

Prof. Dr. G. M. Pastor  
Institut für Theoretische Physik  
Fachbereich Naturwissenschaften  
Universität Kassel

## **Einführung in die mathematischen Methoden der Physik (WS 07/08)**

### **Vorbemerkung**

Das Skript, das Sie in der Hand haben, ist ein Lernhilfsmittel zum ersten Teil der Vorlesung „Rechenmethoden der Physik“, wie sie an der Universität Kassel im WS 07/08 gehalten wird. Es handelt sich keineswegs um ein vollständiges Werk oder um einen richtigen Beitrag zur der zahlreichen auf dem Gebiet vorhandenen Literatur, deren Bearbeitung den Studenten dringend anempfohlen wird.

Darüber hinaus soll die jetzige Fassung nur als ein erster Entwurf betrachtet werden. Dies wird der Leser nicht nur an den „geschickt“ verstreuten Druckfehlern und knappen Formulierungen, sondern auch an einem totalen Mangel an erläuternden Figuren erkennen. Gerade bei den stark geometrisch geprägten Themen, die hier behandelt werden, ist dies eine Schwäche, die wir in späteren Fassungen beheben werden.

Kassel, Dezember 2007

# Inhaltsverzeichnis

<b>1</b>	<b>Komplexe Zahlen</b>	<b>5</b>
1.1	Definition des Körpers $\mathbb{C}$ . . . . .	6
1.2	Elementare Algebra komplexer Zahlen . . . . .	7
1.3	Polar-Darstellung . . . . .	9
1.4	Produkt und Division komplexer Zahlen in Polar-Darstellung . . . . .	9
1.5	Potenzen . . . . .	10
1.6	Wurzeln . . . . .	11
1.7	Eulersche Formel . . . . .	12
1.8	Reelle und komplexe Polynome . . . . .	13
1.9	Wurzeln von Polynomen: Fundamentalsatz der Algebra . . . . .	14
1.10	Die Exponentialfunktion in $\mathbb{C}$ . . . . .	15
1.11	Der gedämpfte harmonische Oszillator . . . . .	18
<b>2</b>	<b>Vektoren</b>	<b>19</b>
2.1	Skalare und vektorielle Größen . . . . .	19
2.2	Komponenten und Richtungswinkel . . . . .	20
2.3	Winkel zwischen zwei Vektoren . . . . .	22
2.4	Summe von Vektoren . . . . .	23
2.5	Multiplikation mit einem Skalar . . . . .	24
2.6	Skalarprodukt . . . . .	25
2.7	Transformationsregel für Vektoren bei Änderungen des Koordinatensystems . . . . .	27
2.8	Die zwei Orientierungen eines Koordinatensystems . . . . .	28
2.9	Vektoren und Pseudovektoren . . . . .	30
2.10	Das Vektorprodukt . . . . .	30
2.11	Kartesische Komponente des Vektorprodukts . . . . .	33
2.12	Matrizen und Determinanten . . . . .	36
2.13	Spatprodukt . . . . .	42
2.14	Doppeltes Vektorprodukt . . . . .	44
2.15	Andere Vektorprodukte . . . . .	45
2.16	Anwendungen der Vektoralgebra . . . . .	46
2.17	Zylinderkoordinaten . . . . .	51
2.18	Kugelkoordinaten . . . . .	52
<b>3</b>	<b>Differentiation und Integration in <math>\mathbb{R}^n</math></b>	<b>54</b>
3.1	Abbildungen . . . . .	54
3.2	Der $n$ -dimensionale Euklidische Raum $\mathbb{R}^n$ . . . . .	55
3.3	Grenzwert und Stetigkeit . . . . .	61
3.4	Ableitung in $\mathbb{R}$ . . . . .	63
3.5	L'Hopitalsche Regel . . . . .	68
3.6	Höhere Ableitungen . . . . .	69
3.7	Taylorisches Approximationspolynom . . . . .	70
3.8	Taylor-Reihe . . . . .	71
3.9	Ableitung und lokale Eigenschaften . . . . .	72
3.10	Vektorwertige Funktionen eines Parameters: Raumkurven . . . . .	73

3.11	Ableitung vektorwertiger Funktionen . . . . .	73
3.12	Ableitung von Produkten von Vektoren . . . . .	74
3.13	Skalarfelder . . . . .	76
3.14	Partielle Ableitung . . . . .	76
3.15	Satz von Schwarz . . . . .	77
3.16	Gradient eines Skalarfeldes . . . . .	78
3.17	Geometrische Eigenschaften des Gradienten . . . . .	80
3.18	Nablaoperator . . . . .	81
3.19	Richtungsableitung . . . . .	81
3.20	Totale Ableitung nach der Zeit . . . . .	82
3.21	Kraftfeld aus einem Potential – Energieerhaltung . . . . .	82
3.22	Vektorfelder: Richtungs- und totale Ableitungen . . . . .	83
3.23	Divergenz . . . . .	84
3.24	Physikalische Interpretation der Divergenz . . . . .	85
3.25	Kontinuitätsgleichung . . . . .	87
3.26	Rotation . . . . .	87
3.27	Physikalische Bedeutung der Rotation . . . . .	89
3.28	Vektorielle Formeln für den Rotoroperator . . . . .	92
3.29	Taylorsche Entwicklung in $\mathbb{R}^n$ . . . . .	94
3.30	Lokale Extreme . . . . .	96
<b>4</b>	<b>Integration in <math>\mathbb{R}^n</math></b>	<b>99</b>
4.1	Riemannsches Integral in $\mathbb{R}$ . . . . .	99
4.2	Eigenschaften des Riemannsches Integrals . . . . .	100
4.3	Hauptsatz der Infinitesimalrechnung . . . . .	101
4.4	Partielle Integration und Substitution . . . . .	102
4.5	Uneigentliche Integrale . . . . .	105
4.6	Integrale von Funktionen mit Parametern . . . . .	106
4.7	Doppel- und Dreifachintegrale . . . . .	107
4.8	Wechsel der Variablen: Krummlinige Koordinaten . . . . .	109
4.9	Kurvenintegrale . . . . .	114
4.10	Oberflächen . . . . .	116
4.11	Oberflächenintegrale . . . . .	118
4.12	Gauss'sches Theorem . . . . .	119
4.13	Stokes'sches Theorem . . . . .	119
<b>5</b>	<b>Differentialgleichungen</b>	<b>121</b>
5.1	Einleitung: Grundbegriffe und Klassifikation . . . . .	121
5.2	Beispiele gewöhnlicher Differentialgleichungen . . . . .	123
5.3	Lineare Differentialgleichungen: Superpositionsprinzip . . . . .	126
5.4	Elementare Lösungsmethoden gewöhnlicher Differentialgleichungen erster Ordnung . . . . .	128
5.5	Exakte Differentialgleichungen . . . . .	132
5.6	Integrierender Faktor . . . . .	132
5.7	Lineare Differentialgleichungen erster Ordnung . . . . .	132
5.8	Lineare Differentialgleichungen zweiter Ordnung . . . . .	132
5.9	Systeme linearer Differentialgleichungen erster Ordnung . . . . .	132
5.10	Differentialgleichungen höherer Ordnung . . . . .	132



# 1 Komplexe Zahlen

In der Physik interessieren wir uns in erster Linie für die Funktionen reeller Variablen. Trotzdem wollen wir von Anfang an die komplexen Zahlen einführen, da sie sehr hilfreich für die Lösung zahlreicher Probleme, wie z.B. harmonische Schwingungen sind. Darüber hinaus sind viele Sätze und Rechenmethoden der reellen Analysis auch für komplexe Funktionen gültig. Damit können wir uns eine spätere Verallgemeinerung ersparen.

Die komplexen Zahlen bilden einen Körper  $\mathbb{C}$ , der den Körper der reellen Zahlen  $\mathbb{R}$  erweitert, d.h.,  $\mathbb{R} \subset \mathbb{C}$ . Mit  $\mathbb{C}$  ist die letzte Stufe des „Turms“ oder der „russischen Puppenstruktur“ der Zahlen erreicht:

$$\mathbb{N} \subset \mathbb{Z} \subset \mathbb{Q} \subset \mathbb{R} \subset \mathbb{C}.$$

	a) + Summe	b) · Multiplikation
Geschlossenheit	i) $x + y \in \mathbb{N}$	i) $x \cdot y \in \mathbb{N}$
Assoziativ	ii) $(x + y) + z = x + (y + z)$	ii) $(x \cdot y) \cdot z = x \cdot (y \cdot z)$
$\exists$ neutrale Elemente	iii) $\exists$ neutrales Element 0 $x + 0 = 0 + x = x$	iii) $\exists$ neutrales Element $1 \neq 0$ (Einselement) $x \cdot 1 = 1 \cdot x = x$
Kommutativ	iv) $x + y = y + x$	vi) $x \cdot y = y \cdot x$
Distributiv	v) a und b: Distributivgesetz $x \cdot (y + z) = (x \cdot y) + (x \cdot z)$ $= x \cdot y + x \cdot z$	
$\exists$ additives Inverses	vi) $\forall x \in \mathbb{Z} \exists -x \in \mathbb{Z}$ so dass $x + (-x) = (-x) + x = 0$	
$\mathbb{Z}$ ist eine kommutative (oder Abelsche) Gruppe bzgl. der Summe. $\mathbb{Z}$ ist ein Ring mit Einselement		
$\exists$ multiplikatives Inverses		vi) $\forall x \neq 0 \in \mathbb{Q} \exists x^{-1} \in \mathbb{Q}$ so dass $x \cdot x^{-1} = x^{-1} \cdot x = 1$
$\mathbb{Q} - \{0\}$ ist eine Abelsche Gruppe bzgl. der Multiplikation. $\mathbb{Q}$ ist ein Körper. $\mathbb{Q}$ ist nicht vollständig, z.B. $\sqrt{2} \neq p/q$ mit $p, q \in \mathbb{Z}$ , $\sqrt{2} \notin \mathbb{Q}$ .		
$\mathbb{R}$ ist vollständig, d.h. $\forall a, b \in \mathbb{R} (a < b)$ ist $[a, b] \subset \mathbb{R}$ und $a < x < b \Rightarrow x \in \mathbb{R}$ . $\mathbb{R}$ ist der kleinste geordnete Körper.		

## 1.1 Definition des Körpers $\mathbb{C}$

Die Konstruktion des Körpers  $\mathbb{C}$  als Erweiterung von  $\mathbb{R}$  ist dadurch motiviert, dass Gleichungen wie

$$x^2 + 1 = 0$$

in  $\mathbb{R}$  keine Lösung haben.

Def.: Wir definieren die Erweiterung  $\mathbb{C}$  allein aus zwei Eigenschaften:

- i)  $\mathbb{C}$  ist ein Körper.
- ii)  $\exists i \in \mathbb{C}$  mit  $i^2 = -1$ .

*Alle anderen Eigenschaften und Rechenregeln der komplexen Zahlen folgen aus den Körper-eigenschaften.*

Aufgrund der Geschlossenheit bezüglich Summe und Multiplikation gilt

$$z = x + iy \in \mathbb{C} \quad \forall x, y \in \mathbb{R}.$$

Dabei sind  $x$  und  $y$  eindeutig; sie werden Realteil und Imaginärteil genannt.

Realteil:  $\operatorname{Re}\{z\} = x$

Imaginärteil:  $\operatorname{Im}\{z\} = y$ .

Um die Eindeutigkeit zu demonstrieren, betrachten wir 2 komplexe Zahlen

$$z = x + iy \quad \text{mit} \quad x, y \in \mathbb{R}$$

und

$$w = u + iv \quad \text{mit} \quad u, v \in \mathbb{R}$$

und nehmen wir an, dass

$$z = w,$$

i.e.,

$$x + iy = u + iv.$$

Damit ist entweder  $y = v$ , was auch  $x = u$  bedeutet, oder es ist  $y \neq v$ , in welchem Fall sich ergibt, dass

$$\frac{x - u}{v - y} = i \quad \text{und damit} \quad i \in \mathbb{R},$$

was ein Widerspruch ist. Also sind 2 komplexe Zahlen gleich, wenn und nur wenn Real- und Imaginärteil gleich sind.

$$z = w \Leftrightarrow \begin{cases} \operatorname{Re}\{z\} = \operatorname{Re}\{w\} & \text{und} \\ \operatorname{Im}\{z\} = \operatorname{Im}\{w\}. \end{cases}$$

Daraus schließt man die wichtige Darstellung der komplexen Zahlen:

$$\mathbb{C} = \{x + iy \in \text{Körper} \mid x, y \in \mathbb{R}\}.$$

$\mathbb{C}$  ist isomorph, d.h., Punkt für Punkt äquivalent zur Ebene

$$\mathbb{R}^2 = \{(x, y) \mid x, y \in \mathbb{R}\}.$$

Jeder Punkt auf der Ebene entspricht einer und nur einer komplexen Zahl.

## 1.2 Elementare Algebra komplexer Zahlen

Aus den Körpereigenschaften ergeben sich folgende Rechenregeln:

Sei  $z = x + iy$  und  $w = u + iv$ .

$$z + w = (x + u) + i(y + v) \quad (1.1)$$

$$-z = -x + i(-y). \quad (1.2)$$

Diese Operationen lassen sich auch geometrisch leicht interpretieren.

Die Multiplikation zweier komplexer Zahlen ergibt:

$$z \cdot w = (x + iy)(u + iv) = xu + xiv + iyu + i^2yv.$$

Da  $i^2 = -1$ , erhalten wir

$$zw = (xu - yv) + i(xv + yu). \quad (1.3)$$

Man beachte die einfache Zusammensetzung des Real- und Imaginärteils des Produkts mittels des Real- und Imaginärteils der Multiplikatanden.

Für  $z \neq 0$  können wir das multiplikative Inverse  $z^{-1}$  wie folgt bestimmen:

$$z^{-1} = \frac{1}{z} = \frac{1}{x + iy} = \frac{x - iy}{(x + iy)(x - iy)} = \frac{x - iy}{x^2 + y^2}.$$

Wir erkennen hier das Betragsquadrat  $|z|^2 = x^2 + y^2$ , das natürlich ungleich Null ist, da

$$|z|^2 = 0 \quad \Leftrightarrow \quad z = 0 \quad (x, y \in \mathbb{R}).$$

Der Ausdruck  $|z| = \sqrt{x^2 + y^2}$  wird *Betrag von  $z$*  genannt, denn für  $z \in \mathbb{R}$  stimmt  $|z|$  mit dem üblichen Betrag in  $\mathbb{R}$  überein. Geometrisch entspricht  $|z|$  dem Abstand zum Ursprung.

Wir erkennen auch das Spiegelbild  $x - iy$  von  $z$  bezüglich der reellen Achse. Man nennt  $x - iy = \bar{z} = z^*$  die zu  $z$  *konjugierte komplexe Zahl*:

$$z = x + iy \xrightarrow{\text{Konjugation}} x - iy = \bar{z} = z^*.$$

Für jede komplexe Zahl  $z \neq 0$  gilt also:

$$z^{-1} = \frac{\bar{z}}{|z|^2}.$$

Folgende Formeln lassen sich schnell verifizieren:

i)  $|z|^2 = z\bar{z} = |\bar{z}|^2$  ( $|z| = \sqrt{z\bar{z}} = \sqrt{x^2 + y^2}$ ).

ii)  $z = \bar{z} \quad \Leftrightarrow \quad z \in \mathbb{R}$ .

iii)  $\bar{\bar{z}} = z$ .

iv)  $\overline{z + w} = \bar{z} + \bar{w}$ .

v)  $\overline{z \bar{w}} = \bar{z} w$ .

vi)  $\operatorname{Re}\{z\} = \frac{z + \bar{z}}{2}; \quad \operatorname{Im}\{z\} = \frac{z - \bar{z}}{2i}$ .

Zum Beispiel haben wir für  $z = x + iy$  und  $w = u + iv$

$$z \bar{z} = (x + iy)(x - iy) = x^2 - (iy)^2 = x^2 + y^2 = |z|^2$$

und

$$\begin{aligned} \overline{z \bar{w}} &= \overline{(x + iy)(u + iv)} = \overline{xu - yv + i(xv + yu)} = \\ &= xu - yv - i(xv + yu) = xu - yv + i[x(-v) + (-y)u] = \bar{z} \bar{w}. \end{aligned}$$

Der absolute Betrag  $|z|$  erfüllt seinerseits folgende Beziehungen:

i)  $|z w| = |z| |w|$ . Dies folgt aus  $|z w|^2 = (z w)(\overline{z w}) = z \bar{z} w \bar{w} = |z|^2 |w|^2$ .

ii)  $|\operatorname{Re}(z)| \leq |z|$  und  $|\operatorname{Im}(z)| \leq |z|$ , da  $x^2 \leq x^2 + y^2$  und  $y^2 \leq x^2 + y^2$  ( $x, y \in \mathbb{R}$ ).

iii)  $|z| \geq 0 \quad \forall z \in \mathbb{C}$  und  $|z| = 0 \Leftrightarrow z = 0$ .

iv)  $|z| - |w| \leq |z + w| \leq |z| + |w|$ . Das ist die sogenannte *Dreiecksungleichung*.

Um dies zu zeigen, schreiben wir einerseits

$$\begin{aligned} |z + w|^2 &= (z + w)\overline{(z + w)} = (z + w)(\bar{z} + \bar{w}) = |z|^2 + |w|^2 + z \bar{w} + \bar{z} w \\ &= |z|^2 + |w|^2 + z \bar{w} + \overline{z \bar{w}} = |z|^2 + |w|^2 + 2 \operatorname{Re}(z \bar{w}) \end{aligned}$$

und andererseits

$$\begin{aligned} (|z| \pm |w|)^2 &= |z|^2 + |w|^2 \pm 2|z| |w| \\ &= |z|^2 + |w|^2 \pm 2|z| |\bar{w}| \\ &= |z|^2 + |w|^2 \pm 2|z \bar{w}|. \end{aligned}$$

Die Dreiecksungleichung folgt dann aus ii), d.h., ( $\operatorname{Re}(z \bar{w}) \leq |z \bar{w}|$ ), wenn man bedenkt, dass für reelle  $a, b > 0$  gilt:  $a < b \Leftrightarrow a^2 < b^2$ .

### 1.3 Polar-Darstellung

Die Analogie bzw. Isomorphismus mit  $\mathbb{R}^2$  legt eine Polar-Darstellung der komplexen Zahlen nahe:

$$z = x + iy = r(\cos \varphi + i \sin \varphi) \quad \text{mit} \quad \begin{cases} r^2 = x^2 + y^2 \\ \sin \varphi = \frac{y}{r} = \frac{y}{\sqrt{x^2+y^2}} \\ \cos \varphi = \frac{x}{r} = \frac{x}{\sqrt{x^2+y^2}} \\ \tan \varphi = \frac{y}{x} \\ \varphi = \arg(z) = \arctan\left(\frac{y}{x}\right) \end{cases}$$

Der Radius  $r = |z|$  ist nichts anderes als der Betrag von  $z$ . Der Winkel  $\varphi$  wird als *Argument* von  $z$  bezeichnet:

$$\varphi = \arg(z).$$

Beispiele:

Sei  $|z| = 1$  und  $\arg(z) = 45^\circ$ , dann ist

$$\begin{aligned} z &= \frac{\sqrt{2}}{2} (1 + i) \\ z^2 &= \frac{1}{2} (1 + i)^2 = \frac{1}{2} (1 + i^2 + 2i) = i \\ z^3 &= \frac{\sqrt{2}}{2} (1 + i) i = \frac{\sqrt{2}}{2} (-1 + i). \end{aligned}$$

Sei  $|z| = 1$  und  $\arg(z) = 30^\circ$ .

$$\begin{aligned} z &= \frac{\sqrt{3}}{2} + \frac{1}{2} i \\ \frac{i}{z} &= i \frac{\bar{z}}{|z|^2} = i \left( \frac{\sqrt{3}}{2} - \frac{1}{2} i \right) = \frac{1}{2} + \frac{\sqrt{3}}{2} i = \cos 60^\circ + i \sin 60^\circ. \end{aligned}$$

Diese Beispiele deuten an, dass bei dem Produkt (Division) von komplexen Zahlen die Argumente addiert (subtrahiert) werden sollen. Dies ist in der Tat so, wie im folgenden Abschnitt gezeigt wird.

### 1.4 Produkt und Division komplexer Zahlen in Polar-Darstellung

Sei  $z_1 = r_1 (\cos \varphi_1 + i \sin \varphi_1)$  und  $z_2 = r_2 (\cos \varphi_2 + i \sin \varphi_2)$ , dann gilt:

$$\begin{aligned} z_1 z_2 &= r_1 r_2 [\cos \varphi_1 \cos \varphi_2 - \sin \varphi_1 \sin \varphi_2 + \\ &\quad + i (\cos \varphi_1 \sin \varphi_2 + \sin \varphi_1 \cos \varphi_2)]. \end{aligned}$$

Nach Einsetzen der aus der Trigonometrie bekannten Beziehungen

$$\begin{aligned}\cos(\alpha + \beta) &= \cos \alpha \cos \beta - \sin \alpha \sin \beta \\ \sin(\alpha + \beta) &= \cos \alpha \sin \beta + \sin \alpha \cos \beta\end{aligned}$$

folgt:

$$z_1 z_2 = r_1 r_2 [\cos(\varphi_1 + \varphi_2) + i \sin(\varphi_1 + \varphi_2)]. \quad (1.4)$$

Zusammenfassend haben wir

$$|z_1 z_2| = |z_1| |z_2| \quad (1.5)$$

und

$$\arg(z_1 z_2) = \arg(z_1) + \arg(z_2). \quad (1.6)$$

Die Polar-Darstellung erlaubt also eine geometrisch anschauliche Visualisierung des Produktes komplexer Zahlen. Aus den Gleichungen (1.5) und (1.6) folgt auch

$$\left| \frac{z_1}{z_2} \right| = \frac{|z_1|}{|z_2|} \quad \text{und} \quad \arg\left(\frac{z_1}{z_2}\right) = \arg(z_1) - \arg(z_2),$$

da  $\arg(z^{-1}) = \arg(\bar{z}) = -\arg(z)$  ist. Dies ergibt sich aus  $z \cdot z^{-1} = 1 \Rightarrow \arg(z \cdot z^{-1}) = \arg(1) = 0 \Rightarrow \arg(z) + \arg(z^{-1}) = 0$ .

## 1.5 Potenzen

Eine sukzessive Anwendung der Gleichungen (1.5) und (1.6) zeigt, dass für  $z = |z| (\cos \varphi + i \sin \varphi)$  gilt:

$$z^n = |z|^n [\cos(n \varphi) + i \sin(n \varphi)].$$

Anders ausgedrückt

$$\begin{aligned}|z^n| &= |z|^n \\ \arg(z^n) &= n \arg(z).\end{aligned}$$

Daraus folgt die Moivresche Formel

$$(\cos \varphi + i \sin \varphi)^n = \cos(n \varphi) + i \sin(n \varphi), \quad (1.7)$$

wie man es leicht sehen kann, wenn man

$$z^n = [|z| (\cos \varphi + i \sin \varphi)]^n = |z|^n (\cos \varphi + i \sin \varphi)^n$$

mit  $z^n = |z|^n [\cos(n \varphi) + i \sin(n \varphi)]$  vergleicht. Die Moivresche Formel erlaubt die Winkelfunktionen des  $n$ -fachen eines Winkels  $n \varphi$  mittels der Winkelfunktionen des Winkels selbst auszudrücken bzw. zu berechnen.

## 1.6 Wurzeln

Die Berechnung der  $n$ -ten Wurzeln komplexer Zahlen ist ein klein wenig subtiler. Die  $n$ -te Wurzel ist durch folgende Beziehung definiert:

$$z^{1/n} = \sqrt[n]{z} = w \Leftrightarrow w^n = z.$$

Gegeben sei  $z = |z|(\cos \varphi + i \sin \varphi)$ . Wir suchen  $w = |w|(\cos \theta + i \sin \theta)$ , so dass

$$w^n = |w|^n (\cos n\theta + i \sin n\theta) = |z|(\cos \varphi + i \sin \varphi).$$

Dies ist dann nur möglich, wenn

$$|w| = \sqrt[n]{|z|}, \quad (1.8)$$

wobei  $\sqrt[n]{\phantom{x}}$  die übliche  $n$ -te Wurzel für positive reelle Zahlen ist, und wenn

$$\begin{aligned} \cos n\theta &= \cos \varphi \\ \sin n\theta &= \sin \varphi. \end{aligned} \quad (1.9)$$

Während die Bestimmung von  $|w|$  eindeutig ist, gibt es  $n$  verschiedene Lösungen für  $\theta$ . Aus den Bedingungen (1.9) folgt

$$n\theta = \varphi + 2\pi k \quad \text{mit} \quad k \in \mathbb{Z},$$

wobei die Gleichung

$$\theta = \frac{\varphi}{n} + \frac{2\pi}{n}k \quad \text{mit} \quad k = 0, 1, \dots, n-1 \quad (1.10)$$

$n$  verschiedene komplexe Zahlen  $w$  liefert.

Zusammenfassend gilt für  $z = |z|(\cos \varphi + i \sin \varphi)$

$$z^{1/n} = \sqrt[n]{|z|} \left[ \cos \left( \frac{\varphi}{n} + \frac{2\pi k}{n} \right) + i \sin \left( \frac{\varphi}{n} + \frac{2\pi k}{n} \right) \right] \quad (1.11)$$

mit  $k = 0, 1, \dots, n-1$ .

Beispiele:

Da  $|i| = 1$  und  $\arg(i) = \pi/2$ , erhalten wir für  $\sqrt{i}$  folgendes Resultat:

$$i^{1/2} = \sqrt{i} = \cos \left( \frac{\pi}{4} + \frac{2\pi}{2}k \right) + i \sin \left( \frac{\pi}{4} + \frac{2\pi}{2}k \right)$$

mit  $k = 0$  oder  $1$ . Also

$$\sqrt{i} = \begin{cases} \cos \frac{\pi}{4} + i \sin \frac{\pi}{4} \\ \cos \frac{3\pi}{4} + i \sin \frac{3\pi}{4} \end{cases}$$

$$\sqrt{i} = \begin{cases} \frac{\sqrt{2}}{2} + i \frac{\sqrt{2}}{2} \\ -\frac{\sqrt{2}}{2} - i \frac{\sqrt{2}}{2} \end{cases}.$$

Auf ähnliche Weise erhalten wir für  $\sqrt[4]{1}$  folgendes:

$$\begin{aligned} \sqrt[4]{1} &= \cos\left(\frac{\pi}{2}k\right) + i \sin\left(\frac{\pi}{2}k\right) \quad \text{mit } k = 0, 1, 2, 3 \\ &= \{1, i, -1, -i\}. \end{aligned}$$

Ein besonders wichtiges Beispiel sind die  $n$ -ten Wurzeln der Einheit, d.h.

$$\sqrt[n]{1} = \left\{ \eta_k = \cos\left(\frac{2\pi}{n}k\right) + i \sin\left(\frac{2\pi}{n}k\right) \quad \text{mit } k = 0, 1, \dots, n-1 \right\},$$

die eine Abelsche Gruppe mit der Multiplikation bilden. Die Wurzeln der Einheit stellen die Phasenunterschiede zwischen den verschiedenen Wurzeln beliebiger komplexer Zahlen dar. Es gilt nämlich

$$\sqrt[n]{z} = \left\{ \xi_0 \eta_k, \quad \text{wobei } \xi_0^n = z \quad \text{und } \eta_k = e^{i \frac{2\pi}{n} k} \quad \text{mit } k = 0, 1, \dots, n-1 \right\}.$$

## 1.7 Eulersche Formel

In diesem Zusammenhang ist die Eulersche Formel zu erwähnen, die eine sehr wichtige Beziehung zwischen den Winkelfunktionen und der Exponentialfunktion für imaginäre Zahlen darstellt. Sie lautet:

$$e^{i\varphi} = \cos \varphi + i \sin \varphi, \tag{1.12}$$

wobei  $e = \sum_{k=0}^{+\infty} \frac{1}{k!} = 2 + 1/2 + 1/6 + 1/24 + \dots = 2.718281\dots$  die Eulersche Zahl ist.

Damit lassen sich komplexe Zahlen auf eine sehr kompakte und äußerst praktische Form darstellen:

$$z = |z| e^{i\varphi} \quad \text{mit } \varphi = \arg(z). \tag{1.13}$$

Wir werden die Eulersche Formel später herleiten, sobald wir die Taylor-Reihen in Kapitel 3 definiert haben. Die Eulersche Schreibweise komplexer Zahlen [Gl. (1.12) und (1.13)] fasst alle Eigenschaften der Multiplikation, Division, Potenzierung und Wurzel zusammen. Dazu braucht man nur wissen, dass

$$e^{z+w} = e^z \cdot e^w \quad \forall z, w \in \mathbb{C},$$

wie bei den reellen Zahlen.

Ob man die kartesische  $z = x + iy$  oder die polare  $z = r(\cos \varphi + i \sin \varphi)$  Darstellung benutzt, hängt natürlich von der spezifischen Anwendung ab. Im allgemeinen sind Summe und Subtraktion einfacher in kartesischer Darstellung, während die Polar-Darstellung am besten für Multiplikation, Division, Potenzen und Wurzeln geeignet ist.

## 1.8 Reelle und komplexe Polynome

Nachdem wir die komplexen Zahlen und ihre Algebra eingeführt haben, können wir den Definitionsbereich der reellen Funktionen auf  $\mathbb{C}$  erweitern. Ein wichtiges Beispiel sind die Polynome in einem unbestimmten  $x$ :

$$p(x) = a_n x^n + a_{n-1} x^{n-1} + \dots + a_1 x + a_0 \quad (1.14)$$

mit  $n \in \mathbb{N}$  und reellen oder komplexen Koeffizienten  $a_k$ . Die Gesamtheit der Polynome bilden einen Ring mit ähnlichen Eigenschaften wie die ganzen Zahlen  $\mathbb{Z}$ . Polynome können also summiert und multipliziert werden. Sie können auch dividiert werden, wobei im Allgemeinen wie bei den ganzen Zahlen ein Rest entsteht.

Die Ordnung  $n$  der höchsten Potenz  $x^n$  mit einem nicht verschwindenden Koeffizienten wird *Grad des Polynoms* genannt. Im Falle von  $p(x)$  aus der Gleichung (1.14) ist

$$\deg(p) = n \quad \text{wenn} \quad a_n \neq 0.$$

Der Grad erfüllt folgende bekannte Beziehungen:

$$\begin{aligned} \deg(p \cdot q) &= \deg(p) + \deg(q), \\ \deg(p + q) &\leq \max \{ \deg(p), \deg(q) \}, \text{ und} \\ \deg(p) > \deg(q) &\Rightarrow \deg(p + q) = \deg(p). \end{aligned}$$

Eine sehr nützliche Eigenschaft ist die Möglichkeit, Polynome mit Rest zu dividieren. Es seien  $a(x)$  und  $q(x)$  Polynome mit  $q(x) \neq 0$ , dann gibt es eine eindeutige Darstellung der Form

$$a(x) = b(x)q(x) + r(x)$$

mit gewissen Polynomen  $b(x)$  und  $r(x)$  wobei  $\deg(r) < \deg(q)$ . Ein ähnliches Theorem gilt für die Division ganzer Zahlen.

Beispiele:

i)  $a(x) = x^4 - 1$     und     $q(x) = x - 1$ .

$$\begin{array}{r} x^4 - 1 \\ - (x^4 - x^3) \\ \hline x^3 - 1 \\ - (x^3 - x^2) \\ \hline x^2 - 1 \\ - (x^2 - x) \\ \hline x - 1 \\ - (x - 1) \\ \hline 0 \end{array} \quad \begin{array}{l} | x - 1 \\ x^3 + x^2 + x + 1 \\ \\ \\ \Rightarrow x^4 - 1 = (x^3 + x^2 + x + 1)(x - 1) \\ b(x) = x^3 + x^2 + x + 1 \\ r(x) = 0 \end{array}$$

ii)  $a(x) = x^3 + 2x^2 + 2x + 1$  und  $q(x) = x^2 + 1$ .

$$\begin{array}{r}
 x^3 + 2x^2 + 2x + 1 \\
 - (x^3 + x) \\
 \hline
 2x^2 + x + 1 \\
 - (2x^2 + 2) \\
 \hline
 x - 1
 \end{array}
 \quad \left| \begin{array}{l}
 x^2 + 1 \\
 x + 2
 \end{array} \right.$$

$$\Rightarrow x^3 + 2x^2 + 2x + 1 = (x^2 + 1)(x + 2) + (x - 1)$$

$$b(x) = x + 2$$

$$r(x) = x - 1$$

### 1.9 Wurzeln von Polynomen: Fundamentalsatz der Algebra

Sei  $\xi \in \mathbb{C}$ . Man sagt, dass  $\xi$  eine Wurzel des Polynoms

$$p(x) = x^n + a_{n-1}x^{n-1} + \dots + a_1x + a_0$$

ist, wenn  $p(\xi) = 0$ . In anderen Worten ist  $\xi$  eine Lösung der algebraischen Gleichung

$$z^n + a_{n-1}z^{n-1} + \dots + a_1z + a_0 = 0.$$

Für reelle Polynome ( $a_k \in \mathbb{R} \ \forall k$ ) gilt:  $\overline{p(\xi)} = p(\bar{\xi})$  und demzufolge gilt  $p(\xi) = 0 \Leftrightarrow p(\bar{\xi}) = 0$ . Anders ausgedrückt: Für reelle Polynome sind die Wurzeln reell oder sie kommen immer in Paaren  $\xi$  und  $\bar{\xi}$ .

Die komplexen Zahlen wurden dazu eingeführt, um die Lösung der speziellen Gleichung

$$x^2 + 1 = 0$$

zu ermöglichen. In der Zwischenzeit haben wir auch gezeigt, dass die Gleichungen der Form

$$z^n + a_0 = 0$$

immer  $n$  Lösungen in  $\mathbb{C}$  haben. Es ist aber durchaus bemerkenswert, dass in  $\mathbb{C}$  alle algebraischen Gleichungen eine Lösung finden. Es gilt nämlich folgender *Fundamentalsatz der Algebra*: Jede Gleichung der Form

$$p(z) = z^n + a_{n-1}z^{n-1} + \dots + a_1z + a_0 = 0$$

mit  $a_k \in \mathbb{C}$  und  $n \geq 1$  besitzt wenigstens eine Lösung in  $\mathbb{C}$ .

Aus dem Satz über die Division von Polynomen folgt das Lemma:

Wenn  $p(\xi) = 0$  ( $\xi \in \mathbb{C}$ ), dann ist  $p(x)$  durch  $(x - \xi)$  teilbar, d.h.,

$$p(x) = b(x)(x - \xi)$$

mit  $\deg(b) = \deg(p) - 1$ . In der Tat gilt

$$p(x) = b(x)(x - \xi) + r(x)$$

mit  $\deg(r) < 1$ , also  $r(x) = r_0 = \text{Konstante}$ . Aber  $p(\xi) = 0 \Rightarrow r_0 = 0$ .

Da  $\deg(b) = \deg(p) - 1$ , können wir den Fundamentalsatz der Algebra bei  $b(x)$  anwenden, solange  $\deg(b) \geq 1$ , und folgenden Satz der *Faktorisierung von Polynomen* beweisen. Sei

$$p(x) = x^n + a_{n-1}x^{n-1} + \dots + a_1x + a_0$$

ein normiertes Polynom ( $a_n = 1$ ). Dann gibt es  $n$  bis auf die Numerierung eindeutige Zahlen  $\xi_1, \xi_2, \dots, \xi_n \in \mathbb{C}$ , so dass

$$p(x) = (x - \xi_1)(x - \xi_2) \dots (x - \xi_n).$$

Es ist klar, dass zwei oder mehr Wurzeln  $\xi_i$  gleich sein können, wie zum Beispiel bei  $p(x) = (x - 1)^2 = x^2 - 2x + 1$ .

## 1.10 Die Exponentialfunktion in $\mathbb{C}$

Als ein weiteres sehr wichtiges Beispiel für Funktionen komplexer Variablen betrachten wir die Exponentialfunktion  $\exp(z)$ , die durch folgende Exponentialreihe definiert ist:

$$\begin{aligned} \exp(z) &= \sum_{k=0}^{+\infty} \frac{z^k}{k!} \quad k \in \mathbb{N} \\ &= \lim_{N \rightarrow +\infty} \sum_{k=0}^N \frac{z^k}{k!}. \end{aligned}$$

Die Exponentialreihe ist ein Spezialfall der Taylor-Reihen, die wir in Kapitel 3 untersuchen werden.

Eine genaue Diskussion der Konvergenz der Exponentialreihe ist leider jenseits unserer jetzigen Möglichkeiten, da wir dazu die Theorie der Folgen und Reihen brauchen würden. Man möge also ohne Beweis akzeptieren, dass  $\exp(z) \forall z \in \mathbb{C}$  konvergent ist ( $|z| < +\infty$ ). Das heißt  $\forall \varepsilon > 0 \exists n_0 \in \mathbb{N}$  mit

$$\left| \sum_{k=n+1}^m \frac{z^k}{k!} \right| < \varepsilon \quad \forall m \geq n > n_0.$$

Sie ist sogar absolut konvergent, d.h.,

$$\exp(|z|) = \sum_{k=0}^{+\infty} \frac{|z|^k}{k!}$$

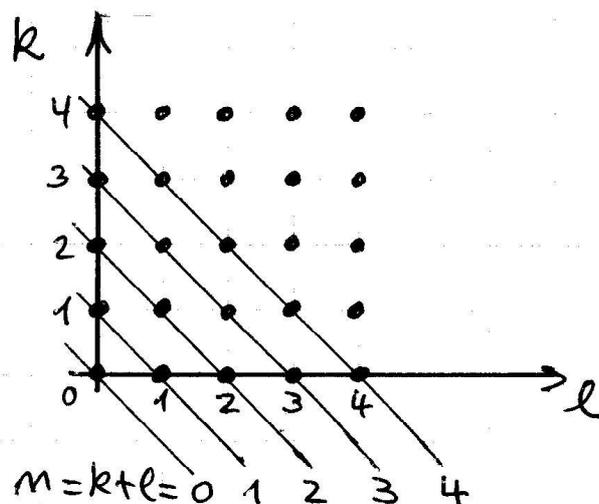
konvergiert. Die fundamentale Eigenschaft der Exponentialfunktion ist

$$\exp(a + b) = \exp(a) \exp(b)$$

$\forall a, b \in \mathbb{C}$ . Um dies zu beweisen, schreiben wir

$$\begin{aligned} \exp(a) \exp(b) &= \left( \sum_{k=0}^{+\infty} \frac{a^k}{k!} \right) \left( \sum_{l=0}^{+\infty} \frac{b^l}{l!} \right) \\ &= \sum_{k=0}^{+\infty} \sum_{l=0}^{+\infty} \frac{a^k}{k!} \frac{b^l}{l!}. \end{aligned}$$

Diese Doppelsumme kann mit Hilfe des folgenden Diagramms umgeschrieben werden:



Sei  $n = k + l$  und also  $l = n - k$ . Dann gilt

$$\begin{aligned} \exp(a) \cdot \exp(b) &= \sum_{n=0}^{+\infty} \sum_{k=0}^n \frac{a^k}{k!} \frac{b^{n-k}}{(n-k)!} \\ &= \sum_{n=0}^{+\infty} \frac{1}{n!} \sum_{k=0}^n \frac{n!}{k!(n-k)!} a^k b^{n-k}. \end{aligned}$$

Wir erkennen die Binomialkoeffizienten  $\binom{n}{k} = \frac{n!}{k!(n-k)!}$  und den binomischen Lehrsatz

$(a + b)^n = \sum_{k=0}^n \binom{n}{k} a^k b^{n-k}$ , so dass

$$\begin{aligned} \exp(a) \exp(b) &= \sum_{n=0}^{\infty} \frac{1}{n!} (a + b)^n = \\ &= \exp(a + b). \end{aligned}$$

Diese wichtige Eigenschaft erlaubt uns, die Verbindung zwischen der mit Hilfe der Exponentialreihe definierten Exponentialfunktion komplexer Zahlen  $\exp(z)$  und der aus den Reellen bekannten Funktion  $e^x$  ( $x \in \mathbb{R}$ ) herzustellen.

Wir betrachten  $\exp(z)$  für  $z = x \in \mathbb{R}$  und zeigen, dass

$$\exp(x) = 1 + x + \frac{x^2}{2!} + \dots > 1 \quad \text{für } x > 0.$$

Darüber hinaus haben wir

$$\exp(x) \cdot \exp(-x) = \exp(x - x) = \exp(0) = 1.$$

Deshalb ist

$$0 < \exp(x) = \frac{1}{\exp(-x)} < 1 \quad \text{für } x < 0.$$

Wir definieren die Eulersche Zahl

$$e = \exp(1) = \frac{1}{0!} + \frac{1}{1!} + \frac{1}{2!} + \dots = 2,7182818$$

und beweisen die Identität

$$\exp(x) = e^x \quad \forall x \in \mathbb{Q}.$$

Sei  $x = \frac{p}{q}$  mit  $p$  und  $q \in \mathbb{Z}$ . Dann haben wir

$$\begin{aligned} \left[ \exp\left(\frac{p}{q}\right) \right]^q &= \underbrace{\exp\left(\frac{p}{q}\right) \cdot \exp\left(\frac{p}{q}\right) \dots \exp\left(\frac{p}{q}\right)}_{q \text{ Faktoren}} = \\ &= \exp\left(\underbrace{\frac{p}{q} + \dots + \frac{p}{q}}_{q \text{ Summanden}}\right) = \exp(p) \\ &= \underbrace{\exp(1) \cdot \exp(1) \dots \exp(1)}_{p \text{ Faktoren}} = e^p. \end{aligned}$$

Da  $\exp\left(\frac{p}{q}\right) > 0$  und  $e^p > 0$  ( $e > 0$ ), können wir die  $q$ -te Wurzel ziehen, und damit ist

$$\exp\left(\frac{p}{q}\right) = \sqrt[q]{e^p} = e^{p/q}.$$

Damit haben wir bewiesen, dass

$$e^z = \exp(z) = \sum_{k=0}^{+\infty} \frac{z^k}{k!}$$

eine Fortsetzung auf  $\mathbb{C}$  der zunächst nur auf  $\mathbb{Q}$  definierten Funktion  $e^x$  ist. Für ein beliebiges  $z \in \mathbb{C}$  gilt also

$$e^z = e^{\operatorname{Re}\{z\} + i \operatorname{Im}\{z\}} = e^{\operatorname{Re}\{z\}} \cdot e^{i \operatorname{Im}\{z\}}.$$

Wir können nun die Eulersche Formel für die Exponentialfunktion imaginärer Zahlen anwenden:

$$e^{i\varphi} = \cos \varphi + i \sin \varphi.$$

Damit haben wir

$$e^z = e^{\operatorname{Re}\{z\}} [\cos(\operatorname{Im}\{z\}) + i \sin(\operatorname{Im}\{z\})].$$

## 1.11 Der gedämpfte harmonische Oszillator

Damit die komplexen Zahlen nicht ohne jegliche physikalische Anwendung bleiben, betrachten wir hier das Problem des eindimensionalen harmonischen Oszillators. Newtons Gleichung lautet

$$m \ddot{x} = -\eta \dot{x} - k x.$$

Hier ist  $x = x(t)$  die Position der Masse  $m$  bezüglich der Gleichgewichtsstelle der Feder,  $\dot{x} = \frac{dx}{dt}$  ist die Geschwindigkeit,  $\ddot{x} = \frac{d^2x}{dt^2}$  die Beschleunigung,  $\eta$  die Konstante der Reibungskraft  $-\eta \dot{x}$ , die entgegengesetzt zur Geschwindigkeit ist, und  $k > 0$  die Federkonstante. Um die Bewegungsgleichung zu lösen, versuchen wir den Ansatz

$$x(t) = e^{zt} \quad \text{mit } z \in \mathbb{C} \text{ eine Konstante}$$

mit Einheit  $[z] = [t]^{-1} = s^{-1}$ . Dies ergibt

$$0 = m \ddot{x} + \eta \dot{x} + k x = (m z^2 + \eta z + k) e^{zt}.$$

Da  $|e^{zt}| > 0 \quad \forall z, t$ , ist dies nur möglich, wenn

$$m z^2 + \eta z + k = 0,$$

was zu

$$z = -\frac{\eta}{2m} \pm \sqrt{\left(\frac{\eta}{2m}\right)^2 - \frac{k}{m}}$$

führt. Wir konzentrieren uns auf den Fall kleiner Reibung, d.h.,  $\left(\frac{\eta}{2m}\right)^2 < \frac{k}{m}$ , bei dem

$$z = -\frac{\eta}{2m} \pm i \sqrt{\frac{k}{m} - \left(\frac{\eta}{2m}\right)^2}$$

und definieren  $\frac{1}{\tau} = \frac{\eta}{2m} > 0$  und  $\omega^2 = \frac{k}{m} - \left(\frac{\eta}{2m}\right)^2 > 0$ . Damit lautet unsere komplexe Lösung der Bewegungsgleichung

$$\begin{aligned} x(t) &= e^{zt} = e^{\frac{-t}{\tau} \pm i \omega t} \\ &= e^{\frac{-t}{\tau}} [\cos(\omega t) \pm i \sin(\omega t)]. \end{aligned}$$

Physikalisch ist also  $\tau$  die Dämpfungszeit oder Halbdauer und  $T = \frac{2\pi}{\omega}$  die Schwingungsperiode. Da die Newton-Gleichung reelle Koeffizienten hat, wird sie nicht nur durch  $x(t)$ , sondern auch durch  $\bar{x}(t)$  erfüllt. Darüber hinaus ist die Gleichung linear und deshalb sind  $\frac{x + \bar{x}}{2} = \operatorname{Re}(x)$  und  $\frac{x - \bar{x}}{2i} = \operatorname{Im}(x)$  beide unabhängige Lösungen. Die allgemeine Form der Position  $x(t)$  als Funktion der Zeit lautet

$$x(t) = e^{\frac{-t}{\tau}} [A \cos(\omega t) + B \sin(\omega t)],$$

wobei  $A$  und  $B$  an die Anfangsbedingungen anzupassen sind.

## 2 Vektoren

In diesem Kapitel betrachten wir die Eigenschaften von Vektoren im dreidimensionalen Raum  $\mathbb{R}^3$  als besonders wichtiges konkretes Beispiel des Euklidischen Raums  $\mathbb{R}^n$ . Hier gehen wir von einer physikalischen Motivation aus und zeigen, dass die physikalischen vektoriellen Größen die allgemeinen Eigenschaften eines Vektorraums im mathematischen Sinne erfüllen. Eine mathematische Darstellung des Euklidischen Raums  $\mathbb{R}^n$  und seiner fundamentalen Eigenschaften wird in Abschnitt 3.2 gegeben.

### 2.1 Skalare und vektorielle Größen

In der Physik (und natürlich auch in der Geometrie) unterscheidet man zwischen skalaren, vektoriellen und im allgemeinen tensoriellen Größen. Eine genaue mathematische Definition dieser Begriffe soll später gegeben werden. Im Moment möchten wir sie auf intuitive Weise einführen.

Ein *Skalar* ist eine Größe, die lediglich durch eine Zahl erfasst wird (typischerweise eine reelle Zahl) und die *invariant bezüglich Drehungen* des Koordinatensystems ist. Da eine Drehung des Koordinatensystems äquivalent zu einer Drehung des physikalischen Systems (in entgegengesetzte Richtung) ist, kann man auch sagen, dass eine Größe ein Skalar ist, wenn sie invariant gegenüber Drehungen des dazugehörigen physikalischen Systems ist.

Beispiele von Skalaren sind

- 1) die gesamte Masse  $M = \sum_i m_i$  eines Systems (z.B. eines Moleküls),
- 2) der Abstand zwischen zwei Punkten eines Systems (z.B. zwischen 2 Sternen oder zwei Atomen in einem Molekül),
- 3) der Betrag der Geschwindigkeit eines Körpers.  $|\vec{v}| = \sqrt{v_x^2 + v_y^2 + v_z^2}$ , und

- 4) die mittlere interne kinetische Energie

$$E_K = \frac{1}{2} \sum_i m_i |\vec{v}_i|^2$$

oder die Temperatur eines Körpers.

*Vektoren* sind Größen, die sowohl einen *Betrag* als auch eine *Richtung* beinhalten. Anders ausgedrückt, man braucht nicht nur eine Zahl, sondern auch eine Richtung, um sie zu erfassen.

Beispiele von Vektoren sind

- 1) die Position eines Punktes  $P$  im Raum bezüglich eines vorgegebenen Ursprungs  $O$ :  
 $\vec{r} = \overrightarrow{OP}$ ,
- 2) die Translation von einem Punkt  $P_1$  zu einem anderen Punkt  $P_2$ , d.h.,  $\vec{r}_{12} = \overrightarrow{P_1P_2}$ .

$$\overrightarrow{OP_1} + \overrightarrow{P_1P_2} = \overrightarrow{OP_2}$$

$$\overrightarrow{P_1P_2} = \overrightarrow{OP_2} - \overrightarrow{OP_1}$$

3) die Geschwindigkeit, d.h., die Änderung der Position per Einheit Zeit

$$\vec{v} = \frac{\overrightarrow{OP}(t + \Delta t) - \overrightarrow{OP}(t)}{\Delta t} = \frac{\overrightarrow{P}(t) \overrightarrow{P}(t + \Delta t)}{\Delta t}, \text{ und}$$

4) die Kräfte, z. B. die Gravitationskraft, mit der ein Planet auf einen Satelliten wirkt.

Diese Beispiele zeigen, dass Vektoren äquivalent (genauer gesagt isomorph) zur Menge der orientierten Segmente oder Pfeile sind. Ein Vektor bzw. eine vektorielle Größe dreht mit dem physikalischen System mit.

Motiviert durch die vorherigen Beispiele werden Vektoren und vektorielle Größen wie folgt definiert:

1. Als *Vektor* bezeichnen wir also jedes orientierte Segment. Der erste Punkt des Segments ist der Anfangspunkt und der zweite der Endpunkt des Vektors.
2. Eine Größe heißt *vektoriell*, wenn die Menge ihrer Werte in einer eins-zu-eins stetigen Beziehung mit der Menge der orientierten Vektoren mit gemeinsamen Anfangspunkten steht.
3. Als *Betrag*  $|\vec{v}|$  eines Vektors  $\vec{v}$  bezeichnet man die Stärke der Größe, die er darstellt. Zum Beispiel ist bei der Position  $\overrightarrow{OP}$  der Betrag

$$\begin{aligned} |\overrightarrow{OP}| &= \text{Abstand zwischen dem Punkt } P \text{ und dem Ursprung,} \\ |\overrightarrow{P_1P_2}| &= \text{Abstand zwischen } P_1 \text{ und } P_2, \\ |\vec{v}| &= \text{Stärke der Geschwindigkeit.} \end{aligned}$$

4. Zwei Vektoren sind also gleich, wenn sie den gleichen Betrag und die gleiche Richtung haben.

Man stelle sich also vor, dass man die Vektoren an einen gemeinsamen Ursprung bringt, selbst wenn sie physikalisch an verschiedenen Punkten im Raum angeknüpft sind (z.B. wie bei der Geschwindigkeit parallel laufender Schifahrer). Dieses Gleichheitskriterium entspricht dem Begriff freier Vektoren, d.h., ohne festen Ursprung. Es ist äquivalent zur Gleichheit von Vektoren in  $\mathbb{R}^n$ .

## 2.2 Komponenten und Richtungswinkel

Wir betrachten ein kartesisches Koordinatensystem mit Ursprung  $O$  und Achsen  $x$ ,  $y$  und  $z$ . Dabei sind es zwei Punkte  $P_1$  mit den Koordinaten  $(x_1, y_1, z_1)$  und  $P_2$  mit den Koordinaten  $(x_2, y_2, z_2)$ , die den Vektor  $\vec{a} = \overrightarrow{P_1P_2}$  von  $P_1$  nach  $P_2$  definieren.

Definition: Als *Komponenten eines Vektors*  $\vec{A}$  bezüglich des Koordinatensystems  $O(x, y, z)$  bezeichnet man die Projektionen von  $\vec{A}$  auf die Achsen  $x$ ,  $y$  und  $z$ , d.h.,

$$\begin{aligned} a_1 &= x_2 - x_1 \\ a_2 &= y_2 - y_1 \\ a_3 &= z_2 - z_1 \end{aligned}$$

und wir schreiben  $\vec{a} = (a_1, a_2, a_3)$ .

Man achte darauf, dass die Koordinaten des Vektors  $\vec{a} = \overrightarrow{P_1P_2}$ , also mit Anfangspunkt  $P_1$  und Endpunkt  $P_2$ , genau der Unterschied der Koordinaten vom Endpunkt minus Anfangspunkt sind. Demzufolge haben entgegengesetzte Vektoren Koordinaten mit gleichem absoluten Wert, aber entgegengesetzten Vorzeichen:

$$\overrightarrow{P_1P_2} = \vec{a} = (a_1, a_2, a_3) \quad \Rightarrow \quad \overrightarrow{P_2P_1} = -\vec{a} = (-a_1, -a_2, -a_3).$$

Aus der Definition der Komponenten folgt, dass zwei Vektoren gleich sind, wenn und nur wenn sie die gleichen Komponenten haben, und zwar bei jedem Koordinatensystem. Dies ist natürlich auch die Definition der Gleichheit zweier Elemente von  $\mathbb{R}^n$  (siehe Abschnitt 3.2).

Mit den Komponenten eines Vektors  $\vec{a} = (a_1, a_2, a_3)$  lässt sich der Betrag  $|\vec{a}| = a$  sofort berechnen. Es gilt

$$|\vec{a}| = a = \sqrt{a_1^2 + a_2^2 + a_3^2}.$$

Beispiele:

- i) Der Translationsvektor von Punkt  $P_1 = (-2, -1, 2)$  und  $P_2 = (1, -1, -2)$  ist

$$\vec{a} = \overrightarrow{P_1P_2} = (1 - (-2), -1 - (-1), -2 - 2) = (3, 0, -4)$$

und

$$a = |\vec{a}| = \sqrt{9 + 0 + 16} = 5.$$

- ii) Ein Vektor liegt auf der  $xy$ -Ebene, wenn seine Projektion auf der  $z$ -Achse null ist, also wenn die dritte Komponente  $a_3 = 0$  ist. Diese Eigenschaft hängt offensichtlich von der Wahl des Koordinatensystems ab.
- iii) Die Vektoren  $(1, 2, -3)$  und  $(-1, -2, 3)$  liegen auf einer Geraden, die durch den Ursprung  $O = (0, 0, 0)$  geht, haben den gleichen Betrag und entgegengesetzte Richtungen.

*Richtungskosinuse*

Definition: Als Richtungskosinuse eines Vektors  $\vec{a}$  bezüglich eines Koordinatensystems  $O(x, y, z)$  bezeichnet man die Kosinuse der Winkel zwischen dem Vektor  $\vec{a}$  und den positiven Richtungen der entsprechenden Koordinatenachsen.

Mittels der Komponenten können wir die Richtungskosinuse wie folgt berechnen:

$$\cos \alpha = \frac{a_1}{a} = \frac{a_1}{\sqrt{a_1^2 + a_2^2 + a_3^2}}, \quad \cos \beta = \frac{a_2}{a}, \quad \text{und} \quad \cos \gamma = \frac{a_3}{a}.$$

Daraus schließt man, dass ein Vektor (in Richtung und Betrag) vollständig definiert ist, wenn man seine Komponenten kennt.

Für jedes  $\vec{a}$  gilt dann die fundamentale Beziehung zwischen den Richtungskosinusen

$$\cos^2 \alpha + \cos^2 \beta + \cos^2 \gamma = 1.$$

In anderen Worten, der Vektor  $\frac{\vec{a}}{|\vec{a}|} = \frac{\vec{a}}{a} = (\cos \alpha, \cos \beta, \cos \gamma)$  hat immer die Länge 1. Nur zwei Richtungskosinuse und das Vorzeichen des dritten sind unabhängig.

Beispiele:

i) Für den Vektor  $\vec{a} = (1, -2, 2)$  ist  $a = \sqrt{1 + 4 + 4} = 3$  und

$$\cos \alpha = \frac{1}{3} \quad \cos \beta = \frac{-2}{3} \quad \cos \gamma = \frac{2}{3}.$$

ii) Wenn ein Vektor den Betrag  $a = 5$  hat und die Richtungskosinuse gleich  $\left(\frac{-1}{9}, \frac{4}{9}, \frac{8}{9}\right)$  sind, dann lauten seine Komponenten  $\vec{a} = \left(\frac{-5}{9}, \frac{20}{9}, \frac{40}{9}\right)$ .

*Parallele Vektoren*

Zwei Vektoren  $\vec{a}$  und  $\vec{b}$  sind parallel und haben die gleiche Richtung, wenn sie die gleichen Richtungskosinuse haben. Wenn die Richtungen entgegengesetzt sind, dann haben die Kosinuse gleiche Beträge und entgegengesetzte Vorzeichen. Das heißt

$$\frac{a_1}{a} = \cos \alpha = \pm \frac{b_1}{b}, \quad \frac{a_2}{a} = \cos \beta = \pm \frac{b_2}{b} \quad \text{und} \quad \frac{a_3}{a} = \cos \gamma = \pm \frac{b_3}{b},$$

wobei  $+$  ( $-$ ) für gleiche (entgegengesetzte) Richtungen gilt. Daraus folgt

$$\pm \frac{a}{b} = \frac{a_1}{b_1} = \frac{a_2}{b_2} = \frac{a_3}{b_3}.$$

Zwei Vektoren sind parallel, wenn und nur wenn ihre Komponenten proportional sind, d.h., wenn

$$\vec{a} = (a_1, a_2, a_3) = (\lambda b_1, \lambda b_2, \lambda b_3) = \lambda (b_1, b_2, b_3) = \lambda \vec{b}.$$

Beispiele:  $\vec{a} = (2, -4, 6)$  und  $\vec{b} = (1, -2, 3)$  oder  $\vec{b} = (-6, 12, -18)$ .

### 2.3 Winkel zwischen zwei Vektoren

Um den Winkel zwischen zwei Vektoren  $\vec{a}$  und  $\vec{b}$  zu berechnen, brauchen wir die Projektion von  $\vec{b}$  auf die Halbgerade, die  $\vec{a}$  enthält. Dazu sollten wir vorher im klaren darüber sein, dass die Projektion der Summe von Vektoren gleich der Summe der Projektionen ist.

Für zwei beliebige Vektoren  $\vec{a} = (a_1, a_2, a_3)$  und  $\vec{b} = (b_1, b_2, b_3)$ , die einen Winkel  $\theta$  bilden, lässt sich die Projektion von  $\vec{b}$  auf  $\vec{a}$  auf zwei verschiedene Weisen berechnen. Einerseits ist sie gleich  $b \cos \theta$ , wobei  $b = |\vec{b}|$  ist. Andererseits können wir  $\vec{b} = (b_1, b_2, b_3) = (b_1, 0, 0) + (0, b_2, 0) + (0, 0, b_3)$  als die Summe seiner Komponenten schreiben und jede Komponente, die natürlich auf den Achsen liegen, auf  $\vec{a}$  projizieren. Da die Kosinuse der Winkel zwischen  $\vec{a}$  und den Achsen gleich der Richtungskosinuse  $\cos \alpha = \frac{a_1}{a}$ ,  $\cos \beta = \frac{a_2}{a}$ , und  $\cos \gamma = \frac{a_3}{a}$  sind, gilt

$$\begin{aligned}
b \cos \theta &= b_1 \cos \alpha + b_2 \cos \beta + b_3 \cos \gamma \\
&= b_1 \frac{a_1}{a} + b_2 \frac{a_2}{a} + b_3 \frac{a_3}{a}.
\end{aligned}$$

Der Winkel zwischen zwei Vektoren ist also durch folgenden Ausdruck gegeben:

$$\cos \theta = \frac{a_1 b_1 + a_2 b_2 + a_3 b_3}{a b}.$$

Man verifiziert sofort, dass  $\cos \theta = \pm 1$  wenn  $\vec{a}$  parallel zu  $\vec{b}$  ist. Aus der Formel für  $\cos \theta$  schließt man auch, dass zwei Vektoren senkrecht zueinander stehen ( $\vec{a} \perp \vec{b}$ ), wenn und nur wenn

$$a_1 b_1 + a_2 b_2 + a_3 b_3 = 0$$

ist.

Beispiele:

i) Für  $\vec{a} = (-1, 0, 1)$  und  $\vec{b} = (2, 1, -3)$  ist

$$\cos \theta = \frac{-2 + 0 - 3}{\sqrt{2} \sqrt{4 + 1 + 9}} = \frac{-5}{2\sqrt{7}}.$$

ii) Die Vektoren  $(1, 1, 1)$ ,  $(1, -1, 0)$  und  $(1, 1, -2)$  stehen senkrecht zueinander.

## 2.4 Summe von Vektoren

Die Summe von Vektoren lässt sich intuitiv als aufeinander folgende Translation oder als die Zusammensetzung von Kräften definieren.

Bei Translationen lassen sich die assoziativen und kommutativen Gesetze “experimentell” prüfen. Dasselbe gilt für komponentenweise Summe von Vektoren. In beiden Fällen erhält man als Resultat, dass die Komponenten von  $\vec{a} + \vec{b}$  die Summe der Komponenten von  $\vec{a}$  und  $\vec{b}$  sind:

$$\vec{a} + \vec{b} = (a_1 + b_1, a_2 + b_2, a_3 + b_3).$$

Daraus folgen die kommutativen und assoziativen Eigenschaften

$$\vec{a} + \vec{b} = \vec{b} + \vec{a} \quad \text{und} \quad \vec{a} + (\vec{b} + \vec{c}) = (\vec{a} + \vec{b}) + \vec{c},$$

wie bei der Summe von Elementen von  $\mathbb{R}^n$ .

Die Subtraktion von Vektoren  $\vec{a} - \vec{b}$  ist nichts anderes als die Summe  $\vec{a} + (-\vec{b})$ . Bei der Summe mehrerer Vektoren  $\vec{a} + \vec{b} + \vec{c} + \dots + \vec{e}$  sind die Komponenten durch die Summe der einzelnen Komponenten gegeben

$$\vec{a} + \vec{b} + \vec{c} + \dots + \vec{e} = (a_1 + b_1 + \dots + e_1, a_2 + b_2 + \dots + e_2, a_3 + b_3 + \dots + e_3).$$

Es gelten, wie bei  $\mathbb{R}^n$  im allgemeinen, die assoziativen und kommutativen Gesetze der reellen Zahlen.

## 2.5 Multiplikation mit einem Skalar

Man definiert

$$\lambda \vec{a} = (\lambda a_1, \lambda a_2, \lambda a_3).$$

Es lässt sich leicht zeigen, dass  $|\lambda \vec{a}| = |\lambda| |\vec{a}|$ , und dass  $\lambda \vec{a}$  immer parallel zu  $\vec{a}$  ist. Die Richtung von  $\lambda \vec{a}$  ist entgegengesetzt zu  $\vec{a}$ , wenn  $\lambda < 0$ . Man kann sich davon leicht überzeugen, wenn man die Richtungskosinuse von  $\lambda \vec{a}$  berechnet

$$\frac{(\lambda \vec{a})_i}{|\lambda \vec{a}|} = \frac{\lambda a_i}{|\lambda| a} = \text{sign}(\lambda) \cdot \frac{a_i}{a},$$

wobei  $\vec{a} = (a_1, a_2, a_3)$  und  $a = \left(\sum_{i=1}^3 a_i^2\right)^{1/2}$ . Damit lässt sich der Einheitsvektor in Richtung  $\vec{a}$  definieren:

$$\hat{a} = \frac{\vec{a}}{|\vec{a}|} = \left(\frac{a_1}{a}, \frac{a_2}{a}, \frac{a_3}{a}\right) = (\cos \alpha, \cos \beta, \cos \gamma).$$

Von besonderem Interesse sind die Einheitsvektoren in den Richtungen der Achse des Koordinatensystems:

$$\begin{aligned}\hat{x} &= \hat{e}_1 = \hat{i} = (1, 0, 0) \\ \hat{y} &= \hat{e}_2 = \hat{j} = (0, 1, 0) \\ \hat{z} &= \hat{e}_3 = \hat{k} = (0, 0, 1).\end{aligned}$$

Sie werden fundamentale Einheitsvektoren genannt. Damit lässt sich jeder Vektor als lineare Kombination von  $\hat{e}_1$ ,  $\hat{e}_2$  und  $\hat{e}_3$  schreiben:

$$\vec{a} = (a_1, a_2, a_3) = a_1 \hat{e}_1 + a_2 \hat{e}_2 + a_3 \hat{e}_3.$$

Beispiele:

i)  $\vec{a} = (-3, 1, 2)$  und  $\vec{b} = (0, 3, -5) \Rightarrow \vec{c} = \vec{a} + \vec{b} = (-3, 4, -3)$ .

ii) Seien 2 Vektoren  $\vec{a}$  und  $\vec{b}$  mit  $|\vec{b}| < 2|\vec{a}|$ . Wann ist  $|\vec{a} + \vec{b}| < |\vec{a}|$ ?

Um dies zu klären, nutzen wir zunächst aus, dass  $|\vec{a} + \vec{b}| < |\vec{a}| \Leftrightarrow |\vec{a} + \vec{b}|^2 < |\vec{a}|^2$ . Die Bedingung lässt sich wie folgt umschreiben:

$$\begin{aligned}(a_1 + b_1)^2 + (a_2 + b_2)^2 + (a_3 + b_3)^2 &< a_1^2 + a_2^2 + a_3^2 \\ a_1^2 + b_1^2 + 2a_1b_1 + a_2^2 + b_2^2 + 2a_2b_2 + a_3^2 + b_3^2 + 2a_3b_3 &< a_1^2 + a_2^2 + a_3^2 \\ b_1^2 + b_2^2 + b_3^2 + 2(a_1b_1 + a_2b_2 + a_3b_3) &< 0 \\ b^2 + 2a b \cos \theta &< 0 \\ \cos \theta &< \frac{-b}{2a}.\end{aligned}$$

Man merke, dass die Bedingung  $\frac{b}{2a} < 1 \Rightarrow \frac{-b}{2a} > -1$ . Das Resultat lässt sich grafisch einfach interpretieren, wenn man die Ebene betrachtet, die die beiden Vektoren enthält ( $\cos \theta = -\sin \alpha$ ).

## 2.6 Skalarprodukt

Das Skalarprodukt oder innere Produkt  $\vec{a} \cdot \vec{b}$  zweier Vektoren  $\vec{a}$  und  $\vec{b}$  ist definiert als

$$\vec{a} \cdot \vec{b} = a b \cos \theta,$$

wobei  $a = |\vec{a}|$ ,  $b = |\vec{b}|$ , und  $\theta$  der Winkel zwischen den zwei Vektoren ist. In Abschnitt 2.3 haben wir gezeigt, dass

$$a b \cos \theta = a_1 b_1 + a_2 b_2 + a_3 b_3,$$

wobei  $\theta$  der Winkel zwischen  $\vec{a} = (a_1, a_2, a_3)$  und  $\vec{b} = (b_1, b_2, b_3)$  ist. Daraus folgt, dass das Skalarprodukt sich mittels der Komponenten der Vektoren wie folgt schreiben lässt:

$$\vec{a} \cdot \vec{b} = a_1 b_1 + a_2 b_2 + a_3 b_3.$$

Folgende fundamentale Eigenschaften lassen sich schnell beweisen:

i) Das Skalarprodukt ist kommutativ:

$$\vec{a} \cdot \vec{b} = \vec{b} \cdot \vec{a}.$$

ii) Das Skalarprodukt ist distributiv:

$$\begin{aligned} \vec{a} \cdot (\vec{b} + \vec{c}) &= \sum_i a_i (b_i + c_i) = \sum_i a_i b_i + \sum_i a_i c_i \\ &= \vec{a} \cdot \vec{b} + \vec{a} \cdot \vec{c}. \end{aligned}$$

iii) Darüber hinaus gilt

$$\begin{aligned} \vec{a} \cdot (\lambda \vec{b}) &= \sum_i a_i (\lambda b_i) = \lambda \sum_i a_i b_i \\ &= \lambda \vec{a} \cdot \vec{b} \quad \forall \lambda \in \mathbb{R}. \end{aligned}$$

Man kann ii) und iii) mit Hilfe von i) zusammenfassen, indem man sagt, dass das Skalarprodukt eine bilineare Funktion zweier Vektoren ist. Das heisst

$$\vec{a} \cdot (\vec{b} + \lambda \vec{c}) = \vec{a} \cdot \vec{b} + \lambda \vec{a} \cdot \vec{c} \quad \forall \lambda \in \mathbb{R}.$$

iv) Es gilt  $\vec{a} \cdot \vec{a} \geq 0 \forall \vec{a}$ , und  $\vec{a} \cdot \vec{a} = 0 \Leftrightarrow \vec{a} = \vec{0}$ . Der Betrag eines Vektors lässt sich wie folgt rechnen:

$$|\vec{a}| = a = \sqrt{\vec{a} \cdot \vec{a}}.$$

v) Zwei Vektoren sind senkrecht zueinander, wenn und nur wenn  $\vec{a} \cdot \vec{b}$  ist.

vi) Für die fundamentalen Einheitsvektoren  $\hat{e}_1 = (1, 0, 0)$ ,  $\hat{e}_2 = (0, 1, 0)$  und  $\hat{e}_3 = (0, 0, 1)$  gilt

$$\hat{e}_i \cdot \hat{e}_j = \delta_{ij} = \begin{cases} 1 & \text{wenn } i = j \\ 0 & \text{wenn } i \neq j. \end{cases}$$

Dabei haben wir das Kronecker-Delta oder den Kronecker-Tensor  $\delta_{ij}$  eingeführt.

vii) Mit Hilfe des Skalarprodukts lässt sich der Winkel zwischen zwei Vektoren als

$$\cos \theta = \frac{\vec{a} \cdot \vec{b}}{a b}$$

berechnen ( $a \neq 0$  und  $b \neq 0$ ).

viii) Die Projektion  $\vec{b}'$  eines Vektors  $\vec{b}$  auf die Halbgerade, die durch  $\vec{a}$  definiert ist, ist also

$$\vec{b}' = \frac{\vec{a}}{a^2} (\vec{a} \cdot \vec{b}),$$

wobei  $b' = b \cos \theta = \frac{\vec{a}}{a} \cdot \vec{b} = \hat{a} \cdot \vec{b}$ .

ix) Man kann also die Komponenten eines Vektors  $\vec{a} = (a_1, a_2, a_3)$  mit Hilfe des Skalarprodukts ausdrücken. Aus

$$\vec{a} = \sum_{i=1}^3 a_i \hat{e}_i$$

folgt

$$\vec{a} \cdot \hat{e}_j = \sum_i a_i \hat{e}_i \cdot \hat{e}_j = \sum_i a_i \delta_{ij} = a_j.$$

x) Es gilt die fundamentale Cauchy-Schwarz-Ungleichung

$$|\vec{a} \cdot \vec{b}| \leq |\vec{a}| |\vec{b}|.$$

Darüber hinaus gilt  $\vec{a} \cdot \vec{b} = |\vec{a}| |\vec{b}|$  nur, wenn  $\vec{a} = \lambda \vec{b}$  mit  $\lambda \in \mathbb{R}$ , d.h., wenn  $\vec{a}$  und  $\vec{b}$  parallel sind.

Dies ist einerseits klar aus der Definition von  $\vec{a} \cdot \vec{b} = a b \cos \theta$ , da  $|\cos \theta| \leq 1$  ist, und  $|\cos \theta| = 1$  nur wenn  $\vec{a} \parallel \vec{b}$  ist. Andererseits ist es instruktiv, sie allein aus den fundamentalen Eigenschaften i) - iv) herzuleiten. Man betrachtet den auxiliären Vektor

$$\vec{x} = (\vec{a} \cdot \vec{b}) \vec{a} - a^2 \vec{b},$$

für den

$$0 \leq \vec{x} \cdot \vec{x} = [(\vec{a} \cdot \vec{b}) \vec{a} - a^2 \vec{b}] \cdot [(\vec{a} \cdot \vec{b}) \vec{a} - a^2 \vec{b}]$$

dank iv) gilt. Mit Hilfe von i) - iii) erhalten wir nun

$$\begin{aligned} 0 &\leq (\vec{a} \cdot \vec{b})^2 a^2 + a^4 b^2 - 2 (\vec{a} \cdot \vec{b}) a^2 (\vec{a} \cdot \vec{b}) \\ 0 &\leq a^4 b^2 - (\vec{a} \cdot \vec{b})^2 a^2 \\ a^2 (\vec{a} \cdot \vec{b})^2 &\leq a^4 b^2. \end{aligned}$$

Angenommen  $a \neq 0$  (bei  $a = 0$  gilt ja die Ungleichung), haben wir

$$(\vec{a} \cdot \vec{b})^2 \leq a^2 b^2 \Leftrightarrow |\vec{a} \cdot \vec{b}| \leq a b.$$

Die Gleichheit  $|\vec{a} \cdot \vec{b}| = a b$  gilt nur, wenn  $\vec{x} \cdot \vec{x} = 0 \Rightarrow \vec{x} = (\vec{a} \cdot \vec{b}) \vec{a} - a^2 \vec{b} = \vec{0}$ . Dies folgt aus iv). Dann ist aber  $(\vec{a} \cdot \vec{b}) \vec{a} = a^2 \vec{b}$ , also  $\vec{a} \parallel \vec{b}$ .

## Beispiele:

i) Für  $\vec{a} = (-1, -1, 3)$  und  $\vec{b} = (-2, 1, -2)$  ist

$$\begin{aligned}\vec{a} \cdot \vec{b} &= 2 - 1 - 6 = -5, \\ a^2 &= \vec{a} \cdot \vec{a} = 1 + 1 + 9 = 11, \\ b^2 &= \vec{b} \cdot \vec{b} = 4 + 1 + 4 = 9 \quad \text{und} \\ \cos \theta &= \frac{\vec{a} \cdot \vec{b}}{ab} = \frac{-5}{3\sqrt{11}}.\end{aligned}$$

ii) Wir betrachten ein Viereck, das aus zwei Vektoren entsteht. Die Diagonalen sind also  $\vec{a} + \vec{b}$  und  $\vec{a} - \vec{b}$ . Wann sind die Diagonalen senkrecht?

Im Allgemeinen gilt  $(\vec{a} + \vec{b}) \cdot (\vec{a} - \vec{b}) = \vec{a} \cdot \vec{a} - \vec{b} \cdot \vec{b} = a^2 - b^2$ . Die Diagonalen sind senkrecht zueinander, wenn  $a = b$  ist, d.h., wenn das Viereck regulär ist.

iii) Gegeben sind zwei Vektoren  $\vec{a} = (1, 1, 1)$  und  $\vec{b} = (2, -1, 0)$ . Man kann  $\vec{b}$  als die Summe von zwei Vektoren schreiben, von denen der eine parallel und der andere senkrecht zu  $\vec{a}$  ist. Man projiziert zunächst  $\vec{b}$  auf  $\frac{\vec{a}}{a}$

$$\begin{aligned}\vec{b}' &= \frac{\vec{a}}{a} \left( \vec{b} \cdot \frac{\vec{a}}{a} \right) = \vec{a} \frac{(\vec{b} \cdot \vec{a})}{a^2} \\ \left. \begin{array}{l} a^2 = 3 \\ \vec{a} \cdot \vec{b} = 1 \end{array} \right\} &\Rightarrow \vec{b}' = \frac{1}{3} (1, 1, 1) \\ \vec{b} &= \vec{b}' + (\vec{b} - \vec{b}') = \left( \frac{1}{3}, \frac{1}{3}, \frac{1}{3} \right) + \left( \frac{5}{3}, \frac{-4}{3}, \frac{-1}{3} \right),\end{aligned}$$

wobei  $\left( \frac{5}{3}, \frac{-4}{3}, \frac{-1}{3} \right) \cdot (1, 1, 1) = 0$  ist. Diese Prozedur lässt sich graphisch leicht interpretieren, wenn man die Ebene betrachtet, die beide Vektoren enthält.

Es gilt im Allgemeinen

$$\vec{b} = \vec{b}_{\parallel} + \vec{b}_{\perp}$$

mit  $\vec{b}_{\parallel} = \frac{(\vec{b} \cdot \vec{a})}{a} \frac{\vec{a}}{a}$  und  $\vec{b}_{\perp} = \vec{b} - \vec{b}_{\parallel}$ .  $\vec{b}_{\parallel}$  ist per Konstruktion parallel zu  $\vec{a}$ , und  $\vec{a} \cdot \vec{b}_{\perp} = \vec{a} \cdot \vec{b} - \vec{a} \cdot \vec{b}_{\parallel} = \vec{a} \cdot \vec{b} - \vec{b} \cdot \vec{a} = 0$ .

## 2.7 Transformationsregel für Vektoren bei Änderungen des Koordinatensystems

Wir betrachten eine Koordinaten-Transformation zwischen einem Koordinatensystem  $\hat{e}_1, \hat{e}_2$  und  $\hat{e}_3$  mit den Positionsvektoren

$$\vec{r} = (x_1, x_2, x_3) = \sum_j x_j \hat{e}_j$$

und ein Koordinatensystem  $\hat{e}'_1, \hat{e}'_2$  und  $\hat{e}'_3$  mit den Positionsvektoren

$$\vec{r}' = (x'_1, x'_2, x'_3) = \sum_k x'_k \hat{e}'_k.$$

Die Invarianz der Position und von vektoriellen Größen im Allgemeinen bezüglich Änderungen des Koordinatensystems bedeutet, dass

$$\vec{r} = \sum_j x_j \hat{e}_j = \sum_k x'_k \hat{e}'_k = \vec{r}'.$$

Die Transformationsregel für die Komponenten von Vektoren erhält man, indem man mit  $\hat{e}_j$  bzw.  $\hat{e}'_i$  skalar multipliziert:

$$x'_i = \sum_j \hat{e}'_i \cdot \hat{e}_j x_j$$

und

$$x_j = \sum_k \hat{e}_j \cdot \hat{e}'_k x'_k.$$

Eine physikalische oder geometrische Größe heißt *vektoriell*, wenn sie aus drei Komponenten  $(v_1, v_2, v_3) = \vec{v}$  besteht, die bei Änderungen des Koordinatensystems wie die Komponenten des Positionsvektors transformieren. Das heißt, wenn

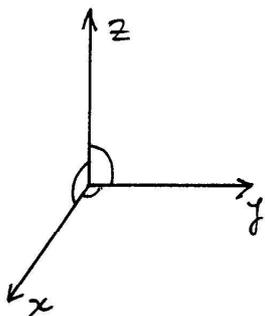
$$v'_i = \sum_j \hat{e}'_i \cdot \hat{e}_j v_j.$$

Eine Größe heißt *skalar*, wenn ihre Werte an jedem Punkt im Raum unverändert bei Koordinatentransformationen bleiben. Das heißt, wenn

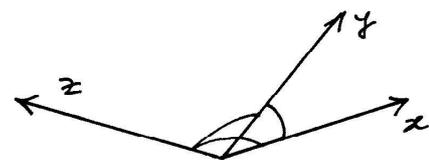
$$\varphi'(x'_1, x'_2, x'_3) = \varphi(x_1, x_2, x_3).$$

## 2.8 Die zwei Orientierungen eines Koordinatensystems

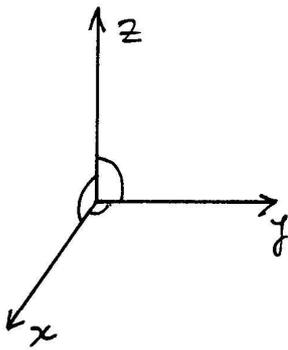
In drei Dimensionen lassen sich die Koordinatensysteme in zwei Gruppen oder Äquivalenzklassen teilen. Man sagt, dass zwei Koordinatensysteme  $O(x, y, z)$  und  $O'(x', y', z')$  mit den Achsen in der vorgegebenen Reihenfolge  $(x, y, z)$  oder  $(x', y', z')$  die gleiche Orientierung haben, wenn sie sich kontinuierlich durch Verzerrungen und Drehungen in einander bringen lassen, ohne dass dabei jemals die drei Achsen in einer Ebene liegen. Zum Beispiel haben



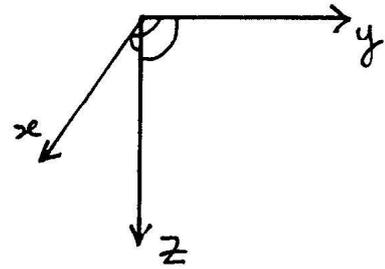
und



die gleiche Orientierung. Dagegen haben



und



(a) Rechtssystem, positive oder direkte Orientierung

(b) Linkssystem, negative oder inverse Orientierung

entgegengesetzte Orientierungen.

Man sagt, dass die Orientierung eines Koordinatensystems wie bei der Figur (a) positiv, direkt oder rechts ist. Man kann sich folgende intuitive Definition einprägen: Wenn die Drehung einer Schraube von der Achse  $x$  nach der Achse  $y$  dazu führt, dass die Schraube in Richtung der positiven  $z$ -Achse vorangeht, dann ist das Koordinatensystem ein Rechtssystem. In den anderen Fällen, wie bei der Figur (b), ist die Orientierung negativ, invers oder links.

Andere äquivalente Definitionen sind folgende: Nehmen Sie Ihre rechte Hand und bilden Sie ein "Koordinatensystem" mit Mittelfinger, Daumen und Zeigefinger senkrecht zueinander auf die "natürlichste Weise". Dann bildet die Reihenfolge Mittelfinger, Daumen, Zeigefinger ein Rechtssystem. Wenn man auf die  $xy$ -Ebene aus dem Halbraum mit positiven  $z$ -Koordinaten schaut, dann ist es so, dass bei Rechtssystemen eine Drehung der positiven  $x$ -Achse in die positive  $y$ -Achse gegen die Drehrichtung einer Uhr geht.

Es ist klar, dass folgende Transformationen eines Koordinatensystems

$$\begin{array}{ll} x \rightarrow -x & x \rightarrow y \\ y \rightarrow -y & \text{oder } y \rightarrow z \\ z \rightarrow -z & z \rightarrow z \end{array}$$

ein Rechtssystem in ein Linkssystem umwandeln. Es handelt sich um Transformationen, die man nicht kontinuierlich führen kann, ohne dass an einem Punkt die drei Achsen in einer Ebene liegen.

Im Gegensatz dazu lassen zyklische Transformationen der Achsen wie

$$\begin{array}{l} x \rightarrow y \\ y \rightarrow z \\ z \rightarrow x \end{array}$$

und die Inverse oder Quadrat davon

$$x \rightarrow z$$

$$y \rightarrow x$$

$$z \rightarrow y$$

die Orientierung unverändert.

## 2.9 Vektoren und Pseudovektoren

Wir können jetzt den Begriff “Vektor” verfeinern, indem wir das Verhalten von Vektoren bezüglich Änderungen der Orientierung des Koordinatensystems betrachten. Wir haben bereits definiert, dass Vektoren physikalische Größen beschreiben (z.B. Position, Geschwindigkeit, Kräfte oder Beschleunigung) und deshalb unabhängig vom Koordinatensystem sind. Ein Vektor bleibt unverändert, wenn man das Koordinatensystem dreht. Ebenfalls bei einer Änderung der Orientierung des Systems (z.B. bei einer Inversion der Koordinatenachsen  $x \rightarrow -x$ ,  $y \rightarrow -y$  und  $z \rightarrow -z$ ) bleiben Vektoren unverändert. Es gibt aber physikalische Größen, die bei einer Änderung der Orientierung des Systems ihre Richtung ändern, und zwar genau entgegengesetzt von  $\vec{a}$  zu  $-\vec{a}$ . Man nennt sie Pseudovektoren.

Definition: Eine vektorielle Größe  $\vec{a}$  ist ein *Pseudovektor* oder *axialer Vektor*, wenn sie sich bei Drehungen des Koordinatensystems wie ein Vektor verhält (d.h., dass sie unverändert bleibt), und wenn sie bei Änderung der Orientierung des Koordinatensystems ihre Richtung von  $\vec{a}$  auf  $-\vec{a}$  ändert.

Pseudovektoren sind also vom Koordinatensystem abhängig. Allerdings, wenn die Orientierung des Koordinatensystems festliegt und unverändert bleibt, sind Pseudovektoren eindeutig definiert und ihre Eigenschaften (Summe, Multiplikation mit Skalaren, etc.) in allen Beziehungen mit denen der Vektoren identisch. Das wichtigste Beispiel eines Pseudovektors ist das Vektorprodukt zweier Vektoren. Wichtige davon abgeleitete physikalische Größen wie das Drehmoment einer Kraft, der Drehimpuls oder die Rotation eines Vektorfelds sind auch Pseudovektoren.

## 2.10 Das Vektorprodukt

Gegeben sei ein orientiertes Koordinatensystem  $O(x, y, z)$ . Wir definieren das Vektorprodukt zweier Vektoren  $\vec{a}$  und  $\vec{b}$  als den Pseudovektor  $\vec{c} = \vec{a} \times \vec{b}$  mit folgenden Eigenschaften:

a) Der Betrag von

$$|\vec{a} \times \vec{b}| = |\vec{a}| |\vec{b}| \sin \theta,$$

wobei  $\theta$  der Winkel zwischen  $\vec{a}$  und  $\vec{b}$  ist.

b) Der Pseudovektor  $\vec{a} \times \vec{b}$  ist senkrecht zur von  $\vec{a}$  und  $\vec{b}$  definierten Ebene. Das heißt,  $\vec{a} \times \vec{b}$  ist senkrecht zu  $\vec{a}$  und zu  $\vec{b}$ .

- c) Die Richtung von  $\vec{a} \times \vec{b}$  ist die, bei der das Dreibein  $\vec{a}, \vec{b}$  und  $\vec{a} \times \vec{b}$  (in dieser Reihenfolge) die gleiche Orientierung wie das Koordinatensystem hat.

Der Vollständigkeit halber sollte man dazu sagen, dass  $\vec{0} \times \vec{a} = \vec{a} \times \vec{0} = \vec{0} \forall \vec{a}$  ist. Die Bedingung c) gibt dem Vektorprodukt seinen Pseudovektor-Charakter.

Fundamentale Eigenschaften des Vektorproduktes:

i)  $\vec{a} \times \vec{b} = -\vec{b} \times \vec{a}$ .

Aus a) folgt  $|\vec{b} \times \vec{a}| = ab \sin \theta$ , und die von  $\vec{a}$  und  $\vec{b}$  aufgespannte Ebene hängt ja nicht von der Reihenfolge der Vektoren  $\vec{a}$  und  $\vec{b}$  ab. Allerdings, wenn  $\vec{a}, \vec{b}$  und  $\vec{a} \times \vec{b}$  eine Orientierung haben (z.B. ein Rechtssystem bilden), dann hat  $\vec{b}, \vec{a}$  und  $\vec{a} \times \vec{b}$  die entgegengesetzte Orientierung. Es muss also  $\vec{b} \times \vec{a} = -\vec{a} \times \vec{b}$  gelten, damit das Dreibein  $\vec{b}, \vec{a}, -\vec{a} \times \vec{b}$  die gleiche Orientierung wie  $\vec{a}, \vec{b}, \vec{a} \times \vec{b}$  und das Koordinatensystem hat. Man sagt, dass das Vektorprodukt antikommutativ ist ( $\vec{a} \times \vec{b} = -\vec{b} \times \vec{a}$ ).

- ii) Für alle Skalare  $\lambda \in \mathbb{R}$  ist

$$(\lambda \vec{a}) \times \vec{b} = \vec{a} \times (\lambda \vec{b}) = \lambda (\vec{a} \times \vec{b}).$$

Aus a) folgt

$$\begin{aligned} |(\lambda \vec{a}) \times \vec{b}| &= |\lambda \vec{a}| |\vec{b}| \sin \theta \\ &= |\lambda| |\vec{a}| |\vec{b}| \sin \theta = |\vec{a}| |\lambda \vec{b}| \sin \theta. \end{aligned}$$

Da  $\sin \theta = \sin(\pi - \theta)$ , ändert sich  $\sin \theta$  selbst bei  $\lambda < 0$  nicht. Deshalb gilt  $|(\lambda \vec{a}) \times \vec{b}| = |\vec{a} \times (\lambda \vec{b})| = |\lambda (\vec{a} \times \vec{b})|$ .

Bei  $\lambda > 0$  ändern sich die Richtungen von  $\lambda \vec{a}, \lambda \vec{b}$  und  $\lambda (\vec{a} \times \vec{b})$  nicht. Bei  $\lambda < 0$  ändern sich die Richtungen von  $\lambda \vec{a}$  und  $\lambda \vec{b}$ , dafür auch die Orientierung des Dreibeins  $\lambda \vec{a}, \lambda \vec{b}$  und  $\vec{a} \times \vec{b}$  oder  $\vec{a}, \lambda \vec{b}$  und  $\vec{a} \times \vec{b}$ . Damit die Orientierung von  $\lambda \vec{a}, \lambda \vec{b}$  und  $(\lambda \vec{a}) \times \vec{b}$  die gleiche bleibt, muss die Richtung von  $(\lambda \vec{a}) \times \vec{b}$  bei  $\lambda < 0$  entgegengesetzt zur Richtung von  $\vec{a} \times \vec{b}$  sein. Die Beziehung

$$\lambda \vec{a} \times \vec{b} = \vec{a} \times \lambda \vec{b} = \lambda \vec{a} \times \vec{b}$$

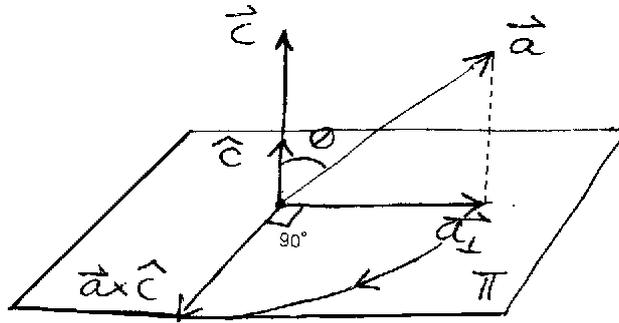
gilt also  $\forall \lambda \in \mathbb{R}$ . Man merkt, wie wichtig es ist, dass die Orientierung von  $\vec{a}, \vec{b}$  und  $\vec{a} \times \vec{b}$  wie in der Definition von  $\vec{a} \times \vec{b}$  durch die Orientierung des Koordinatensystems festgelegt wird, damit diese fundamentale algebraische Eigenschaft gilt.

- iii) Das Vektorprodukt ist distributiv bezüglich der Summe von Vektoren, d.h.,

$$(\vec{a} + \vec{b}) \times \vec{c} = \vec{a} \times \vec{c} + \vec{b} \times \vec{c}.$$

Der Beweis dafür ist instruktiv, da er eine geometrische Sicht auf das Vektorprodukt liefert. Wir betrachten der Einfachheit halber einen auf 1 normierten Vektor  $\hat{c} = \frac{\vec{c}}{|\vec{c}|}$ .

Es ist klar, dass wenn das Distributivgesetz für  $\hat{c}$  gilt, dann gilt es durch Multiplikation mit  $|\vec{c}|$  auch für ein beliebiges  $\vec{c}$ . Die Figur zeigt die Projektion  $\vec{a}_\perp$  von  $\vec{a}$  auf die Ebene  $\pi$ , die senkrecht zu  $\hat{c}$  ist.



$|\vec{a}_\perp| = a \sin \theta = |\vec{a} \times \hat{c}|$ , da  $|\hat{c}| = 1$ . Das zeigt, dass  $\vec{a} \times \hat{c}$  nichts anderes ist als die Projektion  $\vec{a}_\perp$  von  $\vec{a}$  auf die zu  $\hat{c}$  senkrechte Ebene gefolgt von einer  $90^\circ$ -Drehung um  $\hat{c}$ , und zwar so, dass der gedrehte Vektor  $\vec{a} \times \hat{c}$ ,  $\vec{a}_\perp$  und  $\hat{c}$  die selbe Orientierung haben wie das Koordinatensystem.

Andererseits wissen wir, dass die Projektion und die Drehung von Vektoren lineare Abbildungen sind. Das heißt, für jede Projektion  $P_\pi$  auf eine beliebige Ebene  $\pi$  gilt

$$P_\pi(\vec{a} + \vec{b}) = P_\pi(\vec{a}) + P_\pi(\vec{b}).$$

Wenn das nicht so wäre, wären die Bilder im Kino verzerrt. Dieselbe Eigenschaft haben Drehungen, d.h., für jede Drehung  $D_{\varphi, \hat{n}}$  mit einem Winkel  $\varphi$  um eine beliebige Achse  $\hat{n}$ , die durch den Ursprung der Vektoren geht, gilt

$$D_{\varphi, \hat{n}}(\vec{a} + \vec{b}) = D_{\varphi, \hat{n}}(\vec{a}) + D_{\varphi, \hat{n}}(\vec{b}).$$

Es ist auch klar, dass die Komposition von linearen Abbildungen linear ist, d.h.,

$$D[P_\pi(\vec{a} + \vec{b})] = D[P_\pi(\vec{a}) + P_\pi(\vec{b})] = D[P_\pi(\vec{a})] + D[P_\pi(\vec{b})].$$

Deshalb ist auch die Vektormultiplikation mit  $\hat{c}$  linear, d.h.,

$$(\vec{a} + \vec{b}) \times \hat{c} = \vec{a} \times \hat{c} + \vec{b} \times \hat{c}.$$

Das Distributivgesetz ist eine weitere sehr wichtige Eigenschaft des Vektorprodukts, die uns das Herleiten eines geschlossenen Ausdrucks für die kartesischen Koordinaten von  $\vec{a} \times \vec{b}$  mittels der kartesischen Koordinaten von  $\vec{a}$  und  $\vec{b}$  ermöglichen wird.

Als Zusammenfassung von ii) und iii) haben wir

$$(\alpha \vec{a} + \beta \vec{b}) \times (\gamma \vec{c}) = \alpha \gamma (\vec{a} \times \vec{c}) + \beta \gamma (\vec{b} \times \vec{c})$$

$\forall \alpha, \beta, \gamma \in \mathbb{R}$  und  $\forall \vec{a}, \vec{b}$  und  $\vec{c} \in \mathbb{R}^3$ .

iv) Das Vektorprodukt  $\vec{a} \times \vec{b}$  ist gleich dem Nullvektor  $\vec{0}$ , wenn und nur wenn  $\vec{a}$  und  $\vec{b}$  parallel sind.

$$\vec{a} \times \vec{b} = \vec{0} \Leftrightarrow \vec{a} = \lambda \vec{b} \quad \text{mit} \quad \lambda \in \mathbb{R} \quad (\vec{b} \neq \vec{0}).$$

Wenn  $a = 0$  oder  $b = 0$ , gilt die Eigenschaft offensichtlich. Für  $a \neq 0$  und  $b \neq 0$  ist  $|\vec{a} \times \vec{b}| = a b \sin \theta = 0 \Rightarrow \theta = 0$  oder  $\pi$ , und deshalb ist  $\cos \theta = \pm 1$ , d.h.,  $\vec{a}$  parallel zu  $\vec{b}$ .

- v) Der Betrag von  $\vec{a} \times \vec{b}$  ist gleich der Fläche des von  $\vec{a}$  und  $\vec{b}$  gebildeten Parallelogramms (Fläche =  $a \sin \theta b$ ). Damit lässt sich auch iv) beweisen.

## 2.11 Kartesische Komponente des Vektorprodukts

Gegeben sind zwei Vektoren  $\vec{a} = (a_1, a_2, a_3)$  und  $\vec{b} = (b_1, b_2, b_3)$ . Mittels der Einheitsvektoren des Koordinatensystems  $\hat{e}_1 = \hat{x} = (1, 0, 0)$ ,  $\hat{e}_2 = \hat{y} = (0, 1, 0)$  und  $\hat{e}_3 = \hat{z} = (0, 0, 1)$  lassen sich

$$\vec{a} = a_1 \hat{e}_1 + a_2 \hat{e}_2 + a_3 \hat{e}_3 = \sum_{i=1}^3 a_i \hat{e}_i$$

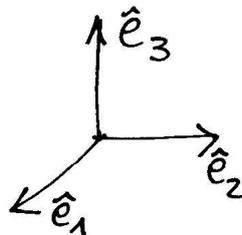
und

$$\vec{b} = \sum_{j=1}^3 b_j \hat{e}_j$$

berechnen. Wenn wir ii) und iii) anwenden, erhalten wir

$$\vec{a} \times \vec{b} = \left( \sum_i a_i \hat{e}_i \right) \times \left( \sum_j b_j \hat{e}_j \right) = \sum_{ij} a_i b_j (\hat{e}_i \times \hat{e}_j).$$

Wir nehmen jetzt an, dass wir es mit einem Rechtssystem zu tun haben, wie in der Figur skizziert.



Ein Blick auf die Figur und die Tatsachen, dass  $|\hat{e}_i| = 1 \forall i$  und  $\sin \theta_{ij} = \sin \left( \frac{\pi}{2} \right) = 1$ , zeigt

$$\begin{aligned} \hat{e}_1 \times \hat{e}_2 &= \hat{e}_3 = -\hat{e}_2 \times \hat{e}_1 \\ \hat{e}_2 \times \hat{e}_3 &= \hat{e}_1 = -\hat{e}_3 \times \hat{e}_2 \\ \hat{e}_3 \times \hat{e}_1 &= \hat{e}_2 = -\hat{e}_1 \times \hat{e}_3 \end{aligned}$$

und natürlich

$$\hat{e}_i \times \hat{e}_i = 0 \forall i.$$

Damit erhalten wir

$$\vec{a} \times \vec{b} = (a_2 b_3 - a_3 b_2) \hat{e}_1 + (a_3 b_1 - a_1 b_3) \hat{e}_2 + (a_1 b_2 - a_2 b_1) \hat{e}_3, \quad (2.15)$$

was die kartesischen Komponenten von  $\vec{a} \times \vec{b}$  als Funktion der Komponenten von  $\vec{a}$  und  $\vec{b}$  angibt. Es lohnt sich, einen Blick auf diese wichtige Rechenregel zu werfen, um folgendes zu bemerken:

- i) Alle Terme haben die Form  $a_i b_j \hat{e}_k$  mit  $i, j$  und  $k$  alle unterschiedlich.
- ii) Wenn die Reihenfolge  $(i, j, k)$  bei  $a_i b_j \hat{e}_k$  gleich  $(1,2,3)$  oder eine zyklische Permutation davon [d.h.,  $(2,3,1)$  und  $(3,1,2)$ ] ist, dann ist das Vorzeichen positiv.
- iii) In den anderen drei Fällen, d.h.,  $(i, j, k) = (2, 1, 3)$  und den zyklischen Permutationen  $(3,2,1)$  und  $(1,3,2)$  davon, ist das Vorzeichen negativ.

Diese Eigenschaften lassen sich kompakt mit Hilfe des Levy-Civita-Symbols  $\varepsilon_{ijk}$ , das für Indizes  $1 \leq i, j, k \leq 3$  wie folgt definiert ist, zusammenfassen:

$$\varepsilon_{ijk} = \begin{cases} 1 & \text{für } (i, j, k) = (1, 2, 3) \text{ oder} \\ & \text{die zyklischen Permutationen } (2, 3, 1) \text{ und } (3, 1, 2) \\ -1 & \text{für } (i, j, k) = (2, 1, 3) \text{ oder} \\ & \text{die zyklischen Permutationen } (1, 3, 2) \text{ und } (3, 2, 1) \\ 0 & \text{in den anderen Fällen, d.h.,} \\ & \text{sobald zwei Indices gleich sind.} \end{cases}$$

$\varepsilon_{ijk}$  ist vollkommen antisymmetrisch, d.h., antisymmetrisch bezüglich der Vertauschung jedes Paares von Indizes:

$$\varepsilon_{ijk} = -\varepsilon_{jik} = -\varepsilon_{ikj} = -\varepsilon_{kji} \quad \forall i, j, k.$$

Daraus folgt natürlich  $\varepsilon_{iik} = \varepsilon_{iki} = \varepsilon_{kii} = 0 \quad \forall i, k$ . Zyklische Permutationen der Indizes, die sich aus 2 Vertauschungen ergeben, lassen  $\varepsilon_{ijk}$  unverändert (z.B.  $\varepsilon_{123} = -\varepsilon_{213} = \varepsilon_{231}$  ).

Man kann leicht verifizieren, dass

$$\hat{e}_i \times \hat{e}_j = \sum_k \varepsilon_{ijk} \hat{e}_k.$$

Daraus folgt

$$\vec{a} \times \vec{b} = \sum_{ijk=1}^3 \varepsilon_{ijk} a_i b_j \hat{e}_k. \quad (2.16)$$

Es wird dem Leser überlassen, sich an dieser Stelle durch explizite Durchführung der Summe davon zu überzeugen, dass der Ausdruck (2.16) mit (2.15) übereinstimmt.

Eine weitere äquivalente Form, das Vektorprodukt zu berechnen, ist die formale Entwicklung der  $3 \times 3$ -Matrix<sup>1</sup>

$$\vec{a} \times \vec{b} = \begin{vmatrix} \hat{e}_1 & \hat{e}_2 & \hat{e}_3 \\ a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \end{vmatrix} = \hat{e}_1 \begin{vmatrix} a_2 & a_3 \\ b_2 & b_3 \end{vmatrix} - \hat{e}_2 \begin{vmatrix} a_1 & a_3 \\ b_1 & b_3 \end{vmatrix} + \hat{e}_3 \begin{vmatrix} a_1 & a_2 \\ b_1 & b_2 \end{vmatrix},$$

wobei für  $2 \times 2$ -Matrizen gilt

$$\begin{vmatrix} a & b \\ c & d \end{vmatrix} = ad - bc, \quad \text{d.h.}, (\searrow) - (\swarrow).$$

Beispiele:

i) Wir betrachten  $\vec{a} = (1, -2, 0)$  und  $\vec{b} = (-3, 1, 4)$ . Dann ist

$$\vec{a} \times \vec{b} = \begin{vmatrix} \hat{e}_1 & \hat{e}_2 & \hat{e}_3 \\ 1 & -2 & 0 \\ -3 & 1 & 4 \end{vmatrix} = -8\hat{e}_1 - 4\hat{e}_2 - 5\hat{e}_3.$$

ii) Sei  $\vec{a} = (1, 1, 0)$  und  $\vec{b} = (-1, 2, 3)$ , dann ist  $a = |\vec{a}| = \sqrt{2}$  und  $b = |\vec{b}| = \sqrt{14}$ .

$$\vec{a} \times \vec{b} = \begin{vmatrix} \hat{e}_1 & \hat{e}_2 & \hat{e}_3 \\ 1 & 1 & 0 \\ -1 & 2 & 3 \end{vmatrix} = (3, -3, 3).$$

Man kann sofort prüfen, dass  $\vec{a} \times \vec{b} \perp \vec{a}$  und  $\perp \vec{b}$  ist, indem man die Skalarprodukte berechnet.

Der Betrag  $|\vec{a} \times \vec{b}| = 3\sqrt{3} = \sqrt{9+9+9}$ . Daraus folgt

$$\sin \theta = \frac{|\vec{a} \times \vec{b}|}{ab} = \frac{3\sqrt{3}}{\sqrt{28}} = 0.98$$

und  $\sin^2 \theta = \frac{27}{28}$ . Andererseits haben wir

$$\begin{aligned} \vec{a} \cdot \vec{b} &= \sum_i a_i b_i = -1 + 2 = 1 \quad \text{und} \\ \frac{\vec{a} \cdot \vec{b}}{ab} &= \cos \theta = \frac{1}{\sqrt{28}} = 0.19 \quad \Rightarrow \quad \cos^2 \theta = \frac{1}{28}, \end{aligned}$$

wie es sein sollte.

<sup>1</sup> Eine kurze elementare Einführung zu Matrizen und Determinanten befindet sich in Abschnitt 2.12.

Man sollte allerdings merken, dass man aus dem Vektorprodukt nicht eindeutig den Winkel bestimmen kann, wohl aber aus dem Skalarprodukt, da  $\sin \theta = \sin(\pi - \theta)$ , während  $\cos \theta$  injektiv für  $0 \leq \theta \leq \pi$  ist.

## 2.12 Matrizen und Determinanten

Für die praktische Berechnung von Vektorprodukt und Spatprodukt ist es nützlich, den Begriff der Matrix und der Determinante einer Matrix einzuführen. Wir werden uns auf die einfachsten Eigenschaften von Matrizen beschränken, nämlich ihrer Eigenschaften als Vektorraum oder linearer Raum über dem Körper der Reellen. Weitere Eigenschaften wie das Produkt von Matrizen und ihre Beziehung zu linearen Abbildungen sollen zu einem späteren Zeitpunkt besprochen werden. Ebenfalls bei der Determinanten quadratischer Matrizen beschränken wir uns auf die Definition und ihre geometrische Interpretation bei  $n = 2$ , und 3 Dimensionen.

### 2.12.1 Matrizen

Eine *Matrix*  $A$  ist definiert als eine Tabelle von reellen Zahlen  $a_{ij}$  mit  $1 \leq i \leq m$  und  $1 \leq j \leq n$ , wobei  $m$  und  $n$  ganze Zahlen sind. Man stellt sie zweckmäßig wie folgt vor:

$$A = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix} = (a_{ij}).$$

Die  $a_{ij}$  werden die Elemente oder Komponenten der Matrix  $A$  genannt. Diese Matrix besteht also aus  $m$  Zeilen  $i$  ( $1 \leq i \leq m$ )

$$(a_{i1} \ a_{i2} \ \dots \ a_{in}) \in \mathbb{R}^n$$

und  $n$  Spalten  $j$  ( $1 \leq j \leq n$ )

$$\begin{pmatrix} a_{1j} \\ a_{2j} \\ \dots \\ a_{mj} \end{pmatrix} \in \mathbb{R}^m.$$

Man sagt auch, dass  $A$  zu  $\mathbb{R}^{m \times n}$  gehört ( $A \in \mathbb{R}^{m \times n}$ ). Die Zeilen und Spalten einer Matrix können als Vektoren in  $\mathbb{R}^n$  und  $\mathbb{R}^m$  betrachtet werden.

Eine Matrix  $A$  heißt *quadratisch*, wenn  $n = m$ , also wenn  $A \in \mathbb{R}^{n \times n}$  (z.B.  $2 \times 2$ ,  $3 \times 3$ ).

Logischerweise definiert man Gleichheit von Matrizen folgendermaßen:

$$A, B \in \mathbb{R}^{m \times n} \quad A = B \Leftrightarrow a_{ij} = b_{ij} \quad \forall i, j.$$

Die Summe zweier Matrizen  $A = (a_{ij})$  und  $B = (b_{ij}) \in \mathbb{R}^{m \times n}$  ist die Matrix  $C \in \mathbb{R}^{m \times n}$  mit den Elementen

$$c_{ij} = a_{ij} + b_{ij} \quad \forall i, j \quad \begin{array}{l} 1 \leq i \leq m \\ 1 \leq j \leq n. \end{array}$$

Es ist klar, dass die Summe zweier Matrizen nur dann Sinn macht, wenn die Matrizen vom selben Typ sind, also mit gleichem  $n$  und  $m$ .

Für jede Matrix  $A = (a_{ij})$  und jedes  $\lambda \in \mathbb{R}$  ist das Produkt  $\lambda A$  die Matrix mit den Elementen

$$(\lambda A)_{ij} = \lambda a_{ij},$$

also

$$\lambda A = \begin{bmatrix} \lambda a_{11} & \lambda a_{12} & \dots & \lambda a_{1n} \\ \lambda a_{21} & \lambda a_{22} & \dots & \lambda a_{2n} \\ \vdots & \vdots & & \vdots \\ \lambda a_{m1} & \lambda a_{m2} & \dots & \lambda a_{mn} \end{bmatrix} \quad \lambda \in \mathbb{R}.$$

Beispiele:

Die Nullmatrix

$$0 = \begin{pmatrix} 0 & \dots & 0 \\ \vdots & & \vdots \\ 0 & \dots & 0 \end{pmatrix} \quad 0_{ij} = 0$$

mit der Eigenschaft  $A + 0 = A \quad \forall a \in \mathbb{R}^{n \times m}$ .

Die Identitätsmatrix oder Einheitsmatrix

$$\mathbb{1} = \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & \dots & 0 & 1 \end{pmatrix} \quad \mathbb{1}_{ij} = \delta_{ij} = \begin{cases} 1 & i = j \\ 0 & i \neq j. \end{cases}$$

Mit diesen Definitionen ist es klar, dass die Mengen der Matrizen  $A \in \mathbb{R}^{m \times n}$  einen Vektorraum oder linearen Raum über den reellen Zahlen bilden. Dieser Vektorraum ist isomorph mit dem  $l$ -dimensionalen Raum  $\mathbb{R}^l$  mit  $l = m \cdot n$ . Matrizen verhalten sich also bezüglich der Summe und der Multiplikation mit reellen Zahlen wie Vektoren eines Raumes höherer Dimension  $l = m \cdot n$ .

Beispiele:

i)

$$A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} \quad \text{und} \quad B = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$$

$$C = A + B = \begin{pmatrix} 1 & 3 \\ 2 & 4 \end{pmatrix}$$

$$2B = \begin{pmatrix} 0 & 2 \\ -2 & 0 \end{pmatrix}.$$

ii) Matrizen können auch aus Vektoren gebildet werden, indem man die Vektoren als die Zeilen der Matrix stellt. Zum Beispiel können wir aus  $\hat{e}_1 = (1, 0, 0)$ ,  $\hat{e}_2 = (0, 1, 0)$  und  $\hat{e}_3 = (0, 0, 1)$

$$A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = \mathbb{1}$$

bilden. Oder wir können aus  $\vec{a} = (1, 2, 3)$ ,  $\vec{b} = (4, 5, 6)$  und  $\vec{c} = (7, 8, 9)$

$$A = \begin{pmatrix} \vec{c} \\ \vec{a} \\ \vec{b} \end{pmatrix} = \begin{pmatrix} 7 & 8 & 9 \\ 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix}$$

bilden.

### 2.12.2 Determinanten

Die Determinante einer quadratischen Matrix ist eine wichtige Abbildung zwischen Matrizen und reellen Zahlen, mit der man einige wichtige Eigenschaften der Matrizen identifizieren kann. Sie stellt auch eine kompakte Form dar, um einige Formeln der Vektorrechnung zusammenzufassen.

Definition: Unter Determinante  $|A|$  oder  $\det(A)$  einer quadratischen  $n \times n$ -Matrix  $A = (a_{ij})$  versteht man eine Abbildung

$$\det(A) : \mathbb{R}^{n \times n} \rightarrow \mathbb{R},$$

die durch folgende rekursive Vorschrift gegeben ist: Bei  $n = 1$  ist  $\det(A) = a_{11}$ .  
Bei  $n \geq 2$  ist

$$\det(A) = |A| = \begin{vmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{vmatrix} =$$

$$= a_{11} A_{11} - a_{12} A_{12} + \dots + (-1)^{1+i} a_{1i} A_{1i} + \dots + (-1)^{1+n} a_{1n} A_{1n},$$

wobei  $A_{ij}$  die Determinante der  $(n-1) \times (n-1)$ -Matrix ist, die man erhält, wenn man bei  $A$  die Zeile  $i$  und die Spalte  $j$  wegnimmt.

Konkret heißt es

i) Bei  $n = 1$

$$A = a_{11} \in \mathbb{R} \Rightarrow \det(A) = a_{11}.$$

ii) Bei  $n = 2$

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$$

$$\det(A) = \det \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} = a_{11} A_{11} - a_{12} A_{21},$$

wobei

$$A_{11} = \det(a_{22}) = a_{22}$$

und

$$A_{12} = \det(a_{21}) = a_{21}.$$

Daraus folgt

$$\det(A) = a_{11} a_{22} - a_{12} a_{21} = (\searrow) - (\swarrow).$$

iii) Bei  $n = 3$

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}$$

$$\begin{aligned}
\det(A) &= a_{11} \begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} - a_{12} \begin{vmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{vmatrix} + a_{13} \begin{vmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{vmatrix} \\
&= a_{11} (a_{22} a_{33} - a_{23} a_{32}) - \\
&\quad - a_{12} (a_{21} a_{33} - a_{23} a_{31}) + \\
&\quad + a_{13} (a_{21} a_{32} - a_{22} a_{31}) \\
\det(A) &= a_{11} a_{22} a_{33} - a_{11} a_{23} a_{32} - \\
&\quad - a_{12} a_{21} a_{33} + a_{12} a_{23} a_{31} + \\
&\quad + a_{13} a_{21} a_{32} - a_{13} a_{22} a_{31}.
\end{aligned}$$

Es sind  $n! = 3! = 6$  Summanden mit jeweils  $n = 3$  Faktoren. Im allgemeinen, bei  $A \in \mathbb{R}^{n \times n}$ , sind es  $n!$  Terme mit  $n$  Faktoren. In der Tat ist

$$\det(A) = \underbrace{a_{11} A_{11} + \dots + (-1)^{1+n} a_{1n} A_{1n}}_{n \text{ Summanden}}$$

und bei  $A_{ij}$  sind es  $(n-1)!$  Summanden, also bei  $\det(A)$  sind es  $n(n-1)!$  Summanden.

Das Vorzeichen ist  $+1$  bei

$$\left. \begin{array}{l} a_{11} \ a_{22} \ a_{33} \\ a_{13} \ a_{21} \ a_{32} \\ a_{12} \ a_{23} \ a_{31} \end{array} \right\} = a_{1i} \ a_{2j} \ a_{3k} \quad \text{mit } (i, j, k) = (1, 2, 3) \text{ oder zyklische Permutationen}$$

und  $-1$  bei

$$\left. \begin{array}{l} a_{11} \ a_{23} \ a_{32} \\ a_{12} \ a_{21} \ a_{33} \\ a_{13} \ a_{22} \ a_{31} \end{array} \right\} = a_{1i} \ a_{2j} \ a_{3k} \quad \text{mit } (i, j, k) \text{ antizyklisch, d.h., } (i, j, k) = (1, 3, 2) \\ \text{und die zyklischen Permutationen davon.}$$

Man kann zeigen, dass

$$\det(A) = \sum_j (-1)^{i+j} a_{ij} A_{ij} = \sum_i (-1)^{i+j} a_{ij} A_{ij}.$$

Das heißt, dass die Determinante nach einer beliebigen Zeile oder Spalte entwickelt werden kann, wenn man die richtige Alternierung des Vorzeichens  $(-1)^{i+j}$  berücksichtigt.

Beispiele:

i)

$$A = \begin{bmatrix} 1 & 2 & 3 \\ -2 & 1 & 4 \\ 0 & 1 & 0 \end{bmatrix} \Rightarrow \det(A) = -10.$$

Es wird dem Leser überlassen zu zeigen, dass man das gleiche Resultat erhält, wenn man nach Zeile 1 oder nach Zeile 3 entwickelt.

ii) Man betrachtet 3 Vektoren  $\vec{a} = (a_1, a_2, a_3)$ ,  $\vec{b} = (b_1, b_2, b_3)$  und  $\vec{c} = (0, 0, 1)$  und man bildet die Matrix

$$A = \begin{pmatrix} \vec{c} \\ \vec{a} \\ \vec{b} \end{pmatrix} = \begin{pmatrix} 0 & 0 & 1 \\ a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \end{pmatrix}$$

$$\det(A) = a_1 b_2 - a_2 b_1 = (\vec{a} \times \vec{b})_3$$

oder mit  $\vec{c} = (0, 1, 0)$

$$A = \begin{pmatrix} 0 & 1 & 0 \\ a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \end{pmatrix}$$

$$\det(A) = a_3 b_1 - a_1 b_3 = (\vec{a} \times \vec{b})_2$$

und mit  $\vec{c} = (1, 0, 0)$

$$\det(A) = \begin{vmatrix} 1 & 0 & 0 \\ a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \end{vmatrix} = a_2 b_3 - a_3 b_2 = (\vec{a} \times \vec{b})_1.$$

Wir können also das Vektorprodukt zweier Vektoren  $\vec{a}$  und  $\vec{b}$  als die formale Entwicklung der Determinante

$$\vec{a} \times \vec{b} = \begin{vmatrix} \hat{e}_1 & \hat{e}_2 & \hat{e}_3 \\ a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \end{vmatrix}$$

schreiben.

## 2.13 Spatprodukt

Als *Spatprodukt* dreier Vektoren  $\vec{a}$ ,  $\vec{b}$  und  $\vec{c}$  bezeichnet man das gemischte Vektor- und Skalarprodukt der Form

$$(\vec{a} \times \vec{b}) \cdot \vec{c} \in \mathbb{R}.$$

Es handelt sich um einen Skalar, genauer gesagt um einen Pseudoskalar, wenn  $\vec{a}$ ,  $\vec{b}$  und  $\vec{c}$  Vektoren sind.

Das Spatprodukt besitzt folgende Eigenschaften:

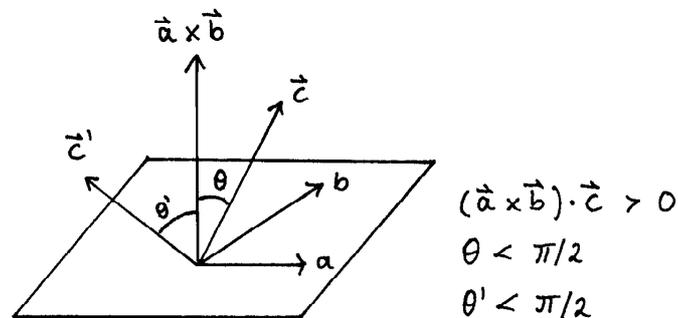
- i) Der Betrag von  $(\vec{a} \times \vec{b}) \cdot \vec{c}$  ist gleich dem Volumen des Parallelepipeds, das von  $\vec{a}$ ,  $\vec{b}$  und  $\vec{c}$  definiert ist.

$$\text{Volumen} = \text{Fläche} \times \text{Höhe} = |\vec{a} \times \vec{b}| \frac{|(\vec{a} \times \vec{b}) \cdot \vec{c}|}{|\vec{a} \times \vec{b}|} = |(\vec{a} \times \vec{b}) \cdot \vec{c}|.$$

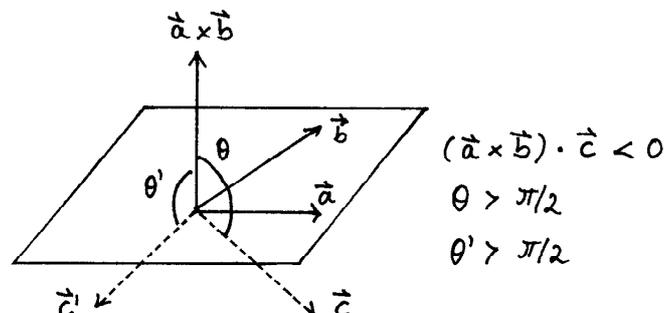
- ii) Was das Vorzeichen angeht, gilt folgendes:

$(\vec{a} \times \vec{b}) \cdot \vec{c} > 0$ , wenn  $\vec{a}$ ,  $\vec{b}$  und  $\vec{c}$  die gleiche Orientierung haben wie das Koordinatensystem. Anderenfalls ist  $(\vec{a} \times \vec{b}) \cdot \vec{c} < 0$ .

Beispielsweise sind für ein Rechtssystem  $(\vec{a} \times \vec{b}) \cdot \vec{c}$  und  $(\vec{a} \times \vec{b}) \cdot \vec{c}'$  in der folgenden Figur positiv [ $(\vec{a} \times \vec{b}) \cdot \vec{c} > 0$ ].



Dagegen sind im folgenden Beispiel  $(\vec{a} \times \vec{b}) \cdot \vec{c} < 0$ .



- iii) Drei Vektoren  $\vec{a}$ ,  $\vec{b}$  und  $\vec{c}$  liegen in derselben Ebene, wenn und nur wenn

$$(\vec{a} \times \vec{b}) \cdot \vec{c} = 0.$$

iv) Mittels der Komponenten von  $\vec{a} = (a_1, a_2, a_3)$ ,  $\vec{b} = (b_1, b_2, b_3)$  und  $\vec{c} = (c_1, c_2, c_3)$  erhalten wir

$$(\vec{a} \times \vec{b}) = (a_2 b_3 - a_3 b_2) \hat{e}_1 + (a_3 b_1 - a_1 b_3) \hat{e}_2 + (a_1 b_2 - a_2 b_1) \hat{e}_3$$

und

$$\begin{aligned} (\vec{a} \times \vec{b}) \cdot \vec{c} &= (a_2 b_3 - a_3 b_2) c_1 + (a_3 b_1 - a_1 b_3) c_2 + (a_1 b_2 - a_2 b_1) c_3 \\ &= \sum_{ijk=1}^3 \varepsilon_{ijk} a_i b_j c_k \\ &= \begin{vmatrix} c_1 & c_2 & c_3 \\ a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \end{vmatrix}. \end{aligned}$$

v) Mit einer einfachen Regruppierung der Terme in der Formel für  $(\vec{a} \times \vec{b}) \cdot \vec{c}$  lässt sich folgende Identität zeigen:

$$(\vec{a} \times \vec{b}) \cdot \vec{c} = (\vec{c} \times \vec{a}) \cdot \vec{b}.$$

In der Tat wissen wir bereits, dass  $\varepsilon_{ijk} = \varepsilon_{kij} \quad \forall i, j, k$  (Zyklische Permutation). Deshalb gilt

$$(\vec{a} \times \vec{b}) \cdot \vec{c} = \sum_{ijk} \varepsilon_{ijk} a_i b_j c_k = \sum_{ijk} \varepsilon_{kij} a_i b_j c_k.$$

Und wenn man die Variablen als  $k \rightarrow i$ ,  $i \rightarrow j$  und  $j \rightarrow k$  umbenennt, folgt

$$(\vec{a} \times \vec{b}) \cdot \vec{c} = \sum_{ijk} \varepsilon_{ijk} a_j b_k c_i = (\vec{c} \times \vec{a}) \cdot \vec{b}.$$

Das heißt, man kann die Reihenfolge der Vektoren im Spatprodukt  $(\vec{a} \times \vec{b}) \cdot \vec{c}$  zyklisch permutieren, ohne dabei das Resultat zu ändern.

vi) Wir wissen, dass das Skalarprodukt kommutativ ist und damit gilt  $(\vec{a} \times \vec{b}) \cdot \vec{c} = \vec{c} \cdot (\vec{a} \times \vec{b})$ . Dies zusammen mit v) führt zu

$$(\vec{a} \times \vec{b}) \cdot \vec{c} = \vec{a} \cdot (\vec{b} \times \vec{c}).$$

Das bedeutet, dass man “ $\times$ ” und “ $\cdot$ ” vertauschen kann.

vii) Da  $\vec{a} \times \vec{b} = -\vec{b} \times \vec{a}$ , gilt auch

$$(\vec{a} \times \vec{b}) \cdot \vec{c} = -(\vec{b} \times \vec{a}) \cdot \vec{c}.$$

Zusammenfassend aus v)-vii):

$$\begin{aligned} (\vec{a} \times \vec{b}) \cdot \vec{c} &= (\vec{c} \times \vec{a}) \cdot \vec{b} = (\vec{b} \times \vec{c}) \cdot \vec{a} = \vec{a} \cdot (\vec{b} \times \vec{c}) = \\ &= -(\vec{b} \times \vec{a}) \cdot \vec{c} = -(\vec{a} \times \vec{c}) \cdot \vec{b} = -(\vec{c} \times \vec{b}) \cdot \vec{a}. \end{aligned}$$

viii) Das Spatprodukt  $(\vec{a} \times \vec{b}) \cdot \vec{c}$  ist ein Pseudoskalar, weil es wie das Vektorprodukt sein Vorzeichen ändert, wenn die Orientierung des Koordinatensystems geändert wird. Bezüglich Drehungen ist es invariant, wie das Skalarprodukt zweier Vektoren.

## 2.14 Doppeltes Vektorprodukt

Es kommt oft vor, dass man das doppelte Vektorprodukt

$$\vec{a} \times (\vec{b} \times \vec{c})$$

berechnen muss. Wir wissen einerseits, dass  $\vec{b} \times \vec{c}$  senkrecht zu  $\vec{b}$  und  $\vec{c}$  ist, und dementsprechend steht  $\vec{b} \times \vec{c}$  senkrecht zur durch  $\vec{b}$  und  $\vec{c}$  definierten Ebene  $\vec{b} \oplus \vec{c}$ . Andererseits ist  $\vec{a} \times (\vec{b} \times \vec{c})$  senkrecht zu  $(\vec{b} \times \vec{c})$  für jedes  $\vec{a}$ . Es folgt dann, dass  $\vec{a} \times (\vec{b} \times \vec{c})$  in der Ebene  $\vec{b} \oplus \vec{c}$  liegt, also

$$\vec{a} \times (\vec{b} \times \vec{c}) = \beta \vec{b} + \gamma \vec{c} \quad \text{mit } \beta, \gamma \in \mathbb{R}.$$

Aus der Orthogonalität zu  $\vec{a}$  folgt

$$0 = \vec{a} \cdot [\vec{a} \times (\vec{b} \times \vec{c})] = \beta \vec{a} \cdot \vec{b} + \gamma \vec{a} \cdot \vec{c}$$

oder

$$\frac{\beta}{\vec{a} \cdot \vec{c}} = \frac{-\gamma}{\vec{a} \cdot \vec{b}}.$$

Wenn man diesen Quotienten  $\varepsilon$  nennt, kann man  $\beta = \varepsilon (\vec{a} \cdot \vec{c})$  und  $\gamma = -\varepsilon (\vec{a} \cdot \vec{b})$  schreiben, und daher ist

$$\vec{a} \times (\vec{b} \times \vec{c}) = \varepsilon [\vec{b}(\vec{a} \cdot \vec{c}) - \vec{c}(\vec{a} \cdot \vec{b})].$$

Damit ist der geometrisch interessante Zusammenhang zwischen  $\vec{a} \times (\vec{b} \times \vec{c})$  und  $\vec{b}$  und  $\vec{c}$  erstellt.

Man kann auch zeigen, dass  $\varepsilon = 1$ , indem man das Koordinatensystem an  $\vec{a}$ ,  $\vec{b}$  und  $\vec{c}$  anpasst. Das ist aber langweilig und man lernt nicht viel dabei.

Ein eleganter Beweis wird im folgenden gegeben, mit dem man die Manipulation von Levy-Civita- und Kronecker-Symbolen trainieren kann. Wir wissen, dass

$$(\vec{b} \times \vec{c}) = \sum_{ijk} \varepsilon_{ijk} b_i c_j \hat{e}_k$$

und

$$\begin{aligned} \vec{a} \times (\vec{b} \times \vec{c}) &= \sum_{lkm} \varepsilon_{lkm} a_l (\vec{b} \times \vec{c})_k \hat{e}_m \\ &= \sum_{lkm} \varepsilon_{lkm} a_l \left( \sum_{ij} \varepsilon_{ijk} b_i c_j \right) \hat{e}_m \\ &= \sum_{lm} \sum_{ij} a_l b_i c_j \hat{e}_m \left( \sum_k \varepsilon_{lkm} \varepsilon_{ijk} \right). \end{aligned}$$

Die Summe über  $k$  lässt sich vereinfachen. Zunächst permutieren wir die Indizes in  $\varepsilon_{lkm} = \varepsilon_{mlk}$  zyklisch und erhalten

$$\sum_{k=1}^3 \varepsilon_{lkm} \varepsilon_{ijk} = \sum_{k=1}^3 \varepsilon_{mlk} \varepsilon_{ijk}.$$

Damit die Summe nicht null ist, muss das Paar  $(m, l) = (i, j)$  oder  $(m, l) = (j, i)$  sein. In der Tat, wenn z. B.  $m \neq i$  und  $m \neq j$ , dann ist entweder  $k = i$  oder  $k = j$ , und damit  $\varepsilon_{ijk} = 0$ , oder es ist  $k = m$ , und damit  $\varepsilon_{mlk} = 0$ . Es gibt also zwei Möglichkeiten.

$$\begin{cases} m = i \\ l = j \end{cases} \Rightarrow \sum_k \varepsilon_{ijk} \varepsilon_{ijk} = \sum_k (\varepsilon_{ikj})^2 = 1$$

oder

$$\begin{cases} m = j \\ l = i \end{cases} \Rightarrow \sum_k \varepsilon_{jik} \varepsilon_{ijk} = - \sum_k (\varepsilon_{ijk})^2 = -1.$$

Zusammenfassend können wir

$$\sum_{k=1}^3 \varepsilon_{lkm} \varepsilon_{ijk} = \delta_{mi} \delta_{lj} - \delta_{mj} \delta_{li}$$

schreiben. Daraus folgt

$$\begin{aligned} \vec{a} \times (\vec{b} \times \vec{c}) &= \sum_{\substack{lm \\ ij}} a_l b_i c_j \hat{e}_m (\delta_{mi} \delta_{lj} - \delta_{mj} \delta_{li}) \\ &= \sum_{ij} a_j b_i c_j \hat{e}_i - \sum_{ij} a_i b_i c_j \hat{e}_j \\ &= \left( \sum_i b_i \hat{e}_i \right) \left( \sum_j a_j c_j \right) - \left( \sum_j c_j \hat{e}_j \right) \left( \sum_i a_i b_i \right). \end{aligned}$$

und damit

$$\vec{a} \times (\vec{b} \times \vec{c}) = \vec{b}(\vec{a} \cdot \vec{c}) - \vec{c}(\vec{a} \cdot \vec{b}).$$

Diese Gleichung wirkt wie eine Art Linearisierung, da sie kompliziertere dreifache Vektorprodukte in Summen bzw. Linearkombinationen von Vektoren umwandelt. Sie ist also sehr nützlich, um komplexere Ausdrücke zu vereinfachen.

Man sollte in diesem Zusammenhang betonen, dass das Vektorprodukt nicht assoziativ ist.  $\vec{a} \times (\vec{b} \times \vec{c})$  liegt in der  $\vec{b} \oplus \vec{c}$ -Ebene, während  $(\vec{a} \times \vec{b}) \times \vec{c}$  in der  $\vec{a} \oplus \vec{b}$ -Ebene liegt!

## 2.15 Andere Vektorprodukte

Der Vollständigkeit halber werden hier ein paar zusätzliche Produkte erwähnt, die man mit Hilfe der Beziehungen für  $(\vec{a} \times \vec{b}) \cdot \vec{c}$  und  $\vec{a} \times (\vec{b} \times \vec{c})$  vereinfachen kann.

i) Wir betrachten zunächst

$$\begin{aligned} (\vec{a} \times \vec{b}) \cdot \underbrace{(\vec{c} \times \vec{d})}_{\vec{e}} &= \vec{a} \cdot (\vec{b} \times (\vec{c} \times \vec{d})) = \vec{a} \cdot [\vec{c}(\vec{b} \cdot \vec{d}) - \vec{d}(\vec{b} \cdot \vec{c})] \\ &= (\vec{a} \cdot \vec{c})(\vec{b} \cdot \vec{d}) - (\vec{a} \cdot \vec{d})(\vec{b} \cdot \vec{c}). \end{aligned}$$

ii) Als Spezialfall für  $\vec{a} = \vec{c}$  und  $\vec{b} = \vec{d}$  erhalten wir

$$(\vec{a} \times \vec{b})^2 = (\vec{a} \times \vec{b}) \cdot (\vec{a} \times \vec{b}) = a^2 b^2 - (\vec{a} \cdot \vec{b})^2.$$

Dies ist allerdings nichts Neues, da  $(\vec{a} \times \vec{b})^2 = |\vec{a} \times \vec{b}|^2 = a^2 b^2 \sin^2 \theta = a^2 b^2 (1 - \cos^2 \theta) = a^2 b^2 - (\vec{a} \cdot \vec{b})^2$ .

iii) Als direkte Anwendung der Formel für das doppelte Vektorprodukt erhält man

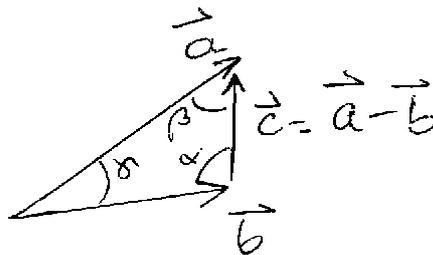
$$\underbrace{(\vec{a} \times \vec{b})}_{\vec{e}} \times (\vec{c} \times \vec{d}) = \vec{c} [(\vec{a} \times \vec{b}) \cdot \vec{d}] - \vec{d} [(\vec{a} \times \vec{b}) \cdot \vec{c}].$$

## 2.16 Anwendungen der Vektoralgebra

In den folgenden Abschnitten werden einige Anwendungen der Vektoralgebra bei der analytischen Geometrie beispielhaft erläutert.

### 2.16.1 Trigonometrie

Wir betrachten ein durch die Vektoren  $\vec{a}$  und  $\vec{b}$  definiertes Dreieck, wobei  $\vec{c} = \vec{a} - \vec{b}$  ist.



Es gilt

$$c^2 = (\vec{a} - \vec{b}) \cdot (\vec{a} - \vec{b}) = a^2 + b^2 - 2\vec{a} \cdot \vec{b}.$$

und damit

$$c^2 = a^2 + b^2 - 2ab \cos \gamma.$$

Dies ist der sogenannte Kosinussatz. Darüber hinaus gilt

$$\begin{aligned} \vec{a} \times \vec{c} &= \vec{a} \times (\vec{a} - \vec{b}) = \vec{b} \times \vec{a} \\ |\vec{a} \times \vec{c}| &= |\vec{b} \times \vec{a}| \end{aligned}$$

$$ac \sin \beta = ab \sin \gamma$$

$$\frac{c}{\sin \gamma} = \frac{b}{\sin \beta}.$$

Andererseits haben wir

$$\vec{b} \times \vec{c} = \vec{b} \times (\vec{a} - \vec{b}) = \vec{b} \times \vec{a}$$

$$|\vec{b} \times \vec{c}| = |\vec{b} \times \vec{a}|$$

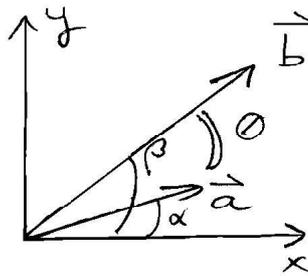
$$bc \sin(\pi - \alpha) = bc \sin \alpha = ab \sin \gamma$$

$$\frac{c}{\sin \gamma} = \frac{a}{\sin \alpha}.$$

Und damit folgt der Sinussatz

$$\frac{a}{\sin \alpha} = \frac{b}{\sin \beta} = \frac{c}{\sin \gamma}.$$

### 2.16.2 Additives Theorem für Winkelfunktionen



$$\vec{a} = a (\cos \alpha, \sin \alpha)$$

$$\vec{b} = b (\cos \beta, \sin \beta).$$

Aus  $\vec{a} \cdot \vec{b} = \sum_i a_i b_i$  folgt

$$\vec{a} \cdot \vec{b} = ab (\cos \alpha \cos \beta + \sin \alpha \sin \beta).$$

Andererseits gilt

$$\vec{a} \cdot \vec{b} = ab \cos \theta = ab \cos(\beta - \alpha).$$

Daraus folgt

$$\cos(\beta - \alpha) = \cos \alpha \cos \beta + \sin \alpha \sin \beta.$$

Wenn man berücksichtigt, dass  $\sin(-\alpha) = -\sin \alpha$ , erhalten wir die bekannte Formel

$$\cos(\alpha + \beta) = \cos \alpha \cos \beta - \sin \alpha \sin \beta.$$

Eine ähnliche Manipulation mit dem Vektorprodukt ergibt einerseits aus der Formel für die Komponenten

$$\vec{a} \times \vec{b} = (0, 0, a_1 b_2 - a_2 b_1) = ab (0, 0, \cos \alpha \sin \beta - \sin \alpha \cos \beta)$$

und andererseits aus der Definition

$$\vec{a} \times \vec{b} = (0, 0, a b \sin(\beta - \alpha)).$$

Daraus folgt

$$\sin(\beta - \alpha) = \cos \alpha \sin \beta - \sin \alpha \cos \beta$$

oder

$$\sin(\alpha + \beta) = \cos \alpha \sin \beta + \sin \alpha \cos \beta.$$

### 2.16.3 Vektorgleichung einer Geraden

Für zwei beliebige Punkte  $\vec{a}$  und  $\vec{b}$  in  $\mathbb{R}^3$  ist die Vektorgleichung für die Punkte  $\vec{r}$  der Geraden, die durch  $\vec{a}$  und  $\vec{b}$  geht, folgende:

$$\vec{r} = \vec{a} + \lambda (\vec{b} - \vec{a}) \quad \text{mit } \lambda \in \mathbb{R}.$$

In Komponentenform schreiben wir  $\vec{r} = (x, y, z)$ ,  $\vec{a} = (a_1, a_2, a_3)$  und  $\vec{b} = (b_1, b_2, b_3)$ . Dann ist die Gerade durch folgende Komponentengleichungen gegeben:

$$\begin{cases} x = a_1 + \lambda (b_1 - a_1) \\ y = a_2 + \lambda (b_2 - a_2) \\ z = a_3 + \lambda (b_3 - a_3). \end{cases}$$

mit beliebigem  $\lambda \in \mathbb{R}$ .

### 2.16.4 Vektorgleichung einer Ebene

Wir betrachten einen Vektor  $\vec{a}$  senkrecht zu dieser Ebene. Wenn die Ebene durch den Koordinatenursprung geht, lautet die Gleichung einfach

$$\vec{a} \cdot \vec{r} = 0.$$

Wenn nicht, ist die Projektion der Punkte in der Ebene auf den normalen Vektor  $\vec{a}$  konstant. Das heißt

$$\frac{\vec{a}}{a} \cdot \vec{r} = h,$$

wobei  $h$  der Abstand zwischen der Ebene und dem Koordinatenursprung ist. In kartesischen Koordinaten haben wir

$$a_1 x + a_2 y + a_3 z = a h = b = \text{Konstante}.$$

Wenn man den Abstand nicht kennt, sondern die Richtung der Normale  $\vec{a}$  und einen beliebigen Punkt  $\vec{r}_0$  auf der Ebene, dann kann man schreiben

$$\vec{a} \cdot \vec{r} = a h \quad \text{und} \quad \vec{a} \cdot \vec{r}_0 = a h,$$

also

$$\vec{a} \cdot (\vec{r} - \vec{r}_0) = 0.$$

Anders ausgedrückt, sind alle Vektoren  $\vec{r} - \vec{r}_0$  in der Ebene (d.h., relativ zu  $\vec{r}_0$ ) senkrecht zu  $\vec{a}$ . In kartesischen Koordinaten haben wir

$$a_1(x - x_0) + a_2(y - y_0) + a_3(z - z_0) = 0,$$

wenn  $\vec{r}_0 = (x_0, y_0, z_0)$ .

Sollte man die Ebene suchen, die durch drei Punkte  $\vec{r}_1, \vec{r}_2$  und  $\vec{r}_3$  geht, dann liegen  $\vec{r}_2 - \vec{r}_1$  und  $\vec{r}_3 - \vec{r}_1$  parallel zur Ebene, und damit ist  $\vec{a} = (\vec{r}_2 - \vec{r}_1) \times (\vec{r}_3 - \vec{r}_1)$  die Normale. Die Vektorgleichung lautet dann

$$\begin{aligned} (\vec{r} - \vec{r}_1) \cdot \vec{a} &= 0, \quad \text{d.h.} \\ (\vec{r} - \vec{r}_1) \cdot [(\vec{r}_2 - \vec{r}_1) \times (\vec{r}_3 - \vec{r}_1)] &= 0 \end{aligned}$$

oder in kartesischen Koordinaten

$$\begin{vmatrix} x - x_1 & y - y_1 & z - z_1 \\ x_2 - x_1 & y_2 - y_1 & z_2 - z_1 \\ x_3 - x_1 & y_3 - y_1 & z_3 - z_1 \end{vmatrix} = 0.$$

### 2.16.5 Abstand eines Punktes zu einer Ebene

Die Gleichung der Ebene  $E$  lautet  $\vec{a} \cdot \vec{r} = b$  und die Koordinaten des Punktes  $\vec{p} = (p_1, p_2, p_3)$ . Die Projektion von  $\vec{p}$  auf  $\vec{a}$  lautet

$$\frac{\vec{a}}{a} \cdot \vec{p},$$

und der Abstand der Ebene zum Ursprung ist

$$\vec{r} \cdot \frac{\vec{a}}{a} = \frac{b}{a} \quad \forall \vec{r} \in E.$$

Der Abstand  $d$  vom Punkt zur Ebene ist

$$\begin{aligned} d &= \frac{\vec{a}}{a} \cdot \vec{p} - \vec{r} \cdot \frac{\vec{a}}{a} \\ d &= \frac{\vec{a} \cdot \vec{p} - b}{a}. \end{aligned}$$

Beispiel:

Wir betrachten den Punkt  $\vec{p} = (6, 3, -2)$  und die Ebene  $2x - 4y + z = 2$ . Das heißt  $\vec{a} \cdot \vec{r} = b$  mit  $\vec{a} = (2, -4, 1) \Rightarrow a = \sqrt{4 + 16 + 1} = \sqrt{21}$ .  
 $b = 2$

$$d = \frac{\vec{a} \cdot \vec{p} - 2}{\sqrt{21}} = \frac{(12 - 12 - 2) - 2}{\sqrt{21}} = \frac{-4}{21}$$

Das Resultat  $d < 0$  bedeutet, dass der Punkt unter der Ebene bezüglich der Richtung von  $\vec{a}$  liegt.

### 2.16.6 Abstand eines Punktes $\vec{p}$ zu einer Geraden

Wir betrachten die Gerade  $\vec{r} = \vec{r}_0 + \lambda \vec{a}$  mit  $\lambda \in \mathbb{R}$  und den Punkt  $\vec{p}$ . Für den Abstand  $d$  gilt

$$d = |\vec{p} - \vec{r}_0| \sin \theta$$

$$d = \frac{|(\vec{p} - \vec{r}_0) \times \vec{a}|}{a}.$$

Beispiel:

Sei  $\vec{p} = (-3, 2, 1)$  und die Gerade  $\frac{x+1}{-2} = \frac{y-3}{4} = \frac{z-1}{-1} = \lambda \in \mathbb{R}$ .

$$(x, y, z) = \vec{r} = \lambda(-2, 4, -1) + (-1, 3, 1).$$

$$\vec{p} - \vec{r}_0 = (-2, -1, 0)$$

$$\vec{a} = (-2, 4, -1) \Rightarrow a = \sqrt{4 + 16 + 1} = \sqrt{21}$$

$$(\vec{p} - \vec{r}_0) \times \vec{a} = \begin{vmatrix} \hat{e}_1 & \hat{e}_2 & \hat{e}_3 \\ -2 & -1 & 0 \\ -2 & 4 & -1 \end{vmatrix} = (1, -2, -10)$$

$$|(\vec{p} - \vec{r}_0) \times \vec{a}| = \sqrt{1 + 4 + 100} = \sqrt{105}$$

$$d = \sqrt{105}/\sqrt{21}.$$

### 2.16.7 Minimaler Abstand zwischen zwei Geraden

$$G_1 : \vec{r} = \vec{r}_1 + \alpha \vec{a} \quad \alpha \in \mathbb{R}$$

$$G_2 : \vec{r} = \vec{r}_2 + \beta \vec{b} \quad \beta \in \mathbb{R}.$$

Der Abstand ist gleich der Projektion von einem beliebigen Vektor, der von der Gerade 1 zur Gerade 2 geht, z.B.  $\vec{r}_2 - \vec{r}_1$ , auf die Richtung, die senkrecht zu beiden Geraden ist. Das heißt

$$d = (\vec{r}_2 - \vec{r}_1) \cdot \frac{(\vec{a} \times \vec{b})}{|\vec{a} \times \vec{b}|} \quad \text{wenn } |\vec{a} \times \vec{b}| \neq 0.$$

Wenn  $|\vec{a} \times \vec{b}| = 0$ , dann ist  $\vec{a} \parallel \vec{b}$ , und damit sind beide Geraden parallel. Es ist also egal, was für einen Punkt man auf die Gerade  $G_2$  nimmt, da alle den gleichen Abstand zur Geraden  $G_1$  haben. Wir nehmen also  $\vec{r}_2$  und erhalten  $d = |\vec{r}_2 - \vec{r}_1| \sin \theta = \frac{|(\vec{r}_2 - \vec{r}_1) \times \vec{a}|}{a}$ . Dies gilt nur für parallele Geraden.

Beispiel:

Wir betrachten die Geraden

$$G_1 : \frac{x-3}{2} = \frac{y+5}{-3} = \frac{z+1}{5}$$

$$G_2 : \frac{x+1}{-4} = \frac{y-1}{3} = \frac{z-3}{-1}.$$

Dies ist äquivalent zu

$G_1 : \vec{r} = \alpha(2, -3, 5) + (3, -5, -1)$  mit  $\alpha \in \mathbb{R}$

$G_2 : \vec{r} = \beta(-4, 3, -1) + (-1, 1, 3)$  mit  $\beta \in \mathbb{R}$ .

$$\vec{r}_2 - \vec{r}_1 = (-1, 1, 3) - (3, -5, -1) = (-4, 6, 4)$$

$$\vec{a} \times \vec{b} = \begin{vmatrix} \hat{e}_1 & \hat{e}_2 & \hat{e}_3 \\ 2 & -3 & 5 \\ -4 & 3 & -1 \end{vmatrix} = (-12, -18, -6) = -6(2, 3, 1).$$

$$|\vec{a} \times \vec{b}| = 6\sqrt{4 + 9 + 1} = 6\sqrt{14}$$

$$\frac{(\vec{r}_2 - \vec{r}_1) \cdot (\vec{a} \times \vec{b})}{|\vec{a} \times \vec{b}|} = \frac{-6(-8 + 18 + 4)}{6\sqrt{14}} = -\sqrt{14} \Rightarrow d = \sqrt{14}.$$

### 2.16.8 Winkel zwischen Geraden und Ebenen

Der Winkel zwischen zwei Geraden  $G_1$ ,  $\vec{r} = \vec{r}_1 + \alpha \vec{a}$ , und  $G_2$ ,  $\vec{r} = \vec{r}_2 + \beta \vec{b}$ , ist durch

$$\cos \theta = \frac{\vec{a} \cdot \vec{b}}{a b}$$

gegeben.

Der Winkel zwischen zwei Ebenen  $E_1$ ,  $\vec{a} \cdot \vec{r} = \alpha$ , und  $E_2$ ,  $\vec{b} \cdot \vec{r} = \beta$ , ist durch

$$\cos \theta = \frac{\vec{a} \cdot \vec{b}}{a b}$$

gegeben.

Der Winkel zwischen einer Geraden  $G$ ,  $\vec{r} = \vec{r}_1 + \alpha \vec{a}$ , und einer Ebene  $E$ ,  $\vec{r} \cdot \vec{b} = \beta$ , ist durch

$$\cos \left( \frac{\pi}{2} - \theta \right) = \sin \theta = \frac{\vec{a} \cdot \vec{b}}{a b}$$

gegeben.

## 2.17 Zylinderkoordinaten

Die Polarkoordinaten  $(\varrho, \varphi)$  sind uns aus den komplexen Zahlen bekannt. Jeder Punkt  $\vec{r} = (x, y, 0)$  auf der  $x, y$ -Ebene lässt sich durch den Abstand  $\varrho$  zur  $z$ -Achse und den Azimutwinkel  $\varphi$  zur positiven  $x$ -Achse eindeutig definieren.

Wir haben also die Beziehungen zwischen polaren und kartesischen Koordinaten:

$$\begin{cases} x = \varrho \cos \varphi & \text{mit } 0 \leq \varrho \\ y = \varrho \sin \varphi & 0 \leq \varphi \leq 2\pi \end{cases}$$

und

$$\begin{cases} \varrho = \sqrt{x^2 + y^2} & \text{mit } x, y \in \mathbb{R} \\ \varphi = \arctan(y/x) & \text{und } 0 \leq \varphi \leq 2\pi. \end{cases}$$

Man sollte hier daran erinnern, dass  $\tan(\varphi) = \tan(\varphi + \pi)$ , und deshalb muss man in jedem Fall prüfen, welcher Wert von  $\varphi$  zu dem Quadrant von  $(x, y)$  passt.

Die Zylinderkoordinaten  $(\varrho, \varphi, z)$  in  $\mathbb{R}^3$  sind auf die Polarkoordinaten basiert. Die Beziehungen zwischen kartesischen Koordinaten  $(x, y, z)$  und Zylinderkoordinaten  $(\varrho, \varphi, z)$  sind

$$\begin{cases} x = \varrho \cos \varphi \\ y = \varrho \sin \varphi \\ z = z \end{cases}$$

oder

$$\begin{cases} \varrho = \sqrt{x^2 + y^2} \\ \varphi = \arctan(y/x) \\ z = z. \end{cases}$$

Die Zylinderkoordinaten sind am besten für Systeme geeignet, die eine kontinuierliche Symmetrieachse haben (z.B. Drähte oder Zylinder).

## 2.18 Kugelkoordinaten

Liegt eine Kugelsymmetrie vor (z.B. bei dem Anziehungskraftfeld der Sonne oder einer Punktladung), ist es viel günstiger, Kugelkoordinaten zu benutzen. Sie sind durch den Radius oder Abstand zum Ursprung  $r = \sqrt{x^2 + y^2 + z^2}$ , den Azimutwinkel  $\varphi$  zwischen der Projektion des Positionsvektors  $\vec{r} = (x, y, z)$  auf die  $xy$ -Ebene und der positiven  $x$ -Achse sowie den Polarwinkel  $\theta$  zwischen Positionsvektor und der positiven  $z$ -Achse gegeben. Die Beziehungen zwischen Kartesischen und Kugelkoordinaten sind

$$\begin{cases} x = r \sin \theta \cos \varphi & 0 \leq 2\pi \leq \varphi \\ y = r \sin \theta \sin \varphi & 0 \leq \theta \leq \pi \\ z = r \cos \theta & r \geq 0 \end{cases}$$

oder

$$\begin{cases} r = \sqrt{x^2 + y^2 + z^2} \\ \varphi = \arctan(y/x) & 0 \leq \varphi \leq 2\pi \\ \theta = \arccos \frac{z}{\sqrt{x^2 + y^2 + z^2}} & 0 \leq \theta \leq \pi. \end{cases}$$

Wie bei den Zylinderkoordinaten muss der Wert von  $\varphi$  aus  $\arctan(y/x)$  an den jeweiligen Quadrant von  $(x, y)$  angepasst sein, da  $\tan(\varphi) = \tan(\varphi + \pi)$ .

Beispiel:

Wir betrachten zwei Punkte auf der Erdoberfläche. Der eine Punkt  $\vec{r}_1$  hat Kugelkoordinaten  $(r, \varphi_1, \theta_1)$ , und der andere Punkt  $\vec{r}_2$  hat Koordinaten  $(r, \varphi_2, \theta_2)$ . Wie groß ist der Abstand zwischen den Punkten  $\vec{r}_1$  und  $\vec{r}_2$  auf der Erdoberfläche?

Der Abstand auf der Kugeloberfläche (also kein direkter gerader Weg) ist  $d = r \gamma$ , wobei  $\gamma$  der Winkel zwischen  $\vec{r}_1$  und  $\vec{r}_2$  ist. Wir brauchen also das Skalarprodukt in Kugelkoordinaten. Dies lässt sich wie folgt berechnen:

$$\begin{aligned}\vec{r}_1 &= r_1 (\sin \theta_1 \cos \varphi_1, \sin \theta_1 \sin \varphi_1, \cos \theta_1) \\ \vec{r}_2 &= r_2 (\sin \theta_2 \cos \varphi_2, \sin \theta_2 \sin \varphi_2, \cos \theta_2).\end{aligned}$$

Damit haben wir

$$\begin{aligned}\vec{r}_1 \cdot \vec{r}_2 &= r_1 r_2 (\sin \theta_1 \sin \theta_2 \cos \varphi_1 \cos \varphi_2 + \\ &\quad + \sin \theta_1 \sin \theta_2 \sin \varphi_1 \sin \varphi_2 + \cos \theta_1 \cos \theta_2) \\ &= r_1 r_2 (\sin \theta_1 \sin \theta_2 \cos (\varphi_1 - \varphi_2) + \cos \theta_1 \cos \theta_2).\end{aligned}$$

Und daraus folgt

$$\cos \gamma = \sin \theta_1 \sin \theta_2 \cos (\varphi_1 - \varphi_2) + \cos \theta_1 \cos \theta_2.$$

### 3 Differentiation und Integration in $\mathbb{R}^n$

#### 3.1 Abbildungen

Unser Anfangspunkt ist der bekannte Begriff der Funktion oder Abbildung als eine allgemeine Beziehung zwischen den Elementen von zwei Mengen: dem Definitionsbereich  $A$  und dem Wertevorrat  $B$ :

$$f: A \rightarrow B \\ a \rightarrow b = f(a), \quad \text{wobei } a \in A \quad \text{und} \quad f(a) \in B.$$

Neben den wohl bekanntesten Funktionen reeller Variablen

$$f: \mathbb{R} \rightarrow \mathbb{R}$$

interessiert man sich in der Physik für eine Reihe anderer, teilweise komplexer, Abbildungen. Zum Beispiel

i) Folgen:  $f: \mathbb{N} \rightarrow \mathbb{R}$ , wobei  $f(n) = a_n$  mit  $n \in \mathbb{N}$  und  $a_n \in \mathbb{R}$ .

ii) Permutationen: Bijektive  $f: [1, n] \rightarrow [1, n]$  im Intervall natürlicher Zahlen  $[1, n] \in \mathbb{N}$ .

$$P = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix} \quad \text{oder} \quad \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix}.$$

iii) Funktionen komplexer Variablen:

$f: \mathbb{C} \rightarrow \mathbb{C}$ , die Definitionsbereich und Wertevorrat reeller Funktionen erweitern.

iv) Kurven oder Trajektorien in  $\mathbb{R}^n$ :

$$r: \mathbb{R} \rightarrow \mathbb{R}^3 \\ t \rightarrow \vec{r}(t) = (x(t), y(t), z(t))$$

Funktionen mehrerer reeller Variablen

v)  $f: \mathbb{R}^2 \rightarrow \mathbb{R}$   
 $(x, y) \rightarrow z = f(x, y)$

Dabei definiert  $(x, y, f(x, y))$  mit  $(x, y) \in \mathbb{R}^2$  eine Fläche in  $\mathbb{R}^3$ .

vi)  $f: \mathbb{R}^3 \rightarrow \mathbb{R}$  ← Skalar-Felder

$$T(x, y, z) \quad \text{Temperatur} \\ \rho(x, y, z) \quad \text{Dichte eines Gases}$$

vii) Allgemeine Funktionen:  $f: \mathbb{R}^n \rightarrow \mathbb{R}$

$E(V, T, N, E, H)$  Interne Energie der Thermodynamik  
 $\Psi(x_1, y_1, z_1, x_2, y_2, z_2, x_3, y_3, z_3, \dots, t)$  Wellenfunktion der Quantenmechanik  
wobei  $(x_i, y_i, z_i)$  die Koordinaten des Elektrons  $i$  sind und  $t$  die Zeit ist.

viii) Vektor-Felder:

$$\begin{aligned} f: \mathbb{R}^2 &\rightarrow \mathbb{R}^2 \\ f: \mathbb{R}^3 &\rightarrow \mathbb{R}^3 \quad \vec{v}(x, y, z) \text{ Geschwindigkeitsfeld} \end{aligned}$$

$$\begin{aligned} \vec{E}(x, y, z) &\text{ Elektrisches Feld} \\ \vec{B}(x, y, z) &\text{ Magnetfeld} \end{aligned}$$

ix) Die Abbildungen von Funktionen auf anderen Funktionen werden Operatoren genannt:

$$\begin{aligned} \mathbb{B} &= \{f: \mathbb{R} \rightarrow \mathbb{R}, \text{ bijektiv}\} \\ I: \mathbb{B} &\rightarrow \mathbb{B} \\ f(x) &\rightarrow f^{-1}(x) \\ e^x &\rightarrow \ln x \\ \mathbb{D} &= \{f: \mathbb{R} \rightarrow \mathbb{R}, \text{ differenzierbar}\} \\ D: \mathbb{D} &\rightarrow \mathbb{D} \quad \text{Ableitungsoperator} \\ D[f(x)] &= f'(x) = \frac{df}{dx} \end{aligned}$$

x) Die Abbildungen von Funktionen auf einer Zahl werden Funktionellen genannt:

$$\begin{aligned} \mathcal{F} &= \{f: \mathbb{R} \rightarrow \mathbb{R}\} \\ I: \mathcal{F} &\rightarrow \mathbb{R} \quad \text{Integral über ein bestimmtes Intervall } [a, b]. \\ I[f] &= \int_a^b f(x) dx \end{aligned}$$

### 3.2 Der $n$ -dimensionale Euklidische Raum $\mathbb{R}^n$

Die Beispiele des vorherigen Absatzes zeigen, dass Funktionen mehrerer Variablen eine zentrale Rolle in der Physik spielen. Es ist also angebracht, die fundamentalen Eigenschaften des Definitionsbereichs dieser Funktionen zu besprechen, d.h., des  $n$ -dimensionalen Raums  $\mathbb{R}^n$ .

Die Elemente  $\vec{x}$  von  $\mathbb{R}^n$  werden Punkte oder Vektoren genannt. Sie sind definiert als  $n$ -Tupel von reellen Zahlen, d.h.,

$$\vec{x} = (x_1, x_2, \dots, x_n),$$

wobei  $x_i \in \mathbb{R}$  ( $i = 1, \dots, n$ ) die Komponenten oder Koordinaten des Vektors  $\vec{x}$  bezüglich der kanonischen Basis oder des Koordinatensystems heißen.

$$\hat{e}_1 = (1, 0, \dots, 0), \hat{e}_2 = (0, 1, 0, \dots, 0), \dots, \hat{e}_n = (0, 0, \dots, 1).$$

Das Symbol  $\hat{\phantom{x}}$  bezeichnet hier Vektoren der Länge 1. Zwei Vektoren sind gleich, nur wenn alle ihre Komponenten gleich sind.

Für jede zwei Vektoren  $\vec{x} = (x_1, \dots, x_n)$  und  $\vec{y} = (y_1, \dots, y_n) \in \mathbb{R}^n$  wird die Addition als

$$\vec{x} + \vec{y} = (x_1 + y_1, \dots, x_n + y_n)$$

definiert. Es handelt sich um eine komponentenweise Addition, d.h., die  $i$ -te Komponente der Summe ist gleich der Summe der  $i$ -ten Komponenten der Summanden. Die Summe entspricht der Summe von Translationen oder Kräften in  $\mathbb{R}^3$ . Damit lässt sich jeder Vektor  $\vec{x} = (x_1, \dots, x_n)$  als Summe von Vektoren entlang der orthogonalen Achsen schreiben:

$$\vec{x} = (x_1, 0, \dots, 0) + (0, x_2, \dots, 0) + \dots + (0, 0, \dots, x_n).$$

Der Vektor  $(0, \dots, x_i, \dots, 0)$  ist also die Projektion von  $\vec{x}$  auf die Koordinatenachse  $i$ .

Die so definierte Addition erfüllt folgende Eigenschaften, die aus der Summe reeller Zahlen übertragen werden:

- 1) Geschlossenheit:  
 $\forall \vec{x}, \vec{y} \in \mathbb{R}^n$  ist  $\vec{x} + \vec{y} \in \mathbb{R}^n$ .
- 2) Assoziativgesetz:  
 $\vec{x} + (\vec{y} + \vec{z}) = (\vec{x} + \vec{y}) + \vec{z}$ .
- 3)  $\exists!$  neutrales Element  $\vec{0} = (0, 0, \dots, 0) \in \mathbb{R}^n$   
mit  $\vec{x} + \vec{0} = \vec{0} + \vec{x} = \vec{x} \quad \forall \vec{x} \in \mathbb{R}^n$ .
- 4)  $\exists$  additives Inverses:  
 $\forall \vec{x} \in \mathbb{R}^n \quad \exists \quad -\vec{x} = (-x_1, -x_2, \dots, -x_n)$   
mit  $\vec{x} + (-\vec{x}) = (-\vec{x}) + \vec{x} = \vec{0}$ .
- 5) Kommutativgesetz:  
 $\vec{x} + \vec{y} = \vec{y} + \vec{x} \quad \forall \vec{x}, \vec{y} \in \mathbb{R}^n$ .

$\mathbb{R}^n$  und die Summe bilden also eine kommutative oder Abelsche Gruppe.

Darüber hinaus lassen sich Vektoren durch beliebige reelle Zahlen  $\lambda \in \mathbb{R}$ , die sogenannten Skalare, multiplizieren. Das  $\lambda$ -fache des Vektors  $\vec{x}$  ist definiert als

$$\lambda \vec{x} = (\lambda x_1, \lambda x_2, \dots, \lambda x_n).$$

Diese Multiplikation eines Vektors mit einem Element des Körpers  $\mathbb{R}$  erfüllt folgende Eigenschaften, die wiederum aus den Eigenschaften von  $\mathbb{R}$  stammen:

- 6)  $(\lambda + \mu) \vec{x} = (\lambda \vec{x}) + (\mu \vec{x}) \quad \forall \lambda, \mu \in \mathbb{R} \quad \forall \vec{x} \in \mathbb{R}^n$ .
- 7)  $\lambda (\vec{x} + \vec{y}) = \lambda \vec{x} + \lambda \vec{y}$ .
- 8)  $1 \vec{x} = \vec{x}$ , wobei  $1 \in \mathbb{R}$  das Einselement des Körpers ist.

Wir können nun einen allgemeinen Vektor  $\vec{x}$  in der Form

$$\begin{aligned} \vec{x} &= x_1 (1, 0, \dots, 0) + x_2 (0, 1, \dots, 0) + \dots + x_n (0, 0, \dots, 1) \\ &= \sum_{i=1}^n x_i \hat{e}_i \end{aligned}$$

schreiben. Man nennt dies eine lineare Kombination der Basisvektoren. Diese Darstellung ist eindeutig, da  $\sum_i x_i \hat{e}_i = \vec{0} \Rightarrow x_i = 0 \quad \forall i$ . Man sagt auch, dass die Vektoren  $\hat{e}_1, \hat{e}_2, \dots, \hat{e}_n$  linear unabhängig sind.

Eine Menge mit den Eigenschaften 1) bis 8) wird ein *linearer Raum* oder *Vektorraum* über dem Körper der reellen Zahlen genannt. In diesem Fall wird  $n$  die *Dimension des Vektorraums* genannt.

Abgesehen vom jetzigen Interesse als Definitionsbereich von Funktionen mit mehreren Variablen spielen Vektorräume eine sehr wichtige Rolle in der Physik. Viele davon sind isomorph mit  $\mathbb{R}^n$  für ein bestimmtes  $n$ , z.B. die Menge aller Polynome mit Grad  $\leq m$ . Es handelt sich um Linearräume endlicher Dimension. Andere lineare Räume sind unendlich dimensional. Das berühmteste Beispiel davon ist sicher der lineare Raum aller Zustände eines quantenmechanischen Systems.

### 3.2.1 Skalarprodukt und Norm

Das Skalarprodukt ist eine sehr wichtige Verknüpfung zweier Vektoren in  $\mathbb{R}^n$ , die uns die Definition einer Metrik oder eines Abstands in  $\mathbb{R}^n$  ermöglichen wird. Für jede zwei Vektoren  $\vec{x}$  und  $\vec{y}$  in  $\mathbb{R}^n$  definieren wir das Skalarprodukt  $\vec{x} \cdot \vec{y} \in \mathbb{R}$  als

$$\vec{x} \cdot \vec{y} = x_1 y_1 + x_2 y_2 + \dots + x_n y_n = \sum_{i=1}^n x_i y_i. \quad (3.1)$$

Die fundamentalen Eigenschaften eines Skalarprodukts  $\vec{x} \cdot \vec{y}$  als Funktion von  $\mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}$  sind folgende:

S1)  $\vec{x} \cdot \vec{x} \geq 0 \quad \forall \vec{x} \in \mathbb{R}^n \quad \text{und} \quad \vec{x} \cdot \vec{x} = 0 \quad \Rightarrow \quad \vec{x} = \vec{0}.$

Dies ist offensichtlich, da

$$\vec{x} \cdot \vec{x} = \sum_i x_i^2 \quad \text{und} \quad x_i \in \mathbb{R}.$$

S2)  $\vec{x} \cdot \vec{y} = \vec{y} \cdot \vec{x}.$

S3)  $\vec{x} \cdot (\vec{y} + \vec{z}) = \vec{x} \cdot \vec{y} + \vec{x} \cdot \vec{z}.$

S4)  $\vec{x} \cdot (\lambda \vec{y}) = \lambda \vec{x} \cdot \vec{y}.$

Das Skalarprodukt ist also eine bilineare positiv definierte Funktion zweier Vektorvariablen. Diese fundamentalen Eigenschaften eines allgemeinen Skalarprodukts lassen sich im Falle der kanonischen Form (3.1) leicht beweisen. Zum Beispiel

$$\begin{aligned} \vec{x} \cdot (\lambda \vec{y} + \vec{z}) &= \sum_i x_i (\lambda y_i + z_i) \\ &= \lambda \sum_i x_i y_i + \sum_i x_i z_i = \lambda \vec{x} \cdot \vec{y} + \vec{x} \cdot \vec{z}. \end{aligned}$$

Aus der Definition (3.1) ergibt sich für die Vektoren  $\hat{e}_1 = (1, 0, \dots, 0)$ ,  $\hat{e}_2 = (0, 1, \dots, 0)$  ...  $\hat{e}_n = (0, 0, \dots, 1)$  der kanonischen Basis

$$\hat{e}_i \cdot \hat{e}_j = \delta_{ij} = \begin{cases} 1 & \text{für } i = j \\ 0 & \text{für } i \neq j. \end{cases} \quad (3.2)$$

Man sagt dann, dass die Koordinatenvektoren  $\hat{e}_1, \hat{e}_2, \dots, \hat{e}_n$  eine *orthonormierte Basis* bilden.

Es ist auch interessant zu bemerken, dass für alle Vektoren der Form

$$\vec{x} = \sum_{i=1}^n x_i \hat{e}_i \quad \text{mit } x_i \in \mathbb{R}$$

die Gleichung (3.2) zusammen mit den allgemeinen Eigenschaften des Skalarprodukts S3 und S4 direkt zu (3.1) und damit zu S1 und S2 führt.

Mittels des Skalarprodukts definieren wir den absoluten Betrag oder Norm eines Vektors als

$$|\vec{x}| = \sqrt{\vec{x} \cdot \vec{x}} = \sqrt{\sum_{i=1}^n x_i^2}.$$

Mit Hilfe des Satzes von Pythagoras wird  $|\vec{x}|$  als die Länge des Vektors  $\vec{x}$  oder als Abstand zum Ursprung interpretiert.

### 3.2.2 Cauchy-Schwarz-Ungleichung

Wir können jetzt die fundamentale Cauchy-Schwarz-Ungleichung beweisen, die die Beziehung zwischen Skalarprodukt und Betrag von Vektoren deutlich macht. Es gilt

$$|\vec{x} \cdot \vec{y}| \leq |\vec{x}| |\vec{y}| \quad \forall \vec{x}, \vec{y} \in \mathbb{R}^n,$$

und zwar besteht die Gleichheit nur dann, wenn einer der beiden Vektoren ein  $\lambda$ -faches des anderen ist, d.h., wenn die Vektoren parallel oder antiparallel sind. Wir werden im Folgenden  $\vec{x} \neq \vec{0}$  annehmen, da für  $\vec{x}$  die Aussage offensichtlich gilt. Für den Beweis betrachten wir den auxiliären Vektor

$$\begin{aligned} \vec{z} &= (\vec{x} \cdot \vec{x}) \vec{y} - (\vec{x} \cdot \vec{y}) \vec{x} \\ &= |\vec{x}|^2 \vec{y} - (\vec{x} \cdot \vec{y}) \vec{x} \end{aligned}$$

und berechnen

$$\begin{aligned} |\vec{z}|^2 &= \vec{z} \cdot \vec{z} = |\vec{x}|^4 \vec{y} \cdot \vec{y} + (\vec{x} \cdot \vec{y})^2 \vec{x} \cdot \vec{x} \\ &\quad - 2 |\vec{x}|^2 (\vec{x} \cdot \vec{y}) (\vec{x} \cdot \vec{y}) \\ &= |\vec{x}|^4 |\vec{y}|^2 - |\vec{x}|^2 (\vec{x} \cdot \vec{y})^2. \end{aligned}$$

Wie bei jedem  $\vec{z} \in \mathbb{R}^n$  ist

$$|z|^2 \geq 0 \quad \Rightarrow \quad |\vec{x} \cdot \vec{y}| \leq |\vec{x}| |\vec{y}|.$$

Darüber hinaus gilt  $|z|^2 = 0$  nur wenn  $\vec{z} = \vec{0}$ , also

$$|\vec{x} \cdot \vec{y}| = |\vec{x}| |\vec{y}| \quad \Rightarrow \quad \vec{0} = \vec{z} = |\vec{x}|^2 \vec{y} - (\vec{x} \cdot \vec{y}) \vec{x}.$$

Das heißt

$$|\vec{x} \cdot \vec{y}| = |\vec{x}| |\vec{y}| \quad \Rightarrow \quad \vec{y} = \lambda \vec{x} \quad (\vec{x} \neq \vec{0}).$$

### 3.2.3 Winkel zwischen Vektoren

Wir betrachten zwei von Null verschiedene Vektoren  $\vec{x}$  und  $\vec{y}$  und berechnen das Skalarprodukt zwischen den entsprechenden auf 1 normierten Vektoren. Es gilt nach Cauchy-Schwarz

$$\left| \frac{\vec{x}}{|\vec{x}|} \cdot \frac{\vec{y}}{|\vec{y}|} \right| \leq 1$$

oder

$$-1 \leq \frac{\vec{x}}{|\vec{x}|} \cdot \frac{\vec{y}}{|\vec{y}|} \leq 1.$$

Dieses Skalarprodukt liegt also zwischen  $-1$  und  $1$ . Darüber hinaus ist

$$\frac{\vec{x}}{|\vec{x}|} \cdot \frac{\vec{y}}{|\vec{y}|} = -1, \text{ nur wenn } \vec{y} = -\vec{x}, \text{ und } \frac{\vec{x}}{|\vec{x}|} \cdot \frac{\vec{y}}{|\vec{y}|} = 1, \text{ nur wenn } \vec{y} = \vec{x}.$$

Wir definieren also den Winkel  $\theta$  zwischen den Vektoren  $\vec{x}$  und  $\vec{y}$  durch

$$\frac{\vec{x}}{|\vec{x}|} \cdot \frac{\vec{y}}{|\vec{y}|} = \cos \theta \quad (0 \leq \theta \leq \pi).$$

Äquivalent dazu ist der Ausdruck

$$\vec{x} \cdot \vec{y} = |\vec{x}| |\vec{y}| \cos \theta. \quad (3.3)$$

Diese abstrakte Definition des Winkels ist allgemein anwendbar. Sie erlaubt, den Winkel zwischen Vektoren einzuführen, selbst bei sehr hochdimensionalen Räumen oder sogar zwischen Elementen komplexerer linearer Räume wie Funktionen. Trotzdem ist es angebracht, die Äquivalenz zwischen Gleichung (3.1) und (3.3) in  $\mathbb{R}^n$  auf eine direkte Weise zu zeigen.

Wir betrachten zwei Vektoren  $\vec{x}$  und  $\vec{y}$  in  $\mathbb{R}^n$ , von denen wir annehmen, dass sie auf der von  $\hat{e}_1$  und  $\hat{e}_2$  aufgespannten Ebene liegen, d.h.,

$$\begin{aligned} \vec{x} &= (x_1, x_2, 0, \dots, 0) \\ \vec{y} &= (y_1, y_2, 0, \dots, 0). \end{aligned}$$

Dies ist eigentlich keine Einschränkung, da Gleichung (3.1) invariant bezüglich Drehungen des Koordinatensystems ist, wie wir es später zeigen werden. Wir können also annehmen, dass  $\hat{e}_1$ ,  $\hat{e}_2$ ,  $\vec{x}$  und  $\vec{y}$  in der selben Ebene liegen. Es folgt dann

$$\begin{aligned} x_1 &= |\vec{x}| \cos \varphi_x \\ x_2 &= |\vec{x}| \sin \varphi_x \end{aligned}$$

und

$$\begin{aligned} y_1 &= |\vec{y}| \cos \varphi_y \\ y_2 &= |\vec{y}| \sin \varphi_y, \end{aligned}$$

wobei  $\varphi_x$  ( $\varphi_y$ ) der Winkel zwischen  $\vec{x}$  ( $\vec{y}$ ) und  $\hat{e}_1$  ist. Aus Gleichung (3.1) erhalten wir

$$\begin{aligned} \vec{x} \cdot \vec{y} &= |\vec{x}| |\vec{y}| (\cos \varphi_x \cos \varphi_y + \sin \varphi_x \sin \varphi_y) \\ &= |\vec{x}| |\vec{y}| [\cos(\varphi_y - \varphi_x) - \sin(\varphi_y - \varphi_x)]. \end{aligned}$$

Gleichung (3.3) folgt, wenn man bedenkt, dass  $\theta = \varphi_y - \varphi_x$  der Winkel zwischen  $\vec{x}$  und  $\vec{y}$  ist.

### 3.2.4 Metrik oder Abstand in $\mathbb{R}^n$

Die Eigenschaften des absoluten Betrags eines Vektors lassen sich wie folgt zusammenfassen:

- i)  $|\lambda \vec{x}| = |\lambda| |\vec{x}| \quad \forall \lambda \in \mathbb{R}.$
- ii)  $|\vec{x}| = 0 \quad \text{und} \quad |\vec{x} = 0| \Leftrightarrow \vec{x} = \vec{0}.$
- iii) Cauchy-Schwarz-Ungleichung  
 $|\vec{x} \cdot \vec{y}| \leq |\vec{x}| |\vec{y}|.$
- iv) Dreiecksungleichung  
 $|\vec{x} + \vec{y}| \leq |\vec{x}| + |\vec{y}|.$

Dies folgt aus der Cauchy-Schwarz-Ungleichung

$$\begin{aligned} |\vec{x} + \vec{y}|^2 &= (\vec{x} + \vec{y}) \cdot (\vec{x} + \vec{y}) \\ &= |\vec{x}|^2 + |\vec{y}|^2 + 2 \vec{x} \cdot \vec{y} \\ &\leq |\vec{x}|^2 + |\vec{y}|^2 + 2 |\vec{x}| |\vec{y}| = (|\vec{x}| + |\vec{y}|)^2. \end{aligned}$$

- v)  $|\vec{x} - \vec{z}| \leq |\vec{x} - \vec{y}| + |\vec{y} - \vec{z}| \quad \forall \vec{x}, \vec{y}, \vec{z} \in \mathbb{R}^n.$

Dies folgt aus iv), wenn man  $\vec{x}$  durch  $\vec{x} - \vec{y}$  und  $\vec{y}$  durch  $\vec{y} - \vec{z}$  ersetzt.

Die Eigenschaften i), ii) und v) erlauben uns, eine Metrik oder Abstand in  $\mathbb{R}^n$  einzuführen. Wir definieren also den Abstand zwischen 2 Vektoren  $\vec{x}$  und  $\vec{y}$  als

$$\varrho(\vec{x}, \vec{y}) = |\vec{x} - \vec{y}| = \left( \sum_{i=1}^n (x_i - y_i)^2 \right)^{1/2}.$$

$\varrho$  ist also eine Abbildung von  $\mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R} \geq 0$  mit folgenden fundamentalen Eigenschaften:

M1)  $\varrho(x, y) \geq 0 \quad \forall \vec{x}, \vec{y} \in \mathbb{R}^n \quad \text{und}$   
 $\varrho(\vec{x}, \vec{y}) = 0 \Leftrightarrow \vec{x} = \vec{y}.$

M2)  $\varrho(\vec{x}, \vec{y}) = \varrho(\vec{y}, \vec{x}).$

M3)  $\varrho(\vec{x}, \vec{z}) \leq \varrho(\vec{x}, \vec{y}) + \varrho(\vec{y}, \vec{z}).$

Eine solche Abbildung  $\varrho$  mit den Eigenschaften M1 bis M3 wird eine Metrik genannt. Die damit bestückten Mengen heißen metrische Räume. Diese Definition erfüllt die intuitiven Bedingungen eines Abstands: Der Abstand zwischen zwei verschiedenen Punkten im Raum ist immer positiv (M1). Der Abstand von  $\vec{x}$  nach  $\vec{y}$  ist gleich dem von  $\vec{y}$  nach  $\vec{x}$  (M2). Darüber hinaus macht der Abstand keine Sprünge, d.h., wenn  $\vec{x}$  nah an  $\vec{y}$  liegt und  $\vec{y}$  nah an  $\vec{z}$ , dann liegt auch  $\vec{x}$  nah an  $\vec{z}$  (M3).

Damit können wir den für die Definition von Stetigkeit, Grenzwert und Ableitung sehr wichtigen Begriff der Umgebung eines Vektors oder Punktes  $\vec{a}$  in  $\mathbb{R}^n$  einführen. Als Umgebung oder  $\varepsilon$ -Umgebung eines Vektors  $\vec{a}$  in  $\mathbb{R}^3$  bezeichnet man die Menge

$$U_\varepsilon(\vec{a}) = \{\vec{x} \in \mathbb{R}^n \quad \text{mit} \quad |\vec{x} - \vec{a}| < \varepsilon\},$$

wobei  $\varepsilon > 0$ . Manchmal ist es auch nützlich, den Punkt  $\vec{a}$  von der Umgebung auszuschließen. Man betrachte dann eine Umgebung in strengem Sinne als  $\dot{U}_\varepsilon(\vec{a}) = \{\vec{x} \in \mathbb{R}^n \text{ mit } 0 < |\vec{x} - \vec{a}| < \varepsilon\}$ . Auf jeden Fall ist es wichtig zu bemerken, dass man sich dem Punkt  $\vec{a}$  aus allen möglichen Richtungen  $\hat{n} \in \mathbb{R}^n$  innerhalb  $U_\varepsilon(\vec{a})$  nähern kann, da  $\forall \hat{n} \in \mathbb{R}^n$  mit  $|\hat{n}| = 1$  ist  $\vec{x} = \vec{a} + \lambda \hat{n} \in U_\varepsilon(\vec{a}) \quad \forall \lambda$  mit  $|\lambda| < \varepsilon$ .

Bevor wir uns mit der Differentiation von Vektorfunktionen befassen, ist es sinnvoll, einige Grundbegriffe der Analysis von Funktionen einer Variablen zu wiederholen.

### 3.3 Grenzwert und Stetigkeit

Der Grenzwert ist der grundlegende Begriff der Analysis, mit dem man die Eigenschaften von Funktionen an Punkten analysieren kann, die mit der normalen Algebra unzugänglich sind. Zum Beispiel, wie groß ist der Wert von  $f(x) = \frac{\sin x}{x}$ , wenn  $x$  gegen null geht?

Intuitiv ist

$$\lim_{x \rightarrow a} f(x) = l,$$

wenn  $f(x)$  sich beliebig an  $l$  nähert, wenn  $x$  gegen  $a$  geht. Mathematisch formal sagt man

$$\lim_{x \rightarrow a} f(x) = l \quad \Leftrightarrow \quad \forall \varepsilon > 0 \quad \exists \delta > 0$$

so dass  $|f(x) - l| < \varepsilon \quad \forall |x - a| < \delta$ .

Wichtig ist zu bemerken, dass man sich dem Punkt  $x = a$  von beiden Seiten nähert ( $x > a$  und  $x < a$ ) und dass  $f(x)$  in beiden Fällen gegen denselben Wert  $l$  konvergiert.

Bei Funktionen mehrerer Variablen wird der Begriff von Grenzwert ähnlich definiert. Sei

$$f : \mathbb{R}^n \rightarrow \mathbb{R}$$

$$f(x_1, x_2, \dots, x_n) = f(\vec{r}) \quad \vec{r} \in \mathbb{R}^n$$

und sei  $\vec{a} = (a_1, a_2, \dots, a_n) \in \mathbb{R}^n$ , dann ist

$$\lim_{\vec{r} \rightarrow \vec{a}} f(\vec{r}) = l \quad \Leftrightarrow \quad \forall \varepsilon > 0 \quad \exists \delta > 0$$

so dass  $|f(\vec{r}) - l| < \varepsilon \quad \forall \vec{r} \in |\vec{r} - \vec{a}| < \delta$ .

Man nennt Umgebung  $\delta$  von  $\vec{a}$  die Menge

$$U_\delta(\vec{a}) = \{\vec{r} \in \mathbb{R}^n \text{ mit } |\vec{r} - \vec{a}| < \delta\}.$$

Man sollte sich im klaren darüber sein, dass man sich  $\vec{a}$  aus allen Richtungen nähert.

Rechenregel: Wir betrachten zwei Funktionen  $f, g : \mathbb{R}^n \rightarrow \mathbb{R}$ , bei denen

$$\lim_{\vec{r} \rightarrow \vec{a}} f(\vec{r}) = l \quad \text{und} \quad \lim_{\vec{r} \rightarrow \vec{a}} g(\vec{r}) = k.$$

Dann gilt

- i)  $\lim_{\vec{r} \rightarrow \vec{a}} (f + \alpha g) = l + \alpha k$  für  $\alpha \in \mathbb{R}$
- ii)  $\lim_{\vec{r} \rightarrow \vec{a}} (fg) = lk$
- iii)  $\lim_{\vec{r} \rightarrow \vec{a}} \frac{f}{g} = \frac{l}{k}$  für  $k \neq 0$ .

Def: Eine Funktion  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  heißt *stetig* am Punkt  $\vec{a} \in \mathbb{R}^n$ , wenn

$$\lim_{\vec{r} \rightarrow \vec{a}} f(\vec{r}) = f(\vec{a}).$$

Intuitiv würde man sagen, dass eine stetige Funktion am Punkt  $\vec{a}$  keinen Sprung hat. Die Voraussetzung ist, dass der  $\lim_{\vec{r} \rightarrow \vec{a}} f(\vec{r})$  existiert.

Beispiele:

i)  $f : \mathbb{R} \rightarrow \mathbb{R}$

$$f(x) = \begin{cases} \frac{x}{|x|} & x \neq 0 \\ 1 & x = 0. \end{cases}$$

Es ist klar, dass  $\lim_{x \rightarrow 0} f(x)$  nicht existiert. Im besten Fall könnte man sagen, dass es zwei verschiedene Grenzwerte gibt, wenn man von links ( $x < 0$ ) oder rechts ( $x > 0$ ) kommt.  $f(x)$  ist also nicht stetig bei  $x = 0$ .

ii)  $f : \mathbb{R} \rightarrow \mathbb{R}$

$$f(x) = \begin{cases} \frac{\sin(x)}{x} & x \neq 0 \\ 1 & x = 0. \end{cases}$$

Um  $\lim_{x \rightarrow 0} \frac{\sin(x)}{x}$  zu berechnen, brauchen wir eine Näherung für  $\sin(x)$  für  $|x| \ll 1$ , da sowohl  $x$  als auch  $\sin(x)$  gegen 0 gehen, wenn  $x \rightarrow 0$  geht. Aus der Graphik sieht man, dass  $\sin(x) \simeq x + O(x^3)$  für  $x \rightarrow 0$ . Daraus folgt

$$\lim_{x \rightarrow 0} \frac{\sin(x)}{x} = \lim_{x \rightarrow 0} \frac{x + O(x^3)}{x} = 1.$$

Damit ist  $f(x)$  stetig für  $x = 0$ .

Ähnlich wie bei den reellen Zahlen kann man den Grenzwert von Funktionen komplexer Variablen definieren. Für  $f : \mathbb{C} \rightarrow \mathbb{C}$  gilt

$$\lim_{z \rightarrow z_0} f(z) = w \in \mathbb{C} \quad \Leftrightarrow \quad \forall \varepsilon > 0 \quad \exists \delta > 0 \\ \text{so dass } |f(z) - w| < \varepsilon \quad \forall |z - z_0| < \delta.$$

Wichtig ist, dass man sich  $z_0$  aus allen Richtungen nähert und dass man immer gegen den gleichen Wert  $w$  konvergiert.

### 3.3.1 Einseitige und uneigentliche Grenzwerte

Bei Funktionen einer Variablen ist es manchmal nützlich, die Grenzwerte von links und von rechts zu definieren:

$$\begin{aligned}\lim_{x \rightarrow a^+} f(x) & \quad x > a \quad \text{rechtsseitiger Grenzwert} \\ \lim_{x \rightarrow a^-} f(x) & \quad x < a \quad \text{linksseitiger Grenzwert.}\end{aligned}$$

Oft betrachtet man den Grenzwert für sehr große positive ( $x \rightarrow +\infty$ ) oder negative ( $x \rightarrow -\infty$ ) Werte der Variablen  $x$ :

$$\lim_{\substack{x \rightarrow +\infty \\ (-\infty)}} f(x) = l \quad \Leftrightarrow \quad \forall \varepsilon > 0 \quad \exists K > 0 \\ \text{so dass } |f(x) - l| < \varepsilon \quad \forall x > K \quad (x < -K).$$

Und manchmal ist es so, dass die Werte der Funktion  $f(x)$  immer größer werden, wenn  $x \rightarrow a$  geht:

$$\lim_{x \rightarrow a} f(x) = +\infty \quad (-\infty) \quad \Leftrightarrow \quad \forall M > 0 \quad \exists \delta > 0 \\ \text{so dass } f(x) > M \quad (f(x) < -M) \quad \forall |x - a| < \delta.$$

Einfache Beispiele sind  $\lim_{x \rightarrow +\infty} e^{-x} = 0$  und  $\lim_{x \rightarrow 0} \frac{1}{x^2} = +\infty$ .

## 3.4 Ableitung in $\mathbb{R}$

Wir betrachten eine Funktion  $f : \mathbb{R} \rightarrow \mathbb{R}$  und berechnen für einen festen Punkt  $x_0 \in \mathbb{R}$  die Änderung  $\Delta f$  der Funktion  $f(x)$  um den Punkt  $x_0$

$$\Delta f = f(x_0 + \Delta x) - f(x_0) \quad \text{mit } \Delta x \in \mathbb{R}.$$

Wir gehen von einer stetigen Funktion aus, deshalb ist

$$\lim_{\Delta x \rightarrow 0} \Delta f = \lim_{\Delta x \rightarrow 0} f(x_0 + \Delta x) - f(x_0) = 0.$$

Die mittlere Steigung der Funktion

$$\frac{\Delta f}{\Delta x} = \frac{f(x_0 + \Delta x) - f(x_0)}{\Delta x} \quad \Delta x \neq 0$$

ist aber endlich, und zwar so, dass der Winkel  $\bar{\alpha}$  der Gerade, die durch  $(x_0, f(x_0))$  und  $(x_0 + \Delta x, f(x_0 + \Delta x))$  geht, durch

$$\tan \bar{\alpha} = \frac{\Delta f}{\Delta x}$$

gegeben ist. Die Ableitung  $f'(x_0)$  der Funktion  $f(x)$  am Punkt  $x = x_0$  ist definiert als die Steigung von  $f(x)$  am Punkt  $(x_0)$ , genauer gesagt als die Steigung der Tangente an den

Graphen  $(x, f(x))$  am Punkt  $(x_0, f(x_0))$ . Das heißt

$$f'(x_0) = \frac{df}{dx}(x_0) = \left. \frac{df}{dx} \right|_{x=x_0} = \lim_{\Delta x \rightarrow 0} \frac{\Delta f}{\Delta x} = \lim_{\Delta x \rightarrow 0} \frac{f(x_0 + \Delta x) - f(x_0)}{\Delta x}.$$

Die Funktion  $f(x)$  heißt differenzierbar am Punkt  $x_0$ , wenn dieser Grenzwert existiert. Manchmal existiert  $\lim_{\Delta x \rightarrow 0} \frac{\Delta f}{\Delta x}$  nicht, weil die rechtsseitigen und linksseitigen Grenzwerte unterschiedlich sind. In diesem Fall spricht man von linksseitiger Ableitung

$$f'(x_0^-) = \left. \frac{df}{dx} \right|_{x=x_0^-} = \lim_{\Delta x \rightarrow 0^-} \frac{\Delta f}{\Delta x}$$

und rechtsseitiger Ableitung

$$f'(x_0^+) = \left. \frac{df}{dx} \right|_{x=x_0^+} = \lim_{\Delta x \rightarrow 0^+} \frac{\Delta f}{\Delta x}.$$

Eine wichtige Eigenschaft der Ableitung ist, dass sie ermöglicht, die Funktion  $f(x)$  durch eine Gerade am Punkt  $x_0$  zu nähern:

$$f(x) = f(x_0) + f'(x_0)(x - x_0) + r(x)(x - x_0)$$

mit  $\lim_{x \rightarrow x_0} r(x) = 0$ . Das heißt, der Fehler  $r(x)(x - x_0)$ , der durch die lineare Näherung entsteht, geht schneller gegen Null als  $(x - x_0)$ .

Man schreibt auch

$$f(x) = f(x_0) + f'(x_0)(x - x_0) + O[(x - x_0)^2]$$

oder

$$\Delta f = f(x) - f(x_0) = f'(x_0)(x - x_0) + O[(x - x_0)^2].$$

Wenn man die Änderung  $\Delta x = (x - x_0)$  der Variablen  $x$  einführt, ergibt sich

$$\Delta f = f'(x_0) \Delta x + O(\Delta x^2).$$

Im Limes  $\Delta x \rightarrow 0$  ist die Änderung der Funktion proportional zur Änderung der Variable, da die Terme proportional zu  $\Delta x^2$  vernachlässigbar werden. Diese Proportionalitätsbeziehung für  $\Delta x \rightarrow 0$  und  $\Delta f \rightarrow 0$  lässt sich auf folgende kompakte Form schreiben:

$$df = f'(x_0) dx$$

oder für beliebiges  $x$  als

$$df = f'(x) dx.$$

$dx$  und  $df$  werden Differentiale genannt. Darunter ist lediglich die Proportionalität zwischen  $\Delta f$  und  $\Delta x$  für  $\Delta x \rightarrow 0$  zu verstehen. Ausdrücke wie  $dx = \lim_{\Delta x \rightarrow 0} \Delta x$  oder  $df = \lim_{\Delta x \rightarrow 0} \Delta f$  machen keinen Sinn, da diese Grenzwerte null sind. Wenn dies berücksichtigt wird, kann man auch schreiben

$$\frac{df}{dx} = f'(x).$$

Die Ableitung lässt sich also als Differentialquotient verstehen.

### 3.4.1 Rechenregeln

- i)  $(\alpha f + \beta g)' = \alpha f' + \beta g'$  für  $\alpha, \beta \in \mathbb{R}$ . Die Ableitung ist ein linearer Operator.
- ii)  $(f \cdot g)' = f'g + fg'$ .
- iii)  $\left(\frac{f}{g}\right)' = \frac{f'g - fg'}{g^2}$ .

Um ii) zu beweisen, betrachten wir die Änderung der Funktion  $f(x)g(x)$ .

$$\begin{aligned}(fg)(x + \Delta x) &= f(x + \Delta x) g(x + \Delta x) \\ &= (f + \Delta f)(g + \Delta g) \\ &= fg + f \Delta g + g \Delta f + \Delta f \Delta g\end{aligned}$$

$$\Delta(fg) = (fg)(x + \Delta x) - (fg)(x) = f \Delta g + g \Delta f + \underbrace{\Delta f \Delta g}_{\ll \Delta f \text{ oder } \Delta g}.$$

Der letzte Term  $\Delta f \Delta g$  ist vernachlässigbar klein im Verhältnis zu den anderen Termen, wenn  $\Delta x \rightarrow 0$ . Es gilt

$$\frac{\Delta(fg)}{\Delta x} = f \frac{\Delta g}{\Delta x} + g \frac{\Delta f}{\Delta x} + \Delta f \frac{\Delta g}{\Delta x}$$

und damit

$$(fg)' = \lim_{\Delta x \rightarrow 0} \frac{\Delta(fg)}{\Delta x} = f'g + fg'.$$

- iv)  $[g(f(x))]' = g'(f(x)) f'(x)$ . Dies ist die sogenannte Kettenregel.

Wir betrachten also zwei Funktionen  $f(x) : \mathbb{R} \rightarrow \mathbb{R}$  und  $g(y) : \mathbb{R} \rightarrow \mathbb{R}$  und suchen die Änderung von  $g(f(x))$ , die durch eine Änderung  $\Delta x$  der Variablen  $x$  verursacht wird. Die lineare Näherung für  $g$  ergibt

$$g(f(x + \Delta x)) = g(f + \Delta f) = g(f(x)) + g'(f(x)) \Delta f + O(\Delta f^2).$$

Andererseits wissen wir, dass

$$\Delta f = f(x + \Delta x) - f(x) = f'(x) \Delta x + O(\Delta x^2).$$

Daraus folgt

$$\begin{aligned}\Delta g &= g(f(x + \Delta x)) - g(f(x)) \\ &= g'(f(x)) f'(x) \Delta x + \underbrace{g'(f(x)) O(\Delta x^2) + O(\Delta f^2)}_{O(\Delta x^2)},\end{aligned}$$

$$\frac{\Delta g}{\Delta x} = g'(f(x)) f'(x) + O(\Delta x)$$

und

$$\lim_{\Delta x \rightarrow 0} \frac{\Delta g}{\Delta x} = [g(f(x))]' = g'(f(x)) f'(x).$$

- v) Eine weitere wichtige Regel für die Ableitung der Umkehrfunktion  $f^{-1}(x)$  lässt sich aus der Kettenregel herleiten. Wir wissen, dass

$$f(f^{-1}(x)) = x \quad \forall x.$$

Nach Ableiten auf beiden Seiten und unter Anwendung der Kettenregel folgt

$$f'(f^{-1}(x)) (f^{-1}(x))' = 1$$

und demzufolge

$$(f^{-1}(x))' = \frac{d}{dx} f^{-1}(x) = \frac{1}{f'(f^{-1}(x))}. \quad (3.4)$$

Damit lassen sich die Ableitungen der üblichen Umkehrfunktionen  $\arcsin(x)$ ,  $\arctan(x)$  oder  $\ln(x)$  leicht berechnen.

Diese Rechenregel begründet eine in der Physik sehr übliche Manipulation von Differentialen. Sei  $y = y(x)$  eine bijektive Funktion von  $x$  und  $x = x(y)$  die Umkehrfunktion. Wenn wir  $y = y(x) = f^{-1}(x)$  und  $x(y) = f(y)$  in die Gleichung (3.4) einsetzen, erhalten wir

$$\frac{dy(x)}{dx} = \frac{df^{-1}(x)}{dx} = \frac{1}{f'(f^{-1}(x))} = \frac{1}{\frac{dx}{dy}(y(x))}$$

oder, wenn man das Risiko eingeht, die Argumente der Funktion wegfassen zu lassen, noch einfacher

$$\frac{dy}{dx} = \frac{1}{\frac{dx}{dy}}.$$

Diese Manipulationen sind zwar praktisch, aber mit Vorsicht zu genießen.

### 3.4.2 Beispiele

- i) Wir betrachten die Funktion  $f(z) = z^2$  mit  $f : \mathbb{C} \rightarrow \mathbb{C}$ . Die Ableitung in  $\mathbb{C}$  ist durch folgenden Grenzwert gegeben:

$$\begin{aligned} f'(z) &= \lim_{\Delta z \rightarrow 0} \frac{(z + \Delta z)^2 - z^2}{\Delta z} = \lim_{\Delta z \rightarrow 0} \frac{2z \Delta z + \Delta z^2}{\Delta z} = \\ &= 2z \lim_{\Delta z \rightarrow 0} \frac{\Delta z}{\Delta z} + \lim_{\Delta z \rightarrow 0} \Delta z = 2z. \end{aligned}$$

- ii) Wir betrachten  $f(z) = z^k$  und wollen zeigen, dass  $f'(z) = k z^{k-1}$ . Wir wissen, dass dies für  $k \leq 2$  gilt. Wir gehen also rekursiv vor, unter der Annahme, dass für  $k' < k$  die Behauptung gilt. Wir schreiben  $f(z) = z z^{k-1}$  und erhalten

$$f'(z) = z^{k-1} + z (z^{k-1})' = z^{k-1} + z (k-1) z^{k-2} = k z^{k-1}.$$

iii) Die Exponentialfunktion komplexer Zahlen ist durch folgende Taylor-Reihe definiert worden:

$$e^z = \sum_{k=0}^{+\infty} \frac{1}{k!} z^k$$

$$\frac{d}{dz} e^z = (e^z)' = \sum_{k=0}^{+\infty} \frac{1}{k!} k z^{k-1} = \sum_{k=1}^{+\infty} \frac{1}{(k-1)!} z^{k-1} = e^z.$$

iv) Die Logarithmusfunktion ist als die Umkehrfunktion der Exponentialfunktion definiert. Wir haben also

$$e^{\ln z} = z \quad \forall z \neq 0, z \in \mathbb{C}.$$

Wir leiten auf beiden Seiten unter Anwendung der Kettenregel ab und erhalten

$$\frac{dz}{dz} = (z)' = 1 = e^{\ln z} (\ln z)' = z (\ln z)' \Rightarrow (\ln z)' = \frac{d}{dz} \ln z = \frac{1}{z}.$$

Wir betrachten nun einige wichtige Funktionen einer reellen Variablen.

v)  $\frac{d}{dx} \sin(x) = \cos(x)$

vi)  $\frac{d}{dx} \cos(x) = \frac{d}{dx} [\sin(x + \frac{\pi}{2})] = \cos(x + \frac{\pi}{2}) = -\sin(x).$

vii)  $\frac{d}{dx} \tan(x) = \frac{d}{dx} \frac{\sin(x)}{\cos(x)} = \frac{\cos^2(x) + \sin^2(x)}{\cos^2(x)} = \frac{1}{\cos^2(x)}.$

$\frac{d}{dx} \sin(x) = \cos(x)$  lässt sich mit Hilfe folgender trigonometrischer Beziehung beweisen:

$$\sin \alpha - \sin \beta = 2 \sin \left( \frac{\alpha - \beta}{2} \right) \cos \left( \frac{\alpha + \beta}{2} \right).$$

$$\frac{\Delta f}{\Delta x} = \frac{\sin(x + \Delta x) - \sin(x)}{\Delta x} = \frac{2 \sin \left( \frac{\Delta x}{2} \right) \cos \left( x + \frac{\Delta x}{2} \right)}{\Delta x}$$

$$\frac{\Delta(\sin(x))}{\Delta x} = \underbrace{\frac{\sin \left( \frac{\Delta x}{2} \right)}{\left( \frac{\Delta x}{2} \right)}}_1 \underbrace{\cos \left( x + \frac{\Delta x}{2} \right)}_{\cos(x)} \xrightarrow{\Delta x \rightarrow 0} \cos(x)$$

viii)  $f(x) = \sqrt{x^2 + 1}.$

$$x \xrightarrow{g} x^2 + 1 \xrightarrow{h} \sqrt{x^2 + 1} \quad f(x) = h(g(x)).$$

$$f'(x) = h'(g(x)) g'(x) = \frac{1}{2\sqrt{x^2 + 1}} 2x = -\frac{x}{\sqrt{x^2 + 1}}$$

ix)  $f(x) = 2\sqrt{\cos(x^3)}$

$$f'(x) = \frac{2}{2\sqrt{\cos(x^3)}} \left[ -\sin(x^3) \right] 3x^2 = \frac{-3x^2 \sin(x^3)}{\sqrt{\cos(x^3)}}.$$

x)  $F(x) = e^{\sqrt{1-x^2}}$

$$x \xrightarrow{f} | - x^2 \xrightarrow{g} \sqrt{1-x^2} \xrightarrow{h} e^{\sqrt{1-x^2}} \quad F(x) = h(g(f(x))).$$

$$F'(x) = h'(g(f(x))) g'(f(x)) f'(x) = e^{\sqrt{1-x^2}} \frac{1}{2\sqrt{1-x^2}} (-2x) = -\frac{x e^{\sqrt{1-x^2}}}{\sqrt{1-x^2}}.$$

xi)  $f(x) = \arctan(x) = \tan^{-1}(x) \quad (\neq \frac{1}{\tan x}!)$

$$\frac{d}{dx} f^{-1}(x) = \frac{1}{\frac{df}{dx}(f^{-1}(x))}$$

$$\frac{d}{dx} \tan^{-1}(x) = \frac{1}{\left. \frac{d \tan(y)}{dy} \right|_{y=\tan^{-1}(x)}}.$$

Wir wissen, dass  $\frac{d}{dy} \tan(y) = \frac{1}{\cos^2 y}$ , und deshalb

$$\left. \frac{d \tan(y)}{dy} \right|_{x=\tan^{-1}(x)} = \frac{1}{\cos^2(\tan^{-1}(x))}.$$

Daraus folgt

$$\begin{aligned} \frac{d}{dx} \tan^{-1}(x) &= \frac{1}{\frac{1}{\cos^2(\tan^{-1}(x))}} = \frac{1}{\frac{\cos^2(\tan^{-1}(x)) + \sin^2(\tan^{-1}(x))}{\cos^2(\tan^{-1}(x))}} \\ &= \frac{1}{1 + \tan^2(\tan^{-1}(x))} = \frac{1}{1 + x^2}. \end{aligned}$$

### 3.5 L'Hopitalsche Regel

Die Ableitung ist auch sehr nützlich, um Grenzwerte der Form

$$\lim_{x \rightarrow a} \frac{f(x)}{g(x)}$$

zu bestimmen, wenn

$$\lim_{x \rightarrow a} f(x) = \lim_{x \rightarrow a} g(x) = \begin{cases} 0 \\ \infty. \end{cases}$$

In diesem Fall gilt nämlich die L'Hopitalsche Regel, die besagt, dass wenn

$$\lim_{x \rightarrow a} \frac{f'(x)}{g'(x)} = l$$

existiert, dann ist

$$\lim_{x \rightarrow a} \frac{f(x)}{g(x)} = \lim_{x \rightarrow a} \frac{f'(x)}{g'(x)} = l.$$

Zum Beispiel

$$\lim_{x \rightarrow 0} \frac{\sin x}{x} = \lim_{x \rightarrow 0} \frac{(\sin x)'}{(x)'} = \lim_{x \rightarrow 0} \frac{\cos x}{1} = 1.$$

Ein weiteres Beispiel ist

$$\lim_{x \rightarrow +\infty} e^{-x} x^n = \lim_{x \rightarrow +\infty} \frac{x^n}{e^x}.$$

Durch  $n$ -fache Anwendung der Regel erhalten wir

$$\lim_{x \rightarrow +\infty} e^{-x} x^n = \lim_{x \rightarrow +\infty} \frac{n!}{e^x} = 0.$$

Die Regel lässt sich leicht begründen für den Fall  $f(x) \rightarrow 0$  und  $g(x) \rightarrow 0$ , wenn  $g'(a) \neq 0$ . In diesem Fall haben wir

$$\begin{aligned} f(x) &= f'(a) \Delta x + O(\Delta x^2). \\ g(x) &= g'(a) \Delta x + O(\Delta x^2) \quad \text{mit } \Delta x = x - a \end{aligned}$$

Daraus folgt

$$\lim_{x \rightarrow a} \frac{f(x)}{g(x)} = \lim_{x \rightarrow a} \frac{f'(a) \Delta x + O(\Delta x^2)}{g'(a) \Delta x + O(\Delta x^2)} = \lim_{\Delta x \rightarrow 0} \frac{\frac{f'(a)}{g'(a)} + \frac{O(\Delta x^2)}{g'(a) \Delta x}}{1 + \frac{O(\Delta x^2)}{g'(a) \Delta x}} = \frac{f'(a)}{g'(a)}.$$

### 3.6 Höhere Ableitungen

Die höheren Ableitungen  $f^{(k)}(x) = \frac{d^k f}{dx^k}$  einer Funktion  $f(x)$  werden rekursiv wie folgt definiert:

$$\frac{d^k f}{dx^k} = f^{(k)}(x) = \left( f^{(k-1)}(x) \right)' = \frac{d}{dx} \left( f^{(k-1)}(x) \right),$$

wobei  $k \geq 0$  und  $f^{(0)}(x) = f(x)$ . Voraussetzung für die Existenz der  $k$ -ten Ableitung ist natürlich die Existenz der  $(k-1)$ -ten Ableitung und die Differenzierbarkeit von  $f^{(k-1)}(x)$ .

### 3.7 Taylorsches Approximationspolynom

Wir betrachten eine am Punkt  $a$   $n$ -mal differenzierbare Funktion  $f(x)$  und wir suchen eine möglichst genaue lokale Näherung der Funktion  $f(x)$  an diesem Punkt mittels eines Polynoms  $n$ -ten Grades. Wie wir zeigen werden, wird dies erreicht, indem man die ersten  $n$  Ableitungen des Polynoms gleich der ersten  $n$  Ableitungen von  $f(x)$  am Punkt  $a$  setzt.

Sei  $P(x) = \sum_{k=0}^n \alpha_k (x-a)^k$  mit den Koeffizienten  $\alpha_k \in \mathbb{R}$ . Wir verlangen also, dass

$$\left. \frac{d^k}{dx^k} P(x) \right|_{x=a} = k! \alpha_k = f^{(k)}(a).$$

Deshalb nimmt das  $n$ -te Taylorsche Approximationspolynom von  $f$  am Punkt  $a$  die Form

$$\mathcal{T}_{n,f,a}(x) = \sum_{k=0}^n \frac{1}{k!} f^{(k)}(a) (x-a)^k$$

an. Ohne Verlust der Allgemeinheit lässt sich die Funktion  $f(x)$  in der Form

$$f(x) = \mathcal{T}_n(x) + \varepsilon(x) \quad \forall x$$

schreiben. Der Fehler, der durch die Taylorsche Näherung entsteht, ist durch die Funktion  $\varepsilon(x)$  gegeben. Die Qualität von  $\mathcal{T}_n(x)$  als lokale Näherung für  $f(x)$  lässt sich dadurch beweisen, dass

$$\lim_{x \rightarrow a} \frac{\varepsilon(x)}{(x-a)^n} = 0$$

ist. Das heißt, der Fehler geht schneller gegen null, wenn  $x \rightarrow a$  geht, als  $(x-a)^n$ . Der Beweis erfolgt durch  $n$ -fache Anwendung der L'Hopitalschen Regel

$$\lim_{x \rightarrow a} \frac{\varepsilon(x)}{(x-a)^n} = \lim_{x \rightarrow a} \frac{d^n \varepsilon(x)}{dx^n} = \lim_{x \rightarrow a} [f^{(n)}(x) - f^{(n)}(a)] = 0.$$

Beispiele:

i) Wir betrachten  $f(x) = \frac{1}{1+x}$ , dessen Ableitungen  $f'(x) = \frac{-1}{(1+x)^2}$  und  $f''(x) = \frac{2}{(1+x)^3}$  sind. Daraus folgt für  $x=0$

$$f(0) = 1, \quad f'(0) = -1, \quad \text{und} \quad f''(0) = 2.$$

Die Taylorsche Näherung zweiter Ordnung lautet

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

ii) Für  $f(x) = \sin x$  haben wir

$$f'(x) = \cos(x), \quad f''(x) = -\sin(x), \quad f^{(3)}(x) = -\cos(x) \quad \text{und} \quad f^{(4)}(x) = \sin(x).$$

Daraus folgt für  $x = 0$

$$f(0) = 0, \quad f^{(1)}(0) = 1, \quad f^{(2)}(0) = 0, \quad f^{(3)}(0) = -1 \quad \text{und} \quad f^{(4)}(0) = 0.$$

Das Taylorsche Approximationspolynom ist dann

$$\sin x \simeq x - \frac{1}{6}x^3 + O(x^5).$$

### 3.8 Taylor-Reihe

Ist die Funktion  $f(x)$  am Punkt  $a$  beliebig oft differenzierbar, so lässt sich dort die Taylor-Reihe von  $f$  konstruieren

$$\mathcal{T}_\infty(x) = \sum_{k=0}^{+\infty} \frac{1}{k!} f^{(k)}(a) (x - a)^k,$$

mit der man Funktionen  $f(x)$  darstellen kann, die gewisse Bedingungen erfüllen

$$f(x) = \sum_{k=0}^{+\infty} \frac{1}{k!} f^{(k)}(a) (x - a)^k.$$

Eine genauere Charakterisierung der Funktionen, die sich durch eine Taylor-Reihe darstellen lassen, und der Umgebung von  $a$ , wo die Reihe konvergiert, soll später im Rahmen der Funktionen komplexer Variablen diskutiert werden. Einige Beispiele werden im Folgenden diskutiert:

Beispiele:

i) Die Taylor-Reihe der Exponentialfunktion:

$$f(x) = e^x \quad \Rightarrow \quad f^{(n)}(x) = e^x \quad \Rightarrow \quad f^{(n)}(0) = 1 \quad \forall n \geq 0.$$

Die Taylor-Reihe

$$e^x = \sum_{k=0}^{\infty} \frac{1}{k!} e^0 x^k = \sum_{k=0}^{\infty} \frac{x^k}{k!}$$

stimmt mit der Definition der Exponentialfunktion  $\forall x \in \mathbb{R}$  überein.

ii) Für  $f(x) = \sin(x)$  haben wir  $f'(x) = \cos(x)$ ,  $f''(x) = -\sin(x)$ ,  $f^{(3)}(x) = -\cos(x)$  und  $f^{(4)}(x) = \sin(x)$ . Daraus folgt  $\forall l \geq 0$

$$\begin{aligned} f^{(4l)}(0) &= 0 \\ f^{(4l+1)}(0) &= 1 \\ f^{(4l+2)}(0) &= 0 \\ f^{(4l+3)}(0) &= -1. \end{aligned}$$

Damit erhalten wir

$$\begin{aligned}\sin x &= x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + \dots \\ &= \sum_{k=0}^{+\infty} (-1)^k \frac{x^{2k+1}}{(2k+1)!} = \sum_{k=0}^{+\infty} i^{2k} \frac{x^{2k+1}}{(2k+1)!} = -i \sum_{k=0}^{+\infty} \frac{(ix)^{2k+1}}{(2k+1)!}.\end{aligned}$$

iii) Für  $f(x) = \cos(x)$  gilt  $f'(x) = -\sin(x)$ ,  $f''(x) = -\cos(x)$ ,  $f^{(3)}(x) = \sin(x)$  und  $f^{(4)}(x) = \cos(x)$ . Daraus folgt  $\forall l \geq 0$

$$\begin{aligned}f^{(4l)}(0) &= 1 \\ f^{(4l+1)}(0) &= 0 \\ f^{(4l+2)}(0) &= -1 \\ f^{(4l+3)}(0) &= 0.\end{aligned}$$

Damit ist die Taylor-Reihe

$$\begin{aligned}\cos x &= 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \frac{x^6}{6!} + \dots \\ &= \sum_{k=0}^{+\infty} (-1)^k \frac{x^{2k}}{(2k)!} = \sum_{k=0}^{+\infty} \frac{(ix)^{2k}}{(2k)!}.\end{aligned}$$

iv) Daraus folgt die Eulersche Formel

$$\cos x + i \sin x = \sum_{k=0}^{+\infty} \frac{(ix)^{2k}}{(2k)!} + \sum_{k=0}^{+\infty} \frac{(ix)^{2k+1}}{(2k+1)!} = \sum_{k=0}^{+\infty} \frac{(ix)^k}{k!} = e^{ix}.$$

### 3.9 Ableitung und lokale Eigenschaften

Die erste und zweite Ableitung einer Funktion an einem Punkt  $a$  geben eine sehr wertvolle Information über die lokalen Eigenschaften der Funktion an diesem Punkt.

Wir betrachten eine zweimal stetig differenzierbare Funktion  $f(x)$  in einem Intervall  $I \in \mathbb{R}$ . Die geometrische Interpretation der Ableitung als die Steigung der Tangente an den Graphen  $(x, f(x))$  im Punkt  $(a, f(a))$  lässt uns folgende Regeln formulieren:

- i)  $f(x)$  steigt (nimmt ab) linear mit zunehmendem  $x \Leftrightarrow f'(x) > 0$  ( $f'(x) < 0$ ).
- ii)  $f''(a) = 0 \Leftrightarrow$  die Tangente horizontal ist. In diesem Fall haben wir einen Extremwert, d.h., ein Maximum, ein Minimum oder einen Sattelpunkt.
- iii)  $f''(x) > 0$  ( $f''(x) < 0$ )  $\Leftrightarrow$  die Steigung der Tangente zunimmt (abnimmt) mit zunehmendem  $x \Leftrightarrow$  die Krümmung des Graphen  $(x, f(x))$  ist nach oben (unten).
- iv)  $f''(a) = 0 \Leftrightarrow$  die Steigung der Tangente einen Extremwert am Punkt  $a$  erreicht. Dies kann eine Änderung der Konvexität bedeuten, wenn  $\text{sign}(f''(a + \delta)) \neq \text{sign}(f''(a - \delta))$ . Es kann sich aber auch um einen Extremwert handeln (z.B.  $x^4$  bei  $x = 0$ ).

Damit lassen sich Funktionen mit Hilfe der ersten und zweiten Ableitung analysieren.

Man sollte in der Praxis darauf achten, dass die Minima und Maxima von Funktionen sich am Rande des Definitionsbereichs oder dort, wo die Ableitung eine Unstetigkeit aufweist, befinden können. Diese Punkte sind gesondert zu untersuchen.

Beispiel:

Das Minimum von  $f(x)$  liegt am Punkt  $b$ , obwohl dort  $f'(b) < 0$ , und das Minimum von  $g(x)$  liegt am Punkt  $c$ , wo die Ableitung unstetig ist (siehe Fig. ...).

### 3.10 Vektorwertige Funktionen eines Parameters: Raumkurven

Wir interessieren uns für Vektoren  $\vec{a}(t) \in \mathbb{R}^3$ , die eine Funktion eines Parameters  $t$  sind, also für Abbildungen der Form

$$\begin{aligned} \vec{a} : [t_1, t_2] &\rightarrow \mathbb{R}^3 \\ t &\rightarrow \vec{a}(t). \end{aligned}$$

Man spricht von einer Parameterdarstellung des Vektors  $\vec{a}$ . Darunter findet man die Raumkurven, die man als stetige vektorwertige Funktionen eines Parameters definieren kann, aber auch andere Vektoren wie ein homogenes elektrisches oder magnetisches Feld, das von der Zeit abhängt. Schon bekannte Beispiele von Kurven sind i) die Gerade:  $\vec{r} = \vec{r}_0 + \lambda \vec{a}$  mit  $\vec{r}_0, \vec{a}$  vorgegebene Vektoren  $\in \mathbb{R}^3$  und  $\lambda$  beliebig  $\in \mathbb{R}$  und ii) der Kreis  $\vec{r}(\varphi) = R(\cos \varphi, \sin \varphi, 0)$  mit  $0 \leq \varphi \leq 2\pi$ .

### 3.11 Ableitung vektorwertiger Funktionen

Wir betrachten eine Punktmasse, die eine gewisse Trajektorie oder Bahn im Raum durchläuft. Zu jedem Zeitpunkt  $t$  gehört eine Position

$$\vec{r}(t) = (x(t), y(t), z(t)).$$

Wir interessieren uns für die instantane Geschwindigkeit dieser Masse am Zeitpunkt  $t$ . In Analogie zur Bewegung in einer Dimension definieren wir die Geschwindigkeit als die Änderung der Position  $\vec{r}(t)$  an zwei nah beieinanderliegenden Zeiten  $t$  und  $t + \Delta t$  geteilt durch das Zeitintervall  $\Delta t$ :

$$\vec{v}(t) = \lim_{\Delta t \rightarrow 0} \frac{\vec{r}(t + \Delta t) - \vec{r}(t)}{\Delta t}.$$

Der Limes  $\Delta t \rightarrow 0$  sorgt dafür, dass  $\vec{v}(t)$  der instantanen Geschwindigkeit zur Zeit  $t$  entspricht. Die Geschwindigkeit ist wohl definiert, wenn dieser Grenzwert existiert. Man merke, dass die Geschwindigkeit ein Vektor ist, da sie der Unterschied zweier Vektoren ist. Zum Beispiel werden, wenn ein Körper gedreht wird, alle Positionsvektoren gedreht, und damit alle Unterschiede  $\Delta \vec{r}$ , also auch  $\vec{v}(t)$ . Die Geschwindigkeit motiviert die Verallgemeinerung des Begriffs der Ableitung auf vektorwertige Funktionen.

Für eine beliebige vektorwertige Funktion eines Parameters

$$\vec{a} : \mathbb{R} \rightarrow \mathbb{R}^3$$

definieren wir die Ableitung  $\frac{d\vec{a}}{dt}$  am Zeitpunkt oder Parameter  $t_0$  als

$$\left. \frac{d\vec{a}}{dt} \right|_{t_0} = \dot{\vec{a}} = \lim_{\Delta t \rightarrow 0} \frac{\vec{a}(t_0 + \Delta t) - \vec{a}(t_0)}{\Delta t}.$$

Wenn  $\vec{a} = (a_1(t), a_2(t), a_3(t))$  ist, dann haben wir

$$\frac{d\vec{a}}{dt} = \left( \frac{da_1}{dt}, \frac{da_2}{dt}, \frac{da_3}{dt} \right).$$

Genau wie bei den Skalarfunktionen einer Variablen lässt sich die Raumkurve  $\vec{a}(t)$  am Zeitpunkt oder Parameter  $t_0$  durch eine Gerade nähern, wobei die Ableitung  $\left. \frac{d\vec{a}}{dt} \right|_{t_0}$  der Richtung der Gerade entspricht:

$$\vec{a}(t) = \vec{a}(t_0) + \left. \frac{d\vec{a}}{dt} \right|_{t_0} (t - t_0) + \vec{\varepsilon}(t),$$

wobei  $\vec{\varepsilon}(t)$  der Fehler oder die Abweichung zwischen Kurve und Gerade ist. Die Gerade ist tangential, weil die Abweichung schneller gegen Null geht als  $(t - t_0)$ . Es gilt nämlich

$$\frac{\vec{\varepsilon}(t)}{(t - t_0)} = \frac{\vec{a}(t) - \vec{a}(t_0)}{t - t_0} - \left. \frac{d\vec{a}}{dt} \right|_{t_0}$$

und deshalb

$$\lim_{t \rightarrow t_0} \frac{\vec{\varepsilon}(t)}{t - t_0} = \vec{0}.$$

### 3.12 Ableitung von Produkten von Vektoren

Für beliebige differenzierbare vektorwertige Funktionen  $\vec{a}(t)$  und  $\vec{b}(t)$  gelten folgende Beziehungen, die man leicht zeigen kann, indem man die Produktregel für die Ableitung von Funktionen einer Variablen anwendet.

i)  $\frac{d}{dt} (\vec{a} + \vec{b}) = \frac{d\vec{a}}{dt} + \frac{d\vec{b}}{dt}.$

ii)  $\frac{d}{dt} (\vec{a} \cdot \vec{b}) = \frac{d\vec{a}}{dt} \cdot \vec{b} + \vec{a} \cdot \frac{d\vec{b}}{dt}.$  Dies folgt aus

$$\frac{d}{dt} \sum_i a_i b_i = \sum_i \left( \frac{da_i}{dt} b_i + a_i \frac{db_i}{dt} \right) = \frac{d\vec{a}}{dt} \cdot \vec{b} + \vec{a} \cdot \frac{d\vec{b}}{dt}.$$

Als Spezialfälle gelten

$$\frac{d}{dt} (\vec{a} \cdot \vec{a}) = \frac{d}{dt} a^2 = 2 \vec{a} \cdot \frac{d\vec{a}}{dt}$$

und auch

$$\frac{da}{dt} = \frac{d}{dt} \sqrt{\vec{a} \cdot \vec{a}} = \frac{1}{2\sqrt{\vec{a} \cdot \vec{a}}} \cdot 2\vec{a} \cdot \vec{a} = \vec{a} \cdot \frac{\vec{a}}{a}$$

$$\frac{da}{dt} = \frac{d|\vec{a}|}{dt} = \frac{\vec{a} \cdot \vec{a}}{|\vec{a}|}.$$

Hieraus kann man eine für die klassische Dynamik interessante Bemerkung machen. Eine Kraft ändert den Betrag der Geschwindigkeit einer Punktmasse nicht, wenn sie senkrecht zur Geschwindigkeit wirkt. Aus dem Newtonschen Gesetz haben wir

$$m \frac{d\vec{v}}{dt} = \vec{F} \quad \text{und} \quad \frac{dv}{dt} = \frac{\vec{v}}{v} \cdot \frac{d\vec{v}}{dt} = \frac{\vec{v}}{v} \cdot \vec{F},$$

und deshalb

$$\frac{dv}{dt} = 0 \quad \Leftrightarrow \quad \vec{v} \cdot \vec{F} = 0 \quad \Leftrightarrow \quad \vec{v} \perp \vec{F}.$$

iii) Wenn  $\lambda(t) : \mathbb{R} \rightarrow \mathbb{R}$  eine Skalarfunktion des Parameters  $t$  ist, gilt

$$\frac{d}{dt} (\lambda \vec{a}) = \frac{d\lambda}{dt} \vec{a} + \lambda \frac{d\vec{a}}{dt}.$$

iv)  $\frac{d}{dt} (\vec{a} \times \vec{b}) = \frac{d\vec{a}}{dt} \times \vec{b} + \vec{a} \times \frac{d\vec{b}}{dt}$ . Dies lässt sich wie folgt zeigen:

$$\frac{d}{dt} \left( \sum_{ijk} \varepsilon_{ijk} a_i b_j \hat{e}_k \right) = \sum_{ijk} \varepsilon_{ijk} \left( \frac{da_i}{dt} b_j + a_i \frac{db_j}{dt} \right) \hat{e}_k = \frac{d\vec{a}}{dt} \times \vec{b} + \vec{a} \times \frac{d\vec{b}}{dt}.$$

v)  $\frac{d}{dt} [\vec{a} \cdot (\vec{b} \times \vec{c})] = \frac{d\vec{a}}{dt} \cdot (\vec{b} \times \vec{c}) + \vec{a} \cdot \left( \frac{d\vec{b}}{dt} \times \vec{c} \right) + \vec{a} \cdot \left( \vec{b} \times \frac{d\vec{c}}{dt} \right)$ . Der Beweis ist ähnlich wie bei iv).

Daraus erhalten wir die Ableitung der Determinante einer Matrix:

$$\frac{d}{dt} \begin{vmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{vmatrix} = \begin{vmatrix} \dot{a}_1 & \dot{a}_2 & \dot{a}_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{vmatrix} + \begin{vmatrix} a_1 & a_2 & a_3 \\ \dot{b}_1 & \dot{b}_2 & \dot{b}_3 \\ c_1 & c_2 & c_3 \end{vmatrix} + \begin{vmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ \dot{c}_1 & \dot{c}_2 & \dot{c}_3 \end{vmatrix}.$$

### 3.13 Skalarfelder

Ein skalares Feld  $\varphi(x, y, z)$  ist eine auf  $\mathbb{R}^3$  definierte reelle Funktion

$$\varphi : \mathbb{R}^3 \rightarrow \mathbb{R},$$

zum Beispiel

$$\varphi(\vec{r}) = \frac{1}{r} = \frac{1}{\sqrt{x^2 + y^2 + z^2}}$$

oder

$$\varphi(\vec{r}) = \vec{a} \cdot \vec{r} = a_1x + a_2y + a_3z.$$

Die Bezeichnung ‘‘Skalar’’ kommt daher, dass die Werte der Funktion an einem Punkt  $\vec{P}$  sich nicht ändern, wenn man eine Transformation des Koordinatensystems (Drehung oder Inversion) durchführt. Beispiele dafür sind die Dichteverteilung  $\varrho(x, y, z)$  oder die Temperatur  $T(x, y, z)$  eines Gases. Skalarfelder können auch von der Zeit explizit abhängen. In diesem Fall haben wir  $\varrho = \varrho(x, y, z, t)$  oder  $T = T(x, y, z, t)$ .

Wenn man das Koordinatensystem ändert, ändert sich im Allgemeinen der Ausdruck der physikalischen Skalarfelder als Funktion der neuen Koordinaten, und zwar so, dass der Wert der Größe an jedem Punkt im Raum unverändert bleibt. Bei der Transformation des Koordinatensystems  $(\hat{e}_1, \hat{e}_2, \hat{e}_3)$  zu dem System  $(\hat{e}'_1, \hat{e}'_2, \hat{e}'_3)$  gilt  $\sum_i x_i \hat{e}_i = \sum_i x'_i \hat{e}'_i$ . Es folgen dann die Beziehungen  $x'_j = \sum_i \hat{e}'_j \cdot \hat{e}_i x_i$  und  $x_i = \sum_k \hat{e}_i \cdot \hat{e}'_k x'_k$  (siehe Abschnitt 2.7). Die Transformation der Koordinaten ist also  $(x_1, x_2, x_3) \rightarrow (x'_1, x'_2, x'_3)$ . Bei der Transformation der Skalarfunktion  $\varphi \rightarrow \varphi'$  gilt

$$\varphi(x_1, x_2, x_3) = \varphi'(x'_1, x'_2, x'_3).$$

Zum Beispiel bei der Funktion  $\varphi(x, y, z) = mgz$  und dem Koordinatenwechsel

$$\begin{aligned} x &\rightarrow y' \\ y &\rightarrow z' \quad \Rightarrow \quad \varphi'(x', y', z') = mgx' \\ z &\rightarrow x' \end{aligned}$$

### 3.14 Partielle Ableitung

Wir betrachten  $f(x, y, z)$  als Funktion der Variable  $x$  bei festgehaltenem, aber beliebigem  $y$  und  $z$ . So betrachtet ist  $\varphi(x, y, z)$  eine einfache Funktion von  $x$ . Wir können also die partielle Ableitung von  $\varphi$  nach  $x$  definieren als

$$\partial_x f = \frac{\partial f}{\partial x} = \left. \frac{\partial f}{\partial x} \right|_{y,z} = \lim_{\Delta x \rightarrow 0} \frac{f(x + \Delta x, y, z) - f(x, y, z)}{\Delta x}.$$

Die partielle Ableitung ist die Steigung der Tangente an der Kurve  $(x, \varphi(x, y, z))$  bei konstantem  $y$  und  $z$ . Es gilt also die lineare Näherung von  $f$  als Funktion der Verschiebung  $\Delta x$ :

$$f(x + \Delta x, y, z) - f(x, y, z) = \frac{\partial f}{\partial x}(x, y, z) \Delta x + O(\Delta x^2).$$

Entsprechend lassen sich die partiellen Ableitungen nach  $y$  und  $z$  definieren

$$\partial_y f = \frac{\partial f}{\partial y} = \left. \frac{\partial f}{\partial y} \right|_{x,z} = \lim_{\Delta y \rightarrow 0} \frac{f(x, y + \Delta y, z) - f(x, y, z)}{\Delta y}$$

mit

$$f(x, y + \Delta y, z) - f(x, y, z) = \frac{\partial f}{\partial y}(x, y, z) \Delta y + O(\Delta y^2).$$

Beispiel:

$$f(x, y) = x^2 \ln y + e^z$$

$$\frac{\partial f}{\partial x} = 2x \ln y \quad \frac{\partial f}{\partial y} = \frac{x^2}{y} \quad \frac{\partial f}{\partial z} = e^z.$$

Sind die ersten partiellen Ableitungen von  $f$  differenzierbar, so lassen sich höhere Ableitungen analog definieren, zum Beispiel

$$\frac{\partial^2 f}{\partial x \partial y} = \frac{\partial}{\partial x} \left( \frac{\partial f}{\partial y} \right).$$

Dabei sollte man merken, dass bei der ersten Ableitung  $\left. \frac{\partial f}{\partial y} \right|_{x,z}$   $x$  und  $z$  festgehalten werden (während man nach  $y$  ableitet), und dass in der zweiten Ableitung  $y$  und  $z$  fest bleiben (während man  $\left. \frac{\partial f}{\partial y} \right|_{x,y}$  nach  $x$  ableitet).

Beispiel:

$$\frac{\partial}{\partial x} \left( \frac{\partial f}{\partial y} \right) = \frac{2x}{y} \quad \text{und} \quad \frac{\partial}{\partial y} \left( \frac{\partial f}{\partial x} \right) = \frac{2x}{y}.$$

### 3.15 Satz von Schwarz

Für zweimal stetig differenzierbare Funktionen  $f(x_1, x_2, \dots, x_n)$  mit  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  gilt der Satz von Schwarz

$$\frac{\partial^2 f}{\partial x_i \partial x_j} = \frac{\partial^2 f}{\partial x_j \partial x_i}.$$

Dies besagt, dass die Reihenfolge der partiellen Ableitungen gleichgültig ist, wenn die partiellen Ableitungen existieren. Für den Beweis führen wir die Notation

$$\frac{\partial f}{\partial x} = \partial_x f$$

ein und zeigen

$$\partial_x(\partial_y f) = \lim_{\Delta x \rightarrow 0} \frac{\partial_y f(x + \Delta x, y) - \partial_y f(x, y)}{\Delta x}$$

$$\lim_{\Delta x \rightarrow 0} \frac{1}{\Delta x} \left( \left\{ \lim_{\Delta y \rightarrow 0} \frac{f(x + \Delta x, y + \Delta y) - f(x + \Delta x, y)}{\Delta y} \right\} - \left\{ \lim_{\Delta y \rightarrow 0} \frac{f(x, y + \Delta y) - f(x, y)}{\Delta y} \right\} \right).$$

Die in der Hypothese verlangte zweimal stetige Differenzierbarkeit der Funktion erlaubt uns, die Grenzwerte zu vertauschen. Daraus folgt

$$\begin{aligned} \partial_x(\partial_y f) &= \lim_{\Delta y \rightarrow 0} \frac{1}{\Delta y} \lim_{\Delta x \rightarrow 0} \frac{1}{\Delta x} \left( \{f(x + \Delta x, y + \Delta y) - f(x, y + \Delta y)\} - \{f(x + \Delta x, y) - f(x, y)\} \right) \\ &= \lim_{\Delta y \rightarrow 0} \frac{1}{\Delta y} \left( \partial_x f(x, y + \Delta y) - \partial_x f(x, y) \right) = \partial_y(\partial_x f). \end{aligned}$$

### 3.16 Gradient eines Skalarfeldes

Mit Hilfe der partiellen Ableitungen lassen sich die Änderungen  $\Delta\varphi$  der Funktion  $\varphi(x, y, z)$  bei beliebigen infinitesimalen Änderungen der Variablen  $\vec{r} = (x, y, z) \rightarrow \vec{r} + \Delta\vec{r} = (x + \Delta x, y + \Delta y, z + \Delta z)$  darstellen.

$$\begin{aligned} \Delta\varphi &= \varphi(x + \Delta x, y + \Delta y, z + \Delta z) - \varphi(x, y, z) = \\ &= \varphi(x + \Delta x, y + \Delta y, z + \Delta z) - \varphi(x, y + \Delta y, z + \Delta z) + \\ &+ \varphi(x, y + \Delta y, z + \Delta z) - \varphi(x, y, z + \Delta z) + \\ &+ \varphi(x, y, z + \Delta z) - \varphi(x, y, z) = \\ &= \frac{\partial\varphi}{\partial x}(x, y + \Delta y, z + \Delta z) \Delta x + \frac{\partial\varphi}{\partial y}(x, y, z + \Delta z) \Delta y + \frac{\partial\varphi}{\partial z}(x, y, z) \Delta z + O(\Delta\vec{r}^2). \end{aligned}$$

Da  $\frac{\partial\varphi}{\partial x}$  und  $\frac{\partial\varphi}{\partial y}$  stetig sind, verschwinden ihre Änderungen bezüglich  $\Delta y$  und  $\Delta z$  (typischerweise proportional zu  $\Delta y$  und  $\Delta z$  oder schneller). Hier steht  $O(\Delta\vec{r}^2)$  für Terme, die proportional zu Produkten zweiter Ordnung der Form  $\Delta x_i \Delta x_j$  mit  $x_i \equiv x, y, z$  sind. Im Limes  $\Delta\vec{r} \rightarrow \vec{0}$  gilt also

$$\begin{aligned} \Delta\varphi &= \varphi(x + \Delta x, y + \Delta y, z + \Delta z) - \varphi(x, y, z) = \\ &= \frac{\partial\varphi}{\partial x} \Delta x + \frac{\partial\varphi}{\partial y} \Delta y + \frac{\partial\varphi}{\partial z} \Delta z + O(\Delta\vec{r}^2). \end{aligned}$$

oder in Differentialform

$$d\varphi = \frac{\partial\varphi}{\partial x} dx + \frac{\partial\varphi}{\partial y} dy + \frac{\partial\varphi}{\partial z} dz.$$

Dies lässt sich als Skalarprodukt

$$d\varphi = \left( \frac{\partial\varphi}{\partial x}, \frac{\partial\varphi}{\partial y}, \frac{\partial\varphi}{\partial z} \right) \cdot d\vec{r}$$

zweier Vektoren schreiben, nämlich der Gradient des Skalarfeldes  $\varphi$  und des Verschiebungsvektors  $d\vec{r}$ .

Wir definieren den *Gradienten eines Skalarfeldes*  $\varphi$  als den Vektor, dessen Komponenten die partiellen Ableitungen von  $\varphi$  sind

$$\begin{aligned}\text{grad } \varphi &= \vec{\nabla} \varphi = \left( \frac{\partial \varphi}{\partial x}, \frac{\partial \varphi}{\partial y}, \frac{\partial \varphi}{\partial z} \right) = \\ &= \frac{\partial \varphi}{\partial x} \hat{x} + \frac{\partial \varphi}{\partial y} \hat{y} + \frac{\partial \varphi}{\partial z} \hat{z} \\ &= \sum_i \frac{\partial \varphi}{\partial x_i} \hat{e}_i.\end{aligned}$$

Um zu zeigen, dass  $\text{grad } \varphi$  tatsächlich ein Vektor ist, müssen wir sehen, wie seine Komponenten nach einer Änderung des Koordinatensystems transformieren. Dies lässt sich mit Hilfe der Kettenregel herleiten

$$\frac{\partial \varphi}{\partial x'_i} = \sum_j \frac{\partial \varphi}{\partial x_j} \frac{\partial x_j}{\partial x'_i}.$$

Für die Koordinatentransformation gilt

$$x'_j = \sum_i \hat{e}'_i \cdot \hat{e}_j x_i \quad \text{und} \quad x_j = \sum_k \hat{e}_j \cdot \hat{e}'_k x'_k$$

und damit

$$\frac{\partial x_j}{\partial x'_i} = \hat{e}_j \cdot \hat{e}'_i.$$

Es folgt

$$\frac{\partial \varphi}{\partial x'_i} = \sum_j \hat{e}_j \cdot \hat{e}'_i \frac{\partial \varphi}{\partial x_j},$$

was genau der Transformationsregel für die Koordinaten eines Vektors entspricht.

Der vektorielle Charakter des Gradienten ist konsistent mit der Tatsache, dass  $\varphi$  ein Skalarfeld ist. Die Änderungen  $d\varphi = \vec{\nabla} \varphi \cdot d\vec{r}$  sind invariant bezüglich Transformationen des Koordinatensystems.

Folgende Eigenschaften lassen sich sofort aus der Definition von  $\text{grad } \varphi$  herleiten:

$$\vec{\nabla}(\varphi + \psi) = \vec{\nabla} \varphi + \vec{\nabla} \psi$$

und

$$\vec{\nabla}(\varphi \psi) = \varphi \vec{\nabla} \psi + \psi \vec{\nabla} \varphi.$$

Beispiele:

i)

$$\begin{aligned}\varphi &= x^3 - 2xz + y^2 - 2z^2 + 1 \\ \vec{\nabla} \varphi &= (3x^2 - 2z, 2y, -2x - 4z).\end{aligned}$$

ii)

$$\begin{aligned}\varphi &= \sqrt{(x - x_0)^2 + (y - y_0)^2 + (z - z_0)^2} = |\vec{r} - \vec{r}_0| \\ \vec{\nabla} \varphi &= \left( \frac{x - x_0}{|\vec{r} - \vec{r}_0|}, \frac{y - y_0}{|\vec{r} - \vec{r}_0|}, \frac{z - z_0}{|\vec{r} - \vec{r}_0|} \right) = \frac{\vec{r} - \vec{r}_0}{|\vec{r} - \vec{r}_0|}.\end{aligned}$$

### 3.17 Geometrische Eigenschaften des Gradienten

Wenn man die Verschiebungen  $d\vec{r}$  des Ortsvektors  $\vec{r}$  in alle möglichen Richtungen bei konstantem Betrag  $|d\vec{r}| = dr$  betrachtet, gibt es bei jedem Punkt  $\vec{r}$  im Raum eine Richtung, in der die Änderung von  $\varphi(\vec{r})$  maximal ist. Da der Betrag von  $\vec{\nabla}\varphi$  unabhängig von  $d\vec{r}$  ist, ist  $d\varphi = \vec{\nabla}\varphi \cdot d\vec{r} = |\vec{\nabla}\varphi| |d\vec{r}| \cos\theta$  maximal, wenn  $d\vec{r}$  parallel zu  $\vec{\nabla}\varphi$  ist ( $\theta = 0$ ). Es gilt also folgende wichtige Eigenschaft:

Die Richtung des Gradienten  $\vec{\nabla}\varphi$  eines Skalarfeldes oder Funktion an einem Punkt  $\vec{r}$  ist die, in der  $\varphi(\vec{r})$  am stärksten steigt. Der Betrag des Gradienten ist gleich der Differentialquotient

$$|\vec{\nabla}\varphi| = \max \frac{|d\varphi|}{|d\vec{r}|},$$

wenn die Verschiebung  $d\vec{r}$  in die Richtung maximaler Änderung stattfindet.

Der Gradient definiert auch die Konstanzflächen oder Niveauflächen  $\varphi(\vec{r}) = K$ , auf denen  $\varphi(\vec{r})$  konstant ist. Die Bedingung

$$d\varphi = \vec{\nabla}\varphi \cdot d\vec{r} = 0$$

besagt, dass  $\vec{\nabla}\varphi$  immer senkrecht zu den Flächen ist, bei denen  $\varphi(\vec{r}) = \text{Konstante}$  ist.

Die Gradientenlinien sind stets parallel zum Gradienten, d.h., für eine Verschiebung  $d\vec{r}$  entlang einer Gradientenlinie gilt

$$(dx, dy, dz) = \lambda \vec{\nabla}\varphi = \lambda \left( \frac{\partial\varphi}{\partial x}, \frac{\partial\varphi}{\partial y}, \frac{\partial\varphi}{\partial z} \right),$$

wobei  $\lambda$  eine Konstante ist. Daraus ergeben sich folgende Bedingungen, die die Gleichungen für die Gradientenlinien definieren:

$$\frac{dx}{\frac{\partial\varphi}{\partial x}} = \frac{dy}{\frac{\partial\varphi}{\partial y}} = \frac{dz}{\frac{\partial\varphi}{\partial z}}.$$

Beispiel:

Wir betrachten die Funktion  $\varphi = x^2 - y^2$  auf der Ebene ( $z = 0$ ). Die Linien konstanter  $\varphi$  sind Hyperbola ( $x^2 - y^2 = K$ ). Der Gradient lautet  $\vec{\nabla}\varphi = (2x, -2y)$ . Die Linien des Gradienten ergeben sich aus

$$\frac{dx}{2x} = \frac{-dy}{2y} \Leftrightarrow y dx + x dy = 0$$

$$\Leftrightarrow d(xy) = 0 \Leftrightarrow xy = \text{Konstante}.$$

Sie sind also auch Hyperbola.

### 3.18 Nablaoperator

Eine sehr nützliche Schreibweise besteht darin,  $\vec{\nabla}$  als Vektoroperator, den sogenannten *Nablaoperator*, zu betrachten:

$$\vec{\nabla} = \hat{x} \frac{\partial}{\partial x} + \hat{y} \frac{\partial}{\partial y} + \hat{z} \frac{\partial}{\partial z} = \left( \frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z} \right).$$

Der Nablaoperator ist ein Vektor, da seine Komponenten  $\frac{\partial}{\partial x_i}$  wie ein Vektor transformieren. Die Anwendung von  $\vec{\nabla}$  auf eine Skalarfunktion  $\varphi$  ergibt den Vektor  $\vec{\nabla}\varphi$ . Der Nablaoperator kann auch auf Vektoren angewandt werden, wie wir es in den folgenden Abschnitten diskutieren werden.

### 3.19 Richtungsableitung

Man definiert die *Richtungsableitung* der Funktion  $\varphi(\vec{r})$  nach der Richtung des Einheitsvektors  $\hat{n}$  durch folgenden Limes:

$$\frac{d\varphi}{d\hat{n}} = \lim_{s \rightarrow 0} \frac{\varphi(\vec{r} + \hat{n}s) - \varphi(\vec{r})}{s}.$$

Da  $d\varphi = \varphi(\vec{r} + \hat{n}s) - \varphi(\vec{r}) = \vec{\nabla}\varphi \cdot \hat{n}s$  ist, gilt

$$\frac{d\varphi}{d\hat{n}} = \hat{n} \cdot \vec{\nabla}\varphi.$$

Die Richtungsableitung ist also der Differentialquotient  $\frac{d\varphi}{ds}$  bei einer Verschiebung  $d\vec{s} = \hat{n} ds$ .

Es ist auch üblich, dass der Vektor  $\vec{v} = v\hat{v}$ , entlang dem die Ableitung berechnet wird, nicht auf 1 normiert ist. In diesem Fall haben wir

$$\lim_{s \rightarrow 0} \frac{\varphi(\vec{r} + \vec{v}s) - \varphi(\vec{r})}{s} = \vec{v} \cdot \vec{\nabla}\varphi = v\hat{v} \cdot \vec{\nabla}\varphi = v \frac{d\varphi}{d\hat{v}}.$$

Wie man erwarten sollte, ist der Differentialquotient  $\frac{d\varphi}{ds}$  bei einer Verschiebung  $d\vec{s} = \vec{v} ds$  umso größer, je größer die Norm von  $\vec{v}$  ist.

Beispiel:

Sei  $\varphi = x^2 + y^2 - z^3$ . Die Richtungsableitung nach  $\vec{v} = (1, -2, 3)$  ist

$$\begin{aligned} \lim_{s \rightarrow 0} \frac{\varphi(\vec{r} + s\vec{v}) - \varphi(\vec{r})}{s} &= \vec{v} \cdot \vec{\nabla}\varphi = \\ &= v_x \frac{\partial\varphi}{\partial x} + v_y \frac{\partial\varphi}{\partial y} + v_z \frac{\partial\varphi}{\partial z} = (1, -2, 3) \cdot (2x, 2y, -3z^2) \\ &= 2x - 4y - 9z^2. \end{aligned}$$

Man kann auch die Ableitung von  $\varphi$  in die Richtung eines Vektorfeldes  $\vec{A}(\vec{r})$  betrachten, das von der Position abhängt. Zum Beispiel ist für  $\vec{A}(\vec{r}) = (x, -y, 3z)$

$$\begin{aligned}\vec{A} \cdot \vec{\nabla} \varphi &= (x, -y, 3z) \cdot (2x, 2y, -3z^2) \\ &= 2x^2 - 2y^2 - 9z^3.\end{aligned}$$

### 3.20 Totale Ableitung nach der Zeit

Wir betrachten ein Skalarfeld, das explizit von  $t$  abhängt  $\varphi = \varphi(x, y, z, t)$  und eine Raumkurve oder Trajektorie als Funktion der Zeit  $\vec{r} = \vec{r}(t)$ . Die Ableitung der Funktion  $\varphi(x(t), y(t), z(t), t)$  nach der Zeit ergibt sich aus der Kettenregel:

$$\frac{d\varphi}{dt} = \frac{\partial\varphi}{\partial x} \cdot \frac{dx}{dt} + \frac{\partial\varphi}{\partial y} \cdot \frac{dy}{dt} + \frac{\partial\varphi}{\partial z} \cdot \frac{dz}{dt} + \frac{\partial\varphi}{\partial t}.$$

Führt man die Geschwindigkeit  $\vec{v} = \dot{\vec{r}}(t)$  entlang der Trajektorie ein, dann ergibt sich

$$\frac{d\varphi}{dt} = \vec{v} \cdot \vec{\nabla} \varphi + \frac{\partial\varphi}{\partial t} = \dot{\vec{r}} \cdot \vec{\nabla} \varphi + \frac{\partial\varphi}{\partial t}.$$

### 3.21 Kraftfeld aus einem Potential – Energieerhaltung

Eine besonders wichtige Skalarfunktion ist die potentielle Energie  $V(\vec{r})$ , aus dem sich das konservative Kraftfeld eines klassischen Systems

$$\vec{F}(\vec{r}) = -\vec{\nabla} V(\vec{r})$$

herleiten lässt. Zum Beispiel haben wir bei der Gravitationskraft

$$V(\vec{r}) = \frac{-k}{r} = \frac{-k}{\sqrt{x^2 + y^2 + z^2}} \quad k > 0$$

und

$$\vec{F}(\vec{r}) = -\vec{\nabla} V(\vec{r}) = -k \frac{\vec{r}}{r^3}.$$

Ein weiteres bekanntes Beispiel ist die potentielle Energie im Gravitationskraftfeld an der Erdoberfläche

$$V(\vec{r}) = mgz \quad \Rightarrow \quad \vec{F} = -\vec{\nabla} V = (0, 0, -mg).$$

Eine interessante Anwendung der totalen Ableitung einer Skalargröße ist die Berechnung der Änderung der kinetischen Energie einer Masse  $m$  entlang ihrer Trajektorie. Damit ist die Trajektorie  $\vec{r}(t)$  gemeint, die diese Masse unter dem Einfluss eines gegebenen Kraftfeldes  $F(\vec{r})$  und bei gewissen Anfangsbedingungen durchläuft.

Die kinetische Energie  $T$  ist die Energie, die mit der Bewegung, also mit der Geschwindigkeit  $\vec{v} = \dot{\vec{r}}$  zusammenhängt. Da der Raum isotrop ist, kann sie nur eine Funktion von  $v^2 = \vec{v} \cdot \vec{v}$  sein. Sie ist definiert als

$$T = \frac{1}{2} m v^2.$$

Die Zeitänderung von  $T$  ist wie folgt gegeben:

$$\frac{dT}{dt} = \frac{1}{2} m 2 \vec{v} \cdot \frac{d\vec{v}}{dt} = \vec{v} \cdot \vec{F}, \quad (3.5)$$

wobei wir die Newton-Gleichung  $m \dot{\vec{v}} = \vec{F}$  eingesetzt haben. Dies gilt für jeden Punkt  $\vec{r} = (x(t), y(t), z(t))$  auf der Trajektorie, selbst wenn  $\vec{F} = \vec{F}(\vec{r}, t)$  eine Funktion der Position und der Zeit ist.

Die differentiale Änderung der Energie bei einer Zeitverschiebung  $dt$  ist also

$$dT = \vec{F} \cdot \vec{v} dt = \vec{F} \cdot d\vec{r}.$$

Mathematisch ist  $\vec{F} \cdot d\vec{r}$  das Differential der *Zirkulation* des Vektorfeldes  $\vec{F}(\vec{r})$  entlang der Kurve  $\vec{r}(t)$ . Wenn es sich wie hier um eine Kraft handelt, wird dies die *Arbeit* bzw. Differential der Arbeit der Kraft  $\vec{F}$  bei einer infinitesimalen Translation  $d\vec{r}$  genannt

$$dT = dW = \vec{F} \cdot d\vec{r}.$$

Wenn mehrere Kräfte  $\vec{F}_i$  ins Spiel kommen, ist natürlich  $\vec{F} = \sum \vec{F}_i$  die vektorielle Summe davon.

Gleichung (3.5) nimmt im Falle von Kraftfeldern  $\vec{F} = -\vec{\nabla}V$ , die aus einem Potential  $V(\vec{r})$  abgeleitet sind, eine besonders transparente Form an. In diesem Fall haben wir

$$\frac{dT}{dt} = -\vec{v} \cdot \vec{\nabla}V = -\frac{dV}{dt}.$$

Dies bedeutet, dass die gesamte Energie  $E = T + V$  eine *Konstante der Bewegung* bzw. eine *Erhaltungsgröße* ist, da  $dE/dt = 0$ . Damit erklärt sich auch, warum die Kräfte, die aus dem Gradienten eines Potentials hergeleitet sind, konservative Kräfte heißen.

### 3.22 Vektorfelder: Richtungs- und totale Ableitungen

Ein Vektorfeld ist ein Vektor, dessen Komponenten eine Funktion der Koordinaten  $x, y, z$  sind. In anderen Worten wird jedem Punkt im Raum  $\vec{r} = (x, y, z)$  ein Vektor  $\vec{A}(\vec{r})$  zugewiesen

$$\begin{aligned} \vec{A} : \mathbb{R}^3 &\rightarrow \mathbb{R}^3 \\ \vec{r} &\rightarrow \vec{A}(\vec{r}) = (A_x(\vec{r}), A_y(\vec{r}), A_z(\vec{r})). \end{aligned}$$

Beispiele sind Kraftfelder  $\vec{F}(\vec{r})$  (z.B. die Gravitationskraft eines Sterns), Geschwindigkeitsfelder  $\vec{v}(\vec{r})$  oder Stromdichten  $\vec{j}(\vec{r}) = \varrho(\vec{r}) \vec{v}(\vec{r})$  in einer Flüssigkeit, und  $\vec{A}(\vec{r}) = \text{grad } \varphi$  mit  $\varphi$  Skalarfeld.

Wie im Falle der Skalarfelder spielen die lokalen Änderungen von Vektorfeldern eine wichtige Rolle in der Physik. Die Richtungsableitung eines Vektorfelds  $\vec{A}(\vec{r})$  gibt ein Maß für die lokale Änderung von  $\vec{A}(\vec{r})$  in die vom Einheitsvektor  $\hat{n}$  vorgegebene Richtung

$$\frac{d\vec{A}}{d\hat{n}} = \lim_{s \rightarrow 0} \frac{\vec{A}(\vec{r} + \hat{n} s) - \vec{A}(\vec{r})}{s} = \left( \frac{\partial A_1}{\partial \hat{n}}, \frac{\partial A_2}{\partial \hat{n}}, \frac{\partial A_3}{\partial \hat{n}} \right).$$

Dies ist ja als Vektorgleichung, d.h., komponentenweise zu verstehen. Wenn  $\vec{A} = (A_1, A_2, A_3)$  und  $\hat{n} = (n_1, n_2, n_3)$  ist, dann gilt

$$dA_i = \frac{\partial A_i}{\partial x} dx + \frac{\partial A_i}{\partial y} dy + \frac{\partial A_i}{\partial z} dz,$$

wobei  $d\vec{r} = (dx, dy, dz) = (n_1, n_2, n_3) ds = \hat{n} ds$ . Wir können also schreiben

$$dA_i = \left( n_1 \frac{\partial A_i}{\partial x} + n_2 \frac{\partial A_i}{\partial y} + n_3 \frac{\partial A_i}{\partial z} \right) ds$$

und damit

$$\frac{d\vec{A}}{d\hat{n}} = n_1 \frac{\partial \vec{A}}{\partial x} + n_2 \frac{\partial \vec{A}}{\partial y} + n_3 \frac{\partial \vec{A}}{\partial z}$$

oder in einer kompakteren Form

$$\frac{d\vec{A}}{d\hat{n}} = (\hat{n} \cdot \vec{\nabla}) \vec{A},$$

wobei das Skalarprodukt

$$\hat{n} \cdot \vec{\nabla} = \sum_i n_i \frac{\partial}{\partial x_i} = n_1 \frac{\partial}{\partial x} + n_2 \frac{\partial}{\partial y} + n_3 \frac{\partial}{\partial z}$$

ein Skalaroperator ist. Damit lässt sich die totale Ableitung eines Vektorfeldes nach der Zeit berechnen. Gegeben ist  $\vec{A} = \vec{A}(x, y, z, t)$  und eine Trajektorie  $\vec{r}(t) = (x(t), y(t), z(t))$ , dann ist

$$\frac{dA_i}{dt} = \frac{\partial A_i}{\partial t} + \frac{\partial A_i}{\partial x} \dot{x} + \frac{\partial A_i}{\partial y} \dot{y} + \frac{\partial A_i}{\partial z} \dot{z},$$

und damit

$$\frac{d\vec{A}}{dt} = \frac{\partial \vec{A}}{\partial t} + (\vec{v} \cdot \vec{\nabla}) \vec{A},$$

wobei  $\vec{v} = (\dot{x}, \dot{y}, \dot{z})$ . Der Skalaroperator  $\vec{v} \cdot \vec{\nabla}$  wirkt auf jede Komponente von  $\vec{A}$ .

### 3.23 Divergenz

Die Divergenz  $\text{div } \vec{A} = \vec{\nabla} \cdot \vec{A}$  eines Vektorfeldes  $\vec{A} = \vec{A}(\vec{r}) = (A_1, A_2, A_3)$  ist das Resultat der Anwendung des Nablaoperators

$$\vec{\nabla} = \left( \frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z} \right)$$

auf den Vektor  $\vec{A}(\vec{r})$ , und zwar in der Form eines Skalarprodukts

$$\text{div } \vec{A} = \vec{\nabla} \cdot \vec{A} = \frac{\partial A_1}{\partial x} + \frac{\partial A_2}{\partial y} + \frac{\partial A_3}{\partial z}.$$

Die Divergenz eines Vektorfeldes ist ein Skalarfeld, weil sie das Skalarprodukt von zwei Vektoren ist.

Beispiele:

i) Für  $\vec{A}(\vec{r}) = \vec{r}$  ist

$$\vec{\nabla} \cdot \vec{A} = \frac{\partial x}{\partial x} + \frac{\partial y}{\partial y} + \frac{\partial z}{\partial z} = 3.$$

ii) Für  $\vec{A} = (3x^2 - y, 2xz, (z - y))$  ist  $\text{div } \vec{A} = 6x + 1$ .

iii) Divergenz eines zentralen Kraftfeldes  $\vec{F}(\vec{r}) = \vec{r} f(r)$  :

$$\begin{aligned}\vec{\nabla} \cdot \vec{F} &= \frac{\partial}{\partial x} (x f(r)) + \frac{\partial}{\partial y} (y f(r)) + \frac{\partial}{\partial z} (z f(r)) \\ &= 3 f(r) + \frac{df}{dr} \left( x \frac{\partial r}{\partial x} + y \frac{\partial r}{\partial y} + z \frac{\partial r}{\partial z} \right) \\ &= 3 f(r) + \left( \frac{x^2}{r} + \frac{y^2}{r} + \frac{z^2}{r} \right) \frac{df}{dr} \\ &= 3 f(r) + r \frac{df}{dr}.\end{aligned}$$

Es wird dem Leser überlassen, herauszufinden, wie  $f(r)$  auszusehen hat, damit  $\vec{\nabla} \cdot \vec{F} = 0 \forall \vec{r} \neq 0$  ist.

Folgende Eigenschaften der Divergenz lassen sich leicht zeigen:

i) Der Nablaoperator ist ein linearer Operator, und von daher gilt

$$\vec{\nabla} \cdot (\alpha \vec{A} + \beta \vec{B}) = \alpha \vec{\nabla} \cdot \vec{A} + \beta \vec{\nabla} \cdot \vec{B}$$

für  $\alpha, \beta \in \mathbb{R}$  konstant und  $\vec{A}, \vec{B} : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ .

ii) Wenn  $\varphi : \mathbb{R}^3 \rightarrow \mathbb{R}$  ein Skalarfeld ist und  $\vec{A}$  ein Vektorfeld, gilt

$$\vec{\nabla} \cdot (\varphi \vec{A}) = \vec{\nabla} \varphi \cdot \vec{A} + \varphi \vec{\nabla} \cdot \vec{A}$$

oder mit der äquivalenten Notation

$$\text{div}(\varphi \vec{A}) = \text{grad}(\varphi) \cdot \vec{A} + \varphi \text{div}(\vec{A}).$$

Dies folgt aus

$$\sum_i \frac{\partial}{\partial x_i} (\varphi A_i) = \sum_i \frac{\partial \varphi}{\partial x_i} A_i + \varphi \sum_i \frac{\partial A_i}{\partial x_i}.$$

### 3.24 Physikalische Interpretation der Divergenz

Um die Divergenz zu interpretieren, brauchen wir den Begriff des *Flusses eines Vektorfelds* durch eine Oberfläche. Der Fluss eines Vektorfeldes  $\vec{A}$  durch eine infinitesimale Fläche  $dS$ , die senkrecht zum Einheitsvektor  $\vec{n}$  am Punkt  $\vec{r}$  liegt, ist definiert als der Skalar

$$d\Phi = \vec{A}(\vec{r}) \cdot \hat{n} dS = \vec{A}(\vec{r}) \cdot \vec{dS}.$$

Das Vorzeichen von  $d\Phi$  hängt natürlich von der Richtung von  $\hat{n}$  ab. Bei geschlossenen Oberflächen (d.h., Oberflächen, die ein Volumen einschließen) zeigt  $\hat{n}$  konventionsgemäß nach außen. Bei offenen Flächen ist die Wahl arbiträr, sie hängt meistens mit der ausgewählten Richtung der Zirkulation entlang der Kurve an der Grenze der Fläche zusammen.

Um die physikalische Bedeutung von  $d\Phi$  zu erläutern, betrachten wir also eine Flüssigkeit mit einer orts- und zeitabhängigen Dichte  $\rho(x, y, z, t)$  und einem Geschwindigkeitsfeld  $\vec{v}(x, y, z)$ . Die Stromdichte  $\vec{j}(x, y, z, t)$  ist durch das Produkt

$$\vec{j} = \rho \vec{v}$$

definiert. Der Fluss von  $\vec{j}$  durch eine infinitesimale orientierte Fläche  $\vec{dS} = \hat{n} dS$  ist durch

$$d\Phi = \vec{j} \cdot \vec{dS} = \vec{j} \cdot \hat{n} dS$$

gegeben. Der Fluss der Stromdichte  $\vec{j}$  entspricht der Menge der Materie, die durch  $\vec{dS}$  pro Einheit Zeit durchläuft. Es gilt nämlich

$$d\Phi = \vec{j} \cdot \hat{n} dS = \rho (\vec{v} \cdot \hat{n}) dS$$

und

$$d\Phi \Delta t = \rho (\vec{v} \cdot \hat{n}) \Delta t dS = \rho dl dS,$$

wobei  $dl dS$  das Volumen der Flüssigkeit ist, das durch  $\vec{dS}$  in der Zeit  $\Delta t$  durchgeht.

Wir betrachten nun ein infinitesimales Volumen  $dV = dx dy dz$  um einen beliebigen Punkt  $x, y, z$  und berechnen den gesamten Stromfluss oder Fluss von Materie, der aus diesem Volumen herausgeht, i.e., positiv, wenn der Fluss nach außen ist. Die Fläche  $ABCD$  hat  $\hat{n} = (-1, 0, 0)$ , und die Fläche  $EFGH$  hat  $\hat{n} = (1, 0, 0)$ . Deshalb ist der gesamte Fluss durch diese beiden Flächen

$$\begin{aligned} d\Phi(EFGH) + d\Phi(ABCD) &= [j_x(x + dx, y, z) - j_x(x, y, z)] dy dz \\ &= \frac{\partial j_x}{\partial x} dx dy dz. \end{aligned}$$

Dieselbe Überlegung für die anderen Flächen ergibt

$$d\Phi(AECCG) + d\Phi(BFDH) = \frac{\partial j_y}{\partial y} dx dy dz$$

und

$$d\Phi(ABEF) + d\Phi(CDGH) = \frac{\partial j_z}{\partial z} dx dy dz.$$

Wenn die Beiträge aller Flächen addiert werden, erhält man

$$\begin{aligned} d\Phi &= \left( \frac{\partial j_x}{\partial x} + \frac{\partial j_y}{\partial y} + \frac{\partial j_z}{\partial z} \right) dV \\ &= \vec{\nabla} \cdot \vec{j} dV. \end{aligned}$$

Der gesamte Fluss ist proportional zum Volumen des Würfels, und der Proportionalitätsfaktor ist nichts anderes als die Divergenz des Vektorfelds an dem Punkt. Anders ausgedrückt ist die Divergenz  $\vec{\nabla} \cdot \vec{j}(\vec{r})$  am Punkt  $\vec{r}$  der Differentialquotient zwischen dem Fluss  $d\Phi$  aus einem infinitesimalen Volumen  $\Delta V$  um den Punkt  $\vec{r}$  und diesem Volumen  $\Delta V$ .

### 3.25 Kontinuitätsgleichung

Der Fluss der Stromdichte  $\vec{j} = \rho \vec{v}$  entspricht der Menge der Materie, die durch eine Oberfläche pro Einheit Zeit hindurchgeht. Im Falle eines infinitesimalen Volumens  $dV$  ist  $d\Phi = \vec{\nabla} \cdot \vec{j} dV$  gleich der Menge der Flüssigkeit, die aus dem infinitesimalen Volumen  $dV$  um den Punkt  $\vec{r}$  pro Einheit Zeit herausgeht. Für ein Zeitintervall  $\Delta t$  ist die Menge der aus  $dV$  herausgeflossenen Materie gleich  $d\Phi \Delta t = \vec{\nabla} \cdot \vec{j} dV \Delta t$ . Da die Materie erhalten bleibt, muss dies mit der Änderung der im Volumen eingeschlossenen Materie übereinstimmen. Daraus folgt

$$\Delta \rho dV = -\vec{\nabla} \cdot \vec{j} dV \Delta t,$$

wobei  $\Delta \rho$  die Änderung der Dichte am Punkt  $\vec{r}$  ist. Das Volumenelement selbst ist ja zeitunabhängig. Achten Sie auf das negative Vorzeichen, welches daher kommt, dass die Menge der Materie abnimmt, wenn der Fluss positiv ist.

Die Änderung der Dichte in der Zeit  $\Delta t$  ist also

$$\Delta \rho = -\vec{\nabla} \cdot \vec{j} \Delta t.$$

Daraus folgt die sogenannte *Kontinuitätsgleichung*

$$\frac{\partial \rho}{\partial t} + \vec{\nabla} \cdot \vec{j} = 0,$$

die den Erhalt der Materie darstellt.

### 3.26 Rotation

Der Nablaoperator  $\vec{\nabla} = \left( \frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z} \right)$  kann auch in der Form eines Vektorproduktes auf Vektorfelder angewandt werden. Damit definiert man die *Rotation* eines Vektorfeldes  $\vec{A}(\vec{r})$  als den Pseudovektor

$$\begin{aligned} \text{rot}(\vec{A}) &= \vec{\nabla} \times \vec{A} = \hat{x} \left( \frac{\partial A_z}{\partial y} - \frac{\partial A_y}{\partial z} \right) + \hat{y} \left( \frac{\partial A_x}{\partial z} - \frac{\partial A_z}{\partial x} \right) + \hat{z} \left( \frac{\partial A_y}{\partial x} - \frac{\partial A_x}{\partial y} \right) \\ &= \begin{vmatrix} \hat{x} & \hat{y} & \hat{z} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ A_x & A_y & A_z \end{vmatrix} = \sum_{ijk} \varepsilon_{ijk} \hat{e}_i \frac{\partial A_k}{\partial x_j}. \end{aligned}$$

Sollte  $\vec{A}$  ein Pseudovektor sein, dann ist  $\vec{\nabla} \times \vec{A}$  ein Vektorfeld. Insbesondere ist die zweifache Anwendung des Rotoroperators an einem Vektor, d.h.,  $\vec{\nabla} \times \vec{\nabla} \times \vec{A} = \text{rot}(\text{rot}(\vec{A}))$ , wieder ein Vektor.

Für jedes differenzierbare Vektorfeld bildet  $\vec{\nabla} \times \vec{A}$  ein weiteres Pseudovektorfeld, das Rotationsfeld von  $\vec{A}$  genannt wird. Die Feldlinien, die entlang  $\vec{\nabla} \times \vec{A}$  laufen, werden Wirbellinien genannt. Sie sind durch folgende Gleichungen gegeben:

$$\frac{dx}{(\vec{\nabla} \times \vec{A})_x} = \frac{dy}{(\vec{\nabla} \times \vec{A})_y} = \frac{dz}{(\vec{\nabla} \times \vec{A})_z}$$

$$\frac{\frac{dx}{\partial A_z} - \frac{\partial A_y}{\partial z}}{\frac{\partial A_z}{\partial y} - \frac{\partial A_y}{\partial z}} = \frac{\frac{dy}{\partial A_x} - \frac{\partial A_z}{\partial x}}{\frac{\partial A_x}{\partial z} - \frac{\partial A_z}{\partial x}} = \frac{\frac{dz}{\partial A_y} - \frac{\partial A_x}{\partial y}}{\frac{\partial A_y}{\partial x} - \frac{\partial A_x}{\partial y}}.$$

Beispiele:

i) Sei  $\vec{A} = (x^2, -x + y, z^2 + \sin x)$ .

$$\vec{\nabla} \times \vec{A} = \begin{vmatrix} \hat{x} & \hat{y} & \hat{z} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ x^2 & -x + y & z^2 + \sin x \end{vmatrix} = (0, -\cos x, -1).$$

ii)  $\vec{A} = f(r) \vec{r}$  mit  $\vec{r} = (x, y, z)$  und  $r = |\vec{r}|$ .

$$\begin{aligned} (\vec{\nabla} \times \vec{A})_x &= \frac{\partial A_z}{\partial y} - \frac{\partial A_y}{\partial z} = f' \frac{\partial r}{\partial y} z + f(r) \frac{\partial z}{\partial y} - \left( f' \frac{\partial r}{\partial z} y + f(r) \frac{\partial y}{\partial z} \right) \\ &= f' \frac{(y z - z y)}{r} = 0. \end{aligned}$$

Die anderen Komponenten sind ebenfalls null, wie man es durch zyklische Permutation sehen kann. Damit ist  $\vec{\nabla} \times \vec{A} = 0$ . Ein sphärisch symmetrisches Radialfeld hat also eine überall verschwindende Rotation.

Zum Vergleich ist es interessant, die Divergenz dieses Feldes zu berechnen:

$$\vec{\nabla} \cdot \vec{A} = 3 f(r) + x f'(r) \frac{x}{r} + y f'(r) \frac{y}{r} + z f'(r) \frac{z}{r} = 3 f(r) + r f'(r).$$

Man stellt fest, dass  $\vec{\nabla} \cdot \vec{A}$  im Allgemeinen nicht null ist, mit einer einzigen Ausnahme: Es gilt  $\vec{\nabla} \cdot \vec{A} = 0 \quad \forall r \neq 0$  wenn und nur wenn  $3 f(r) + r f'(r) = 0 \Leftrightarrow f(r) = \frac{\alpha}{r^3}$ , wobei  $\alpha$  eine Konstante ist. In diesem Fall ist  $\vec{A} = \alpha \frac{\vec{r}}{r^3}$ , wie beim Gravitationskraftfeld oder elektrischen Coulomb-Feld. Man sollte allerdings merken, dass selbst in diesem Fall, wo  $\vec{\nabla} \cdot \vec{A} = 0 \quad \forall r \neq 0$  die Divergenz am Ursprung ungleich null bleibt. Sie divergiert sogar, da der Fluss  $\Phi$  von  $\vec{A}$  durch eine Kugel mit Radius  $r$  unabhängig von  $r$  ist:

$$\Phi = \int_s \vec{A} \cdot d\vec{s} = r^2 \int_0^{2\pi} d\varphi \int_0^\pi \sin \theta d\theta \alpha \frac{\vec{r}}{r^3} \cdot \frac{\vec{r}}{r} = 4\pi \alpha.$$

Hier haben wir benutzt, dass das Flächendifferential auf einer Kugel mit Radius  $r$  durch  $ds = r \sin \theta d\varphi \times r d\theta$  gegeben ist.<sup>2</sup> Das bedeutet, dass das Integral des

<sup>2</sup> Näheres dazu wird im Abschnitt 4.8 gegeben.

Skalarfeldes  $\vec{\nabla} \cdot \left( \alpha \frac{\vec{r}}{r^3} \right)$  in einer Sphäre um den Koordinatenursprung endlich ist, obwohl  $\vec{\nabla} \cdot \left( \alpha \frac{\vec{r}}{r^3} \right) = 0 \quad \forall r \neq 0$ . Skalarfelder dieser Art werden Delta-Funktionen genannt.

iii) Wir betrachten das Vektorfeld

$$\vec{A}(\vec{r}) = f(\varrho) (-y, x, 0),$$

wobei  $\varrho = \sqrt{x^2 + y^2}$  der Abstand zur  $z$ -Achse ist.  $\vec{A}$  ist unabhängig von  $z$  und

$$\begin{aligned} \vec{\nabla} \times \vec{A} &= \begin{vmatrix} \hat{x} & \hat{y} & \hat{z} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ -f(\varrho)y & f(\varrho)x & 0 \end{vmatrix} \\ &= \left( 0, 0, f'(\varrho) \frac{x^2}{\varrho} + f(\varrho) + f'(\varrho) \frac{y^2}{\varrho} + f(\varrho) \right) = (0, 0, 2f(\varrho) + \varrho f'(\varrho)). \end{aligned}$$

Die Wirbellinien von  $\vec{A}$  sind also parallel zur  $z$ -Achse und  $\vec{\nabla} \times \vec{A}$  ist im Allgemeinen nicht null. Dies ist sicherlich das einfachste Beispiel eines Rotationsfeldes.

Die Bedingung, damit  $\vec{\nabla} \times \vec{A} = 0$  für  $\varrho \neq 0$  lautet  $\varrho f'(\varrho) + 2f(\varrho) = 0 \Leftrightarrow f(\varrho) = \frac{\alpha}{\varrho^2} \Leftrightarrow \vec{A}(\vec{r}) = \frac{\alpha}{\varrho^2} (-y, x, 0)$ . Allerdings bleibt selbst in diesem Fall  $\vec{\nabla} \times \vec{A}$  nicht nur ungleich null am Ursprung, sondern sogar divergent.

Zum Vergleich ist es interessant, die Divergenz von  $\vec{A} = f(\varrho) (-y, x, 0)$  zu berechnen. Man erhält

$$\vec{\nabla} \cdot \vec{A} = -y \frac{\partial f(\varrho)}{\partial x} + x \frac{\partial f(\varrho)}{\partial y} = -y f'(\varrho) \frac{x}{\varrho} + x f'(\varrho) \frac{y}{\varrho} = \frac{f'(\varrho)}{\varrho} (xy - yx) = 0.$$

Dies bedeutet, dass der Fluss von  $\vec{A}$  durch die Oberfläche jedes infinitesimalen oder endlichen Volumens verschwindet. Dies lässt sich durch die azimuthale Richtung der Feldlinien von  $\vec{A}$  und die Unabhängigkeit des Betrags von  $\vec{A}$  bezüglich  $\varrho$  geometrisch leicht erklären.

### 3.27 Physikalische Bedeutung der Rotation

Um die Rotation zu interpretieren, brauchen wir den Begriff der *Zirkulation eines Vektorfelds* entlang einer Kurve  $\mathcal{C}$ , dem wir eigentlich bereits im Abschnitt 3.21 kurz begegnet sind. Die Zirkulation eines Vektorfeldes  $\vec{A}$  entlang einer infinitesimalen Verschiebung  $d\vec{r}$  am Punkt  $\vec{r}$  ist definiert als der Skalar

$$d\Gamma = \vec{A}(\vec{r}) \cdot d\vec{r}.$$

Bei Kurven endlicher Länge ist die Zirkulation durch das Integral von  $d\Gamma$  entlang der Kurve gegeben:

$$\Gamma = \int_{\mathcal{C}} d\Gamma = \int_{\mathcal{C}} \vec{A}(\vec{r}) \cdot d\vec{r}.$$

Wenn  $\vec{A}$  ein Kraftfeld  $\vec{F}$  ist, entspricht  $d\Gamma = \vec{F} \cdot d\vec{r}$  dem Differential der Arbeit des Kraftfeldes  $\vec{F}$  bei einer infinitesimalen Verschiebung  $d\vec{r}$  und  $\Gamma$  der gesamten Arbeit der Kraft  $\vec{F}$  entlang der Kurve oder Trajektorie  $\mathcal{C}$ .

Liegt die Kurve  $\mathcal{C}$  in einer Parameterdarstellung  $\vec{r} = \vec{r}(t) = (x(t), y(t), z(t))$  mit  $\vec{r}(t) : [t_1, t_2] \rightarrow \mathbb{R}^3$  vor, dann ist

$$d\vec{r} = \frac{d\vec{r}}{dt} dt$$

und

$$\Gamma = \int_{\mathcal{C}(\vec{r}_1 \rightarrow \vec{r}_2)} \vec{A}(\vec{r}) \cdot d\vec{r} = \int_{t_2}^{t_1} \vec{A}(\vec{r}) \cdot \frac{d\vec{r}}{dt} dt,$$

wobei  $\vec{r}_1 = \vec{r}(t_1)$  der Anfang und  $\vec{r}_2 = \vec{r}(t_2)$  das Ende der Kurve  $\mathcal{C}$  sind.

Um die physikalische Bedeutung der Rotation zu erläutern, betrachten wir eine infinitesimale geschlossene Kurve  $\mathcal{C} \equiv ABCDA$  um einen beliebigen Punkt  $\vec{r} = (x, y, z)$  und berechnen die Zirkulation des Vektorfeldes  $\vec{A}$  entlang  $\mathcal{C}$ . Der Einfachheit halber bildet  $\mathcal{C}$  ein Rechteck der Länge  $dx$  mal  $dy$  in der  $xy$ -Ebene, und zwar sind die Koordinaten der Ecken  $A = (x, y, z)$ ,  $B = (x + dx, y, z)$ ,  $C = (x + dx, y + dy, z)$  und  $D = (x, y + dy, z)$  (siehe Abb. ...). Die Wege  $AB$  und  $CD$  sind parallel bzw. antiparallel zu  $\hat{x}$ , während die Wege  $BC$  und  $DA$  parallel bzw. antiparallel zu  $\hat{y}$  sind. Die Zirkulation von  $\vec{A}$  auf den verschiedenen infinitesimalen Strecken lautet

$$\begin{aligned} d\Gamma_{AB} &= \vec{A}(x, y, z) \cdot d\vec{r} = A_x(x, y, z) dx \\ d\Gamma_{CD} &= \vec{A}(x, y + dy, z) \cdot d\vec{r} = -A_x(x, y + dy, z) dx \\ d\Gamma_{BC} &= \vec{A}(x + dx, y, z) \cdot d\vec{r} = A_y(x + dx, y, z) dy \\ d\Gamma_{DA} &= \vec{A}(x, y, z) \cdot d\vec{r} = -A_y(x, y, z) dy. \end{aligned}$$

Für die gesamte Zirkulation erhalten wir

$$\begin{aligned} d\Gamma_{ABCD} &= d\Gamma_{AB} + d\Gamma_{BC} + d\Gamma_{CD} + d\Gamma_{DA} \\ &= [A_y(x + dx, y, z) - A_y(x, y, z)] dy - [A_x(x, y + dy, z) - A_x(x, y, z)] dx \\ &= \frac{\partial A_y}{\partial x} dx dy - \frac{\partial A_x}{\partial y} dy dx \\ &= \left( \frac{\partial A_y}{\partial x} - \frac{\partial A_x}{\partial y} \right) dx dy \\ &= (\vec{\nabla} \times \vec{A})_z dx, dy. \end{aligned}$$

Analoge Resultate lassen sich für die anderen Komponenten der Rotation herleiten. Es gilt folgende wichtige Beziehung: Die Komponente  $(\vec{\nabla} \times \vec{A}) \cdot \hat{n}$  der Rotation eines Vektorfeldes  $\vec{A}$  am Punkt  $\vec{r} = (x, y, z)$  ist gleich der Zirkulation von  $\vec{A}$  entlang einer infinitesimalen geschlossenen Kurve  $\mathcal{C}$ , die senkrecht zu  $\hat{n}$  um den Punkt  $\vec{r}$  liegt, geteilt durch die Fläche, die die Kurve einschließt. Dies lässt sich in Vektorform wie folgt schreiben:

$$d\Gamma(\mathcal{C}) = \oint_{\mathcal{C}} \vec{A} \cdot d\vec{r} = (\vec{\nabla} \times \vec{A}) \cdot \hat{n} ds(\mathcal{C}) = (\vec{\nabla} \times \vec{A}) \cdot d\vec{s}(\mathcal{C}),$$

wobei  $d\vec{s}(\mathcal{C})$  das orientierte Flächenelement ist, das von der infinitesimalen geschlossenen Kurve  $\mathcal{C}$  eingeschlossen wird. Anders ausgedrückt, das Differential  $d\Gamma$  der Zirkulation von  $\vec{A}$  entlang  $\mathcal{C}$  ist gleich das Differential des Flusses von  $\vec{\nabla} \times \vec{A}$  durch  $d\vec{s}(\mathcal{C})$ . Diese wichtige Beziehung gilt für jede infinitesimal kleine, stetig differenzierbare geschlossene Kurve, unabhängig von ihrer Form oder Orientierung.

Es ist instruktiv, diese Beziehung nochmals herzuleiten, indem man die Zirkulation von  $\vec{A}$  um einen Kreis mit Radius  $r$  im Limes  $r \rightarrow 0$  berechnet. Der Einfachheit halber nehmen wir an, dass der Kreis am Koordinatenursprung in der  $xy$ -Ebene liegt. Die Parameterdarstellung der Kurve ist bekanntlich

$$\vec{r}(\varphi) = r (\cos \varphi, \sin \varphi, 0)$$

mit  $0 \leq \varphi \leq 2\pi$ , und daher

$$d\vec{r} = r (-\sin \varphi, \cos \varphi, 0) d\varphi.$$

Die Zirkulation  $\Gamma$  ist

$$\begin{aligned} \Gamma &= \int_{\odot} \vec{A} \cdot d\vec{r} = r \int_0^{2\pi} \vec{A} \cdot (-\sin \varphi, \cos \varphi, 0) d\varphi \\ &= r \int_0^{2\pi} (-A_x \sin \varphi + A_y \cos \varphi) d\varphi. \end{aligned}$$

Um  $\Gamma$  zu berechnen, brauchen wir die Komponenten von  $A_x$  und  $A_y$  von  $\vec{A}$  in der unmittelbaren Umgebung des Ursprungs. Diese können wir mit Hilfe einer Taylor-Entwicklung erster Ordnung bezüglich  $x$  und  $y$  gewinnen:

$$\begin{aligned} A_x(x, y, 0) &= A_x(\vec{0}) + x \partial_x A_x(\vec{0}) + y \partial_y A_x(\vec{0}) + O(r^2) \\ A_y(x, y, 0) &= A_y(\vec{0}) + x \partial_x A_y(\vec{0}) + y \partial_y A_y(\vec{0}) + O(r^2) \end{aligned}$$

Dies ist nichts anderes als die lineare Entwicklung einer Funktion  $\varphi(x, y, z)$  mittels des Gradienten  $\vec{\nabla}\varphi$  angewandt auf die Komponenten  $A_x$  und  $A_y$  eines Vektorfeldes. Wenn man die kompaktere Notation  $A_x^0 = A_x(\vec{0})$ ,  $\partial_x A_x^0 = \partial_x A_x(\vec{0})$  etc. einführt, erhält man

$$\begin{aligned} \Gamma &= r \int_0^{2\pi} (-A_x^0 \sin \varphi + A_y^0 \cos \varphi) d\varphi + \\ &\quad + r^2 \int_0^{2\pi} (-\partial_x A_x^0 \cos \varphi \sin \varphi - \partial_y A_x^0 \sin^2 \varphi + \partial_x A_y^0 \cos^2 \varphi + \partial_y A_y^0 \sin \varphi \cos \varphi) d\varphi \\ &= r^2 \left( \partial_x A_y^0 \int_0^{2\pi} \cos^2 \varphi d\varphi - \partial_y A_x^0 \int_0^{2\pi} \sin^2 \varphi d\varphi \right) \\ &= \pi r^2 (\partial_x A_y^0 - \partial_y A_x^0) = \pi r^2 (\vec{\nabla} \times \vec{A})_z. \end{aligned}$$

Dabei haben wir benutzt, dass  $\int_0^{2\pi} \sin \varphi d\varphi = \int_0^{2\pi} \cos \varphi d\varphi = 0$ ,  $\int_0^{2\pi} \cos \varphi \sin \varphi d\varphi = \int_0^{2\pi} \sin^2 \varphi d\varphi = 0$  und  $\int_0^{2\pi} \cos^2 \varphi d\varphi = \int_0^{2\pi} \sin^2 \varphi d\varphi = \frac{1}{2} \left( \int_0^{2\pi} (\cos^2 \varphi + \sin^2 \varphi) d\varphi \right) = \pi$ .

Wir können uns jetzt  $\vec{A}$  als ein Kraft- oder Geschwindigkeitsfeld vorstellen, und den Kreis  $\odot$  als eine infinitesimale Windmühle oder ein Rad, das sich in einer Ebene unter dem Einfluss von  $\vec{A}$  ohne Reibung drehen kann. Dort, wo die Rotation nicht null ist, wird das kleine Testrad sich drehen, und zwar in die eine oder andere Richtung in Abhängigkeit vom Vorzeichen von  $(\vec{\nabla} \times \vec{A}) \cdot \hat{n}$ .

### 3.28 Vektorielle Formeln für den Rotoroperator

Sei  $\varphi$  ein Skalarfeld und  $\vec{A}$  und  $\vec{B}$  Vektorfelder. Folgende Identitäten lassen sich direkt aus der Definition der Rotation herleiten. Es ist wahrscheinlich nützlich für den Leser, einige dieser Identitäten durch explizite Anwendung der Definition von  $\vec{\nabla} \times \vec{A}$  in Komponentenform zu beweisen, ohne auf das hier benutzte Levy-Civita-Symbol zurückzugreifen.

i)  $\vec{\nabla} \times (\alpha \vec{A} + \beta \vec{B}) = \alpha \vec{\nabla} \times \vec{A} + \beta \vec{\nabla} \times \vec{B} \quad \forall \alpha, \beta$  Konstanten in  $\mathbb{R}$ . Der Nablaoperator und die Vektor-Multiplikation sind lineare Abbildungen.

ii)  $\vec{\nabla} \times (\varphi \vec{A}) = (\vec{\nabla} \varphi) \times \vec{A} + \varphi \vec{\nabla} \times \vec{A}$ . Der Beweis geht wie folgt:

$$\begin{aligned} \vec{\nabla} \times (\varphi \vec{A}) &= \sum_{ijk} \varepsilon_{ijk} \hat{e}_i \frac{\partial}{\partial x_j} (\varphi A_k) = \\ &= \sum_{ijk} \varepsilon_{ijk} \hat{e}_i \left( \frac{\partial \varphi}{\partial x_j} A_k + \varphi \frac{\partial A_k}{\partial x_j} \right) \\ &= \sum_{ijk} \varepsilon_{ijk} \hat{e}_i (\vec{\nabla} \varphi)_j A_k + \varphi \sum_{ijk} \varepsilon_{ijk} \hat{e}_i \frac{\partial A_k}{\partial x_j} \\ &= (\vec{\nabla} \varphi) \times \vec{A} + \varphi \vec{\nabla} \times \vec{A}. \end{aligned}$$

iii)  $\vec{\nabla} \times (\vec{\nabla} \varphi) = 0$ . Der Beweis ist folgender:

$$\begin{aligned} \vec{\nabla} \times (\vec{\nabla} \varphi) &= \sum_{ijk} \varepsilon_{ijk} \hat{e}_i \frac{\partial}{\partial x_j} (\vec{\nabla} \varphi)_k = \\ &= \sum_{ijk} \varepsilon_{ijk} \hat{e}_i \frac{\partial \varphi}{\partial x_j \partial x_k} = \sum_{ijk} \varepsilon_{ijk} \hat{e}_i \frac{\partial \varphi}{\partial x_k \partial x_j} = 0. \end{aligned}$$

Die Summe über  $j$  und  $k$  ist null für alle  $i$ , da  $\frac{\partial \varphi}{\partial x_j \partial x_k} = \frac{\partial \varphi}{\partial x_k \partial x_j}$  und  $\varepsilon_{ijk} = -\varepsilon_{ikj}$ . Wir haben hiermit gezeigt, dass wenn ein Vektorfeld  $\vec{E}$  proportional zum Gradienten eines Skalarfeldes ist (z.B.  $\vec{E} = -\vec{\nabla} \varphi$ ), dann ist  $\vec{\nabla} \times \vec{E} = \vec{0}$ . Dies ist eine sehr wichtige Eigenschaft von konservativen Kraftfeldern  $\vec{F} = -\vec{\nabla} V$  und von statischen elektrischen Feldern  $\vec{E} = -\vec{\nabla} \varphi$ . Wie wir später zeigen werden, gilt auch die Eigenschaft in die andere Richtung: Sei  $\vec{E}(\vec{r})$  ein rotationsfreies Vektorfeld, d.h.,  $\vec{\nabla} \times \vec{E} = \vec{0} \quad \forall \vec{r} \in \mathbb{R}^3$ , dann lässt sich  $\vec{E}$  als

$$\vec{E} = -\vec{\nabla} \varphi$$

schreiben, wobei  $\varphi$  ein Skalarfeld ist, das Skalarpotential genannt wird.

iv)  $\vec{\nabla} \cdot (\vec{\nabla} \times \vec{A}) = 0$ , d.h., die Divergenz eines Rotationsfeldes  $\vec{\nabla} \times \vec{A}$  ist null. Dies lässt sich wie folgt zeigen:

$$\begin{aligned}\vec{\nabla} \cdot (\vec{\nabla} \times \vec{A}) &= \sum_i \frac{\partial}{\partial x_i} (\nabla \times A)_i = \sum_i \frac{\partial}{\partial x_i} \left( \sum_{jk} \varepsilon_{ijk} \frac{\partial A_k}{\partial x_j} \right) \\ &= \sum_{ijk} \varepsilon_{ijk} \frac{\partial A_k}{\partial x_i \partial x_j} = \sum_{ijk} \varepsilon_{ijk} \frac{\partial A_k}{\partial x_j \partial x_i} = 0.\end{aligned}$$

Die Summe über  $i$  und  $j$  ist null, da  $\frac{\partial A_k}{\partial x_i \partial x_j} = \frac{\partial A_k}{\partial x_j \partial x_i}$  und  $\varepsilon_{ijk} = -\varepsilon_{jik}$ .

Wir haben hiermit gezeigt, dass wenn ein Vektorfeld  $\vec{B}$  ein Rotationsfeld ist (d.h.,  $\vec{B} = \vec{\nabla} \times \vec{A}$ ), dann ist  $\vec{\nabla} \cdot \vec{B} = 0$ . Dies ist eine wichtige Eigenschaft von *stationären* Stromdichten  $\vec{j}$  ( $\vec{\nabla} \cdot \vec{j} = -\frac{\partial \rho}{\partial t} = 0$ ) und von magnetischen Feldern  $\vec{B}$ . Wie wir später zeigen werden, gilt auch die Eigenschaft in die andere Richtung: Sei  $\vec{B}(\vec{r})$  ein divergenzfreies Feld, d.h.,  $\vec{\nabla} \cdot \vec{B} = 0 \quad \forall \vec{r} \in \mathbb{R}^3$ , dann lässt sich  $\vec{B}$  als

$$\vec{B} = \vec{\nabla} \times \vec{A}$$

schreiben, wobei  $\vec{A}$  ein Vektorfeld ist, das Vektorpotential genannt wird.

Die Eigenschaften iii) und iv) lassen sich qualitativ interpretieren oder einfach im Gedächtnis festhalten, indem der Nablaoperator  $\vec{\nabla} = \left( \frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z} \right)$  als Vektor betrachtet wird. In diesem Sinne ist  $\vec{\nabla} \parallel \vec{\nabla}$ , und dementsprechend soll das Vektorprodukt  $\vec{\nabla} \times \vec{\nabla} = \vec{0}$  sein, also  $\vec{\nabla} \times \vec{\nabla} \varphi = 0$ . Darüber hinaus kann man sich vorstellen, dass  $\vec{\nabla} \times \vec{A} \perp \vec{\nabla}$ , und damit soll das Skalarprodukt  $\vec{\nabla} \cdot (\vec{\nabla} \times \vec{A}) = 0$  sein.

Obwohl der Nablaoperator  $\vec{\nabla} = \left( \frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z} \right)$  ein Vektoroperator ist und die Analogie mit Vektoren richtig und meistens sehr nützlich ist, bleibt es riskant, oft sogar irreführend, die Identitäten für mehrfache Vektorprodukte, die wir in Kapitel 2 für konstante Vektoren hergeleitet haben, an Vektorfeldern und Nablaoperatoren ohne weiteres anzuwenden. Dies liegt daran, dass die partiellen Ableitungen, d.h., die Komponenten von  $\vec{\nabla}$ , nicht mit den Komponenten von Vektorfeldern kommutieren, wie die reellen Zahlen es tun (z.B.  $\frac{\partial}{\partial x} (A_y B_z) \neq A_y \frac{\partial}{\partial x} B_z \neq A_y B_z \frac{\partial}{\partial x}$ ). Wenn der  $\vec{\nabla}$ -Operator auf ein Produkt von zwei Feldern  $\vec{A}$  und  $\vec{B}$  wirkt, entstehen zwei Terme, wie bei der Produktregel der elementaren Ableitung. In dem einen wirkt  $\vec{\nabla}$  auf  $\vec{A}$  als wäre  $\vec{B}$  konstant, und in dem zweiten wirkt  $\vec{\nabla}$  auf  $\vec{B}$  als wäre  $\vec{A}$  konstant  $\left( \frac{\partial}{\partial x_i} (A_j B_k) = \frac{\partial A_j}{\partial x_i} B_k + A_j \frac{\partial B_k}{\partial x_i} \right)$ .

In der folgenden Identität erkennt man einen solchen kniffligen Fall:

$$\vec{\nabla} \times (\vec{A} \times \vec{B}) = \text{rot}(\vec{A} \times \vec{B}) = (\vec{B} \cdot \vec{\nabla}) \vec{A} - \vec{B} (\vec{\nabla} \cdot \vec{A}) + \vec{A} (\vec{\nabla} \cdot \vec{B}) - (\vec{A} \cdot \vec{\nabla}) \vec{B}, \quad (3.6)$$

wobei  $(\vec{B} \cdot \vec{\nabla})$  der Skalaroperator

$$\left( B_x \frac{\partial}{\partial x} + B_y \frac{\partial}{\partial y} + B_z \frac{\partial}{\partial z} \right)$$

ist, und analog für den Operator  $(\vec{A} \cdot \vec{\nabla})$ . Wie immer bei Skalaroperatoren, die auf einen Vektor wirken, ist hier die Anwendung auf jede Komponente gemeint:

$$(\vec{B} \cdot \vec{\nabla}) \vec{A} = \sum_i \hat{e}_i (\vec{B} \cdot \vec{\nabla}) A_i = \sum_i \hat{e}_i \left( B_x \frac{\partial A_i}{\partial x} + B_y \frac{\partial A_i}{\partial y} + B_z \frac{\partial A_i}{\partial z} \right).$$

Der Beweis geht wie folgt:

$$\vec{\nabla} \times (\vec{A} \times \vec{B}) = \vec{\nabla}_A \times (\vec{A} \times \vec{B}) + \vec{\nabla}_B \times (\vec{A} \times \vec{B}).$$

Wir können jetzt die Formel

$$\begin{aligned} \vec{A} \times (\vec{B} \times \vec{C}) &= \vec{B} (\vec{A} \cdot \vec{C}) - \vec{C} (\vec{A} \cdot \vec{B}) \\ &= (\vec{A} \cdot \vec{C}) \vec{B} - (\vec{A} \cdot \vec{B}) \vec{C} \end{aligned}$$

anwenden, allerdings ohne  $\vec{\nabla}$  mit  $\vec{A}$  und  $\vec{B}$  zu vertauschen (d.h. die Reihenfolge zu ändern). Man erhält dann

$$\vec{\nabla} \times (\vec{A} \times \vec{B}) = (\vec{\nabla}_A \cdot \vec{B}) \vec{A} - (\vec{\nabla}_A \cdot \vec{A}) \vec{B} + (\vec{\nabla}_B \cdot \vec{B}) \vec{A} - (\vec{\nabla}_B \cdot \vec{A}) \vec{B}.$$

Jetzt können  $\vec{\nabla}_A$  und  $\vec{B}$  einerseits und  $\vec{\nabla}_B$  und  $\vec{A}$  andererseits vertauscht werden, da  $\vec{\nabla}_A$  ( $\vec{\nabla}_B$ ) nur auf  $\vec{A}$  ( $\vec{B}$ ) wirkt. Es folgt

$$\vec{\nabla} \times (\vec{A} \times \vec{B}) = (\vec{B} \cdot \vec{\nabla}_A) \vec{A} - \vec{B} (\vec{\nabla}_A \cdot \vec{A}) + \vec{A} (\vec{\nabla}_B \cdot \vec{B}) - (\vec{A} \cdot \vec{\nabla}_B) \vec{B},$$

was äquivalent zu Gl. (3.6) ist.

### 3.29 Taylorsche Entwicklung in $\mathbb{R}^n$

Das Taylorsche Approximationspolynom von Funktionen einer Variable  $x \in \mathbb{R}$  ermöglicht, eine lokale Näherung von  $f(x)$  um einen Punkt  $a$  mittels der ersten  $n$  Ableitungen

$f^{(k)}(a) = \frac{d^k f}{dx^k}(a)$  am Punkt  $a$  zu gewinnen:

$$f(x) = \sum_{k=0}^n \frac{1}{k!} f^{(k)}(a) (x-a)^k + \varepsilon(x),$$

wobei der Fehler  $\varepsilon(x)$  die Größenordnung  $O(x-a)^{n+1}$  hat, d.h.,  $\lim_{x \rightarrow a} \frac{\varepsilon(x)}{(x-a)^n} = 0$  (siehe Abschnitt 3.7).

Eine ähnliche Entwicklung gilt für Funktionen oder Felder mehrerer Variablen. Um sie herzuleiten, betrachten wir eine Funktion  $\varphi(x, y) : \mathbb{R}^2 \rightarrow \mathbb{R}$  mit stetigen partiellen Ableitungen am Punkt  $\vec{r}_0 = (x_0, y_0)$  und entwickeln  $\varphi(x, y)$  zunächst nach  $x$  bei festem  $y$ :

$$\varphi(x, y) = \varphi(x_0, y) + \frac{\partial \varphi}{\partial x}(x_0, y) (x - x_0) + \frac{1}{2} \frac{\partial^2 \varphi}{\partial x^2}(x_0, y) (x - x_0)^2 + O(x - x_0)^3. \quad (3.7)$$

Die Funktionen  $\varphi(x_0, y)$ ,  $\frac{\partial \varphi}{\partial x}(x_0, y)$ ,  $\frac{\partial^2 \varphi}{\partial x^2}(x_0, y)$  etc. lassen sich ihrerseits bezüglich  $y$  am Punkt  $y_0$  entwickeln. Man erhält:

$$\varphi(x_0, y) = \varphi(x_0, y_0) + \frac{\partial \varphi}{\partial y}(x_0, y_0) (y - y_0) + \frac{1}{2} \frac{\partial^2 \varphi}{\partial y^2}(x_0, y_0) (y - y_0)^2 + O(y - y_0)^3,$$

$$\frac{\partial \varphi}{\partial x}(x_0, y) = \frac{\partial \varphi}{\partial x}(x_0, y_0) + \frac{\partial^2 \varphi}{\partial y \partial x}(x_0, y_0) (y - y_0) + O(y - y_0)^2$$

und

$$\frac{\partial^2 \varphi}{\partial x^2}(x_0, y) = \frac{\partial^2 \varphi}{\partial x^2}(x_0, y_0) + O(y - y_0).$$

Setzt man dies in Gleichung (3.7) ein, ergibt sich

$$\begin{aligned} \varphi(x, y) = & \varphi(x_0, y_0) + \frac{\partial \varphi}{\partial x}(x_0, y_0) (x - x_0) + \frac{\partial \varphi}{\partial y}(x_0, y_0) (y - y_0) + \\ & + \frac{1}{2} \left[ \frac{\partial^2 \varphi}{\partial x^2}(x_0, y_0) (x - x_0)^2 + 2 \frac{\partial^2 \varphi}{\partial y \partial x}(x_0, y_0) (x - x_0) (y - y_0) + \right. \\ & \left. \frac{\partial^2 \varphi}{\partial y^2}(x_0, y_0) (y - y_0)^2 \right] + O(|\vec{r} - \vec{r}_0|^3). \end{aligned}$$

Für Funktionen oder Skalarfelder  $\varphi : \mathbb{R}^n \rightarrow \mathbb{R}$ , die von  $n$  Variablen  $\vec{r} = (x_1, \dots, x_n) \in \mathbb{R}^n$  abhängen, und auch für die Komponenten von Vektorfeldern gilt

$$\varphi(\vec{r}) = \varphi(\vec{a}) + \sum_{i=1}^n \frac{\partial \varphi}{\partial x_i}(\vec{a}) (x - x_i) + \frac{1}{2} \sum_{i,j=1}^n \frac{\partial^2 \varphi}{\partial x_i \partial x_j}(\vec{a}) (x_i - a_i) (x_j - a_j) + O(|\vec{r} - \vec{a}|^3), \quad (3.8)$$

wobei  $\vec{a} = (a_1, \dots, a_n)$  der Punkt ist, an dem die Funktion und die partiellen Ableitungen evaluiert werden. Der Beweis folgt aus der Taylor-Entwicklung der auxiliären Funktion

$$f(t) = \varphi(\vec{a} + t \vec{x})$$

bezüglich der Variable  $t \in \mathbb{R}$  bei  $t = 0$  für beliebige  $\vec{a}$  und  $\vec{x}$ . Die Ableitung nach  $t$  lautet

$$f'(t) = \frac{df}{dt} = \sum_{i=1}^n \frac{\partial \varphi}{\partial x_i} \frac{d(a_i + t x_i)}{dt} = \sum_{i=1}^n \frac{\partial \varphi}{\partial x_i} x_i,$$

und damit erhalten wir für die  $k$ -te Ableitung nach  $t$  ( $a_i$  und  $x_i$  fest  $\forall i$ )

$$f^{(k)}(t) = \frac{d^k f}{dt^k} = \sum_{i_1 \dots i_k=1}^n \frac{\partial^k \varphi}{\partial x_{i_1} \partial x_{i_2} \dots \partial x_{i_k}} x_{i_1} x_{i_2} \dots x_{i_k}.$$

Die Substitution  $t \vec{x} = \vec{r} - \vec{a}$  bei der Taylor-Entwicklung

$$\begin{aligned} f(t) = \varphi(\vec{a} + t \vec{x}) &= \sum_{k=0}^m \frac{1}{k!} f^{(k)}(0) t^k + O(t^{m+1}) \\ &= \sum_{k=0}^m \frac{1}{k!} \sum_{i_1 \dots i_k=1}^n \frac{\partial^k \varphi}{\partial x_{i_1} \dots \partial x_{i_k}}(\vec{a}) x_{i_1} \dots x_{i_k} t^k + O(t^{m+1}) \end{aligned}$$

von  $f(t)$  bei  $t = 0$  bis zur Ordnung  $m$  ergibt

$$\varphi(\vec{r}) = \sum_{k=0}^m \frac{1}{k!} \sum_{i_1 \dots i_k=1}^n \frac{\partial^k \varphi}{\partial x_{i_1} \dots \partial x_{i_k}}(\vec{a}) (r_{i_1} - a_{i_1}) \dots (r_{i_k} - a_{i_k}) + O(|\vec{r} - \vec{a}|^{m+1}),$$

was die Erweiterung von Gleichung (3.8) für eine beliebige Ordnung  $m$  ist.

### 3.30 Lokale Extreme

Besonders wichtig für die Anwendungen ist die Taylor-Entwicklung bis zur zweiten Ordnung, mit der man den größten Teil der Extremaufgaben bei Funktionen mehrerer Variablen lösen kann. Wir haben

$$\varphi(\vec{r}) = \varphi(\vec{a}) + \vec{\nabla} \varphi(\vec{a}) \cdot (\vec{r} - \vec{a}) + \frac{1}{2} \sum_{i,j=1}^n \underbrace{\frac{\partial^2 \varphi}{\partial x_i \partial x_j}}_{H_{ij}}(\vec{a}) (x_i - a_i) (x_j - a_j) + O(|\vec{r} - \vec{a}|^3),$$

wobei  $\vec{\nabla} \varphi(\vec{a})$  der Gradient von  $\varphi$  ist und  $H_{ij} = \frac{\partial^2 \varphi}{\partial x_i \partial x_j}(\vec{a})$  die sogenannte Hesse-Matrix der partiellen zweiten Ableitungen ist. Zum Beispiel ist für  $n = 2$

$$H = \begin{pmatrix} \frac{\partial^2 \varphi}{\partial x^2} & \frac{\partial^2 \varphi}{\partial x \partial y} \\ \frac{\partial^2 \varphi}{\partial y \partial x} & \frac{\partial^2 \varphi}{\partial y^2} \end{pmatrix}.$$

Die Hesse-Matrix ist wohl symmetrisch, d.h.,

$$H_{ij} = \frac{\partial^2 \varphi}{\partial x_i \partial x_j} = \frac{\partial^2 \varphi}{\partial x_j \partial x_i} = H_{ji}.$$

Die Funktion  $\varphi$  ist stationär bei  $\vec{a}$  wenn

$$\vec{\nabla} \varphi(\vec{a}) = 0.$$

Man sagt auch, dass  $\vec{a}$  ein *kritischer* oder *stationärer Punkt* von  $\varphi(\vec{r})$  ist. Lokale Extreme (Maxima und Minima) sind unbedingt kritische Punkte, denn bei  $\vec{\nabla}\varphi(\vec{a}) \neq 0$  würde  $\varphi(\vec{r}) - \varphi(\vec{a})$  in jede zu  $\vec{\nabla}\varphi(\vec{a})$  nicht senkrechte Richtung sein Vorzeichen am Punkt  $\vec{a}$  ändern. Es gilt also

$$\vec{\nabla}\varphi(\vec{a}) = 0$$

bei lokalen Extremen  $\vec{a}$ .

Ob  $\vec{a}$  ein Maximum, Minimum oder Sattelpunkt ist, entscheidet das Verhalten der *quadratischen Form*

$$\Delta\varphi = \frac{1}{2} \sum_{ij} \frac{\partial^2\varphi}{\partial x_i \partial x_j} (x_i - a_i) (x_j - a_j).$$

Es gelten folgende Bedingungen:

- i) Wenn  $\Delta\varphi$  *positiv definit* ist, d.h.,  $\Delta\varphi > 0 \quad \forall \vec{r} = (x_1, \dots, x_n) \neq \vec{0}$ , dann hat  $\varphi$  bei  $\vec{a}$  ein Minimum.
- ii) Wenn  $\Delta\varphi$  *negativ definit* ist, d.h.,  $\Delta\varphi < 0 \quad \forall \vec{r} \neq \vec{0}$ , dann hat  $\varphi$  ein Maximum.
- iii) Wenn  $\Delta\varphi$  *indefinit* ist (also sowohl  $\Delta\varphi > 0$  als auch  $\Delta\varphi < 0$  in Abhängigkeit von  $\vec{x}$ ), so besitzt  $\varphi$  bei  $\vec{a}$  kein Extremum, sondern einen Sattelpunkt.

Das Verhalten von  $\Delta\varphi$  hängt lediglich von den Elementen der Hesse-Matrix  $H_{ij} = \frac{\partial^2\varphi}{\partial x_i \partial x_j}(\vec{a})$

ab. Eine praktische Charakterisierung von positiv definiten quadratischen Formen kann man mit Hilfe der Determinante und Unterdeterminanten von  $H$  vornehmen.

Bei  $n = 2$  gilt

$$\text{i) } \frac{\partial^2\varphi}{\partial x^2} > 0 \quad \text{und} \quad \det(H) = \begin{vmatrix} \frac{\partial^2\varphi}{\partial x^2} & \frac{\partial^2\varphi}{\partial x \partial y} \\ \frac{\partial^2\varphi}{\partial x \partial y} & \frac{\partial^2\varphi}{\partial y^2} \end{vmatrix} > 0 \Rightarrow \Delta\varphi > 0 \Rightarrow \text{Minimum.}$$

$$\text{ii) } \frac{\partial^2\varphi}{\partial x^2} < 0 \quad \text{und} \quad \det(H) > 0 \Rightarrow \Delta\varphi < 0 \Rightarrow \text{Maximum.}$$

$$\text{iii) } \det(H) < 0 \Rightarrow \Delta\varphi \text{ indefinit} \Rightarrow \text{Sattelpunkt.}$$

Bei  $n = 3$

$$\text{i) } \frac{\partial^2\varphi}{\partial x^2} > 0, \quad \begin{vmatrix} \frac{\partial^2\varphi}{\partial x^2} & \frac{\partial^2\varphi}{\partial x \partial y} \\ \frac{\partial^2\varphi}{\partial y \partial x} & \frac{\partial^2\varphi}{\partial y^2} \end{vmatrix} > 0 \quad \text{und} \quad \det(H) > 0 \Rightarrow \Delta\varphi > 0 \Rightarrow \text{Minimum.}$$

$$\text{ii) } \frac{\partial^2 \varphi}{\partial x^2} < 0, \quad \begin{vmatrix} \frac{\partial^2 \varphi}{\partial x^2} & \frac{\partial^2 \varphi}{\partial x \partial y} \\ \frac{\partial^2 \varphi}{\partial y \partial x} & \frac{\partial^2 \varphi}{\partial y^2} \end{vmatrix} > 0 \quad \text{und} \quad \det(H) < 0 \Rightarrow \Delta \varphi < 0 \Rightarrow \text{Maximum.}$$

iii) In allen anderen Fällen, solange  $\det(H) \neq 0$ , ist  $\Delta \varphi$  indefinit und  $\varphi$  hat einen lokalen Sattelpunkt.

Im Allgemeinen (für  $n$  Variablen) gilt, dass  $\Delta \varphi$  positiv definit ist, wenn die  $n$  Unterdeterminanten auf der Diagonale der Hesse-Matrix, die man mit den ersten  $j$  Variablen bilden kann ( $1 \leq j \leq n$ ), positiv sind, d.h.,

$$H_{11} > 0, \quad \begin{vmatrix} H_{11} & H_{12} \\ H_{21} & H_{22} \end{vmatrix} > 0, \quad \dots \quad \begin{vmatrix} H_{11} & \dots & H_{1j} \\ \vdots & & \vdots \\ H_{j1} & \dots & H_{jj} \end{vmatrix} > 0, \quad \dots \quad \text{und} \quad \begin{vmatrix} H_{11} & \dots & H_{1n} \\ \vdots & & \vdots \\ H_{n1} & \dots & H_{nn} \end{vmatrix} > 0.$$

Da dieses Kriterium das einfachste ist, sollte man in der Praxis verifizieren, ob  $H$  oder  $-H$  es erfüllen. Im ersten Fall ist  $\Delta \varphi > 0$  und  $\varphi$  hat ein Minimum bei  $\vec{a}$ , im anderen Fall hat  $\varphi$  ein Maximum bei  $\vec{a}$ . Ansonsten ist  $\vec{a}$  ein Sattelpunkt.

Sollte  $\det H = 0$  sein, dann kann man mit den zweiten partiellen Ableitungen nicht entscheiden, ob  $\varphi$  ein Minimum oder Maximum bei  $\vec{a}$  hat. In dem Fall braucht man höhere Ableitungen.

## 4 Integration in $\mathbb{R}^n$

Das Integral einer Funktion  $f(x)$  über ein Domain  $A \in \mathbb{R}^n$  ist der Grenzwert der Summe der Produkte der mittleren Werte der Funktion in den verschiedenen Unterteilungen  $a_i$  von  $A$  ( $\bigcup_{i=1}^k a_i \equiv A$ ) mal die Ausdehnung oder das  $n$ -dimensionale Volumen  $\delta a_i$  von  $a_i$ , wenn die Anzahl  $k$  der Unterteilungen nach  $+\infty$  geht und damit  $\delta a_i \rightarrow 0$  geht. Der Begriff des Integrals ist also unabhängig von der Anzahl der Dimensionen ( $\mathbb{R}$ ,  $\mathbb{R}^2$ , etc.) oder der Natur des Integrationsdomains  $A$  (Intervall  $[a, b]$  in  $\mathbb{R}$ , kubisches, sphärisches oder zylindrisches Volumen in  $\mathbb{R}^3$ , Kurve  $\vec{r}(t)$ , Fläche  $\vec{r}(u, v)$ , etc.). Um dies konkret zu formulieren, betrachten wir zunächst den eigentlich schon bekannten Fall von  $\mathbb{R}$ .

### 4.1 Riemannsches Integral in $\mathbb{R}$

Wir betrachten zunächst finite Funktionen  $f : [a, b] \rightarrow \mathbb{R}$ , d.h., Funktionen, die beschränkt sind:

$$|f(x)| \leq M \quad \forall x \in [a, b] \subseteq \mathbb{R}.$$

Das Integrationsintervall  $[a, b]$  wird in  $2^r$  Stücke  $[x_i, x_{i+1}]$  zerlegt mit  $r \in \mathbb{N}$ ,  $0 \leq i \leq 2^r$ , der Länge  $|\Delta x_r| = |x_{i+1} - x_i| = |b - a| / 2^r$ ,  $x_0 = a$ ,  $x_{2^r} = b$  und  $x_{i+1} = x_i + \frac{b-a}{2^r}$ . Wenn  $b \geq a$ , ist  $\Delta x_r = x_{i+1} - x_i \geq 0$ , sonst ist  $\Delta x_r < 0$ .

Die Idee von Riemann ist, zwei Summen zu betrachten, mit denen man das Integral eingrenzen kann. Wir definieren

$$\underline{f}_{ir} = \inf\{f(x) \mid x \in [x_{i-1}, x_i]\}$$

und

$$\overline{f}_{ir} = \sup\{f(x) \mid x \in [x_{i-1}, x_i]\},$$

wobei  $i = 1, \dots, 2^r$ . Die Infima und Maxima existieren, da  $|f(x)| \leq M$ . Die entsprechende Riemannsche *Untersumme* und *Obersumme* sind

$$\underline{S}_r = \sum_{i=1}^{2^r} \underline{f}_{ir} \Delta x_r$$

und

$$\overline{S}_r = \sum_{i=1}^{2^r} \overline{f}_{ir} \Delta x_r.$$

Sie haben die wichtige Eigenschaft

$$\underline{S}_0 \leq \underline{S}_1 \leq \dots \leq \underline{S}_r \leq \overline{S}_r \leq \overline{S}_{r-1} \leq \dots \leq \overline{S}_0$$

für alle  $r$ . Die Untersumme (Obersumme) bildet eine steigende (abnehmende) Folge, die von oben (unten) beschränkt ist. Demzufolge sind beide konvergent. Die entsprechenden Grenzwerte heißen *unteres und oberes Riemannsches Integral*:

$$\int_a^b f(x) dx = \lim_{r \rightarrow \infty} \sum_{i=1}^{2^r} \underline{f}_{ir} \Delta x_r$$

und

$$\int_a^{\overline{b}} f(x) dx = \lim_{r \rightarrow \infty} \sum_{i=1}^{2^r} \overline{f}_{ir} \Delta x_r.$$

Die Funktion  $f$  heißt *Riemann-integrierbar* im Intervall  $[a, b]$ , wenn beide Grenzwerte gleich sind. In diesem Fall haben wir

$$\int_a^b f(x) dx = \int_a^{\underline{b}} f(x) dx = \int_a^{\overline{b}} f(x) dx.$$

## 4.2 Eigenschaften des Riemannsches Integrals

i) Jede finite stückweise stetige Funktion ist integrierbar. Dies folgt daraus, dass bei stetigen Funktionen  $\underline{f}_{ir}$  und  $\overline{f}_{ir}$  beliebig nah aneinander kommen, wenn  $\Delta x_r \rightarrow 0$  geht. Eine endliche Anzahl von Unstetigkeiten spielt bei den unteren und oberen Integralen keine Rolle.

ii) Seien  $f$  und  $g$  integrierbar in  $[a, b]$  und  $\alpha, \beta \in \mathbb{R}$ . Dann gilt

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

Das Integral ist ein lineares Funktionell.

iii) Für  $f$  und  $g$  integrierbar ist auch  $fg$  integrierbar.

iv) Wenn  $f(x) \geq 0 \quad \forall x \in [a, b] \Rightarrow \int_a^b f(x) dx \geq 0$ . Daraus folgt, dass

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

wenn  $f(x) \leq g(x) \quad \forall x \in [a, b]$ .

v) Dreiecksgleichung: Ist  $f$  integrierbar, so ist  $|f|$  auch integrierbar, und es gilt

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

Dies folgt aus iv), da  $-|f(x)| \leq f(x) \leq |f(x)|$ .

vi) Ist  $f$  integrierbar in  $[a, b]$ , so ist  $f$  integrierbar in  $[a, c]$  und  $[c, b]$  für alle  $c \in [a, b]$ , und es gilt

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

Der Grund dafür ist, dass die Riemannsches Summen  $\underline{S}_r$  und  $\overline{S}_r$  sich problemlos in zwei Summen teilen lassen; die eine für  $x_i \in [a, c]$  und die andere für  $x_i \in [c, b]$ . Die Beziehung gilt auch für  $c$  außerhalb  $[a, b]$  angenommen  $f$  ist integrierbar im größeren Intervall.

$$\text{vii) } \int_a^b f(x) dx = - \int_b^a f(x) dx.$$

Dies folgt aus der Definition von  $\underline{S}_r$  und  $\overline{S}_r$ , da  $\Delta x_r = (b-a)/2^r$  sein Vorzeichen ändert, wenn man die Grenzen des Integrals vertauscht. Daraus folgt

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

viii) Mittelwertsatz der Integralrechnung: Ist die Funktion  $f(x)$  stetig in  $[a, b]$ , so gibt es ein  $\xi \in [a, b]$  so dass

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

Man kann sich davon schnell überzeugen, wenn man folgendes bedenkt. Seien  $f_{\min}$  und  $f_{\max}$  die minimalen und maximalen Werte von  $f(x)$  in  $[a, b]$ . Dann gilt

$$f_{\min} \leq f(x) \leq f_{\max} \quad \Rightarrow \quad (b-a) f_{\min} \leq \int_a^b f(x) dx \leq (b-a) f_{\max}.$$

Aus der Stetigkeit von  $f(x)$  folgt, dass ein  $\xi \in [a, b]$  existiert mit  $f(\xi) = \frac{1}{b-a} \int_a^b f(x) dx$ .

### 4.3 Hauptsatz der Infinitesimalrechnung

Die Riemannsche Formulierung des Integrals entspricht zwar streng der physikalischen Motivation des Integrals als Summe infinitesimaler Beiträge, aber sie bietet an sich keine praktische allgemeine Rechenmethode. Dieser Mangel wird mit folgendem Satz behoben:

Sei  $f : [a, b] \rightarrow \mathbb{R}$  eine stetige Funktion. Da  $f$  in allen Subintervallen  $[a, x]$  mit  $x \in [a, b]$  integrierbar ist, definieren wir die Funktion

$$F(x) = \int_a^x f(x') dx' \quad \forall x \in [a, b]. \quad (4.1)$$

Die Funktion  $F(x) : [a, b] \rightarrow \mathbb{R}$  ist differenzierbar und es gilt

$$\frac{dF}{dx} = \frac{d}{dx} \int_a^x f(x') dx' = f(x).$$

Um dies zu beweisen, betrachten wir

$$F(x+h) - F(x) = \int_a^{x+h} f(x') dx' - \int_a^x f(x') dx'.$$

Aus vi) folgt

$$F(x+h) - F(x) = \int_x^{x+h} f(x') dx' = f(\xi) h,$$

wobei  $x \leq \xi \leq x+h$  (Mittelwertsatz).

$$\frac{dF}{dx} = \lim_{h \rightarrow 0} \frac{F(x+h) - F(x)}{h} = \lim_{h \rightarrow 0} f(\xi) = f(x).$$

Die letzte Gleichheit folgt aus der Stetigkeit von  $f(x)$ , da  $x \leq \xi \leq x+h$ .

Die Funktion  $F(x)$  heißt *Stammfunktion* oder *unbestimmtes Integral* von  $f$ . Sie ist durch die Bedingung

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

definiert. Jede stetige Funktion  $f$  hat eine Stammfunktion (innerhalb ihrer Stetigkeitsdomain). Eigentlich gibt es unendlich viele Stammfunktionen, die sich nur durch eine Konstante voneinander unterscheiden. In der Tat  $\frac{dF}{dx} = \frac{dG}{dx} = f(x) \Rightarrow \frac{d}{dx}(F - G) = 0 \Rightarrow F - G = \text{Konstante}$ .

Die durch Gleichung (4.1) definierte Stammfunktion erfüllt die Randbedingung  $F(a) = 0$ . Bei einer beliebigen Stammfunktion  $F(x)$  gilt

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

Die Berechnung von Riemannschen Integralen reduziert sich also auf die nicht immer einfache Berechnung von Stammfunktionen. Die Menge aller Stammfunktionen wird unbestimmtes Integral von  $f$  genannt und als  $\int f(x) dx$  oder einfach  $\int f dx$  bezeichnet.

#### 4.4 Partielle Integration und Substitution

Zu jeder Differentiationsregel gehört eine Regel oder Methode unbestimmt zu integrieren. Einige davon sollen hier erwähnt werden.

i)  $\int f(x) f'(x) dx = \frac{1}{2} f^2(x).$

ii)  $\int \frac{f'(x)}{f(x)} dx = \ln(|f(x)|)$  für  $f(x) \neq 0$ .

iii) Partielle Integration: Sind  $u$  und  $v$  differenzierbar, dann gilt

$$(uv)' = u'v + uv'$$

und damit

$$\int uv' dx = uv - \int u'v dx.$$

Man nennt dies eine *partielle Integration*, weil man zunächst den Faktor  $v'$  und dann  $u'v$  integriert. Diese Regel ist sehr nützlich, um Funktionen wie  $x^n e^x$ ,  $x^n \cos x$ ,  $x^n \sin x$  oder  $\ln x$  zu integrieren.

Beispiel:  $I_n = \int x^n e^x dx$  ( $n \in \mathbb{N}$ ). Für  $n = 0$  ist  $I_0 = e^x$ . Für  $n \geq 1$  gilt

$$\int \underbrace{x^n}_u \underbrace{e^x}_{v'} dx = \underbrace{x^n}_u \underbrace{e^x}_v - \int \underbrace{nx^{n-1}}_{u'} \underbrace{e^x}_v = x^n e^x - nI_{n-1}.$$

Die Regel gilt natürlich auch für bestimmte Integrale

$$\int_a^b uv' dx = uv \Big|_a^b - \int_a^b u'v dx.$$

Es kommt oft vor, dass eine der Funktionen (z. B.  $u(x)$ ) an beiden Integrationsgrenzen verschwindet (z. B. für  $x \rightarrow \pm\infty$  bei  $\int_{-\infty}^{+\infty} uv' dx$  oder weil sie eine besondere Randbedingung wie  $u(a) = u(b) = 0$  erfüllen müssen). In diesem Fall gilt

$$\int_a^b uv' dx = - \int_a^b u'v dx,$$

wenn  $u(a) = u(b) = 0$  oder allgemeiner  $u(a)v(a) = u(b)v(b)$ .

iv) Substitution: Aus der Kettenregel für die Ableitung erfolgt eine sehr nützliche Integrationsmethode. Wir interessieren uns zunächst für ein Integral der Form

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

mit  $f : \mathbb{R} \rightarrow \mathbb{R}$  stetig und  $u : \mathbb{R} \rightarrow \mathbb{R}$  differenzierbar. Es gilt

$$\int f(u(x)) \frac{du}{dx} dx = \int f(u) du \Big|_{u=u(x)} = F(u(x)),$$

wobei  $F(u)$  die Stammfunktion von  $f(u)$  ist, wenn  $f$  als Funktion der Variable  $u$  betrachtet wird. In der Tat

$$\frac{dF(u(x))}{dx} = \frac{dF}{du} \Big|_{u=u(x)} \frac{du}{dx} = f(u(x)) \frac{du}{dx}.$$

In der Praxis muss man, um ein Integral der Form

$$\int f(u(x)) \frac{du}{dx} dx$$

zu berechnen, folgende Schritte durchführen:

1. Formale Substitution von  $u(x) = u$  und  $\frac{du}{dx} dx = du$ .
2. Integration nach  $u \rightarrow$  Stammfunktion  $F(u)$ .
3. Substitution von  $u$  durch  $u(x)$  in der Stammfunktion  $F(u)$ .

Beispiel:  $\int x e^{x^2} dx = \int e^u \frac{du}{2} \Big|_{u=x^2} = \frac{e^{x^2}}{2}.$   
 $u = x^2 \rightarrow du = 2x dx.$

Eine alternative Anwendung der Substitutionsregel ist folgende:

$$\int f(x) dx = \int f(\varphi(t)) \frac{d\varphi}{dt} dt \Big|_{t=\varphi^{-1}(x)}.$$

Hier haben wir die Variable  $x$  durch eine bijektive Funktion  $\varphi(t)$  der Variable  $t$  ersetzt. Nachdem man die Stammfunktion  $\int f(\varphi(t)) \frac{d\varphi}{dt} dt$  gefunden hat, muss man  $t$  durch  $\varphi^{-1}(x)$  ersetzen, um die Stammfunktion  $\int f(x) dx$  zu bekommen.

In der Praxis ist die Prozedur, um  $\int f(x) dx$  zu berechnen, folgende:

1. Wahl einer passenden bijektiven Hilfsfunktion  $\varphi(t)$ .
2. Formale Substitution  $x = \varphi(t)$  und  $dx = \frac{d\varphi}{dt} dt$ .
3. Integration nach  $t$ .
4. Substitution von  $t = \varphi^{-1}(x)$  durch die inverse Funktion.

Beispiele:

- i.  $\int \frac{x^2 + 1}{\sqrt{x+1}} dx$ . Man versucht, das Integral zu vereinfachen, indem man die Wurzeln substituiert. Man wählt also  $t = \sqrt{x+1} \Leftrightarrow x = t^2 - 1$ . Daraus folgt  $dx = 2t dt$  und

$$\begin{aligned} \int \frac{x^2 + 1}{\sqrt{x+1}} dx &= \int \frac{[(t^2 - 1)^2 + 1]}{t} 2t dt \Big|_{t=\sqrt{x+1}} = 2 \int (t^4 - 2t + 2) dt \\ &= 2 \left( \frac{t^5}{5} - t^2 + 2t \right) \Big|_{t=\sqrt{x+1}}. \end{aligned}$$

- ii. Ein weiteres bekanntes Beispiel ist  $\int (1+x^2)^{-1} dx$ . In diesem Fall empfiehlt sich die Substitution  $x = \tan(t) \Rightarrow dx = \frac{1}{\cos^2 t} dt$ ,  $t = \arctan(x)$ . Man erhält dann

$$\begin{aligned} \int \frac{1}{1+x^2} dx &= \int \frac{1}{1 + \frac{\sin^2 t}{\cos^2 t}} \cdot \frac{1}{\cos^2 t} dt = \int dt \Big|_{t=\arctan(x)} = t \Big|_{t=\arctan(x)} \\ &= \arctan(x). \end{aligned}$$

Bei bestimmten Integralen braucht man die zweite Substitution  $t = \varphi^{-1}(x)$  nicht durchzuführen. In diesem Fall ist es praktischer, die Integrationsgrenzen zu ersetzen. Es gilt

$$\int_a^b f(x) dx = \int_{\varphi^{-1}(a)}^{\varphi^{-1}(b)} f(\varphi(t)) \frac{d\varphi}{dt} dt = F(b) - F(a),$$

wobei  $\frac{dF}{dx} = f(x)$ . Dies ist äquivalent zu

$$\int_a^b f(u(x)) \frac{du}{dx} dx = \int_{u(a)}^{u(b)} f(u) du = F(u(b)) - F(u(a)),$$

wobei  $\frac{dF}{du} = f(u)$ .

## 4.5 Uneigentliche Integrale

Das Riemannsches Integral ist für finite Funktionen ( $f(x) \leq M$ ) über ein endliches Intervall  $[a, b]$  definiert. Es gibt grundsätzlich zwei wichtige Erweiterungen. In dem einen Fall wird die Integrationsdomain unendlich, z.B.

$$\int_0^{+\infty} e^{-x} dx = -e^{-x} \Big|_0^{+\infty} = 1,$$

und in dem anderen divergiert die Funktion an eine oder beide Integrationsgrenzen, z.B.

$$\int_0^1 \frac{1}{\sqrt{x}} dx = 2\sqrt{x} \Big|_0^1 = 2.$$

In beiden Fällen werden die Integrale durch die entsprechenden Grenzwerte definiert. Die Existenz des Integrals hängt natürlich mit der der Grenzwerte zusammen.

Man definiert also bei integrierbarem  $f(x)$  in  $[a, b] \forall b > a$ :

$$\int_a^{+\infty} f(x) dx = \lim_{b \rightarrow +\infty} \int_a^b f(x) dx.$$

Wenn  $f(x)$  finit und integrierbar für alle  $x \in [a, b[$  ( $a \leq x < b$ ) ist, definiert man

$$\int_a^b f(x) dx = \lim_{x \rightarrow b^-} \int_a^x f(x') dx'.$$

Als Beispiele betrachten wir die Integrale

$$\int_0^1 \frac{dx}{x^\alpha} \quad \text{und} \quad \int_1^{+\infty} \frac{dx}{x^\alpha}.$$

Für  $\alpha = 1$  ist  $\int \frac{dx}{x} = \ln x$ . Man erhält

$$\int_0^1 \frac{dx}{x} = \ln x \Big|_0^1 = - \lim_{x \rightarrow 0^+} \ln x = +\infty$$

und

$$\int_1^{+\infty} \frac{dx}{x} = \ln x \Big|_1^{+\infty} = \lim_{x \rightarrow +\infty} \ln x = +\infty.$$

Beide Integrale sind also divergent.

Für  $\alpha \neq 1$  ist  $\int \frac{dx}{x^\alpha} = \frac{x^{1-\alpha}}{1-\alpha}$ . Die gesuchten Integrale lauten

$$\int_0^1 \frac{dx}{x^\alpha} = \frac{x^{1-\alpha}}{1-\alpha} \Big|_0^1 = \frac{1}{1-\alpha} - \lim_{x \rightarrow 0^+} \frac{x^{1-\alpha}}{1-\alpha} = \begin{cases} \frac{1}{1-\alpha} & \alpha < 1 \\ +\infty & \alpha > 1 \end{cases}$$

und

$$\int_1^{+\infty} \frac{dx}{x^\alpha} = \frac{x^{1-\alpha}}{1-\alpha} \Big|_1^{+\infty} = \lim_{x \rightarrow +\infty} \frac{x^{1-\alpha}}{1-\alpha} - \frac{1}{1-\alpha} = \begin{cases} \frac{1}{\alpha-1} & \alpha > 1 \\ +\infty & \alpha < 1. \end{cases}$$

Diese Integrale sind besonders nützlich im Zusammenhang mit dem folgenden *Vergleichskriterium*:

Sei  $|f(x)| \leq g(x) \forall x \in [a, b[$  ( $a \leq x < b$ ). Wenn das uneigentliche Integral  $\int_a^b g(x) dx$  konvergiert, dann konvergiert auch  $\int_a^b f(x) dx$ .

#### 4.6 Integrale von Funktionen mit Parametern

Wir betrachten Funktionen  $f(x, \lambda) : \mathbb{R}^2 \rightarrow \mathbb{R}$ , das heißt, Funktionen einer Variable  $x$  und eines (oder mehrerer) Parameter  $\lambda \in \mathbb{R}$ . Wir gehen davon aus, dass  $f(x, \lambda)$  stetig bezüglich  $x \forall \lambda$  ist und dass  $f(x, \lambda)$  partiell differenzierbar nach  $\lambda \forall x$  ist mit  $\frac{\partial f}{\partial \lambda}(x, \lambda)$  eine stetige Funktion von  $x$  und  $\lambda$ . Dann gilt die sogenannte *Leibnizsche Regel*:

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

Dies wird oft dazu benutzt, eine Beziehung zwischen Integralen herzustellen, die die Berechnung des einen als die Ableitung des anderen ermöglicht. Zum Beispiel

$$\int_0^{+\infty} x^2 e^{-x^2} dx = -\frac{d}{d\lambda} \int_0^{+\infty} e^{-\lambda x^2} dx \Big|_{\lambda=1}.$$

Manchmal hängt die Funktion  $f(x)$  nicht von  $\lambda$  ab, aber dafür die Integrationsgrenzen. Zum Beispiel

$$I(\lambda) = \int_a^{b(\lambda)} f(x) dx.$$

Wenn  $f(x)$  stetig in  $[a, b(\lambda)] \forall \lambda$  ist und  $b(\lambda)$  differenzierbar ist, gilt

$$\frac{dI}{d\lambda} = \frac{dI}{db} \cdot \frac{db}{d\lambda} = f(b(\lambda)) \cdot \frac{db}{d\lambda}.$$

Im letzten Schritt wurde der Hauptsatz der Infinitesimalrechnung angewandt.

Es ist nicht auszuschließen, dass sowohl die Funktion als auch die Grenzen von einem externen Parameter abhängen. In diesem Fall und unter den üblichen Voraussetzungen für das Wohlverhalten von  $f(x, \lambda)$  und  $b(\lambda)$  gilt:

$$\frac{d}{d\lambda} \int_a^{b(\lambda)} f(x, \lambda) dx = f(b(\lambda), \lambda) \frac{db}{d\lambda} + \int_a^{b(\lambda)} \frac{\partial f}{\partial \lambda}(x, \lambda) dx. \quad (4.2)$$

Man kann sich davon schnell überzeugen:

$$\begin{aligned}
 I(\lambda + \Delta\lambda) - I(\lambda) &= \int_0^{b(\lambda+\Delta\lambda)} f(\lambda + \Delta\lambda) dx - \int_0^{b(\lambda)} f(\lambda) dx = \\
 &= \int_0^{b(\lambda+\Delta\lambda)} f(\lambda + \Delta\lambda) dx - \int_0^{b(\lambda+\Delta\lambda)} f(\lambda) dx + \\
 &\quad + \int_0^{b(\lambda+\Delta\lambda)} f(\lambda) dx - \int_0^{b(\lambda)} f(\lambda) dx \\
 &= \int_0^{b(\lambda+\Delta\lambda)} [f(\lambda + \Delta\lambda) - f(\lambda)] dx + \int_{b(\lambda)}^{b(\lambda+\Delta\lambda)} f(\lambda) dx.
 \end{aligned}$$

Gleichung (4.2) entsteht, wenn man durch  $\Delta\lambda$  teilt und den Limes  $\Delta\lambda \rightarrow 0$  nimmt.

Integrale der Form  $\int_a^b f(x, \lambda) dx$  können natürlich auch bezüglich  $\lambda$  integriert werden. Wenn  $f(x, \lambda)$  stetig bezüglich  $x$  und  $\lambda$  ist, darf man die Reihenfolge der Integrationen miteinander vertauschen:

$$\int_\alpha^\beta d\lambda \int_a^b f(x, \lambda) dx = \int_a^b dx \int_\alpha^\beta f(x, \lambda) d\lambda.$$

Voraussetzung dafür ist, dass die Integrationsgrenzen  $a$  und  $b$  ( $\alpha$  und  $\beta$ ) unabhängig von  $\lambda$  ( $x$ ) sind.

## 4.7 Doppel- und Dreifachintegrale

In diesem Abschnitt erweitern wir den Begriff des Integrals auf 2, 3 oder  $n$  Variablen. Typische Beispiele für Anwendungen sind die Berechnung des Volumens unter einer Kurve  $z = f(x, y)$  in einem abgegrenzten Bereich  $x_1 \leq x \leq x_2$  und  $y_1 \leq y \leq y_2$  in  $\mathbb{R}^2$  oder die Berechnung der gesamten Masse eines dreidimensionalen Körpers mit der Dichte  $\rho(x, y, z)$  innerhalb eines Volumens  $x_1 \leq x \leq x_2$ ,  $y_1 \leq y \leq y_2$  und  $z_1 \leq z \leq z_2$  in  $\mathbb{R}^3$ .

Zur Erläuterung betrachten wir zunächst den zweidimensionalen Fall. Das Integrationsgebiet  $A$  wird normalerweise in der Form  $x_1 \leq x \leq x_2$  und  $y_1(x) \leq y \leq y_2(x)$  mit  $x_1$  und  $x_2$  Konstanten und  $y_1(x)$  und  $y_2(x) : \mathbb{R} \rightarrow \mathbb{R}$  stetige Funktionen, oder in der Form  $y_1 \leq y \leq y_2$  und  $x_1(y) \leq x \leq x_2(y)$  geschrieben. Bei einer komplexeren Topologie des Integrationsbereiches ist man dazu gezwungen, mehrere solcher einfachen Gebiete zu betrachten. Das Integral der Funktion  $f(x, y) : \mathbb{R}^2 \rightarrow \mathbb{R}$  über das Gebiet  $A \subseteq \mathbb{R}^2$  wird in Analogie zum eindimensionalen Fall durch folgenden Grenzwert definiert:

$$\int_A f(x, y) da = \lim_{n_x, n_y \rightarrow +\infty} \sum_{i=1}^{n_x} \sum_{j=1}^{n_y} f(x_i, y_j) \Delta a_{ij},$$

wobei  $\Delta a_{ij}$  das Flächenelement am Punkt  $(x_i, y_j) \in A$  ist ( $n_x, n_y \in \mathbb{N}$ ). Da  $\bigcup_{ij} \Delta a_{ij} \equiv A \forall n_x$  und  $n_y$ , geht  $\Delta a_{ij} \rightarrow 0$ , wenn  $n_x$  oder  $n_y \rightarrow +\infty$  geht. Eigentlich hätte man die untere und obere Riemannsche Summe betrachten müssen, daraus untere und obere Integrale herleiten und das Integral erst definieren, wenn die unteren und oberen Grenzwerte gleich sind.

Da die  $x$ - und  $y$ -Achsen senkrecht zueinander sind (kartesische Koordinaten), ist  $\Delta a_{ij} = \Delta x_i \Delta y_j$ , wobei  $\Delta x_i = x_{i+1} - x_i$  und  $\Delta y_j = y_{j+1} - y_j$  mit ihrem Vorzeichen die Zerlegung in  $x$ - und  $y$ -Richtung angeben. Damit können wir schreiben

$$\int_A f(x, y) da = \lim_{n_x \rightarrow \infty} \sum_{i=1}^{n_x} \Delta x_i \lim_{n_y \rightarrow \infty} \sum_{j=1}^{n_y} f(x_i, y_j) \Delta y_j.$$

Wir erkennen dabei das eindimensionale Integral bezüglich  $y$ :

$$\lim_{n_y \rightarrow \infty} \sum_{j=1}^{n_y} f(x_i, y_j) \Delta y_j = \int_{y_1(x_i)}^{y_2(x_i)} f(x_i, y) dy.$$

Daraus folgt

$$\int_A f(x, y) da = \lim_{n_x \rightarrow \infty} \sum_{i=1}^{n_x} \Delta x_i \int_{y_1(x_i)}^{y_2(x_i)} f(x_i, y) dy$$

und

$$\int_A f(x, y) da = \int_A f(x, y) dx dy = \int_{x_1}^{x_2} dx \int_{y_1(x)}^{y_2(x)} f(x, y) dy.$$

Das Integral in zwei Dimensionen lässt sich also in der Form zweier aufeinanderfolgender eindimensionaler Integrale schreiben. Die Voraussetzungen dafür sind, dass  $f(x, y)$  als Funktion von  $y$  für alle  $x$  [ $(x, y) \in A$ ] Riemann-integrierbar ist und dass  $F(x) = \int_{y_1(x)}^{y_2(x)} f(x, y) dy$  als Funktion von  $x$  in  $[x_1, x_2]$  Riemann-integrierbar ist. Diese Voraussetzungen werden natürlich erfüllt, wenn  $f(x, y)$  eine stückweise stetige Funktion von  $(x, y) \in A$  ist.

Ist das Integrationsgebiet in der Form  $A = \{(x, y) \in \mathbb{R}^2 \mid y_1 \leq y \leq y_2 \text{ und } x_1(y) \leq x \leq x_2(y)\}$  gegeben, dann gilt

$$\int_A f(x, y) da = \int_{y_1}^{y_2} dy \int_{x_1(y)}^{x_2(y)} f(x, y) dx.$$

Sind die Integrationsgrenzen unabhängig von  $x$  und  $y$ , dann kann man die Reihenfolge der Integrale vertauschen:

$$\int_{x_1}^{x_2} dx \int_{y_1}^{y_2} dy f(x, y) = \int_{y_1}^{y_2} dy \int_{x_1}^{x_2} dx f(x, y) = \int_{x_1}^{x_2} \int_{y_1}^{y_2} f(x, y) dx dy.$$

Dies liegt daran, dass man die Limes für  $n_x \rightarrow \infty$  und  $n_y \rightarrow \infty$  miteinander vertauschen kann, wenn beide existieren. Voraussetzung dafür ist, dass beide Integrale konvergieren, was bei manchen uneigentlichen Integralen nicht unbedingt der Fall ist.

Im besonderen Fall, dass sich  $f(x, y) = g(x) h(y)$  als Produkt zweier Funktionen einer Variable schreiben lässt, ist das Doppelintegral durch das Produkt zweier eindimensionaler Integrale gegeben. Es gilt

$$\int_{x_1}^{x_2} dx \int_{y_1}^{y_2} g(x) h(y) dy = \left[ \int_{x_1}^{x_2} g(x) dx \right] \left[ \int_{y_1}^{y_2} h(y) dy \right].$$

Beispiel: Sei  $f(x, y) = x^2 y$  und  $A$  ein Halbkreis auf der positiven  $y$ -Halbebene, d.h.,  $A = \{(x, y) \in \mathbb{R}^2 \mid -R \leq x \leq R \text{ und } 0 \leq y \leq \sqrt{R^2 - x^2}\}$ . Das Integral von  $f$  über  $A$  ist

$$\begin{aligned} \int_A f(x, y) da &= \int_{-R}^R dx \int_0^{\sqrt{R^2 - x^2}} x^2 y dy = \int_{-R}^R dx x^2 \left( \frac{y^2}{2} \right) \Big|_{y=0}^{y=\sqrt{R^2 - x^2}} \\ &= \int_{-R}^R x^2 \left[ \frac{(\sqrt{R^2 - x^2})^2}{2} - \frac{0^2}{2} \right] dx = \frac{1}{2} \int_{-R}^R (R^2 - x^2) x^2 dx \\ &= \frac{2}{15} R^5. \end{aligned}$$

Man hätte aber genauso gut das Gebiet in der Form  $A = \{(x, y) \in \mathbb{R}^2 \mid 0 \leq y \leq R \text{ und } -\sqrt{R^2 - y^2} \leq x \leq \sqrt{R^2 - y^2}\}$  betrachten können. Dann hat man

$$\begin{aligned} \int_A f(x, y) da &= \int_0^R y dy \int_{-\sqrt{R^2 - y^2}}^{\sqrt{R^2 - y^2}} x^2 dx = \int_0^R y dy \left( \frac{x^3}{3} \right) \Big|_{x=-\sqrt{R^2 - y^2}}^{x=\sqrt{R^2 - y^2}} \\ &= \frac{2}{3} \int_0^R y (R^2 - y^2)^{3/2} dy = \frac{2}{3} \left[ -\frac{1}{5} (R^2 - y^2)^{5/2} \right] \Big|_0^R \\ &= \frac{2}{15} R^5. \end{aligned}$$

Die Erweiterung auf drei bzw.  $n$  Dimensionen ist selbstverständlich. Zum Beispiel ist das Volumenintegral der Funktion  $f(x, y, z) : \mathbb{R}^3 \rightarrow \mathbb{R}$  über das Volumen  $V = \{(x, y, z) \in \mathbb{R}^3 \mid x_1 \leq x \leq x_2, y_1(x) \leq y \leq y_2(x) \text{ und } z_1(x, y) \leq z \leq z_2(x, y)\}$  durch das dreifache Integral

$$\int_V f(x, y, z) dv = \int_{x_1}^{x_2} dx \int_{y_1(x)}^{y_2(x)} dy \int_{z_1(x, y)}^{z_2(x, y)} dz f(x, y, z)$$

gegeben. Ein einfaches Beispiel eines Integrationsvolumens in  $\mathbb{R}^3$  ist die Kugel:  $-r \leq x \leq R$ ,  $-\sqrt{R^2 - x^2} \leq y \leq \sqrt{R^2 - x^2}$  und  $-\sqrt{R^2 - x^2 - y^2} \leq z \leq \sqrt{R^2 - x^2 - y^2}$ .

Bei festen Integrationsgrenzen und unter den üblichen Voraussetzungen über das Wohlverhalten von  $f(x, y, z)$  kann man natürlich die Integrale bezüglich  $x$ ,  $y$  und  $z$  in beliebiger Reihenfolge durchführen.

#### 4.8 Wechsel der Variablen: Krummlinige Koordinaten

Wie das vorherige Beispiel es ahnen lässt, können mehrfachdimensionale Integrale in kartesischen Koordinaten  $x$ ,  $y$  und  $z$  selbst bei einfachen Integrationsgebieten wie Kreisen, Kugeln oder Zylindern furchtbar kompliziert aussehen. Die Komplexität des Ausdrucks für das Integrationsgebiet der Kugel in kartesischen Koordinaten scheint völlig unnötig zu sein, wenn man ihre Definition in Kugelkoordinaten betrachtet, nämlich  $0 \leq r \leq R$ ,  $0 \leq \varphi \leq 2\pi$  und  $0 \leq \theta \leq \pi$ . Dazu kommt, dass viele für die Physik relevante Funktionen sphärische oder zylindrische Symmetrie zeigen. Es ist klar, dass eine Integration von  $f : \mathbb{R}^3 \rightarrow \mathbb{R}$  mit  $f = f(r)$  als Funktion von  $r = \sqrt{x^2 + y^2 + z^2}$  extrem viel einfacher ist als drei Integrale von  $f = f(\sqrt{x^2 + y^2 + z^2})$  bezüglich  $x$ ,  $y$  und  $z$ . Um davon

profitieren zu können, müssen wir krummlinige Koordinaten und die Handhabung von Variablenwechseln bei Integralen besprechen.

Bevor wir den Wechsel von Variablen in Doppel- und Dreifachintegralen betrachten, lohnt es sich, nochmals den eindimensionalen Fall kurz zu diskutieren. Wir gehen von einem Integral der Form

$$\int_a^b f(x) dx = \lim_{n \rightarrow \infty} \sum_{i=1}^n f(x_i) \Delta x_i \quad (4.3)$$

aus, wobei rechts des Limes die Riemannsche Summe angedeutet wurde. Wir geben eine bijektive Funktion  $x : [u_a, u_b] \rightarrow [a, b]$  an, mit der  $x = x(u)$  als Funktion der neuen Variable  $u$  ausgedrückt wird [ $x(u_a) = a$  und  $x(u_b) = b$ ]. Um das Integral als Funktion von  $u$  zu berechnen, ersetzen wir das Längenelement  $\Delta x_i = \frac{dx}{du} \Delta u_i$  auf der rechten Seite von Gleichung (4.3) durch seinen Wert als Funktion des Längenelements  $\Delta u_i$  der neuen Variable und erhalten

$$\begin{aligned} \int_a^b f(x) dx &= \lim_{n \rightarrow \infty} \sum_{i=1}^n f(x(u_i)) \frac{dx}{du}(u_i) \Delta u_i = \int_{u_a}^{u_b} f(x(u)) \frac{dx}{du}(u) du \\ &= \int_{x^{-1}(a)}^{x^{-1}(b)} f(x(u)) \frac{dx}{du} du. \end{aligned}$$

Dies stimmt mit der bekannten Formel der Substitutionsmethode überein. Der zusätzliche Faktor  $\frac{dx}{du}$  ergibt sich aus der Beziehung  $dx = \frac{dx}{du} du$  zwischen den Längenelementen in den beiden Variablen. Bei mehreren Variablen geht man genauso vor, d.h., man braucht den Ausdruck für die entsprechenden Flächen- oder Volumenelemente als Funktion der neuen Variablen.

Zum Beispiel ist der Variablenwechsel für  $n = 2$  durch eine bijektive Funktion  $\vec{r} : \mathbb{R}^2 \rightarrow \mathbb{R}^2$  gegeben, die die ursprünglichen Koordinaten  $x = x(u, v)$  und  $y = y(u, v)$  als Funktion der neuen angibt [ $\vec{r} = (x(u, v), y(u, v))$ ]. Wir gehen davon aus, dass  $x(u, v)$  und  $y(u, v)$  stetige partielle Ableitungen haben. Als *parametrische Kurven* der Koordinatentransformation bezeichnet man die Kurven  $\vec{r}(u, v_0) = (x(u, v_0), y(u, v_0))$  und  $\vec{r}(u_0, v) = (x(u_0, v), y(u_0, v))$ , bei denen nur eine der beiden Variablen variiert wird. Da  $\vec{r} = \vec{r}(u, v)$  bijektiv ist, geht nur eine Kurve jeder Art ( $\vec{r}(u, v_0)$  und  $\vec{r}(u_0, v)$ ) durch jeden Punkt  $\vec{r}(u_0, v_0)$  hindurch. Die Tangenten-Vektoren  $\vec{t}_u$  und  $\vec{t}_v$  zu den Kurven  $\vec{r}(u, v_0)$  und  $\vec{r}(u_0, v)$  sind durch entsprechende partielle Ableitungen  $\vec{t}_u = \frac{\partial \vec{r}}{\partial u}$  und  $\vec{t}_v = \frac{\partial \vec{r}}{\partial v}$  gegeben:

$$\vec{r}(u_0 + \Delta u, v_0) = \vec{r}(u_0, v_0) + \frac{\partial \vec{r}}{\partial u}(u_0, v_0) \Delta u + O(\Delta u^2)$$

$$\vec{r}(u_0, v_0 + \Delta v) = \vec{r}(u_0, v_0) + \frac{\partial \vec{r}}{\partial v}(u_0, v_0) \Delta v + O(\Delta v^2).$$

Die parametrischen Kurven  $\vec{r}(u, v_0)$  und  $\vec{r}(u_0, v)$  definieren einen Satz krummliniger Koordinaten auf der Ebene.

Das wichtigste Beispiel krummliniger Koordinaten auf der Ebene sind die Polarkoordinaten

$$\begin{aligned} u &= \varrho & 0 &\leq \varrho & x &= \varrho \cos \varphi \\ v &= \varphi & 0 &\leq \varphi \leq 2\pi & y &= \varrho \sin \varphi. \end{aligned}$$

Die parametrischen Kurven sind radial oder azimuthal und immer senkrecht zueinander. Die entsprechenden Tangenten-Vektoren lauten  $\vec{t}_\rho = \frac{\partial \vec{r}}{\partial \rho} = (\cos \varphi, \sin \varphi)$  und  $\vec{t}_\varphi = \rho (-\sin \varphi, \cos \varphi)$ .

Ein zweidimensionales Integral über ein Gebiet  $A \subseteq \mathbb{R}^2$  ist durch folgenden Grenzwert definiert:

$$\int_A f(x, y) dx dy = \lim_{n \rightarrow \infty} \sum_{ij=1}^n f(x_i, y_j) \Delta a_{ij}, \quad (4.4)$$

wobei  $\Delta a_{ij}$  das Flächenelement um den Punkt  $(x_i, y_j)$  ist. Um dieses Flächenelement als Funktion der krummlinigen Koordinaten  $u$  und  $v$  zu berechnen, betrachten wir die Umgebung eines Punktes  $\vec{r}_0 = (x(u_0, v_0), y(u_0, v_0))$  bezüglich der neuen Koordinaten. Wir interessieren uns für die infinitesimale Fläche um  $\vec{r}_0$ , die durch die Kreuzung der parametrischen  $u$ -Kurven  $\vec{r}(u, v_0)$  und  $\vec{r}(u, v_0 + \Delta v)$  mit den  $v$ -Kurven  $\vec{r}(u_0, v)$  und  $\vec{r}(u_0 + \Delta u, v)$  entsteht. Die Vektoren, die das Flächenelement  $\Delta a$  definieren, sind

$$\Delta \vec{r}_u = \vec{t}_u \Delta u = \left( \frac{\partial x}{\partial u}, \frac{\partial y}{\partial u} \right) \Delta u$$

und

$$\Delta \vec{r}_v = \vec{t}_v \Delta v = \left( \frac{\partial x}{\partial v}, \frac{\partial y}{\partial v} \right) \Delta v.$$

In zwei Dimensionen ist die Fläche des von  $\Delta \vec{r}_u$  und  $\Delta \vec{r}_v$  definierten Parallelogramms durch die Determinante der Matrix gegeben, die man mit den beiden Vektoren bildet. Wir erhalten also

$$\Delta a = \begin{vmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v} \end{vmatrix} \Delta u \Delta v,$$

oder in Differentialform

$$da = \left( \frac{\partial x}{\partial u} \frac{\partial y}{\partial v} - \frac{\partial x}{\partial v} \frac{\partial y}{\partial u} \right) du dv.$$

Substitution in Gleichung (4.4) ergibt

$$\int_A f(x, y) dx dy = \int_A f(x(u, v), y(u, v)) J(u, v) du dv,$$

wobei

$$J(u, v) = \frac{\partial x}{\partial u} \frac{\partial y}{\partial v} - \frac{\partial x}{\partial v} \frac{\partial y}{\partial u} \quad (4.5)$$

die sogenannte *Funktionaldeterminante* oder *Jacobi-Determinante* der Koordinatentransformation  $x = x(u, v)$  und  $y = y(u, v)$  ist.

Es lohnt sich an dieser Stelle, die geometrische Bedeutung der Funktionaldeterminante zu diskutieren.  $J(u, v)$  korrigiert die Distortionen bzw. Krümmungen der Linien konstanter  $u$  und  $v$ . Im kartesischen Koordinatensystem sind diese Linien stets parallel zueinander und mit gleichem Skalierungsfaktor 1 für alle Achsen. Deshalb ist das Flächenelement

$\Delta a = \Delta x \Delta y$ . Bei krummlinigen Koordinaten ist dies nicht mehr der Fall. Was parallel und harmlos im  $(u, v)$ -Raum ist (Die Linien konstanter  $u$  und  $v$  laufen ja parallel im  $(u, v)$ -Referenzsystem), wird durch die Abbildung  $\vec{r}(u, v) = (x(u, v), y(u, v))$  in den kartesischen  $(x, y)$ -Raum skaliert und gekrümmt.  $J(u, v)$  berücksichtigt sowohl die lokalen Änderungen der Skalen in beiden Richtungen als auch die Änderungen der Winkel der parametrischen Kurven.

Die Beziehung

$$dx dy = J(u, v) du dv$$

zeigt auch, dass  $dx dy$  kein übliches Produkt von Differentialen ist. Es ist nicht ohne Grund, dass die Mathematiker immer daran erinnern, dass “ $dx$ ” und “ $dx dy$ ” eher eine symbolische Rolle bei der Schreibweise der Integrale hat. Man kann sich damit amüsieren, das Produkt der Differentiale

$$dx = \frac{\partial x}{\partial u} du + \frac{\partial x}{\partial v} dv$$

und

$$dy = \frac{\partial y}{\partial u} du + \frac{\partial y}{\partial v} dv$$

zu bilden. Man würde in dem Fall etwas erhalten, was nichts mit dem Flächendifferential  $da = J(u, v) du dv$  zu tun hat.

Die Jacobi-Determinante ist negativ, wenn die Orientierung des Raumes durch die Transformation geändert wird. Die Punkte, bei denen  $J(u, v) = 0$  ist, werden singuläre Punkte genannt. Dies ist eher unüblich, da es bedeuten würde, dass die Tangenten der  $u$ - und  $v$ -Kurven an dem Punkt parallel sind. Bei den für uns wichtigen krummlinigen Koordinaten sind die parametrischen Kurven immer senkrecht zueinander (z. B. Polar-, Zylindrische und Kugelkoordinaten).

Als Beispiel betrachten wir die Polarkoordinaten  $x = \varrho \cos \varphi$  und  $y = \varrho \sin \varphi$ . Aus der Gleichung (4.5) erhalten wir

$$J(\varrho, \varphi) = \begin{vmatrix} \frac{\partial x}{\partial \varrho} & \frac{\partial x}{\partial \varphi} \\ \frac{\partial y}{\partial \varrho} & \frac{\partial y}{\partial \varphi} \end{vmatrix} = \begin{vmatrix} \cos \varphi & -\varrho \sin \varphi \\ \sin \varphi & \varrho \cos \varphi \end{vmatrix} = \varrho.$$

Es gilt also  $dx dy = \varrho d\varrho d\varphi$ .

Die Transformation der Volumenelemente in 3 Dimensionen ist völlig analog. Man geht von einer Koordinatentransformation

$$(u, v, w) \rightarrow \begin{cases} x(u, v, w) \\ y(u, v, w) \\ z(u, v, w) \end{cases}$$

aus. Eine Änderung  $\Delta u$  der Koordinate  $u$  bei festem  $v$  und  $w$  führt zu einer Änderung

$$\Delta \vec{r}_u = \left( \frac{\partial x}{\partial u}, \frac{\partial y}{\partial u}, \frac{\partial z}{\partial u} \right) \Delta u$$

des Positionsvektors  $\vec{r}(u, v, w)$  im kartesischen System. Analog dazu ist im Falle einer Variation  $\Delta v$  bei festem  $u$  und  $w$

$$\Delta \vec{r}_v = \left( \frac{\partial x}{\partial v}, \frac{\partial y}{\partial v}, \frac{\partial z}{\partial v} \right) \Delta v,$$

und in Folge einer Variation  $\Delta w$  bei festem  $u$  und  $v$  ist

$$\Delta \vec{r}_w = \left( \frac{\partial x}{\partial w}, \frac{\partial y}{\partial w}, \frac{\partial z}{\partial w} \right) \Delta w.$$

Das dabei entstehende Volumenelement  $\Delta V$  ist durch das Spatprodukt

$$\Delta V = \Delta \vec{r}_u \cdot (\Delta \vec{r}_v \times \Delta \vec{r}_w)$$

gegeben. Wir können also schreiben

$$dV = dx dy dz = J(u, v, w) du dv dw,$$

wobei

$$J(u, v, w) = \begin{vmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} & \frac{\partial x}{\partial w} \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v} & \frac{\partial y}{\partial w} \\ \frac{\partial z}{\partial u} & \frac{\partial z}{\partial v} & \frac{\partial z}{\partial w} \end{vmatrix} \quad (4.6)$$

die Funktional- oder Jacobi-Determinante für die Koordinatentransformation  $(u, v, w) \rightarrow (x, y, z)$  ist.

Beispiele:

- i) Zylinderkoordinaten:  $x = \varrho \cos \varphi$ ,  $y = \varrho \sin \varphi$  und  $z = z$ . Aus Gleichung (4.6) erhalten wir

$$J(\varrho, \varphi, z) = \begin{vmatrix} \cos \varphi & -\varrho \sin \varphi & 0 \\ \sin \varphi & \varrho \cos \varphi & 0 \\ 0 & 0 & 1 \end{vmatrix} = \varrho.$$

Es gilt also  $dx dy dz = \varrho d\varrho d\varphi dz$ .

- ii) Kugelkoordinaten:  $x = r \sin \theta \cos \varphi$ ,  $y = r \sin \theta \sin \varphi$  und  $z = r \cos \theta$ . Aus Gleichung (4.6) erhalten wir

$$\begin{aligned} J(r, \theta, \varphi) &= \begin{vmatrix} \sin \theta \cos \varphi & r \cos \theta \cos \varphi & -r \sin \theta \sin \varphi \\ \sin \theta \sin \varphi & r \cos \theta \sin \varphi & r \sin \theta \cos \varphi \\ \cos \theta & -r \sin \theta & 0 \end{vmatrix} \\ &= \cos \theta (r^2 \cos \theta \sin \theta) + r \sin \theta (r \sin^2 \theta) \\ &= r^2 \sin \theta. \end{aligned}$$

Es gilt dann

$$dx dy dz = r^2 \sin \theta dr d\varphi d\theta = -r dr d\varphi d\cos \theta.$$

Als eine einfache Anwendung berechnen wir das Volumen einer Kugel:

$$V = \int_0^R r^2 dr \int_0^{2\pi} d\varphi \int_0^\pi \sin \theta d\theta = \int_0^R r^2 dr \int_0^{2\pi} d\varphi \int_{-1}^1 d\cos \theta = \frac{4\pi}{3} R^3.$$

## 4.9 Kurvenintegrale

Eine Kurve in  $\mathbb{R}^3$  ist definiert durch eine Vektorfunktion  $\vec{r}(t) : \mathbb{R} \rightarrow \mathbb{R}^3$  eines Parameters  $t \in \mathbb{R}$  oder äquivalenterweise mittels drei sogenannter parametrischer Gleichungen  $x(t)$ ,  $y(t)$  und  $z(t)$  mit  $\vec{r} = (x, y, z)$ . Betrachten wir zwei nah aneinanderliegende Punkte  $\vec{r}(t)$  und  $\vec{r}(t + \Delta t)$  auf der Kurve, dann gilt

$$\Delta \vec{r} = \vec{r}(t + \Delta t) - \vec{r}(t) = \frac{d\vec{r}}{dt} \Delta t + O(\Delta t^2).$$

Im Limes  $\Delta t \rightarrow 0$  ist der Vektor

$$\frac{d\vec{r}}{dt} = \dot{\vec{r}}(t) = (\dot{x}(t), \dot{y}(t), \dot{z}(t))$$

tangent an der Kurve am Punkt  $\vec{r}(t)$ . Das Bogenelement  $ds$  entlang der Kurve ist gleich dem Betrag des Vektors

$$d\vec{r} = \dot{\vec{r}}(t) dt,$$

d.h.,

$$ds = \sqrt{\dot{x}(t)^2 + \dot{y}(t)^2 + \dot{z}(t)^2} dt = \sqrt{dx^2 + dy^2 + dz^2} \quad (4.7)$$

und der tangente Einheitsvektor

$$\hat{s} = \frac{\dot{\vec{r}}(t)}{|\dot{\vec{r}}(t)|} = \frac{(\dot{x}, \dot{y}, \dot{z})}{\sqrt{\dot{x}^2 + \dot{y}^2 + \dot{z}^2}}.$$

Ein Kurvenintegral ist der Limes der Riemannschen Summen auf der Kurve  $\mathcal{C}$ :

$$\int_{\mathcal{C}} \varphi(\vec{r}) ds = \lim_{n \rightarrow \infty} \sum_{i=1}^{\infty} \varphi(\vec{r}_i) \Delta s_i,$$

wobei  $\varphi(\vec{r}) : \mathbb{R}^3 \rightarrow \mathbb{R}$  die zu integrierende Funktion und  $\Delta s_i$  die Längenelemente auf  $\mathcal{C}$  sind. Wenn der Anfangs- und Endpunkt der Kurve  $\vec{r}_1 = \vec{r}(t_1)$  und  $\vec{r}_2 = \vec{r}(t_2)$  sind, und nachdem wir Gleichung (4.7) einsetzen, erhalten wir

$$\int_{\mathcal{C}(\vec{r}_1 \rightarrow \vec{r}_2)} \varphi(\vec{r}) ds = \int_{t_1}^{t_2} \varphi(x(t), y(t), z(t)) \sqrt{\dot{x}^2 + \dot{y}^2 + \dot{z}^2} dt.$$

Damit lassen sich Kurvenintegrale in der Form gewöhnlicher eindimensionaler Integrale berechnen.

Manchmal ist der Parameter  $t$  einer der Koordinaten, z.B. wenn die Kurve durch die Gleichungen  $y = y(x)$  und  $z = z(x)$  gegeben ist. In diesem Fall gilt

$$\int_{\mathcal{C}(\vec{r}_1 \rightarrow \vec{r}_2)} \varphi(\vec{r}) ds = \int_{x_1}^{x_2} \varphi(x, y, z) \sqrt{1 + \left(\frac{dy}{dx}\right)^2 + \left(\frac{dz}{dx}\right)^2} dx,$$

wobei  $\vec{r}_1 = (x_1, y(x_1), z(x_1))$  und  $\vec{r}_2 = (x_2, y(x_2), z(x_2))$ .

Beispiel: Die Länge der Helix  $x(t) = \cos(t)$ ,  $y(t) = \sin(t)$  und  $z(t) = kt$  ( $k \in \mathbb{R}$ ) zwischen  $\vec{r}_1 = (1, 0, 0)$  und  $\vec{r}_2 = \left(0, 1, \frac{k\pi}{2}\right)$  ist

$$\int_{\vec{r}_1}^{\vec{r}_2} ds = \int_0^{\frac{\pi}{2}} \sqrt{1 + k^2} dt = \frac{\pi}{2} \sqrt{1 + k^2}.$$

Ein für die Physik besonders wichtiges Kurvenintegral ist die *Zirkulation*  $\Gamma$  eines Vektorfeldes  $\vec{A}(\vec{r})$  entlang  $\mathcal{C}$ :

$$\Gamma = \int_{\mathcal{C}(\vec{r}_1 \rightarrow \vec{r}_2)} \vec{A} \cdot \hat{s} ds = \int_{\mathcal{C}} \vec{A} \cdot d\vec{r} = \int_{t_1}^{t_2} \vec{A} \cdot \frac{d\vec{r}}{dt} dt.$$

Wenn  $\vec{A} = \vec{\nabla}\varphi$  der Gradient eines Skalarfeldes oder Potentials  $\varphi$  mit stetigen partiellen Ableitungen ist, gilt

$$\begin{aligned} \Gamma &= \int_{\mathcal{C}(\vec{r}_1 \rightarrow \vec{r}_2)} \vec{A} \cdot d\vec{r} = \int_{t_1}^{t_2} \left( \frac{\partial\varphi}{\partial x} \dot{x} + \frac{\partial\varphi}{\partial y} \dot{y} + \frac{\partial\varphi}{\partial z} \dot{z} \right) dt \\ &= \int_{t_1}^{t_2} \frac{d\varphi(\vec{r}(t))}{dt} dt = \varphi(\vec{r}(t_2)) - \varphi(\vec{r}(t_1)) \\ &= \varphi(\vec{r}_2) - \varphi(\vec{r}_1). \end{aligned} \tag{4.8}$$

Das bedeutet, dass die Zirkulation  $\Gamma$  gleich dem Unterschied im Skalarpotential am End- und Anfangspunkt von  $\mathcal{C}$  ist, also unabhängig der Form und Zwischenpunkte der Kurve  $\mathcal{C}$ .  $\Gamma$  ist also null auf jede geschlossene Kurve.

Die reziproke Aussage ist ebenfalls wahr. Sei  $\vec{A}(\vec{r})$  ein Vektorfeld, dessen Komponenten  $A_x(\vec{r})$ ,  $A_y(\vec{r})$  und  $A_z(\vec{r})$  stetige Funktionen von  $\vec{r}$  sind, und bei dem die Zirkulation über jede geschlossene Kurve  $\mathcal{C}$  null ist, d.h.,

$$\int_{\mathcal{C}} \vec{A} \cdot d\vec{r} = 0 \quad \forall \mathcal{C} \text{ geschlossen.}$$

Wir können dann die Funktion  $\varphi : \mathbb{R}^3 \rightarrow \mathbb{R}$  als

$$\varphi(\vec{r}) = \int_{\vec{r}_0}^{\vec{r}} \vec{A}(\vec{r}') \cdot d\vec{r}'$$

definieren, wobei  $\vec{r}_0$  ein fester Punkt ist (z.B.  $\vec{r}_0 =$  Koordinatenursprung).  $\varphi(\vec{r})$  ist wohldefiniert, da  $\int_{\vec{r}_0}^{\vec{r}} \vec{A}(\vec{r}') d\vec{r}'$  unabhängig vom Weg ist. Man kann leicht zeigen, dass  $\varphi(\vec{r})$  bis

auf eine irrelevante Konstante das Skalarpotential von  $\vec{A}$  ist, d.h.,  $\vec{\nabla}\varphi = \vec{A}$ . Für  $h \rightarrow 0$  gilt nämlich

$$\varphi(\vec{r} + h \hat{e}_i) - \varphi(\vec{r}) = \int_{\vec{r}}^{\vec{r}+h\hat{e}_i} \vec{A}(\vec{r}') \cdot d\vec{r}' = \int_0^h \vec{A}(\vec{r} + h' \hat{e}_i) \cdot \hat{e}_i dh' = A_i(\vec{r}) h + O(h^2).$$

Daraus folgt

$$\lim_{h \rightarrow 0} \frac{\varphi(\vec{r} + h \hat{e}_i) - \varphi(\vec{r})}{h} = \frac{\partial \varphi}{\partial x_i} = A_i \quad \text{für } i = 1, 2 \text{ und } 3$$

oder

$$\vec{\nabla}\varphi = \vec{A} \quad \forall \vec{r}.$$

Damit haben wir das folgende wichtige Theorem gezeigt: Die nötige und ausreichende Bedingung dafür, dass die Zirkulation eines Vektorfeldes  $\vec{A} : \mathbb{R}^3 \rightarrow \mathbb{R}^3$  entlang allen geschlossenen Kurven in einer Domäne  $D \subseteq \mathbb{R}^3$  null ergibt, ist, dass  $\vec{A}$  der Gradient eines Skalarfeldes  $\varphi$  mit stetigen partiellen Ableitungen in  $D$  sei. In dem Fall ist natürlich die Zirkulation zwischen zwei beliebigen Punkten unabhängig des Weges und durch Gleichung (4.8) gegeben.

Bemerkung: Die Bedingung, dass  $\varphi$  stetige Ableitungen in  $D$  haben soll, ist *wichtig*. Betrachten wir z. B.  $\varphi(x, y, z) = \arctan\left(\frac{y}{x}\right)$ , dessen Gradient  $\vec{\nabla}\varphi = \frac{(-y, x, 0)}{x^2 + y^2} = \vec{A}$  auf der  $z$ -Achse ( $x = y = 0$ ) nicht definiert ist. Die Zirkulation von  $A$  entlang einer geschlossenen Kurve  $\mathcal{C}$  ist null, nur wenn die Kurve die  $z$ -Achse nicht umdreht. Zum Beispiel haben wir für einen Kreis in der  $xy$ -Ebene

$$\int_{\mathcal{C}} \vec{A} \cdot d\vec{r} = \int_0^{2\pi} \frac{(-\varrho \sin \varphi, \varrho \cos \varphi, 0) \cdot (-\varrho \sin \varphi, \varrho \cos \varphi, 0)}{\varrho^2} d\varphi = \int_0^{2\pi} d\varphi = 2\pi,$$

wobei wir den Kreis wie üblich als  $\vec{r}(\varphi) = \varrho(\cos \varphi, \sin \varphi, 0)$  parametrisiert haben.

## 4.10 Oberflächen

Oberflächen sind definiert durch Parametergleichungen, die Funktionen zweier Variablen  $u$  und  $v$  sind:

$$x = x(u, v), \quad y = y(u, v) \quad \text{und} \quad z = z(u, v),$$

oder in Vektorform

$$\vec{r} = \vec{r}(u, v) = x(u, v) \hat{e}_1 + y(u, v) \hat{e}_2 + z(u, v) \hat{e}_3.$$

Die Kurven  $\vec{r} = \vec{r}(u, v_0)$  als Funktion von  $u$  bei festem  $v_0$  und  $\vec{r} = \vec{r}(u_0, v)$  als Funktion von  $v$  bei festem  $u_0$  werden parametrische Kurven der Oberfläche genannt. Sie definieren ein System krummliniger Koordinaten auf der Oberfläche. Die Vektoren

$$\vec{t}_u = \frac{\partial \vec{r}}{\partial u}(u, v)$$

und

$$\vec{t}_v = \frac{\partial \vec{r}}{\partial v}(u, v)$$

sind tangent an den parametrischen Kurven und an der Oberfläche am Punkt  $\vec{r}(u, v)$ . Der Einheitsvektor  $\hat{n}(u, v)$ , der senkrecht zur Oberfläche am Punkt  $\vec{r}(u, v)$  steht, ist durch das Vektorprodukt

$$\hat{n} = \frac{\vec{t}_u \times \vec{t}_v}{|\vec{t}_u \times \vec{t}_v|}$$

gegeben. Ob  $\vec{t}_u \times \vec{t}_v$  oder  $\vec{t}_v \times \vec{t}_u$  die passende senkrechte Richtung zeigt (z.B. die Richtung nach außen eines eingeschlossenen Volumens), hängt natürlich von der Wahl der Reihenfolge der Parameter ab. Die Norm von  $\vec{t}_u \times \vec{t}_v$  lässt sich mit Hilfe der Lagrange'schen Identität

$$(\vec{A} \times \vec{B}) \cdot (\vec{C} \times \vec{D}) = (\vec{A} \cdot \vec{C})(\vec{B} \cdot \vec{D}) - (\vec{B} \cdot \vec{C})(\vec{A} \cdot \vec{D})$$

als

$$|\vec{t}_u \times \vec{t}_v| = \sqrt{|\vec{t}_u|^2 |\vec{t}_v|^2 - (\vec{t}_u \cdot \vec{t}_v)^2} = \sqrt{EG - F^2},$$

wobei wir die sogenannten *fundamentalen Koeffizienten* der Oberfläche

$$E = \vec{t}_u \cdot \vec{t}_u, \quad F = \vec{t}_u \cdot \vec{t}_v \quad \text{und} \quad G = \vec{t}_v \cdot \vec{t}_v$$

eingeführt haben.

Um Oberflächenintegrale zu berechnen, brauchen wir die Beziehung zwischen dem Flächenelement  $ds$  an der Oberfläche und dem doppelten Differenzial  $du dv$  im Parameterraum  $(u, v)$ . Dazu betrachten wir die infinitesimalen Verschiebungen  $d\vec{r}_u$  und  $d\vec{r}_v$  des Positionsvektors  $\vec{r}(u, v)$  auf der Oberfläche entlang der parametrischen Kurven  $\vec{r} = \vec{r}(u, v_0)$  und  $\vec{r} = \vec{r}(u_0, v)$ :

$$d\vec{r}_u = \vec{r}(u + du, v) - \vec{r}(u, v) = \vec{t}_u du = \frac{\partial \vec{r}}{\partial u} du$$

und

$$d\vec{r}_v = \vec{r}(u, v + dv) - \vec{r}(u, v) = \vec{t}_v dv = \frac{\partial \vec{r}}{\partial v} dv.$$

Das Flächenelement  $dS$  ist gleich dem Betrag des Vektorprodukts

$$d\vec{r}_u \times d\vec{r}_v = (\vec{t}_u \times \vec{t}_v) du dv = \frac{\partial \vec{r}}{\partial u} \times \frac{\partial \vec{r}}{\partial v} du dv,$$

d.h.,

$$dS = |\vec{t}_u \times \vec{t}_v| du dv = \sqrt{EG - F^2} du dv.$$

Mit Hilfe des normalen Einheitsvektors  $\hat{n}$  können wir ein orientiertes Flächendifferential definieren als

$$d\vec{S} = \hat{n} dS = (\vec{t}_u \times \vec{t}_v) du dv = \frac{\partial \vec{r}}{\partial u} \times \frac{\partial \vec{r}}{\partial v} du dv.$$

Beispiele:

- i) Die Oberfläche einer Kugel mit Radius  $r$  ist durch folgende Gleichungen charakterisiert:

$$\vec{r}(\varphi, \theta) = r(\sin \theta \cos \varphi, \sin \theta \sin \varphi, \cos \theta)$$

$$\vec{t}_\varphi = \frac{\partial \vec{r}}{\partial \varphi} = r(-\sin \theta \sin \varphi, \sin \theta \cos \varphi, 0)$$

$$\vec{t}_\theta = \frac{\partial \vec{r}}{\partial \theta} = r(\cos \theta \cos \varphi, \cos \theta \sin \varphi, -\sin \theta)$$

$$E = \vec{t}_\varphi \cdot \vec{t}_\varphi = r^2 \sin^2 \theta$$

$$F = \vec{t}_\varphi \cdot \vec{t}_\theta = 0$$

$$G = \vec{t}_\theta \cdot \vec{t}_\theta = r^2$$

$$ds = \sqrt{EG - F^2} d\varphi d\theta = r^2 \sin \theta d\varphi d\theta$$

$$\vec{t}_\varphi \times \vec{t}_\theta = r^2 (-\sin^2 \theta \cos \varphi, -\sin^2 \theta \sin \varphi, -\sin \theta \cos \theta) = -r^2 \sin \theta \frac{\vec{r}}{r}$$

$$\hat{n} = \frac{\vec{t}_\varphi \times \vec{t}_\theta}{\sqrt{EG - F^2}} = -\frac{\vec{r}}{r}.$$

Wie erwartet, ist  $\hat{n} = \frac{\vec{r}}{r}$  die nach außen zeigende Normale.

- ii) Die Oberfläche eines Zylinders ist parametrisiert als

$$\vec{r}(\varphi, z) = (\varrho \cos \varphi, \varrho \sin \varphi, z)$$

$$\vec{t}_\varphi = \varrho (-\sin \varphi, \cos \varphi, 0)$$

$$\vec{t}_z = (0, 0, 1)$$

$$E = \varrho^2, \quad F = 0 \quad \text{und} \quad G = 1$$

$$ds = \sqrt{EG - F^2} d\varphi dz = \varrho d\varphi dz$$

$$\vec{t}_\varphi \times \vec{t}_z = \varrho (\cos \varphi, \sin \varphi, 0)$$

$$\hat{n} = \frac{\vec{t}_\varphi \times \vec{t}_z}{\sqrt{EG - F^2}} = (\cos \varphi, \sin \varphi, 0).$$

#### 4.11 Oberflächenintegrale

Wir betrachten eine Oberfläche  $S$ , die durch die Vektorgleichung

$$\vec{r}(u, v) = (x(u, v), y(u, v), z(u, v))$$

gegeben ist. Das Oberflächenintegral einer Funktion  $\varphi(\vec{r}) = \varphi(x, y, z)$  auf  $S$  oder einem Teil davon ist durch folgenden Ausdruck gegeben

$$\int_S \varphi(\vec{r}) dS = \int_{u_1}^{u_2} du \int_{v_1}^{v_2} dv \varphi(\vec{r}(u, v)) \sqrt{EG - F^2},$$

wobei  $u_1 \leq u \leq u_2$  und  $v_1 \leq v \leq v_2$  die Parameterintervalle sind, die die Oberfläche definieren.

Beispiel: Die Fläche einer Kugel mit Radius  $r$  ist

$$\int_{\text{Kugel}} dS = \int_0^{2\pi} d\varphi \int_0^\pi d\theta r^2 \sin \theta = 4\pi r^2.$$

Ein für die Physik besonders wichtiges Oberflächenintegral ist der Fluss  $\Phi$  eines Vektorfeldes durch eine Oberfläche  $S$ , der wie folgt definiert ist:

$$\Phi = \int_S \vec{A} \cdot d\vec{S} = \int_S \vec{A} \cdot \hat{n} dS.$$

Beispiel: Sei  $\vec{E} = q \frac{\vec{r}}{r^3}$  das elektrische Feld, das durch eine Ladung  $q$  am Ursprung erzeugt wird. Der Fluss von  $\vec{E}$  durch eine am Ursprung zentrierte Kugel mit Radius  $r$  ist

$$\Phi = \int_{\text{Kugel}} \vec{E} \cdot d\vec{S} = q r^2 \int_0^{2\pi} d\varphi \int_0^\pi \sin \theta d\theta \frac{\vec{r}}{r^3} \cdot \frac{\vec{r}}{r} = 4\pi q.$$

Im Fall des Coulombfeldes ist das Resultat unabhängig von  $r$ .

## 4.12 Gauss'sches Theorem

Wir betrachten ein Vektorfeld  $\vec{E} : \mathbb{R}^3 \rightarrow \mathbb{R}^3$  mit stetigen partiellen Ableitungen  $\frac{\partial E_i}{\partial x_j}$  auf ein Volumen  $V$  und auf die Oberfläche  $S(V)$ , die das Volumen abgrenzt. Unter der Annahme, dass die Oberfläche sich als eine finite Anzahl von Teilen darstellen lässt, bei denen der normale Einheitsvektor kontinuierlich variiert, gilt folgende Integralbeziehung:

$$\int_{S(V)} \vec{E} \cdot d\vec{S} = \int_V \vec{\nabla} \cdot \vec{E} dx dy dz.$$

Dies besagt, dass der Fluss eines Vektorfeldes  $\vec{E}$  durch eine geschlossene Oberfläche  $S$  gleich dem Integral der Divergenz  $\vec{\nabla} \cdot \vec{E}$  über das von der Oberfläche  $S$  abgegrenzte Volumen ist.

Das Gauss'sche Theorem stellt in Integralform dar, was wir bei der Interpretation der Divergenz in Differentialform bereits bewiesen haben.

## 4.13 Stokes'sches Theorem

Wir betrachten eine geschlossene Kurve  $\mathcal{C}$ , die aus einer finiten Anzahl von Bögen besteht, bei denen die Tangente stetig ist, und eine orientierbare Oberfläche  $S$ , deren Rand  $\mathcal{C}$  ist. Die Oberfläche  $S$  besteht auch aus einer finiten Anzahl von Teilen, bei denen der

senkrechte Einheitsvektor stetig ist, und ist durch eine Parametrisierung gegeben, die stetige partielle Ableitungen hat. Wenn  $\vec{A} : \mathbb{R}^3 \rightarrow \mathbb{R}^3$  ein Vektorfeld mit stetigen partiellen Ableitungen auf  $S$  und  $\mathcal{C}$  ist, gilt

$$\int_{\mathcal{C}} \vec{A} \cdot d\vec{r} = \int_{S(\mathcal{C})} (\vec{\nabla} \times \vec{A}) \cdot d\vec{S}.$$

Anders ausgedrückt ist die Zirkulation eines Vektorfeldes  $\vec{A}$  entlang einer Kurve  $\mathcal{C}$  gleich dem Fluss von  $\vec{\nabla} \times \vec{A}$  durch eine Oberfläche  $S(\mathcal{C})$ , deren Rand  $\mathcal{C}$  ist.

Das Stokes'sche Theorem stellt in Integralform dar, was wir bei der Interpretation der Rotation in Differentialform bereits bewiesen haben.

## 5 Differentialgleichungen

### 5.1 Einleitung: Grundbegriffe und Klassifikation

Eine *Differentialgleichung* ist eine Beziehung zwischen einer Funktion und einigen ihrer Ableitungen. Beispiele von Differentialgleichungen sind

$$y' = \alpha y \quad (5.1)$$

oder

$$y^2 + (y')^2 - \alpha = 0 \quad (5.2)$$

für  $y : \mathbb{R} \rightarrow \mathbb{R}$  und  $\alpha \in \mathbb{R}$ ,

$$m \ddot{x} + k x = 0 \quad (5.3)$$

für  $x = x(t)$  und  $m, k \in \mathbb{R} > 0$ , und

$$\frac{\partial^2 \varphi}{\partial x^2} + \frac{\partial^2 \varphi}{\partial y^2} + \frac{\partial^2 \varphi}{\partial z^2} = -4\pi \varrho(x, y, z) \quad (5.4)$$

für das Potential  $\varphi : \mathbb{R}^3 \rightarrow \mathbb{R}$  und die Ladungsdichte  $\varrho : \mathbb{R}^3 \rightarrow \mathbb{R}$ .

Man unterscheidet zwischen *gewöhnlichen* und *partiellen* Differentialgleichungen. Bei einer gewöhnlichen Differentialgleichung treten nur Ableitungen nach einer einzigen Variable auf [z.B. (5.1)-(5.3)], während bei partiellen Differentialgleichungen partielle Ableitungen nach mehreren Variablen auftreten [z.B. (5.4)].

Eine Lösung einer Differentialgleichung ist eine hinreichend oft differenzierbare Funktion, die in einem gewissen Definitionsbereich die Gleichung erfüllt. Zum Beispiel ist die Funktion  $y(x) = a e^{\alpha x}$  bekanntlich eine Lösung der Gleichung (5.1) für beliebiges  $a \in \mathbb{R}$ , und

$x(t) = a \cos\left(\sqrt{\frac{k}{m}} t\right) + b \sin\left(\sqrt{\frac{k}{m}} t\right)$  ist eine Lösung von (5.3) für beliebige  $a, b \in \mathbb{R}$ .

Im Falle der Gleichung (5.2) sind  $y(x) = \pm\sqrt{\alpha} \cos x$  und  $y(x) = \pm\sqrt{\alpha} \sin x$  Lösungen für  $\alpha > 0$ . Für  $\alpha = 0$  gibt es nur die Lösung  $y = 0$ , und für  $\alpha < 0$  gibt es gar keine Lösung  $y : \mathbb{R} \rightarrow \mathbb{R}$ . Wir sehen, dass es manchmal unendlich viele Lösungen gibt, manchmal ein paar oder eine einzige und manchmal überhaupt keine.

Die Differentialgleichungen von physikalischem Interesse haben in der Regel unendlich viele allgemeine Lösungen [z.B. Gleichung (5.3)]. Die spezifische eindeutige Lösung eines gegebenen Problems wird durch zusätzliche Bedingungen festgelegt. Das sind die Anfangsbedingungen in einem zeitabhängigen Problem (z.B. die Position und Geschwindigkeit eines Teilchens am Zeitpunkt  $t = t_0$  bei der Newton-Gleichung) oder die Randbedingungen bei Funktionen der Koordinaten (z.B. die Spannung zwischen den metallischen Platten eines Kondensators bei der Berechnung des elektrostatischen Potentials  $\varphi$  in einem ladungsfreien Gebiet, wo  $\nabla^2 \varphi = 0$ ). Betrachten wir beispielsweise die allgemeine Lösung des Oszillators  $x(t) = a \cos(\omega t) + b \sin(\omega t)$  mit  $\omega = \sqrt{k/m}$  und geben wir die Anfangsbedingung  $x(0) = x_0$  und  $\dot{x}(0) = 0$ . Dann ist die Lösung eindeutig  $x(t) = x_0 \cos(\omega t)$ .

Die *Ordnung* einer Differentialgleichung ist der Grad der höchsten Ableitung, die in der Gleichung vorkommt. So sind die Gleichungen (5.1) und (5.2) erster Ordnung und die Gleichungen (5.3) und (5.4) zweiter Ordnung.

Die gewöhnlichen Differentialgleichungen  $n$ -ter Ordnung lassen sich im Allgemeinen in einer impliziten Form

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

oder, wenn sich die höchste Ableitung  $y^{(n)}$  auflösen lässt, in der expliziten Form

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

darstellen. Explizite und implizite Formen findet man auch bei partiellen Differentialgleichungen. Die Gleichungen (5.1) und (5.4) sind explizit, während (5.2) und (5.3) implizit sind.

Man unterscheidet auch zwischen *linearen* und *nichtlinearen* Differentialgleichungen. Treten die unbekannt Funktionen  $y(x)$  oder  $\varphi(x_1, x_2, \dots, x_n)$  und ihre Ableitungen nur in der ersten Potenz ohne Produkte der Form  $y y'$  oder  $\frac{\partial \varphi}{\partial x_i} \frac{\partial \varphi}{\partial x_j}$  auf, so heißt die Gleichung linear.

In den anderen Fällen spricht man von nichtlinearen Differentialgleichungen. Beispielsweise sind die Gleichungen (5.1), (5.3) und (5.4), aber auch der erzwungene Oszillator  $m \ddot{x} + k x = F(t)$  mit einer äußeren Kraft  $F(t)$ , lineare Gleichungen. Im Gegensatz dazu ist Gleichung (5.2) nichtlinear. Lineare Gleichungen lassen sich in der Form

$$\mathcal{L}(x, y, y', \dots, y^{(n)}) = \hat{\mathcal{L}}(y) = g(x) \quad (5.5)$$

oder

$$\mathcal{L}\left(x_1, \dots, x_n, \varphi, \frac{\partial \varphi}{\partial x_1}, \dots, \frac{\partial \varphi}{\partial x_n}, \dots, \frac{\partial^n \varphi}{\partial x_1, \dots, \partial x_n}\right) = \hat{\mathcal{L}}(\varphi) = g(x_1, \dots, x_n) \quad (5.6)$$

schreiben, wobei  $\hat{\mathcal{L}}$  eine kompakte Schreibweise für den linearen Operator ist, der auf die Funktion  $y = y(x)$  bzw.  $\varphi = \varphi(x_1, \dots, x_n)$  wirkt. Das bedeutet, dass für beliebige hinreichend differenzierbare Funktionen  $y_1, y_2 : \mathbb{R} \rightarrow \mathbb{R}$  und  $\varphi_1, \varphi_2 : \mathbb{R}^n \rightarrow \mathbb{R}$  und für beliebige Koeffizienten  $\alpha, \beta \in \mathbb{R}$  folgende grundsätzliche Beziehungen gelten:

$$\hat{\mathcal{L}}(\alpha y_1 + \beta y_2) = \alpha \hat{\mathcal{L}}(y_1) + \beta \hat{\mathcal{L}}(y_2)$$

und

$$\hat{\mathcal{L}}(\alpha \varphi_1 + \beta \varphi_2) = \alpha \hat{\mathcal{L}}(\varphi_1) + \beta \hat{\mathcal{L}}(\varphi_2).$$

Die Terme  $g(x)$  in Gleichung (5.5) und  $g(x_1, \dots, x_n)$  in Gleichung (5.6), bei denen weder die Funktion noch ihre Ableitungen auftreten, werden Inhomogenitäten genannt.

Die Gleichungen, bei denen dieser Term ausfällt ( $g = 0$ ), werden als *homogene Differentialgleichungen* bezeichnet. In den anderen Fällen heißt die Gleichung *inhomogen*. Die linearen Gleichungen (5.1) und (5.3) sind homogen, während Gleichung (5.4) oder der erzwungene Oszillator  $m \ddot{x} + k x = F(t)$  inhomogen sind.

Zu jeder inhomogenen linearen Differentialgleichung

$$\hat{\mathcal{L}}(y) = g(x) \quad (5.7)$$

gehört eine homogene lineare Gleichung der Form

$$\hat{\mathcal{L}}(y) = 0, \quad (5.8)$$

die man dadurch erhält, dass man den Term  $g(x)$ , bei dem die Funktion und ihre Ableitungen nicht auftreten, gleich null setzt ( $g(x) = 0$ ). Diese einfachere homogene Gleichung (5.8) spielt eine sehr wichtige Rolle bei der Lösung der ursprünglichen inhomogenen Gleichung (5.7). Näheres dazu werden wir in den kommenden Abschnitten besprechen.

Neben den Differentialgleichungen mit einer unbekanntem Funktion  $y(x)$  finden *Systeme von gekoppelten Differentialgleichungen* eine Reihe von physikalisch relevanten Anwendungen. Diese können gewöhnliche Differentialgleichungen wie bei der klassischen Bewegung von Punktteilchen in  $\mathbb{R}^3$  oder partielle Differentialgleichungen wie bei der Beschreibung von zeitabhängigen elektrischen und magnetischen Feldern  $\vec{E}(x, y, z, t)$  und  $\vec{B}(x, y, z, t)$  sein (Newton- und Maxwell-Gleichungen).

Beispiele für Systeme von Differentialgleichungen sind die Newton'schen Gleichungen für  $n$  geladene Teilchen im Vakuum

$$m \ddot{\vec{r}}_i = \sum_{j \neq i} \frac{q_i q_j}{|\vec{r}_i - \vec{r}_j|} \quad i = 1, \dots, n$$

oder für ein System von zwei gekoppelten Oszillatoren

$$m_1 \ddot{x}_1 = -k_1 x_1 - k_3(x_1 - x_2)$$

$$m_2 \ddot{x}_2 = -k_2 x_2 - k_3(x_2 - x_1).$$

Dabei wurden  $x_1$  und  $x_2$  bezüglich der Gleichgewichtspositionen  $x_1^0$  und  $x_2^0$  der Feder  $k_1$  und  $k_2$  gemessen. Die Gleichgewichtslänge von  $k_3$  ist hier  $l_3 = x_2^0 - x_1^0$ .

## 5.2 Beispiele gewöhnlicher Differentialgleichungen

### 5.2.1 Wachstum einer Population

Wir betrachten eine Population mit  $N(t)$  Individuen und interessieren uns für die Zeitabhängigkeit von  $N(t)$ . Die Änderung von  $N(t)$  kann in erster Näherung als proportional zur Anzahl der Individuen betrachtet werden:

$$\Delta N(t) = N(t + \Delta t) - N(t) = \alpha N(t) \Delta t.$$

Der Koeffizient  $\alpha$  ist die Vermehrungsrate ( $\alpha > 0$ ) oder Zerfallsrate ( $\alpha < 0$ ), die die Zahl der Vermehrungen bzw. Zerfälle pro Mitglied der Population und pro Zeiteinheit angibt. Im Allgemeinen hängt  $\alpha$  von  $N$  und eventuell auch  $t$  ab. Allerdings kann  $\alpha$ , wenn die Dichte der Population niedrig ist und es ausreichende Ressourcen gibt, als eine Konstante angenommen werden.

Im Limes  $\Delta t \rightarrow 0$  erhalten wir die Gleichung

$$\frac{dN}{dt} = \alpha N,$$

die die bekannte allgemeine Lösung

$$N(t) = A e^{\alpha t}$$

besitzt. Die Konstante  $A$  wird durch die Anfangsbedingung  $N(t_0) = N_0$  festgelegt. Die eindeutige, der Anfangsbedingung angepasste Lösung lautet

$$N(t) = N_0 e^{\alpha(t-t_0)}.$$

Diese beschreibt ein exponentielles Wachstum für  $\alpha > 0$  oder einen exponentiellen Zerfall, wie beim radioaktiven Zerfall von Kernen, wenn  $\alpha < 0$  ist.

Die Annahme, dass die Wachstumsrate  $\alpha > 0$  unabhängig von  $N(t)$  ist, kann bei einem endlichen System mit endlichen Ressourcen nicht beliebig lange gelten. Eine realistischere Beschreibung wäre es, anzunehmen, dass  $\alpha$  mit zunehmendem  $N$  abnimmt und sogar negativ wird, wenn  $N$  eine kritische Größe  $N_c$  überschreitet. Wir schreiben also  $\alpha(N) = \alpha_0(1 - N/N_c)$ , und die Differentialgleichung lautet

$$\frac{dN}{dt} = \alpha_0 \left(1 - \frac{N}{N_c}\right) N.$$

Diese Gleichung kann mit der Methode der Trennung von Variablen gelöst werden:

$$\begin{aligned} \frac{dN}{\left(1 - \frac{N}{N_c}\right) N} &= \alpha_0 dt \\ \int_{N_0}^N \frac{dN'}{\left(1 - \frac{N'}{N_c}\right) N'} &= \alpha_0 \int_{t_0}^t dt' = \alpha_0(t - t_0). \end{aligned}$$

Das Integral auf der linken Seite hat die Form  $\int \frac{dx}{(1-x)x}$  mit  $x = \frac{N'}{N_c} < 1$  und kann mit der Substitution  $u = \frac{1}{x}$  ( $du = -dx/x^2$ ,  $u > 1$ ) gelöst werden:

$$\int \frac{dx}{(1-x)x} = \int \frac{dx}{\left(\frac{1}{x} - 1\right) x^2} = - \int \frac{du}{(u-1)} = - \ln(u-1) = - \ln\left(\frac{1}{x} - 1\right) = \ln\left(\frac{x}{1-x}\right).$$

Wir erhalten dann

$$\begin{aligned} \ln \frac{x}{1-x} - \ln \frac{x_0}{1-x_0} &= \alpha_0 t \\ \frac{x}{1-x} &= \frac{x_0}{1-x_0} e^{\alpha_0 t} = c e^{\alpha_0 t} \\ x(1 + c e^{\alpha_0 t}) &= c e^{\alpha_0 t} \\ x(t) &= \frac{N(t)}{N_c} = \frac{c e^{\alpha_0 t}}{1 + c e^{\alpha_0 t}}. \end{aligned}$$

Die Lösung  $x(t) = N(t)/N_c$  zeigt das erwartete Sättigungsverhalten mit  $x(t) \rightarrow 1$  für  $t \rightarrow +\infty$ .

### 5.2.2 Entladung eines Kondensators

$$\begin{aligned} I &= \frac{dQ}{dt} \\ IR &= -\frac{Q}{C} \\ \frac{dQ}{dt} &= -\frac{1}{RC} Q \quad \Rightarrow \quad Q(t) = Q_0 e^{-\frac{t}{RC}}. \end{aligned}$$

### 5.2.3 Freier Fall

Die Newtonsche Gleichung

$$m \ddot{z}(t) = -m g$$

lässt sich sofort integrieren:

$$\begin{aligned} \dot{z}(t) &= -g t + v_0 \\ z(t) &= -\frac{1}{2} g t^2 + v_0 t + z_0. \end{aligned}$$

Wenn wir den Luftwiderstand berücksichtigen wollen, führen wir eine Reibungskraft  $\vec{F}_r = -\alpha \vec{v}$  mit  $\alpha > 0$  ein, die proportional zur Geschwindigkeit angenommen wird:

$$\ddot{z}(t) = -g - \frac{\alpha}{m} \dot{z}(t).$$

Wenn wir die Substitution  $\dot{z}(t) = v(t)$  einführen, erhalten wir die inhomogene lineare Gleichung erster Ordnung

$$\dot{v}(t) + \frac{\alpha}{m} v(t) = -g. \quad (5.9)$$

Eine spezielle Lösung  $v_p$  dieser Gleichung erhalten wir, indem wir einen zeitunabhängigen Ansatz machen. Die konstante Speziallösung lautet

$$\frac{\alpha v_p}{m} = -g \quad \Leftrightarrow \quad v_p = -\frac{m g}{\alpha}.$$

Diese Speziallösung kann nur in ganz besonderen Fällen der Bewegung einer fallenden Masse entsprechen. Trotzdem ist sie sehr hilfreich, um das allgemeine Problem zu lösen, aufgrund folgender Eigenschaft linearer Gleichungen:

Sind zwei Funktionen  $v_1(t)$  und  $v_2(t)$  Lösungen einer inhomogenen Differentialgleichung  $\hat{\mathcal{L}}(v) = g$ , dann ist der Unterschied  $v_2(t) - v_1(t)$  eine Lösung der assoziierten homogenen Gleichung  $\hat{\mathcal{L}}(v) = 0$ .

Demzufolge lässt sich die allgemeine Lösung der inhomogenen Gleichung (5.9) als

$$v(t) = v_p + v_h(t)$$

schreiben, wobei  $v_p$  eine Speziallösung der inhomogenen Gleichung ist und  $v_h(t)$  eine allgemeine Lösung der homogenen Gleichung

$$\dot{v}(t) + \frac{\alpha}{m} v(t) = 0$$

ist. Die homogene Gleichung ist uns bekannt, und die allgemeine Lösung lautet

$$v_h(t) = A e^{-\frac{\alpha}{m} t}.$$

Damit ist die allgemeine Lösung der Gleichung (5.9)

$$v(t) = -\frac{m g}{\alpha} + A e^{-\frac{\alpha}{m} t}.$$

Mit Hilfe der Anfangsbedingung  $v(0) = v_0$  definieren wir die eindeutige Lösung

$$v(t) = \left( v_0 + \frac{m g}{\alpha} \right) e^{-\frac{\alpha}{m} t} - \frac{m g}{\alpha}.$$

Die Speziallösung entspricht der Grenzggeschwindigkeit  $v(\infty) = -\frac{m g}{\alpha}$ .

Für die Position erhalten wir

$$z(t) = z_0 - \frac{m g}{\alpha} t + \frac{m}{\alpha} \left( v_0 + \frac{m g}{\alpha} \right) \left( 1 - e^{-\frac{\alpha}{m} t} \right).$$

Im Limes  $t \rightarrow \infty$  ist die Bewegung uniform, d.h.,  $z(t) \simeq z_0 + \frac{1}{2} \left( v_0 + \frac{m g}{\alpha} \right) t - \frac{m g}{\alpha} t$  für  $t \rightarrow \infty$ , mit der Grenzggeschwindigkeit  $\dot{z}(\infty) = -\frac{m g}{\alpha}$ .

### 5.3 Lineare Differentialgleichungen: Superpositionsprinzip

Lineare gewöhnliche und partielle Differentialgleichungen spielen eine sehr wichtige Rolle in der Physik. Zum Beispiel die Schrödinger-Gleichung

$$-\frac{\hbar^2}{2m} \nabla^2 \Psi + v(\vec{r}) \Psi = i \hbar \frac{\partial \Psi}{\partial t}$$

der nichtrelativistischen Quantentheorie ist eine homogene lineare partielle Differentialgleichung zweiter Ordnung ( $\nabla^2 = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}$ ) für die Wellenfunktion  $\Psi(x, y, z, t)$  eines Elektrons im Potential  $v(\vec{r})$ .

Die Maxwell-Gleichungen für elektrische und magnetische Felder  $\vec{E}$ ,  $\vec{B}$ ,  $\vec{D}$  und  $\vec{H}$ :

$$\vec{\nabla} \cdot \vec{D} = 4\pi \varrho, \quad \vec{\nabla} \times \vec{H} = 4\pi \vec{j} + \frac{\partial \vec{D}}{\partial t}, \quad \vec{\nabla} \times \vec{E} + \frac{\partial \vec{B}}{\partial t} = 0 \quad \text{und} \quad \vec{\nabla} \cdot \vec{B} = 0,$$

zusammen mit den Beziehungen  $\vec{D} = \varepsilon_0 \vec{E}$  und  $\vec{B} = \mu_0 \vec{H}$  im Vakuum bilden ein inhomogenes System partieller Differentialgleichungen erster Ordnung. Die Inhomogenitäten  $\varrho(\vec{r}, t)$  und  $\vec{j}(\vec{r}, t)$  sind die Quellen der elektrischen und magnetischen Felder, nämlich die Ladungsdichte  $\varrho(\vec{r}, t)$  und die Stromdichte  $\vec{j}(\vec{r}, t)$ .

Lineare Differentialgleichungen werden auch als eine erste nicht triviale Näherung zur Beschreibung von Systemen gekoppelter Teilchen nah an ihrer Gleichgewichtsposition benutzt. Damit können die Schwingungsmoden gekoppelter Oszillatoren berechnet werden,

solange die Abweichungen von der Gleichgewichtsposition klein genug sind, so dass die Kräfte proportional (d.h., linear) zu den Abweichungen sind. Zu dieser Kategorie gehören auch die Probleme des klassischen Elektromagnetismus, solange die Dielektrizitätskonstante  $\varepsilon$  und die magnetische Permeabilität  $\mu$  in den Materialgleichungen  $\vec{D} = \varepsilon \vec{E}$  und  $\vec{B} = \mu \vec{H}$  unabhängig von  $\vec{E}$  und  $\vec{H}$  sind. Ein wichtiges Beispiel dafür ist die Berechnung der Zeitabhängigkeit von Ladung und Strom in Stromkreisen, die aus Kondensatoren, Spulen und Resistoren bestehen.

Die zentrale Eigenschaft linearer Differentialgleichungen ist das sogenannte *Superpositionsprinzip* der homogenen linearen Gleichungen: Wenn  $\varphi_1$  und  $\varphi_2$  Lösungen der Gleichung  $\hat{\mathcal{L}}(\varphi) = 0$  sind, dann sind alle Superpositionen dieser beiden Lösungen,  $\varphi_1 + \varphi_2$  und im Allgemeinen  $\alpha \varphi_1 + \beta \varphi_2$  für beliebige  $\alpha, \beta \in \mathbb{C}$ , auch eine Lösung. Dies folgt sofort aus der Linearität von  $\hat{\mathcal{L}}$ , da

$$\hat{\mathcal{L}}(\alpha \varphi_1 + \beta \varphi_2) = \alpha \hat{\mathcal{L}}(\varphi_1) + \beta \hat{\mathcal{L}}(\varphi_2) = 0.$$

Insbesondere ist die Funktion  $\varphi \equiv 0$  immer eine Lösung der homogenen Gleichung. Dies bedeutet, dass die allgemeine Lösung einer homogenen Gleichung sich als *lineare Kombination*

$$\varphi = \sum_{i=1}^n \alpha_i \varphi_i$$

von  $n$  unabhängigen Lösungen schreiben lässt. Die konstanten Koeffizienten  $\alpha_i \in \mathbb{C}$  werden durch die Anfangs- oder Randbedingungen bestimmt. Eine homogene lineare gewöhnliche Differentialgleichung  $n$ -ter Ordnung hat im Allgemeinen  $n$  solche unabhängigen Lösungen  $\varphi_i$ . Jeder solche Satz von  $n$  unabhängigen Funktionen, mit dem man eine beliebige Lösung darstellen kann, wird als *fundamentales oder vollständiges System* bezeichnet.

Man sagt, dass  $n$  Funktionen oder Lösungen *linear unabhängig* sind, wenn keine dieser Funktionen sich als lineare Superposition der anderen  $n-1$  schreiben lässt. Genauer gesagt sind  $\{\varphi_1, \varphi_2, \dots, \varphi_n\}$  linear unabhängig ( $\varphi_i : \mathbb{R} \rightarrow \mathbb{R} \ \forall i$ ), wenn

$$\sum_i^n \alpha_i \varphi_i(x) = 0 \ \forall x \quad \Rightarrow \quad \alpha_i = 0 \ \forall i = 1, \dots, n.$$

Lineare inhomogene Systeme  $\hat{\mathcal{L}}(\varphi) = g$  haben folgende wichtige Eigenschaft: Die allgemeine Lösung einer inhomogenen linearen Gleichung hat die Form

$$\varphi = \varphi_p + \sum_{i=1}^n \alpha_i \varphi_i^h, \tag{5.10}$$

wobei  $\varphi_p$  eine spezielle Lösung der inhomogenen Gleichung ist (d.h.,  $\hat{\mathcal{L}}(\varphi_p) = g$ ) und die  $\varphi_i^h$  ( $i = 1, \dots, n$ ) ein vollständiger Satz von Lösungen der homogenen Gleichung sind ( $\alpha_i \in \mathbb{C}$ ). Man kann sich schnell davon überzeugen. Einerseits ist  $\varphi$  aus Gleichung (5.10) offensichtlich eine Lösung:

$$\hat{\mathcal{L}}(\varphi) = \hat{\mathcal{L}}(\varphi_p) + \sum_{i=1}^n \alpha_i \hat{\mathcal{L}}(\varphi_i) = \hat{\mathcal{L}}(\varphi_p) = g.$$

Andererseits gilt für eine beliebige Lösung  $\varphi$  der inhomogenen Gleichung

$$\hat{\mathcal{L}}(\varphi - \varphi_p) = \hat{\mathcal{L}}(\varphi) - \hat{\mathcal{L}}(\varphi_p) = g - g = 0.$$

Das bedeutet, dass  $\varphi - \varphi_p$  eine Lösung der homogenen Gleichung und damit in der Form

$$\varphi - \varphi_p = \sum_{i=1}^n \alpha_i \varphi_i^h$$

darstellbar ist.

Zusammenfassend, um eine inhomogene gewöhnliche oder partielle lineare Differentialgleichung zu lösen, aber auch für Systeme gekoppelter linearer Differentialgleichungen, brauchen wir einen vollständigen Satz von Lösungen der homogenen Gleichung und nur eine Speziallösung der inhomogenen Gleichung zu bestimmen. Die allgemeine Lösung erhält man dann als Linearkombination (5.10).

Eine weitere wichtige Variante des Superpositionsprinzips ist folgende: Betrachten wir eine lineare Gleichung mit einer Inhomogenität  $g = g_1 + g_2$  und zwei Funktionen  $\varphi_1$  und  $\varphi_2$ , die Lösungen der inhomogenen Gleichungen

$$\hat{\mathcal{L}}(\varphi_1) = g_1$$

und

$$\hat{\mathcal{L}}(\varphi_2) = g_2$$

sind. Dann ist die Superposition

$$\varphi = \varphi_1 + \varphi_2$$

eine Lösung der Gleichung

$$\hat{\mathcal{L}}(\varphi) = \hat{\mathcal{L}}(\varphi_1) + \hat{\mathcal{L}}(\varphi_2) = g_1 + g_2 = g.$$

Dieses Prinzip wird sehr oft im Elektromagnetismus benutzt. Zum Beispiel ist das elektrostatische Feld, das bei einer beliebigen Überlagerung von Punktladungen erzeugt wird, gleich der Summe der elektrischen Felder, die von jeder einzelnen separaten Ladung erzeugt werden. Ähnliches gilt bei der Response eines Stromkreises unter verschiedenen äußeren Spannungen oder Strömen.

An dieser Stelle sollte man auf die Rolle der Anfangs- oder Randbedingungen aufmerksam machen. Gibt es keine Randbedingungen oder sind Randbedingungen von  $\varphi_1$ ,  $\varphi_2$  und  $\varphi$  erfüllt, dann gibt es nichts dazu zu sagen. Ein Beispiel dafür ist die Poissongleichung  $\nabla^2 \phi = -4\pi \rho$  mit der Randbedingung  $\phi(\vec{r}) \rightarrow 0$  für  $r \rightarrow \infty$ . In anderen Fällen ist es so, dass weder  $\varphi_1$  noch  $\varphi_2$  die gewünschte Randbedingung erfüllen, aber doch  $\varphi = \varphi_1 + \varphi_2$ . Die Lösungsstrategie besteht manchmal darin, dass man  $g_2$  und die daraus resultierende  $\varphi_2$  so aussucht, dass  $\varphi = \varphi_1 + \varphi_2$  die vorgegebene Randbedingung erfüllt. Dies ist z.B. bei der Methode der Bildladungen in der Elektrostatik der Fall.

#### 5.4 Elementare Lösungsmethoden gewöhnlicher Differentialgleichungen erster Ordnung

Eine allgemeine Lösungstheorie gewöhnlicher Differentialgleichungen gibt es nicht. Für einige Klassen von Differentialgleichungen sind Existenz- und Eindeigkeitssätze bekannt, aber nur in wenigen Fällen gibt es geschlossene Methoden zur analytischen Auffindung der Lösung. Trotzdem sind viele für die Physik grundsätzliche Differentialgleichungen

analytisch lösbar. Andere wichtige Gleichungen, wie z.B. die Schrödinger-Gleichung mit vielen miteinander wechselwirkenden Elektronen und Ionen, können nur mit viel Mühe numerisch näherungsweise gelöst werden. Sie bleiben auf dem heutigen Stand der Forschung offene Probleme.

In den folgenden Abschnitten werden einige elementare Methoden kurz besprochen, die auf die manchmal iterative Suche von Stammfunktionen reduziert werden können. Wir gehen davon aus, dass dieser elementare, allerdings oft nicht einfache Schritt ohne weiteres durchgeführt werden kann.

#### 5.4.1 Trennung der Variablen

Eine wichtige Klasse nichtlinearer Differentialgleichungen erster Ordnung bilden die Gleichungen der Form

$$y' = f(x) g(y), \quad (5.11)$$

bei denen die Variablen  $x$  und  $y$  trennbar sind. Solange die Lösung  $y(x)$  die Bedingung  $g(y) \neq 0$  erfüllt, können wir schreiben

$$\frac{1}{g(y)} \frac{dy}{dx} = f(x)$$

oder

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

Integration auf beiden Seiten ergibt

$$\int_{y_0}^{y(x)} \frac{dy'}{g(y')} = \int_{x_0}^x dx' f(x'), \quad (5.12)$$

wobei wir die Anfangsbedingung  $y(x_0) = y_0$  eingesetzt haben. Die Lösung  $y(x)$  bestimmt man aus dieser Gleichung. Zur Verifikation kann man Gleichung (5.12) nach  $x$  ableiten. Nach Anwendung des Hauptsatzes und der Kettenregel bekommt man

$$\frac{1}{g(y(x))} y'(x) = f(x),$$

was mit Gleichung (5.11) übereinstimmt.

Beispiel:

$$\begin{aligned} y' = xy^2 &\Rightarrow \frac{dy}{y^2} = x dx \Rightarrow \left(-\frac{1}{y}\right) \Big|_{y_0}^y = \frac{x^2}{2} \Big|_{x_0}^x \\ &\Rightarrow \left(\frac{1}{y_0} - \frac{1}{y}\right) = \frac{1}{2} (x^2 - x_0^2) \Rightarrow y = \left[\frac{1}{y_0} - \frac{1}{2} (x^2 - x_0^2)\right]^{-1}. \end{aligned}$$

Zwei Spezialfälle von Gleichungen der Form (5.11) sollten hier separat erwähnt werden. Zunächst die bekannte Gleichung

$$y' = f(x). \quad (5.13)$$

In diesem Fall ergibt die Methode der Trennung von Variablen

$$dy = f(x) dx \Rightarrow y = y_0 + \int_{x_0}^x f(x') dx',$$

was der bekannte Ausdruck für die Stammfunktion ist. Interessanter ist der Fall

$$y' = g(y), \tag{5.14}$$

aus dem folgt

$$\frac{dy}{g(y)} = dx \Rightarrow x = x_0 + \int_{y_0}^y \frac{dy'}{g(y')} = x(y).$$

Die Lösung  $y(x)$  erhält man durch Inversion des Ausdrucks  $x = x(y)$ .

#### 5.4.2 Substitution

Viele andere auf den ersten Blick nicht trennbare Gleichungen lassen sich in der Form (5.11) durch eine passende Substitution von Variablen umschreiben. Zum Beispiel die Gleichung

$$y' = f(ax + by + c) \quad \text{mit } a, b, c \text{ Konstanten und } b \neq 0. \tag{5.15}$$

Wir betrachten die neue Variable

$$u = ax + by + c,$$

die folgende Gleichung erfüllt:

$$u' = a + by' = a + bf(ax + by + c) = a + bf(u),$$

welche die Form (5.14) hat. Die Lösung der Gleichung (5.15) ist dann

$$y(x) = [u(x) - ax - c] / b.$$

Eine ähnliche Methode kann man bei Gleichungen der Form

$$y' = f(y/x) \tag{5.16}$$

anwenden. Hier setzt man  $u = y/x$  für  $x \neq 0$ , welches die Gleichung

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

erfüllt. Diese Gleichung hat die Form (5.11) und ist also trennbar.

Durch Substitution lässt sich auch die Gleichung

$$y' = f\left(\frac{ax + by + c}{\alpha x + \beta y + \gamma}\right) \quad \text{mit } a, b, c, \alpha, \beta \text{ und } \gamma \text{ konstant} \tag{5.17}$$

lösen. In diesem Fall führt man zwei neue Variablen  $\eta = y - y_0$  und  $\xi = x - x_0$  ein, wobei  $(x_0, y_0)$  die Lösung des linearen Systems

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

ist. Wir gehen hier davon aus, dass diese Lösung eindeutig ist, was die Bedingung  $\begin{vmatrix} a & b \\ \alpha & \beta \end{vmatrix} \neq 0$  verlangt. Die Funktion  $\eta(\xi) = y(\xi + x_0) - y_0$  erfüllt dann eine Gleichung der Form (5.16).

### 5.4.3 Bernoulli-Gleichung

Die Bernoulli-Gleichung lautet

$$y' + g(x)y + h(x)y^\alpha = 0 \quad (5.18)$$

mit  $\alpha \neq 1$ . Wir teilen die Gleichung durch  $y^\alpha$  und erhalten

$$\begin{aligned} y^{-\alpha}y' + g(x)y^{(1-\alpha)} + h(x) &= 0 \\ \frac{1}{(1-\alpha)} [y^{(1-\alpha)}]' + g(x)y^{(1-\alpha)} + h(x) &= 0 \\ [y^{(1-\alpha)}]' + (1-\alpha)g(x)y^{(1-\alpha)} + h(x) &= 0. \end{aligned}$$

Das ist eine lineare Gleichung erster Ordnung für  $u = y^{(1-\alpha)}$ , die man mit den Methoden des Abschnitts 5.7 lösen kann.

### 5.4.4 Riccati-Gleichung

Die Riccati-Gleichung

$$y' + g(x)y + h(x)y^2 = k(x) \quad (5.19)$$

ist eine inhomogene nichtlineare Differentialgleichung. Ihre allgemeine Lösung lässt sich in der Form

$$y = u + y_p$$

auffinden, wobei  $y_p$  eine spezielle Lösung von (5.19) und  $u$  die Lösung einer homogenen Bernoulli-Gleichung mit  $\alpha = 2$  ist. Setzt man  $y$  und  $y_p$  in die Gleichung (5.19) ein und subtrahiert sie, erhält man

$$u' + g(x)u + h(x)(y^2 - y_p^2) = 0.$$

Schreibt man  $y^2 - y_p^2 = (y + y_p)(y - y_p) = (u + 2y_p)u$ , führt dies zu

$$u' + g(x)u + h(x)(u + 2y_p)u = 0$$

$$u' + [g(x) + 2y_p(x)]u + h(x)u^2 = 0.$$

Dies hat die Form einer Bernoulli-Gleichung (5.18) mit  $\alpha = 2$ .

### 5.4.5 D'Alembert-Gleichung

Die Gleichung lautet

$$y = x f(y') + g(y'). \quad (5.20)$$

Da die Abhängigkeit als Funktion von  $y'$  beliebig ist, lohnt es sich, den Variablenwechsel

$$u = y'(x)$$

einzuführen. Solange  $y''(x) \neq 0$  ist, können wir diese Beziehung nach  $x$  auflösen und  $x = x(u)$  schreiben. Dann ist auch  $y = y(x) = y(x(u))$  eine Funktion der neuen Variable  $u$ . Wir suchen nach der Gleichung, die  $x(u)$  erfüllen muss, damit  $y = y(x(u))$  Gleichung (5.20) erfüllt. Dazu leiten wir (5.20) nach  $u$  ab und erhalten

$$\frac{dy}{du} = \frac{dx}{du} f(u) + x f'(u) + g'(u),$$

wobei  $f'(u) = \frac{df}{du}$  und  $g'(u) = \frac{dg}{du}$  ist. Andererseits wissen wir, dass  $y = y(x(u))$  und daher

$$\frac{dy}{du} = \frac{dy}{dx} \cdot \frac{dx}{du} = y'(x) \frac{dx}{du} = u \frac{dx}{du}.$$

Gleichsetzen beider Ausdrücke für  $\frac{dy}{du}$  ergibt

$$\begin{aligned} u \frac{dx}{du} &= \frac{dx}{du} f(u) + x f'(u) + g'(u) \\ \frac{dx}{du} (f(u) - u) + x f'(u) + g'(u) &= 0. \end{aligned}$$

Dies ist eine lineare Differentialgleichung für  $x = x(u)$ , die man mit Standardmethoden lösen kann (siehe Abschnitt 5.7). Ist  $x = x(u)$  bekannt, muss man die Abbildung invertieren, um  $u = u(x)$  und damit  $y = x f(u(x)) + g(u(x))$  zu erhalten.

## 5.5 Exakte Differentialgleichungen

## 5.6 Integrierender Faktor

## 5.7 Lineare Differentialgleichungen erster Ordnung

## 5.8 Lineare Differentialgleichungen zweiter Ordnung

### 5.8.1 Struktur der allgemeinen Lösung

### 5.8.2 Konstante Koeffizienten: Exponentialansatz

## 5.9 Systeme linearer Differentialgleichungen erster Ordnung

## 5.10 Differentialgleichungen höherer Ordnung

## 6 Literaturhinweise

1. G. B. Arfken and H. Weber, *Mathematical Methods for Physicists*, 6th edition, (Elsevier Academic Press, 2005).
2. S. Großmann, *Mathematischer Einführungskurs für die Physik*, Teubner Studienbücher Physik, (B. G. Teubner, Stuttgart, 1991).
3. Ch. Blatter, *Analysis I, II und III*, Heidelberger Taschenbücher, (Springer Verlag, Berlin, 1977).
4. E. Kamke, *Differentialgleichungen. Lösungsmethoden und Lösungen*, 7. Auflage, (Akademische Verlagsgesellschaft Geest & Portig, Leipzig, 1961)

### Danksagung

Die Vorbereitung dieses Skriptes ist nur mit Hilfe der sorgfältigen Arbeit und des interessierten Einsatzes von Frau Andrea Wecker möglich gewesen, wofür der Autor und hoffentlich auch der Leser sehr dankbar sind.