(-\phi) = a\cos(\omega_0 t) + b\sin(\omega_0 t)$  (6.16) mit  $A\cos(\phi) \equiv a$  und  $A\sin(\phi) \equiv -b$  und damit  $A = \sqrt{a^2 + b^2}$  und  $\tan(\phi) = -b/a$  (s.

mit  $A\cos(\phi) \equiv a$  und  $A\sin(\phi) \equiv -b$  und damit  $A = \sqrt{a^2 + b^2}$  und  $\tan(\phi) = -b/a$  (s. Additionstheorem im Abschnitt über komplexe Zahlen 5.5).

Für inhomogene Gleichungen gilt etwas ähnliches: Sei  $y_1(x)$  eine Lösung einer inhomogenen Gleichung  $\mathcal{L} y_1 = f$  und  $y_2(x)$  eine Lösung der dazugehörigen homogenen Gleichung  $\mathcal{L} y_2 = 0$ , dann ist  $y_1(x) + a y_2(x)$  eine weitere Lösung der inhomogenen Gleichung:

$$\mathcal{L}(y_1 + a \, y_2) = \mathcal{L}y_1 + a \, \mathcal{L} \, y_2 = f \, . \tag{6.17}$$

**Beispiel:** Diese Eigenschaft linearer DGL haben wir bereits bei den Beispielen (6.1) und (6.2) kennengelernt. Die Lösung der inhomogenen Gleichung  $\frac{d^2y}{dt^2} = g$  ist  $y(t) = gt^2/2$ , Lösungen der dazugehörigen homogenen  $\frac{d^2y}{dt^2} = 0$  Gleichung sind y(t) = 1 und y(t) = t, ihre Linearkombination at + b löst ebenso die homogene Gleichung. Erst die Summe der Lösung der inhomogenen Gleichung und die der homogene Gleichung,  $y(t) = gt^2/2 + at + b$ , ist dann die vollständige Lösung der DGL. Die Koeffizienten a und b werden aus der Anfangsbedingung (Ort und Geschwindigkeit) bestimmt.

Komplexe Lösungen linearer DGL Für lineare DGL sind Realteil und Imaginärteil einer komplexen Lösung wiederum Lösungen der DGL. Diese Eigenschaft wird ihren enormen Wert zeigen, wenn wir in Abschnitt 6.4.3 die lineare Bewegungsgleichung des harmonischen Oszillators lösen. Sei  $\mathcal{L}$  ein reeller Operator (die Koeffizienten aller Ableitungen sind reellwertig),  $z(x) = z_R(x) + i z_I(x)$  aber eine komplexe Lösung der (homogenen) DGL. Aus  $\mathcal{L}z = 0 = 0 + i 0$  und der Linearität des Differentialoperators  $\mathcal{L}$  folgt durch Vergleich von Real- und Imaginärteil

$$0 = \mathcal{L}(z_R + i z_I) = \mathcal{L}z_R + i \mathcal{L}z_I$$

$$(6.18)$$

$$\sim \mathcal{L}z_R = 0, \ \mathcal{L}z_I = 0.$$

**Beispiel:** Als Beispiel betrachten wir wieder die lineare DGL  $\frac{d^2y}{dt^2} + \omega_0^2 y(t) = 0$ . Sie wird gelöst durch die komplexe Funktion  $z(t) = e^{i\omega_0 t}$ , da  $\frac{d^2y}{dt^2} = -\omega_0^2 e^{i\omega_0 t} = -\omega_0^2 z(t)$ . Realund Imaginärteil von  $z(t) = \cos(\omega_0 t) + i\sin(\omega_0 t)$  sind ebenso Lösungen der DGL.

#### 6.4.1 Gewöhnliche lineare DGL erster Ordnung: homogene Gleichung

$$\frac{dy}{dx} + a(x)y(x) = 0, \qquad \mathcal{L}y = 0 \qquad \mathcal{L} \equiv \frac{d}{dx} + a(x) . \tag{6.19}$$

In Analogie mit der DGL  $\frac{dy}{dx} + \lambda y(x) = 0$ , die durch die Exponentialfunktion  $y(x) = y(0)e^{-\lambda x}$  gelöst wird, versuchen wir zuerst den Ansatz  $y(x) = y(0)e^{-a(x)x}$ . Mit

$$\frac{dy}{dx} = -y(0) \frac{d}{dx} (a(x)x) e^{-a(x)x}$$
(6.20)

ergibt er allerdings keine Lösung der DGL. Der zweite Versuch mit  $y(x) = y(0) \exp \left\{-\int_0^x dx' a(x')\right\} \equiv y(0) e^{-A(x)}$  ist allerdings erfolgreich

$$y(x) = y(0) \exp\left\{-\int_0^x dx' \, a(x')\right\} \equiv y(0) \, e^{-A(x)}$$
$$\frac{dy}{dx} = y(0) \, \frac{d}{dx} \left(-\int_0^x dx' \, a(x')\right) e^{-A(x)}$$
$$= -y(0) \, a(x) \, e^{-A(x)} = -a(x) \, y(x) \; . \tag{6.21}$$

#### 6.4.2 Gewöhnliche lineare DGL erster Ordnung: inhomogene Gleichung

Als Ansatz zur Lösung der inhomogenen DGL

$$\frac{dy}{dx} + a(x)y(x) = f(x) \tag{6.22}$$

wählen wir  $y(x) = c(x) e^{-A(x)}$  mit einer noch zu bestimmenden Funktion c(x). Bei der Ableitung dieses dieses Ansatzes nach x erhalten wir nach der Produktregel 2 Terme:  $\frac{dy}{dx} = -a(x) c(x) e^{-A(x)} + c'(x) e^{-A(x)}$ . Der erste Term und a(x) y(x) heben sich gegenseitig weg. Der zweite Term muss dann  $c'(x) e^{-A(x)}$  gleich der rechten Seite von (6.22) sein:

$$\sim \frac{dy}{dx} + a(x) y(x) = c'(x) e^{-A(x)} \stackrel{\text{DGL}}{=} f(x)$$

$$\sim c'(x) = e^{A(x)} f(x)$$

$$c(x) = \int_0^x dx' e^{A(x')} f(x')$$

$$y(x) = c(x) e^{-A(x)} = \int_0^x dx' e^{-A(x) + A(x')} f(x')$$

$$= \int_0^x dx' G(x, x') f(x')$$

$$\text{mit } G(x, x') = e^{-A(x) + A(x')}$$

$$(6.23)$$

G(x, x') heißt die **Greensche Funktion** dieser DGL. Die einzelnen Werte von f(x') im Intervall  $0 \le x' \le x$  bestimmen also y(x) bei einem gegebenen Wert von x. Die Greenfunktion G(x, x') gibt an, wie y(x) von f(x') abhängt. Ist x die Zeit t, dann beschreibt die Greensche Funktion G(t, t') den Einfluss von f(t') auf die Lösung zu einem späteren Zeitpunkt t.

Bei dieser Lösung fällt auf, daß es keinen freien Parameter zur Wahl einer Anfangsbedingung gibt, bei x = 0 ist y = 0. Addieren wir allerdings zur Lösung der inhomogenen DGL Linearkombinationen der Lösung von  $\mathcal{L}y = 0$ , erhalten wir weitere Lösung der inhomogenen DGL (siehe oben). Die Lösung für beliebige Anfangsbedingung ist dann

$$y(x) = y(0) e^{-A(x)} + \int_0^x dx' e^{-A(x) + A(x')}$$
(6.24)

Der erste Term in diesem Ausdruck wird als **allgemeine Lösung** bezeichnet (da er nicht von der speziellen Form der Inhomogenität abhängt), der zweite Term als **spezielle Lösung**. Bei x = 0 ist der zweite Term gleich null und die Exponentialfuntion im ersten Term gleich eins, y(x = 0) ist damit gleich der frei gewählten Anfangsbedingung. Analog, wenn die Anfangsbedingung nicht bei x = 0, sondern bei einem anderen Wert von x = a vorliegt wählen wir als untere Integralgrenzen den entsprechenden Wert von x und den Vorfaktor des ersten Term als y(a).

Diese Methode zur Lösung von Differentialgleichungen wird (etwas paradoxerweise) als **Variation der Konstanten** bezeichnet. Den Vorfaktor c(x) der Exponentialfunktion, der in der homogenen Lösung eine Konstante ist, lässt man nun variieren, um eine Lösung einer inhomogenen Gleichung zu finden.

## 6.4.3 Gewöhnliche lineare DGL zweiter Ordnung

Gewöhnliche lineare DGL zweiter Ordnung können in die Form

$$\frac{d^2y}{dx^2} + a_1(x)\frac{dy}{dx} + a_0(x)y(x) = f(x)$$
(6.25)

gebracht werden (indem wir beide Seiten durch einen möglichen Koeffizienten  $a_2(x)$  teilen). Wir beschränken uns hier auf konstante Koeffizienten (obwohl es, wie bei der DGL erster Ordnung, wieder Lösungen zu allgemeinen Koeffizienten gibt, die Greensche Funktion für diesen Fall ist aber komplizerter). DGL mit konstanten Koeffizienten sind sowohl von den physikalischen Anwendungen, als auch vom Lösungsansatz her interessant.

**Beispiel:** Eine Masse m an einer Feder mit Federstärke k bewege sich in einer Dimension durch eine viskose Flüssigkeit (Reibungskraft proportional zur Geschwindigkeit) und wird von einer externen Kraft  $F_{\text{ext}}(t)$  angetrieben. Aus der Newtonschen Bewegungsgleichung  $m \frac{d^2x}{dt^2} = -kx - \gamma \frac{dx}{dt} + F_{\text{ext}}(t)$  erhalten wir mit  $b = \gamma/m, c = k/m, f(t) = F_{\text{ext}}/m$ 



#### 6.4.4 Gedämpfter hamonischer Oszillator (homogene Gleichung)

Wir beginnen mit der Analyse des Falles ohne externe Antriebskraft, f(t) = 0.

$$\frac{d^2x}{dt} + b\frac{dx}{dt} + cx(t) = 0.$$
(6.27)

Wir erwarten generell als Lösung eine sinusförmige Schwingung mit exponentiell abfallender Amplitude und wählen als Ansatz  $x(t) = Ae^{-\xi t} \cos(\omega t - \phi)$ . Einsetzen dieses Ansatzes in die DGL (6.27) führt allerdings auf eine aufwändige trigonometrische Rechnung (ausprobieren!). Wir erinnern uns: bei linearen DGL sind Realteile und Imaginärteile von komplexen Lösungen ebenso Lösungen der DGL. Der komplexer Ansatz  $x(t) = Ae^{i\omega t} A, \omega \in \mathbb{C}$  beschreibt sowohl die Oszillation, als auch den exponentiellen Abfall von x mit t. Mit der Zerlegung eines komplexen  $\omega$  in Realteil und Imaginärteil  $\omega = \omega_R + i\omega_I (\omega_R, \omega_I \in \mathbb{R})$  ist nämlich

$$e^{i\omega t} = e^{i\omega_R t} e^{-\omega_I t} = (\cos(\omega_R t) + i\sin(\omega_R t))e^{-\omega_I t} , \qquad (6.28)$$

der Realteil von  $e^{i\omega t}$  ist also eine exponentiell abfallende Kosinusfunktion. Eine Phasenverschiebung dieser Funktion lässt sich über eine komplexe Amplitude  $A = \tilde{A}e^{-i\phi}$   $(\tilde{A}, \phi \in \mathbb{R})$  beschreiben,

$$\Re[Ae^{i\omega_R t}] = \Re[\tilde{A}e^{-i\phi+i\omega_R t}]$$
  
=  $\tilde{A}\Re[\cos(\omega_R t - \phi) + \sin(\omega_R t - \phi)]$   
=  $\tilde{A}\cos(\omega_R t - \phi)$ . (6.29)

Der komplexe Exponentialansatz ist die wichtigste Anwendung der Eulerformel (5.10) in der Physik. Die Ableitungen von  $A e^{i\omega t}$  nach t sind jetzt nämlich einfach

$$\begin{aligned} x(t) &= A e^{i\omega t} \\ \frac{dx}{dt} &= i\omega A e^{i\omega t} \\ \frac{d^2x}{dt^2} &= -\omega^2 A e^{i\omega t} \end{aligned}$$

Für die DGL (6.27) ergibt der Ansatz  $x(t) = A e^{i\omega t}$  dann

$$(-\omega^{2} + bi\omega + c) A e^{i\omega t} = 0 \qquad \forall t$$
  

$$A = 0 \quad \text{(triviale Lösung)}$$
  

$$-\omega^{2} + ib\omega + c = 0$$

Die ersten Lösung A ist uninteressant und entspricht einem ruhenden Pendel. Die zweite Lösung führt auf eine quadratische Gleichung in  $\omega$  mit den beiden Lösungen

$$\omega_{+/-} = \frac{-ib \pm \sqrt{-b^2 + 4c}}{-2} = \frac{i}{2}b \pm \frac{1}{2}\sqrt{4c - b^2} .$$
(6.30)

Mithilfe des Exponentialansatzes  $A e^{i\omega t}$  ist es also gelungen, die lineare Differentialgleichung (6.27) auf eine algebraische Gleichung zu reduzieren. Die Lösung der homogenen DGL (6.27) ist der Realteil von  $A_+e^{i\omega_+t} + A_-e^{i\omega_-t}$ . (Bei linearen DGL beliebiger Ordnung mit konstanten Koeffizienten wäre der Exponentialansatz übrigens genauso erfolgreich.) Für den Spezialfall ohne Reibung b = 0 erhalten wir die beiden Lösungen

$$\omega_{+/-} = \pm \omega_0 \equiv \pm \sqrt{c} = \pm \sqrt{k/m},$$

und somit  $z(t) = A_+e^{i\omega_0 t} + A_-e^{-i\omega_0 t}$   $(A_+, A_- \in \mathbb{C})$  mit einem Realteil der Form  $x(t) = \tilde{A}\cos(\omega_0 t - \phi)$ . Diese Lösung hat zwei freie Parameter,  $\tilde{A}$  und  $\phi$ , die noch aus den Anfangsbedingungen (Position und Geschwindigkeit) zu bestimmen sind. Die Auslenkung des Oszillators zum Zeitpunkt t = 0 sei  $x_0$ , seine Geschwindigkeit sei  $v_0$ . Setzen wir in der Lösung x(t) und ihrer Ableitung nach der Zeit diese Anfangsbedingungen ein, erhalten wir

$$\tilde{A}\cos(-\phi) = x_0$$

$$-\omega_0 \tilde{A}\sin(-\phi) = v_0 .$$
(6.31)

Teilen wir die zweite Gleichung durch  $\omega_0$  und nutzen  $\cos^2(-\phi) + \sin^2(-\phi) = 1$  erhalten wir  $\tilde{A} = \sqrt{x_0^2 + v_0^2/\omega_0^2}$ . Teilen wir die zweite Gleichung durch die erste, erhalten wir tan(-phi) =  $v_0/(x_0\omega_0)$ , und damit die Phasenverschiebung  $\phi$ .

Für  $b \geq 0$  führen wir eine Fallentscheidung für steigendes b (steigende Dämpfung) durch

# 1. Gedämpfte harmonische Schwingung $4c - b^2 > 0$

$$\begin{split} \omega_{+/-} &= \pm \omega_R + i\omega_I \qquad \omega_R = \frac{1}{2}\sqrt{4c - b^2} \quad , \quad \omega_I = b/2 \\ z(t) &= A_+ e^{i\omega_+ t} + A_- e^{i\omega_- t} \\ \text{mit Realteil } A e^{-\omega_I t} \cos(\omega_R t - \phi) \; , \end{split}$$

die Kreisfrequenz  $w_R = \frac{1}{2}\sqrt{4c - b^2}$  nimmt mit zunehmender Dämpfung *b* ab, die Abklingrate  $\omega_I = b/2$  wächst mit *b*.

# 2. Kritische Dämpfung $4c - b^2 = 0$

$$\omega_{\pm} = \frac{i}{2}b$$
  
ergibt  $x(t) = A_{\pm}e^{i\omega_{\pm}t}$ 

Da  $\omega$  rein imaginär ist fällt  $x_R(t)$  ohne Oszillation exponentiell ab. Überraschend ist, dass der Exponentialansatz nur eine Lösung findet. Wir benötigen aber 2 freie Parameter, um alle Anfangsbedingungen umsetzen zu können. Die zweite Lösung findet sich durch Betrachtung des Grenzfalls  $b \to b_c \equiv 2\sqrt{c}$ . Wir betrachten eine Linearkombination der beiden Lösungen  $e^{i\omega_+t}$  und  $e^{i\omega_-t}$  aus (6.30) und führen den den Grenzfall  $\omega_+ \to \omega_-$  aus

$$\lim_{\omega_+ \to \omega_-} \frac{e^{i\omega_+ t} - e^{i\omega_- t}}{i\omega_+ - i\omega_-} = \lim_{\omega_+ \to \omega_-} e^{i\omega_+ t} \frac{1 - e^{i(\omega_- - \omega_+)t}}{i\omega_+ - i\omega_-}$$
$$= \lim_{\omega_+ \to \omega_-} e^{i\omega_+ t} \frac{1 - (1 + (i\omega_- - i\omega_+)t)}{i\omega_+ - i\omega_-} = te^{i\omega_- t}$$

und erhalten eine weitere Lösung, die sich nicht einfach aus dem Exponentialansatz ergibt. Die allgemeine Lösung ist dann die Linearkombination der beiden Lösungen

$$x(t) = A e^{i\omega t} + Bt e^{i\omega t} \qquad i\omega = -b/2$$

mit zwei freien Parametern (Realteile von A, B).

3. Überdämpfter Oszillator  $4c - b^2 < 0$ 

x(t)

ť

$$\omega_{+/-} = \frac{i}{2}b \pm \frac{i}{2}\sqrt{|4c - b^2|} = \frac{i}{2}\left(b \pm \sqrt{b^2 - 4c}\right)$$

Die allgemeine Lösung ist dann gegeben durch

$$x(t) = A_{+}e^{i\omega_{+}t} + A_{-}e^{i\omega_{-}t} = A_{+}e^{-\omega_{+}t} + A_{-}e^{-\omega_{-}t}$$

Wegen  $\omega_+ > \omega_-$  entsprechen die beiden Terme einem (relativ) schnellen und einem langsamen Abklingprozess, beide ohne Oszillation. Für lange Zeiten ist diese Lösung durch den zweiten (langsamen) Term dominiert.

Zum Schluss vergleichen wir noch das Abklingverhalten unter überkritischer Dämpfung und kritischer Dämpfung. Wegen

$$w_{-I} = \frac{b - \sqrt{b^2 - 4c}}{2} < \frac{b}{2} = \omega_{cI} \quad \text{(kritisches Abklingen)}$$

führt kritische Dämpfung mit  $b = b_c$  zum schnellsten Abklingen, bei der nach einer Auslenkung des Oszillators x(t) so schnell wie möglich zur Ruhelage x = 0 zurückkehrt.

# 6.4.5 Gedämpfter harmonischer Oszillator (inhomogene Gleichung)

$$\mathcal{L}x = f$$
  $\mathcal{L} = \frac{d^2}{dt^2} + b\frac{d}{dt} + \omega_0^2$ 

Physikalisch beschreibt die Inhomogenität f(t) eine externe Kraft, die den Oszillator antreibt. Wir beschränken uns auf zwei Spezialfälle. Wie wir im nächsten Kapitel sehen werden, ist deren Lösung bereits ausreichend, um die Lösung des harmonischen Oszillators unter einer beliebigen periodischen Antriebskraft zu berechnen.

Der erste Spezialfall ist eine konstante externe Kraft f(t) = f. Als Ansatz für die spezielle Lösung wählen wir x(t) = A, einsetzen in die DGL ergibt A = f/c.

#### 6.4.6 Sinusförmige externe Kraft

Der zweite Spezialfall ist eine sinusförmige externe Kraft. Dieser Fall ist wichtig für eine Vielzahl von Anwendungen (regelmässiges Anschubsen einer Schaukel, regelmässige Windstöße gegen eine Brücke). Im nächsten Kapitel werden wir sehen, wie sich auch nichtsinusförmige Kräfte aus Sinusschwingungen unterschiedlicher Frequenz zusammensetzen lassen. Die Kreisfrequenz  $\Omega$  und Antriebsamplitude F sind vorgegeben, sie hängen auch nicht von den Eigenschaften des Oszillators ab. Gesucht ist also die Lösung der DGL

$$\frac{d^2}{dt^2}x + b\frac{dx}{dt} + \omega_0^2 x(t) = F \cos\Omega t \tag{6.33}$$

Wir nutzen wieder die Linearität der DGL um eine komplexe Lösung zu finden. Hierbei lassen wir auch eine komplexe Inhomogenität zu und schreiben die Antriebskraft als  $Fe^{i\Omega t}$ (mit reellem  $\Omega$  und F). Denn für einen reellen Differentialoperator  $\mathcal{L}$  und die DGL  $\mathcal{L}z = f_z = f_R + i f_I$  ist der Realteil einer Lösung von  $\mathcal{L}z = f_z$  eine Lösung von  $\mathcal{L}y = f_R$ 

$$f_R + i f_I = \mathcal{L}(z_R + i z_I) = \mathcal{L}z_R + i\mathcal{L}z_I \quad \curvearrowright \mathcal{L}z_R = f_R .$$
(6.34)

Wir suchen also die Lösung von  $\frac{d^2}{dt^2}x + b\frac{dx}{dt} + cx(t) = Fe^{i\Omega t}$  und nutzen wieder einen komplexen Ansatz  $z(t) = Ae^{i\omega t}$ . Wir erhalten die Gleichung

$$(-\omega^2 + i \, b \, \omega + \omega_0^2) \, A e^{i \, \omega \, t} = F \, e^{i \, \Omega \, t} \,, \tag{6.35}$$

die für alle Zeiten t erfüllt sein muss. Dazu muss  $\omega = \Omega$  gelten, d.h. der Oszillator folgt der Kreisfrequenz der externen Kraft. Für die (komplexe) Amplitude A erhalten wir

$$A = |A|e^{-i\phi} = \frac{F}{\omega_0^2 - \Omega^2 + i\Omega b}$$

$$(6.36)$$

mit Betrag  $|A| = F((\omega_0^2 - \Omega^2)^2 + \Omega^2 b^2)^{-1/2}$  und Phase  $\phi = \arctan(\frac{\Omega b}{\omega_0^2 - \Omega^2})$ . Der Realteil dieser komplexen Lösung

$$\Re(Ae^{i\Omega t}) = \Re(|A|e^{i\Omega t - i\phi}) = |A|\cos(\Omega t - \phi)$$
(6.37)

ist die Lösung der inhomogenen Gleichung.

R = |A|/F gibt das Verhältnis der Antriebsamplitude F und der Amplitude der inhomogenen Lösung an und hängt ab von der Kreisfrequenz der Antriebskraft  $\Omega$ , der Kreisfrequenz des ungedämpften harmonischen Oszillators  $\omega_0 = \sqrt{k/m}$ , und dem Dämpfungsparameter b. Bei kleinen Werten der Dämpfung kann R für  $\Omega \approx \omega_0$  sehr groß werden. Dieses



107

Phänomen wird als **Resonanz** bezeichnet. Die Abbildung zeigt R für  $\omega_0 = 1$  und b = 0 (rot) und  $b = 0.2, 0.4, \ldots$  Der Effekt lässt sich am einfachsten mit dem Kind auf der Schaukel illustrieren: Aufgabe ist mit kleiner Antriebsamplitude A eine große Schwingung der Schaukel zu erreichen. Dies erzielt ein externer Schaukelanschubser, wenn er der Schaukel immer einen Schubs gibt, wenn sich die Schaukel gerade ihm weg bewegt. Die Kreisfrequenz der Schaukel muss also gleich der des Antriebs sein, der Antrieb muss in Phase mit der Bewegung der Schau-

kel sein. Ist die Kreisfrequenz des Antriebs  $\Omega$  gleich der Kreisfrequenz des gedämpften harmonischen Oszillators  $\sqrt{\omega_0^2 - b^2/4}$  erreicht R sein Maximum. (Ohne Dämpfung liegt dieses Maximum bei Unendlich. In einem realistischen System würde der Oszillator allerdings bei großen Auslenkungen Effekte der Nichtlinearität zeigen.)

Um die vollständige Lösung zu bestimmen, müssen wir zu dieser Lösung der inhomogenen Gleichung noch die allgemeine Lösung der homogenen Gleichung addieren. Die beiden freien Koeffizienten der allgemeinen Lösung sind dann aus den Anfangsbedingungen zu bestimmen. Für  $b \neq 0$  fällt die Amplitude der allgemeinen Lösung allerdings exponentiell ab, für grosse Zeiten ist x(t) also asymptotisch durch die Lösung der inhomogenen Gleichung gegeben. Die Amplitude von von x(t) ist dann konstant |A|, über eine ganze Periode ist die Energie die durch die externe Kraft dem Pendel zugeführt wird genau gleich der Energiedissipation durch Reibung.

# 6.5 Partielle DGL (PDG)

DGL mit Ableitungen nach unterschiedlichen Variablen (partielle Ableitungen) werden als **partielle DGL (PDG)** bezeichnet. Die Klassifizierung nach Ordung der DGL, linear/nichtlinear ist analog zu gewöhnlichen DGL. PDG treten bei der räumlichen oder raumzeitlichen Beschreibung von Feldern auf.
Diffusionsgleichung

$$\frac{\partial \rho}{\partial t} = D \frac{\partial^2 \rho}{\partial x^2}$$
(s. auch Skizze unten)

#### Laplacegleichung der Elektrostatik

$$\frac{\partial^2 \phi}{\partial x^2} + \frac{\partial^2 \phi}{\partial y^2} + \frac{\partial^2 \phi}{\partial z^2} = 0$$
 für das elektr. Potential  $\phi(x, y, z)$  i

für die Teilchendichte  $\rho(x,t)$ 

Wellengleichung 
$$\frac{\partial^2}{\partial x^2}\phi = \frac{1}{c^2}\frac{\partial^2}{\partial t^2}\phi$$

#### Schrödingergleichung der Quantenmechanik

$$i\hbar\frac{\partial}{\partial t}\psi = -\frac{\hbar^2}{2m}\frac{\partial^2}{\partial x^2}\psi(x,t) + V(x)\psi(x,t) \qquad \qquad \text{für Wellenfunktion }\psi(x,t) \text{ eines Terms}$$



PDG sind generell viel schwerer zu lösen als gewöhnliche Differentialgleichungen, außerdem ist die Theorie zur Existenz und Eindeutigkeit von Lösungen sehr viel komplexer als bei gewöhnlichen Differentialgleichungen (auch wenn wir bei PDG die die Natur beschreiben davon ausgehen, dass Lösungen dieser Gleichungen existieren und eindeutig sind).

### 6.5.1 Randbedingungen

Zusätzlich zu den Anfangsbedingungen schränken bei PDF oft Randbedingungen die Lösung ein. Als Beispiel betrachten wir die Wellengleichung in einer Dimension,  $\frac{\partial^2 \phi}{dx^2} = \frac{1}{c^2} \frac{\partial^2 \phi}{\partial t^2}$ . Sie beschreibt zum Beispiel die (kleine) Auslenkung eines Gummibandes, das zwischen zwei Punkten aufgespannt ist. An diesen Punkten, z.B. x = 0 und x = L ist dann  $\phi(x = 0, t) = 0, \phi(x = L, t) = 0$  für alle Zeiten t. Analog in 2D: Eine über eine Öffnung gespannte Membran (Trommel) mit Auslenkung  $\phi(x, y, t)$  gehorcht der Wellengleichung  $\frac{\partial^2 \phi}{dx^2} + \frac{\partial^2 \phi}{dy^2} = \frac{1}{c^2} \frac{\partial^2 \phi}{\partial t^2}$ . Dann ist auf der Öffnung (auf der wir die Membran festgezurrt haben)  $\phi = 0$  für alle Zeiten t.

Unsere Strategie zur Lösung von linearen homogenen PDG (wie die obigen Beispiele) ist analog zu gewöhnlichen DGL. Wir suchen zunächst unterschiedliche Lösungen (ggf. mit Einschränkungen durch Randbedingungen), und suchen dann die Linearkombination dieser Lösungen, die die Anfangsbedingungen erfüllt.

### 6.5.2 Separationsansatz

Der Separationsansatz erlaubt eine Vielzahl (aber bei weitem nicht alle) PDG auf gewöhnliche DGL zurückzuführen, die dann verhältnismäßig leicht gelöst werden können. Sämtliche oben aufgeführte PDG lassen sich durch diesen Ansatz lösen. Vor allem in der Quantenmechanik stellt sich dieser Ansatz als sehr mächtig heraus. Als konkretes Beispiel betrachten wir die Wellengleichung in 1D

 $\frac{\partial^2 \phi}{\partial x^2} = \frac{1}{c^2} \frac{\partial^2 \phi}{\partial t^2} \tag{6.38}$ 

mit Randbedingung  $\phi(x = 0, t) = \phi(x = L, t) = 0$   $\forall t$  (z.B. ein Seil, das zwischen zwei Punkten bei x = 0 und x = L aufgespannt ist, mit kleiner Auslenkung  $\phi(x, t)$  in einer Richtung rechtwinklig zur x-Achse).

Als Ansatz für die Lösung nutzen wir das Produkt einer (noch zu bestimmenden) Funktion von x und einer (ebenso noch zu bestimmenden) Funktion von t;  $\phi(x,t) = f(x) g(t)$ . Diesen Ansatz setzen wir in die Wellengleichung ein und teilen auf beiden Seiten durch f(x) g(t)

$$\frac{\partial^2}{\partial x^2} (f(x) g(t)) = \frac{1}{c^2} \frac{\partial^2}{\partial t^2} (f(x) g(t))$$

$$g(t) \frac{d^2}{dx^2} f(x) = \frac{1}{c^2} f(x) \frac{d^2}{dt^2} g(t) \qquad \text{(Linearität der PDG)}$$

$$\underbrace{\frac{1}{f(x)} \frac{d^2}{dx^2} f(x)}_{\text{Funktion von } x} = \underbrace{\frac{1}{c^2} \frac{1}{g(t)} \frac{d^2}{dt^2} g(t)}_{\text{Funktion von } t} \qquad \text{(durch } \phi(x, t) \text{ geteilt)}$$

Diese Gleichung muss für alle x, t gelten. Die linke Seite ist eine Funktion von x, die rechte Seite eine Funktion von t. x und t sind unabhängige Variablen. Die einzige Möglichkeit ist, dass beide Seiten konstante Funktionen sind

Die Konstante  $\alpha$  ist noch zu bestimmen. Aus der PDF in zwei Variablen hat der Separationsansatz zwei *gewöhnliche* DGL gemacht, die jetzt noch zu lösen sind:

$$\frac{d^2 f}{dx^2} = \alpha f(x) \qquad f(x) = A \sin kx \qquad kL = \pi n \quad , \quad n \in \mathbb{N}$$

$$\frac{d^2 g}{dt^2} = c^2 \alpha g(t) \qquad g(t) = B e^{ikct}$$

$$= B e^{i\omega t}$$

$$\operatorname{mit} \omega = kc \qquad \curvearrowright \alpha = -k^2 \; . \tag{6.39}$$

Jedes  $k=\frac{\pi n}{L}$  mit  $n\in\mathbb{N}$ löst die DGL mit der obigen Randbedingung und ergibt eine Lösung

$$\phi(x,t) = C_n \sin\left(\frac{\pi xn}{L}\right) e^{i\frac{\pi n}{L}ct} .$$
(6.40)

Die Konstante  $\alpha$  in (6.39) ist dann durch  $\alpha = -k^2$  gegeben. Prinzipiell könnte  $\alpha$  auch positiv sein, dann wäre f(x) aber eine Exponentialfunktion  $e^{\pm kx}$  und die Randbedingung f(0) = 0, f(L) = 0 liesse sich nicht implementieren.

 $k = \pi n/L = 2\pi/\lambda$  wird als **Kreiswellenzahl** der Lösung bezeichnet. n gibt also an, wieviele halbe Perioden in das Intervall [0, L] passen. Lösungen zu hohen Werten von n haben eine kürzere Wellenlänge  $\lambda$ . Die Kreisfrequenz  $\omega = ck$  hängt ebenso von n ab,  $e^{i\frac{\pi n}{L}ct} = e^{i\omega t}$  mit  $\omega = \pi cn/L$ . Eine Lösung der Form (6.40) zeichnet sich dadurch aus, dass  $\phi(x,t)$  für alle Punkte x mit der gleichen Frequenz oszillieren, bei konstanter Amplitude und relativer Phase zueinander. Eine solche Bewegung bezeichnet man als **Normalmode** des Systems.

Da die Wellengleichung (6.38) eine lineare homogene Gleichung ist, lösen auch Linearkombinationen von Lösungen mit unterschiedlichem n die Wellengleichung, die allgemeine Lösung der Wellengleichung ist also

$$\phi(x,t) = \sum_{n=1}^{\infty} C_n \sin\left(\frac{\pi xn}{L}\right) e^{i\frac{\pi n}{L}ct}$$
(6.41)

Diese allgemeine Lösung ist nicht mehr von der Form f(x) g(t), unterliegt also nicht der Einschränkung des Separationsansatzes. Die einzelnen Terme in (6.40) sind die Normalmoden der Wellengleichung. Die Koeffizienten  $C_n \equiv |C_n|e^{i\phi}$  geben die Amplitude und Phase der Normalmoden an.

Die Abbildung zeigt links eine Periode der ersten Mode (Zeit läuft von oben nach unten), n = 1, mit  $C_1 = 1$  und  $C_n = 0$  sonst. Die zweite Mode mit doppelt so hoher Kreisfrequenz (Mitte) hat  $C_2 = 1$  und  $C_n = 0$  sonst. Eine sogenannte **Überlagerung** oder **Superposition** der beiden ersten Moden mit  $C_1 = C_2 = 1$  ist rechts gezeigt.



Grundsätzlich müssen die Koeffizienten  $C_n$  der Normalmoden aus den Anfangsbedingungen bestimmt werden. In unserem Beispiel der Wellengleichung ist dies die Auslenkung zum Zeitpunkt t = 0,  $\phi(x, t = 0) \equiv \phi_0(x)$ , und die Geschwindigkeit jedes Punktes des Gummibandes zum Zeitpunkt t = 0,  $\partial \phi / \partial t(x, t = 0) = v_0(x)$ . Die Koeffizienten  $C_n$  müssen also so gewählt werden, dass

$$\sum_{n=1}^{\infty} C_n \sin\left(\frac{\pi xn}{L}\right) = \phi_0(x) . \qquad (6.42)$$

Eine zweite Gleichung erhalten wir, wenn wir die Ableitung von (6.41) nach der Zeit gleich der Geschwindigkeit  $v_0(x)$  setzen. Aber läßt sich überhaupt jede Funktion  $\phi_0(x)$ , die die Randbedingung  $\phi_0(0) = \phi_0(L) = 0$  erfüllt, als eine Summe von (prinzipiell unendliche vielen) Sinustermen schreiben? Dieser Frage werden wir im nächsten Kapitel nachgehen.

# Sieben

## Fourierreihen

Periodische Funktionen lassen sich zerlegen in Sinus- und Kosinusfunktionen unterschiedlicher Frequenzen, lassen sich also als eine Reihe von Sinus- und Kosinusfunktionen darstellen. Wir lernen die Koeffizienten dieser Reihe zu berechnen und interpretieren die Sinus- und Kosinusfunktionen als eine Basis, die den Vektorraum periodischer Funktionen aufspannt.

Viele periodische Funktionen lassen sich als Linearkombination von Sinus- und Kosinusfunktion schreiben, also als eine Reihe von Sinus- und Kosinusfunktionen unterschiedlicher Frequenz. Diese Behauptung wurde 1807 von Joseph Fourier aufgestellt (er meinte sogar, jede periodische Funktion ließe sich als eine solche Reihe schreiben). Vorläufer der sogenannten Fourierreihen finden sich wohl aber schon in den Arbeiten babylonischer Astronomen.

Wir betrachten zunächst eine Funktion f(x) mit Periode  $2\pi$ ,  $f(x+2\pi) = f(x) \forall x$ . (Durch Transformation der Variablen x lassen sich dann alle andere Perioden erreichen.) Die Funktionen  $\cos(nx)$  und  $\sin(nx)$  mit  $n \in \mathbb{N}$  haben alle Periode  $2\pi$ , unterscheiden sich allerdings durch ihre kleinste Periode  $2\pi/n$ , oder durch die Wellenkreiszahl k = n. Der Ansatz

$$f(x) = \frac{1}{2}a_0 + a_1\cos(x) + a_2\cos(2x) + a_3\cos 3x + \dots$$
(7.1)  
+  $b_1\sin(x) + b_2\sin(2x) + b_3\sin(3x) + \dots$ 

ist eine Überlagerung (Summe, Linearkombination) aller Sinus- und Kosinusfunktionen mit kleinster Periode  $2\pi/n$ ,  $n \in \mathbb{N}$ . Sie heißt **Fourierreihe** einer Funktion f(x) mit Periode  $2\pi$ . Die sogenannten **Fourierkoeffizienten**  $\{a_n, b_n\}$  sind noch zu bestimmen. Ebenso werden wir die Frage stellen, wann diese unendliche Reihe konvergiert, und ob ihr Grenzwert in der gesamten Domäne von f(x) mit der Funktion übereinstimmt. **Info:** Vergleichen Sie diesen Ansatz mit der Taylorreihe (5.5). Die Taylorreihe ist die Zerlegung einer Funktion f(x) in unterschiedliche Potenzen von x, die Fourierreihe ist eine Zerlegung in Sinus- und Kosinusterme unterschiedlicher Frequenz.

### Anwendungen

• getriebener harmonische Oszillator mit periodischer Antriebskraft, siehe 6.4.6. Da die DGL des harmonischen Oszillators linear ist, können wir die Lösung für eine beliebige periodische Inhomogenität finden, indem wir Linearkombinationen aus Lösungen für sinusförmige externe Antriebskraft bilden: Aus der Bewegungsgleichung (6.33) für den getriebenen harmonischen Oszillator folgt durch Zerlegung der externen Kraft in Fourierkomponenten

$$\mathcal{L} y(t) = f(t) \qquad \text{mit } \mathcal{L} = \frac{d^2}{dt^2} + b \frac{d}{dt} + c$$
$$= \frac{1}{2}a_0 + a_1 \cos(t) + a_2 \cos(2t) + \dots$$
$$+ b_1 \sin(t) + b_2 \sin(2t) + \dots$$
(7.2)

-0

Mit (6.37) kennen wir die Lösungen der DGL für sinus/kosinusförmige externe Kraft unterschiedlicher Antriebsfrequenz  $\mathcal{L} y_0(t) = 1$ ,  $\mathcal{L} y_1^{(a)} = \cos t$ ,  $\mathcal{L} y_2^{(a)} = \cos 2t$ , etc. Dann ist aufgrund der Linearität von  $\mathcal{L}$  (einsetzen!) die Lösung von  $\mathcal{L} y(t) = f(t)$ 

$$y(t) = \frac{1}{2}a_0y_0(t) + a_1y_1^{(a)}(t) + a_2y_2^{(a)}(t) + \dots + b_1y_1^{(b)}(t) + b_2y_1^{(b)}(t) + \dots$$
(7.3)

Die Lösung der DGL mit allgemeiner (aber periodischer) Antriebskraft ergibt sich also durch Zerlegung der Antriebstkraft in ihre Fourierreihe. Die Lösung setzt sich als Linearkombination aus den Lösungen der DGL für Sinus/Kosinusförmige Antriebskraft zusammen, die Koeffizienten dieser Linearkombination sind die Koeffizienten der Fourierreihe.

- Anfangsbedingung der Wellengleichung zerlegt in die verschiedenen Moden der Gleichung, siehe 6.5.2.
- periodische Funktionen als Elemente eines Vektorraumes mit  $\sin(nx), \cos(nx)$  als Basis

## 7.1 Fourierreihe der Rechteckfunktion

Die Berechnung der Koeffizienten der Fourierreihe lässt sich am besten zunächst an einem konkreten Beispiel zeigen. Wir betrachten die sogenannte Rechteckfunktion

$$f(x) = -1 -\pi < x < 0 f(x) = +1 0 < x < \pi$$

und Periode  $2\pi$ .



Unser Ziel ist die Koeffizienten der Fourierreihe für diese Funktion zu bestimmen

$$f(x) = \frac{1}{2}a_0 + a_1 \cos(x) + a_2 \cos(2x) + \dots + b_1 \sin(x) + b_2 \sin(2x) + \dots$$
(7.4)

### 1. Bestimmung von $a_0$

Wir integrieren rechte und linke Seite von (7.4) über das Intervall  $[-\pi,\pi]$ 

$$\int_{-\pi}^{\pi} \mathrm{d}x \ f(x) = \int_{-\pi}^{\pi} \mathrm{d}x \left[ \frac{1}{2} a_0 + a_1 \cos x + a_2 \cos 2x + \dots + b_1 \sin x + b_2 \sin 2x \dots \right] \,.$$
(7.5)

linke Seite 
$$\int_{-\pi}^{\pi} dx f(x) = 0$$
(7.6)  
rechte Seite 
$$\int_{-\pi}^{\pi} dx \left[ \frac{1}{2} a_0 + a_1 \cos x + a_2 \cos 2x + \dots b_1 \sin x + b_2 \sin 2x \dots \right]$$
$$= \frac{1}{2} 2\pi a_0 + a_1 0 + a_2 0 + \dots b_1 0 + b_2 0 \dots$$
$$= \pi a_0 \qquad \curvearrowright a_0 = \frac{1}{\pi} \int_{-\pi}^{\pi} dx f(x) = 0$$

Dieses Ergebnis ist nicht unerwartet, da $\cos(kx)$ ,  $\sin(kx)$  im Mittel über eine oder mehrere Perioden Null sind, muss  $a_0/2$  gleich dem Mittelwert von f(x) über eine Periode sein.

## **2. Bestimmung von** $a_n, b_n$ n > 0

Wir werden die folgende Integrale benötigen

$$\int_{-\pi}^{\pi} dx \, \sin^2(nx) = \int_{-\pi}^{\pi} dx \, \cos^2(nx) = \frac{1}{2} \, 2\pi = \pi \qquad (\text{da} \, \cos^2(nx) + \sin^2(nx) = 1)$$

$$\int_{-\pi}^{\pi} dx \, \sin(nx) \, \cos(mx) = 0 \qquad n, m > 0$$

$$\int_{-\pi}^{\pi} dx \, \sin(nx) \, \sin(mx) = 0 \qquad n \neq m$$

$$\int_{-\pi}^{\pi} dx \, \cos(nx) \, \cos(mx) = 0 \qquad n \neq m \qquad (7.7)$$

Die letzten 3 Integrale lassen sich mit Hilfe der Eulerformel (5.10) leicht berechnen. Zum Beispiel  $\int_{-\pi}^{\pi} dx \cos(nx) \cos(mx) = \int_{-\pi}^{\pi} dx \frac{1}{2} \left(e^{inx} + e^{-inx}\right) \frac{1}{2} \left(e^{imx} + e^{-imx}\right)$ . Durch Ausmultiplizieren erhält man vier Terme, die alle von der Form sind

$$\int_{-\pi}^{\pi} \mathrm{d}x \ e^{ilx} = \frac{1}{il} \left[ e^{ilx} \right]_{-\pi}^{\pi} = \frac{1}{il} \left( e^{il\pi} - e^{-il\pi} \right) = \frac{2}{l} \sin(l\pi) = 0 \quad \forall l = \pm n \pm m \neq 0.$$
(7.8)

Zur Bestimmung der Koeffizienten der Kosinusterme  $a_n$  multiplizieren wir jetzt rechte und linke Seite von (7.4) mit  $\cos(nx)$  und integrieren wieder über  $[-\pi, \pi]$ 

$$\int_{-\pi}^{\pi} \mathrm{d}x \, \cos(nx) f(x) = \int_{-\pi}^{\pi} \mathrm{d}x \cos(nx) \left[ \frac{1}{2} a_0 + a_1 \, \cos x + a_2 \, \cos 2x + \dots + b_1 \, \sin x + b_2 \, \sin 2x \dots \right] \,. \tag{7.9}$$

Von der linken Seite erhalten wir  $\int_{-\pi}^{\pi} dx f(x) \cos(nx) = 0$ , da (hier) f(x) eine ungerade Funktion von x ist,  $\cos(nx)$  aber gerade ist. Auf der rechten Seite stehen Integrale über  $\cos(nx), \cos(nx)\cos(mx), \text{ und } \cos(nx)\sin(mx)$ . Nach (7.7) integrieren alle diese Terme zu null, bis auf einen einzelnen Term  $\cos(nx)\cos(mx)$  mit n = m, der  $a_n \int_{-\pi}^{\pi} dx \cos^2(nx) = a_n \pi$  ergibt. Damit ist  $a_n$  bestimmt:

$$0 = \int_{-\pi}^{\pi} \mathrm{d}x \cos(nx) f(x) = a_n \pi \qquad \frown a_n = 0 \qquad \forall n > 0 .$$
 (7.10)

Zur Bestimmung der Koeffizienten der Sinusterme  $b_n$  gehen wir analog vor und multiplizieren wir rechte und linke Seite von (7.4) mit  $\sin(nx)$  und integrieren wieder

$$\int_{-\pi}^{\pi} \mathrm{d}x \,\sin(nx)f(x) = \int_{-\pi}^{\pi} \mathrm{d}x\sin(nx) \left[\frac{1}{2}a_0 + a_1\,\cos x + a_2\,\cos 2x + \dots + b_1\,\sin x + b_2\,\sin 2x\dots\right] \,. \tag{7.11}$$

Von der linken Seite erhalten wir

$$\int_{-\pi}^{\pi} dx \, \sin(nx) f(x) = -\int_{-\pi}^{0} dx \, \sin(nx) + \int_{0}^{\pi} dx \, \sin(nx)$$
$$= 2 \int_{0}^{\pi} dx \, \sin(nx) = -2 \left[\frac{1}{n} \cos(nx)\right]_{0}^{\pi}$$
$$= \begin{cases} \frac{4}{n} & n \text{ ungerade} \\ 0 & n \text{ gerade} \end{cases}.$$

Von der rechten Seite integrieren wieder alle Terme zu Null, bis auf den Term  $\sin(nx)\sin(mx)$ mit n = m, er gibt  $b_n \int_{-\pi}^{\pi} dx \sin^2(nx) = \pi b_n$ . Damit ist

$$b_n = \begin{cases} \frac{4}{\pi n} & \text{wenn n ungerade} \\ 0 & \text{wenn l gerade.} \end{cases}$$

somit ist die Fourierreihe der Rechteckfunktion

$$f(x) = \frac{4}{\pi} \left( \sin x + \frac{1}{3} \sin 3x + \frac{1}{5} \sin 5x + \cdots \right)$$
(7.12)

Die ersten 3 und die ersten 6 Terme dieser Reihe sind in blau und grün gezeigt.



### 7.2 Fourierkoeffizienten einer allgemeinen periodische Funktion

Zur Darstellung einer Funktion mit allgemeiner Periode 2L im Intervall [-L, L] gehen wir analog vor. Die Argumente der Sinus- und Kosinusfunktionen der Fourierreihe sind um einen Faktor  $\pi/L$  gestreckt

$$f(x) = \frac{a_0}{2} + a_1 \cos(\frac{\pi x}{L}) + a_2 \cos(\frac{2\pi x}{L}) + a_3 \cos(\frac{3\pi x}{L}) + \dots + b_1 \sin(\frac{\pi x}{L}) + b_2 \sin(\frac{2\pi x}{L}) + \dots ,$$
(7.13)

so dass jetzt das Intervall [L, L] ein ganzes Vielfaches der Periode der Sinus- und Kosinus-funktionen ist.

**Berechnung von**  $a_0$ : Wir multiplizieren (7.13) mit 1 und integrieren über das Intervall [-L, L]. Die Sinus und Kosinusterme integrieren zu Null, es bleibt

$$\int_{-L}^{L} dx f(x) = \left(\frac{1}{2}a_{0}\right) \int_{-L}^{L} dx \, 1 = \frac{1}{2}a_{0} \times 2L = a_{0}L$$

$$a_{0} = \frac{1}{L} \int_{-L}^{L} dx \, f(x) \qquad (7.14)$$

**Berechnung von**  $a_n, b_n$  (n > 0): Wir multiplizieren (7.13) mit  $\cos(\frac{\pi nx}{L})$ , bzw.  $\sin(\frac{\pi nx}{L})$ 

$$\int_{-L}^{L} dx \cos\left(\frac{\pi n}{L}x\right) f(x) = a_n \int_{-L}^{L} dx \cos^2\left(\frac{\pi n}{L}x\right)$$
$$x' \stackrel{=\pi x}{=} a_n \frac{L}{\pi} \int_{-\pi}^{\pi} dx' \cos^2(nx') = a_n \frac{L}{\pi} \frac{1}{2} \times 2\pi = a_n L$$
$$\curvearrowright a_n = \frac{1}{L} \int_{-L}^{L} dx \cos\left(\frac{\pi n}{L}x\right) f(x)$$
$$b_n = \frac{1}{L} \int_{-L}^{L} dx \sin\left(\frac{\pi n}{L}x\right) f(x)$$
(7.15)

#### Intuitiv

Die Formeln für die Koeffizienten der Fourierreihe lassen sich als Maß für die Ähnlichkeit zwischen der Funktion f(x) und den Termen der Fourierreihe interpretieren. Z.B. gibt das Integral zur Bestimmung von  $b_n$  an, wie "ähnlich" f(x) der Funktion sin  $\left(\frac{\pi nx}{L}\right)$  ist; ist f(x)dort positiv, wo auch die Kosinus-Funktion positiv ist ergibt das einen positiven Beitrag zum Integral.



 $\int_{-L}^{L} dx f(x) \sin \frac{\pi xn}{L}$  wird als **Überlapp** der Funktion f(x) und der Funktion  $\sin \frac{\pi xn}{L}$  bezeichnet. Terme der Fourrierreihe mit großem Überlapp mit f(x) haben dann entsprechend

große Koeffizienten. Diese Interpretation der Ergebnisse (7.15) lässt sich im Rahmen des folgenden Bildes noch vertiefen.

#### Interpretation von Funktionen als Elemente eines Vektorraumes

Funktionen lassen sich addieren und mit Zahlen multiplizieren, das Ergebnis ist wieder eine Funktion. Funktionen bilden also einen Vektorraum. Betrachten wir f(x) als Element eines Vektorraumes, der aus Funktionen mit Periode 2L besteht, kann man  $1, \cos \frac{\pi xl}{L}, \sin \frac{\pi xl}{L}$  als seine Basisvektoren auffassen.

$$f(x) = \frac{a_0}{2} + a_1 \cos\left(\frac{\pi x}{L}\right) + a_2 \cos\left(\frac{2\pi x}{L}\right) + \cdots + b_1 \sin\left(\frac{\pi x}{L}\right) + b_2 \sin\left(\frac{2\pi x}{L}\right) + \cdots$$
  
Vergl.  $\vec{v} = v_1 \vec{e}_1 + v_2 \vec{e}_2 + v_3 \vec{e}_3 \cdots$  (7.16)

$$a_n = \frac{1}{L} \int_{-L}^{L} dx f(x) \cos\left(\frac{\pi n}{L}x\right)$$
  

$$b_n = \frac{1}{L} \int_{-L}^{L} dx f(x) \sin\left(\frac{\pi n}{L}x\right)$$
  
Vergl.  $v_i = \vec{e}_i \cdot \vec{v}$ 
(7.17)

Die auftretenden Integrale können als Verallgemeinerung des euklidischen Skalarproduktes  $\vec{v} \cdot \vec{w} = v_1 w_1 + v_2 w_2 + v_3 w_3 + \cdots = \sum_i v_i w_i$  mit  $\sum_i \longrightarrow \frac{1}{L} \int_{-L}^{L} dx$  aufgefasst werden. Dabei ist es hilfreich eine (beliebige) Funktion f(x) als Vektor  $\mathbf{f}$  mit Komponenten  $f(x_1), f(x_2), f(x_3), \ldots$  zu denken. Je enger die Stützstellen  $x_1, x_2, x_3 \ldots$  liegen, desto besser beschreibt dieser Vektor die Funktion f(x). Das euklidische Skalarprodukt zweier solcher Vektoren  $\mathbf{f}$  und  $\mathbf{g}$  ist  $f(x_1)g(x_1) + f(x_2)g(x_2) + f(x_3)g(x_3) + \ldots$  Multipliziert man mit dem Abstand der Stützstellen  $\Delta x$  erhält man im Grenzfall  $\Delta x \to 0$  das Integral  $\int dx f(x)g(x)$ .



Die Integrale (7.7) zeigen außerdem, dass die Kosinus- und Sinusfunktionen unterschiedlicher Wellenzahl n nach diesem Skalarprodukt orthogonal sind. (7.17) kann dann verstanden

werden als Bestimmung der Koeffizienten der Basisvektoren  $\cos\left(\frac{\pi n}{L}x\right), \sin\left(\frac{\pi n}{L}x\right)$  durch Berechnung des Skalarproduktes zwischen f(x) und  $\cos\left(\frac{\pi n}{L}x\right)$  bezw.  $\sin\left(\frac{\pi n}{L}x\right)$ , analog zu (1.14).

### 7.2.1 Konvergenz der Fourierreihe: Die Dirichlet Bedingung

Die Fourierreihe hatten wir als einen Ansatz konstruiert, (7.1) enthält alle Terme die mit der Periode einer Funktion kompatibel sind. Aber lässt sich wirklich jede periodische Funktion als Fourierreihe darstellen? Wann und wo ist eine allgemeine periodische Funktion f(x) gleich ihrer Fourierreihe?

Die allgemeinste Anwort gibt folgendes Theorem: Sei f(x) im Intervall [-L, L] quadratintegrabel, d.h.  $\frac{1}{L} \int_{-L}^{L} dx |f(x)|^2$  ist endlich, dann ist an fast allen Punkten x f(x) gleich seiner Fourierreihe.

Von Lennart Carleson 1966 (!) bewiesen, gilt dieser Satz als einer der Höhepunkte der Analysis des letzten Jahrhunderts. An fast allen Punkten bedeutet (grob) überall, außer möglicherweise an einzelnen isolierten Punkten. Welche einzelnen Punkte problematisch sein können, sagt die Dirichlet- Bedingung (1829), die hier ebenso ohne Beweis zitiert wird: Ist f(x) periodisch mit Periode 2L, absolut integrierbar, d.h.  $\frac{1}{L} \int_{-L}^{L} dx |f(x)|$  ist endlich, nach oben und unten beschränkt und lässt sich in endlich viele Intervalle zerlegen, in denen f(x) stetig und monoton ist und existiert an den Intervallgrenzen rechter und linker Grenzwert  $f_+$  und  $f_-$ , dann konvergiert die Fourierreihe von f(x) punktweise gegen

f(x)	wenn $f$ in $x$ stetig ist
$(f_{-}+f_{+})/2$	wenn $f$ in $x$ nicht stetig ist.

Wenn f(x) in einer Periode also eine *endliche* Zahl von Maxima, Minima oder Sprüngen aufweist, konvergiert die Fourierreihe. An einer Sprungstelle konvergiert die Fourierreihe zum arithmetischen Mittel der Funktion rechts und links der Sprungstelle. *Definiert* man den Wert von f(x) am Punkt x des Sprunges z.B. gleich dem rechten Grenzwert, konvergiert die Fourierreihe dennoch zu  $(f_- + f_+)/2$ . Dann ist f(x) an fast allen Punkten gleich seiner Fourierreihe, nämlich überall außer am Punkt des Sprunges x.

Das Gegenteil des Satzes von Dirichlet gilt übrigens nicht; es gibt Funktionen die die Dirichlet-Bedingung verletzen und dennoch konvergente Fourierreihen haben.

### 7.2.2 Das Gibbs-Phänomen

Approximiert man f(x) durch eine endliche Fourierreihe, bricht also den Ansatz (7.1) nach einer endlichen Zahl von Termen ab, beobachtet man in der Nähe von Sprüngen deutliche Abweichungen zwischen f(x) und der endlichen Reihe. Die Rechteckfunktion ist ein gutes Beispiel, die Abbildung zeigt die Summer der ersten 20 und 50 Terme der Reihe (7.12) Die Abbildung zeigt die ersten 5 Terme (grün), 12 Terme (rot) und 25 Terme (blau) der Fourierreihe der Recheckfunktion (7.12) in der Nähe des Sprungs bei x = 0. Wir bebachten ein Überschwingen der endlichen Fourierreihe; in der Nähe von Sprüngen sind die Ausschläge der endlichen Reihe grösser als der eigentliche Sprung. Das liegt daran, dass bei *endlichen* Frequenzen  $k = n\pi/L$ eine große Steigung in einem bestimmten Bereich von x einhergeht mit einer großen Amplitude. Die endliche Fourierreihe erkauft also den korrekten Wert der Steigung am Sprung mit einer Abweichung zwischen der endlichen Reihe und f(x) kurz vor und hinter dem Sprung selbst.



Die Überschwing-Breite nimmt mit zunehmendender Zahl der Terme in der Fourierreihe ab (geht gegen Null), die Überschwing-Amplitude bleibt jedoch endlich.

**Info:** Nutzt man die Fourierreihe zur Darstellung eines Bildes (f(x, y)) beschreibt z.B. Grauwerte), führt das Gibbs-Phänomen in der Nähe von scharfen Kontrastgrenzen zu sogenannten "ringing artefacts"

(http://en.wikipedia.org/wiki/File:Asterisk\_with\_jpg-artefacts.png). Ein Analogon in der Klangverarbeitung sind sogenannte "pre-echos", die als Artefakte bei Aufnahmen von Perkussionsinstrumenten in MP3-Dateien auftreten, ein Beispiel finden sie auf http://en.wikipedia.org/wiki/Pre-echo.

Wer selbst mit der Fourierreihe experimentieren möchte, sei auf Wolfram Alpha (www.wolframalpha.com) verwiesen. Der Befehl 4/Pi Sum[1/n Sin[n x],n,1,9,2] in das Eingabefenster eingegeben zeigt das Ergebnis der Reihe (7.12) bis zum Term  $1/9\sin(9x)$ .

#### 7.2.3 Die komplexe Fourierreihe

Die Euler-Formel (5.10) führt auf eine einfache und kompakte Form der Fourierreihe. Wir beschränken uns wieder auf Funktionen mit Periode  $2\pi$ . Aus der Euler-Formel  $e^{ix}$  =  $\cos(x) + i\sin(x)$  erhalten wir  $\cos(nx) = \frac{1}{2}(e^{inx} + e^{-inx})$  und  $\sin(nx) = \frac{1}{2}(e^{inx} - e^{-inx})$ . Setzt man diese Ausdrücke in die Fourierreihe (7.1) ein erhält man

$$f(x) = c_0 + c_1 e^{ix} + c_{-1} e^{-ix} + c_2 e^{2ix} + c_{-2} e^{-2ix} + \dots$$
(7.18)

mit  $\frac{1}{2}a_0 = c_0$ ,  $\frac{a_n}{2} + \frac{b_n}{2i} = c_n$  und  $\frac{a_n}{2} - \frac{b_n}{2i} = c_{-n}$  für n > 0. Wenn f(x) reell ist, erhalten wir ausserdem  $c_{-n} = c_n^*$ , denn dann ist  $c_n e^{ixn} + c_{-n} e^{-inx} = c_n e^{ixn} + c_n^* e^{-inx} \in \mathbb{R}$ , d.h. der Beitrag der Terme n und -n zur Fourierreihe ist eine reelle Zahl.

Die Koeffzienten  $c_n$  können auch direkt aus der Funktion f(x) bestimmt werden, die Berechnung folgt dem Schema von (7.15), ist allerdings erheblich einfacher.

Berechung von  $c_0$ :

$$\int_{-\pi}^{\pi} dx f(x) = c_0 \int_{-\pi}^{\pi} dx 1 = 2\pi c_0$$
(7.19)
$$\sim c_0 = \frac{1}{2\pi} \int_{-\pi}^{\pi} dx f(x) .$$

Berechung von  $c_n, n \neq 0$ : Wir multiplizeren (7.18) mit  $e^{-inx}$  und integrieren über das Intervall  $[-\pi, \pi]$ 

$$\int dx f(x) e^{-inx} = \int_{-\pi}^{\pi} dx e^{-inx} \left[ c_0 + c_1 e^{ix} + c_{-1} e^{-ix} + c_2 e^{2ix} + c_{-2} e^{-2ix} + \dots \right] = c_n \int_{-\pi}^{\pi} dx = 2\pi c_n$$

$$(7.20)$$

Damit ist die komplexe Fourierreihe

$$f(x) = \sum_{n=-\infty}^{\infty} c_n e^{ixn}$$

$$c_n = \frac{1}{2\pi} \int_{-\pi}^{\pi} \mathrm{d}x \ e^{-inx} f(x) \ . \tag{7.21}$$

Die Interpretation des Integrals (7.21) als Skalarprodukt beinhaltet hier (Minuszeichen in  $e^{-inx}$ !) für komplexe Funktionen

$$\langle f,g\rangle \equiv \int_{-\pi}^{\pi} \mathrm{d}x f^*(x)g(x) \ . \tag{7.22}$$

Diese Definition des Skalarproduktes für komplexe Argumente stellt sicher, dass das Quadrat der Norm eines komplexwertigen Vektors (hier einer Funktion) stets nicht-negativ ist,  $\langle f, f \rangle = \int_{-\pi}^{\pi} |f(x)|^2$ , und dass der einzige Vektor mit Norm Null die Nullfunktion  $f(x) = 0 \forall x$  ist.

### 7.2.4 Parsevals Theorem

verbindet  $\frac{1}{2\pi} \int_{-\pi}^{\pi} dx |f(x)|^2$  mit dem Betragsquadraten der Fourierreihe von f(x). Wir gehen aus von der komplexen Form der Fourierreihe,

$$f(x) = \sum_{n = -\infty}^{\infty} c_n e^{inx} ,$$

bilden von beiden Seiten das Betragsquadrat und integrieren über eine Periode:

links: 
$$\int_{-\pi}^{\pi} dx |f(x)|^{2}$$
  
rechts: 
$$\int_{-\pi}^{\pi} dx \left(\sum_{n=-\infty}^{\infty} c_{n} e^{inx}\right) \left(\sum_{m=-\infty}^{\infty} c_{m}^{*} e^{-imx}\right)$$
$$= \sum_{n,m=-\infty}^{\infty} c_{n} c_{m}^{*} \int_{-\pi}^{\pi} dx e^{inx-imx}$$
$$= \sum_{n=-\infty}^{\infty} |c_{n}|^{2} 2\pi ,$$

wobei wir im letzten Schritt wieder  $\int_{-\pi}^{\pi} dx \, e^{imx} \stackrel{m \neq 0}{=} 0$  genutzt haben, s. (7.8). Setzen wir linke und rechte Seite gleich, erhalten wir **Parsevals Theorem** 

$$\int_{-\pi}^{\pi} \mathrm{d}x |f(x)|^2 = 2\pi \sum_{n=-\infty}^{\infty} |c_n|^2$$
(7.23)

Mit der Interpretation von f(x) als Vektor, und dem Skalarprodukt zwischen zwei Vektoren (7.22) ist die linke Seite von Parsevals Theorem das Quadrat der Norm von f(x), die rechte Seite ist die Summe der Quadrate der Komponenten von f(x) in einer quadratischen Basis, analog zu  $||\mathbf{v}||^2 = v_1^2 + v_2^2 + \dots$ 

### 7.3 Ausblick: die Fouriertransformation

Dieser Ausblick (nur wenig redigiert!) war nicht Teil der Vorlesung. Dennoch haben Sie jetzt die nötigen Grundlagen, um den nächsten Schritt nach der Fourier-Reihe zu nehmen. Daher für alle, die wissen wollen wie es weitergeht:

Ausgangspunkt der Fourierreihe (7.1) war eine Überlagerung aus allen trigonometrischen Funktion die im Intervall  $2\pi$ , bezw. L periodisch sind. Das Ergebnis waren Fourierkoeffzienten  $a_n$  und  $b_n$  (bezw.  $c_n$  für die komplexe Reihe). Diese Koeffizienten geben an, wie stark eine Oszillation mit Kreisfrequenz  $k = n\pi/L$  ( $n \in \mathbb{N}$ ) in der Fourierreihe vertreten ist.

Natürlich sind nicht alle interessanten oder relevanten Funktionen periodisch und wiederholen sich nach einem endlichen Intervall 2L. Betrachtet man die Fourierreihe

$$f(x) = \sum_{n = -\infty}^{\infty} c_n e^{inx\frac{\pi}{L}}$$
(7.24)

einer Funktion mit großer Periode 2L, dann liegen die Kreisfrequenzen  $k = n\pi/L$  umso dichter, je grösser die Periode ist. Um nicht-periodische Funktion zu beschreiben, machen wir einen Ansatz analog zur Fourierreihe (7.24), bei dem die Kreisfrequenzen allerdings dicht im Frequenzraum liegen

$$f(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \mathrm{d}k \ c(k) \ e^{ikx} \ .$$
 (7.25)

Dieser Ansatz kann als Überlagerung der  $e^{ikx}$  für beliebige Kreisfrequenzen k verstanden werden. Die Funktion c(k) gibt an, wie stark eine Kreisfrequenz k in (7.25) vertreten ist. Im Gegensatz zur Fourierreihe einer periodischen Funktion gibt es im Allgemeinen aber keine niedrigste Frequenz  $k = \pi/L$  unterhalb der c(k) verschwindet, und die Kreisfrequenzen k in (7.25) sind im Allgemeinen auch keine Multiplen einer solchen niedrigsten Frequenz.  $\tilde{f}(k) \equiv c(k)$  heißt **Fouriertransformierte** von f(x). Wie bei der Fourierreihe ist noch zu klären, für welche Funktionen f(x) eine solche Fouriertransformierte  $\tilde{f}(k)$  existiert.

### Das Gaußsche Wellenpaket

Um den Zusammenhang zwischen einer Funktion und ihrer Fouriertransformierten (FT) zu verstehen, betrachten wir zunächst ein konkretes Beispiel

$$c(k) = \tilde{f}(k) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2}(k-k_0)^2/\sigma^2}$$
(7.26)

Diese sogenannte **Gaußsche Normalverteilung** beschreibt eine Verteilung der unterschiedlichen Wellenzahlen k mit Mittelwert  $k_0$ . Der Parameter  $\sigma$  gibt an, wie breit die Verteilung ist, bei kleinen Werten von  $\sigma$  ist c(k) stark um  $k_0$  herum lokalisiert, d.h. bei Werten von k weitab von  $k_0$  ist c(k) klein. In diesem Beispiel leisten Oszillation mit Wellenzahl  $k = k_0$  den größten Beitrag zur Funktion f(x).



Ziel ist es jetzt die Funktion f(x) zu dieser konkreten Fouriertransformierten  $\tilde{f}(k) = c(k)$ 

zu berechnen. Mit (7.25) ist

$$f(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \mathrm{d}k \; e^{ikx} \; \tilde{f}(k) \tag{7.27}$$

$$= \frac{1}{2\pi\sigma} \int_{-\infty}^{\infty} \mathrm{d}k \, \exp\left\{ikx - \frac{1}{2\sigma^2}(k-k_0)^2\right\} \,.$$
(7.28)

Das Argument der Exponentialfunktion lässt sich schreiben als

$$ikx - \frac{1}{2\sigma^2}(k - k_0)^2 = -\frac{1}{2\sigma^2}(k^2 - 2k_0k + k_0^2 - 2i\sigma^2kx)$$
  
=  $-\frac{1}{2\sigma^2}(k^2 - 2k[k_0 + i\sigma^2x] + [k_0 + i\sigma^2x]^2 - [k_0 + i\sigma^2x]^2 + k_0^2)$   
=  $-\frac{1}{2\sigma^2}(k - [k_0 + i\sigma^2x])^2 + ik_0x - \frac{1}{2}\sigma^2x^2$ 

Mit dieser sogenannten quadratischen Ergänzung ist

$$f(x) = \frac{1}{2\pi\sigma} \int_{-\infty}^{\infty} \mathrm{d}k \, \exp\left\{-\frac{1}{2\sigma^2} (k - [k_0 + i\sigma x])^2\right\} \times e^{ik_0 x - \frac{1}{2}\sigma^2 x^2} \tag{7.29}$$

$$= \frac{1}{2\pi} e^{ik_0 x - \frac{1}{2}\sigma^2 x^2} \int_{-\infty}^{\infty} \mathrm{d}k' \ e^{-\frac{1}{2}k'^2} , \qquad (7.30)$$

wobei wir im letzten Schritt die Variablentransformation  $k' = (k - [k_0 + i\sigma x])/\sigma$  durchgeführt haben.

Info: Der letzte Schritt ist die Auswertung des Integrals über k'. Integrale über die Gaußsche Normalverteilung lassen sich mithilfe des folgenden Tricks berechnen: wir berechnen zunächst das Quadrat des Integrals, also

$$\left(\int_{-\infty}^{\infty} \mathrm{d}x \ e^{-\frac{1}{2}x^2}\right)^2 = \iint_{-\infty}^{\infty} \mathrm{d}x \ dy \ e^{-\frac{1}{2}(x^2+y^2)} = \int_{0}^{\infty} \mathrm{d}r \ \int_{0}^{2\pi} \mathrm{d}\phi \ r \ e^{-\frac{1}{2}r^2} = 2\pi [-e^{-\frac{1}{2}r^2}]_{0}^{\infty} = 2\langle \overline{a} \rangle.31$$

Im dritten Schritt haben wir zu Polarkoordinaten  $x = r \cos \phi, y = r \sin \phi$  gewechselt. In kartesischen Koordinaten ist der Flächeninhalt eines kleinen Rechtecks mit Kantenlängen  $\Delta x$  und  $\Delta y$  das Produkt  $\Delta x \Delta y$ . Das durch eine kleine Änderung des Winkels  $\phi$ und des Radius r definierte (näherungsweise) Rechteck hat die Fläche  $r\Delta\phi$   $\Delta r$ , woraus sich der zusätzliche Faktor r im Integral ergibt. Damit ist  $\int_{-\infty}^{\infty} dx \ e^{-x^2/2} = \sqrt{2\pi}$  und  $\frac{1}{\sqrt{2\pi\sigma^2}} \int dx \ ; e^{-x^2/(2\sigma^2)} = 1.$ Nach Auswertung des Integrals über k' in (7.29) erhalten wir

$$f(x) = \frac{1}{\sqrt{2\pi}} \exp\left\{-\frac{1}{2}\sigma^2 x^2 + ik_0 x\right\}$$
(7.32)

Diese Überlagerung f(x) von Oszillationen unterschiedlicher Frequenzen, bei denen c(k)der Gaußverteilung folgt, heiß Gaußsches Wellenpaket. Für  $k_0 = 0$  ist f(x) auch wieder eine Gauß-Kurve, mit Maximum bei x = 0. Allerdings ist die Breite nun durch  $\sigma^{-1}$  gegeben; je breiter die Verteilung der Wellenzahlen c(k), desto schmaler ist also die Funktion f(x), und umgekehrt. Bei endlichem  $k_0$  oszilliert f(x) mit Wellenzahl  $k_0$ , die Amplitude dieser Oszillation fällt mit  $e^{\sigma^2 x^2/2}$  ab. Im Fall  $\sigma = 0$  beschreibt f(x) ungedämpfte Oszillationen mit Wellenzahl  $k_0$  (so wie wir es für ein scharf bei  $k_0$  lokalisertes c(k) auch erwarten).

#### Die $\delta$ -Funktion

Im Fall von großen Werten von  $\sigma$  ist c(k) eine breite Verteilung, f(x) in (7.27) enhält also Beiträge von vielen unterschiedlichen Wellenzahlen k. Wir betrachten den Fall großer Werte von  $\sigma$  für  $\tilde{f}(k) = \frac{1}{\sqrt{2\pi}}e^{-\frac{k^2}{2\sigma^2}}$ 

$$f(x) = \lim_{\sigma \to \infty} \frac{1}{2\pi} \int dk \ e^{ikx - \frac{1}{2\sigma^2}k^2} = \lim_{\sigma \to \infty} \sqrt{\frac{\sigma^2}{2\pi}} e^{-\frac{1}{2}\sigma^2 x^2} - \delta(x)^{"}$$
(7.33)

 $\delta(x)$  ist eine "Funktion", die überall Null ist, bis auf x = 0. Dennoch ist das Integral von  $\sqrt{\frac{\sigma^2}{2\pi}}e^{-\frac{1}{2}\sigma^2x^2}$  für beliebige  $\sigma$  gleich 1 und damit  $\int dx \, \delta(x) = 1$ .

 $\delta(x)$  wird als **Delta-Funktion** bezeichnet, obwohl  $\delta(x)$  bei x = 0 unendlich ist (und sonst überall null) und damit keine wohldefinierte Funktion ist.  $\delta(x)$  wird verwendet um punktförmige Objekte zu beschreiben, z.B. die Masseverteilung einer Punktmasse bei x =0, oder die Ladungsverteilung einer Punktladung. Deren Masse- oder Ladungsdichte ist überall Null, bis auf einen Punkt, ebenso ist das Integral über die Dichte endlich. Wir können die  $\delta$ -Funktion als ein Objekt auffassen, das nur innerhalb eines Integrals Sinn macht (bei einzelnen Rechenschritten mag das Integral dann unterschlagen werden, so dass man in der Praxis mit  $\delta(x)$  wie mit einer Funktion rechnet). Zentrale Eigenschaft ist  $\delta(x) = 0$  überall ausser bei x = 0 und  $\int dx$ ;  $\delta(x) = 1$ , und damit für eine Funktion g(x) $\int dx$ ;  $\delta(x)g(x) = \int dx$ ;  $\delta(x)[g(0) + xg'(x) + ...] = g(0)$ 

**Info:** Der mathematisch korrekte Zugang basiert auf sogenannten **Funktionalen**, Abbildungen von Funktionen auf Zahlen. Ein Beispiel für ein Funktional ist die Abbildung von der Funktion f(x) auf die Zahl  $\int dx f(x)g(x)$ . Eine Funktion g(x) definiert dadurch ein Funktional. Das  $\delta$ -Funktional ist die Abbildung von f(x) auf die Zahl f(0).

### 7.3.1 Fouriertransformation und inverse Fouriertransformation

Wir nutzen

$$\frac{1}{2\pi} \int_{-\infty}^{\infty} dk \ e^{ik(x-x_0)} = \frac{1}{2\pi} \lim_{\sigma \to 0} \frac{1}{2\pi} \int_{-\infty}^{\infty} dk \ e^{-\frac{1}{2}\sigma^2 k^2 + ik(x-x_0)}$$
$$= \lim_{\sigma \to \infty} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left\{-\frac{1}{2\sigma^2}(x-x_0)^2\right\}$$

als Integraldarstellung von  $\delta(x - x_0)$ , also einer  $\delta$ -Funktion, die überall Null ist bis auf  $x_0$ . Damit zeigen wir nun, dass  $\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \mathrm{d}x \; e^{-ikx} \; f(x)$  die Fouriertransformierte von f(x) ist

$$\begin{aligned} \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \mathrm{d}k \; e^{ikx} \; \tilde{f}(k) \; &= \; \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \mathrm{d}k \; e^{ikx} \; \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \mathrm{d}x \; e^{ikx'} \; f(x') \\ &= \; \int_{-\infty}^{\infty} \mathrm{d}x' \; f(x') \; \underbrace{\frac{1}{2\pi} \int \mathrm{d}k \; e^{ik(x-x')}}_{\delta(x-x')} \\ &= \; \int_{-\infty}^{\infty} \mathrm{d}x' \; f(x') \; \delta(x-x') = f(x) \; . \end{aligned}$$

Die Fourier-Transformation von f(x) zu  $\tilde{f}(k)$  und die inverse Fouriertransformation von  $\tilde{f}(k)$  zu f(x) zeichnen sich also durch eine hohes Maß an Symmetrie aus

$$\tilde{f}(k) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} dx \ e^{-ikx} \ f(x)$$
Fourier-Transformation  
$$f(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} dk \ e^{ikx} \ \tilde{f}(k)$$
inverse Fourier-Transformation . (7.34)

Diese Ausdrücke sind das Analogon zur komplexen Fourierreihe (7.21). In der Tat ist mit (7.34) die Fouriertransformierte von  $e^{ik'x}$  die  $\delta$ -Funktion  $\delta(k - k')$ , und die Fouriertransformierte einer periodischen Funktion eine Reihe von  $\delta$ -förmigen Maxima bei  $\pi/L, 2\pi/L, 3\pi/L, \ldots$ 

Durch die Definition (7.25) treten in beiden Ausdrücken an gleicher Stelle Faktoren von  $(2\pi)^{-1/2}$  auf. (Vorsicht, manche Texte definieren  $(2\pi)^{-1/2}\tilde{f}(k)$  als die Fouriertransformierte von f(x).)

### 7.3.2 Anwendungen

Kern aller (?) Anwendungen der Fourier-Transformation in der Physik ist die Eigenschaft, dass die Fourier-Transformation von  $\frac{df}{dx} = f'(x)$  gleich  $ik \ \tilde{f}(k)$  ist.

$$\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \mathrm{d}x \, \left(\frac{d}{\mathrm{d}x} f(x)\right) \, e^{-ikx} = \frac{1}{\sqrt{2\pi}} \left[f(x) \, e^{-ikx}\right]_{-\infty}^{\infty} - \frac{1}{\sqrt{2\pi}} \int \, \mathrm{d}x \, f(x) \, \frac{d}{\mathrm{d}x} \, e^{-ikx} \\ \hookrightarrow 0, \, \text{wenn } f(x) \text{ quadratintegrabel}$$

$$= \frac{ik}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \mathrm{d}x \ f(x) \ e^{-ikx} = ik \ \tilde{f}(k)$$

Damit ist auch die FT der *n*-ten Ableitung von f(x) gleich  $(ik)^n \tilde{f}(k)$  Ableitungen von f(x) sind also rein algebraische Operationen der Fouriertransformierten. Lineare Differentialgleichungen lassen sich also durch FT leicht lösen

Beispiel: Wir lösen die DGL

$$f'' + b f' + c f(x) = g(x)$$
(7.35)

für eine allgemeine Funktion g(x). Mit  $x \longrightarrow t$  erhalten wir die Bewegungsgleichung des harmonischen Oszillators mit zeitabhhängiger Antriebskraft g(x). Für  $g(x) = e^{-Kx}$  hatten wir in Sektion 6.4.6 diese Gleichung bereits gelöst.

Fourier-Transformation von rechter und linker Seite dieser Gleichung ergibt

$$-k^2 \tilde{f}(k) + ikb \tilde{f}(k) + c \tilde{f}(k) = \tilde{g}(k) \qquad \curvearrowright \tilde{f}(k) = \frac{\tilde{g}(k)}{c + ikb - k^2}$$
(7.36)

Die inverse Fourier-Transformation von  $\tilde{f}(k)$  ergibt nun die gesuchte Lösung f(x) der Differentialgleichung.

Den Nenner  $c+ikb-k^2$  in diesem Ergebnis haben wir bereits in der Lösung des getriebenen harmonischen Oszillators (6.37) gesehen. Das ist kein Zufall, die Fouriertransformation zerlegt g(x) in eine Kombination von von Oszillationen mit Wellenzahl k, für jede einzelne Wellenzahl löst (7.36) die Bewegungsgleichung, die inverse Fourier-Transformation von  $\frac{\tilde{g}(k)}{c+ikb-k^2}$  setzt die Beiträge von jeder Oszillationsfrequenz zur speziellen Lösung der Bewegungsgleichung für ein gegebenes g(x) zusammen.