Andreas Schadschneider

# Mathematische Methoden der Physik

Version: 6. Februar 2009

Wintersemester 2008/09

# Vorbemerkungen

Das vorliegende Skript zu Vorlesung *Mathematische Methoden* ersetzt nicht den regelmässigen Besuch der Vorlesungen. Es ist als Ergänzung gedacht, zum Nacharbeiten oder zur Vorbereitung auf Klausuren und Prüfungen. Deshalb sollten alle Formeln und Aussagen immer kritisch betrachtet werden, es könnten noch Druckfehler enthalten sein!
Wesentlicher Bestandteil der Vorlesung *Mathematische Methoden* sind die Übungen. Gerade in den ersten Semestern ist es unbedingt erforderlich, den Stoff durch eigenständiges Bearbeiten von Übungsaufgaben zu vertiefen.
Ziel der Vorlesung und der dazu gehörigen Übungen ist die Vermittlung grundlegender mathematischer Techniken und Fähigkeiten, die zur Lösung physikalischer Aufgabenstellungen benötigt werden. Im ersten Teil werden die mathematischen Techniken und Methoden eingeführt, die zum Verständnis der Vorlesungen “Experimentalphysik I+II” notwendig sind. Schwerpunkt des zweiten Blocks ist die Vorbereitung auf die Kursvorlesungen der Theoretischen Physik. Die meisten der hier vorgestellten Methoden sind nicht Gegenstand der Mathematikvorlesung.
Wesentliches Ziel der Vorlesung ist die praktische Beherrschung der vorgestellten Verfahren. Notwendigerweise bleibt im Vergleich zu einer ‘echten’ Mathematikvorlesung die mathematische Strenge etwas auf der Strecke. Es werden daher höchstens Beweisideen oder Beweisskizzen präsentiert. Außerdem soll speziell auf den besonderen Umgang der Physiker mit der Mathematik hingewiesen werden.

Für Fehlermeldungen und Verbesserungsvorschläge bin ich jederzeit dankbar. Sie können auch per email an mich (`as@thp.uni-koeln.de`) geschickt werden. Die jeweils aktuellste Version des Skripts ist im Internet über meine Homepage

`http://www.thp.uni-koeln.de/~as`

verfügbar.

Andreas Schadschneider

2

# Literaturempfehlungen

Im folgenden finden Sie eine kommentierte Auswahl der populärsten Lehrbücher. Die Vorlesung orientiert sich nicht speziell an einem Buch. Ich empfehle Ihnen deshalb, sich vor einem eventuellen Kauf zunächst die einzelnen Werke gründlich anzusehen. Alle sind in der Studentenbibliothek vorhanden.

- S. Großmann: *Mathematischer Einführungskurs für die Physik* (Teubner-Verlag)

  Preiswerte Einführung in die wichtigsten mathematischen Techniken. Kann während des gesamten Studiums verwendet werden, insbesondere als Nachschlagewerk. Es werden allerdings nicht alle Themen der Vorlesung behandelt.
  Preis: ca. 30 Euro

- C.B. Lang, N. Pucker: *Mathematische Methoden in der Physik* (Spektrum)

  Ein neueres Lehrbuch, das alle Themen der Vorlesung umfasst. Enthält auch zahlreiche Übungsaufgaben mit Lösungen.
  Preis: ca. 45 Euro

- K. F. Riley, M. P. Hobson, S. J. Bence: *Mathematical Methods for Physics and Engineering* (Cambridge University Press)

  Ein englischsprachiges Lehrbuch, das sehr ausführlich ist (mehr als 1300 Seiten!). Ergänzend finden sich zahlreiche Übungsaufgaben zum Selbststudium.

- H. Fischer, H. Kaul: *Mathematik für Physiker* (Teubner), 3 Bände
  Eine dreibändiger Klassiker, der den Stoff der Vorlesung weitgehend abdeckt und darüberhinaus weitere Themen sehr detailliert diskutiert. Ist auch für die Mathematik-Vorlesung nützlich.
  Preis: ca. 40 Euro pro Band

- T. Arens, F. Hettlich, C. Karpfinger, U. Kockelkorn, K. Lichtenegger, H. Stachel: *Mathematik* (Spektrum Verlag)

  Eine Empfehlung der Studierenden! Sehr umfangreiches Mathematik-Buch, das sich aber auch um die Anschauung bemüht und daher sehr gut zur Begleitung der Vorlesung geeignet ist.
  Preis: ca. 70 Euro

# Inhaltsverzeichnis


**1 Vektoren** **7**
- 1.1 Vektorräume . . . 8
- 1.2 Vektoralgebra . . . 9
- 1.3 Basis eines Vektorraumes . . . 11
- 1.4 Koordinatensysteme . . . 12
  - 1.4.1 Kartesische Koordinaten . . . 12
  - 1.4.2 (Ebene) Polarkoordinaten . . . 13
  - 1.4.3 Zylinderkoordinaten . . . 13
  - 1.4.4 Kugelkoordinaten (sphärische Polarkoordinaten) . . . 14

**2 Matrizen und Lineare Gleichungssysteme** **17**
- 2.1 Lineare Abbildungen . . . 17
- 2.2 Matrizen . . . 18
- 2.3 Determinanten . . . 21
- 2.4 Lineare Gleichungssysteme . . . 23
  - 2.4.1 Gauß-Algorithmus . . . 24

**3 Funktionen und Differentiation** **27**
- 3.1 Eigenschaften von Funktionen . . . 27
- 3.2 Elementare Funktionen . . . 30
  - 3.2.1 Potenzfunktion . . . 30
  - 3.2.2 Exponentialfunktion . . . 31
  - 3.2.3 Logarithmus . . . 32
  - 3.2.4 Trigonometrische Funktionen . . . 34
- 3.3 Differentiation . . . 39
- 3.4 Potenzreihen und Taylor-Entwicklung . . . 41
  - 3.4.1 Potenzreihen . . . 41
  - 3.4.2 Taylor-Entwicklung . . . 42

**4 Integration** **45**
- 4.1 Stammfunktion . . . 45
- 4.2 Bestimmtes Integral . . . 46

4.3 Integrationsverfahren . . . 48
4.3.1 Partielle Integration . . . 49
4.3.2 Substitutionsregel . . . 50
4.3.3 Ableitung nach Parameter . . . 51
4.4 Uneigentliche Integrale . . . 52
4.4.1 Integration über ein unbeschränktes Intervall . . . 53
4.4.2 Integration über Polstellen . . . 54

**5 Komplexe Zahlen 55**
5.1 Definition und Rechenregeln . . . 55
5.2 Komplexe Ebene . . . 56
5.3 Eulersche Formel . . . 58
5.4 Wurzeln . . . 59

**6 Differentialgleichungen 61**
6.1 Klassifikation von DGL . . . 61
6.2 Lineare DGL mit konstanten Koeffizienten . . . 62
6.2.1 Homogener Fall . . . 62
6.2.2 Inhomogener Fall . . . 65
6.3 Allgemeine lineare DGL 1. Ordnung . . . 66
6.3.1 Homogener Fall . . . 66
6.3.2 Inhomogener Fall . . . 67
6.4 Nichtlineare DGL . . . 68
6.4.1 DGL mit getrennten Variablen . . . 68
6.4.2 Bernoulli-Gleichung . . . 69

**7 Vektoranalysis 71**
7.1 Partielle Ableitung und totales Differential . . . 72
7.2 Gradient . . . 75
7.3 Divergenz . . . 77
7.4 Rotation . . . 78

**8 Wegintegral und mehrdimensionale Integration 81**
8.1 Wegintegrale . . . 81
8.2 Wegunabhängigkeit und Potential . . . 84
8.3 Mehrdimensionale Integrale . . . 86
8.4 Integration in krummlinigen Koordinaten . . . 88
8.5 Flächenintegrale und Vektorfluss . . . 92

**9 Die Integralsätze von Gauß und Stokes 95**
9.1 Der Greensche Satz in der Ebene . . . 95
9.2 Integraldarstellung der Rotation . . . 97
9.3 Integralsatz von Stokes . . . 98

9.4 Integraldarstellung der Divergenz . . . 100
9.5 Gaußscher Satz . . . 102

**10 Operatoren und Eigenwerte 105**
10.1 Eigenwerte von Matrizen . . . 105
10.1.1 Theoreme zum Eigenwertproblem . . . 110
10.2 Operatoren . . . 111
10.3 Eigenwertproblem für Operatoren . . . 113
10.4 Operatoren in der Quantenmechanik . . . 114

**11 Fourierreihen und Integraltransformationen 117**
11.1 Funktionenräume . . . 117
11.2 Fourierreihen . . . 119
11.3 Fourier-Transformation . . . 124
11.3.1 Delta-Funktion . . . 129
11.3.2 Anwendung: DGL . . . 131

**12 Systeme von Differentialgleichungen 133**
12.1 Lineare Differentialgleichungssysteme . . . 134
12.2 Dynamische Systeme . . . 136

**13 Symmetrien und Gruppen 139**
13.1 Symmetrien . . . 139
13.2 Gruppen . . . 140
13.3 Wichtige Gruppen . . . 141
13.3.1 Permutationsgruppe . . . 141
13.3.2 Matrixgruppen . . . 142
13.4 Darstellungen . . . 144
13.5 Kontinuierliche Gruppen . . . 145
13.6 Generatoren und Lie-Algebra . . . 146
13.7 Tensorrechnung . . . 148
13.7.1 Definition . . . 149
13.7.2 Rechenregeln für Tensoren . . . 152

**14 Differentialgleichungen III 153**
14.1 DGL als Eigenwertproblem . . . 153
14.2 Spezielle DGL . . . 155
14.2.1 Legendre'sche DGL . . . 155
14.2.2 Kugelflächenfunktionen . . . 156
14.2.3 Bessel'sche DGL . . . 158
14.2.4 Hermite'sche DGL . . . 158
14.2.5 Laguerre'sche DGL . . . 158
14.3 Partielle DGL . . . 159

14.3.1 Wichtige pDGL der Physik . . . 159
14.4 Lösungsverfahren für pDGL; Green'sche Funktionen . . . 161
14.4.1 Integraldarstellung . . . 161
14.4.2 Integraltransformation . . . 161
14.4.3 Green'sche Funktion . . . 162
14.4.4 Separation der Variablen . . . 164

**15 Koordinatensysteme 165**
15.1 Gebräuchliche Koordinatensysteme . . . 165
15.1.1 Zylinderkoordinaten . . . 167
15.1.2 Kugelkoordinaten . . . 168
15.2 Bestimmung von Vektorkomponenten . . . 169
15.3 Differentialoperatoren in krummlinigen Koordinaten . . . 172
15.3.1 Gradient . . . 172
15.3.2 Divergenz . . . 173
15.3.3 Rotation . . . 174
15.3.4 Laplace-Operator . . . 175
15.4 Bogen-, Flächen-, Volumenelemente . . . 176
15.4.1 Raumkurven . . . 176
15.4.2 Flächen- und Volumenelemente . . . 177

# Kapitel 1

# Vektoren

Vektoren sind gerichtete Größen. Sie beinhalten Informationen über *Länge* (Zahl) und *Richtung*. Vektoren können auch dimensionsbehaftete Größen sein. Dann haben sie auch eine Einheit.

Ein typischer Vektor ist eine Verschiebung (siehe Abb. 1.0.1). Damit lautet eine erste, intuitive Definition für einen Vektor:

Größen, die sich wie eine Verschiebung verhalten, bezeichnet man als **Vektoren**.

Man schreibt hierfür $\vec{a}, \underline{a}, \mathbf{a}, \ldots$.

Im Rahmen der Vorlesung und in diesem Skript verwenden wir die Notation $\underline{a}$. Dies ist vor allem dann bequem, wenn man mit der Hand schreibt.

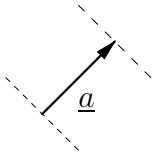



Abbildung 1.0.1: Verschiebungen lassen sich durch Vektoren charakterisieren.

Der Angriffspunkt spielt bei der Charakterisierung einer Verschiebung keine Rolle. Daher sind Vektoren parallel verschiebbar.

Typische Beispiele für Vektoren in der Physik sind der Ortsvektor $\underline{r}$, der die Position eines Teilchens angibt, die Geschwindigkeit $\underline{v}$ oder eine Kraft $\underline{F}$.

Eine physikalische Größe, die sich durch Angabe einer einfachen Zahl beschreiben lässt, bezeichnet man als **Skalar**. Wichtige Beispiele hierfür sind die Masse $m$, die Temperatur $T$ oder die Zeit $t$.

## 1.1 Vektorräume

Eine mathematisch formalere Definition eines Vektors ist folgende:

**Definition 1.1.1** (Vektorraum, Vektor)**.**
Eine Menge von Elementen $V = \{\underline{a}, \underline{b}, \dots\}$ heißt **Vektorraum** und $\underline{a}, \underline{b}, \dots$ **Vektoren**, wenn für alle $\underline{a}, \underline{b}, \underline{c}, \dots \in V$ gilt:

1. Abgeschlossenheit: $\underline{a}+\underline{b} \in V$ und $\lambda\underline{a} \in V$ (für alle $\lambda \in \mathbb{R}$; $\lambda$ bezeichnet man auch als Skalar)
2. Assoziativität: $\underline{a} + (\underline{b} + \underline{c}) = (\underline{a} + \underline{b}) + \underline{c}$
3. Neutrales Element (Nullvektor): Es existiert $\underline{0} \in V$ mit $\underline{a} + \underline{0} = \underline{a}$ für alle $\underline{a} \in V$
4. Inverse: Für alle $\underline{a} \in V$ existiert $\tilde{\underline{a}} \in V$ mit $\underline{a} + \tilde{\underline{a}} = \underline{0}$. $\tilde{\underline{a}}$ bezeichnet man als inverses Element. Man schreibt auch $-\underline{a}$.
5. Kommutativität: $\underline{a} + \underline{b} = \underline{b} + \underline{a}$
6. Für alle $\lambda, \mu \in \mathbb{R}$ gilt:
   (a) $(\lambda + \mu)\underline{a} = \lambda\underline{a} + \mu\underline{a}$
   (b) $(\lambda\mu)\underline{a} = \lambda(\mu\underline{a})$
   (c) $\lambda(\underline{a} + \underline{b}) = (\lambda\underline{a}) + (\lambda\underline{b})$
7. $1 \cdot \underline{a} = \underline{a}$

Die meisten dieser Eigenschaften sind offensichtlich, wenn man an Verschiebungen denkt!

Die Addition zweier Vektoren $\underline{a}$ und $\underline{b}$ kann man sich anschaulich als Hintereinanderausführung zweier Verschiebungen vorstellen, die sich natürlich durch eine einzige Verschiebung realisieren lässt. Diese bezeichnet man dann mit $\underline{a} + \underline{b}$. Grafisch lässt sich das wie folgt veranschaulichen (siehe Abb. 1.1.1a): Der Vektor $\underline{b}$ wird parallel an das Ende des Vektors $\underline{a}$ verschoben. $\underline{a} + \underline{b}$ ist dann die Verbindung des Anfangspunktes von $\underline{a}$ mit dem Endpunkt des verschobenen Vektors $\underline{b}$.

Die anschauliche Bedeutung des Multiplikation mit einem Skalar werden wir später in diesem Abschnitt diskutieren.

Abb. 1.1.1b veranschaulicht das inverse Element $\tilde{\underline{a}} = -\underline{a}$. $-\underline{a}$ hat die gleiche Länge, aber die entgegengesetzte Richtung von $\underline{a}$. Damit läßt sich auch die Subtraktion von Vektoren als Addition des Inversen definieren: $\underline{a} - \underline{b} := \underline{a} + (-\underline{b}) = \underline{a} + \tilde{\underline{b}}$.

Abbildung 1.1.1: Veranschaulichung der Addition von Vektoren (links) und des inversen Elements $\tilde{\underline{a}} = -\underline{a}$ (rechts).

**Beispiel 1.1.1.**
$\mathbb{R}^2$ mit $\underline{a} = \begin{pmatrix} a_1 \\ a_2 \end{pmatrix}$, wobei $a_1, a_2 \in \mathbb{R}$

mit der Addition und skalaren Multiplikation: $\boxed{\underline{a} + \lambda\underline{b} := \begin{pmatrix} a_1 + \lambda b_1 \\ a_2 + \lambda b_2 \end{pmatrix}}$

Verallgemeinerung auf $\mathbb{R}^n$: analoge komponentenweise Definition
Speziell für $n = 1$ erhält man die bekannten Rechenregeln für die reellen Zahlen $\mathbb{R}$.

**Definition 1.1.2** (Betrag eines Vektors)**.**
Den **Betrag** bzw. die **Länge** oder auch **Norm** eines Vektors $\underline{a}$ bezeichnet man mit

$$|\underline{a}| = a.$$

Später werden wir immer die Schreibweise $a$ verwenden[1], sofern keine Verwechslungsgefahr besteht.
Speziell gilt: $|\underline{0}| = 0$.

Für das obige Beispiel 1.1 des $\mathbb{R}^2$ definiert man den Betrag als $\left|\begin{pmatrix} a_1 \\ a_2 \end{pmatrix}\right| = \sqrt{a_1^2 + a_2^2}$. Allgemein gilt für den $\mathbb{R}^n$: $\boxed{\left|\begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix}\right| = \sqrt{a_1^2 + \cdots + a_n^2}}$

# 1.2 Vektoralgebra

Im folgenden wollen wir drei verschiedene Multiplikationen von Vektoren definieren.

**Definition 1.2.1** (Skalare Multiplikation)**.** $\underline{b} = \lambda\underline{a}$ (Vektor=Skalar · Vektor)

Die **skalare Multiplikation** kennen wir bereits aus den Vektorraum-Axiomen, die auch ihre Regeln festgelegen. Man kann das Ergebnis einfach geometrisch interpretieren:
a) $\lambda > 0$ : $\underline{a}$ und $\underline{b}$ zeigen in die gleiche Richtung (parallel): $\underline{a} \parallel \underline{b}$
b) $\lambda < 0$ : $\underline{a}$ und $\underline{b}$ entgegengesetzt (antiparallel).

[1] D.h. wir lassen einfach das Vektorsymbol _ weg.

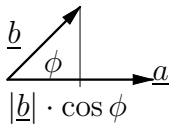



Abbildung 1.2.1: Die Projektion von $\underline{b}$ auf $\underline{a}$.

Für den Betrag gilt $|\underline{b}| = |\lambda| \cdot |\underline{a}|$. Daher:
a) $|\lambda| > 1$: Streckung
b) $|\lambda| < 1$: Stauchung

**Definition 1.2.2** (Skalarprodukt)**.** Vektor·Vektor=Skalar (Zahl)

Das **Skalarprodukt** zweier Vektoren $\underline{a}$, $\underline{b}$ liefert einen Skalar.
Im $\mathbb{R}^2$ ist es definiert als: $\underline{a} \cdot \underline{b} := a_1 b_1 + a_2 b_2$.

Die Verallgemeinerung auf den $\mathbb{R}^n$ ist offensichtlich: $\boxed{\underline{a} \cdot \underline{b} := a_1 b_1 + \ldots + a_n b_n}$

Wir können das Skalarprodukt auch geometrisch definieren: $\boxed{\underline{a} \cdot \underline{b} = |\underline{a}| \cdot |\underline{b}| \cdot \cos\phi}$

Diese geometrische Deutung impliziert nun:
$\underline{a} \cdot \underline{b} = \underline{0} \Leftrightarrow \phi = \pm\pi/2 \Leftrightarrow \underline{a}$ senkrecht auf $\underline{b}$ $\quad (\underline{a} \perp \underline{b})$

Die **Projektion** von $\underline{b}$ auf $\underline{a}$ ist $|\underline{b}| \cos\phi$ (siehe Abb. 1.2.1). Diese Projektion lässt sich auch direkt mit Hilfe des Skalarproduktes ausdrücken als $\frac{\underline{a} \cdot \underline{b}}{|\underline{a}|}$ oder $\hat{a} \cdot \underline{b}$, wobei wir den Einheitsvektor $\hat{a} = \frac{\underline{a}}{|\underline{a}|}$ in Richtung von $\underline{a}$ eingeführt haben (s.u., Def. 1.3.3).

**Definition 1.2.3** (Kreuzprodukt (Vektorprodukt, äußeres Produkt))**.** Vektor $\times$ Vektor $=$ Vektor

Im Gegensatz zur skalaren Multiplikation und dem Skalarprodukt ist das **Kreuzprodukt** $\underline{c} = \underline{a} \times \underline{b} \in \mathbb{R}^3$ zweier Vektoren $\underline{a}, \underline{b} \in \mathbb{R}^3$ nur im $\mathbb{R}^3$ definiert!

Anschauliche Definition: $\underline{c} = \underline{a} \times \underline{b}$ hat den Betrag $\boxed{|\underline{c}| = |\underline{a}| \cdot |\underline{b}| \cdot \sin\phi}$ und steht senkrecht auf $\underline{a}$ und $\underline{b}$: $\boxed{\underline{c} \perp \underline{a},\ \underline{c} \perp \underline{b}.}$ Dabei bilden $\underline{a}, \underline{b}, \underline{c}$ ein Rechtssystem (Rechte-Hand-Regel).
Bem: Daraus folgt sofort: $\underline{a} \times \underline{b} = -\underline{b} \times \underline{a}$ (das Kreuzprodukt ist antikommutativ).

Komponentenweise Definition: $\boxed{\underline{a} \times \underline{b} := \begin{pmatrix} a_2 b_3 - a_3 b_2 \\ a_3 b_1 - a_1 b_3 \\ a_1 b_2 - a_2 b_1 \end{pmatrix}}$

Ein weiteres Produkt, das nur im $\mathbb{R}^3$ definiert ist, ist das Spatprodukt.

**Definition 1.2.4** (Spatprodukt)**.**
Als **Spatprodukt** bezeichnet man folgendes Produkt dreier Vektoren $\underline{a}$, $\underline{b}$, $\underline{c} \in \mathbb{R}^3$: $\quad \underline{a} \cdot (\underline{b} \times \underline{c})$.

Weitere Rechenregeln für Skalar-und Kreuzprodukt lernen Sie in den Übungen kennen!

Wichtige Bemerkung: Im Gegensatz zur Multiplikation von reellen Zahlen ($\neq 0$) lassen sich die Multiplikationen von Vektoren i.a. nicht umkehren.
M.a.W.: **Es gibt keine Division durch Vektoren!**
Ein Grund hierfür ist, dass man z.B. aus der Kenntnis von $\underline{a} \cdot \underline{b}$ und $\underline{a}$ nicht *eindeutig* auf $\underline{b}$ zurückschliessen kann. Tatsächlich gibt es hier i.a. unendlich viele Lösungen. Man kann nur durch Beträge von Vektoren dividieren!

## 1.3 Basis eines Vektorraumes

**Definition 1.3.1** (Linearkombination, Erzeugendensystem)**.**

Es ist zweckmässig, in einem Vektorraum gewisse Vektoren $\underline{e}_1, \underline{e}_2, \ldots, \underline{e}_n$ auszuzeichnen, mit deren Hilfe sich alle anderen Vektoren $\underline{a} \in V$ als

$$\underline{a} = \sum_{j=1}^{n} \lambda_j \underline{e}_j = \lambda_1 \underline{e}_1 + \lambda_2 \underline{e}_2 + \cdots + \lambda_n \underline{e}_n$$

darstellen lassen. Man sagt: $\underline{a}$ ist eine **Linearkombination** der Vektoren $\underline{e}_1, \underline{e}_2, \ldots, \underline{e}_n$. Ist es bei fest vorgegebenen $\underline{e}_1, \underline{e}_2, \ldots, \underline{e}_n$ möglich, für *alle* Vektoren $\underline{a} \in V$ geeignete Skalare $\lambda_j \in \mathbb{R}$ zu finden, so dass sich $\underline{a}$ als Linearkombination dieser Vektoren ergibt, so bilden die Vektoren $\underline{e}_1, \underline{e}_2, \ldots, \underline{e}_n$ ein sog. **Erzeugendensystem** von V.

**Definition 1.3.2** (Basis, Dimension)**.**

Kann man keinen der Vektoren eines Erzeugendensystems weglassen, ohne dass diese Eigenschaft verloren geht, so heißt $\underline{e}_1, \underline{e}_2, \ldots, \underline{e}_n$ eine **Basis** des Vektorraumes $V$. Die Zahl $n$ der Vektoren der Basis heißt die **Dimension** von $V$. Man schreibt hierfür auch $n = \dim V$.
$n$ kann auch unendlich sein! In diesem Fall spricht man von einem unendlich-dimensionalen Vektorraum und schreibt $\dim V = \infty$. In der Linearen Algebra wird gezeigt, dass $n$ eindeutig bestimmt ist, d.h. alle Basen eines Vektorraums bestehen aus gleich vielen Elementen.

Die $\lambda_j$ sind die **Komponenten** von $\underline{a}$ bzgl. der Basis $\underline{e}_1, \underline{e}_2, \ldots, \underline{e}_n$.

**Beispiel 1.3.1.** Die Vektoren $\underline{e}_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$, $\underline{e}_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$, $\underline{e}_3 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$ bilden ein Erzeugendensystem des $\mathbb{R}^2$. $\underline{e}_1$, $\underline{e}_2$ bilden sogar eine Basis, denn für einen beliebigen Vektor $\underline{a} = \begin{pmatrix} a_1 \\ a_2 \end{pmatrix}$ gilt offensichtlich $\underline{a} = a_1 \underline{e}_1 + a_2 \underline{e}_2$. Somit hat $\underline{a}$ die Komponenten $a_1$ und $a_2$ bzgl. der Basis $\underline{e}_1$, $\underline{e}_2$. Außerdem folgt $\dim \mathbb{R}^2 = 2$.
Nach Hinzunahme des Vektors $\underline{e}_3$ kann man immer noch alle Elemente des $\mathbb{R}^2$ als Linearkombination darstellen. Man setzt einfach $\lambda_3 = 0$. Allerdings ist $\underline{e}_1$, $\underline{e}_2$, $\underline{e}_3$ keine Basis mehr.

**Beispiel 1.3.2.**

$$\underline{e}_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \quad \underline{e}_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \quad \underline{e}_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$$

ist eine Basis des $\mathbb{R}^3$.
Allgemein hat der $\mathbb{R}^n$ die Dimension $n$.

Häufig ist es zweckmässig, die Basisvektoren möglichst "einfach" zu wählen. I.a. normiert man ihre Länge auf 1 (d.h. $\underline{e}_j \cdot \underline{e}_j = 1$ bzw. $|\underline{e}_j| = 1$) und wählt die Vektoren paarweise orthogonal ($\underline{e}_j \perp \underline{e}_l$): $\underline{e}_j \cdot \underline{e}_l = 0$ für alle $j \neq l$. Man spricht dann von einer **Orthonormalbasis**.
Ist $|\underline{e}_j| \neq 1$ für mindestens ein $j$, so hat man es mit einer **Orthogonalbasis** zu tun.

**Definition 1.3.3** (Einheitsvektor)**.**
**Einheitsvektoren** sind Vektoren der Länge 1. Zu einem beliebigen Vektor $\underline{a} \in V$ ($\underline{a} \neq \underline{0}$) kann man einen Einheitsvektor gleicher Richtung definieren:

$$\hat{a} = \frac{\underline{a}}{|\underline{a}|} \qquad \curvearrowright \quad |\hat{a}| = 1$$

Offensichtlich sind $\hat{a}$ und $\underline{a}$ parallel zueinander.

Speziell für die Einheitsvektoren in $x, y, z$-Richtung schreibt man: $\hat{x}, \hat{y}, \hat{z}$ oder $\underline{e}_x, \underline{e}_y, \underline{e}_z$.

Die Wahl einer Basis eines Vektorraumes ist nicht eindeutig. Man wählt i.a. die Basis, die für ein gegebenes Problem am zweckmässigsten erscheint. In der Regel sind dies Orthonormalbasen.

## 1.4 Koordinatensysteme

In der Physik tritt sehr häufig das Problem auf, die Lage von Punkten in einem Raum zu beschreiben[2]. Dies geschieht mit Hilfe von Koordinatensystemen. Ein **Koordinatensystem** erhält man durch Festlegung eines Punktes, dem **Ursprung**, und einer Basis.

Im folgenden wollen wir die wichtigsten Koordinatensysteme kurz vorstellen. Wir werden später darauf zurückkommen und z.B. die Basisvektoren der Systeme explizit definieren.

Wir betrachten nun einen Punkt $P$, dessen Lage im Raum wir auf verschiedene Arten charakterisieren wollen. Wir stellen uns daher vor, dass wir einen ausgezeichneten Punkt als Ursprung gewählt haben. Der Punkt $P$ kann dann durch den Vektor $\underline{a}$ charakterisiert werden. Dieser entspricht der Verschiebung vom Ursprung nach $P$.

### 1.4.1 Kartesische Koordinaten

Wir betrachten zunächst den Fall eines zweidimensionalen Raumes, den wir mit dem $\mathbb{R}^2$ identifizieren. Die einfachste Wahl einer Basis des $\mathbb{R}^2$ führt auf das sog. **kartesische Koordinatensystem**. Hier wählt man (orthogonalen!) Einheitsvektoren $\underline{e}_x$, $\underline{e}_y$ in $x-$ und $y-$Richtung als Basisvektoren (siehe Abb. 1.4.1) und kommt so (durch Projektion auf die $x-$ bzw. $y-$Achse) zu der bekannten Darstellung $\underline{a} = a_x \underline{e}_x + a_y \underline{e}_y$ bzw. $\underline{a} = \begin{pmatrix} a_x \\ a_y \end{pmatrix}$.
Hier sind also die Koordinaten $(a_x, a_y)$ des Punktes $P$ gleich den Komponenten des Vektors $\underline{a}$.

[2] Der 'Raum' kann dabei durchaus sehr abstrakt sein!

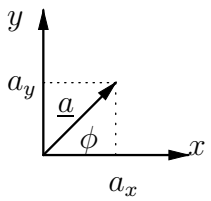



Abbildung 1.4.1: Kartesisches Koordinatensystem und Definition der ebenen Polarkoordinaten.

### 1.4.2 (Ebene) Polarkoordinaten

Das zweite häufig verwendete Koordinatensystem in zwei Dimensionen sind die **(ebenen) Polarkoordinaten** $(|\underline{a}|, \phi)$. Die Koordinaten sind hier nicht durch die Projektionen auf die $x-$ und $y-$Achse gegeben, sondern durch die Länge $a = |\underline{a}|$ des Vektors $\underline{a}$ und den Winkel $\phi$, den er mit der $x-$Achse einschließt (siehe Abb. 1.4.1). Daraus ergibt sich der folgende Zusammenhang zwischen den kartesischen und den Polarkoordinaten:

$$\left.\begin{array}{rcl} a_x &=& a\cos\phi \\ a_y &=& a\sin\phi \end{array}\right\} \quad \Longleftrightarrow \quad \left\{\begin{array}{rcl} a &=& \sqrt{a_x^2 + a_y^2} \\ \tan\phi &=& \frac{a_y}{a_x} \end{array}\right.$$

Die Koordinaten können folgende Werte annehmen: $a \geq 0$ und $0 \leq \phi < 2\pi$.

Man beachte, dass $\phi$ durch die Bedingung $\tan\phi = a_y/a_x$ nicht eindeutig bestimmt wird! Dies wird in Kap. 5.2 genauer diskutiert!

Bem: Ein Vektor $\underline{a}$ an sich ist unabhängig vom Koordinatensystem. Erst die Komponenten eines Vektors sind von der Wahl des Koordinatensystems abhängig.
Die Polarkoordinaten stimmen nicht mehr mit den Komponenten des Vektors überein!

### 1.4.3 Zylinderkoordinaten

Üblicherweise arbeitet man im $\mathbb{R}^3$ mit einem dreidimensionalen kartesischen Koordinatensystem. Bei Problemen mit bestimmten Rotationssymmetrien können jedoch daran angepasste Koordinatensysteme weiterhelfen.
In **Zylinderkoordinaten** werden zweidimensionale ebene Polarkoordinaten in der $x$-$y$-Ebene mit einer $z$-Koordinate kombiniert (siehe Abb. 1.4.2). $\underline{a}' = (a_x, a_y, 0)$ ist die Projektion von $\underline{a}$ in die $x-y-$Ebene und $\phi$ der Winkel zwischen $\underline{a}'$ und der $x-$Achse. Dabei hat man zu beachten, dass das urprüngliche kartesische Koordinatensystem rechtshändig ist, d.h. die Rechte-Hand-Regel erfüllt[3]. Als Koordinaten erhält man $a' = |\underline{a}'|$, $\phi$ und $a_z$ mit folgender Umrechnungsvorschrift:

[3]Der Daumen zeigt in Richtung der $x-$Achse, der Zeigefinger in $y-$Richtung und der Mittelfinger in Richtung der $z-$Achse.

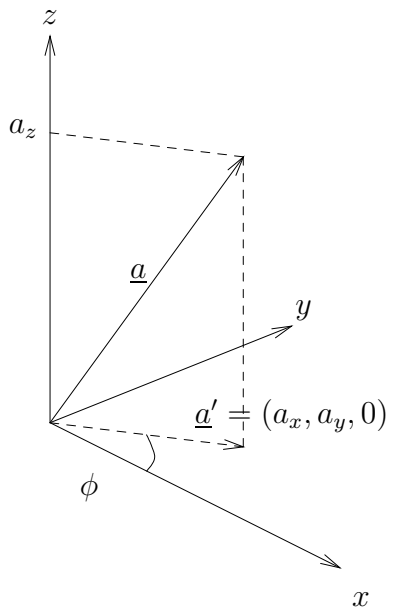



Abbildung 1.4.2: Zylinderkoordinaten

$$\left.\begin{array}{lcl} a_x &=& a'\cos\phi \\ a_y &=& a'\sin\phi \\ a_z & & \end{array}\right\} \quad \Longleftrightarrow \quad \left\{\begin{array}{lcl} a' &=& \sqrt{a_x^2+a_y^2} \\ \tan\phi &=& \frac{a_y}{a_x} \\ a_z & & \end{array}\right.$$

Die Koordinaten können folgende Werte annehmen: $a' \geq 0$, $0 \leq \phi < 2\pi$ und $-\infty < a_z < \infty$.

Bem.: Man beachte, dass $a'$ nicht der Betrag von $\underline{a}$ ist, also dessen Länge. Stattdessen ist $a'$ der Abstand von der $z-$Achse.

### 1.4.4 Kugelkoordinaten (sphärische Polarkoordinaten)

In **Kugelkoordinaten** benutzt man außer der Länge $a = |\underline{a}|$ des Vektors $\underline{a}$ zwei Winkel $\phi$, $\theta$ (siehe Abb. 1.4.3). Der Winkel $\theta$ ist dabei der Winkel zwischen dem Vektor $\underline{a}$ und der $z-$Achse. Um von den Zylinderkoordinaten zu den Kugelkoordinaten überzugehen, muß man die Projektionen $\underline{a}'$ in die $x-y-$Ebene und $a_z$ auf die $z-$Achse durch die Länge $a$ und den Winkel $\theta$ ausdrücken. Mit $a' = a\sin\theta$ und $a_z = a\cos\theta$ erhält man:

$$\left.\begin{array}{lcl} a_x &=& a\sin\theta\cos\phi \\ a_y &=& a\sin\theta\sin\phi \\ a_z &=& a\cos\theta \end{array}\right\} \quad \Longleftrightarrow \quad \left\{\begin{array}{lcl} a &=& \sqrt{a_x^2+a_y^2+a_z^2} = |\underline{a}| \\ \tan\phi &=& \frac{a_y}{a_x} \\ \tan\theta &=& \frac{\sqrt{a_x^2+a_y^2}}{a_z} \end{array}\right.$$

Die Koordinaten können folgende Werte annehmen: $r \geq 0$, $0 \leq \phi < 2\pi$ und $0 \leq \theta \leq \pi$.

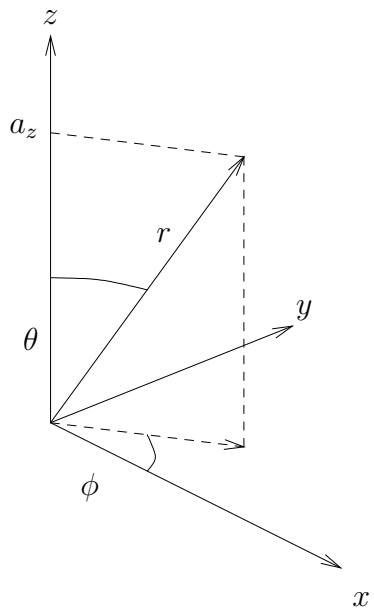



Abbildung 1.4.3: Kugelkoordinaten

Wir werden später sehen, dass man durch geschickte Wahl des Koordinatensystems oftmals ein Problem stark vereinfachen kann. Dabei hat man sich an den Symmetrien des Problems zu orientieren. Bei einer Kreisbewegung ist z.B. der Radius konstant und nur der Winkel $\phi$ verändert sich in einer Polarkoordinatendarstellung. In kartesischen Koordinaten würden sich dagegen sowohl die $x-$ als auch die $y-$Koordinate zeitlich ändern.

# Kapitel 2

# Matrizen und Lineare Gleichungssysteme

## 2.1 Lineare Abbildungen

Wir betrachten einen $n$-dimensionalen Vektorraum $V$ mit der Basis $\{\underline{e}_1, \ldots, \underline{e}_n\}$ und einen $n'$-dimensionalen Vektorraum $V'$ mit der Basis $\{\underline{e}'_1, \ldots, \underline{e}'_n\}$.
Eine Abbildung $f : V \to V'$ heißt **linear**, falls für alle $\underline{v}_1, \underline{v}_2 \in V$ und alle $\alpha \in \mathbb{R}$ gilt:

$$f(\underline{v}_1 + \alpha \underline{v}_2) = f(\underline{v}_1) + \alpha f(\underline{v}_2)\,.$$

Hieraus folgt offenbar sofort, dass $f(\underline{0}) = \underline{0}$ sein muss!

Man kann lineare Abbildungen auch anders charakterisieren, nämlich über das Verhalten der Komponenten bzw. Koordinatenvektoren. Dazu betrachten wir das Verhalten der Basisvektoren aus $V$ unter der Abbildung $f$. Da es sich um Vektoren in $V'$ handelt, lassen sich die Bildvektoren in der Basis von $V'$ darstellen:

$$f(\underline{e}_j) = \sum_{l=1}^{n'} M_{lj} \underline{e}'_l\,.$$

Diese Gleichungen definieren die $n \cdot n'$ reellen (oder komplexen) Zahlen $M_{lj}$.
Wir werden nun zeigen, dass die Kenntnis von $M_{lj}$ ausreicht, um den Bildvektor eines beliebigen Vektors $\underline{v} \in V$ zu bestimmen. Dazu zerlegen wir $\underline{v}$ zunächst bezüglich der Basis in $V$:

$$\underline{v} = \sum_{j=1}^{n} v_j \underline{e}_j\,.$$

Unter Ausnutzung der Linearität von $f$ und der Definition der $M_{lj}$ ergibt sich dann:

$$\begin{aligned} f(\underline{v}) &= f\left(\sum_{j=1}^{n} v_j \underline{e}_j\right) = \sum_{j=1}^{n} v_j f(\underline{e}_j) = \sum_{j=1}^{n} v_j \sum_{l=1}^{n'} (M_{lj} \underline{e}'_l) \\ &= \sum_{l=1}^{n'} \left(\sum_{j=1}^{n} M_{lj} v_j\right) \underline{e}'_l \stackrel{!}{=} \sum_{l=1}^{n'} w_l \underline{e}'_l\,. \end{aligned}$$

Somit kann man die Komponenten $w_l$ des Bildvektors $f(\underline{v})$ durch die Komponenten von $\underline{v}$ und die $M_{lj}$ ausdrücken:

$$w_l = \sum_{j=1}^{n} M_{lj} v_j \,.$$

Die Zahlen $M_{lj}$ enthalten also die vollständige Information über die lineare Abbildung $f$. Dies motiviert die Definitionen im folgenden Abschnitt.

## 2.2 Matrizen

**Definition 2.2.1** (Matrizen)**.**
Die folgende Anordnung von $m \times n$ (reellen) Zahlen $a_{jl}$ (mit $j = 1, \ldots, m$ und $l = 1, \ldots, n$) in $m$ Zeilen und $n$ Spalten

$$\underline{\underline{A}} = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix}$$

bezeichnen wir als $m \times n$ **Matrix**.
Oft schreibt man auch einfach $\underline{\underline{A}} = (a_{jl})$, wenn man die **Elemente** $a_{jl}$ nicht explizit angeben möchte.
Die Menge aller $m \times n$-Matrizen bezeichnen wir im folgenden auch mit $\mathcal{M}(m, n)$.

Wir wollen die wichtigsten Regeln für das **Rechnen mit Matrizen** zusammenstellen. Seien

$$\underline{\underline{A}} = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix} \quad \text{und} \quad \underline{\underline{\tilde{A}}} = \begin{pmatrix} \tilde{a}_{11} & \cdots & \tilde{a}_{1n} \\ \vdots & \ddots & \vdots \\ \tilde{a}_{m1} & \cdots & \tilde{a}_{mn} \end{pmatrix}$$

$m \times n$-Matrizen und

$$\underline{\underline{B}} = \begin{pmatrix} b_{11} & \cdots & b_{1n} \\ \vdots & \ddots & \vdots \\ b_{m1} & \cdots & b_{mn} \end{pmatrix}$$

eine $n \times p$-Matrix.

1. $\underline{\underline{A}} = \underline{\underline{\tilde{A}}} \quad \Longleftrightarrow \quad a_{jl} = \tilde{a}_{jl}$ für alle $j, l$.
   Zwei Matrizen sind also gleich, wenn alle ihre Elemente übereinstimmen. Insbesondere müssen beide Matrizen gleich groß sein!

2. Die Addition zweier Matrizen ist elementweise definiert durch

$$\underline{\underline{A}} + \underline{\underline{\tilde{A}}} := \begin{pmatrix} a_{11} + \tilde{a}_{11} & \cdots & a_{1n} + \tilde{a}_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} + \tilde{a}_{m1} & \cdots & a_{mn} + \tilde{a}_{mn} \end{pmatrix} = (a_{jl} + \tilde{a}_{jl}) \,.$$

   Man beachte, dass die Summe nur für Matrizen der gleichen Größe definiert ist !

3. Die Multiplikation einer Matrix mit einem Skalar $\lambda \in \mathbb{R}$ ist komponentenweise durch

$$\lambda \underline{\underline{A}} := \begin{pmatrix} \lambda a_{11} & \cdots & \lambda a_{1n} \\ \vdots & \ddots & \vdots \\ \lambda a_{m1} & \cdots & \lambda a_{mn} \end{pmatrix} = (\lambda a_{jl}) \,.$$

   definiert.

4. Das Matrixprodukt $\underline{\underline{C}} := \underline{\underline{A}} \cdot \underline{\underline{B}}$ ist eine $m \times p$-Matrix mit den Elementen

$$\boxed{c_{ij} = \sum_{l=1}^{n} a_{il} b_{lj}.}$$

   Damit man zwei Matrizen multiplizieren kann, muß die Zeilenzahl der zweiten Matrix gleich der Zahl der Spalten der ersten Matrix sein ! Bei der Berechnung des $(i,j)$-ten Elements des Produktes muß man also die Produkte der $l$-ten Elemente der $i$-ten Zeile von $\underline{\underline{A}}$ mit dem $l$-ten Element der $j$-ten Spalte von $\underline{\underline{B}}$ aufsummieren.

   *Bemerkung.* Im allgemeinen ist

$$\underline{\underline{A}} \cdot \underline{\underline{C}} \neq \underline{\underline{C}} \cdot \underline{\underline{A}}.$$

   Z.B. ist das rechte Produkt für $m \neq p$ überhaupt nicht definiert.

5. Für die Matrixmultiplikation gilt das Distributivgesetz

$$\left(\underline{\underline{A}} + \underline{\underline{\tilde{A}}}\right) \cdot \underline{\underline{B}} = \underline{\underline{A}} \cdot \underline{\underline{B}} + \underline{\underline{\tilde{A}}} \cdot \underline{\underline{B}}.$$

6. Das Nullelement in $\mathcal{M}(m,n)$ ist die Matrix

$$\underline{\underline{0}} = \begin{pmatrix} 0 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & 0 \end{pmatrix}.$$

7. Als $n \times n$-**Einheitsmatrix** in $\mathcal{M}(n,n)$ bezeichnet man die Matrix

$$\underline{\underline{E}} := \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix} = (\delta_{jl})_{j,l=1,\ldots,n} =: \mathbb{1} =: \underline{\underline{I}}$$

   mit dem sog. **Kronecker-Symbol**

$$\delta_{jl} = \begin{cases} 1 & \text{falls } j = l \\ 0 & \text{sonst} \end{cases}$$

Die Einheitsmatrix ist eine spezielle **Diagonalmatrix**, bei denen nur die Elemente $a_{jj}$ auf der Diagonalen von Null verschieden sein können.
Man überlegt sich leicht, dass für eine beliebige Matrix $\underline{\underline{A}} \in \mathcal{M}(n,n)$ gilt:

$$\underline{\underline{A}} \cdot \underline{\underline{E}} = \underline{\underline{E}} \cdot \underline{\underline{A}} = \underline{\underline{A}}.$$

8. Die **transponierte Matrix** $\underline{\underline{A}}^t$ einer Matrix $\underline{\underline{A}} \in \mathcal{M}(m,n)$ ist die $n \times m$-Matrix, die aus $\underline{\underline{A}}$ durch Vertauschung von Zeilen und Spalten hervorgeht:

$$\underline{\underline{A}} = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix} \quad \Longrightarrow \quad \underline{\underline{A}}^t = \begin{pmatrix} a_{11} & \cdots & a_{m1} \\ \vdots & \ddots & \vdots \\ a_{1n} & \cdots & a_{nm} \end{pmatrix}, \qquad \text{d.h.} \quad a^t_{jl} = a_{lj}.$$

   Für quadratische Matrizen ($n = m$) entspricht dies einer Spiegelung an der Diagonalen.

9. **Inverse Matrix**: Sei $\underline{\underline{A}}$ eine $n \times n$-Matrix. Existiert eine Matrix $\underline{\underline{A}}^{-1}$ mit

$$\underline{\underline{A}} \cdot \underline{\underline{A}}^{-1} = \underline{\underline{E}} = \underline{\underline{A}}^{-1} \cdot \underline{\underline{A}},$$

   so bezeichnet man diese als die zu $\underline{\underline{A}}$ inverse Matrix.

*Wichtiger Spezialfall*: $n \times 1$-Matrizen, also Matrizen mit nur einer Spalte, identifiziert man mit Vektoren. Die Multiplikation eines Vektors $\underline{b} \in \mathcal{M}(n,1)$ mit einer Matrix $\underline{\underline{A}} \in \mathcal{M}(n,n)$ ist dann gegeben durch

$$\underline{\underline{A}} \cdot \underline{b} = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix} \cdot \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix} = \begin{pmatrix} a_{11}b_1 + a_{12}b_2 + \cdots + a_{1n}b_n \\ \vdots \\ a_{m1}b_1 + a_{m2}b_2 + \cdots + a_{mn}b_n \end{pmatrix}$$

Für $n = m = 2$ entspricht das gerade unserem oben betrachteten Beispiel. Jetzt sollte auch die Matrixdarstellung von linearen Gleichungssystemen klar sein!

Somit lassen sich lineare Abbildungen auch durch Matrizen charakterisieren. Nach Wahl von Basen in Ursprungs- und Bildraum, kann man die Matrix $\underline{\underline{M}}$ bestimmen. Damit kann man für einen beliebigen Vektor $\underline{v}$ die Kompenten des Bildvektors $\underline{w}$ bestimmen:

$$w_l = \sum_j M_{lj} v_j \qquad \text{bzw.} \qquad \underline{w} = \underline{\underline{M}}\,\underline{v}\,.$$

## 2.3 Determinanten

**Definition 2.3.1** (Determinante)**.**
Als **Determinante** einer $n \times n$-Matrix $\underline{\underline{A}} = (a_{jl})$ bezeichnen wir folgende reelle Zahl:

$$\det \underline{\underline{A}} = \det \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix} := \sum_P (-1)^P a_{1P(1)} a_{2P(2)} \cdots a_{nP(n)}$$

Diese Definition wird auch als *Leibniz-Formel* bezeichnet. Die Summe läuft dabei über alle Permutationen $P$ der Zahlen $1, 2, \ldots, n$. Als **Permutation** bezeichnet dabei jede Umordung der Reihenfolge. Z.B. für $n = 3$ hat man folgende Permutationen: 123, 132, 213, 231, 312, 321. Für die zweite Permutation 132 bedeutet dies genauer: $P(1) = 1$, $P(2) = 3$ und $P(3) = 2$.
Die Zahl der Permutationen wächst sehr schnell mit $n$ an. Sie ist gegeben durch[1] $n! := n(n-1)(n-2)\cdots 2 \cdot 1$. Damit ist $3! = 6$, aber für $n = 6$ haben wir schon $6! = 720$.

$(-1)^P$ bezeichnet das sog. **Signum (Vorzeichen)** der Permutation. Manchmal wird auch das Symbol $\mathrm{sign(P)}$ verwendet. Eine Permutation heißt **gerade** (**ungerade**), wenn sie durch eine gerade (bzw. ungerade) Anzahl an Paarvertauschungen aus $1, \ldots, n$ hervorgeht. Manchmal schreibt man auch $\mathrm{sgn}(P)$ statt $(-1)^P$.
Zum Beispiel ist die Permutation 132 ungerade, da sie aus 123 durch Vertauschung von 2 und 3 hervorgeht. Dagegen ist 231 gerade, denn hier benötigt man zwei Paarvertauschungen (z.B. erst 1 mit 2, was 213 liefert, und dann 1 mit 3).

Obige Definition der Determinanten lässt sich folgendermaßen interpretieren: Man hat alle möglichen Produkte zu bilden, bei denen aus jeder Zeile und jeder Spalte genau ein Element vorkommt. Diese Produkte sind dann zu addieren, wobei man das Vorzeichen aus dem Signum der entsprechenden Permutation beachten muß.

Speziell im Fall $n = 2$ gilt:

$$\det \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} = a_{11}a_{22} - a_{12}a_{21}\,.$$

Der erste Summand kommt von der identischen Permutation mit $P(1) = 1$ und $P(2) = 2$, der zweite von $P(1) = 2$, $P(2) = 1$. Letzere ist ungerade und daher ist $(-1)^P = -1$.

Im folgenden geben wir einige wichtige Eigenschaften von Determinanten ohne Beweis an:

1. $\det \underline{\underline{A}} = 0$, falls eine Zeile oder eine Spalte verschwindet (d.h. dort alle Elemente $= 0$ sind).

2. $\det \underline{\underline{A}} = 0$, falls zwei Zeilen oder zwei Spalten gleich oder proportional sind (d.h. es gibt $i$ und $k$ so dass $a_{ij} = \lambda a_{kj}$ bzw. $a_{ji} = \lambda a_{jk}$ für alle $j$ mit einer geeigneten reellen Zahl $\lambda$).

[1] Gesprochen: $n$ Fakultät

3. $$\det \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ \lambda a_{l1} & \cdots & \lambda a_{ln} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix} = \lambda \det \underline{\underline{A}},$$

   d.h. man kann einen gemeinsamen Faktor aus einer Zeile (oder Spalte) vor die Determinante ziehen.

4. Vertauscht man zwei Zeilen oder zwei Spalten, so wechselt das Vorzeichen der Determinante.

5. $\det\left(\underline{\underline{A}}^t\right) = \det \underline{\underline{A}}$.

6. $\det \underline{\underline{A}}$ ändert sich nicht, wenn zu einer Zeile (oder Spalte) das Vielfache einer *anderen* Zeile (bzw. Spalte) addiert wird.

Die ursprüngliche Definition ist nicht besonders gut zur praktischen Berechnung von Determinanten geeignet. Wir wollen hier daher ein praktischeres Verfahren vorstellen, die sog. Laplace-Entwicklung. Hierbei handelt es sich um ein rekursives Verfahren, bei dem die Berechnung von $n \times n$-Determinanten auf die Berechnung von $(n-1) \times (n-1)$-Determinanten zurückgeführt wird.
Als Vorbereitung geben wir zunächst zwei Definitionen an. Dazu betrachten wir eine $n \times n$-Matrix $\underline{\underline{A}} = (a_{ij})$.
Als **Minor** von $a_{ij}$ bezeichnet man dann folgende reelle Zahl:

$$M_{ij} := \det \begin{pmatrix} a_{11} & \cdots & \widehat{a_{1j}} & \cdots & a_{1n} \\ \vdots & & \vdots & \cdots & \vdots \\ \widehat{a_{i1}} & \cdots & \widehat{a_{ij}} & \cdots & \widehat{a_{in}} \\ \vdots & & \vdots & & \vdots \\ a_{n1} & \cdots & \widehat{a_{1j}} & \cdots & a_{nn} \end{pmatrix}$$

wobei die Elemente $\widehat{a_{ij}}$, die durch $\widehat{\ldots}$ gekennzeichnet sind, aus der Matrix zu streichen sind. Der Minor ist also die Determinante der Matrix, die aus $\underline{\underline{A}}$ durch Streichen der $i$-ten Zeile und $j$-ten Spalte entsteht.

$C_{ij} := (-1)^{i+j} M_{ij}$ heißt **Kofaktor** von $a_{ij}$.

Die **Laplace-Entwicklung** einer Determinante ist nun definiert durch

$$\boxed{\det \underline{\underline{A}} = \sum_{j=1}^{n} a_{ij} C_{ij} = \sum_{i=1}^{n} a_{ij} C_{ij},}$$

wobei die erste Summe einer Entwicklung nach der $i$-ten Zeile entspricht, die zweite der Entwicklung nach der $j$-ten Spalte.

Wir wollen dies am Beispiel einer $3 \times 3$-Matrix veranschaulichen. Eine Entwicklung nach der ersten Spalte liefert:

$$\begin{aligned}\det \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} &:= a_{11} \cdot \det \begin{pmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{pmatrix} - a_{21} \cdot \det \begin{pmatrix} a_{12} & a_{13} \\ a_{32} & a_{33} \end{pmatrix} \\ &+ a_{31} \cdot \det \begin{pmatrix} a_{12} & a_{13} \\ a_{22} & a_{23} \end{pmatrix}.\end{aligned}$$

*Bemerkung.* Neben der Determinante gibt es noch eine weitere einfache Zahl, die sich aus einer Matrix bilden läßt und die später noch eine wichtige Rolle spielen wird.
Die **Spur** einer $n \times n$-Matrix $\underline{\underline{A}}$ ist definiert durch

$$\operatorname{Spur} \underline{\underline{A}} := \sum_{j=1}^{n} A_{jj},$$

d.h. als Summe der Diagonalelemente.

## 2.4 Lineare Gleichungssysteme

In vielen praktischen Probleme hat es mit **linearen Gleichungssystemen** zu tun. Hier möchte man zunächst einmal wissen, ob sie überhaupt eine Lösung haben und wie man diese gegebenenfalls bestimmen kann. Dabei heißt ein Gleichungssystem linear, wenn keine Potenzen oder Produkte der Unbekannten auftreten.
Als ein typisches Beispiel betrachten wir folgendes System aus zwei Gleichungen mit zwei Unbekannten und seine Lösung:

$$\begin{array}{rcll} x + 3y & = & 6 & \Longrightarrow x + 3(2x - 5) = 6 \quad \Longrightarrow 7x = 21 \quad \Longrightarrow x = 3 \\ 2x - y & = & 5 & \Longrightarrow y = 2x - 5 \qquad\qquad\qquad\qquad\qquad\quad \Longrightarrow y = 1. \end{array}$$

Eine alternative Darstellung des Gleichungssystem mit Hilfe von Vektoren $\begin{pmatrix} x \\ y \end{pmatrix}$ und $\begin{pmatrix} 6 \\ 5 \end{pmatrix}$ im $\mathbb{R}^2$, die die Unbekannten $x$, $y$ bzw. Inhomogenitäten zusammenfassen, ist durch

$$\begin{pmatrix} 1 & 3 \\ 2 & -1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 6 \\ 5 \end{pmatrix}$$

gegeben. Die $2 \times 2$-Matrix, die hier auftritt, bezeichnet man auch als **Koeffizienten-Matrix**.
Allgemein lassen sich lineare Gleichungssysteme aus $m$ Gleichungen mit $n$ Unbekannten immer in der Form

$$\underline{\underline{A}} \cdot \underline{x} = \underline{b}$$

schreiben, mit dem Vektor $\underline{x} \in \mathbb{R}^n$, der die $n$ Unbekannten als Komponenten enthält, der Inhomogenität $\underline{b} \in \mathbb{R}^m$ und einer Koeffizientenmatrix $\underline{\underline{A}} \in \mathcal{M}(m, n)$.

Ist $m = n$ und kann $\underline{\underline{A}}$ invertiert werden, so kann man die Lösung mit Hilfe der Inversen bestimmen:

$$\underline{x} = \underline{\underline{A}}^{-1} \cdot \underline{b}$$

### 2.4.1 Gauß-Algorithmus

Wir betrachten ein lineares Gleichungssystem aus $n$ Gleichungen mit $n$ Unbekannten:

$$\begin{array}{ccccc} a_{11}x_1 & + \cdots & +a_{1n}x_n & = b_1 \\ \vdots & & \vdots & \vdots \\ a_{n1}x_1 & + \cdots & +a_{nn}x_n & = b_n \end{array}$$

bzw. $\underline{\underline{A}} \cdot \underline{x} = \underline{b}$. Dann gilt:
Ist $\det \underline{\underline{A}} \neq 0$, dann hat $\underline{\underline{A}} \cdot \underline{x} = \underline{b}$ eine eindeutige Lösung. Außerdem existiert $\underline{\underline{A}}^{-1}$. Im homogenen Fall, d.h. $\underline{b} = \underline{0}$ kann man die Lösung dann sofort angeben: In diesem Fall muß $\underline{x} = \underline{0}$ sein!
Dieses stellen wir mit Hilfe des folgenden Schemas symbolisch dar:

$$\left(\begin{array}{ccc|c} a_{11} & \cdots & a_{1n} & b_1 \\ \vdots & \ddots & \vdots & \vdots \\ a_{n1} & \cdots & a_{nn} & b_n \end{array}\right)$$

Ziel ist es nun, dieses Matrixschema durch Äquivalenzumformungen auf Dreiecksgestalt zu bringen. Mit *Äquivalenzumformungen* meinen wir dabei Operationen, die den Lösungsraum nicht verändern. Es handelt sich dabei um die bekannten Operationen, die man bei der Lösung "zu Fuß" (siehe das Beispiel zu Beginn dieses Kapitels) anwendet, also z.B. die Addition zweier Gleichungen oder die Multiplikation mit einer reellen Zahl.

Wir gehen nun folgendermaßen vor: Im ersten Schritt subtrahieren wir von allen Zeilen $i$ (außer der ersten!) die jeweils mit $a_{i1}/a_{11}$ multiplizierte erste Zeile. Hierdurch werden die ersten Elemente aller Zeile (außer der ersten!) zu Null gemacht. Man erhält dann folgendes Schema:

$$\left(\begin{array}{cccc|c} a_{11} & a_{12} & \cdots & a_{1n} & b_1 \\ 0 & a_{22} - \frac{a_{21}}{a_{11}}a_{12} & \cdots & a_{2n} - \frac{a_{21}}{a_{11}}a_{1n} & b_2 - \frac{a_{21}}{a_{11}}b_1 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & a_{n2} - \frac{a_{n1}}{a_{11}}a_{12} & \cdots & a_{nn} - \frac{a_{n1}}{a_{11}}a_{1n} & b_n - \frac{a_{n1}}{a_{11}}b_1 \end{array}\right)$$

Dies wiederholt man nun für die Untermatrix, die in der 2. Zeile und 2. Spalte beginnt etc. Am Ende hat man dann folgendes Schema vorliegen:

$$\left(\begin{array}{cccc|c} a_{11} & a_{12} & \cdots & a_{1n} & b_1 \\ 0 & \tilde{a}_{22} & \cdots & \tilde{a}_{2n} & \tilde{b}_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & \tilde{a}_{nn} & \tilde{b}_n \end{array}\right)$$

Die explizite Form der Elemente $\tilde{a}_{ij}$ geben wir hier nicht, sie könnte aber im Prinzip bestimmt werden.
Ausgeschrieben lautet die letzte Zeile $\tilde{a}_{nn}x_n = \tilde{b}_n$. Damit haben wir $x_n$ explizit bestimmt: $x_n = \frac{\tilde{b}_n}{\tilde{a}_{nn}}$. Dies können wir nun in die vorletzte Zeile einsetzen, so dass dort nur noch $x_{n-1}$ als Unbekannte explizit auftaucht usw. Damit lassen sich alle $x_j$ rekursiv ablesen.
Alternativ könnte man auch noch weitere Äquivalenzumformungen vornehmen und das Schema auf Diagonalform bringen:

$$\left(\begin{array}{cccc|c} a_{11} & 0 & \cdots & 0 & \breve{b}_1 \\ 0 & \breve{a}_{22} & \cdots & 0 & \breve{b}_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & \breve{a}_{nn} & \breve{b}_n \end{array}\right)$$

Hieraus kann man die Lösung direkt ablesen: $x_j = \breve{b}_j/\breve{a}_{jj}$. Um diese Diagonalform zu erreichen, muss man geeignete Vielfache der letzten Zeile von den anderen abziehen und so die letzte Spalte zu Null machen etc.

Bem.: Manchmal kann ein Diagonalelement in einem Zwischenschritt den Wert $0$ haben. Dann muß man Zeilen oder Spalten vertauschen, was u.U. einer Umnummerierung der Variablen entspricht.

Wir wollen das Vorgehen an einem konkreten Beispiel demonstrieren.

**Beispiel 2.4.1.**

$$\begin{array}{rcl} x + 2y + z &=& 3 \\ x + y - z &=& -1 \\ 2x - 2y + 4z &=& 10 \end{array}$$

Dies liefert folgendes Schema:

$$\left(\begin{array}{ccc|c} 1 & 2 & 1 & 3 \\ 1 & 1 & -1 & -1 \\ 2 & -2 & 4 & 10 \end{array}\right) \longrightarrow \left(\begin{array}{ccc|c} 1 & 2 & 1 & 3 \\ 0 & -1 & -2 & -4 \\ 0 & -6 & 2 & 4 \end{array}\right) \longrightarrow \left(\begin{array}{ccc|c} 1 & 2 & 1 & 3 \\ 0 & 1 & -2 & -4 \\ 0 & 0 & 14 & 28 \end{array}\right) .$$

Die letzte Zeile lautet explizit $14z = 28$ also $z = 2$. Setzt man dies in die vorletzte Zeile $y - 2z = -4$ ein, so folgt $y = 0$. Aus der ersten Zeile erhält man schließlich $x = 1$.
Alternativ kann man auch weitere Umformungen vornehmen und den Koeffiziententeil auf Diagonalgestalt bringen. Dabei geht man analog vor, wobei nun geeignete Vielfache der letzten Zeile von den anderen Zeilen subtrahiert werden. Außerdem kann man sich zunutze machen, dass die Multiplikation einer Zeile mit einer reellen Zahl ebenfalls eine Äquivalenzumformung darstellt. So erhält man dann folgende weitere Umformungsschritte: Dies liefert folgendes Schema:

$$\longrightarrow \left(\begin{array}{ccc|c} 1 & 2 & 1 & 3 \\ 0 & 1 & -2 & -4 \\ 0 & 0 & 1 & 2 \end{array}\right) \longrightarrow \left(\begin{array}{ccc|c} 1 & 2 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 2 \end{array}\right) \longrightarrow \left(\begin{array}{ccc|c} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 2 \end{array}\right) .$$

Man sieht, dass man nun die Lösung direkt in der rechten Spalte ablesen kann, wenn man den Koeffiziententeil in Diagonalform gebracht hat!

Zum Abschluss einige Bemerkungen:

1. Obiges Vorgehen funktioniert auch mit mehreren Inhomogenitäten gleichzeitig. Dies tritt z.B. beim Invertieren einer Matrix $\underline{\underline{A}}$ auf:

$$\left(\begin{array}{ccc|ccc} a_{11} & \cdots & a_{1n} & 1 & \cdots & 0 \\ \vdots & & \vdots & & \ddots & \\ a_{n1} & \cdots & a_{nn} & 0 & \cdots & 1 \end{array}\right) \longrightarrow \cdots \longrightarrow \left(\begin{array}{ccc|c} 1 & \cdots & 0 & \\ & \ddots & \vdots & \underline{\underline{A}}^{-1} \\ 0 & \cdots & 1 & \end{array}\right)$$

2. Die vorgestellte Lösungsstrategie kann auch auf allgemeinere Systeme aus $n$ Gleichungen mit $m$ Unbekannten (wobei $n \neq m$) angewendet werden.

3. Für quadratische Systeme ($n = m$) gilt: Ist $\det \underline{\underline{A}} \neq 0$, so hat $\underline{\underline{A}}\underline{x} = \underline{b}$ eine eindeutige Lösung. Außerdem existiert $\underline{\underline{A}}^{-1}$.
   Hieraus folgt, dass im Fall $\underline{b} = \underline{0}$ (homogenes Gleichungssystem) $\underline{x} = \underline{0}$ die eindeutige Lösung ist.

4. Ist in einer Zeile der Koeffiziententeil überall Null, das entsprechende Element des inhomogene Teils (rechte Spalte) aber nicht, so besteht offensichtlich ein Widerspruch. Das Gleichungssystem ist dann nicht lösbar.

5. Ist in eine Zeile identisch Null (inklusive der Inhomogenität), so kann man eine Variable beliebig wählen und das System hat unendlich viele Lösungen.

# Kapitel 3

# Funktionen und Differentiation

**Definition 3.0.1** (Funktion)**.**
Eine **Funktion** ist eine Vorschrift $f$, die jedem Element einer Menge $D$, dem **Definitionsbereich**, ein Element $y =: f(x)$ einer anderen Menge $Z$, der **Zielmenge**, zuordnet. Man spricht auch von einer **Abbildung von** $D$ **nach** $Z$ und bezeichnet diese mit

$$f: \ D \to Z$$

oder charakterisiert die Zuordnung durch

$$x \mapsto f(x) \qquad \text{mit } \ x \in D, \ f(x) \in Z.$$

Bemerkungen:

1. $x$ und/oder $f(x)$ können auch Vektoren sein!
2. Statt $y = f(x)$ schreibt man oft auch einfach $y = y(x)$. Die Schreibweise $y(x)$ soll andeuten, dass man in der Physik häufig nicht zwischen der Funktion $f$ (also der Zuordnungsvorschrift) und der abhängigen Variablen $y$ unterscheidet.

## 3.1 Eigenschaften von Funktionen

Funktionen $f : D \to Z$ lassen sich auf verschiedene Arten und Weisen charakterisieren. In einer Kurvendiskussion kann man folgende Eigenschaften untersuchen.

1. **Injektivität**: Die Funktion $f$ ist injektiv, wenn $f(x_1) \neq f(x_2)$ für alle $x_1, x_2 \in D$ mit $x_1 \neq x_2$.
2. **Surjektivität**: Die Funktion $f$ ist surjektiv, wenn die Bildmenge $f(D) := \{f(x)/x \in D\}$ mit der Zielmenge übereinstimmt: $f(D) = Z$.
3. **Bijektivität**: Eine Funktion, die injektiv und surjektiv ist, bezeichnet man als bijektiv oder auch eineindeutige Abbildung. Dies wird wichtig, wenn man die Umkehrfunktion bilden möchte.

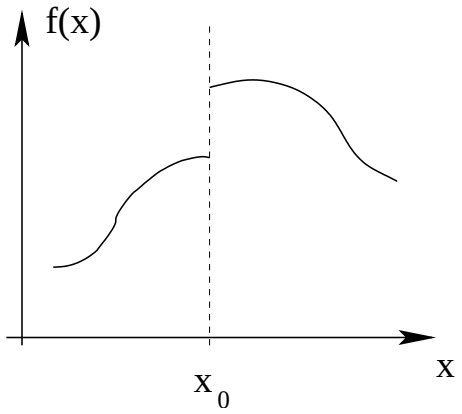



Abbildung 3.1.1: Die dargestellte Funktion ist unstetig bei $x_0$, da sie dort einen Sprung macht.

4. **Monotonie**: Die Funktion $f$ heißt monoton wachsend (bzw. fallend), wenn für alle $x_1 < x_2$ gilt $f(x_1) \leq f(x_2)$ (bzw. $f(x_1) \geq f(x_2)$). Gilt sogar die strenge Ungleichheit, so heißt die Funktion streng monoton wachsend (bzw. fallend).

5. **Extrema**: $f$ nimmt im Punkt $x_0$ ein (lokales) Maximum (bzw. Minimum) an, falls es ein Intervall $]x_0 - \delta, x_0 + \delta[$ um $x_0$ gibt ($\delta > 0$), in dem $f(x) \leq f(x_0)$ (bzw. $f(x) \geq f(x_0)$) ist. Das größte lokale Maximum (bzw. kleinste lokale Minimum) heißt absolutes Maximum (bzw. absolutes Minimum)[1].

6. **Asymptotik**: In der Physik interessiert man sich oft für das Verhalten von Funktionen für große oder kleine Argumente bzw. allgemeiner am Rand des Definitionesbereiches. Dieses lässt sich oft durch einfachere Funktionen charakterisieren. Z.B. sagt man, dass sich $f(x)$ für große $x$ asymptotisch wie $x^n$ verhält, wenn $f(x)/x^n \to 1$ für $x \to \infty$.

7. **Symmetrie (gerade/ungerade Funktionen)**: Eine Funktion $f$ heißt gerade, falls für alle $x \in D$ gilt: $f(-x) = f(x)$.
   Die Funktion heißt ungerade, falls für alle $x \in D$ gilt: $f(-x) = -f(x)$.
   Man beachte, dass in der Regel Funktionen weder gerade noch ungerade sind!

**Definition 3.1.1** (Stetigkeit)**.**
Eine stetige Funktion hat einen Kurvenverlauf ohne Sprung[2]. Der Graph der Funktion kann dann ohne abzusetzen gezeichnet werden (siehe Abb. 3.1.1).

Stellen, an denen eine Funktion ein "außergewöhnliches" Verhalten zeigt, bezeichnet man auch als **Singularitäten**.
Zwei wichtige Typen von singulärem Verhalten sind:

[1] Vorsicht! Hier müssen evtl. Funktionswerte an den Rändern des Definitionsbereichs separat betrachtet werden!
[2] Wir verzichten hier auf eine streng mathematische Definition zu Gunsten der intuitiven Vorstellung.

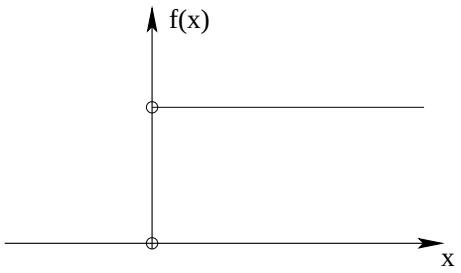



Abbildung 3.1.2: Der Graph der Heaviside'schen Sprungfunktion.

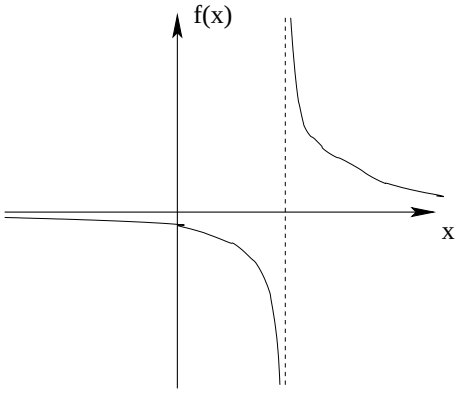



Abbildung 3.1.3: Funktion mit einer Singularität im engeren Sinne.

1. Unbestimmtheit: Hiermit meint man einen speziellen Fall von Nichtstetigkeit, den wir am Einfachsten anhand einer oft gebrauchten Funktion illustrieren, der sog. **Heaviside'schen Sprungfunktion**

$$\Theta(x) := \begin{cases} 1 & \text{für } x > 0 \\ 0 & \text{für } x < 0 \end{cases}.$$

   Man sagt: "$\Theta$ ist singulär bei $x = 0$". $\Theta(0)$ wird i.a. per Konvention festgelegt. Auf welchen Wert hängt meist von der Anwendung ab. Die gebräuchlichsten Konventionen sind $\Theta(0) = 1$, $\Theta(0) = 0$ und $\Theta(0) = 1/2$.
   Der Graph der Sprungfunktion ist in Abb. 3.1.2 dargestellt.

2. Unendlichkeit (Divergenz): Der zweite wichtige Typ von Singularitäten sind Stellen, an denen die Funktionen keinen endlichen Wert annehmen, d.h. divergieren. Dies sind Singularitäten im engeren Sinne.
   Ein Beispiel ist in Abb. 3.1.3 dargestellt, nämlich die Funktion $f(x) = \frac{1}{x-1}$. Sie ist offensichtlich für $x = 1$ nicht definiert, da dort der Nenner eine Nullstelle hat. Man sagt auch, dass $f$ bei $x = 1$ *singulär* ist bzw. genauer, dass $f$ dort einen **Pol** (bzw. eine **Polstelle**) hat.

In der Physik spielen Singularitäten eine wichtige Rolle, z.B. in der Theorie der Phasenübergänge. Diese lassen sich an Hand ihres Verhaltens in der Nähe von Singularitäten charakterisieren!

Bemerkung: Er gibt auch Singularitäten, die sich "beheben" lassen. Ein Beispiel ist die Funktion

$$f(x) = \frac{x-2}{x^2-3x+2} = \frac{1}{x-1}.$$

Formal hat der Nenner die Nullstellen $x = 1$ und $x = 2$, d.h. dort sollten Singularitäten vorliegen. Da die Nullstelle $x = 2$ von der entsprechenden Nullstelle des Zählers kompensiert wird, merkt man aber von ihr nichts. Die Funktion kann an der Stelle $x = 2$ stetig ergänzt werden durch die Festlegung $f(2) = 1$. Formal entspricht das dem Kürzen des Linearfaktors $x-2$. Man bezeichnet diese Singularität daher als **hebbare Singularität**.

## 3.2 Elementare Funktionen

Im folgenden wollen wir die wichtigsten Funktionenklassen, die in der Physik immer wieder vorkommen, vorstellen und ihre wesentlichen Eigenschaften aufzählen. Später und in den Übungen werden wir noch weitere wichtige Funktionen kennenlernen.

### 3.2.1 Potenzfunktion

Die Potenzfunktion ist definiert durch

$$f(x) = x^a.$$

$x$ bezeichnet man als **Basis** und $a$ als **Exponenten**.
Für Exponenenten $a = n \in \mathbb{N}$ sind die Potenzfunktionen gerade Funktionen, während sie für ungerade $n$ ungerade Funktionen sind (siehe Abb. 3.2.1).
Folgende Rechenregeln gelten für die Potenzfunktion (mit $x \in \mathbb{R}_0^+$ und $a, b \in \mathbb{R}$):

$$\boxed{x^a x^b = x^{a+b}, \qquad (x^a)^b = x^{ab}, \qquad x^{-a} = \frac{1}{x^a}.}$$

Außerdem gilt $x^0 = 1$, wobei aber $0^0$ nicht definiert ist!

Als **Polynom $n$-ter Ordnung** bezeichnet man Summen der Form

$$f_n(x) = \sum_{j=0}^{n} a_j x^j = a_0 + a_1 x + a_2 x^2 + \cdots + a_n x^n$$

mit der natürlichen Zahl $n \in \mathbb{N}$ und $a_j \in \mathbb{R}$.

Eine **rationale Funktion** ist Quotient zweier Polynome $f_m$ und $g_n$:

$$h(x) = \frac{f_m(x)}{g_n(x)}.$$

In den Übungen werden wir die sog. Partialbruchzerlegung diskutieren, mit der sich rationale Funktionen in eine Standardform bringen lassen.

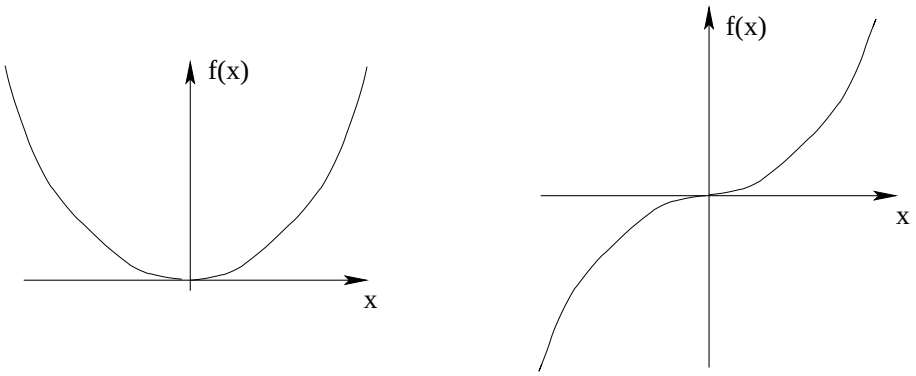



Abbildung 3.2.1: Graph der Potenzfunktion $f(x) = x^n$ für gerades $n$ (links) und ungerades $n$ (rechts).

### 3.2.2 Exponentialfunktion

Bei der Potenzfunktion ist die Basis variabel und der Exponent fest. Bei den Exponentialfunktionen ist es genau anders herum:

$$f_a : \mathbb{R} \to \mathbb{R}^+ \qquad \text{mit} \qquad f_a(x) = a^x.$$

Spezielle Exponentialfunktionen sind

$$f_2(x) = 2^x, \qquad f_{10}(x) = 10^x, \qquad f_e(x) = e^x = \exp(x),$$

wobei $e$ die sogenannte **Eulersche Zahl** ist: $e = 2.71828\ldots$.
Man bezeichnet die Exponentialfunktion zur Basis $e$ auch als *die* Exponentialfunktion oder als **natürliche Exponentialfunktion**.

Die Exponentialfunktionen sind streng monoton wachsend und stetig. Für $a > 1$ wachsen sie für große $x$ sehr schnell an und fallen für große negative $x$ sehr schnell ab (Abb. 3.2.2).
Spezielle Funktionswerte sind

$$f_a(0) = 1, \qquad f_a(1) = a, \qquad f_a(-1) = \frac{1}{a}.$$

Für die Exponentialfunktionen gelten folgende Rechenregeln:

$$\boxed{a^x b^x = (ab)^x, \qquad a^x a^y = a^{x+y}.}$$

Sie sind also durch folgende Funktionalgleichung charakterisiert:

$$\boxed{f_a(x) f_a(y) = f_a(x+y).}$$

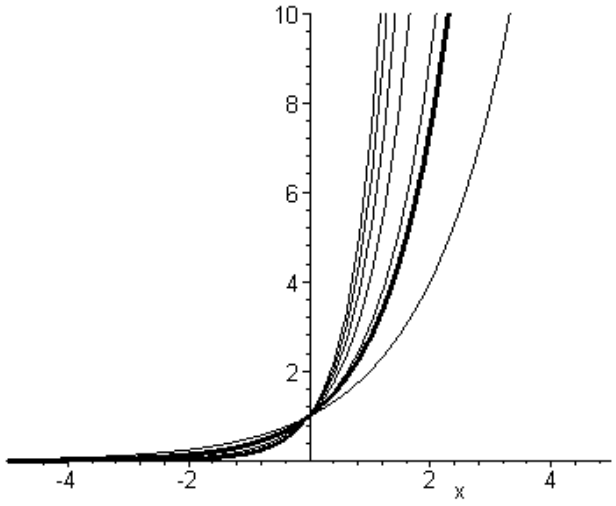



Abbildung 3.2.2: Graph der Exponentialfunktion $f_a$ für verschiedene Basen $a = 2, 3, \cdots, 7$. Die natürliche Exponentialfunktion ($a = e$) entspricht dem fett gezeichneten Graphen.

Die Exponentialfunktion ist eine der wichtigsten Funktionen im Rahmen der Naturwissenschaften. Sie beschreibt z.B. viele Wachstumsprozesse (Zellteilung, Kernspaltung) und Zerfallsprozesse (Radioaktivität).

Wir haben schon erwähnt, dass die Exponentialfunktionen sehr schnell anwachsen. Etwas präziser: Sie wachsen stärker an als jede Potenz, was sich mathematisch folgendermaßen ausdrücken lässt:

$$\lim_{x\to\infty} \frac{x^n}{f_a(x)} = 0 \qquad (a > 1, n \geq 0)$$

Entsprechend fallen die Funktionen für $x \to -\infty$ sehr schnell ab, was man sich mit $a^{-x} = 1/a^x$ leicht klar macht. Mit dieser Beziehung kann man sich auch überlegen, was im Fall $0 < a < 1$ passiert!

### 3.2.3 Logarithmus

Der Logarithmus ist die Umkehrfunktion der Exponentialfunktion.

**Definition 3.2.1** (Umkehrfunktion)**.**
Ist $f : X \to Y$ eine bijektive Abbildung, so existiert die **Umkehrfunktion** (oder auch **inverse Funktion**) $f^{-1} : Y \to X$ charakterisiert durch

$$\boxed{f\left(f^{-1}(y)\right) = y \qquad \text{und} \qquad f^{-1}\left(f(x)\right) = x \qquad \text{für alle } x \in X, y \in Y.}$$

Hierbei darf man die Umkehrfunktion $f^{-1}(x)$ nicht mit dem Kehrwert $(f(x))^{-1} = 1/f(x)$ verwechseln! In Bücher etc. wird hier die Notation oft nicht ganz sauber verwendet. Man muß dann

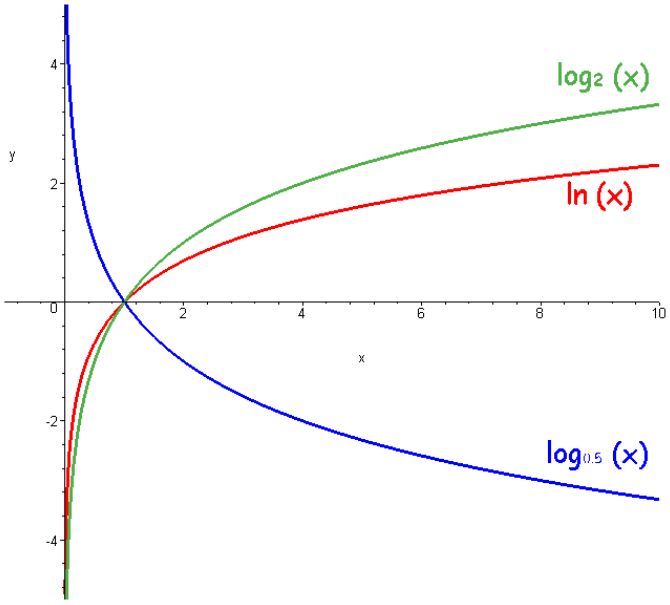



Abbildung 3.2.3: Graph der Logarithmusfunktion $\log_a$ für verschiedene Basen $a = 2$, $a = e$ und $a = 1/2$.

dem Zusammenhang entnehmen, was gemeint ist! Manchmal schreibt man daher auch $f^{-1}$ für die Umkehrfunktion.

Die Umkehrfunktion der Exponentialfunktion $f_a(x) = a^x$ (mit $a > 0$) bezeichnet auch man als **Logarithmus (zur Basis $a$)** $\log_a y : \mathbb{R}^+ \to \mathbb{R}$.
Die Logarithmusfunktion ist streng monoton wachsend und stetig. Ihren Graphen erhält man, wie allgemein für Umkehrfunktionen, durch Spiegelung des Graphen von $f_a$ an der Diagonalen $y = x$ (siehe Abb. 3.2.3). Es gilt

$$\boxed{x = \log_a(a^x), \qquad y = a^{\log_a y}.}$$

Für die Fälle $a = 10$ und $a = e$ haben sich spezielle Bezeichnungen eingebürgert:

$$\ln x := \log_e x, \qquad \log x := \log_{10} x.$$

Bem.: Wir hatten im vorigen Abschnitt gesehen, dass die Exponentialfunktion für große $x$ schneller als jede Potenz anwächst. Als Konsequenz daraus, wächst der Logarithmus langsamer als jede Potenz, d.h.

$$\lim_{x\to\infty} \frac{\log_a x}{x^\gamma} = 0 \qquad \text{für } \gamma > 0.$$

Auch die Rechenregeln ergeben sich direkt aus den entsprechenden Regeln für die Exponentialfunktion. Wir stellen sie hier kurz zusammen:

$$\boxed{\begin{aligned} \log_a(AB) &= \log_a A + \log_a B \\ \log_a(A^r) &= r \log_a A \\ \log_a\left(\frac{1}{A}\right) &= -\log_a A \end{aligned}}$$

mit $A, B > 0$ und $r \in \mathbb{R}$.
Mit Hilfe des Logarithmus können wir nun Exponentialfunktionen mit unterschiedlichen Basen ineinander umrechnen. Insbesondere gilt:

$$\boxed{a^x = \left(e^{\ln a}\right)^x = e^{x \ln a}.}$$

Dies zeigt, dass es tatsächlich genügt, die natürliche Exponentialfunktion[3] $\exp(x) = e^x$ zu kennen. In den Übungen werden wir sehen, dass eine analoge Aussage auch für den natürlichen Logarithmus $\ln x$ gilt!

### 3.2.4 Trigonometrische Funktionen

Zur Definition der trigonometrischen Funktionen betrachten wir einen beliebigen Punkt auf dem Einheitskreis. Diese lassen sich vollständig durch Angabe des Winkels $\phi$ charakterisieren.

Winkel werden in der Physik in der Regel im Bogenmaß (Radiant) gemessen. Die Umrechnung in Grad geschieht folgendermassen:

$$\text{Winkel in Grad} = \frac{360^\circ}{2\pi} \cdot \text{Winkel in Radiant}.$$

Daher entspricht der Winkel $\pi$ dem Winkel $\frac{360^\circ}{2\pi} \cdot \pi = 180^\circ$.

Die Koordinaten eines Punktes auf dem Einheitskreis sind durch $x = \cos\phi$ und $y = \sin\phi$ gegeben. Dies entspricht der klassischen geometrischen Definition der Winkelfunktionen

$$\cos\phi = \frac{\text{Ankathete}}{\text{Hypothenuse}}, \qquad \sin\phi = \frac{\text{Gegenkathete}}{\text{Hypothenuse}}$$

da die Länge der Hypothenuse hier gleich 1 ist.
Wenn wir nun den Winkel $\phi$ variieren, ändern sich auch $\cos\phi$ und $\sin\phi$. Wir können daher Sinus und Cosinus als Funktionen von $\phi$ auffassen.
Bevor wir die trigonometrischen Funktionen im Detail diskutieren, noch eine Definition:

**Definition 3.2.2** (periodische Funktionen)**.**
Eine Funktion $f$ heißt **periodisch mit Periode** $T$, wenn für alle $x$ gilt:

$$f(x+T) = f(x).$$

Man beachte, dass die Periode $T > 0$ nicht eindeutig ist, da mit $T$ immer auch $2T$, $3T$ etc. Perioden sind. Die kleinste Wert von $T > 0$, der obige Identität erfüllt, heißt *kleinste Periode* oder einfach nur *die Periode* von $f(t)$.

[3] Oder jede beliebige andere Exponentialfunktion!

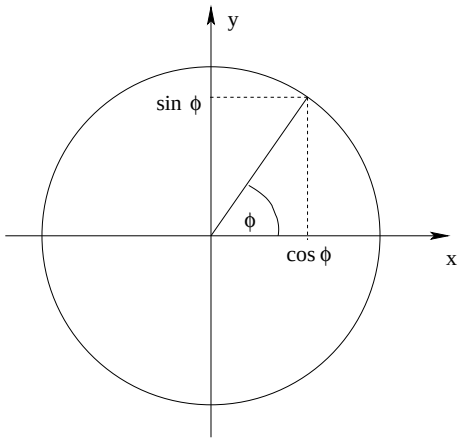



Abbildung 3.2.4: Definition von Sinus und Cosinus am Einheitskreis.

Abb. 3.2.5 zeigt den Graphen von $\cos\phi$. Der Cosinus ist als Funktion auf ganz $\mathbb{R}$ definiert. Aus der geometrischen Interpretation ist klar, dass die Funktionswerte im Intervall $[-1,1]$ liegen müssen. Sie zeigt außerdem, dass der Cosinus eine gerade Funktion ist, die zudem $2\pi$-periodisch ist. Zusammengefasst gilt also:

$$\begin{aligned} \cos &: \mathbb{R} \to [-1,1] \\ \cos\phi &= \cos(-\phi) \\ \cos\phi &= \cos(\phi + 2\pi) \\ \text{Nullstellen}: &\quad (2n+1)\,\frac{\pi}{2} \qquad (n \in \mathbb{Z}) \end{aligned}$$

Analoge Überlegungen können für den Sinus angestellt werden. Sein Graph ist in Abb. 3.2.5 gezeigt. Allerdings ist der Sinus eine ungerade Funktion. Es gilt:

$$\begin{aligned} \sin &: \mathbb{R} \to [-1,1] \\ \sin\phi &= -\sin(-\phi) \\ \sin\phi &= \sin(\phi + 2\pi) \\ \text{Nullstellen}: &\quad n\pi \qquad (n \in \mathbb{Z}) \end{aligned}$$

Zwischen der Sinus- und Cosinusfunktion bestehen verschiedene wichtige Zusammenhänge:

$$\boxed{\begin{aligned} 1 &= \sin^2\phi + \cos^2\phi, \\ \cos\phi &= \sin\left(\phi + \frac{\pi}{2}\right), \\ \sin\phi &= \cos\left(\phi - \frac{\pi}{2}\right). \end{aligned}}$$

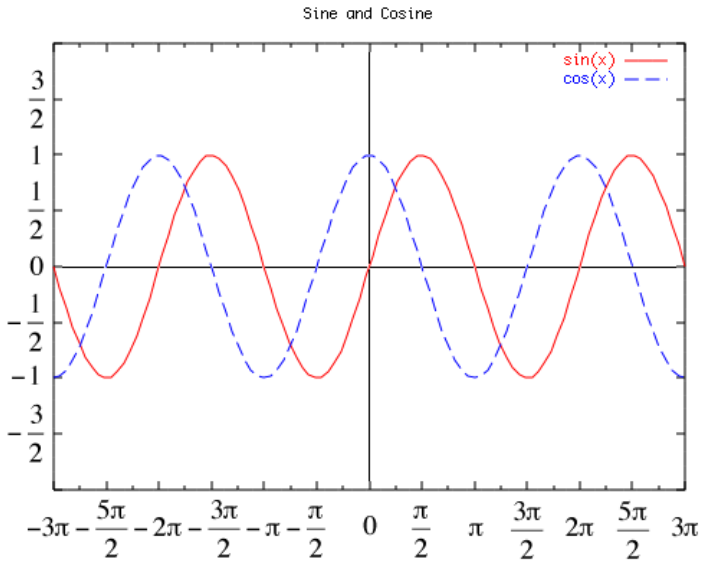



Abbildung 3.2.5: Graph der Cosinus- und Sinusfunktion.

Die erste Beziehung folgt sofort aus dem Satz von Pythagoras und der Definition am Einheitskreis. Die anderen beiden Identitäten besagen, dass Sinus und Cosinus um $\frac{\pi}{2}$ gegeneinander verschoben sind (siehe Abb. 3.2.5).

Aus Sinus und Cosinus lassen sich weitere trigonometrische Funktionen definieren, die in Anwendungen häufig auftreten. Zunächst ist dies der **Tangens**

$$\tan\phi := \frac{\sin\phi}{\cos\phi}$$

mit

$$\begin{aligned} \tan &: \mathbb{R}\setminus\left\{(2n+1)\frac{\pi}{2}\middle| n\in\mathbb{Z}\right\}\to\mathbb{R} \\ \tan\phi &= -\tan(-\phi) \\ \tan\phi &= \tan(\phi+\pi) \\ \text{Nullstellen}: &\quad n\pi \qquad (n\in\mathbb{Z}) \\ \text{Pole}: &\quad (2n+1)\frac{\pi}{2} \qquad (n\in\mathbb{Z}) \end{aligned}$$

Im Gegensatz zu Sinus und Cosinus ist der Tangens also $\pi$-periodisch. Seine Nullstellen stimmen mit denen des Sinus überein. Außerdem hat er Polstellen an den Nullstellen des Cosinus (siehe Abb. 3.2.6). Im Gegensatz zur Sinus- und Cosinusfunktion ist der Tangens nicht beschränkt, sein Wertebereich sind die gesamten reellen Zahlen $\mathbb{R}$.
Der **Cotangens** ist der Kehrwert des Tangens:

$$\cot\phi := \frac{1}{\tan\phi} = \frac{\cos\phi}{\sin\phi}$$

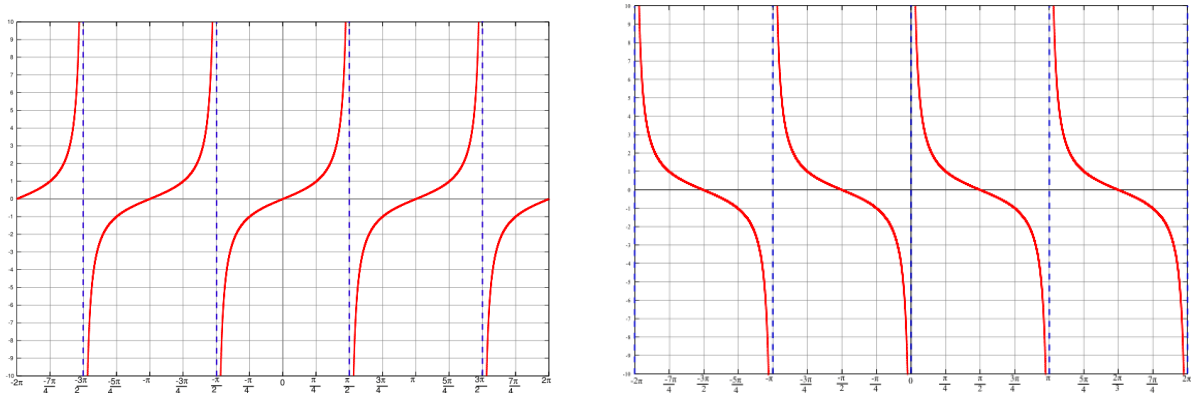


Abbildung 3.2.6: Graph der Tangens- (links) und Cotangensfunktion (rechts).

und hat die folgenden Eigenschaften:

$$\begin{aligned} \cot &: \mathbb{R} \setminus \{n\pi \,|\, n \in \mathbb{Z}\} \to \mathbb{R} \\ \cot\phi &= -\cot(-\phi) \\ \cot\phi &= \cot(\phi+\pi) \\ \text{Nullstellen}: &\quad (2n+1)\frac{\pi}{2} \qquad (n \in \mathbb{Z}) \\ \text{Pole}: &\quad n\pi \qquad (n \in \mathbb{Z}) \end{aligned}$$

Der Cotangens ist auch $\pi$-periodisch. Er hat Pole an den Nullstellen des Sinus und seine Nullstellen stimmen mit denen des Cosinus überein. Wie der Tangens ist auch der Cotangens nicht beschränkt, sein Wertebereich sind die gesamten reellen Zahlen $\mathbb{R}$.

Neben den trigonometrischen Funktionen sollte man auch deren Umkehrfunktionen gut kennen. Da die trigonometrischen Funktionen alle periodisch sind, muß man ihren Definitionsbereich einschränken, so dass sie bijektiv werden und man eine Umkehrfunktion überhaupt definieren kann. Dies führt auf die Hauptwerte der Funktionen. Im folgenden sind diese zusammengestellt:

$$\begin{aligned} \cos: \quad [0,\pi] &\longrightarrow [-1,1] \\ &\longleftarrow \quad : \arccos \quad (\text{Arcuscosinus}) \end{aligned}$$

$$\begin{aligned} \sin: \quad \left[-\tfrac{\pi}{2},\tfrac{\pi}{2}\right] &\longrightarrow [-1,1] \\ &\longleftarrow \quad : \arcsin \quad (\text{Arcussinus}) \end{aligned}$$

$$\begin{aligned} \tan: \quad \left]-\tfrac{\pi}{2},\tfrac{\pi}{2}\right[ &\longrightarrow \ ]-\infty,\infty \\ &\longleftarrow \quad : \arctan \quad (\text{Arcustangens}) \end{aligned}$$

$$\begin{aligned} \cot: \quad ]0,\pi[ &\longrightarrow \ ]-\infty,\infty[ \\ &\longleftarrow \quad : \operatorname{arccot} \quad (\text{Arcuscotangens}) \end{aligned}$$

Abb. 3.2.7 und 3.2.8 zeigen die Graphen der inversen trigonometrischen Funktionen.

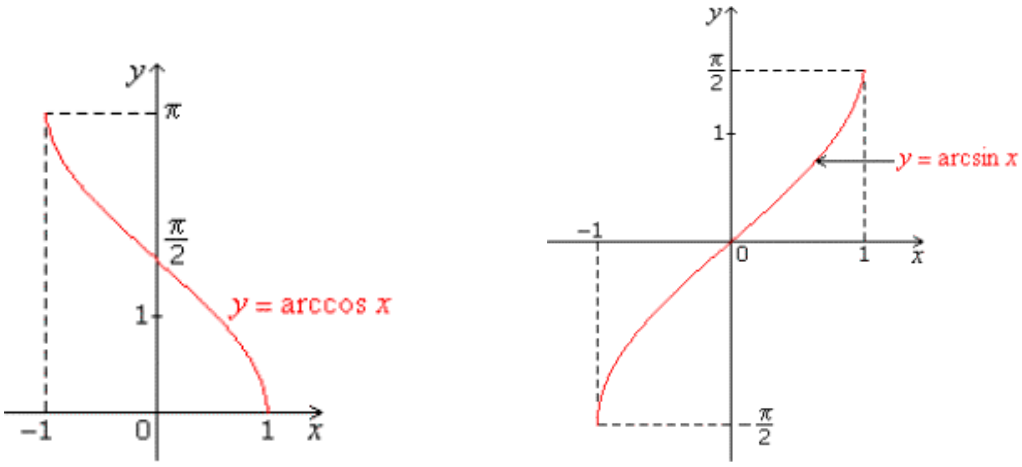



Abbildung 3.2.7: Graph der Umkehrfunktionen $\arccos$ (links) und $\arcsin$ (rechts) der Sinus- und Cosinusfunktion.

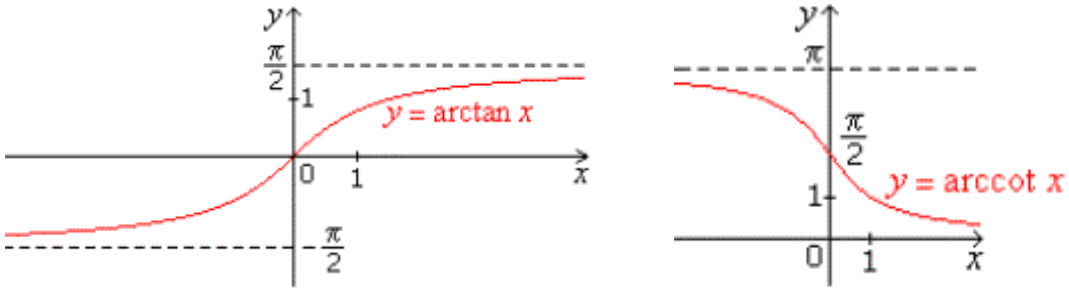



Abbildung 3.2.8: Graph der Umkehrfunktionen $\arctan$ (links) und $\operatorname{arccot}$ (rechts) der Tangens- und Cosinusfunktion.

## 3.3 Differentiation

Ein typisches physikalisches Problem ist die Bestimmung der Geschwindigkeit, wenn die zurückgelegte Entfernung als Funktion der Zeit bekannt ist. Dies gilt insbesondere, wenn die Geschwindigkeit nicht konstant ist. Man kann nach der momentanen Geschwindigkeit zu einem gegebenen Zeitpunkt fragen. Anschaulich entspricht dies der Bestimmung der Steigung des Graphen einer vorgegebenen Funktion in einem beliebigen Punkt.

**Definition 3.3.1** (Ableitung, Differentiation)**.**
Der Zuwachs[4] $y \to y + \Delta y$ einer Funktion $y(x)$ bei Veränderung des Arguments $x \to x + \Delta x$ ist ein Maß für die Veränderung einer Funktion (siehe Abb. 3.3.1):

$$\frac{\Delta y}{\Delta x} = \frac{y(x + \Delta x) - y(x)}{\Delta x}$$

Dieser **Differenzenquotient** ist also die Steigung der Geraden durch die Punkte $(x, y(x))$ und $(x + \Delta x, y(x + \Delta x))$ (siehe Abb. 3.3.1), der sog. **Sekante**. Im oben angegebenen Beispiel entspricht dies der Durchschnittsgeschwindigkeit im 'Zeitintervall' $[x, x + \Delta x]$.
Die **Ableitung** $y'(x)$ einer Funktion $y(x)$ ist dann die *momentane* Veränderung der Funktion, die durch den Grenzwert

$$y'(x) := \lim_{\Delta x \to 0} \frac{\Delta y}{\Delta x}$$

gegeben ist. Man schreibt für diesen **Differentialquotienten** auch symbolisch

$$y' = \frac{dy}{dx}$$

mit den **Differentialen** $dx$, $dy$.
Höhere Ableitungen sind rekursiv definiert: $y'' = (y')'$, $y^{(n+1)} = (y^{(n)})'$, etc. Man schreibt auch $y^{(n)}(x) = \frac{d^n y}{dx^n}$.

*Bemerkung:* Streng genommen ist $\frac{dy}{dx}$ kein Quotient zweier Größen $dy$ und $dx$ und kann nicht auseinandergerissen werden. Trotzdem macht man dies in der Physik häufig! Dahinter steckt die Vorstellung, daß man mit den Veränderungen $\Delta y$, $\Delta x$ rechnet und am Ende erst den Grenzübergang $\Delta x \to 0$ vollzieht.

**Beispiel 3.3.1.**

$$y(x) = x^n \quad \curvearrowright \quad y'(x) = nx^{n-1}$$

Dies kann man mit Hilfe der binomischen Entwicklung einsehen, denn es gilt $(x + \Delta x)^n = x^n + nx^{n-1}\Delta x + O((\Delta x)^2)$. Setzt man dies in den Differenzenquotienten ein, so erhält man $\frac{y(x+\Delta x)-y(x)}{\Delta x} = nx^{n-1} + O(\Delta x)$, woraus im Limes $\Delta x \to 0$ das Ergebnis folgt.

Im folgenden stellen wir die wichtigsten Rechenregeln zusammen, mit denen sich aus bekannten Ableitungen weitere Ableitungen bestimmen lassen:

[4] Dieser 'Zuwachs' kann auch negativ sein!

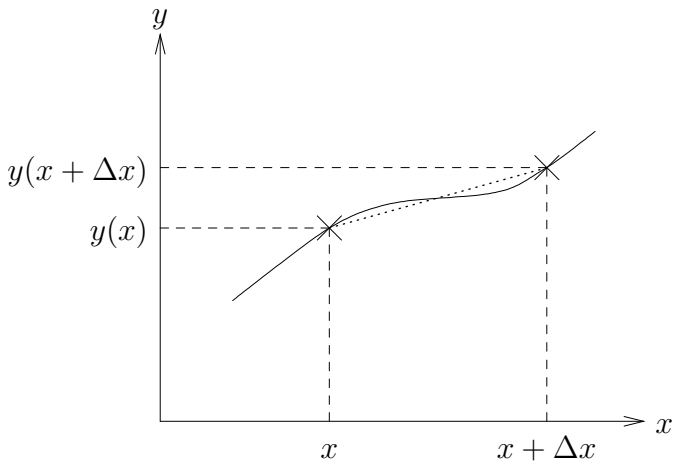



Abbildung 3.3.1: Zur Definition der Ableitung

1. $y(x) = u(x)v(x) \quad \Rightarrow \quad \boxed{y'(x) = u'(x)v(x) + u(x)v'(x)}$ (Produktregel)

2. $y(x) = \frac{u(x)}{v(x)} \quad \Rightarrow \quad \boxed{y'(x) = \frac{u'(x)v(x) - u(x)v'(x)}{v(x)^2}}$ (Quotientenregel)

3. $y = f(u), u = u(x) \quad \Rightarrow \quad y(x) = f\left(u\left(x\right)\right)$
   $\Rightarrow \; y' = \frac{dy}{dx} \underset{\text{“erweitern”}}{=} \frac{dy}{du}\frac{du}{dx} = f'\left(u\left(x\right)\right)u'(x)$ d.h. $\boxed{y' = f'\left(u\left(x\right)\right)u'(x)}$ (Kettenregel)

4. *Ableitung der Umkehrfunktion* $x = f^{-1}(y)$ *von* $y = f(x)$*:*

$$\boxed{\left(f^{-1}\right)'(x) = \frac{1}{f'(f^{-1}(x))}}, \qquad \text{bzw. in Kurzform} \qquad y' = \frac{dy}{dx} = \frac{1}{\frac{dx}{dy}}.$$

   Dies beweist man z. B. über die Kettenregel, da $f(f^{-1}(x)) = x$. Man beachte, dass man die Ableitung von $f'$ an der Stelle $f^{-1}(x)$ zu nehmen hat (siehe folgendes Beispiel).

**Beispiel 3.3.2.**
Die Umkehrfunktion der Exponentialfunktion $f(x) = e^x$ ist bekanntlich der (natürliche) Logarithmus $f^{-1}(x) = \ln x$. Da $f'(x) = e^x$ erhält man als Ableitung des Logarithmus

$$\frac{d\ln x}{dx} = \left(f^{-1}\right)'(x) = \frac{1}{f'(f^{-1}(x))} = \frac{1}{e^{\ln x}} = \frac{1}{x}.$$

Im Prinzip genügen die oben angegebenen Rechenregeln zusammen mit der Kenntnis einige weniger Ableitung, um fast alle wichtigen Funktionen differenzieren zu können. Die wichtigsten Funktionen, deren Ableitung man auswendig kennen sollte, sind in folgender Tabelle zusammengestellt:

| $f(x)$ | $f'(x)$ |
|---|---|
| $a$ | $0$ |
| $x^a$ | $ax^{a-1}$ |
| $e^x$ | $e^x$ |
| $\ln x$ | $\frac{1}{x}$ |
| $\sin x$ | $\cos x$ |
| $\cos x$ | $-\sin x$ |

Dabei ist $a \in \mathbb{R}$ eine Konstante.
Im Prinzip könnte man diese Liste noch verkürzen, da man z.B. die Ableitung des Logarithmus wie in Beispiel 4.3.2 aus der der Exponentialfunktion bestimmen könnte. Auch die Ableitung des Sinus könnte aus der des Kosinus hergeleitet werden:

$$\frac{d\sin x}{dx} = \frac{d}{dx}\cos\left(x - \frac{\pi}{2}\right) = \sin\left(x - \frac{\pi}{2}\right) = \cos x.$$

Dabei haben wir neben den bekannten Identitäten für trigonometrische Funktionen nur die Kettenregel benutzt.

**Definition 3.3.2** (Ableitung von Vektoren)**.**
Die **Ableitung von Vektoren** ist komponentenweise erklärt: Sei $\underline{a}(t) = \begin{pmatrix} a_1(t) \\ a_2(t) \end{pmatrix}$ eine vektorwertige Funktion und $\underline{a}(t + \Delta t) = \underline{a}(t) + \Delta\underline{a}(t)$ deren Änderung. Dann ist

$$\frac{d\underline{a}}{dt} := \lim_{\Delta t \to 0} \frac{\Delta\underline{a}(t)}{\Delta t} = \begin{pmatrix} \frac{da_1}{dt} \\ \frac{da_2}{dt} \end{pmatrix} = \begin{pmatrix} a_1'(t) \\ a_2'(t) \end{pmatrix}.$$

# 3.4 Potenzreihen und Taylor-Entwicklung

Viele Funktionen, die in der Physik auftreten, sind nicht in geschlossener Form angebbar, d.h. lassen sich nicht auf einfache Weise durch elementare Funktionen ausdrücken. Stattdessen gibt man sie in der Form von Reihen an.

## 3.4.1 Potenzreihen

**Definition 3.4.1** (Potenzreihen)**.**
Eine **Potenzreihe** mit Koeffenzienten $a_n$ hat die Form

$$f(x) = \sum_{n=0}^{\infty} a_n x^n.$$

Man beachte, dass alle auftretenden Potenzen $n$ ganzzahlig und nicht-negativ sind. Wir betrachten $x$ als Variable und haben somit eine Funktion von $x$ definiert.

Formal sind Potenzreihen Grenzwert der Folge $\left(\sum_{n=0}^{N} a_n x^n\right)_{N\in\mathbb{N}}$ der Partialsummen.
Die Reihe konvergiert evtl. nur in einem endlichen Bereich $|x| < R$. Dann heißt $R$ auch **Konvergenzradius** der Potenzreihe.

**Beispiel 3.4.1.** Aus der Schule dürfte ihnen die **geometrische Reihe**

$$\sum_{n=0}^{\infty} x^n = \frac{1}{1-x} \qquad \text{für} |x| < 1$$

bekannt sein. Diese kann man als Potenzreihe interpretieren, wenn man $x$ als variabel betrachtet. Die Koeffizienten dieser Reihe sind $a_n = 1$ für alle $n$. Bekanntermaßen hat die Reihe den Konvergenzradius 1.

Wir sehen an den Beispiel, dass man manchmal Reihen durch einfache Funktionen ausdrücken kann. Genauer gesagt, stimmt die geometrische Reihe im Intervall $]-1,1[$ mit der Funktion $\frac{1}{1-x}$ überein.

### 3.4.2 Taylor-Entwicklung

Es stellt sich daher die Frage, ob man eine beliebige Funktion $f(x)$ in Form einer Potenzreihe darstellen kann. Wir nehmen dazu an, dass dies möglich ist und versuchen die Koeffizienten $a_n$ zu bestimmen. Sei also $f(x) = \sum_{n=0}^{\infty} a_n x^n$. Da man Potenzreihe gliedweise differenziert[5], folgt:

$$\begin{aligned}
f(x) &= a_0 + a_1 x + a_2 x^2 + a_3 x^3 + \cdots + a_n x^n + \cdots \\
f'(x) &= a_1 + 2a_2 x + 3a_3 x^2 + \cdots + n a_n x^{n-1} + \cdots \\
f''(x) &= 2a_2 + 6a_3 x + \cdots + n(n-1) a_n x^{n-2} + \cdots \\
f'''(x) &= 6a_3 + \cdots + n(n-1)(n-2) a_n x^{n-3} + \cdots
\end{aligned}$$

Hieraus liest man ab:

$$f(0) = a_0, \quad f'(0) = a_1, \quad f''(0) = 2a_2, \quad f'''(0) = 6a_3, \cdots, f^{(n)}(0) = n! a_n, \cdots$$

und somit

$$a_n = \frac{1}{n!} f^{(n)}(0).$$

Somit erhalten wir die **Taylor-Reihe** (oder **Taylor-Entwicklung**) von $f$ um $x = 0$

$$\boxed{f(x) = \sum_{n=0}^{\infty} \frac{1}{n!} f^{(n)}(0) x^n.}$$

[5] Über die Frage, wann das erlaubt ist, wollen wir uns hier keine Gedanken machen! Zumindest muss die gliedweise differenzierte Reihe auch wieder konvergieren!

Eine große Klasse von Funktionen kann durch Taylor-Reihen dargestellt werden. Diese Funktionen heißen **analytisch** (in $x = 0$).
Dies gilt aber nicht für alle Funktionen. Ein Gegenbeispiel ist

$$f(x) = e^{-1/x^2}.$$

Für diese Abbildung gilt

$$f^{(n)}(0) = 0 \qquad \text{für alle } n.$$

Somit ist die Taylor-Reihe identisch Null!
Ein anderes Beispiel ist die Funktion $f(x) = \frac{1}{x}$, die in $x = 0$ gar nicht definiert ist, genauso wie alle ihre Ableitungen.

Eine wichtige Anwendung der Taylor-Entwicklung ist die Bestimmung einer einfachen Näherung einer Funktion in der Nähe von $x = 0$, z.B.

$$f(x) = f(0) + f'(0)x + \frac{1}{2}f''(0)x^2 + \mathcal{O}(x^3).$$

Dabei bezeichnet $\mathcal{O}(x^3)$ Terme, die für $x \to 0$ mindestens so schnell wie $x^3$ verschwinden. Eine allgemeine mathematische Definition hierfür lautet:

**Definition 3.4.2.**
Seien $f(x)$ und $g(x)$ vorgegebene Funktionen und $x_0 \in \mathbb{R} \cup \{\pm\infty\}$. Dann definiert man

$$f(x) = \mathcal{O}(g(x)) \quad :\Longleftrightarrow \quad \text{es gibt ein } M \in \mathbb{R} \text{ mit } \left|\frac{f(x)}{g(x)}\right| < M \ \text{ für } x \to x_0\,.$$

Für den Fall $\lim_{x\to x_0} f(x) = 0$ bedeutet dies, dass $f(x)$ für $x \to x_0$ mindestens so schnell verschwindet wie $g(x)$. Für den Fall $\lim_{x\to x_0} |f(x)| = \infty$, wächst $|f(x)|$ höchstens von derselben Ordnung wie $|g(x)|$.
Analog definiert man:

$$f(x) = o(g(x)) \quad :\Longleftrightarrow \quad \lim_{x\to x_0} \frac{f(x)}{g(x)} = 0,$$

d.h. $f(x)$ verschwindet schneller als $g(x)$ für $x \to x_0$, falls $\lim_{x\to x_0} f(x) = 0$. Für den Fall $\lim_{x\to x_0} |f(x)| = \infty$ wird $|g(x)|$ für $x \to x_0$ schneller wachsen als $|f(x)|$.

Oft ist es nützlich, die Taylor-Entwicklung von $f$ nicht um Null, sondern einen anderen Punkt $x_0$ zu betrachten. Diese können wir nun leicht bestimmen. Dazu betrachten wir die Funktion

$$g(\Delta x) := f(x_0 + \Delta x) = f(x),$$

d.h. $g$ ist um $x_0$ gegenüber $f$ verschoben: $g(x - x_0) = f(x)$. Wir entwickeln nun $g$ um Null:

$$g(\Delta x) = \sum_{n=0}^{\infty} \frac{1}{n!} g^{(n)}(0)(\Delta x)^n.$$

Nach Kettenregel gilt aber für alle $n$

$$g^{(n)}(0) = f^{(n)}(x_0)$$

und somit

$$\boxed{f(x) = \sum_{n=0}^{\infty} \frac{1}{n!} f^{(n)}(x_0)(x - x_0)^n.}$$

Dies ist die Taylor-Entwicklung von $f$ um $x_0$. Voraussetzung ist natürlich, dass $f$ in $x_0$ analytisch ist.
Im folgenden wollen wir einmal die wichtigsten Potenzreihendarstellung auflisten:

$$\begin{aligned}
\exp(x) &= e^x = \sum_{n=0}^{\infty} \frac{1}{n!} x^n \qquad (x \in \mathbb{R}) \\
\sin x &= \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n+1)!} x^{2n+1} \qquad (x \in \mathbb{R}) \\
\cos x &= \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n)!} x^{2n} \qquad (x \in \mathbb{R}) \\
\ln(1+x) &= \sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n} x^n \qquad (x \in ]-1, 1]) \\
(1+x)^p &= \sum_{n=0}^{\infty} \left( \begin{array}{c} p \\ n \end{array} \right) x^n \qquad (p \in \mathbb{R},\ x \in ]-1, 1[)
\end{aligned}$$

mit dem Binomialkoeffizienten

$$\left( \begin{array}{c} p \\ n \end{array} \right) := \frac{p(p-1)(p-2)\cdots(p-n+1)}{n!}.$$

Später werden wir noch sehen, dass die Darstellung einer Funktion als Taylor-Reihe i.a. Ausgangspunkt ihrer Erweiterung ins Komplexe ist. Für beliebige $z \in \mathbb{C}$ ist die Funktion dann dadurch definiert, dass man in der Taylor-Reihe $x$ durch $z$ ersetzt. Dies ist sinnvoll, da die Funktion so für reelle Argumente mit ihrer ursprünglichen Definition übereinstimmt und die Berechnung der Reihe im Prinzip nur die Grundrechenarten erfordert.

# Kapitel 4

# Integration

Eine häufig auftauchende Frage lautet: Wie bestimmt man $x(t)$ wenn $v(t) = \frac{dx}{dt}$ bekannt ist ?
In physikalischen Problemen ist dabei z.B. $v(t)$ der zeitliche Geschwindigkeitsverlauf einer Bewegung und $x(t)$ die bis zur Zeit $t$ zurückgelegte Strecke.
Die zur Beantwortung dieser Frage nötige Umkehrung der Differentiation bezeichnet man als **Integration**.

## 4.1 Stammfunktion

**Definition 4.1.1** (Stammfunktion)**.**
$F(x)$ heißt **Stammfunktion** der Funktion $f(x)$, wenn $F'(x) = f(x)$. Man schreibt auch

$$\boxed{F(x) = \int f(x)dx}$$

und bezeichnet dies als das **unbestimmte Integral** von $f$. In der Physik schreibt man auch häufig $\int dx f(x)$, was später bei mehrdimensionalen Integralen eine nützliche Konvention ist.

$F$ ist nicht eindeutig, denn mit $F(x)$ ist auch $F(x) + a$ mit einer beliebigen reellen Konstanten $a$ eine Stammfunktion. Genauer bezeichnet daher $\int f(x)dx$ die Menge aller Stammfunktionen von $f$. $a$ heißt auch **Integrationskonstante**.

Ähnlich wie die Ableitung können wir uns eine Tabelle mit den wichtigsten Stammfunktionen zusammenstellen:

| $f(x)$ | $F(x)$ |
|---|---|
| $x^r$ | $\frac{1}{r+1}x^{r+1}$ |
| $\frac{1}{x}$ | $\ln \lvert x \rvert$ |
| $e^x$ | $e^x$ |
| $\sin x$ | $-\cos x$ |
| $\cos x$ | $\sin x$ |

Dabei ist zu beachten, dass bei jeder Stammfunktion additiv noch Integrationskonstante $a$ auftritt. Im ersten Beispiel $x^r$ ist $r \in \mathbb{R} \setminus \{-1\}$. Ist $r \notin \mathbb{N}$, so muß $x > 0$ sein.

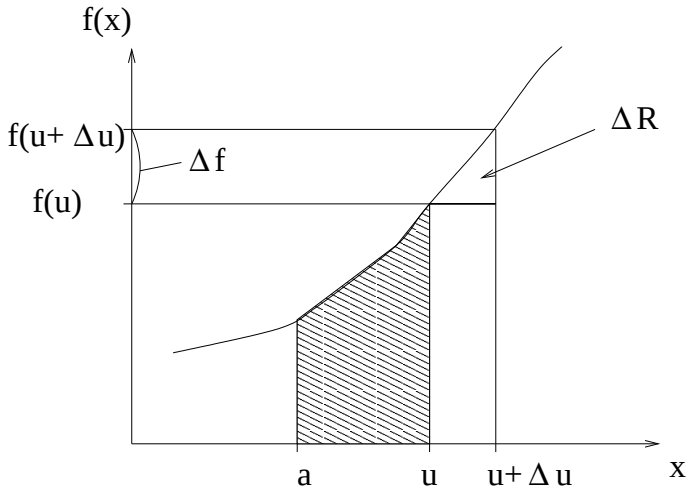



Abbildung 4.2.1: Das Integral

## 4.2 Bestimmtes Integral

Eine andere Motivation der Integration ist die Flächenberechnung. Man spricht auch von *bestimmten Integralen*. Für die Fläche die vom Graphen $f(x)$ der Funktion $f$ mit der $x-$Achse zwischen $x = a$ und $x = u$ eingeschlossene Fläche (schraffierter Bereich in Abb. 4.2.1) schreibt man auch:

$$F(u) = \int_a^u f(x)dx.$$

Man fragt sich nun: Wie sieht $F(u)$ aus, wenn $f(x)$ bekannt ist?
Auf den ersten Blick ist natürlich nicht offensichtlich, dass dieses Problem etwas mit dem im vorigen Abschnitt eingeführten Integral zu tun hat. Die Notation deutet aber einen solchen Zusammenhang schon an, den wir uns im folgenden klar machen wollen.
Zu Bestimmung der Funktion $F(u)$ vergrößern wir den Integrationsbereich um ein Stück $\Delta u$ (siehe Abb. 4.2.1):

$$\begin{aligned} F(u + \Delta u) &= F(u) + f(u)\Delta u + \Delta R, \\ \Delta R &\approx \Delta u \Delta f. \end{aligned}$$

Die Fläche $\Delta R$ haben wir nur sehr grob durch $\Delta u \Delta f$ approximiert. I.a. wird hier noch ein Koeffizient $\alpha$ auftreten, d.h. $\Delta R = \alpha \Delta u \Delta f$. In dem Beispiel in Abb. 4.2.1 wäre eine Approximation des Bereichs $\Delta R$ durch ein Dreieck (d.h. $\alpha = 1/2$) sicher genauer. Wir werden aber gleich sehen, dass es darauf gar nicht ankommt. Wichtig ist nur, dass $\Delta R$ gegen Null geht, wenn $\Delta u$ oder $\Delta f$ gegen Null gehen.
Nun ergibt sich durch einfache Umformung:

$$\frac{F(u + \Delta u) - F(u)}{\Delta u} = \frac{f(u)\Delta u + \Delta R}{\Delta u} \approx f(u) + \Delta f$$

und somit für $\Delta u \to 0$:

$$\lim_{\Delta u \to 0} \frac{F(u+\Delta u) - F(u)}{\Delta u} = f(u)$$

Dabei haben wir ausgenutzt, dass $\lim_{\Delta u \to 0} \Delta f = 0$.

Insgesamt haben wir also:

$$f(u) = \frac{dF(u)}{du} = F'(u),$$

d.h. $F$ ist eine Stammfunktion von $f$ bzw. $F = \int f(x)dx$.

Die Bedingung $F(u = a) = 0$ legt die Integrationskonstante fest. So kommt man zum **bestimmten Integral:**

$$\boxed{\int_a^u f(x)dx = F(u) - F(a) =: F(x)\Big|_a^u}$$

Das bestimmte Integral von $f$ zwischen $a$ und $u$ ist also eine Zahl!

Den Zusammenhang zwischen bestimmtem Integral und Stammfunktion bezeichnet man auch als den **Hauptsatz der Differential- und Integralrechnung**.

Bei der Berechung kann eine *beliebige* Stammfunktion gewählt werden, da die additive Konstante bei der Bildung der Differenz $F(u) - F(a)$ herausfällt.

**Beispiel 4.2.1.**
Als einfaches Beispiel betrachten wir die Funktion $x^n$. Das unbestimmte Integral ist gegeben durch

$$F(x) = \int x^n dx = \frac{x^{n+1}}{n+1} + c.$$

Das bestimmte Integral zwischen $a$ und $u$ ist

$$\int_a^u x^n dx = \frac{u^{n+1}}{n+1} - \frac{a^{n+1}}{n+1}.$$

Speziell für $a = 1$ und $u = 2$ ergibt sich

$$\int_1^2 x^n dx = \frac{2^{n+1} - 1}{n+1}.$$

Bei der Interpretation des bestimmten Integrals als eingeschlossene Fläche muss man vorsichtig sein, wenn die Funktion negativ werden kann. In dem Beispiel in Abb. 4.2.2 wäre die eingeschlossene Fläche durch $F_1 + F_2 + F_3$ gegeben, wobei $F_j > 0$. Das Integral liefert aber

$$\int_a^b f(x)dx = F_1 - F_2 + F_3.$$

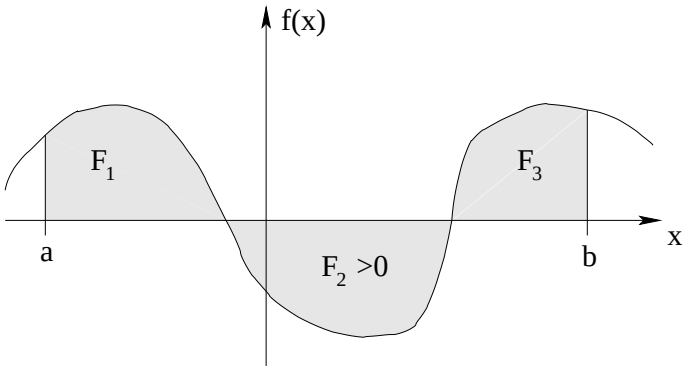



Abbildung 4.2.2: Zur Interpretation des bestimmten Integrals als Fläche.

Im folgenden stellen wir einige einfache Rechenregeln zusammen. Zunächst einmal ist die Integration *linear*, d.h. für beliebige (integrierbare) Funktionen $f$ und $g$ und reelle Zahlen $r$ gilt

$$\int_a^b (f(x) + rg(x))\, dx = \int_a^b f(x)dx + r\int_a^b g(x)dx.$$

Außerdem ist die Integration additiv, d.h.

$$\int_a^c f(x)dx + \int_c^b f(x)dx = \int_a^b f(x)dx \qquad (\text{mit } a < c < b).$$

Außerdem gilt

$$\int_a^b f(x)dx = F(b) - F(a) = -\left(F(a) - F(b)\right) = -\int_b^a f(x)dx,$$

wobei $F$ eine Stammfunktion von $f$ ist. Mit der letzten Regel überlegt man sich leicht, was bei der Additivität in dem Fall passiert, in dem $c$ nicht zwischen $a$ und $b$ liegt!

## 4.3 Integrationsverfahren

In der Praxis muß man eine gewisse Menge elementarer (unbestimmter) Integrale auswendig können. Aus diesen kann man sich durch Anwendung geeigneter Regeln viele andere Integrale herleiten. Sehr nützlich sind auch Integrationstabelle, z.B. das Buch von Gradstein/Ryshik. Heutzutage gibt es auch Computeralgebraprogramme wie *Mathematica* oder *Maple*, die sogar unbestimmte Integrale bestimmen können.
Im folgenden wollen wir einige nützliche Integrationsverfahren vorstellen.

### 4.3.1 Partielle Integration

Die partielle Integration ist gewissermaßen die Umkehrung der Produktregel der Differentiation. Sei $f(x) = g'(x)h(x)$, d.h. $f$ ist Produkt der Ableitung einer Funktion $g$, die wir kennen, und einer Funktion $h$. Da nach Produktregel $\frac{d}{dx}(gh) = g'h + gh'$, folgt:

$$\int f(x)dx = \int \left[\frac{d}{dx}(gh) - gh'\right] dx = gh - \int gh'dx$$

bzw. als bestimmtes Integral

$$\boxed{\int_a^b g'(x)h(x)dx = g(x)h(x)\Big|_a^b - \int_a^b g(x)h'(x)dx.}$$

Der Nutzen dieser Regel liegt darin, dass manchmal das Integral $\int gh'dx$ einfacher auszurechnen ist als $\int g'hdx$.
Wir wollen dies an einigen Beispielen illustrieren.

**Beispiel 4.3.1.**

1. Wir wollen $\int_a^b xe^x dx$ bestimmen. Dazu identifizieren wir $e^x$ mit $g$, also $g'(x) = e^x$, und $h(x) = x$, d.h. $h'(x) = 1$. Damit ergibt sich aus der Regel der partiellen Integration:

   $$\int_a^b xe^x dx = xe^x\Big|_a^b - \int_a^b 1 \cdot e^x dx = (xe^x - e^x)\Big|_a^b.$$

   Natürlich hätte man auch die Identifikation $g'(x) = x$ und $h(x) = e^x$ wählen können. Dann hätte aber die partielle Integration zu keiner Vereinfachung geführt, da immer höhere Potenzen von $x$ auftreten würden.

2. Als zweites Beispiel untersuchen wir ein Integral über eine trigonometrische Funktion, nämlich $\int_a^b \sin^2 x dx$. Hier hilft die Identifikation $g'(x) = h(x) = \sin x$ weiter. Dann ist $g(x) = -\cos x$ und $h'(x) = \cos x$:

   $$\begin{aligned}\int_a^b \sin^2 x dx &= \int_a^b \sin x \sin x dx = -\cos x \sin x\Big|_a^b - \int_a^b (-\cos^2 x)dx \\ &= -\cos x \sin x\Big|_a^b + \int_a^b (1 - \sin^2 x)dx \\ &= -\cos x \sin x\Big|_a^b + x|_a^b - \int_a^b \sin^2 x dx.\end{aligned}$$

   Hier haben wir beim Übergang zur zweiten Zeile die Beziehung $\sin^2 x + \cos^2 x = 1$ benutzt. Zunächst sieht es so aus, als ob wir uns im Kreis gedreht haben, da auf der rechten Seite wieder das gleiche Integral auftaucht, das wir ausrechnen wollten. Man beachte allerdings das andere Vorzeichen! Wir können daher den Term auf die andere Seite bringen und erhalten somit

   $$\int_a^b \sin^2 x dx = \frac{1}{2}\left[(x - \cos x \sin x)\Big|_a^b\right].$$

3. Als letztes Beispiel betrachten wir $\int_a^b \ln x dx$. Hier hilft die Identifikation $g'(x) = 1$ und $h(x) = \ln x$:

$$\int_a^b \ln x dx = x \ln x\Big|_a^b - \int_a^b x \cdot \frac{1}{x} dx = x \ln x\Big|_a^b - \int_a^b dx = (x \ln x - x)\Big|_a^b.$$

Die Stammfunktion von $\ln x$ ist also $x \ln x - x$.

### 4.3.2 Substitutionsregel

Die Substitutionsregel ist die Umkehrung der Kettenregel der Substitution.
Sei $y = g(x)$ invertierbar mit Umkehrfunktion $g^{-1}(y)$. Dann gilt:

$$\boxed{\int_a^b f\left(g(x)\right) dx = \int_{g(a)}^{g(b)} f(y) \frac{dg^{-1}(y)}{dy} dy.}$$

Dies sieht man folgendermaßen ein: Sei $h(x) := f\left(g(x)\right)$ und $H$ eine Stammfunktion von $h$:

$$H(x)\Big|_a^b = \int_a^b f\left(g(x)\right) dx.$$

Andererseits gilt auch

$$H(x)\Big|_a^b = H(g^{-1}(y))\Big|_{g(a)}^{g(b)},$$

da $g^{-1}$ die Umkehrfunktion von $g$ ist und somit z.B. $g^{-1}(g(a)) = a$. Weiter gilt nach dem Hauptsatz der Differential- und Integralrechnung und Kettenregel

$$H(g^{-1}(y))\Big|_{g(a)}^{g(b)} = \int_{g(a)}^{g(b)} H'\left(g^{-1}(y)\right) \frac{dg^{-1}(y)}{dy} dy = \int_{g(a)}^{g(b)} h\left(g^{-1}(y)\right) \frac{dg^{-1}(y)}{dy} dy,$$

wobei wir im letzten Schritt ausgenutzt haben, dass $H$ Stammfunktion von $h$ ist. Da $h(x) := f\left(g(x)\right)$, folgt hieraus die Behauptung.

Wir geben noch zwei äquivalente Formen der Substitionsregel an, die manchmal nützlich sind[1]:

$$\boxed{\int_a^b f(x) dx = \int_{g^{-1}(a)}^{g^{-1}(b)} f(g(x)) g'(x) dx.}$$

oder auch

$$\boxed{\int_{g(a)}^{g(b)} f(x) dx = \int_a^b f(g(x)) g'(x) dx.}$$

[1]Es genügt natürlich, eine der Regeln zu kennen, nämlich die Form, die man sich am leichtesten merken kann.

Die Form für unbestimmte Integrale

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

kann man sich leicht merken, wenn man $u'(x)$ als Differentialquotient $\frac{du}{dx}$ ausdrückt und dann quasi $dx$ "wegkürzt". Man beachte, dass in dem rechten Integral $u$ nicht mehr für eine Funktion steht, sondern die Integrationsvariable bezeichnet. Wir werden dies gleich in den Beispielen noch explizit sehen.
Die Substitutionsregel kann man in beide Richtungen (von links nach rechts oder von rechts nach links) anwenden. Manchmal ist es nützlich, durch Substitution mit einer geeigneten Funktion $u(x)$ zum scheinbar schwierigeren Integral auf der linken Seite überzugehen.

**Beispiel 4.3.2.**

1. $$\int e^{\sin x}\cos x dx \underset{\substack{u=\sin x\\ du=\cos x dx}}{=} \int e^u du = e^u = e^{\sin x}$$

   Offensichtlich bietet sich hier die Identifikation $u = \sin x$ an. Wegen $\frac{du}{dx} = \cos x$, ist $du = \cos x dx$.

2. $$\int \frac{u'(x)}{u(x)}dx \underset{du=u'dx}{=} \int \frac{du}{u} = \ln|u| = \ln|u(x)|.$$

   Ähnlich wie im ersten Beispiel haben wir hier "$dx$" quasi weggekürzt.

3. Ein Beispiel mit bestimmtem Integral:

   $$\int_1^7 \frac{1}{\sqrt{2x+2}}dx = \int_{g(1)}^{g(7)} \frac{1}{\sqrt{y}}\cdot\frac{1}{2}dy = \frac{1}{2}\int_4^{16} y^{-1/2} = \frac{1}{2}2y^{1/2}\Big|_4^{16} = 2.$$

   Hier haben wir mit $y = g(x) = 2x+2$ substituiert. Damit ist die Umkehrfunktion $g^{-1}(y) = \frac{1}{2}(y-2)$ und $\left(g^{-1}(y)\right)'(y) = \frac{1}{2}$.

### 4.3.3 Ableitung nach Parameter

Als weiteres nützliches Verfahren wollen wir einen Trick verstellen, der von R.P. Feynman häufig gewinnbringend eingesetzt wurde. Dieser *Feynman-Trick* beruht auf folgendem Ergebnis:

**Satz 4.3.1** (Ableitung nach einem Parameter)**.**
Wir betrachten das Integral

$$I(\alpha) := \int_a^b f(x,\alpha)dx,$$

bei dem der Integrand neben der Variablen $x$ von einem weiteren Parameter $\alpha$ abhängt, über den nicht integriert wird. Die so definierte Funktion $I(\alpha)$ von $\alpha$ ist unter relativ schwachen Voraussetzungen differenzierbar und es gilt:

$$I'(\alpha) = \frac{d}{d\alpha}\int_a^b f(x,\alpha)dx = \int_a^b \frac{\partial f(x,\alpha)}{\partial \alpha}dx.$$

Mit anderen Worten: Man kann die Ableitung nach $\alpha$ "unter das Integral ziehen". Dann wird sie zur partiellen Ableitung, da $f$ ja von zwei Variablen abhängt.

Wir wollen den Nutzen dieser Tatsache für die Berechnung von Integralen an einem Beispiel konkretisieren.

**Beispiel 4.3.3.**

$$\begin{aligned}\int_{-1}^1 xe^{\alpha x}dx &= \int_{-1}^1 \frac{\partial e^{\alpha x}}{\partial \alpha}dx \\ &= \frac{d}{d\alpha}\int_{-1}^1 e^{\alpha x}dx = \frac{d}{d\alpha}\left(\frac{1}{\alpha}e^{\alpha x}\big|_{-1}^1\right) = \frac{d}{d\alpha}\left(\frac{1}{\alpha}\left(e^{\alpha}-e^{-\alpha}\right)\right) \\ &= \frac{d}{d\alpha}\left(\frac{2}{\alpha}\sinh\alpha\right) = -\frac{2}{\alpha^2}\sinh\alpha + \frac{2}{\alpha}\cosh\alpha.\end{aligned}$$

Dabei haben wir in der ersten Zeile erkannt, dass sich der Integrand als Ableitung einer einfacheren Funktion, nämlich $f(x,\alpha) = e^{\alpha x}$ schreiben läßt, auf die wir den Satz von der Ableitung nach einem Parameter angewendet haben. Bei der Auswertung des Integrals haben wir dann die in den Übungen untersuchte hyperbolische Sinusfunktion $\sinh\alpha$ eingeführt, was die Berechnung der Ableitung im letzten Schritt etwas vereinfacht. Hierbei wurden die Produkt- und Quotientenregel verwendet.

In dem obigen Beispiel war relativ klar, nach welchem Parameter man abzuleiten hat. Die eigentliche Kunst liegt darin, geeignete Parameter einzuführen! Stellen Sie sich z.B. vor, wir hätten nur das Integral $\int_{-1}^1 xe^x dx$ berechnen müssen! Der eigentliche Trick bestünde dann darin zu erkennen, dass es viel einfacher ist, das allgemeinere Integral $\int_{-1}^1 xe^{\alpha x}dx$ zu bestimmen und am Ende dann $\alpha = 1$ zu setzen. Gerade darin bestand die Meisterschaft von Feynman!

## 4.4 Uneigentliche Integrale

In der Physik steht man häufig vor dem Problem, entweder Integrale über unbeschränkte Intervalle zu berechnen, bei denen eine oder beide Grenzen unendlich sind, oder aber dass der Integrand Polstellen hat. Man spricht dann auch von **uneigentlichen Integralen**. Im folgenden wollen wir uns ansehen, wie man in diesen Fällen vorgeht.

### 4.4.1 Integration über ein unbeschränktes Intervall

Wir betrachten eine Funktion $f(x)$ und fragen uns, ob das Integral $\int_a^\infty f(x)dx$ über das unbeschränkte Intervall $[a,\infty[$ existiert. Dazu fassen wir es als den Grenzwert

$$\int_a^\infty f(x)dx = \lim_{b\to\infty}\int_a^b f(x)dx$$

auf. Wir müssen also untersuchen, ob der Grenzwert überhaupt existiert. Eine offensichtliche Forderung, die $f(x)$ erfüllen muss, ist, dass die Funktion für $x\to\infty$ hinreichend schnell abfällt. Was das bedeutet, wollen wir an einem konkreten Beispiel untersuchen, das oft als Vergleichsfall herangezogen wird, nämlich die Potenzfunktion $x^r$.

**Beispiel 4.4.1.**

$$\begin{aligned}\int_a^\infty x^r dx &= \lim_{b\to\infty}\int_a^b x^r dx \underset{r\neq -1}{=} \lim_{b\to\infty}\left(\frac{1}{r+1}x^{r+1}\Big|_a^b\right)\\ &= \lim_{b\to\infty}\left[\frac{1}{r+1}\left(b^{r+1}-a^{r+1}\right)\right] = \frac{1}{r+1}\lim_{b\to\infty} b^{r+1} - \frac{1}{r+1}a^{r+1}.\end{aligned}$$

Der Grenzwert existiert nur, falls $r+1<0$ ist. Dann ist nämlich $b^{r+1}=\frac{1}{b^{|r+1|}}$, was für große $b$ gegen Null strebt. Für $r<-1$ gilt also

$$\int_a^\infty x^r dx = -\frac{1}{r+1}a^{r+1} = \frac{1}{|r+1|}\frac{1}{a^{|r+1|}}.$$

Für $r>-1$ ist $b^{r+1}$ divergent und das uneigentliche Integral existiert nicht.

Zur Vollständigkeit sei noch der Fall $r=-1$ angegeben:

$$\int_a^\infty \frac{1}{x}dx = \lim_{b\to\infty}\int_a^b \frac{1}{x}dx = \lim_{b\to\infty}\ln x\Big|_a^b.$$

Auch hier existiert das Integral nicht, da $\ln b$ für $b\to\infty$ divergiert.

Zusammenfassend können wir also feststellen, dass das uneigentliche Integral der Potenzfunktion $x^r$ über $[a,\infty[$ existiert, falls $r<-1$ ist.
Wie schon erwähnt, dient die Potenzfunktion als Vergleich, um die Existenz von uneigentlichen Integralen entscheiden zu können. Wir könnnen daher festhalten, dass $\int_a^\infty f(x)dx$ existiert, falls $f(x)$ für große $x$ schneller als $1/x$ abfällt!

Zum Abschluss noch eine wichtige Bemerkung zum Fall, dass beide Grenzen unendlich sind, d.h. für die Integration über $]-\infty,\infty[$. Bei solchen Integralen, die an beiden Grenzen uneigentlich sind, ist die Konvergenz an beiden Grenzen *unabhängig* erforderlich!
Betrachten wir hierzu das Beispiel $\int_{-\infty}^\infty \frac{x}{1+x^2}dx$. Offensichtlich gilt

$$\lim_{a\to\infty}\int_{-a}^a \frac{x}{1+x^2}dx = 0,$$

da der Integrand eine ungerade Funktion ist, die über das symmetrische Intervall $[-a, a]$ integriert wird. Allerdings existiert das Integral nicht, denn die unabhängige Konvergenz ist nicht erfüllt, da $\ln(1+x^2)$ Stammfunktion von $\frac{x}{1+x^2}$ ist (siehe das 2. Beispiel in Beispiel 5.3.2) und somit

$$\int_a^b \frac{x}{1+x^2} dx = \frac{1}{2}\left[\ln(1+b^2) - \ln(1+a^2)\right].$$

Dies divergiert offensichtlich für $b \to \infty$ als auch für $a \to -\infty$.

### 4.4.2 Integration über Polstellen

Wir betrachten nun eine Funktion, die eine Polstelle bei $x_0$ hat. Auf Grund der Additivität des Integrals können die Überlegungen leicht auf den Fall mehrerer Polstellen verallgemeinert werden. Als Beispiel dient uns wieder die Potenzfunktion $x^r$, die für $r < 0$ einen Pol bei $x_0 = 0$ hat. Dazu definieren wir

$$\int_0^b x^r dx = \lim_{a \to 0+} \int_a^b x^r dx.$$

Dabei bedeutet $a \to 0+$, dass man den Grenzwert bildet, indem man sich nur von der rechten Seite, also $a > 0$ der Null nähert. Eine andere Schreibweise hierfür ist $a \searrow 0$.

Eine analoge Rechnung wie in Abschnitt 4.4.1 liefert dann, dass $\int_a^b x^r dx$ für $a \to 0+$ konvergiert, falls $r > -1$, und divergiert, falls $r \leq -1$.

Für den Fall mehrerer Polstellen oder in Kombination mit einer Integration über unbeschränkte Intervalle gelten wieder ähnliche Aussagen wie in Abschnitt 4.4.1, d.h. die Konvergenz muss unabhängig vorliegen!

# Kapitel 5

# Komplexe Zahlen

Beim Addieren, Subtrahieren, Multiplizieren, Dividieren und Potenzieren von reellen Zahlen bleiben wir im Bereich der reellen Zahlen. Beim Wurzelziehen stoßen wir auf einen neuen Zahlentyp.

## 5.1 Definition und Rechenregeln

Problem: $x^2 = -1$ hat keine reelle Lösung $x$

Man definiert daher[1] (Euler 1777): $i := \sqrt{-1}$.
$i$ wird als **imaginäre Einheit** bezeichnet.

Eine allgemeine **imaginäre Zahl** ist dann von der Form $b \cdot i$ mit reellem $b$, also z.B. $2i$, $\pi i$, etc.

Wir übertragen nun die üblichen Rechenregeln der reellen Zahlen auf die imaginären Zahlen (unter Beachtung von $i^2 = -1$):

$$b_1 i + b_2 i = (b_1 + b_2)i$$
$$(b_1 i)(b_2 i) = b_1 b_2 i^2 = -b_1 b_2$$
$$i^3 = i^2 i = -i \qquad \text{etc.}$$

**Definition 5.1.1** (komplexe Zahlen)**.**
Allgemeine **komplexe Zahlen** sind von der Form

$$z = x + iy$$

mit reellen Zahlen $x$ und $y$. $x$ bezeichnet man auch als den **Realteil** von $z$ und $y$ als den **Imaginäerteil**. Man schreibt dann auch:

$$z = \mathrm{Re}(z) + i\mathrm{Im}(z).$$

[1] Man beachte, dass hier streng genommen die "positive" Wurzel gemeint ist.

Wie für die rein imaginären Zahlen übertragen wir auch für die komplexen Zahlen die bekannten Rechenregeln. Man erhält so den Körper $\mathbb{C}$ der komplexen Zahlen.

$$\begin{aligned} \text{Addition: } z_3 &= z_1 + z_2 = (x_1 + iy_1) + (x_2 + iy_2) = (x_1 + x_2) + i(y_1 + y_2) \\ & \quad \mathrm{Re}(z_3) = \mathrm{Re}(z_1) + \mathrm{Re}(z_2) \quad \text{und} \quad \mathrm{Im}(z_3) = \mathrm{Im}(z_1) + \mathrm{Im}(z_2) \end{aligned}$$

$$\begin{aligned} \text{Multiplikation: } z_1 z_2 &= (x_1 + iy_1)(x_2 + iy_2) = (x_1 x_2) + i(x_1 y_2) + i(y_1 x_2) + i^2(y_1 y_2) \\ &= (x_1 x_2 - y_1 y_2) + i(x_1 y_2 + x_2 y_1) \end{aligned}$$

Speziell für reelles $\lambda$ gilt:

$$\lambda z = \lambda(x + iy) = (\lambda x) + i(\lambda y)$$

**Definition 5.1.2** (Komplexe Konjugation)**.**
$z^* := x - iy$ heißt **komplex-konjugiert** zu $z = x + iy$, d.h. die Operation $^*$ bedeutet Vorzeichenwechsel beim Imaginärteil. Manchmal schreibt man auch $\bar{z}$ statt $z^*$.
Offensichtlich gilt: $\mathrm{Re}(z) = \frac{1}{2}(z + z^*)$ und $\mathrm{Im}(z) = \frac{1}{2i}(z - z^*)$

Für das Produkt einer komplexen Zahl $z$ mit ihrem komplex-konjugierten gilt

$$z \cdot z^* = (x + iy)(x - iy) = x^2 + y^2.$$

Es ist also immer reell und größer gleich Null.

Nützlich ist das Komplex-Konjugierte auch bei der Division durch komplexe Zahlen:

$$\frac{1}{z} = \frac{z^*}{z \cdot z^*} = \frac{x - iy}{x^2 + y^2} = \frac{x}{x^2 + y^2} - i\frac{y}{x^2 + y^2}.$$

Hiermit ist die Division in $\mathbb{C}$ auf eine Multiplikation zurückgeführt.

**Definition 5.1.3** (Betrag)**.**
Man definiert den **Betrag** $|z|$ einer komplexen Zahl $z$ durch $\boxed{|z| = \sqrt{zz^*}}$.
Für ihn gelten analoge Rechenregeln wie für den Betrag von reellen Zahlen, z.B. die Dreiecksungleichung.

## 5.2 Komplexe Ebene

Es bietet sich an, Real- und Imaginärteil als Komponenten eines zweidimensionalen Vektors zu interpretieren. Man kommt so zu der zweidimensionalen Darstellung in der sog. **komplexen Ebene** (siehe Abb. 5.2.1):

$$z = x + iy \in \mathbb{C} \to \begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^2,$$

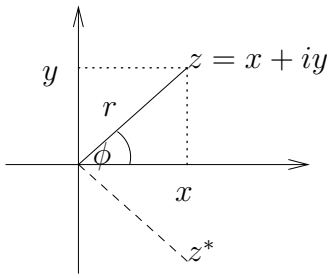



Abbildung 5.2.1: Die komplexe Ebene. Die Darstellung wird auch als **Argand-Diagramm** bezeichnet.

wobei nun $\begin{pmatrix} x \\ y \end{pmatrix}$ ein zweidimensionaler Vektor ist. Nun kann man wieder zu ebenen Polarkoordinaten $x = r\cos\phi$ und $y = r\sin\phi$ übergehen und erhält

$$z = x + iy = r(\cos\phi + i\sin\phi).$$

Eine einfache geometrische Überlegung liefert $r = \sqrt{x^2 + y^2}$, d.h. $r = |z|$. Man bezeichnet den Betrag von $z$ in diesem Zusammenhang auch manchmal als **Modul** oder **Modulus** von $z$ und schreibt $r = \text{mod}(z)$.
Den Winkel $\phi$ in der Polarkoordinatendarstellung bezeichnet man auch als **Argument** von $z$: $\phi = \text{arg}(z)$.
Insgesamt hat man daher folgende Darstellung einer komplexen Zahl:

$$\boxed{z = |z|(\cos\phi + i\sin\phi).}$$

Das Argument wird wie bei den ebenen Polarkoordinaten über $\phi = \arctan\frac{y}{x}$ bestimmt. Hierbei ist aber Vorsicht geboten, da der Wertebereich[2] des $\arctan$ das Intervall $[-\frac{\pi}{2}, \frac{\pi}{2}]$ ist. I.a. sind daher Zusatzüberlegungen zur Bestimmung des Winkels nötig und man darf nicht blind der Ausgabe eines Taschenrechners vertrauen!
Streng genommen gilt nur $\tan\phi = \frac{y}{x}$ und man muß die Periodizität $\tan(\phi + \pi) = \tan\phi$ des Tanges berücksichtigen. Wir wollen dies an einem Beispiel verdeutlichen.
Dazu betrachten wir $z = -1 - i$. In der komplexen Ebene liegt $z$ im 3. Quadranten, d.h. es gilt $\pi \leq \phi \leq \frac{3\pi}{2}$. Es gilt aber $\arctan\frac{-1}{-1} = \arctan 1 = \frac{\pi}{4}$. Auf Grund der $\pi$-Periodizität des Tangens können wir hierzu Vielfache von $\pi$ addieren, ohne den Wert des Tangens zu ändern. Um in den 3. Quadranten zu gelangen, müssen wir $\pi$ addieren, d.h. es gilt $\text{arg}(-1 - i) = \frac{5\pi}{4}$.

Auch mit diesem Verfahren ist $\phi$ nicht eindeutig bestimmt, da wir immer Vielfache $n = 0, \pm1, \pm2, \ldots$ von $2\pi$ addieren können. Dies entspricht dann $n$ Umrundungen des Ursprungs. Die Konvention ist aber, dass $\phi \in [0, 2\pi[$. Dies bezeichnet man als den Hauptwert. Manchmal ist es auch zweckmässig, $\phi \in ]-\pi, \pi]$ zu wählen.

[2]Genauer: Der Wertebereich des Hauptzweiges.

## 5.3 Eulersche Formel

Mit komplexen Zahlen kann man komplexe Funktionen bilden: $f(z) \in \mathbb{C}$, z.B. $f(z) = 2z^2 + z$.

**Beispiel 5.3.1.** Wir betrachten im folgenden ein für die Praxis sehr wichtiges Beispiel, die komplexwertige Exponentialfunktion.
Für reelle Argumente $x$ ist die Exponentialfunktion durch ihre Reihenentwicklung gegeben (für eine Definition, siehe Kap. 3.2.2):

$$\exp(x) = e^x = 1 + \frac{x}{1!} + \frac{x^2}{2!} + \frac{x^3}{3!} + \cdots = \sum_{n=0}^{\infty} \frac{1}{n!} x^n.$$

Betrachten wir diese Reihe für ein imaginäres Argument $i\phi$ mit reellem $\phi$, dann gilt

$$\begin{aligned} e^{i\phi} &= 1 + \frac{i\phi}{1!} + \frac{(i\phi)^2}{2!} + \frac{(i\phi)^3}{3!} + \cdots \\ &= 1 + i\phi - \frac{1}{2!}\phi^2 - i\frac{1}{3!}\phi^3 + \frac{1}{4!}\phi^4 + i\frac{1}{5!}\phi^5 + \cdots \\ &= 1 - \frac{1}{2!}\phi^2 + \frac{1}{4!}\phi^4 - \cdots + i\left(\phi - \frac{1}{3!}\phi^3 + \frac{1}{5!}\phi^5 - \cdots\right) \\ &= \cos\phi + i\sin\phi, \end{aligned}$$

wobei wir im letzten Schritt die Potenzreihen von $\sin$ und $\cos$ verwendet haben.

Dies liefert die **Eulersche Formel**:

$$\boxed{e^{i\phi} = \cos\phi + i\sin\phi}$$

d.h. die komplexe Zahl $e^{i\phi}$ mit $\phi \in \mathbb{R}$ liegt auf dem Einheitskreis (da $|e^{i\phi}|^2 = \cos^2\phi + \sin^2\phi = 1$) mit Winkel $\phi$.

Wir können daher für reelle $\phi$ den Real- und Imaginärteil der Exponentialfunktion bestimmen:

$$\begin{aligned} \mathrm{Re}(e^{i\phi}) &= \cos\phi, \\ \mathrm{Im}(e^{i\phi}) &= \sin\phi. \end{aligned}$$

Mit der Eulerschen Formel können wir nun eine beliebige komplexe Zahl darstellen als

$$\boxed{z = |z|e^{i\phi}.}$$

Hiermit macht man sich die folgenden Rechenregeln für den Betrag und das Argument klar:

$$z_1 \cdot z_2 = |z_1|e^{i\phi_1} \cdot |z_2|e^{i\phi_2} = |z_1 z_2| e^{i(\phi_1+\phi_2)}.$$

Speziell für das Komplex-Konjugierte folgt:

$$z^* = |z|(\cos\phi - i\sin\phi) = |z|(\cos(-\phi) + i\sin(-\phi)) = |z|e^{-i\phi}$$

Allgemein gilt folgende Aussage: Ist $f(x)$ für alle $x \in \mathbb{R}$ reellwertig, so gilt: $f(z^*) = (f(z))^*$.

Unter Benutzung der Rechenregeln für die Exponentialfunktion, die analog für komplexe Argumente gelten, können wir die Eulersche Formel auf beliebige $z = x + iy$ verallgemeinern:

$$e^z = e^{x+iy} = e^x e^{iy} = e^x \left(\cos y + i \sin y\right).$$

Man kann nun auch die Winkelfunktionen durch die komplexe Exponentialfunktion darstellen. Da $e^{i\phi} = \cos\phi + i\sin\phi$ und $e^{-i\phi} = \cos(-\phi) + i\sin(-\phi) = \cos\phi - i\sin\phi$, folgt

$$\boxed{\cos\phi = \frac{1}{2}(e^{i\phi} + e^{-i\phi}) \qquad \text{und} \qquad \sin\phi = \frac{1}{2i}(e^{i\phi} - e^{-i\phi}).}$$

Diese Identitäten sind enorm nützlich. Zum Beispiel lassen sich mit ihnen die bekannten Additionstheoreme für die trigonometrischen Funktionen einfach beweisen. Letztlich führt man sie auf die Eigenschaften der Expontialfunktion zurück. Dies werden wir in den Übungen genauer untersuchen.

## 5.4 Wurzeln

Lösungen von $z = w^n$ bzw. $z^{1/n} = w$ heißen **n-te Wurzeln** von $z$.

Die Zahl der Lösungen hängt offensichtlich von $n$ ab:

$$\begin{aligned} w^2 = 1 \quad &\Longrightarrow \quad w = 1,\ w = -1, \\ w^4 = 1 \quad &\Longrightarrow \quad w^2 = 1,\ w^2 = -1 \quad \Longrightarrow \quad w = 1,\ w = -1,\ w = i,\ w = -i. \end{aligned}$$

Allgemein können wir die allgemeine Lösung für dieses Problem elegant mit Hilfe der Polardarstellung bestimmen.

$$z^{1/n} \;=\; \left(|z|e^{i\phi}\right)^{1/n} = |z|^{1/n} e^{i\phi/n} = \sqrt[n]{|z|}\left(\cos\frac{\phi}{n} + i\sin\frac{\phi}{n}\right).$$

Die Nichteindeutigkeit des Winkels $\phi$,

$$e^{i\phi} = e^{i(\phi+2\pi k)} \qquad (k = 0, \pm 1, \pm 2, \ldots)$$

der nur bis auf Vielfache von $2\pi$ bestimmt ist, führt auf verschiedene Wurzeln, die für verschiedene Werte von $k$ auftreten können (solange $\frac{\phi+2\pi k}{n} \in [0, 2\pi[$).
Somit finden wir, dass $z = w_k^n$ mit

$$w_k := \sqrt[n]{|z|}\left(\cos\frac{\phi + 2\pi k}{n} + i\sin\frac{\phi + 2\pi k}{n}\right) \qquad \text{mit } k = 0, 1, \ldots, n-1,$$

d.h. es gibt $n$ verschiedene $n$-te Wurzeln. Es ist wichtig, dies immer im Hinterkopf zu behalten. Später werden Sie sehen, dass in physikalischen Problemen nicht immer die reelle Lösung die

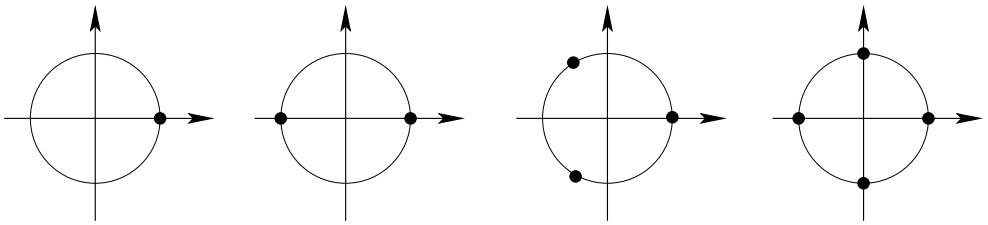



Abbildung 5.4.1: Lage der Einheitswurzeln für $n = 1, 2, 3, 4$.

interessanteste ist!

Insbesondere gibt es auch $n$ sogenannte **$n$-te Einheitswurzeln**

$$e^{i2\pi k/n} \qquad (k = 0, 1, \ldots, n-1).$$

Sie liegen auf dem Einheitskreis (siehe Abb. 5.4.1).

Es sei davor gewarnt, die aus dem Reellen bekannten Rechenregeln für Wurzel einfach ins Komplexe zu übertragen. Dies geht, gerade auf Grund der Nichteindeutigkeit der Wurzeln, oft schief! In den Übungen werden wir dies näher beleuchten.

# Kapitel 6

# Differentialgleichungen

Bei der Integration haben wir es mit dem Problem zu tun, eine Funktion $y(x)$ aus ihrer bekannten Ableitung $y'(x) = f(x)$ zu bestimmen. Als Verallgemeinerung der Integration werden uns im folgenden immer wieder sogenannte **Differentialgleichungen (DGL)** begegnen.

In erster Linie werden wir es mit zwei Typen zu tun haben, den sog.
DGL 1. Ordnung: $y'(x) = f(x, y)$,
DGL 2. Ordnung: $y''(x) = f(x, y, y')$.

Die fundamentale Gleichung der Mechanik, die *Newtonsche Bewegungsgleichung*

$$m\ddot{x}(t) = F(x, t)$$

ist ein Beispiel für eine DGL 2. Ordnung. Dabei ist $\ddot{x} = \frac{d^2x}{dt^2}$ die Beschleunigung und $F(x, t)$ die zur Zeit $t$ auf die am Ort $x$ befindliche Masse $m$ wirkende Kraft.

## 6.1 Klassifikation von DGL

Differentialgleichungen lassen sich auf unterschiedliche Arten klassifizieren.

- **gewöhnliche DGL**: Eine gewöhnliche DGL enthält keine partiellen Ableitungen, es gibt nur eine unabhängige Variable.
- **partielle DGL**: Eine partielle DGL spezifiziert Funktionen von mehreren Variablen. Sie enthält partielle Ableitungen nach den verschiedenen unabhängigen Variablen.
- **Ordnung der DGL**: Die Ordnung einer DGL ist durch die Ordnung der höchsten vorkommenden Ableitung gegeben.
- **explizite DGL**: Explizite DGL haben die Form $y^{(n)}(x) = f(x, y, y', \ldots, y^{(n-1)})$, d.h. sie drücken die höchste vorkommende Ableitung als Funktion der anderen Ableitungen aus.
- **implizite DGL**: Implizite DGL haben die Form $f(x, y, y', \ldots, y^{(n)}) = 0$.

- **lineare DGL**: In linearen DGL treten keine Potenzen von $y$, $y', \ldots$ auf, ebensowenig Produkte wie $yy'$ oder $y'y''$ etc. Potenzen der unabhängigen Variablen $x$ sind aber erlaubt. Ein Beispiel einer linearen DGL ist $y' + x^2 + y\cos x = 0$.

- **nichtlineare DGL**: Nichtlineare DGL enthalten Potenzen oder Produkte von $y, y', \ldots$, z.B. $yy'' + (y')^2 = 0$.

- **Anfangswertproblem**: Bei einem Anfangswertproblem sucht man die Lösung einer DGL, die zusätzlich die sog. **Anfangsbedingung** $y(x_0) = y_0$ erfüllt, wobei $x_0$ und $y_0$ vorgegebene Konstanten sind. Dies tritt häufig bei physikalischen Fragestellungen auf, wo z.B. $x_0$ die Anfangszeit und $y_0$ der Anfangsort sein kann. Bei DGL höherer Ordnung sind i.a. weitere Anfangsbedingungen vorgegeben, wie die Anfangsgeschwindigkeit $y'(x_0) = v_0$.

*Bemerkung:* I.a. wird die Lösung einer DGL nicht eindeutig sein! Wie beim Integrieren treten Integrationskonstanten auf, und zwar in der Regel $n$ Stück, wenn $n$ die Ordnung der DGL ist. Diese werden dann durch die Anfangsbedingungen fixiert.

## 6.2 Lineare DGL mit konstanten Koeffizienten

Als erstes wichtiges Beispiel betrachten wir nun eine **lineare DGL 2. Ordnung mit konstanten Koeffizienten**:

| | | | | |
|---|---|---|---|---|
| $(i)$ | $ay'' + by' + cy$ | $=$ | $0$ | homogene DGL |
| $(ii)$ | $ay'' + by' + cy$ | $=$ | $f(x)$ | inhomogene DGL |

mit Konstanten $a, b, c \in \mathbb{R}$.

Wir können o.B.d.A.[1] annehmen, dass $a = 1$ ist. Dies können wir immer erreichen, indem wir die Gleichung durch $a$ dividieren. Dabei ist der Fall $a = 0$ uninteressant, da es sich dann nicht um eine DGL 2. Ordnung handelt!

Eine wichtige Anwendung dieser Gleichung haben Sie schon in der Experimentalphysik kennengelernt. Die homogene Gleichung tritt bei der Beschreibung einer harmonischen gedämpften Schwingung auf. Wirkt zusätzlich eine äussere Kraft auf den Oszillator, so führt dies auf die inhomogene Gleichung. In beiden Fällen ist $y(x)$ die Auslenkung aus der Ruhelage zur Zeit $x$.

### 6.2.1 Homogener Fall

Wir untersuchen daher zunächst die homogene Gleichung

$$y'' + by' + cy = 0.$$

Zur Lösung machen wir den Ansatz

$$y(x) = Ae^{\lambda x},$$

[1] D.h. "ohne Beschränkung der Allgemeinheit"

den wir in die DGL einsetzen. Als "Ansatz" bezeichnet man allgemein eine Lösung, deren Struktur man gewissermaßen geraten hat, obwohl man einige Details der Lösung noch nicht kennt. In unserem Fall ist ein Versuch mit einer Exponentialfunktion naheliegend, da diese bei Ableitung reproduziert wird. Wir müssen nun überprüfen, ob wir die noch nicht bestimmten Konstanten $A$ und $\lambda$ so wählen können, dass $Ae^{\lambda x}$ tatsächlich eine Lösung der Gleichung wird.
Zunächst benötigen wir die Ableitungen $y'(x) = \lambda A e^{\lambda x}$ und $y''(x) = \lambda^2 A e^{\lambda x}$. Setzt man diese in die DGL ein, so folgt

$$Ae^{\lambda x}(\lambda^2 + b\lambda + c) = 0.$$

Da $A \neq 0$ (wir sind schließlich an einer nichttrivialen Lösung interessiert!) und $e^{\lambda x} \neq 0$, folgt

$$\lambda^2 + b\lambda + c = 0$$

und somit

$$\lambda_\pm = -\frac{b}{2} \pm \sqrt{\frac{b^2}{4} - c}.$$

Da $b, c$ reell sind, ist

$$\sqrt{\frac{b^2}{4} - c} \qquad \begin{cases} \text{reell}, & \text{falls } \frac{b^2}{4} - c \geq 0, \\ \text{imaginär}, & \text{falls } \frac{b^2}{4} - c < 0. \end{cases}$$

Deshalb sind $\lambda_\pm$ für $\frac{b^2}{4} - c < 0$ komplex[2]:

$$\lambda_\pm = -\frac{b}{2} \pm i\sqrt{c - \frac{b^2}{4}},$$

d.h. insbesondere ist $\lambda_+ = (\lambda_-)^*$.
Für $\frac{b^2}{4} - c > 0$ haben wir zwei verschiedene reelle Lösungen $\lambda_+$ und $\lambda_-$.
Für $\frac{b^2}{4} - c = 0$ ist $\lambda_+ = \lambda_-$. Dann existiert aber eine weitere Lösung $y = xe^{\lambda x}$ (nachrechnen!). Diesen Fall wollen wir im folgenden nicht näher betrachten.

Wir haben also zwei Lösungen der homogenen Gleichung gefunden, nämlich

$$y_\pm(x) = A_\pm e^{\lambda_\pm x},$$

wovon man sich leicht durch Nachrechnen überzeugt. Dabei ist $\lambda_\pm$ durch die Koeffizienten $b$ und $c$ der DGL bestimmt, während $A_\pm$ beliebige Integrationskonstanten sind!

Da die DGL *linear* ist, hat sie die angenehme Eigenschaft, dass die Summe $y = y_+ + y_-$ zweier Lösungen $y_+$, $y_-$ wieder eine Lösung liefert[3]:

$$\begin{aligned} y'' + by' + cy &= (y_+ + y_-)'' + b(y_+ + y_-)' + c(y_+ + y_-) \\ &= (y_+'' + by_+' + cy_+) + (y_-'' + by_-' + cy_-) = 0, \end{aligned}$$

[2] Vergleiche auch mit Aufgabe 6g der Übungen!
[3] Dies gilt leider nicht allgemein!

wobei im letzten Schritt beide Klammerausdrücke separat verschwinden, da $y_+$, $y_-$ Lösungen sind.

Zusammenfassend lautet also die allgemeine Lösung der homogenen Gleichung im Fall $\frac{b^2}{4} \neq c$

$$\boxed{y(x) = A_+ e^{\lambda_+ x} + A_- e^{\lambda_- x}}$$

mit zwei Integrationskonstanten $A_\pm$, die durch die Anfangsbedingungen $y(x_0) = y_0$ und $y'(x_0) = v_0$ festgelegt werden und $\lambda_\pm = -\frac{b}{2} \pm \sqrt{\frac{b^2}{4} - c}$.

Ist $\frac{b^2}{4} = c$, so lautet die allgemeine Lösung

$$y(x) = A_1 e^{-\frac{b}{2}x} + A_2 x e^{-\frac{b}{2}x}.$$

Wir sehen, dass in beiden Fällen zwei unabhängige Integrationskonstanten auftreten. Da es sich um eine DGL 2. Ordnung handelt, haben wir die allgemeinste Lösung gefunden.

Wenn $y$ (z.B. als eine Auslenkung aus der Ruhelage) aus physikalischen Gründen reell sein muß, können wir im Fall $\frac{b^2}{4} < c$ den Realteil (oder auch den Imaginärteil) der obigen Lösung nehmen. Auf Grund der Linearität der Gleichung und der Bildung des Realteiles bzw. der Ableitung sind diese im Falle reeller Koeffizienten $b, c$ auch eine Lösung der DGL, denn aus $y'' + by' + cy = 0$ folgt

$$\begin{aligned} 0 &= \text{Re}\,(y'' + by' + cy) \\ &= \text{Re}(y'') + b\text{Re}(y') + c\text{Re}(y) \\ &= (\text{Re}\; y)'' + b\,(\text{Re}\; y)' + c\text{Re}(y), \end{aligned}$$

wobei wir beim Übergang zur zweiten Zeile die Linearität der Realteilbildung ausgenutzt haben und die Tatsache, dass $b$ und $c$ reell sind, und beim Übergang zur dritten Zeile, dass man Differentiation und Realteilbildung vertauschen kann, da die Ableitung einer komplexwertigen Funktion gerade als Summe der Ableitungen von Real- und Imaginärteil definiert ist.

Somit können wir also im Fall $\frac{b^2}{4} < c$ aus der komplexen Lösung die neuen reellen Lösungen

$$y_1(x) = \text{Re}(y(x)) \qquad \text{und} \qquad y_2(x) = \text{Im}(y(x))$$

gewinnen. Explizit ergibt sich

$$\begin{aligned} y_1(x) &= \text{Re}\left(A_+ e^{\lambda_+ x} + A_- e^{\lambda_- x}\right) \\ &= A_1 e^{-\frac{b}{2}x} \cos \omega x, \end{aligned}$$

mit $A_1 = A_+ + A_-$ und $\omega = \sqrt{\left|\frac{b^2}{4} - c\right|} = \sqrt{c - \frac{b^2}{4}} = \text{Im}(\lambda_+)$ bzw. für den Imaginärteil

$$\begin{aligned} y_2(x) &= \text{Im}\left(A_+ e^{\lambda_+ x} + A_- e^{\lambda_- x}\right) \\ &= A_2 e^{-\frac{b}{2}x} \sin \omega x, \end{aligned}$$

mit $A_2 = A_+ - A_-$.
Die allgemeinste reelle Lösung im Fall $\frac{b^2}{4} < c$ ist dann gegeben durch

$$\boxed{y(x) \;=\; A_1 e^{-\frac{b}{2}x}\cos\omega x + A_2 e^{-\frac{b}{2}x}\sin\omega x}$$

mit den beiden (reellen) Integrationskonstanten $A_1$ und $A_2$, die durch die Anfangsbedingungen festgelegt werden.

Wir sehen, dass sich die Lösungen für die Fälle $\frac{b^2}{4} > c$ und $\frac{b^2}{4} < c$ qualitativ unterscheiden. Im ersten Fall haben wir als Lösung reine Exponentialfunktionen, während die Lösungen im zweiten Fall periodische Funktionen mit der Periode $T = 2\pi/\omega$ enthalten. Man spricht auch von quasi-periodischen Lösungen.

Wir wollen noch zwei wichtige Spezialfälle erwähnen, die in der Physik immer wieder auftreten. Dies ist zum Einen der Fall $b = 0$ und $c > 0$. In diesem Fall tritt die lineare DGL häufig in folgender Form auf:

$$y'' + \omega^2 y = 0,$$

mit $\omega^2 = c$. Diese Gleichung heißt auch **Schwingungsgleichung**. Deren allgemeine Lösung sollte man unbedingt kennen:

$$y(x) = A\cos\omega x + B\sin\omega x$$

mit den beiden Integrationskonstanten $A$ und $B$.

Der andere wichtige Spezialfall ist die lineare DGL 1. Ordnung, also der Fall $a = 0$, den wir bisher ausgeschlossen hatten. Diese lässt sich dann so schreiben:

$$y' = \lambda y.$$

Eine solche Gleichung tritt z.B. bei der Beschreibung des radioaktiven Zerfalls auf. Die allgemeine Lösung ist

$$y(x) = Ae^{\lambda x}.$$

Ist $\lambda > 0$ so beschreibt die Lösung **exponentielles Anwachsen**, im Falle $\lambda < 0$ einen **exponentiellen Zerfall**.

Zum Schluss noch der Hinweis auf den allgemeinen Fall einer homogenen linearen DGL mit konstanten Koeffizienten $n$-ter Ordnung. Diese kann analog mit dem gleichen Exponentialansatz $y = e^{\lambda x}$ gelöst werden. Einsetzen in die DGL liefert dann ein Polynom $n$-ter Ordnung in $\lambda$, dessen Nullstellen die erlaubten $\lambda$-Werte bestimmen. Man beachte, dass es für $n > 4$ kein systematisches Verfahren (sprich eine Verallgemeinerung der $p$-$q$-Formel) mehr gibt, um diese Nullstellen zu bestimmen!

### 6.2.2 Inhomogener Fall

Wir betrachten nun die inhomogene DGL. Sei daher $\overline{y}$ eine beliebige Lösung der inhomogenen DGL, die wir uns auf irgendeine Art und Weise verschafft haben, z.B. durch Raten oder Variation der Konstanten (s.u.). Dann ist die allgemeine Lösung der inhomogenen DGL gegeben durch

$$\boxed{y(x) = \overline{y}(x) + y_{\text{hom}}(x),}$$

wobei $y_{\text{hom}}(x)$ die allgemeine Lösung der homogenen DGL ist (wie in Abschnitt 6.2.1 diskutiert). Das $y(x)$ eine tatsächlich eine Lösung der DGL ist, macht man sich wieder leicht klar. Die Rechnung ist analog zum Beweis, dass im homogenen Fall die Summe zweier Lösungen wieder eine Lösung der homogenen Gleichung ist. Dabei ist die Linearität der DGL wichtig.

Oft findet man leicht eine spezielle Lösung der inhomogenen DGL (z.B. $\overline{y} =$ konst.). Diese erfüllt i.a. aber nicht die gewünschten Anfangsbedingungen. Deshalb benötigt man die allgemeine Lösung.

Allgemein kann man folgende *Faustregel* festhalten: Hat man eine Lösung einer DGL $n$-ter Ordnung gefunden, die $n$ *unabhängige* Integrationskonstanten enthält, so ist dies die allgemeinste Lösung. Dies gilt für den homogenen und inhomogenen Fall und auch im Falle nichtlinearer DGL.

Damit bleibt für die inhomogenen Gleichungen die Frage, wie man sich eine spezielle Lösung verschafft. Ein in vielen Fällen probates Mittel ist tatsächlich *Raten*. Z.B. kann man leicht testen, ob $\overline{y}(x)$ eine Lösung ist. Ein systematisches Verfahren, die sog. *Variation der Konstanten*, werden wir später kennenlernen.

## 6.3 Allgemeine lineare DGL 1. Ordnung

### 6.3.1 Homogener Fall

Die allgemeine lineare DGL 1. Ordnung hat im homogenen Fall folgende Form:

$$\boxed{y'(x) = a(x)y(x).}$$

Wir werden in Abschnitt 6.4.1 sehen, dass es sich hierbei um einen Spezialfall einer DGL mit getrennten Variablen handelt.

Im Prinzip kann man die Lösung fast raten! Offensichtlich wird bei der Ableitung die Funktion selbst reproduziert, mit einem zusätzlichen Faktor. Das ist genau das, was bei der Ableitung der Verkettung der Exponentialfunktion mit einer anderen Funktion passiert. Die Lösung der DGL lautet daher

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

Genauer gesagt ist dies die Lösung der DGL, die die Anfangsbedingung $y(x_0) = y_0$ erfüllt. Da die angegebene Lösung eine Integrationskonstante enthält, nämlich $y_0$, ist dies bereits die allgemeine Lösung.

Wir können die Lösung auch “herleiten”, und zwar folgendermaßen: Zunächst haben wir

$$\frac{dy}{dx} = a(x)y.$$

Nun trennen wir die Variablen folgendermaßen:

$$\frac{dy}{y} = a(x)dx.$$

Hier haben wir wieder den Differentialquotienten auseinandergerissen, um alle $y$-abhängigen Größen auf die linke, alle $x$-abhängigen Größen auf die rechte Seite zu schaffen.
Diese Gleichung wird nun integriert, d.h. wir bestimmen die Stammfunktionen auf beiden Seiten:

$$\int \frac{dy}{y} = \int a(x)dx + \tilde{c}.$$

$c$ ist dabei eine Integrationskonstante[4]. Die linke Seite können wir integrieren ($\int \frac{dy}{y} = \ln y$) und so erhalten wir nach Exponentieren

$$y = \exp\left(\int a(x)dx + \tilde{c}\right) = c\exp\left(\int a(x)dx\right)$$

mit $c = e^{\tilde{c}}$.

### 6.3.2 Inhomogener Fall

Die allgemeinste Form der inhomogenen linearen DGL 1. Ordnung lautet

$$\boxed{y'(x) = a(x)y(x) + b(x).}$$

Hier wollen wir die Lösung mit Hilfe der Lösung der homogenen Gleichung bestimmen. Dazu benutzen wir ein Verfahren, dass als **Variation der Konstanten** bezeichnet wird. Die grundlegende Idee dabei ist, die Lösung der homogenen Gleichung als Ansatz zu wählen, wobei die auftretenden Integrationskonstanten nun Funktionen sein können. Dann versucht man diese Funktionen so zu bestimmen, dass man eine Lösung der inhomogenen DGL erhält. In unserem Beispiel lautet daher der Ansatz

$$y(x) = \alpha(x)\exp\left(A(x)\right) \qquad \text{mit der Abkürzung} \quad A(x) := \int_{x_0}^{x} a(t)dt.$$

Wir haben also die Integrationskonstante $y_0$ durch die Funktion $\alpha(x)$ ersetzt, die nun geeignet zu bestimmen ist. Dazu müssen wir den Ansatz in die inhomogene DGL einsetzen. Zunächst bestimmen wir mit Produkt- und Kettenregel die Ableitung unseres Ansatzes:

$$y' = \alpha' e^A + \alpha A' e^A = \alpha' e^A + \alpha a e^A,$$

wobei wir im zweiten Schritt ausgenutzt haben, dass per Definition $A$ Stammfunktion von $a$ ist. Einsetzen liefert dann

$$\alpha' e^A + \alpha a e^A = a\alpha e^A + b$$

[4]Eigentlich tritt auf beiden Seite eine Integrationskonstante auf, die wir zu einer zusammenfassen können!

woraus folgt

$$\alpha' = be^{-A}.$$

Dies kann man leicht integrieren und erhält so die gesuchte Funktion $\alpha(x)$:

$$\alpha(x) = y_0 + \int_{x_0}^{x} b(s)e^{-A(s)}ds.$$

Damit können wir die allgemeine Lösung der inhomogenen DGL angeben

$$y(x) = y_0 e^{A(x)} + \int_{x_0}^{x} b(s)e^{A(x)-A(s)}ds \qquad \text{mit} \quad A(x) := \int_{x_0}^{x} a(t)dt.$$

Diese Lösung enthält eine freie Integrationskonstante, nämlich $y_0$.

## 6.4 Nichtlineare DGL

Nichtlineare Gleichungen sind i.a. nicht so leicht lösbar. Wir diskutieren hier aber zwei wichtige Fälle, in denen man ein allgemeines Lösungsverfahren angeben kann.

### 6.4.1 DGL mit getrennten Variablen

Wir betrachten eine DGL der Form

$$\boxed{y'(x) = g(x)h(y).}$$

Dies ist offensichtlich eine Verallgemeinerung der homogenen linearen DGL 1. Ordnung (siehe Kap. 6.3.1), die man für die Wahl $h(y) = y$ erhält. Das Vorgehen ist analog zu dem oben besprochenen Fall, d.h. wir lösen die Gleichung durch **Trennung der Variablen**: Mit $y' = \frac{dy}{dx}$ folgt

$$\frac{dy}{h(y)} = g(x)dx.$$

Somit erhält man

$$\int \frac{dy}{h(y)} = \int g(x)dx + c$$

und daraus nach Ausführung der Integrationen und Auflösen nach $y = y(x)$ die Lösung. Im Gegensatz zum linearen Fall können wir hier das Integral über $y$ nicht explizit ausrechnen, da die Funktion $h(y)$ beliebig ist. Wir wollen uns das Verfahren daher an einem konkreten Beispiel klarmachen.

**Beispiel 6.4.1.**
Wir betrachten folgende DGL mit getrennten Variablen: $y'(x) = -2xy^2(x)$.
Wir haben

$$\int \frac{dy}{y^2} = \int (-2)xdx = -x^2 + c$$

Andererseits:

$$\int \frac{dy}{y^2} = -\frac{1}{y} \qquad \Rightarrow \qquad \frac{1}{y(x)} = x^2 + c \qquad \curvearrowright \qquad y(x) = \frac{1}{x^2 + c}.$$

Mit Anfangsbedingung $y(1) = 1$ ergibt sich die Integrationskonstante zu $c = 0$ und man erhält die Lösung $y(x) = \frac{1}{x^2}$. Dass dies tatsächlich eine Lösung der DGL $y'(x) = -2xy^2(x)$ überprüft man leicht durch Differenzieren.

### 6.4.2 Bernoulli-Gleichung

Als Beispiel dafür, dass man manchmal DGL durch geeignete Transformationen auf einfachere DGL zurückführen kann, betrachten wir die sog. **Bernoulli-Gleichung**

$$\boxed{y'(x) = a(x)y(x) + b(x)y^{\alpha}(x)}$$

mit $\alpha \in \mathbb{R}$.
Zunächst multiplizieren wir die Gleichung mit $y^{-\alpha}$. Man erhält

$$y'y^{-\alpha} = a(x)y^{1-\alpha} + b(x).$$

Nun führen wir eine neue Funktion $z(x)$ ein, die durch $z := y^{1-\alpha}$ definiert ist. Dann ist $z' = (1-\alpha)y^{-\alpha}y'$ und die Gleichung lautet daher

$$z'(x) = (1-\alpha)a(x)z(x) + (1-\alpha)b(x).$$

Dies ist aber die bereits bekannte inhomogene lineare DGL 1. Ordnung, die wir systematisch gelöst haben (Kap. 6.3.2). Somit können wir $z$ explizit bestimmen und hieraus über $y = z^{1/(1-\alpha)}$ die gesuchte Funktion $y(x)$.

# Kapitel 7

# Vektoranalysis

Wir betrachten nun Funktionen, die entweder von mehreren Variablen abhängen oder die vektorwertig sind. Als unabhängige Variablen kommen dabei in der Physik vor allem die Zeit $t$ und der Ort $\underline{r} = \begin{pmatrix} x \\ y \\ z \end{pmatrix}$ in Frage.

**Definition 7.0.1** (Skalarfeld)**.**

Ein **Skalarfeld** $f$ ordnet jedem Raumpunkt $\underline{r} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in \mathbb{R}^n$ einen Skalar $f(x_1, x_2, \ldots, x_n)$ zu.

Beispiele für physikalische Skalarfelder sind das Temperaturfeld $T(\underline{r}, t)$, dass jedem Punkt $\underline{r}$ die dortige Temperatur zur Zeit $t$ zuordnet, oder Höhenprofile $h(x, y)$, die die Höhe eines Gebirges am Ort $(x, y)$ angeben.

**Definition 7.0.2** (Vektorfeld)**.**

Ein **Vektorfeld** $\underline{F}(\underline{r})$ ordnet jedem Raumpunkt $\underline{r} \in \mathbb{R}^n$ einen Vektor

$$\underline{F}(\underline{r}) = \begin{pmatrix} F_1(x_1, x_2, \ldots, x_n) \\ \vdots \\ F_m(x_1, x_2, \ldots, x_n) \end{pmatrix} \in \mathbb{R}^m$$

zu. Dabei kann $m \neq n$ sein. In der Physik ist i.a. $m = 3$ und $n = 3$ (zeitunabhängige Probleme) oder $n = 4$ (zeitabhängige Probleme).

Beispiele für physikalische Vektorfelder sind Kraftfelder $\underline{F}(\underline{r}, t)$ oder die sog. Bahnkurve $\underline{r}(t)$, die die Position eines Teilchens zur Zeit $t$ angibt.

Zusammenfassend sei noch einmal betont, dass Skalar- und Vektorfelder eigentlich nichts anderes als spezielle Funktionen sind, bei denen Definitions- und/oder Zielmenge mehrdimensionale Vektorräume sind!

## 7.1 Partielle Ableitung und totales Differential

**Definition 7.1.1** (Partielle Ableitungen)**.**
Die Ableitung von $y = f(x_1, x_2)$ nach $x_1$, die man auch als **partielle Ableitung** von $y$ nach $x_1$ bezeichnet, wird berechnet, indem man $x_2$ festhält und dann wie oben mit $x = x_1$ ableitet. Schreibweise:

$$\frac{\partial f(x_1, x_2)}{\partial x_1} \quad \text{bzw.} \quad \frac{\partial f(x_1, x_2)}{\partial x_2}$$

**Beispiel 7.1.1.**
Wir betrachten die Funktion $f(x_1, x_2) = x_1 x_2^2$ der Variablen $x_1$ und $x_2$. Als partielle Ableitungen erhält man dann

$$\frac{\partial f}{\partial x_1} = x_2^2 \quad \text{und} \quad \frac{\partial f}{\partial x_2} = 2x_1 x_2.$$

Wir können auch höhere partielle Ableitungen definieren:

$$\begin{aligned}
\frac{\partial^2 f}{\partial x_1^2} &:= \frac{\partial}{\partial x_1}\left(\frac{\partial f}{\partial x_1}\right), \\
\frac{\partial^2 f}{\partial x_2 \partial x_1} &:= \frac{\partial}{\partial x_2}\left(\frac{\partial f}{\partial x_1}\right), \\
\frac{\partial^2 f}{\partial x_1 \partial x_2} &:= \frac{\partial}{\partial x_1}\left(\frac{\partial f}{\partial x_2}\right),
\end{aligned}$$

was sich leicht auf mehrere Variablen $x_1, \ldots, x_n$ und höhere Ableitungen verallgemeinern lässt.

Eine wichtige Eigenschaft ist die Vertauschbarkeit von partiellen Ableitungen:

$$\boxed{\frac{\partial^2 f}{\partial x_2 \partial x_1} \stackrel{!}{=} \frac{\partial^2 f}{\partial x_1 \partial x_2}.}$$

Analoges gilt natürlich für die anderen Ableitungen. Es ist also egal, in welcher Reihenfolge man partiell nach zwei verschiedenen unabhängigen Variablen differenziert.

Wir wollen uns davon noch einmal an dem Beispiel $f(x_1, x_2) = x_1 x_2^2$ überzeugen. Die ersten Ableitungen hatten wir bereits bestimmt. Für die zweiten Ableitungen folgt daraus:

$$\frac{\partial^2 f}{\partial x_2 \partial x_1} = \frac{\partial}{\partial x_2}\left(\frac{\partial f}{\partial x_1}\right) = 2x_2 = \frac{\partial}{\partial x_1}\left(\frac{\partial f}{\partial x_2}\right) = \frac{\partial^2 f}{\partial x_1 \partial x_2}.$$

Außerdem gilt

$$\frac{\partial^2 f}{\partial x_1^2} = 0 \qquad \text{und} \qquad \frac{\partial^2 f}{\partial x_2^2} = 2x_1.$$

Man beachte, dass natürlich in der Regel $\frac{\partial^2 f}{\partial x_1^2} \neq \frac{\partial^2 f}{\partial x_2^2}$ ist.

Die partielle Ableitung von $f(x_1, x_2)$ misst die Stärke der Änderung der Funktion in $x_1$-Richtung. Analoges gilt für $\frac{\partial f}{\partial x_2}$. Es stellt sich die Frage, wie man den Gesamtzuwachs der Funktion charakterisieren kann. Wir gehen dabei analog zum Fall einer Variablen vor und betrachten die Änderung $f(x_1 + \Delta x_1, x_2 + \Delta x_2) - f(x_1, x_2)$. Wir nehmen dabei wieder an, dass $\Delta x_1$ und $\Delta x_2$ klein sind[1]. Daher können wir $f(x_1 + \Delta x_1, x_2 + \Delta x_2)$ durch eine Taylorentwicklung approximieren. Zunächst machen wir eine solche Entwicklung für die Funktion $g_1(x_1 + \Delta x_1) := f(x_1 + \Delta x_1, x_2 + \Delta x_2)$. Wir stellen uns also vor, dass $x_2 + \Delta x_2$ konstant gehalten wird und entwickeln um $x_1$. Bis zur ersten Ordnung erhalten wir dann:

$$f(x_1 + \Delta x_1, x_2 + \Delta x_2) = f(x_1, x_2 + \Delta x_2) + \frac{\partial f}{\partial x_1}(x_1, x_2 + \Delta x_2)\Delta x_1 + \mathcal{O}\left((\Delta x_1)^2\right).$$

Als nächstes untersuchen wir den ersten Term weiter. Wir machen nun eine Taylor-Entwicklung der Funktion $g_2(x_2 + \Delta x_2) := f(x_1, x_2 + \Delta x_2)$ um $x_2$ , d.h. nun halten wir $x_1$ fest:

$$f(x_1, x_2 + \Delta x_2) = f(x_1, x_2) + \frac{\partial f}{\partial x_2}(x_1, x_2)\Delta x_2 + \mathcal{O}\left((\Delta x_2)^2\right).$$

Als Letztes entwickeln wir $g_3(x_2 + \Delta x_2) := \frac{\partial f}{\partial x_1}(x_1, x_2 + \Delta x_2)$. Hier genügt für unsere Zwecke eine Entwicklung nullter Ordnung um $x_2$:

$$\frac{\partial f}{\partial x_1}(x_1, x_2 + \Delta x_2) = \frac{\partial f}{\partial x_1}(x_1, x_2) + \mathcal{O}\left(\Delta x_2\right).$$

Setzen wir alle diese Entwicklungen ein, so erhalten wir für den Gesamtzuwachs der Funktion

$$f(x_1 + \Delta x_1, x_2 + \Delta x_2) - f(x_1, x_2) = \frac{\partial f}{\partial x_1}(x_1, x_2)\Delta x_1 + \frac{\partial f}{\partial x_2}(x_1, x_2)\Delta x_2 + \mathcal{O}\left((\Delta x)^2\right).$$

Dabei steht $\mathcal{O}\left((\Delta x)^2\right)$ für Korrekturen, die mindestens quadratisch in den $\Delta x_j$ sind, also Terme der Form $(\Delta x_1)^2$, $(\Delta x_2)^2$ und $\Delta x_1 \Delta x_2$.
Machen wir nun $\Delta x_1$ und $\Delta x_2$ infinitesimal klein, so erhalten wir für $df := f(x_1 + dx_1, x_2 + dx_2) - f(x_1, x_2)$

$$\boxed{df = \frac{\partial f}{\partial x_1}(x_1, x_2)dx_1 + \frac{\partial f}{\partial x_2}(x_1, x_2)dx_2.}$$

Dies bezeichnet man als das **totale Differential** von $f$.
Für Funktionen von $n$ Variablen $x_1, \ldots, x_n$ gilt analog

$$\boxed{df = \sum_{j=1}^{n} \frac{\partial f}{\partial x_j}(x_1, \ldots, x_n)dx_j.}$$

[1] Später werden wir sie wieder gegen Null gehen lassen!

Wir betrachten als Beispiel unsere bekannte Funktion $f(x_1, x_2) = x_1 x_2^2$, für die wir ja bereits die partiellen Ableitungen bestimmt hatten. Damit ergibt sich

$$df = x_2^2 dx_1 + 2x_1 x_2 dx_2$$

für das totale Differential.

Es sei hier angemerkt, dass nicht jede Größe der Form $g_1(x_1, x_2)dx_1 + g_2(x_1, x_2)dx_2$ ein totales Differential ist! Hierzu müsste es nämlich eine Funktion $f(x_1, x_2)$ geben, so dass $\frac{\partial f}{\partial x_1} = g_1$ und $\frac{\partial f}{\partial x_2} = g_2$ ist.

Als Beispiel hierfür betrachten wir $x_2 dx_1 + 2x_1 x_2 dx_2$. Dies ist kein totales Differential, denn es gibt keine Funktion $f(x_1, x_2)$ mit $\frac{\partial f}{\partial x_1} = x_2$ und $\frac{\partial f}{\partial x_2} = 2x_1 x_2$. Um dies einzusehen, nehmen wir einmal an, es gäbe eine solche Funktion. Auf Grund der Vertauschbarkeit der partiellen Ableitungen müsste dann auch $\frac{\partial^2 f}{\partial x_2 \partial x_1} = \frac{\partial^2 f}{\partial x_1 \partial x_2}$ sein. Dies ist aber offensichtlich nicht erfüllt, da $\frac{\partial^2 f}{\partial x_2 \partial x_1} = 1$ und $\frac{\partial^2 f}{\partial x_1 \partial x_2} = 2x_2$ ist.
Solche Größen heißen **inexakte Differentiale** und spielen z.B. in der Thermodynamik eine wichtige Rolle.

Zum Abschluss wollen wir noch zwei Anwendungen des totalen Differentials diskutieren:

1. Die partiellen Ableitungen geben uns an, wie stark sich die Funktion $f(x_1, x_2)$ in die $x_1$- bzw. $x_2$-Richtung ändert. Dies entspricht der Wahl $dx_2 = 0$ bzw. $dx_1 = 0$ im totalen Differential. Andere Wahlen von $dx_1$ und $dx_2$ liefern die Veränderung von $f$ in andere Richtungen. Dies werden wir im nächsten Abschnitt mit Hilfe der sog. Richtungsableitung präzisieren.

2. Das totale Differential erlaubt es, die Kettenregel auf Skalarfelder zu verallgemeinern. Wir betrachten dazu eine Funktion $g(t) = f(x(t), y(t), z(t))$, d.h. die Argumente $x, y, z$ hängen selbst wieder von einer unabhängigen Variablen $t$ ab. Hierbei kann $t$ z.B. die Zeit sein, $x(t), y(t), z(t)$ die Koordinaten eines Massenpunktes zur Zeit $t$ und $f$ der Betrag der auf das Teilchen wirkenden Kraft. Diese ist i.a. als Funktion des Ortes $\underline{r}$ gegeben, wird aber bei Einsetzen der Bahnkurve $\underline{r}(t)$ zu einer Funktion $f(t)$ der Zeit.

   Die Ableitung von $g$ nach dem Parameter $t$ bestimmt sich dann aus

$$\frac{dg}{dt} = \frac{\partial f}{\partial x}\frac{dx}{dt} + \frac{\partial f}{\partial y}\frac{dy}{dt} + \frac{\partial f}{\partial z}\frac{dz}{dt}.$$

   Wie erwartet gehen dabei natürlich auch die $t-$Abhängigkeiten der Funktionen $x(t)$,$y(t)$, $z(t)$ ein. Die Änderung von $g$ hängt sowohl davon ab, wie stark sich $g$ mit $x, y, z$ ändert, als auch davon, wie stark diese Größen sich mit $t$ ändern.

Wir werden nun eine weitere Größe einführen, die uns wichtige Informationen über die lokalen Änderungen eines Skalarfeldes liefert.

## 7.2 Gradient

**Definition 7.2.1** (Gradient)**.**
Zu einem skalaren Feld $f(\underline{r}) = f(x, y, z)$ definieren wir

$$\boxed{\operatorname{grad} f := \begin{pmatrix} \frac{\partial f}{\partial x} \\ \frac{\partial f}{\partial y} \\ \frac{\partial f}{\partial z} \end{pmatrix},}$$

d.h. die $j$-te Komponente des Gradienten von $f$ ist gerade die partielle Ableitung nach der $j$-ten unabhängigen Variablen. Dieses Vektorfeld bezeichnet man als **Gradienten von** $f$.
Eine analoge Definition gilt für beliebige Skalarfelder $f(x_1, \ldots, x_n)$. Hier hat der Gradient $n$ Komponenten.
Die angegebene Definition gilt in kartesischen Koordinaten. Die Form des Gradienten in anderen Koordinatensystemen werden wir später kennenlernen.

Mit Hilfe des Gradienten läßt sich das totale Differential von $f$ schreiben als

$$df = (\operatorname{grad} f) \cdot d\underline{r} \qquad \text{mit} \quad d\underline{r} = \begin{pmatrix} dx \\ dy \\ dz \end{pmatrix}$$

**Definition 7.2.2** (Nabla-Operator)**.**
Der sog. **Nabla-Operator** $\underline{\nabla}$ ist definiert als

$$\boxed{\underline{\nabla} := \begin{pmatrix} \frac{\partial}{\partial x} \\ \frac{\partial}{\partial y} \\ \frac{\partial}{\partial z} \end{pmatrix}}$$

Allgemein bezeichnet ein **Operator** eine Abbildung, die eine Funktionenmenge als Definitions- und Wertebereich hat. Ein Operator ist also eine Funktion, die auf Funktionen wirkt und als Ergebnis wieder eine Funktion liefert. Der Nabla-Operator ist ein **Vektoroperator**, da er als Ergebnis einen Vektor liefert, dessen Komponenten Funktionen sind.
Damit kann man den Gradienten von $f$ auch schreiben als

$$\operatorname{grad} f = \underline{\nabla} f.$$

Wir werden im folgenden beide Schreibweisen verwenden.

Der Gradient $\operatorname{grad} f$ eines skalaren Feldes $f$ hat auch eine anschauliche Bedeutung.

1. Wir betrachten die Menge aller Punkte $\underline{r}$, für die gilt$f(\underline{r}) = c$, wobei $c$ eine vorgegebene Konstante ist. Diese Menge bezeichnet man als **Niveaufläche**[2]. Auf dieser Fläche ist per

[2]In der Physik treten sie z.B. als *Äquipotentialflächen* auf.

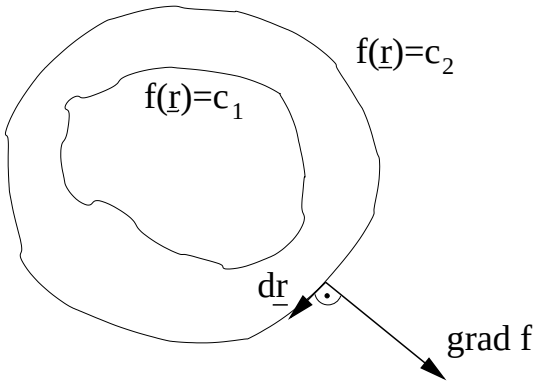



Abbildung 7.2.1: $\operatorname{grad} f$ steht senkrecht auf der Niveaufläche $f(\underline{r}) = \text{const.}$

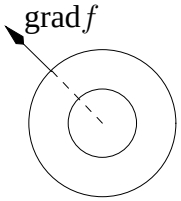
Definition $df = 0$ und somit $(\operatorname{grad} f) \cdot d\underline{r} = 0$. Dabei verbindet $d\underline{r}$ zwei infinitesimal benachbarte Punkte $\underline{r}$ und $\underline{r} + d\underline{r}$ der Niveaufläche, bezeichnet also die Tangente an diese Fläche. Somit können wir schließen:
$\operatorname{grad} f$ steht senkrecht auf den Niveauflächen der Funktion $f$ (siehe Abb. 7.2.1).

2. Der Vektor $\operatorname{grad} f(\underline{r}) = (\operatorname{grad} f)(\underline{r})$ zeigt in Richtung des stärksten Anstieges von $f$ am Punkt $\underline{r}$.
   Dazu betrachten wir die **Richtungsableitung** in Richtung $\underline{e}$ mit $|\underline{e}| = 1$. Diese ist definiert durch $\underline{e} \cdot \underline{\nabla} f$. Offensichtlich wird das Skalarprodukt $\underline{e} \cdot \underline{\nabla} f = |\underline{\nabla} f| \cos\alpha$ (wobei $\alpha$ der Winkel zwischen $\underline{e}$ und $\underline{\nabla} f$ ist) genau dann maximal, wenn $\underline{\nabla} f$ parallel zu $\underline{e}$ ist. Dies heißt aber nichts anderes, als dass die Richtungsableitung in Richtung von $\underline{\nabla} f$ maximal wird, d.h. $\underline{\nabla} f$ zeigt in Richtung des stärksten Anstiegs.

3. Der Betrag $|\operatorname{grad} f|$ misst die Stärke der Änderung von $f$ senkrecht zu den Flächen $f(\underline{r}) = \text{const.}$

**Beispiel 7.2.1.**
Die obigen Eigenschaften kann man sich an folgendem Beispiel unmittelbar klar machen:

$$\begin{aligned} f(x,y) &= x^2 + y^2 \\ \operatorname{grad} f &= 2\begin{pmatrix} x \\ y \end{pmatrix} = 2\underline{r} \end{aligned}$$




Die Kurven $f(\underline{r}) = c = \text{const.}$ sind hier Kreise vom Radius $R = \sqrt{c}$. Das Gradientenfeld $\operatorname{grad} f = 2\underline{r}$ zeigt radial nach außen und steht senkrecht auf diesen Kurven.

Für den Gradienten gelten eine Reihe von Rechenregeln. Die meisten folgen direkt aus denen für die partielle Ableitung, z.B. die Linearität $\operatorname{grad}(f + \alpha g) = \operatorname{grad} f + \alpha \operatorname{grad} g$ mit $\alpha \in \mathbb{R}$.

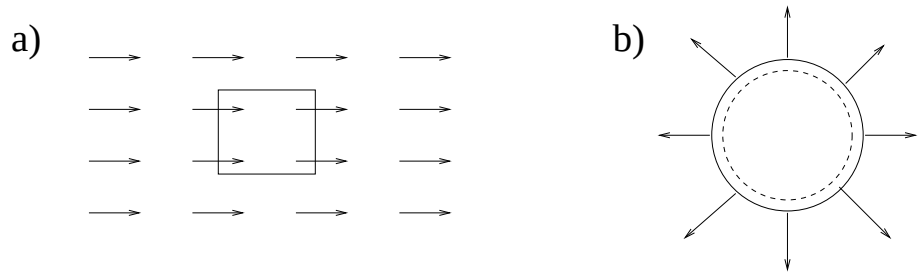



Abbildung 7.3.1: Zur Interpretation der Divergenz eines Vektorfeldes.

## 7.3 Divergenz

**Definition 7.3.1** (Divergenz)**.**
Die **Divergenz** eines Vektorfeldes $\underline{F}(\underline{r})$ ist (in kartesischen Koordinaten) definiert als

$$\boxed{\operatorname{div} \underline{F} := \underline{\nabla} \cdot \underline{F} = \frac{\partial F_x}{\partial x}(\underline{r}) + \frac{\partial F_y}{\partial y}(\underline{r}) + \frac{\partial F_z}{\partial z}(\underline{r}).}$$

Eine analoge Definition gilt für die Divergenz eines Vektorfeldes in beliebigen Dimensionen. Man erhält also $\operatorname{div} \underline{F}$, indem man formal das Skalarprodukt des Nabla-Vektors mit dem Vektorfeld bildet! Die Divergenz eines Vektorfeldes ist also ein Skalarfeld.

Die Divergenz gibt Auskunft über die Quellen und Senken eines Vektorfeldes. Ist $\operatorname{div} \underline{F} = 0$, so nennt man $\underline{F}$ auch **quellenfrei**.
Zur Interpretation stellt man sich $\underline{F}$ z.B. als Strömungsfeld einer Flüssigkeit vor, d.h. $\underline{F} = \rho \underline{v}$, wobei $\rho$ die Dichte und $\underline{v}$ die Geschwindigkeit der Flüssigkeit am Ort $\underline{r}$ ist. Wir betrachten dann ein kleines Volumen $\Delta V$ um den Punkt $\underline{r}$. $\operatorname{div} \underline{F}$ gibt dann an, wieviel Masse *effektiv* in das Volumen hineinfließt, also die Differenz von Ein- und Ausfluss.

**Beispiel 7.3.1.** Wir betrachten zwei Beispiele zur Interpretation der Divergenz:

1. Zunächst untersuchen wir ein Feld $\underline{F}(\underline{r})$, das in Betrag und Richtung konstant ist (siehe Abb. 7.3.1a). Anschaulich sollte klar sein, dass der Einfluss und der Ausfluss aus dem eingezeichneten Rechteck um den Punkt $\underline{r}$ in diesem Fall gleich sind. Im gezeigten Fall gibt es Einfluss nur an der linken Seite, Ausfluss nur auf der rechten. Man kann zeigen, dass auch für beliebig geformte Flächen Ein- und Ausfluss immer gleich sind. Natürlich gilt das für beliebige Punkte $\underline{r}$. Somit verschwindet die Divergenz. Dies ist konsistent mit der expliziten Berechnung $\operatorname{div} \underline{F} = \frac{\partial F_x}{\partial x} + \frac{\partial F_y}{\partial y} = 0$, da beide partiellen Ableitungen wegen der Konstanz des Feldes verschwinden.

2. Als zweites Beispiel betrachten wir ein radiales Feld, z.B. $\underline{F}(\underline{r}) = \underline{r}$ (siehe Abb. 7.3.1b). Als Fläche betrachten wir einen dünnen Kreisring. Offensichtlich ist der Einfluss kleiner als der Ausfluss, da das Feld im inneren Kreis (gestrichelt) einen kleineren Betrag hat als auf dem äußeren. Somit ist $\operatorname{div} \underline{F} > 0$. Dies ist konsistent mit der expliziten Berechung $\operatorname{div} \underline{F} = \frac{\partial F_x}{\partial x} + \frac{\partial F_y}{\partial y} = 2$, da $F_x = x$ und $F_y = y$.

Später (in Kap. 8.5) werden wir lernen, wie man den Fluss durch eine Fläche explizit berechnet.

Für die Divergenz gelten einige einfache Rechenregeln, die sich aus denen für die partiellen Ableitungen ergeben, insbesondere die Linearität $\operatorname{div}(\underline{F} + \alpha\underline{G}) = \operatorname{div}\underline{F} + \alpha\ \operatorname{div}\underline{G}$ mit $\alpha \in \mathbb{R}$.

Aus Divergenz und Gradient können wir einen weiteren wichtigen Operator bilden, der häufig in der Physik eine Rolle spielt.

**Definition 7.3.2** (Laplace-Operator)**.**
Der **Laplace-Operator** $\triangle$ ordnet einem Skalarfeld $f(\underline{r})$ ein anderes Skalarfeld zu gemäß

$$\boxed{\Delta f := \operatorname{div}(\operatorname{grad} f) = \underline{\nabla} \cdot \underline{\nabla} f = \underline{\nabla}^2 f.}$$

bzw. explizit (was man durch Nachrechnen leicht prüft)

$$\boxed{\Delta f = \frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2} + \frac{\partial^2 f}{\partial z^2}.}$$

Einen analogen Ausdruck erhält man für Skalarfelder $f(x_1, \ldots, x_n)$.

Manchmal ist es beim Laplace-Operator praktisch, wenn man seine Wirkung auf ein Vektorfeld $\underline{F}(\underline{r})$ gemäß

$$\Delta\underline{F} := \begin{pmatrix} \Delta F_x \\ \Delta F_y \\ \Delta F_z \end{pmatrix}$$

definiert. In diesem Fall ist die Wirkung also komponentenweise zu verstehen!

Wie schon erwähnt, tritt der Laplace-Operator in zahlreichen Anwendungen in der Physik auf. Hier seien nur die *Wellengleichung* und die *Diffusionsgleichung* erwähnt.

## 7.4 Rotation

**Definition 7.4.1** (Rotation)**.**
In 3 Dimensionen (und nur dort!) können wir die **Rotation** eines Vektorfeldes $\underline{F}(\underline{r})$ (in kartesischen Koordinaten) durch

$$\boxed{\operatorname{rot}\underline{F} = \underline{\nabla} \times \underline{F} = \begin{pmatrix} \frac{\partial F_z}{\partial y} - \frac{\partial F_y}{\partial z} \\ \frac{\partial F_x}{\partial z} - \frac{\partial F_z}{\partial x} \\ \frac{\partial F_y}{\partial x} - \frac{\partial F_x}{\partial y} \end{pmatrix}}$$

definieren. Mit Hilfe des sog. Levi-Cevità-Symbols können wir dies auch etwas kompakter als

$$(\operatorname{rot}\underline{F})_i = \sum_{j,k=1}^{3} \epsilon_{ijk} \frac{\partial}{\partial x_j} F_k$$

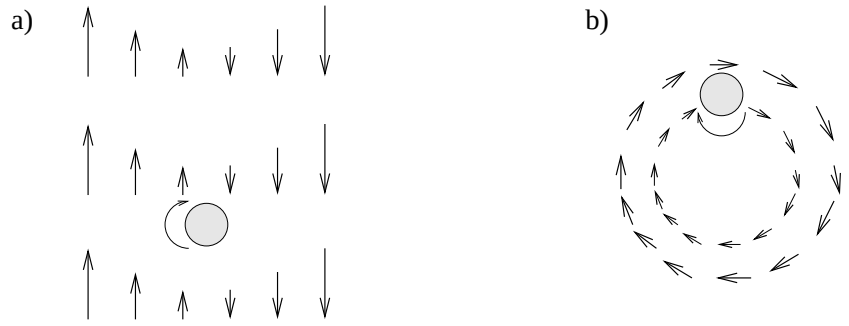



Abbildung 7.4.1: Zur Interpretation der Rotation eines Vektorfeldes: In beiden Feldern wird ein kleines Testobjekt vom Strömungsfeld in Drehung versetzt.

schreiben. Dabei ist $\epsilon_{123} = \epsilon_{231} = \epsilon_{312} = 1$, $\epsilon_{132} = \epsilon_{213} = \epsilon_{321} = -1$ und alle anderen $\epsilon_{ijk} = 0$. Formal ergibt sie sich also aus dem Kreuzprodukt des Nabla-Operators mit dem Vektorfeld. Die Rotation eines Vektorfeldes ist also wieder ein Vektorfeld.

Die Rotation eines Vektorfeldes gibt Auskunft über die Wirbel des Feldes. $\mathrm{rot}\,\underline{F}$ bezeichnet man auch als **Wirbelfeld**[3].

Zur Interpretation stelle man sich $\underline{F}$ als das Geschwindigkeitsfeld einer strömenden Flüssigkeit vor. Wird ein kleines Objekt bei $\underline{r}$ von diesem Feld gedreht, so ist $\mathrm{rot}\,\underline{F}(\underline{r}) \neq 0$. Dies passiert z.B. bei den beiden Feldern, die in Abb. 7.4.1 gezeigt sind.
Später werden wir noch sehen, wie man die Wirbelstärke explizit berechnen kann!

Es gibt eine Reihe von wichtigen Identitäten, die die Wirkung unterschiedlicher Operatoren miteinander verknüpfen. Wir werden diese hier nicht alle aufführen, da Sie diese in den meisten Lehrbüchern und Formelsammlungen finden. Einige werden wir in den Übungen diskutieren.
Zwei Identitäten wollen wir aber explizit nennen, die später eine wichtige Rolle spielen werden:

$$\boxed{\mathrm{rot}(\mathrm{grad}\, f) = 0, \qquad \text{und} \qquad \mathrm{div}(\mathrm{rot}\,\underline{F}) = 0}$$

Diese Identitäten gelten allgemein für beliebige Skalarfelder $f(\underline{r})$ bzw. Vektorfelder $\underline{F}(\underline{r})$.

Die folgende Tabelle fasst die verschiedenen Ableitungen, die wir in diesem Kapitel kennengelernt haben, noch einmal zusammen:

| Operator | wirkt auf | liefert |
|---|---|---|
| $\frac{\partial}{\partial x}$ | Skalarfeld | Skalarfeld |
| grad | Skalarfeld | Vektorfeld |
| div | Vektorfeld | Skalarfeld |
| $\Delta$ | Skalarfeld | Skalarfeld |
| rot | Vektorfeld (in 3d) | Vektorfeld (in 3d) |

[3]Manchmal bezeichnet man auch Felder mit $\mathrm{rot}\,\underline{F} \neq 0$ als Wirbelfelder.

Hier sei noch einmal darauf hingewiesen, dass man die Anwendung des Laplace-Operators auf Vektorfelder komponentenweise definieren kann!

# Kapitel 8

# Wegintegral und mehrdimensionale Integration

Integrale über Vektorfelder $\underline{F}(t)$, die nur von einer Variablen $t$ abhängen, sind komponentenweise definiert:

$$\int \underline{F}(t)dt := \begin{pmatrix} \int \underline{F}_x(t)dt \\ \int \underline{F}_y(t)dt \\ \int \underline{F}_z(t)dt \end{pmatrix}.$$

Im folgenden wollen wir einige andere Fälle von mehrdimensionalen Integrale betrachten.

## 8.1 Wegintegrale

Zunächst wollen wir ein Vektorfeld $\underline{F}(\underline{r})$ längs eines Weges integrieren.

**Definition 8.1.1** (Wegintegral)**.**
Das Integral eines Vektorfeldes $\underline{F}(\underline{r})$ längs eines Weges $c$ vom Punkt $P$ zum $Q$, das man auch als **Wegintegral** (oder als **Kurvenintegral** bzw. **Linienintegral**) bezeichnet, ist definiert durch

$$\int_P^Q \underline{F}(\underline{r}) \cdot d\underline{r} \quad = \quad \lim_{\Delta r_i \to 0} \sum_i \underline{F}(\underline{r}_i) \cdot \Delta \underline{r}_i = \lim_{\Delta r_i \to 0} \sum_i F(\underline{r}_i) \Delta r_i \cos\phi_i.$$

Im Gegensatz zum früher definierten Integral über eine vektorwertige Funktion werden hier also nur die Projektionen $\underline{F}(\underline{r}_i) \cdot \Delta \underline{r}_i$ des Vektorfeldes $\underline{F}(\underline{r})$ auf die Richtung $\Delta \underline{r}_i$ des Weges aufsummiert (siehe Abb. 8.1.1). Man muss dabei beachten, dass der Wert des Integrales i.a. vom genauen Verlauf des Weges $c$ zwischen $P$ und $Q$ abhängt, ganz im Gegensatz zum eindimensionalen Integral, wo der Hauptsatz der Differential- und Integralrechnung die Wegunabhängigkeit garantiert. Der Weg muss daher beim Wegintegral immer spezifiziert werden.
Bem.: Statt $\int_P^Q \underline{F}(\underline{r}) \cdot d\underline{r}$ müsste man präziser $\int_c \underline{F}(\underline{r}) \cdot d\underline{r}$ schreiben, da der Wert des Wegintegrals i.a. nicht nur von den Anfangs- und Endpunkten abhängt, sondern vom genauen Verlauf des Weges $c$ von $P$ nach $Q$. Dies ist anders als beim gewöhnlichen Integral, wo der Hauptsatz der

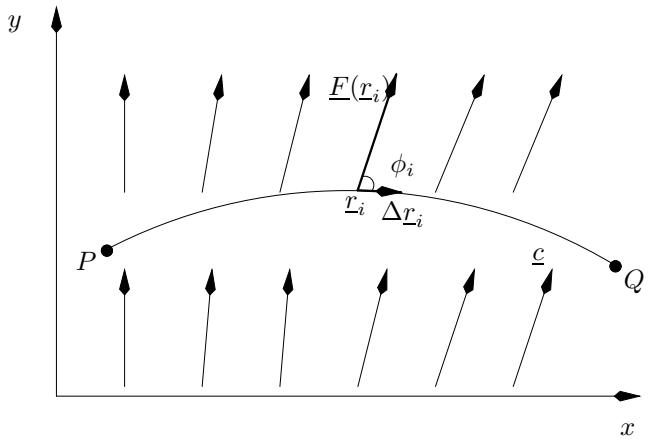



Abbildung 8.1.1: Wegintegral

Differential- und Integralrechung besagt, dass der Wert des Integrals durch die Differenz der Funktionswerte der Stammfunktion an den Grenzen des Integrationsintervalls gegeben ist.

Die praktische Berechnung von Wegintegralen $\int_P^Q \underline{F} \cdot d\underline{r}$ geschieht i.a. in zwei Schritten so:
1) Bestimme eine Parametrisierung $\underline{c}(t)$ des Weges von $P$ nach $Q$:

$$\begin{array}{lcl} \underline{c}(t) & = & (c_x(t), c_y(t)) \\ \text{mit} & & \underline{c}(t_0) = \underline{r}_P,\ \underline{c}(t_1) = \underline{r}_Q \\ \text{und} & & \underline{c}(t) \text{ liegt auf dem Weg für alle } t \in [t_0, t_1]. \end{array}$$

Man kann sich den Parameter $t$ als Zeit vorstellen. Dann ist $\underline{c}(t)$ der Ort auf dem Weg, an dem man sich zur Zeit $t$ befindet. 2) Diese Parametrisierung wird dann in den Integranden eingesetzt und das Wegintegral somit zu einem Integral über den Parameter $t$:

$$\boxed{\int_P^Q \underline{F}(\underline{r}) \cdot d\underline{r} = \int_{t_0}^{t_1} \underline{F}(\underline{c}(t)) \cdot \dot{\underline{c}}(t)\, dt}$$

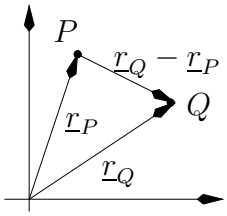
Dabei ist $\underline{F}(\underline{c}(t))$ die Feldstärke auf dem Weg und $\dot{\underline{c}}(t) = \frac{d}{dt}\underline{c}(t)$ die "Geschwindigkeit", mit der der Weg durchlaufen wird. Man beachte, dass das Skalarprodukt zu bilden ist, denn das Wegintegral liefert als Ergebnis einen Skalar!
Im zweidimensionalen Fall lautet diese Gleichung expliziter

$$\int_P^Q \underline{F}(\underline{r}) \cdot d\underline{r} \;=\; \int_{t_0}^{t_1} \left[ F_x(c_x(t), c_y(t)) \frac{dc_x}{dt}(t) + F_y(c_x(t), c_y(t)) \frac{dc_y}{dt}(t) \right]$$

Das Vorgehen in $n$ Dimensionen ist vollkommen analog zu diesem zweidimensionalen Beispiel! Damit haben wir die Berechung von Wegintegrale auf die Berechnung eines eindimensionalen Integrals zurückgeführt. Dies kann man als Verallgemeinerung der Substitutionsregel ansehen, bei der ja auch die Ableitung der substituierten Funktion auftritt.

**Beispiel 8.1.1.** Im folgenden geben wir die wichtigsten Beispiele für Parametrisierungen an:

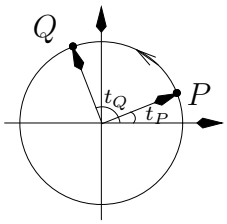
1. Gerade von $P$ nach $Q$:




$$\underline{c}(t) = \underline{r}_P + (\underline{r}_Q - \underline{r}_P)\, t$$
$$(t \in [0,1])$$

Offensichtlich beschreibt dies eine Gerade mit $\underline{c}(0) = \underline{r}_P$ und $\underline{c}(1) = \underline{r}_Q$.

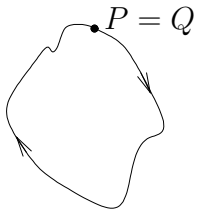
2. Kreisbogen:




$$\underline{c}(t) = R(\cos t, \sin t)$$
$$\text{mit } \; R = |\underline{r}_P| = |\underline{r}_Q| \; \text{ und } \; t \in [t_P, t_Q]$$

Dabei sind $t_P$ und $t_Q$ die Winkel, die die Ortsvektoren $\underline{r}_P$ und $\underline{r}_Q$ von $P$ und $Q$ mit der $x-$Achse bilden, und $R$ der Radius des Kreisbogens. Bei solchen Wegen müssen natürlich $P$ und $Q$ den gleichen Abstand $R$ vom Ursprung haben.

**Definition 8.1.2** (Geschlossener Weg)**.**
Bei einem Wegintegral können Anfangs- und Endpunkt auch identisch sein ($P = Q$). Für solche Wegintegrale führt man ein spezielles Symbol ein:




$$\oint \underline{F} \cdot d\underline{r}$$

I.a. wird $\oint \underline{F} \cdot d\underline{r} \neq 0$ sein. Für den Weg, der nur aus dem Punkt $P = Q$ besteht, ist natürlich immer $\oint \underline{F} \cdot d\underline{r} = 0$.

**Beispiel 8.1.2.**

Wir wollen das Wegintegral des Vektorfeldes $\underline{F}(\underline{r}) = \begin{pmatrix} z \\ xy \\ x \end{pmatrix}$ längs eines geraden Weges von $\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$ nach $\begin{pmatrix} 3 \\ 1 \\ 0 \end{pmatrix}$ berechnen. Zunächst bestimmen wir eine Parametrisierung des Weges:

$$\underline{c}(t) = t \begin{pmatrix} 3 \\ 1 \\ 0 \end{pmatrix}, \qquad \dot{\underline{c}}(t) = \begin{pmatrix} 3 \\ 1 \\ 0 \end{pmatrix}.$$

Durch Einsetzen dieser Parametrisierung folgt dann:

$$\begin{aligned} \int_P^Q \underline{F}(\underline{r}) \cdot d\underline{r} &= \int_0^1 \underline{F}(\underline{c}(t)) \cdot \dot{\underline{c}}\, dt = \int_0^1 \underline{F}(3t, t, 0) \cdot \dot{\underline{c}}\, dt \\ &= \int_0^1 \begin{pmatrix} 0 \\ 3t \cdot t \\ 3t \end{pmatrix} \cdot \begin{pmatrix} 3 \\ 1 \\ 0 \end{pmatrix} dt = \int_0^1 (0 + 3t^2 + 0)\, dt = t^3|_0^1 = 1\,. \end{aligned}$$

Hier sieht man noch einmal explizit, dass das Wegintegral als Wert eine reelle Zahl liefert!

Wegintegrale sind additiv, d.h. wenn man erst längs des Weges $\underline{c}_1(t)$ von $A$ nach $B$ geht und dann längs $\underline{c}_2(t)$ von $B$ nach $C$ , dann gilt für das Integral längs des kombinierten Weges $\underline{c}(t)$ von $A$ nach $C$, den man suggestiv auch als $\underline{c}(t) = \underline{c}_1(t) + \underline{c}_2(t)$ schreibt,

$$\int_{\underline{c}_1 + \underline{c}_2} \underline{F}(\underline{r}) \cdot d\underline{r} = \int_{\underline{c}_1} \underline{F}(\underline{r}) \cdot d\underline{r} + \int_{\underline{c}_2} \underline{F}(\underline{r}) \cdot d\underline{r}.$$

Durchläuft man den Weg $\underline{c}(t)$ in umgekehrter Richtung, so bezeichnet man diesen umgekehrten Weg auch als $-\underline{c}(t)$. In diesem Fall gilt

$$\int_{-\underline{c}} \underline{F}(\underline{r}) \cdot d\underline{r} = -\int_{\underline{c}} \underline{F}(\underline{r}) \cdot d\underline{r}$$

in Analogie zur Integrationsregel im eindimensionalen Fall[1].

## 8.2 Wegunabhängigkeit und Potential

Wie bereits erwähnt, hängt der Wert des Wegintegrals $\int_P^Q \underline{F} \cdot d\underline{r}$ i.a. vom gewählten Weg zwischen $P$ und $Q$ ab. Die Frage, wann das Integral unabhängig vom expliziten Weg ist, ist eng verknüpft mit der Frage nach der Existenz eines sog. Potentials.

[1] Vorzeichenänderung bei Vertauschung der Integrationsgrenzen!

**Definition 8.2.1** (Potential)**.**
Existiert zu einem Vektorfeld $\underline{F}(\underline{r})$ ein Skalarfeld $f$ mit $\underline{F} = -\text{grad} f$, dann nennt man $f$ ein **Potential** von $\underline{F}$.
Das Minuszeichen in der Definition ist eine Konvention, die sich in der Physik als sinnvoll herausstellt.

Ein Potential $f$ ist nicht eindeutig bestimmt, da auch $f + c$ mit einer beliebigen Konstanten $c$ ein Potential ist. Häufig wählt man $c$ geeignet, um z.B. gewisse Nebenbedingungen zu erfüllen.

Eine wichtige Frage lautet nun: Wann existiert zu vorgegebenem Vektorfeld ein Skalarfeld $f$ mit $\underline{F} = -\text{grad} f$ ? Wir geben nun (ohne vollständigen Beweis) wichtige Kriterien für die Wegunabhängigkeit bzw. die Existenz von Potentialen an:

**Satz 8.2.1.** Folgende vier Aussagen sind äquivalent:

1) $\int_P^Q \underline{F} \cdot d\underline{r}$ ist für beliebige, feste $P$ und $Q$ wegunabhängig

2) $\oint_{\underline{c}} \underline{F} \cdot d\underline{r} = 0$ für alle geschlossenen Wege.

3) Es existiert ein Potential $f$ von $\underline{F}$, d.h. $\underline{F} = -\text{grad} f$

4) $\text{rot}\, \underline{F} = 0$ (falls der Definitionsbereich von $\underline{F}$ einfach-zusammenhängend ist)

Aussage 1) ist dabei nicht so zu verstehen, dass alle Wegintegrale den gleichen Wert haben. Wählt man zwei andere Punkte $\tilde{P}$ und $\tilde{Q}$, so kann das Wegintegral zwischen diesen Punkten einen anderen Wert haben als das zwischen $P$ und $Q$!
Aus Aussage 4) kann man nur dann auf die anderen Aussagen schließen, wenn eine zusätzliche Bedingung an den Definitionsbereich des Vektorfeldes erfüllt ist. Diese werden wir weiter unter diskutieren.

Die Äquivalenz von 1) und 2) sieht man leicht an Hand der oben erwähnten Additivitätseigenschaften bzgl. des Integrationsweges ein. Der Beweis zu "3) $\Rightarrow$ 4)" ist leicht, denn $\text{rot}(\text{grad} f) = 0$ gilt (siehe Übungen) für beliebige $f$. Die anderen Beweise wollen wir hier nicht angeben.
Existiert ein Potential von $\underline{F}$, so kann man es aus

$$\boxed{f(\underline{r}) = f(\underline{r}_0) - \int_{\underline{r}_0}^{\underline{r}} \underline{F} \cdot d\underline{r}}$$

bestimmen, wobei der Integrationsweg zwischen den fest gewählten (beliebigen) Punkten $\underline{r}_0$ und $\underline{r}$ beliebig ist.

**Definition 8.2.2** (einfach-zusammenhängend)**.**
Eine Menge $M \subset \mathbb{R}^n$ heißt **einfach-zusammenhängend**, wenn sich jeder geschlossene Weg durch kontinuierliche Deformation ganz in $M$ auf einen Punkt zusammenziehen lässt. Außerdem muss $M$ wegzusammenhängend sein, d.h. zwei beliebige Punkte müssen sich durch einen Weg ganz in $M$ miteinander verbinden lassen.

Diese Definition wollen wir durch zwei Beispiele illustrieren:

**Beispiel 8.2.1.**

1. Sei $M = \mathbb{R}^2 \setminus \{0\}$, also die $x$-$y$-Ebene ohne den Ursprung. Betrachten wir einen geschlossenen Weg, der den Ursprung umschließt, z.B. einen Kreis vom Radius $R$. Dieser lässt sich nicht auf einen Punkt zusammenziehen, da man beim Zusammenziehen immer am Ursprung $0$ "hängenbleibt", der ja nicht zu $M$ gehört. $M$ ist also nicht einfachzusammenhängend.

2. Sei $M = \mathbb{R}^3 \setminus \{0\}$, also der dreidimensionale Raum ohne den Ursprung. Hier lassen sich die im vorigen Beispiel betrachteten Wege auf einen Punkt zusammenziehen, da man sie zunächst aus der $x$-$y$-Ebene anheben kann. Man kann zeigen, dass $M = \mathbb{R}^3 \setminus \{0\}$ einfachzusammenhängend ist. In drei Dimensionen muss man typischerweise eindimensionale Gebiete (z.B. Geraden) entfernen, damit der einfache Zusammenhang verloren geht.

In den Übungen wird ein Beispiel diskutiert, das zeigt, dass die Bedingung über den einfachen Zusammenhang des Definitionsbereichs tatsächlich erfüllt sein muss, um aus rot $\underline{F} = 0$ auf die Existenz eines Potentials schließen zu können. Das Feld $\underline{F}(x, y, z) = (-\frac{y}{x^2+y^2}, \frac{x}{x^2+y^2}, 0)$ hat eine verschwindende Rotation, aber das Wegintegral längs eines Kreises um den Ursprung verschwindet nicht!

## 8.3 Mehrdimensionale Integrale

**Definition 8.3.1** (Volumenintegral)**.**
Mehrdimensionale **Volumenintegrale** sind durch den Grenzübergang zu infinitesimalen Volumenelementen definiert:

$$\int\limits_{\text{Volumen } V} f(\underline{r})dV = \lim_{\Delta V_k \to 0} \sum_k f(\underline{r}_k) \cdot \Delta V_k.$$

Dies ist in vollkommener Analogie zum gewöhnlichen Integral zu verstehen. Man unterteilt das Volumen (statt des Intervalls) in kleine Teilvolumina $\Delta V_k$ und approximiert die Funktion in diesem Teilvolumen z.B. durch den Wert $f(\underline{r}_k)$ "in der Mitte". Dann macht man die Teilvolumina infinitesimal. Natürlich müßte man zeigen, dass das Ergebnis unabhängig von der Art der Aufteilung ist, aber dies ist Aufgabe der Mathematik.

Solche Integrale treten z.B. bei der Berechnung von Volumina (dann ist $f(\underline{r}) = 1$) oder Massen (dann ist $f(\underline{r}) = \rho(\underline{r})$ die Massendichte des betrachteten Körpers) auf.

Analog dazu definiert man **Flächenintegrale**

$$\int\limits_{\text{Fläche } F} f(\underline{r})dF.$$

Analog sind Integrale in beliebigen Dimensionen definiert. Statt $dV$ bzw. $dF$ schreibt man auch $d^3r$ bzw. $d^2r$ oder allgemein in $d$ Dimensionen $d^dr$.

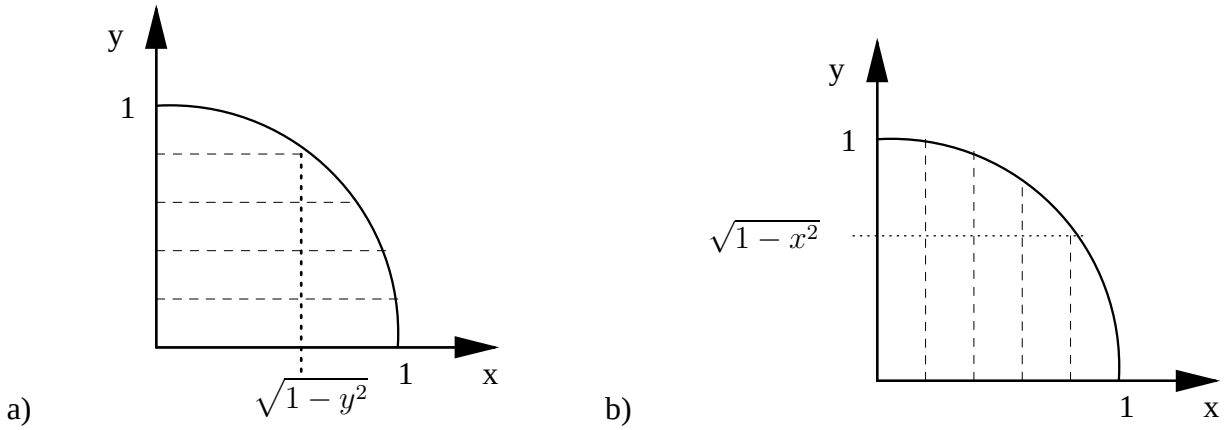



Abbildung 8.3.1: Veranschaulichung der Integration über einen Viertelkreis für verschiedene Integrationsreihenfolgen. a) Erst $x-$Integration, dann $y-$Integration; b) erst $y-$Integration, dann $x-$Integration

In kartesischen Koordinaten schreibt man Volumenintegrale ausführlicher als

$$\iiint\limits_{\text{Volumen}} f(x,y,z)dxdydz = \int\limits_{z_0}^{z_1} \left[ \int\limits_{y_0(z)}^{y_1(z)} \left\{ \int\limits_{x_0(y,z)}^{x_1(y,z)} f(x,y,z)dx \right\} dy \right] dz$$

Man beachte, dass die Integrationsgrenzen noch von den Koordinaten abhängen können, über die noch nicht integriert wurde. Nach der ersten Integration über $x$ erhält man eine Funktion $g_1(y,z) = \int\limits_{x_0(y,z)}^{x_1(y,z)} f(x,y,z)dx$, die nur noch von $y$ und $z$ abhängt. Diese wird dann zunächst über $y$ integriert, so dass man eine nur noch $z$ abhängige Funktion $g_2(z) = \int_{y_0(z)}^{y_1(z)} g_1(y,z)dz$ erhält. Diese wird dann über $z$ integriert.

Die Reihenfolge der Integrationen ist dabei i.a. unwichtig. Man hat aber zu beachten, dass sich die Integrationsgrenzen natürlich ändern können, wenn man die Reihenfolge vertauscht (siehe das folgende Beispiel).

**Beispiel 8.3.1** (Viertelkreis in 2 Dimensionen)**.**
Als Beispiel wollen wir das Integral der Funktion $f(x,y) = x$ über einen Viertelkreis integrieren. Dabei wollen wir verschiedene Integrationsreihenfolgen betrachten. Ein solches Integral ist z.B. bei der Bestimmung des Schwerpunktes des Viertelkreises zu bestimmen.

a) Zuerst soll $y$ von $0$ bis $1$ laufen und $x$ dementsprechend gewählt werden. Zu festem $y$ läuft dann die $x-$Integration bis $\sqrt{1-y^2}$ (siehe Abb. 8.3.1).

$$\begin{aligned}\int\limits_{\text{Viertelkreis}} x\,dx\,dy &= \int_0^1 \left[\int_0^{\sqrt{1-y^2}} x\,dx\right] dy = \int_0^1 \left[\frac{1}{2}x^2\Big|_0^{\sqrt{1-y^2}}\right] dy \\ &= \int_0^1 \frac{1}{2}(1-y^2)dy = \frac{1}{2}(y-\frac{1}{3}y^3)\Big|_0^1 = \frac{1}{2}\cdot\frac{2}{3} = \frac{1}{3}\end{aligned}$$

b) Mit der umgekehrten Reihenfolge der Integrationen (erst $y-$, dann $x-$Integration) ergibt sich analog:

$$\begin{aligned}\int\limits_{\text{Viertelkreis}} x\,dx\,dy &= \int_0^1 \left[\int_0^{\sqrt{1-x^2}} x\,dy\right] dx = \int_0^1 \left[x\cdot y\Big|_0^{\sqrt{1-x^2}}\right] dx \\ &= \int_0^1 x\sqrt{1-x^2}\,dx = -\frac{1}{3}(1-x^2)^{3/2}\Big|_0^1 = \frac{1}{3}.\end{aligned}$$

Hier fixiert man erst ein $x$ und bestimmt dann für dieses $x$ den Bereich der erlaubten $y$. Dies bestimmt die Grenzen der $y$-Integration, die zuerst ausgeführt wird.

Das Beispiel zeigt, dass es zum einen nicht auf die Integrationsreihenfolge ankommt. Allerdings führt eine Änderung der Reihenfolge auch zu einer Änderung der Integrationsgrenzen. Zum Anderen sehen wir, dass die Details der Rechnung von der Reihenfolge der Integrationen abhängen. Man kann also meistens durch geschickte Wahl der Reihenfolge ein Problem vereinfachen!

## 8.4 Integration in krummlinigen Koordinaten

Die allgemeine Vorgehensweise für die Berechnung von Flächen- und Volumenintegralen ist in Abb. 8.4.1a) dargestellt. Dort wird die Form der infinitesimalen Flächen- bzw. Volumenelemente mit Hilfe der sog. **Koordinatenlinien** bestimmt. Diese erhält man, wenn man alle Koordinaten bis auf eine fixiert und diese freie Koordinate alle erlaubten Werte durchlaufen lässt.

In kartesischen Koordinaten findet eine Zerlegung der Fläche in infinitesimale Rechtecke der Fläche $dx\,dy$ statt. Diese erhält man aus der Fläche, die von infinitesimal benachbarten Koordinatenlinien begrenzt wird (siehe Abb. 8.4.1a)).
Bei der Integration in kartesischen Koordinaten ergeben sich i.a. Probleme mit der Wahl der Integrationsgrenzen, wenn der Körper nicht rechteckig ist. Durch passende Wahl des Koordinatensystemes läßt sich dieses Problem in vielen Fällen umgehen, z. B. mit Polarkoordinaten:

$$\int f(x,y)\,dx\,dy = \int f(r\cos\phi, r\sin\phi)\,r\,dr\,d\phi$$

Wir wollen dies an unserem Beispiel etwas genauer untersuchen.

**Beispiel 8.4.1** (Viertelkreis in Polarkoordinaten)**.**
Ideal geeignet für die Integration über den Viertelkreis sind Polarkoordinaten. Wir wollen wieder $\int\limits_{\text{Viertelkreis}} x\,dF$ berechnen, wobei in Polarkoordinaten $x = r\cos\phi$ ist. Durch eine analoge

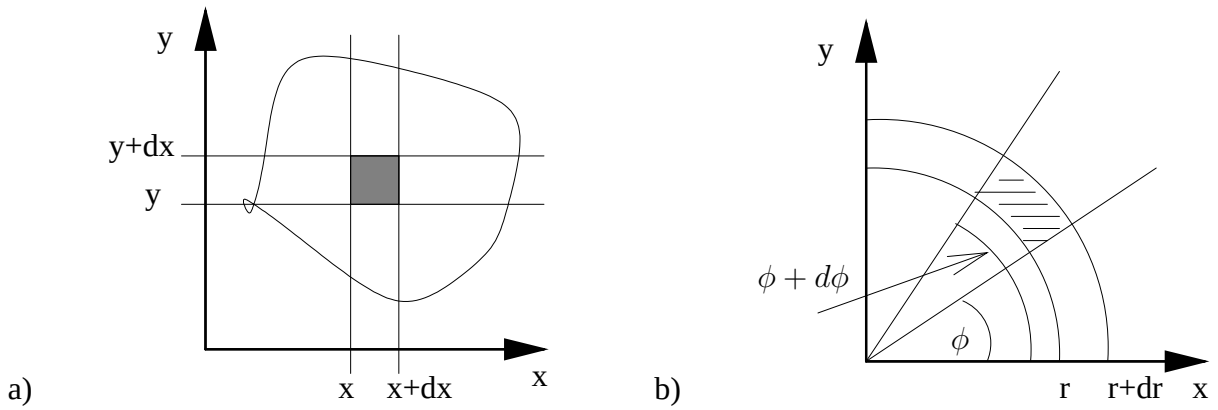



Abbildung 8.4.1: Flächenzerlegung in a) kartesischen Koordinaten bzw. b) ebenen Polarkoordinaten. Die Fläche, die durch die infinitesimal benachbarte Koordinatenlinien begrenzt wird, entspricht dem elementaren Flächenelement.

Konstruktion[2] wie bei der Bestimmung des Flächenelements $dF = dx\,dy$ in kartesischen Koordinaten erkennt man, dass in Polarkoordinaten gilt (siehe Abb. 8.4.1b))

$$dF = r dr\, d\phi.$$

Dies entspricht der schraffierten Fläche, die einem Kreissegment mit Bogenlänge $r d\phi$ und "Tiefe" $dr$ entspricht.

Um in Polarkoordinaten einen Viertelkreis vom Radius $1$ zu beschreiben, muss $r$ zwischen $0$ und $1$ variieren und $\phi$ zwischen $0$ und $\frac{\pi}{2}$. Man beachte dabei, dass für jedes $r$ der Winkel $\phi$ über das volle Intervall variieren muss (und umgekehrt), d.h. die Integrationsgrenzen sind unabhängig von $r$ und $\phi$. Damit ergibt sich:

$$\int\limits_{\text{Viertelkreis}} r\cos\phi\; r\, d\phi\, dr = \int_0^1 \int_0^{\pi/2} r^2 \cos\phi\, d\phi\, dr = \int_0^1 r^2 \sin\phi \Big|_0^{\pi/2} dr = \frac{1}{3} r^3 \Big|_0^1 = \frac{1}{3}.$$

Das Ergebnis stimmt natürlich mit dem in kartesischen Koordinaten überein. Der große Vorteil bei der obigen Integration ist natürlich, dass das zweidimensionale Integral in zwei eindimensionale Integrale faktorisiert, da die Integrationsgrenzen unabhängig von den Integrationsvariablen sind:

$$\int_0^1 \int_0^{\pi/2} r^2 \cos\phi\, d\phi\, dr \underset{\text{Grenzen fest}}{=} \left[\int_0^1 r^2\, dr\right] \left[\int_0^{\pi/2} \cos\phi\, d\phi\right].$$

Das gilt natürlich nicht allgemein, siehe z.B. die entsprechende Integration in kartesischen Koordinaten.

[2] d.h. durch infinitesimale Variation von $r$ und $\phi$ erhält man wieder $dF$ als eingeschlossen Fläche, siehe Abb. 8.4.1

**Satz 8.4.1** (Transformationssatz)**.**
Eine Verallgemeinerung der Substitutionsregel für eindimensionale Integrale auf mehrdimensionale Integrale ist der **Transformationssatz**. Gegeben sei eine Koordinatentransformation

$$(x, y, z) \to (u, v, w) \quad \text{mit} \quad x = x(u, v, w) \quad \text{etc.}$$

Dann transformiert sich ein Volumenintegral entsprechend

$$\int\limits_V f(x,y,z)dx\,dy\,dz = \int\limits_{\tilde{V}} f(x(u,v,w), y(u,v,w), z(u,v,w)) \cdot \left|\det\left(\frac{\partial(x,y,z)}{\partial(u,v,w)}\right)\right| \, du\,dv\,dw$$

mit der **Funktionaldeterminante** (bzw. **Jacobi-Determinante**)

$$\det\left(\frac{\partial(x,y,z)}{\partial(u,v,w)}\right) := \det\begin{pmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} & \frac{\partial x}{\partial w} \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v} & \frac{\partial y}{\partial w} \\ \frac{\partial z}{\partial u} & \frac{\partial z}{\partial v} & \frac{\partial z}{\partial w} \end{pmatrix}$$

Wie bei der bekannten Substitutionsregel ändert sich natürlich der Bereich, über den zu integrieren ist (hier von $V$ nach $\tilde{V}$). In einer Dimension ändert sich i.a. nur die Länge des Integrationsintervalls, in höheren Dimensionen ändert sich darüber hinaus i.a. die Form des Volumens. Da bei der Substitutionsregel auch die Ableitung der substituierten Funktion auftritt, ist es nicht überraschend, dass dies hier auch passiert. Da wir es mit Funktionen mehrerer Variablen zu tun haben, treten natürlich alle partiellen Ableitung auf. Warum diese gerade in einer speziellen Kombination, nämlich der Funktionaldeterminante, auftreten, soll hier nicht diskutiert werden.

Wir wollen noch darauf hinweisen, dass der Transformationssatz natürlich in analoger Form in beliebigen Dimensionen gilt. Für Funktionen von einer Variablen reduziert er sich auf die Substitutionsregel!

Im Transformationssatz tritt der Betrag der Jacobi-Determinante auf. Im folgenden werden wir nur Transformationen betrachten, bei der sich die Händigkeit des Koordinatensystems nicht ändert. In diesem Fall kann er weggelassen werden.

**Beispiel 8.4.2** (Ebene Polarkoordinaten)**.**
Wir wollen uns am Beispiel der ebenen Polarkoordinaten davon überzeugen, dass das transformierte Flächenelement tatächlich genau die Form hat, die wir vorher schon anschaulich aus Abb. 8.4.1 abgeleitet haben. Mit $x = r\cos\phi$ und $y = r\sin\phi$ folgt für die Funktionaldeterminante

$$\begin{aligned} \det\left(\frac{\partial(x,y)}{\partial(r,\phi)}\right) &= \det\begin{pmatrix} \frac{\partial x}{\partial r} & \frac{\partial x}{\partial \phi} \\ \frac{\partial y}{\partial r} & \frac{\partial y}{\partial \phi} \end{pmatrix} = \det\begin{pmatrix} \cos\phi & -r\sin\phi \\ \sin\phi & r\cos\phi \end{pmatrix} \\ &= r\cos^2\phi + r\sin^2\phi = r \end{aligned}$$

und damit

$$dF = \det\left(\frac{\partial(x,y)}{\partial(r,\phi)}\right) dr\,d\phi = r dr\,d\phi.$$

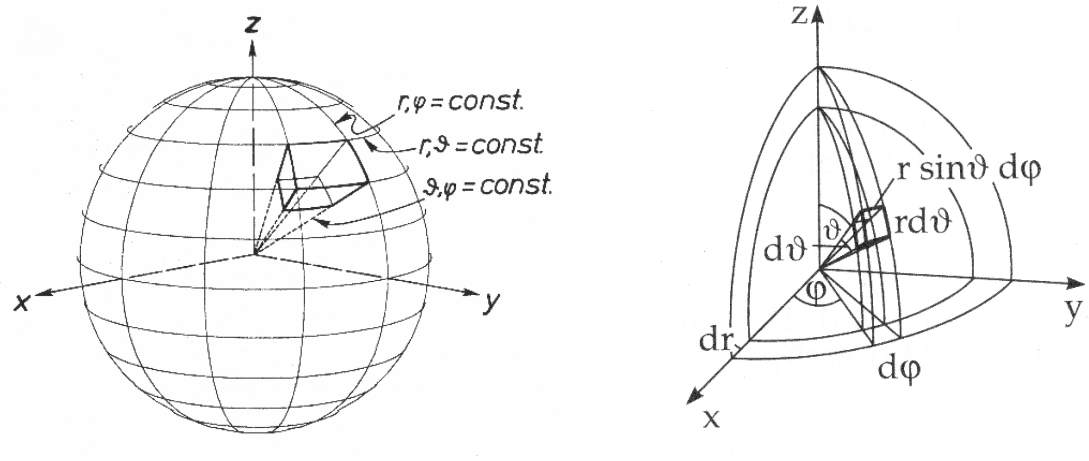



Abbildung 8.4.2: Infinitesimales Volumenelement in Kugelkoordinaten.

Die Determinante $\det\left(\frac{\partial(x,y,z)}{\partial(u,v,w)}\right)$ beschreibt die Verzerrung des infinitesimalen Volumenelementes unter der Koordinatentransformation. Speziell für die bereits bekannten Koordinatensysteme ergibt sich:

| | | |
|---|---|---|
| Kartesisch: | $dV = dx\,dy\,dz$ | $\curvearrowright \int f(x,y,z)dx\,dy\,dz$ |
| Ebene Polarkoordinaten: | $dF = r\,dr\,d\phi$ | $\curvearrowright \int f(r\cos\phi, r\sin\phi)\; r\,dr\,d\phi$ |
| Zylinderkoordinaten: | $dV = r\,dr\,d\phi\,dz$ | $\curvearrowright \int f(r\cos\phi, r\sin\phi, z)\; r\,dr\,d\phi\,dz$ |
| Kugelkoordinaten: | $dV = r^2\sin\theta\; dr\,d\phi\,d\theta$ | $\curvearrowright \int \tilde{f}(r,\phi,\theta)\; r^2\;\sin\theta\,dr\,d\phi\,d\theta$ |

wobei $\tilde{f}(r,\phi,\theta) = f(r\cos\phi\sin\theta, r\sin\phi\sin\theta, r\cos\theta)$.
Die Form des infinitesimalen Volumenelements in Zylinderkoordinaten kann man sich leicht vorstellen, z.B. durch (kartesische) Fortsetzung des Flächenelements in ebenen Polarkoordinaten (Abb. 8.4.1b). Es ähnelt einem Tortenstück, von dem man die Spitze abgeschnitten hat. Das infinitesimale Volumenelement in Kugelkoordinaten ist schwerer vorzustellen. Es ist in Abb. 8.4.2 dargestellt[3].

Ein Problem bei der Anwendung des Transformationssatzes ist die Bestimmung des neuen Integrationsbereiches $\tilde{V}$. Wie bei der Substitutionsregel besteht natürlich ein formaler Zusammenhang zwischen $V$ und $\tilde{V}$ über die Transformationsfunktionen $u$, $v$ und $w$. Statt diesen explizit auszunutzen, ist es in der Praxis meist einfacher, sich die neuen Integrationsgrenzen in den neuen Koordinaten anschaulich zu überlegen, genau so wie wir das in Beispiel 8.4.1 gemacht haben!

[3]Im Druck gibt es verschiedene Darstellungen für die griechischen (Klein-)Buchstaben phi ($\phi = \varphi$) und theta ($\theta = \vartheta$).

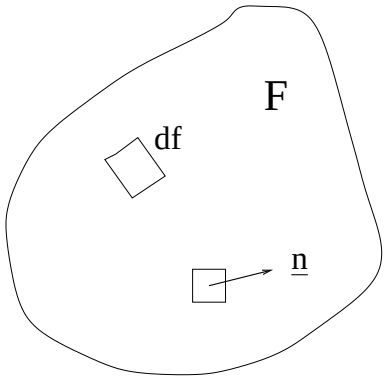



Abbildung 8.5.1: Eine Fläche $F$ bestehend aus (infinitesimalen) Flächenelementen $df$ mit Normalenvektoren $\underline{n}$.

## 8.5 Flächenintegrale und Vektorfluss

In Kapitel 8.4 haben wir gelernt, wie wir Skalarfelder in beliebigen Dimensionen integrieren können. Analog zum Wegintegral (siehe Kapitel 8.1), bei dem Vektorfelder längs eines Weges integriert werden, können wir auch Vektorfelder über (gerichtete) Flächen integrieren. Dies wird z.B. benötigt, um den Fluss durch eine Fläche zu bestimmen. Dies führt auf das Konzept des *Flussintegrals*, das man als zweidimensionale Verallgemeinerung des Wegintegrals auffassen kann. Die Bezeichnung "Fluss" ist durch die Hydrodynamik motiviert. Man kann sich tatsächlich eine durch die Oberfläche strömende Flüssigkeit vorstellen.
Wir betrachten eine Fläche im dreidimensionalen Raum[4], die beliebig geformt sein kann. Ihr Flächeninhalt (die Oberfläche) sei $F$. Wir stellen uns vor, dass die Fläche aus infinitesimalen Flächenelementen $df$ besteht (siehe Abb. 8.5.1), so dass

$$F = \int df.$$

**Definition 8.5.1** (Normale, Fächenvektor)**.**
Auf jedem infinitesimalen Flächenelement $df$ errichten wir einen senkrecht stehenden Einheitsvektor $\underline{n}$. $\underline{n}$ heißt auch **Normale**.
$\underline{df} = \underline{n}df$ heißt (infinitesimaler) **Flächenvektor**.

Bemerkung: Bisher sind für die Normale noch zwei Richtungen möglich. Für geschlossene Flächen (z.B. Kugeloberflächen) zeigt die Normale jedoch per Konvention immer nach außen.

**Definition 8.5.2** (Vektorfluss)**.**
Seien nun ein Vektorfeld $\underline{A}(\underline{r})$ und eine Fläche $F$ gegeben. Dann bezeichnet man

$$\int_F \underline{A} \cdot \underline{df} = \int_F \underline{A} \cdot \underline{n} df$$

[4]Die meisten Überlegungen lassen sich auf Hyperflächen beliebiger Dimension übertragen.

als den **Fluss von $\underline{A}$ durch** $F$. Dabei wird das Integral über die Oberfläche genommen.
Der Vektorfluss ist eine skalare Größe, die von $\underline{A}$ und der Art und Lage der Fläche bestimmt wird (siehe Abb. 8.5.2).
Die Darstellung auf der rechten Seite zeigt, dass man letztlich ein gewöhnliches zweidimensionales Integral über ein Skalarfeld zu berechnen hat. Man beachte aber dabei, dass der Normalenvektor $\underline{n}$ i.a. keine Konstante ist, sondern sich von Position zu Position auf der Fläche ändert!

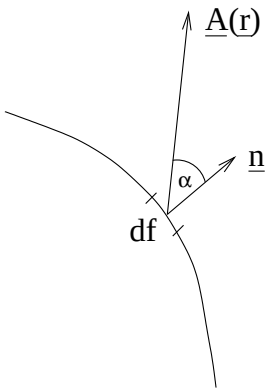



Abbildung 8.5.2: Fluss des Vektorfeldes $\underline{A}$ durch die Fläche $F$.

Sei $\alpha$ der Winkel zwischen $\underline{A}$ und $\underline{df}$ (bzw. $\underline{n}$) (siehe Abb. 8.5.2). Dann ist $\underline{A} \cdot \underline{df} = |\underline{A}| \cos \alpha df$. Ist $\underline{A}$ an der betrachteten Stelle senkrecht zu $\underline{df}$, d.h. $\underline{A}$ liegt in der Fläche, so ist $\underline{A} \cdot \underline{df} = 0$. Zum Vektorfluss trägt also nur der durch die (gekrümmte) Fläche jeweils senkrecht hindurchtretende Anteil von $\underline{A}$ bei.

**Beispiel 8.5.1.**
Als Beispiel betrachten wir eine Kugelschale $K_R$ vom Radius $R$ um den Ursprung. Der Normalenvektor zeigt dann immer radial nach außen, d.h. wir haben $\underline{n} = \hat{r} = \underline{r}/r$. Das gerichtete Flächenelement ist daher (in Kugelkoordinaten) gegeben durch

$$\underline{df} = R^2 \sin\theta d\theta d\varphi \, \hat{r} = R^2 d\Omega \, \hat{r},$$

wobei wir das **Raumwinkelelement**

$$d\Omega = \sin\theta d\theta d\phi$$

eingeführt haben. Die Form des Flächenelements macht man sich leicht aus der allgemeinen Form des Volumenselements in Kugelkoordinaten klar, da das Flächenelement keine "Ausdehnung" in radialer Richtung hat.
Wir wollen nun ein Radialfeld der Form

$$\underline{A}(\underline{r}) = A(r)\hat{r}$$

über die Kugelschale integrieren. Felder solcher Form nennt man **Radialfelder**. Sie zeigen immer in radiale Richtung und ihr Betrag hängt nur vom Abstand $r$ ab, nicht von der Richtung. Dieser Fall tritt in der Physik sehr häufig auf, z.B. bei Gravitations- oder elektrischen Feldern. Für das Flussintegral erhalten wir nun:

$$\begin{aligned}\int_{K_R} \underline{A}(\underline{r}) \cdot \underline{df} &= \int_0^{2\pi} d\varphi \int_0^{\pi} d\theta R^2 \sin\theta \underline{A}(\underline{r}) \cdot \hat{r} \\ &= A(R)R^2 \int_0^{2\pi} d\varphi \int_0^{\pi} d\theta R^2 \sin\theta \\ &= 4\pi R^2 A(R).\end{aligned}$$

Im Integral der ersten Zeilen haben wir die Parametrisierung der Fläche in Kugelkoordinaten benutzt, in der zweiten Zeile die spezielle Form des Vektorfeldes, insbesondere die Tatsache, dass sein Betrag auf der Kugelschale konstant ist. Letzteres macht auch das Endergebnis klar, denn offensichtlich ist der Wert des Flussintegrals gegeben durch das Produkt der Kugelfläche und dem (konstanten) Wert des Feldes auf dieser.

# Kapitel 9

# Die Integralsätze von Gauß und Stokes

In diesem Kapitel werden wir die wichtigsten Integralsätze kennenlernen. Diese stellen Zusammenhänge zwischen Integralen in verschiedenen Dimensionen her. Die Integralsätze von Gauß und Stokes spielen eine wichtige Rolle in der Elektrodynamik, im Zusammenhang mit den Maxwellschen Gleichungen. Sie erlauben die Ausnutzung von Symmetrien und vereinfachen daher in vielen Fällen die Berechnung von elektrischen und magnetischen Feldern drastisch.
Zunächst wollen wir uns allerdings den Greenschen Satz anschauen, mit dessen Hilfe wir dann allgemeinere Sätze ableiten können.

## 9.1 Der Greensche Satz in der Ebene

Wir betrachten eine Fläche $F$, deren Rand durch die Randkurve $\mathcal{C} = \partial F$ gegeben ist. Wir nehmen o.B.d.A. an, dass es für jeden Wert von $x$ genau eine obere und eine untere Grenze gibt[1]. Dies definiert die untere bzw. obere Grenzkurve $f_1(x)$ bzw. $f_2(x)$ (siehe Abb. 9.1.1) und für alle $x \in [a, b]$ gilt $f_1(x) \leq y \leq f_2(x)$, falls der Punkt $(x, y)$ in der Fläche $F$ liegt.

Es sei nun $V(x, y)$ eine Funktion, die auf $F$ stetig partiell differenzierbar sei. Dann gilt:

$$\begin{aligned}
\int_F dxdy\, \frac{\partial V(x,y)}{\partial y} &= \int_a^b dx \left[ \int_{f_1(x)}^{f_2(x)} dy \frac{\partial V(x,y)}{\partial y} \right] \\
&= \int_a^b dx \left[ V(x, f_2(x)) - V(x, f_1(x)) \right] \\
&= -\int_b^a dx V(x, f_2(x)) - \int_a^b dx V(x, f_1(x)) \\
&= -\oint_{\mathcal{C}} dx V(x,y)\,,
\end{aligned}$$

d.h. das Flächenintegral kann in ein Wegintegral umgeschrieben werden!

[1] Ansonsten muß man die kompliziertere Fläche aus solchen einfachen zusammensetzen.

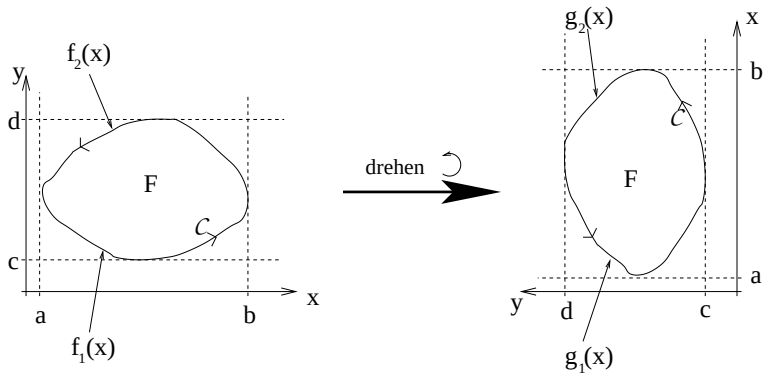



Abbildung 9.1.1: Zur Herleitung des Greenschen Satzes.

Wir drehen nun die Fläche, wie in Fig. 9.1.1 dargestellt und gehen in der gedrehten Version analog vor:

$$\int_F dxdy \, \frac{\partial V(x,y)}{\partial x} = \oint_{\mathcal{C}} dy V(x,y) \, .$$

Wir fassen nun diese beiden Teilergebnisse zusammen:

**Satz 9.1.1** (Greenscher Satz)**.**
Der **Satz von Green** für Funktionen $V_1(x,y)$, $V_2(x,y)$ die auf der zweidimensionalen Fläche $F$ stetig partiell differenzierbar sind, lautet

$$\boxed{\int_F dxdy \left[ \frac{\partial V_2(x,y)}{\partial x} - \frac{\partial V_1(x,y)}{\partial y} \right] = \oint_{\mathcal{C}} \left[ V_1(x,y)dx + V_2(x,y)dy \right]}$$

wobei $\mathcal{C} = \partial F$ die Randkurve von $F$ ist.

Wir wollen zwei Anwendungen dieses Satzes diskutieren.

1. Sei
$$V_1(x,y) = \frac{\partial \phi}{\partial x} \, , \qquad V_2(x,y) = \frac{\partial \phi}{\partial y} \, .$$
Dann verschwindet das Flächenintegral im Greenschen Satz, da
$$\frac{\partial V_2}{\partial x} = \frac{\partial^2 \phi}{\partial x \partial y} = \frac{\partial^2 \phi}{\partial y \partial x} = \frac{\partial V_1}{\partial y}$$
ist. Außerdem gilt mit $\underline{r} = \begin{pmatrix} x \\ y \end{pmatrix}$:
$$V_1 dx + V_2 dy = d\phi = \mathrm{grad}\, \phi \cdot d\underline{r} \, .$$

Somit lautet die Aussage des Greenschen Satzes in diesem Fall

$$\oint_{\mathcal{C}} d\phi = 0\,.$$

Dies wissen wir natürlich schon, da $\underline{V} := \begin{pmatrix} V_1 \\ V_2 \end{pmatrix}$ per Definition ein konservatives Feld mit Potential $-\phi$ ist.

2. Für $\underline{V} := \begin{pmatrix} V_1 \\ V_2 \end{pmatrix}$ gilt

$$\oint_{\mathcal{C}} \underline{V} \cdot \underline{dr} = \int_F \operatorname{rot} \underline{V} \cdot \underline{df}\,,$$

wobei $\underline{df} = \underline{n} dxdy$ senkrecht zur $x-y$-Ebene steht und auf der rechten Seite $\underline{V}$ als dreidimensionaler Vektor mit 3. Komponente $V_3 = 0$ zu interpretieren ist.

3. Wir können mit Hilfe des Greenschen Satzes Flächen berechnen. Dazu betrachten wir die Funktionen $V_1(x,y) = -y$ und $V_2(x,y) = x$. Dann gilt:

$$\oint_{\mathcal{C}} [-ydx + xdy] = \int\int_F dxdy \left[\frac{\partial x}{\partial x} - \left(-\frac{\partial y}{\partial y}\right)\right] = 2\int\int_F dxdy = 2F\,.$$

Als Beispiel bestimmen wir die Fläche einer Ellipse, die durch

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1 \tag{9.1.1}$$

definiert ist[2]. Eine Parameterdarstellung ist durch $\underline{r}(t) = (a\cos t, b\sin t)$ gegeben ($t \in [0, 2\pi]$). Dann gilt:

$$F_{\text{Ellipse}} = \frac{1}{2}\oint_{\text{Ellipse}} [-ydx + xdy] = \frac{1}{2}\int_0^{2\pi} [-b\sin t(-a\sin t) + a\cos t b\cos t]\, dt = ab\pi\,.$$

## 9.2 Integraldarstellung der Rotation

Die Rotation eines Vektorfeldes $\underline{A}(\underline{r})$ läßt sich als Wegintegral ausdrücken. Diese Darstellung wird in den folgenden Abschnitten sehr nützlich sein.
Wir betrachten eine geschlossenen Weg $\mathcal{C}$ um den Punkt $\underline{r}$, an dem wir die Rotation bestimmen wollen. Dieser Weg schließt die Fläche $\Delta F$ ein. Ist diese hinreichend klein, so können wir sie als eben ansehen und durch den Normalenvektor $\underline{n}$ charakterisieren, der senkrecht auf ihr steht. Die Komponente der Rotation parallel zu $\underline{n}$ lässt sich dann folgendermaßen bestimmen:

$$\boxed{\underline{n} \cdot \operatorname{rot} \underline{A}(\underline{r}) = \lim_{\Delta F \to 0} \frac{1}{\Delta F} \oint_{\mathcal{C}} \underline{A} \cdot d\underline{r}.}$$

[2]D.h. alle $(x,y)$, die diese Gleichung erfüllen, liegen auf der Ellipse.

Um die Rotation vollständig zu bestimmen, kann man z.B. Wege nehmen, die in der $x-y$-, $x-z$- und $y-z$-Ebene liegen, wobei dann $\underline{n}$ der Einheitsvektor in $z$-, $y$- und $x$-Richtung ist.

Zum Beweis dieser Integraldarstellung betrachten wir o.B.d.A. ein infinitesimales rechteckiges Flächenelement (Abb. 9.2.1) in der $y-z$-Ebene. Das Wegintegral längs des Weges $1 \to 2 \to 3 \to 4 \to 1$ läßt sich dann folgendermaßen berechnen:

$$\begin{aligned}
\oint_{\mathcal{L}} \underline{A} \cdot d\underline{r} &= A_y(x,y,z-\frac{1}{2}dz)dy + A_z(x,y+\frac{1}{2}dy,z)dz \\
&- A_y(x,y,z+\frac{1}{2}dz)dy - A_z(x,y-\frac{1}{2}dy,z)dz \\
&= \frac{\partial A_z(\underline{r})}{\partial y}dydz - \frac{\partial A_y(\underline{r})}{\partial z}dydz \\
&= \left(\frac{\partial A_z}{\partial y} - \frac{\partial A_y}{\partial z}\right)\Delta F = (\operatorname{rot}\underline{A})_x \,\Delta F.
\end{aligned}$$

Dabei haben wir angenommen, dass die Kantenlängen $dy$ und $dz$ infinitesimal sind und deshalb der Wert das Vektorfeldes auf einer Kante konstant ist. Hierbei haben wir den Funktionswert in der Kantenmitte gewählt. Für die Kante $1 \to 2$ sind die $x$- und $z$-Koordinaten konstant (nämlich $x$ und $z-\frac{1}{2}dz$) und die $y$-Koordinate variiert von $y-\frac{1}{2}dy$ nach $y+\frac{1}{2}dy$. Die Kantenmitte ist also bei $(x,y,z-\frac{1}{2}dz)$ und deshalb ist $A_y(x,y,z-\frac{1}{2}dz)$ zu wählen[3]. Beim Übergang zur dritten Zeile sind wir analog vorgegangen wie beim Beweis des Satzes von Gauß. Hier wurden die auftretenden Beiträge Taylor-entwickelt, wobei sich die konstanten Terme wegheben.

Somit haben wir für $\underline{n} = \hat{x}$ gezeigt

$$\lim_{\Delta F \to 0} \frac{1}{\Delta F} \oint_{\Delta\mathcal{L}} \underline{A} \cdot d\underline{r} = (\operatorname{rot}\underline{A})_x = \underline{n} \cdot \operatorname{rot}\underline{A}.$$

Mit einem analogen Argument läßt sich diese Aussage auch für die $y-$ und $z-$Komponenten beweisen, womit $(**)$ gezeigt wäre, und damit der Stokessche Integralsatz.

## 9.3 Integralsatz von Stokes

Es sei $\underline{A}(\underline{r})$ ein beliebiges Vektorfeld. Aus Kapitel 8.2 wissen wir, dass gilt (falls der Definitionsbereich von $\underline{A}$ einfach-zusammenhängend ist):

$$\underline{A} = -\operatorname{grad}\varphi \quad \Longleftrightarrow \quad \operatorname{rot}\underline{A} = 0 \quad \Longleftrightarrow \quad \oint_{\mathcal{C}} \underline{A} \cdot d\underline{r} = 0,$$

wobei die rechte Seite für alle geschlossenen Wege $\mathcal{C}$ gelten soll.

[3] Es trägt nur die $y$-Komponente des Feldes bei, da sich auf dem Weg nur die $y$-Koordinate ändert und deshalb $\underline{A} \cdot d\underline{r} = A_y dy$ ist.

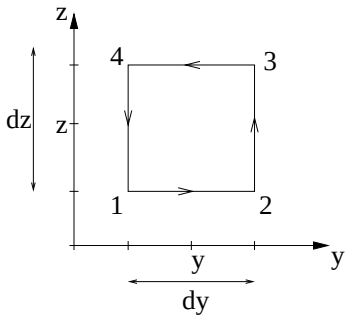



Abbildung 9.2.1: Zum Beweis der Integraldarstellung der Rotation.

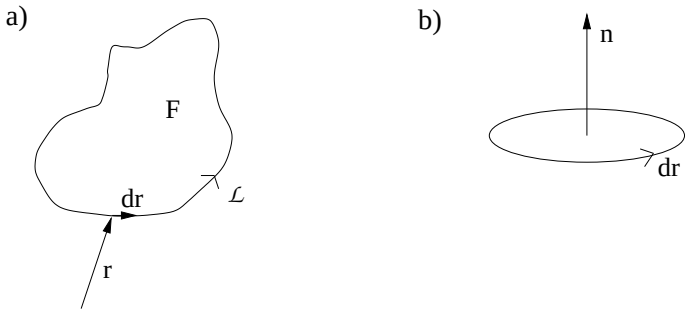



Abbildung 9.3.1: Zum Satz von Stokes. a) Fläche $F$ mit Rand $\mathcal{C}$. b) Orientierung der Flächennormalen (rechte-Hand-Regel).

Was passiert für $\mathrm{rot}\,\underline{A} \neq 0$ ? Man erwartet, dass dann i.a. auch $\oint_{\mathcal{L}} \underline{A} \cdot d\underline{r} \neq 0$ sein wird. Dies bringt der Integralsatz von Stokes (in quantitativer Form) zum Ausdruck.

**Satz 9.3.1** (Stokes'scher Integralsatz)**.**
Sei $F$ eine beliebige (gekrümmte) Fläche im $\mathbb{R}^3$ mit Rand $\mathcal{C}$, d.h. der Rand ist eine geschlossene Kurve (siehe Abb. 9.3.1a). Dann gilt:

$$\boxed{\oint_{\mathcal{C}} \underline{A} \cdot d\underline{r} = \int_F \mathrm{rot}\,\underline{A} \cdot \underline{df}.}$$

Dabei ist das infinitesimale Flächenelement $\underline{df} = \underline{n} df$ gemäß einer rechten-Hand-Regel (siehe Abb. 9.3.1b) orientiert.
Der Satz von Stokes besagt also, dass das Wegintegral von $\underline{A}$ längs des geschlossenen Weges $\mathcal{C}$ gleich dem Fluss von $\mathrm{rot}\,\underline{A}$ durch die von $\mathcal{C}$ berandete Fläche $F$ ist.

Beim Beweis gehen wir ähnlich vor wie beim Gaußschen Satz. Zunächst unterteilen wir die Fläche $F$ in zwei Teile $F_1$ und $F_2$. Diese Teile werden von den geschlossenen Kurven $\mathcal{C}_1$ und $\mathcal{C}_2$

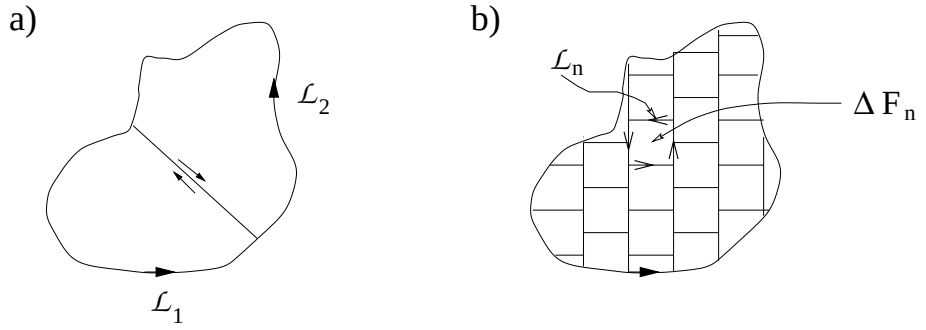



Abbildung 9.3.2: Zum Beweis des Stokesschen Integralsatzes.

berandet (Abb. 9.3.2a). Dann gilt:

$$\oint_{\mathcal{C}} \underline{A} \cdot d\underline{r} = \oint_{\mathcal{C}_1} \underline{A} \cdot d\underline{r} + \oint_{\mathcal{C}_2} \underline{A} \cdot d\underline{r}.$$

Dabei haben wir schon ausgenutzt, dass die Schnittfläche von $\mathcal{C}_1$ und $\mathcal{C}_2$ in unterschiedlichen Richtungen durchlaufen wird und sich die entsprechenden Beiträge deshalb wegheben.
Die obigen Teilungsprozedur setzen wir nun immer weiter fort (Abb. 9.3.2b). Dann erhalten wir:

$$\begin{aligned} \oint_{\mathcal{C}} \underline{A} \cdot d\underline{r} &= \sum_n \oint_{\mathcal{C}_n} \underline{A} \cdot d\underline{r} = \sum_n \left[ \frac{1}{\Delta F_n} \oint_{\mathcal{C}_n} \underline{A} \cdot d\underline{r} \right] \Delta F_n \\ &\rightarrow \int_F \operatorname{rot} \underline{A} \cdot \underline{n} df \end{aligned}$$

wobei wir beim Übergang zur zweiten Zeile die Summe wieder als Zwischensumme eines Integrals interpretiert haben.
Der Integralsatz von Stokes hat zahlreiche wichtige Anwendungen in der Physik, insbesondere in der Elektrodynamik im Zusammenhang mit den Maxwellschen Gleichungen.

## 9.4 Integraldarstellung der Divergenz

Eine andere Interpretation der Divergenz erhält man durch Betrachtung des Vektorflusses durch die Oberfläche $\Delta F$ eines Volumenelements $\Delta V$ (siehe Abb. 9.4.1):

$$\boxed{\operatorname{div} \underline{A} = \lim_{\Delta V \to 0} \frac{1}{\Delta V} \int_{\Delta F} \underline{A} \cdot d\underline{f}.}$$

Die Divergenz des Vektorfeldes kann also als Grenzfall des Flusses durch die Oberfläche eines kleinen Volumenelements aufgefasst werden. Dies quantifiziert unsere anschauliche Interpretation der Divergenz als “Quellstärke” des Feldes. Befindet sich eine Quelle in dem betrachteten

Abbildung 9.4.1: Fluß durch $\Delta V$

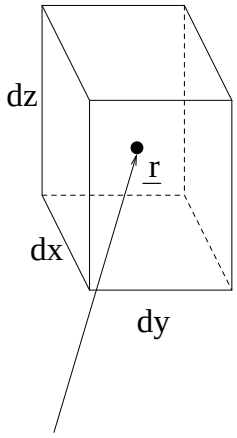



Abbildung 9.4.2: Zum Gaußschen Satz: Fluss durch einen infinitesimalen Würfel mit den Kantenlängen $dx$, $dy$ und $dz$. Das Zentrum des Würfels befindet sich bei $\underline{r} = (x, y, z)$.

Volumen, so strömt mehr heraus als herein. Dann ist $\mathrm{div}\,\underline{A} > 0$. Entsprechend ist bei einer Senke $\mathrm{div}\,\underline{A} < 0$.

Zum Beweis der Integraldarstellung betrachten wir den Fluss durch die Oberfläche eines Würfels mit den Kantenlängen $dx$, $dy$ und $dz$ um $\underline{r} = (x, y, z)$ (siehe Abb. 9.4.2). Auf der Vorderseite $F_1$ zeige der Normalenvektor $\underline{n}$ in $x$-Richtung, auf der Rückseite $F_2$ in $-x$-Richtung. Die Flüsse durch $F_1$ und $F_2$ sind dann

$$\begin{aligned} \text{Fluss durch } F_1 &= A_x(x + \frac{1}{2}dx, y, z)dydz, \\ \text{Fluss durch } F_2 &= A_x(x - \frac{1}{2}dx, y, z)dydz \end{aligned}$$

wobei $A_x(\underline{r}) = A_x(x, y, z)$ die $x$-Komponente des Vektorfeldes am Ort $\underline{r}$, dem Zentrum des Würfels, bezeichnet. Wir haben hier angenommen, dass das Vektorpotential auf den Flächen als konstant angesehen werden kann, wobei der Wert durch den Wert in der Mitte gegeben ist.

Somit erhält man für den Fluss durch die beiden Flächen $F_1$ und $F_2$:

$$\begin{aligned}\int_{F_1+F_2} \underline{A}\cdot\underline{df} &= \left[A_x(x+\frac{1}{2}dx,y,z)-A_x(x-\frac{1}{2}dx,y,z)\right]dydz\\ &= \left[\left(A_x(x,y,z)+\frac{\partial A_x}{\partial x}\,\frac{1}{2}dx\right)-\left(A_x(x,y,z)-\frac{\partial A_x}{\partial x}\,\frac{1}{2}dx\right)\right]dydz\\ &= \frac{\partial A_x(\underline{r})}{\partial x}\,dxdydz.\end{aligned}$$

Dabei haben wir beim Übergang zur zweiten Zeile ausgenutzt, dass der Würfel infinitesimal ist und wir deshalb die Feldwerte durch eine Taylor-Entwicklung erster Ordnung approximieren können.
Analog erhalten wir dann für den Fluss in $y$- und $z$-Richtung:

$$\begin{aligned}y-\text{Richtung:}\qquad &\int_{F_3+F_4}\underline{A}\cdot\underline{df}=\frac{\partial A_y(\underline{r})}{\partial y}\,dxdydz,\\ z-\text{Richtung:}\qquad &\int_{F_5+F_6}\underline{A}\cdot\underline{df}=\frac{\partial A_z(\underline{r})}{\partial z}\,dxdydz.\end{aligned}$$

Fasst man nun diese drei Teilergebnisse zusammen, so folgt die angegebene Integraldarstellung der Divergenz.

## 9.5 Gaußscher Satz

Ziel des Gaußschen Integralsatzes ist es, eine andere Darstellung für ein Flussintegral der Form $\oint \underline{A}\cdot\underline{df}$ anzugeben.

**Satz 9.5.1** (Gaußscher Integralsatz)**.**
Sei $\underline{A}(\underline{r})$ ein Vektorfeld und $F$ eine beliebige geschlossene Fläche, die ein Volumen $V$ umschließt. Dann gilt:

$$\boxed{\oint_F \underline{A}\cdot\underline{df}=\int_V \operatorname{div}\underline{A}\,dV.}$$

Der **Gaußsche Integralsatz** erlaubt also die Umwandlung eines Flussintegrals in ein Volumenintegral über die Divergenz des Vektorfeldes über das von der Fläche eingeschlossene Volumen.

Zum Beweis teilen wir zunächst das Volumen $V$ in zwei (beliebige) Teilvolumina $V_1$ und $V_2$ (siehe Abb. 9.5.1). Die neue Oberfläche der beiden Teilvolumina bezeichnen wir mit $F_1$ bzw. $F_2$. Man beachte, dass eine gemeinsame Trennfläche zwischen den beiden Teilvolumina entsteht. Da die Normalen jeweils aus dem Volumen herauszeigen, ist auf der gemeinsamen Trennfläche $\underline{n}_1=-\underline{n}_2$ (siehe Abb. 9.5.1) und somit

$$\oint_F \underline{A}\cdot\underline{df}=\oint_{F_1}\underline{A}\cdot\underline{df}+\oint_{F_2}\underline{A}\cdot\underline{df},$$

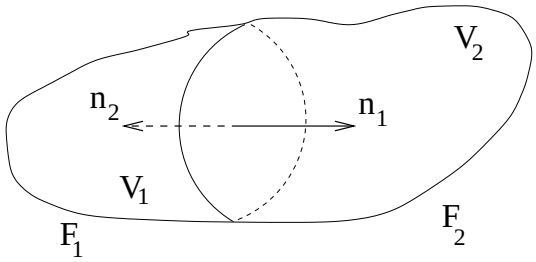



Abbildung 9.5.1: Zum Gaußschen Satz: Unterteilung des Volumens $V$ in zwei Teilvolumina $V_1$ und $V_2$. $\underline{n}_1$, $\underline{n}_2$ sind die Normalenvektoren auf der Trennfläche.

da sich die Beiträge auf der Trennfläche wegheben.

Wir setzen nun diese Teilungsprozedur immer weiter fort und erhalten so eine Einteilung in viele kleine Volumina $\Delta V_n$ mit Oberfläche $F_n$:

$$\begin{aligned} \oint_F \underline{A} \cdot \underline{df} \quad &= \quad \sum_{n=1}^{N} \oint_{F_n} \underline{A} \cdot \underline{df} = \sum_{n=1}^{N} \left[ \frac{1}{\Delta V_n} \oint_{F_n} \underline{A} \cdot \underline{df} \right] \Delta V_n \\ &\underset{\Delta V \to 0}{\to} \quad \int_V \left( \lim_{\Delta V \to 0} \frac{1}{\Delta V} \oint_{F(\Delta V)} \underline{A} \cdot \underline{df} \right) dV = \int_V \operatorname{div} \underline{A} \, dV \end{aligned}$$

wobei $F(\Delta V)$ die Oberfläche des Würfels $\Delta V$ um den Punkt $\underline{r}$ ist. Im letzten Schritt haben wir die Integraldarstellung der Divergenz aus Abschnitt 9.4 verwendet.

Eine wichtige Anwendung des Gaußschen Integralsatzes ist die Herleitung der sogenannten **Kontinuitätsgleichung**

$$\boxed{\frac{\partial \rho}{\partial t} + \operatorname{div} \underline{j} = 0.}$$

Gleichungen dieses Typs treten in der Physik immer wieder auf, z.B. in der Hydrodynamik, denn sie der mathematische Ausdruck für ein lokales (d.h. von $\underline{r}$ abhängiges) Erhaltungsgesetz. $\rho(\underline{r}, t)$ bezeichnet dabei immer eine Dichte (z.B. eine Massen- oder Ladungsdichte) und $\underline{j}(\underline{r}, t)$ die zugehörige Stromdichte (z.B. den Massen- oder elektrischen Strom). In den Übungen werden wir die Kontinuitätsgleichung aus sehr allgemeinen Überlegungen ableiten.

# Kapitel 10

# Operatoren und Eigenwerte

## 10.1 Eigenwerte von Matrizen

**Definition 10.1.1** (Eigenwert, Eigenvektor)**.**
Es sei $\underline{\underline{A}}$ eine quadratische Matrix. Gilt dann

$$\boxed{\underline{\underline{A}} \cdot \underline{v} = \lambda \underline{v}}$$

so heißt $\lambda$ **Eigenwert** von $\underline{\underline{A}}$ und $\underline{v}$ **Eigenvektor** (zum Eigenwert $\lambda$)[1]. Dabei ist $\lambda$ eine reelle oder komplexe Zahl.

Bem.: Die Eigenvektoren zum Eigenwert $\lambda$ sind nicht eindeutig festgelegt, da mit $\underline{v}$ auch jedes Vielfache $\alpha\underline{v}$ (mit $\alpha \neq 0$) Eigenvektor zum gleichen Eigenwert ist.
Die Menge aller Eigenwerte einer Matrix bezeichnet man als **Spektrum** der Matrix, die Menge aller Eigenwerte und Eigenvektoren als **Eigensystem**.

**Beispiel 10.1.1.**

1. Bei Drehungen gehen Punkte auf der Drehachse in sich über: $\underline{\underline{A}}\underline{v} = \underline{v}$.

2. Wir betrachten die Matrix $\underline{\underline{A}} = \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix}$. Dann gilt

$$\underline{\underline{A}} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad \text{und} \quad \underline{\underline{A}} \begin{pmatrix} 0 \\ 1 \end{pmatrix} = 2 \begin{pmatrix} 0 \\ 1 \end{pmatrix} .$$

   Die Matrix $\underline{\underline{A}}$ hat also die Eigenwerte 1 und 2 mit den zugehörigen Eigenvektoren $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$ bzw. $\begin{pmatrix} 0 \\ 1 \end{pmatrix}$.

Wir wollen uns nun mit der Frage beschäftigen, wie man die Eigenwerte und -vektoren konkret berechnen kann. Dies bezeichnet man als das **Eigenwertproblem**.

[1] Man beachte, dass man im Englischen von *eigenvalue* und *eigenvector* spricht!

Die Eigenwerte und -vektoren sind als Lösungen des linearen Gleichungssystems

$$\left(\underline{\underline{A}} - \lambda \mathbb{1}\right) \underline{v} = 0$$

definiert. Man beachte, dass man hierbei auch das $\lambda$ a priori nicht kennt! Das Gleichungssystem hat nicht-triviale Lösungen, falls

$$\boxed{\det\left(\underline{\underline{A}} - \lambda \mathbb{1}\right) = 0}$$

ist. Ansonsten existiert nur die triviale Lösung $\underline{v} = \underline{0}$. Diese Gleichung bezeichnet man auch als **Säkulargleichung**. Ist $\underline{\underline{A}}$ eine $n \times n$-Matrix, so ist liefert die Determinante ein Polynom $n$-ten Grades in $\lambda$, das sog. **charakteristische Polynom** $p(\lambda)$.
Nach dem Fundamentalsatz der Algebra hat ein Polynom $n$-ten Grades $n$ Nullstellen, die aber komplex sein können. Wir bezeichnen die Nullstellen von $p(\lambda)$ mit $\lambda_1, \lambda_2, \ldots, \lambda_n$. Sind einige der $\lambda_j$ gleich, so spricht man von **Nullstellen höherer Ordnung** bzw. **Entartung**. Man charakterisiert diese durch die **algebraische Multiplizität** oder **algebraische Vielfachheit**. Ist z.B. $\lambda_1 = \lambda_2$, so hat der Eigenwert $\lambda_1$ die algebraische Multiplizität 2.

Sind die Eigenwerte $\lambda_j$ bekannt, so kann man die zugehörigen Eigenvektoren $\underline{v}_j$ als Lösungen des linearen Gleichungssystems

$$\underline{\underline{A}} \cdot \underline{v}_j = \lambda_j \underline{v}_j$$

explizit bestimmen.

**Beispiel 10.1.2.** Als Beispiel betrachten wir die Matrix

$$\underline{\underline{A}} = \begin{pmatrix} 10 & -3 \\ -3 & 2 \end{pmatrix}.$$

Zunächst bestimmen wir mit Hilfe der Säkulargleichung das charakteristische Polynom:

$$p(\lambda) := \det\left(\underline{\underline{A}} - \lambda \mathbb{1}\right) = \det \begin{pmatrix} 10-\lambda & -3 \\ -3 & 2-\lambda \end{pmatrix} = (10-\lambda)(2-\lambda) - (-3)(-3) = \lambda^2 - 12\lambda + 11.$$

$p(\lambda)$ hat die Nullstellen $\lambda_1 = 1$ und $\lambda_2 = 11$, die also jeweils die algebraische Vielfachheit 1 haben und die Eigenwerte von $\underline{\underline{A}}$ sind.
Einen Eigenvektor zu $\lambda_1$ finden wir durch Lösung des linearen Systems

$$\left(\underline{\underline{A}} - \lambda_1 \mathbb{1}\right) \underline{v}_1 = 0.$$

Explizit lautet dieses System

$$\begin{pmatrix} 9 & -3 \\ -3 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ y_1 \end{pmatrix} = 0.$$

Die Lösung ist $x_1 = t$ und $y_1 = 3t$ mit beliebigem $t \in \mathbb{R}$, $t \neq 0$. Somit haben wir die Eigenvektoren zum Eigenwert $\lambda_1 = 1$ bestimmt:

$$\underline{v}_1 = \begin{pmatrix} t \\ 3t \end{pmatrix}.$$

Bei der Bestimmung des Eigenvektors zum Eigenwert $\lambda_2 = 11$ gehen wir analog vor. Hier haben wir das System

$$\left(\underline{\underline{A}} - \lambda_2 \mathbb{1}\right) \underline{v}_2 = \begin{pmatrix} -1 & -3 \\ -3 & -9 \end{pmatrix} \begin{pmatrix} x_2 \\ y_2 \end{pmatrix} = 0$$

zu lösen. Wir erhalten $x_2 = t$ und $y_2 = -\frac{t}{3}$.
In der Regel gibt man die normierten Einheitsvektoren an, also hier

$$\underline{v}_1 = \frac{1}{\sqrt{10}} \begin{pmatrix} 1 \\ 3 \end{pmatrix}, \qquad \underline{v}_2 = \frac{1}{\sqrt{10}} \begin{pmatrix} 3 \\ -1 \end{pmatrix}.$$

Man beachte, dass offensichtlich $v_1$ und $v_2$ senkrecht aufeinander stehen. Wir werden später Kriterien dafür angegeben, wann dies der Fall ist.

*Bemerkung.*

1. Allgemein sind die Eigenvektoren zu verschiedenen Eigenwerte linear unabhängig!

2. Es gibt nicht immer gleich viele Eigenvektoren wie Eigenwerte. Zum Beispiel hat die Matrix
$$\underline{\underline{A}} = \begin{pmatrix} 0 & 1 \\ -1 & -2 \end{pmatrix}$$
den entarteten Eigenwert $\lambda = -1$ (mit algebraischer Multiplizität 2), aber nur einen Eigenvektor $\underline{v} = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$. Die Zahl der Eigenvektoren zu einem bestimmten Eigenwert nennt man seine **geometrische Multiplizität** oder **geometrische Vielfachheit**. Sie ist kleiner oder gleicher der algebraischen Multiplizität.

Wenn wir die Eigenwerte und Eigenvektoren einer Matrix $\underline{\underline{A}}$ kennen, können wir die Matrix unter bestimmten Umständen **diagonalisieren**, d.h. durch eine **Ähnlichkeitstransformation**[2] in Diagonalgestalt bringen. Dies entspricht einem Basiswechsel.
Wir betrachten hierzu eine $n \times n$-Matrix $\underline{\underline{A}}$, deren Eigenwerte $\lambda_j$ und zugehörigen Eigenvektoren $\underline{v}_j$ wir kennen:

$$\underline{\underline{A}} \cdot \underline{v}_j = \lambda_j \underline{v}_j \,.$$

Wir nehmen nun an, dass alle Eigenwerte reell und die Eigenvektoren $\underline{v}_j$ paarweise orthogonal sind, d.h. $\overline{\underline{v}_j} \cdot \underline{v}_l = \delta_{jl}$. Der linke Faktor ist komplex zu konjugieren, in Übereinstimmung mit der Definition des Skalarproduktes in Abschnitt 11.1. Außerdem sollen die Eigenvektoren normiert sein, d.h. $|\underline{v}_j| = 1$.
Wir definieren nun die Matrix

$$\underline{\underline{U}} := (\underline{v}_1, \ldots, \underline{v}_n) \,,$$

deren Spalten aus den Eigenvektoren gebildet werden, d.h. es gilt $U_{jl} = (\underline{v}_l)_j$.

[2] Das sind Transformationen der Form $\underline{\underline{M}}^{-1} \cdot \underline{\underline{A}} \cdot \underline{\underline{M}}$.

Für diese Matrix gilt

$$\underline{\underline{U}} \cdot \underline{\underline{U}}^\dagger = \underline{\underline{U}}^\dagger \cdot \underline{\underline{U}} = \begin{pmatrix} \cdots & \overline{\underline{v}_1} & \cdots \\ \cdots & \overline{\underline{v}_2} & \cdots \\ & \vdots & \\ \cdots & \overline{\underline{v}_n} & \cdots \end{pmatrix} \begin{pmatrix} \vdots & \vdots & \vdots \\ \underline{v}_1 & \underline{v}_2 & \underline{v}_n \\ \vdots & \vdots & \vdots \end{pmatrix} = \begin{pmatrix} \overline{\underline{v}_1} \cdot \underline{v}_1 & \overline{\underline{v}_1} \cdot \underline{v}_2 & & \cdots \\ \overline{\underline{v}_2} \cdot \underline{v}_1 & \overline{\underline{v}_2} \cdot \underline{v}_2 & & \cdots \\ \vdots & \vdots & \ddots & \\ & & & \overline{\underline{v}_n} \cdot \underline{v}_n \end{pmatrix} = \mathbb{1} \,,$$

wobei

$$\underline{\underline{U}}^\dagger := \left(\underline{\underline{\bar{U}}}\right)^T \,,$$

die **adjungierte Matrix** (oder auch **hermitesch konjugierte Matrix**) ist. Sie entsteht durch Bildung der Transponierten der Matrix, die aus den komplex-konjugierten Elementen $\overline{U_{jl}}$ der Matrix $\underline{\underline{U}}$ besteht. Ist die Matrix reell, so ist die Adjungierte gleich der Transponierten.
Daher gilt für die oben definierte Matrix $\underline{\underline{U}}$, dass

$$\underline{\underline{U}}^{-1} = \underline{\underline{U}}^\dagger \,.$$

Eine solche Matrix nennt man **unitär**.
Außerdem sieht man schnell ein, dass per Konstruktion von $\underline{\underline{U}}$ aus den Eigenvektoren von $\underline{\underline{A}}$ gilt:

$$\underline{\underline{A}} \cdot \underline{\underline{U}} = (\lambda_1 \underline{v}_1, \ldots, \lambda_n \underline{v}_n) = \underline{\underline{U}} \cdot \underline{\underline{\Lambda}}$$

mit der Diagonalmatrix

$$\underline{\underline{\Lambda}} := \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix} .$$

Somit gilt wegen der Unitarität von $\underline{\underline{U}}$

$$\underline{\underline{U}}^\dagger \cdot \underline{\underline{A}} \cdot \underline{\underline{U}} = \underline{\underline{\Lambda}} \qquad \text{bzw.} \qquad \underline{\underline{A}} = \underline{\underline{U}} \cdot \underline{\underline{\Lambda}} \cdot \underline{\underline{U}}^\dagger \,.$$

Man sagt auch: $\underline{\underline{U}}$ diagonalisiert $\underline{\underline{A}}$.
Komponentenweise lautet die letzte Gleichung

$$(\underline{\underline{A}})_{jm} = \sum_k \lambda_k (\underline{v}_k)_j (\overline{\underline{v}_k})_m \,.$$

Die Matrix $\underline{\underline{A}}$ läßt sich also mit Hilfe ihrer Eigenwerte und Eigenvektoren darstellen. Dies bezeichnet man als **Spektralzerlegung** oder **Spektraldarstellung**.
Speziell für $\underline{\underline{A}} = \mathbb{1}$ gilt dann

$$\sum_k (\underline{v}_k)_j (\overline{\underline{v}_k})_m = \delta_{jm} \,.$$

Dies bezeichnet man auch als **Zerlegung der Einheit** oder **Vollständigkeitsrelation**.

*Bemerkung.* Beachte, dass sich nicht alle Matrizen auf diese Art diagonalisieren lasssen! Für eine spezielle Klasse von Matrizen läßt sich dies aber allgemein beweisen.

**Definition 10.1.2** (Normale Matrizen)**.**
Gilt

$$\left[\underline{\underline{A}}, \underline{\underline{A}}^\dagger\right] := \underline{\underline{A}} \cdot \underline{\underline{A}}^\dagger - \underline{\underline{A}}^\dagger \cdot \underline{\underline{A}} = 0\,,$$

so heißt die Matrix $\underline{\underline{A}}$ **normal**.
Dabei ist $[A, B] = AB - BA$ der sogenannte **Kommutator**.

Insbesondere sind also hermitesche ($\underline{\underline{A}}^\dagger = \underline{\underline{A}}$) und symmetrische ($\underline{\underline{A}}$ reell mit $\underline{\underline{A}}^T = \underline{\underline{A}}$) Matrizen normal.

Es gilt: Normale Matrizen sind unitär diagonalisierbar (d.h. $\underline{\underline{U}}^\dagger \cdot \underline{\underline{A}} \cdot \underline{\underline{U}} = \underline{\underline{\Lambda}}$) und haben eine Spektraldarstellung.

*Bemerkung.* Die Diagonalisierung entspricht dem Übergang zu einem neuen Koordinatensystem, d.h. einem Basiswechsel. Es gilt

$$\underline{\underline{U}}^\dagger \cdot \underline{\underline{A}} \cdot \underline{\underline{U}} \left(\underline{\underline{U}}^\dagger \underline{v}\right) = \lambda \left(\underline{\underline{U}}^\dagger \underline{v}\right),$$

also mit $\underline{\underline{U}}^\dagger \cdot \underline{\underline{A}} \cdot \underline{\underline{U}} = \underline{\underline{\Lambda}}$ und der Abkürzung $\underline{\tilde{v}} := \underline{\underline{U}}^\dagger \underline{v}$

$$\underline{\underline{\Lambda}} \cdot \underline{\tilde{v}} = \lambda \underline{\tilde{v}}.$$

In der Basis der Eigenvektoren $\underline{\tilde{v}}$ hat $\underline{\underline{A}}$ also Diagonalform!
Da sich die Spur und die Determinante bei Basiswechseln nicht ändern, gilt

$$\begin{aligned} \operatorname{Spur} \underline{\underline{A}} &= \sum_{j=1}^{n} \lambda_j = \sum_j \lambda_j p_j\,, \\ \det \underline{\underline{A}} &= \prod_{j=1}^{n} \lambda_j = \prod_j \lambda_j^{p_j}\,. \end{aligned}$$

Dabei ist $p_j$ die algebraische Vielfachheit des Eigenwerts[3] $\lambda_j$. Die **Spur** einer $n \times n$-Matrix $\underline{\underline{A}}$ ist definiert durch

$$\operatorname{Spur} \underline{\underline{A}} := \sum_{j=1}^{n} A_{jj},$$

d.h. als Summe der Diagonalelemente.

*Bemerkung.* Die Eigenwerte der Matrix $\underline{\underline{A}}$ und ihrer Transponierten $\underline{\underline{A}}^T$ stimmen überein. Dies folgt aus der allgemeingültigen Identität $\det \underline{\underline{A}}^T = \det \underline{\underline{A}}$, die die Gleichheit der charakteristischen Polynome impliziert. I.a. stimmen aber die zugehörigen Eigen*vektoren* nicht überein! Dies führt zu folgender Definition:

[3] In der entsprechenden Darstellung ist nur noch über die *unterschiedlichen* Eigenwerte zu summieren bzw. zu multiplizieren!

**Definition 10.1.3** (Rechte und linke Eigenvektoren)**.**
Die bisher betrachteten Eigenvektoren $\underline{v}$ sind **rechte Eigenvektoren** für die gilt

$$\underline{\underline{A}} \cdot \underline{v} = \lambda \underline{v}.$$

In Analogie hierzu definiert man **linke Eigenvektoren** $\underline{\tilde{v}}$ als die Eigenvektoren von $\underline{\underline{A}}^T$:

$$\underline{\underline{A}}^T \underline{\tilde{v}} = \lambda \underline{\tilde{v}}\,.$$

Die Bezeichnung erklärt sich aus folgender Identität:

$$\underline{\tilde{v}}^T \underline{\underline{A}} = \left(\underline{\underline{A}}^T \cdot \underline{\tilde{v}}\right)^T = (\lambda \underline{\tilde{v}})^T = \lambda \underline{\tilde{v}}^T\,,$$

d.h. ein linker Eigenvektor wird bei Multiplikation von links an die Matrix $\underline{\underline{A}}$ bis auf einen Faktor reproduziert.

Für normale Matrizen, also insbesondere für hermitesche und symmetrische Matrizen, stimmen rechter und linker Eigenvektor überein.

### 10.1.1 Theoreme zum Eigenwertproblem

Wir geben im folgenden eine Sammlung wichtiger und nützlicher Aussagen zum Eigenwertproblem an, die wir aber nicht beweisen wollen.

**Satz 10.1.1** (Eigenwertprobleme)**.**

1. Eigenvektoren zu unterschiedlichen Eigenwerten sind linear unabhängig.
2. **Hermitesche Matrizen**, d.h. Matrizen für die $\underline{\underline{A}}^\dagger = \underline{\underline{A}}$ gilt, haben *reelle* Eigenwerte und die (normierten) Eigenvektoren bilden eine *Orthonormalbasis*. Sie sind mit Hilfe der unitären Matrix $\underline{\underline{U}}$ der Eigenvektoren diagonalisierbar und haben eine Spektraldarstellung.
3. Eine analoge Aussage gilt für **symmetrische Matrizen**, d.h. alle Elemente $A_{jl}$ sind reell und es gilt $\underline{\underline{A}}^T = \underline{\underline{A}}$. In diesem Fall kann $\underline{\underline{A}}$ mit Hilfe einer **orthogonalen Matrix** $\underline{\underline{U}}$ diagonalisiert werden, d.h. $\underline{\underline{U}}^{-1} = \underline{\underline{U}}^T$.
4. Kommutierende, hermitesche Matrizen $\underline{\underline{A}}$, $\underline{\underline{B}}$ (mit $[\underline{\underline{A}}, \underline{\underline{B}}] = 0$) haben ein gemeinsames Eigenvektorensystem $\underline{v}_1, \ldots, \underline{v}_n$, d.h. es gilt

   $$\underline{\underline{A}} \cdot \underline{v}_j = \lambda_j \underline{v}_j \qquad \text{und} \qquad \underline{\underline{B}} \cdot \underline{v}_j = \tilde{\lambda}_j \underline{v}_j.$$

   Man beachte, dass die Eigenwerte $\lambda_j$ und $\tilde{\lambda}_j$ aber i.a. verschieden sind!
5. Die (normierten) Eigenvektoren einer normalen Matrix bilden eine Orthonormalbasis. Sie sind wie hermitesche Matrizen unitär diagonalisierbar und haben eine Spektraldarstellung.
6. Alle normalen Matrizen sind diagonalisierbar, aber nicht alle diagonalisierbaren Matrizen sind normal.
7. Nicht alle diagonalisierbaren Matrizen haben eine Spektraldarstellung.

## 10.2 Operatoren

Wir betrachten zwei Vektorräume $V$ und $W$. Allgemein versteht man unter einem **Operator** $\hat{A}$ eine Abbildung[4] $\hat{A} : V \to W$. Im folgenden werden wir, sofern nicht anders angegeben oder aus dem Zusammmenhang offensichtlich, vor allem den Fall $V = W$ behandeln. Außerdem wollen wir vor allem **lineare Operatoren** betrachten, die durch

$$\hat{A}(x + \alpha y) = \hat{A}x + \alpha \hat{A}y \qquad \text{für alle } x, y \in V \text{ und } \alpha \in \mathbb{R} \text{ oder } \mathbb{C}$$

charakterisiert sind. Hierbei haben wir schon eine häufig verwendete Konvention benutzt, nämlich die, dass man bei Operatoren oft $\hat{A}x$ statt $\hat{A}(x)$ schreibt. Dies kennnen natürlich schon von Differentialoperatoren wie $\frac{d}{dx}$ oder $\mathrm{grad}$.
Im folgenden werden wir die Notation für den Fall eines komplexen Vektorraumes benutzen. Wenn nicht anders angegeben, gelten analoge Ergebnisse für den reellen Fall, man muß lediglich die komplexe Konjugation weglassen!
Bisher haben wir den Fall $V = \mathbb{R}^n$ betrachtet, wo wir $\hat{A}$ durch eine Matrix $\underline{\underline{A}}$ darstellen können. Die für Matrizen angestellten Überlegungen lassen sich dann übertragen, wenn wir ein Skalarprodukt $(x, y)$ (mit $x, y \in V$) zur Verfügung haben (siehe Kapitel 11.1) und folgende Identifikationen benutzen:

$$\underline{\overline{x}} \cdot \underline{y} \leftrightarrow (x, y), \qquad \underline{\overline{x}} \cdot \underline{\underline{A}} \cdot \underline{y} \leftrightarrow (x, \hat{A}y), \qquad \text{etc.}$$

Dies werden wir gleich weiter ausarbeiten. Zunächst wollen wir Beispiele für lineare Operatoren kennenlernen.

**Beispiel 10.2.1.**

1. Das Skalarprodukt mit einem festen $x_0 \in V$ definiert einen linearen Operator: $\hat{A}x := (x_0, x)$. In diesem Fall ist $W = \mathbb{R} \neq V$.
2. Die Ableitung $\hat{A}f = f'$ ist ein linearer Operator.

**Definition 10.2.1** (adjungierte und hermitesche Operatoren)**.**
Ist $\hat{A} : V \to V$ ein linearer Operator, so bezeichnet man den Operator $\hat{A}^\dagger : V \to V$ charakterisiert durch

$$(x, \hat{A}y) = (\hat{A}^\dagger x, y) \qquad \text{für alle } x, y \in V$$

als den **adjungierten Operator** oder auch **hermitesch konjugierten Operator**.
Ist $\hat{A}^\dagger = \hat{A}$, so nennt man $\hat{A}$ auch **hermitesch** oder **selbstadjungiert**[5].

Wir können nun Darstellungen des Operators $\hat{A}$ angegeben. Dazu sei $\{\varphi_k\}$ eine Orthonormalbasis von $V$. Beliebige Elemente $x \in V$ lassen sich daher als Linearkombinationen $x = \sum_k c_k \varphi_k$ darstellen. Wir können hiermit z.B. die Identität $\hat{A}x = y$ in eine Matrixgleichung übersetzen:

$$\sum_k \tilde{c}_k \varphi_k = y = \hat{A}x = \hat{A} \sum_k c_k \varphi_k = \sum_k c_k \hat{A} \varphi_k \,.$$

[4] Operatoren kennzeichnen wir oft durch ein $\hat{}$, z.B. $\hat{A}$.

[5] In der Mathematik gibt es einen winzigen Unterschied zwischen *hermitesch* und *selbstadjungiert*, der in der Physik aber in der Regel unwichtig ist.

Nun bilden wir das Skalarprodukt dieser Identität mit $\varphi_j$:

$$\tilde{c}_j = \sum_k \tilde{c}_k \delta_{jk} = \sum_k \tilde{c}_k (\varphi_j, \varphi_k) = \sum_k c_k (\varphi_j, \hat{A}\varphi_k) =: \sum_k c_k A_{jk}$$

wobei wir die Abkürzung $A_{jk} := (\varphi_j, \hat{A}\varphi_k)$ eingeführt haben. Somit läßt sich die Identität nach Wahl einer Basis auch in Matrixform darstellen als

$$\sum_k A_{jk} c_k = \tilde{c}_j \qquad \text{bzw.} \qquad \underline{\underline{A}} \cdot \underline{c} = \underline{\tilde{c}}\,,$$

wobei die Elemente $A_{jk}$ zu einer Matrix zusammengefasst und die $c_k$, $\tilde{c}_j$ als Komponenten eines (Spalten-)Vektors aufgefasst haben. Man beachte, dass Matrix und Vektoren im Prinzip unendlich sein können, wenn es sich bei $V$ nicht um einen endlichdimensionalen Vektorraum handelt.
Weiterhin gilt:

$$(x, \hat{A}y) \;=\; \left( \sum_j c_j \varphi_j, \hat{A} \sum_l c'_l \varphi_l \right) = \sum_{j,l} \overline{c_j} c'_l (\varphi_j, \hat{A}\varphi_l) = \sum_{j,l} \overline{c_j} A_{jl} c'_l\,.$$

Für $x = y$ nennt man dies auch **Erwartungswert** von $\hat{A}$.

Es sei nun $\{\psi_k\}$ eine andere Orthonormalbasis von $V$. Der Basis- oder Darstellungswechsel $\{\varphi_k\} \to \{\psi_k\}$ wird dann durch einen unitären Operator $\hat{U}$ vermittelt:

$$\psi_j = \hat{U}\varphi_j \qquad \text{mit} \qquad \hat{U}^\dagger \hat{U} = \hat{U}\hat{U}^\dagger = \mathbb{1}$$

mit der Matrixdarstellung

$$U_{jl} = (\varphi_j, \hat{U}\varphi_l) = (\varphi_j, \psi_l)\,.$$

Es gilt wegen $\hat{U}^\dagger \hat{U} = \mathbb{1}$ für alle $x, y \in V$:

$$(x, y) = (x, \hat{U}^\dagger \hat{U} y) = (\hat{U}x, \hat{U}y)\,.$$

Skalarprodukte bleiben daher bei einem solchen Basiswechsel unverändert (*invariant*), weshalb man ihn als Drehung im Vektorraum interpretieren kann.
In Komponenten gilt:

$$\begin{aligned}
(\psi_k, x) &= \sum_l (\psi_k, \varphi_l) = \sum_l \left(U^\dagger\right)_{kl} c_l = \sum_l \left(U^\dagger\right)_{kl} (\varphi_l, x)\,, \\
(\psi_j, \hat{A}\psi_l) &= (U^\dagger A U)_{jl} = \sum_{k,m} \left(U^\dagger\right)_{jk} (\varphi_k, \hat{A}\varphi_m) U_{ml}.
\end{aligned}$$

## 10.3 Eigenwertproblem für Operatoren

Analog zum Fall von Matrizen können wir auch bei Operatoren Eigenwerte und Eigenvektoren suchen. Das entsprechende Eigenwertproblem lautet dann

$$\hat{A}f = \lambda f \,.$$

Im folgenden werden wir uns vorstellen, dass der Vektorraum aus Funktionen besteht. Dies haben wir hier schon durch die Notation '$f$' statt '$x$' für den Eigenvektor angedeutet. Natürlich gelten aber alle folgenden Aussagen für beliebige zugrundeliegende Vektorräume.

**Beispiel 10.3.1.** Ein typisches Beispiel ist der Differentialoperator

$$\hat{A} = \frac{d^2}{dx^2} \,.$$

Dieser hat die Eigenvektoren $f_k := \sin(kx)$ mit den zugehörigen Eigenwerten $\lambda_k = -k^2$.

Im vorigen Kapitel haben wir gesehen, dass wir ein Operatorproblem in ein Matrixproblem übersetzen können. Daher ist es nicht überraschend, dass sich die Aussagen für das Eigenwertproblem von Matrizen auf Operatoren übertragen lassen. Insbesondere gilt:

1. Eigenvektoren zu unterschiedlichen Eigenwerten sind linear unabhängig.

2. Normale Operatoren ($[\hat{A}, \hat{A}^\dagger] = 0$, d.h. $\hat{A}\hat{A}^\dagger = \hat{A}^\dagger\hat{A}$) besitzen eine Orthonormalbasis $\{\varphi_j\}$ aus Eigenvektoren: $\hat{A}\varphi_j = \lambda_j\varphi_j$.

3. Selbstadjungierte Operatoren haben reelle Eigenwerte:

   $$\lambda_j = (\varphi_j, \lambda_j\varphi_j) = (\varphi_j, \hat{A}\varphi_j) = (\hat{A}\varphi_j, \varphi_j) = (\lambda_j\varphi_j, \varphi_j) = \overline{\lambda_j}\,(\varphi_j, \varphi_j) = \overline{\lambda_j} \,.$$

   Dabei haben wir, neben der Tatsache dass $\{\varphi_j\}$ eine Orthonormalbasis aus Eigenvektoren ist, im dritten Schritt die Selbstadjungiertheit von $\hat{A}$ ausgenutzt.

4. Unitäre Operatoren beschreiben Darstellungswechsel. Im Eigenvektorsystem wird der Operator diagonalisiert:

   $$(\varphi_j, \hat{A}\varphi_l) = \lambda_l(\varphi_j, \varphi_l) = \lambda_l\delta_{jl} \,.$$

   Der Operator transformiert sich daher wie

   $$\underbrace{(\hat{U}^\dagger\hat{A}\hat{U})}_{=\hat{\Lambda}}(\hat{U}^\dagger f) = \lambda(\hat{U}^\dagger f)$$

   wobei $\hat{\Lambda}$ diagonal ist.

5. Hermitesche, miteinander kommutierende Operatoren haben ein gemeinsames Eigenvektorsystem.
   Bem.: Dies wird in der Quantenmechanik wichtig, wo kommutierende Operatoren Größen entsprechen, die man gleichzeitig "scharf" messen kann.

Es sei $\hat{A}$ ein hermitescher Operator und $f = \sum_j c_j \varphi_j$ ein beliebiger Vektor, wobei $\{\varphi\}$ eine Orthonormalbasis aus Eigenvektoren von $\hat{A}$ ist. Dann ist der **Erwartungswert** von $\hat{A}$ (bzgl. $f$) gegeben durch

$$\begin{aligned}(f, \hat{A}f) &= \sum_{j,l} \overline{c_j} c_l (\varphi_j, \hat{A}\varphi_l) = \sum_{j,l} \overline{c_j} c_l \lambda_l (\varphi_j, \varphi_l) = \sum_{j,l} \overline{c_j} c_l \lambda_l \delta_{jl} \\ &= \sum_j \lambda_j |c_j|^2 .\end{aligned}$$

Es tragen also i.a. alle Eigenvektoren und Eigenwerte zum Erwartungswert bei!

## 10.4 Operatoren in der Quantenmechanik

Wir haben in Kapitel 11.1 schon den Begriff des Hilbertraumes kennengelernt. Hierbei handelt es sich um einen Vektorraum, auf dem ein Skalarprodukt definiert ist und der darüberhinaus vollständig ist.
In der mathematischen Beschreibung der Quantenmechanik wird ein *quantenmechanischer Zustand* mit einem Vektor eines geeigneten Hilbertraumes identifiziert. *Meßgrößen (Observablen)* entsprechen dann selbstadjungierten linearen Operatoren in diesem Hilbertraum. Die möglichen *Meßwerte* sind dann die Eigenwerte des entsprechenden Operators.

**Definition 10.4.1** (Dirac-Notation)**.**
Dirac hat eine sehr elegante und zweckmässige Notation, die sog. **Dirac-Notation**, erfunden, mit der sich besonders elegant rechnen läßt.
Vektoren im Hilbertraum schreibt man dabei als $|a\rangle$, den zu ihm hermitesch adjungierten Vektor, den man auch als **dualen Vektor** bezeichnet als $\langle a| := (|a\rangle)^\dagger$. Im $\mathbb{R}^n$ entspricht $|a\rangle$ einem Spaltenvektor und $\langle a|$ dem zugehörigen Zeilenvektor[6].
Man bezeichnet $\langle a|$ auch als **Bra(-Vektor)** und $|a\rangle$ als **Ket(-Vektor)**. Zusammengenommen ergibt dies (fast) das englische Wort *bracket* für Klammer.

In Dirac-Notation lautet das Eigenwertproblem für eine Operator $\hat{A}$

$$\hat{A}|a\rangle = a|a\rangle .$$

Hierbei haben wir schon die Standardkonvention benutzt, dass man den Eigenwert mit dem gleichen Buchstaben wie den Vektor bezeichnet. Auf Grund der Notation ist hier eine Verwechslung fast ausgeschlossen.
Das Skalarprodukt zweier Vektoren ist durch

$$\langle a|b\rangle$$

[6] Dies sollte man im folgenden immer im Hinterkopf behalten!

gegeben. Dies ist analog zum Standardskalarprodukt im $\mathbb{R}^n$. In Vektorschreibweise hatten wir dieses immer als $\underline{a} \cdot \underline{b}$ geschrieben. Es gilt aber

$$\underline{a} \cdot \underline{b} = \sum_{j=1}^{n} a_j b_j = \underline{a}^T \underline{b}$$

wobei wir die rechte Seite als Matrixmultiplikation der $1 \times n$-Matrix $\underline{a}^T$ mit der $n \times 1$-Matrix $\underline{b}$ zu interpretieren haben.
Mit dem Skalarprodukt können wir auch die Norm angeben:

$$||\underline{a}||^2 := \langle a|a\rangle \,.$$

Der Erwartungswert im Zustand $|a\rangle$ ist durch $\langle a|\hat{A}|a\rangle$ gegeben, falls $|a\rangle$ normiert ist.

Der Bra-Vektor $|a\rangle$ entspricht der *abstrakten* Darstellung des quantenmechanischen Zustandes. Wählt man eine geeignete Basis, so kommt man z.B. zu den sog. *Wellenfunktionen*.

Es sei nun wieder $\hat{A}$ ein selbstadjungierter Operator mit Eigensystem $\hat{A}|a_n\rangle = a_n|a_n\rangle$. Ein beliebiger Vektor (Zustand) läßt sich dann schreiben als

$$|f\rangle = \sum_n f_n|a_n\rangle \qquad \text{mit} \qquad f_n = \langle a_n|f\rangle \,.$$

Die Darstellung der Entwicklungskoeffizienten $f_n$ folgt wie früher, da $\{|a_n\rangle\}$ eine Orthonormalbasis ist.
Wir können jetzt die Spektraldarstellung von $\hat{A}$ in sehr eleganter Weise angeben:

$$\hat{A} = \sum_n a_n|a_n\rangle\langle a_n| \,.$$

Dies sieht man folgendermaßen. Für eine beliebiges $|f\rangle = \sum_n f_n|a_n\rangle$ gilt:

$$\begin{aligned}\left(\sum_n a_n|a_n\rangle\langle a_n|\right)|f\rangle &= \sum_n a_n|a_n\rangle \underbrace{\langle a_n|f\rangle}_{=f_n} = \sum_n f_n a_n|a_n\rangle = \sum_n f_n \hat{A}|a_n\rangle \\ &= \hat{A}|f\rangle \,.\end{aligned}$$

Diese Rechnung wesentlichen Prinzipien des Arbeitens mit der Dirac-Notation. Zunächst haben wir alle Bra's und Ket's gewissermaßen ausmultipliziert, unter Beachtung der Reihenfolge der Faktoren. Danach haben wir nach auftretenden Größen der Form $\langle b|a\rangle$ gesucht und diese als Skalarprodukte interpretiert (und berechnet).
Speziell für den Einheitsoperator $\mathbb{1}$ erhalten wir wieder die **Zerlegung der Einheit** oder **Vollständigkeitsrelation**:

$$\mathbb{1} = \sum_n |a_n\rangle\langle a_n| \,.$$

# Kapitel 11

# Fourierreihen und Integraltransformationen

## 11.1 Funktionenräume

Funktionen können oft auch als Vektoren, also Elemente eines Vektorraumes, interpretiert werden. Dies ist oft sehr nützlich, z.B. in der Quantenmechanik. Dabei sind Addition und Multiplikation mit einem Skalar wie üblich definiert:

$$h = f + \alpha g\,, \qquad \text{d.h. } h(x) = f(x) + \alpha g(x) \text{ für alle } x\,.$$

Ein Beispiel ist die Menge

$$C^m(X) := \{m - \text{fach stetig differenzierbaren Funktionen auf } X\}\ .$$

Die bekannten Definitionen und Konzepte wie 'Basis' etc. lassen sich übertragen. Oft sind Funktionenräume unendlich-dimensional. Außerdem gibt es weitere nützliche Strukturen, die wir im Folgenden vorstellen wollen.

**Definition 11.1.1** (normierter Raum, Norm)**.**
Eine Abbildung $||\ldots|| : V \to \mathbb{R}$ auf einem Vektorraum $V$ heißt **Norm**, wenn für alle $x, y, z \in V$ gilt:

1. $||x|| \geq 0$ (Nichtnegativität)
2. $||x|| = 0$ genau dann, wenn $x = \underline{0}$ (Eindeutigkeit)
3. $||\lambda x|| = |\lambda|\ ||x||$ mit $\lambda \in \mathbb{R}$ (Skalierung)
4. $||x + y|| \leq ||x|| + ||y||$ (Dreiecksungleichung)

In diesem Fall bezeichnet man $V$ auch als **normierten Raum**.

**Beispiel 11.1.1.** Ein wichtiges Beispiel, das z.B. in der Quantenmechanik eine zentrale Rolle spielt, ist der Raum der quadratisch-integrablen Funktionen auf $X \subset \mathbb{R}$,

$$L^2(\mathbb{R}) := \left\{ f : X \to \mathbb{C} \middle| \int_X |f(x)|^2 \, dx < \infty \right\}$$

mit der Norm

$$||f||_2 := \sqrt{\int_X |f(x)|^2 \, dx} \, .$$

**Definition 11.1.2** (Skalarprodukt)**.**
Es sei $X$ ein Vektorraum über $\mathbb{C}$. Ein **Skalarprodukt** $(\ldots,\ldots)$ ist eine Abbildung $X \times X \to \mathbb{C}$ mit folgenden Eigenschaften (für alle $x, y, z \in X$ und $\lambda \in \mathbb{C}$):

1. $(x, y) = \overline{(y, x)}$,
2. $(\lambda x, y) = \bar{\lambda}(x, y)$,
   $(x, \lambda y) = \lambda(x, y)$,
3. $(x + y, z) = (x, z) + (y, z)$,
4. $(x, x) \geq 0$,
5. $(x, x) = 0 \;\; \Rightarrow \;\; x = 0$.

Dabei bezeichnet $\bar{a}$ das komplex-Konjugierte von $a$.

Eine unmittelbare Folgerung aus Eigenschaft 1. ist, dass $(x, x) \in \mathbb{R}$ ist.
Eigenschaft 2. bedeutet, dass das Skalarprodukt antilinear im ersten Argument ist und linear im zweiten. Dies ist eine Konvention, die in der Physik häufig benutzt wird.
Aus Eigenschaft 3. folgt wegen 1., dass auch $(x, y + z) = (x, y) + (x, z)$ gilt.

Weitere Folgerungen aus den Eigenschaften sind die

- **Cauchy-Schwarz-Ungleichung**

$$|(x, y)|^2 \leq (x, x)(y, y) \, ,$$

- **Minkowski-Ungleichung**

$$\sqrt{(x + y, x + y)} \leq \sqrt{(x, x)} + \sqrt{(y, y)} \, .$$

**Definition 11.1.3** (Hilbertraum)**.**
Ein Vektorraum mit Skalarprodukt heisst auch **Prähilbertraum**. Ist der Raum vollständig, d.h. jede Cauchy-Folge konvergiert gegen ein Element des Raumes, so heißt er **Hilbertraum**.

*Bemerkung.*

1. Durch $||f|| := \sqrt{(f,f)}$ ist eine Norm definiert.
2. Die Definitionen gelten analog für reelle Räume, wobei man überall die komplexe Konjugation weglässt. Für die Quantenmechanik sind aber komplexe Räume wichtig.

**Beispiel 11.1.2.**
Wir können in $L^2(X)$ durch

$$(f,g) := \int_X \bar{f}(x)g(x)dx$$

ein Skalarprodukt definieren.
Eine Verallgemeinerung hiervon ist

$$(f,g) := \int_X w(x)\bar{f}(x)g(x)dx$$

mit einer nicht-negativen Gewichtsfunktion $w(x)$ (z.B. $w(x) = e^{-x^2}$).

## 11.2 Fourierreihen

Im diesem Abschnitt wollen wir untersuchen, inwieweit sich periodische Funktionen nach "Teilfrequenzen" zerlegen lassen. Dieses Verfahren ist auch unter dem Namen **Fourier-Analyse** bekannt.
Zunächst wollen wir noch einmal an die Definition von periodischen Funktionen erinnern (siehe Kap. 3.2.4):

**Definition 11.2.1** (Periodische Funktionen)**.**
Die $f(t)$ ist heißt **periodisch mit der Periode** $T$ ($T > 0$), wenn für alle $t$ gilt:

$$f(t+T) = f(t).$$

Der kleinste Wert von $T > 0$, der dies erfüllt, heißt **kleinste Periode** oder einfach nur **die Periode** von $f(t)$.

Prototypen periodischer Funktionen sind die komplexe Exponentialfunktion $e^{i\omega t}$ und die trigonometrischen Funktionen $\sin\omega t$ und $\cos\omega t$, die jeweils die Periode $T = \frac{2\pi}{\omega}$ haben.
Im folgenden werden wir zur Vereinfachung $x := \omega t$ als Variable wählen, d.h. wir betrachten nur die normierte Periode $T = 2\pi$ der Funktionen $\sin x$, $\cos x$. Am Ende des Abschnittes werden wir dann der Vollständigkeit halber die allgemeinen Formeln für beliebige Perioden $T$ angeben.

**Definition 11.2.2** (Fourierreihen)**.**
Eine Reihe der Form

$$\frac{a_0}{2} + \sum_{n=1}^{\infty} [a_n \cos nx + b_n \sin nx]$$

bezeichnet man als **Fourierreihe**. Man beachte, dass diese Reihe $2\pi$-periodisch ist.

Im folgenden wollen wir die Frage untersuchen, wann sich eine periodische Funktion als Fourierreihe darstellen läßt.

**Satz 11.2.1** (Fourierreihendarstellung periodischer Funktionen)**.**
Die Funktion $f(x)$ sei im Intervall $[-\pi, \pi]$ definiert. Weiterhin sollen die sog. **Dirichlet-Bedingungen** erfüllt sein:

a) $f(x)$ ist $2\pi$-periodisch, d.h. $f(x + 2\pi) = f(x)$

b) $f(x)$ und $f'(x)$ stückweise stetig in $[-\pi, \pi]$

Durch die Bedingung a) ist $f(x)$ für alle reellen Zahlen $x$ eindeutig definiert. Bedingung b) bedeutet, dass $f(x)$ und $f'(x)$ im Intervall $[-\pi, \pi]$ bis auf endlich viele Punkte stetig sind.

Sind die Dirichlet-Bedingungen erfüllt, dann gilt: $f(x)$ ist darstellbar als **Fourierreihe**, d.h.

$$\boxed{f(x) = \frac{a_0}{2} + \sum_{n=1}^{\infty} [a_n \cos nx + b_n \sin nx]}$$

mit den Koeffizienten (**Fourier-Koeffizienten**)

$$\boxed{\begin{aligned} a_n &= \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \cos nx \, dx \qquad (n = 0, 1, \dots), \\ b_n &= \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \sin nx \, dx \qquad (n = 1, 2, \dots). \end{aligned}}$$

Das bedeutet präziser: Die Fourierreihe konvergiert gegen den Wert

$$\frac{a_0}{2} + \sum_{n=1}^{\infty} [a_n \cos nx + b_n \sin nx] \quad \rightarrow \quad \begin{cases} f(x), & \text{wenn } f \text{ stetig bei } x \text{ ist} \\ \frac{f(x+0)+f(x-0)}{2}, & \text{wenn } f \text{ unstetig bei } x \text{ ist.} \end{cases}$$

Dabei bezeichnen $f(x+0)$ und $f(x-0)$ den rechts- bzw. linksseitigen Grenzwert (siehe Abb. 11.2.1).

Bemerkungen:

1. Die Dirichlet-Bedingungen sind hinreichend, aber nicht notwendig. Für die physikalische Praxis reicht dies i.a. aus.

2. Als Periodenintervall kann jedes Intervall $[x_0, x_0 + 2\pi]$ gewählt werden, z. B. auch $[0, 2\pi]$.

3. Das Intervall $[-\pi, \pi]$ ist aber zweckmäßig, da symmetrisch um $0$. Denn man sieht sofort:

$$\begin{aligned} a_n = 0, &\qquad \text{wenn } f(x) = -f(-x) \quad \text{(ungerade Funktion)} \\ b_n = 0, &\qquad \text{wenn } f(x) = f(-x) \quad \text{(gerade Funktion).} \end{aligned}$$

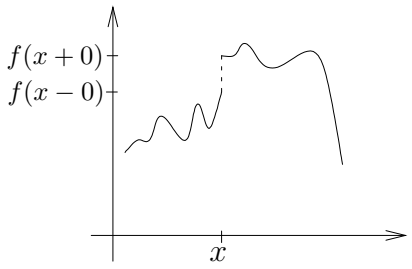



Abbildung 11.2.1: Unstetige Funktion

Das Interessante an obigem Satz ist ist, dass sich *jede* periodische Funktion (mit Dirichlet-Bedingungen) durch eine Reihe mit $\cos nx$ und $\sin nx$ darstellen läßt (Beweis: $\hookrightarrow$ Mathematik).

Eine andere Interpretation ist die über den Funktionenraum

$$\left\{ \frac{1}{\sqrt{2}}, \sin nx, \cos nx \middle| n = 1, 2, \ldots \right\}.$$

Dieser bildet ein Orthonormalsystem für periodische Funktionen in $L^2([-\pi, \pi])$ mit dem Skalarprodukt

$$(f, g) := \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) g(x) dx$$

Wir nehmen nun an, eine periodische Funktion lässt sich als Fourierreihe mit noch unbekannten Koeffizienten $a_n$, $b_n$ darstellen. Diese können wir dann folgendermassen berechnen:

$$\begin{aligned}
(f, \cos mx) &= \left( \frac{a_0}{2} + \sum_{n=1}^{\infty} \left[ a_n \cos nx + b_n \sin nx \right], \cos mx \right) \\
&= \left( \frac{a_0}{2}, \cos mx \right) + \sum_{n=1}^{\infty} a_n (\cos nx, \cos mx) + \sum_{n=1}^{\infty} b_n (\sin nx, \cos mx) \\
&= a_m.
\end{aligned}$$

Im letzten Schritt haben wir ausgenutzt, dass die Funktionen orthonormiert sind, d.h. insbesondere ist $(\sin nx, \cos mx) = 0$ und $(\cos nx, \cos mx) = \delta_{n,m}$. Wenn man nun das Skalarprodukt $(f, \cos mx)$ explizit als Integral ausschreibt, erhält man die oben angegebene Darstellung der Koeffizienten. Analog bestimmt man die anderen Koeffizienten zu $b_m = (f, \sin mx)$.
Den Beweis, dass sich die periodische Funktion $f$ tatsächlich in der angegebenen Reihenform schreiben lässt, überlassen wir der Mathematik. Hier haben wir gezeigt, wie sich die Koeffizienten bestimmen lassen, wenn dies der Fall ist.
Die Tatsache, dass sich beliebige periodische Funktionen durch die Elemente des oben angegebenen Orthonormalsystems darstellen lassen, gibt Anlass zu folgender Definition:

**Definition 11.2.3** (vollständiges Funktionensystem)**.**
Da sich beliebige $2\pi$-periodische Funktionen als Reihe darstellen lassen, bezeichnet man die

Menge $\{\cos nx, \sin nx \ (n = 0, 1, \dots)\}$ als **vollständig** (zur Darstellung periodischer Funktionen).

Man kann auch die Taylorentwicklung analytischer Funktionen in diesem Sinne verstehen. Statt nach den (vollständigen) periodischen Funktionen $\cos nx$ und $\sin nx$ zu entwickeln, entwickelt man hier nach den Funktion $x^n$.

Vielfach ist eine *komplexe Darstellung* der Fourierreihe zweckmäßig und einfacher. Mit den bekannten Beziehungen

$$\begin{aligned} e^{\pm inx} &= \cos nx \pm i \sin nx, \\ \cos nx &= \frac{1}{2}(e^{inx} + e^{-inx}), \qquad \sin nx = \frac{1}{2i}(e^{inx} - e^{-inx}), \end{aligned}$$

sieht man sofort ein, dass auch $\{e^{inx} | n = -\infty, \dots, \infty\}$ vollständig ist. Es gilt:

$$\boxed{f(x) = \sum_{n=-\infty}^{\infty} c_n e^{inx}}$$

mit

$$\boxed{c_n = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(x) e^{-inx} dx.}$$

Die reelle und die komplexe Darstellung lassen sich leicht ineinander umrechnen:

$$\begin{aligned} f(x) &= \sum_{n=-\infty}^{\infty} c_n e^{inx} = c_0 + \sum_{n=1}^{\infty} [c_n e^{inx} + c_{-n} e^{-inx}] \\ &= c_0 + \sum_{n=1}^{\infty} [c_n(\cos nx + i \sin nx) + c_{-n}(\cos nx - i \sin nx)] \\ &= c_0 + \sum_{n=1}^{\infty} [(c_n + c_{-n}) \cos nx + i(c_n - c_{-n}) \sin nx)] \end{aligned}$$

woraus folgt

$$c_0 = \frac{1}{2} a_0, \qquad a_n = c_n + c_{-n}, \qquad b_n = i(c_n - c_{-n}).$$

Die Funktionen $e^{inx}$ bilden ein Orthonormalsystem bzgl. des Skalarproduktes

$$(f, g) := \frac{1}{2\pi} \int_{-\pi}^{\pi} \bar{f}(x) g(x) dx,$$

denn es gilt:

$$n \neq m: \qquad \left(e^{imx}, e^{inx}\right) = \frac{1}{2\pi} \frac{e^{i(n-m)x}}{i(n-m)}\Big|_{-\pi}^{\pi} = \frac{e^{i(n-m)\pi} - e^{-i(n-m)\pi}}{i(n-m)} = 0$$

$$n = m: \qquad \left(e^{inx}, e^{inx}\right) = \frac{1}{2}\int_{-\pi}^{\pi} dx = 1.$$

Dieses Ergebnis läßt sich mit Hilfe des Kronecker-Symbols kompakt zusammenfassen:

$$\left(e^{imx}, e^{inx}\right) = \delta_{n,m}.$$

Wir wollen nun noch einmal die Fourierdarstellung beweisen, diesmal in der komplexen Variante. Wieder nehmen wir an, dass sich die Funktion $f(x)$ als eine Fourierreihe schreiben läßt. Wir zeigen jetzt, das die Entwicklungskoeffizienten gerade die in dem obigen Satz behauptete Form haben. Dabei verwenden wir wieder die Orthonormalität der Funktionen $e^{inx}$:

$$\frac{1}{2}\int_{-\pi}^{\pi} dx f(x) e^{-imx} = \left(e^{imx}, f\right) = \left(e^{imx}, \sum_{n=-\infty}^{\infty} c_n e^{inx}\right) = \sum_{n=-\infty}^{\infty} c_n \left(e^{imx}, e^{inx}\right) = c_m.$$

**Beispiel 11.2.1.** Wir betrachten die Funktion $f(x) = x$ in $[0, 2\pi]$, die wir uns $2\pi$-periodisch auf ganz $\mathbb{R}$ fortgesetzt denken (siehe Abb. 11.2.2)[1]. Der Graph von $f(x)$ ähnelt dann einem Sägezahn.

$$n \neq 0: \quad c_n = \frac{1}{2\pi}\int_0^{2\pi} x e^{-inx} dx = \frac{x e^{-inx}}{-2\pi i n}\Big|_0^{2\pi} + \frac{1}{2\pi i n} \underbrace{\int_0^{2\pi} e^{-inx} dx}_{=0} = \frac{1}{-in} = \frac{i}{n}$$

$$n = 0: \quad c_0 = \frac{1}{2\pi}\int_0^{2\pi} x dx = \frac{x^2}{4\pi}\Big|_0^{2\pi} = \pi$$

Also haben wir folgende Darstellung von $f$ als Fourierreihe:

$$x\Big|_{2\pi\text{-periodisch}} = \pi + \sum_{n\neq 0} \frac{i}{n} e^{inx} = \pi - \sum_{n=1}^{\infty} \frac{2}{n} \sin(nx),$$

wobei wir bei der letzten Umformung die Terme zu $n$ und $-n$ zusammengefaßt haben. Man beachte, dass die Reihe an den Unstetigkeitspunkten $x = 0 \mod 2\pi$ (d. h. $x$ ist ein Vielfaches von $2\pi$) tatsächlich den Wert $\frac{f(x+0)+f(x-0)}{2} = \frac{0+2\pi}{2} = \pi$ interpoliert!

Bemerkung:
Verwendet man statt der vollen Fourierreihe nur endlich viele Funktionen $e^{inx}$, so liefert dies eine Approximation der periodischen Funktion. Diese wird umso schlechter, je schneller sich die Funktion bei $x$ ändert. Insbesondere an Unstetigkeiten tritt das sogenannte Gibbs-Phänomen auf. Ein Beispiel ist in Abb. 11.2.3 gezeigt.
Zum Abschluss wollen wir noch die Fourierdarstellung einer beliebigen $T-$periodischen Funktion ($T > 0$) angeben.

[1] Natürlich ist $f$ auch periodisch mit Periodizitätsintervall $[-\pi, \pi]$. Die folgenden Rechnungen sind aber einfacher in $[0, 2\pi]$.

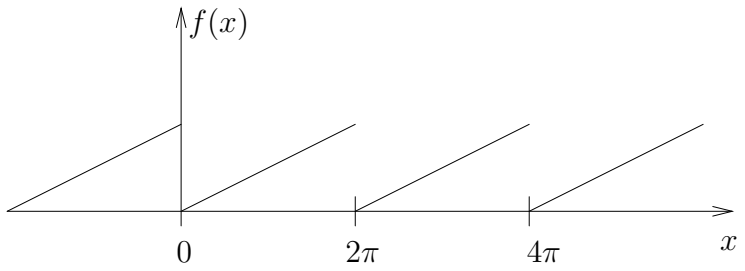



Abbildung 11.2.2: Sägezahn

**Satz 11.2.2** (Fourierreihendarstellung $T-$periodischer Funktionen)**.**
Die Funktion $f(x)$ sei $T-$periodisch und erfülle die Dirichlet-Bedingungen im Intervall $\left[-\frac{T}{2}, \frac{T}{2}\right]$. Dann hat $f(x)$ folgende Fourierdarstellung:

$$\boxed{\begin{aligned} f(x) &= \frac{a_0}{2} + \sum_{n=1}^{\infty} \left[ a_n \cos \frac{2\pi n x}{T} + b_n \sin \frac{2\pi n x}{T} \right] \\ &= \sum_{n=-\infty}^{\infty} c_n e^{i\frac{2\pi n x}{T}} \end{aligned}}$$

mit den Fourier-Koeffizienten

$$\boxed{\begin{aligned} a_n &= \frac{2}{T} \int_{-T/2}^{T/2} f(x) \cos\left(\frac{2\pi n x}{T}\right) dx \qquad (n = 0, 1, \dots), \\ b_n &= \frac{2}{T} \int_{-T/2}^{T/2} f(x) \sin\left(\frac{2\pi n x}{T}\right) dx \qquad (n = 1, 2, \dots), \\ c_n &= \frac{1}{T} \int_{-T/2}^{T/2} f(x) e^{-i\frac{2\pi n x}{T}} dx \qquad (n = -\infty, \dots, \infty). \end{aligned}}$$

Wie man sieht, geht dies für den Spezialfall $T = 2\pi$ in die bekannten Formeln über.

## 11.3 Fourier-Transformation

Im vorigen Abschnitt haben wir gesehen, dass sich periodische Funktionen als Fourierreihen darstellen lassen. Was passiert aber für Funktionen, die gar nicht periodisch sind ? Dieser Fall entspricht gerade dem Grenzübergang $T \to \infty$.

Wir betrachten zunächst die Funktion $\cos n\omega t = \cos(\frac{2\pi n}{T} t)$, die periodisch mit Periode $T$ ist. Wenn $T \to \infty$ liegen die Punkte $\omega_n = \frac{2\pi n}{T}$ immer dichter. Der Limes $T \to \infty$ entspricht dann

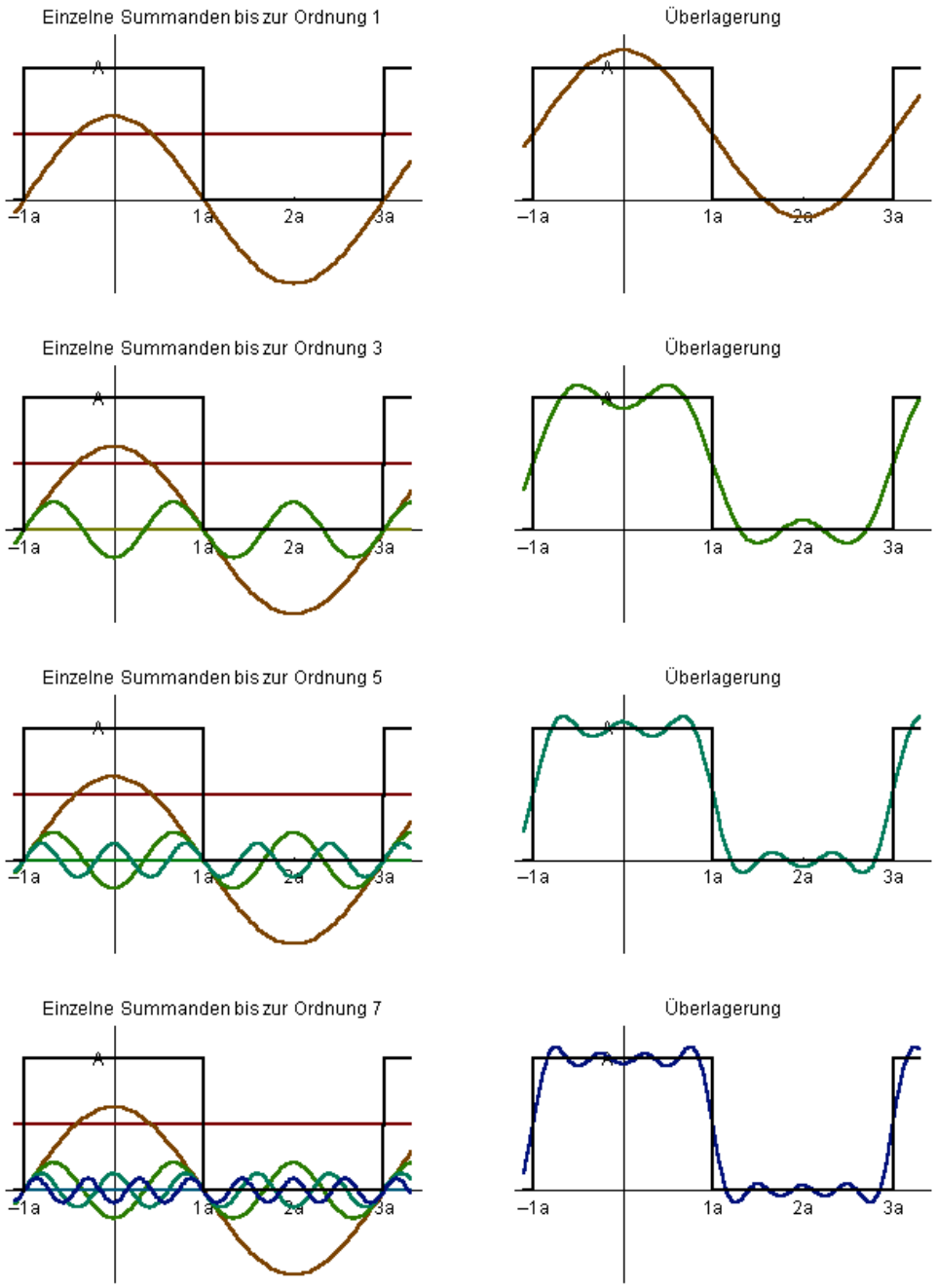



Abbildung 11.2.3: Die Abbildung zeigt die ersten Beiträge zur Fourierreihe der Kastenfunktion, die auf dem Intervall $[-a, a]$ den Wert $A$ hat und auf $[a, 3a]$ den Wert 0, also die Periode $4a$. In der linken Spalte sind jeweils die bis zur Ordnung $n = 1, 3, 5, 7$ beitragenden Funktionen (mit ihren Amplituden) dargestellt und in der rechten Spalte ihre Überlagerung, also die Fourierreihe bis zur Ordnung $n$. Man sieht, dass die Reihe relativ schnell konvergiert und bereits für $n = 7$ eine gute Übereinstimmung besteht. In der Nähe der Unstetigkeiten, also bei $x = \pm a$ erkennt man das Gibbs-Phänomen.

dem Übergang von einer Fourier-Summe zu einem Fourier-Integral. Dies wollen wir im folgenden genauer betrachten.

**Satz 11.3.1** (Fourier-Transformation)**.**
Die Funktion $f(x)$ erfülle folgende Voraussetzungen:

1) $f(x)$ genügt den Dirichlet-Bedingungen im Intervall $]-\infty, \infty[$,

2) $\int_{-\infty}^{\infty} |f(x)|^2 dx < \infty$.

Unter diesen Voraussetzungen gilt:

$$\boxed{f(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} F(\omega) e^{i\omega x} d\omega}$$

mit der sog. **Fourier-Transformierten**

$$\boxed{F(\omega) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f(x) e^{-i\omega x} dx\,.}$$

Man beachte, dass die oben genannten Voraussetzungen hinreichend, aber nicht notwendig sind. In der Praxis sind sie aber meist erfüllt.

*Bemerkung.* $F$ ist Fourier-Transformierte von $f$ und $f$ ist Fouriertransformierte von $F$ !

*Bemerkung.* Wir haben hier eine symmetrische Aufteilung der Koeffizienten (d. h. $\frac{1}{\sqrt{2\pi}}$) vor dem Integral gewählt. Häufig werden andere Konventionen verwendet, d. h.

$$\begin{aligned} f(x) &= \int_{-\infty}^{\infty} d\omega \ldots \quad \text{und} \quad F(\omega) = \frac{1}{2\pi} \int_{-\infty}^{\infty} dx \ldots \\ \text{oder} \quad f(x) &= \frac{1}{2\pi} \int_{-\infty}^{\infty} d\omega \ldots \quad \text{und} \quad F(\omega) = \int_{-\infty}^{\infty} dx \ldots \end{aligned}$$

Die entsprechenden Fourier-Transformierten unterscheiden sich um einen Faktor ($\frac{1}{\sqrt{2\pi}}$ bzw. $\sqrt{2\pi}$). Dies muß man z. B. beim Nachschlagen in Tabellen immer beachten.
Ebenso, ob die Fourier-Transformierte als Integral über $e^{-i\omega x}$ oder $e^{+i\omega x}$ definiert ist. In letzterem Fall ändert sich natürlich auch das entsprechende Vorzeichen in der Fourier-Darstellung von $f(x)$ !

Im folgenden zeigen wir, wie sich die Fourier-Transformation aus dem Grenzfall $T \to \infty$ der Fourier-Reihe

$$\begin{aligned} f(x) &= \sum_{n=-\infty}^{\infty} c_n e^{i\omega_n x} \qquad \text{mit} \quad \omega_n = \frac{2\pi n}{T}, \\ c_n &= \frac{1}{T} \int_{-\frac{T}{2}}^{\frac{T}{2}} f(x) e^{-i\omega_n x} dx \end{aligned}$$

ergibt, wobei $f(x) = f(x+T)$.
Wir machen nun den Grenzübergang $T \to \infty$ und $n \to \infty$ so, dass $\omega_n = \omega$ konstant gehalten wird. Dann ist auch $\Delta\omega_n = \omega_{n+1} - \omega_n = \frac{2\pi}{T} =: \Delta\omega$ konstant und beim Grenzübergang geht $\Delta\omega \to d\omega$. Für die Fourierreihe folgt:

$$\begin{aligned} f(x) &= \sum_{n=-\infty}^{\infty} c_n \frac{T}{2\pi} e^{i\omega_n x} \Delta\omega = \frac{1}{\sqrt{2\pi}} \sum_{n=-\infty}^{\infty} F(\omega_n) e^{i\omega_n x} \Delta\omega \\ &\to \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} F(\omega) e^{i\omega x} d\omega \end{aligned}$$

wobei wir definiert haben

$$F(\omega_n) := \frac{T}{\sqrt{2\pi}} c_n = \frac{1}{\sqrt{2\pi}} \int_{-\frac{T}{2}}^{\frac{T}{2}} f(x) e^{-i\omega_n x} dx \quad \overset{t,n\to\infty}{\longrightarrow} \quad F(\omega).$$

Damit haben wir gezeigt, dass sich die Fourier-Transformierte als Grenzfall der Fourierreihe einer Funktion mit unendlicher Periode ergibt.

*Bemerkung.* In der Physik interessieren uns häufig nur reelle Funktionen $f(x)$.
Dann gilt: $F(\omega)^* = F(-\omega)$
Beweis:

$$\begin{aligned} F(\omega)^* &= \left( \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f(x) e^{i\omega x} \right)^* = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \underbrace{f^*(x)}_{f(x)} \underbrace{\left(e^{-i\omega x}\right)^*}_{e^{i\omega x}} dx \\ &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f(x) e^{i\omega x} dx = F(-\omega) \end{aligned}$$

Weitere solcher nützlicher Identitäten werden wir in den Übungen kennenlernen.
Als Beispiel wollen wir die Fouriertransformierte der Blockfunktion (bzw. der charakterischen Funktion des Intervalls $[-a, a]$) (siehe Abb. 11.3.1)

$$f(x) = \begin{cases} 1 & |x| < a \\ 0 & |x| > a \end{cases}$$

berechnen:

$$\begin{aligned} F(\omega) &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f(x) e^{i\omega x} dx = \frac{1}{\sqrt{2\pi}} \int_{-a}^{a} e^{i\omega x} dx \\ &= \frac{1}{\sqrt{2\pi}} \int_{0}^{a} \cos(\omega x) dx \\ &= \sqrt{\frac{2}{\pi}} \frac{\sin(\omega x)}{\omega} \Big|_0^a = \sqrt{\frac{2}{\pi}} \frac{\sin(\omega a)}{\omega}. \end{aligned}$$

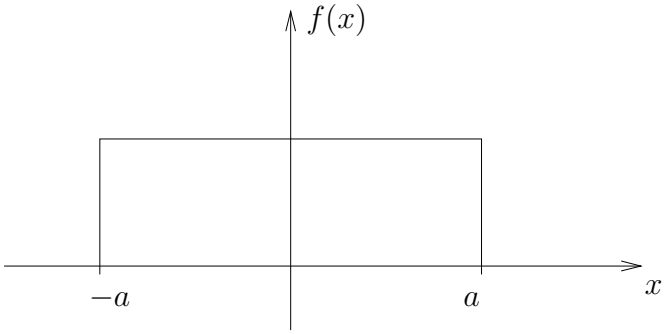



Abbildung 11.3.1: Blockfunktion

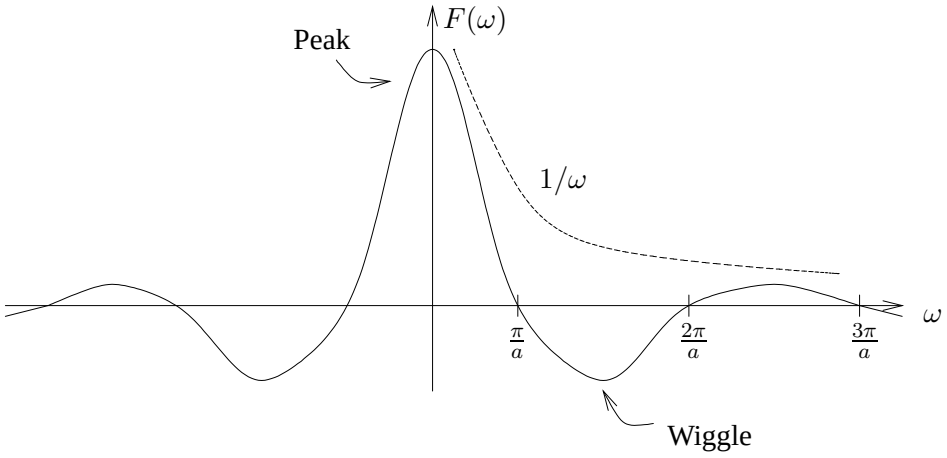



Abbildung 11.3.2: Fouriertransformierte der Blockfunktion

Beim Übergang zur zweiten Zeile haben wir ausgenutzt, dass das Integral über den Imaginärteil $\sin(\omega x)$ verschwindet, da der Sinus ungerade ist. Abb. 11.3.2 zeigt die Form der Fouriertransformierten. Das Maximum bei $x = 0$ bezeichnet man manchmal auch mit dem englischen Ausdruck **Peak** und die kleineren Maxima im oszillierenden Teil als **Wiggle**.
Speziell für $x = 0$ haben wir:

$$\begin{aligned} 1 &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} F(\omega) d\omega = \frac{1}{\pi} \int_{-\infty}^{\infty} \frac{\sin(\omega a)}{\omega} d\omega = \frac{2}{\pi} \int_{0}^{\infty} \frac{\sin(\omega a)}{\omega} d\omega \\ &\overset{u=\omega a}{=} \frac{2}{\pi} \int_{0}^{\infty} \frac{\sin u}{u} du \end{aligned}$$

und somit

$$\int_{0}^{\infty} \frac{\sin u}{u} du = \frac{\pi}{2}.$$

Ähnlich wie die Fourierreihe bei der Berechnung komplizierter Reihen hilfreich sein kann, hilft die Fouriertransformation manchmal bei der Berechnung von Integralen.

### 11.3.1 Delta-Funktion

Wir haben bereits häufig das Kronecker-Symbol

$$\delta_{jl} = \begin{cases} 1 & \text{falls } j = l \\ 0 & \text{falls } j \neq l \end{cases}$$

gewinnbringend verwendet. Eine ähnliche Größe kann man im kontinuierlichen Fall definieren.

**Definition 11.3.1** (Delta-Funktion)**.**
Wir definieren die sog. **Delta-Funktion** durch folgende Eigenschaften:

$$\begin{aligned} \delta(x - x_0) &= 0 \qquad \text{für alle } x \neq x_0, \\ \text{und} \qquad \int_{-\infty}^{\infty} f(x)\delta(x - x_0)dx &= f(x_0) \end{aligned}$$

wobei die zweite Eigenschaft für alle Funktionen $f(x)$ gelten soll, die stetig bei $x_0$ sind.

Bem.: Für analytische Funktionen $f(x)$ kann die zweite Bedingung abgeschwächt werden zu $\int_{-\infty}^{\infty} \delta(x - x_0)dx = 1$.

Dies ist eine heuristische Definition!! Eigentlich ist $\delta$ keine Funktion, sondern eine sogenannte **Distribution** ($\rightarrow$ Mathematik): Der Funktion $f$ wird ihr Funktionswert $f(x_0)$ an der Stelle $x_0$ zugeordnet. Als Funktion ist $\delta(x - x_0)$ hochgradig singulär, da sie nur an einem Punkt $x_0$ von Null verschieden ist, aber trotzdem ein endliches Integral besitzt. $\delta(x - x_0)$ kann deshalb in $x_0$ keinen endlichen Wert annehmen, sondern wird dort unendlich. In Abb. 11.3.3 ist angedeutet, wie man sich $\delta(x - x_0)$ vorzustellen hat. Man kann die Delta-Funktion als das kontinierliche Analogon des Kronecker-Deltas ansehen, insbesondere wenn man beachtet das $\delta_{a,b} = \delta_{a-b,0}$.

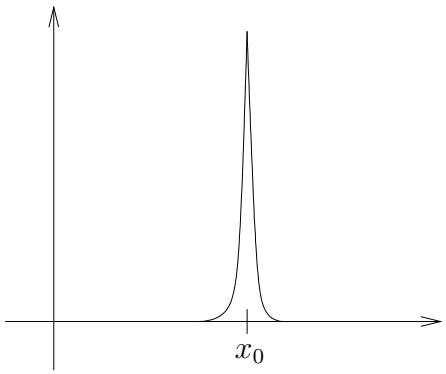



Abbildung 11.3.3: „Graph“ von $\delta(x - x_0)$

Die Delta-Funktion spielt eine wichtige Rolle in der Physik, denn sie hat viele Anwendungen. Ein Beispiel hierfür liefern die Punktmassen, von denen wir sehr oft gesprochen haben.
Für die Massendichte $\rho(x)$ einer Punktmasse $m$ am Ort $x_0$ gilt:

$$\begin{aligned} \rho(x) &= 0 \qquad \text{für} \quad x \neq x_0, \\ \text{aber} \qquad m &= \int_{-\infty}^{\infty} \rho(x)dx. \end{aligned}$$

Somit sieht man, dass die Massendichte einer Punktmasse durch

$$\rho = m\delta(x - x_0)$$

gegeben ist. Mit Hilfe der Deltafunktion kann man also im Prinzip diskrete und kontinuierliche Systeme in einheitlicher Weise beschreiben.

Das Fourier-Integral führt auf eine „Darstellung“ der $\delta$-Funktion. Setzt man nämlich die Fouriertransformierte explizit in die Fourierdarstellung ein, so folgt:

$$\begin{aligned} f(x) &= \int_{-\infty}^{\infty} \frac{d\omega}{\sqrt{2\pi}} e^{i\omega x} F(\omega) = \int_{-\infty}^{\infty} \frac{d\omega}{\sqrt{2\pi}} e^{i\omega x} \int_{-\infty}^{\infty} \frac{dy}{\sqrt{2\pi}} e^{-i\omega y} f(y) \\ &= \int_{-\infty}^{\infty} f(y) \left[ \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{i\omega(x-y)} d\omega \right] dy \stackrel{!}{=} f(x). \end{aligned}$$

Hieraus liest man das wichtige Ergebnis

$$\boxed{\frac{1}{2\pi} \int_{-\infty}^{\infty} e^{i\omega(x-y)} d\omega = \delta(x - y)}$$

ab. Dies kann man als Verallgemeinerung von $\frac{1}{2\pi} \int_{-\pi}^{\pi} e^{i(n-m)x} dx = \delta_{n,m}$ ansehen.

Zum Abschluss wollen wir noch eine weitere Darstellung von $\delta(x)$ angeben. Die Funktionenfolge $(\delta_\alpha)_\alpha$ mit

$$\delta_\alpha(x) = \frac{1}{2\pi}\int_{-\infty}^{\infty} d\omega e^{i\omega x} e^{-\alpha|\omega|} = \frac{1}{2\pi}\frac{2\alpha}{\alpha^2+x^2}$$

konvergiert nämlich für $\alpha \to 0$ gegen die Deltafunktion $\delta(x)$. Dies kann man zeigen, indem man verifiziert, dass für $\alpha \to 0$ die beiden definierenden Eigenschaften der Deltafunktion erfüllt sind. Dies wird in den Übungen genauer betrachtet, zusammen mit anderen Darstellungen und den wichtigsten Rechenregeln.

### 11.3.2 Anwendung: DGL

Als physikalische Anwendung der Fourier-Transformation betrachten wir die DGL für eine erzwungene Schwingung:

$$\ddot{x} + a\dot{x} + bx = f(t).$$

Dabei ist $x(t)$ die Auslenkung aus der Ruhelage zur Zeit $t$ und $f(t)$ die (zeitabhängige) äußere Kraft. Die Terme auf der linken Seite repräsentieren die Beschleunigung, die Dämpfung durch eine geschwindigkeitsabhängige Reibungskraft und die Rückstellkraft (Hooke'sches Gesetz).
Wir wollen nun diese DGL mit Hilfe der Fourier-Transformation ganz allgemein lösen. Dazu stellen wir die Auslenkung $x(t)$ durch ihre Fourier-Transformierte $\tilde{x}(\omega)$ dar:

$$x(t) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} e^{i\omega t}\tilde{x}(\omega).$$

Nun können wir die Fourier-Transformierten der Zeitableitungen bestimmen:

$$\dot{x}(t) = \frac{1}{\sqrt{2\pi}}\frac{d}{dt}\int_{-\infty}^{\infty} e^{i\omega t}\tilde{x}(\omega) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}\frac{\partial}{\partial t}\left(e^{i\omega t}\tilde{x}(\omega)\right) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} i\omega e^{i\omega t}\tilde{x}(\omega).$$

Hieraus lesen wir ab:

$$\mathrm{FT}(\dot{x}) = i\omega\,\mathrm{FT}(x),$$

wobei $\mathrm{FT}(\dot{x})$ die Fourier-Transformierte von $\dot{x}$ bezeichnet. Analog folgt

$$\mathrm{FT}(\ddot{x}) = -\omega^2\,\mathrm{FT}(x).$$

Wir sehen hier die wesentliche Vereinfachung, die wir durch Betrachtung der Fourier-Transformierten erzielt haben: Nach Fouriertransformation werden aus Ableitungen einfache Multiplikationen (mit $i\omega$)!!!
Wir wenden nun die Fouriertransformation auf die gesamte DGL an, d.h. wir multiplizieren beide Seiten der Gleichung mit $e^{-i\omega t}$ und integrieren dann. Unter Ausnutzung der Linearität sowohl der DGL als auch der Fourier-Transformation erhalten wir dann:

$$\begin{aligned} F(\omega) := \mathrm{FT}(f(t)) &= \mathrm{FT}(\ddot{x} + a\dot{x} + bx) = \mathrm{FT}(\ddot{x}) + a\,\mathrm{FT}(\dot{x}) + b\,\mathrm{FT}(x) \\ &= -\omega^2\,\mathrm{FT}(x) + ai\omega\,\mathrm{FT}(x) + b\,\mathrm{FT}(x) = (b + ai\omega - \omega^2)\,\mathrm{FT}(x) \\ &= (b + ai\omega - \omega^2)\tilde{x}(\omega). \end{aligned}$$

Somit können wir die Fourier-Transformierte $\tilde{x}(\omega)$ von $x(t)$ explizit bestimmen:

$$\tilde{x}(\omega) = \mathrm{FT}(x) = \frac{F(\omega)}{b + ai\omega - \omega^2}.$$

Durch Rücktransformation erhalten wir dann die gesuchte Lösung

$$x(t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} d\omega e^{i\omega t} \tilde{x}(\omega) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} d\omega \frac{F(\omega) e^{i\omega t}}{b + ai\omega - \omega^2}.$$

Zur expliziten Berechnung muss natürlich die antreibende Kraft $f(t)$ (und damit ihre Fourier-Transformierte $F(\omega)$) angegeben werden. Später werden wir sehen, wie wir solche Intregale mit Hilfe der Funktionentheorie berechnen können.

# Kapitel 12

# Systeme von Differentialgleichungen

In diesem Kapitel wollen wir uns weiter mit Differentialgleichungen beschäftigten. In Kapitel 6 hatten wir die Grundlagen der Theorie der gewöhnlichen DGL kennengelernt. Zunächst wollen wir diese Thema etwas vertiefen und uns mit Systemen von DGL und einigen speziellen DGL der Physik genauer beschäftigen. Zum Abschluß werden wir dann partielle Differentialgleichungen betrachten.

In der Physik hat man es sehr oft mit Systemen von DGL zu tun, d.h. mehreren DGL, die untereinander gekoppelt sind. So etwas passiert z.B. immer dann, wenn man Systeme aus mehreren Teilchen betrachtet, die miteinander wechselwirken (d.h. Kräfte aufeinander ausüben). Über diese Wechselwirkungskräfte sind die Newtonschen Bewegungsleichungen der Teilchen dann miteinander verkoppelt.
Es gibt aber auch andere Gründe, warum man an Systemen von DGL interessiert sein kann. Als weitere Motivation betrachten wir die folgende DGL 3. Ordnung:

$$y''' - 3y'' - 9y' - 5y = 0\,.$$

Wir definieren nun neue Funktionen durch

$$y_1(x) := y(x), \qquad y_2(x) := y'(x), \qquad y_3(x) := y''(x)\,.$$

Für dies gilt dann offensichtlich folgendes System von DGL:

$$\begin{aligned} y_1' &= y_2\,, \\ y_2' &= y_3\,, \\ y_3' &= 3y_3 + 9y_2 + 5y_1\,, \end{aligned}$$

wobei wir die letzte Gleichung aus der ursprünglichen DGL 3. Ordnung erhalten, wenn wir sie nach $y''' = y_3'$ auflösen und die auftretenden Ableitungen durch die neu definierten Funktionen ersetzen.

In Matrixform können wir dieses System kompakt schreiben als

$$\begin{pmatrix} y_1' \\ y_2' \\ y_3' \end{pmatrix} = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 5 & 9 & 3 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix}.$$

Dies ist jetzt ein DGL-system, allerdings ist es von 1. Ordnung, höhere Ableitungen treten nicht mehr auf.
Allgemein kann man mit diesem Trick eine DGL n-ter Ordnung in $n$ gekoppelte DGL 1. Ordnung verwandeln.

Ein allgemeines Differentialgleichungssystem (1. Ordnung) hat die Form

$$\underline{y}' = \underline{f}(\underline{y}, x) \qquad \text{mit} \quad \underline{y}(x) = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}.$$

Hinzu kommt u.U. noch eine Anfgangsbedinung $\underline{y}(x_0) = \underline{y}_0$.
Ist $\underline{f}(\underline{y}, x) = \underline{f}(\underline{y})$, so spricht man von einem **autonomen System**.

Im Fall $n = 1$ reduziert sich dies auf $y' = f(y, x)$, die allgemeine DGL 1. Ordnung wie wir sie bereits in Kapitel 6 kennengelernt haben.

## 12.1 Lineare Differentialgleichungssysteme

Zunächst betrachten wir den Fall eines autonomen Systems, bei dem die Funktion $\underline{f}(\underline{y})$ linear ist. Dann können wir das System auch als Matrixgleichung der Form

$$\underline{y}' = \underline{\underline{A}} \cdot \underline{y} + \underline{B}$$

mit $n \times n$-Matrizen $\underline{\underline{A}}$ und $\underline{\underline{B}}$ schreiben, deren Elemente Konstanten sind.

Zunächst betrachten wir wieder den homogenen Fall

$$\underline{y}' = \underline{\underline{A}} \cdot \underline{y}$$

und machen analog zum Fall $n = 1$ den Ansatz

$$\underline{y}_j(x) = e^{\lambda_j x} \underline{u}_j .$$

Dann ist $\underline{y}_j' = \lambda_j \underline{y}_j$ und somit

$$\lambda_j \underline{y}_j = \underline{\underline{A}} \cdot \underline{y}_j .$$

Dies ist eine Eigenwertgleichung! Somit liefert unser Ansatz eine Lösung des Systems, falls $\lambda_j$ ein Eigenwert von $\underline{\underline{A}}$ ist und $\underline{u}_j$ der zugehörige Eigenvektor.

Gibt es $n$ unterschiedliche Eigenwerte und linear unabhängige Eigenvektoren, so lautet die allgemeine Lösung

$$\underline{y}(x) = \sum_{j=1}^{n} c_j \underline{y}_j(x)$$

mit Konstanten $c_j$, die durch die Anfangsbedingung $\underline{y}(x_0) = \underline{y}_0$ bestimmt sind.

**Beispiel 12.1.1.** Als ein Beispiel betrachten wir die Schwingungsgleichung

$$\ddot{y} + \omega^2 y = 0\,.$$

Mit $y_1 := y$ und $y_2 = \dot{y}$ erhält man das System

$$\dot{\underline{y}} = \underline{\underline{A}} \cdot \underline{y} \qquad \text{mit} \quad \underline{\underline{A}} = \begin{pmatrix} 0 & 1 \\ -\omega^2 & 0 \end{pmatrix}\,.$$

Die Eigenwerte von $\underline{\underline{A}}$ sind

$$\lambda_{1,2} = \pm i\omega \qquad \text{mit Eigenvektor} \quad \underline{u}_{1,2} = \begin{pmatrix} 1 \\ \pm i\omega \end{pmatrix}\,.$$

Damit lautet die allgemeine Lösung

$$\underline{y}(t) = c_1 \begin{pmatrix} 1 \\ i\omega \end{pmatrix} e^{i\omega t} + c_2 \begin{pmatrix} 1 \\ -i\omega \end{pmatrix} e^{-i\omega t}\,.$$

Die erste Kompente entspricht der Lösung, die wir früher schon bestimmt hatten. Die zweite ist deren Ableitung, entspricht also der Geschwindigkeit, wenn man $y_1(t)$ als Auslenkung interpretiert.

*Bemerkung.* Wenn zwei Eigenwerte $\lambda_j = \lambda_l$ übereinstimmen, also entartet sind, und nur ein Eigenvektor $\underline{u}_j$ existiert, dann macht man den Ansatz

$$\underline{y}_l(x) = x\underline{y}_l(x) + e^{\lambda_j x}\underline{v} = e^{\lambda_j x}\left(x\underline{u}_l + \underline{v}\right)\,.$$

Eine kurze Rechnung liefert dann folgende Bedingung an den noch unbekannten Vektor $\underline{v}$:

$$\left(\underline{\underline{A}} - \lambda_j \mathbb{1}\right) \cdot \underline{v} = \underline{u}_j.$$

Diese Gleichung hat eine nicht-triviale Lösung für $\underline{v}$ (und damit $\underline{y}_l$), da ja $\det\left(\underline{\underline{A}} - \lambda_j \mathbb{1}\right) = 0$ ist, da $\lambda_j$ Eigenwert von $\underline{\underline{A}}$ ist.
Bei höheren Entartungen macht man dann den Ansatz

$$\underline{y}_k(x) = e^{\lambda_j x}\left(x^2\underline{u}_j + x\underline{v} + \underline{w}\right)$$

etc.

Um die Lösung des inhomogenen Falls ($\underline{\underline{B}} \neq 0$) zu erhalten, geht man wie früher im Fall $n = 1$ vor. Man bestimmt eine *spezielle* Lösung des inhomogenen Systems, z.B. durch Raten, Variation der Konstanten etc. Die allgemeine Lösung erhält man dann als Summe der speziellen und der allgemeinen Lösung des homogenen Problems (wie oben angegeben).

## 12.2 Dynamische Systeme

Wir betrachten nun ein allgemeines DGLsystem $\dot{\underline{y}} = \underline{f}(\underline{y}, t)$ und wollen die unabhängige Variable $t$ als 'Zeit' interpretieren. Dabei sei die Funktion $\underline{f}(\underline{y}, t)$ nicht notwendig linear. Ein solches Problem bezeichnet man auch als **dynamisches System**. Die Lösungen im nichtlinearen Fall haben häufig interessante Eigenschaften (Fraktale etc.)!

Wir wollen uns hier nur mit dem autonomen Fall

$$\dot{\underline{y}} = \underline{f}(\underline{y})$$

beschäftigen. Wenn wir $\underline{y}$ als den Ortsvektor eines Teilchens interpretieren, dann liefert die DGL die zu einer Position gehörige Geschwindigkeit. Dies macht die Bezeichnung "dynamisches System" plausibel.
Was passiert nun, falls an einem Punkt $\underline{y}_0$ gilt: $\underline{f}(\underline{y}_0) = 0$ ? Offensichtlich ist dann $\dot{\underline{y}} = 0$ und daher ändert sich $\underline{y}$ nicht mehr, wenn das Teilchen einmal den Punkt $\underline{y}_0$ erreicht hat! Solche Punkte bezeichnet man daher auch als **Fixpunkte**. I.a. werden sie erst nach langer Zeit, d.h. im Grenzfall $t \to \infty$ erreicht.
Je nach Verhalten in der Nähe des Fixpunktes unterscheidet man

- **anziehende** oder **attraktive Fixpunkte**,
- **abstoßende** oder **repulsive Fixpunkte**.

Man spricht in diesem Zusammenhang auch von der **Stabilität** des Fixpunktes und nennt attraktive Fixpunkte **stabil** und repulsive Fixpunkte **instabil**. Über die Stabilität eines Fixpunktes erhalten wir Aufschluß durch eine Taylorentwicklung in seiner Nähe. Da $\underline{f}(\underline{y}_0) = 0$ ist, folgt:

$$\underline{f}(\underline{y}) \approx \underline{\underline{A}} \cdot (\underline{y} - \underline{y}_0) \qquad \text{mit} \quad A_{jl} = \left.\frac{\partial f_j}{\partial y_l}\right|_{\underline{y}=\underline{y}_0} .$$

Mit der Abkürzung $\underline{z} := \underline{y} - \underline{y}_0$ erhält man dann die Gleichung

$$\dot{\underline{z}} = \underline{\underline{A}} \cdot \underline{z} \,,$$

die das Verhalten des dynamischen Systems in der Nähe des Fixpunktes beschreibt. $\underline{\underline{A}}$ heißt **Stabilitätsmatrix**, ihre Eigenwerte charakterisieren das Verhalten nahe dem Fixpunkt.

**Beispiel 12.2.1.** Als Beispiel betrachten wir im Fall $n = 1$ die **logististische Gleichung**

$$\dot{y} = \alpha y(1 - y)$$

mit $\alpha > 0$, die z.B. in der Populationsdynamik auftritt. Sie enthält zwei konkurrierende Terme $y$ und $-y^2$, die zum Wachstum bzw. Schrumpfen (Zerfall) von $y$ führen. Ist $y$ klein, so kann man $y^2$ gegenüber $y$ vernachlässigen. Dann dominiert das Wachstum. Ist $y$ aber groß, so dominiert der quadratische Term, d.h. der Zerfall.

Die Funktion $f(y) = \alpha y(1-y)$ hat zwei Fixpunkte, die man direkt ablesen kann. Es sind $y_0 = 0$ und $y_1 = 1$. Führt man die oben beschriebene Entwicklung durch, so erhält man für $y_0$:

$$A_0 = f'(y = 0) = \alpha(1 - 2y)|_{y=0} = \alpha$$

und somit die linearisierte Gleichung für $z = y - y_0 = y$

$$\dot{z} = A_0 z = \alpha z.$$

Deren Lösung ist

$$z_{\text{lin}}^{(0)}(t) = Ce^{\alpha t}\,.$$

Analog erhält man für $y_1$ zunächst

$$A_1 = f'(y = 1) = \alpha(1 - 2y)|_{y=1} = -\alpha$$

und hieraus die linearisierte Gleichung für $z = y - y_1 = y - 1$

$$\dot{z} = A_1 z = -\alpha z\,,$$

deren Lösung durch

$$z_{\text{lin}}^{(1)}(t) = Ce^{-\alpha t}\,.$$

gegeben ist.
Da $\alpha > 0$, wachsen Abweichungen vom Fixpunkt $y_0 = 0$ schnell (exponentiell!) an, während Abweichungen vom Fixpunkt $y_1 = 1$ schnell kleiner werden. Daher ist $y_0$ abstoßend, also instabil, und $y_1$ anziehend und somit stabil.

# Kapitel 13

# Symmetrien und Gruppen

## 13.1 Symmetrien

Symmetrien oder Invarianzen spielen in der Physik eine extrem wichtige Rolle, da sie z.B. eine Vereinfachung der Beschreibung erlauben. Dabei gibt es unterschiedliche Klassen von Symmetrien, die eine Rolle spielen.

1. **Symmetrien von Objekten**: Z.B. ist ein Kreis invariant unter beliebigen Drehungen um seinen Mittelpunkt. Ein Quadrat ist dagegen nur invariant unter Drehungen um den Mittelpunkt, wenn der Drehwinkel ein Vielfaches von $\frac{\pi}{2}$ beträgt. Ein gleichseitiges Dreieck schließlich ist invariant unter Drehungen um den Schwerpunkt um Vielfache von $\frac{2\pi}{3}$.

2. **Symmetrien der Naturgesetze**: Wir nehmen an, dass die Naturgesetze zeitlich unveränderlich sind, d.h. ein (ideales) Experiment liefert heute und nächste Woche das gleiche Ergebnis. Die Naturgesetze sind daher invariant unter zeitlichen Verschiebungen.
   In analoger Weise sind sie auch unter räumlichen Verschiebungen invariant, d.h. die Naturgesetze in Köln und New York sind die gleichen.
   Offensichtlich besteht auch Invarianz unter beliebigen Kombinationen von zeitlichen und räumlichen Verschiebungen. Eine wichtige Konsequenz solcher Invarianzen sind sog. *Erhaltungssätze*, z.B. impliziert die Invarianz unter zeitlichen Verschiebungen den Energieerhaltungssatz.

Allgemein versteht man unter einer **Symmetrie** eine Operation, die ein Objekt in eine von ihm nicht unterscheidbare Form überführt. Diese abstrakte Definition wird von den oben angegebenen Beispiele gut illustriert.

Grundsätzlich ist es eine Bestimmung aller dieser Operation (Rotationen, Verschiebungen, Spiegelungen etc.) bei jedem Problem sehr wichtig.

Die Gruppentheorie erlaubt nun die Untersuchung der Gesamtheit solcher Operationen. Die Elemente der (mathematischen) Gruppe entsprechen dabei den Symmetrieoperationen.

## 13.2 Gruppen

**Definition 13.2.1** (Gruppe)**.**
Eine **Gruppe** ist eine (nichtleere) Menge $G$, für deren Elemente eine Verknüpfung '$\cdot$' definiert ist, die üblicherweise als (Gruppen-)Multiplikation[1] bezeichnet wird.
Die Gruppenmultiplikation muß folgende Eigenschaften haben (Man beachte, dass statt $a \cdot b$ oft einfach $ab$ geschrieben wird!):

1. Abgeschlossenheit: Für alle $a, b \in G$ ist auch $ab \in G$.
2. Assoziativität: Sind $a, b, c \in G$, so gilt: $(ab)c = a(bc)$.
3. Einheitselement (neutrales Element): Es gibt ein $e \in G$, so dass für alle $a \in G$ gilt: $ae = a$ und $ea = a$.
4. Inverses Element: Zu jedem $a \in G$ existiert ein $b \in G$ mit $ab = e$ und $ba = e$. Man bezeichnet dieses Element auch als das zu $a$ Inverse und schreibt $b = a^{-1}$.
   Bem.: Wie man leicht überprüft, gilt $(ab)^{-1} = b^{-1}a^{-1}$.

Die Anzahl der Elemente einer Gruppe heißt **Ordnung** der Gruppe. **Diskrete Gruppen** haben endlich oder abzählbar unendlich viele Elemente. **Kontinuierliche Gruppen** haben dagegen überabzählbar viele Elemente.

Gilt zusätzlich

**5.** Kommutativität: Für alle $a, b \in G$ ist $ab = ba$.

so heißt die Gruppe **kommutativ** oder **abelsch**.

**Beispiel 13.2.1.** Wir betrachten einige Beispiele für Gruppen.

1. Die ganzen Zahlen $\mathbb{Z}$ bilden bzgl. der Addition eine diskrete, abelsche Gruppe.
2. $\{-1, 1\}$ bzgl. der (üblichen) Multiplikation.
3. $\left\{e^{i\alpha} \middle| \alpha \in \mathbb{R}\right\}$ mit der Multiplikation.
4. Alle Vektorräume sind bzgl. der Vektoraddition abelsche Gruppen.
5. Die natürlichen Zahlen $\mathbb{N}$ mit der Addition sind bilden keine Gruppe, da die Inversen nicht enthalten sind (z.B. $-2$, das Inverse von $2$ bzgl. der Addition).

Endliche Gruppen kann man auch durch eine **Gruppenmultiplikationstabelle** oder **Gruppentafel** definieren. Für Beispiel 2) lautet diese z.B.

| $\cdot$ | $1$ | $-1$ |
|---|---|---|
| $1$ | $1$ | $-1$ |
| $-1$ | $-1$ | $1$ |

[1] Selbst dann, wenn es sich eigentlich um eine Addition handelt!

Gruppen mit einer analogen Gruppentafel bezeichnet man mit $Z_2$.
Die Gruppe $Z_4$ ist durch die Gruppentafel

| · | 1 | a | b | c |
|---|---|---|---|---|
| 1 | 1 | a | b | c |
| a | a | b | c | 1 |
| b | b | c | 1 | a |
| c | c | 1 | a | b |

definiert.
Anhand solcher Gruppentafeln lassen sich die Gruppeneigenschaften leicht überprüfen.

**Definition 13.2.2** (Untergruppe)**.**
Ist eine Teilmenge $H \subset G$ selbst eine Gruppe, so nennt man sie auch **Untergruppe von G**.

Bem.: Man muß nicht alle Gruppeneigenschaften nachweisen. Es genügt, neben der Abgeschlossenheit von $H$, zu zeigen, dass das neutrale Element und die Inversen *in* $H$ enthalten sind. Ihre Existenz ist bereits klar, da $G$ eine Gruppe ist.

**Definition 13.2.3** (direktes Produkt)**.**
Aus zwei Gruppen $G_1$ und $G_2$ kann man eine neue Gruppe $G := G_1 \times G_2$ konstruieren, das **direkte Produkt** von $G_1$ und $G_2$, mit den Elementen $g := (a, b)$, wobei $a \in G_1$ und $b \in G_2$, und der Multiplikation

$$g_1 g_2 = (a_1, b_1)(a_2, b_2) := (a_1 a_2, b_1 b_2).$$

Auf Grund dieser Definition "erbt" $G$ die Gruppeneigenschaften seiner Bausteine $G_1$ und $G_2$.

## 13.3 Wichtige Gruppen

### 13.3.1 Permutationsgruppe

**Definition 13.3.1** (Permutationsgruppe $S_n$)**.**
Die Menge aller möglichen Permutationen von $n$ Elementen ist eine Gruppe der Ordnung $n!$, die sog. **Permutationsgruppe** $S_n$. Dabei ist die Gruppenmultiplikation als Hintereinanderausführung zweier Permutationen definiert.

Wir hatten schon in Kapitel 2.3 einige Eigenschaften von Permutationen kennengelernt. Ausgehend von der Identität $(1\ 2\ 3\ 4)$ (im Falle $n = 4$) können wir z.B. die Permutation $1 \to 1$, $2 \to 3$, $3 \to 4$, $4 \to 2$ beschreiben durch die Aufzählung der Bildmenge $(1\ 3\ 4\ 2)$ oder ausführlicher durch $\begin{pmatrix} 1 & 2 & 3 & 4 \\ 1 & 3 & 4 & 2 \end{pmatrix}$.
Die Multiplikation, also Hintereinanderausführung zweier Permutationen sieht dann z.B. folgendermaßen aus ($n = 3$):

$$\begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix} \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix}.$$

Dabei wird der rechte 'Faktor' zuerst angewendet und auf dessen Ergebnis dann der linke Faktor. Z.B. wird die '1' zunächst auf die '2' abgebildet und diese dann auf die '3'.
Vertauschen wir die Reihenfolge der Faktoren, so erhalten wir

$$\begin{pmatrix}1&2&3\\2&3&1\end{pmatrix}\begin{pmatrix}1&2&3\\1&3&2\end{pmatrix}=\begin{pmatrix}1&2&3\\2&1&3\end{pmatrix}.$$

Die Permutationsgruppe $S_n$ ist also nicht kommutativ!

Die Menge (siehe auch Kap. 2.3)

$$A_n := \{P \in S_n \big| P \text{ ist gerade Permutation}\}$$

ist eine Untergruppe von $S_n$, die sog. **alternierende Gruppe**.

Die Menge

$$A_n := \{P \in S_n \big| P \text{ ist zyklische Permutation}\}$$

ist ebenfalls eine Untergruppe von $S_n$, die man als **zyklische Gruppe** bezeichnet. Genauer ist $A_n$ eine abelsche Untergruppe mit $n$ Elementen.
Z.B. im Fall $n = 4$ sind ist die zyklische Gruppe gegeben durch

$$A_n := \{(1234), (4123), (3412), (2341)\}.$$

Die Menge $\{0, 1, 2, \ldots, n-1\}$ mit der Addition modulo $n$ (d.h. man betrachtet nur den Rest bei Division durch $n$, z.B. ist dann $(n-1)+3 = n+2 \equiv 2 \mod n$) ist eine zyklische Gruppe, die man $Z_n$ nennt. Die Multiplikationstabellen im Fall $n = 2$ und $n = 4$ hatten wir schon in Kapitel 13.2 angegeben.

### 13.3.2 Matrixgruppen

**Definition 13.3.2** (Matrixgruppen)**.**
Eine Reihe wichtiger Gruppen sind Teilmengen der Menge $\mathcal{M}(n, n)$ aller $n \times n$-Matrizen. In Kapitel 13.4 werden wir sehen, dass sie auch Repräsentanten von abstrakten Gruppen mit der gleichen Struktur sind.

- Die **orthogonale Gruppe**

$$\boxed{O(n) := \left\{\underline{\underline{M}} \in \mathcal{M}(n,n) \big| \underline{\underline{M}}^T \cdot \underline{\underline{M}} = \mathbb{1} = \underline{\underline{M}} \cdot \underline{\underline{M}}^T\right\}}$$

aller orthogonalen $n \times n$-Matrizen. Dabei heißt eine Matrix $\underline{\underline{M}}$ **orthogonal**, falls gilt

$$\underline{\underline{M}}^{-1} = \underline{\underline{M}}^T,$$

d.h. die transponierte Matrix ist gleich der Inversen.

Orthogonale Matrizen lassen Skalarprodukte invariant:

$$(\underline{\underline{M}}\,\underline{a})\cdot(\underline{\underline{M}}\,\underline{b}) = (\underline{\underline{M}}\,\underline{a})^T\cdot(\underline{\underline{M}}\,\underline{b}) = \underline{a}^T\underline{\underline{M}}^T\underline{\underline{M}}\,\underline{b} = \underline{a}^T\underline{b} = \underline{a}\cdot\underline{b}.$$

Dabei haben wir ausgenutzt, dass sich das Skalarprodukt $\underline{a}\cdot\underline{b}$ zweier Vektoren auch als Matrixmultiplikation $\underline{a}^T\underline{b}$ interpretieren läßt und das allgemein gilt $(\underline{\underline{A}}\,\underline{\underline{B}})^T = \underline{\underline{B}}^T\underline{\underline{A}}^T$. Anschaulich umfasst daher die Gruppe $O(n)$ alle Drehungen und Drehspiegelungen im $\mathbb{R}^n$. Es gilt

$$1 = \det(\mathbb{1}) = \det\left(\underline{\underline{M}}^T\cdot\underline{\underline{M}}\right) = \det(\underline{\underline{M}}^T)\det(\underline{\underline{M}}) = \left(\det\underline{\underline{M}}\right)^2.$$

Somit haben orthogonale Matrizen die Determinante $\det\underline{\underline{M}} = \pm 1$. Matrizen mit $\det\underline{\underline{M}} = 1$ beschreiben reine Drehungen, die mit $\det\underline{\underline{M}} = -1$ Drehspiegelungen.

- Die **spezielle orthogonale Gruppe**

$$\boxed{SO(n) := \left\{\underline{\underline{M}} \in O(n)\middle|\,\det(\underline{\underline{M}}) = 1\right\}}$$

beschreibt reine Drehungen im $\mathbb{R}^n$.

- Die **unitäre Gruppe**

$$\boxed{U(n) := \left\{\underline{\underline{U}} \in \mathcal{M}(n,n)\middle|\underline{\underline{U}}^\dagger\cdot\underline{\underline{U}} = \mathbb{1} = \underline{\underline{U}}\cdot\underline{\underline{U}}^\dagger\right\}}$$

aller unitären $n\times n$-Matrizen. Dabei heißt eine Matrix $\underline{\underline{U}}$ **unitär**, falls gilt

$$\underline{\underline{U}}^{-1} = \underline{\underline{U}}^\dagger,$$

wobei die **adjungierte Matrix** (oder auch **hermitesch konjugierte Matrix**) definiert ist durch

$$\underline{\underline{U}}^\dagger := \left(\bar{\underline{\underline{U}}}\right)^T,$$

d.h. man bildet die Transponierte der Matrix, die aus den komplex-konjugierten Elementen der Matrix $\underline{\underline{U}}$ besteht. Ist die Matrix reell, so ist die Adjungierte gleich der Transponierten.

Wie im Fall der orthogonalen Gruppe folgt wieder, dass $\det(\underline{\underline{U}}) = \pm 1$ ist.

- Die **spezielle unitäre Gruppe**

$$\boxed{SU(n) := \left\{\underline{\underline{U}} \in U(n)\middle|\,\det(\underline{\underline{U}}) = 1\right\}.}$$

- Die **allgemeine lineare Gruppe** (engl. “general linear group”)

$$\boxed{GL(n,\mathbb{R}) := \left\{\underline{\underline{M}} \in \mathcal{M}(n,n)\middle|\,\det(\underline{\underline{M}}) \neq 0\right\}}$$

umfaßt die allgemeinen, nicht singulären (d.h. invertierbaren), linearen Transformationen $\underline{x}\mapsto\underline{x}' = \underline{\underline{M}}\,\underline{x}$ im $\mathbb{R}^n$.

## 13.4 Darstellungen

Gruppen können auf verschiedene Weise realisiert werden. Erfolgt die Realisierung durch Matrizen, so spricht man von einer **Darstellung** der Gruppe.

**Beispiel 13.4.1.** Wir hatten schon am Anfang dieses Kapitels argumentiert, dass die Drehungen im $\mathbb{R}^n$ eine Gruppe bilden. Wählt man nun eine Basis, so entsprechen den Vektoren als $n$-komponentige Spaltenvektoren und deren Drehung der Multiplikation mit einer orthogonalen $n \times n$-Matrix (mit Determinante 1). In diesem Sinne ist die Matrixgruppe $SO(n)$ als Darstellung einer abstrakten Gruppe zu interpretieren, die den Drehungen im $\mathbb{R}^n$ entspricht.

**Definition 13.4.1** (Darstellung)**.**
Eine **Darstellung** $D$ der Gruppe $G$ ist eine Abbildung

$$D:\ G \to GL(n,\mathbb{C}) \qquad \text{mit} \quad \underline{\underline{D}}(a)\underline{\underline{D}}(b) = \underline{\underline{D}}(ab).$$

Eine Abbildung mit dieser Eigenschaft bezeichnet man auch als **Gruppenhomomorphismus**. Offensichtlich folgt hieraus bereits, dass $\underline{\underline{D}}(e) = \mathbb{1}$ sein muß.
Darstellungen sind nicht eindeutig, man kann immer äquivalente Darstellungen $D'$ durch eine **Ähnlichkeitstransformation**

$$\underline{\underline{D}}'(a) := \underline{\underline{V}}\ \underline{\underline{D}}(a)\underline{\underline{V}}^{-1}$$

mit einer nicht-singulären Matrix $\underline{\underline{V}}$ erzeugen. Das diese die gleichen Gruppeneigenschaften hat, sieht man folgendermaßen: Aus $\underline{\underline{D}}(a)\underline{\underline{D}}(b) = \underline{\underline{D}}(ab)$ folgt

$$\underline{\underline{D}}'(ab) = \underline{\underline{V}}\ \underline{\underline{D}}(ab)\underline{\underline{V}}^{-1} = \underline{\underline{V}}\ \underline{\underline{D}}(a)\underline{\underline{D}}(b)\underline{\underline{V}}^{-1} = \underline{\underline{V}}\ \underline{\underline{D}}(a)\underline{\underline{V}}^{-1}\underline{\underline{V}}\,\underline{\underline{D}}(b)\underline{\underline{V}}^{-1} = \underline{\underline{D}}'(a)\underline{\underline{D}}'(b).$$

Wir betrachten noch einige Spezialfälle:

- Bei einer **unitären Darstellung** sind die Matrizen $\underline{\underline{D}}(a)$ unitär. Dann gilt auch

$$\underline{\underline{D}}(a^{-1}) = \left(\underline{\underline{D}}(a)\right)^{\dagger} = \left(\underline{\underline{D}}(a)\right)^{-1}.$$

- Bei einer **treuen Darstellung** entspricht jedem Gruppenelement genau eine Matrix und umgekehrt.

- Eine **reguläre Darstellung** einer endlichen Gruppe der Ordnung $n$ ist durch eine $n \times n$-Matrixdarstellung gegeben. Reguläre Darstellungen sind auch treu.

- Eine **reduzible Darstellung** hat z.B. die Form

$$\underline{\underline{D}}(a) = \begin{pmatrix} \underline{\underline{A}}(a) & 0 \\ 0 & \underline{\underline{B}}(a) \end{pmatrix},$$

  d.h. sie zerfällt in unabhängige Teile, die jeweils für sich eine Gruppe darstellen. Ist das nicht der Fall, so spricht man von einer **irreduziblen Darstellung**.

## 13.5 Kontinuierliche Gruppen

**Definition 13.5.1** (kontinuierlichen Gruppen)**.**
Bei einer **kontinuierlichen Gruppe** $G$ können Gruppenelemente $g(t) \in G$ durch stetige Änderungen eines Parameters $t$ ineinander übergeführt werden. Dabei kann $t$ auch ein Vektor sein, der für mehrere reelle Parameter steht.
Auf Grund der Gruppeneigenschaft gilt

$$g(t_1)g(t_2) = g(t_3) \qquad \text{mit } t_3 = h(t_1, t_2).$$

**Beispiel 13.5.1.** Ein Beispiel ist die unitäre Gruppe $U(1)$, deren Elemente die Form $g(\varphi) = e^{i\varphi}$ haben: $U(1) = \{e^{i\varphi} | \varphi \in \mathbb{R}\}$. In diesem Fall ist wegen

$$g(\varphi_1)g(\varphi_2) = e^{i\varphi_1}e^{i\varphi_2} = e^{i(\varphi_1+\varphi_2)} = g(\varphi_1 + \varphi_2)$$

$h$ durch

$$h(\varphi_1, \varphi_2) = \varphi_1 + \varphi_2$$

gegeben.

**Definition 13.5.2** (Lie-Gruppe)**.**
Bei einer **Lie-Gruppe** ist die Funktion $h(t_1, t_2)$ analytisch in ihren Argumenten.

**Beispiel 13.5.2.** Die Elemente $g(\varphi)$ der speziellen orthogonalen Gruppe $SO(2)$ beschreiben Drehungen um die $z$-Achse um den Winkel $\varphi$ und können allgemein in der Form (vgl. Aufg. 34)

$$g(\varphi) = \begin{pmatrix} \cos\varphi & \sin\varphi \\ -\sin\varphi & \cos\varphi \end{pmatrix}$$

mit $\varphi \in [0, 2\pi[$ geschrieben werden.
Von Drehungen um die gleiche Achse erwartet man, dass $g(\varphi_1)g(\varphi_2) = g(\varphi_1 + \varphi_2)$ gilt. Dies kann man explizit an Hand der angegeben Darstellung nachprüfen[2]. Daher ist wieder $h(\varphi_1, \varphi_2) = \varphi_1 + \varphi_2$.
Offensichtlich kommutieren die Drehungen miteinander. Dies gilt aber nicht mehr, wenn man Drehungen um *unterschiedliche* Achsen betrachtet!

**Beispiel 13.5.3.** Ein weiteres Beispiel sind die Translationen

$$T(a) : \; x \mapsto x' = x + a$$

für die offensichtlich auch

$$T(a)T(b) = T(a + b)$$

gilt. Es handelt sich also um eine abelsche Gruppe.

[2] Bei der Rechnung gehen die Additionstheoreme für Winkelfunktionen ein!

## 13.6 Generatoren und Lie-Algebra

Um zu betonen, dass der Parameter $t$ auch ein Vektor sein kann, schreiben wir im folgenden $g(\underline{a})$ statt $g(t)$, mit einem $n$-komponentigen Vektor $\underline{a}$. Außerdem nehmen wir o.B.d.A. an, dass $g(\underline{a} = \underline{0}) = e$, dem neutralen Element der Gruppe $G$, ist.

**Definition 13.6.1** (Generatoren einer Lie-Gruppe)**.**
Dann kann man $g(\underline{a})$ nach Taylor entwickeln (bis 1. Ordnung):

$$g(\underline{a}) = g(\underline{0}) + \sum_{k=1}^{n} a_k \frac{\partial g(\underline{a})}{\partial a_k}\bigg|_{\underline{a}=\underline{0}} + \mathcal{O}(a^2).$$

Dann definiert man die **Generatoren** $X_k$ der Lie-Gruppe durch

$$iX_k := \frac{\partial g(\underline{a})}{\partial a_k}\bigg|_{\underline{a}=\underline{0}},$$

wobei der Faktor $i$ einer verbreiteten Konvention entspricht.
Ist $g$ Darstellungsmatrix einer Gruppe, dann sind auch die Generatoren Matrizen. Hierzu werden wir gleich ein Beispiel kennenlernen.

Speziell für unitäre oder orthogonale Gruppen gilt:

$$\begin{aligned} e = g(\underline{a})g(\underline{a})^{-1} &= g(\underline{a})g(\underline{a})^{\dagger} \\ &= \left(e + i\sum_k a_k X_k + \mathcal{O}(a^2)\right)\left(e - i\sum_k a_k X_k^{\dagger} + \mathcal{O}(a^2)\right) \\ &= e + i\sum_k a_k (X_k - X_k^{\dagger}) + \mathcal{O}(a^2). \end{aligned}$$

In der Ordnung $\mathcal{O}(a)$ gilt daher

$$X_k = X_k^{\dagger},$$

d.h. die Generatoren sind hermitesch.
Weiter gilt für unitäre oder orthogonale Gruppen

$$\boxed{g(\underline{a}) = \exp\left(i\sum_{k=1}^{n} a_k X_k\right)}$$

wobei die rechte Seite durch die Reihendarstellung der Exponentialfunktion definiert ist.
Aus der allgemeinen Identität für Matrizen (bzw. Operatoren)[3]

$$\det(\exp(\underline{\underline{A}})) = \exp(\mathrm{Spur}\,\underline{\underline{A}}),$$

[3]Die man z.B. durch Diagonalisierung beweisen kann.

wobei die **Spur** einer $n \times n$-Matrix $\underline{\underline{A}}$ definiert ist durch

$$\mathrm{Spur}\,\underline{\underline{A}} := \sum_{j=1}^{n} A_{jj},$$

folgt speziell für die Wahl[4] $\underline{\underline{A}} = \ln \underline{\underline{Q}}$

$$1 = \det g = \exp(\mathrm{Spur}\ln g) = \exp\left[\mathrm{Spur}(i\sum_k a_k X_k)\right].$$

Die erste Identität gilt auf Grund der Unitarität von $g$. Somit erhalten wir

$$\mathrm{Spur}\, X_k = 0.$$

Wir fassen die obigen Ergebnisse zusammen:

**Satz 13.6.1.** Die Generatoren von unitären und orthogonalen Gruppen sind hermitesch und spurlos.

Wir wollen diese abstrakten Definitionen an einigen konkreten Beispielen diskutieren.

**Beispiel 13.6.1.**

1. Für die Gruppe $U(1)$ mit der Darstellung $g(\varphi) = e^{i\varphi}$ ist $n = 1$ und es gibt daher nur einen Generator. Wegen
$$\left.\frac{\partial g(\varphi)}{\partial \varphi}\right|_{\varphi=0} = \left. ie^{i\varphi}\right|_{\varphi=0} = i$$
ist dieser durch
$$X = 1$$
gegeben.

2. Für die spezielle orthogonale Gruppe $SO(2)$ ist
$$g(\varphi) = \begin{pmatrix} \cos\varphi & \sin\varphi \\ -\sin\varphi & \cos\varphi \end{pmatrix}$$
und somit wieder $n = 1$ und
$$iX = \left.\frac{\partial g(\varphi)}{\partial \varphi}\right|_{\varphi=0} = \begin{pmatrix} \cos\varphi & \sin\varphi \\ -\sin\varphi & \cos\varphi \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$$
also
$$X = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}.$$

[4] Wobei die rechte Seite wieder über die Reihenentwicklung definiert ist.

Hiermit können wir jetzt auch das oben angegebene allgemeine Resultat

$$g(\varphi) = \exp(i\varphi X)$$

überprüfen:

$$\begin{aligned}
\exp(i\varphi X) &= \sum_{k=0}^{\infty} \frac{\varphi^k}{k!}(iX)^k = \mathbb{1}\sum_{k=0}^{\infty}\frac{(-1)^k\varphi^{2k}}{(2k)!} + iX\sum_{k=0}^{\infty}\frac{(-1)^k\varphi^{2k+1}}{(2k+1)!} \\
&= \mathbb{1}\cos\varphi + iX\sin\varphi = \begin{pmatrix}\cos\varphi & 0\\ 0 & \cos\varphi\end{pmatrix} + \begin{pmatrix}0 & \sin\varphi\\ -\sin\varphi & 0\end{pmatrix} \\
&= \begin{pmatrix}\cos\varphi & \sin\varphi\\ -\sin\varphi & \cos\varphi\end{pmatrix} = g(\varphi),
\end{aligned}$$

wobei wir $(iX)^2 = \mathbb{1}$ ausgenutzt haben, was man leicht verifiziert.

3. Für die Gruppe der Translationen $T(a):\ x \mapsto x' = x + a$ gilt:

$$X = -i\frac{d}{dx}\,.$$

4. In den Übungen (Aufgabe 35) untersuchen wir den Fall der $SU(2)$, die drei Generatoren besitzt.

**Definition 13.6.2** (Lie-Algebra)**.**
Für die Generatoren einer Lie-Gruppe gilt

$$[X_j, X_k] = i\sum_l c_{jk}^l X_l,$$

wobei der **Kommutator** definiert ist als

$$[A, B] := AB - BA.$$

Die Zahlen $c_{jk}^l$ heißen **Strukturkonstanten**.
Die Generatoren bilden also einen Vektorraum bzw. sogar eine Algebra, die man **Lie-Algebra** der Gruppe bezeichnet.

## 13.7 Tensorrechnung

Neben Skalaren und Vektoren treten in vielen Bereichen der Physik sog. *Tensoren* auf. Diese sind durch ihr Verhalten unter orthogonalen Transformationen charakterisiert. In der Experimentalphysik-Vorlesung haben Sie schon den *Trägheitstensor* kennengelernt. Eine besondere Rolle spielen Tensoren aber in allgemeinen Relativitätstheorie.

### 13.7.1 Definition

Bevor wir eine genaue Definition geben, wollen wir ein Beispiel betrachten.

**Beispiel 13.7.1.** Wir betrachten die lineare Vektorfunktion $\underline{b} = T\underline{a}$, bei der jedem Vektor $\underline{a}$ in linearer Weise ein Vektor $\underline{b}$ zugeordnet wird. Für den Fall, dass es sich jeweils um Elemente des $\mathbb{R}^3$ handelt, schreibt man auch komponentenweise

$$\begin{aligned} b_1 &= t_{11}a_1 + t_{12}a_2 + t_{13}a_3\,, \\ b_2 &= t_{21}a_1 + t_{22}a_2 + t_{23}a_3\,, \\ b_3 &= t_{31}a_1 + t_{32}a_2 + t_{33}a_3\,, \end{aligned}$$

oder kürzer

$$b_i = t_{ij}a_j := \sum_{j=1^3} t_{ij}a_j\,.$$

Dabei wurde die **Einstein'sche Summenkonvention** verwendet. Diese besagt, dass über doppelt vorkommende Indizes zu summieren ist. Die Summationsgrenze ergeben sich dabei aus dem Zusammenhang. Diese Konvention ist sehr praktisch, wenn man es, wie in der allgemeinen Relativitätstheorie, mit vielen Matrizen, Tensoren etc. zu tun hat. Man muß allerdings sehr aufpassen, z.B. kann man in einem Ausdruck wie $t_{ij}a_j/(s_{ij}a_j)$ nicht einfach $a_j$ kürzen! In diesem Kapitel werden wir die Summenkonvention durchgehend verwenden!
In dem Beispiel haben wir die lineare Abbildung $T$ durch eine Matrixmultiplikation dargestellt. Unser Tensor ist daher nichts anderes als eine Matrix[5].
In der Physik charakterisiert man allgemein Tensoren durch ihr Verhalten unter orthogonalen Transformationen, z.B. Drehungen. Wir wenden daher auf $\underline{a}$ und $\underline{b}$ eine orthogonalen Transformation $\underline{\underline{R}} = (r_{ij})$, d.h. $\underline{\underline{R}}^{-1} = \underline{\underline{R}}^t$ bzw.

$$\underline{\underline{R}} \cdot \underline{\underline{R}}^t = \mathbb{1}\,, \qquad \text{bzw.} \quad r_{ij}r_{il} = \delta_{jl}\,,$$

(wobei wir die Summenkonvention verwendet haben) also eine Drehung oder Drehspiegelung, an:

$$a'_l := r_{lj}a_j\,, \qquad b'_k := r_{ki}b_i\,,$$

wobei die transformierten ("gedrehten") Vektoren durch einen Strich gekennzeichnet sind, bzw. die Umkehrung

$$a_j = r_{lj}a'_l\,, \qquad b_i = r_{ki}b'_k\,.$$

Diese Umkehrung folgt aus der Orthogonalität der Matrix $\underline{\underline{R}}$.
Wir fordern nun, dass im gedrehten System der gleiche formale Zusammenhang zwischen den Vektoren wie im ursprünglichen System bestehen soll, d.h.

$$b'_k = t'_{kl}a'_l\,.$$

[5] Wegen der späteren Verallgemeinerung des Tensorbegriffes schreiben wir hier $T$ statt $\underline{\underline{T}}$.

Es stellt sich dann die Frage, wie der transformierte Tensor $T' = (t'_{kl})$ mit dem ursprünglichen $T = (t_{kl})$ zusammenhängt. Dies zeigt folgende kurze Rechnung:

$$t'_{kl}a'_l = b'_k = r_{ki}b_i = r_{ki}\left(t_{ij}a_j\right) = r_{ki}t_{ij}\left(r_{lj}a'_l\right) = \left(r_{ki}r_{lj}t_{ij}\right)a'_l\,.$$

Somit lesen wir ab:

$$t'_{kl} = r_{ki}r_{lj}t_{ij}\,.$$

Dies definiert das Transformationsverhalten eines **Tensors 2. Stufe**, d.h. alle Größen, die sich unter Drehungen genau so transformieren, sind Tensoren.
Man beachte, dass dieses Transformationsverhalten analog zu dem des Produktes $b_ia_j$ zweier Vektoren[6] ist:

$$b'_ka'_l = r_{ki}r_{lj}b_ia_j\,.$$

**Definition 13.7.1** (Tensoren)**.**

Ein **Tensor 2. Stufe** transformiert sich unter orthogonalen Transformationen $\underline{\underline{R}} = (r_{ij})$ wie das Produkt zweier Vektorkomponenten:

$$\boxed{t'_{kl} = r_{ki}r_{lj}t_{ij}\,.}$$

Allgemein ist ein **Tensor n-ter Stufe** (vom Rang $n$ oder der Ordnung $n$) eine Menge von Termen mit $n$ Indizes und analogem Transformationsverhalten

$$\boxed{t'_{klmn\ldots} = r_{ka}r_{lb}r_{mc}r_{nd}\cdots t_{abcd\ldots}\,,}$$

d.h. mit einem Faktor $r_{ij}$ für jeden Index. Speziell haben wir

$$\begin{aligned}
n = 0 &: \quad \text{Skalare} && (1 \text{ Komponente im } \mathbb{R}^3)\,,\\
n = 1 &: \quad \text{Vektoren} && (3 \text{ Komponenten im } \mathbb{R}^3)\,,\\
n = 2 &: \quad \text{Matrizen} && (3^2 = 9 \text{ Komponenten im } \mathbb{R}^3)\,,
\end{aligned}$$

d.h. die Fälle $n = 1, 2, 3$ reduzieren sich auf bereits bekannte Objekte.
Etwas formaler sind Tensoren multilineare Abbildungen des Raumes $V_1 \otimes V_2 \otimes \cdots \otimes V_n$**,** dem **direkten Produkt** oder **Tensorprodukt** der Vektorräume $V_1, \ldots, V_n$**,** in die reellen oder komplexen Zahlen. Der Tensor "erbt" dann das Transformationsverhalten der einzelnen Räume.

Man unterscheidet häufig zwischen **symmetrischen** und **antisymmetrischen Tensoren**

$$t_{ik} = \pm t_{ki}\,,$$

wobei das obere (untere) Vorzeichen für den symmetrischen (antisymmetrischen) Fall steht. Bei Tensoren höherer Stufe muß man zusätzlich angeben, bzgl. welcher Indizes eine (Anti-)-Symmetrie besteht. So bedeutet

$$t_{ijkl\ldots} = \pm\, t_{jikl\ldots}$$

[6] Genauer müßte man sagen, zum Transformationsverhalten des Produktes zweier Vektorkomponenten!

(Anti-)Symmetrie in den ersten beiden Indizes.
Man kann Tensoren 2. Stufe immer in einen symmetrischen und einen antisymmetrischen Anteil aufspalten:

$$t_{ij} = \underbrace{\frac{1}{2}\left(t_{ij} + t_{ji}\right)}_{\text{symmetrisch}} + \underbrace{\frac{1}{2}\left(t_{ij} - t_{ji}\right)}_{\text{antisymmetrisch}} .$$

Für Tensoren höherer Stufe geht das für zwei beliebige Indizes.

**Beispiel 13.7.2.** Wir geben einige wichtige Beispiele für Tensoren.

1. Das Kronecker-Delta $\delta_{ij}$ ist ein symmetrischer Tensor 2. Stufe mit
$$\delta'_{ij} = r_{is} r_{jt} \delta_{st} = r_{is} r_{js} = \delta_{ij} ,$$
d.h. $\delta_{ij}$ ist invariant unter orthogonalen Transformationen.

2. Im $\mathbb{R}^2$ haben wir mit $\sigma_{12} = -\sigma_{21} = 1$ und $\sigma_{11} = \sigma_{22} = 0$ den sog. antisymmetrischen Tensor 2. Stufe
$$\sigma = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} .$$

3. Aus den Übungen (vgl. Aufgabe 7) kennen wir bereits das **Levi-Cevita-Symbol**
$$\epsilon_{ijk} := \begin{cases} 1, & (ijk) \text{ gerade Permutation von (123)}, \\ -1, & (ijk) \text{ ungerade Permutation von (123)}, \\ 0, & \text{sonst (d.h. mind. zwei gleiche Indizes)}. \end{cases}$$
Es ist also $\epsilon_{123} = \epsilon_{231} = \epsilon_{312} = 1$ und $\epsilon_{132} = \epsilon_{213} = \epsilon_{321} = -1$, alle 21 anderen Komponenten verschwinden.

   Das Levi-Cevita-Symbol bezeichnet man auch als $\epsilon$-**Tensor** oder den **vollständig antisymmetrischen Tensor 3. Stufe**.

4. Als physikalischen Beispiel wollen wir den **Trägheitstensor** $I = (I_{ij})$ nennen, der den Zusammenhang zwischen Drehimpuls $\underline{L}$ und Winkelgeschwindigkeit $\underline{\omega}$ für einen starren Körper beschreibt:
$$L_i = I_{ij}\omega_j = \int_V dV\, \rho(x_1, x_2, x_3) \left(r^2 \delta_{ij} - x_i x_j\right) \omega_j ,$$
wobei $\rho(\underline{r})$ die Massendichte des Körpers ist und $r^2 = x_l x_l = x_1^2 + x_2^2 + x_3^2$. Explizit hat er die Form
$$I = \int_V dV\, \rho(x_1, x_2, x_3) \begin{pmatrix} x_2^2 + x_3^2 & -x_1 x_2 & -x_1 x_3 \\ -x_2 x_1 & x_1^2 + x_3^2 & -x_2 x_3 \\ -x_3 x_1 & -x_3 x_2 & x_1^2 + x_2^2 \end{pmatrix} ,$$
d.h. der Trägheitstensor ist ein symmetrischer Tensor 2. Stufe. Das Trägheitsmoment um eine Achse (durch den Ursprung) in Richtung $\underline{n}$ ist dann durch $I_{\underline{n}} = n_i I_{ij} n_j$ gebeben.

### 13.7.2 Rechenregeln für Tensoren

Wir stellen hier kurz die wichtigsten Rechenregeln für Tensoren zusammen.
Die Addition (Subtraktion) zweier Vektoren (gleicher Stufe!) ist analog zur Addition von Matrizen komponentenweise definiert:

$$c_{ik} = a_{ik} \pm b_{ik} \,.$$

Entsprechendes gilt für Tensoren beliebiger Stufe.

Im Gegensatz zur Addition unterscheidet sich die Multiplikation von Tensoren von der von Vektoren oder Matrizen.

Das **direkte Produkt** oder **Tensorprodukt**

$$r_{jklm} := a_{jk} \otimes b_{lm} := a_{jk} b_{lm}$$

führt nämlich auf Tensoren höherer Stufe[7]. Es ist nicht nur für Tensoren gleicher Stufe definiert, die Definition läßt sich analog auf beliebige Produkte verallgemeinern. Allgemein gilt:

$$n\text{-Stufe} \otimes m\text{-Stufe} = (n+m)\text{-Stufe} \,.$$

Eine weitere neue Operation ist die sogenannte **Verjüngung** oder **Kontraktion**. Hierbei werden Indizes aussummiert, z.B. im Falle eines Tensors 4. Stufe, bei dem die ersten beiden Indizes verjüngt werden:

$$s_{lm} = r_{jjlm} \quad \left( = \sum_j r_{jjlm} \right) \,.$$

Die Verjüngung, die über zwei beliebige Indizes erfolgen kann, führt also auf einen Tensor der Stufe $n-2$.

Die Matrixmultiplikation können wir nun als Tensorprodukt mit anschließender Verjüngung darstellen, z.B. für das Produkt einer Matrix und eines Vektors:

$$t_{ij} \otimes a_k = t_{ij} a_k =: s_{ijk} \qquad \longrightarrow \qquad b_i = t_{ij} a_j = s_{ijj} \,.$$

Dies bezeichnet man auch als **Überschiebung** zweier Tensoren.
Analog ist das Skalarprodukt die Überschiebung zweier Tensoren 1. Stufe:

$$a_i \otimes b_j = a_i b_j \qquad \longrightarrow \qquad a_i b_i = \underline{a} \cdot \underline{b} \,.$$

Im Fall eines Tensors 2. Stufe entspricht die Verjüngung der Spurbildung:

$$\mathrm{Spur}(t_{jk}) = t_{jj} \,.$$

[7] Mit Ausnahme des Falls Tensoren 0. Stufe.

# Kapitel 14

# Differentialgleichungen III

## 14.1 DGL als Eigenwertproblem

Differentialgleichungen haben manchmal die Form von Eigenwertproblemen. Deren typische Form ist

$$\hat{D}y(x) = \lambda y(x)\,,$$

wobei $\hat{D}$ ein Operator ist, der Ableitungen enthält. In der Regel fordert man noch zusätzliche Randbedingungen, so dass die DGL nicht unbedingt für alle Werte von $\lambda$ eine Lösung hat. Existiert eine Lösung $y_\lambda(x)$, so heißt das entsprechende $\lambda$ Eigenwert und $y_\lambda(x)$ Eigenfunktion oder **Eigenlösung** von $\hat{D}$.

Ein wichtiges Beispiel ist die Schwingungsgleichung, die sich für $\hat{D} = \frac{d^2}{dx^2}$ ergibt. Wir werden darauf am Ende dieses Abschnitts zurückkommen. Vorher wollen wir uns mit einem allgemeineren Problem beschäftigen.

**Definition 14.1.1** (Sturm-Liouville-Problem)**.**
Wir betrachten die Funktionen $q(x)$, $p(x)$ und $r(x)$, die jeweils im Intervall $[a, b]$ definiert sind. Dabei sei $q(x)$ stetig, $p(x) > 0$ und zweimal stetig differenzierbar und $r(x) > 0$ und stetig in $]a, b[$.
Das Eigenwertproblem

$$\boxed{\hat{S}y(x) + \lambda r(x)y(x) = 0}$$

bezeichnet man als **Sturm-Liouville-Problem**, wobei $\hat{S}$ mit

$$\boxed{\hat{S}y(x) := \frac{d}{dx}\left(p(x)y'(x)\right) + q(x)y(x)}$$

als **Sturm-Liouville-Differentialoperator** bezeichnet wird.
Zusätzlich fordert man die sog. **Sturm-Liouville-Randbedingungen**. Hierbei soll für zwei beliebige, unterschiedliche Lösungen $u(x)$ und $v(x)$ des Sturm-Liouville-Problems gelten:

$$p(x)\left[u(x)v'(x) - u'(x)v(x)\right]\Big|_a^b = 0\,.$$

Dabei ist $...|_a^b$ wie bei der Integration zu interpretieren als die Differenz der Funktionswerte an der Stelle $b$ und $a$. Der Grund für diese recht ungewöhnliche Wahl der Randbedingungen wird gleich etwas klarer werden.

Man kann zeigen, dass das Sturm-Liouville-Problem unter den oben angegebenen Voraussetzungen (inklusive der Randbedingungen) reelle Eigenwerte $\lambda_n$ hat mit zugehörigen Eigenlösungen $y_n(x)$. Diese erfüllen folgende Orthogonalitätsbedingung

$$(y_n, y_m) := \int_a^b dx\, r(x) y_n(x) y_m(x) = 0 \quad \text{für } n \neq m\,.$$

Dies kann man relativ einfach beweisen. Zunächst gilt, da $y_n(x)$ eine Lösung der Sturm-Liouville-DGL ist,

$$(p(x) y_n'(x))' = -\left[q(x) + \lambda_n r(x)\right] y_n(x)\,.$$

Nun multiplizieren wir diese Gleichung mit $y_m(x)$ und integrieren dann über das Intervall $[a, b]$. Eine analoge Beziehung erhalten wir, wenn wir in obiger Vorgehensweise $n$ und $m$ vertauschen. Bildet man die Differenz dieser beiden Gleichungen, so folgt

$$p(x)\left[y_m(x) y_n'(x) - y_m'(x) y_n(x)\right]\Big|_a^b = (\lambda_m - \lambda_n) \int_a^b dx\, r(x) y_n(x) y_m(x)\,.$$

Die linke Seite verschwindet auf Grund der Sturm-Liouville-Randbedingungen. Da $\lambda_m \neq \lambda_n$ ist, muß das Integral verschwinden. An dieser Stelle erkennen wir einen Grund für die ungewöhnliche Wahl der Randbedingungen.

**Beispiel 14.1.1.** Als ein wichtiges Beispiel wollen wir auf die Schwingungsgleichung zurückkommen. Hier ist offensichtlich

$$p(x) = 1, \quad q(x) = 0, \quad r(x) = 1\,.$$

Als Intervall wählen wir $[-\pi, \pi]$, d.h. $b = -a = \pi$. Mit der Wahl $y(\pm\pi) = 0$ sind offensichtlich die Sturm-Liouville-Randbedingungen erfüllt.
Die Lösungen dieses Sturm-Liouville-Problems sind durch

$$\lambda_n = n^2 \qquad \text{und} \qquad y_n(x) = A \sin(nx)$$

gegeben. Dies sieht man folgendermaßen: Ohne Berücksichtigung der Randbedingungen lautet die allgemeine Lösung für $\lambda \geq 0$

$$y(x) = A \sin(\sqrt{\lambda} x) + B \cos(\sqrt{\lambda} x)\,.$$

Die Randbedingungen lassen sich nur erfüllen, falls $B = 0$ ist und $\sqrt{\lambda} = n \in \mathbb{N}$. Daher haben die Eigenwerte in diesem Fall die Form $\lambda_n = n^2$ mit $n \in \mathbb{N}$.
Die Eigenlösungen erfüllen die Orthogonalitätsbedingungen

$$(y_n, y_m) = \int_{-\pi}^{\pi} \sin(nx) \sin(mx) dx = 0 \quad \text{für } n \neq m\,,$$

was wir früher schon explizit gezeigt hatten.

## 14.2 Spezielle DGL

### 14.2.1 Legendre'sche DGL

Die **Legendre'sche DGL** lautet

$$\boxed{(1-x^2)y''(x) - 2xy'(x) + l(l+1)y(x) = 0\,,}$$

wobei $x \in [-1,1]$ und $l \in \mathbb{Z}$ sein soll[1]. Die Lösungen dieser Gleichung bezeichnet man als **Legendre-Polynome** $P_l(x)$ zu den Eigenwerten $l(l+1)$.
Dies kann man explizit nachprüfen mit dem Potenzreihenansatz $y(x) = \sum_{n=0}^{\infty} a_n x^n$. Explizit lauten die ersten Legendre-Polynome

$$P_0(x) = 1, \qquad P_1(x) = x, \qquad P_2(x) = \frac{3}{2}x^2 - \frac{1}{2}, \qquad P_3(x) = \frac{5}{2}x^3 - \frac{3}{2}x\,.$$

Für $l < 0$, $l \in \mathbb{Z}$ gilt:

$$P_{-l-1} = P_l\,.$$

Die Legendre-Polynome genügen der Orthogonalitätsbeziehung

$$\int_{-1}^{1} dx\, P_l(x) P_n(x) = \frac{1}{2l+1}\delta_{ln}\,.$$

Man kann die Legendre-Polynome durch Ableitung aus einer sog. **erzeugenden Funktion** erhalten. Diese ist definiert durch

$$\phi(x,h) := \sum_{l=0}^{\infty} P_l(x) h^l\,.$$

Damit gilt dann

$$P_n(x) = \frac{1}{n!}\frac{\partial^n}{\partial h^n}\phi(x,h)\big|_{h=0}\,.$$

Man kann zeigen, dass $\phi(x,h)$ explizit gegeben ist durch

$$\phi(x,h) = \frac{1}{\sqrt{1-2xh+h^2}} \qquad (|h|<1)\,.$$

Hieraus kann man dann durch Berechnung der $n$-ten Ableitung an der Stelle $h = 0$ das $n$-te Legendre-Polynom bestimmen.
Eine andere Möglichkeit zur Berechnung von $P_n(x)$ liefert die Rekursionsgleichung

$$(n+1)P_{n+1}(x) = (2n+1)xP_n(x) - nP_{n-1}(x)\,.$$

Hiermit kann man aus $P_0(x) = 1$ und $P_1(x) = x$ rekursiv alle höheren Legendre-Polynome bestimmen.

Neben der Legendre-DGL tritt auch manchmal die **verallgemeinerte Legendresche DGL**

[1] Es gibt auch Lösung für $l \in \mathbb{R}$. Diese haben die Form von Potenzreihen und werden hier nicht betrachtet.

$$\boxed{\frac{d}{dx}\left((1-x^2)y'(x)\right) - \left(-\frac{m^2}{1-x^2} + l(l+1)\right) y(x) = 0\,,}$$

auf, wobei $m \in \mathbb{Z}$ mit $m^2 \leq l^2$ ist. Für $m = 0$ reduziert sich dies zur Legendreschen DGL.
Die Lösungen der verallgemeinerten Legendreschen DGL sind durch die **zugeordneten Legendre-Polynome**

$$P_l^m(x) := (-1)^m (1-x^2)^{m/2} \frac{d^m}{dx^m} P_l(x) \qquad (0 \leq m \leq l)$$

gegeben. Für $-l < m < l$ gilt dabei

$$P_l^{-m}(x) = (-1)^m \frac{(l-m)!}{(l+m)!}\, P_l^m(x)\,.$$

### 14.2.2 Kugelflächenfunktionen

Eine sehr wichtige Gleichung der Physik ist die **Laplace-Gleichung**

$$\boxed{\triangle f = 0\,,}$$

die z.B. im Zusammenhang mit Wellenphänomenen auftaucht. Streng genommen handelt es sich um eine partielle DGL (siehe Kapitel 14.3), die wir aber im folgenden auf gewöhnliche DGLen zurückführen werden.
Wir wollen speziell die Lösung der Laplace-Gleichungen in Kugelkoordinaten $f(r, \vartheta, \varphi)$ untersuchen. Dazu machen wir den (Separations-)Ansatz (siehe später in Kapitel 14.4.4)

$$f(r, \vartheta, \varphi) = R(r)S(\vartheta)T(\varphi)\,,$$

d.h. wir suchen Lösungen, die sich als Produkt von Funktionen nur einer Variablen schreiben lassen!
Zunächst setzen wir diesen Ansatz in die Laplace-Gleichung ein, wobei wir den Laplace-Operator $\triangle$ zweckmäßigerweise in Kugelkoordinaten darstellen (siehe Kapitel 15.3.4):

$$\begin{aligned}
0 &= \triangle f(r, \vartheta, \varphi) \\
&= \left[\frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial}{\partial r}\right) + \frac{1}{r^2 \sin\vartheta}\frac{\partial}{\partial \vartheta}\left(\sin\vartheta\frac{\partial}{\partial \vartheta}\right) + \frac{1}{r^2\sin^2\vartheta}\frac{\partial^2}{\partial\varphi^2}\right] f(r, \vartheta, \varphi) \\
&= \frac{RST}{r^2}\left[\frac{1}{R}\frac{d}{dr}\left(r^2\frac{dR}{dr}\right) + \frac{1}{\sin^2\vartheta}\left(\frac{\sin\vartheta}{S}\frac{d}{d\vartheta}\left(\sin\vartheta\frac{dS}{d\vartheta}\right) + \frac{1}{T}\frac{d^2T}{d\varphi^2}\right)\right]\,.
\end{aligned}$$

Wir betrachten nun die beiden Terme in der eckigen Klammer. Der erste hängt nur von $r$ ab, der zweite nur von $\vartheta$ und $\varphi$. Daher kann die Gleichung für beliebige Werte von $r$, $\vartheta$ und $\varphi$ nur dann erfüllt sein, wenn beide konstant sind[2], genauer

$$\frac{1}{R}\frac{d}{dr}\left(r^2\frac{dR}{dr}\right) = \alpha \qquad \text{und} \qquad \frac{1}{\sin^2\vartheta}\left(\frac{\sin\vartheta}{S}\frac{d}{d\vartheta}\left(\sin\vartheta\frac{dS}{d\vartheta}\right) + \frac{1}{T}\frac{d^2T}{d\varphi^2}\right) = -\alpha$$

[2]Dahinter steckt folgende Aussage: Gilt $f(x) + g(y) = 0$ für alle $x$ und $y$, so ist $f(x) = a$ und $g(y) = -a$ mit einer Konstanten $a$. Die Aussage zeigt man leicht durch Differenzieren nach $x$ bzw. $y$.

mit einer noch unbekannten Konstanten $\alpha$.
Die Gleichung für $r$ wollen wir hier nicht weiter betrachten. Für die zweite Gleichung gilt nach der elementaren Umformung in

$$\left(\frac{\sin\vartheta}{S}\frac{d}{d\vartheta}\left(\sin\vartheta\frac{dS}{d\vartheta}\right)+\alpha\sin^2\vartheta\right)+\frac{1}{T}\frac{d^2T}{d\varphi^2}=0$$

ein ähnliches Argument: Der erste Teil hängt nur von $\vartheta$ ab, der zweite nur von $\varphi$. Daher gilt

$$\frac{\sin\vartheta}{S}\frac{d}{d\vartheta}\left(\sin\vartheta\frac{dS}{d\vartheta}\right)+\alpha\sin^2\vartheta=\beta \qquad \text{und} \qquad \frac{1}{T}\frac{d^2T}{d\varphi^2}=-\beta\,,$$

mit einer Konstanten $\beta$.
Die zweite Gleichung kann direkt gelöst werden:

$$T(\varphi)=t_1e^{i\sqrt{\beta}\varphi}+t_2e^{-i\sqrt{\beta}\varphi}\,.$$

Die erste Gleichung formen wir zunächst mit Hilfe der Substitution $x:=\cos\vartheta$ um in

$$\frac{d}{dx}\left((1-x^2)\frac{dS}{dx}\right)+\left(\alpha-\frac{\beta}{1-x^2}\right)S(x)=0\,.$$

Dies ist eine verallgemeinerte Legendre-DGL. Damit eine Lösung existiert, müssen $\alpha$ und $\beta$ die Form

$$\alpha=l(l+1) \qquad \text{und} \qquad \beta=m^2$$

haben mit ganzzahligen $l$ und $m$. Es gilt dann

$$S(\vartheta)=P_l^m(\cos\vartheta)\,.$$

Somit haben wir den Winkelanteil der Lösung der Laplace-Gleichung in Kugelkoordinaten bestimmt. Er wird erzeugt durch die sog. Kugelflächenfunktionen.

**Definition 14.2.1** (Kugelflächenfunktionen)**.**
Die **Kugelflächenfunktionen** sind definiert durch

$$\boxed{Y_{lm}(\vartheta,\varphi):=\sqrt{\frac{(2l+1)}{4\pi}\frac{(l-m)!}{(l+m)!}}\,P_l^m(\cos\vartheta)e^{im\varphi}}$$

Man beachte hierbei, dass es in der Literatur durchaus unterschiedliche Konventionen für das Vorzeichen und die Normierung gibt!
Explizit lauten die ersten Kugelflächenfunktionen

$$Y_{00}=\sqrt{\frac{1}{4\pi}},\quad Y_{11}=-\sqrt{\frac{3}{8\pi}}\sin\vartheta e^{i\varphi},\quad Y_{10}=\sqrt{\frac{3}{4\pi}}\cos\vartheta.$$

Außerdem gilt

$$Y_{l,-m}=(-1)^m\overline{Y_{lm}}\,.$$

Die Kugelflächenfunktionen erfüllen die Orthogonalitätsbeziehungen

$$\int d\Omega\overline{Y_{l'm'}}(\vartheta,\varphi)Y_{lm}(\vartheta,\varphi)=\int_0^{2\pi}d\varphi\int_0^{\pi}d\vartheta\,\sin\vartheta\,\overline{Y_{l'm'}}(\vartheta,\varphi)Y_{lm}(\vartheta,\varphi)=\delta_{ll'}\delta_{mm'}\,.$$

### 14.2.3 Bessel'sche DGL

Macht man einen analogen Separationsansatz für die Laplace-Gleichung in Zylinderkoordinaten, so kommt man für den $\rho$-abhängigen Teil der Lösung zur **Bessel'sche DGL**

$$\boxed{x\,\frac{d}{dx}\left(x\,\frac{dy}{dx}\right)+\left(x^2-p^2\right)y(x)=0\,.}$$

Eine Klasse von Lösungen sind die **Besselfunktionen 1. Art der Ordnung p**:

$$J_p(x)=\sum_{n=0}^{\infty}\frac{(-1)^n}{\Gamma(n+1)\Gamma(n+p+1)}\left(\frac{x}{2}\right)^{2n+p}\,.$$

Dabei ist $\Gamma(x)$ die **Gamma-Funktion**. Für positive, ganzzahlige $x=n+1$ gilt $\Gamma(n+1)=n!$. Weitere Lösungen sind durch die **Besselfunktionen 2. Art**

$$Y_p(x)=\lim_{\nu\to p}\frac{\cos(\pi\nu)J_\nu(x)-J_{-\nu}(x)}{\sin(\pi\nu)}$$

gegeben.

### 14.2.4 Hermite'sche DGL

Die **Hermite'sche DGL**

$$\boxed{\frac{d}{dx}\left(e^{-x^2}\frac{dy}{dx}\right)+2ne^{-x^2}y(x)=0}$$

wird durch die **Hermite-Polynome** $H_n(x)$ gelöst. Diese erhält man aus der erzeugenden Funktion

$$\phi(x,h)=\sum_{n=0}^{\infty}H_n(x)\frac{h^n}{n!}=e^{-h^2+2hx}\,.$$

Sie genügen der Orthogonalitätsrelation

$$\int_{-\infty}^{\infty}dxe^{-x^2}H_n(x)H_m(x)=\sqrt{\pi}2^n n!\delta_{nm}\,.$$

### 14.2.5 Laguerre'sche DGL

Die **Laguerre'sche DGL** lautet

$$\boxed{\frac{d}{dx}\left(xe^{-x}\frac{dy}{dx}\right)+ne^{-x}y(x)=0\,.}$$

Ihre Lösungen sind die **Laguerre-Polynome** $L_n$, die durch die erzeugende Funktion

$$\phi(x,h)=\sum_{n=0}^{\infty}L_n(x)h^n=\frac{1}{1-h}e^{-\frac{xh}{1-h}}$$

bestimmt sind.

## 14.3 Partielle DGL

**Partielle Differentialgleichungen** (pDGL) enthalten (partielle) Ableitungen nach mehreren Variablen, z.B. $\frac{\partial u(x,t)}{\partial x}$ und $\frac{\partial u(x,t)}{\partial t}$. Sie sind in der Regel deutlich schwieriger als gewöhnliche DGL und selbst für Gleichungen 1. Ordnung (bei denen nur erste Ableitungen auftauchen) gibt es keine so allgemeine Theorie wie für gewöhnliche DGL.

In der Physik am wichtigsten sind partielle DGL 2. Ordnung, bei der höchstens 2. Ableitungen auftreten. Ihre allgemeine Form ist im Falle von zwei unabhängigen Variablen $x$ und $y$:

$$a\frac{\partial^2 u}{\partial x^2} + b\frac{\partial^2 u}{\partial x \partial y} + c\frac{\partial^2 u}{\partial y^2} + d\frac{\partial u}{\partial x} + e\frac{\partial u}{\partial y} + fu = R(x,y)\,.$$

Dabei können alle auftretenden Koeffizienten wie die gesuchte Funktion $u(x,y)$ ebenfalls Funktionen von $x$ und $y$ sein, also $a = a(x,y)$ etc.

Man unterscheidet drei Typen : **elliptische**, **parabolische** und **hyperbolische** pDGL 2. Ordnung. Diese liegen vor, wenn folgende Bedingungen erfüllt sind:

$$\begin{aligned} \text{hyperbolisch:} \quad & b^2 > 4ac\,, \\ \text{parabolisch:} \quad & b^2 = 4ac\,, \\ \text{elliptisch:} \quad & b^2 < 4ac\,. \end{aligned}$$

Da $a$, $b$ und $c$ Funktionen sind, kann der Typ in verschiedenen Bereichen des Definitionsbereiches unterschiedlich sein!

Häufig hat man es mit einem **Randwertproblem** zu tun. Hierbei muß die gesuchte Lösung $u$ noch am Rand $\partial A$ des Definitionsbereiches $A$ gewisse Bedingungen erfüllen. Man unterscheidet dabei folgende Typen:

- **Dirichlet-Randbedingungen:** Hier sind die Funktionswerte am Rand vorgegeben, also $u(\partial A)$.
- **Neumann-Randbedingungen:** Hier ist die Ableitung der Funktion am Rand in Richtung der Normalen vorgegeben, d.h. $\underline{n} \cdot \operatorname{grad} u\big|_{\partial A}$.
- **Cauchy-Randbedingungen:** Hierbei handelt es sich um eine (gewichtete) Kombination von Dirichlet- und Neumann-Randbedingungen.

### 14.3.1 Wichtige pDGL der Physik

Im folgenden wollen wir die wichtigsten pDGL der Physik aufzählen und klassifizieren.

- Die **Laplace-Gleichung**

$$\boxed{\triangle u = 0}$$

haben wir bereits kennengelernt. In zwei Dimensionen lautet sie expliziter

$$\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = 0\,.$$

Die zugehörige inhomogene Gleichung

$$\boxed{\triangle u = \rho(x,y)}$$

heißt **Poisson-Gleichung**. Sie beschreibt z.B. das elektrische Potential einer Ladungsverteilung $\rho(x,y)$.

Die Poisson-Gleichung ist vom elliptischen Typ.

- Die **Diffusions-Gleichung** oder auch **Wärmeleitungsgleichung** lautet

$$\boxed{\triangle u(x,y,t) = \frac{1}{\kappa}\frac{\partial u}{\partial t}\,.}$$

Dabei ist $\kappa > 0$ eine Konstante.

Es handelt sich um eine parabolische DGL. Für $t \to \infty$ erhält man wegen

$$\lim_{t\to\infty} \frac{\partial u}{\partial t} = 0$$

eine statische (d.h. zeitunabhängige) Lösung. Diese erfüllt dann die Laplace-Gleichung!

- Die **Wellengleichung**

$$\boxed{\triangle u(x,y,t) - \frac{1}{v^2}\frac{\partial^2 u}{\partial t^2} = 0}$$

beschreibt die Ausbreitung die Ausbreitung von Wellen mit der Fortpflanzungsgeschwindigkeit $v$. Sie ist vom hyperbolischen Typ.

- Die **Schrödinger-Gleichung**

$$\boxed{-\frac{\hbar^2}{2m}\triangle u(\underline{r},t) + V(\underline{r})u(\underline{r},t) = i\hbar\frac{\partial u}{\partial t}}$$

ist die grundlegende Gleichung der Quantenmechanik. Sie beschreibt die quantenmechanische Wellenfunktion $u(\underline{r},t) = u(x,y,z,t)$ eines (nichtrelativistischen) Teilchens der Masse $m$, das sich im Potential $V(\underline{r})$ bewegt. $\hbar = h/2\pi$ ist das Wirkungsquantum.

In den obigen Beispielen haben wir in der Regel den zweidimensionalen Fall (zwei Raumdimensionen $x$ und $y$) angegeben. Die Verallgemeinerung auf den $n$-dimensionalen Fall ist in allen Beispielen offensichtlich.

## 14.4 Lösungsverfahren für pDGL; Green'sche Funktionen

Wie schon erwähnt, gibt es für pDGL weniger allgemeine Aussagen oder Lösungsverfahren als für gewöhnliche DGL. In der Praxis werden daher häufig numerische Methoden wie Relaxationsverfahren oder Fast Fourier Transformation (FFT) eingesetzt. Hier wollen wir aber einige Beispiele für analytische Methoden vorstellen.

### 14.4.1 Integraldarstellung

Wir betrachten die eindimensionale Diffusionsgleichung[3]

$$\frac{\partial^2 u}{\partial x^2} = \frac{\partial u}{\partial t}$$

mit der Rand- bzw. Anfangsbedingung $u(x,0) = f(x)$. Sie beschreibt die zeitliche Entwicklung der zur Zeit $t = 0$ vorgegebenen Wärmeverteilung $u(x,t)$.
In den Übungen (Aufgabe 32) werden wir mit Hilfe der Fourier-Transformation zeigen, dass sich Lösung in folgender Form schreiben läßt:

$$\boxed{u(x,t) = \frac{1}{2\sqrt{\pi t}} \int_{-\infty}^{\infty} dy\, e^{-\frac{(x-y)^2}{4t}} f(y)\,.}$$

Den Faktor $e^{-\frac{(x-y)^2}{4t}}$ bezeichnet man in diesem Zusammenhang auch als **Hitzekern** ("heat kernel").

### 14.4.2 Integraltransformation

Wir betrachten die Poisson-Gleichung in drei Raumdimensionen:

$$\triangle u(\underline{r}) = -4\pi q \delta^{(3)}(\underline{r})\,,$$

wobei

$$\delta^{(3)}(\underline{r}) := \delta(x)\delta(y)\delta(z)\,.$$

Physikalisch beschreibt dann $u(\underline{r})$ das elektrische Potential einer Punktladung $q$ am Ursprung $\underline{r} = 0$. Wir werden später sehen, wie man aus der Lösung dieses speziellen Problems zur Lösung für eine beliebige Inhomogenität $\rho(\underline{r})$ kommt.
Zur Lösung dieser Gleichung wenden wir eine Fouriertransformation auf alle drei Variablen $x$, $y$, $z$ an, z.B.

$$\tilde{u}(p_x,p_y,p_z) := \mathrm{FT}(u(\underline{r})) = \frac{1}{\sqrt{2\pi}^3} \int_{-\infty}^{\infty} dx \int_{-\infty}^{\infty} dy \int_{-\infty}^{\infty} dz\, e^{-ip_x x} e^{-ip_y y} e^{-ip_z z} u(x,y,z)\,.$$

[3]Wobei wir o.B.d.A. $\kappa = 1$ setzen.

Man beachte, dass es sich um drei getrennte Transformationen handelt. Die Variablen im Fourierraum haben wir dabei mit $p_x$, $p_y$ und $p_z$ bezeichnet. Wenn wir sie als Komponenten eines Vektors $\underline{p}$ interpretieren, können wir dies auch kompakter schreiben als

$$\tilde{u}(\underline{p}) = \frac{1}{\sqrt{2\pi}^3} \int d^3r \, e^{-i\underline{p}\cdot\underline{r}} \, u(\underline{r}) \, .$$

Wendet man nun die Fouriertransformation auf die obige Poisson-Gleichung an, so erhält man nach kurzer Rechnung

$$-p^2 \tilde{u}(\underline{p}) = -4\pi q \, \frac{1}{\sqrt{2\pi}^3} \, ,$$

und somit

$$\tilde{u}(\underline{p}) = \sqrt{\frac{2}{\pi}} \, \frac{q}{p^2} \, .$$

Der Faktor $-p^2 = -p_x^2 - p_y^2 - p_z^2$ entsteht dabei durch Anwendung des Laplace-Operators $\triangle = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}$ auf die Fouriertransformierte.
Durch Rücktransformation können wir nun die Lösung explizit bestimmen:

$$\begin{aligned} u(\underline{r}) &= \frac{1}{\sqrt{2\pi}^3} \int d^3p \, \tilde{u}(\underline{p}) e^{i\underline{p}\cdot\underline{r}} \\ &= \frac{q}{2\pi^2} \int_0^\infty dp \int_0^{2\pi} d\varphi \int_0^\pi d\vartheta \, p^2 \sin\vartheta \, \frac{1}{p^2} e^{ipr\cos\vartheta} \\ &= \frac{q}{r} \, . \end{aligned}$$

Hier haben wir die Details der Rechnung ausgelassen. Beim Übergang zur zweiten Zeile wurden für die Berechnung des Integrals Kugelkoordinaten $(p = |p|, \vartheta, \varphi)$ im $p$-Raum eingeführt. Dabei wurde die $p_z$-Achse parallel zur Richtung von $\underline{r}$ gewählt. Bei der Berechnung des Integrals wurde außerdem $\int_0^\infty \frac{\sin y}{y} dy = \frac{\pi}{2}$ verwendet.
Aus physikalischer Sicht ist das Ergebnis natürlich nicht überraschend, denn es handelt sich um das bekannte Coulomb-Potential einer Punktladung $q$ im Ursprung!

### 14.4.3 Green'sche Funktion

Wir wollen jetzt ein wichtiges Verfahren diskutieren, das zur Lösung allgemeiner inhomogener DGL mit gegebenen Randbedingungen, insbesondere aber auch von pDGL.
Wir wollen hier speziell den eindimensionalen Fall betrachten, und zwar für ein Sturm-Liouville-Problem vom Typ

$$\hat{D}y(x) = f(x) \qquad \text{mit } \; y(a) = y(b) = 0$$

und dem Sturm-Liouville-Operator

$$\hat{D}y = (p(x)y'(x))' + a(x)y \, .$$

Die Wahl der Randbedingungen ist keine einschneidende Einschränkung. Für den Fall $y(a), y(b) \neq 0$ löst man zuerst die homogene Gleichung $\hat{D}y_h = 0$ mit $y_h(a) = y_a$, $y_h(b) = y_b$. Ist $y(x)$ die Lösung des oben angegebenen Problems mit $y(a) = y(b) = 0$, so ist

$$\hat{y}(x) = y(x) + y_h(x)$$

die Lösung von $\hat{D}y(x) = f(x)$ mit $\hat{y}(a) = y_a$ und $\hat{y}(b) = y_b$.

Wir kommen nun zur Lösung von $\hat{D}y(x) = f(x)$ mit $y(a) = y(b) = 0$.

**Definition 14.4.1** (Green'sche Funktion)**.**
Die **Green'sche Funktion** (oder *Einflussfunktion*) der DGL $\hat{D}y(x) = f(x)$ mit $y(a) = y(b) = 0$ ist die Lösung $G(x, z)$ der Gleichung

$$\boxed{\begin{aligned} &\hat{D}G(x,z) = \delta(x-z) \\ &G(a,z) = G(b,z) = 0 \qquad (a < x, z < b)\,. \end{aligned}}$$

$G(x, z)$ stellt eine Art *Elementarlösung* dar, aus der sich die Lösung von $\hat{D}y(x) = f(x)$ mit $y(a) = y(b) = 0$ bestimmen läßt. Dies sieht man folgendermaßen: Zunächst gilt, nach Definition der Green'schen Funktion,

$$(pG')' + aG = \delta(x - z)\,.$$

Diese Gleichung multiplizieren wir mit $y(x)$ und subtrahieren hiervon die betrachtete Sturm-Liouville-Gleichung,

$$G(py')' + ayG = G(x, z)f(x)\,,$$

die wir noch mit $G(x, z)$ multipliziert haben. Dies ergibt zunächst

$$[p(yG' - y'G)]' = y\delta(x - z) - Gf$$

und nach Integration über $x$ über das Intervall $[a, b]$

$$p(yG' - y'G)\Big|_a^b = \int_a^b dx\, y(x)\delta(x - z) - \int_a^b dx\, G(x, z)f(x) = y(z) - \int_a^b dx\, G(x, z)f(x)\,.$$

Die linke Seite verschwindet auf Grund der Randbedingungen. Somit erhalten wir folgende Darstellung der Lösung des Sturm-Liouville-Problems $\hat{D}y(x) = f(x)$, $y(a) = y(b) = 0$ durch die Green'sche Funktion:

$$\boxed{y(x) = \int_a^b dx\, G(x, z)f(x)}$$

für $z \in ]a, b[$. Dies erlaubt uns, bei Kenntnis der Green'schen Funktion, die DLG für beliebige Inhomogenitäten $f(x)$ explizit zu lösen!

**Beispiel 14.4.1.** Als Beispiel betrachten wir die Poisson-Gleichung[4]

$$\triangle u = -4\pi\rho(x,y)\,.$$

Die Green'sche Funktion hatten wir im Prinzip schon in Kapitel 14.4.2 bestimmt, wenn man dort $q = 1$ setzt:

$$G(\underline{r},\underline{r}') = \frac{1}{|\underline{r}-\underline{r}'|}\,.$$

Somit lautet die allgemeine Lösung der Poisson-Gleichung

$$u(\underline{r}) = \int_{\mathbb{R}^3} d^3r\,\frac{\rho(\underline{r}')}{|\underline{r}-\underline{r}'|}\,.$$

### 14.4.4 Separation der Variablen

Ein Verfahren, mit dem sich in vielen Fällen Lösungen von pDGL bestimmen lassen, beruht auf einem **Separationsansatz**, z.B.

$$u(x,y,t) = X(x)Y(y)T(t)\,.$$

Wir haben dies bereits bei der Herleitung der Kugelflächenfunktionen in Kapitel 14.2.2 angewendet. Das dort beschriebene Vorgehen ist typisch für dieses Verfahren. Wir wollen uns daher auf ein paar allgemeine Bemerkungen beschränken.

- Ein Separationssatz bietet sich vor allem dann an, wenn der Rand durch Koordinatenlinien oder -flächen darstellbar ist. Daher ist beim Separationsansatz die Wahl geeigneter Koordinaten, in denen man die Separation durchführt, wichtig.
- Durch den Ansatz wird die pDGL in mehrere gewöhnliche DGL zerlegt, die oft die Form von Eigenwertproblemen haben. Diese DGL sind durch gemeinsame Konstanten miteinander verbunden.
- Die allgemeine Lösung erhält man als Linearkombination von Produkten aller Teillösungen.
- Randbedingungen schränken die Lösungen ein, z.B. legen sie die erlaubten Eigenwerte fest oder bestimmen die Koeffizienten der Linearkombinationen.
- Manchmal gibt es nur für bestimmte Parameterwerte der DGL zulässige Lösungen.

[4]Wir haben hier die Inhomogenität etwas anders normiert, um die Poisson-Gleichung in eine 'physikalische' Form zu bringen.

# Kapitel 15

# Koordinatensysteme

## 15.1 Gebräuchliche Koordinatensysteme

Bisher haben wir im $\mathbb{R}^3$ die kartesischen Koordinaten $x, y, z$ (bzw. $x_1, x_2, x_3$), die Zylinderkoordinaten $\rho, \varphi, z$ und die Kugelkoordinaten $r, \varphi, \vartheta$ kennengelernt. Daneben gibt es noch viele weitere Koordinaten, die in einzelnen Fällen hilfreich sein können!
Streng genommen haben wir aber meist nicht wirklich Zylinder- oder Kugelkoordinaten verwendet, d.h. Vektoren in der *Basis* dieser Koordinatensysteme dargestellt, sondern nur die *Komponenten* der kartesischen Koordinaten mit Hilfe der Variablen $\rho, \varphi, z$ bzw. $r, \varphi, \vartheta$ *parametrisiert*.
Wir wollen in diesem Kapitel ganz allgemein herleiten, wie sich Größen in unterschiedlichen Koordinatensystemen beschreiben lassen. Wir konzentrieren uns auf den Fall des $\mathbb{R}^3$, aber alle Überlegungen lassen sich in offensichtlicher Weise auf beliebige Dimensionen verallgemeinern.
Bei unseren Betrachtungen werden uns die kartesischen Koordinaten als Referenzsystem dienen. Zur Vereinfachung bezeichnen wir sie im folgenden mit $x_1, x_2, x_3$ statt $x, y, z$, d.h. der Ortsvektor ist durch $\underline{r} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$ gegeben.
Es seien nun $u_1, u_2, u_3$ beliebige andere Koordinaten. Wir stellen die kartesischen Koordinaten als Funktion dieser neuen Koordinaten dar:

$$\begin{aligned} x_1 &= x_1(u_1, u_2, u_3), \\ x_2 &= x_2(u_1, u_2, u_3), \\ x_3 &= x_3(u_1, u_2, u_3), \end{aligned}$$

wobei wir den Funktionsnamen gleich der entsprechenden kartesischen Komponente gewählt haben! Dieses Gleichungssystem ist nach den $u_j$ auflösbar, falls für die Jacobi- (bzw. Funktional-) Determinante (vgl. Kap. 8.4) gilt:

$$\frac{\partial(x_1, x_2, x_3)}{\partial(u_1, u_2, u_3)} := \det \left( \frac{\partial x_j}{\partial u_l} \right)_{jl} \neq 0.$$

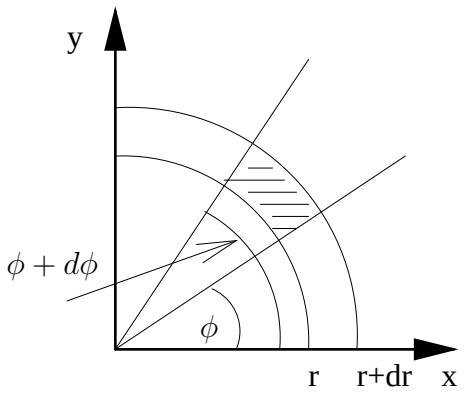



Abbildung 15.1.1: Koordinatenlinien in ebenen Polarkoordinaten. Die $r$-Linien sind radiale Gerade mit Winkel $\varphi_0$ und die $\varphi$-Linien Kreise vom Radius $r_0$.

Wir betrachten nun die **Koordinatenlinien** (vgl. Kap. 8.4), die man erhält wenn man alle Koordinaten bis auf eine fixiert und diese freie Koordinate alle erlaubten Werte durchlaufen lässt. Die Koordinatenlinien durch den festen Punkt $\underline{r}^{(0)} = \underline{r}\left(u_1^{(0)}, u_2^{(0)}, u_3^{(0)}\right)$ nennen wir

$$\underline{r}_1(u_1) = \underline{r}(u_1, u_2^{(0)}, u_3^{(0)}), \quad \underline{r}_2(u_1) = \underline{r}(u_1^{(0)}, u_2, u_3^{(0)}), \quad \underline{r}_3(u_1) = \underline{r}(u_1^{(0)}, u_2^{(0)}, u_3).$$

Schneiden sich die Koordinatenlinien paarweise in einem rechten Winkel, so spricht man von **rechtwinkligen** oder **orthogonalen Koordinaten**.
In kartesischen Koordinaten sind die Koordinatenlinien Geraden parallel zur korrespondierenden Achse, z.B. ist $\underline{r}_1(x_1) = \underline{r}(x_1, x_2^{(0)}, x_3^{(0)})$ eine Parallele zur $x$-Achse.
Abb. 15.1.1 zeigt die Koordinatenlinien $\underline{r}_1(r) = \underline{r}(r, \varphi_0)$ ($r$-Linien) und $\underline{r}_2(r) = \underline{r}(r_0, \varphi)$ ($\varphi$-Linien) der ebenen Polarkoordinaten.
Die Basisvektoren der kartesischen Koordinaten sind die Einheitsvektoren $\underline{e}_1, \underline{e}_2, \underline{e}_3$ in Richtung der Koordinatenachsen. Sie stimmen mit den Tangentenvektoren an die Koordinatenlinien überein! Dies wollen wir auf beliebige Koordinaten übertragen und definieren daher als Basisvektoren $\underline{e}_{u_j}$ der Koordinaten $u_1, u_2, u_3$

$$\boxed{\underline{e}_{u_j} := \frac{1}{h_{u_j}} \frac{\partial \underline{r}}{\partial u_j}\bigg|_{\underline{r}=\underline{r}_0} \qquad \text{mit} \quad h_{u_j} = \left|\frac{\partial \underline{r}}{\partial u_j}\right|.}$$

Damit ist $\underline{e}_{u_j}$ also ein normierter Tangentenvektor $\underline{T}_j$ an die $u_j$-Linie durch $\underline{r}_0$. Man beachte, dass daher die $\underline{e}_{u_j}$ noch vom Punkt $\underline{r}_0$ abhängen können! Die Normierungsfaktoren $h_{u_j}$ werden wir später explizit benötigen!
Für orthogonale Koordinaten, auf die wir uns hier beschränken wollen, gilt

$$\underline{e}_{u_j} \cdot \underline{e}_{u_l} = \delta_{jl}$$

Daher bilden die $\{\underline{e}_{u_j}\}$ eine Orthonormalbasis des $\mathbb{R}^3$. I.a. benutzt man die Konvention, dass dieses System rechtshändig ist, d.h. es gilt

$$\underline{e}_{u_1} = \underline{e}_{u_2} \times \underline{e}_{u_3} \qquad + \text{zyklische Indexvertauschungen}.$$

Under construction !

Abbildung 15.1.2: Die Richtung der Basisvektoren $\underline{e}_{u_j}$ ändert sich i.a. mit der Position.

An dieser Stelle sei nochmals betont, dass sich i.a. die Richtung der Basisvektoren $\underline{e}_{u_j}$, und damit die Orientierung der Basis, als Funktion von $\underline{r}_0$ ändert (siehe Abb. 15.1.2). Man mache sich das am Beispiel der ebenen Polarkoordinaten klar (siehe Abb. 15.1.1). Der Basisvektor $\underline{e}_r$ zeigt als Tangentenvektor an die $r$-Linie immer in radiale Richtung, die aber mit $\varphi_0$ variiert. Die kartesischen Koordinaten bilden hier eine Ausnahme, da die Koordinatenlinien durch verschiedenen Punkte parallel sind!

Wir wollen diese allgemeinen Betrachtungen auf die bereits bekannten Fälle der Zylinder- und Kugelkoordinaten spezialisieren. Die Ergebnisse werden wir dann verwenden, um z.B. die Form der Differentialoperatoren in diesen Systemen herzuleiten.

### 15.1.1 Zylinderkoordinaten

Wir hatten die Zylinderkoordinaten in Kapitel 1.4.3 eingeführt. Präziser gesagt, haben wir dort nur die 'Zylindervariablen' $\rho, \varphi, x_3$ definiert und mit ihnen die kartesischen Koordinaten parametrisiert. In unserer jetzigen Notation (mit $u_1 = \rho$, $u_2 = \varphi$ und $u_3 = x_3$) lauten die dort abgeleiten Beziehungen:

$$\underline{r} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} \rho\cos\varphi \\ \rho\sin\varphi \\ x_3 \end{pmatrix} \qquad \text{mit} \qquad \begin{array}{rcll} \rho & = & \sqrt{x_1^2 + x_2^2} & \in [0, \infty[ \\ \varphi & = & \arctan\frac{x_2}{x_1} & \in [0, 2\pi[ \\ x_3 & \in & \mathbb{R} & \end{array} .$$

Die Koordinatenlinien lassen sich analog zum oben betrachteten Fall der ebenen Polarkoordinaten leicht bestimmen. Die $\rho$-Linien

$$\underline{r}_1(\rho) = \begin{pmatrix} \rho\cos\varphi_0 \\ \rho\sin\varphi_0 \\ x_3^{(0)} \end{pmatrix}$$

sind radiale Halbgeraden in Höhe $x_3^{(0)}$ im Winkel $\varphi_0$. Die $\varphi$-Linien sind Kreise um den Ursprung und die $x_3$-Linien sind Parallelen zur $z$-Achse.

Die Basisvektoren der Zylinderkoordinaten sind daher gegeben durch

$$\boxed{\begin{aligned}
\underline{e}_\rho &= \frac{1}{h_\rho}\frac{\partial \underline{r}}{\partial \rho} = \begin{pmatrix}\cos\varphi\\ \sin\varphi\\ 0\end{pmatrix}, & h_\rho &= 1,\\
\underline{e}_\varphi &= \frac{1}{h_\varphi}\frac{\partial \underline{r}}{\partial \varphi} = \begin{pmatrix}-\sin\varphi\\ \cos\varphi\\ 0\end{pmatrix}, & h_\varphi &= \rho,\\
\underline{e}_{x_3} &= \frac{1}{h_{x_3}}\frac{\partial \underline{r}}{\partial x_3} = \begin{pmatrix}0\\ 0\\ 1\end{pmatrix} = \underline{e}_3, & h_{x_3} &= 1.
\end{aligned}}$$

Man beachte, dass wir hier die Basisvektoren durch ihre Darstellung in kartesischen Koordinaten (als 3-komponentige Vektoren) angegeben haben.
Da

$$\underline{e}_\rho \times \underline{e}_\varphi = \underline{e}_{x_3},$$

wie man leicht nachrechnet, bilden $(\underline{e}_\rho, \underline{e}_\varphi, \underline{e}_{x_3})$ ein Rechtssystem.

### 15.1.2 Kugelkoordinaten

Die Kugelkoordinaten $\rho, \vartheta.\varphi$ hatten wir ebenfalls in Kapitel 1.4.4 kennengelernt[1]. Auch sie hatten wir bisher im Wesentlichen zur Parametrisierung der kartesischen Komponenten benutzt:

$$\underline{r} = \begin{pmatrix}x_1\\ x_2\\ x_3\end{pmatrix} = \begin{pmatrix}r\sin\vartheta\cos\varphi\\ r\sin\vartheta\sin\varphi\\ r\cos\vartheta\end{pmatrix} \qquad \text{mit} \qquad \begin{aligned}
\rho &= \sqrt{x_1^2+x_2^2+x_3^2} && \in [0,\infty[\\
\vartheta &= \arctan\frac{\sqrt{x_1^2+x_2^2+x_3^2}}{x_3} && \in [0,\pi]\\
\varphi &= \arctan\tfrac{x_2}{x_1} && \in [0,2\pi[
\end{aligned}$$

Wir können nun die Basisvektoren der Kugelkoordinaten im kartesischen System angeben:

$$\boxed{\begin{aligned}
\underline{e}_r &= \frac{1}{h_r}\frac{\partial \underline{r}}{\partial r} = \begin{pmatrix}\sin\vartheta\cos\varphi\\ \sin\vartheta\sin\varphi\\ \cos\vartheta\end{pmatrix}, & h_r &= 1,\\
\underline{e}_\vartheta &= \frac{1}{h_\vartheta}\frac{\partial \underline{r}}{\partial \vartheta} = \begin{pmatrix}\cos\vartheta\cos\varphi\\ \cos\vartheta\sin\varphi\\ -\sin\vartheta\end{pmatrix}, & h_\vartheta &= r,\\
\underline{e}_\varphi &= \frac{1}{h_\varphi}\frac{\partial \underline{r}}{\partial \varphi} = \begin{pmatrix}-\sin\varphi\\ \cos\varphi\\ 0\end{pmatrix}, & h_\varphi &= r\sin\vartheta.
\end{aligned}}$$

Da

$$\underline{e}_r \times \underline{e}_\vartheta = \underline{e}_\varphi,$$

bilden $(\underline{e}_r, \underline{e}_\vartheta, \underline{e}_\varphi)$ ein Rechtssystem.

[1] Siehe auch Aufgabe 3 der Übungen.

## 15.2 Bestimmung von Vektorkomponenten

**Definition 15.2.1** ((Vektor-)Komponenten)**.**
Da die Vektoren $\{\underline{e}_{u_j}\}$ eine Basis bilden, kann man beliebige Vektoren $\underline{F}$ in ihr darstellen:

$$\underline{F} = \sum_j F_{u_j} \underline{e}_{u_j}.$$

Die Zahlen $F_{u_j}$ nennt man die **Komponenten** von $\underline{F}$ bzgl. der Basis $\{\underline{e}_{u_j}\}$.

Für orthogonale Koordinaten ist $\{\underline{e}_{u_j}\}$ sogar eine Orthogonalbasis. Dann können wir die Komponente $F_{u_l}$ einfach als Projektion von $\underline{F}$ auf die Richtung $\underline{e}_{u_l}$ (die Richtung der $u_l$-Koordinatenlinie) berechnen:

$$\underline{F} \cdot \underline{e}_{u_l} = \sum_j F_{u_j} \underline{e}_{u_j} \cdot \underline{e}_{u_l} = \sum_j F_{u_j} \delta_{jl} = F_{u_l}.$$

Die kartesischen Komponenten $F_j$, d.h. die Komponenten bzgl. der Basis $\{e_x, e_y, e_z\}$, sind offensichtlich ein Spezialfall der obigen Definition. Meist sind Vektoren in kartesischen Koordinaten gegeben:

$$\underline{F} = \sum_j F_j \underline{e}_j,$$

wobei wir vereinfachend $\underline{e}_j$ für $\underline{e}_{x_j}$ schreiben. Die kartesischen Komponenten fasst man dann auch als Spaltenvektor $\underline{F} = \begin{pmatrix} F_1 \\ F_2 \\ F_3 \end{pmatrix}$ zusammen.

**Beispiel 15.2.1.**

1. Die Komponenten des Ortsvektors $\underline{r}$ in Zylinderkoordinaten sind gegeben durch:

$$\begin{aligned} r_\rho &= \underline{r} \cdot \underline{e}_\rho = \cos\varphi(\rho\cos\varphi) + \sin\varphi(\rho\sin\varphi) = \rho(\cos^2\varphi + \sin^2\varphi) = \rho, \\ r_\varphi &= \underline{r} \cdot \underline{e}_\varphi = (-\sin\varphi)\rho\cos\varphi + \cos\varphi(\rho\sin\varphi) = 0, \\ r_{x_3} &= \underline{r} \cdot \underline{e}_{x_3} = x_3. \end{aligned}$$

   Dabei haben wir die explizite Form der Basisvektoren $\underline{e}_\rho, \underline{e}_\varphi, \underline{e}_{x_3}$ ausgenutzt, die wir in Kapitel 15.1.1 angegeben hatten. Somit erhalten wir also die Darstellung von $\underline{r}$ in Zylinderkoordinaten als

$$\underline{r} = \rho\underline{e}_\rho + x_3\underline{e}_{x_3}.$$

2. Analog können wir nun mit Hilfe der Ergebnisse aus Kapitel 15.1.2 den Ortsvektor in Kugelkoordinaten darstellen. Man sieht schnell, dass $\underline{r} \cdot \underline{e}_\varphi = 0 = \underline{r} \cdot \underline{e}_\vartheta$ ist und so ergibt sich

$$\underline{r} = r\underline{e}_r.$$

   Die Darstellung in Zylinder- und insbesondere in Kugelkoordinaten ist also viel einfacher, als die Darstellung $\underline{r} = x_1\underline{e}_{x_1} + x_2\underline{e}_{x_2} + x_3\underline{e}_{x_3}$ in kartesischen Koordinaten!

3. Abschließend wollen wir noch den Vektor $\underline{F} = x_3\underline{e}_{x_1} = \begin{pmatrix} x_3 \\ 0 \\ 0 \end{pmatrix}$ in Kugelkoordinaten darstellen.
   Mit $x_3 = r\cos\vartheta$ folgt:

$$\begin{aligned} F_r &= \underline{F}\cdot\underline{e}_r = \begin{pmatrix} r\cos\vartheta \\ 0 \\ 0 \end{pmatrix}\cdot\begin{pmatrix} \sin\vartheta\cos\varphi \\ \sin\vartheta\sin\varphi \\ \cos\vartheta \end{pmatrix} = r\sin\vartheta\cos\vartheta\cos\varphi, \\ F_\vartheta &= \underline{r}\cdot\underline{e}_\vartheta = r\cos^2\vartheta\cos\varphi, \\ F_\varphi &= \underline{r}\cdot\underline{e}_\varphi = -r\cos\vartheta\sin\varphi. \end{aligned}$$

   Die Darstellung von $\underline{F}$ in Kugelkoordinaten ist also wesentlich komplizierter als die in kartesischen Koordinaten.

Als Nächstes wollen wir uns mit dem Differenzieren von Vektoren beschäftigen. Da die Koordinatenlinien i.a. keine Geraden sind (außer im kartesischen Fall), ändern die Basisvektoren ihre Richtung und müssen daher auch differenziert werden, nicht nur die Komponenten, wie im kartesischen Fall. Bei Ableitung nach den Koordinaten $u_l$ gilt daher:

$$\frac{\partial\underline{F}}{\partial u_j} = \sum_l \left(\frac{\partial F_{u_l}}{\partial u_j}\underline{e}_{u_l} + F_{u_l}\frac{\partial\underline{e}_{u_l}}{\partial u_j}\right).$$

Bei der Ableitung nach anderen Parametern (z.B. der Zeit) geht man unter Berücksichtigung der Kettenregel analog vor.

**Beispiel 15.2.2.** Wir wollen die Geschwindigkeit $\underline{v}(t)$ aus $\underline{r}(t)$ bestimmen.
In kartesischen Koordinaten sind die Basisvektoren $\underline{e}_j$ konstant und wir haben die bekannte Formel

$$\underline{v}(t) = \dot{\underline{r}}(t) = \sum_j \dot{x}_j(t)\underline{e}_j,$$

d.h. wir erhalten die Komponenten der Geschwindigkeit durch Zeitableitung der Komponenten von $\underline{r}$.
In Zylinderkoordinaten ist das nicht so! Wir verwenden die in Beispiel 15.2.1 abgeleitete Darstellung des Ortsvektors in Zylinderkoordinaten und erhalten zunächst

$$\begin{aligned} \underline{v}(t) &= \dot{\underline{r}}(t) = \frac{d}{dt}\left(\rho\underline{e}_\rho + x_3\underline{e}_{x_3}\right) \\ &= \dot\rho\underline{e}_\rho + \rho\dot{\underline{e}}_\rho + \dot{x}_3\underline{e}_{x_3} + x_3\dot{\underline{e}}_{x_3}. \end{aligned}$$

Wir benötigen nun die Zeitableitungen der Basisvektoren. Diese können wir uns aus den Ergebnissen in Kapitel 15.1.1 beschaffen, denn es gilt z.B.

$$\underline{e}_\rho = \frac{d}{dt}\begin{pmatrix} \cos\varphi \\ \sin\varphi \\ 0 \end{pmatrix} = \begin{pmatrix} -\dot\varphi\sin\varphi \\ \dot\varphi\cos\varphi \\ 0 \end{pmatrix} = \dot\varphi\underline{e}_\varphi\,.$$

Offensichtlich gilt $\dot{\underline{e}}_{x_3} = 0$ und somit erhalten wir für die Geschwindigkeit

$$\underline{v}(t) = \dot{\rho}\underline{e}_\rho + \rho\dot{\varphi}\underline{e}_\varphi + \dot{x_3}\underline{e}_{x_3}.$$

Dieses Ergebnis kann man alternativ auch folgendermaßen herleiten: Zunächst drückt man $\underline{r}$ im kartesischen System mit Hilfe der Zylindervariablen aus:

$$\underline{r} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} \rho\cos\varphi \\ \rho\sin\varphi \\ x_3 \end{pmatrix}.$$

Hieraus folgt

$$\dot{\underline{r}} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} \dot{\rho}\cos\varphi - \rho\dot{\varphi}\sin\varphi \\ \dot{\rho}\sin\varphi + \rho\dot{\varphi}\cos\varphi \\ \dot{x}_3 \end{pmatrix}.$$

Die Projektionen $\dot{\underline{r}} \cdot \underline{e}_\rho$ etc. ins Basissystem der Zylinderkoordinaten liefern dann wieder $\underline{v}(t) = \dot{\rho}\underline{e}_\rho + \rho\varphi\underline{e}_\varphi + \dot{x_3}\underline{e}_{x_3}$.

Bei der obigen Rechnung haben wir die Ableitungen der Basisvektoren benötigt. Allgemein ist folgendes Ergebnis nützlich:
Es sei $\underline{e}(t)$ ein beliebiger Einheitsvektor, dessen Richtung von der Zeit abhängt. Dann gilt wegen $1 = \underline{e}(t) \cdot \underline{e}(t) = e^2(t)$ (wobei $e(t) = |\underline{e}(t)|$ den Betrag des Vektors bezeichnet)

$$0 = \frac{d}{dt}(\underline{e}(t) \cdot \underline{e}(t)) = 2\underline{e}(t) \cdot \dot{\underline{e}}(t)$$

und somit ist

$$\dot{\underline{e}}(t) \perp \underline{e}(t).$$

Die Ableitung eines Einheitsvektors steht immer senkrecht auf diesem!
In obigem Beispiel wissen wir daher, dass sich z.B. $\dot{\underline{e}}_\rho$ darstellen lassen muß als $\dot{\underline{e}}_\rho = a\dot{\underline{e}}_\varphi + b\dot{\underline{e}}_{x_3}$ mit Koeffizienten $a, b$. Diese lassen sich manchmal durch andere Überlegungen bestimmen.

*Bemerkung.*

- Es besteht ein Unterschied zwischen der bloßen Verwendung krummliniger Koordinaten (genauer: Variablen) und der Darstellung eines Vektors in einem solchen Basissystem!
- Die Spezifikation eines Vektors durch seine Komponenten macht nur bei Angabe des Basissystems Sinn! Wir wir an den Beispielen gesehen haben, sind die Komponenten eines Vektors in kartesischen Koordinaten als die in Kugelkoordinaten.

**Beispiel 15.2.3.** Wir wollen diese wichtigen Bemerkungen noch einmal an Hand eines Beispiels diskutieren. Dazu betrachten wir den Vektor $\underline{a} = 5x_1\underline{e}_1 + 5x_2\underline{e}_2 + 5x_3\underline{e}_3$ in der kartesischen Basis $\{\underline{e}_1, \underline{e}_2, \underline{e}_3\}$. Dort hat der die Komponenten $(5x_1, 5x_2, 5x_3)$. Wir können nun diese Komponenten (!) auch in Kugelvariablen ausdrücken als $(5r\sin\vartheta\cos\varphi, 5r\sin\vartheta\sin\varphi, 5r\cos\vartheta)$. Man beachte:

Dies sind immer noch die *kartesischen* Komponenten, nun aber parametrisiert durch die Variablen der Kugelkoordinaten!
In der Basis $\{\underline{e}_r, \underline{e}_\vartheta, \underline{e}_\varphi\}$ der Kugelkoordinaten (und in den entsprechenden Variablen ausgedrückt) haben wir aber $\underline{a} = 5r\underline{e}_r$ mit den Komponenten $(5r, 0, 0)$. Diese können wir auch noch mit den kartesischen Variablen ausdrücken als $(5\sqrt{x_1^2 + x_2^2 + x_3^2}, 0, 0)$.

*Bemerkung.* Wenn nicht anders angegeben, sind mit "Komponenten" i.a. die kartesischen Komponenten gemeint, z.B. in der komponentenweisen Definition eines Vektors.

## 15.3 Differentialoperatoren in krummlinigen Koordinaten

Als nächstes wollen wir die Frage klären, wie die Differentialoperatoren Gradient, Rotation, Divergenz und der Laplace-Operator in anderen orthogonalen Basissystemen aussehen.

### 15.3.1 Gradient

Es sei $\phi(\underline{r})$ ein Skalarfeld. Die Komponente des Gradienten $\operatorname{grad}\phi$ in $\underline{e}_{u_j}$-Richtung ist gegeben durch

$$\begin{aligned}(\operatorname{grad}\phi)_{u_j} &= (\operatorname{grad}\phi)\cdot\underline{e}_{u_j} = (\operatorname{grad}\phi)\cdot\left(\frac{1}{h_{u_j}}\frac{\partial\underline{r}}{\partial u_j}\right)\\ &= \frac{1}{h_{u_j}}\sum_l \frac{\partial\phi}{\partial x_l}\frac{\partial x_l}{\partial u_j} = \frac{1}{h_{u_j}}\frac{\partial\phi}{\partial u_j},\end{aligned}$$

und somit gilt im neuen Basissystem

$$\boxed{\operatorname{grad} = \sum_j \underline{e}_{u_j}\frac{1}{h_{u_j}}\frac{\partial}{\partial u_j}}\,.$$

Speziell für die uns bekannten Koordinaten erhalten wir hieraus explizit

| | |
|---|---|
| kartesisch: | $\operatorname{grad}\phi(x_1,x_2,x_3) = \sum_j \underline{e}_j\,\frac{\partial\phi}{\partial x_j}\,,$ |
| Zylinder: | $\operatorname{grad}\phi(\rho,\varphi,x_3) = \left[\underline{e}_\rho\,\frac{\partial}{\partial\rho} + \frac{1}{\rho}\,\underline{e}_\varphi\,\frac{\partial}{\partial\varphi} + \underline{e}_{x_3}\,\frac{\partial}{\partial x_3}\right]\phi(\rho,\varphi,x_3)\,,$ |
| Kugel: | $\operatorname{grad}\phi(r,\vartheta,\varphi) = \left[\underline{e}_r\,\frac{\partial}{\partial r} + \frac{1}{r}\,\underline{e}_\vartheta\,\frac{\partial}{\partial\vartheta} + \frac{1}{r\sin\vartheta}\,\underline{e}_\varphi\,\frac{\partial}{\partial\varphi}\right]\phi(r,\vartheta,\varphi)\,.$ |

### 15.3.2 Divergenz

Aus dem allgemeinen Ergebnis für den Gradienten angewandt auf die Koordinaten selbst folgt zunächst eine Darstellung der Basisvektoren durch den Gradienten:

$$h_{u_j} \operatorname{grad} u_j = h_{u_j} \sum_l \underline{e}_{u_l} \frac{1}{h_{u_l}} \frac{\partial u_j}{\partial u_l} = h_{u_j} \sum_l \underline{e}_{u_l} \frac{1}{h_{u_l}} \delta_{jl} = \underline{e}_{u_j} .$$

Wir betrachten nun ein Skalarfeld $\underline{A} = \sum_j A_{u_j}(u_1, u_2, u_3)\underline{e}_{u_j}$, dessen Komponenten $A_{u_j}$ in den krummlinigen Koordinaten bekannt sind.
Zur Vereinfachung untersuchen wir zunächst nur die Wirkung der Divergenz auf den ersten Term $A_{u_1}\underline{e}_{u_1}$:

$$\begin{aligned}
\operatorname{div}\left(A_{u_1}\underline{e}_{u_1}\right) &= \operatorname{div}\left(\left(\underline{e}_{u_2} \times \underline{e}_{u_3}\right) A_{u_1}\right) \\
&= \operatorname{div}\left(h_{u_2} h_{u_3} A_{u_1} \left(\operatorname{grad} u_2 \times \operatorname{grad} u_3\right)\right) \\
&= \operatorname{grad}(h_{u_2} h_{u_3} A_{u_1}) \cdot (\operatorname{grad} u_2 \times \operatorname{grad} u_3) + h_{u_2} h_{u_3} A_{u_1} \operatorname{div}\left(\operatorname{grad} u_2 \times \operatorname{grad} u_3\right) \\
&= \frac{1}{h_{u_2} h_{u_3}} \underline{e}_{u_1} \cdot \operatorname{grad}(h_{u_2} h_{u_3} A_{u_1}) \\
&= \frac{1}{h_{u_2} h_{u_3}} \frac{\partial (h_{u_2} h_{u_3} A_{u_1})}{\partial u_1}.
\end{aligned}$$

Dabei haben wir im ersten Schritt ausgenutzt, dass die $\underline{e}_{u_j}$ ein rechtshändiges Orthonormalsystem bilden, und im zweiten Schritt die oben angegebene Darstellung der Basisvektoren durch den Gradienten. Im dritten Schritt wurde die allgemeine Produktregel

$$\operatorname{div}(\phi \underline{A}) = \underline{A} \cdot \operatorname{grad} \phi + \phi \operatorname{div} \underline{A},$$

die für beliebige Skalarfelder $\phi$ und Vektorfelder $\underline{A}$ gilt, benutzt. Eine ähnliche Identität für die Divergenz eines Kreuzproduktes wurde beim Übergang zur 4. Zeile ausgenutzt:

$$\operatorname{div}(\underline{A} \times \underline{B}) = \underline{B} \cdot \operatorname{rot} \underline{A} - \underline{A} \cdot \operatorname{rot} \underline{B} .$$

Speziell für $\underline{A} = \operatorname{grad} u_2$ und $\underline{B} = \operatorname{grad} u_3$ folgt wegen $\operatorname{rot}(\operatorname{grad} \phi) = 0$ (siehe Aufgabe 15):

$$\operatorname{div}(\operatorname{grad} u_2 \times \operatorname{grad} u_3) = 0.$$

Die Rechnung für die anderen Terme geht analog. Als Endergebnis für die Divergenz in einem beliebigen orthogonalen Koordinatensystem erhalten wir daher

$$\boxed{\operatorname{div} \underline{A}(u_1, u_2, u_3) = \frac{1}{h_{u_1} h_{u_2} h_{u_3}} \left[ \frac{\partial}{\partial u_1}(h_{u_2} h_{u_3} A_{u_1}) + \frac{\partial}{\partial u_2}(h_{u_1} h_{u_3} A_{u_2}) + \frac{\partial}{\partial u_3}(h_{u_1} h_{u_2} A_{u_3}) \right] .}$$

Speziell für die uns bekannten Koordinaten erhalten wir hieraus explizit

| | |
|---|---|
| kartesisch: | $\operatorname{div} \underline{A} = \sum_j \dfrac{\partial A_j}{\partial x_j}\,,$ |
| Zylinder: | $\operatorname{div} \underline{A}(\rho,\varphi,x_3) = \dfrac{1}{\rho}\dfrac{\partial(\rho A_\rho)}{\partial\rho} + \dfrac{1}{\rho}\dfrac{\partial A_\varphi}{\partial\varphi} + \dfrac{\partial A_{x_3}}{\partial x_3}\,,$ |
| Kugel: | $\operatorname{div} \underline{A}(r,\vartheta,\varphi) = \dfrac{1}{r^2}\dfrac{\partial(r^2 A_r)}{\partial r} + \dfrac{1}{r\sin\vartheta}\dfrac{\partial(\sin\vartheta A_\vartheta)}{\partial\varphi} + \dfrac{1}{r\sin\vartheta}\dfrac{\partial A_\varphi}{\partial\varphi}\,.$ |

Dabei haben wir das Vektorfeld jeweils in der entsprechenden Basis dargestellt, z.B. für Zylinderkoordinaten als $\underline{A}(\rho,\varphi,x_3) = A_\rho \underline{e}_\rho + A_\varphi \underline{e}_\varphi + A_{x_3}\underline{e}_{x_3}$.

**Beispiel 15.3.1.** Als Beispiel für die Nützlichkeit dieser Ergebnisse wollen wir die Divergenz des Vektorfeldes $\underline{F} = \frac{1}{r^2}\underline{e}_r$ berechnen, das also in der Basis der Kugelkoordinaten gegeben ist. Es beschreibt z.B. die Gravitationskraft zwischen zwei Punktmassen.
Die Komponenten von $\underline{F}$ in Kugelkoordinaten sind $F_r = \frac{1}{r^2}$ und $F_\vartheta = F_\varphi = 0$. Somit gilt

$$\operatorname{div} \underline{F} = \frac{1}{r^2}\frac{\partial(r^2 F_r)}{\partial r} = 0\,.$$

### 15.3.3 Rotation

Wir betrachten zunächst nur den Anteil $A_{u_1}\underline{e}_{u_1}$ des Vektorfeldes $\underline{A} = \sum_j A_{u_j}\underline{e}_{u_j}$ und gehen ähnlich vor wie bei der Berechnung der Divergenz:

$$\begin{aligned}
\operatorname{rot}(A_{u_1}\underline{e}_{u_1}) &= \operatorname{rot}(h_{u_1} A_{u_1} \operatorname{grad} u_1) \\
&= h_{u_1} A_{u_1} \operatorname{rot}(\operatorname{grad} u_1) + (\operatorname{grad}(h_{u_1} A_{u_1})) \times \operatorname{grad} u_1 \\
&= -\operatorname{grad} u_1 \times (\operatorname{grad}(h_{u_1} A_{u_1})) \\
&= -\frac{1}{h_{u_1}}\,\underline{e}_{u_1} \times \left( \sum_l \frac{1}{h_{u_l}}\,\frac{\partial(h_1 A_{u_1})}{\partial u_l}\,\underline{e}_{u_l} \right) \\
&= \frac{1}{h_{u_1} h_{u_3}}\,\frac{\partial(h_1 A_{u_1})}{\partial u_3}\,\underline{e}_{u_2} - \frac{1}{h_{u_1} h_{u_2}}\,\frac{\partial(h_1 A_{u_1})}{\partial u_2}\,\underline{e}_{u_3}\,.
\end{aligned}$$

Dabei haben wir die Identität

$$\operatorname{rot}(\phi \underline{A}) = \phi \operatorname{rot} \underline{A} + \operatorname{grad}\phi \times \underline{A}$$

benutzt, die für beliebige Skalarfelder $\phi$ und Vektorfelder $\underline{A}$ gilt.
Die Rechnung für die anderen Beiträge geht analog. Insgesamt kann man die Rotation kompakt in folgender Form als Determinante schreiben:

$$\operatorname{rot} A(u_1,u_2,u_3) = \begin{vmatrix} h_{u_1}\underline{e}_{u_1} & h_{u_2}\underline{e}_{u_2} & h_{u_3}\underline{e}_{u_3} \\ \partial/\partial u_1 & \partial/\partial u_2 & \partial/\partial u_3 \\ h_{u_1}A_{u_1} & h_{u_2}A_{u_2} & h_{u_3}A_{u_3} \end{vmatrix}\,.$$

Diese Regel ist analog zur bekannten Merkregel für das Kreuzprodukt zu interpretieren: Man entwickelt die Determinante formal nach der ersten Zeile, die ja Vektoren als Elemente enthält, und erhält so die Komponenten der Rotation. Bei der Entwicklung ist zu beachten, dass man die Reihenfolge der Faktoren nicht vertauschen darf, da die zweite Zeile Differentialoperatoren enthält, die auf die entsprechenden Elemente der dritten Zeile wirken.
Alernativ können wir das allgemeine Ergebnis auch mit Hilfe des Levi-Cevità-Symbols $\epsilon_{ijk}$ (vgl. Auf. 7) schreiben als

$$(\mathrm{rot}\, A(u_1,u_2,u_3))_{u_i} = \sum_{j,k} \epsilon_{ijk} h_{u_i} \frac{\partial (h_{u_k} A_{u_k})}{\partial u_j}.$$

Dabei ist dann $\mathrm{rot}\, A = \sum_i (\mathrm{rot}\, A)_{u_i} \underline{e}_{u_i}$.

Wir spezialisieren dieses allgemeine Resultat wieder auf die wichtigsten Fälle:

kartesisch:
$$\left(\mathrm{rot}\, \underline{A}(x_1,x_2,x_3)\right)_i = \sum_{j,k} \epsilon_{ijk} \frac{\partial A_k}{\partial x_j}\,,$$

Zylinder:
$$\mathrm{rot}\, \underline{A}(\rho,\varphi,x_3) = \left(\frac{1}{\rho}\frac{\partial A_{x_3}}{\partial \varphi} - \frac{\partial A_\varphi}{\partial x_3}\right)\underline{e}_\rho + \left(\frac{\partial A_\rho}{\partial x_3} - \frac{\partial A_{x_3}}{\partial \rho}\right)\underline{e}_\varphi + \frac{1}{\rho}\left(\frac{\partial(\rho A_\varphi)}{\partial \rho} - \frac{\partial A_\rho}{\partial \varphi}\right)\underline{e}_{x_3}\,,$$

Kugel:
$$\mathrm{rot}\, \underline{A}(r,\vartheta,\varphi) = \frac{1}{r\sin\vartheta}\left(\frac{\partial(\sin\vartheta A_\varphi)}{\partial\vartheta} - \frac{\partial A_\vartheta}{\partial\varphi}\right)\underline{e}_r + \left(\frac{1}{r\sin\vartheta}\frac{\partial A_r}{\partial\varphi} - \frac{1}{r}\frac{\partial(rA_\varphi)}{\partial r}\right)\underline{e}_\vartheta + \frac{1}{r}\left(\frac{\partial(rA_\vartheta)}{\partial r} - \frac{1}{r}\frac{\partial A_r}{\partial\vartheta}\right)\underline{e}_\varphi\,.$$

Dabei bezeichnet wieder $\epsilon_{ijk}$ das Levi-Cevità-Symbol.

### 15.3.4 Laplace-Operator

Abschließend wollen wir noch die Ergebnisse für den Laplace-Operator ohne Rechnung angeben. Im Prinzip kann man sie aus den bereits abgeleiteten Darstellungen des Gradienten und der Divergenz über $\triangle\phi = \mathrm{div}(\mathrm{grad}\,\phi)$ gewinnen.
Das allgemeine Resultat ist

$$\begin{aligned}\triangle\phi(u_1,u_2,u_3) &= \mathrm{div}\,(\mathrm{grad}\,\phi(u_1,u_2,u_3)) \\ &= \frac{1}{h_{u_1}h_{u_2}h_{u_3}}\left[\frac{1}{\partial u_1}\left(\frac{h_{u_2}h_{u_3}}{h_{u_1}}\frac{\partial\phi}{\partial u_1}\right) + \frac{1}{\partial u_2}\left(\frac{h_{u_1}h_{u_3}}{h_{u_2}}\frac{\partial\phi}{\partial u_2}\right) + \frac{1}{\partial u_3}\left(\frac{h_{u_1}h_{u_2}}{h_{u_3}}\frac{\partial\phi}{\partial u_3}\right)\right].\end{aligned}$$

Under construction !

Abbildung 15.4.1: Parametrisierung einer Raumkurve durch die Bogenlänge.

Speziell für die drei wichtigsten Koordinatensysteme bedeutet dies:

kartesisch: $$\triangle\phi(x_1,x_2,x_3)=\sum_j\frac{\partial^2\phi}{\partial x_j^2}\,,$$

Zylinder: $$\triangle\phi(\rho,\varphi,x_3)=\left[\frac{\partial^2}{\partial\rho^2}+\frac{1}{\rho}\frac{\partial}{\partial\rho}+\frac{1}{\rho^2}\frac{\partial^2}{\partial\varphi^2}+\frac{\partial^2}{\partial x_3^2}\right]\phi(\rho,\varphi,x_3)\,,$$

Kugel: $$\triangle\phi(r,\vartheta,\varphi)=\frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial\phi}{\partial r}\right)+\frac{1}{r^2\sin\vartheta}\frac{\partial}{\partial\vartheta}\left(\sin\vartheta\frac{\partial\phi}{\partial\vartheta}\right)+\frac{1}{r^2\sin^2\vartheta}\frac{\partial^2\phi}{\partial\varphi^2}\,.$$

## 15.4 Bogen-, Flächen-, Volumenelemente

### 15.4.1 Raumkurven

Wir betrachten eine Raumkurve, die durch die Parametrisierung $\underline{r}(t)$ beschrieben wird. Beispiele hatten wir schon im Abschnitt über Wegintegrale kennengelernt.
Eine natürliche Parametrisierung ist durch die Bogenlänge gegeben, d.h. dem zurückgelegten Weg entlang der Kurve. Dieser wird relativ zu einem festen Bezugspunkt $\underline{r}_0=\underline{r}(t=0)$ gemessen (Abb. 15.4.1).
Die infinitesimale Bogenlänge $ds$ können wir einfach in kartesischen Koordinaten ausdrücken:

$$(ds)^2=(dx_1)^2+(dx_2)^2+(dx_3)^2\,,$$

woraus folgt

$$s(t)=\int_0^t ds(t')=\int_0^t dt'\sqrt{\left(\frac{dx_1}{dt}\right)^2+\left(\frac{dx_2}{dt}\right)^2+\left(\frac{dx_3}{dt}\right)^2}\,.$$

Nach Definition der Geschwindigkeit gilt außerdem für ihren Betrag

$$v(t)=\frac{ds}{dt}\,.$$

Der (normierte) **Tangentenvektor** an die Kurve ist

$$\underline{T}(t)=\frac{\underline{v}(t)}{v(t)}$$

gegeben, was sich mit Hilfe der Bogenlänge in der Form

$$\underline{T}=\frac{d\underline{r}}{ds}$$

schreiben läßt.
Neben der Tangenten gibt es noch zwei weitere Vektoren, die die Raumkurve lokal charakterisieren. Der erste ist der **Hauptnormalenvektor** (oder **Krümmungsvektor**)

$$\underline{H} = \frac{1}{\kappa}\,\frac{d\underline{T}}{ds}$$

wobei $\kappa$ durch die Forderung $|\underline{H}| = 1$ bestimmt ist. $\kappa$ nennt man auch **Krümmung**. Man kann sie auch über

$$\kappa = \frac{|\dot{\underline{r}} \times \ddot{\underline{r}}|}{|\dot{\underline{r}}|}$$

bestimmen.
Aus der Normiertheit von $\underline{T}$ folgt

$$0 = \frac{d}{ds}\underline{T}^2 = 2\underline{T} \cdot \frac{d\underline{T}}{ds}$$

d.h. der Hauptnormalenvektor steht senkrecht auf der Tangenten.
Um ein lokales Orthogonalsystem zu erhalten, benötigen wir noch einen Vektor, der senkrecht auf $\underline{T}$ und $\underline{H}$ steht. Diesen nennt man auch **Binormalenvektor**:

$$\underline{B} := \underline{T} \times \underline{H}\,.$$

Wir können die Bogenlänge auch in beliebigen Koordinaten $u_j$ ausdrücken:

$$ds = \sqrt{(h_{u_1} du_1)^2 + (h_{u_2} du_2)^2 + (h_{u_3} du_3)^2}\,.$$

Speziell in Zylinderkoordinaten haben wir z.B.

$$(ds)^2 = (d\rho)^2 + \rho^2 (d\varphi)^2 + (dz)^2\,.$$

### 15.4.2 Flächen- und Volumenelemente

Wir wollen nun noch die allgemeine Form von Flächen- und Volumenelementen in krummlinigen (orthogonalen) Koordinaten ableiten.
Zunächst betrachten wir eine Koordinatenfläche, die durch die Parametrisierung $\underline{r}(u_1, u_2, u_3^{(0)})$ gegeben sei. Dann ist das Flächenelement gegeben durch

$$dA(u_1, u_2) = \left|(h_{u_1}\underline{e}_{u_1}) \times (h_{u_2}\underline{e}_{u_3})\right| du_1 du_2.$$

Analoge Ergebnisse erhält man, wenn man andere Koordinaten als $u_3$ fixiert.
Ein Beispiel ist der Zylindermantel, der sich als Koordinatenfläche ergibt, wenn man $\rho = \rho_0$ fixiert. Dann gilt

$$dA(\varphi, x_3) = \rho_0 d\varphi dx_3\,.$$

Das Volumenelement ist allgemein durch

$$dV = |h_{u_1} h_{u_2} h_{u_3}| du_1 du_2 du_3\,.$$

gegeben. Hieraus erhält man dann für die Kugel- und Zylinderkoordinaten die in Kap. 8.4 mit Hilfe des Transformationssatzes abgeleiteten Ergebnisse.