Mathematische Ergänzung zu Experimentalphysik II

Skript einer zweistündigen Vorlesung
(gehalten im SS 2005, SS 2006, SS 2007)

Prof. Dr. Barbara Schrempp

Empfohlene elementare Literatur

- "Mathematik für Physiker I und II", 12. Auflage, K. Weltner, Springer Verlag, 2001
- "Mathematischer Einführungskurs für die Physik", 8. Auflage, S. Großmann, Teubner Verlag, 2000
- "Physik mit Bleistift", H. Schulz, Harri Deutsch Verlag, 2004

Empfohlene weiterführende Literatur

- "Mathematische Methoden in der Physik", C.B. Lang und N. Pucker, Spektrum Akademischer Verlag, 1998 (Hochschultaschenbuch)
- "Mathematik I und II – Geschrieben für Physiker", K. Jänich, Springer Verlag, 2001 (von einem Mathematiker, der den Stoff der Analysis und der Linearen Algebra zeitnah zum Stoff der Experimentalphysik I und II anordnet)
- E. Kamke, "Differentialgleichungen, Bd. 1, Gewöhnliche Differentialgleichungen", Teubner Verlag, 2002.

Danksagung

Ich danke Prof. Dr. E. Pehlke dafür, dass er mir das handschriftliche Skript seiner Vorlesung "Elementare Mathematische Methoden der Physik I" zur Verfügung gestellt hat. Es hat neben vielfältiger weiterer Literatur zur Stoffauswahl und Stoffgestaltung dieser Vorlesung beigetragen.

Inhaltsverzeichnis

1 Koordinatensysteme und Koordinatentransformationen

In der Vorlesung "Mathematische Ergänzung zu Experimentalphysik I" wurde bereits das rechtwinklige, sogenannte kartesische Koordinatensystem eingeführt. Dort wurde auch darauf hingewiesen, dass es sinnvoll ist, die Lage von Punkten, Vektoren und Raumkurven im dreidimensionalen Raum in einem auf die Problemstellung zugeschnittenen Koordinatensystem zu beschreiben. Es soll auch hier noch einmal betont werden, dass die Naturgesetze der Physik unabhängig vom Bezugssystem, d.h. der Wahl des Koordinatensystems, sind. Eine sinnvolle Wahl eines Koordinatensystems kann jedoch ganz wesentlich den Rechenaufwand bei der Lösung eines physikalischen Problems reduzieren.

Hier sollen zunächst weitere geläufige Typen von sogenannten *krummlinigen* Koordinatensystemen eingeführt werden, die je nach den Symmetrieeigenschaften eines physikalischen Problems geeigneter für seine Behandlung sein können als das kartesische Koordinatensystem. Diese werden o.B.d.A. relativ zu einem kartesischen Koordinatensystem eingeführt. Die Einführung kann zwar allgemeiner durchgeführt werden, wird jedoch dadurch erleichtert und ist deshalb auch durchaus gängig in der Literatur. Anschliessend werden die Transformationseigenschaften eines kartesischen Koordinatensystems unter Transformationen wie Translationen und (speziellen) Drehungen hergeleitet.

1.1 Polarkoordinaten in der Ebene

1.1.1 Definition

Wie in Abb. 1 veranschaulicht, wird ein Punkt P, der im ebenen kartesischen Koordinatensystem durch das Wertepaar (x, y) beschrieben wird, in ebenen Polarkoordinaten durch das Wertepaar (r, ϕ) von zwei neuen Variablen r und ϕ gekennzeichnet.

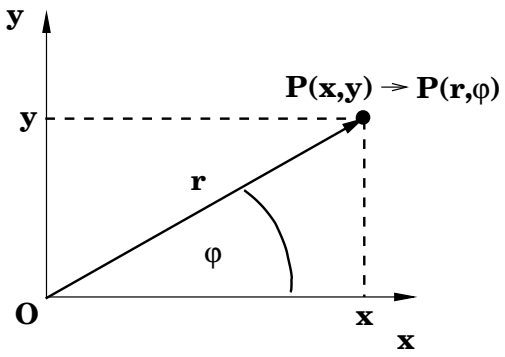


Abbildung 1:

ebene Polarkoordinaten sind durch zwei Variablen r und ϕ gekennzeichnet.

- Die Variable r ist die Länge $|\vec{r}|$ des Ortsvektors $\vec{r}$ vom Ursprung, dem Nullpunkt O, zum Punkt P.
- Die Variable ϕ ist der Winkel zwischen der positiven x-Achse und dem Ortsvektor $\vec{r}$.

Der Wertebereich des Winkels ϕ

- wird in der Regel zu $0^o \leq \phi < 360^o$ oder äquivalent zu $0 \leq \phi < 2\pi$ definiert, d.h. er springt bei 2π wieder auf 0 zurück;
- es gibt jedoch gelegentlich Anwendungen, in denen man ϕ unbeschränkt läßt, wie zum Beispiel bei der Beschreibung eines rotierenden Rades mit Winkelgeschwindigkeit ω, in dem $\phi = \omega t$ als Funktion der Zeit t kontinuierlich wächst.

Ebene Polarkoordinaten sind geeignet für physikalische Anwendungen, die auf die Ebene beschränkt sind und *symmetrisch unter Drehungen in der Ebene* um den Ursprung O sind. Dies wird an folgendem Beispiel deutlich.

1.1.2 Beispiel

Die Beschreibung aller Punkte eines Kreises um den Ursprung O mit Radius R ist besonders einfach in ebenen Polarkoordinaten

$$r = R, \quad 0 \leq \phi < 2\pi. \tag{1}$$

Hier bekommt auch die Bezeichnung "krummlinige" Koordinaten zum ersten Mal eine anschauliche Bedeutung. Während im ebenen kartesischen Koordinatensystem die Gleichung $x = x_0$ eine *Gerade* parallel zur y-Achse beschreibt, entspricht in ebenen Polarkoordinaten die Gleichung $r = R$ der "krummlinigen" Kurve eines *Kreises* mit Radius R.

1.1.3 Umrechnung von ebenen Polarkoordinaten in ebene kartesische Koordinaten und umgekehrt

Aus Abb. 1 liest man unmittelbar die Beziehungen zwischen den kartesischen und den Polarkoordinaten ab

$$x = r \cos\phi, \tag{2}$$
$$y = r \sin\phi. \tag{3}$$

Die Umkehrtransformation ist

$$r = +\sqrt{x^2 + y^2}, \tag{4}$$
$$\tan\phi = \frac{y}{x}. \tag{5}$$

Die Umkehrfunktion des Tangens in (5) ist nicht eindeutig. Wenn man sich auf den Hauptwert des $\arctan y/x$ beschränkt, $-\pi/2 < \arctan y/x < \pi/2$, dann erhält man die folgende Umkehrfunktion von (5), wie in Abb. 2 veranschaulicht,

$$\phi = \arctan\frac{y}{x} \qquad \text{für} \quad x > 0,\ y \geq 0 \tag{6}$$
$$\phi = \frac{\pi}{2} \qquad \text{für} \quad x = 0,\ y \geq 0 \tag{7}$$

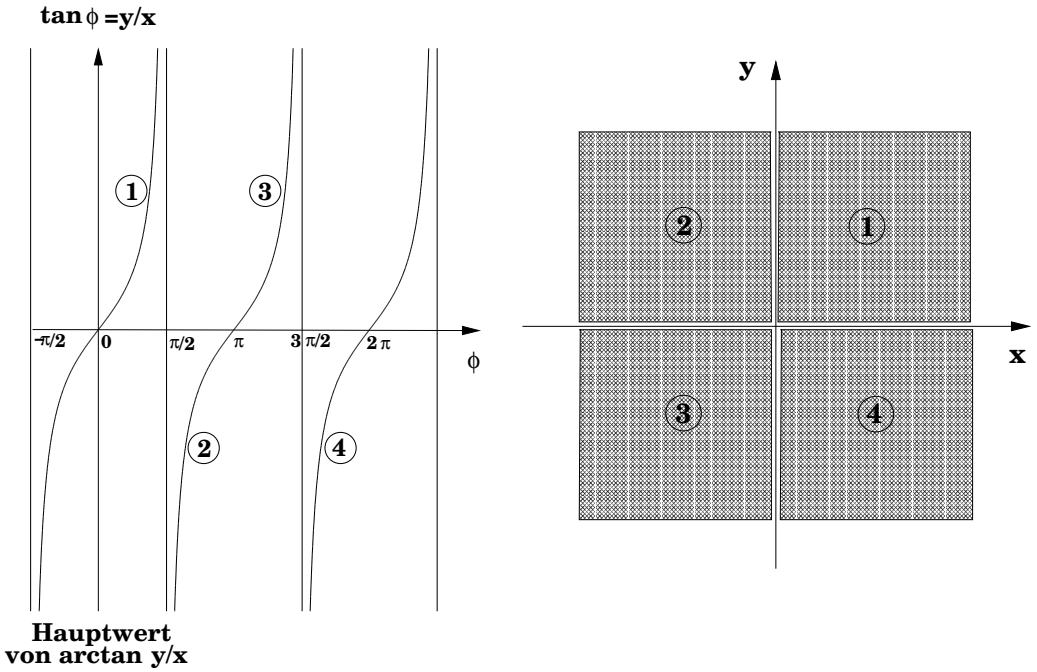


Abbildung 2:

$$\phi = \arctan\frac{y}{x} + \pi \qquad \text{für} \quad x < 0,\ y \geq 0 \tag{8}$$

$$\phi = \pi \qquad \text{für} \quad x < 0,\ y = 0 \tag{9}$$

$$\phi = \arctan\frac{y}{x} + \pi \qquad \text{für} \quad x < 0,\ y \leq 0 \tag{10}$$

$$\phi = \frac{3\pi}{2} \qquad \text{für} \quad x = 0,\ y \leq 0 \tag{11}$$

$$\phi = \arctan\frac{y}{x} + 2\pi \qquad \text{für} \quad x > 0,\ y \leq 0 \tag{12}$$

Die Zuordnung der einzelnen Halbzweige von ϕ zu den einzelnen Quadranten in der x-y-Ebene läßt sich in Abb. 2 ablesen.

1.2 Zylinderkoordinaten

1.2.1 Definition

Wie in Abb. 3 veranschaulicht, wird ein Punkt P, der im dreidimensionalen kartesischen Koordinatensystem durch das Wertetriplet (x, y, z) beschrieben wird, in Zylinderkoordinaten durch das Wertetriplet (r, ϕ, z) gekennzeichnet, das neben der kartesischen Koordinate z zwei neuen Variablen r und ϕ enthält.

Zylinderkoordinaten sind durch drei Variablen gekennzeichnet, durch z sowie durch r und ϕ. r und ϕ sind die ebenen Polarkoordinaten (siehe Kap. 1.1) der Projektion $P'(x, y, 0)$ des Punktes $P(x, y, z)$ in die x-y-Ebene ; r ist dementsprechend der Abstand sowohl von $P'(x, y, 0)$ als auch von $P(x, y, z)$ von der z-Achse.

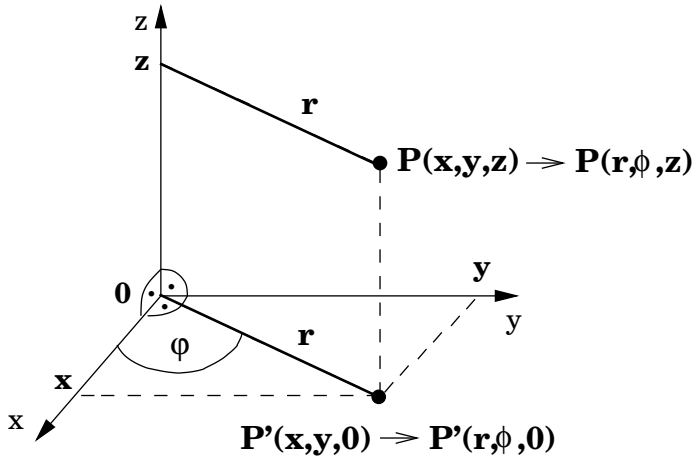


Abbildung 3:

Zylinderkoordinaten sind geeignet in physikalischen Anwendungen, die im dreidimensionalen Raum *symmetrisch unter Drehungen um eine Achse* sind, die o.B.d.A. entlang der z-Achse gelegt werden kann. Dies wird an folgendem Beispiel deutlich.

1.2.2 Beispiel

Die Beschreibung der Oberfläche eines z.B. unendlich langen Zylinders mit Radius R und Symmetrieachse längs der z-Achse ist

$$-\infty < z < \infty, \tag{13}$$
$$r = R, \tag{14}$$
$$0 \leq \phi < 2\pi. \tag{15}$$

1.2.3 Umrechnung von Zylinderkoordinaten in kartesische Koordinaten und umgekehrt

Hier kann auf die Umwandlung der ebenen Polarkoordinaten in ebene kartesische Koordinaten verwiesen werden, (2) und (3),

$$x = r\cos\phi, \tag{16}$$
$$y = r\sin\phi, \tag{17}$$
$$z = z. \tag{18}$$

Die Umkehrrelationen sind analog zu (4) und (5)

$$r = +\sqrt{x^2+y^2}, \tag{19}$$
$$\tan\phi = \frac{y}{x}, \tag{20}$$
$$z = z. \tag{21}$$

Dabei sind wieder die Konsequenzen der Vieldeutigkeit (6)-(12) der arctan-Funktion zu beachten.

1.3 Kugelkoordinaten

1.3.1 Definition

Wie in Abb. 4 veranschaulicht, wird ein Punkt P, der im dreidimensionalen kartesischen Koordinatensystem durch das Wertetriplet (x, y, z) beschrieben wird, in Kugelkoordinaten durch das Wertetriplet (r, θ, ϕ) von drei neuen Variablen r, θ und ϕ gekennzeichnet.

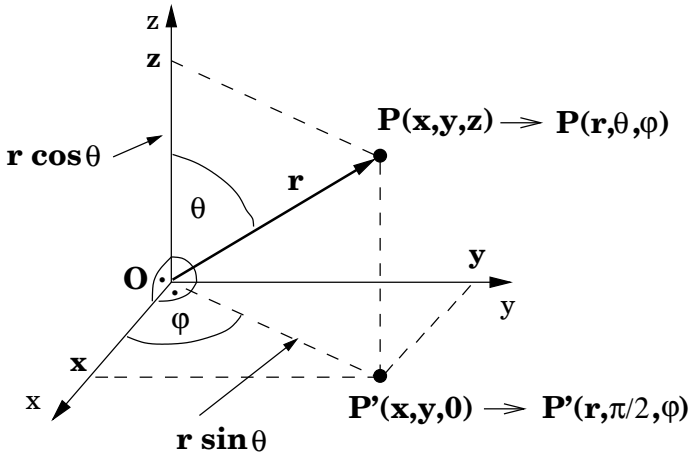


Abbildung 4:

Kugelkoordinaten, auch **sphärische Polarkoordinaten** genannt, sind durch drei Variablen gekennzeichnet, r, θ und ϕ.

- r ist der Betrag $|\vec{r}|$ des Ortsvektors $\vec{r}$, d.h. der Abstand des Punktes P vom Ursprung, dem Nullpunkt O,
- der sogenannte *Polarwinkel* θ ist der Winkel zwischen der positiven z-Achse und dem Ortsvektor $\vec{r}$
- der sogenannte *Azimuthwinkel* ϕ ist der Winkel zwischen der positiven x-Achse und dem Strahl vom Ursprung durch den Punkt P', der Projektion von P auf die x-y-Ebene.

Der Wertebereich der Winkelvariablen in Kugelkoordinaten ist

$$0 \leq \theta \leq \pi, \tag{22}$$

$$0 \leq \phi < 2\pi. \tag{23}$$

Kugelkoordinaten sind geeignet in physikalischen Anwendungen, die im dreidimensionalen Raum *symmetrisch unter beliebigen Drehungen um den Nullpunkt* sind, sogenannten *radialsymmetrischen* Funktionen. Dies gilt für Beschreibungen von Funktionen, die nur vom Radius r, dem Betrag des Ortsvektors, abhängen. Ein Beispiel ist die quantenmechanische Beschreibung des Wasserstoffatoms; zwischen dem negativ geladenen Elektron und dem positiv geladenen Kern herrscht die Coulombkraft, die nur vom Abstand r des Elektrons vom Kern abhängt.

Es ist zu betonen, dass die Variable r in ebenen Polarkordinaten, in Zylinderkoordinaten und in Kugelkoordinaten jeweils eine andere Definition hat, obwohl sie mit dem gleichen Symbol bezeichnet wird: in ebenen Polarkoordinaten ist sie der Betrag des Ortsvektors $\vec{r}$ von P im zweidimensionalen Koordinatensystem, in Zylinderkoordinaten ist sie der Betrag des Ortsvektors des Hilfspunktes P', der Projektion von P in die x-y-Ebene, oder in anderen Worten der Abstand des Punktes P von der z-Achse, in Kugelkoordinaten ist sie der Betrag des Ortsvektors $\vec{r}$ von P im dreidimensionalen Koordinatensystem.

1.3.2 Beispiel

Die Beschreibung der Oberfläche einer Kugel mit Radius R um den Ursprung O ist

$$r = R, \tag{24}$$
$$0 \leq \theta \leq \pi, \tag{25}$$
$$0 \leq \phi < 2\pi. \tag{26}$$

Dabei beschreiben die ersten beiden Wertebereiche einen halben Großkreis der Kugel, der bei Drehung um die z-Achse, bei der der volle Winkelbereich von ϕ von 0 bis 2π überstrichen wird, die Oberfläche der gesamten Kugel überstreicht. Dies illustriert auch sehr anschaulich, warum der Winkelbereich in θ von 0 bis π reicht (und nicht bis 2π).

1.3.3 Umrechnung von Kugelkoordinaten in kartesische Koordinaten und umgekehrt

Aus Abb. 4 läßt sich ablesen, dass die folgenden Beziehungen zwischen den kartesischen und den Kugelkoordinaten bestehen

$$x = r\sin\theta\cos\phi, \tag{27}$$
$$y = r\sin\theta\sin\phi, \tag{28}$$
$$z = r\cos\theta. \tag{29}$$

Hilfreich ist dabei die Feststellung, dass die Länge des Radiusvektors des Hilfspunktes P', der Projektion des Punktes P auf die x-y-Ebene, $r\sin\theta$ ist.

Die Umkehrtransformationen sind

$$r = |\vec{r}| = +\sqrt{x^2+y^2+z^2}, \tag{30}$$
$$\tan\theta = \frac{+\sqrt{x^2+y^2}}{z}, \tag{31}$$
$$\tan\phi = \frac{y}{x}. \tag{32}$$

Hier sind wieder die Konsequenzen der Vieldeutigkeit (6)-(12) der auftretenden arctan-Funktionen zu beachten.

1.4 Koordinatentransformationen kartesischer Koordinaten

1.4.1 Translation eines dreidimensionalen Koordinatensystems

Die einfachste Koordinatentransformation ist eine *Verschiebung*, eine sogenannte *Translation* des Koordinatensystems um einen beliebigen Vektor $\vec{r}_0$ mit den Komponenten x_0, y_0, z_0. In Abb. 5

sind zwei dreidimensionale kartesische Koordinatensysteme eingezeichnet, eines mit den Achsen x, y, z und ein parallel dazu um den Vektor $\vec{r}_0$ verschobenes Koordinatensystem mit den Achsen x', y', z'. Ein beliebiger Punkt P im Raum, der von dieser Verschiebung unberührt bleibt, hat im x-y-z-Koordinatensystem den Ortsvektor $\vec{r}$ mit den Komponenten x, y, z und im x'-y'-z'-Koordinatensystem den Ortsvektor $\vec{r}^{\,'}$ mit den Komponenten x', y', z'. Es gilt offensichtlich für jeden Punkt P im Raum

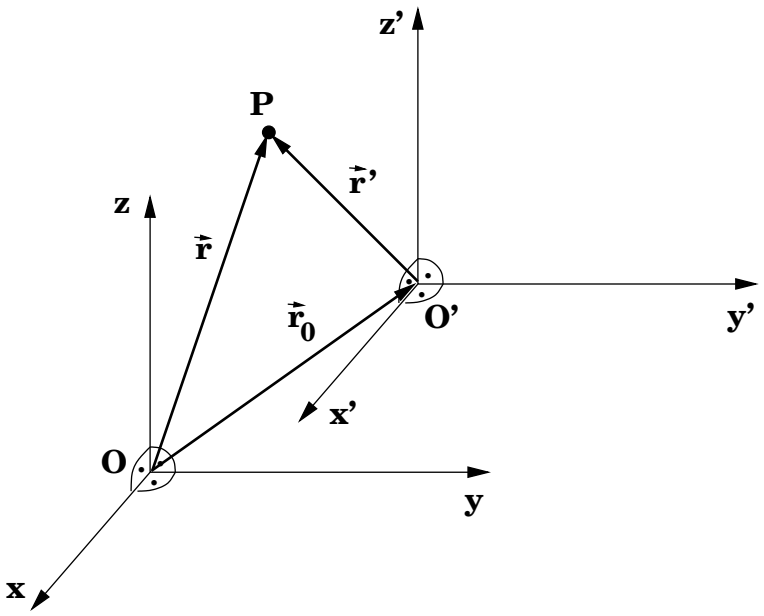


Abbildung 5:

$$\vec{r}^{\,'} = \vec{r} - \vec{r}_0 \tag{33}$$

und für die Komponenten

$$\begin{aligned} x' &= x - x_0, \\ y' &= y - y_0, \\ z' &= z - z_0. \end{aligned} \tag{34}$$

Die Gleichungen (33) bzw. (34) sind die Transformationsgleichungen für die Verschiebung oder Translation eines kartesischen Koordinatensystems.

1.4.2 Drehung eines ebenen Koordinatensystems

Gegeben sei ein ebenes x-y-Koordinatensystem, dargestellt in Abb. 6, in dem ein beliebiger Punkt P mit seinem Ortsvektor $\vec{r}$ eingezeichnet ist.

Die vektorielle Zerlegung des Ortsvektors in Vektoren parallel zur x-Achse und zur y-Achse unter Benutzung der Einheitsvektoren $\vec{e}_x$ bzw. $\vec{e}_y$,

$$\vec{r} = (x, y) = x\,\vec{e}_x + y\,\vec{e}_y, \tag{35}$$

ist ebenfalls dargestellt.

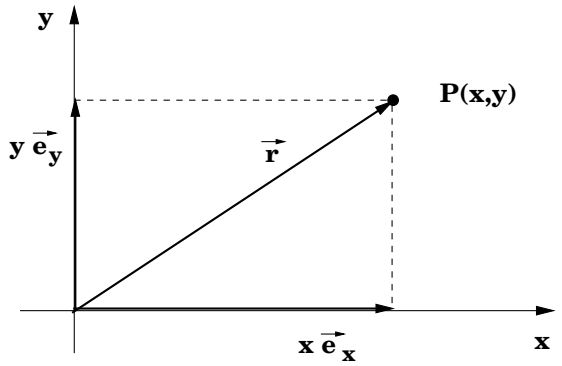


Abbildung 6:

In Abb. 7 sind zwei zweidimensionale kartesische Koordinatensysteme eingezeichnet, eines mit den Achsen x, y und ein um den Nullpunkt O um den Winkel ϕ gedrehtes Koordinatensystem mit den Achsen x', y'. Der Punkt P, der von dieser Drehung unberührt bleibt, hat im x-y-Koordinatensystem die Koordinaten x, y und im x'-y'-Koordinatensystem die Koordinaten x', y'.

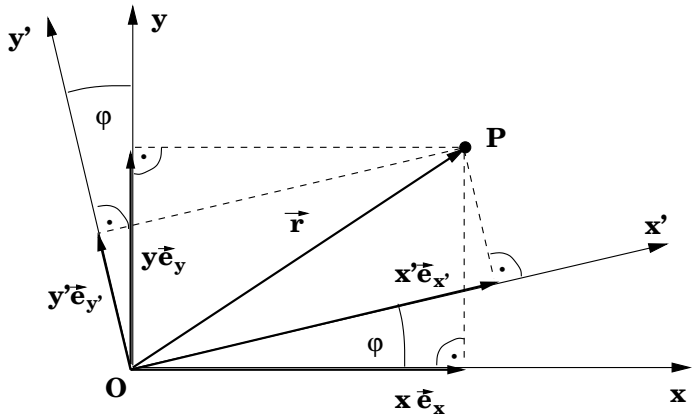


Abbildung 7:

In den beiden Abb. 8 sind relevante Ausschnitte der Abb. 7 dargestellt, die die jeweilige vektorielle Zerlegung von $x\vec{e}_x$ bzw. $y\vec{e}_y$ in die Vektoren $x'\vec{e}_{x'}$ parallel zur x'-Achse und $y'\vec{e}_{y'}$ parallel zur y'-Achse darstellen.

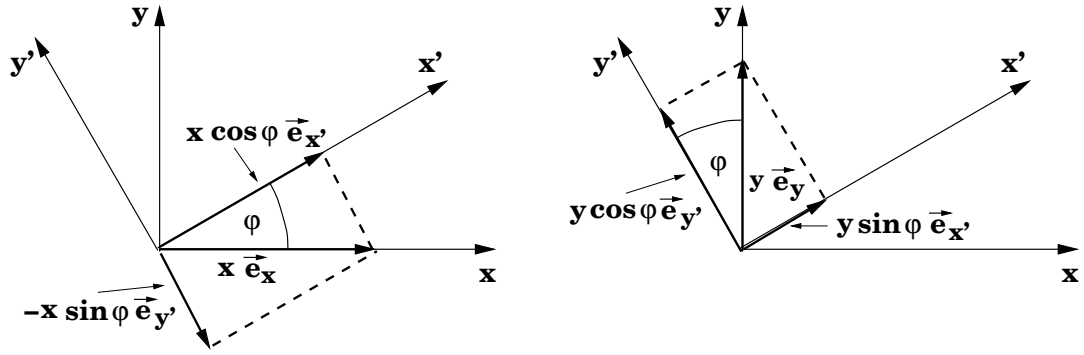


Abbildung 8:

Daraus liest man ab

$$x\,\vec{e}_x \;=\; x\cos\phi\,\vec{e}_{x'} - x\sin\phi\,\vec{e}_{y'}, \tag{36}$$

$$y\,\vec{e}_y \;=\; y\sin\phi\,\vec{e}_{x'} + y\cos\phi\,\vec{e}_{y'}. \tag{37}$$

Der Ortsvektor $\vec{r}$, der die beiden Punkte O und P miteinander verbindet, läßt sich also nach (35)-(37) im x'-y'-Koordinatensystem schreiben als

$$\vec{r} = x\,\vec{e}_x + y\,\vec{e}_y = (x\cos\phi + y\sin\phi)\,\vec{e}_{x'} + (-x\sin\phi + y\cos\phi)\,\vec{e}_{y'}. \tag{38}$$

Dies ist zu identifizieren mit

$$\vec{r} = x'\,\vec{e}_{x'} + y'\,\vec{e}_{y'}. \tag{39}$$

Daraus folgen die

Transformationsgleichungen für die Koordinaten eines beliebigen Punktes $P(x, y)$ bei der **Drehung eines zweidimensionalen Koordinatensystems** um den Winkel ϕ

$$x' \;=\; \cos\phi\,x + \sin\phi\,y, \tag{40}$$
$$y' \;=\; -\sin\phi\,x + \cos\phi\,y. \tag{41}$$

1.4.3 Spezielle Drehungen des Koordinatensystems im dreidimensionalen Raum

Es läßt sich aus der Herleitung und den Ergebnissen von Kap. 1.4.2 leicht die spezielle Drehung eines dreidimensionalen Koordinatensystems um den Winkel ϕ um die z-Achse angeben. Die beiden speziellen Drehungen um die x-Achse bzw. um die y-Achse, jeweils um einen Winkel ϕ, erhält man durch zyklische Vertauschungen der Koordinatensätze $x,\ y,\ z$ und $x',\ y',\ z'$.

Transformationsgleichungen für die Koordinaten eines beliebigen Punktes $P(x, y, z)$ unter drei **speziellen Drehungen eines dreidimensionalen Koordinatensystems** um den Winkel ϕ.

- Drehung um den Winkel ϕ um die z-Achse, wobei der Winkel ϕ in der x-y -Ebene von der positiven x-Achse aus in Richtung der positiven y-Achse gerechnet wird.

$$\begin{aligned} x' &= \cos\phi\, x + \sin\phi\, y, && (42)\\ y' &= -\sin\phi\, x + \cos\phi\, y, && (43)\\ z' &= z; && (44) \end{aligned}$$

- Drehung um den Winkel ϕ um die x-Achse, wobei der Winkel ϕ in der y-z-Ebene von der positiven y-Achse aus in Richtung der positiven z-Achse gerechnet wird.

$$\begin{aligned} x' &= x, && (45)\\ y' &= \cos\phi\, y + \sin\phi\, z, && (46)\\ z' &= -\sin\phi\, y + \cos\phi\, z; && (47) \end{aligned}$$

- Drehung um den Winkel ϕ um die y-Achse, wobei der Winkel ϕ in der z-x-Ebene von der positiven z-Achse aus in Richtung der positiven x-Achse gerechnet wird.

$$\begin{aligned} x' &= \cos\phi\, x - \sin\phi\, z, && (48)\\ y' &= y, && (49)\\ z' &= \sin\phi\, x + \cos\phi\, z; && (50) \end{aligned}$$

Diese drei speziellen Drehungen spielen eine wichtige Rolle, da jede beliebige Drehung eines Koordinatensystems durch drei aufeinander folgende spezielle Drehungen um die x-Achse, die y-Achse und die z-Achse beschrieben werden kann. Dies wird in Kap. 9.1 weiter ausgeführt.

2 Komplexe Zahlen

2.1 Definition und Rechenregeln für komplexe Zahlen

2.1.1 Definition der imaginären Zahlen

Beim Lösen von quadratischen Gleichungen trifft man auf die Notwendigkeit, komplexe Zahlen einzuführen, die eine erfolgreiche Erweiterung der reellen Zahlen darstellen.

Die spezifische quadratische Gleichung

$$x^2 = -1 \tag{51}$$

hat die Lösungen

$$x_{1,2} = \pm\sqrt{-1}. \tag{52}$$

$\sqrt{-1}$ ist keine reelle Zahl. Für diese "neuartige Zahl" wurde von Euler die Bezeichnung

$$i \text{ für } \sqrt{-1} \tag{53}$$

eingeführt.

Die quadratische Gleichung

$$x^2 = -4 \tag{54}$$

hat die Lösungen

$$x_{1,2} = \pm 2\sqrt{-1} = \pm 2\,i. \tag{55}$$

Diese beiden Beispiele motivieren die folgende Definition.

Definition: Die Zahl i mit der Eigenschaft $i^2 = -1$ ist **die Einheit der imaginären Zahlen**. Das Produkt aus einer reellen Zahl y und i ist eine imaginäre Zahl $y\,i$; es gilt $i\,y = y\,i$.

Die imaginäre Einheit i spielt eine analoge Rolle für die imaginären Zahlen wie die Zahl 1 für die reellen Zahlen.

2.1.2 Definition der komplexen Zahlen

Ein weiterführendes Beispiel ist die quadratische Gleichung

$$x^2 + x + 1 = 0; \tag{56}$$

sie hat die Lösungen

$$x_{1,2} = -\frac{1}{2} \pm \frac{\sqrt{3}}{2}\sqrt{-1} = -\frac{1}{2} \pm \frac{\sqrt{3}}{2}\,i, \tag{57}$$

die Summen aus einer reellen Zahl und einer imaginären Zahl sind.

Definition: Die Summe z aus einer reellen Zahl x und einer imaginären Zahl $i\,y$ ist eine **komplexe Zahl**

$$z = x + i\,y \quad \text{mit} \;\; x, y \;\in\; \mathbb{R}, \text{ der Menge der reellen Zahlen.} \tag{58}$$

Eine beliebige komplexe Zahl wird also beschrieben durch ein reelles Zahlenpaar (x,y). Die Menge der komplexen Zahlen wird in der Mathematik mit $\mathbb{C}$ bezeichnet. Komplexe Zahlen werden gern durch die Kleinbuchstaben z, w, ... vom Ende des Alphabets bezeichnet.

x heißt *Realteil von* z, y heißt *Imaginärteil von* z, in Kurzschreibweise

$$x = \mathrm{Re}\, z, \quad y = \mathrm{Im}\, z \quad \text{oder} \tag{59}$$

$$x = \Re\, z, \quad y = \Im\, z. \tag{60}$$

In der Vorlesung sollen die Bezeichnungen (59) verwendet werden. Es sei betont, dass der Imaginärteil einer Zahl reell ist und *nicht* – wie der Name suggerieren könnte – imaginär.

Definition: Zu jeder komplexen Zahl $z = x + i\,y$ ist die **konjugiert komplexe Zahl** definiert

$$z^* = x - i\,y \quad \text{mit} \quad (z^*)^* = z, \tag{61}$$

die eine vergleichbar wichtige Rolle spielt wie z.

2.1.3 Physikalische Motivation

Die Messwerte aller physikalischer Größen sind reell. Kein Messergebnis in der Physik wird durch eine komplexe Zahl beschrieben. Umso erstaunlicher ist die große Rolle, die komplexe Zahlen in Konzepten der Physik spielen.

- Komplexe Zahlen vereinfachen Lösungen von gewöhnlichen und partiellen Differentialgleichungen, die Schwingungen bzw. Wellen beschreiben; d.h. sie kommen sowohl in der klassischen Mechanik als auch in der Elektrodynamik zum Einsatz.
- In der Wechselstromlehre vereinfachen sie die Berechnung von Phasenverschiebungen zwischen Strom und Spannung.
- Sie treten in der speziellen Relativitätstheorie auf.
- Vor allem spielen sie eine wichtige Rolle in der Quantenmechanik, zu deren Beschreibung komplexe Wellenfunktionen oder alternativ Vektoren und Matrizen mit komplexen Elementen erforderlich sind.

2.1.4 Rechenregeln für komplexe Zahlen

In der Mathematik lernt man, dass die reellen Zahlen einen Körper bilden; auch die komplexen Zahlen bilden einen Körper. Die zentralen Eigenschaften werden hier ohne Bezug auf den mathematischen Begriff eines Körpers genannt.

Seien

$$z_1 = x_1 + i\,y_1, \quad z_2 = x_2 + i\,y_2, \quad z_3 = x_3 + i\,y_3 \tag{62}$$

beliebige komplexe Zahlen. Es ist

$$\begin{aligned} z_1 &= z_2 \text{ genau dann, wenn } x_1 = x_2 \text{ und } y_1 = y_2, && (63)\\ z &= 0 \text{ genau dann, wenn } x = 0 \text{ und } y = 0 && (64)\\ -z &= -x - i\,y. && (65) \end{aligned}$$

Definition der **Addition komplexer Zahlen**

$$z_1 + z_2 = x_1 + x_2 + i\,(y_1 + y_2), \tag{66}$$

d.h.

$$\mathrm{Re}\,(z_1 + z_2) = \mathrm{Re}\,z_1 + \mathrm{Re}\,z_2 = x_1 + x_2, \quad \mathrm{Im}\,(z_1 + z_2) = \mathrm{Im}\,z_1 + \mathrm{Im}\,z_2 = y_1 + y_2. \tag{67}$$

Es gilt das *Kommutativgesetz*

$$z_1 + z_2 = z_2 + z_1 \tag{68}$$

sowie das *Assoziativgesetz*

$$(z_1 + z_2) + z_3 = z_1 + (z_2 + z_3). \tag{69}$$

Definition der **Multiplikation komplexer Zahlen**

$$z_1\, z_2 = x_1\, x_2 - y_1\, y_2 + i\,(x_1\, y_2 + x_2\, y_1), \tag{70}$$

d.h.

$$\mathrm{Re}\, z_1\, z_2 = x_1\, x_2 - y_1\, y_2, \quad \mathrm{Im}\, z_1\, z_2 = x_1\, y_2 + x_2\, y_1. \tag{71}$$

Es gilt das *Kommutativgesetz*

$$z_1\, z_2 = z_2\, z_1, \tag{72}$$

das *Assoziativgesetz*

$$(z_1\, z_2)\, z_3 = z_1\, (z_2\, z_3) \tag{73}$$

sowie das *Distributivgesetz*

$$(z_1 + z_2)\, z_3 = z_1\, z_3 + z_2\, z_3. \tag{74}$$

(70) kann durch einfaches Ausmultiplizieren von $(x_1 + i\, y_1)(x_2 + i\, y_2) = x_1\, x_2 + + i\, x_1\, y_2 + i\, y_1\, x_2 = i^2\, y_1\, y_2$ unter Berücksichtigung von $i^2 = -1$ verifiziert werden.

Alle oben genannten Gesetze lassen sich leicht auf die entsprechenden Gesetze für reelle Zahlen zurückführen.

Das spezielle Produkt

$$\begin{aligned} z\, z^* &= (x + i\, y)\, (x - i\, y) = x^2 + i\, y\, x - i\, x\, y - i^2\, y^2 && (75)\\ &= x^2 + y^2 && (76) \end{aligned}$$

ist *reell* und ≥ 0; insbesondere gilt

$$i\, i^* = i(-i) = 1. \tag{77}$$

Es gilt offensichtlich

$$i^0 = 1, \quad i^1 = i, \quad i^2 = -1, \quad i^3 = i^2\, i = -i, \quad i^4 = (i^2)^2 = (-1)^2 = 1, \ \ldots \tag{78}$$

Die Division ergibt sich mit Hilfe der Multiplikation. Für $z \neq 0$ ist $\dfrac{1}{z} = \dfrac{1}{x + i\, y}$ folgendermaßen zu behandeln: der Bruch wird mit z^* erweitert; dadurch erhält man den reellen Nenner $z\, z^*$. Das Ergebnis ist

$$\frac{1}{z} = \frac{z^*}{z\, z^*} = \frac{x - i\, y}{x^2 + y^2} = \frac{x}{x^2 + y^2} - i\, \frac{y}{x^2 + y^2} \tag{79}$$

Insbesondere gilt

$$\frac{1}{i} = -i. \tag{80}$$

Analog gilt für $z_2 \neq 0$

$$\frac{z_1}{z_2} = \frac{z_1\, z_2^*}{z_2\, z_2^*} = \frac{x_1\, x_2 + y_1\, y_2}{x_2^2 + y_2^2} - i\, \frac{x_1\, y_2 - y_1\, x_2}{x_2^2 + y_2^2}. \tag{81}$$

2.2 Gaußsche Zahlenebene und Polardarstellung komplexer Zahlen

Die komplexe Zahl $z = x + i\,y$, beschrieben durch das reelle Zahlenpaar (x, y), wird – in Anlehnung an die Darstellung eines Punktes $P(x, y)$ mit den Koordinaten x und y in einem zweidimensionalen kartesisches Koordinatensystem – als *Punkt in der zweidimensionalen Gaußschen Zahlenebene*, Abb. 9, dargestellt. Letztere wird auch *komplexe z-Ebene* genannt.

Die "x"-Achse heißt *reelle Achse*, die "y"-Achse *imaginäre Achse*. Die Projektionen des Punktes z auf die relle bzw. imaginäre Achse geben als Achsenabschnitte den Realteil $x = \mathrm{Re}\, z$ bzw. Imaginärteil $y = \mathrm{Im}\, z$ an.

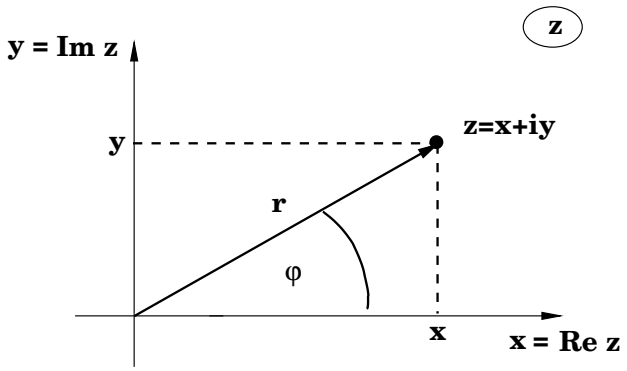


Abbildung 9:

Analog zum Übergang von zweidimensionalen kartesischen Koordinaten (x, y) zu zweidimensionalen Polarkoordinaten r, ϕ gibt es

Definition: die **Polardarstellung** der komplexen Zahl z in der Gaußschen Zahlenebene durch einen "Radiusvektor" mit *Betrag* von z

$$r = |z| = \sqrt{x^2 + y^2}, \quad \text{mit} \quad |z| = \sqrt{z\,z^*} \geq 0 \tag{82}$$

und *Winkel* ϕ mit der positiven reellen Achse

$$\phi = \arg z = \arctan \frac{y}{x}, \tag{83}$$

dem sogenannten *Argument* von z.

Die arctan-Funktion ist, wie bereits im Kap. 1.1 diskutiert, vieldeutig. Man kann sich auf den Bereich

$$-\pi < \arg z \leq \pi \tag{84}$$

beschränken; negative Winkel werden von der positiven x-Achse aus nach unten abgetragen.

Aus Abb. 9 liest man ab

$$\mathrm{Re}\, z = x \;=\; r\,\cos\phi \quad \text{oder} \quad \mathrm{Re}\, z = x = |z|\,\cos\,\arg z, \tag{85}$$

$$\mathrm{Im}\, z = y \;=\; r\,\sin\phi \quad \text{oder} \quad \mathrm{Im}\, z = y = |z|\,\sin\,\arg z \tag{86}$$

bzw.

$$z = r\,(\cos\phi + i\,\sin\phi) \quad \text{oder} \quad z = |z|\,(\cos\arg z + i\,\sin\arg z). \tag{87}$$

Der zu z konjugiert komplexe Punkt z^* erscheint in der Gaußschen Zahlenebene als der *an der reellen Achse gespiegelte Punkt*

$$z^* = r\,(\cos\phi - i\,\sin\phi) \quad \text{oder} \quad z^* = |z|\,(\cos\arg z - i\,\sin\arg z), \tag{88}$$

wie in Abb. 10 dargestellt. Dabei wurde $\cos\phi = \cos(-\phi)$ und $\sin\phi = -\sin(-\phi)$ benutzt.

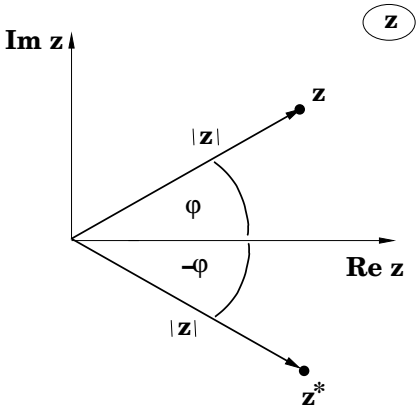


Abbildung 10:

2.3 Euler-Formel

Es gilt die folgende Identität

Euler-Formel:

$$e^{i\,\phi} = \cos\phi + i\,\sin\phi \quad \text{mit} \tag{89}$$

$$\mathrm{Re}\,e^{i\,\phi} = \cos\phi, \quad \mathrm{Im}\,e^{i\,\phi} = \sin\phi. \tag{90}$$

Zum Beweis wird an die Taylorentwicklungen für reelle x erinnert

$$e^x = \sum_{n=0}^{\infty} \frac{1}{n!}\,x^n, \qquad \cos x = \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n)!}\,x^{2n}, \qquad \sin x = \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n+1)!}\,x^{2n+1}. \tag{91}$$

Diese Enwicklungen lassen sich auch auf komplexe z erweitern (der mathematische Fachausdruck ist *in die komplexe z-Ebene analytisch fortsetzen*). Davon wird noch in Kap. 2.5 die Rede sein. Hier sollen sie zunächst auf ein imaginäres Argument $i\,\phi$, ϕ reell, erweitert werden

$$e^{i\,\phi} = \sum_{n=0}^{\infty} \frac{1}{n!}\,(i\,\phi)^n = \sum_{n=0}^{\infty} i^n\,\frac{1}{n!}\,\phi^n \tag{92}$$

$$= 1 + i\,\frac{1}{1!}\,\phi + i^2\,\frac{1}{2!}\,\phi^2 + i^3\,\frac{1}{3!}\,\phi^3 + i^4\,\frac{1}{4!}\,\phi^4 + \ldots \tag{93}$$

$$= 1 + i\,\frac{1}{1!}\,\phi - \frac{1}{2!}\,\phi^2 - i\,\frac{1}{3!}\,\phi^3 + \frac{1}{4!}\,\phi^4 + \ldots \tag{94}$$

$$= \sum_{n=0}^{\infty} \left(\frac{(-1)^n}{(2n)!}\, \phi^{2n} + i\, \frac{(-1)^n}{(2n+1)!}\, \phi^{2n+1} \right) \tag{95}$$

$$= \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n)!}\, \phi^{2n} + i \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n+1)!}\, \phi^{2n+1} \tag{96}$$

$$= \cos\phi + i\, \sin\phi. \tag{97}$$

Der Betrag von $e^{i\phi}$ ist

$$|e^{i\phi}| = \sqrt{\cos^2\phi + \sin^2\phi} = 1 \tag{98}$$

für alle Phasenwinkel ϕ. Es gilt offensichtlich

$$e^{-i\,\phi} = \cos\phi - i\, \sin\phi \tag{99}$$

und zusammen mit (89)

$$\cos\phi = \frac{1}{2}\,(e^{i\,\phi} + e^{-i\,\phi}), \tag{100}$$

$$\sin\phi = \frac{1}{2i}\,(e^{i\,\phi} - e^{-i\,\phi}); \tag{101}$$

es ist zu beachten, dass im Nenner von (101) ein i steht.

Eine wichtige Eigenschaft von $e^{i\,\phi}$ läßt sich aus der Periodizität von $\cos\phi$ und $\sin\phi$

$$\cos\phi = \cos(\phi + 2\pi n), \quad \sin\phi = \sin(\phi + 2\pi n) \quad \text{für ganze Zahlen n} \tag{102}$$

ablesen

$$e^{i\,\phi} = e^{i\,(\phi + 2\pi n)} = e^{i\,\phi + 2\pi n i} = e^{i\,\phi}\, e^{2\pi n i} \quad \text{für ganze Zahlen n.} \tag{103}$$

Daraus folgt die Identität

$$e^{2\pi n i} = 1 \quad \text{für ganze Zahlen } n. \tag{104}$$

Es gilt weiterhin

$$e^{\pi n i} = \cos\pi n + i\, \sin\pi n = (-1)^n \quad \text{für ganze Zahlen } n. \tag{105}$$

Von Interesse sind auch die sogenannten Einheitswurzeln

$$e^{2\pi \frac{k}{n} i} \quad \text{für} \quad k = 0,\, 1,\, ...,\, n-1, \tag{106}$$

die in den Übungen diskutiert werden sollen.

2.4 Rechnen in der Gaußschen Zahlenebene und mit der Euler-Formel

- Mit Hilfe der Euler-Formel läßt sich jede komplexe Zahl z nach (85) und (86) schreiben als

$$z = |z|\,(\cos\,\arg z + i\, \sin\,\arg z) = |z|\, e^{i\,\arg z}. \tag{107}$$

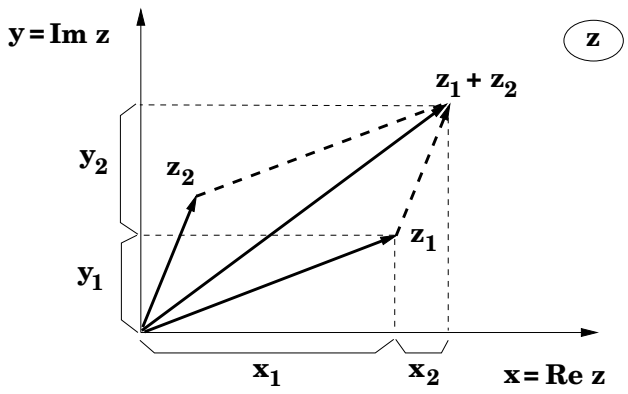


Abbildung 11:

- **Addition**

  Die Addition von zwei komplexen Zahlen $z_1 = x_1 + i\,y_1$ und $z_2 = x_2 + i\,y_2$ in der Gaußschen Zahlenebene entspricht der Addition von zwei Vektoren in der Ebene; aus Abb. 11 läßt sich nachvollziehen, dass sich bei der Vektoraddition jeweils die beiden Realteile und die beiden Imaginärteile addieren.

- **Multiplikation**

  Für die Multiplikation von zwei komplexen Zahlen $z_1 = x_1 + i\,y_1$ und $z_2 = x_2 + i\,y_2$ benutzt man die Euler-Formel (89)

$$z_1 = |z_1|\,(\cos\arg z_1 + i\,\sin\arg z_1) = |z_1|\,e^{i\arg z_1}, \tag{108}$$
$$z_2 = |z_2|\,(\cos\arg z_2 + i\,\sin\arg z_2) = |z_2|\,e^{i\arg z_2}. \tag{109}$$

  Ihr Produkt ist dann

$$z_1\,z_2 = |z_1|\,|z_2|\,e^{i\arg z_1}\,e^{i\arg z_2} = |z_1|\,|z_2|\,e^{i\,(\arg z_1 + \arg z_2)}. \tag{110}$$

  Es gilt also

$$|z_1\,z_2| = |z_1|\,|z_2| \quad \text{und} \tag{111}$$
$$\arg z_1\,z_2 = \arg z_1 + \arg z_2 \mod 2\pi \quad \text{gesprochen: modulo } 2\pi. \tag{112}$$

  Hier ist die mathematische Schreibweise für die Vieldeutigkeit (103) eingeführt worden; sie impliziert, dass $\arg z_1\,z_2$ nur bis auf additive ganze Vielfache von 2π bekannt ist und dass gegebenenfalls ganze Vielfache von 2π zu $arg\,z_1 z_2$ addiert werden müssen, um $\arg z_1 z_2$ im Wertebereich (84) zu halten.

- **Inverses** für $z \neq 0$

$$\frac{1}{z} = \frac{z^*}{|z|^2} = \frac{|z|\,e^{-i\arg z}}{|z|^2} = \frac{1}{|z|}\,e^{-i\arg z} \tag{113}$$

  oder

$$\left|\frac{1}{z}\right| = \frac{1}{|z|} \quad \text{und} \quad \arg\Big(\frac{1}{z}\Big) = -\arg z. \tag{114}$$

- **Division** für $z_2 \neq 0$

$$\frac{z_1}{z_2} = \frac{z_1\, z_2^*}{|z_2|^2} = \frac{|z_1|\,|z_2|\, e^{i(\arg z_1 - \arg z_2)}}{|z_2|^2} = \frac{|z_1|}{|z_2|}\, e^{i(\arg z_1 - \arg z_2)} \tag{115}$$

oder

$$\left|\frac{z_1}{z_2}\right| = \frac{|z_1|}{|z_2|} \quad \text{und} \quad \arg\left(\frac{z_1}{z_2}\right) = \arg z_1 - \arg z_2 \bmod 2\pi. \tag{116}$$

2.5 Einige Funktionen von komplexen Variablen

Viele der bekannten Funktionen, die in der Physik eine Rolle spielen, lassen sich auf den Definitionsbereich $\mathbb{C}$ komplexer Zahlen erweitern (der Fachausdruck ist: in die komplexe z-Ebene analytisch fortsetzen), wie e^z, $\sin z$, $\cos z$, ..., $\sinh z$, $\cosh z$, ... Sie werden zu komplexen Funktionen eines komplexen Arguments und bilden $\mathbb{C}$ auf $\mathbb{C}$ ab.

- e^z

$$e^z = e^{(x+i\,y)} = e^x\, e^{i\,y} = e^x\,(\cos y + i\,\sin y) \tag{117}$$

mit

$$\mathrm{Re}\, e^z = e^x\,\cos y, \quad \mathrm{Im}\, e^z = e^x\,\sin y. \tag{118}$$

- $\sin z$

$$\begin{aligned} \sin z &= \frac{1}{2\,i}\left(e^{i\,z} - e^{-i\,z}\right) = \frac{1}{2\,i}\left(e^{i\,x-y} - e^{-i\,x+y}\right) && (119)\\ &= \frac{-i}{2}\left(e^{-y}(\cos x + i\,\sin x) - e^{y}(\cos x - i\,\sin x)\right) && (120)\\ &= \sin x\,\cosh y + i\,\cos x\,\sinh y && (121) \end{aligned}$$

mit

$$\mathrm{Re}\,\sin z = \sin x\,\cosh y, \quad \mathrm{Im}\,\sin z = \cos x\,\sinh y. \tag{122}$$

- $\sinh z$

$$\begin{aligned} \sinh z &= \frac{1}{2}\left(e^{z} - e^{-z}\right) = \frac{1}{2}\left(e^{x+i\,y} - e^{-x-i\,y}\right) && (123)\\ &= \frac{1}{2}\left(e^{x}(\cos y + i\,\sin y) - e^{-x}(\cos y - i\,\sin y)\right) && (124)\\ &= \sinh x\,\cos y + i\,\cosh x\,\sin y && (125) \end{aligned}$$

mit

$$\mathrm{Re}\,\sinh z = \sinh x\,\cos y, \quad \mathrm{Im}\,\sinh z = \cosh x\,\sin y. \tag{126}$$

In allen Fällen kann man leicht nachvollziehen, dass im Grenzwert $\mathrm{Im}\, z = y \to 0$ die erwartete Reduktion auf die entsprechende reelle Funktion der reellen Variablen x eintritt.

2.6 Die komplexe Wurzel

Für reelle Variablen $x > 0$ hat die Gleichung

$$y^2 = x \tag{127}$$

die beiden Lösungen

$$y_1 = +\sqrt{x} \quad \text{und} \quad y_2 = -\sqrt{x}, \tag{128}$$

wobei $\sqrt{x}$ als die *positive* Wurzel definiert ist.

Für komplexe Variablen z ist der Ausgangspunkt

$$z = |z|\, e^{i\,\phi + 2\pi n\, i}, \quad n \text{ ganze Zahl.} \tag{129}$$

Der Exponent $2\pi n\, i$ berücksichtigt die Vieldeutigkeit des Phasenwinkels.

Definition der **komplexen Wurzel**

$$\sqrt{z} \;=\; z^{\frac{1}{2}} = \sqrt{|z|}\, e^{i\,\frac{\phi}{2} + \pi n\, i} \tag{130}$$

$$\;=\; (-1)^n \sqrt{|z|}\, e^{i\,\frac{\phi}{2}} = \begin{cases} +\sqrt{|z|}\, e^{i\,\frac{\phi}{2}} & \text{für } n \text{ gerade} \\ -\sqrt{|z|}\, e^{i\,\frac{\phi}{2}} & \text{für } n \text{ ungerade} \end{cases} \tag{131}$$

Die komplexe Wurzel ist – wie nicht anders erwartet – *zweideutig*.

2.6.1 Eindeutige Definition der Wurzel

Die Wurzel kann (willkürlich) eindeutig definiert werden, indem man den Wertebereich des Phasenwinkels ϕ einschränkt, z.B. durch

$$-\pi < \phi \leq \pi. \tag{132}$$

Jede Definition $\alpha < \phi \leq \alpha + 2\pi$ würde dem gleichen Zweck dienen.

Eindeutige Definition der **komplexen Wurzel** sei für ϕ im Bereich (132), $-\pi < \phi \leq \pi$,

$$\sqrt{z} = z^{\frac{1}{2}} = \sqrt{|z|}\, e^{i\,\frac{\phi}{2}}. \tag{133}$$

Die Folge dieser Definition ist eine *Unstetigkeit* längs der *negativen rellen Achse* der komplexen z-Ebene. Dies ist in Abb. 12 verdeutlicht.

Wählt man ϕ infinitesimal kleiner als π, ist die Einschränkung (132) erfüllt und dementsprechend (133) $\sqrt{z} = \sqrt{|z|}\, e^{i\,\frac{\phi}{2}}$. Wählt man ϕ infinitesimal größer als π, ist die Einschränkung (132) nicht mehr erfüllt; sie muss erzwungen werden, indem von ϕ der Betrag 2π subtrahiert wird; der resultierende Phasenwinkel $\phi - 2\pi$ führt entsprechend (133) auf $\sqrt{z} = \sqrt{|z|}\, e^{i\,\frac{\phi}{2} - \pi\, i} = -\sqrt{|z|}\, e^{i\,\frac{\phi}{2}}$.

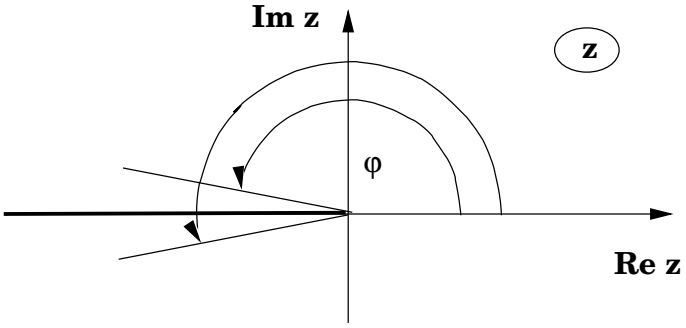


Abbildung 12:

Insgesamt erfolgt beim Überschreiten der negativen rellen Achse *ein unstetiger Vorzeichenwechsel*. Analoges gilt, wenn die negative reelle Achse von unten überschritten wird. Dieser unstetige Vorzeichenwechsel ist der Preis für die künstliche Einschränkung (132).

Ein zweiter Preis ist, dass nicht immer $\sqrt{z_1\,z_2} = \sqrt{z_1}\sqrt{z_2}$ gilt. Davon kann man sich an einem Beispiel überzeugen.

Sei $z_1 = z_2 = e^{\frac{3}{4}\pi\,i}$. Da $\phi = \frac{3}{4}\pi$ im erlaubten Bereich (132) liegt, ist

$$\sqrt{z_1}\sqrt{z_2} = e^{\frac{3}{8}\pi\,i}\,e^{\frac{3}{8}\pi\,i} = e^{\frac{3}{4}\pi\,i}. \tag{134}$$

Im Produkt $z_1\,z_2 = e^{\frac{3}{4}\pi\,i}\,e^{\frac{3}{4}\pi\,i} = e^{\frac{3}{2}\pi\,i}$ verläßt der Phasenwinkel $\frac{3}{2}\pi$ den erlaubten Bereich (132) und fällt erst nach Subtraktion von 2π wieder in den erlaubten Bereich, $\frac{3}{2}\pi - 2\pi = -\frac{1}{2}\pi$.

$$\sqrt{z_1\,z_2} = \sqrt{e^{-\frac{1}{2}\pi\,i}} = e^{-\frac{1}{4}\pi\,i} = -e^{\frac{3}{4}\pi\,i}. \tag{135}$$

Es gilt also in diesem Beispiel

$$\sqrt{z_1\,z_2} = -\sqrt{z_1}\sqrt{z_2}. \tag{136}$$

2.6.2 Riemann-Flächen am Beispiel der Wurzelfunktion

Es gibt viele mehrdeutige komplexe Funktionen. So ist z.B. die Funktion $z^{\frac{1}{n}}$ n-deutig. In den Übungen wird die Funktion $\log z$ diskutiert, die abzählbar unendlich vieldeutig ist. Eine elegante Lösung des Problems vieldeutiger Funktionen besteht in der Einführung von sogenannten *Riemann-Flächen*. Dieses Thema übersteigt zwar den Rahmen dieser Vorlesung, soll aber hier am Beispiel der komplexen Quadratwurzel kurz skizziert werden.

Der Definitionsbereich des Phasenwinkels ϕ wird anstelle des Definitionsbereichs (132), $-\pi < \phi \leq \pi$, ausgedehnt auf den doppelten Definitionsbereich

$$-\pi < \phi \leq 3\pi. \tag{137}$$

Dabei überdeckt der Winkelbereich $-\pi < \phi \leq \pi$ die komplexe z-Ebene einmal – dies ist das sogenannte *erste Blatt der Riemann-Fläche* – und der anschließende Winkelbereich $\pi < \phi \leq 3\pi$ überdeckt sie ein zweites Mal– dies ist das sogenannte *zweite Blatt der Riemann-Fläche*. Die beiden Blätter sind in Abb. 13 so eingezeichnet, dass das erste Blatt voll sichtbar ist und das zweite unvollständig verdeckt darunter liegt. Die beiden Blätter sind nur längs der negativen reellen Achse miteinander verbunden, dem sogenannten *Verzweigungsschnitt*. Die so zusammenhängenden zwei Blätter bilden die *Riemann-Fläche der komplexen Wurzel*.

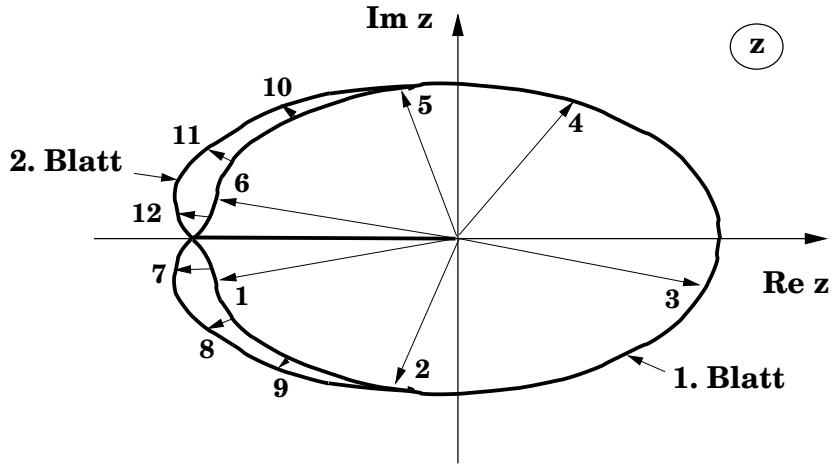


Abbildung 13:

Natürlich haben beide Blätter eine Ausdehnung bis $|z| \to \infty$, die endliche Ausdehnung in Abb. 13 dient nur der Veranschaulichung.

Der Weg, der bei Überstreichen des Winkelbereichs (137) durchlaufen wird, ist ebenfalls in Abb. 13 skizziert, indem einzelne, repräsentative Stationen des Weges mit den Zahlen 1 bis 12 durchnummeriert wurden. Man beginnt auf dem ersten Blatt bei einem Winkel, der infinitesimal größer als $-\pi$ ist, durchläuft über die Stationen 1 bis 6 das erste Blatt, gelangt bei $\phi = \pi$ auf die negative reelle Achse und läuft dort *ohne Unstetigkeit* in das zweite Blatt hinüber, wo die Stationen 7 bis 12 durchlaufen werden bis zur erneuten Ankunft an der negativen reellen Achse beim Winkel $\phi = 3\,\pi$. Jenseits von $\phi = 3\,\pi$ gelangt man ohne Unstetigkeit wieder auf das erste Blatt usw.

Auf der gesamten Riemann-Fläche gilt für die komplexe Wurzel

$$\sqrt{z} = \sqrt{|z|}\, e^{i\,\frac{\phi}{2}}. \tag{138}$$

Zur Beschreibung der Funktion $z^{\frac{1}{n}}$ ist eine n-blättrige Riemann-Fläche notwendig.

3 Gewöhnliche Differentialgleichungen II

Hier wird Bezug genommen auf das Vorwissen aus der Vorlesung "Mathematische Ergänzung zu Experimentalphysik I". Im Folgenden soll stets die Lösbarkeit der Differentialgleichung vorausgesetzt werden d.h. insbesondere die Existenz der auftretenden Integrale; auf zeitraubende Einschränkungen im Definitions- oder Wertebereich soll verzichtet und auf die Mathematikvorlesung oder auf Lehrbücher verwiesen werden.

3.1 Gewöhnliche Differentialgleichungen erster Ordnung

Definition: Die **gewöhnliche Differentialgleichung erster Ordnung** hat die allgemeine Form

$$y'(x) = f(y(x), x) \;\text{ oder kurz }\; y' = f(y, x) \;\text{ mit }\; y' = y'(x) = \frac{\mathrm{d}\,y}{\mathrm{d}\,x} = \frac{\mathrm{d}\,y(x)}{\mathrm{d}\,x}. \tag{139}$$

Gesucht ist die allgemeine Lösung $y(x)$ der Gleichung, gegebenenfalls mit der *Anfangsbedingung*

$$y(x_0) = y_0, \tag{140}$$

wobei das Wertepaar x_0, y_0 vorgegeben wird.

3.1.1 Die Differentialgleichung $\mathbf{y' = f(x)}$

Die Differentialgleichung lautet ausführlich

$$y'(x) = f(x). \tag{141}$$

Hier reduziert sich das Problem auf eine einfache Integration. Die allgemeine Lösung ist die Stammfunktion von $f(x)$, d.h. das Integral

$$y(x) = \int_{\xi}^{x} f(x')\,\mathrm{d}x'. \tag{142}$$

Charakteristisch für die Stammfunktion ist, dass ihre unabhängige Variable x in der Integrationsgrenze auftaucht. Dementsprechend sollte man der Integrationsvariablen eine andere Bezeichnung geben; hier wurde x' gewählt. Die untere Integrationsgrenze ξ ist beliebig wählbar. Hier sei zur Erinnerung noch einmal wiederholt, dass die Freiheit in der Wahl von ξ der freien Wahl einer Integrationskonstanten entspricht. Eine andere Wahl ξ' für die untere Integrationsgrenze führt auf eine Lösung, die sich von der ersten Lösung um das bestimmte Intregral $\int_{\xi}^{\xi'} f(x')\,\mathrm{d}x'$ unterscheidet,

$$\int_{\xi}^{x} f(x')\,\mathrm{d}x' = \int_{\xi'}^{x} f(x')\,\mathrm{d}x' + \int_{\xi}^{\xi'} f(x')\,\mathrm{d}x'. \tag{143}$$

Dieses bestimmte Integral ist eine Konstante und spielt die Rolle der Integrationskonstanten.

Mit der Anfangsbedingung $y(x_0) = y_0$ wird die Lösung

$$y(x) = y_0 + \int_{x_0}^{x} f(x')\,\mathrm{d}x'; \tag{144}$$

die Anfangsbedingung ist offensichtlich erfüllt, da für $x = x_0$ das Integral verschwindet. Es sei daran erinnert, dass das Wertepaar (x_0, y_0) beliebig vorgegeben werden kann, d.h. dass sowohl der Wert der unabhängigen Variablen x_0 als auch der Funktionswert $y(x_0) = y_0$ frei wählbar sind (innerhalb gewisser Grenzen, die von der Funktion $f(x)$ abhängen).

3.1.2 Die Differentialgleichung $\mathbf{y' = f(y)}$

Die Differentialgleichung lautet ausführlich

$$y'(x) = f(y(x)). \tag{145}$$

Hier kommt die Methode der *Separation der Variablen* zum Einsatz. Die Differentialgleichung läßt sich mit Hilfe von $y' = \dfrac{\mathrm{d}\,y}{\mathrm{d}\,x}$ umschreiben in eine Form, deren linke Seite nur von der Variablen y und deren rechte Seite nur von der Variablen x abhängt

$$\frac{\mathrm{d}\,y}{f(y)} = \mathrm{d}\,x. \tag{146}$$

Es können also unabhängig auf beiden Seiten die jeweiligen Stammfunktionen bestimmt werden

$$\int_{y_0}^{y} \frac{\mathrm{d}\,\tilde{y}}{f(\tilde{y})} = \int_{x_0}^{x} \mathrm{d}x' = x - x_0. \tag{147}$$

In die Lösung wurde – der Einfachheit halber – bereits die Anfangsbedingung $y(x_0) = y_0$ implementiert; die rechte Seite verschwindet für $x = x_0$, die linke für $y = y_0$. Da die linke Seite über die obere Integrationsgrenze eine Funktion von y ist, muss die Integrationsvariable eine andere Bezeichnung erhalten. Passend wäre y' gewesen; um jedoch der Gefahr einer Verwechslung mit der Ableitung y' vorzubeugen, wurde $\tilde{y}$ gewählt. Die gesuchte Lösungsfunktion $y(x)$ sitzt in der oberen Integrationsgrenze

$$x = x_0 + \int_{y_0}^{y} \frac{\mathrm{d}\,\tilde{y}}{f(\tilde{y})}. \tag{148}$$

In praktischen Anwendungen wird man in vielen Fällen in der Lage sein, die Integration analytisch auszuführen und die resultierende Gleichung nach $y(x)$ aufzulösen.

Beispiel: $y' = a\sqrt{y}$

Gleichung (147) wird zu

$$\int_{y_0}^{y} \frac{\mathrm{d}\,\tilde{y}}{\sqrt{\tilde{y}}} = a\,(x - x_0). \tag{149}$$

Mit $\displaystyle\int_{y_0}^{y} \frac{\mathrm{d}\,\tilde{y}}{\sqrt{\tilde{y}}} = 2\Big(\sqrt{y} - \sqrt{y_0}\Big)$ erhält man die Lösung

$$y(x) = \Big(\sqrt{y_0} + \frac{a}{2}\,(x - x_0)\Big)^2. \tag{150}$$

3.1.3 Die Differentialgleichung $\mathbf{y' = f(x)\,g(y)}$

Die Differentialgleichung lautet ausführlich

$$y'(x) = f(x)\,g\Big(y(x)\Big). \tag{151}$$

Hier führt Separation der Variablen sofort zu

$$\frac{\mathrm{d}\,y}{g(y)} = f(x)\,\mathrm{d}\,x \tag{152}$$

und nach Integration beider Seiten unter Implementation der Anfangsbedingung $y(x_0) = y_0$ erhält man

$$\int_{y_0}^{y} \frac{\mathrm{d}\,\tilde{y}}{g(\tilde{y})} = \int_{x_0}^{x} f(x')\,\mathrm{d}x'. \tag{153}$$

Hier ist die Lösung $y(x)$ wieder implizit gegeben; sie steht in der oberen Integrationsgrenze. In praktischen Anwendungen muss man sehen, ob man die beiden Integrationen analytisch ausführen kann und ob man die resultierende Gleichung explizit nach $y(x)$ auflösen kann.

3.1.4 Die Differentialgleichung $\mathbf{y' = f(A\,x + B\,y + C)}$

Die Differentialgleichung lautet ausführlich

$$y'(x) = f(A\,x + B\,y(x) + C) \tag{154}$$

mit beliebigen vorgegebenen Konstanten A, B und C.

Hier soll am Beispiel eine Methode vorgestellt werden, die sogenannte *Substitution*, die beim Lösen von Differentialgleichungen allgemein sehr wertvoll sein kann.

Im vorliegenden Beispiel empfiehlt es sich, den Ausdruck $A\,x+B\,y(x)+C$ durch eine Hilfsfunktion $u(x)$ zu substituieren

$$u(x) = A\,x + B\,y(x) + C. \tag{155}$$

Das erste Ziel ist nun, eine Differentialgleichung für die Hilfsfunktion $u(x)$ aufzustellen und diese zu lösen.

Aus (155) folgt

$$u'(x) = A + B\,y'(x). \tag{156}$$

Setzt man für $y'(x)$ die Differentialgleichung in der Form $y'(x) = f(u(x))$ ein, erhält man eine Differentialgleichung für $u(x)$

$$u'(x) = A + B\,f(u(x)). \tag{157}$$

Diese löst man wieder nach der Methode der Separation der Variablen

$$\int_{u_0}^{u} \frac{\mathrm{d}\,\tilde{u}}{A + B\,f(\tilde{u})} = \int_{x_0}^{x} \mathrm{d}\,x' = x - x_0 \quad \text{mit} \quad u_0 = A\,x_0 + b\,y_0 + C. \tag{158}$$

Die Lösungsfunktion $u(x)$ tritt wieder in der oberen Integrationsgrenze auf.

Nach erfolgter Integration ersetzt man $u(x)$ durch $A\,x + B\,y(x) + C$ und u_0 durch $A\,x_0 + b\,y_0 + C$ und löst nach $y(x)$ auf.

3.1.5 Die Differentialgleichung $\mathbf{y' = f\left(\frac{y}{x}\right)}$

Die Differentialgleichung lautet ausführlich

$$y'(x) = f\Big(\frac{y(x)}{x}\Big). \tag{159}$$

Der erste Schritt ist die Substitution

$$u(x) = \frac{y(x)}{x}. \tag{160}$$

Die Ableitung läßt sich mit Hilfe der Differentialgleichung $y'(x) = f(u(x))$ auswerten zu

$$u'(x) = \frac{x\,y'(x) - y(x)}{x^2} = \frac{1}{x}\Big(f(u(x)) - u(x)\Big). \tag{161}$$

Dies ist wieder eine Differentialgleichung für die Hilfsfunktion $u(x)$, die mit Hilfe der Separation der Variablen gelöst werden kann, wobei die Lösungsfunktion $u(x)$ wieder in der oberen Integrationsgrenze auftritt,

$$\int_{u_0}^{u} \frac{\mathrm{d}\,\tilde{u}}{f(\tilde{u}) - \tilde{u}} = \int_{x_0}^{x} \frac{\mathrm{d}\,x'}{x'} = \log\frac{x}{x_0}. \tag{162}$$

In diese Lösung wird nach Integration $u(x) = \dfrac{y(x)}{x}$ und $u_0 = \dfrac{y_0}{x_0}$ eingesetzt und nach der Lösungsfunktion $y(x)$ aufgelöst.

3.1.6 Die lineare Differentialgleichung erster Ordnung

Die lineare inhomogene Differentialgleichung erster Ordnung ist

$$y'(x) + f(x)\,y(x) + g(x) = 0 \quad \text{oder kurz} \quad y' + f(x)\,y + g(x) = 0. \tag{163}$$

mit beliebigen integrierbaren Funktionen $f(x)$ und $g(x)$. Sie ist von zentraler Bedeutung, da sie in vielen praktischen Anwendungen in Physik und Technik auftritt. Sie ist linear in dem Sinne, dass sowohl die Funktion $y(x)$ als auch ihre Ableitung $y'(x)$ linear auftreten. Dieser Ansatz schließt nichtlineare Terme wie z.B. $y'\,y$ oder y^2 aus. Sie ist inhomogen, da sie neben einem Term in $y'(x)$ und einem in $y(x)$ auch den inhomogenen Term $g(x)$ enthält.

Hier wird im Rückblick auf die Vorlesung "Mathematische Ergänzung zu Experimentalphysik I" daran erinnert, dass die Lösung einer inhomogenen linearen Differentialgleichung (163) folgendermaßen gewonnen wird.

- Es wird zunächst die allgemeine Lösung $y_h(x)$ der zugehörigen homogenen Differentialgleichung

  $$y'(x) + f(x)\,y(x) = 0 \tag{164}$$

  bestimmt. Für eine lineare Differentialgleichung erwartet man *eine* Fundamentallösung multipliziert mit einer frei wählbaren Konstanten.

- Dann wird eine spezielle Lösung $y_s(x)$ der inhomogenen Differentialgleichung (163) bestimmt.

- Die allgemeine Lösung der inhomogenen Differentialgleichung ist dann

  $$y(x) = y_h(x) + y_s(x). \tag{165}$$

  Es sei daran erinnert, dass dies impliziert, dass zwei verschiedene spezielle Lösungen der inhomogenen Gleichung sich um eine Lösung der homogenen Gleichung unterscheiden.

Die allgemeine Lösung der homogenen Differentialgleichung (164) mit der Anfangsbedingung $y(x_0) = y_0$ erhält man mit Hilfe der Separation der Variablen

$$\int_{y_0}^{y} \frac{\mathrm{d}\,\tilde{y}}{\tilde{y}} = -\int_{x_0}^{x} f(x')\,\mathrm{d}x' \quad \text{und auf der linken Seite ausintegriert} \quad \log\frac{y}{y_0} = -\int_{x_0}^{x} f(x')\,\mathrm{d}x', \tag{166}$$

d.h. die allgemeine Lösung der homogenen Differentialgleichung (164) ist

$$y_h(x) = y_0\, e^{-\int_{x_0}^{x} f(x')\,\mathrm{d}x'}. \tag{167}$$

Zur speziellen Lösung $y_s(x)$ der inhomogenen Differentialgleichung (163) kommt man, indem man in einem ersten Schritt die Lösung (167) umschreibt

$$y_h(x) = C\, e^{-\int^{x} f(x')\,\mathrm{d}x'} \tag{168}$$

mit einer frei wählbaren multiplikativen Konstanten C. Als nächstes wählt man die Methode der *Variation der Konstanten*, die bereits in der Vorlesung "Mathematische Ergänzung zu Experimentalphysik I" zum Einsatz kam: der folgende *Lösungsansatz*

$$y_s(x) = C(x)\, e^{-\int^{x} f(x')\,\mathrm{d}x'} \tag{169}$$

mit einer vorerst noch *unbekannten Funktion* $C(x)$ wird zum Erfolg führen.

Die Ableitung nach x erhält man nach der Produktregel

$$y_s'(x) = C'(x)\, e^{-\int^x f(x')\,\mathrm{d}x'} + C(x)\,\Big(-f(x)\Big)\, e^{-\int^x f(x')\,\mathrm{d}x'}; \tag{170}$$

hier wurde benutzt, dass $\dfrac{\mathrm{d}}{\mathrm{d}x}\displaystyle\int^x f(x')\,\mathrm{d}x' = f(x)$ ist.

Setzt man den Ansatz (169) für $y_s(x)$ und den Ausdruck (170) für $y_s'(x)$ in die inhomogene Differentialgleichung (163) ein, erhält man

$$C'(x) = -g(x)\, e^{\int^x f(x')\,\mathrm{d}x'}. \tag{171}$$

Auf beiden Seiten der Gleichung wird die Stammfunktion bestimmt unter angemessener Umbenennung der Integrationsvariablen

$$C(x) = \int^x C'(x')\,\mathrm{d}x' = -\int^x \mathrm{d}x'\, g(x')\; e^{\int^{x'} f(x'')\,\mathrm{d}x''}. \tag{172}$$

Die Variable x tritt entsprechend jeweils in der oberen Integrationsgrenze auf. Es ist zu beachten, dass auf der rechten Seite die Integration in der Exponentialfunktion bis zur oberen Integrationsgrenze x' geht, der Integrationsvariablen der äußeren Integration.

Der Ansatz (169) der Variation der Konstanten hat sich als erfolgreich erwiesen, da mit (172) eine Funktion $C(x)$ bestimmt werden konnte, die zu einer speziellen Lösung $y_s(x)$ der inhomogenen Differentialgleichung (163) führt; hier wird jetzt überall x_0 als untere Integrationsgrenze gewählt, was offensichtlich $y_s(x_0) = 0$ entspricht

$$y_s(x) = -e^{-\int_{x_0}^x f(x')\,\mathrm{d}x'} \int_{x_0}^x \mathrm{d}x'\, g(x')\; e^{\int_{x_0}^{x'} f(x'')\,\mathrm{d}x''}. \tag{173}$$

Dies führt auf die

allgemeine Lösung der linearen inhomogenen Differentialgleichung

$$y'(x) + f(x)\, y(x) + g(x) = 0 \tag{174}$$

$$y(x) = y_h(x) + y_s(x) = e^{-\int_{x_0}^x f(x')\,\mathrm{d}x'} \left(y_0 - \int_{x_0}^x g(x')\; e^{\int_{x_0}^{x'} f(x'')\,\mathrm{d}x''}\,\mathrm{d}x'\right), \tag{175}$$

die der Anfangsbedingung

$$y(x_0) = y_0 \tag{176}$$

genügt.

Beispiel

Die Koeffizientenfunktionen seien Konstanten a und b, d.h. die Differentialgleichung sei

$$y' + a\, y + b = 0. \tag{177}$$

Die zwei wichtigen Integrale in der Lösungsfunktion (175) sind mit $f(x) = a$ und $g(x) = b$

$$e^{-\int_{x_0}^{x} f(x')\,\mathrm{d}x'} = e^{-a\int_{x_0}^{x}\mathrm{d}x'} = e^{-a\,(x-x_0)}, \tag{178}$$

$$\int_{x_0}^{x} g(x')\,e^{\int_{x_0}^{x'} f(x'')\,\mathrm{d}x''}\,\mathrm{d}x' = \int_{x_0}^{x} b\,e^{a\,(x'-x_0)}\,\mathrm{d}x' = \frac{b}{a}\left(e^{a\,(x-x_o)} - 1\right). \tag{179}$$

Damit wird die Lösung (175)

$$y(x) = e^{-a\,(x-x_0)}\left(y_0 + \frac{b}{a}\right) - \frac{b}{a}. \tag{180}$$

3.2 Nachtrag zu gewöhnlichen Differentialgleichungen zweiter Ordnung

Es handelt sich um einen Nachtrag zur Vorlesung "Mathematische Ergänzung zu Experimentalphysik I", in der die lineare homogene Differentialgleichung zweiter Ordnung mit konstanten Koeffizienten behandelt wurde

$$y''(x) + a\,y'(x) + b\,y(x) = 0. \tag{181}$$

Die Behandlung läßt sich mit Hilfe der mittlerweile vermittelten Kenntnisse der komplexen Zahlen eleganter durchführen.

Der Lösungsansatz

$$y(x) = e^{\lambda\,x} \tag{182}$$

führt auf die charakteristische Gleichung

$$\lambda^2 + a\,\lambda + b = 0 \tag{183}$$

mit den beiden Lösungen

$$\lambda_{1,2} = -\frac{a}{2} \pm \sqrt{\frac{a^2}{4} - b}. \tag{184}$$

Die allgemeine Lösung von (181) für $a^2/4 - b \neq 0$ ergibt sich damit zu

$$y(x) = C_1\,e^{\lambda_1\,x} + C_2\,e^{\lambda_2\,x}. \tag{185}$$

mit beliebig wählbaren Konstanten C_1 und C_2.

Für positive Diskriminante, $a^2/4 - b > 0$, erhält man das wohlbekannte Resultat

$$y(x) = e^{-\frac{a}{2}\,x}\left(C_1\,e^{\sqrt{\frac{a^2}{4}-b}\,x} + C_2\,e^{-\sqrt{\frac{a^2}{4}-b}\,x}\right). \tag{186}$$

Für negative Diskriminante, $a^2/4 - b < 0$, erhält man Exponentialfunktionen mit imaginären Exponenten

$$y(x) = e^{-\frac{a}{2}\,x}\left(C_1\,e^{i\,\sqrt{b-\frac{a^2}{4}}\,x} + C_2\,e^{-i\,\sqrt{b-\frac{a^2}{4}}\,x}\right). \tag{187}$$

Die Benutzung der Eulerformel, $e^{\pm i\,\phi} = \cos\phi \pm i\,\sin\phi$, führt auf die komplexe Lösung

$$y(x) = e^{-\frac{a}{2}\,x}\left((C_1 + C_2)\,\cos\sqrt{b - \frac{a^2}{4}}\,x + i\,(C_1 - C_2)\,\sin\sqrt{b - \frac{a^2}{4}}\,x\right). \tag{188}$$

mit beliebig wählbaren komplexen Konstanten C_1, C_2.

Konzentriert man sich, wie in physikalischen Anwendungen typisch angebracht, auf reelle Lösungen, dann kommt man mit

$$\tilde{C}_1 = \mathrm{Re}\,(C_1 + C_2) \quad \text{und} \quad \tilde{C}_2 = \mathrm{Re}\Big(i\,(C_1 - C_2)\Big) \tag{189}$$

auf die Lösung

$$y(x) = e^{-\frac{a}{2}x}\left(\tilde{C}_1\,\cos\sqrt{\frac{a^2}{4} - b}\,x + \tilde{C}_2\,\sin\sqrt{\frac{a^2}{4} - b}\,x\right) \tag{190}$$

mit beliebig wählbaren reellen Konstanten $\tilde{C}_1$ und $\tilde{C}_2$.

Der Ausdruck (185) für die allgemeine Lösung mit Exponentialfunktionen führt also elegant über komplexe Lösungen auf die bereits bekannten Ergebnisse (186) und (190).

4 Partielle Differentialgleichungen – die Wellengleichung

4.1 Definition von partiellen Differentialgleichungen und physikalische Motivation

Ein Charakteristikum der bisher behandelten gewöhnlichen Differentialgleichungen ist, dass nur *eine einzige unabhängige Variable* auftritt, im Allgemeinen mit x oder t bezeichnet. Bei gewöhnlichen Differentialgleichungen tritt eine Funktion von x samt ihren Ableitungen auf

$$y(x),\ y'(x),\ y''(x),\ ..., \tag{191}$$

bei Systemen von gekoppelten gewöhnlichen Differentialgleichungen ein Satz von solchen Funktionen und ihren Ableitungen

$$y_1(x),\ y_1'(x),\ ... \tag{192}$$

$$y_2(x),\ y_2'(x),\ ... \tag{193}$$

$$\vdots$$

Mit einer unabhängigen Variablen kommt man in der Physik nicht aus; unsere Welt ist nicht eindimensional und im Allgemeinen nicht stationär. In vielen Fällen hat man es mit Feldern zu tun, die von den drei Ortsvariablen x_1, x_2, x_3 und der Zeit t abhängen. Da physikalische Gesetze vorzugsweise in Form von Differentialgleichungen formuliert sind, wird man so auf partielle Differentialgleichungen geführt, die tatsächlich einen großen Bereich der physikalischen Gesetzmäßigkeiten beherrschen.

Definition: **Eine partielle Differentialgleichung** ist eine Differentialgleichung für eine Funktion von mehreren Variablen, d.h. für ein Feld, in der **partielle Ableitungen nach mehreren Variablen** vorkommen. Die Ordnung der partiellen Differentialgleichung ist durch die Ordnung der höchsten vorkommenden Ableitung bestimmt.

Als Beispiel sei die sogenannte *eindimensionale Wellengleichung* vorweggenommen. Man hat es mit einer Funktion $f(x,t)$ zu tun, also einem Feld mit zwei Variablen, einer Raumdimension x und der Zeit t. Die Wellengleichung lautet

$$\frac{\partial^2}{\partial\, t^2}\, f(x,t) = v^2\, \frac{\partial^2}{\partial\, x^2}\, f(x,t), \tag{194}$$

wobei v eine Konstante mit der Dimension einer Geschwindigkeit ist. Die Wellengleichung stellt eine Beziehung her zwischen den zweifachen partiellen Ableitungen des Feldes $f(x,t)$ nach dem Ort und nach der Zeit, sie ist eine partielle Differentialgleichung zweiter Ordnung.

Bei gewöhnlichen Differentialgleichungen sind in vielen Fällen allgemeine Lösungen bestimmbar mit beliebig wählbaren Konstanten, die durch Anfangsbedingungen festgelegt werden können. Dies gilt nicht mehr für partielle Differentialgleichungen. Während es für partielle Differentialgleichungen erster Ordnung noch Standardverfahren gibt, muss man bei partiellen Differentialgleichungen zweiter Ordnung, die in der Physik die Hauptrolle spielen, individuell vorgehen. Ein Grund dafür ist, dass ihre Lösungen stark abhängig von der Wahl sogenannter *Randbedingungen* sind; da man es mit Feldern zu tun hat, die von mehreren Variablen abhängen, hat man es auch mit einem mehrdimensionalen Definitionsgebiet zu tun; die Randbedingungen schreiben ein Verhalten der Lösung auf den Rändern dieses Definitionsgebietes vor. In der Physik ergeben sich die Randbedingungen jeweils aus der spezifischen Problemstellung.

Das klingt relativ abstrakt. Einfache Beispiele aus der Physik können das näherbringen.

- Eine schwingungsfähige Membran sei an ihrem äußeren Rand in einen Rahmen eingespannt; bei einer Trommel ist er kreisförmig; man kann sich abstrakt aber auch z.B. einen ellipsenförmigen oder einen quadratischen Rahmen vorstellen. Jeder dieser Rahmen stellt eine andere Randbedingung für die eindimensionale Begrenzung der zur Schwingung angeregten zweidimensionalen Membran dar. Die Schwingungsamplitude ist als Funktion der zwei Raumdimensionen x_1, x_2 und der Zeit t ein Feld $f(x_1, x_2, t)$. Obwohl in allen diesen Fällen dieselbe partielle Differentialgleichung für das Feld $f(x_1, x_2, t)$ zu lösen ist, fallen aufgrund der verschiedenen Randbedingungen die Lösungen in Form von Paaren von "Eigenfrequenzen" und Knotenlinien (Linien auf der Membran, längs denen die Schwingungsamplitude verschwindet) jeweils verschieden aus. Anfangsbedingungen, z.B. durch den Anschlag der Membran mit einem Klöppel zur Zeit $t = t_0$ realisiert, hängen nicht nur von der Stärke und Geschwindigkeit des Anschlags sondern auch vom Anschlagsort auf der Membran ab. Beim Anschlag eines Schlagzeugs mit einem Besen hat man es mit vielen solchen Anschlagsorten zu tun.

- Eine schwingungsfähige dünne Metallplatte mit einer kreisförmigen oder quadratischen Begrenzung sei in ihrer Mitte eingespannt und gleichmäßig mit feinem Sand bestreut. Wird sie zu Schwingungen angeregt, sammelt sich der Sand in den Schwingungsknotenlinien; diese werden dadurch sichtbar gemacht. Die entstehenden sogenannten Chladnyschen Klangfiguren sind – bei identischer partieller Differentialgleichung – wieder charakteristisch für die Randbedingungen, d.h. für die Form der Begrenzung der schwingenden Platte.

- Verschieden geformte elektrisch geladene Leiter seien im Raum verteilt. Die Randbedingungen an das skalare elektromagnetische Potential $V(x_1, x_2, x_3)$ auf der Oberfläche der Leiter bestimmen sehr stark die Lösung der zugehörigen, aus der entsprechenden Maxwellgleichung abgeleiteten partiellen Differentialgleichung für $V(x_1, x_2, x_3)$.

- Besonders faszinierend ist der Effekt von Randbedingungen in der Quantenmechanik; Lösungen der relevanten partiellen Differentialgleichung (Schrödingergleichung), die kontinuierliche Werte für Größen wie Energie oder Drehimpuls zulassen, werden unter dem Einfluss von physikalischen Randbedingungen auf Lösungen mit gequantelten Energiewerten bzw. Drehimpulsen reduziert.

4.2 Physikalische Motivation für die Wellengleichung

Unter einer *Welle* versteht man *eine charakteristische räumliche Ausbreitung* eines Feldes als Funktion der Zeit. Zugrunde liegen partielle Differentialgleichungen mit Wellenlösungen.

- In der klassischen Mechanik kennt man z.B. die Seilwelle, die Wasserwelle, die Schallwelle, bei der Druckschwankungen der Luft durch den Raum laufen.
- In der Elektrodynamik tritt elektromagnetische Strahlung auf, bei der sich der elektrische Feldstärkevektor $\vec{E}(\vec{x},t)$ und der magnetische Feldstärkevektor $\vec{B}(\vec{x},t)$ als Funktionen der Zeit im Raum ausbreiten (z.B. Hertzscher Dipol).
- In der Quantenmechanik regiert der Teilchen-Welle-Dualismus. Ein Teilchen wird in der Schrödingergleichung durch eine Wellenfunktion beschrieben, (wenngleich die Schrödingergleichung nicht die Standardform der klassischen Wellengleichung hat).

4.3 Eindimensionale harmonische Welle am Beispiel der Seilwelle

Eine heuristische Einführung soll am Beispiel einer Seilwelle gegeben werden. Ein Seil werde idealisiert rechts "im Unendlichen" reflexionsfrei fixiert und am freien linken Ende harmonisch, d.h. entsprechend einer Kosinus-Funktion, auf und ab bewegt. Die Funktion $f(x,t)$, die die Auslenkung des Seils an einem beliebigen Ort x zur Zeit t beschreibt, ist in Form von drei Schnappschüssen zu aufeinanderfolgenden Zeiten t_0, $t_1 > t_0$ und $t_2 > t_1$ in Abb. 14 als Funktion von x dargestellt. Es breitet sich als Funktion der Zeit eine harmonische Welle aus; die Ausbreitungsrichtung ist in diesem Beispiel die positive x-Richtung.

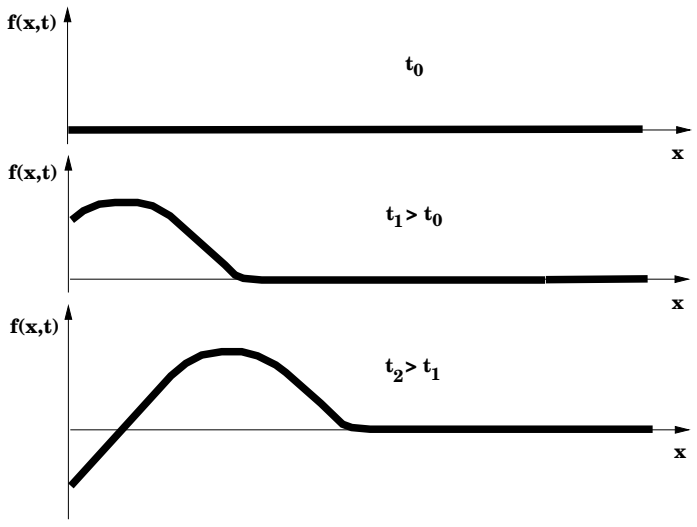


Abbildung 14:

Folgende Größen gehen in die Bescheibung einer eindimensionalen harmonischen Welle ein

- die *Wellenlänge* λ, die gleich dem örtlichen Abstand zweier aufeinanderfolgender Maxima bei konstanter (festgehaltener) Zeit ist,
- die *Schwingungsdauer* T, die Dauer einer Schwingungsperiode an einem konstanten (festgehaltenen) Ort, sowie die aus beiden abgeleiteten Größen,

- die *Wellenzahl* $k = \dfrac{2\,\pi}{\lambda}$,
- die *Frequenz* $\nu = \dfrac{1}{T}$ oder
- die *Kreisfrequenz* $\omega = 2\,\pi\,\nu = \dfrac{2\,\pi}{T}$ und
- die *Phasengeschwindigkeit*, die sog. *Ausbreitungsgeschwindigkeit*, mit der sich eine ausgezeichnete Stelle der Welle, z.B. ein Maximum, in x-Richtung fortbewegt

$$v = \frac{\lambda}{T} = \lambda\nu = \frac{\omega}{k}. \tag{195}$$

Diese eindimensionale harmonische Welle wird beschrieben durch die

eindimensionale harmonische Wellenfunktion einer **nach rechts** laufenden Welle

$$\begin{aligned} f(x,t) &= A\,\cos\left(2\pi(\frac{t}{T} - \frac{x}{\lambda}) + \phi_0\right) && (196)\\ &= A\,\cos(\omega t - kx + \phi_0), && (197)\end{aligned}$$

wobei A die Maximalamplitude der Welle ist und ϕ_0 eine beliebige konstante Phase.

Die entsprechende Wellenfunktion einer *nach links* laufenden Welle erhält man durch die Ersetzung $x \to -x$

$$f(x,t) = A\,\cos(\omega t + kx + \phi_0). \tag{198}$$

Wenn es Rechenvorteile bringt, wie sehr oft, kann man die Wellenfunktion auch komplex ergänzen

Komplexe Ergänzung der eindimensionalen harmonischen Wellenfunktion

$$f(x,t) = A\,e^{i\,(\omega t \pm kx + \phi_0)}. \tag{199}$$

Die eindimensionale Wellenfunktion ist eine Funktion von zwei unabhängigen Variablen, des Orts x und der Zeit t, also ein Feld.

4.4 Die eindimensionale Wellengleichung

4.4.1 Herleitung

Die eindimensionale Wellengleichung soll am Beispiel der Seilwelle hergeleitet werden. Die Seilkraft in einem gekrümmten Seil hat überall denselben Betrag F_0, ist jedoch in jedem Punkt tangentiell gerichtet. Die zur Zeit t im Punkt x angreifende Seilkraft bilde mit der Horizontalen den Winkel α, der eine Funktion von x und t ist, $\alpha(x,t)$. Zunächst wird die Zeit t als konstant betrachtet. Wie in Abb. 15 abgebildet, läßt sich die tangentielle Seilkraft in jedem Punkt x in eine horizontale Komponente $F_\parallel$ und eine vertikale Komponente $F_\perp$ zerlegen.

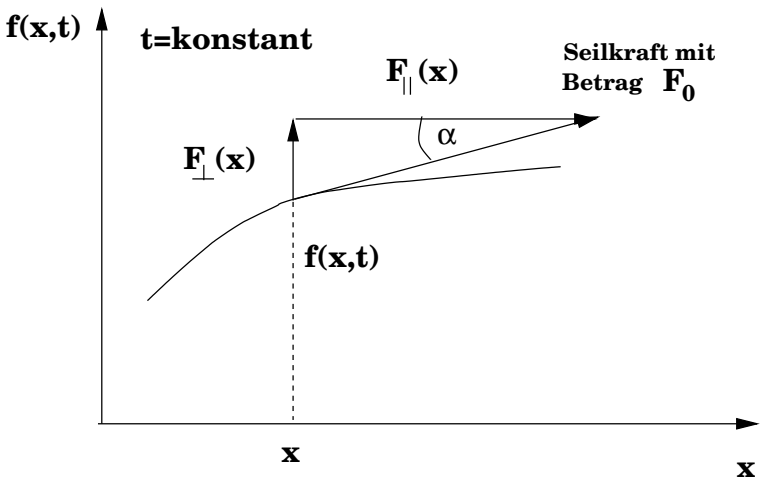


Abbildung 15:

In Abb. 16 wird in einem Schnappschuss zur Zeit t ein (inifinitesimales) Seilelement zwischen x und $x+\mathrm{d}\,x$ betrachtet. Aus den in x und $x+\mathrm{d}\,x$ angreifenden Kräften ergibt sich eine resultierende vertikale Kraft $\mathrm{d}F_\perp$, die zur Ruhelage zurücktreibt,

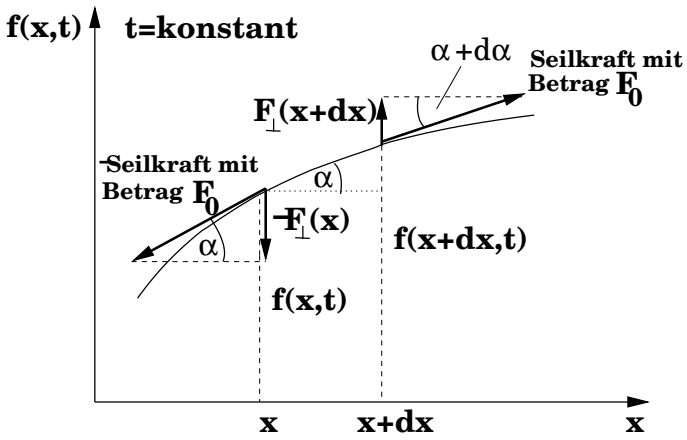


Abbildung 16:

$$\mathrm{d}F_\perp = F_\perp(x+\mathrm{d}\,x) - F_\perp(x) = F_0\,\Big(\sin(\alpha+\mathrm{d}\,\alpha) - \sin\alpha\Big). \tag{200}$$

In der Approximation kleiner Auslenkung, d.h. kleiner Winkel α, gilt $\sin\alpha \approx \alpha$, d.h. approximativ

$$\mathrm{d}\,F_\perp = F_0\,\mathrm{d}\,\alpha. \tag{201}$$

Nun soll $\alpha(x,t)$ mit Hilfe der Amplitude $f(x,t)$ der Seilwelle ausgedrückt werden. Bei konstant gehaltener Zeit t liest man die Steigung $\tan\alpha$ aus Abb. 16 ab

$$\frac{\partial\, f(x,t)}{\partial\, x} = \lim_{\mathrm{d}\,x\to 0} \frac{f(x+\mathrm{d}\,x,t) - f(x,t)}{\mathrm{d}\,x} = \tan\alpha \approx \alpha. \tag{202}$$

Für das Differential $\mathrm{d}\alpha$ zum Zeitpunkt t gilt schließlich nach (202)

$$\mathrm{d}\,\alpha = \frac{\partial\,\alpha(x,t)}{\partial\, x}\,\mathrm{d}\,x = \frac{\partial^2 f(x,t)}{\partial\, x^2}\,\mathrm{d}\,x. \tag{203}$$

Auf das betrachtete Seilelement wirkt also nach (201) die rücktreibende Kraft

$$\mathrm{d}\,F_\perp = F_0\,\frac{\partial^2 f(x,t)}{\partial\,x^2}\,\mathrm{d}\,x. \tag{204}$$

Andererseits ist die rücktreibende Kraft $F_\perp$ entsprechend dem zweiten Newtonschen Gesetz gleich dem Produkt aus der Masse und der Beschleunigung des Seilelements. Die Masse des Seilelements bei konstanter Massendichte $\mathrm{d}\,m/\mathrm{d}\,x = \rho$ und Länge $\mathrm{d}\,x$ ist $\mathrm{d}\,m = \rho\,\mathrm{d}\,x$; die Beschleunigung ist die zweite Ableitung der Auslenkung $f(x,t)$ *nach der Zeit*. Damit wird

$$\mathrm{d}\,F_\perp = \rho\,\mathrm{d}\,x\,\frac{\partial^2 f(x,t)}{\partial\,t^2}. \tag{205}$$

Die Gleichheit der beiden Ausdrücke (204) und (205) führt auf die Wellengleichung der Seilwelle

$$\frac{\partial^2 f(x,t)}{\partial\,t^2} = \frac{F_0}{\rho}\,\frac{\partial^2 f(x,t)}{\partial\,x^2} \tag{206}$$

oder in der Schreibweise der Physiker

$$\ddot{f}\,(x,t) = \frac{F_0}{\rho}\,f''(x,t). \tag{207}$$

F_0/ρ ist eine konstante Größe, eine Materialkonstante des Seils, mit der Dimension eines Geschwindigkeitsquadrats. Allgemein lautet die

Eindimensionale Wellengleichung

$$\frac{\partial^2 f(x,t)}{\partial\,t^2} = v^2\,\frac{\partial^2 f(x,t)}{\partial\,x^2}. \tag{208}$$

Dies ist eine partielle Differentialgleichung zweiter Ordnung in den Variablen x und t. Die Konstante v hat die Dimension einer Geschwindigkeit.

4.4.2 Spezielle Lösungen

Die harmonische Welle (197) bzw. (198) ist eine spezielle Lösung dieser Wellengleichung. Mit dem Lösungsansatz

$$f(x,t) = A\,\cos(\omega\,t \pm k\,x + \phi_0) \tag{209}$$

werden

$$\frac{\partial^2 f(x,t)}{\partial\,t^2} = -A\,\omega^2\,\cos(\omega\,t \pm k\,x + \phi_0), \tag{210}$$

$$\frac{\partial^2 f(x,t)}{\partial\,x^2} = -A\,k^2\,\cos(\omega\,t \pm k\,x + \phi_0); \tag{211}$$

also ist mit

$$v = \frac{\omega}{k} = \frac{\lambda}{T} \tag{212}$$

(209) eine Lösung der Wellengleichung (208). Der Parameter v in (208) hat dann die Interpretation der Phasengeschwindigkeit oder *Ausbreitungsgeschwindigkeit* der Welle.

Zusammengefaßt

Die **harmonischen Wellenfunktionen**

$$f(x,t) = A\,\cos(\omega\,t \pm k\,x + \phi_0) \tag{213}$$

sind **spezielle Lösungen der eindimensionalen Wellengleichung**

$$\frac{\partial^2 f(x,t)}{\partial\,t^2} = v^2\,\frac{\partial^2 f(x,t)}{\partial\,x^2} \tag{214}$$

mit Ausbreitungsgeschwindigkeit der Welle

$$v = \frac{\omega}{k}. \tag{215}$$

Die allgemeine Form der Lösung der Wellengleichung (208)

$$f(x,t) = F(v\,t \pm x) \tag{216}$$

mit einer beliebigen zweimal differenzierbaren Funktion F wird in den Übungen hergeleitet.

4.5 Randbedingungen, stehende eindimensionale Wellen und Anfangsbedingungen

4.5.1 Einführung von Randbedingungen

Hier soll der starke Einfluss von Randbedingungen auf Lösungen von partiellen Differentialgleichungen beispielhaft vorgeführt werden.

Es wird ein Beispiel aus der Musik von Saiteninstrumenten gewählt: eine schwingende Saite der Länge L, die an beiden Enden, $x = 0$ und $x = L$, eingespannt ist. Dieses physikalische Problem ist durch die folgenden Angaben mathematisch definiert.

Die Auslenkung $f(x,t)$ der Saite am Ort x zur Zeit t

- genügt der **partiellen Differentialgleichung** (208), der Wellengleichung,
- sie erfüllt die **Randbedingungen**

$$f(0,t) \equiv 0 \quad \text{und} \quad f(L,t) \equiv 0 \quad \textit{für alle } t. \tag{217}$$

Es wird ein Lösungsansatz in Form eines Produktes

$$f(x,t) = g(x)\,h(t) \tag{218}$$

gewählt, der sich als besonders geeignet ist für die anschließende Implementation der Randbedingungen erweist. Er führt folgendermaßen auf die *Separation der Variablen* x und t

$$\frac{\partial^2 f(x,t)}{\partial\,t^2} \;=\; g(x)\,\frac{\mathrm{d}^2}{\mathrm{d}\,t^2}\,h(t) \tag{219}$$

$$\frac{\partial^2 f(x,t)}{\partial x^2} = h(t)\,\frac{\mathrm{d}^2}{\mathrm{d}x^2}\,g(x). \tag{220}$$

Eingesetzt in (208) führt dies auf

$$g(x)\;\ddot{h}\,(t) = v^2\,g''(x)\,h(t) \tag{221}$$

bzw. auf

$$\frac{1}{v^2}\,\frac{\ddot{h}\,(t)}{h(t)} = \frac{g''(x)}{g(x)}. \tag{222}$$

Die linke Seite ist unabhängig von x, die rechte Seite ist unabhängig von t; die Gleichung muß jedoch für alle x und t aus dem Definitionsbereich gelten. Beide Seiten müssen also gleich derselben Konstanten sein. Die Konstante werde negativ gewählt, $-K$ mit $K > 0$. Diese Wahl führt auf periodische Lösungen und erlaubt, die Randbedingungen (217) zu erfüllen; (man kann sich leicht davon überzeugen, dass dies für eine positive Konstante nicht möglich ist).

Damit reduziert sich die partielle Differentialgleichung (208) auf zwei entkoppelte gewöhnliche Differentialgleichungen, eine in t und eine in x

$$\ddot{h}\,(t) + K\,v^2\,h(t) = 0, \tag{223}$$
$$g''(x) + K\,g(x) = 0 \tag{224}$$

mit den allgemeinen Lösungen

$$h(t) = A\,\cos\left(v\,\sqrt{K}\,t\right) + B\,\sin\left(v\,\sqrt{K}\,t\right) \tag{225}$$
$$g(x) = C\,\cos\left(\sqrt{K}\,x\right) + D\,\sin\left(\sqrt{K}\,x\right), \tag{226}$$

wobei, A, B, C und D beliebig wählbare Konstanten sind. Es sei betont, dass dies Lösungen für jeden positiven Wert von K sind; wie sich im Folgenden erweist, sind es die Randbedingungen, die diese Lösungsmannigfaltigkeit stark einschränken werden.

Die Randbedingungen (217) führen im Lösungsansatz (218) auf $g(0) = 0$ und $g(L) = 0$ und diese auf

$$C = 0, \quad \text{und} \tag{227}$$
$$\sin\left(\sqrt{K}\,L\right) = 0, \quad \text{d.h.} \quad \sqrt{K}\,L = n\,\pi, \tag{228}$$

wobei n eine ganze Zahl ist. Es gibt also nicht mehr Lösungen für jeden positiven Wert von K, sondern die Randbedingungen zeichnen *abzählbar unendlich viele* Lösungen aus mit den spezifischen Werten

$$K_n = \left(\frac{n\,\pi}{L}\right)^2 \quad \text{für ganzzahlige Werte von } n \tag{229}$$

mit den zugehörigen Lösungsfunktionen

$$\begin{aligned} f_n(x,t) &= h_n(t)\,g_n(x) \\ &= \left(A_n\,\cos\left(\frac{vn\pi}{L}\,t\right) + B_n\,\sin\left(\frac{vn\pi}{L}\,t\right)\right)\,\sin\left(\frac{n\pi}{L}\,x\right) \\ &= \tilde{A}_n\,\cos\left(\frac{vn\pi}{L}\,t + \phi_n\right)\,\sin\left(\frac{n\pi}{L}\,x\right). \end{aligned} \tag{230}$$

Die frei wählbaren Konstanten A_n, B_n oder äquivalent $\tilde{A}_n$, ϕ_n können durch Anfangsbedingungen festgelegt werden. Dies wird in Kap. 4.5.4 an einem Beispiel vorgeführt.

4.5.2 Diskussion der Lösungen

Die eingespannte Saite kann also nicht mit beliebiger Kreisfrequenz schwingen, sondern nur mit den sogenannten

Eigenfrequenzen oder charakteristischen Frequenzen:

$$\omega_n = \frac{vn\pi}{L}. \tag{231}$$

Die zugehörigen charakteristischen Wellenzahlen bzw. Wellenlängen sind

$$k_n = \frac{n\pi}{L} \quad \text{bzw.} \quad \lambda_n = \frac{2L}{n}. \tag{232}$$

Dieses Phänomen nennt man *stehende Wellen*, da die Randbedingungen der Saite Eigenschwingungen aufzwingen, die durch die Ortsfunktion $\sin\left(\frac{n\pi}{L}\,x\right)$ festgelegt sind und deren Amplitude als Funktion der Zeit harmonisch zwischen dem jeweiligen Maximalwert $\tilde{A}_n$ und dem jeweiligen Minimalwert $-\tilde{A}_n$ variiert.

Man unterscheidet unter diesen sog. *Normalschwingungen* die

$$\text{Grundschwingung für} \quad n = 1, \tag{233}$$

$$\text{erste Oberschwingung für} \quad n = 2, \tag{234}$$

$$\text{zweite Oberschwingung für} \quad n = 3, \quad \text{usw.} \tag{235}$$

Die Frequenz bzw. Wellenzahl der n-ten Oberschwingung sind also durch das jeweils n-Fache der Frequenz bzw. Wellenzahl der Grundschwingung gegeben.

Die Ortsfunktionen $\sin\left(\frac{n\pi}{L}\,x\right)$ sind in Abb. 17 für $n = 1,\ 2,\ 3$ dargestellt. Sie erfüllen die Randbedingungen der an den Enden eingespannten Saite offensichtlich und haben zusätzliche ortsfeste, zeitunabhängige Nullstellen, sogenannte *Knoten* oder *Schwingungsknoten* der Schwingung bei

$$x = \frac{m}{n}\,L, \quad \text{für} \quad m = 1,\ ...,\ n-1. \tag{236}$$

Zwischen je zwei benachbarten Nullstellen liegen die sogenannten *Schwingungsbäuche*, die mit der Maximalamplitude schwingen. Eine weitergehende Diskussion wird in den Übungen stattfinden.

Die allgemeine Lösung des durch die partielle Differentialgleichung (208) und die Randbedingungen (217) definierten Problems ist je nach Anfangsbedingungen eine Überlagerung, d.h. Linearkombination, der Normalschwingungungslösungen (230).

Allgemeine Lösung der an beiden Enden eingespannten schwingenden Saite:

$$f(x,t) = \sum_{n=1}^{\infty} \left(A_n \cos\left(\frac{vn\pi}{L}\,t\right) + B_n \sin\left(\frac{vn\pi}{L}\,t\right)\right) \sin\left(\frac{n\pi}{L}\,x\right) \tag{237}$$

$$= \sum_{n=1}^{\infty} \tilde{A}_n \cos\left(\frac{vn\pi}{L}\,t + \phi_n\right) \sin\left(\frac{n\pi}{L}\,x\right). \tag{238}$$

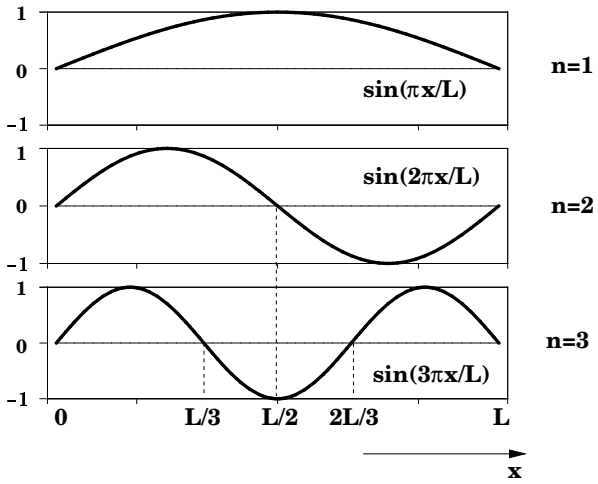


Abbildung 17:

Dabei sind

$$A_n = \tilde{A}_n \cos(\phi_n), \tag{239}$$
$$B_n = -\tilde{A}_n \sin(\phi_n). \tag{240}$$

Beim Geigenspiel oder Gitarrenspiel ist die aktive Länge der Saite und damit die Frequenzen der Grundschwingung des erzeugten Tones und aller Oberschwingungen durch den Abstand L zwischen dem Steg und dem auf dem Griffbrett aufgesetzten Finger gegeben. Verschiedene Überlagerungen von Grund- und Oberschwingungen äussern sich in der Musik in verschiedenen Klangfarben eines Tones.

4.5.3 Zusammenhang zwischen stehenden und laufenden Wellen

Die stehende Welle läßt sich als Überlagerung einer nach links und einer nach rechts laufenden harmonischen Welle gleicher Frequenz, Wellenzahl, Phasenverschiebung und Amplitude folgendermaßen verstehen

$$f_n(x,t) = \frac{\tilde{A}_n}{2}\Big(\sin(\frac{vn\pi}{L}\,t + \frac{n\pi}{L}\,x + \phi_n) - \sin(\frac{vn\pi}{L}\,t - \frac{n\pi}{L}\,x + \phi_n)\Big) \tag{241}$$
$$= \frac{\tilde{A}_n}{2}\Big(\sin(\frac{vn\pi}{L}\,t + \phi_n)\,\cos(\frac{n\pi}{L}\,x) + \cos(\frac{vn\pi}{L}\,t + \phi_n)\,\sin(\frac{n\pi}{L}\,x) \tag{242}$$
$$- \sin(\frac{vn\pi}{L}\,t + \phi_n)\,\cos(\frac{n\pi}{L}\,x) + \cos(\frac{vn\pi}{L}\,t + \phi_n)\,\sin(\frac{n\pi}{L}\,x)\Big) \tag{243}$$
$$= \tilde{A}_n\,\cos(\frac{vn\pi}{L}\,t + \phi_n)\,\sin(\frac{n\pi}{L}\,x). \tag{244}$$

4.5.4 Anfangsbedingungen

Mit Hilfe von *Anfangsbedingungen* können die Konstanten A_n und B_n in der allgemeinen Lösung (237) bzw. $\tilde{A}_n$ und ϕ_n in der äquivalenten Form (238) festgelegt werden. Anders als bei gewöhnlichen Differentialgleichungen muss dazu eine *Vorgabe der gesamten x-Abhängigkeit* von $f(x,t_0)$ und von $\frac{\partial}{\partial t} f(x,t)_{|t=t_0}$ zu einem Zeitpunkt $t = t_0$ gemacht werden.

Als Beispiel sei die Anfangsbedingung gewählt, dass die Saite zur Zeit $t = 0$ in der Mitte gezupft, d.h. bis zu einer Höhe h angehoben, und dann losgelassen wird; $f(x, 0)$ ist in Abb. 18 dargestellt. Mathematisch heißt dies

$$f(x,0) = \begin{cases} \frac{2h}{L}\, x & \text{für } 0 \le x \le \frac{L}{2} \\ 2h - \frac{2h}{L}\, x & \text{für } \frac{L}{2} \le x \le L \end{cases}, \tag{245}$$

$$\frac{\partial}{\partial t}\, f(x,t)_{|t=0} \equiv 0 \quad \text{für alle } x. \tag{246}$$

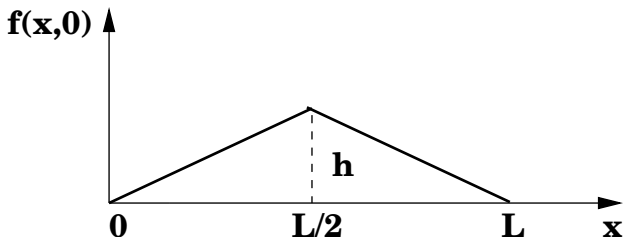


Abbildung 18:

Die Anfangsbedingung muss natürlich konform mit der Randbedingung (217) sein, d.h. $f(0, 0) = 0$ und $f(L, 0)$ erfüllen.

Die Anfangsbedingung (246) führt in der allgemeinen Lösung (237) zu $B_n = 0$ für alle n; die Anfangsbedingung (245) führt – ohne Beweis – in der allgemeinen Lösung (237) zu

$$A_n = \frac{8h}{\pi^2}\, \frac{(-1)^n}{(2n+1)^2}. \tag{247}$$

Der Beweis kann mit den an dieser Stelle noch nicht vorhandenen Kenntnissen aus Kap. 6 geführt werden.

4.6 Die dreidimensionale Wellengleichung

4.6.1 Definition

Die dreidimensionale Wellengleichung findet vielfach Anwendung in der Physik. Prominente Beispiele werden im folgenden Kap. 4.7 vorgestellt.

Ausgangspunkt ist ein skalares Feld $f(\vec{r}, t)$ der vier Variablen

$$\vec{r} = (x_1,\, x_2,\, x_3) = \text{Ortsvektor} \quad \text{und} \quad t = \text{Zeit}. \tag{248}$$

Die **dreidimensionale Wellengleichung** ist

$$\frac{\partial^2 f(\vec{r},t)}{\partial t^2} = v^2\, \Delta\, f(\vec{r},t), \tag{249}$$

wobei an die Definitionen

$$\Delta = \vec{\nabla} \cdot \vec{\nabla} = \text{div}\,\text{grad} = \frac{\partial^2}{\partial\, x_1^2} + \frac{\partial^2}{\partial\, x_2^2} + \frac{\partial^2}{\partial\, x_3^2} \tag{250}$$

und

$$\vec{\nabla} = (\frac{\partial}{\partial\, x_1}, \frac{\partial}{\partial\, x_2}, \frac{\partial}{\partial\, x_3}) \tag{251}$$

erinnert wird.

Dies ist eine partielle Differentialgleichung zweiter Ordnung in den vier Variablen x_1, x_2, x_3 und t und eine offensichtliche Verallgemeinerung der eindimensionalen Wellengleichung (208).

4.6.2 Ebene Wellenlösungen

In Analogie zur speziellen harmonischen Lösung (197) der eindimensionalen Wellengleichung (208) wird der folgende spezielle Lösungsansatz für die dreidimensionale Wellengleichung (249) gemacht

$$f(\vec{r}, t) = A\, \cos(\omega\, t - \vec{k} \cdot \vec{r} + \phi_0). \tag{252}$$

Der Vektor $\vec{k}$ ist der *Wellenzahlvektor* mit drei Komponenten $\vec{k} = (k_1, k_2, k_3)$; sein Betrag $|\vec{k}| = \sqrt{k_1^2 + k_2^2 + k_3^2} = k$ ist die *Wellenzahl*, die durch

$$k = |\vec{k}| = \frac{2\pi}{\lambda} \tag{253}$$

zur *Wellenlänge* in Beziehung steht. Das auftretende Produkt $\vec{k} \cdot \vec{r}$ ist das Skalarprodukt

$$\vec{k} \cdot \vec{r} = k_1\, x_1 + k_2\, x_2 + k_2\, x_2. \tag{254}$$

Die relevanten zweifachen partiellen Ableitungen sind

$$\frac{\partial^2}{\partial\, x_i^2}\, f(\vec{r}, t) \;=\; -k_i^2\, A\, \cos(\omega\, t - \vec{k} \cdot \vec{r} + \phi_0) \quad \text{für} \quad i = 1,\, 2,\, 3, \tag{255}$$

$$\frac{\partial^2}{\partial\, t^2}\, f(\vec{r}, t)) \;=\; -\omega^2\, A\, \cos(\omega\, t - \vec{k} \cdot \vec{r}\omega\, t + \phi_0). \tag{256}$$

Offensichtlich gilt:

die sogenannte **ebene Wellenlösung**

$$f(\vec{r}, t) = A\, \cos(\omega\, t - \vec{k} \cdot \vec{r} + \phi_0) \tag{257}$$

ist eine **spezielle Lösung der dreidimensionalen Wellengleichung**

$$\frac{\partial^2\, f(\vec{r}, t)}{\partial\, t^2} = v^2\, \Delta\, f(\vec{r}, t) \tag{258}$$

mit Ausbreitungsgeschwindigkeit der Welle

$$v = \frac{\omega}{\sqrt{\vec{k}^2}} = \frac{\omega}{\sqrt{k_1^2 + k_2^2 + k_3^2}} = \frac{\omega}{k}. \tag{259}$$

Die Lösungen (252) bzw. (257) heissen *ebene Wellenlösungen* aus folgendem anschaulichen Grund. Man fragt, welche räumlich-zeitlichen Gebilde entstehen, wenn man die gesamte Phase in (252) konstant wählt,

$$\omega\, t - \vec{k} \cdot \vec{r} = k\,(v\, t - \vec{r} \cdot \frac{\vec{k}}{k}) = \text{konstant}. \tag{260}$$

In einem Schnappschuss zu einem festen Zeitpunkt t erhält man parallele Ebenen im dreidimensionalen Ortsraum; sie bestehen aus all denjenigen Punkten, für die die Projektion des Ortsvektors $\vec{r}$ auf den Einheitsvektor $\vec{k}/k$ in Richtung von $\vec{k}$ denselben Wert hat. Dies sind offensichtlich *Ebenen senkrecht zu* $\vec{k}$. Schaltet man die Zeitabhängigkeit ein, bewegen sich diese Ebenen konstanter Phase mit der Geschwindigkeit v in Richtung des Einheitsvektors $\vec{k}/k$, *der Ausbreitungsrichtung der Welle im dreidimensionalen Raum.* In Abb. 19 ist dies in einer zweidimensionalen Welt verdeutlicht; die parallelen Ebenen schrumpfen dort zu parellelen Geraden.

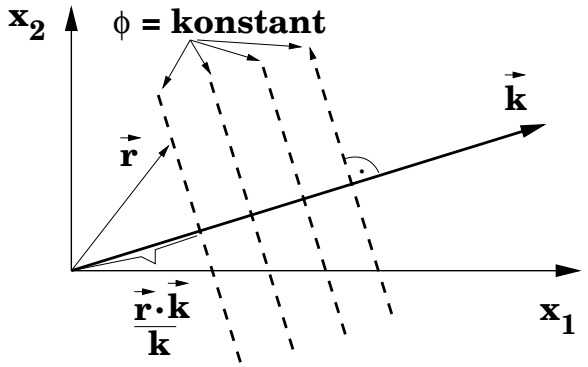


Abbildung 19:

Die ebenen Wellenlösungen spielen eine wichtige Rolle in der Physik.

4.7 Homogene partielle Differentialgleichungen in der Physik

Hier sollen partielle Differentialgleichungen verschiedenen Typs vorgestellt werden, die in der Physik eine große Rolle spielen. In der Physik treten sie in der Regel begleitet von bestimmten Rand- und Anfangsbedingungen auf, die die Lösungen entscheidend prägen. Der Schwierigkeitsgrad steckt in der Regel nicht in der Bestimmung von Lösungen der partiellen Differentialgleichung sondern in der Implementation der Rand- und Anfangsbedingungen. Dieser Aspekt kann prinzipiell nicht systematisch behandelt werden, sondern muss an Beispielen illustriert werden. Aus Zeitgründen muss auf derartige Beispiele im Folgenden verzichtet werden. Auch die Lösung von vereinfachten Formen der vorkommenden partiellen Differentialgleichungen wird in die Übungen verlagert. In allen betrachteten Fällen ist die Methode der *Separation der Variablen* erfolgreich.

Die partiellen Differentialgleichungen zweiter Ordnung, die in der Physik eine hervorragende Rolle spielen, sind in drei Kategorien einteilbar

- elliptisch,
- parabolisch,
- hyperbolisch.

Ihre Namen haben sie auf Grund der relativen Vorzeichen der Ableitungen erhalten, die an die impliziten Gleichungen der entsprechenden Kegelschnitte erinnern. Im folgenden werden zentrale Beispiele der Physik zu diesen Typen vorgeführt.

Es werden Felder der vier Variablen

$$\vec{r} = (x_1,\, x_2,\, x_3) = \text{Ortsvektor} \quad \text{und} \quad t = \text{Zeit} \tag{261}$$

betrachtet.

4.7.1 Die Diffusionsgleichung der Thermodynamik

Die Diffusions- oder Wärmeleitungsgleichung der Thermodynamik, die z.B. eine zeitabängige Temperaturverteilung in einem Körper beschreibt, ist eine

partielle Differentialgleichung vom parabolischen Typ

$$\Delta\, f(\vec{r},t) = \text{Konstante} \cdot \frac{\partial}{\partial\, t}\, f(\vec{r},t) \quad \text{mit positiver Konstante,} \tag{262}$$

wobei noch einmal an

$$\Delta = \vec{\nabla}\cdot\vec{\nabla} = \text{div}\,\text{grad} = \frac{\partial^2}{\partial\, x_1^2} + \frac{\partial^2}{\partial\, x_2^2} + \frac{\partial^2}{\partial\, x_3^2} \tag{263}$$

und

$$\vec{\nabla} = (\frac{\partial}{\partial\, x_1},\, \frac{\partial}{\partial\, x_2},\, \frac{\partial}{\partial\, x_3}) \tag{264}$$

erinnert werden soll. Dies ist eine Differentialgleichung, in der die örtlichen partiellen Ableitungen zweifach und die zeitliche Ableitung einfach auftreten.

4.7.2 Elektrodynamik – Maxwellgleichungen

Die zentralen Maxwellgleichungen für die elektrische Feldstärke $\vec{E}(\vec{r},t)$ und die magnetische Feldstärke $\vec{B}(\vec{r},t)$ sind

$$\vec{\nabla}\cdot\vec{E}(\vec{r},t) = \text{div}\,\vec{E}(\vec{r},t) = \frac{1}{\epsilon_0}\,\rho(\vec{r},t), \tag{265}$$

$$\vec{\nabla}\times\vec{E}(\vec{r},t) = \text{rot}\,\vec{E}(\vec{r},t) = -\kappa\frac{\partial}{\partial\, t}\,\vec{B}(\vec{r},t), \tag{266}$$

$$\vec{\nabla}\cdot\vec{B}(\vec{r},t) = \text{div}\,\vec{B}(\vec{r},t) = 0, \tag{267}$$

$$\vec{\nabla}\times\vec{B}(\vec{r},t) = \text{rot}\,\vec{B}(\vec{r},t) = \kappa\mu_0\Big(\vec{j}(\vec{r},t) + \epsilon_0\,\frac{\partial}{\partial\, t}\,\vec{E}(\vec{r},t)\Big) \tag{268}$$

mit Konstanten ϵ_0, κ, μ_0 und Inputfeldern, der Ladungsdichte $\rho(\vec{r},t)$ und der Stromdichte $\vec{j}(\vec{r},t)$, deren physikalische Bedeutung hier nicht diskutiert werden soll. Es gilt der Zusammenhang $\kappa^2\mu_0\epsilon_0 = 1/c^2$, wobei c die Lichtgeschwindigkeit ist.

Sie wurden bereits als Beispiele in der Vorlesung "Mathematische Ergänzung zu Experimentalphysik I" vorgestellt.

Die Maxwellgleichungen sind gekoppelte partielle Differentialgleichungen erster Ordnung in den Orts- und in der Zeitvariablen für die Vektorfelder $\vec{E}(\vec{r},t)$ und $\vec{B}(\vec{r},t)$.

4.7.3 Elektrodynamik – Elektrostatik

Die Elektrostatik beschreibt eine stationäre physikalische Situation ohne Magnetfeld $\vec{B}(\vec{r},t)$ mit zeitunabhängigem elektrischem Feld $\vec{E}(\vec{r})$. Es soll der einfachste Fall betrachtet werden, in dem auch die stationäre Ladungsdichte $\rho(\vec{r})$ verschwindet. Die relevanten Maxwellgleichungen sind (265) und (266). Die Gleichung (266) wird, wie in der Vorlesung "Mathematische Ergänzung zu Experimentalphysik I" gezeigt wurde, durch einen Ansatz $\vec{E}(\vec{r}) = -\vec{\nabla}\,\Phi(\vec{r})$ mit einem Skalarfeld $\Phi(\vec{r})$, dem sogenannten skalaren Potential, gelöst. Setzt man diesen Ansatz in (265) ein, erhält man

eine **partielle Differentialgleichung vom elliptischen Typ**, die sog. **Laplacegleichung**

$$\Delta\,\Phi(\vec{r}) = 0. \tag{269}$$

Sie ist eine partielle Differentialgleichung zweiter Ordnung in allen drei Ortsvariablen.

4.7.4 Elektrodynamik – Maxwellgleichungen im Vakuum

Im Vakuum verschwinden Ladungsdichte $\rho(\vec{r},t)$ und Stromdichte $\vec{j}(\vec{r},t)$. In diesem Fall lassen sich die beiden Maxwellgleichungen (265) und (266) auf die Form

$$\Delta\,\vec{E}(\vec{r},t) - \frac{1}{c^2}\frac{\partial^2}{\partial\,t^2}\,\vec{E}(\vec{r},t) = 0 \tag{270}$$

bringen. Dies wird erreicht, indem auf die linke und die rechte Seite von (266) der Operator rot angewandt wird. Die linke Seite wird mit Hilfe der Identität $\vec{\nabla}\times\Big(\vec{\nabla}\times\vec{E}(\vec{r},t)\Big) = \vec{\nabla}\Big(\vec{\nabla}\cdot\vec{E}(\vec{r},t)\Big) - \Delta\,\vec{E}(\vec{r},t)$ und mit (265) gleich $-\Delta\,\vec{E}(\vec{r},t)$. Die rechte Seite wird gleich $-\kappa\,\vec{\nabla}\times\Big(\frac{\partial}{\partial\,t}\,\vec{B}(\vec{r},t)\Big) = -\kappa\,\frac{\partial}{\partial\,t}\Big(\vec{\nabla}\times\vec{B}(\vec{r},t)\Big)$ und mit Hilfe von (268) gleich $-\kappa^2\mu_0\epsilon_0\,\frac{\partial}{\partial\,t}\Big(\frac{\partial}{\partial\,t}\,\vec{E}(\vec{r},t)\Big) = -\frac{1}{c^2}\,\frac{\partial^2}{\partial\,t^2}\,\vec{E}(\vec{r},t)$, wobei c die Lichtgeschwindigkeit ist. Analog kann die entsprechende Gleichung für $\vec{B}(\vec{r},t)$ hergeleitet werden. In dieser Form entkoppeln die Differentialgleichungen für das elektrische und das magnetische Feld.

Die Gleichung (270) für $\vec{E}(\vec{r},t)$ und die entsprechende Gleichung für $\vec{B}(\vec{r},t)$ sind

partielle Differentialgleichungen zweiter Ordnung vom hyperbolischen Typ

$$\frac{\partial^2}{\partial t^2}\,\vec{E}(\vec{r},t) = c^2\,\Delta\,\vec{E}(\vec{r},t) \tag{271}$$

$$\frac{\partial^2}{\partial t^2}\,\vec{B}(\vec{r},t) = c^2\,\Delta\,\vec{B}(\vec{r},t). \tag{272}$$

Es handelt sich offensichtlich um *dreidimensionale Wellengleichungen* vom Typ (249) für jede Komponente von $\vec{E}$ und von $\vec{B}$. Die auftretende Konstante mit Dimension einer Geschwindigkeit ist die Lichtgeschwindigkeit c. Sie haben **ebene Wellenlösungen** vom Typ (252) mit **Dispersionsrelation**

$$\omega = c\,k. \tag{273}$$

Diese ebenen Wellenlösungen beschreiben elektromagnetische Strahlung (Hertzscher Dipol).

4.7.5 Schrödingergleichung

Wie bereits mehrfach erwähnt, wird im Rahmen der Teilchen-Welle-Dualität der Quantenmechanik ein Teilchen der Masse m durch eine komplexe Wellenfunktion $\psi(\vec{r},t)$ beschrieben. Für ein kräftefreies Teilchen lautet die Schrödingergleichung

$$i\,\hbar\,\frac{\partial}{\partial t}\,\psi(\vec{r},t) = -\frac{\hbar^2}{2m}\,\Delta\,\psi(\vec{r},t), \tag{274}$$

wobei $h = 2\pi\hbar$ das Plancksche Wirkungsquantum ist. Dies ist eine partielle Differentialgleichungen, die zweifache partielle Ableitungen in den Ortsvariablen und eine einfache partielle Ableitung in der Zeit enthält. Im Gegensatz zur Diffusionsgleichung der Thermodynamik enthält sie jedoch die imaginäre Zahl i und hat deshalb auch ebene Wellen-Lösungen.

Der Ansatz einer ebenen Welle, diesmal in der komplexen Ergänzung in Erweiterung von (199),

$$\psi(\vec{r},t) = A\,e^{i(\vec{k}\,\vec{r}-\omega\,t)} \tag{275}$$

erfüllt die Schrödingergleichung (274) wegen

$$\Delta\,\psi(\vec{r},t) = -\vec{k}^2\,A\,e^{i(\vec{k}\,\vec{r}-\omega\,t)} \tag{276}$$

$$i\,\frac{\partial}{\partial t}\,\psi(\vec{r},t) = \omega\,A\,e^{i(\vec{k}\,\vec{r}-\omega\,t)} \tag{277}$$

für die folgende Beziehung zwischen der Kreisfrequenz ω und der Wellenzahl $k = |\vec{k}|$,

die **Dispersionsrelation**

$$\omega = \frac{\hbar}{2m}\,k^2. \tag{278}$$

5 Uneigentliche Integrale

5.1 Integrale mit unendlichen Integrationsgrenzen

Integrale mit unendlichen Integrationsgrenzen heißen **uneigentliche Integrale**. Das uneigentliche Integral heißt **konvergent**, wenn sein Wert endlich ist, und **divergent**, wenn sein Wert unendlich ist.

Dies sei an zwei Beispielen illustriert.

Beispiel für ein konvergentes uneigentliches Integral

Ausgangspunkt ist das Integral mit endlichen Integrationsgrenzen a und b und $b > a > 0$

$$A(b) = \int_a^b \frac{1}{x^2}\,\mathrm{d}\,x = -\frac{1}{x}\Big|_a^b = \frac{1}{a} - \frac{1}{b}, \tag{279}$$

betrachtet als Funktion einer variablen oberen Grenze b. Der Integrand ist endlich und stetig im endlichen Intervall $[a, b]$, also integrierbar.

Im Grenzwert $b \to \infty$ wird das Integral (279) zu einem konvergenten uneigentlichen Integral

$$\lim_{b\to\infty} A(b) = \lim_{b\to\infty} \int_a^b \frac{1}{x^2}\,\mathrm{d}\,x = \frac{1}{a} - \lim_{b\to\infty} \frac{1}{b} = \frac{1}{a}, \tag{280}$$

d.h.

$$\int_a^\infty \frac{1}{x^2}\,\mathrm{d}\,x = \frac{1}{a}, \tag{281}$$

wobei $\int_a^\infty \frac{1}{x^2}\,\mathrm{d}\,x$ durch den Grenzwert $\lim_{b\to\infty} \int_a^b \frac{1}{x^2}\,\mathrm{d}\,x$ definiert ist.

Abb. 20 illustriert, dass das hinreichend schnelle Verschwinden des Integranden für $b \to \infty$ verantwortlich für die Konvergenz ist.

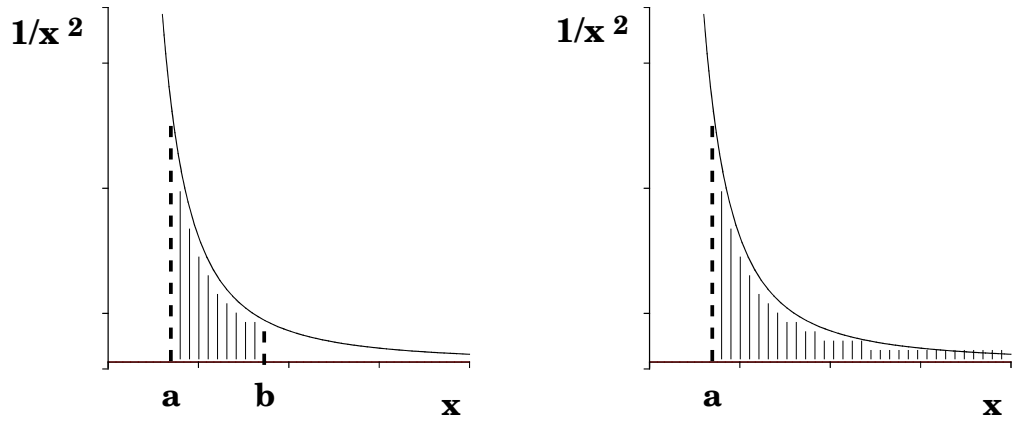


Abbildung 20:

Analog wird

$$\int_{-\infty}^b \frac{1}{x^2}\,\mathrm{d}\,x = -\frac{1}{b} \quad \text{für} \quad b < 0. \tag{282}$$

Allgemein läßt sich analog zeigen, dass das Integral

$$\int_a^\infty \frac{1}{x^\alpha}\,\mathrm{d}\,x \quad \text{für } a>0 \quad \text{und} \quad \int_{-\infty}^b \frac{1}{x^\alpha}\,\mathrm{d}\,x \quad \text{für } b<0 \tag{283}$$

für $\alpha > 1$ ein konvergentes uneigentliches Integral ist.

Beispiel für ein divergentes uneigentliches Integral

Wählt man statt dessen als Ausgangspunkt

$$B(b) = \int_a^b \frac{1}{x}\,\mathrm{d}\,x = \log x\Big|_a^b = \log b - \log a, \tag{284}$$

wieder mit $b > a > 0$, dann wird das Integral im Grenzwert $b \to \infty$ divergent, da

$$\lim_{b\to\infty} \log b \tag{285}$$

unendlich wird.

5.2 Integrale mit Unendlichkeitsstelle des Integranden im Integrationsweg

Der Integrand $f(x)$ besitze an einer Stelle x_0 innerhalb des Integrationsintervalls $[a, b]$ eine Singularität, wie z.B. in Abb. 21 generisch dargestellt.

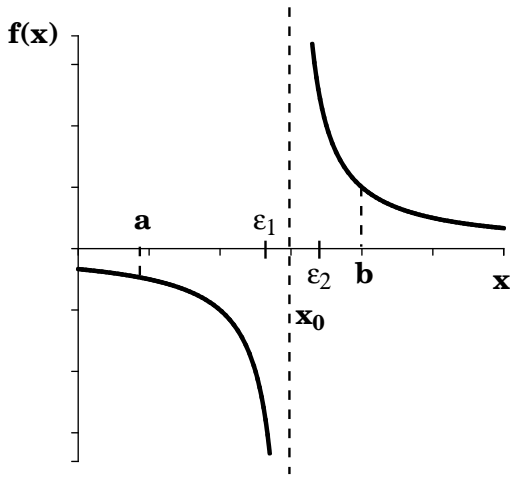


Abbildung 21:

Über eine solche Unendlichkeitsstelle darf man nicht einfach hinwegintegrieren, da in der Riemann-Summe $f(x_0)$ und damit der Wert Unendlich auftreten kann.

Ein Integral mit einer Unendlichkeitsstelle des Integranden bei x_0 im Integrationsweg, $a \leq x_0 \leq b$, heißt **uneigentliches Integral**

$$\int_a^b f(x)\,\mathrm{d}\,x, \tag{286}$$

falls der Grenzwert

$$\lim_{\epsilon_1 \to 0^+} \int_a^{x_0-\epsilon_1} f(x)\,\mathrm{d}\,x + \lim_{\epsilon_2 \to 0^+} \int_{x_0+\epsilon_2}^b f(x)\,\mathrm{d}\,x \tag{287}$$

existiert und zwar unabhängig davon, wie die separaten Limites $\epsilon_1 \to 0^+$ und $\epsilon_2 \to 0^+$ durchgeführt werden. Das uneigentliche Integral (286) ist durch den Grenzwert (287) definiert.

Beispiel

$$\int_{-1}^1 \frac{1}{\sqrt{|x|}}\,\mathrm{d}\,x \tag{288}$$

Der Integrand hat eine Singularität bei $x = 0$ im Integrationsweg. Es sind also die beiden separaten unabhängigen Grenzwerte

$$\lim_{\epsilon_1 \to 0^+} \int_{-1}^{-\epsilon_1} \frac{1}{\sqrt{|x|}}\,\mathrm{d}\,x = -\lim_{\epsilon_1 \to 0^+} \int_1^{\epsilon_1} \frac{1}{\sqrt{|x|}}\,\mathrm{d}\,x = \lim_{\epsilon_1 \to 0^+} \int_{\epsilon_1}^1 \frac{1}{\sqrt{|x|}}\,\mathrm{d}\,x \tag{289}$$

$$= \lim_{\epsilon_1 \to 0^+} \int_{\epsilon_1}^1 \frac{1}{\sqrt{x}}\,\mathrm{d}\,x = \lim_{\epsilon_1 \to 0^+} 2\sqrt{x}\,\Big|_{\epsilon_1}^1 = 2 - 2 \lim_{\epsilon_1 \to 0^+} \epsilon_1 = 2, \tag{290}$$

wobei im ersten Schritt die Integrationsvariable x durch $-x$ substituiert wurde, und

$$\lim_{\epsilon_2 \to 0^+} \int_{\epsilon_2}^1 \frac{1}{\sqrt{|x|}}\,\mathrm{d}\,x = \lim_{\epsilon_2 \to 0^+} \int_{\epsilon_2}^1 \frac{1}{\sqrt{x}}\,\mathrm{d}\,x \tag{291}$$

$$= \lim_{\epsilon_2 \to 0^+} 2\sqrt{x}\,\Big|_{\epsilon_2}^1 = 2 - 2 \lim_{\epsilon_2 \to 0^+} \epsilon_2 = 2 \tag{292}$$

durchzuführen und zu addieren. Damit existiert das uneigentliche Integral

$$\int_{-1}^1 \frac{1}{\sqrt{|x|}}\,\mathrm{d}\,x = 4, \tag{293}$$

obwohl es eine Singularität im Integrationsweg hat.

Eine verschärfte Situation entsteht, wenn die beiden Limites

$$\lim_{\epsilon_1 \to 0^+} \int_a^{x_0-\epsilon_1} f(x)\,\mathrm{d}\,x \quad \text{und} \quad \lim_{\epsilon_2 \to 0^+} \int_{x_0+\epsilon_2}^b f(x)\,\mathrm{d}\,x \tag{294}$$

nicht separat existieren, jedoch als gekoppelte Limites für $\epsilon_1 = \epsilon_2$ existieren.

Ein Integral mit einer Unendlichkeitsstelle des Integranden bei x_0 im Integrationsweg, $a \leq x_0 \leq b$, heißt **Hauptwertintegral**, bezeichnet mit

$$P \int_a^b f(x)\,\mathrm{d}\,x, \tag{295}$$

falls der Grenzwert

$$\lim_{\epsilon \to 0^+} \left(\int_a^{x_0-\epsilon} f(x)\,\mathrm{d}\,x + \int_{x_0+\epsilon}^b f(x)\,\mathrm{d}\,x \right) \tag{296}$$

existiert, der Grenzwert der einzelnen Summanden jedoch nicht. Das Hauptwertintegral (295) ist durch den Grenzwert (296) definiert.

Das Symbol P vor dem Integral (benannt nach dem englischen Wort "principal value" für Hauptwert) weist darauf hin, dass das Integral nur im Fall der gekoppelten Grenzwertbildung existiert.

Beispiel

Gegeben seien der Integrand $1/x$ und die Integrationsgrenzen $-a$ und b mit $-a < 0 < b$. Der Integrand hat bei $x = 0$ eine Unendlichkeitsstelle im Integrationsweg. Die separaten Limites

$$\lim_{\epsilon_1 \to 0^+} \int_{-a}^{-\epsilon_1} \frac{1}{x}\,\mathrm{d}\,x = -\lim_{\epsilon_1 \to 0^+} \int_{\epsilon_1}^{a} \frac{1}{x}\,\mathrm{d}\,x = \lim_{\epsilon_1 \to 0^+} \log \epsilon_1 - \log a \tag{297}$$

$$\lim_{\epsilon_2 \to 0^+} \int_{\epsilon_2}^{b} \frac{1}{x}\,\mathrm{d}\,x = \log b - \lim_{\epsilon_2 \to 0^+} \log \epsilon_2 \tag{298}$$

sind $-\infty$ bzw. ∞; das uneigentliche Integral existiert also nicht. Es läßt sich jedoch das Hauptwertintegral

$$P \int_{-a}^{b} \frac{1}{x}\,\mathrm{d}\,x = \lim_{\epsilon \to 0^+} \left(\int_{-a}^{-\epsilon} \frac{1}{x}\,\mathrm{d}\,x + \int_{\epsilon}^{b} \frac{1}{x}\,\mathrm{d}\,x \right) \tag{299}$$

$$= \lim_{\epsilon \to 0^+} (\log \epsilon - \log a + \log b - \log \epsilon) = \log b - \log a \tag{300}$$

berechnen; es ist endlich. Das Ergebnis wird in Abb. 22 verdeutlicht. Im Hauptwertintegral heben sich die Beiträge der beiden vertikal schraffierten Flächen für jeden Wert von ϵ gegenseitig auf – und dafür ist der gekoppelte Grenzwert $\epsilon \to 0$ essentiell – und es bleibt nur der Beitrag der horizontal schraffierten Fläche.

6 Fourierreihen und Fourierintegrale

6.1 Physikalische Motivation

Hier wird zunächst an die Entwicklung einer Funktion in eine Taylorreihe angeknüpft. Es kann sinnvoll sein, eine Funktion $f(x)$ in eine Taylorreihe z.B. um $x = 0$ zu entwickeln

$$f(x) = \sum_{n=0}^{\infty} a_n\, x^n \tag{301}$$

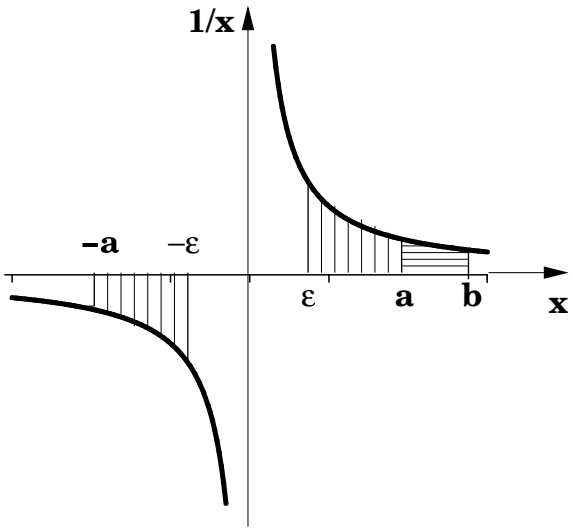


Abbildung 22:

oder um eine Stelle x_0

$$f(x) = \sum_{n=0}^{\infty} \tilde{a}_n \, (x - x_0)^n. \tag{302}$$

Dies sind Entwicklungen nach Potenzfunktionen x^n bzw. $(x - x_0)^n$, die sich leicht differenzieren und integrieren lassen und deren erste wenige Terme aufsummiert bereits gute Näherungen geben für $|x| \ll 1$ bzw. $|x - x_0| \ll 1$ (innerhalb des Konvergenzradius).

Ein ebenso vielversprechender Ansatzpunkt ist die Frage: kann man z.B. eine *periodische* Funktion $f(x)$ in eine unendliche Reihe von einfachen periodischen Funktionen wie trigonometrische Funktionen entwickeln?

In der Physik gibt es viele periodische Funktionen, alle Schwingungs- und Wellenphänomene in der klassischen Mechanik, Optik, Akustik, Elektrodynamik, Quantenmechanik und Astrophysik, für die solche sogenannten *Fourierentwicklungen* Anwendungen finden.

6.2 Entwicklung einer periodischen Funktion in eine Fourierreihe

6.2.1 Definition

Gegeben sei der Einfachheit halber zunächst eine periodische Funktion mit der Periode 2π

$$f(x) = f(x + 2\pi) \tag{303}$$

für alle x; dann gilt auch

$$f(x) = f(x + 2n\pi) \tag{304}$$

für alle ganzzahligen Werte von n.

Als *Basisfunktionen*, nach denen sinnvoll entwickelt werden kann, bieten sich an

$$\sin(n\,x) \text{ und } \cos(n\,x). \tag{305}$$

$$\tag{306}$$

Definition: Die Reihenentwicklung einer periodischen Funktion $f(x)$ mit Periode 2π

$$f(x) = \frac{a_0}{2} + \sum_{n=1}^{\infty} \Big(a_n \cos(n\,x) + b_n \sin(n\,x) \Big) \tag{307}$$

heißt **Fourierentwicklung**; die Reihe heißt **Fourierreihe**; die Koeffizienten a_0, a_n und b_n sind die **Fourierkoeffizienten**.

Damit ist $f(x+2\pi) = a_0/2 + \sum_{n=1}^{\infty} \Big(a_n \cos(n\,x + 2\pi n) + b_n \sin(n\,x + 2\pi n) \Big) = f(x)$ automatisch erfüllt.

6.2.2 Mathematische Voraussetzungen für die Entwicklung in eine Fourierreihe

Die mathematischen Voraussetzungen für eine Fourierentwicklung sind im *Satz von Dirichlet* zusammengefasst:

- Eine Funktion $f(x)$ habe die Periode 2π.
- $f(x)$ und $f'(x)$ seien *stückweise stetig*, d.h. $f(x)$ und $f'(x)$ haben höchstens endlich viele Unstetigkeitsstellen in einer Periode.

Dann konvergiert die Fourierreihe an allen Stetigkeitsstellen gegen $f(x)$. An einer Unstetigkeitsstelle x_0 ist der Wert der Fourierreihe gleich dem *arithmetischen Mittel* aus dem linksseitigen und dem rechtsseitigen Grenzwert der Funktion, d.h.

$$\frac{1}{2}\Big(\lim_{\Delta x \to 0} f(x_0 + \Delta\,x) + \lim_{\Delta x \to 0} f(x_0 - \Delta\,x) \Big). \tag{308}$$

In allen folgenden Beispielen kann man sich davon überzeugen, dass jeweils die Voraussetzungen und Folgerungen des Satzes von Dirichlet erfüllt sind.

6.2.3 Bestimmung der Fourierkoeffizienten

Die Bestimmung der Fourierkoeffizienten a_0, a_n, b_n geschieht in drei Schritten.

Bestimmung des Koeffizienten a_0:

Dazu wird $f(x)$ über eine ganze Periode integriert, die o.B.d.A. von $-\pi$ bis π gewählt wird

$$\int_{-\pi}^{\pi} f(x)\,\mathrm{d}\,x = a_0\,\pi + \sum_{n=1}^{\infty} \Big(a_n \int_{-\pi}^{\pi} \cos(n\,x)\,\mathrm{d}\,x + b_n \int_{-\pi}^{\pi} \sin(n\,x)\,\mathrm{d}\,x \Big); \tag{309}$$

hier sind Integration und Summation vertauscht worden. Da über eine ganze Periode integriert wird, gilt

$$\int_{-\pi}^{\pi} \cos(n\,x)\,\mathrm{d}\,x = 0, \quad \int_{-\pi}^{\pi} \sin(n\,x)\,\mathrm{d}\,x = 0. \tag{310}$$

und somit

$$a_0 = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x)\, \mathrm{d}\, x. \tag{311}$$

Bestimmung der Koeffizienten a_n:

Dazu ist das folgende Integral nützlich, wobei m eine natürliche Zahl ist,

$$\int_{-\pi}^{\pi} f(x)\, \cos(m\, x) \mathrm{d}\, x = \frac{a_0}{2} \int_{-\pi}^{\pi} \cos(m\, x) \mathrm{d}\, x + \Sigma_1 + \Sigma_2 \quad \text{mit} \tag{312}$$

$$\Sigma_1 = \sum_{n=1}^{\infty} a_n \int_{-\pi}^{\pi} \cos(n\, x)\, \cos(m\, x)\, \mathrm{d}\, x, \tag{313}$$

$$\Sigma_2 = \sum_{n=1}^{\infty} b_n \int_{-\pi}^{\pi} \sin(n\, x)\, \cos(m\, x)\, \mathrm{d}\, x. \tag{314}$$

$$\tag{315}$$

Das erste Integral auf der rechten Seite von (312) verschwindet; in (313) wird die trigonometrische Identität

$$\cos(n\, x)\, \cos(m\, x) = \frac{1}{2} \cos\Big((n+m)\, x\Big) + \frac{1}{2} \cos\Big((n-m)\, x\Big) \tag{316}$$

benutzt

$$\int_{-\pi}^{\pi} \cos(n\, x)\, \cos(m\, x)\, \mathrm{d}\, x = \frac{1}{2} \int_{-\pi}^{\pi} \cos\Big((n+m)\, x\Big)\, \mathrm{d}\, x + \frac{1}{2} \int_{-\pi}^{\pi} \cos\Big((n-m)\, x\Big)\, \mathrm{d}\, x. \tag{317}$$

Auf der rechten Seite verschwindet das erste Integral, da $n+m \neq 0$, das zweite Integral verschwindet außer für $m = n$. Das Ergebnis ist

$$\int_{-\pi}^{\pi} \cos(n\, x)\, \cos(m\, x) \mathrm{d} = \begin{cases} \pi & \text{für } m = n \\ 0 & \text{sonst} \end{cases} = \pi\, \delta_{nm}. \tag{318}$$

Damit besteht die Summe Σ_1 nur aus einem Summanden

$$\Sigma_1 = \pi\, a_m. \tag{319}$$

In Σ_2 führt die trigonometrische Identität

$$\sin(n\, x)\, \cos(m\, x) = \frac{1}{2} \sin\Big((n+m)\, x\Big) + \frac{1}{2} \sin\Big((n-m)\, x\Big) \tag{320}$$

für alle natürlichen n, m auf

$$\int_{-\pi}^{\pi} \sin(n\, x)\, \cos(m\, x) \mathrm{d} = 0, \quad \text{d.h. } \Sigma_2 = 0. \tag{321}$$

Damit wird

$$a_n = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x)\, \cos(n\, x)\, \mathrm{d}\, x \quad n = 1,\, 2,\, \ldots\,. \tag{322}$$

Bestimmung der Koeffizienten $\mathbf{b_n}$:

Analog führt die Berechnug von

$$\int_{-\pi}^{\pi} f(x)\, \sin(m\, x)\, \mathrm{d}\, x \tag{323}$$

für natürliche m zu

$$b_n = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x)\, \sin(n\, x)\, \mathrm{d}\, x \quad n = 1,\, 2,\, \ldots\,. \tag{324}$$

6.3 Beispiele für Fourierreihen

6.3.1 Gerade und ungerade Funktionen

Ist $f(x)$ eine **gerade Funktion** von x,

$$f(-x) = f(x), \tag{325}$$

dann gilt wegen $\sin(-n\, x) = -\sin(n\, x)$

$$f(x) = \frac{a_0}{2} + \sum_{n=1}^{\infty} a_n \cos(n\, x). \tag{326}$$

Ist $f(x)$ eine **ungerade Funktion** von x,

$$f(-x) = -f(x), \tag{327}$$

dann gilt wegen $\cos(-n\, x) = \cos(n\, x)$

$$f(x) = \sum_{n=1}^{\infty} b_n \sin(n\, x). \tag{328}$$

Diese Einsicht bringt häufig eine Rechenerleichterung.

6.3.2 Rechteckschwingung

Die Variable sei – motiviert durch viele physikalische Anwendungen – o.B.d.A. als die Zeit t bezeichnet mit Periodendauer $T = 2\pi$

Definition: Eine **Rechteckschwingung** liegt vor für die periodische Funktion $f(t)$ mit Periodendauer $T = 2\pi$

$$f(t) = \begin{cases} -1 & \text{für} \quad -\pi < t \le -\frac{\pi}{2} \\ 1 & \text{für} \quad -\frac{\pi}{2} \le t \le \frac{\pi}{2} \\ -1 & \text{für} \quad \frac{\pi}{2} \le t \le \pi \end{cases} \tag{329}$$

Sie ist in Abb. 23 skizziert.

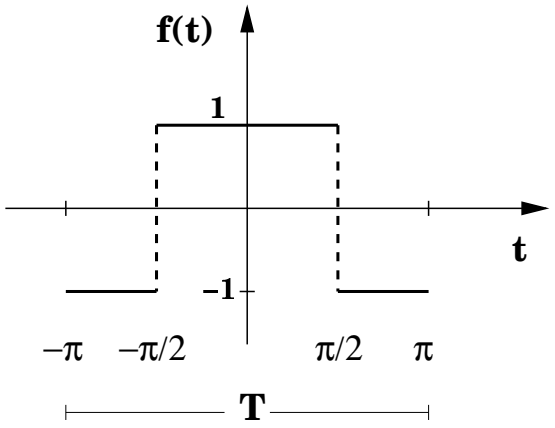


Abbildung 23:

$f(t)$ ist offensichtlich eine gerade Funktion von t mit zwei Unstetigkeitsstellen $t = \pm\pi/2$ innerhalb $-\pi < t \le \pi$. Entsprechend (326) sind nur die Koeffizienten a_0 und a_n zu berechnen. Nach (311) bzw. (322) gilt

$$a_0 = \frac{1}{\pi}\int_{-\pi}^{\pi} f(t)\,\mathrm{d}\,t = \frac{1}{\pi}\Big(-\int_{-\pi}^{-\pi/2}\mathrm{d}\,t + \int_{-\pi/2}^{\pi/2}\mathrm{d}\,t - \int_{\pi/2}^{\pi}\mathrm{d}\,t\Big) = 0, \tag{330}$$

$$a_n = \frac{1}{\pi}\left(-\int_{-\pi}^{-\pi/2}\cos(n\,t)\,\mathrm{d}\,t + \int_{-\pi/2}^{\pi/2}\cos(n\,t)\,\mathrm{d}\,t - \int_{\pi/2}^{\pi}\cos(n\,t)\,\mathrm{d}\,t\right) \tag{331}$$

$$= \frac{4}{n\pi}\sin\Big(\frac{n\pi}{2}\Big), \tag{332}$$

wobei zu beachten ist, dass

$$\sin(\frac{\pi}{2}) = 1, \quad \sin(\frac{2\pi}{2}) = 0, \quad \sin(\frac{3\pi}{2}) = -1, \quad \sin(\frac{4\pi}{2}) = 0, \ldots. \tag{333}$$

Die **Fourierreihe der Rechteckschwingung** ist

$$f(t) = \frac{4}{\pi}\sum_{n=1}^{\infty}\frac{1}{n}\sin\Big(\frac{n\pi}{2}\Big)\cos(n\,t). \tag{334}$$

In Abb. 24 sind der erste Term, mit $f_1(t)$ bezeichnet, die Summe der ersten zwei nichtverschwindenden Terme, mit $f_2(t)$ bezeichnet, die Summe der ersten drei nichtverschwindenden Terme,

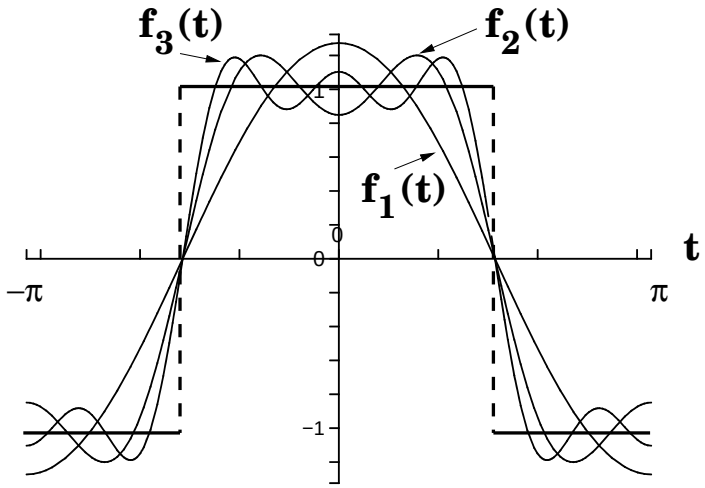


Abbildung 24:

mit $f_3(t)$ bezeichnet, im Vergleich mit der Rechteckschwingung $f(t)$ eingezeichnet; dies gibt einen guten Eindruck von der schnellen Konvergenz der Fourierreihe.

6.3.3 Kippschwingung

Definition: Eine **Kippschwingung** liegt vor für die periodische Funktion $f(t)$ mit Periodendauer $T = 2\pi$

$$f(t) = \frac{t}{\pi} \quad \text{für} \quad -\pi < t \leq \pi. \tag{335}$$

Sie ist in Abb. 25 skizziert; zur grafischen Illustration der Periodizität ist die Funktion auch jenseits des Periodizitätsintervalls $-\pi < t \leq \pi$ angedeutet.

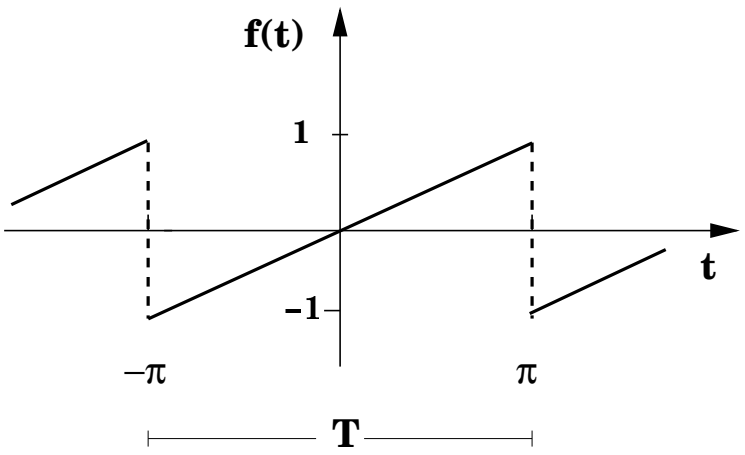


Abbildung 25:

$f(t)$ ist eine ungerade Funktion von t mit einer Unstetigkeitsstelle $t = \pi$ innerhalb von $-\pi < t \leq \pi$. Entsprechend (328) sind nur die Koeffizienten b_n zu berechnen. Nach (324) gilt

$$b_n = \frac{1}{\pi} \int_{-\pi}^{\pi} \frac{t}{\pi} \sin(n\,t)\,\mathrm{d}\,t. \tag{336}$$

Wegen der Symmetrie des Integranden unter $t \to -t$ erhält man

$$b_n = \frac{2}{\pi^2} \int_0^{\pi} t\,\sin(nt)\,\mathrm{d}\,t \tag{337}$$

und mit Hilfe der partiellen Integration

$$\begin{aligned} b_n &= \frac{2}{\pi^2}\,t\,\Big(\frac{-1}{n}\,\cos(n\,t)\Big)\Big|_0^{\pi} - \frac{2}{\pi^2} \int_0^{\pi} \Big(\frac{-1}{n}\,\cos(n\,t)\Big)\,\mathrm{d}\,t & (338)\\ &= -\frac{2}{n\,\pi}\,\cos(n\,\pi) = \frac{2}{n\,\pi}\,(-1)^{n+1}. & (339) \end{aligned}$$

Die **Fourierreihe der Kippschwingung** ist

$$f(t) = \frac{2}{\pi} \sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n}\,\sin(n\,t). \tag{340}$$

6.3.4 Dreieckschwingung

Definition: Eine **Dreieckschwingung** liegt vor für die periodische Funktion $f(t)$ mit Periodendauer $T = 2\pi$

$$f(t) = \begin{cases} -t & \text{für} \quad -\pi < t \leq 0 \\ t & \text{für} \quad 0 \leq t \leq \pi \end{cases} \tag{341}$$

Sie ist in Abb. 26 skizziert; das Verhalten jenseits des Periodizitätsintervalls $-\pi \leq t \leq \pi$ ist wieder angedeutet.

$f(t)$ ist offensichtlich eine gerade Funktion von t mit zwei Unstetigkeitsstellen $t = 0,\ \pi$ innerhalb von $-\pi < t \leq \pi$. Entsprechend (326) sind nur die Koeffizienten a_0 und a_n zu berechnen. Nach (311) bzw. (322) gilt

$$\begin{aligned} a_0 &= \frac{1}{\pi} \int_{-\pi}^{\pi} f(t)\,\mathrm{d}\,t = \frac{1}{\pi}\Big(\int_{-\pi}^{0} (-t)\,\mathrm{d}\,t + \int_0^{\pi} t\,\mathrm{d}\,t\Big) = \frac{2}{\pi} \int_0^{\pi} t\,\mathrm{d}\,t = \pi & (342)\\ a_n &= \frac{1}{\pi}\Big(\int_{-\pi}^{0} (-t)\,\cos(n\,t)\,\mathrm{d}\,t + \int_0^{\pi} t\,\cos(n\,t)\,\mathrm{d}\,t\Big) = \frac{2}{\pi} \int_0^{\pi} t\,\cos(n\,t)\,\mathrm{d}\,t & (343)\\ &= \frac{2}{\pi}\,t\,\Big(\frac{\sin(n\,t)}{n}\Big)\Big|_0^{\pi} - \frac{2}{\pi} \int_0^{\pi} \Big(\frac{\sin(n\,t)}{n}\Big)\,\mathrm{d}\,t & (344)\\ &= 0 + \frac{2}{\pi\,n^2}\,\cos(n\,t)\Big|_0^{\pi} = \frac{2}{\pi\,n^2}\,\Big(\cos(n\,\pi) - 1\Big) = \frac{2}{\pi\,n^2}\,\Big((-1)^n - 1\Big), & (345) \end{aligned}$$

wobei zur Berechnung der Koeffizienten a_n partielle Integration benutzt wurde.

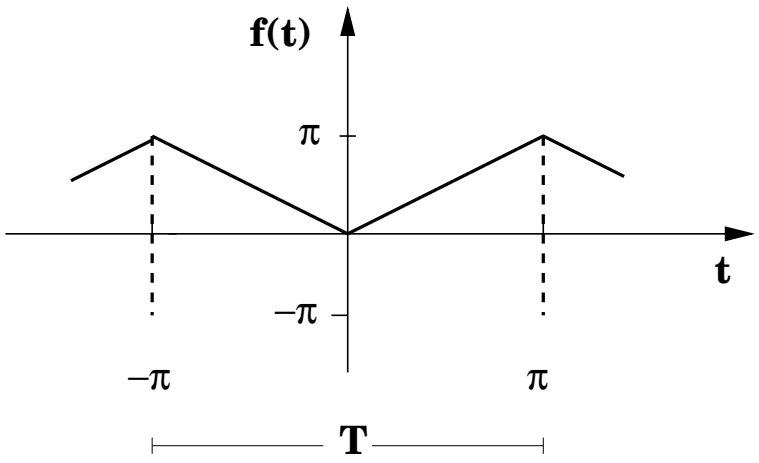


Abbildung 26:

Die **Fourierreihe der Dreieckschwingung** ist

$$f(t) = \frac{\pi}{2} + \frac{2}{\pi} \sum_{n=1}^{\infty} \frac{(-1)^n - 1}{n^2} \cos(n\,t). \tag{346}$$

6.4 Fourierreihen für Funktionen beliebiger Periode

Bisher waren periodische Funktionen mit der Periode 2π betrachtet worden, genauer Funktionen $f(t)$ der Zeit t, definiert im Intervall $-\pi < t \leq \pi$, mit der Periodizität $f(t) = f(t + 2\pi)$ und Periodendauer 2π. Auch weiterhin soll im Hinblick auf viele Anwendungen in Physik und Technik als Variable die Zeit t beibehalten werden.

Die Aussagen können leicht auf Funktionen $f(t)$ der Zeit t erweitert werden mit beliebig wählbarer Periodendauer T und entsprechender Periodizität

$$f(t) = f(t + T) \tag{347}$$

für alle t und Definitionsintervall

$$-\frac{T}{2} < t \leq \frac{T}{2}. \tag{348}$$

Es ist ersichtlich, dass die Fourierentwicklung dieser Funktionen aus (307) mit Koeffizienten (311), (322) und (324) durch die Ersetzungen

$$x \;\rightarrow\; \frac{2\pi}{T}\,t \quad \text{und entsprechend} \quad \mathrm{d}\,x \;\rightarrow\; \frac{2\pi}{T}\,\mathrm{d}\,t \tag{349}$$

$$\text{sowie der Integralgrenzen} \quad -\pi \;\rightarrow\; -T/2 \quad \text{bzw.} \quad \pi \;\rightarrow\; T/2 \tag{350}$$

hervorgeht.

Fourierreihe einer periodischen Funktion mit Periode T

$$f(t) = \frac{a_0}{2} + \sum_{n=1}^{\infty} \left(a_n \cos\left(2\pi n \frac{t}{T}\right) + b_n \sin\left(2\pi n \frac{t}{T}\right)\right) \tag{351}$$

mit Koeffizienten

$$a_0 = \frac{2}{T} \int_{-\frac{T}{2}}^{\frac{T}{2}} f(t)\,\mathrm{d}\,t \tag{352}$$

$$a_n = \frac{2}{T} \int_{-\frac{T}{2}}^{\frac{T}{2}} f(t) \cos\left(2\pi n \frac{t}{T}\right) \mathrm{d}\,t \tag{353}$$

$$b_n = \frac{2}{T} \int_{-\frac{T}{2}}^{\frac{T}{2}} f(t) \sin\left(2\pi n \frac{t}{T}\right) \mathrm{d}\,t. \tag{354}$$

6.5 Fourieranalyse

Ausgangspunkt sei o.B.d.A. die zu (351) äquivalente sog. *spektrale Darstellung* der Fourierreihe

$$f(t) = \frac{A_0}{2} + \sum_{n=1}^{\infty} A_n \cos\left(2\pi n \frac{t}{T} + \phi_n\right), \tag{355}$$

in der jeder Term durch *eine* Fourierkomponente mit *zwei* Parametern A_n und ϕ_n dargestellt wird. Trägt man die Amplituden A_n gegen die Kreisfrequenz $\omega_n = 2n\pi/T$ für die Werte $n = 1, 2, ...$ auf, wie generisch in Abb. 27 dargestellt, erhält man das *Amplitudenspektrum*, das auch *Fourierspektrum* oder *Frequenzspektrum* genannt wird. Trägt man die Phasenwinkel ϕ_n der Fourierkomponenten gegen die Kreisfrequenz auf, spricht man von einem *Phasenspektrum*. Die Ermittlung des Amplituden- und Phasenspektrums einer gegebenen periodischen Funktion $f(t)$ heißt *Fourieranalyse* oder *Frequenzanalyse*.

6.6 Fourierintegrale

6.6.1 Übergang von der Fourierreihe zum Fourierintegral

In der Fourierentwicklung werden *periodische* Funktionen nach trigonometrischen Funktionen entwickelt.

Im Folgenden geht es um eine Erweiterung zu einer Darstellung einer *nichtperiodischen Funktion* $f(t)$ mit Hilfe von trigonometrischen Funktionen. Dies kommt in Physik und Technik typisch zur Anwendung, wenn ein *zeitlich begrenztes* Signal vorliegt, wie generisch in Abb. 28 dargestellt. Auch in der Quantenmechanik spielt diese sog. Fouriertransformation eine hervorragende Rolle.

Die Erweiterung kann man sich am einfachen Beispiel des Rechtecksignals in Abb. 29 verdeutlichen.

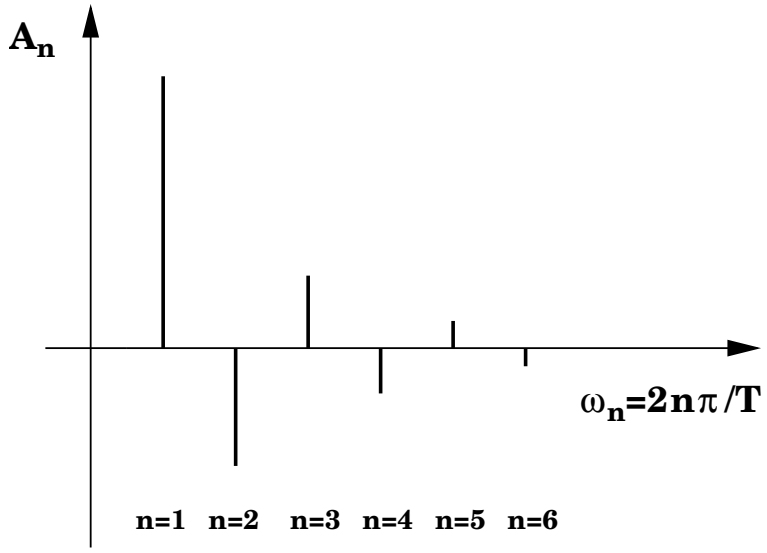


Abbildung 27:

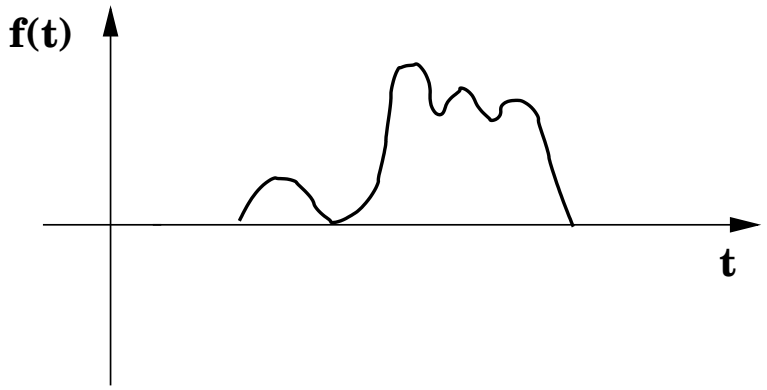


Abbildung 28:

Ausgangspunkt ist zunächst eine *periodische* Rechteckfunktion mit Periode T und Signaldauer $t_0 < T$. Die Fourierentwicklung läßt sich für diese gerade Funktion leicht herleiten zu

$$f(t) = \frac{t_0}{T} + \frac{2}{\pi} \sum_{n=1}^{\infty} \frac{1}{n} \sin\left(\frac{n\pi}{T}\, t_0\right) \cos\left(\frac{2\pi n}{T}\, t\right). \tag{356}$$

Ziel ist es nun, die Periode T gegen unendlich gehen zu lassen und dabei die Zeitdauer t_0 des Signals festzuhalten

$$T \to \infty, \quad t_0 \text{ fest.} \tag{357}$$

Wie Abb. 29 verdeutlicht, wird in diesem Grenzwert die Periodizität der Funktion bedeutungslos und man hat es in der Tat mit einer nichtperiodischen Funktion, einem begrenzten Rechtecksignal der Zeitdauer t_0, zu tun.

Zunächst wird in jedem Term der Reihe die Variable n durch die Variable

$$\omega = \frac{2\pi}{T}\, n \tag{358}$$

ersetzt. Außerdem wird jeder Term der Reihe mit $\Delta\, n$ multipliziert, der Differenz zweier aufeinanderfolgender Werte von n, die identisch gleich $\Delta\, n = 1$ ist. Die entsprechende Kreisfrequenzdifferenz ist

$$\Delta\, \omega = \frac{2\pi}{T}\, \Delta\, n = \frac{2\pi}{T}. \tag{359}$$

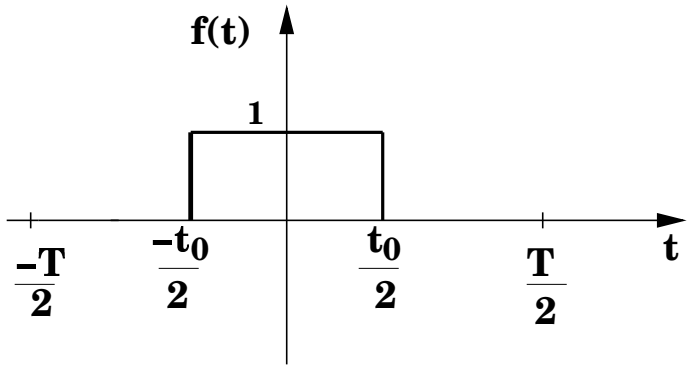


Abbildung 29:

Der Grenzübergang (357) übersetzt sich bei festgehaltenem $\Delta n = 1$ in

$$\Delta\omega \to 0 \quad \text{für} \quad T \to \infty. \tag{360}$$

In diesem Grenzwert geht der erste Summand, t_0/T, in (356) gegen Null; die Summe über n geht über in ein Integral über ω wobei $\Delta\omega$ zum Differential $\mathrm{d}\,\omega$ wird

$$f(t) = \frac{2}{\pi}\int_0^\infty \frac{1}{\omega}\sin\left(\omega\,\frac{t_0}{2}\right)\cos(\omega\,t)\,\mathrm{d}\,\omega \tag{361}$$

$$= \int_0^\infty A(\omega)\,\cos(\omega\,t)\,\mathrm{d}\,\omega \quad \text{mit} \quad A(\omega) = \frac{2}{\omega\,\pi}\,\sin\left(\frac{t_0}{2}\,\omega\right). \tag{362}$$

Aus den diskreten Koeffizienten a_n für $n = 1, 2, \ldots$ wird in diesem Grenzwert eine Funktion $A(\omega)$, die Integrationsgrenzen werden 0 und ∞.

Dies ist das spezifische Ergebnis für die Rechteckfunktion, die eine gerade Funktion der Zeit t ist.

Diese heuristische Betrachtung an einem Beispiel muss natürlich durch eine saubere allgemeine mathematische Beweisführung unterlegt werden, auf die hier jedoch aus Zeitmangel verzichtet wird.

6.6.2 Allgemeine Fouriertransformation

Gegeben sei eine nichtperiodische Funktion $f(t)$, die integrierbar ist; dies ist gewährleistet, wenn sie stückweise stetig ist und im Unendlichen hinreichend schnell verschwindet. Die Integrierbarkeit kann man als Verallgemeinerung des Begriffs eines zeitlich begrenzten Signals interpretieren.

Die allgemeine **Fouriertransformation** ist gegeben durch

$$f(t) = \int_0^\infty \Big(A(\omega)\, \cos(\omega\, t) + B(\omega)\, \sin(\omega\, t) \Big)\, \mathrm{d}\, \omega \tag{363}$$

mit dem **Amplitudenspektrum**

$$A(\omega) = \frac{1}{\pi} \int_{-\infty}^{\infty} f(t)\, \cos(\omega\, t)\, \mathrm{d}\, t, \tag{364}$$

$$B(\omega) = \frac{1}{\pi} \int_{-\infty}^{\infty} f(t)\, \sin(\omega\, t)\, \mathrm{d}\, t. \tag{365}$$

Für eine **gerade Funktion** $f(t)$ wird offensichtlich $B(\omega) \equiv 0$, für eine **ungerade Funktion** $f(t)$ wird $A(\omega) \equiv 0$.

Das Mirakel der Fourierdarstellung ist, dass sie erlaubt, eine *nichtperiodische* Funktion durch *periodische* trigonometrischen Funktionen darzustellen. Der Preis ist, dass die Reihenentwicklung durch eine Integraldarstellung ersetzt wird.

6.6.3 Komplexe Darstellung der Fouriertransformation

Ausgangspunkt ist wieder die *reell-wertige* Funktion $f(t)$. Die Fouriertransformation läßt sich sehr elegant und kompakt unter Hinweis auf die Euler-Formel (89), $e^{i\phi} = \cos(\phi) + i\, \sin(\phi)$, erweitern auf

die **komplexe Darstellung der Fouriertransformation**

$$f(t) = \int_{-\infty}^{\infty} F(\omega)\, e^{i\,\omega\, t}\, \mathrm{d}\, \omega. \tag{366}$$

Dabei ist $F(\omega)$ eine *komplex-wertige* Funktion $F(\omega) = \mathrm{Re}\, F(\omega) + i\, \mathrm{Im}\, F(\omega)$,

$$F(\omega) = \frac{1}{2\pi} \int_{-\infty}^{\infty} f(t)\, e^{-i\,\omega\, t}\, \mathrm{d}\, t. \tag{367}$$

Man beachte, dass im Gegensatz zur reellen Darstellung das Integral (366) von $-\infty$ und nicht von 0 an läuft und demensprechend $F(\omega)$ einen zusätzlichen Faktor $1/2$ gewinnt.

Setzt man (367) zurück in (366) ein, erwartet man als Ergebnis wieder $f(t)$

$$\int_{-\infty}^{\infty} F(\omega)\, e^{i\,\omega\, t}\, \mathrm{d}\, \omega \tag{368}$$

$$= \frac{1}{2\pi} \int_{-\infty}^{\infty} \mathrm{d}\, \omega\, e^{i\,\omega\, t} \int_{-\infty}^{\infty} f(t')\, e^{-i\,\omega\, t'}\, \mathrm{d}\, t' \tag{369}$$

$$= \frac{1}{2\pi} \int_{-\infty}^{\infty} f(t')\, \mathrm{d}\, t' \int_{-\infty}^{\infty} e^{i\,\omega (t-t')}\, \mathrm{d}\, \omega \tag{370}$$

$$= \int_{-\infty}^{\infty} f(t')\,\delta(t-t')\,\mathrm{d}\,t' \tag{371}$$
$$= f(t). \tag{372}$$

Dabei wurden die t'- und die ω-Integrationen vertauscht und es wurde Gebrauch von der *Distribution* $\delta(x)$ gemacht, die in der Vorlesung "Vorkurs zur Theoretischen Physik" eingeführt wurde. Ihre Definition soll hier vorausgesetzt werden und an die für das obige Argument benutzten Eigenschaften

$$\delta(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{i\,\omega\,x}\,\mathrm{d}\,\omega, \tag{373}$$

$$\int_a^b f(x)\,\delta(x-x_0)\,\mathrm{d}\,x = f(x_0) \tag{374}$$

soll erinnert werden.

7 Reelle Matrizen

7.1 Physikalische Motivation für Matrizen

- Räumliche Materialeigenschaften, wie sie im Trägheitstensor oder im Torsionstensor in der klassischen Mechanik oder im Polarisationstensor, im Dielektrizitätstensor oder im Tensor der Leitfähigkeit in der Elektrodynamik kodiert sind, erfordern eine Beschreibung mit Hilfe von Matrizen.
- Lineare Abbildungen im dreidimensionalen Raum lassen sich mit Hilfe von Matrizen wesentlich eleganter behandeln. Im Rahmen der speziellen Relativitätstheorie hat man es mit einem vierdimensionalen Koordinatensystem, bestehend aus den drei Raumkoordinaten und einer Zeitkoordinate, und dessen linearen Transformationen zu tun; hier spielt der Matrizenkalkül eine analoge Rolle.
- Matrizen spielen eine Rolle, wann immer in der Physik ein lineares Gleichungssystem zu lösen ist.
- Im Rahmen der Elektrodynamik tritt der elektromagnetische Feldstärketensor auf, eine Matrix, in die die Komponenten der elektrischen und magnetischen Feldstärken eingehen.
- Matrizen spielen häufig eine Rolle, wenn in der Physik ein System von gekoppelten Differentialgleichungen zu lösen ist. Darauf soll in Kap. 10.7 zurückgekommen werden.
- Eine hervorragende Rolle spielen Matrizen in der theoretischen Quantenmechanik: eine Formulierung der Quantenmechanik von Heisenberg ordnet jeder physikalischen Messgröße eine Matrix zu. Im Lauf dieses Kapitels und in Kap. 10 soll diese Anwendung noch weiter präzisiert werden.

7.2 Verallgemeinerung des Vektorbegriffs: der normierte reelle Vektorraum $\mathbb{R}^{\mathbf{n}}$

Bisher hatten wir uns in der Vorlesung auf dreidimensionale Vektoren im physikalischen Ortsraum $\mathbb{R}^3$ konzentriert. Um Matrizen beliebiger Dimension zu diskutieren, braucht man die abstrakte Erweiterung auf n-dimensionale Vektoren. Dazu werden als mathematisches Unterfutter

der n-dimensionale normierte reelle Vektorraum $\mathbb{R}^n$ und später – im Hinblick auf die Quantenmechanik – der n-dimensionale komplexe Hilbertraum benötigt. Diese werden ausführlich in den Mathematikvorlesungen behandelt. In dieser Vorlesung wird die Behandlung vom mathematischen Standpunkt aus unvollständig bleiben; die im Folgenden wichtigen Eigenschaften werden nach und nach in folgenden Kapiteln eingeführt.

Ein reeller n-dimensionaler Vektorraum $\mathbb{R}^n$ besteht aus Vektoren mit n reellen Komponenten; die Vektoren können mit einer reellen skalaren Zahl multipliziert werden und sie können miteinander addiert werden. Das Ergebnis ist jeweils wieder ein Vektor desselben Vektorraums. Als *Basis* eines Vektorraums bezeichnet man eine Menge von Vektoren, aus denen sich alle anderen Vektoren des Vektorraums mittels Skalarmultiplikation und Addition konstruieren lassen, also einen vollständigen Satz von linear unabhängigen Vektoren. Die Anzahl der Basisvektoren ist gleich der *Dimension* n des Vektorraums. Ein n-dimensionaler Vektorraum sei durch einen Satz von orthonormierten Basisvektoren $\vec{e}_1$, $\vec{e}_2$, ..., $\vec{e}_n$ aufgespannt. Der Begriff Orthonormalität setzt die Definition eines Skalarproduktes zweier Vektoren voraus, das weiter unten nachgeholt werden wird. Jeder Vektor des Vektorraums läßt sich dann folgendermaßen durch einen Spaltenvektor charakterisieren

$$\vec{x} = x_1\,\vec{e}_1 + x_2\,\vec{e}_2 + ... + x_n\,\vec{e}_n = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \tag{375}$$

mit den reellen Komponenten x_1, x_2, ..., x_n. Bei der Addition zweier n-dimensionaler Vektoren $\vec{x}$ und $\vec{y}$ werden die Komponenten addiert

$$\vec{x} + \vec{y} = (x_1 + y_1)\,\vec{e}_1 + (x_2 + y_2)\,\vec{e}_2 + ... + (x_n + y_n)\,\vec{e}_n = \begin{pmatrix} x_1 + y_1 \\ x_2 + y_2 \\ \vdots \\ x_n + y_n \end{pmatrix}. \tag{376}$$

Bei der Multiplikation eines Vektors $\vec{x}$ mit einem reellen Skalar λ wird jede Komponente mit λ multipliziert

$$\lambda\,\vec{x} = \lambda\,x_1\,\vec{e}_1 + \lambda\,x_2\,\vec{e}_2 + ... + \lambda\,x_n\,\vec{e}_n = \begin{pmatrix} \lambda\,x_1 \\ \lambda\,x_2 \\ \vdots \\ \lambda\,x_n \end{pmatrix}. \tag{377}$$

Im Matrixkalkül wird es wichtig sein, Spalten- und Zeilenvektoren zu unterscheiden. Deshalb wird bereits hier die sogenannte *Transposition* eingeführt, die sowohl auf Vektoren als auch auf Matrizen anwendbar ist. Sie vertauscht allgemein Spalten- und Zeilenvektoren. Gegeben der n-dimensionale Spaltenvektor $\vec{x}$, wird daraus der entsprechende n-dimensionale Zeilenvektor durch Transposition gewonnen und mit $\vec{x}^{\,T}$ bezeichnet

$$\vec{x}^{\,T} = (x_1, x_2, ..., x_n) \quad \text{mit} \quad (\vec{x}^{\,T})^{\,T} = \vec{x}. \tag{378}$$

Das Skalarprodukt von zwei n-dimensionalen Vektoren $\vec{x}$ und $\vec{y}$ mit Komponenten x_1, x_2, ..., x_n bzw. y_1, y_2, ..., y_n, das in Zukunft aus Konsistenzgründen immer als Produkt eines Zeilenvektors mit einem Spaltenvektor zu verstehen ist, ist definiert durch

$$\vec{x}^{\,T}\,\vec{y} = \vec{y}^{\,T}\,\vec{x} = x_1\,y_1 + x_2\,y_2 + ... + x_n\,y_n = \sum_{i=1}^{n} x_i\,y_i. \tag{379}$$

Das Ergebnis ist offensichtlich ein reeller Skalar. Die Reihenfolge "Zeilenvektor multipliziert mit Spaltenvektor" ist bei der Bildung des Skalarproduktes (379) essenziell. (Das Produkt eines Spaltenvektors mit einem Zeilenvektor, das sogenannte *dyadische Produkt*, das in Kap. 9.2 vorgestellt werden soll, führt auf eine Matrix.)

Das Skalarprodukt eines Vektors mit sich selbst,

$$\vec{x}^{\,T}\,\vec{x} = x_1^2 + x_2^2 + ... + x_n^2 = \sum_{i=1}^{n} x_i^2, \tag{380}$$

ist positiv semidefinit, d.h.

$$\vec{x}^{\,T}\,\vec{x} \geq 0 \quad \forall x \in \mathbb{R}^n. \tag{381}$$

Es führt auf die sogenannte *Norm* $\sqrt{\vec{x}^{\,T}\,\vec{x}}$, eine Verallgemeinerung des geometrischen Begriffs der Länge eines Vektors im dreidimensionalen Ortsraum,

$$\sqrt{\vec{x}^{\,T}\,\vec{x}} = \sqrt{x_1^2 + x_2^2 + ... + x_n^2} = \sqrt{\sum_{i=1}^{n} x_i^2}. \tag{382}$$

$\sqrt{\vec{x}^{\,T}\,\vec{x}} = 0$ gilt genau dann, wenn der Vektor $\vec{x}$ der n-komponentige Nullvektor

$$\vec{0} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \tag{383}$$

ist.

In Zukunft werde – wie bereits oben angekündigt – eine *orthonormale Basis* von n paarweise orthogonalen, auf 1 normierten Einheitsvektoren $\vec{e}_1$, $\vec{e}_2$, ..., $\vec{e}_n$ betrachtet

$$\vec{e}_i^{\,T}\,\vec{e}_j = \delta_{ij} \quad \text{für} \quad i,\, j \in \{1,\, 2,\, ...,\, n\} \tag{384}$$

$$= \begin{cases} 1 & \text{für} \quad i = j, \\ 0 & \text{sonst.} \end{cases} \tag{385}$$

Sie spannen den normierten reellen n-dimensionalen Vektorraum $\mathbb{R}^n$ auf. Die orthonormierte Basis

$$\vec{e}_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \quad \vec{e}_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \quad ..., \quad \vec{e}_n = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix} \tag{386}$$

liegt der Definition (375) zugrunde.

7.3 Definition einer Matrix und besondere Matrizen

7.3.1 Definition einer Matrix

Definition: Eine **reelle Matrix** A ist ein rechteckiges Zahlenschema

$$A = \left(a_{ij}\right) = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix} \tag{387}$$

von $m \times n$ reellen Zahlen a_{ij}, den sogenannten **Matrixelementen**; m und n sind beliebige natürliche Zahlen. Die Matrix soll auch alternativ durch ihre in Klammern gesetzten Elemente, $\left(a_{ij}\right)$, gekennzeichnet werden. Die Matrix heißt "vom Typ $m \times n$" oder man spricht von einer $m \times n$ - Matrix.

Matrizen werden üblicherweise mit lateinischen Großbuchstaben A, B, ... bezeichnet.

Die Matrix A besteht aus m horizontalen Zahlenreihen mit je n Komponenten, die mit dem Index $i = 1, 2, ..., m$ von oben nach unten durchnumeriert werden und die man als *Zeilen* der Matrix bezeichnet. Im Folgenden erweist es sich als sinnvoll, sie als Zeilenvektoren aufzufassen

$$(a_{i1}, a_{i2}, ..., a_{in}) \quad \text{für} \quad i = 1, \ ..., \ m. \tag{388}$$

Alternativ kann man die Matrix auffassen als bestehend aus n vertikalen Zahlenreihen mit je m Komponenten, die mit dem Index $j = 1, 2, ..., n$ von links nach rechts durchnumeriert werden und die man als *Spalten* der Matrix bezeichnet. Es wird sich auch hier als sinnvoll erweisen, die Spalten als Spaltenvektoren aufzufassen

$$\begin{pmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{mj} \end{pmatrix} \quad \text{für} \quad j = 1, \ ..., \ n. \tag{389}$$

Das Matrixelement a_{ij} befindet sich am "Kreuzungspunkt" der i-ten Zeile und der j-ten Spalte.

Spezialfälle sind quadratische Matrizen, d.h. Matrizen vom Typ $n \times n$, mit denen wir es später vorwiegend zu tun haben werden.

Einfache Beispiele sind die

- 2×2 - Matrix

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}, \quad \text{z.B.} \quad A = \begin{pmatrix} 2 & 7 \\ 33 & -8 \end{pmatrix}, \tag{390}$$

- 3×3 - Matrix

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}, \tag{391}$$

  wie sie z.B. bei dreidimensionalen Koordinatendrehungen oder allgemeiner bei dreidimensionalen linearen Abbildungen auftreten.

7.3.2 Besondere Matrizen

- Die $m \times n$-*Nullmatrix*

$$O = \Big(o_{ij}\Big) = \begin{pmatrix} 0 & \dots & 0 \\ \vdots & & \vdots \\ 0 & \dots & 0 \end{pmatrix} \quad \text{mit} \quad o_{ij} = 0 \ \text{ für alle } \ i = 1, 2, ..., m, \ j = 1, 2, ..., n. \tag{392}$$

- Die *Einheitsmatrix*, die nur für Matrizen vom Typ $n \times n$ definiert ist,

$$\mathbf{1} = \Big(\delta_{ij}\Big) = \begin{pmatrix} 1 & 0 & 0 & \dots & 0 \\ 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & & 0 \\ \vdots & \vdots & & \ddots & \vdots \\ 0 & 0 & 0 & \dots & 1 \end{pmatrix}, \tag{393}$$

  wobei zur Erinnerung das Kronecker-Symbol δ_{ij} definiert ist als

$$\delta_{ij} = \begin{cases} 1 & \text{für } \ i = j, \\ 0 & \text{sonst.} \end{cases} \tag{394}$$

  Die Einheitsmatrix hat nur nichtverschwindene Elemente auf der sogenannten *Hauptdiagonalen*, die alle den Wert 1 haben.

- Auch die *Diagonalmatrix* tritt nur bei quadratischen, d.h. $n \times n$-Matrizen auf. Sie hat ebenfalls nur nichtverschwindende Elemente auf der Hauptdiagonalen, die jedoch beliebig reell sein können,

$$D = \Big(d_{ii}\,\delta_{ij}\Big) = \begin{pmatrix} d_{11} & 0 & 0 & \dots & 0 \\ 0 & d_{22} & 0 & \dots & 0 \\ 0 & 0 & d_{33} & & 0 \\ \vdots & \vdots & & \ddots & \vdots \\ 0 & 0 & 0 & \dots & d_{nn} \end{pmatrix}. \tag{395}$$

7.4 Erste Rechenregeln mit Matrizen

Zwei Matrizen $A = \Big(a_{ij}\Big)$ und $B = \Big(b_{ij}\Big)$ vom gleichen Typ $m \times n$ sind gleich, wenn alle ihre Matrixelemente übereinstimmen

$$A = B \quad \text{falls} \quad a_{ij} = b_{ij} \quad \text{für alle} \quad i = 1, 2, ..., m, \ j = 1, 2, ..., n. \tag{396}$$

7.4.1 Addition von Matrizen

Definition: **Die Summe zweier Matrizen** $A = \left(a_{ij}\right)$ und $B = \left(b_{ij}\right)$ vom gleichen Typ $m \times n$ ist wieder eine $m \times n$-Matrix $C = \left(c_{ij}\right)$,

$$C = A + B = \begin{pmatrix} a_{11} + b_{11} & \dots & a_{1n} + b_{1n} \\ \vdots & & \vdots \\ a_{m1} + b_{m1} & \dots & a_{mn} + b_{mn} \end{pmatrix}, \tag{397}$$

d.h. die Matrixelemente mit dem gleichen Indexpaar werden addiert

$$c_{ij} = a_{ij} + b_{ij} \quad \text{für alle} \quad i = 1, 2, ..., m, \; j = 1, 2, ..., n. \tag{398}$$

Für die Matrixaddition gilt das **Kommutativgesetz**

$$A + B = B + A \tag{399}$$

sowie das **Assoziativgesetz**

$$(A + B) + C = A + (B + C). \tag{400}$$

7.4.2 Multiplikation einer Matrix mit einem Skalar

Definition: Die **Multiplikation einer Matrix mit einem Skalar**, einer rellen Zahl λ, erfolgt elementweise

$$\lambda\, A = \begin{pmatrix} \lambda\, a_{11} & \dots & \lambda\, a_{1n} \\ \vdots & & \vdots \\ \lambda\, a_{m1} & \dots & \lambda\, a_{mn} \end{pmatrix}. \tag{401}$$

7.5 Multiplikation einer Matrix mit einem Vektor

7.5.1 Hinführung mit Hilfe der dreidimensionalen linearen Abbildung

Im Hinblick auf n-dimensionale Verallgemeinerungen soll der Ortsvektor im dreidimensionalen Raum durch den Vektor

$$\vec{x} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \tag{402}$$

bezeichnet werden. Eine sogenannte lineare Abbildung z.B. in drei Dimensionen, bildet den Vektor $\vec{x}$ folgendermaßen auf den Vektor

$$\vec{x}\,' = \begin{pmatrix} x_1' \\ x_2' \\ x_3' \end{pmatrix} \quad \text{ab} \tag{403}$$

$$\begin{aligned} x_1' &= a_{11}\,x_1 + a_{12}\,x_2 + a_{13}\,x_3 \\ x_2' &= a_{21}\,x_1 + a_{22}\,x_2 + a_{23}\,x_3 \\ x_3' &= a_{31}\,x_1 + a_{32}\,x_2 + a_{33}\,x_3, \end{aligned} \tag{404}$$

wobei die Koeffizienten a_{ij} beliebige reelle Zahlen sind.

Lineare Abbildungen sind ganz allgemein von hervorragender Bedeutung in der Physik. Häufig ist eine Störung eines Systems klein genug, dass das System näherungsweise durch eine lineare Abhängigkeit von den Variablen beschrieben werden kann. Solche Situationen wurden bereits im Zusammenhang mit der Taylorentwicklung diskutiert. In diesem Kapitel spielen lineare Abbildungen eine zusätzliche wichtige Rolle bei der Einführung und anschaulichen Interpretation der Matrixmultiplikation.

Die Koeffizienten in (404) können sinnvollerweise in einer 3×3 - Matrix A zusammengefasst werden

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}. \tag{405}$$

Ziel ist es nun, die Multiplikation der Matrix A mit dem Vektor $\vec{x}$ so zu definieren, dass das Gleichungssystem (404) in der folgenden Vektorgleichung zusammengefasst werden kann

$$\vec{x}\,' = A\,\vec{x}. \tag{406}$$

7.5.2 Definition

Es gibt zwei verschiedene Möglichkeiten, eine Matrix mit einem Vektor zu multiplizieren, die *Rechtsmultiplikation* mit einem Spaltenvektor und die *Linksmultiplikation* mit einem Zeilenvektor.

- Definition: **Multiplikation mit einem Vektor von rechts**

  Die $m \times n$ - Matrix A wird *von rechts* mit einem n-dimensionalen Spaltenvektor $\vec{x}$ multipliziert, ausgeschrieben

$$A\,\vec{x} = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} \sum_{j=1}^{n} a_{1j}\,x_j \\ \sum_{j=1}^{n} a_{2j}\,x_j \\ \vdots \\ \sum_{j=1}^{n} a_{mj}\,x_j \end{pmatrix}. \tag{407}$$

  Hier wird daran erinnert, dass die Matrix A als aus m Zeilenvektoren (388) aufgebaut interpretiert werden kann, die mit $i = 1, ..., m$ durchnumeriert werden. Das Produkt $A\,\vec{x}$ der $m\times n$ - Matrix A mit dem *n-dimensionalen Spaltenvektor* $\vec{x}$ ist wieder ein Vektor und zwar ein

m-dimensionaler Spaltenvektor, dessen Komponenten durch die Skalarprodukte $\sum_{j=1}^{n} a_{ij}\, x_j$ zwischen den Zeilenvektoren (388) von A und dem Spaltenvektor $\vec{x}$ gegeben sind und durch den Index $i = 1, .., m$ durchnumeriert werden. Es wird klar, dass die Zeilen von A, als Zeilenvektoren (388) interpretiert, dieselbe Dimension n haben müssen wie der Vektor $\vec{x}$.

Definition: **Multiplikation mit einem Vektor von links**

Die $m \times n$ - Matrix A wird von links mit einem m-dimensionalen Zeilenvektor $\vec{x}^{\,T}$ multipliziert, ausgeschrieben

- $$\vec{x}^{\,T} A = \left(x_1, x_2, ..., x_m \right) \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix} \tag{408}$$

$$= \left(\sum_{i=1}^{m} a_{i1}\, x_i, \sum_{i=1}^{m} a_{i2}\, x_i, ..., \sum_{i=1}^{m} a_{in}\, x_i \right). \tag{409}$$

Hier wird der Vollständigkeit halber wieder daran erinnert, dass die Matrix A als aus n Spaltenvektoren (389) aufgebaut interpretiert werden kann, die mit $j = 1...n$ durchnumeriert werden. Das Produkt $\vec{x}^{\,T} A$ des *m-dimensionalen Zeilenvektors* $\vec{x}^{\,T}$ mit der $m \times n$ - Matrix A ist wieder ein Vektor und zwar ein *n-dimensionaler Zeilenvektor*, dessen Komponenten durch die Skalarprodukte $\sum_{i=1}^{m} a_{ij}\, x_i$ zwischen dem Zeilenvektor $\vec{x}^{\,T}$ und den Spaltenvektoren (389) gegeben sind und durch den Index $j = 1, .., n$ durchnumeriert werden. Wieder müssen die Spalten von A, als Spaltenvektoren (389) interpretiert, dieselbe Dimension m haben wie der Vektor $\vec{x}^{\,T}$.

Als Ergebnis sieht man, dass in der *Rechtsmultiplikation* eines Vektors mit einer Matrix der Vektor konsistent als Spaltenvektor auftritt und in der *Linksmultiplikation* konsistent als Zeilenvektor. Es ist zu beachten, dass die *Rechtsmultiplikation* auf ein anderes Ergebnis als die *Linksmultiplikation* führt. Dies reflektiert bereits eine fundamentale Eigenschaft der Multiplikation von Matrizen, die in Kap. 7.6.3 diskutiert wird. Beide, sowohl die Rechtsmultiplikation als auch die Linksmultiplikation, reduzieren sich in den erforderlichen Rechenoperationen auf die Bildung von Skalarprodukten von Zeilenvektoren mit Spaltenvektoren. Es ist auch hier wieder essenziell, dass diese Reihenfolge "Zeilenvektoren multipliziert mit Spaltenvektoren" im Skalarprodukt eingehalten wird.

Das in Kap. 7.5.1 definierte Ziel (406) ist offensichtlich mit Hilfe der Definition der Rechtsmultiplikation erreicht.

7.6 Multiplikation von Matrizen

7.6.1 Hinführung mit Hilfe der dreidimensionalen linearen Abbildung

Ein m-dimensionaler Spaltenvektor ist eine $m \times 1$ - Matrix, ein n-dimensionaler Zeilenvektor ist eine $1 \times n$ - Matrix. Die allgemeine Matrixmultiplikation muss dementsprechend die Rechts-oder Linksmultiplikation einer Matrix mit einem Vektor als Spezialfall enthalten. Genau genommen hätte man deshalb auf die separate Einführung der Multiplikation einer Matrix mit einem Vektor verzichten können; es schien jedoch aus pädagogischen Gründen angebracht.

Weitergehend ist die Hinführung zur Matrixmultiplikation, wenn man die Hintereinanderausführung zweier z.B. dreidimensionaler linearer Abbildungen vom Typ (404) betrachtet,

$$\vec{x}\,' = \begin{pmatrix} x_1' \\ x_2' \\ x_3' \end{pmatrix} \quad \text{mit} \quad \begin{aligned} x_1' &= a_{11}\,x_1 + a_{12}\,x_2 + a_{13}\,x_3 \\ x_2' &= a_{21}\,x_1 + a_{22}\,x_2 + a_{23}\,x_3 \\ x_3' &= a_{31}\,x_1 + a_{32}\,x_2 + a_{33}\,x_3 \end{aligned} , \tag{410}$$

zusammengefasst in der Gleichung

$$\vec{x}\,' = A\vec{x} \quad \text{mit der Matrix} \quad A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}, \tag{411}$$

sowie

$$\vec{x}\,'' = \begin{pmatrix} x_1'' \\ x_2'' \\ x_3'' \end{pmatrix} \quad \text{mit} \quad \begin{aligned} x_1'' &= b_{11}\,x_1' + b_{12}\,x_2' + b_{13}\,x_3' \\ x_2'' &= b_{21}\,x_1' + b_{22}\,x_2' + b_{23}\,x_3' \\ x_3'' &= b_{31}\,x_1' + b_{32}\,x_2' + b_{33}\,x_3' \end{aligned} , \tag{412}$$

zusammengefasst in der Gleichung

$$\vec{x}\,'' = B\vec{x}\,' \quad \text{mit der Matrix} \quad B = \begin{pmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \\ b_{31} & b_{32} & b_{33} \end{pmatrix}. \tag{413}$$

Die Hintereinanderausführung kann durch Einsetzen auf die Form gebracht werden

$$\vec{x}\,'' = B\,(A\,\vec{x}), \tag{414}$$

in der die auftretenden Matrizen jeweils mit einem Vektor multipliziert werden und zwar zunächst die Matrix A mit dem Vektor $\vec{x}$ und dann die Matrix B mit dem Vektor $A\,\vec{x}$. Schreibt man (414) um in der Form

$$\vec{x}\,'' = (B\,A)\,\vec{x}, \tag{415}$$

dann ist das Ziel, die Multiplikation $B\,A$ einer Matrix B mit einer Matrix A so zu definieren, dass (415) mit (414) übereinstimmt.

7.6.2 Multiplikation einer $m \times n$ - Matrix mit einer $n \times l$ - Matrix

Definition: **Matrixmultiplikation**

Die $m \times n$ - Matrix A wird von *von rechts* mit einer $n \times l$ - Matrix B multipliziert

$$A\,B = \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \dots & a_{mn} \end{pmatrix} \begin{pmatrix} b_{11} & \dots & b_{1l} \\ \vdots & & \vdots \\ b_{n1} & \dots & b_{nl} \end{pmatrix} = \Big(\sum_{k=1}^{n} a_{ik} b_{kj}\Big), \tag{416}$$

d.h. das ij-te Element des Matrixproduktes $A\,B$ ist das Skalarprodukt der i-ten Zeile von A, als n-dimensionaler Zeilenvektor

$$(a_{i1}, ..., a_{in}) \tag{417}$$

aufgefasst, mit der j-ten Spalte von B, als n-dimensionaler Spaltenvektor

$$\begin{pmatrix} b_{1j} \\ \vdots \\ b_{nj} \end{pmatrix} \tag{418}$$

aufgefasst. Das Ergebnis ist eine $m \times l$ - Matrix. Es ist ersichtlich, dass zwei Matrizen nur dann miteinander multipliziert werden können, wenn die Zahl der Spalten der linken Matrix gleich der Zahl der Zeilen der rechten Matrix ist.

7.6.3 Nichtkommutativität der Matrixmultiplikation

Die Matrixmultiplikation ist **nichtkommutativ**, d.h. im Allgemeinen gilt für das Produkt zweier beliebiger Matrizen A und B

$$AB \neq BA. \tag{419}$$

Dies sieht man an folgendem einfachen

Beispiel:

- $$\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} \begin{pmatrix} 5 & 6 \\ 7 & 8 \end{pmatrix} = \begin{pmatrix} 1\cdot 5+2\cdot 7 & 1\cdot 6+2\cdot 8 \\ 3\cdot 5+4\cdot 7 & 3\cdot 6+4\cdot 8 \end{pmatrix} = \begin{pmatrix} 19 & 22 \\ 43 & 50 \end{pmatrix}, \tag{420}$$
- $$\begin{pmatrix} 5 & 6 \\ 7 & 8 \end{pmatrix} \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} = \begin{pmatrix} 5\cdot 1+6\cdot 3 & 5\cdot 2+6\cdot 4 \\ 7\cdot 1+8\cdot 3 & 7\cdot 2+8\cdot 4 \end{pmatrix} = \begin{pmatrix} 23 & 34 \\ 31 & 46 \end{pmatrix}. \tag{421}$$

Die beiden Matrixprodukte sind klar ungleich.

Definition: Zwei Matrizen heißen **vertauschbar**, wenn gilt

$$AB = BA \tag{422}$$

So sind zum Beispiel die Einheitsmatrix $\mathbf{1}$ oder allgemeiner eine Diagonalmatrix D mit jeder Matrix A vertauschbar

$$\mathbf{1}\, A = A\, \mathbf{1} \quad \text{und} \quad DA = AD. \tag{423}$$

Die folgende Definition ist sinnvoll.

Definition: Der **Kommutator** zweier Matrizen A und B ist

$$[A, B] = AB - BA. \tag{424}$$

Für vertauschbare Matrizen gilt

$$[A, B] = 0. \tag{425}$$

Der Kommutator spielt eine hervorragende Rolle in der Heisenbergschen Matrix-Formulierung der Quantenmechanik: Messgrößen werden Matrizen zugeordnet und das Kriterium, ob zwei Messgrößen gleichzeitig scharf meßbar sind, also keiner Unschärferelation unterworfen sind, ist die Vertauschbarkeit der zugehörigen Matrizen. Unschärfe besteht bei Nichtvertauschbarkeit. Dies läßt sich zum Beispiel auch auf die berühmte Heisenbergsche Unschärferelation anwenden

$$\Delta x \, \Delta p \geq h, \tag{426}$$

wobei h das Plancksche Wirkungsquantum ist, Δx die prinzipielle Unschärfe einer Ortsmessung und Δp die prinzipielle Unschärfe einer Impulsmessung. In der Matrixformulierung sind den Messgrößen Ort bzw. Inpuls jeweils (unendlich-dimensionale) Matrizen zugeordnet, deren Kommutator ungleich Null und zwar proportional zu h ist.

7.6.4 Assoziativ- und Distributivgesetze der Matrixmultiplikation

Es gelten die folgenden **Assoziativ- und Distributivgesetze**

$$\Big(AB\Big)\,C = A\,\Big(BC\Big), \tag{427}$$

$$\Big(A+B\Big)\,C = A\,C + B\,C, \tag{428}$$

$$C\,\Big(A+B\Big) = C\,A + C\,B. \tag{429}$$

Wegen der Nichtkommutativität der Matrixmultiplikation kommt es überall auf die Reihenfolge an, in der die Matrizen im Produkt auftreten.

7.7 Transposition von Matrizen

Gegeben sei eine $m \times n$ - Matrix

$$A = \Big(a_{ij}\Big) = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix}. \tag{430}$$

Definition: Die **transponierte Matrix** A^T erhält man durch Vertauschung jeder Zeile mit der entsprechenden Spalte

$$A^T = \Big(a_{ji}\Big) = \begin{pmatrix} a_{11} & a_{21} & \dots & a_{m1} \\ a_{12} & a_{22} & \dots & a_{m2} \\ \vdots & \vdots & & \vdots \\ a_{1n} & a_{2n} & \dots & a_{mn} \end{pmatrix}. \tag{431}$$

A^T ist dementsprechend eine $n \times m$ - Matrix; ihre Matrixelemente a_{ji} erhält man aus den Matrixelementen a_{ij} von A durch Vertauschung der Indices i und j. Es gilt offensichtlich, dass zweimalige Transposition wieder zum Ausgangspunkt zurückführt

$$(A^T)^T = A. \tag{432}$$

Für die Transposition des Produktes AB einer $m \times n$ - Matrix A mit einer $n \times l$ - Matrix B gilt

$$(AB)^T = B^T A^T, \tag{433}$$

wie man sich leicht auf dem Niveau der Matrixelemente herleiten kann

$$(AB)^T = \Big(\sum_{k=1}^{n} a_{ik} b_{kj}\Big)^T = \Big(\sum_{k=1}^{n} a_{jk} b_{ki}\Big) = \Big(\sum_{k=1}^{n} b_{ki} a_{jk}\Big) = B^T A^T. \tag{434}$$

Die Produkte $A\,\vec{x}$ einer $m \times n$ - Matrix A mit einem n-dimensionalen Spaltenvektor $\vec{x}$ und $\vec{y}^T A$ eines m-dimensionalen Zeilenvektor $\vec{y}^T$ mit einer $m \times n$ - Matrix A verhalten sich entsprechend unter Transposition wie folgt

$$\begin{aligned} (A\,\vec{x})^T &= \vec{x}^T A^T & (435)\\ (\vec{y}^T A)^T &= A^T\,(\vec{y}^T)^T = A^T\,\vec{y}. & (436) \end{aligned}$$

Der Skalar $\vec{y}^T A\,\vec{x}$ bleibt erwartungsgemäß invariant, d.h. unverändert, unter der Transposition

$$(\vec{y}^T A\,\vec{x})^T = \vec{x}^T\,A^T\,\vec{y} = \sum_{i=1}^{n}\sum_{j=1}^{m} x_i a_{ji} y_j \overset{i\leftrightarrow j}{=} \sum_{j=1}^{n}\sum_{i=1}^{m} x_j\,a_{ij}\,y_i = \sum_{i=1}^{m}\sum_{j=1}^{n} y_i\,a_{ij}\,x_j = \vec{y}^T A\,\vec{x}. \tag{437}$$

8 Matrixinversion und Determinanten

8.1 Inverse Matrizen I – Fragestellung und Motivation

Der Ausgangspunkt ist eine lineare Abbildung des normierten reellen n-dimensionalen Vektorraums auf sich selbst

$$\begin{aligned} x_1' &= a_{11}\,x_1 + ... + a_{1n}\,x_n, \\ &\vdots \\ x_n' &= a_{n1}\,x_1 + ... + a_{nn}\,x_n \end{aligned} \tag{438}$$

oder kompakt geschrieben

$$\begin{pmatrix} x_1' \\ \vdots \\ x_n' \end{pmatrix} = \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \dots & a_{nn} \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \quad \text{bzw.} \tag{439}$$

$$\vec{x}\,' = A\,\vec{x}. \tag{440}$$

A ist eine $n \times n$ - Matrix. Gesucht ist die Rücktransformation oder *Inversion*; d.h. man fasst die lineare Abbildung (404) als inhomogenes Gleichungssystem in den unabhängigen Variablen

x_1, ..., x_n auf und fragt nach der Auflösung des Gleichungssystem nach diesen Variablen. Gesucht ist also nach einer $n \times n$ - Matrix A^{-1}, die von links an beide Seiten der Gleichung (440) multipliziert auf

$$A^{-1}\,\vec{x}\,' = A^{-1}\,A\,\vec{x} = \vec{x} \tag{441}$$

führt. Dies soll für alle Vektoren $\vec{x}$ gelten, so dass die Matrixgleichung

$$A^{-1}\,A = \mathbf{1} \tag{442}$$

gelten muss, wobei $\mathbf{1}$ die Einheitsmatrix (393) ist. Zunächst muss man einige Fragen klären: unter welchen Umständen existiert zu einer vorgegebenen Matrix A eine solche *inverse Matrix* A^{-1}? Falls sie existiert, ist sie eindeutig? Falls sie existiert (und eindeutig ist), wie berechnet man sie?

Zunächst können einige Vorbetrachtungen gemacht werden. Vorausgesetzt A^{-1} existiert und ist eindeutig und es gilt (442) auch für $(A^{-1})^{-1}$, dann folgt

$$(A^{-1})^{-1}\;A^{-1} = \mathbf{1}. \tag{443}$$

Multipliziert man beide Seiten der Gleichung von rechts mit A und benutzt (442), dann findet man – nicht verwunderlich –

$$(A^{-1})^{-1} = A; \tag{444}$$

eingesetzt in (443) führt dies zu $A\,A^{-1} = \mathbf{1}$.

Also gilt insgesamt für die

inverse Matrix A^{-1}

$$A\,A^{-1} = A^{-1}\,A = \mathbf{1}, \tag{445}$$

vorausgesetzt A^{-1} existiert und ist eindeutig.

Einfaches Beispiel:

Gegeben sei das Gleichungssystem

$$\begin{aligned} x_1' &= a_{11}\,x_1 + a_{12}\,x_2 \\ x_2' &= a_{21}\,x_1 + a_{22}\,x_2 \end{aligned} \tag{446}$$

bzw.

$$\vec{x}\,' = A\,\vec{x} \;\text{ mit }\; \vec{x} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix},\;\; \vec{x}\,' = \begin{pmatrix} x_1' \\ x_2' \end{pmatrix},\;\; A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}. \tag{447}$$

Das Gleichungssystem (446) hat die Lösung

$$\vec{x} = A^{-1}\,\vec{x}\,' \;\text{ mit} \tag{448}$$

$$x_1 = \frac{1}{a_{11}\,a_{22} - a_{12}\,a_{21}}\,(a_{22}\,x_1' - a_{12}\,x_2') \tag{449}$$

$$x_2 = \frac{1}{a_{11}\,a_{22} - a_{12}\,a_{21}}\,(-a_{21}\,x_1' + a_{11}\,x_2'). \tag{450}$$

Daraus liest man ab

$$A^{-1} = \frac{1}{a_{11}\,a_{22} - a_{12}\,a_{21}} \begin{pmatrix} a_{22} & -a_{12} \\ -a_{21} & a_{11} \end{pmatrix}. \tag{451}$$

Offensichtlich existiert in diesem Beispiel A^{-1} nur, wenn die skalare Größe $\det A$

$$\det A = |A| = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}\, a_{22} - a_{12}\, a_{22} \neq 0 \tag{452}$$

ungleich Null ist; dann ist A^{-1} auch eindeutig. Von der Gültigkeit der Beziehungen (445) kann man sich leicht überzeugen. Hier ist bereits die sogenannte *Determinante* einer 2×2 - Matrix in drei Schreibweisen vorgestellt, definiert sowie als entscheidende Größe isoliert worden, mit deren Hilfe sich ein Kriterium für die eindeutige Existenz einer inversen Matrix formulieren läßt. Das nächste Unterkapitel wird sich diesem recht umfangreichen Thema für $n \times n$ - Matrizen widmen.

8.2 Determinante einer Matrix

Die Determinante ist nur für quadratische, d.h. $n \times n$ - Matrizen definiert.

8.2.1 Definition der Determinante

Eine erste Definition der **Determinante einer n × n - Matrix** ist

$$\begin{aligned} \det A &= \begin{vmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \dots & a_{nn} \end{vmatrix} \\ &= a_{11}\, A_{11} - a_{12}\, A_{12} + a_{13}\, A_{13} - + \ldots a_{1n}\, A_{1n} = \sum_{j=1}^{n} (-1)^{1-j}\, a_{1j}\, A_{1j}, \end{aligned} \tag{453}$$

die sich später als spezifische *Entwicklung der Determinante nach der 1. Zeile* herausstellen wird.

Die folgende *Entwicklung der Determinante nach der 1. Spalte* wird sich mit Hilfe der Leibnizformeln (461) und (462) als äquivalent zu (453) herausstellen

$$\begin{aligned} \det A &= \begin{vmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \dots & a_{nn} \end{vmatrix} \\ &= a_{11}\, A_{11} - a_{21}\, A_{21} + a_{31}\, A_{31} - + \ldots a_{n1}\, A_{n1} = \sum_{i=1}^{n} (-1)^{1-i}\, a_{i1}\, A_{i1} \end{aligned} \tag{454}$$

Dabei gilt allgemein

A_{ij} ist die **Determinante** derjenigen $(n-1)\times(n-1)$-**Untermatrix**, die aus der $n\times n$-Matrix A durch Streichung der i-ten Zeile und der j-ten Spalte hervorgeht

$$A_{ij} = \begin{vmatrix} a_{11} & \dots & a_{1j} & \dots & a_{1n} \\ \vdots & & \vdots & & \vdots \\ a_{i1} & \dots & a_{ij} & \dots & a_{in} \\ \vdots & & \vdots & & \vdots \\ a_{n1} & \dots & a_{nj} & \dots & a_{nn} \end{vmatrix} . \tag{455}$$

Beispiel n = 2:

$$\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}\,A_{11} - a_{12}\,A_{12} = a_{11}\,a_{22} - a_{12}\,a_{21}. \tag{456}$$

Beispiel n = 3:

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = a_{11}\,A_{11} - a_{12}\,A_{12} + a_{13}\,A_{13} \tag{457}$$

$$= a_{11} \begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} - a_{12} \begin{vmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{vmatrix} + a_{13} \begin{vmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{vmatrix} \tag{458}$$

$$= a_{11}\,a_{22}\,a_{33} - a_{11}\,a_{23}\,a_{32} - a_{12}\,a_{21}\,a_{33} + a_{12}\,a_{23}\,a_{31} + a_{13}\,a_{21}\,a_{32} - a_{13}\,a_{22}\,a_{31} \tag{459}$$

$$= \sum_P (-1)^\tau\, a_{1j_1}\, a_{2j_2}\, a_{3j_3} = \sum_P (-1)^\tau\, a_{i_1 1}\, a_{i_2 2}\, a_{i_3 3}. \tag{460}$$

Die Berechnung der 2×2-Unterdeterminanten erfolgt rekursiv nach (453) bzw. (454). In den Summen in (460) durchlaufen die Spaltenindices $j_1,\ j_2,\ j_3$ bzw. die Zeilenindices $i_1,\ i_2,\ i_3$ alle möglichen *Permutationen* von 1, 2, 3. Die Zahl τ gibt an, wieviele Vertauschungen von zwei der drei Indices, sogenannte *Transpositionen*, in $j_1,\ j_2,\ j_3$ bzw. in $i_1,\ i_2,\ i_3$ notwendig sind, um sie von der vorgegebenen Permutation in die natürliche Reihenfolge 1, 2, 3 zu bringen.

Permutation P von $1,2,3$	τ	$(-1)^\tau$
1 2 3	0	+1
1 3 2	1	−1
2 1 3	1	−1
2 3 1	2	+1
3 1 2	2	+1
3 2 1	1	−1

Es gibt $3! = 3\cdot 2\cdot 1 = 6$ Permutationen. Es ist zu beachten, dass jeder Summand in (460) genau ein Matrixelement aus jeder Zeile bzw. genau ein Matrixelement aus jeder Spalte besitzt.

Allgemein führt die Berechnung der Determinante der $n\times n$-Matrix A nach (453) zunächst auf die Entwicklung nach $(n-1)\times(n-1)$-Unterdeterminanten, die ihrerseits nach (453) jeweils nach $(n-2)\times(n-2)$-Unterdeterminanten entwickelt werden können, usw. Diese rekursive Prozedur führt auf die sogenannte

Leibnizformel in zwei äquivalenten Varianten:

$$\det A = \begin{vmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \dots & a_{nn} \end{vmatrix} = \sum_P (-1)^\tau \, a_{1j_1} \, a_{2j_2} ... \, a_{nj_n} \tag{461}$$

$$= \sum_P (-1)^\tau \, a_{i_1 1} \, a_{i_2 j} ... \, a_{i_n n}, \tag{462}$$

wobei in diesen Summen die Spaltenindices j_1, j_2, ..., j_n bzw. die Zeilenindices i_1, i_2, ..., i_n jeweils die $n!$ möglichen Permutationen P der Zahlenfolge 1, 2, ..., n durchlaufen. τ ist die Zahl der Transpositionen, d.h. Vertauschungen von zwei der n Indices, die notwendig sind, um die vorgegebene Permutation j_1, j_2, ..., j_n bzw. i_1, i_2, ..., i_n auf die natürliche Reihenfolge 1, 2, ..., n zu bringen. Es gilt natürlich

$$(-1)^\tau = \begin{cases} +1 & \text{für } \tau \text{ gerade} \\ -1 & \text{für } \tau \text{ ungerade.} \end{cases} \tag{463}$$

Der Beweis der Leibnizformeln (461) und (462) kann mit Hilfe der vollständigen Induktion geführt werden.

In manchen Lehrbüchern enthält die Leibnizformel statt des Faktors $(-1)^\tau$ einen Faktor $(-1)^\pi$, wobei π die Zahl der Vertauschungen von jeweils *zwei benachbarten Indices* ist, die notwendig sind, um j_1, j_2, ..., j_n bzw. i_1, i_2, ..., i_n auf die natürliche Reihenfolge 1, 2, ..., n zu bringen. Es ist jedoch

$$(-1)^\tau = (-1)^\pi, \tag{464}$$

da *eine* Vertauschung *zweier beliebiger* Indices, zwischen denen k weitere Indices stehen, sich durch *eine Abfolge von* $2k+1$ Vertauschungen von *benachbarten Indexpaaren* erzeugen läßt, wobei $2k+1$ stets ungerade ist.

Aus der Leibnizformel (461) oder (462) liest man ab, dass jeder Summand in seinem Produkt genau ein Matrixelement aus jeder Zeile und genau ein Matrixelement aus jeder Spalte enthält. In keinem Summanden tritt also ein Matrixelement zweimal auf.

8.2.2 Eigenschaften von Determinanten

1. Unter der Transposition werden die Zeilen einer Matrix A mit den entsprechenden Spalten vertauscht. Auf Grund der Gleichheit der Leibnizformeln (461) und (462), in denen die Rolle der Spalten mit der Rolle der Zeilen vertauscht ist, gilt

$$\det A = \det A^T. \tag{465}$$

2. Werden zwei beliebige Spalten bzw. zwei beliebige Zeilen einer Matrix A miteinander vertauscht, ändert sich das Vorzeichen der Determinante $\det A$,

   da dann jeder Term in der Leibnizformel (461) bzw. (462) auf Grund des Faktors $(-1)^\tau$ sein Vorzeichen wechselt. Man kann die Determinante $\det A$ also nach jeder ihrer Zeilen bzw.

nach jeder ihrer Spalten entwickeln. Die Entwicklung nach der i-ten Zeile ergibt sich zu

$$\det A = \sum_{j=1}^{n} (-1)^{i-j}\, a_{ij}\, A_{ij} \quad \text{für} \quad i = 1,\, 2,\, ...,\, n, \tag{466}$$

die Entwicklung nach der j-ten Spalte ergibt sich zu

$$\det A = \sum_{i=1}^{n} (-1)^{i-j}\, a_{ij}\, A_{ij} \quad \text{für} \quad j = 1,\, 2,\, ...,\, n. \tag{467}$$

3. Wird *eine* Zeile oder *eine* Spalte einer Matrix mit einem Skalar λ multipliziert, entspricht dies einer Multiplikation der Determinante mit λ

$$\begin{vmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ \lambda a_{i1} & \dots & \lambda a_{in} \\ \vdots & & \vdots \\ a_{n1} & \dots & a_{nn} \end{vmatrix} = \begin{vmatrix} a_{11} & \dots & \lambda a_{1j} & \dots & a_{1n} \\ \vdots & & \vdots & & \\ a_{n1} & \dots & \lambda a_{nj} & \dots & a_{nn} \end{vmatrix} = \lambda \begin{vmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \dots & a_{nn} \end{vmatrix}, \tag{468}$$

da jeder Summand in den Leibnizformeln (461) und (462) genau ein Element aus jeder Zeile bzw. aus jeder Spalte enthält. Es gilt dementsprechend

$$\det(\lambda A) = \lambda^n \det A, \tag{469}$$

da jede Zeile bzw. jede Spalte mit λ multipliziert wird.

4. Sind zwei Spalten oder zwei Zeilen von A gleich, verschwindet die Determinante $\det A$.

Vertauscht man zwei gleiche Zeilen oder Spalten einer Matrix A, ist die Determinante der resultierenden Matrix A'

$$\text{einerseits} \qquad \det A' = \det A, \quad \text{da } A' = A, \tag{470}$$

$$\text{andererseits} \qquad \det A' = -\det A, \quad \text{nach 2.}, \tag{471}$$

$$\text{woraus folgt} \qquad \det A = 0. \tag{472}$$

Dies gilt auch, wenn zwei Zeilen oder zwei Spalten proportional zueinander sind, da dann die Determinante mit Hilfe von (468) auf eine Determinante mit zwei gleichen Zeilen bzw. Spalten reduziert werden kann.

5. Eine Determinante ist additiv in Bezug auf die Elemente *einer* Zeile oder *einer* Spalte, da nach (461) und (461) jeder Summand genau ein Element aus jeder Zeile bzw. aus jeder Spalte enthält. So gilt z.B.

$$\begin{vmatrix} a_{11} + a'_{11} & \dots & a_{1n} + a'_{1n} \\ a_{21} & \dots & a_{2n} \\ \vdots & & \vdots \\ a_{n1} & \dots & a_{nn} \end{vmatrix} = \begin{vmatrix} a_{11} & \dots & a_{1n} \\ a_{21} & \dots & a_{2n} \\ \vdots & & \vdots \\ a_{n1} & \dots & a_{nn} \end{vmatrix} + \begin{vmatrix} a'_{11} & \dots & a'_{1n} \\ a_{21} & \dots & a_{2n} \\ \vdots & & \vdots \\ a_{n1} & \dots & a_{nn} \end{vmatrix} \tag{473}$$

6. Eine Determinante ändert sich nicht, wenn man zu einer Zeile ein Vielfaches einer anderen Zeile addiert bzw. zu einer Spalte ein Vielfaches einer anderen Spalte addiert. Dies folgt aus den Eigenschaften 3., 4. und 5. Dies führt auf die allgemeine Eigenschaft

   Sind die Zeilen- oder Spaltenvektoren einer Matrix **linear abhängig, verschwindet die Determinante**.

7. Es gilt

$$\det(AB) = \det A \, \det B. \tag{474}$$

8. Für eine Diagonalmatrix (395) mit Diagonalelementen d_{11}, d_{22}, ..., d_{nn} gilt

$$\det D = d_{11} \, d_{22} \ldots , d_{nn}. \tag{475}$$

   Für die Einheitsmatrix gilt entsprechend

$$\det \mathbf{1} = 1^n = 1. \tag{476}$$

Die Eigenschaften 1. bis 6. sind sehr nützlich bei der Berechnung von Determinanten kleinerer Matrizen. Dies wird in den Übungen ausführlich vorgeführt.

8.2.3 Alternative Berechnung einer Determinante

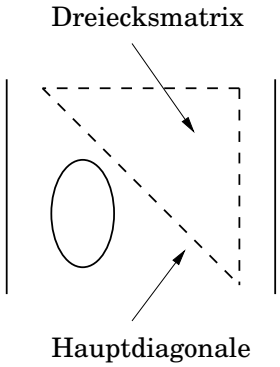
Zur Berechnung der Determinante einer $n \times n$-Matrix A

$$\det A = \begin{vmatrix} a_{11} & \ldots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \ldots & a_{nn} \end{vmatrix} \tag{477}$$

soll die Eigenschaft extensiv und systematisch benutzt werden, dass die Determinante invariant unter der Addition eines beliebigen Vielfachen einer beliebigen Zeile zu einer anderen Zeile ist.

Das Ziel ist, mit Hilfe dieser Methode die Determinante auf die folgende generische Form zu bringen

$$\det A = \tag{478}$$



Die Matrixelemente unterhalb der Hauptdiagonalen sollen zum Verschwinden gebracht werden, was die Berechnung der Determinante enorm vereinfacht.

- In einem ersten Schritt ist es notwendig, dass das Element a_{11} ungleich Null ist; gegebenenfalls muss dazu eine Umordnung der Zeilen vorgenommen und der entsprechende Vorzeichenwechsel der Determinante beachtet werden. Dann werden alle Matrixelemente in der Spalte unterhalb von a_{11} auf Null gebracht, indem

  - das a_{21}/a_{11}-fache der 1-ten Zeile von der 2-ten Zeile abgezogen wird,
    $\vdots$
  - das a_{n1}/a_{11}-fache der 1-ten Zeile von der n-ten Zeile abgezogen wird.

  Die resultierenden Matrixelemente in der $(n-1)\times(n-1)$ - Untermatrix für $i,\, j \in \{2,\, 3,\, ...,\, n\}$ werden umbenannt in $a_{ij}^{(2)}$

$$\det A = \det \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} - \frac{a_{21}}{a_{11}} a_{11} & a_{22} - \frac{a_{21}}{a_{11}} a_{12} & \dots & a_{2n} - \frac{a_{21}}{a_{11}} a_{1n} \\ \vdots & & \vdots & \\ a_{n1} - \frac{a_{n1}}{a_{11}} a_{11} & a_{n2} - \frac{a_{n1}}{a_{11}} a_{1n} & \dots & a_{nn} - \frac{a_{n1}}{a_{11}} a_{1n} \end{pmatrix} \tag{479}$$

$$= \det \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ 0 & a_{22}^{(2)} & \dots & a_{2n}^{(2)} \\ \vdots & \vdots & & \vdots \\ 0 & a_{n2}^{(2)} & \dots & a_{nn}^{(2)} \end{pmatrix} \tag{480}$$

- Mit der analogen Prozedur werden alle Matrixelemente unterhalb $a_{22}^{(2)}$ zu Null gemacht, wobei die Nullen in der ersten Spalte erhalten bleiben.

- Das Endergebnis ist die Determinante einer Matrix vom Typ (478), bei der nur noch die Elemente auf der Hauptdiagonalen von Interesse sind

$$\det A = \det \begin{pmatrix} a_{11} & . & . & \dots & . \\ 0 & a_{22}^{(2)} & . & \dots & . \\ 0 & 0 & a_{33}^{(3)} & \dots & . \\ \vdots & \vdots & & & \vdots \\ 0 & 0 & 0 & \dots & a_{nn}^{(n)} \end{pmatrix}. \tag{481}$$

  Die Diagonalelemente $a_{ii}^{(i)}$ für $i = 2,\, ...,\, n$ sind wohldefinierte Ausdrücke in den Matrixelemente a_{ij}.

- Diese Determinante wird nach (454) nach der ersten Spalte entwickelt, die nur ein einziges nichtverschwindendes Element, a_{11}, besitzt. Die zugehörige Unterdeterminante wird ebenfalls nach ihrer ersten Spalte entwickelt, die wieder nur ein einziges nichtverschwindendes Element, $a_{22}^{(2)}$, besitzt usw. Diese Prozedur führt auf das Produkt der Haupdiagonalelemente von (481)

$$\det A = a_{11}\, a_{22}^{(2)}\, a_{33}^{(3)} \,...\, a_{nn}^{(n)}. \tag{482}$$

8.2.4 Sarrussche Regel für Determinanten dritten Grades

Für 3×3 - Matrizen läßt sich die Determinante memnonisch einfach bestimmen nach der

Sarrusschen Regel

$$\det A = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{23} & a_{33} \end{vmatrix} = \left| \begin{array}{ccc} \mathbf{a_{11}} & \mathbf{a_{12}} & \mathbf{a_{13}} \\ \mathbf{a_{21}} & \mathbf{a_{22}} & \mathbf{a_{23}} \\ \mathbf{a_{31}} & \mathbf{a_{32}} & \mathbf{a_{33}} \end{array} \right| - \left| \begin{array}{ccc} \mathbf{a_{11}} & \mathbf{a_{12}} & \mathbf{a_{13}} \\ \mathbf{a_{21}} & \mathbf{a_{22}} & \mathbf{a_{23}} \\ \mathbf{a_{31}} & \mathbf{a_{32}} & \mathbf{a_{33}} \end{array} \right| \tag{483}$$

wobei die beiden Summanden folgendermaßen zu interpretieren sind. Im ersten Summanden werden drei Produkte aus jeweils drei Matrixelementen, die in gleichartigen Boxen *parallel zur Hauptdiagonalen* liegen, mit *positivem* Vorzeichen addiert zur Summe

$$a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{21}a_{32}a_{13}. \tag{484}$$

Im zweiten Summanden werden drei Produkte aus jeweils drei Matrixelementen, die in gleichartigen Boxen *parallel zur Nebendiagonalen* liegen, mit *negativen* Vorzeichen versehen addiert zur Summe

$$-a_{31}a_{22}a_{13} - a_{21}a_{12}a_{33} - a_{11}a_{32}a_{23}. \tag{485}$$

Die Summe der beiden Ausdrücke ergibt die Determinante.

8.2.5 Vektorprodukt und Determinante

Die folgende Schreibweise ist hilfreich für die Memnonik des Vektorproduktes. Seien $\vec{a} = (a_1,\, a_2,\, a_3)$ und $\vec{b} = (b_1,\, b_2,\, b_3)$ zwei dreidimensionale Vektoren. Ihr Vektorprodukt ist der Vektor

$$\begin{aligned} \vec{a} \times \vec{b} &= (a_2b_3 - a_3b_2,\, a_3b_1 - a_1b_3,\, a_1b_2 - a_2b_1) && (486)\\ &= (a_2b_3 - a_3b_2)\,\vec{e}_1 + (a_3b_1 - a_1b_3)\,\vec{e}_2 + (a_1b_2 - a_2b_1)\,\vec{e}_3 && (487)\\ &= \begin{vmatrix} a_2 & a_3 \\ b_2 & b_3 \end{vmatrix} \vec{e}_1 - \begin{vmatrix} a_1 & a_3 \\ b_1 & b_3 \end{vmatrix} \vec{e}_2 + \begin{vmatrix} a_1 & a_2 \\ b_1 & b_2 \end{vmatrix} \vec{e}_3 && (488)\\ &\overset{\text{formal}}{=} \begin{vmatrix} \vec{e}_1 & \vec{e}_2 & \vec{e}_2 \\ a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \end{vmatrix}. && (489) \end{aligned}$$

Die drei Vektoren $\vec{e}_1$, $\vec{e}_2$, $\vec{e}_3$ sind die Einheitsvektoren, die den dreidimensionalen Ortsraum aufspannen. Es wurde die Entwicklung (453) der Determinante nach ihrer ersten Zeile benutzt. Bei dieser *formalen* Umschreibung des Vektorproduktes in Form einer Determinante einer dreidimensionalen Matrix ist wie stets auf die Reihenfolge der Zeilen zu achten.

Die Umschreibung (489) ist nur sinnvoll für die Memnonik; mathematisch ist eine derartige Verquickung von Vektoren und Skalaren nicht definiert.

8.2.6 Spatprodukt als Determinante

Das Spatprodukt aus drei dreidimensionalen Vektoren $\vec{a}_1 = (a_{11},\, a_{12},\, a_{13})$, $\vec{a}_2 = (a_{21},\, a_{22},\, a_{23})$, $\vec{a}_3 = (a_{31},\, a_{32},\, a_{33})$ läßt sich analog als Determinante einer dreidimensionalen Matrix umschreiben

$$\vec{a}_1 \cdot (\vec{a}_2 \times \vec{a}_3) = a_{11} \begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} - a_{12} \begin{vmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{vmatrix} + a_{13} \begin{vmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{vmatrix} \tag{490}$$

$$= \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}. \tag{491}$$

Dies ist eine echte Identität. Es wurde wieder die Entwicklung (453) der Determinante nach der ersten Zeile benutzt.

In den Übungen war gezeigt worden, dass das Volumen eines von den drei Vektoren im dreidimensionalen Raum aufgespannten Parallelepipeds, Abb. 30, durch den Betrag der Determinante (491) gegeben ist.

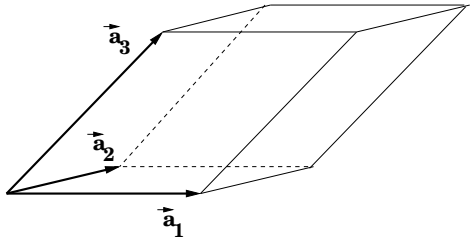


Abbildung 30:

Diese Aussage läßt sich auf n-dimensionale Vektorräume verallgemeinern. Gegeben seien n n-dimensionale Vektoren

$$\begin{aligned} \vec{a}_1 &= (a_{11} \ldots a_{1n}) \\ &\vdots \\ \vec{a}_n &= (a_{n1} \ldots a_{nn}). \end{aligned} \tag{492}$$

Das n-dimensionale Volumen des durch die n Vektoren aufgespannten verallgemeinerten Parallelepipeds ist der Betrag der Determinante

$$\begin{vmatrix} a_{11} & \ldots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \ldots & a_{nn} \end{vmatrix}. \tag{493}$$

Das Ergebnis ist nur von Null verschieden, wenn die n Vektoren linear unabhängig sind.

8.3 Inverse Matrizen II

8.3.1 Bestimmung der inversen Matrix

Gegeben sei eine $n \times n$-Matrix A

$$A = \begin{pmatrix} a_{11} & \ldots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \ldots & a_{nn} \end{pmatrix}. \tag{494}$$

Es sei aus Kap. 8.1 daran erinnert, dass für die zu A inverse Matrix A^{-1} gilt

$$A\,A^{-1} = A^{-1}\,A = \mathbf{1}, \tag{495}$$

vorausgesetzt sie existiert und ist eindeutig. Es gilt nun

Inverse Matrix

$$A^{-1} = \left(\tilde{a}_{ij}\right) \text{ mit} \tag{496}$$

$$\tilde{a}_{ij} = \frac{(-1)^{j-i} A_{ji}}{\det A}, \tag{497}$$

wobei A_{ij} die Determinante (455) derjenigen $(n-1) \times (n-1)$ - Untermatrix von A ist, die aus A durch Streichung der i-ten Zeile und der j-ten Spalte hervorgeht. Es ist zu beachten, dass in (497) die Unterdeterminante in der Form A_{ji}, d.h. mit vertauschten Indices, auftritt. Aus (497) kann man direkt ablesen:

die **inverse Matrix** A^{-1} **existiert** und ist **eindeutig** genau dann, wenn

$$\det A \neq 0. \tag{498}$$

Zum Beweis der Behauptung (496) und (497) wird das ij-te Element des Produktes AA^{-1}, das Skalarprodukt aus dem i-ten Zeilenvektor von A mit dem j-ten Spaltenvektor von A^{-1}, betrachtet

$$\sum_{k=1}^{n} a_{ik}\, \tilde{a}_{kj} = \frac{\sum_{k=1}^{n} (-1)^{j-k}\, a_{ik}\, A_{jk}}{\det A}. \tag{499}$$

Die folgende Fallunterscheidung zeigt, dass die rechte Seite sich auf das Kroneckersymbol δ_{ij}, reduziert, das ij-te Matrixelement der Einheitsmatrix.

- $\mathbf{i = j}$: im Zähler von (499) steht $\det A$ in der Form als Entwicklung (466) nach der i-ten Zeile, so dass das Ergebnis erwartungsgemäß 1 ist für alle $i = 1, 2, ..., n$;
- $\mathbf{i \neq j}$: den Zähler identifiert man als Determinante einer fiktiven Matrix in der Form einer Entwicklung (466) nach der j-ten Zeile. Diese fiktive Matrix wird aus A gewonnen, indem die j-te Zeile durch die i-te Zeile ersetzt wird. Dabei ist zu beachten, dass die Größen A_{jk} im Zähler unberührt von dieser Ersetzung bleiben. Die Determinante der fiktiven Matrix verschwindet, da sie zwei gleiche Zeilen hat. Dies gilt für alle $i,\ j \in \{1, 2, ..., n\}$ mit $i \neq j$.

8.3.2 Eigenschaften invertierbarer Matrizen

- Invertierbare Matrizen (mit $\det A \neq 0$) heißen *regulär*, Matrizen mit Determinante $\det A = 0$ heißen *singulär*.
- Seien A und B zwei invertierbare Matrizen, so gilt

$$(AB)^{-1} = B^{-1} A^{-1}, \tag{500}$$

denn

$$(AB)^{-1}\, A\, B = B^{-1} A^{-1}\, A\, B \overset{A^{-1}A=\mathbf{1}}{=} B^{-1} B \overset{B^{-1}B=\mathbf{1}}{=} \mathbf{1}. \tag{501}$$

- Sei A eine invertierbare Matrix, dann gilt

$$\det A^{-1} = \frac{1}{\det A}. \tag{502}$$

Dies folgt aus (474), $\det(A\,B) = \det A \det B$, für $B = A^{-1}$ und aus $\det(\mathbf{1})$=1.

8.4 Lösung von linearen inhomogenen Gleichungssystemen

Gegeben sei das System von n linearen, inhomogenen Gleichungen

$$\begin{aligned} x_1' &= a_{11}x_1 + \ldots + a_{1n}x_n \\ &\vdots \\ x_n' &= a_{n1}x_1 + \ldots + a_{nn}x_n \end{aligned} \tag{503}$$

oder in Kurzform

$$\vec{x}\,' = A\,\vec{x}. \tag{504}$$

Gegeben sei der n-dimensionale Vektor $\vec{x'}$ und die $n \times n$ - Matrix A; gesucht ist der n-dimensionale Vektor $\vec{x}$ mit Komponenten x_1, ..., x_n, die Lösung des linearen inhomogenen Gleichungssystems (503). Diese Aufgabe kann elegant durch die Matrixinversion gelöst werden

$$\vec{x} = A^{-1}\,\vec{x}\,'. \tag{505}$$

Es gibt zwei Verfahren zur Lösung des Gleichungssystems (503).

8.4.1 Cramersche Regel zur Inversion

Cramersche Regel: Die Komponenten x_i des Lösungsvektors sind

$$x_i = \frac{\det A^{(i)}}{\det A}, \tag{506}$$

wobei die $n \times n$ - Matrix $A^{(i)}$ aus A durch Ersetzung der i-ten Spalte durch den Spaltenvektor

$$\vec{x}\,' = \begin{pmatrix} x_1' \\ \vdots \\ x_n' \end{pmatrix} \tag{507}$$

gewonnen wird. Diese Prozedur erfordert die Bestimmung von $n+1$ Determinanten von $n \times n$-Matrizen.

Zu zeigen ist entsprechend (506), dass $\det A^{(i)} = x_i \det A$ gilt. Der Beweis wird geführt für $i = 1$ und läuft für die Werte $i = 2$, ..., n analog.

$$\det A^{(1)} = \begin{vmatrix} x_1' & a_{12} & \ldots & a_{1n} \\ \vdots & \vdots & & \vdots \\ x_n' & a_{n2} & \ldots & a_{nn} \end{vmatrix} \tag{508}$$

$$= \begin{vmatrix} a_{11}x_1 + \ldots + a_{1n}x_n & a_{12} & \ldots & a_{1n} \\ \vdots & \vdots & & \vdots \\ a_{n1}x_1 + \ldots + a_{nn}x_n & a_{n2} & \ldots & a_{nn} \end{vmatrix} \tag{509}$$

$$= x_1 \begin{vmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \ldots & a_{nn} \end{vmatrix} \tag{510}$$

$$= x_1 \det A. \tag{511}$$

Hier wurde (503) eingesetzt, die Eigenschaft 6. von Kap. 8.2.2 benutzt, nach der die Subtraktion des x_j-fachen der j-ten Spalte von der ersten Spalte für $j = 2, \ldots, n$ den Wert der Determinante unverändert läßt, sowie (468) verwendet.

8.4.2 Gaußsches Eliminationsverfahren

Um das lineare inhomogene Gleichungssystem (503) bzw. (504), $\vec{x}\,' = A\,\vec{x}$, zu lösen, unterwirft man die beiden Seiten des Gleichungssystems (503) einer Abfolge von Manipulationen, die die Gleichung auf die Form

$$\vec{\tilde{x}}\,' = \tilde{A}\,\vec{x} \tag{512}$$

bringen, in der die Matrix $\tilde{A}$ Dreiecksform (478) gewinnt, deren Elemente unterhalb der Hauptdiagonalen verschwinden.

Dabei kann man sich weitgehend am Vorgehen in Kap. 8.2.3 orientieren. In einem ersten Schritt werden die Gleichungen wieder so umgeordnet und umbenannt, dass $a_{11} \neq 0$; in einem zweiten Schritt folgt für $i = 2, \ldots, n$

$$\text{die Subtraktion des } \frac{a_{i1}}{a_{11}} - \text{fachen der 1. Gleichung von der i-ten Gleichung} \tag{513}$$

jeweils für die linken und rechten Seiten separat. Dies führt auf ein Gleichungssystem, in dem ab der 2. Gleichung auf der rechten Seite die Variable x_1 eliminiert ist. Im nächsten

Schritt wird analog die Variable x_2 ab der 3. Gleichung eliminiert usw. Das Ergebnis ist schematisch

$$\begin{pmatrix} \tilde{x}\,'_1 \\ . \\ . \\ . \\ \tilde{x}\,'_n \end{pmatrix} = \begin{pmatrix} a_{11} & .. & .. & \ldots & .. \\ 0 & \tilde{a}_{22} & .. & \ldots & .. \\ 0 & 0 & \tilde{a}_{33} & \ldots & .. \\ \vdots & & & & \vdots \\ 0 & 0 & 0 & \ldots & \tilde{a}_{nn} \end{pmatrix} \begin{pmatrix} x_1 \\ . \\ . \\ . \\ x_n \end{pmatrix}. \tag{514}$$

Dieses Gleichungssystem lässt sich iterativ lösen, beginnend bei x_n:

$$\text{aus } \tilde{x}\,'_n = \tilde{a}_{nn}\,x_n \qquad \text{folgt} \qquad x_n = \frac{\tilde{x}\,'_n}{\tilde{a}_{nn}} \tag{515}$$

$$\text{aus } \tilde{x}\,'_{n-1} = \tilde{a}_{n-1,n-1}\,x_{n-1} + \tilde{a}_{n-1,n}\,x_n \quad \text{folgt mit (515)} \quad x_{n-1} = \frac{\tilde{x}\,'_{n-1} - \frac{\tilde{a}_{n-1,n}}{\tilde{a}_{nn}}\,\tilde{x}\,'_n}{\tilde{a}_{n-1,n-1}} \tag{516}$$

$$\vdots$$

Diese Methode wird auch zur numerischen Lösung von großen linearen inhomogenen Gleichungssystemen benutzt.

9 Spezielle reelle Matrizen: Räumliche Drehungen und Dyadisches Produkt

9.1 Allgemeine räumliche Drehungen

In Kap. 1.4.3 waren nur spezielle Drehungen um die z-, x-, y-Achsen, (42)-(50), vorgestellt worden. Hier soll nun die Parametrisierung einer beliebigen Drehung im dreidimensionalen Ortsraum hergeleitet werden. Dies musste zurückgestellt werden, da dazu der Umgang mit Matrizen erforderlich ist.

Das physikalische Interesse an dieser Frage ist evident. Sie ist relevant für die Formulierung der Unabhängigkeit der physikalischen Gesetzmäßigkeiten von Drehungen des Koordinatensystems und gegebenenfalls für die Wahl eines geeigneten Koordinatensystems zur Erleichterung des Rechenaufwands.

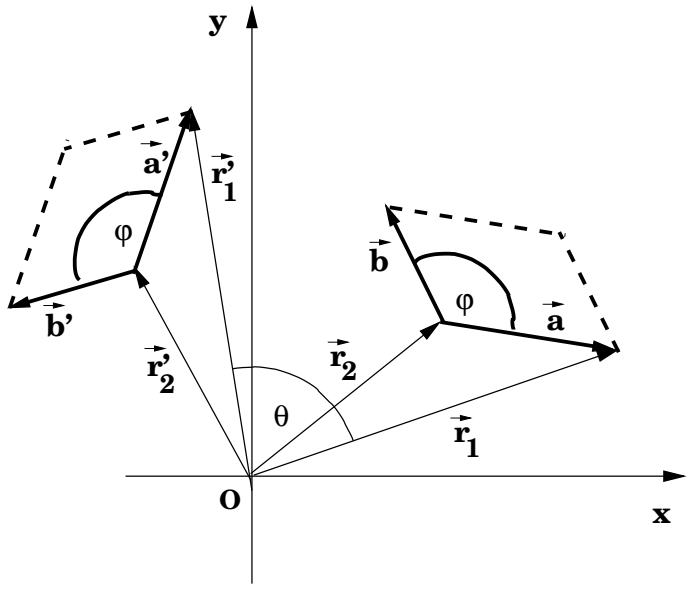


Abbildung 31:

Gegeben seien zwei beliebige reelle Vektoren im dreidimensionalen Ortsraum

$$\vec{a} = \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix} \text{ mit Betrag } |\vec{a}| = \sqrt{a_1^2 + a_2^2 + a_3^2} \text{ und } \vec{b} = \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix} \text{ mit Betrag } |\vec{b}| = \sqrt{b_1^2 + b_2^2 + b_3^2}, \tag{517}$$

die den Winkel ϕ einschließen.

Eine räumliche Drehung dieser zwei Vektoren $\vec{a}$ und $\vec{b}$ um den beliebigen Winkel θ ist in Abb. 31 dargestellt; aus pädadogischen Gründen wurde die Drehung in der x-y-Ebene gezeichnet. Es gilt offensichtlich

Unter einer beliebigen **räumlichen Drehung** bleiben **Längen** der Vektoren und **Winkel** zwischen Vektoren **erhalten**.

Aus der Längen- und Winkelerhaltung unter räumlichen Drehungen folgt für das Skalarprodukt zweier Vektoren $\vec{a}$ und $\vec{b}$

$$\vec{a}^{\,T}\vec{b} = |\vec{a}|\,|\vec{b}|\cos\phi = |\vec{a}\,'|\,|\vec{b}\,'|\cos\phi' = \vec{a}\,'^{\,T}\,\vec{b}\,', \text{d.h.} \tag{518}$$

unter einer beliebigen **räumlichen Drehung bleibt das Skalarprodukt** zweier beliebiger Vektoren $\vec{a}$ und $\vec{b}$ **erhalten**, $\vec{a}^{\,T}\vec{b} = \vec{a}\,'^{\,T}\,\vec{b}\,'$.

Macht man den folgenden Ansatz einer linearen Abbildung für eine beliebige solche Drehung

$$\vec{a}\,' = R\,\vec{a}, \quad \vec{b}\,' = R\,\vec{b} \;\text{ mit der Drehmatrix }\; R = \begin{pmatrix} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{23} \\ r_{31} & r_{32} & r_{33} \end{pmatrix} \tag{519}$$

mit reellen Elementen r_{ij} für $i, j =$ 1, 2, 3, dann folgt mit der Erhaltung des Skalarproduktes

$$\vec{a}\,'^{T}\vec{b}\,' = \vec{a}^{\,T} R^T R\,\vec{b} = \vec{a}^{\,T}\vec{b}. \tag{520}$$

Da dies für alle Vektorpaare $\vec{a}$, $\vec{b}$ gilt, folgt

$$R^T R = R\,R^T = \mathbf{1} \;\text{ mit }\; R^{-1} = R^T; \tag{521}$$

Drehmatrizen sind orthogonale Matrizen, die durch die Eigenschaft (521) definiert sind.

Die Drehmatrix R hat in ihrem allgemeinen Ansatz (519) neun reelle Parameter, die Matrixelemente r_{ij} für $i, j \in \{1, 2, 3\}$. Die Matrixgleichung (521) entspricht sechs Nebenbedingungen an diese Parameter

$$\sum_{i=1}^{3} r_{ji}\, r_{ij} \;=\; 1, \quad \text{für} \;\; j = 1, 2, 3, \tag{522}$$

$$\sum_{i=1}^{3} r_{ji}\, r_{ik} \;=\; 0, \quad \text{für} \;\; j < k; \tag{523}$$

die entsprechenden Bedingungen für $j > k$ stellen sich als identisch mit (523) heraus. Insgesamt werden also orthogonale Drehmatrizen durch drei frei wählbare Parameter beschrieben.

Hier kann man auf die speziellen Drehungen (42)-(50), die Drehungen um die z-Achse, x-Achse und y-Achse beschreiben, zurückgreifen. Es soll eine Drehung um den Winkel ϕ_x um die x-Achse mit R_x, eine Drehung um den Winkel ϕ_y um die y-Achse mit R_y und eine Drehung um den Winkel ϕ_z um die z-Achse mit R_z bezeichnet werden. Das Produkt $R = R_x R_y R_z$ ist wegen $R^T R = (R_x R_y R_z)^T\, R_x R_y R_z = R_z^T R_y^T R_x^T R_x R_y R_z = R_z^T R_y^T R_y R_z = R_z^T R_z = \mathbf{1}$ eine orthogonale Matrix, die durch drei frei wählbare Parameter, die Winkel ϕ_x, ϕ_y und ϕ_z, beschrieben wird.

Eine **beliebige räumliche Drehung** kann als **Hintereinanderausführung** von drei **speziellen Drehungen** R_x, R_y, R_z um die x- bzw. y- bzw. z-Achse beschrieben werden. Die Drehmatrix wird durch die drei zugehörigen Drehwinkel ϕ_x, ϕ_y und ϕ_z parametrisiert.

Der Beweis wird hier nicht durchgeführt.

9.2 Dyadisches Produkt

9.2.1 Definition

Das dyadische Produkt soll für einen reellen n-dimensionalen Vektorraum formuliert werden. Seien $\vec{a}$, $\vec{b}$, $\vec{x}$ und $\vec{x}\,'$ beliebige n-dimensionale Vektoren. Es werde die spezielle lineare Abbildung

$$\vec{x}\,' = \vec{b}\left(\vec{a}^{\,T}\,\vec{x}\right) \tag{524}$$

betrachtet, in der das Skalarprodukt $\vec{a}^{\,T}\,\vec{x}$ als Produkt aus einem Zeilenvektor $\vec{a}^{\,T}$ mit einem Spaltenvektor $\vec{x}$

$$\vec{a}^{\,T}\,\vec{x} = \begin{pmatrix} a_1 \,...\, a_n \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \sum_{i=1}^{n} a_i\, x_i. \tag{525}$$

auftritt. Der Spaltenvektor $\vec{x}\,'$ wird

$$\vec{x}\,' = \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix} \left(\begin{pmatrix} a_1 \,...\, a_n \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \right) = \begin{pmatrix} b_1 \sum_{i=1}^{n} a_i\, x_i \\ \vdots \\ b_n \sum_{i=1}^{n} a_i\, x_i \end{pmatrix}. \tag{526}$$

In einem nächsten Schritt soll die lineare Abbildung (526) umgeschrieben werden in die Form

$$\vec{x}\,' = A\,\vec{x} = \left(\begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix} \begin{pmatrix} a_1 \,...\, a_n \end{pmatrix} \right) \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}. \tag{527}$$

Definition: Die Matrix A ist als Produkt aus dem Spaltenvektor $\vec{b}$ mit dem Zeilenvektor $\vec{a}^{\,T}$ definiert, dem sogenannten **dyadischen Produkt**,

$$A = \left(\begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix} \begin{pmatrix} a_1\, a_2 \,...\, a_n \end{pmatrix} \right) = \vec{b}\,\vec{a}^{\,T} = \begin{pmatrix} b_1a_1 & b_1a_2 & \dots & b_1a_n \\ b_2a_1 & b_2a_2 & \dots & b_2a_n \\ \vdots & \vdots & & \vdots \\ b_na_1 & b_na_2 & \dots & b_na_n \end{pmatrix}. \tag{528}$$

Das Produkt (528) wurde offensichtlich so definiert, dass das Ergebnis von (527) mit (526) übereinstimmt.

9.2.2 Physikalische Motivation: die Projektionsmatrix

In der Physik und insbesondere in der theoretischen Quantenmechanik spielen Projektionsmatrizen auf bestimmte Richtungen entweder im Ortsraum oder in einem abstrakten n-dimensionalen Vektorraum eine Rolle.

Als Beispiel sei die Projektion auf die x_3-Achse im Ortsraum gewählt. Der Einheitsvektor in x_3-Richtung ist $\vec{e}_3$. Die Projektion eines beliebigen Ortsvektors $\vec{r}$ auf die x_3-Achse läßt sich einerseits schreiben als

$$\vec{e}_3 \left(\vec{e}_3{}^T \vec{r} \right) = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \left(\begin{pmatrix} 0\,0\,1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \right) = \begin{pmatrix} 0 \\ 0 \\ x_3 \end{pmatrix}. \tag{529}$$

und andererseits als

$$\left(\vec{e}_3 \, \vec{e}_3{}^T \right) \vec{r} = \left(\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \begin{pmatrix} 0\,0\,1 \end{pmatrix} \right) \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = P\,\vec{x} = \begin{pmatrix} 0 \\ 0 \\ x_3 \end{pmatrix} \tag{530}$$

mit

$$P = \vec{e}_3 \, \vec{e}_3{}^T = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \begin{pmatrix} 0\,0\,1 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}. \tag{531}$$

$P = \vec{e}_3 \, \vec{e}_3{}^T$ ist also eine Matrix, die jeden Ortsvektor $\vec{r}$ auf die Richtung des Vektors $\vec{e}_3$ projiziert. Dies läßt sich verallgemeinern.

Die Projektionsmatrix auf einen Einheitsvektor $\vec{n}$ mit $\vec{n}^T \vec{n} = 1$ in beliebiger Richtung im n-dimensionalen Vektorraum ist als dyadisches Produkt darstellbar

$$P = \vec{n} \, \vec{n}^T \tag{532}$$

und erfüllt die definierenden Eigenschaften einer Projektionsmatrix

$$P = P^T, \qquad P^2 = P \quad \text{und} \quad \det P = 0, \tag{533}$$

da $P^T = (\vec{n} \, \vec{n}^T)^T = (\vec{n}^T)^T \, \vec{n}^T = \vec{n} \, \vec{n}^T = P$ und $P^2 = (\vec{n} \, \vec{n}^T)(\vec{n} \, \vec{n}^T) = \vec{n} \, (\vec{n}^T \, \vec{n}) \, \vec{n}^T = \vec{n} \, \vec{n}^T = P$ gilt.

10 Komplexe Matrizen

10.1 Physikalische Motivation

Die Beschreibung der Quantenmechanik als sogenannte Matrizenmechanik geht auf Heisenberg zurück. Im Rahmen der Matrizenmechanik hat man es mit einem speziellen *komplexen Vektorraum* zu tun, dem sogenannten *Hilbertraum*, in dem Vektoren und Matrizen komplexe Elemente haben. *Physikalische Messgrößen* wie Energie, Impuls, Drehimpuls, Ort, ... werden durch sogenannte *hermitesche* komplexe Matrizen beschrieben. Die sogenannten Eigenwerte und Eigenvektoren von hermiteschen Matrizen, die im Folgenden behandelt werden sollen, spielen dabei eine hervorragende Rolle. Die in der Quantenmechanik auftretenden Matrizen sind häufig unendlich-dimensional; dieser Aspekt übersteigt jedoch den Rahmen dieser Vorlesung. Die hier vorgestellte Behandlung endlich-dimensionaler komplexer Matrizen bietet jedoch eine unerläßliche Grundlage.

10.2 Komplexe, adjungierte, hermitesche und reell-symmetrische Matrizen

Der bisher betrachtete reelle n-dimensionale Vektorraum $\mathbb{R}^n$ soll auf einen komplexen n-dimensionalen Vektorraum – eingeschränkt auf einen Hilbertraum – erweitert werden. Auf eine saubere Definition des Hilbertraumes muss hier verzichtet werden, da dies über den Rahmen der Vorlesung weit hinaus gehen würde. Es werden nur die Eigenschaften eingeführt, die im Folgenden wichtig sind.

Vektoren $\vec{x}$ sollen komplexe Komponenten x_i, $i = 1, 2, ..., n$, haben

$$\vec{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} \mathrm{Re}\, x_1 + i\, \mathrm{Im}\, x_1 \\ \vdots \\ \mathrm{Re}\, x_n + i\, \mathrm{Im}\, x_n \end{pmatrix}. \tag{534}$$

Die Rolle des transponierten Zeilenvektors $\vec{x}^{\,T}$ des reellen Vektorraums übernimmt der komplexe Zeilenvektor $\vec{x}^{\,\dagger}$, der aus $\vec{x}$ durch Transposition und komplexe Konjugation aller Komponenten entsteht und mit dem Symbol † gekennzeichnet wird, der sogenannte *hermitesch konjugierte* oder *adjungierte* Vektor,

$$\vec{x}^{\,\dagger} = \vec{x}^{\,*T} = \vec{x}^{\,T*} = \Big(\mathrm{Re}\, x_1 - i\, \mathrm{Im}\, x_1,\ ...,\ \mathrm{Re}\, x_n - i\, \mathrm{Im}\, x_n\Big) \quad \text{mit} \quad (\vec{x}^{\,\dagger})^{\,\dagger} = \vec{x}. \tag{535}$$

Ist $\vec{x}$ reell, reduziert sich $\vec{x}^{\,\dagger}$ auf $\vec{x}^{\,T}$.

Das Skalarprodukt eines komplexen n-dimensionalen adjungierten Vektors $\vec{x}^{\,\dagger}$ mit einem komplexen n-dimensionalen Vektor $\vec{y}$ ist definiert als

$$\vec{x}^{\,\dagger}\, \vec{y} = x_1^*\, y_1 + x_2^*\, y_2 + ... + x_n^*\, y_n. \tag{536}$$

Dies ist ein komplexer Skalar, d.h. eine komplexe Zahl, und es gilt

$$\vec{x}^{\,\dagger}\, \vec{y} = \Big(\vec{y}^{\,\dagger}\, \vec{x}\Big)^*. \tag{537}$$

Diese Definition des Skalarproduktes impliziert die besondere Eigenschaft, dass das Skalarprodukt eines Vektors mit sich selbst

$$\vec{x}^{\,\dagger}\, \vec{x} = |x_1|^2 + |x_2|^2 + ... + |x_n|^2 = \sum_{i=1}^{n} |x_1|^2 \tag{538}$$

reell und ≥ 0 ist. Die *Norm* des n-dimensionalen komplexen Vektorraumes, $\sqrt{\vec{x}^{\,\dagger}\, \vec{x}}$, ist wieder *positiv semi-definit*. Diese Eigenschaft des Hilbertraumes ist essentiell für die Beschreibung der Quantenmechanik, da dem Skalarprodukt eines Vektors mit sich selbst die physikalische Bedeutung einer Wahrscheinlichkeit zugeordnet wird.

Orthonormalität eines Satzes von n komplexen Basisvektoren $\vec{e}^{\,(i)}$ für $i = 1, 2, ..., n$ ist entsprechend folgendermaßen definiert

$$\vec{e}^{\,(i)\,\dagger}\, \vec{e}^{\,(j)} = \delta_{ij} \quad \text{für} \quad i,\, j \in \{1, ..., n\}. \tag{539}$$

(Hier zählen die Indices i und j die Basisvektoren durch und nicht Komponenten von Vektoren.) Die Basisvektoren spannen den komplexen normierten n-dimensionalen Vektorraum auf.

Zu betrachtende komplexe $n \times n$-Matrizen haben komplexe Matrixelemente

$$A = \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \dots & a_{nn} \end{pmatrix} = \Big(a_{ij}\Big) \tag{540}$$

mit

$$a_{ij} = \mathrm{Re}\, a_{ij} + i\, \mathrm{Im}\, a_{ij} \quad \text{für } i,\, j \in \{1, \dots n\}. \tag{541}$$

Definition: Die zur Matrix A **adjungierte Matrix** ist

$$A^\dagger = A^{*T} = A^{T*} \tag{542}$$

$$= \begin{pmatrix} a^*_{11} & \dots & a^*_{n1} \\ \vdots & & \vdots \\ a^*_{1n} & \dots & a^*_{nn} \end{pmatrix} = \Big(a^*_{ji}\Big) \quad \text{mit} \tag{543}$$

$$\text{mit} \quad a^*_{ji} = \mathrm{Re}\, a_{ji} - i\, \mathrm{Im}\, a_{ji} \quad \text{für } i,\, j \in \{1, \dots n\}. \tag{544}$$

Dabei bedeutet A^* komplexe Konjugation für jedes Element von A und es wird daran erinnert, dass A^T die Transposition von A bezeichnet, bei der jede Zeile mit der entsprechenden Spalte vertauscht wird. Vertauschung von Zeilen mit entsprechenden Spalten resultiert im Austausch der Indices der Matrixelemente, $a_{ij} \to a_{ji}$.

Definition: Für eine **hermitesche Matrix** A gilt

$$A = A^\dagger, \quad \text{d.h. } a_{ij} = a^*_{ji} \quad \text{für } i,\, j \in \{1, \dots n\}. \tag{545}$$

Dies impliziert, dass die Diagonalelemente von hermiteschen Matrizen reell sind.

Für *reelle* Matrizen A gilt

$$A^\dagger = A^T \tag{546}$$

und hermitesche Matrizen reduzieren sich auf *symmetrische* Matrizen, die *reell-symmetrisch* genannt werden.

Definition: Für eine **reell-symmetrische Matrix** A, die definitionsgemäß reelle Matrixelemente hat, gilt

$$A = A^T \quad \text{d.h. } a_{ij} = a_{ji} \quad \text{für } i,\, j \in \{1, \dots n\}. \tag{547}$$

Die Multiplikation $A\,\vec{x}$ für alle Vektoren $\vec{x}$ bildet den komplexen Vektorraum auf sich selbst ab.

10.3 Eigenwertgleichung, Eigenwerte, Eigenvektoren und das Eigenwertproblem

Betrachtet werden lineare Abbildungen im n-dimensionalen Raum

$$A\,\vec{x} = \vec{x}\,', \tag{548}$$

wobei A eine komplexe $n \times n$ - Matrix ist und $\vec{x}$, $\vec{x}\,'$ komplexe n-dimensionale Spaltenvektoren sind. Solche Abbildungen sind in der Regel (selbst für reelle Matrizen) geometrisch schwer darstellbar, da ein Vektor $\vec{x}$ auf einen anderen Vektor $\vec{x}\,'$ abgebildet wird, der im Allgemeinen eine andere Richtung und einen anderen Betrag hat.

Es gibt besonders ausgezeichnete Richtungen $\vec{x}$ im n-dimensionalen Raum, in denen sich die Wirkung von A auf $\vec{x}$ auf die *Multiplikation von $\vec{x}$ mit einem Skalar λ* reduziert, wobei dieser Skalar in der Regel komplex ist. Ist der Skalar λ reell, besteht die Wirkung von A in diese ausgezeichnete Richtung sogar nur in einer Längenänderung. Solche ausgezeichnete Richtungen sollen Gegenstand der nächsten Unterkapitel sein.

Definition: Die **Eigenwertgleichung** für eine vorgegebene Matrix A lautet

$$A\,\vec{x} = \lambda\,\vec{x}. \tag{549}$$

Der Skalar λ heißt **Eigenwert von** A, der Vektor $\vec{x}$ **der zum Eigenwert λ gehörige Eigenvektor**. Die Aufgabe, alle Eigenwerte und zugehörigen Eigenvektoren einer vorgegebenen Matrix A zu bestimmen, nennt man das **Eigenwertproblem**.

Die Lösung von Eigenwertproblemen spielt eine hervorragende Rolle sowohl in der Mathematik als auch in der Physik, vor allem in der Quantenmechanik.

Es sei noch einmal betont, dass in der Eigenwertgleichung ein Vektor auf der linken Seite mit einer Matrix und auf der rechten Seite mit einer Zahl multipliziert wird, ausgeschrieben

$$\begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \dots & a_{nn} \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \lambda \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \quad \text{bzw.} \tag{550}$$

$$\sum_{i=1}^{n} a_{1i}\,x_i = \lambda\,x_1, \tag{551}$$

$$\vdots$$

$$\sum_{i=1}^{n} a_{ni}\,x_i = \lambda\,x_n, \tag{552}$$

wobei der Skalar λ in allen n Gleichungen (551)-(552) *denselben Wert* hat. Das ist eine höchst nichttrivialer Satz von Beziehungen.

Die Eigenwertgleichung (549) kann folgendermaßen in ein homogenes Gleichungssystem umgeschrieben werden, wobei $\mathbf{1}$ die $n \times n$ Einheitsmatrix (393) ist

$$\Big(A - \lambda\,\mathbf{1}\Big)\,\vec{x} = \vec{0}. \tag{553}$$

Falls $\det\Big(A - \lambda\,\mathbf{1}\Big) \neq 0$ ist, hat diese Gleichung nur die triviale Lösung $\vec{x} = \vec{0}$, die offensichtlich nicht von Interesse ist. Es gibt nur nichttriviale Lösungen $\vec{x} \neq \vec{0}$, falls

Definition: die **charakteristische Gleichung**

$$\det\Big(A-\lambda\,\mathbf{1}\Big)=\det\begin{pmatrix} a_{11}-\lambda & a_{12} & a_{13} & \dots & a_{1n} \\ a_{21} & a_{22}-\lambda & a_{23} & \dots & a_{2n} \\ a_{31} & a_{32} & a_{33}-\lambda & \dots & a_{3n} \\ \vdots & \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & a_{n3} & \dots & a_{nn}-\lambda \end{pmatrix} = p_n(\lambda) \;=\; 0xxxx \qquad (554)$$

erfüllt ist. Die Determinante ist ein Polynom,

Definition: das **charakteristische Polynom** n-ten Grades in der Variablen λ, dessen erste beiden Terme sowie der letzte Term einfach anzugeben sind

$$p_n(\lambda) = (-\lambda)^n + (-\lambda)^{n-1}\,\mathrm{Sp}\,A + ... + \det A. \qquad (555)$$

Der Koeffizient $\mathrm{Sp}\,A$ von $(-\lambda)^{n-1}$ ist die sogenannte *Spur der Matrix A*, die Summe der Diagonalelemente,

$$\mathrm{Sp}\,A = a_{11} + a_{22} + ... + a_{nn}. \qquad (556)$$

Definition: Die n komplexen Lösungen λ_i für $i = 1, ..., n$ der charakteristischen Gleichung werden die **Eigenwerte** der Matrix A genannt.

Bisher wurden allgemein komplexe Matrizen angenommen. Eine wichtige Beobachtung ist, dass selbst *im Fall reeller Matrizen Eigenwerte und Eigenvektoren komplex sein können*. Deshalb kann man eine Diskussion des Eigenwertproblems nicht auf reelle Vektorräume beschränken.

Dies kann man an einem einfachen Beispiel sehen. Gegeben sei die reelle 2×2 - Matrix

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}. \qquad (557)$$

Die charakteristische Gleichung

$$\det\Big(A-\lambda\,\mathbf{1}\Big) = \begin{vmatrix} a_{11}-\lambda & a_{12} \\ a_{21} & a_{22}-\lambda \end{vmatrix} = (a_{11}-\lambda)(a_{22}-\lambda) - a_{12}a_{21} \qquad (558)$$

$$= \lambda^2 - (a_{11}+a_{22})\lambda + a_{11}a_{22} - a_{12}a_{21} = 0 \qquad (559)$$

führt auf die beiden Eigenwerte

$$\lambda_{1,2} = \frac{a_{11}+a_{22}}{2} \pm \sqrt{\frac{(a_{11}-a_{22})^2}{4} + a_{12}a_{21}}. \qquad (560)$$

Offensichtlich wird der Radikant der Wurzel für negative Werte von $a_{12}a_{21}$ und $|a_{12}a_{21}| > (a_{11} - a_{22})^2/4$ negativ. Obwohl die Matrix A reell ist, können also ihre Eigenwerte in Form von konjugiert komplexen Zahlenpaaren komplex werden; dementsprechend werden auch die zugehörigen Eigenvektoren komplex. Mit einer Beobachtung soll bereits etwas vorgegriffen werden. Die Eigenwerte $\lambda_{1,2}$ werden reell für $a_{12} = a_{21}$, d.h. wenn die reelle Beispielmatrix A symmetrisch ist.

10.4 Eigenwerte komplexer hermitescher und reell-symmetrischer Matrizen

Für komplexe hermitesche und damit auch für reell-symmetrische Matrizen läßt sich beweisen, dass ihre Eigenwerte reell sind. Ausgangspunkt ist eine komplexe hermitesche Matrix

$$A = A^\dagger = \begin{pmatrix} a_{11} & \operatorname{Re} a_{12} + i \operatorname{Im} a_{12} & \dots & \operatorname{Re} a_{1n} + i \operatorname{Im} a_{1n} \\ \operatorname{Re} a_{12} - i \operatorname{Im} a_{12} & a_{22} & \dots & \operatorname{Re} a_{2n} + i \operatorname{Im} a_{2n} \\ \vdots & \vdots & & \vdots \\ \operatorname{Re} a_{1n} - i \operatorname{Im} a_{1n} & \operatorname{Re} a_{2n} - i \operatorname{Im} a_{2n} & \dots & a_{nn} \end{pmatrix} \tag{561}$$

und komplexe Vektoren $\vec{x}$, wie in (534) definiert.

Die Eigenwertgleichung (549) für A wird auf beiden Seiten von links mit $\vec{x}^\dagger$ multipliziert

$$\vec{x}^\dagger A \vec{x} = \lambda \vec{x}^\dagger \vec{x}. \tag{562}$$

Beide Seiten der Gleichung (562) werden dabei auf eine komplexe Zahl reduziert. Bildet man den konjugiert komplexen Ausdruck beider Seiten der Gleichung, benutzt man weiterhin deren Invarianz unter Transposition und die Eigenschaft, dass die Transposition eines Produktes gleich dem Produkt aus den transponierten Multiplikanden in umgekehrter Reihenfolge ist, erhält man für die linke Seite

$$\left(\vec{x}^\dagger A \vec{x}\right)^* = \left(\vec{x}^\dagger A \vec{x}\right)^{T*} = \left(\vec{x}^\dagger A \vec{x}\right)^\dagger = \vec{x}^\dagger A^\dagger (\vec{x}^\dagger)^\dagger = \vec{x}^\dagger A^\dagger \vec{x} \tag{563}$$

und für die rechte Seite

$$\lambda^* \left(\vec{x}^\dagger \vec{x}\right)^* = \lambda^* \left(\vec{x}^\dagger \vec{x}\right)^{T*} = \lambda^* \left(\vec{x}^\dagger \vec{x}\right)^\dagger = \lambda^* \left(\vec{x}^\dagger (\vec{x}^\dagger)^\dagger\right) = \lambda^* \left(\vec{x}^\dagger \vec{x}\right). \tag{564}$$

Hermitizität der Matrix A, $A = A^\dagger$, macht den Ausdruck (563) gleich der linken Seite der Gleichung (562), damit muss der Ausdruck (564) gleich der rechten Seite von (562) sein, d.h. es gilt

$$\lambda = \lambda^*, \tag{565}$$

die **Eigenwerte einer komplexen hermiteschen Matrix sind reell**.

Da reell-symmetrische Matrizen Spezialfälle von komplexen hermiteschen Matrizen sind, gilt insbesondere

Die **Eigenwerte einer reell-symmetrischen Matrix sind reell**.

Wie bereits mehrfach angedeutet, werden in der Quantenmechanik den physikalischen Messgrößen hermitesche Matrizen zugeordnet. Deren Eigenwerte stellen die möglichen Messergebnisse für die Messgrößen dar. Da Messgrößen stets reell sind, konnte eine solche Formulierung nur mit hermiteschen Matrizen gelingen.

10.5 Eigenvektoren komplexer hermitescher und reell-symmetrischer Matrizen

Hier soll das Eigenwertproblem von hermiteschen $n \times n$ - Matrizen

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{12}^* & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{1n}^* & a_{2n}^* & \dots & a_{nn} \end{pmatrix} \quad \text{mit} \quad A = A^\dagger \tag{566}$$

behandelt werden. Eine hermitesche Matrix hat n reelle Eigenwerte λ_1, ..., λ_n als Lösungen der charakteristischen Gleichung (554). Die Eigenwerte sind nicht notwendig alle verschieden. Falls zwei oder mehr Eigenwerte zusammenfallen, spricht man von *Entartung*.

Es gibt n Eigenvektoren, die im Allgemeinen komplex sind. Der jeweils zum Eigenwert λ_i gehörige Eigenvektor soll mit $\vec{x}^{\,(i)}$ bezeichnet werden. Die n Eigenwertgleichungen lauten

$$A\,\vec{x}^{\,(i)} = \lambda_i\,\vec{x}^{\,(i)} \quad \text{für} \quad i = 1, ..., n. \tag{567}$$

(Der Index (i) in $\vec{x}^{\,(i)}$ zählt die Eigenvektoren durch und nicht die Komponenten eines Vektors.)

Es gilt nun die folgende elegante Aussage, deren Beweis skizziert werden soll.

Die **Eigenvektoren** $\vec{x}^{\,(i)}$ einer hermiteschen Matrix bilden ein **Orthonormalsystem**, d.h. es gilt

$$\vec{x}^{\,(i)\dagger}\,\vec{x}^{\,(j)} = \delta_{ij} \quad \text{für} \quad i,\, j \in \{1, ..., n\}, \tag{568}$$

d.h. sie sind alle auf 1 normiert und stehen "paarweise senkrecht" aufeinander.

In der Eigenwertgleichung (567) sind die Eigenvektoren offensichtlich bis auf eine multiplikative Konstante bestimmt. Da $\vec{x}^{\,(i)\dagger}\,\vec{x}^{\,(i)} = \sqrt{x_1^{(i)2} + ... + x_n^{(i)2}}$ für $i = 1, ..., n$ reell und ≥ 0 ist, kann man sie o.B.d.A. auf 1 normieren

$$\vec{x}^{\,(i)\dagger}\,\vec{x}^{\,(i)} = 1 \quad \text{für} \quad i = 1, ..., n. \tag{569}$$

Der Beweis für die "Orthogonalität" der Eigenvektoren läuft folgendermaßen. Die Eigenwertgleichung (567) wird von links mit $\vec{x}^{\,(j)\dagger}$ multipliziert

$$\vec{x}^{\,(j)\dagger} A\,\vec{x}^{\,(i)} = \lambda_i\,\vec{x}^{\,(j)\dagger}\,\vec{x}^{\,(i)}. \tag{570}$$

Von beiden Seiten der Gleichung werden die entsprechenden Seiten einer Gleichung subtrahiert, in der die Indices i und j vertauscht wurden und die komplexe Konjugation vorgenommen wurde,

$$\vec{x}^{\,(j)\dagger} A\,\vec{x}^{\,(i)} - \left(\vec{x}^{\,(i)\dagger} A\,\vec{x}^{\,(j)}\right)^* = \lambda_i\,\vec{x}^{\,(j)\dagger}\,\vec{x}^{\,(i)} - \lambda_j^*\left(\vec{x}^{\,(i)\dagger}\,\vec{x}^{\,(j)}\right)^*. \tag{571}$$

Beide Seiten sind Skalare, also invariant unter Transposition. Die linke Seite von (571) verschwindet auf Grund von $A = A^\dagger$

$$\begin{aligned} \vec{x}^{\,(j)\dagger} A\,\vec{x}^{\,(i)} - \left(\vec{x}^{\,(i)\dagger} A\,\vec{x}^{\,(j)}\right)^* &= \vec{x}^{\,(j)\dagger} A\,\vec{x}^{\,(i)} - \left(\vec{x}^{\,(i)\dagger} A\,\vec{x}^{\,(j)}\right)^{*T} && (572) \\ &= \vec{x}^{\,(j)\dagger} A\,\vec{x}^{\,(i)} - \left(\vec{x}^{\,(i)\dagger} A\,\vec{x}^{\,(j)}\right)^\dagger && (573) \\ &= \vec{x}^{\,(j)\dagger} A\,\vec{x}^{\,(i)} - \vec{x}^{\,(j)\dagger} A^\dagger\,(\vec{x}^{\,(j)\dagger})^\dagger = 0. && (574) \end{aligned}$$

Die rechte Seite von (571) muss demnach ebenfalls verschwinden; nach einer Umformung erhält man mit $\lambda_j = \lambda_j^*$

$$0 = \lambda_i \, \vec{x}^{\,(j)\dagger} \, \vec{x}^{\,(i)} - \lambda_j^* \left(\vec{x}^{\,(i)\dagger} \, \vec{x}^{\,(j)} \right)^* = \lambda_i \, \vec{x}^{\,(j)\dagger} \, \vec{x}^{\,(i)} - \lambda_j \left(\vec{x}^{\,(i)\dagger} \, \vec{x}^{\,(j)} \right)^{*T} \tag{575}$$

$$= \lambda_i \, \vec{x}^{\,(j)\dagger} \, \vec{x}^{\,(i)} - \lambda_j \left(\vec{x}^{\,(i)\dagger} \, \vec{x}^{\,(j)} \right)^{\dagger} = \lambda_i \, \vec{x}^{\,(j)\dagger} \, \vec{x}^{\,(i)} - \lambda_j \, \vec{x}^{\,(j)\dagger} \, (\vec{x}^{\,(i)\dagger})^{\dagger} \tag{576}$$

$$= (\lambda_i - \lambda_j) \, \vec{x}^{\,(j)\dagger} \, \vec{x}^{\,(i)}. \tag{577}$$

Zur weiteren Auswertung muss man die folgende Fallunterscheidung machen.

- Sind die Eigenwerte λ_i nicht entartet, d.h. $\lambda_i - \lambda_j \neq 0$ für alle Paare von i, j mit $i \neq j$, dann folgt aus (575)-(577) die paarweise Orthogonalität aller Eigenvektoren

$$\vec{x}^{\,(j)\dagger} \, \vec{x}^{\,(i)} = 0 \quad \text{für} \quad i \neq j. \tag{578}$$

- Liegt Entartung vor, $\lambda_i - \lambda_j = 0$ für Paare von i und j mit $i \neq j$, dann kann man aus (577) nicht mehr auf die Orthogonalität der zugehörigen Eigenvektoren schließen. In diesem Fall hat man jedoch genügend Freiheit in der Wahl der zugehörigen Eigenvektoren, dass Orthogonalität (578) erfüllt werden kann. Zum Beispiel für den Fall, dass zwei Eigenwerte gleich werden – o.B.d.A. sei $\lambda_1 = \lambda_2$ angenommen – sind mit $\vec{x}^{\,(1)}$ und $\vec{x}^{\,(2)}$ auch alle Linearkombinationen

$$\alpha \, \vec{x}^{\,(1)} + \beta \, \vec{x}^{\,(2)}, \qquad \alpha, \ \beta = \text{beliebige komplexe Konstanten}, \tag{579}$$

  zulässige Eigenvektoren. In der durch (579) definierten "Ebene" können dann zwei orthogonale Eigenvektoren frei gewählt werden, z.B.

$$\vec{x}^{\,(1)} \quad \text{und} \quad \vec{x}^{\,(2)} - \left(\vec{x}^{\,(1)\dagger} \, \vec{x}^{\,(2)} \right) \vec{x}^{\,(1)}, \tag{580}$$

  und anschließend normiert werden.

Damit ist der Beweis von (568) skizziert.

Das Orthonormalsystem von n Eigenvektoren einer hermiteschen Matrix bildet eine Orthonormalbasis von linear unabhängigen Vektoren des betrachteten komplexen normierten Vektorraumes, d.h. alle Vektoren $\vec{x}$ des Vektorraumes können als Linearkombinationen der Eigenvektoren $\vec{x}^{\,(i)}$ einer beliebigen hermiteschen Matrix geschrieben werden

$$\vec{x} = \sum_{i=1}^{n} \xi_i \, \vec{x}^{\,(i)}. \tag{581}$$

Durch Multiplikation der Gleichung von links mit $\vec{x}^{\,(j)\dagger}$, unter Ausnutzung der Orthonormalität (568) der Eigenvektoren $\vec{x}^{\,(i)}$ und nach Vertauschung der Indices i und j lassen sich die komplexen Entwicklungskoeffizienten ξ_i bestimmen zu

$$\xi_i = \vec{x}^{\,(i)\dagger} \, \vec{x}. \tag{582}$$

Da reell-symmetrische Matrizen spezielle hermitesche Matrizen sind, gelten für sie alle Ergebnisse dieses Kapitels.

Beispiel

Gegeben sei als spezifische hermitesche Matrix die reell-symmetrische Matrix

$$\begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 2 \end{pmatrix}. \tag{583}$$

Die charakteristische Gleichung ist

$$\det\Big(A-\lambda\,\mathbf{1}\Big) = \det\begin{pmatrix} -\lambda & 1 & 0 \\ 1 & -\lambda & 0 \\ 0 & 0 & 2-\lambda \end{pmatrix} = (2-\lambda)\,(\lambda^2-1) = (2-\lambda)\,(\lambda-1)\,(\lambda+1) = 0. \tag{584}$$

Die resultierenden Eigenwerte sind erwartungsgemäß reell

$$\lambda_1 = 2, \quad \lambda_2 = 1, \quad \lambda_3 = -1. \tag{585}$$

Die zugehörigen Eigenvektoren werden bestimmt

- zu $\lambda_1 = 2$ aus

$$\Big(A-\lambda_1\,\mathbf{1}\Big)\,\vec{x}^{\,(1)} = \begin{pmatrix} -2 & 1 & 0 \\ 1 & -2 & 0 \\ 0 & 0 & 0 \end{pmatrix}\begin{pmatrix} x_1^{(1)} \\ x_2^{(1)} \\ x_3^{(1)} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}, \tag{586}$$

- zu $\lambda_2 = 1$ aus

$$\Big(A-\lambda_2\,\mathbf{1}\Big)\,\vec{x}^{\,(2)} = \begin{pmatrix} -1 & 1 & 0 \\ 1 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix}\begin{pmatrix} x_1^{(2)} \\ x_2^{(2)} \\ x_3^{(2)} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}, \tag{587}$$

- zu $\lambda_3 = -1$ aus

$$\Big(A-\lambda_3\,\mathbf{1}\Big)\,\vec{x}^{\,(3)} = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 3 \end{pmatrix}\begin{pmatrix} x_1^{(3)} \\ x_2^{(3)} \\ x_3^{(3)} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}. \tag{588}$$

Das Ergebnis ist mit der Normierungsbedingung $\vec{x}^{\,(i)\dagger}\,\vec{x}^{\,(i)} = 1$ für $i \in \{1,2,3\}$

$$\vec{x}^{\,(1)} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}, \quad \vec{x}^{\,(2)} = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \quad \vec{x}^{\,(3)} = \frac{1}{\sqrt{2}}\begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix}. \tag{589}$$

Man kann sich leicht von der Orthonormalität der Eigenvektoren (589) überzeugen.

10.6 Koordinatentransformationen, Basiswechsel und Hauptachsentransformation

10.6.1 Basiswechsel

Ausgangspunkt ist die lineare Abbildung (548) im komplexen n-dimensionalen Raum

$$\vec{x}\,' = A\,\vec{x} \tag{590}$$

mit einer beliebigen komplexen $n \times n$-Matrix A. Weiterhin sei eine orthonormierte Basis mit den Basisvektoren

$$\vec{e}_1,\, ...,\, \vec{e}_n \tag{591}$$

gegeben, in der sich die Vektoren $\vec{x}$, $\vec{x}\,'$ aus (590) folgendermaßen darstellen lassen

$$\vec{x} \;=\; x_1\vec{e}_1 + ... + x_n\vec{e}_n = \sum_{i=1}^{n} x_i\,\vec{e}_i, \tag{592}$$

$$\vec{x}\,' \;=\; x'_1\vec{e}_1 + ... + x'_n\vec{e}_n = \sum_{i=1}^{n} x'_i\,\vec{e}_i \tag{593}$$

mit komplexen Komponenten x_i bzw. x'_i. Der Übergang zu einer neuen Basis von linear unabhängigen Basisvektoren $\vec{v}^{\,(i)}$, $i = 1,\, ...,\, n$, wird mit dem Begriff *Basiswechsel* bezeichnet. In dieser neuen Basis lassen sich die Vektoren $\vec{x}$, $\vec{x}\,'$ entwickeln

$$\vec{x} \;=\; \xi_1\vec{v}^{\,(1)} + ... + \xi_n\vec{v}^{\,(n)} = \sum_{j=1}^{n} \xi_j\,\vec{v}^{\,(j)}, \tag{594}$$

$$\vec{x}\,' \;=\; \xi'_1\vec{v}^{\,(1)} + ... + \xi'_n\vec{v}^{\,(n)} = \sum_{j=1}^{n} \xi'_j\,\vec{v}^{\,(j)}, \tag{595}$$

wieder mit komplexen Koeffizienten ξ_j bzw. ξ'_j.

Jeder Vektor $\vec{v}^{\,(j)}$ der neuen Basis läßt sich nach der alten Basis entwickeln

$$\vec{v}^{\,(j)} = v_1^{(j)}\,\vec{e}_1 + ... + v_n^{(j)}\,\vec{e}_n = \sum_{i=1}^{3} v_i^{(j)}\,\vec{e}_i \quad \text{für} \quad j = 1,\, ...,\, n, \tag{596}$$

wobei der komplexe Skalar $v_i^{(j)}$ die i-te Komponente des j-ten Vektors $\vec{v}^{\,(j)}$ ist für $j, j = 1,\, ...,\, n$. Die Basisvektoren der neuen Basis können als Spalten einer $n \times n$-Matrix V aufgefasst werden

$$V = \left(\vec{v}^{\,(1)},\, ...,\, \vec{v}^{\,(n)}\right) = \begin{pmatrix} v_1^{(1)} & \dots & v_1^{(n)} \\ \vdots & & \vdots \\ v_n^{(1)} & \dots & v_n^{(n)} \end{pmatrix} = \left(v_i^{(j)}\right) \quad \text{mit Elementen} \quad V_{ij} = v_i^{(j)}. \tag{597}$$

Anwendung von (594), (596) und (597) sowie Vertauschung der Summationsreihenfolge führen auf

$$\vec{x} = \sum_{j=1}^{n} \xi_j\,\vec{v}^{\,(j)} = \sum_{j=1}^{n} \xi_j \sum_{i=1}^{n} V_{ij}\,\vec{e}_i = \sum_{i=1}^{n} \left(\sum_{j=1}^{n} V_{ij}\,\xi_j\right)\vec{e}_i. \tag{598}$$

Nach Vergleich mit (592) ergeben sich die Vektoren $\vec{x}$, $\vec{x}\,'$ ausgedrückt in der neuen Basis

$$x_i = \sum_{j=1}^{n} V_{ij}\,\xi_j \quad \text{bzw.} \quad \vec{x} = V\,\vec{\xi} \quad \text{und entsprechend} \quad \vec{x}\,' = V\,\vec{\xi}\,'. \tag{599}$$

Die lineare Abbildung (590) läßt sich in der neuen Basis folgendermaßen ausdrücken

$$V\,\vec{\xi}\,' = A\,V\,\vec{\xi}. \tag{600}$$

Da die neue Basis aus linear unabhängigen Vektoren besteht, ist $\det V \neq 0$ und es existiert das Inverse V^{-1} von V. Multipliziert man (600) von links mit V^{-1}, erhält man für die lineare Abbildung (590) in der neuen Basis

$$\vec{\xi}\,' = V^{-1}\,A\,V\,\vec{\xi}. \tag{601}$$

Vergleicht man das mit der Standardform der linearen Abbildung

$$\vec{\xi}\,' = \tilde{A}\,\vec{\xi}, \tag{602}$$

so ergibt sich folgendes Transformationsverhalten für die Matrix A unter dem Basiswechsel

$$\tilde{A} = V^{-1}\,A\,V. \tag{603}$$

Bisher war von einem allgemeinen Basiswechsel die Rede. Besonders interessant ist der Basiswechsel für eine lineare Abbildung mit einer reell-symmetrischen oder einer hermiteschen Matrix A, da dann als Basis die Eigenvektoren der Matrix A gewählt werden können und sich die transformierte Matrix $\tilde{A}$, wie im nächsten Unterkapitel gezeigt werden wird, auf eine Diagonalmatrix reduziert.

Dies kommt in der Physik massiv zum Einsatz in der Matrixformulierung der Quantenmechanik, wenn es darum geht, einen Satz von Messgrößen zu beschreiben, deren gleichzeitige Messbarkeit nicht von einer Unschärferelation verhindert wird; es stellt sich dann heraus, dass alle solchen Messgrößen zugeordneten hermiteschen komplexen Matrizen durch einen gemeinsamen Basiswechsel diagonalisiert werden können.

10.6.2 Hauptachsentransformation bei hermiteschen komplexen Matrizen

Gegeben sei eine hermitesche komplexe Matrix A mit $A = A^{\dagger}$. Wie in Kap. 10.4 und Kap. 10.5 gezeigt wurde, führt das Eigenwertproblem

$$A\,\vec{x}^{\,(i)} = \lambda_i\,\vec{x}^{\,(i)} \tag{604}$$

auf reelle Eigenwerte λ_i und ein orthonormiertes System von Eigenvektoren $\vec{x}^{\,(i)}$. Es bietet sich an, einen Basiswechsel von der orthonormierten Basis $\vec{e}_1$, ..., $\vec{e}_n$ zur neuen orthonormierten Basis $\vec{x}^{(1)}$, ..., $\vec{x}^{(n)}$, der Eigenbasis von A, durchzuführen.

Definition: Diese sogenannte **Hauptachsentransformation** transformiert auf die sogenannte **Eigenbasis** der hermiteschen komplexen Matrix A,

$$\vec{x} = \sum_{i=1}^{n} \xi_i\,\vec{x}^{\,(i)}. \tag{605}$$

Offensichtlich hat man es für jede hermitesche Matrix A mit einer anderen Hauptachsentransformation zu tun. Anstelle der allgemeinen, in (597) definierten Transformationsmatrix V tritt nun die Matrix X auf, deren Spalten aus den Eigenvektoren von A bestehen,

$$X = \left(\vec{x}^{\,(1)}, ..., \vec{x}^{\,(n)}\right). \tag{606}$$

Die adjungierte Matrix $X^\dagger$ hat Zeilen, die aus den adjungierten Eigenvektoren $\vec{x}^{\,(i)\dagger}$ bestehen

$$X^\dagger = \begin{pmatrix} \vec{x}^{\,(1)\dagger} \\ \vdots \\ \vec{x}^{\,(n)\dagger} \end{pmatrix}. \tag{607}$$

Das Produkt $X^\dagger X$ erweist sich auf Grund der Orthonormalität (568) der Eigenvektoren als die Einheitsmatrix

$$X^\dagger\, X = \left(\vec{x}^{\,(i)\dagger}\, \vec{x}^{\,(j)}\right) = \left(\delta_{ij}\right) = \mathbf{1}. \tag{608}$$

Dementsprechend ist

$$X^\dagger = X^{-1} \;\text{ und es gilt }\; X^\dagger X = X X^\dagger = \mathbf{1}. \tag{609}$$

Eine komplexe Matrix X mit dieser Eigenschaft heißt eine *unitäre Matrix*.

Ist die Matrix A reell-symmetrisch, dann ist $X^\dagger = X^T$; entsprechend wird (609) zu

$$X^T = X^{-1} \;\text{ und es gilt }\; X^T X = X X^T = \mathbf{1}; \tag{610}$$

Eine reelle Matrix X mit dieser Eigenschaft ist eine sogenannte *orthogonale* Matrix. (Dreidimensionale räumliche Drehmatrizen sind Beispiele für orthogonale 3×3-Matrizen).

In der Eigenbasis nimmt die in der linearen Abbildung (602)

$$\vec{\xi}\,' = \tilde{A}\,\vec{\xi}, \tag{611}$$

auftretende transformierte Matrix $\tilde{A}$ nach (603), (609), der Eigenwertgleichung (604) und der Orthogonalität (568) der Eigenvektoren die folgende Form an

$$\tilde{A} \;=\; X^\dagger\, A\, X = X^\dagger\, A \left(\vec{x}^{\,(1)}, ..., \vec{x}^{\,(n)}\right) = X^\dagger \left(\lambda_1\, \vec{x}^{\,(1)}, ..., \lambda_n\, \vec{x}^{\,(n)}\right) \tag{612}$$

$$= \left(\lambda_j\, \vec{x}^{\,(i)\dagger}\, \vec{x}^{\,(j)}\right) = \left(\lambda_j\, \delta_{ij}\right) = \begin{pmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \dots & \lambda_n \end{pmatrix}. \tag{613}$$

Eine hermitesche komplexe Matrix A wird durch ihre Hauptachsentransformation in eine **Diagonalmatrix** transformiert, **deren Diagonalelemente durch ihre Eigenwerte gegeben sind**. Dies gilt a forteriori auch für reell-symmatrische Matrizen A.

Am Beispiel der reell-symmetrischen dreidimensionalen Matrix, die in Kap. 10.5 behandelt wurde, läßt sich dieses Ergebnis nachvollziehen

$$\tilde{A} = X^T\, A\, X = \begin{pmatrix} 0 & \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ 0 & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ 1 & 0 & 0 \end{pmatrix}^T \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 2 \end{pmatrix} \begin{pmatrix} 0 & \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ 0 & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ 1 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 2 & 0 & 0 \\ & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix}, \tag{614}$$

wobei als Diagonalelemente die drei Eigenwerte (585) auftreten.

10.7 Nachtrag: Ein System von linearen, homogenen, gekoppelten Differentialgleichungen erster Ordnung mit konstanten Koeffizienten

In der Vorlesung "Mathematische Ergänzung I" war bereits der Begriff gekoppelter Differentialgleichungen eingeführt worden und auf ein System zweier linearer gekoppelter Differentialgleichungen zweiter Ordnung mit konstanten Koeffizienten angewandt worden; dies war nötig, um schnell den mathematischen Unterbau für wichtige gekoppelte Systeme in der klassischen Mechanik bereitzustellen.

Mittlerweile liegt auch das notwendige mathematische Rüstzeug für die Behandlung eines Systems von n linearen, homogenen, gekoppelten Differentialgleichungen erster Ordnung mit konstanten Koeffizienten vor.

Gesucht sind die n Lösungsfunktionen $y_1(x)$, $y_2(x)$, ..., $y_n(x)$ des folgenden Differentialsgleichungssystems

$$\begin{aligned} y_1'(x) &= a_{11}\, y_1(x) + a_{12}\, y_2(x) + ... + a_{1n}\, y_n(x), \\ y_2'(x) &= a_{21}\, y_1(x) + a_{22}\, y_2(x) + ... + a_{2n}\, y_n(x), \\ &\vdots \\ y_n'(x) &= a_{n1}\, y_1(x) + a_{n2}\, y_2(x) + ... + a_{nn}\, y_n(x). \end{aligned} \tag{615}$$

Es seien reelle Konstanten a_{ij} für $i,\ j \in \{1, 2, , ..., n\}$ angenommen. Mit $y_i'(x)$ ist in allen Fällen die Ableitung $\mathrm{d}\, y_i(x)/\mathrm{d}x$ bezeichnet. Dieses Differentialgleichungssystem ist "maximal" gekoppelt in dem Sinn, dass in der Gleichung für $y_i'(x)$ auf der rechten Seite nicht nur ein Term proportional zur Funktion $y_i(x)$ auftritt sondern dass alle anderen Funktionen $y_j(x)$ mit $j \neq i$ ebenfalls auftreten und zwar linear.

Das Differentialgleichungssystem läßt sich elegant umschreiben, indem man die Vektoren

$$\vec{y}(x) = \begin{pmatrix} y_1(x) \\ \vdots \\ y_n(x) \end{pmatrix} \quad \text{und} \quad \vec{y}\,'(x) = \begin{pmatrix} y_1'(x) \\ \vdots \\ y_n'(x) \end{pmatrix} = \frac{\mathrm{d}}{\mathrm{d}\, x}\, \vec{y}(x) \tag{616}$$

und die reelle Matrix der Koeffizienten

$$A = \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \dots & a_{nn} \end{pmatrix} \tag{617}$$

einführt.

Das **gekoppelte System von Differentialgleichungen erster Ordnung** (615) nimmt dann die elegante Form an

$$\vec{y}\,'(x) = A\, \vec{y}(x) \quad \text{mit} \quad \vec{y}\,'(x) = \frac{\mathrm{d}}{\mathrm{d}\, x}\, \vec{y}(x). \tag{618}$$

Als erfolgreicher Lösungsansatz wird sich

$$y_i(x) = Y_i\, e^{\lambda\, x} \quad \text{für} \quad i = 1, ..., n \quad \text{bzw. kurz} \quad \vec{y}(x) = \vec{Y}\, e^{\lambda\, x} \tag{619}$$

erweisen, wobei Y_i, die Komponenten des Vektors $\vec{Y}$, Konstanten sind. Es ist zu beachten, dass in diesem Ansatz für alle Komponenten derselbe Wert des Parameters λ angesetzt ist. Der Ansatz führt auf

$$y_i'(x) = \lambda\, Y_i\, e^{\lambda\, x} \quad \text{für} \quad i = 1,\, ...,\, n \quad \text{bzw. kurz} \quad \vec{y}\,'(x) = \lambda\, \vec{Y}\, e^{\lambda\, x}. \tag{620}$$

Setzt man den Ansatz (619) für $\vec{y}(x)$ und den Ausdruck (620) für $\vec{y}\,'(x)$ in das Differentialgleichungssystem (618), dividiert man beide Seiten durch $e^{\lambda\, x}$ und vertauscht man die Seiten, erhält man die *Eigenwertgleichung* für die Matrix A

$$A\, \vec{Y} = \lambda\, \vec{Y}. \tag{621}$$

Man steht also vor dem Eigenwertproblem, d.h. vor der Aufgabe, die Eigenwerte λ und die zugehörigen Eigenvektoren $\vec{Y}$ der Matrix A zu bestimmen. Nach Kap. 10.3 erhält man die n Eigenwerte als Lösung der charakteristischen Gleichung

$$\det\Big(A - \lambda\, \mathbf{1}\Big) = 0. \tag{622}$$

Sie sind reell oder treten in komplex konjugierten Paaren auf. Ist A reell-symmetrisch, $A = A^T$, sind alle Eigenwerte reell. Die zu den einzelnen Eigenwerten λ_ν für $\nu = 1,\, ...,\, n$ gehörigen Eigenvektoren $\vec{Y}^{\,(\nu)}$ sind Lösungen der Gleichungen

$$A\, \vec{Y}^{\,(\nu)} = \lambda_\nu\, \vec{Y}^{\,(\nu)} \quad \text{für} \quad \nu = 1,\, ...,\, n. \tag{623}$$

Im spezifischen Fall von reell-symmetrischen Matrizen A bilden die Eigenvektoren nach Kap. 10.4 sogar ein Orthonormalsystem.

Es gibt also n Fundamentallösungen des gekoppelten Systems von Differentialgleichungen (618)

$$\vec{Y}^{\,(\nu)}\, e^{\lambda_\nu\, x} \quad \text{für} \quad \nu = 1,\, ...,\, n. \tag{624}$$

Die allgemeine Lösung ist durch eine beliebige Linearkombination dieser **Fundamentallösungen** gegeben

$$\vec{y}(x) \;=\; \sum_{\nu=1}^{n} C_\nu\, \vec{Y}^{\,(\nu)}\, e^{\lambda_\nu\, x} \quad \text{bzw.} \tag{625}$$

$$y_i(x) \;=\; \sum_{\nu=1}^{n} C_\nu\, Y_i^{(\nu)}\, e^{\lambda_\nu\, x}, \tag{626}$$

wobei $Y_i^{(\nu)}$ die i-te Komponente des ν-ten Eigenvektors ist und C_ν beliebig wählbare Konstanten sind.

n Anfangsbedingungen

$$y_i(x_0) = y_{i0} \quad \text{für} \quad i = 1,\, ...,\, n \tag{627}$$

legen die n Konstanten C_ν fest; dazu ist das System von n linearen Gleichungen in den n Konstanten C_ν

$$y_{i0} = \sum_{\nu=1}^{n} C_\nu\, Y_i^{(\nu)}\, e^{\lambda_\nu\, x_0} \tag{628}$$

nach den Konstanten C_ν zu lösen.

Gekoppelte lineare Differentialgleichungen erster Ordnung spielen eine wichtige Rolle bei der Lösung gewöhnlicher linearer Differentialgleichungen n-ter Ordnung. Der Grund ist, dass sich eine lineare gewöhnliche Differentialgleichung n-ter Ordnung mit konstanten Koeffizienten in ein System von n gekoppelten linearen Differentialgleichungen erster Ordnung mit konstanten Koeffizienten umschreiben läßt. Dies soll an einem einfachen Beispiel illustriert werden.

Beispiel für n = 2:

Die lineare homogene Differentialgleichung zweiter Ordnung mit konstanten Koeffizienten

$$y'' + a\,y' + b\,y = 0, \tag{629}$$

die ausführlich in der Vorlesung "Mathematische Ergänzung zu Experimentalphysik I" behandelt worden war, hat die allgemeine Lösung

$$y(x) = C_1\,e^{\lambda_1\,x} + C_2\,e^{\lambda_2\,x} \quad (\text{für} \quad \lambda_1 \neq \lambda_2), \tag{630}$$

wobei die $\lambda_{1,2}$ die Lösungen

$$\lambda_{1,2} = -\frac{a}{2} \pm \sqrt{\frac{a^2}{4} - b} \tag{631}$$

der charakteristischen Gleichung

$$\lambda^2 + a\,\lambda + b = 0 \tag{632}$$

sind.

Diese Ergebnisse sollen nun anhand eines Satzes von zwei linearen, homogenen gekoppelten Differentialgleichungen erster Ordnung mit konstanten Koeffizienten illustriert werden. Ein geeigneter Ansatz ist

$$y_1' = y_2, \tag{633}$$

$$y_2' = -b\,y_1 - a\,y_2; \tag{634}$$

oder zusammengefasst

$$\begin{pmatrix} y_1' \\ y_2' \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ -b & -a \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} \quad \text{bzw. kurz} \quad \vec{y}\,' = A\,\vec{y} \quad \text{mit} \quad A = \begin{pmatrix} 0 & 1 \\ -b & -a \end{pmatrix}. \tag{635}$$

Differenziert man (633), erhält man $y_1'' = y_2'$. Ersetzt man entsprechend y_2' auf der linken Seite von (634) durch y_1'' und setzt man $y_2 = y_1'$ aus (633) auf der rechten Seite von (634) ein, erhält man die Differentialgleichung (629) zweiter Ordnung für y_1.

Die Eigenwerte der Matrix A ergeben sich aus der charakteristischen Gleichung

$$\det \begin{pmatrix} -\lambda & 1 \\ -b & -a-\lambda \end{pmatrix} = \lambda^2 + a\,\lambda + b = 0, \tag{636}$$

die identisch mit (632) ist; sie führt dementsprechend auf die beiden Eigenwerte (631). Die zugehörigen Eigenvektoren können leicht bestimmt werden und führen nach Umdefinition der freien Konstanten ersichtlich auf die allgemeine Lösung (630) für $y_1(x)$ (für $\lambda_1 \neq \lambda_2$).