

Theoretische Physik Ia: Mathematische  
Methoden der klassischen Mechanik

Frank Wilhelm-Mauch

2. Januar 2018

Fachrichtung Physik, Universität des Saarlandes

# Vorbemerkungen

Dieses Skript dient zur Erleichterung der Mitschrift in der Vorlesung Theoretische Physik Ia, Mathematische Methoden der klassischen Mechanik, wie sie an der Universität des Saarlandes für Studierende des ersten Semesters angeboten wird. Diese Vorlesung verfolgt mehrere Ziele: Einerseits liefert sie die Mathematik, die zum Folgen der parallel angebotenen Vorlesung Experimentalphysik I benötigt wird. Sie entwickelt diese aber andererseits deutlich weiter, so dass bereits ein solides Fundament für den Zyklus der Kursvorlesungen in theoretischer Physik gelegt wird - wobei Sie auch dort Konzepte aus der aktuellen Veranstaltung weiterentwickeln werden.

Der Anspruch dieser Vorlesung ist es, Ihnen die entscheidenden Rechenmethoden der Physik der ersten Studienjahre an die Hand zu geben. Das Niveau dieser Methoden hat universitären Anspruch - aber es ersetzt nicht die Mathematikausbildung in den dort vorgesehenen Veranstaltungen: In denen lernen Sie nämlich die Präzision und Strenge, die die Mathematik auszeichnet, Sie definieren und beweisen, während wir in dieser Veranstaltung herleiten und argumentieren und gerade nicht die Strenge des mathematischen Beweises. Wir orientieren uns stattdessen an Lösungsstrategien für physikalische Probleme, wie sie z.B. in der Mechanik auftauchen. Damit greifen wir den Mathematikvorlesungen auch weit vor - dass bereits zu Studienbeginn in Physik eine umfangreiche mathematische Sprache benötigt wird ist eine Herausforderung, die schon zumindest besteht seit unter Arnold Sommerfeld das Physikstudium in seiner heutigen Form definiert wurde.

Wenn Sie Bücher mit zur Vorlesung passenden Titeln anschauen, die schon etwas älter sind, werden Sie oft feststellen, dass die nicht so richtig zu dieser Vorlesung passen. Wenn sie nämlich vor der Umstellung auf Bachelor- und Masterabschlüsse entworfen worden, also für das Studium mit dem Diplom nach 10 Semestern als Erstabschluss, dann gehen sie oft davon aus, dass man *zuerst* umfangreiche Mathematikveranstaltungen besucht und dann in die theoretische Physik eintaucht. Diese Vorlesung orientiert sich im Stil (aber nicht in der Reihenfolge!) an einem Skript das in München und Köln entsteht und hoffentlich bald als Lehrbuch erscheint, und das in der Notation um einiges klarer und deutlich systematischer ist, als ältere Werke.

Dennoch erstellen wir ein separates Skript! Das hat vor allen Dingen etwas mit der Stoffreihenfolge zu tun - nicht mal einer der Autoren des neuen Buches (Jan von Delft) arbeitet sein Werk in der publizierten Reihenfolge ab. Wir

bauen immer wieder Beispiele aus der klassischen Mechanik ein um die Rechen-techniken zu motivieren. Das ist sogar sehr natürlich: Tatsächlich ist die Physik oft Motivation für Entwicklung der Mathematik, und umgekehrt. Wir versuchen auch, das in der Experimentalphysik benötigte Material einzuführen *bevor* es angewandt wird. Dafür nehmen wir einige Sprünge im Stoffaufbau in Kauf.

Das Skript ersetzt weder Buch noch Vorlesungsbesuch. Vielmehr soll es Ihnen lediglich bei der Niederschrift helfen. In der ersten Auflage (WS 17/18) verzichten wir auf Zeichnungen: Die werden in der Vorlesung entwickelt. Wir gehen auch davon aus, dass Sie den Abiturstoff etwa auf dem Niveau beherrschen, das im Vorkurs erarbeitet wird - falls nicht arbeiten Sie bitte im Skript des Vorkurses nach.

## Die unvernünftige Effektivität der Mathematik in den Naturwissenschaften<sup>1</sup>

Physik zeichnet sich unter den Naturwissenschaften als diejenige aus, die immer quantitativ ist: Es geht also nicht um das *was* eines Naturphänomens, sondern auch um das *wieviel*. Um dies zum Ausdruck zu bringen wird Mathematik verwendet. Mathematik ist darum die *Sprache* der Physik die zu erlernen eine Grundlage des Physikstudiums ist, so wie Sie zum Studium der Anglistik erst mal Englisch können müssen.

Wenn man die rigorose Vorgehensweise der modernen Mathematik anschaut und ihre manchmal idiosynkratischen Begrifflichkeiten kann man sich die Frage stellen, ob es sich eigentlich um eine Naturwissenschaft oder eine Abteilung der Philosophie handelt <sup>2</sup>. Dennoch funktioniert Mathematik in der Beschreibung der Natur! Das hat seinen Grund in der Wechselbeziehung von Mathematik und ihren Anwendern - traditionell in Physik und Technik, in den letzten Jahrzehnten auch in der Informatik.

Saarbrücken, im Herbst 2017

Frank Wilhelm-Mauch

---

<sup>1</sup>„The unreasonable effectiveness of mathematics in the natural sciences“ ist ein Artikel des ungarisch-amerikanischen Physikers Eugene Wigner

<sup>2</sup>eine Frage die ernsthafte Philosophen wie Bertrand Russel und Ludwig Wittgenstein durchaus intensiv beschäftigt hat

# Inhaltsverzeichnis

<b>1</b>	<b>Eindimensionale Analysis</b>	<b>8</b>
1.1	Funktionsbegriff und einfache Eigenschaften . . . . .	8
1.1.1	Funktionsbegriff . . . . .	8
1.1.2	Eindeutigkeit und Umkehrbarkeit . . . . .	9
1.1.3	Stetigkeit . . . . .	9
1.2	Ableitungen . . . . .	10
1.2.1	Physikalische Motivation . . . . .	10
1.2.2	Ableitungsregeln . . . . .	11
1.2.3	Ableitungen elementarer Funktionen . . . . .	11
1.3	Taylorreihen und Näherungen . . . . .	12
1.3.1	Die komplexe Exponentialfunktion . . . . .	13
1.4	Integration . . . . .	14
1.4.1	Was ist ein Integral? . . . . .	14
1.4.2	Integration über reelle Intervalle . . . . .	15
1.4.3	Integrationsregeln und -methoden . . . . .	16
1.4.4	Integration in der klassischen Mechanik . . . . .	17
<b>2</b>	<b>Vektorrechnung und Kraftfelder</b>	<b>19</b>
2.1	Motivation . . . . .	19
2.2	Standardvektoren . . . . .	20
2.3	Allgemeine Vektorräume . . . . .	20
2.3.1	Definition . . . . .	20
2.3.2	Beispiele für Vektorräume . . . . .	21
2.3.2.1	Der Standardvektorraum . . . . .	21
2.3.2.2	Affine Räume . . . . .	21
2.3.2.3	Funktionenräume . . . . .	21
2.3.2.4	Komplexe Zahlen . . . . .	22
2.4	Basis und Dimension . . . . .	22
2.4.1	Untervektorräume . . . . .	22
2.4.2	Lineare Unabhängigkeit . . . . .	23
2.4.3	Vollständigkeit und Basis . . . . .	24
2.4.4	Beispiele . . . . .	25
2.4.5	Anwendung: Kräftezerlegung . . . . .	25
2.5	Skalarprodukt, Länge, Euklidische Geometrie . . . . .	26

2.5.1	Skalarprodukt in Vektorräumen über $\mathbb{R}$ . . . . .	26
2.5.2	Das Standardskalarprodukt im $\mathbb{R}^n$ . . . . .	27
2.5.3	Orthogonalität und lineare Unabhängigkeit . . . . .	28
2.6	Das Vektorprodukt . . . . .	29
2.6.1	Geometrische Beschreibung . . . . .	29
2.6.2	Algebraische Beschreibung . . . . .	30
<b>3</b>	<b>Grundbegriffe der Vektoranalysis</b>	<b>32</b>
3.1	Kurven im Raum . . . . .	32
3.2	Ableitungen von Kurven . . . . .	33
3.3	Kurvenlänge . . . . .	34
3.4	Vektorfelder und Linienintegrale . . . . .	35
3.5	Der Gradient eines Skalarfelds . . . . .	37
3.6	Gradientenfelder und Wegintegrale . . . . .	38
3.7	Flächenintegrale . . . . .	40
3.8	Der Satz von Stokes . . . . .	41
3.9	Volumenintegrale und der Satz von Gauß . . . . .	43
3.10	Krummlinige Koordinaten . . . . .	44
3.10.1	Polarkoordinaten . . . . .	45
3.10.2	Zylinderkoordinaten . . . . .	46
3.10.3	Kugelkoordinaten . . . . .	46
<b>4</b>	<b>Lineare Abbildungen und Matrizen</b>	<b>48</b>
4.1	Lineare Abbildungen . . . . .	48
4.1.1	Lineare Abbildungen und die Basisdarstellung . . . . .	49
4.2	Grundeigenschaften von Matrizen . . . . .	50
4.3	Matrizen und lineare Gleichungssysteme . . . . .	52
4.3.1	Die inverse Matrix . . . . .	53
4.3.2	Invertierbarkeit . . . . .	54
4.3.3	Beispiele . . . . .	55
4.4	Basiswechsel . . . . .	56
4.4.1	Beispiele . . . . .	58
4.4.2	Die Spur einer Matrix . . . . .	59
4.5	Determinante . . . . .	59
4.5.1	Permutationen und ihr Signum . . . . .	59
4.5.2	Die Determinante . . . . .	60
4.5.3	Spezialfälle . . . . .	61
4.5.4	Eigenschaften . . . . .	61
4.6	Diagonalisierung und Eigenwerte . . . . .	62
4.6.1	Eigenvektoren und Eigenwerte . . . . .	62
4.6.2	Das charakteristische Polynom . . . . .	62
4.6.3	Eigenschaften von Eigenwerten und -vektoren . . . . .	63
4.6.3.1	Nullstellen des Charakteristischen Polynoms . . . . .	63
4.6.3.2	Determinante und Invertierbarkeit . . . . .	64
4.6.3.3	Lineare Unabhängigkeit verschiedener Eigenvektoren . . . . .	64

4.6.4	Diagonalisierbarkeit und Diagonalisierung . . . . .	64
4.6.4.1	Nichtentartetes Spektrum . . . . .	65
4.6.4.2	Entartetes Spektrum . . . . .	65
4.6.4.3	Spektralsatz . . . . .	66
4.6.4.4	Diagonalisierung . . . . .	66
4.6.5	Beispiele und spezielle Matrizen . . . . .	66
4.6.5.1	Drehmatrizen . . . . .	66
4.6.5.2	Orthogonale Matrizen . . . . .	67
4.6.5.3	Symmetrische Matrizen . . . . .	68
4.7	Matrizen und lineare Abbildungen in komplexen Vektorräumen .	68
4.7.1	Eigenwerte hermitescher Matrizen . . . . .	69
<b>5</b>	<b>Gewöhnliche Differenzialgleichungen</b>	<b>70</b>
5.1	Einführung . . . . .	70
5.2	Differenzialgleichungen mit getrennten Veränderlichen . . . . .	71
5.2.1	Autonome Differenzialgleichungen . . . . .	73
5.3	Lineare Differenzialgleichungen . . . . .	74
5.3.1	Homogene lineare Differenzialgleichungen . . . . .	74
5.3.2	Inhomogene lineare Differenzialgleichungen . . . . .	75
5.3.3	Lineare Differenzialgleichungen mit konstanten Koeffizien- enzen . . . . .	77
5.4	Systeme von Differenzialgleichungen . . . . .	77
5.4.1	Definitionen . . . . .	77
5.4.2	Lineare Systeme mit konstanten Koeffizienten . . . . .	78
5.4.3	Matrixexponentiale . . . . .	78
5.4.4	Lösung des Systems mit konstanten Koeffizienten . . . . .	79
5.4.5	Beispiel: Der gedämpfte harmonische Oszillator . . . . .	80
5.4.6	Physikalische Grenzfälle . . . . .	82
5.4.6.1	Ungedämpft . . . . .	82
5.4.6.2	Unterdämpft . . . . .	83
5.4.6.3	Überdämpft . . . . .	83
5.4.6.4	Kritisch gedämpft . . . . .	83
5.4.7	Der angetriebene harmonische Oszillator . . . . .	84
5.4.7.1	Spezielle Lösung, komplexer Ansatz . . . . .	84
5.4.7.2	Gesamtlösung . . . . .	84
5.4.7.3	Resonanz . . . . .	85
5.5	Fortgeschrittene Techniken zur Lösung von linearen Differenzial- gleichungen . . . . .	86
5.5.1	Überlagerung von Antrieben beim harmonischen Oszillator	86
5.5.2	Das Dirac-Delta . . . . .	87
5.5.3	Annäherung durch Funktionenfolgen . . . . .	87
5.5.4	Fourierzerlegung . . . . .	88
5.5.5	Anwendung auf den getriebenen Oszillator . . . . .	90
5.6	Greensche Funktionen . . . . .	90
5.6.1	Definition und Lösungsformel . . . . .	91

5.6.2	Greensche Funktion des getriebenen harmonischen Oszil- lators . . . . .	91
5.6.3	Anwendung . . . . .	92
5.7	Komplexe Differenzierbarkeit . . . . .	92
5.7.1	Analytische Funktionen . . . . .	93
5.7.2	Integrale über geschlossene Wege . . . . .	93
5.7.3	Der Residuensatz . . . . .	94



# Kapitel 1

## Eindimensionale Analysis

Die Analysis in einer Dimension haben Sie schon in der Schule kennengelernt. Dieses Kapitel dient als Erinnerung und illustriert Anwendung in der eindimensionalen Kinematik. Wir empfehlen, falls Sie am Vorkurs nicht teilgenommen haben, Ihr Wissen mit dem Skript (und den Übungen) des Vorkurses abzugleichen.

### 1.1 Funktionsbegriff und einfache Eigenschaften

#### 1.1.1 Funktionsbegriff

Eine Funktion ist eine Abbildung von einer Definitionsmenge auf eine Wertemenge,  $f : D \mapsto W, x \mapsto f(x)$ . Die Funktion ist wohldefiniert, wenn jedem Element von  $D$  genau ein Element von  $W$  zugeordnet wird. Uns interessieren zunächst Funktionen einer reellen Variablen, also der Fall dass  $D \subseteq \mathbb{R}$  ist. Dabei interessieren uns vor allen Dingen Intervalle wie das geschlossene und das offene

$$[a, b] = \{x \in \mathbb{R} | a \leq x \leq b\} \quad (a, b) = \{x \in \mathbb{R} | a < x < b\} \quad (1.1)$$

sowie die halboffenen  $(a, b]$  und  $[a, b)$ . In der Experimentalphysik begegnet Ihnen z.B. bei der eindimensionalen Bewegung der Ort als Funktion der Zeit,  $x(t)$  als Funktion. Die Wohldefiniertheit bedeutet, dass wir zu jeder Zeit  $t$  eindeutig genau einen Ort  $x(t)$  ablesen können. Physikalische Bewegungen finden normalerweise zwischen festen Zeiten statt, also in abgeschlossenen Intervallen - manchmal ist es aber auch sinnvoll zuzulassen, dass diese Zeiten gegen unendlich gehen. Da am Ende Zeiten und Orte immer Messgrößen sind kann man argumentieren, dass in der Physik höchstens rationale Zahlen vorkommen - da diese aber reelle Zahlen beliebig gut nähern, und da reelle Zahlen mehr Rechnungen ermöglichen, stört uns dieser Unterschied nicht.

### 1.1.2 Eindeutigkeit und Umkehrbarkeit

Betrachten wir den Wertebereich  $W$  genauer - dieser kann zunächst mal großzügig gewählt werden, z.B.  $W = \mathbb{R}$ . Wenn wir aber eine Funktion haben, für die zu jedem  $y \in W$  auch ein Urbild existiert, also ein  $x \in D$  so, dass  $f(x) = y$ , so heißt die Funktion surjektiv. Wenn wir eine Funktion haben, zu der es zu jedem  $y \in W$  nicht mehr als ein Urbild gibt, wenn also aus  $f(x_1) = y$  und  $f(x_2) = y$  automatisch  $x_1 = x_2$  folgt, so heißt die Funktion injektiv. Eine Funktion, die injektiv und surjektiv ist heißt bijektiv (oder ein-eindeutig): In ihr wird jedem Punkt der Wertemenge genau ein Punkt der Definitionsmenge zugeordnet. Damit ist die Funktion umkehrbar, die Umkehrfunktion  $f^{-1} : W \mapsto D$   $f^{-1}(f(x)) = x$  ist wohldefiniert.

Für die eindimensionale Bewegung  $x(t)$  bedeutet Surjektivität, dass wir die Menge aller erreichten Punkte korrekt als  $W$  identifiziert haben - das ist mehr eine Buchhaltungsaufgabe als eine physikalische Aussage. Injektivität bedeutet, dass während der Bewegung jede Ortsmarkierung nur einmal aufgesucht wird, d.h. die Bewegung ändert niemals die Richtung und bleibt auch nie für eine messbare Zeit stehen (prüfen Sie mal  $x = x_0(t/t_0)^3$ ). Wenn wir eine bijektive Bewegung haben, wenn also die Wertemenge ganz abgedeckt ist und nicht umkehren, dann ist die Funktion umkehrbar, d.h. dann können wir durch Beobachten des Ortes die Zeit bestimmen. Das wäre z.B. bei einer ideal pünktlichen Eisenbahn der Fall.

### 1.1.3 Stetigkeit

Ein weiterer Grundbegriff ist der der Stetigkeit. Stetigkeit in  $x_0 \in D$  bedeutet: Egal, wie wir uns  $x_0$  mit einer Zahlenfolge  $\{x_n, n \in \mathbb{N}\}$  in  $D$  nähern, die resultierende Zahlenfolge  $\{f(x_n)\}$  hat immer den gleichen Grenzwert, nämlich  $f(x_0)$ . Praktisch bedeutet das, dass wir die Funktion kontinuierlich zeichnen können ohne den Stift abzusetzen. Eine Unstetigkeitsstelle kann ein Sprung oder eine Singularität sein, so ist z.B.  $f(x) = 1/x$  bei  $x_0 = 0$  auch dann unstetig, wenn wir der Funktion dort willkürlich einen Funktionswert zuordnen. Beispiele für Unstetigkeitsstellen sind bei der Heavisidefunktion

$$H(x) = \begin{cases} 0 & x < 0 \\ 1/2 & x = 0 \\ 1 & x > 0 \end{cases} \quad (1.2)$$

ist unstetig bei  $x = 0$ . Gleiches gilt für die Vorzeichenfunktion  $\text{sgn}(x) = |x|/x = 2H(x) - 1$ . Ein etwas konstruiertes Beispiel dafür, dass uns die Aussage „kontinuierlich zeichnen können“ nicht weiterhilft ist  $f(x) = \sin 1/x$ . Für  $x_n = 1/n\pi$  ist  $f(x_n) = 0$  und für  $x_n = 1/(2n + 1/2)\pi$  ist  $f(x_n) = 1$ , beide Folgen gehen aber gegen 0 - damit ist die Funktion unstetig bei  $x = 0$ .

Unsere eindimensionalen Bahnen  $x(t)$  sind stetig - wenn nicht, dann würde der Massenpunkt spontane Sprünge machen, wofür eine unendliche Kraft notwendig wäre. Haben wir es in der Physik immer mit stetigen Funktionen

zu tun? Das ist eine Frage der Perspektive - es ist oft praktisch, auch un stetige Funktionen zuzulassen, wenn eine bestimmte Auflösung nicht gefragt ist - das Feld einer Punktladung ist unstetig (aber wenn wir einem physikalischen Teilchen zu nahe kommen sehen wir, dass es vielleicht kein Punkt ist), an der Oberfläche eines transparenten Mediums ist der Brechungsindex unstetig (aber wenn wir mikroskopisch hinschauen haben wir da eine Wolke von Atomen) usw. Dieses Finden der „richtigen Auflösung“ gehört zu den wichtigsten Schritten in der Bildung physikalischer Theorien.

## 1.2 Ableitungen

### 1.2.1 Physikalische Motivation

Wir interessieren uns nicht nur für Funktionen, sondern auch für ihre Änderungsrate. Wenn wir beispielsweise in der eindimensionalen Bewegung den Ort als Funktion der Zeit  $x(t)$  kennen interessiert uns die Geschwindigkeit. Praktisch können wir uns mit Metermaß und Stoppuhr ausrüsten und an zwei Zeiten  $t_2$  und  $t_1$  den Ort messen und daraus die Durchschnittsgeschwindigkeit

$$\bar{v} = \frac{x(t_2) - x(t_1)}{t_2 - t_1} = \frac{\Delta x}{\Delta t} \quad (1.3)$$

bestimmen. Das ist nicht immer besonders aussagekräftig - wenn Sie in Saarbrücken vom St. Johanner Markt zur Uni und wieder zurück mit dem Fahrrad fahren, dann geht es auf dem Hinweg den Berg hoch und den Berg herunter - Ihre Durchschnittsgeschwindigkeit fahren Sie eigentlich nie. Die Sache ist schon aussagekräftiger, wenn Sie die Durchschnittsgeschwindigkeit für den Hin- und den Rückweg separat berechnen

$$\bar{v}_{\text{hin}} = \frac{x_{\text{Uni}} - x_{\text{Markt}}}{t_{\text{Uni}} - t_{\text{Markt1}}} \quad \bar{v}_{\text{rück}} = \frac{x_{\text{Markt}} - x_{\text{Uni}}}{t_{\text{Markt2}} - t_{\text{Uni}}} \quad (1.4)$$

und typischerweise dürfte (ohne Hilfsmotor)  $\bar{v}_{\text{rück}} > \bar{v}_{\text{hin}}$  sein. Wenn Sie jetzt aber den Weg zur Uni betrachten, dann sind Sie kurz vor dem Parkhaus wo es steil ist noch etwas langsamer, also möchten Sie eigentlich die Geschwindigkeit noch in kleineren Intervallen messen usw. Darum definieren wir die momentane Geschwindigkeit als

$$v(t) = \lim_{\Delta t \rightarrow 0} \frac{x(t + \Delta t) - x(t)}{\Delta t}. \quad (1.5)$$

Wenn man dies an einer  $x(t)$  Kurve aufzeichnet (Vorlesung) sind die Durchschnittsgeschwindigkeiten jeweils Steigungen von Sekanten und die Momentangeschwindigkeit die Steigung einer Tangente. Wir schreiben für die Ableitung einer Funktion  $f(x)$  auch  $f'(x) = \frac{df}{dx}$  wenn keine Möglichkeit zur Verwechslung besteht. In der Physik haben wir aber die Konvention, dass wir die Ableitung nach der Zeit extra durch einen Punkt beschreiben, also  $v(t) = \dot{x}(t)$ .

Ähnlich können wir mit höheren Ableitungen verfahren: Die zweite Ableitung des Ortes ist die Beschleunigung

$$\ddot{x}(t) = \dot{v}(t) = a(t). \quad (1.6)$$

## 1.2.2 Ableitungsregeln

Zunächst einmal müssen wir uns vergewissern, dass die Ableitung überhaupt existiert, dass also der Grenzwert aus Gl. (1.5) existiert. Da der Nenner beim Grenzübergang gegen null geht, muss der Zähler dies auch tun - die Funktion muss also mindestens einmal stetig sein. Stetigkeit reicht aber nicht aus, so ist z.B. die Betragsfunktion  $f(x) = |x|$  stetig, aber im Ursprung nicht differenzierbar - je nachdem ob man den Grenzwert von der positiven oder der negativen Seite bildet kommt  $\pm 1$  heraus.

Wenn wir aber einmal eine Ableitung haben, dann gelten eine Reihe nützlicher Regeln, die wir hier auflisten (Herleitung und Diskussion in der Schule, nachzulesen im Skript des Vorkurses):

1. Linearität

$$\frac{d}{dx} (\lambda f(x) + \mu g(x)) = \lambda f' + \mu g' \quad (1.7)$$

2. Produktregel:

$$(fg)' = f'g + fg' \quad (1.8)$$

.

3. Kettenregel:

$$\frac{df(g(x))}{dx} = \frac{df(g(x))}{dg} \frac{dg}{dx} \quad (1.9)$$

4. Quotientenregel

$$\frac{d}{dx} \left( \frac{f(x)}{g(x)} \right) = \frac{gf' - fg'}{g^2} \quad (1.10)$$

5. Umkehrfunktion

$$\frac{df^{-1}}{dx} = \frac{1}{f'(f^{-1}(x))} \quad (1.11)$$

## 1.2.3 Ableitungen elementarer Funktionen

Auch dies wurde in der Schule behandelt und wir listen nur auf

1. Einfache Potenzen

$$\frac{dx^r}{dx} = rx^{r-1} \quad (1.12)$$

2. Polynome

$$\frac{d}{dx} \sum_{k=0}^n a_k x^k = \sum_{k=1}^{n-1} k a_k x^{k-1}. \quad (1.13)$$

3. Exponentialfunktionen

$$\frac{d}{dx} e^x = e^x \quad \frac{d}{dx} a^x = \ln a \cdot a^x \quad (1.14)$$

4. Natürlicher Logarithmus

$$\frac{d \ln x}{dx} = \frac{1}{x} \quad (1.15)$$

5. Sinus und Kosinus

$$\frac{d \sin x}{dx} = \cos x \quad \frac{d \cos x}{dx} = -\sin x \quad (1.16)$$

6. Tangens

$$\frac{d \tan x}{dx} = \frac{1}{1+x^2} \quad (1.17)$$

7. Arcusfunktionen

$$\frac{d \arcsin x}{dx} = \frac{1}{\sqrt{1-x^2}} \quad \frac{d \arccos x}{dx} = -\frac{1}{\sqrt{1-x^2}} \quad (1.18)$$

### 1.3 Taylorreihen und Näherungen

Mittels wiederholter Ableitung lassen sich Funktionen darstellen. Der Satz von Taylor (hier ohne Beweis sagt), dass wir eine unendlich oft differenzierbare Funktion  $f$  um den Wert  $x_0$  entwickeln können gemäß

$$f(x) = \sum_{n=0}^{\infty} \frac{1}{n!} \frac{d^n f(x_0)}{dx^n} (x-x_0)^n. \quad (1.19)$$

Dies nennt man die *Taylorentwicklung* von  $f$  um  $x_0$ . Ein solches „Polynom unendlicher“ Ordnung nennt man Potenzreihe und diese nennt man dann entsprechend Taylorreihe. Wir sehen also, dass wir die Information über diese Funktion in die Ableitungen von  $f$  stecken können - der Rest der Funktion ist durch die Forderung, dass diese Funktion unendlich oft differenzierbar ist schon festgelegt.

Die Näherung über die Taylorreihe gilt in einer Umgebung von  $x_0$  (wie groß die ist lernen Sie in der Mathematik). Wenn wir sehr nahe an  $x_0$  sind, wenn  $x-x_0$  also extrem klein ist, dann fallen die Terme der Reihendarstellung schnell ab (und die Fakultät im Nenner tut ihr übrigens). Wir können dann nähern bis zur Ordnung  $N$  (hier geschrieben für den Fall  $x_0 = 0$ )

$$f(x) = \sum_{n=0}^N \frac{1}{n!} \frac{d^n f(x_0)}{dx^n} x^n + O(x^{N+1}) \quad (1.20)$$

Wir deklarieren also, ab welcher Ordnung wir Restglieder weglassen. Das Symbol  $O$  heißt *Landausymbol* und es beschreibt eine Menge:

$$O(f(x)) = \left\{ g(x) \mid \lim_{x \rightarrow 0} \frac{g(x)}{f(x)} \text{ existiert} \right\}. \quad (1.21)$$

Wir betrachten einige Beispiele. Das folgende heißt *geometrische Reihe*

$$\frac{1}{1-x} = \sum_{k=0}^{\infty} x^k = 1 + x + x^2 + \dots \quad |x| < 1 \quad (1.22)$$

Wir können also Dinge schreiben wie

$$\frac{1}{1-x} = 1 + x + O(x^2). \quad (1.23)$$

Die Exponentialfunktion ist ihre eigene Ableitung, wir haben also

$$e^x = \sum_{k=0}^{\infty} \frac{x^k}{k!} \quad (1.24)$$

Beim Sinus sehen wir, dass wir durch wiederholtes Differenzieren einen Zyklus durchlaufen  $\sin x \mapsto \cos x \mapsto -\sin x \mapsto -\cos x \mapsto \sin x$ . Damit ist bei  $x = 0$  die Funktion und jede gerade Ableitung gleich null und die ungeraden alternieren zwischen  $\pm 1$ . Wir schreiben das kompakt als

$$\sin x = \sum_{k=0}^{\infty} \frac{(-1)^k}{(2k+1)!} x^{2k+1} = x - \frac{x^3}{6} + \dots \quad \cos x = \sum_{k=0}^{\infty} \frac{(-1)^k}{2k!} x^{2k} = 1 - \frac{x^2}{2} + \dots \quad (1.25)$$

Der Logarithmus ist für  $x \leq 0$  nicht definiert da der Wertebereich der Exponentialfunktion nur nichtnegative Zahlen umfasst. Was man aber schreiben kann ist

$$\log(1+x) = - \sum_{k=1}^{\infty} \frac{(-x)^k}{k} \quad |x| < 1 \quad (1.26)$$

### 1.3.1 Die komplexe Exponentialfunktion

Die komplexen Zahlen haben einige von Ihnen in der Schule schon kennengelernt und wurden im Vorkurs behandelt. Es handelt sich um die Erweiterung der reellen Zahlen um die Dimension der imaginären Einheit  $i$  mit  $i^2 = -1$ . Die Rechenregeln finden Sie im Vorkursskript. Wir können leicht sehen, dass  $i^0 = 1$ ,  $i^1 = i$ ,  $i^2 = -1$ ,  $i^3 = -i$ ,  $i^4 = 1$  etc. Damit haben wir eine ähnliche „Uhr“ wie bei der Differenziation von Sinus und Kosinus und können schreiben für natürliche  $k$   $i^{2k} = (-1)^k$  und  $i^{2k+1} = i(-1)^k$ . Damit können wir die Ausdrücke Gl. (1.24) und (1.25) verbinden zu

$$e^{ix} = \sum_{n=0}^{\infty} \frac{i^n x^n}{n!} = \sum_{k=0}^{\infty} \frac{(-1)^k x^{2k}}{2k!} + i \sum_{k=0}^{\infty} \frac{(-1)^k x^{2k+1}}{(2k+1)!} = \cos x + i \sin x \quad (1.27)$$

Dieser Zusammenhang zwischen trigonometrischer und Exponentialfunktion heißt auch Eulergleichung und ist ein wichtiges strukturierendes Element der Analysis. Im Fall einer komplexen Exponentialfunktion können wir schreiben

$$e^z = e^{\operatorname{Re}z} e^{i\operatorname{Im}z} = e^{\operatorname{Re}z} (\cos(\operatorname{Im}z) + i \sin(\operatorname{Im}z)) \quad (1.28)$$

Dies motiviert die Polardarstellung der komplexen Zahlen (Zeichnung siehe Vorlesung) und die Definitionen

$$|z| = \sqrt{(\operatorname{Re}z)^2 + (\operatorname{Im}z)^2} = \sqrt{z\bar{z}} \quad \operatorname{arg}z = \arctan \frac{\operatorname{Im}z}{\operatorname{Re}z} \quad z = |z|e^{i\operatorname{arg}z}. \quad (1.29)$$

Wir können Gleichung (1.27) umdrehen zu

$$\cos x = \frac{e^{ix} + e^{-ix}}{2} \quad \sin x = \frac{e^{ix} - e^{-ix}}{2i}. \quad (1.30)$$

Wir können so die Rechenregeln für die Exponentialfunktion benutzen, um Additionstheoreme für die trigonometrischen Funktionen zu beweisen. Kostproben sind zB

$$\begin{aligned} \cos(x+y) &= \frac{1}{2} (e^{ix}e^{iy} + e^{-ix}e^{-iy}) \\ &= \frac{1}{2} [( \cos x + i \sin x)(\cos y + i \sin y) + (\cos x - i \sin x)(\cos y - i \sin y)] \\ &= \cos x \cos y - \sin x \sin y \end{aligned}$$

und

$$\begin{aligned} \cos x + \cos y &= \frac{1}{2} [e^{ix} + e^{iy} + e^{-ix} + e^{-iy}] \\ &= \frac{1}{2} [e^{i\frac{x+y}{2}} e^{i\frac{x-y}{2}} + e^{i\frac{x+y}{2}} e^{i\frac{-x+y}{2}} + c.c.] \\ &= e^{i\frac{x+y}{2}} \cos \frac{x-y}{2} + c.c. = 2 \cos \frac{x+y}{2} \cos \frac{x-y}{2} \end{aligned}$$

## 1.4 Integration

Integration ist eine entscheidende Technik der Mathematikanwendung in der Physik. Sie beinhaltet das Aufstellen und Lösen von Integralen. Manche sagen, Ableiten sein Handwerk und Integration sein Kunst - eine Kunst, die durchaus beherrschbar ist. Wir machen aber zunächst eine Anmerkung zum Integralbegriff per se und beschäftigen uns dann mit eindimensionalen Integralen - allgemeinere Integrale werden uns über das Semester hinaus beschäftigen.

### 1.4.1 Was ist ein Integral?

Integrale tauchen als kontinuierliche Verallgemeinerungen von Summen auf. Eine Summe benötigt einen Bereich, über den die Summe läuft und (reelle oder komplexe) Größen, über die summiert wird. Betrachten wir z.B. eine Massenverteilung  $\rho(x, y)$  auf einer Fläche  $A$  und versuchen, die Gesamtmasse zu schätzen. Dabei kann  $A$  sehr kompliziert sein, d.h. wir wissen evtl. gar nicht a priori den Flächeninhalt. Dazu können wir z.B. die Fläche in Rechtecke der Fläche  $\delta x \delta y$  (Zeichnung in der Vorlesung) ein und werten in jedem mit  $(n, m)$  markierten

Rechtecke in der Mitte aus: Wenn der Mittelpunkt in  $A$  liegt zählen wir den Beitrag mit (setzen  $A_{nm} = 1$  und sonst  $= 0$ ) und gewichten ihn mit  $\rho_{nm}$  - dem Wert der Dichte in diesem Mittelpunkt. Damit ist eine Näherung an die Gesamtmasse die Summe der Beiträge der Masse  $\rho_{nm}\delta x\delta y$

$$M_{\delta x\delta y} = \sum_{(n,m)} A_{nm}\rho_{nm}\delta x\delta y \quad (1.31)$$

Dies kann im Zweifelsfall schon eine ganz ordentliche Abschätzung an die Gesamtmasse geben, allerdings entstehen Fehler dadurch, dass die Massendichte über den Körper hinweg schwanken kann. Dies wird dadurch reduziert, dass man  $\delta x$  und  $\delta y$  möglichst klein wählt, die Fläche also möglichst fein rastert, wodurch automatisch die Zahl der Summanden proportional steigt. Wir haben im Grenzfall die präzise Gesamtmasse

$$M = \lim_{\delta x, \delta y \rightarrow 0} M_{\delta x\delta y} \equiv \int_A \rho(x, y) dx dy \quad (1.32)$$

Ein Integral mit zwei Integrationsvariablen haben Sie vermutlich in der Schule nicht gesehen - wir wollen hier erst einmal einen Integralbegriff in einer Diskussion klären, in der Rechenstechniken noch nicht im Mittelpunkt stehen.

### 1.4.2 Integration über reelle Intervalle

Jetzt begeben wir uns auf etwas vertrauterem Terrain - wir integrieren über ein eindimensionales Intervall  $[a, b]$ . Folgend unserer vorherigen Diskussion definieren wir (veranschaulichene Zeichnung in der Vorlesung)

$$F(b) = \int_a^b f(x) dx \equiv \lim_{n \rightarrow \infty} \frac{b-a}{n} \sum_{k=1}^n f\left(a + (b-a) \frac{k-1/2}{n}\right) \quad (1.33)$$

Hier haben wir die Punkte der Funktionsauswertung jeweils in die Mitte der  $n$  Stück Intervalle der Länge  $\frac{b-a}{n}$  gelegt. Die Summe auf der rechten Seite nennen wir auch *Riemannsches Summe*. Diesen Grenzwert von Hand auszuwerten ist meistens sehr mühevoll. Bei der Auswertung hilft uns der *Hauptsatz der Integral- und Differenzialrechnung*: Dazu berechnen wir

$$\frac{dF(b)}{db} = \lim_{h \rightarrow 0} \left[ \int_a^{b+h} f(x) dx - \int_a^b f(x) dx \right] = f(b) \quad (1.34)$$

wobei wir im letzten Schritt ausnutzen, dass wir die beiden involvierten Grenzübergänge in beliebiger Reihenfolge und auf beliebigem Weg ausführen können. Das Integral einer Funktion ist also eine Stammfunktion des Integranden. Die Diskussion bisher hat die untere Integrationsgrenze nicht tangiert. Da Stammfunktionen bis auf eine additive Konstante eindeutig sind, ist damit für eine beliebige Stammfunktion  $F$  immer

$$\int_a^b f(x) dx = F(b) - F(a). \quad (1.35)$$



Beispiele für die Stammfunktionen elementarer Funktionen finden sich im Vorkurs oder lassen sich aus dem Ableitungskapitel dieses Skriptes leicht konstruieren.

### 1.4.3 Integrationsregeln und -methoden

Wie schon im Ableitungskapitel listen wir hier die wichtigsten Integrationsregeln zusammen mit Beispielen. Details finden Sie in Ihrem Schulstoff oder im Skript des Vorkurses.

- Integration ist linear

$$\int (\lambda f(x) + \mu g(x)) dx = \lambda \int f(x) dx + \mu \int g(x) dx \quad (1.36)$$

- Integrationsgrenzen

$$\int_a^b f(x) dx = - \int_b^a f(x) dx \quad \int_a^b f(x) dx + \int_b^c f(x) dx = \int_a^c f(x) dx \quad (1.37)$$

- Substitution (hier muss  $g$  auf  $[a, b]$  umkehrbar sein)

$$\int_a^b f(g(x)) dx = \int_{g(a)}^{g(b)} \frac{f(g)}{g'(g^{-1}(g))} dg \quad (1.38)$$

- Spezialfall: Logarithmische Integration

$$\int_a^b \frac{f'(x)}{f(x)} dx = \log f(x) \quad (1.39)$$

- Produktregel / Partielle Integration

$$\int_a^b f'(x)g(x) dx = f(x)g(x)|_a^b - \int_a^b f(x)g'(x) dx \quad (1.40)$$

Als Beispiel betrachten wir einmal das Gaußintegral. Unter der Vorgabe (Beweis auf einem späteren Übungsblatt) dass

$$\int_{-\infty}^{\infty} e^{-x^2} dx = \sqrt{\pi} \quad (1.41)$$

berechnen wir für  $a > 0$

$$\int_{-\infty}^{\infty} \exp(-ax^2 + bx) dx = \exp\left(\frac{b^2}{4a}\right) \int_{-\infty}^{\infty} \exp\left[-a\left(x - \frac{b}{2a}\right)^2\right] dx \quad (1.42)$$

Wir substituieren  $g = \sqrt{a} \left(x - \frac{b}{2a}\right)$  und finden für dieses Integral

$$\exp\left(\frac{b^2}{4a}\right) \int_{-\infty}^{\infty} \frac{e^{-g^2}}{\sqrt{a}} dg = \sqrt{\frac{\pi}{a}} e^{b^2/4a}. \quad (1.43)$$

Eine etwas komplexere Substitution gibt es für

$$\int_a^b \frac{dx}{1+x^2} \quad (1.44)$$

Wir substituieren  $x = \tan y$  (das ist umgekehrt geschrieben wie in der Formel oben!) und überzeugen uns von

$$1+x^2 = 1+\tan^2 y = \frac{1}{\cos^2 y}$$

$$\frac{dx}{dy} = \frac{1}{\cos^2 y}$$

sodass sich das obige Integral umschreiben lässt zu

$$\int dy = \arctan x \quad (1.45)$$

Als nächstes Beispiel betrachten wir eine partielle Integration

$$\int \log x \, dx = \int 1 \log x \, dx = x \log x - \int dx = x(\log x - 1) \quad (1.46)$$

Sowie

$$\int \sin^2 x \, dx = -\cos x \sin x + \int \cos^2 x \, dx = x - \cos x \sin x - \int \sin^2 x \, dx \quad (1.47)$$

also

$$\int \sin^2 x \, dx = \frac{1}{2}(x - \cos x \sin x). \quad (1.48)$$

#### 1.4.4 Integration in der klassischen Mechanik

Integration tritt in der Physik und auch in der klassischen Mechanik an vielen Stellen auf. Ein Beispiel ist die Rekonstruktion des Ortes aus der Geschwindigkeit und dem Anfangsort: Gegeben  $x(t_0)$  und  $v(t > t_0)$  ist

$$x(t) = x(t_0) + \int_{t_0}^t v(\tau) d\tau. \quad (1.49)$$

Die additive Integrationskonstante  $x(t_0)$  hat eine klare physikalische Bedeutung - die Angabe der Geschwindigkeit allein erlaubt nicht die Rekonstruktion des Ortes, man braucht die Anfangskoordinate. Wenn stattdessen die Beschleunigung

$a(t)$  und eine Anfangsgeschwindigkeit vorgegeben ist, machen wir diesen Schritt zweimal und erhalten

$$x(t) = x(t_0) + v(t_0)(t - t_0) + \int_{t_0}^t d\tau_1 \int_{t_0}^{\tau_1} d\tau_2 a(\tau_2). \quad (1.50)$$

Ein Beispiel das Sie kennen ist die konstante Beschleunigung  $a(\tau_2) = \text{const.}$ . Dann haben wir

$$x(t) = x(t_0) + v(t_0)(t - t_0) + \int_{t_0}^t d\tau_1 a(\tau_1 - t_0) = x(t_0) + v(t_0)(t - t_0) + \frac{a}{2}(t - t_0)^2. \quad (1.51)$$

Das Lösen der Newtonschen Bewegungsgleichung ist hier sehr einfach, weil die Beschleunigung nicht vom Ort abhängt - was man macht, wenn das der Fall ist, lernen wir später im Semester.

Ansonsten sehen Sie Integration auch bei der Berechnung von Massen ausgedehnter Körper, der Trägheit bei Drehbewegungen, Strömungen und vielem anderen mehr. Wir werden mehrfach Integration in neuen Gewand kennenlernen - über Integrationsbereiche die nicht Intervalle sondern Kurven, Flächen, und Volumen sind und als Integranden sehen wir Zahlen und Vektoren.

## Kapitel 2

# Vektorrechnung und Kraftfelder

Vektorrechnung haben Sie in der Schule kennengelernt. Dort haben Sie Vektoren mit Pfeilen bzw. mit Zahlenpaaren / -tripeln gleichgesetzt und damit konnten Sie die Geometrie mit den Mitteln von Algebra und Analysis kombinieren. Allerdings ist natürlich nicht jedes Zahlentripel ein Vektor - und umgekehrt gibt es Vektoren, die zunächst ganz anders aussehen.

### 2.1 Motivation

Viele physikalische Größen (Masse, Temperatur, Druck ...) lassen sich einfach durch Zahlen ausdrücken. Diese nennt man *Skalare*. Manchmal reicht das aber nicht.

Stellen Sie sich vor, Sie vermieten Ihr WG-Zimmer unter, können aber keine persönliche Übergabe machen (nur zu empfehlen wenn Sie die Zwischenmieterin kennen). Jetzt wollen Sie beschreiben, wo verschiedene Dinge dort sind. Eine Angabe wie „das Sofa ist bei 2 Meter“ ist nicht aussagekräftig. Um aussagekräftig zu werden müssen Sie

- zwei Koordinaten angeben
- sagen, was die beiden Koordinaten bedeuten: Sie brauchen Koordinatenachsen (das nennen wir später eine Basis)
- einen Referenzpunkt angeben (z.B. in einem Zimmereck): Wir werden später sehen, dass wir dies für einen affinen Raum brauchen

So können Sie Ihrer Zwischenmieterin sagen, dass das Sofa z.B. bei (141, 141) steht. Das bedeutet, dass man erst 141 cm entlang einer Wand geht und dann 141 cm parallel zur anderen. Damit haben Sie zwei Vektoren addiert.

## 2.2 Standardvektoren

In der Schule haben wir bereits gelernt, Vektoren als Zahlentupel darzustellen. Solche Vektoren stammen aus der Menge (die sich gleich als Vektorraum entpuppen wird)

$$\mathbb{R}^n = \left\{ \vec{r} = \begin{pmatrix} r_1 \\ r_2 \\ \vdots \\ r_n \end{pmatrix} \mid r_1, r_2, \dots, r_n \in \mathbb{R} \right\} \quad (2.1)$$

Wenn die Einträge komplexe Zahlen sind, dann reden wir entsprechend über den  $\mathbb{C}^n$ . Wir können diese Vektoren addieren und mit reellen Zahlen (Skalaren) multiplizieren. Das sieht dann so aus

$$\lambda \vec{q} + \mu \vec{r} = \lambda \begin{pmatrix} q_1 \\ q_2 \\ \vdots \\ q_n \end{pmatrix} + \mu \begin{pmatrix} r_1 \\ r_2 \\ \vdots \\ r_n \end{pmatrix} = \begin{pmatrix} \lambda q_1 + \mu r_1 \\ \lambda q_2 + \mu r_2 \\ \vdots \\ \lambda q_n + \mu r_n \end{pmatrix} \in \mathbb{R}^n. \quad (2.2)$$

Wir werden sehen, dass Vektorräume eigentlich etwas abstrakteres sind, sich aber immer auf solche Zahlentupel abbilden lassen.

## 2.3 Allgemeine Vektorräume

### 2.3.1 Definition

Wir haben oben schon gesehen, dass die Skalare, mit denen wir einen Vektor multiplizieren können, und die Vektoren in enger Beziehung stehen. Der Begriff eines Vektorraums hat also etwas mit der zugehörigen Menge an Skalaren<sup>1</sup> (für uns  $\mathbb{R}$  oder  $\mathbb{C}$ ) zu tun. Wir nennen dann die Menge  $V = \{\vec{v}\}$  einen Vektorraum wenn für eine geeignete Skalarmultiplikation und eine geeignete Summe gilt

1. Für  $\vec{v}, \vec{w} \in V$  ist auch  $\vec{v} + \vec{w} \in V$  (Abgeschlossenheit bezüglich Addition)
2. Für einen beliebigen Skalar  $\lambda$  und  $\vec{v} \in V$  ist auch  $\lambda \vec{v} \in V$
3. Das neutrale Element der Addition  $\vec{0}$  liegt in  $V$ . Es ist gekennzeichnet dadurch dass  $\forall \vec{v} \in V$  gilt  $\vec{v} + \vec{0} = \vec{v}$ .
4. (Eigentlich folgt das aus Punkt 2 mit  $\lambda = -1$ ) : Für  $\vec{v} \in V$  ist auch  $-\vec{v} \in V$  so dass  $\vec{v} + (-\vec{v}) = \vec{0}$  ist

Die zugrundeliegenden Verknüpfungen müssen folgenden Voraussetzungen genügen

---

<sup>1</sup>diese Menge muss bestimmten Anforderungen genügen - sie muss ein *Körper* sein

1. Distributivgesetz für Skalare  $(\mu + \eta) \vec{v} = \mu \vec{v} + \eta \vec{v}$
2. Distributivgesetz für Vektoren  $\mu (\vec{v} + \vec{w}) = \mu \vec{v} + \mu \vec{w}$
3. Assoziativgesetz  $\mu(\nu \vec{v}) = (\mu\nu) \vec{v}$

Das ist alles und recht abstrakt. Das wird sich als Stärke des Vektorraumkonzepts - viele Ergebnisse, die daraus folgen, lassen sich auf recht heterogene mathematische Objekte anwenden.

### 2.3.2 Beispiele für Vektorräume

Wir wollen diese Definitionen durch Beispiele mit Leben erfüllen

#### 2.3.2.1 Der Standardvektorraum

Der  $\mathbb{R}^n$  ist ein Vektorraum mit Körper  $\mathbb{R}$ . Die Rechenregeln ergeben sich aus den Rechenregeln für reelle Zahlen

#### 2.3.2.2 Affine Räume

In der analytischen Geometrie mit ihren Anwendungen in der Physik identifizieren wir Punkte mit ihren Ortsvektoren. Um dies eindeutig zu machen müssen wir (wie in der Motivation oben illustriert) noch einen Ursprung festlegen und die Ortsvektoren hängen von dieser Wahl des Ursprungs ab. Dies können wir mit dem Konzept des *affinen Raumes* einfangen - ein affiner Raum besteht aus Punkten, deren Ort relativ zum Ursprung durch einen Ortsvektor bestimmt wird, und die Menge aller Ortsvektoren ist ein Vektorraum.

Als Beispiel betrachten wir eine Ebene im dreidimensionalen Raum. Wenn diese Ebene den Ursprung enthält können wir sie mit dem  $\mathbb{R}^2$  identifizieren. Wenn sie den Ursprung nicht enthält, erfüllen die Ortsvektoren aber die Vektorraumeigenschaft nicht mehr. Wenn wir aber einen Aufpunkt in der Ebene definieren und nur die Vektoren von diesem Aufpunkt zu den Punkten der Ebene betrachten, dann haben wir wieder einen zu  $\mathbb{R}^2$  gleichwertigen Vektorraum. Einen solchen Vektorraum relativ zu einem Aufpunkt nennen wir einen affinen Raum. Wenn der Vektorraum ein  $\mathbb{R}^n$  ist. Den Punktraum der zum  $\mathbb{R}^n$  heißt auch euklidischer Raum  $\mathbb{E}^n$ .

#### 2.3.2.3 Funktionenräume

Die Menge aller Funktionen von einem Intervall  $I$  auf  $\mathbb{R}$  ist ein Vektorraum über  $\mathbb{R}$ . Die Linearkombination von zwei Funktionen ist wieder eine Funktion und  $f(x) = 0$  dient als neutrales Element. Zu gegebenem  $g$  ist  $-g$  das inverse und die Rechenregeln sind von den reellen Zahlen geerbt.

Ein weiterer Funktionenraum kann man konstruieren, wenn man Funktionen von  $[0, 1] \rightarrow \mathbb{R}$  betrachtet, sich ein  $n \in \mathbb{N}$  vorgibt und nur Funktionen der Form

(Zeichnung in der Vorlesung)

$$f(x) = f_k, \quad \frac{k-1}{n} \leq x < \frac{k}{n} \quad k = 1 \dots n \quad (2.3)$$

betrachten - Funktionen die intervallweise konstant sind. Dieser ist sogar äquivalent zum  $\mathbb{R}^n$ . Dieser Raum nähert sich dem Raum aller Funktionen auf  $[0, 1]$  an, wenn wir  $n \rightarrow \infty$  gehen lassen.

Ein alternatives Beispiel für einen Funktionenraum ist der Raum aller Polynome vom Grad  $\leq n$ . Auch dieser ist wie gewünscht abgeschlossen.

### 2.3.2.4 Komplexe Zahlen

Wenn wir die komplexen Zahlen einfach als Vektorraum betrachten können wir uns ebenfalls überzeugen, dass die Vektorraumaxiome erfüllt sind, wobei wir als Skalare die reellen oder die komplexen Zahlen verwenden können.

## 2.4 Basis und Dimension

Wir sind bereits daran gewöhnt, den  $\mathbb{R}^n$  als  $n$ -dimensionalen Vektorraum zu bezeichnen. Was bedeutet dieser Begriff in abstrakteren Räumen?

### 2.4.1 Untervektorräume

Wir betrachten einen Vektorraum  $V$  und eine Menge von Vektoren darin  $S \subset V$ ,  $S = \{\vec{v}_1, \dots, \vec{v}_n\}$ . Ferner definieren wir den Spann von  $S$  als Menge

$$\text{span}(S) = \left\{ \sum_{i=1}^n \lambda_i \vec{v}_i \mid \lambda_1 \dots \lambda_n \in \mathbb{R} \right\}. \quad (2.4)$$

Wir können uns davon überzeugen, dass  $\text{span}(S)$  ein Vektorraum ist. Die Eigenschaften der skalaren Multiplikation sind erfüllt, da diese ja vom Vektorraum  $V$  geerbt waren. Wir benötigen nur noch die Abgeschlossenheit, die sich aus der Konstruktion ergibt - wir haben ja gerade alle Linearkombinationen die wir erzeugen können bereits mitgenommen. Konkret: Wenn

$$\vec{w}_{1/2} = \sum \lambda_{i,1/2} \vec{v}_i \in \text{span}(S) \quad (2.5)$$

ist, dann ist auch

$$\mu \vec{w}_1 + \eta \vec{w}_2 = \sum (\mu \lambda_{i,1} + \eta \lambda_{i,2}) \vec{v}_i \in \text{span}(S). \quad (2.6)$$

Klarerweise ist  $\text{span}(S) \subseteq V$  und wir nennen eine Teilmenge eines Vektorraums die selber ein Vektorraum ist einen *Unter(vektor)raum* von  $V$ .

Wir versuchen jetzt  $S$  und den durch  $S$  erzeugten Untervektorraum genauer zu betrachten und in Beziehung zu setzen. Dazu drehen wir die Frage herum: Gegeben ein Vektorraum  $U$  und eine Menge  $S$  so dass  $U = \text{span}(S)$ : Ist dann  $S$  bereits „optimal“ im Sinne davon, eine möglichst kleine Menge an Vektoren zu sein, die  $U$  aufspannen.

### 2.4.2 Lineare Unabhängigkeit

Als ersten Begriff benötigen wir den der *linearen Unabhängigkeit*. Die Menge  $S$  heißt *linear unabhängig*, wenn folgender logischer Schluss gilt (Definition A)

$$\sum_{i=1}^n \lambda_i \vec{v}_i = \vec{0} \Rightarrow \lambda_i = 0, \quad i = 1 \dots n. \quad (2.7)$$

Anders gesagt: Die einzige Möglichkeit den Nullvektor als Linearkombination zu konstruieren ist es, alle Element mit der Null zu gewichten (das geht natürlich immer). Diese Definition ist gleichwertig zu der Aussage, dass keiner der Vektoren  $\vec{v}_i \in S$  als Linearkombination der anderen geschrieben werden kann (Definition B). Wir wollen dies beweisen:

Zunächst mal ist die Sache trivial, wenn  $\vec{0} \in S$  ist - dann ist die Menge nach beiden Kriterien linear abhängig. Wir beschränken uns auf den Fall  $\vec{0} \notin S$

Wenn A gilt, dann existiert eine Summe mit mindestens zwei nichtverschwindenden Termen  $j$  und  $k$  (also  $\lambda_j, \lambda_k \neq 0$ )

$$\sum_{i=1}^n \lambda_i \vec{v}_i = \vec{0}. \quad (2.8)$$

Wir lösen auf indem wir den Term  $j$  heraushaben und finden

$$\vec{v}_j = \sum_{i \neq j} \left( -\frac{\lambda_i}{\lambda_j} \right) \vec{v}_i \quad (2.9)$$

und haben damit Definition B erfüllt.

Umgekehrt wenn B gilt, dann ex. ein  $j$  so dass

$$\vec{v}_j = \sum_{i \neq j} \lambda_i \vec{v}_i \quad (2.10)$$

Wenn wir  $\lambda_j = -1$  setzen ist dies gleichwertig zu

$$\sum_{i=1}^n \lambda_i \vec{v}_i = \vec{0}. \quad (2.11)$$

und zumindest  $\lambda_j \neq 1$ .

Als Beispiel betrachten wir im  $\mathbb{R}^2$  die Menge

$$S = \left\{ \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 2 \end{pmatrix}, \begin{pmatrix} -1 \\ -1 \end{pmatrix} \right\} \quad (2.12)$$

diese ist linear abhängig, es ist

$$\begin{pmatrix} 1 \\ 0 \end{pmatrix} + \begin{pmatrix} 1 \\ 2 \end{pmatrix} - 2 \begin{pmatrix} -1 \\ -1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \quad (2.13)$$

Wir sind nicht überrascht, da wir drei Vektoren in einem zweidimensionalen Raum haben - das wollen wir jetzt formalisieren.



### 2.4.3 Vollständigkeit und Basis

Umgekehrt betrachtet nennen wir eine Menge  $S$  vollständig im Vektorraum  $U$  wenn  $\text{span}S = U$  ist.

Wir nennen insbesondere  $S$  eine Basis von  $U$ , wenn  $S$  linear unabhängig ist und vollständig ist. Dies ermöglicht folgendes Konzept:

Gegeben ein Vektor  $\vec{u} \in U$  folgt aus der Vollständigkeit, dass wir immer Koeffizienten  $\lambda_i$  finden können so, dass

$$\vec{u} = \sum_i \lambda_i \vec{v}_i. \quad (2.14)$$

Aus der linearen Unabhängigkeit folgt, dass diese Koeffizienten eindeutig sind: Wenn wir weitere Koeffizienten  $\mu_i$ , die das gleiche leisten, finden, dann ist

$$\vec{0} = \vec{u} - \vec{u} = \sum_i (\lambda_i - \mu_i) \vec{v}_i \quad (2.15)$$

und daraus folgt dann für alle  $i$  dass  $\lambda_i = \mu_i$ .

Die Basis eines Vektorraumes ist nicht eindeutig bestimmt - wenn Sie an eine Ebene denken dann formen jegliche zwei nicht-kollineare Vektoren in der Ebene eine Basis. Allerdings ist die Zahl der Basisvektoren eindeutig bestimmt, dies nennen wir die Dimension  $\dim U$  des Vektorraumes.

Beide Eigenschaften gemeinsam erlauben uns eine wichtige Identifikation zwischen den beiden verschiedenen Vektorbegriffen, die wir haben: Wenn wir einmal eine Basis  $S$  des  $n$ -dimensionalen Vektorraumes  $U$  festgelegt haben, dann charakterisieren die  $\lambda_i$  jeden Vektor  $\vec{v}$  eindeutig. Wir können daraus einen Vektor im  $\mathbb{R}^n$  basteln

$$\vec{v} \equiv (\lambda_i)_{i=1\dots n} = \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_i \\ \vdots \\ \lambda_n \end{pmatrix}. \quad (2.16)$$

Wenn wir Linearkombinationen von zwei solchen Basisdarstellungen bilden, dann erhalten wir die Basisdarstellung der Linearkombination: Mit der Notation von Gl. (2.5) ist

$$\vec{v}_1 + \vec{v}_2 \equiv \begin{pmatrix} \mu\lambda_{1,1} + \eta\lambda_{1,2} \\ \vdots \\ \mu\lambda_{i,1} + \eta\lambda_{i,2} \\ \vdots \\ \mu\lambda_{n,1} + \eta\lambda_{n,2} \end{pmatrix} \quad (2.17)$$

also das Gleiche, als hätten wir beide Darstellungen addiert. Wir haben, wenn wir einmal eine Basis festgelegt haben, also eine eineindeutige Korrespondenz zwischen Vektoren auf  $U$  und der Basisdarstellung. Diese erlaubt es und, Rechnungen für beliebige Vektoren in  $U$  durch Operationen im  $\mathbb{R}^n$  durchzuführen. Der Name dieser wichtigen Abbildung ist *kanonischer Isomorphismus*.

### 2.4.4 Beispiele

Neben seiner Eigenschaft als universelle Darstellung von Vektoren ist der  $\mathbb{R}^n$  selber ein Vektorraum. Grundsätzlich ist jede linear unabhängige Menge von  $n$  Vektoren im  $\mathbb{R}^n$  eine Basis (die oben genannte Eigenschaft der Vollständigkeit wird erst bei unendlicher Dimension so richtig interessant). Die *Standardbasis* des  $\mathbb{R}^n$  wird aber gebildet von den Vektoren

$$\hat{e}_i = (\delta_{ij})_j = \begin{pmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{pmatrix} \leftarrow \text{Zeile } i \quad (2.18)$$

Im ersten Schritt haben wir zum Zweck der kompakten Notation das *Kronecker-delta* eingeführt

$$\delta_{ij} = \begin{cases} 1 & i = j \\ 0 & \text{sonst} \end{cases} \quad (2.19)$$

Für den oben eingeführten Raum intervallweise Konstanter Funktionen ist die Menge  $H = \{H_k(x)\}$  mit

$$H_k(x) = \begin{cases} 1 & \frac{k-1}{n} \leq x < \frac{k}{n} \\ 0 & \text{sonst} \end{cases} \quad (2.20)$$

eine geeignete Basis.

Für den Raum aller Polynome vom Grad  $\leq n$  ist die Menge

$$\{x^k, k = 1 \dots n\} \quad (2.21)$$

eine geeignete Basis.

### 2.4.5 Anwendung: Kräftezerlegung

In der Mechanik sind Sie oft mit dem Problem der Kräftezerlegung konfrontiert: Sie haben eine gegebene Kraft  $\vec{F}$  die sie, um ihre physikalische Wirkung zu ergründen, in zwei Komponenten entlang vorgegebener Richtungen  $\vec{v}_i$  zerlegen möchten. Sie suchen also die Basisdarstellung

$$\vec{F} = \sum_{i=1}^n F_i \vec{v}_i \quad (2.22)$$

Um die Garantie zu haben, dass dies immer funktioniert, müssen Sie zunächst einmal sicher gehen, dass die  $\{\vec{v}_i\}$  eine Basis bilden. Zur Bestimmung der Komponenten  $F_i$  müssen Sie dann ein lineares Gleichungssystem lösen. Wir betrachten das (etwas konstruierte - dazu später) Beispiel, dass Sie eine Kraft

$$\vec{F} = \begin{pmatrix} 0 \\ F \end{pmatrix} \quad (2.23)$$

zerlegen wollen entlang den Basisvektoren

$$\vec{v}_1 = \begin{pmatrix} \cos \alpha \\ \sin \alpha \end{pmatrix} \quad \vec{v}_2 = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad (2.24)$$

Wir müssen also lösen

$$\begin{aligned} F_1 \cos \alpha + F_2 &= 0 \\ F_2 \sin \alpha &= F \end{aligned}$$

und erhalten  $F_2 = F/\sin \alpha$  und  $F_1 = -F \tan \alpha$ .

Was hier noch willkürlich ist, ist die Längenbestimmung von Vektoren. Außerdem geht diese Übung einfacher, wenn die Basisvektoren aufeinander senkrecht stehen. Die Begriffe von Länge und Orthogonalität müssen wir aber erst noch sauber bilden.

## 2.5 Skalarprodukt, Länge, Euklidische Geometrie

Sie haben in der Schule schon Längenbestimmung in der Vektorrechnung bis zum  $\mathbb{R}^3$  kennengelernt - wir wollen das jetzt auf eine breitere Basis stellen. Wir haben ja bisher nicht gelernt, wie man zwei Vektoren multipliziert und was das überhaupt bedeutet.

### 2.5.1 Skalarprodukt in Vektorräumen über $\mathbb{R}$

Das Skalarprodukt in einem Vektorraum  $V$  über  $\mathbb{R}$  ist eine Abbildung  $\langle \cdot, \cdot \rangle : V \times V \rightarrow \mathbb{R}$  (also eine Abbildung von Paaren von Vektoren auf reelle Zahlen) mit den Eigenschaften

1. Die Abbildung ist bilinear, d.h. für beliebige Vektoren und Skalare gilt  $\langle \vec{u}, \lambda \vec{v} + \eta \vec{w} \rangle = \lambda \langle \vec{u}, \vec{v} \rangle + \eta \langle \vec{u}, \vec{w} \rangle$  und  $\langle \lambda \vec{v} + \eta \vec{w}, \vec{u} \rangle = \lambda \langle \vec{v}, \vec{u} \rangle + \eta \langle \vec{w}, \vec{u} \rangle$
2. Die Abbildung ist symmetrisch,  $\langle \vec{u}, \vec{v} \rangle = \langle \vec{v}, \vec{u} \rangle$
3. Die Abbildung ist positiv definit, d.h. für alle  $\vec{v} \neq \vec{0}$  ist  $\langle \vec{v}, \vec{v} \rangle > 0$ .

Auf Räumen von nichtsingulären stetigen Funktionen von einem Intervall  $I$  auf  $\mathbb{R}$  ist folgendes ein Skalarprodukt

$$\langle f, g \rangle = \int_I dx f(x)g(x) \quad (2.25)$$

Die erste Eigenschaft folgt aus der Linearität der Integration, die zweite aus der Kommutativität der Multiplikation, die dritte aus der Betrachtung von

$$\langle f, f \rangle = \int_I dx f^2(x) \quad (2.26)$$

Man kann sich (Vorlesung) davon überzeugen, dass dies bei stetigen Funktionen nur dann  $= 0$  ist, wenn  $f \equiv 0$  ist und sonst  $> 0$ .

Eine grundlegende Eigenschaft des Skalarprodukts ist die Cauchy-Schwarzsche Ungleichung (CSU)

$$|\langle \vec{v}, \vec{w} \rangle| \leq |\vec{v}| |\vec{w}| \quad (2.27)$$

wobei wir sinnvollerweise  $|\vec{v}| = \sqrt{\langle \vec{v}, \vec{v} \rangle}$  gesetzt haben.

Zum Beweis definieren wir die Zahl

$$a = \frac{\langle \vec{v}, \vec{w} \rangle}{|\vec{w}|^2}. \quad (2.28)$$

Aus der Definition des Skalarproduktes wissen wir

$$\begin{aligned} 0 &\leq \langle \vec{v} - a\vec{w}, \vec{v} - a\vec{w} \rangle = \langle \vec{v}, \vec{v} \rangle - 2a \langle \vec{v}, \vec{w} \rangle + a^2 \langle \vec{w}, \vec{w} \rangle \\ &= |\vec{v}|^2 - \frac{|\langle \vec{v}, \vec{w} \rangle|^2}{|\vec{w}|^2}. \end{aligned}$$

Auflösen und Wurzelziehen ergibt die Behauptung.

### 2.5.2 Das Standardskalarprodukt im $\mathbb{R}^n$

Im Vektorraum  $\mathbb{R}^n$  haben Sie in der Schule schon (zumindest für  $n \leq 3$ ) ein Skalarprodukt kennengelernt. Dieses schreiben wir als

$$\langle \vec{v}, \vec{u} \rangle \equiv \vec{v} \cdot \vec{u} = \sum_{i=1}^n v_i u_i \quad (2.29)$$

Wie schon oben erwähnt können wir mittels des Skalarprodukts eine Länge (in allgemeinen Vektorräumen spricht man von einer *Metrik*) definieren durch

$$u \equiv |\vec{u}| = \sqrt{\vec{u} \cdot \vec{u}} = \sqrt{\sum_{i=1}^n u_i^2} \quad (2.30)$$

was der normalen Länge im Anschauungsraum entspricht. Damit können wir der CSU eine Interpretation angeben. Sie sagt dass

$$c \equiv \frac{|\vec{v} \cdot \vec{u}|}{|\vec{v}| |\vec{u}|} \leq 1 \quad (2.31)$$

ist. Damit existiert ein  $\alpha$  so, dass

$$\cos \alpha = \frac{\vec{v} \cdot \vec{u}}{uv} \quad (2.32)$$

ist.  $\alpha$  ist der von den Vektoren eingeschlossene Winkel. Diese zunächst einmal willkürliche Definition können wir überprüfen, indem wir sie mit dem Kosinussatz an der Schule in Beziehung setzen. Wir definieren ein Dreieck dessen Seiten die Vektoren  $\vec{u}$ ,  $\vec{v}$  und  $\vec{w} = \vec{v} - \vec{u}$  sind (Zeichnung siehe Vorlesung). Die so interpretierte CSU sagt uns, dass

$$w^2 = |\vec{v} - \vec{u}|^2 = v^2 + u^2 - 2\vec{u} \cdot \vec{v} = v^2 + u^2 - 2uv \cos \alpha \quad (2.33)$$

was genau dem Kosinussatz entspricht.

Ein Beispiel für die Anwendung des Skalarprodukts ist die Berechnung der Arbeit gegen eine konstante Kraft  $\vec{F}$ . Hier benötigen wir bei Verschiebung um  $\vec{r}$  nur die zu  $\vec{F}$  parallele Komponente, also gerade  $W = \vec{F} \cdot \vec{r}$ .

### 2.5.3 Orthogonalität und lineare Unabhängigkeit

Orthogonalität ist nützlich bei der Konstruktion von Basen. Wir überzeugen uns erst davon, dass eine wechselseitig orthogonale Menge von Vektoren

$$M = \{\vec{v}_i | \vec{v}_i \cdot \vec{v}_j = v_i^2 \delta_{ij}, v_i^2 = 0\} \quad (2.34)$$

auch linear unabhängig ist. Dazu nehmen wir an, dass dies nicht der Fall ist, d.h. es existieren  $\lambda_i$  so, dass

$$\sum_i \lambda_i \vec{v}_i = \vec{0} \quad (2.35)$$

wenn wir diese Gleichung skalar mit  $\vec{v}_j$  multiplizieren erhalten wir

$$\vec{v}_j \cdot \vec{0} = 0 = \sum_i \lambda_i \vec{v}_j \cdot \vec{v}_i = \lambda_j v_j^2 \quad (2.36)$$

daraus folgt  $\lambda_j = 0$ . Da dies für alle  $j$  gilt, ist die Menge linear unabhängig.

Andererseits ist eine linear unabhängige Menge nicht zwingend orthogonal, wie das Beispiel

$$\left\{ \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \end{pmatrix} \right\} \quad (2.37)$$

zeigt.

Wenn uns eine Menge linear unabhängiger Vektoren

$$L = \{\vec{w}_i, i = 1 \dots n\} \quad (2.38)$$

haben, so können wir daraus eine orthogonale Menge  $\{\vec{u}_i\}$  konstruieren mit dem rekursiven Gram-Schmidt-Orthogonalisierungsverfahren. Es funktioniert folgendermaßen:

1.  $\vec{u}_1 = \vec{w}_1$
2. Für gegebenes  $n$  setze  $\vec{u}_n = \vec{w}_n - \sum_{k=0}^{n-1} (\vec{w}_n \cdot \vec{u}_{n-1}) \frac{\vec{u}_{n-1}}{u_{n-1}^2}$ . Das heißt also, wir ziehen die Projektion auf die vorherigen Vektoren ab und arbeiten mit dem, was übrig bleibt

Wir können wenn wir einmal dabei sind, die Vektoren auch normieren, also ersetzen

$$\vec{w}_n \rightarrow \frac{\vec{w}_n}{w_n} \quad (2.39)$$

Daraus folge eine wichtige Schlussfolgerung: Aus jeder Basis eines Vektorraumes können wir eine Basis aus orthonormierten Vektoren machen, eine Orthonormalbasis (ONB)

$$\{\hat{e}_i | \hat{e}_i \cdot \hat{e}_j = \delta_{ij}\} \quad (2.40)$$

Orthonormalbasen sind für viele in der Physik auftretende Vektorräume die natürliche Basis, insbesondere weil sich die Basisdarstellung hier einfach berechnen lässt

$$\vec{r} = \sum_{i=1}^n (\vec{r} \cdot \hat{e}_i) \hat{e}_i \quad (2.41)$$

Wenn Sie z.B. in der Mechanik eine Kräftezerlegung bzgl. einer Orthonormalbasis berechnen müssen, müssen Sie nicht wie oben umständlich ein lineares Gleichungssystem lösen, es reicht wenn Sie die gegebene Kraft skalar mit den normierten und orthogonalen Richtungsvektoren multiplizieren.

## 2.6 Das Vektorprodukt

Es gibt auch eine multiplikative Verknüpfung zwischen Vektoren im  $\mathbb{R}^3$  (und in der hier beschriebenen Form auch nur dort!), die wieder einen Vektor liefert

$$\vec{c} = \vec{a} \times \vec{b} \quad (2.42)$$

In der Schule und im Vorkurs haben wir die Definition dieses Vektorproduktes (oder auch Kreuzproduktes) genommen und seine Eigenschaften diskutiert. Es ist empfohlen, sich dies mittels Schulunterlagen oder Skript ins Gedächtnis zu rufen.

### 2.6.1 Geometrische Beschreibung

Wir fordern vom Vektorprodukt zweier Vektoren folgende Eigenschaften

1.  $\vec{a} \times \vec{b} \perp \vec{a}, \vec{b}$  - das Vektorprodukt steht senkrecht auf seinen Komponenten
2. rechte Hand Regel: Das Vektorprodukt bildet mit seinen Faktoren ein Rechtssystem
3.  $|\vec{a} \times \vec{b}| = ab \sin \alpha$ . Der Betrag des Vektorprodukts ist die Fläche des von den Faktoren gebildeten Parallelogramms. Hier ist  $\alpha$  der Winkel zwischen den Vektoren

Ein Beispiel für das Vektorprodukt ist das Drehmoment. Greift die Kraft  $\vec{F}$  mittels Hebel  $\vec{l}$  an einer Achse an, dann wirkt an dieser das Drehmoment  $\vec{N} = \vec{r} \times \vec{F}$

Aus dieser Definition ergeben sich sofort einige Korrolare

1. Aus Flächensatz und Rechter-Hand-Regel folgt dass das Vektorprodukt antisymmetrisch ist  $\vec{a} \times \vec{b} = -\vec{b} \times \vec{a}$
2. Flächen sind additiv, darum folgt aus dem Flächensatz das Distributivgesetz  $\vec{a} \times (\vec{b} + \vec{c}) = \vec{a} \times \vec{b} + \vec{a} \times \vec{c}$

3. Das Vektorprodukt ist nicht assoziativ:  $\vec{a} \times (\vec{b} \times \vec{c})$  ist senkrecht auf  $\vec{a}$  aber  $(\vec{a} \times \vec{b}) \times \vec{c}$  steht senkrecht auf  $\vec{c}$ .

Wir bezeichnen das Vektorprodukt als Vektor und nach allen Vektoreigenschaften, die wir bisher behandelt haben ist das auch richtig. Im für die Physik wichtigen mathematischen Gebiet der Differentialgeometrie fordert man aber noch zusätzlich, dass sich Vektoren unter linearen Abbildungen alle gleich verhalten. Das ist hier nicht der Fall: Bei einer Punktspiegelung am Ursprung geht  $\vec{a} \rightarrow -\vec{a}$  und  $\vec{b} \rightarrow -\vec{b}$  und damit  $\vec{a} \times \vec{b} \rightarrow \vec{a} \times \vec{b}$

## 2.6.2 Algebraische Beschreibung

Wir versuchen jetzt anhand der geometrischen Eigenschaften eine algebraische Darstellung des Vektorproduktes zu finden. Wir gehen komponentenweise vor. Da ein Produkt bilinear in den Faktoren sein muss, können wir ansetzen dass

$$\left(\vec{a} \times \vec{b}\right)_i = \sum_{j,k=1}^3 \epsilon_{ijk} a_j b_k \quad (2.43)$$

ist. Wir benötigen jetzt die Elements von  $\epsilon_{ijk}$ . Klarerweise folgt aus der Antisymmetrie dass  $\epsilon_{ijk} = -\epsilon_{jik}$  ist (und damit  $\epsilon_{ijj} = 0$ ). Um sicherzustellen, dass das Vektorprodukt auch senkrecht auf den Komponenten steht ist auch

$$0 = \vec{a} \cdot (\vec{a} \times \vec{b}) = \sum_{ijk} \epsilon_{ijk} a_i a_j b_k = \sum_k b_k \sum_{ij} \epsilon_{ijk} a_i a_j \quad (2.44)$$

Dies kann nur für alle  $\vec{b}$  und  $\vec{a}$  gelten, wenn auch  $\epsilon_{ijk} = -\epsilon_{jik}$ . Da das Vektorprodukt senkrecht auf  $\vec{b}$  stehen muss folgt auch  $\epsilon_{ijk} = -\epsilon_{kji}$  ist. Damit ist  $\epsilon$  total antisymmetrisch - es ändert sein Vorzeichen unter einem beliebigen Indextausch. Wenn man den Index zwei mal tauscht, dann ändert sich nichts. Dies entspricht zyklischen Permutationen  $\epsilon_{ijk} = \epsilon_{jki} = \epsilon_{kij}$ . Damit ist  $\epsilon$  bis auf einen multiplikativen Faktor festgelegt, z.B.  $\epsilon_{123}$ . Wenn wir z.B.

$$1 = |\hat{e}_z| = |\hat{e}_x \times \hat{e}_y| = \epsilon_{123} \quad (2.45)$$

ausrechnen haben wir das auch geleistet. Das so definierte Symbol nennt man *Levi-Civita Symbol* oder auch  $\epsilon$ -Tensor.

Explizit können wir das Vektorprodukt schreiben als

$$\vec{a} \times \vec{b} = \begin{pmatrix} a_y b_z - a_z b_y \\ a_z b_x - a_x b_z \\ a_x b_y - a_y b_x \end{pmatrix} \quad (2.46)$$

Dies kann man sich mit der in der Vorlesung angeschriebenen „Jägerzaunregel“ merken.

Für das Levi-Civita Symbol können wir, unter Zuhilfenahme auch des Kronecker-symbols

$$\delta_{ij} = \begin{cases} 1 & i = j \\ 0 & \text{sonst} \end{cases} \quad (2.47)$$

zeigen dass

$$\sum_i \epsilon_{ijk} \epsilon_{ilm} = \delta_{jl} \delta_{km} - \delta_{jm} \delta_{kl}. \quad (2.48)$$

Damit können wir folgende Relationen überprüfen

1. Zyklisches Spatprodukt (das ist das Volumen des von  $\vec{a}$ ,  $\vec{b}$ ,  $\vec{c}$  aufgespannten Körpers)  $\vec{a} \cdot (\vec{b} \times \vec{c}) = \vec{b} \cdot (\vec{c} \times \vec{a}) = \vec{c} \cdot (\vec{a} \times \vec{b})$
2. Jacobi-Identität:  $\vec{a} \times (\vec{b} \times \vec{c}) + \vec{b} \times (\vec{c} \times \vec{a}) + \vec{c} \times (\vec{a} \times \vec{b}) = \vec{0}$
3. Grassmann-Formel (dreifaches Vektorprodukt):  $\vec{a} \times (\vec{b} \times \vec{c}) = \vec{b}(\vec{a} \cdot \vec{c}) - \vec{c}(\vec{a} \cdot \vec{b})$
4. Lagrange-Identität:  $(\vec{a} \times \vec{b}) \cdot (\vec{c} \times \vec{d}) = (\vec{a} \cdot \vec{c})(\vec{b} \cdot \vec{d}) - (\vec{a} \cdot \vec{d})(\vec{b} \cdot \vec{c})$



## Kapitel 3

# Grundbegriffe der Vektoranalysis

Die Bahnkurve einer Punktmasse lebt in drei Dimensionen - sie ordnet jeder Zeit einen Ortsvektor zu. Daraus können wir Geschwindigkeit und Beschleunigung berechnen. Wir müssen also Analysis und Vektoren miteinander verbinden.

### 3.1 Kurven im Raum

Mathematisch ist (wobei der „Raum“ hier erst einmal durchaus  $n$  Dimensionen haben darf) eine Funktion

$$\vec{k} : [a, b] \rightarrow \mathbb{R}^n, t \mapsto \vec{k}(t) \quad (3.1)$$

die *Parameterdarstellung* einer Kurve. Wir nennen  $t$  den *Kurvenparameter*. Die Bezeichnung  $t$  suggeriert, dass wir oft die Zeit mit dem Kurvenparameter identifizieren - mathematisch ist das aber nicht zwingend. Die Wertemenge dieser Abbildung

$$K = \left\{ \vec{r} \in \mathbb{R}^n \mid \exists t \in [a, b] : \vec{r} = \vec{k}(t) \right\} \quad (3.2)$$

nennen wir *Kurve*. Sie sehen, dass verschiedene Parameterdarstellungen zur gleichen Kurve führen können.

Diesen wichtigen Unterschied möchten wir am Beispiel eines Kreises in der Ebene illustrieren. Wir betrachten die verkettete Funktion

$$\vec{r}_i(t) = \begin{pmatrix} \cos f_i(t) \\ \sin f_i(t) \end{pmatrix} \quad (3.3)$$

Diese Punkte liegen alle auf dem Einheitskreis und solange  $f_i(t)$  ein Intervall der Länge  $2\pi$  durchläuft wird auch der gesamte Einheitskreis abgedeckt. Damit beschreibt die Wahl

$$f_1(t) = 2\pi t \quad (3.4)$$

eine Parameterisierung in der der Kreis gleichmäßig durchlaufen wird während dies bei

$$f_2(t) = \pi(1 - \cos \pi t) \quad (3.5)$$

ungleichmäßig ist. Die Fälle  $i = 1, 2$  sind also unterschiedliche Parameterisierungen der gleichen Kurve.

## 3.2 Ableitungen von Kurven

Ähnlich wie schon in einer Dimension können wir die Geschwindigkeit als Vektor aus der Bewegung ableiten. Wir bezeichnen

$$\vec{v}(t) = \dot{\vec{r}}(t) = \frac{d\vec{r}}{dt} = \lim_{\Delta t \rightarrow 0} \frac{\vec{r}(t + \Delta t) - \vec{r}(t)}{\Delta t} \quad (3.6)$$

als die Kurvengeschwindigkeit (und wenn  $t$  die Zeit ist schlicht als die Geschwindigkeit). Die Geschwindigkeit zeigt immer tangential an die Bahnkurve. Die zweite Ableitung  $\vec{a}(t) = \dot{\vec{v}}(t) = \ddot{\vec{r}}(t)$  ist die Beschleunigung.

Praktisch wird die Ableitung einfach durch Differenzieren der einzelnen Komponenten ausgerechnet und dafür gelten die Ableitungsregeln, die wir schon aus einer Dimension kennen. Zusammengefasst in Vektornotation haben wir damit

1. Linearität

$$\frac{d}{dt} (\lambda \vec{r}(t) + \mu \vec{s}(t)) = \lambda \dot{\vec{r}}(t) + \mu \dot{\vec{s}}(t) \quad (3.7)$$

2. Produktregel für die Skalarmultiplikation

$$\frac{d}{dt} (a(t)\vec{r}(t)) = \dot{a}(t)\vec{r}(t) + a(t)\dot{\vec{r}}(t) \quad (3.8)$$

3. Produktregel für das Skalarprodukt

$$\frac{d}{dt} (\vec{r}(t) \cdot \vec{s}(t)) = \dot{\vec{r}}(t) \cdot \vec{s}(t) + \vec{r}(t) \cdot \dot{\vec{s}}(t) \quad (3.9)$$

4. Produktregel für das Vektorprodukt

$$\frac{d}{dt} (\vec{r}(t) \times \vec{s}(t)) = \dot{\vec{r}}(t) \times \vec{s}(t) + \vec{r}(t) \times \dot{\vec{s}}(t) \quad (3.10)$$

Als Beispiel betrachten wir die oben angeführten Parameterisierungen einer Kreisbahn

$$\dot{\vec{r}}_i(t) = \begin{pmatrix} -\sin f_i(t) \\ \cos f_i(t) \end{pmatrix} \dot{f}_i(t) \quad (3.11)$$

Für die gleichförmige Bewegung,  $i = 1$  ergibt sich

$$\dot{\vec{r}}_1(t) = 2\pi \begin{pmatrix} -\sin f_1(t) \\ \cos f_1(t) \end{pmatrix} \quad (3.12)$$

und für die ungleichmäßige  $i = 2$  haben wir

$$\dot{\vec{r}}_2(t) = \pi \begin{pmatrix} -\sin f_2(t) \\ \cos f_2(t) \end{pmatrix} \sin \pi t \quad (3.13)$$

Beide diese Geschwindigkeiten zeigen tangential an die Bahnkurve wie es sein muss, aber die Beträge sind unterschiedlich.

Zur Berechnung der Beschleunigen differenzieren wir ein weiteres Mal und erhalten mit der Produkt- und der Kettenregel

$$\begin{aligned} \ddot{\vec{r}}_i(t) &= - \begin{pmatrix} \cos f_i(t) \\ \sin f_i(t) \end{pmatrix} \dot{f}_i^2(t) + \begin{pmatrix} -\sin f_i(t) \\ \cos f_i(t) \end{pmatrix} \ddot{f}_i(t) \\ &= -\vec{r}_i(t) \dot{f}_i^2(t) + \dot{\vec{r}}(t) \frac{\ddot{f}_i(t)}{\dot{f}_i(t)} \end{aligned}$$

Wir sehen also zwei Komponenten: Eine die antiparallel zum Ort zeigt - die Zentripetalbeschleunigung. Sie ist gegeben durch das Quadrat der momentanen Winkelgeschwindigkeit. Darüber hinaus haben die eine tangenziale Beschleunigung. Im Fall der gleichförmigen Kreisbewegung tritt nur erstere auf

$$\ddot{\vec{r}}_1(t) = -4\pi^2 \vec{r}(t) \quad (3.14)$$

Die zweite Bewegung ist nicht gleichförmig, hat also auch eine Tangentialbeschleunigung

$$\ddot{\vec{r}}_2(t) = \pi^2 \sin^2(\pi t) \vec{r}(t) + \pi^3 \cos \pi t \begin{pmatrix} -\sin f_i(t) \\ \cos f_i(t) \end{pmatrix}. \quad (3.15)$$

### 3.3 Kurvenlänge

Gegeben eine Kurve (also die Punktmenge!) interessiert uns, wie lang diese ist - was würde in Kurvenlineal oder die Abtastung mit einem Rad messen? Hier kommt unser Integralbegriff zum Zug!

Betrachten wir eine Parameterisierung der Kurve  $\vec{r}(t)$  und ein Intervall  $[t, t + \Delta t]$ . Wir nähern die Kurvenlänge durch die Länge der Sekante (Zeichnung in der Vorlesung)

$$\Delta L = |\vec{r}(t + \Delta t) - \vec{r}(t)| \quad (3.16)$$

Zur Bestimmung der Kurvenlänge unterteilen wir die Kurvenparameterisierung in Intervalle  $[t_i, t_i + \Delta t]$  summieren wir über diese Sekantenlängen. Dann eliminieren wir den Fehler der Sekantennäherung, indem wir das Intervall  $\Delta t$  gegen null laufen lassen

$$L = \lim_{\Delta t \rightarrow 0} \sum_i \Delta L_i = \lim_{\Delta t \rightarrow 0} \sum_i |\vec{r}(t_i + \Delta t) - \vec{r}(t_i)| \frac{\Delta t}{\Delta t} = \int_a^b |\dot{\vec{r}}(t)| dt. \quad (3.17)$$

Die Kurvenlänge ist also das Integral über die Längen der Ableitungen der Parametrisierung. Wenn wir eine  $K$  als Punktmenge gegeben haben, sollte die

Länge aber nicht von der Parameterisierung abhängen. Das gilt nicht immer - es gilt aber, wenn beide Parameterisierungen die Kurve genau einmal durchlaufen. Solche Parameterisierungen  $\vec{r}_{1/2}(t_{1/2})$  bilden ihre Definitionsmengen  $[a_{1/2}, b_{1/2}]$  bijektiv auf die Punkte der Kurve ab, damit existiert auch eine Bijektion zwischen ihnen. Damit können wir die Substitutionsregel anwenden und schreiben

$$L = \int_{a_1}^{b_1} |\dot{\vec{r}}(t_1)| dt_1 = \int_{a_2}^{b_2} |\dot{\vec{r}}(t_2)| \frac{dt_1}{dt_2} dt_2 \quad (3.18)$$

und wie gewünscht hängt die Länge nicht von der Parameterisierung ab.

Wir überprüfen das mit der Berechnung der Länge des Kreisbogens. Mit der ersten Parameterisierung ist

$$L = \int_0^1 dt \, 2\pi = 2\pi \quad (3.19)$$

und mit der zweiten

$$L = \int_0^1 \pi \sin \pi t \, dt = -\cos \pi t \Big|_0^1 = 2\pi \quad (3.20)$$

Als Bogenlänge bezeichnen wir das Integral als Funktion der obersten Grenze

$$s(u) = \int_a^u dt \, |\dot{\vec{r}}|. \quad (3.21)$$

$s$  ist als Integral über eine nichtnegative Größe eine wachsende Funktion von  $u$  und  $s(b) = L$ . Damit ist auch  $s$  eine Parameterisierung - genannt die natürliche Parameterisierung - der Kurve. Ihr Definitionsbereich ist  $[0, L]$ . Wenn wir auf die Bogenlänge substituieren erhalten wir für den Betrag der Kurvengeschwindigkeit

$$\left| \frac{d\vec{r}(s)}{ds} \right| = \left| \frac{d\vec{r}(t)}{dt} \frac{dt}{ds} \right| = \left| \frac{d\vec{r}(t)}{dt} \right| \left| \frac{ds}{dt} \right|^{-1} = 1 \quad (3.22)$$

wobei wir im letzten Schritt Gleichung (3.21) nach der oberen Integrationsgrenze differenziert haben. Damit entspricht die Parameterisierung nach der Bogenlänge gerade dem Durchlaufen der Kurve mit betragsmäßig konstanter Geschwindigkeit.

### 3.4 Vektorfelder und Linienintegrale

Der Begriff des Kraftfeldes ist in der klassischen Mechanik sehr präsent. Damit meinen wir, dass wir an jedem Punkt des Raumes oder eines Raumbereichs eine Kraft anschreiben. Formal ist dies also eine Funktion

$$\vec{F} : \mathbb{R}^3 \rightarrow \mathbb{R}^3 : \vec{r} \rightarrow \vec{F}(\vec{r}) \quad (3.23)$$

Eine solche Funktion nennen wir ein Vektorfeld. Darüber hinaus gibt es Funktionen  $\mathbb{R}^3 \rightarrow \mathbb{R}$  (z.B. die Temperaturverteilung) - diese nennt man Skalarfelder.

Wenn sich jetzt ein Massenpunkt durch ein Kraftfeld am Punkt  $\Delta\vec{r}$  über einen Weg  $\Delta\vec{r}$  bewegt dann leistet dies an ihm eine Arbeit  $\Delta W = \Delta\vec{r} \cdot \vec{F}(\vec{r})$ . Das ist leicht handhabbar wenn  $\vec{F}$  konstant ist.

Wenn  $\vec{F}$  konstant ist können wir die Arbeit, die entlang einer Kurve  $K$  geleistet wird können wir die Arbeit ausrechnen, indem wir  $K$  in kleine Abschnitte  $\Delta\vec{r}_i$  zerlegen und diese Beiträge aufsummieren und dies wiederum beliebig genau machen. Eine Abschätzung bei Zerlegung in  $N$  Abschnitte ist

$$\Delta W = \sum_i \vec{F}(\vec{r}_i) \cdot (\vec{r}_{i+1} - \vec{r}_i) \rightarrow \int_K \vec{F}(\vec{r}) \cdot d\vec{r} = W \quad (3.24)$$

Um diesen Grenzübergang kontrolliert durchführen zu können benötigen wir eine Parameterisierung  $\vec{r}(t) : [0, T] \rightarrow \mathbb{R}^3$  von  $K$ . Mit der können wir den Intervallen Parameterwerte zuordnen schreiben

$$\Delta W = \sum_i \vec{F}(\vec{r}(t_i)) \cdot (\vec{r}(t_{i+1}) - \vec{r}(t_i)) \frac{\Delta t}{\Delta t} \rightarrow \int_0^T \vec{F}(\vec{r}(t)) \cdot \dot{\vec{r}}(t) dt. \quad (3.25)$$

Damit steht nach allem Einsetzen auf der rechten Seite ein ganz normales ein-dimensionales Integral.

Als Beispiel betrachten wir die Kurve  $K$  parameterisiert durch

$$\vec{r}(t) : [0, 1] \rightarrow \left( \begin{array}{c} t \\ \frac{t}{1+t^2} \end{array} \right) \quad (3.26)$$

und das Kraftfeld

$$\vec{F} = \left( \begin{array}{c} y \\ -x \end{array} \right) \quad (3.27)$$

Wir berechnen

$$\begin{aligned} \int_K \vec{F}(\vec{r}) \cdot d\vec{r} &= \int_0^1 \left( \begin{array}{c} \frac{t}{1+t^2} \\ -t \end{array} \right) \cdot \left( \begin{array}{c} 1 \\ \frac{1-t^2}{(1+t^2)^2} \end{array} \right) dt = \int_0^1 \left[ \frac{t}{1+t^2} - \frac{t(1-t^2)}{(1+t^2)^2} \right] dt \\ &= \int_0^1 \frac{2t^3}{(1+t^2)^2} dt = \frac{\ln 4 - 1}{2} \end{aligned}$$

Aus der Definition ergeben sich automatisch zwei wichtige Regeln: Wenn wir eine Kurve  $K$  rückwärts durchlaufen, symbolisiert als  $-K$  ändert das Linienintegral sein Vorzeichen

$$\int_K d\vec{r} \cdot \vec{F}(\vec{r}) = - \int_{-K} d\vec{r} \cdot \vec{F}(\vec{r}) \quad (3.28)$$

Wenn wir zwei Wege (die nicht zusammenhängen müssen, aber können) zu einem Gesamtweg kombinieren,  $K = K_1 \cup K_2$  werden diese einfach addiert

$$\int_K d\vec{r} \cdot \vec{F} = \int_{K_1} d\vec{r} \cdot \vec{F} + \int_{K_2} d\vec{r} \cdot \vec{F} \quad (3.29)$$

Dennoch ist der Weg über die Parameterisierung natürlich immer recht mühevoll. Manchmal geht es aber besser.

### 3.5 Der Gradient eines Skalarfelds

Wir können uns die Frage stellen, ob es zu Vektorfeldern eine Stammfunktion gibt. Die Antwort wird sein, dass das manchmal der Fall ist, aber weit weniger selbstverständlich als in einer Dimension.

Wir betrachten zunächst ein skalares Feld  $\phi : \mathbb{R} \rightarrow \mathbb{R}^3$ ,  $\vec{r} \rightarrow \phi(\vec{r})$ . Die Funktion hängt also von drei Variablen ab, ausgeschrieben als  $\phi(x, y, z)$ . Wenn wir jetzt zwei dieser Variablen festhalten und nur nach der anderen differenzieren, so nennt man dies eine *parzielle Ableitung*, geschrieben mit einem kursiven Symbol

$$\frac{\partial \phi}{\partial x} \quad (3.30)$$

etc. Als Beispiel betrachten wir  $\phi(\vec{r}) = x \exp y + \log z$ . Dort haben wir

$$\frac{\partial \phi}{\partial x} = \exp y \quad \frac{\partial \phi}{\partial y} = x \exp y \quad \frac{\partial \phi}{\partial z} = \frac{1}{z}. \quad (3.31)$$

aus den drei Ableitungen lässt sich ein Vektor formen, der Gradient der Funktion

$$\text{grad} \phi = \begin{pmatrix} \frac{\partial \phi}{\partial x} \\ \frac{\partial \phi}{\partial y} \\ \frac{\partial \phi}{\partial z} \end{pmatrix}. \quad (3.32)$$

Als Abkürzung für die Ausführung aller drei partiellen Ableitungen führen wir einen Vektor aus Ableitungsoperationen ein, den Nablaoperator

$$\nabla \equiv \begin{pmatrix} \frac{\partial}{\partial x} \\ \frac{\partial}{\partial y} \\ \frac{\partial}{\partial z} \end{pmatrix} \quad \nabla \phi = \text{grad} \phi. \quad (3.33)$$

Welche Bedeutung hat dieses aus  $\phi$  erzeugte Vektorfeld? Wir können mit den Partiellen Ableitungen das Verhalten bei Veränderung der drei Veränderlichen beschreiben. Wenn wir nur  $x$  infinitesimal verändern, dann ändert sich der Funktionswert um  $\frac{\partial f}{\partial x} dx$  usw. Wenn wir alle Richtungen erlauben wird daraus das *totale Differenzial*

$$d\phi = \frac{\partial \phi}{\partial x} dx + \frac{\partial \phi}{\partial y} dy + \frac{\partial \phi}{\partial z} dz. \quad (3.34)$$

Eine Anwendung ist (Zeichnung in der Vorlesung) das Konzept der *Richtungsableitung*. Wir wählen einen beliebigen Einheitsvektor  $\hat{e}$  und betrachten die Funktion entlang der durch diesen Vektor definierten Richtung. Dieser Schnitt definiert eine Funktion  $f : \mathbb{R} \rightarrow \mathbb{R} : t \rightarrow f(t) = \phi(\vec{r} + t\hat{e})$ . Die Richtungsableitung von  $\phi$  im Punkt  $\vec{r}$  in Richtung  $\hat{e}$  ist dann definiert als

$$\begin{aligned} f'(0) &= \lim_{t \rightarrow 0} \frac{f(t) - f(0)}{t} = \lim_{t \rightarrow 0} \frac{\phi(\vec{r} + t\hat{e}) - \phi(\vec{r})}{t} \\ &= \frac{\partial \phi}{\partial x} \frac{dx}{dt} + \frac{\partial \phi}{\partial y} \frac{dy}{dt} + \frac{\partial \phi}{\partial z} \frac{dz}{dt} = \text{grad} \phi \cdot \hat{e}. \end{aligned}$$

Damit ist die Richtungsableitung das Skalarprodukt aus Gradient und Richtungs-Einheitsvektor. Wir können dies durch den Winkel  $\alpha$  zwischen dem Gradienten (der ja ein Vektor ist) und der Richtung  $\hat{e}$  ausdrücken als  $f'(0) = |\text{grad}\phi| \cos\alpha$  und folgende Eigenschaften ablesen:

- der Gradientenvektor zeigt in die Richtung der größten Richtungsableitung, also den steilsten Anstieg
- dem entgegengesetzt ist die Richtung des steilsten Abfalls (konsistent damit, dass wir an  $\phi$  eine Tangenzialebene legen
- senkrecht auf den Gradienten ist die Funktion konstant. Im Fall  $n = 2$  steht der Gradient also senkrecht auf den Höhenlinien. Für größeres  $n$  spricht man von Hyperflächen konstanten  $\phi$  - eine Hyperfläche ist ein geometrisches Objekt der Dimension  $n - 1$
- die Länge des Gradienten beschreibt, wie steil der steilste Anstieg ist

### 3.6 Gradientenfelder und Wegintegrale

Es besteht ein wichtiger Zusammenhang zwischen Gradientenfeldern und Wegintegralen. Hier gilt der wichtige Satz: Das Wegintegral über ein Feld  $\vec{F}$  hängt genau dann nur von den Endpunkten, nicht aber von dem sie verbindenden Weg ab, wenn  $\vec{F} = \nabla\phi$  ein Gradientenfeld ist.

Wir zeigen die beiden Richtungen dieser Äquivalenz. Sei zunächst  $\vec{F}$  ein Gradientenfeld. Dann können wir für einen beliebigen Weg  $K$  von  $\vec{r}_1$  nach  $\vec{r}_2$  das Wegintegral schreiben als

$$\int_K d\vec{r} \cdot \vec{F} = \int_K d\vec{r} \cdot \nabla\phi = \int dt \dot{\vec{r}} \cdot \nabla\phi \quad (3.35)$$

wobei wir eine Parameterisierung des Weges angeschaut haben. Wir benutzen die Kettenregel und das totale Differenzial um uns davon zu überzeugen, dass

$$\frac{d}{dt}\phi(\vec{r}(t)) = \sum_{i=1}^3 \frac{\partial\phi}{\partial r_i} \dot{r}_i = \dot{\vec{r}} \cdot \nabla\phi, \quad (3.36)$$

Damit ist

$$\int_K d\vec{r} \cdot \vec{F} = \phi(\vec{r}_2) - \phi(\vec{r}_1) \quad (3.37)$$

was nur von den Endpunkten abhängt, nicht aber vom Weg dazwischen.

Für die andere Richtung gehen wir davon aus, dass das Integral nur von den Endpunkten abhängt. Wir wählen einen festen Aufpunkt  $\vec{r}_0$  und *definieren*

$$\phi(\vec{r}) = \int_K d\vec{r} \cdot \vec{F} \quad (3.38)$$

wobei  $K$  von  $\vec{r}_0$  nach  $\vec{r}$  geht. Wir müssen jetzt zeigen, dass  $\vec{F} = \nabla\phi$  ist. Wir wählen dazu einen parameterisierten Pfad zu  $\vec{r}$  und betrachten  $\phi$  als Funktion des Kurvenparameters

$$\phi(\vec{r}(t)) = \int_0^t d\tau \dot{\vec{r}} \cdot \vec{F}(\vec{r}(\tau)) \quad (3.39)$$

Wir differenzieren auf beiden Seiten nach  $t$  und erhalten

$$\nabla\phi \cdot \dot{\vec{r}}(t) = \dot{\vec{r}}(t) \cdot \vec{F}(\vec{r}(t)) \quad (3.40)$$

wobei wir links wiederum die Kettenregel und rechts den Hauptsatz der Integral- und Differenzialrechnung angewandt haben.

Wir haben also etabliert, dass das Linienintegral genau dann wegunabhängig ist, wenn es sich um ein Gradientenfeld handelt. Äquivalent dazu ist die Aussage, dass das Integral über jeden geschlossenen Weg verschwindet

$$\oint_K d\vec{r} \cdot \vec{F} = 0 \quad (3.41)$$

Dies kann man dadurch sehen, dass man zwei Wege zwischen gleichen Endpunkten  $K_{1/2}$  so zusammensetzt, dass im zusammengesetzten Weg  $K$  der Weg  $K_1$  und der Weg  $K_2$  rückwärts durchlaufen wird - der Weg  $K$  ist dann geschlossen (Zeichnung in der Vorlesung).

Wir haben eine sehr weitreichende Aussage nachgewiesen - die aber ziemlich nutzlos ist. Was tun wir, wenn jemand uns ein Vektorfeld gibt, und wir herausfinden sollen, ob es ein Gradientenfeld ist?

Als erstes haben wir einen Hinweis im Sinn eines *notwendigen* Kriteriums. Wenn  $\vec{F}$  ein Gradientenfeld ist, dann können wir für ein beliebiges Paar von Indizes hernehmen und berechnen

$$\frac{\partial F_j}{\partial r_i} = \frac{\partial^2 \phi}{\partial r_i \partial r_j} = \frac{\partial^2 \phi}{\partial r_j \partial r_i} = \frac{\partial F_i}{\partial r_j} \quad (3.42)$$

Das heißt wir können diese vertauschten Ableitungen testen und wenn die Gleichung *nicht* erfüllt ist, haben wir *kein* Gradientenfeld. Ist dies auch hinreichend? Hier ist Vorsicht geboten:

Betrachten wir als Beispiel das zweidimensionale Vektorfeld

$$\vec{F} = \frac{1}{x^2 + y^2} \begin{pmatrix} -y \\ x \end{pmatrix} \quad (x, y) \neq (0, 0) \quad (3.43)$$

Wir können ableiten

$$\begin{aligned} \frac{\partial F_x}{\partial y} &= \frac{-(x^2 + y^2) + y \cdot 2y}{(x^2 + y^2)^2} = \frac{y^2 - x^2}{(x^2 + y^2)^2} \\ \frac{\partial F_y}{\partial x} &= \frac{(x^2 + y^2) - x \cdot 2x}{(x^2 + y^2)^2} = \frac{y^2 - x^2}{(x^2 + y^2)^2}. \end{aligned}$$



Wenn wir das Feld malen sehen wir aber, dass die Vektoren einen Wirbel bilden, d.h. wenn wir entlang eines Kreises um den Ursprung integrieren haben wir immer das gleiche Vorzeichen im Integranden, das Integral kann also nicht verschwinden. Als Beispiel berechnen wir das Wegintegral um den Einheitskreis, parameterisiert als

$$\vec{r}(\phi) = \begin{pmatrix} \cos \phi \\ \sin \phi \end{pmatrix} \quad \dot{\vec{r}} = \begin{pmatrix} -\sin \phi \\ \cos \phi \end{pmatrix}. \quad (3.44)$$

Wir berechnen das Wegintegral

$$\oint_K d\vec{r} \cdot \vec{F} = \int_0^{2\pi} d\phi \begin{pmatrix} -\sin \phi \\ \cos \phi \end{pmatrix} \cdot \begin{pmatrix} -\sin \phi \\ \cos \phi \end{pmatrix} = 2\pi. \quad (3.45)$$

Warum passiert das? Wir werden später sehen (im Rahmen des Stokesschen Integralsatzes), dass das daran liegt, dass der Definitionsbereich ein Loch im Ursprung hat - wir können den Kreis nicht einfach auf einen Punkt zusammenziehen. Das Gebiet der Mathematik, das sich mit solchen Fragen - was ändert sich aufgrund von solchen Eigenschaften - auseinandersetzt heißt (algebraische) Topologie. Wir werden das nicht behandeln können - aber später in Form des Integralsatzes von Stokes zumindest diesen Aspekt der Lochfreiheit klar identifizieren können.

### 3.7 Flächenintegrale

Wir haben gelernt, über Linien zu integrieren. Unser Integralbegriff war aber zunächst über Flächenintegrale motiviert. Gegeben eine Fläche  $A$ , die wir in Rechtecke  $\delta x \delta y$  unterteilen können und gegeben eine Flächendichte  $f(x, y)$  wollen wir bestimmen, wie groß die Gesamtmasse auf der Fläche ist. Wir summieren über alle Rechtecke, die in der Fläche liegen und machen sie klein

$$F \equiv \int_A d^2r f(x, y) = \lim_{\delta x \delta y \rightarrow 0} \sum_{(i,j) \in A} f(x_i, y_i) \delta x_i \delta y_i. \quad (3.46)$$

Wie gehen wir in der Praxis zur Berechnung eines solchen Integrals vor? Nehmen wir zunächst an,  $A = [x_1, x_2] \times [y_1, y_2]$  ist ein Quadrat, dann integrieren wir

$$F = \int_A d^2r f(x, y) = \int_{y_1}^{y_2} dy \int_{x_1}^{x_2} dx f(x, y). \quad (3.47)$$

Wenn wir eine erste Stammfunktion finden so, dass  $\frac{\partial g}{\partial x}(x, y) = f(x, y)$  ist, dann haben wir

$$F = \int_{y_1}^{y_2} dy (g(x_2, y) - g(x_1, y)) \quad (3.48)$$

und für die  $y$ -Integration müssen wir jetzt noch einmal eine Stammfunktion berechnen. Wir rechnen das Integral also „Streifenweise“ aus.

Wenn wir kein Rechteck haben so können wir (wenn eventuell Stückweise) wenigstens definieren, über welches Intervall  $[x_1(y), x_2(y)]$  es sich ausdehnt. Damit wird aus dem obigen Integral

$$F = \int_{y_1}^{y_2} dy [g(x_2(y), y) - g(x_1(y), y)] \quad (3.49)$$

d.h. die zweite Integrationsvariable taucht auch in den Grenztermen auf.

Als Beispiel betrachten wir als Integrationsbereich das rechtwinklige Dreieck  $A = \{(x, y) | 0 \leq x \leq 1, 0 \leq y \leq x\}$  und  $f(x, y) = xy + e^y$ . Wir integrieren

$$\begin{aligned} \int_A d^2r f(x, y) &= \int_0^1 dx \int_0^x dy (xy + x^2 + e^y) = \int_0^1 dx \left[ \frac{x^3}{2} + x^3 + e^x - 1 \right] \\ &= \frac{3}{8} + e - 1 - 1 = e - \frac{13}{8}. \end{aligned}$$

### 3.8 Der Satz von Stokes

Wir haben in den Beispielen des letzten Abschnittes gesehen, wie wichtig der Rand des Integrationsbereiches ist. Gibt es eine Beziehung zwischen einem Integral über eine Fläche  $A$  und den Funktionswerten auf dem Rand  $\partial A$  die wir ausnutzen können? Die Antwort gibt der Satz von Stokes. Ähnlich wie im Fall von Gradientenfeldern ist die aber an eine Bedingung geknüpft.

Wir konstruieren das Argument in einer Fläche  $A$  der  $xy$  Ebene und kommentieren, wie das auf allgemeine (im Allgemeinen gekrümmte) Flächen verallgemeinert werden kann. Wir betrachten ein Vektorfeld  $\vec{F}$  in der Ebene. Wir berechnen das geschlossene Linienintegral über den Rand der Fläche

$$J = \oint_{\partial A} d\vec{r} \cdot \vec{F} \quad (3.50)$$

Wir zerlegen jetzt die Fläche in kleine Rechtecke der Größe  $\delta x \delta y$  nummeriert durch Indizes  $(i, j)$  und betrachten die geschlossenen Wegintegrale um jedes dieser Rechtecke. Alle inneren Wege - Wege, die in zwei benachbarten Rechtecken durchlaufen werden, werden einmal in jeder Richtung durchlaufen und heben sich weg (Zeichnung in der Vorlesung), nur der Rand von  $A$  bleibt. Damit ist

$$J = \lim_{\delta x \delta y \rightarrow 0} \sum_{ij} \oint_{\delta A_{ij}} d\vec{r} \cdot \vec{F} \quad (3.51)$$

Jetzt berechnen wir den Beitrag eines einzelnen Rechtecks im Limes kleiner

Größe. Dann ist (Zeichnung in der Vorlesung)

$$\begin{aligned}
 \oint_{\delta A_{ij}} d\vec{r} \cdot \vec{F} &= \delta x F_x \left( x + \frac{\delta x}{2}, y \right) + \delta y F_y \left( x + \delta x, y + \frac{\delta y}{2} \right) - \\
 &\quad - \delta x F_x \left( x + \frac{\delta x}{2}, y + \delta y \right) - \delta y F_y \left( x, y + \frac{\delta y}{2} \right) + O(\delta^3) \\
 &= F_x \left( x + \frac{\delta x}{2}, y \right) \delta x + \delta y F_y \left( x, y + \frac{\delta y}{2} \right) + \delta x \delta y \frac{\partial F_y}{\partial x} \left( x, y + \frac{\delta y}{2} \right) \\
 &\quad - \delta x F_x \left( x + \frac{\delta x}{2}, y \right) - \delta x \delta y \frac{\partial F_x}{\partial y} \left( x + \frac{\delta x}{2}, y \right) - \delta y F_y \left( x, y + \frac{\delta y}{2} \right) + O(\delta^3) \\
 &= \delta x \delta y \left[ \frac{\partial F_y}{\partial x} - \frac{\partial F_x}{\partial y} \right] + O(\delta^3)
 \end{aligned}$$

Wir sehen also, dass nur die Terme mit einer weiteren Ableitung auftauchen - diese verhindern, dass sich die Beiträge gegenüberliegender Seiten herausheben. Damit bekommt die Summendarstellung eine andere Interpretation, nämlich

$$\begin{aligned}
 \oint_{\partial A} d\vec{r} \cdot \vec{F} &= J = \lim_{\delta x \delta y \rightarrow 0} \delta x \delta y \left[ \frac{\partial F_y}{\partial x} \left( x, y + \frac{\delta y}{2} \right) - \frac{\partial F_x}{\partial y} \left( x + \frac{\delta x}{2}, y \right) \right] \\
 &= \int_A d^2 r \left[ \frac{\partial F_y}{\partial x} - \frac{\partial F_x}{\partial y} \right].
 \end{aligned}$$

Wir können also aus dem Linienintegral über einen geschlossenen Weg ein Flächenintegral über die eingeschlossene Fläche machen. Wir sehen auch, dass dies ein Zugang zum Kriterium ist, dass es sich bei  $\vec{F}$  um ein Gradientenfeld handelt.

In der Darstellung bisher haben wir uns auf Kurven in der  $xy$  Ebene beschränkt. Ohne Beweis (den lernen Sie bei den Kolleg(inn)en aus der Mathematik) geben wir die allgemeine Form an. Dazu definieren wir die Rotation des Vektorfeldes  $\vec{F}$  als

$$\operatorname{rot} \vec{F} = \nabla \times \vec{F} = \begin{pmatrix} \frac{\partial F_z}{\partial y} - \frac{\partial F_y}{\partial z} \\ \frac{\partial F_x}{\partial z} - \frac{\partial F_z}{\partial x} \\ \frac{\partial F_y}{\partial x} - \frac{\partial F_x}{\partial y} \end{pmatrix}. \quad (3.52)$$

Wir sehen also, diese Ableitung ist wieder ein Vektor. Wir führen außerdem den normierten Normalenvektor  $\hat{n}$  auf die Fläche von Interesse ein. Damit ist die allgemeine Form des obigen Ausdrucks der *Integralsatz von Stokes*

$$\oint_{\partial A} d\vec{r} \cdot \vec{F} = \int_A d^2 r \hat{n} \cdot (\nabla \times \vec{F}) \quad (3.53)$$

Als Beispiel betrachten wir

$$\vec{F} = \frac{1}{2} \begin{pmatrix} -y \\ x \\ 0 \end{pmatrix} \quad \nabla \times \vec{F} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}. \quad (3.54)$$

Damit können wir das Linienintegral über den Einheitskreis einerseits berechnen als

$$\oint_{r=1} d\vec{r} \cdot \vec{F} = \frac{1}{2} \int_0^{2\pi} dt \begin{pmatrix} -\sin t \\ \cos t \\ 0 \end{pmatrix} \cdot \begin{pmatrix} -\sin t \\ \cos t \\ 0 \end{pmatrix} = \pi \quad (3.55)$$

andererseits aber auch die rechte Seite des Satzes von Stokes

$$\int_{r<1} d^2r \hat{e}_z \cdot \hat{e}_z = \pi \quad (3.56)$$

### 3.9 Volumenintegrale und der Satz von Gauß

Ähnlich wie bei Flächenintegralen können wir Volumenintegrale definieren, z.B. zur Berechnung der Gesamtmasse eines dreidimensionalen Objektes. Auch hier können wir das Objekt erst in Schichten und die Schichten dann in Streifen zerlegen, einfach mit einer Schachtelebene mehr als bei Flächenintegralen.

Auch hier interessieren wir uns auf die Möglichkeit, ein Integral über ein Volumen  $V$  durch ein Integral über seine Randfläche  $\partial V$  ausdrücken. Wir haben ja schon Flächenintegrale über die Normalkomponente eines Vektorfeldes

$$J = \oint_{\partial V} d^2r \hat{n} \cdot \vec{F}(\vec{r}) \quad (3.57)$$

Wenn wir uns vorstellen, dass das Vektorfeld z.B. die Flussgeschwindigkeit in einer Flüssigkeit darstellt, ist dieser Integral der *Fluss* des Vektorfeldes  $\vec{F}$  durch die Oberfläche  $\partial V$ . Wir zerlegen das eingeschlossene Volumen wiederum in Quader, numeriert durch  $(i, j, k)$  mit Volumen  $\delta x \delta y \delta z$ . Die Flüsse, die von einem dieser Quader in den anderen fließen heben sich weg ( $\hat{n}$  ist immer nach außen orientiert). Wir definieren damit

$$J = \lim_{\delta x \delta y \delta z \rightarrow 0} \sum_{i,j,k} \oint_{\partial V_{ijk}} d^2r \hat{n} \cdot \vec{F}(\vec{r}). \quad (3.58)$$

Wir berechnen wiederum das Integral das in einem der Summanden auftaucht in niedrigster nichtverschwindender Ordnung, indem wir über die sechs Deckflächen des Quaders summieren

$$\begin{aligned} J_{ijk} &= \delta y \delta z \left( F_x \left( x + \delta x, y + \frac{\delta y}{2}, z + \frac{\delta z}{2} \right) - F_x \left( x + \delta x, y + \frac{\delta y}{2}, z + \frac{\delta z}{2} \right) \right) + \\ &+ \delta x \delta z \left( F_y \left( x + \frac{\delta x}{2}, y + \delta y, z + \frac{\delta z}{2} \right) - F_y \left( x + \delta x, y + \frac{\delta y}{2}, z + \frac{\delta z}{2} \right) \right) + \\ &+ \delta x \delta y \left( F_z \left( x + \frac{\delta x}{2}, y + \frac{\delta y}{2}, z + \delta z \right) - F_z \left( x + \delta x, y + \frac{\delta y}{2}, z + \frac{\delta z}{2} \right) \right) + O(\delta^4) \\ &= \delta x \delta y \delta z \left[ \frac{\partial F_x}{\partial x} + \frac{\partial F_y}{\partial y} + \frac{\partial F_z}{\partial z} \right] + O(\delta^4) \end{aligned}$$

Der Ausdruck in der eckigen Klammer definiert die Divergenz des Vektorfeldes  $\vec{F}$

$$\operatorname{div} \vec{F} \equiv \nabla \cdot \vec{F} = \frac{\partial F_x}{\partial x} + \frac{\partial F_y}{\partial y} + \frac{\partial F_z}{\partial z} \quad (3.59)$$

Damit ergibt sich für das gesamte Integral der Satz von Gauß

$$\oint_{\partial V} d^2 r \hat{n} \cdot \vec{F}(\vec{r}) = \int_V d^3 r \nabla \cdot \vec{F} \quad (3.60)$$

Wir können die Divergenz als Quellendichte des Vektorfeldes interpretieren - wenn mehr Feldlinien aus dem Volumen austreten als eindringen dann befindet sich dort eine Quelle.

Wir betrachten zunächst ein einfaches Beispiel: Wir integrieren über einen Quader die Funktion

$$\vec{F}(\vec{r}) = \begin{pmatrix} xy \\ x^2 - z^2 \\ z^2 + x^2 \end{pmatrix} \quad V = [0, 1]^3 \quad (3.61)$$

Die Divergenz dieses Feldes ist  $\vec{F} = y + 2z$ . Das Volumenintegral

$$\begin{aligned} \int_V d^3 r (y + 2z) &= \int_0^1 dx \int_0^1 dy \int_0^1 dz (y + 2z) = \int_0^1 dx \int_0^1 dy (y + 1) \\ &= \int_0^1 dx \left( \frac{1}{2} + 1 \right) = \frac{3}{2}. \end{aligned}$$

Für das Oberflächenintegral müssen wir über die quadratischen Deckel integrieren

$$\begin{aligned} \int_{\partial V} d^2 r \hat{n} \cdot \vec{F} &= \int_0^1 dy \int_0^1 dz y + \int_0^1 dx \int_0^1 dz (x^2 - z^2 - (x^2 - z^2)) \\ &\quad + \int_0^1 dx \int_0^1 dy (1 + x^2 - x^2) \\ &= \frac{1}{2} + 1 = \frac{3}{2} \end{aligned}$$

### 3.10 Krummlinige Koordinaten

Oftmals in der Physik sind die Koordinaten bzgl. einer Orthonormalbasis nicht die einfachste oder praktischste Methode, einen geometrischen Sachverhalt zu beschreiben. In der Geometrie, in die wir uns jetzt begeben, nennen wir so ein Koordinatensystem auch *kartesisch*. Die Koordinatenachsen in einem kartesischen Koordinatensystem sind Geraden, die Einheitsvektoren sind demzufolge überall die Gleichen. Die Festlegung der Koordinatenachsen ist damit recht willkürlich und einfache Symmetrien wie z.B. die Drehsymmetrie lassen sich in diesem System recht schwer ausdrücken.

### 3.10.1 Polarkoordinaten

Um andere, *krummlinige*, Koordinatensysteme zu erkunden, beschränken wir uns zunächst auf die  $xy$ -Ebene. Ein Koordinatensystem ordnet vernünftigerweise jedem Punkt der Ebene einen eindeutigen Satz von Parametern zu - man kann das als bijektive Abbildung zwischen dem neuen und dem kartesischen Koordinatensystem betrachten. In der Ebene interessieren wir uns für Koordinaten, die die Rotationssymmetrie um den Ursprung einfach zu behandeln machen, Polarkoordinaten. Ein Punkt mit Polarkoordinaten  $(\rho, \phi) \in \mathbb{R}_0^+ \times [0, 2\pi)$  hat die kartesischen Koordinaten

$$x = \rho \cos \phi \quad y = \rho \sin \phi. \quad (3.62)$$

Die Linien mit konstantem  $\rho$  sind damit Kreise um den Ursprung, mit konstantem  $\phi$  Strahlen aus dem Ursprung (Zeichnung: Vorlesung). Wir wollen darum auch die Koordinateneinheitsvektoren entlang dieser Linien zeigen lassen. Dazu erinnern wir uns an das Konzept der Richtungsableitung und definieren

$$\hat{e}_\rho = \frac{\partial \vec{r}}{\partial \rho} = \begin{pmatrix} \cos \phi \\ \sin \phi \end{pmatrix} \quad \hat{e}_\phi = \frac{1}{\rho} \frac{\partial \vec{r}}{\partial \phi} = \begin{pmatrix} -\sin \phi \\ \cos \phi \end{pmatrix} \quad (3.63)$$

wobei wir im zweiten Ausdruck von Hand normiert haben. Wir sehen, dass diese Einheitsvektoren jetzt koordinatenabhängig sind. Insbesondere ist der Ortsvektor eines Punktes immer  $\vec{r} = \rho \hat{e}_\rho$ . Wenn wir jedoch wie bei einer Drehbewegung eine Bahnkurve haben, deren Koordinaten zeitabhängig sind  $(\rho(t), \phi(t))$  dann können wir die Geschwindigkeit und Beschleunigung schreiben als

$$\dot{\vec{r}}(t) = \dot{\rho} \hat{e}_\rho + \rho \dot{\phi} \hat{e}_\phi \quad \ddot{\vec{r}}(t) = (\ddot{\rho} - \rho \dot{\phi}^2) \hat{e}_\rho + (2\dot{\rho} \dot{\phi} + \rho \ddot{\phi}) \hat{e}_\phi \quad (3.64)$$

Wir sehen hier eine Reihe Terme, die Sie in der Experimentalphysik bei Rotationen auch schon gesehen haben:  $\dot{\phi}$  ist die Winkelgeschwindigkeit. Im radialen Term tritt auch bei konstantem Radius  $\rho$  eine Zentripetalbeschleunigung auf.

Wenn wir eine Funktion in der Ebene in Zylinderkoordinaten gegeben haben,  $f(\rho, \phi)$  und über sie integrieren möchten, dann müssen wir das Flächenintegral in Polarkoordinaten ausdrücken. Formal kann das über eine zweidimensionale Substitution geleistet werden. Wir wählen aber einen elementaren, anschaulichen Zugang: In der Definition des Flächenintegrals haben die Funktion über ein Raster an Punkten summiert und über die Fläche jedes Rasterelements summiert. Das können wir auch in Zylinderkoordinaten machen: Das Segment  $[\rho, \rho + d\rho] \times [\phi, \phi + d\phi]$  ist ein Kreisringsegment (Zeichnung in der Vorlesung) mit der Fläche  $\rho d\rho d\phi$ . Kurz schreiben wir  $d^2r = \rho d\rho d\phi$ . Damit ist das Flächenintegral

$$\int_A d^2r f(\rho, \phi) = \int_A d\rho d\phi \rho f(\rho, \phi). \quad (3.65)$$

Wenn wir das Integral in dieser Reihenfolge ausführen, dann integrieren wir zuerst über alle Winkel bei festem Radius und dann über alle Kreisringe. Bei umgekehrter Integrationsreihenfolge integrieren wir zuerst bei festem Winkel über ein Tortenstück und dann im Kreis.

### 3.10.2 Zylinderkoordinaten

Wir wenden uns jetzt dreidimensionalen Situationen zu. Eine einfache Verallgemeinerung von Polarkoordinaten, die insbesondere dann praktisch ist, wenn wir eine einzige Symmetrieachse in unserem Aufbau haben. Zylinderkoordinaten  $(\rho, \phi, z)$  übersetzen sich in kartesische Koordinaten gemäß

$$x = \rho \cos \phi \quad y = \rho \sin \phi \quad z = z. \quad (3.66)$$

Die Einheitsvektoren hier sind

$$\hat{e}_\rho = \frac{\partial \vec{r}}{\partial \rho} = \begin{pmatrix} \cos \phi \\ \sin \phi \\ 0 \end{pmatrix} \quad \hat{e}_\phi = \frac{1}{\rho} \frac{\partial \vec{r}}{\partial \phi} = \begin{pmatrix} -\sin \phi \\ \cos \phi \\ 0 \end{pmatrix} \quad \hat{e}_z = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}. \quad (3.67)$$

Als Beispiel für Integration in Zylinderkoordinaten berechnen wir das Volumen eines Kreiskegels mit Basisradius  $R$  und Höhe  $h$ . Wir legen die  $z$ -Achse auf die Symmetrieachse des Kegels und die Grundfläche in die  $xy$ -Ebene. Damit ist der Radius einer Kreisscheibe  $\rho(z) = R \left(1 - \frac{z}{h}\right)$ . Das Volumen ist dann

$$\begin{aligned} V &= \int_V d^3r = \int_0^h dz \int_0^{\rho(z)} \rho d\rho \int_0^{2\pi} d\phi = 2\pi \int_0^h dz \frac{\rho^2}{2} \Big|_0^{\rho(z)} = \\ &= \pi R \int_0^{2h} dz \left(1 - \frac{z}{h}\right)^2 \end{aligned}$$

Sie sehen, wie wir erst die Fläche einer Kreisscheibe ausrechnen und dann über die Kreisscheiben integrieren. Mit der Substitution  $s = 1 - \frac{z}{h}$  bekommen wir

$$V = \pi R^2 h \int_0^1 ds s^2 = \frac{\pi h R^2}{3}. \quad (3.68)$$

### 3.10.3 Kugelkoordinaten

In Situationen mit voller Rotationssymmetrie bietet sich die Benutzung von Kugelkoordinaten  $(r, \theta, \phi) \in \mathbb{R}_0^+ \times [0, \pi] \times [0, 2\pi)$  an. Die Übersetzung in kartesische Koordinaten ist

$$x = r \sin \theta \cos \phi \quad y = r \sin \theta \sin \phi \quad z = r \cos \theta \quad (3.69)$$

Die Koordinatenachsen mit konstanten Winkeln sind wiederum Ursprungsgeraden. Die Linien mit konstantem  $r$  und  $\theta$  sind Breitenkreise und mit konstantem  $r$  und  $\phi$  sind Längenkreise auf der Kugel mit  $z$  als Achse. Anders als die Konvention in der Geographie laufen die Breitenkreise aber von 0 bis  $\pi$ , der „Nordpol“ ist also bei Breite 0 und der „Südpol“ bei Breite 180 Grad. Die Einheitsvektoren sind

$$\hat{e}_r = \begin{pmatrix} \sin \theta \cos \phi \\ \sin \theta \sin \phi \\ \cos \theta \end{pmatrix} \quad \hat{e}_\theta = \begin{pmatrix} \cos \theta \cos \phi \\ \cos \theta \sin \phi \\ -\sin \theta \end{pmatrix} \quad \hat{e}_\phi = \begin{pmatrix} -\sin \phi \\ \cos \phi \\ 0 \end{pmatrix} \quad (3.70)$$

Die Geschwindigkeit ist damit

$$\dot{\vec{r}} = \dot{r}\hat{e}_r + r\dot{\theta}\hat{e}_\theta + r\sin\theta\dot{\phi}\hat{e}_\phi. \quad (3.71)$$

Zur Volumenintegration konstruieren wir wieder die Größe eines Volumenelements. Dieses ist aus einem Breitenkreis mit Radius  $r\sin\theta$  ausgeschnitten und hat eine Bogenlänge  $rd\phi$ . Insgesamt haben wir also

$$d^3r = r^2 dr d\phi \sin\theta d\theta = r^2 dr d\phi d\cos\theta. \quad (3.72)$$

Als Beispiel berechnen wir die Gesamtmasse einer Kugel vom Radius  $R$  mit einer Massendichte  $\rho(x, y, z) = \rho_0(x^2 + y^2 + 3z^2)$ . Wir konvertieren die Massendichte in Kugelkoordinaten

$$\rho = \rho_0 r^2 (1 + 2\cos^2\theta) \quad (3.73)$$

Wir berechnen die Gesamtmasse

$$\begin{aligned} M &= \int_V d^3r \rho(x, y, z) = \int_0^R r^2 dr \int_{-1}^1 d\cos\theta \int_0^{2\pi} d\phi \rho_0 r^2 (1 + 2\cos^2\theta) \\ &= 2\pi\rho_0 \int_0^R r^4 dr \int_{-1}^1 d\cos\theta (1 + 2\cos^2\theta) = \frac{2\pi R^5 \rho_0}{5} \left(2 + \frac{4}{3}\right) = \frac{10\pi}{3} \rho_0 R^5. \end{aligned}$$

Ein wichtiges Oberflächenintegral in der Physik ist das über das Feld  $\vec{F}(\vec{r}) = \alpha\hat{e}_r/r^2$  über eine Kugel vom Radius  $R$ . Der Normalenvektor hier ist  $\hat{e}_r$  und ein Flächenelement hat die Größe  $r^2 d\phi d\cos\theta$ . Damit ist

$$\oint_{r=R} d^2r \hat{n} \cdot \vec{F} = 4\pi\alpha \quad (3.74)$$

unabhängig von  $R$ !



## Kapitel 4

# Lineare Abbildungen und Matrizen

Wir haben schon an verschiedenen Stellen, z.B. bei der Differenziation und Integration, gesehen, dass verschiedene Operationen *linear* sind. Diese Abbildungen verdienen darum unsere gesonderte Aufmerksamkeit. Wir konstruieren sehr weit gehende Techniken und Eigenschaften, die aus der Linearität der Abbildungen folgen.

### 4.1 Lineare Abbildungen

Seien  $V$  und  $W$  Vektorräume. Eine Abbildung  $A : V \rightarrow W$  heißt *linear*, wenn für eine beliebige Linearkombination gilt

$$A(\lambda \vec{a} + \mu \vec{b}) = \lambda A(\vec{a}) + \mu A(\vec{b}) \quad (4.1)$$

Bei linearen Abbildungen lassen wir die Klammernotation oft weg und schreiben sie als multiplikative Operation (was später noch deutlich besser motiviert wird), also  $A(\vec{a}) \equiv A\vec{a}$ . Wir betrachten einige Beispiele.

Die *orthogonale Projektion* tritt auf, wenn wir z.B. ein Foto machen. Wenn das Foto in der  $xy$ -Ebene liegt, können wir  $F : \mathbb{R}^3 \rightarrow \mathbb{R}^2$  definieren als

$$A \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} x \\ y \end{pmatrix} \quad (4.2)$$

Dies ist klarerweise eine (nicht-invertierbare) lineare Abbildung.

Eine *Drehung um den Ursprung* ist eine lineare Abbildung  $\mathbb{R}^2 \rightarrow \mathbb{R}^2$ . Wir drehen einen beliebigen Vektor indem wir das Koordinatensystem drehen. Damit ist

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} x \cos \alpha + y \sin \alpha \\ -x \sin \alpha + y \cos \alpha \end{pmatrix} \quad (4.3)$$

Die *Punktspiegelung* am Ursprung  $\vec{r} \rightarrow -\vec{r}$  ist eine lineare Abbildung. Ebenso die Spiegelung an einer den Ursprung enthaltenden Ebene. Ist dies z.B. die  $xy$ -Ebene dann ist diese Abbildung

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} \rightarrow \begin{pmatrix} x \\ y \\ -z \end{pmatrix}. \quad (4.4)$$

Aber auch in Funktionenräumen haben wir lineare Abbildungen kennengelernt, z.B. die Differenziation

$$(\lambda f(x) + \mu g(x))' = \lambda f'(x) + \mu g'(x). \quad (4.5)$$

#### 4.1.1 Lineare Abbildungen und die Basisdarstellung

Wir haben bereits gelernt, dass eine kompakte Darstellung von Vektoren einen  $n$  dimensionalen Vektorraumes durch die Darstellung des Vektors bzgl. einer Basis  $\mathcal{B} = \{\vec{b}_i, i = 1 \dots n\}$  gegeben ist

$$\vec{w} = \sum_{i=1}^n \lambda_i \vec{b}_i \quad (4.6)$$

Wenn wir jetzt eine lineare Abbildung anwenden, so folgt aus der Linearität

$$A\vec{w} = A \sum_{i=1}^n \lambda_i \vec{b}_i = \sum_{i=1}^n \lambda_i A\vec{b}_i \quad (4.7)$$

Dieses Ergebnis können wir interpretieren als die Entwicklung des Vektors nach dem Bild der Basis  $A\mathcal{B} = \{A\vec{b}_i, \vec{b}_i \in \mathcal{B}\}$  aber mit unveränderten Entwicklungskoeffizienten. Wenn wir also die Abbildung effizient berechnen wollen, müssen wir sie  $n$  Mal im Schweiße unseres Angesichts anwenden um  $A\mathcal{B}$  zu bestimmen, danach brauchen wir nur noch die Basisdarstellung.

Dies geht besonders effektiv, wenn wir nach der Bestimmung der Basis den kanonischen Isomorphismus anwenden. Zur Erinnerung: Dies bedeutet, dass wir den Vektorraum mit dem  $\mathbb{R}^n$  identifizieren und die Basisvektoren auf die Standardbasis  $\{\hat{e}_i\}$  abbilden. Anwendung der Abbildung liefert dann das Bild der Standardbasis, auch wieder im  $\mathbb{R}^n$  dargestellt  $A\mathcal{B} = \{A\hat{e}_i, i = 1 \dots n\}$  und daraus können wir die gesamte Abbildung rekonstruieren. Idealerweise interessieren uns auch wieder die Komponenten unseres Bildvektors bzgl. der Standardbasis rekonstruieren. Dazu definieren wir die Zahlen

$$A_{ij} = (A\hat{e}_j)_i \quad (4.8)$$

also die  $i$ -te Komponente des Bildes des Basisvektors  $j$ . Wenn wir diese haben, dann ist die  $i$ -te Komponente eines Bildvektors

$$(A\vec{w})_i = \left( \sum_j w_j A\hat{e}_j \right)_i = \sum_j A_{ij} w_j \quad (4.9)$$

Dies können wir in Komponentenschreibweise besonders kompakt darstellen

$$A\vec{w} = \begin{pmatrix} (A\vec{w})_1 \\ (A\vec{w})_2 \\ \vdots \\ (A\vec{w})_n \end{pmatrix} = \begin{pmatrix} A_{11} & A_{12} & \cdots & A_{1m} \\ A_{21} & A_{22} & \cdots & A_{2m} \\ \vdots & \ddots & \ddots & \vdots \\ A_{n1} & A_{n2} & \cdots & A_{nm} \end{pmatrix} \begin{pmatrix} w_1 \\ w_2 \\ \vdots \\ w_m \end{pmatrix} \quad (4.10)$$

Die tabellarische Anordnung von Zahlen nennen wir eine Matrix. Als Kurzschreibweise bezeichnen wir sie als

$$\mathbf{A} = (A_{ij}) \in \mathbb{R}^{n \times m} \quad (4.11)$$

Die Größe der Matrix ist durch die per Abbildung verbundenen Vektorräume vorgegeben: Wir bilden ab  $\mathbb{R}^m \mapsto \mathbb{R}^n$ . Sehr häufig ist der Fall  $n = m$  - die Matrix ist dann quadratisch. Die Matrix charakterisiert die lineare Abbildung vollständig. Die Anwendung auf der Matrix auf den Vektor ist, dass wir in der  $i$  ten Zeile der Ergebnisses das Skalarprodukt der  $i$  ten Zeile mit dem Vektor bilden. Eine gute Merkregel, die sich daraus ergibt ist, dass in den Spalten der Matrix die Bilder der Standardbasisvektoren des  $\mathbb{R}^m$  stehen.

## 4.2 Grundeigenschaften von Matrizen

Wir betrachten jetzt Eigenschaften von Matrizen, die sich in die eigenschaften linearer Abbildungen umsetzen. Gegeben eine Matrix  $\mathbf{A} = (A_{ij}) \in \mathbb{R}^{n \times m}$  bezeichnen wir die Matrix

$$\mathbf{A}^T = (A_{ji}) \in \mathbb{R}^{m \times n} \quad (4.12)$$

als die zu  $\mathbf{A}$  transponierte Matrix. Wir erkennen das ein Vektor eigentlich eine  $n \times 1$  Matrix ist - die Transponierte eines solchen Spaltenvektors ist dann ein Zeilenvektor.

Das Produkt von zwei Matrizen  $\mathbf{A} \in \mathbb{R}^{n \times m}$  und  $\mathbf{B} \in \mathbb{R}^{m \times p}$  ist eine Matrix  $\mathbf{AB} \in \mathbb{R}^{n \times p}$ . Wenn wir unter einen Standard-Basisvektor unter dieser verketteten Abbildung abbilden, dann erhalten wir sein Bild durch Anwendung von  $\mathbf{A}$  auf das Bild des Vektors unter  $\mathbf{B}$ . Da das Bild des  $i$  ten Standard-Basisvektors gerade die  $i$ te Spalte von  $\mathbf{B}$  ist müssen wir also  $\mathbf{A}$  auf die  $i$  te Spalte von  $\mathbf{B}$  anwenden. Die Multiplikation von Matrizen passiert darum durch die sukzessive Anwendung der Matrix auf die Spaltenvektoren.

Wir können das kurz schreiben als

$$(\mathbf{AB})_{ij} = \sum_{k=1}^m (\mathbf{A})_{ik} (\mathbf{B})_{kj} \quad (4.13)$$

Für das Element in der  $i$  ten Zeile und  $j$  ten Spalte der Produktmatrix nehmen wir also das Skalarprodukt der  $i$  ten Zeile von  $\mathbf{A}$  mit der  $j$ -ten Spalte von  $\mathbf{B}$ . In diesem Sinn können wir z.B. das Skalarprodukt von zwei Vektoren schreiben als  $\vec{u}^T \vec{v}$ .

Das Matrixprodukt hat folgende Eigenschaften:

1. Das Matrixprodukt ist linear, insbesondere gilt das Distributivgesetz
2. Das Matrixprodukt ist assoziativ

$$(\mathbf{ABC})_{ij} = \sum_{kl} \mathbf{A}_{ik} \mathbf{B}_{kl} \mathbf{C}_{lj} = \sum_k \mathbf{A}_{ik} (\mathbf{BC})_{kj} = \sum_l (\mathbf{AB})_{il} \mathbf{C}_{lj} \quad (4.14)$$

3. Das Matrixprodukt ist nicht kommutativ. Beispielsweise ist für

$$\mathbf{A} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \quad \mathbf{B} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \quad \mathbf{AB} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} = -\mathbf{BA} \quad (4.15)$$

Diese Eigenschaften finden wir auch bei linearen Abbildungen wieder - z.B. ist das Drehen und das Strecken nicht kommutativ (Zeichnung in der Vorlesung).

Eine besonderer Rolle spielen Diagonalmatrizen. Dies sind Matrizen  $\mathbf{D} \in \mathbb{R}^{n \times n}$  der Form

$$\mathbf{D} = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_n) = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix} \quad (4.16)$$

Den Spezialfall  $\text{diag}(1, 1, \dots, 1) = \hat{\mathbf{1}}$  nennt man Einheitsmatrix. Wenn man eine beliebige Matrix  $\mathbf{A}$  von links mit einer Diagonalmatrix multipliziert erhält man

$$\begin{aligned} \mathbf{DA} &= \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix} \begin{pmatrix} A_{11} & A_{12} & \cdots & A_{1m} \\ A_{21} & A_{22} & \cdots & A_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ A_{n1} & A_{n2} & \cdots & A_{nm} \end{pmatrix} \\ &= \begin{pmatrix} \lambda_1 A_{11} & \lambda_1 A_{12} & \cdots & \lambda_1 A_{1m} \\ \lambda_2 A_{21} & \lambda_2 A_{22} & \cdots & \lambda_2 A_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ \lambda_n A_{n1} & \lambda_n A_{n2} & \cdots & \lambda_n A_{nm} \end{pmatrix} \end{aligned}$$

d.h. die Matrix  $\mathbf{A}$  wird von zeilenweise mit den Diagonalelementen multipliziert. Umgekehrt

$$\begin{aligned} \mathbf{AD} &= \begin{pmatrix} A_{11} & A_{12} & \cdots & A_{1m} \\ A_{21} & A_{22} & \cdots & A_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ A_{n1} & A_{n2} & \cdots & A_{nm} \end{pmatrix} \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix} \\ &= \begin{pmatrix} \lambda_1 A_{11} & \lambda_2 A_{12} & \cdots & \lambda_n A_{1m} \\ \lambda_1 A_{21} & \lambda_2 A_{22} & \cdots & \lambda_n A_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ \lambda_1 A_{n1} & \lambda_2 A_{n2} & \cdots & \lambda_n A_{nm} \end{pmatrix} \end{aligned}$$

Interessant sind auch Matrizen, die bis auf einen Tausch in Zeilen / Spalten  $nm$  Einheitsmatrizen sind

$$\mathbf{S}_{nm} = \begin{pmatrix} 1 & \cdots & 0 & \cdots & 0 & \cdots & 0 \\ \vdots & \ddots & \vdots & \vdots & 0 & \cdots & 0 \\ 0 & \cdots & 0 & \cdots & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & \cdots & 1 & \cdots & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & \cdots & 0 & \cdots & 1 \end{pmatrix} \quad (4.17)$$

Sie können sich leicht davon überzeugen, dass Multiplikation von  $\mathbf{S}_{nm}$  von links auf  $\mathbf{A}$  die Zeilen  $n$  und  $m$  vertauscht, von rechts aber die Zeilen. Außerdem ist noch die Matrix

$$(\mathbf{T}_{nm})_{ij} = \delta_{ij} + \delta_{in}\delta_{jm} \quad (4.18)$$

nützlich - das ist die Einheitsmatrix, die in Zeile  $n$  und Spalte  $m$  noch eine zusätzliche Eins hat. Von links multipliziert addiert sie Zeilen  $m$  auf Zeile  $n$  (und bei Multiplikation rechts gilt das für die Zeilen).

### 4.3 Matrizen und lineare Gleichungssysteme

Ein lineares Gleichungsproblem mit  $n$  Stück Gleichungen für  $m$  Stück Unbekannte kann geschrieben werden als

$$\begin{aligned} A_{11}x_1 + A_{12}x_2 + \cdots + A_{1m}x_m &= y_1 \\ A_{21}x_1 + A_{22}x_2 + \cdots + A_{2m}x_m &= y_2 \\ &\cdots + \cdots + \cdots + \cdots = \cdots \\ A_{n1}x_1 + A_{n2}x_2 + \cdots + A_{nm}x_m &= y_n \end{aligned}$$

Damit können wir dieses Gleichungssystem als Vektorgleichung umschreiben

$$\begin{pmatrix} A_{11} & A_{12} & \cdots & A_{1m} \\ A_{21} & A_{22} & \cdots & A_{2m} \\ \vdots & \ddots & \ddots & \vdots \\ A_{n1} & A_{n2} & \cdots & A_{nm} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_m \end{pmatrix} = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{pmatrix}. \quad (4.19)$$

Damit entsteht eine Verbindung zwischen linearen Gleichungssystemen und Matrizen. Sie haben in der Schule bzw. im Vorkurs bereits den Gaußalgorithmus zur Lösung von linearen Gleichungssystemen gelernt. In diesem eliminieren Sie sukzessive Koeffizienten des LGS, bis Sie die Lösung ablesen können. Die Schritte, die Sie zur Verfügung haben sind das Durchmultiplizieren, das Addieren, und das Tauschen von Zeilen - also gerade Matrixmultiplikationen mit den im letzten Kapiteln konstruierten speziellen Matrizen.

### 4.3.1 Die inverse Matrix

Wie können wir diesen Zusammenhang zwischen LGSen und Matrixrechnung nutzen um lineare Abbildungen besser zu verstehen, oder um LGS besser zu nutzen? Hierzu definieren wir die zur quadratischen Matrix  $\mathbf{A}$  inverse Matrix  $\mathbf{A}^{-1}$  durch die Eigenschaft

$$\mathbf{A}^{-1}\mathbf{A} = \hat{1} \tag{4.20}$$

ohne zu wissen, ob bzw. unter welchen Bedingungen sie existiert. Wenn sie aber existiert, dann können wir zeigen, dass es hier nicht auf die Reihenfolge ankommt. Wir überprüfen das, indem wir zusätzlich definieren  $\mathbf{A}\tilde{\mathbf{A}}^{-1} = \hat{1}$ . Wir können zeigen

$$\mathbf{A}^{-1} = \mathbf{A}^{-1}\hat{1} = \mathbf{A}^{-1}(\mathbf{A}\tilde{\mathbf{A}}^{-1}) = (\mathbf{A}^{-1}\mathbf{A})\tilde{\mathbf{A}}^{-1} = \tilde{\mathbf{A}}^{-1} \tag{4.21}$$

damit sind die „rechtsinverse“ und die „linksinverse“ die gleiche Matrix.

Bevor wir die Existenzfrage behandeln schauen wir uns an, wie wir die Inverse berechnen. Wir nehmen uns dazu die Rechtsinverse vor und schreiben die Definitionsgleichung aus

$$\begin{pmatrix} A_{11} & A_{12} & \cdots & A_{1m} \\ A_{21} & A_{22} & \cdots & A_{2m} \\ \vdots & \ddots & \ddots & \vdots \\ A_{n1} & A_{n2} & \cdots & A_{nm} \end{pmatrix} \begin{pmatrix} A_{11}^{-1} & A_{12}^{-1} & \cdots & A_{1m}^{-1} \\ A_{21}^{-1} & A_{22}^{-1} & \cdots & A_{2m}^{-1} \\ \vdots & \ddots & \ddots & \vdots \\ A_{n1}^{-1} & A_{n2}^{-1} & \cdots & A_{nm}^{-1} \end{pmatrix} = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix}. \tag{4.22}$$

Wenn wir jetzt die  $k$ -te Spalte der rechten Seite herausgreifen dann sehen wir aufgrund der Vorgehensweise bei der Matrixinversion, dass es dabei auch nur auf die  $k$ -te Spalte der Inversen ankommt. Um diese  $k$ -te Spalte zu finden müssen wir also das LGS

$$\begin{pmatrix} A_{11} & A_{12} & \cdots & A_{1m} \\ A_{21} & A_{22} & \cdots & A_{2m} \\ \vdots & \ddots & \ddots & \vdots \\ A_{n1} & A_{n2} & \cdots & A_{nm} \end{pmatrix} \begin{pmatrix} A_{1k}^{-1} \\ A_{2k}^{-1} \\ \vdots \\ A_{nk}^{-1} \end{pmatrix} = \hat{e}_k \tag{4.23}$$

lösen. Abtasten von  $k$  liefert das gewünschte Ergebnis. Wenn wir diese  $n$  Stück LGS mit der jeweils gleichen Matrix  $A$  lösen machen wir in der Gaußelimination immer die gleichen Schritte. Wir können also Arbeit sparen, indem wir alle  $\hat{e}_k$  gleichzeitig behandeln. Wie das aussieht, studieren wir am Beispiel des Invertierens von

$$\mathbf{A} = \begin{pmatrix} 1 & 1 & -1 \\ 1 & 0 & -\frac{1}{2} \\ -1 & -1 & \frac{3}{2} \end{pmatrix} \tag{4.24}$$

Wir schreiben linke und rechte Seiten auf als

$$\left( \begin{array}{ccc|ccc} 1 & 1 & -1 & 1 & 0 & 0 \\ 1 & 0 & -\frac{1}{2} & 0 & 1 & 0 \\ -1 & -1 & \frac{3}{2} & 0 & 0 & 1 \end{array} \right). \tag{4.25}$$

Ausräumen können wir, indem wir die erste Zeile auf die dritte addieren und von der zweiten subtrahieren

$$\left( \begin{array}{ccc|ccc} 1 & 1 & -1 & 1 & 0 & 0 \\ 0 & 1 & -\frac{1}{2} & 1 & -1 & 0 \\ 0 & 0 & \frac{1}{2} & 1 & 0 & 1 \end{array} \right). \quad (4.26)$$

Jetzt subtrahieren wir die zweite Zeile von der Ersten und multiplizieren die letzte Zeile mit 2

$$\left( \begin{array}{ccc|ccc} 1 & 0 & -\frac{1}{2} & 0 & 1 & 0 \\ 0 & 1 & -\frac{1}{2} & 1 & -1 & 0 \\ 0 & 0 & 1 & 2 & 0 & 2 \end{array} \right). \quad (4.27)$$

Und schließlich addieren wir die Hälfte der letzten Zeile auf die beiden Oberen

$$\left( \begin{array}{ccc|ccc} 1 & 0 & 0 & 1 & 1 & 1 \\ 0 & 1 & 0 & 2 & -1 & 1 \\ 0 & 0 & 1 & 2 & 0 & 2 \end{array} \right). \quad (4.28)$$

Und nach Konstruktion steht jetzt im rechten Block die Inverse Matrix

$$\mathbf{A}^{-1} = \begin{pmatrix} 1 & 1 & 1 \\ 2 & -1 & 1 \\ 2 & 0 & 2 \end{pmatrix} \quad (4.29)$$

Wir können leicht verifizieren dass

$$\mathbf{A}^{-1}\mathbf{A} = \begin{pmatrix} 1 & 1 & 1 \\ 2 & -1 & 1 \\ 2 & 0 & 2 \end{pmatrix} \begin{pmatrix} 1 & 1 & -1 \\ 1 & 0 & -\frac{1}{2} \\ -1 & -1 & \frac{3}{2} \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad (4.30)$$

Im Prinzip könnten wir auch alle die speziellen Matrizen des Gaußalgorithmus miteinander multiplizieren, um die Inverse zu erhalten, aber das wollen wir nicht ausschreiben müssen.

Zurückgehend auf die Bedeutung der Matrix als Darstellung einer linearen Abbildung im  $\mathbb{R}^n$  ist  $\mathbf{A}^{-1}$  die Umkehrabbildung und die Spalten dieser Matrix sind die Bilder der Einheitsbasisvektoren unter der Umkehrung, also die Urbilder unter der ursprünglichen Abbildung.

Wenn wir die Inverse Matrix haben, können wir alle LGS der Form  $\mathbf{A}\vec{x} = \vec{y}$  lösen zu  $\vec{y} = \mathbf{A}^{-1}\vec{x}$ . Dies bedeutet eine große Vereinfachung und es lohnt sich damit, die Inverse zu bestimmen. Damit ist die Lösung des LGS das Urbild von  $\vec{y}$  unter der von  $\mathbf{A}$  beschriebenen linearen Abbildung.

### 4.3.2 Invertierbarkeit

Jetzt können wir zu der Frage zurückkommen, wann eine Matrix invertierbar ist, und was wir daraus lernen können, insbesondere aus dem Zusammenhang mit linearen Abbildungen. Hemdsärmelig können wir sagen, dass wenn auch immer das LGS uns immer eine eindeutige Lösung liefert, ist  $\mathbf{A}$  invertierbar. Folgende Aussagen erkennen wir als äquivalent an

1.  $\mathbf{A}$  ist invertierbar.
2. Die Zeilen von  $\mathbf{A}$  sind linear unabhängig
3. Die Spalten von  $\mathbf{A}$  sind linear unabhängig
4. Die einzige Lösung von  $\mathbf{A}\vec{x} = \vec{0}$  ist der Nullvektor,  $\vec{x} = 0$
5.  $\mathbf{A}^T$  ist invertierbar

Warum stimmen diese Aussagen?

- $1 \Leftrightarrow 2$  : Der Gaußalgorithmus bildet Linearkombinationen der Zeilen von  $\mathbf{A}$  und wird genau dann zum Ziel führen, wenn dabei nicht ausversehen eine Nullzeile entsteht - entsprechend der Definition linearer Unabhängigkeit
- $1 \Leftrightarrow 3$  : Die Spalten von  $\mathbf{A}$  sind die Bilder der Einheitsbasisvektoren. Wenn diese wieder eine Basis bilden, dann kann ich jeden Vektor aus  $\mathbb{R}^n$  nach ihnen entwickeln und das Urbild bestimmen - die Abbildung ist invertierbar
- 5 ist äquivalent zu 2 und 3 mit vertauschten Rollen
- 4 ist äquivalent zu 3 per Definition

Wenn wir Invertierbarkeit überprüfen wollen, können wir also eines dieser Kriterien überprüfen. Wir werden noch ein deutlich einfacheres kennenlernen.

### 4.3.3 Beispiele

Die Matrix der Drehung um den Ursprung

$$\mathbf{R}_\phi = \begin{pmatrix} \cos \phi & -\sin \phi \\ \sin \phi & \cos \phi \end{pmatrix} \quad (4.31)$$

hat als Inverses die umgekehrte Drehung

$$\mathbf{R}_\phi^{-1} = \mathbf{R}_{-\phi} = \begin{pmatrix} \cos \phi & \sin \phi \\ -\sin \phi & \cos \phi \end{pmatrix} \quad (4.32)$$

Die Spiegelung an der  $xy$  - Ebene

$$\mathbf{S} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix} \quad (4.33)$$

ist ihr eigenes Inverses.



Betrachten wir jetzt den Vektorraum der Polynome von höchstens drittem Grad mit der Basis aus Monomen  $\mathcal{B} = \{1, x, x^2, x^3\}$ , dargestellt durch den  $\mathbb{R}^4$ , d.h. wir identifizieren

$$a_0 + a_1x + a_2x^2 + a_3x^3 = \begin{pmatrix} a_0 \\ a_1 \\ a_2 \\ a_3 \end{pmatrix} \quad (4.34)$$

Ableiten ist eine lineare Abbildung und durch die Matrix dargestellt

$$\mathbf{D} = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \\ 0 & 0 & 0 & 0 \end{pmatrix} \quad (4.35)$$

Diese Matrix ist nicht invertierbar, ihre Spalten sind nicht linear unabhängig. Dies entspricht unserer Erfahrung: Die unbestimmte Integration als Umkehrabbildung der Ableitung ist nur bis auf eine Konstante eindeutig.

Für eine  $2 \times 2$  Matrix

$$\mathbf{A} = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \quad (4.36)$$

können wir die Inverse sofort angeben zu

$$\mathbf{A}^{-1} = \frac{1}{ad - bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix} \quad (4.37)$$

Der Nenner im Vorfaktor ist ein Beispiel einer *Determinanten* (die allgemeine Theorie von Determinanten lernen wir bald) und die muss von Null verschieden sein. Für die Einträge können wir uns merken, dass wir auf der Diagonalen tauschen und auf der Nebendiagonalen ein Minuszeichen eintragen. Auch für  $3 \times 3$  Matrizen existiert eine solche Formel, die wir aber erst mit der Theorie der Determinanten behandeln werden.

Die Inverse einer Diagonalmatrix  $\mathbf{D} = \text{diag}(d_1, d_2, \dots, d_n)$  existiert wenn alle Diagonalelemente von null verschieden sind und ist  $\mathbf{D}^{-1} = \text{diag}(d_1^{-1}, d_2^{-1}, \dots, d_n^{-1})$

## 4.4 Basiswechsel

Das Nutzen von Basen, um Vektoren auf Spaltenvektoren zu delegieren und lineare Abbildungen auf Matrizen ist sehr praktisch und erlaubt es, Rechenaufwand vielerorts auf das Nötigste zu reduzieren. Allerdings ist die Festlegung, welche der vielen möglichen Basen man nimmt, oft recht willkürlich und kann den Blick auf die Essenz einer Operation verstellen. Betrachten Sie die oben eingeführte Darstellung  $\mathbf{S}$  der Spiegelung an der  $xy$  Ebene. Die ist sehr einfach - aber wie ist es, wenn ich an einer beliebigen Ebene spiegele? Wie arbeite ich damit besonders effektiv? Wir werden diese Aufgabe in mehreren Schritten angehen, die auf das Konzept der Eigenwertzerlegung herausläuft.

Zunächst betrachten wir aber die Frage, wie wir überhaupt Basisdarstellungen ineinander umrechnen. Seien  $\mathcal{B} = \{\vec{v}_i, i = 1 \dots n\}$  und  $\mathcal{B}' = \{\vec{v}'_i, i = 1 \dots n\}$  Basen eines  $n$ - Dimensionalen reellen Vektorraums. Ein beliebiger Vektor  $\vec{w}$  aus diesem Raum kann dargestellt werden mit Koeffizienten

$$\vec{w} = \sum_{i=1}^n w_i \vec{v}_i = \sum_{i=1}^n w'_i \vec{v}'_i \quad (4.38)$$

Beim Basiswechsel interessiert uns eine möglichst effiziente Methode, die Koeffizienten  $w'_i$  aus den  $w_i$  zu bestimmen. Dazu beschaffen wir uns die Entwicklung der alten Basis in den Vektoren der neuen

$$\vec{v}_i = \sum_{j=1}^n T_{ji} \vec{v}'_j \quad (4.39)$$

und setzen dies in Gleichung (4.38) ein

$$\sum_{i=1}^n w_i \vec{v}_i = \sum_{i,j=1}^n w_i T_{ji} \vec{v}'_j = \sum_{j=1}^n w'_j \vec{v}'_j \quad (4.40)$$

Aufgrund der linearen Unabhängigkeit von  $\mathcal{B}'$  können wir Koeffizientenvergleich machen, d.h. es folgt

$$w'_j = \sum_{i=1}^n T_{ji} w_i \quad (4.41)$$

Dies können wir als Matrixgleichung schreiben als  $\mathbf{w}' = \mathbf{T}\mathbf{w}$ . Die Koeffizienten  $T_{ji}$  bilden die *Basiswechselmatrix*. Der Basiswechsel wird damit mathematisch genau so bewerkstelligt, wie eine lineare Abbildung. Man spricht von einer *passiven Abbildung*: Es wird nicht der Vektor verändert, sondern das Bezugssystem gedreht. Wie können wir uns die Gestalt dieser Matrix merken? Wir starten vom Spezialfall  $\vec{w} = \vec{v}_k$ : Dann ist  $w_i = \delta_{ik}$  und wir sehen dass in dem Produkt gerade die  $k$ -te Spalte von  $\mathbf{B}$  herausgegriffen wird - dort steht also die Darstellung des  $k$ -ten Basisvektors der alten Basis. Wenn wir jetzt von  $\mathcal{B}'$  auf  $\mathcal{B}$  zurückgehen wollen benötigen wir entsprechend die inverse Matrix  $\mathbf{T}^{-1}$ . Da wir Basen auf Basen abbilden existiert diese Matrix auf jeden Fall.

Was machen wir jetzt mit linearen Abbildungen? Wenn die Matrixdarstellung  $\mathbf{A}$  einer Linearen Abbildung in der Basis  $\mathcal{B}$  bekannt ist, wir die aber auf die Matrixdarstellung in der Basis  $\mathcal{B}'$  anwenden wollen, dann müssen wir diese zuerst umrechnen, dann anwenden, dann zurückrechnen. Es ist also

$$\mathbf{A}' = \mathbf{T}\mathbf{A}\mathbf{T}^{-1} \quad \mathbf{A} = \mathbf{T}^{-1}\mathbf{A}'\mathbf{T}. \quad (4.42)$$

### 4.4.1 Beispiele

Behandeln wir einige Beispiele. Es sei  $\mathcal{B}$  die Standardbasis und  $\mathcal{B}'$  verknüpft mit ihr durch

$$\begin{aligned}\hat{e}_x &= \frac{3}{4}\vec{v}'_1 + \frac{1}{3}\vec{v}'_2 \\ \hat{e}_y &= -\frac{1}{8}\vec{v}'_1 + \frac{1}{2}\vec{v}'_2\end{aligned}\quad (4.43)$$

Damit haben wir die Transformationsmatrix

$$\mathbf{T} = \begin{pmatrix} \frac{3}{4} & -\frac{1}{8} \\ \frac{1}{3} & \frac{1}{2} \end{pmatrix}\quad (4.44)$$

Wenn wir z.B. den Vektor

$$\vec{w} = \begin{pmatrix} 1 \\ 2 \end{pmatrix}\quad (4.45)$$

bzgl. der neuen Basis ausdrücken wollen, dann rechnen wir

$$\mathbf{w} = \mathbf{T}\vec{w} = \begin{pmatrix} \frac{1}{2} \\ \frac{4}{3} \end{pmatrix}.\quad (4.46)$$

In der Tat können wir mit Gl. (4.43) nachprüfen, dass

$$\hat{e}_x + 2\hat{e}_y = \left(\frac{3}{4} - \frac{1}{4}\right)\vec{v}'_1 + \left(\frac{1}{3} + 1\right)\vec{v}'_2.\quad (4.47)$$

Oftmals ist es viel interessanter, aus einer gegebenen Transformation die umgekehrte Richtung zu betrachten. Hierzu benötigen die Inverse

$$\mathbf{T}^{-1} = \frac{1}{\frac{3}{8} + \frac{1}{24}} \begin{pmatrix} \frac{1}{2} & \frac{1}{8} \\ -\frac{1}{3} & \frac{3}{4} \end{pmatrix} = \begin{pmatrix} \frac{6}{5} & \frac{3}{10} \\ -\frac{4}{5} & \frac{9}{5} \end{pmatrix}\quad (4.48)$$

Wir können an den Spalten dann sofort die beiden neuen Basisvektoren anschauen.

Ein weiteres Beispiel ist der Fall, dass wir eine Ebenenspiegelung nicht an der  $xy$ -Ebene sondern an einer um 30 Grad dazu um die  $y$ -Achse gekippte Ebene durchführen wollen. Dazu kippen wir die Basis entsprechend mit

$$\mathbf{T} = \begin{pmatrix} \frac{\sqrt{3}}{2} & 0 & \frac{1}{2} \\ 0 & 1 & 0 \\ -\frac{1}{2} & 0 & \frac{\sqrt{3}}{2} \end{pmatrix} \quad \mathbf{T}^{-1} = \begin{pmatrix} \frac{\sqrt{3}}{2} & 0 & -\frac{1}{2} \\ 0 & 1 & 0 \\ \frac{1}{2} & 0 & \frac{\sqrt{3}}{2} \end{pmatrix}\quad (4.49)$$

und multiplizieren aus zur neuen Transformation

$$\begin{aligned}\mathbf{A} &= \mathbf{T}^{-1}\mathbf{A}'\mathbf{T} = \begin{pmatrix} \frac{\sqrt{3}}{2} & 0 & -\frac{1}{2} \\ 0 & 1 & 0 \\ \frac{1}{2} & 0 & \frac{\sqrt{3}}{2} \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix} \begin{pmatrix} \frac{\sqrt{3}}{2} & 0 & \frac{1}{2} \\ 0 & 1 & 0 \\ -\frac{1}{2} & 0 & \frac{\sqrt{3}}{2} \end{pmatrix} \\ &= \begin{pmatrix} \frac{\sqrt{3}}{2} & 0 & -\frac{1}{2} \\ 0 & 1 & 0 \\ \frac{1}{2} & 0 & \frac{\sqrt{3}}{2} \end{pmatrix} \begin{pmatrix} \frac{\sqrt{3}}{2} & 0 & \frac{1}{2} \\ 0 & 1 & 0 \\ \frac{1}{2} & 0 & -\frac{\sqrt{3}}{2} \end{pmatrix} = \begin{pmatrix} \frac{1}{2} & 0 & \frac{\sqrt{3}}{2} \\ 0 & 1 & 0 \\ \frac{\sqrt{3}}{2} & 0 & -\frac{1}{4} \end{pmatrix}.\end{aligned}$$

Dieser Matrix sieht man nicht ohne weiteres an, dass es eine Ebenenspiegelung ist. Das würden wir aber gerne können! Wir werden in den nächsten Kapiteln lernen, wie man aus Matrizen basisunabhängige Eigenschaften berechnet.

#### 4.4.2 Die Spur einer Matrix

Die Spur einer Matrix  $\mathbf{A} \in \mathbb{R}^{n \times n}$  ist die Summe der Diagonalelemente. Sie wird in der Literatur deutsch oder englisch abgekürzt

$$\text{Tr} \mathbf{A} = \text{Sp} \mathbf{A} = \sum_{i=1}^n A_{ii}. \quad (4.50)$$

Sie hat eine große Bedeutung u.a. in der statistischen Physik. Uns interessiert sie als Fallbeispiel für eine basisunabhängige Eigenschaft. Zunächst überzeugen wir uns davon, dass die Spur zyklisch invariant ist

$$\text{Tr}(\mathbf{AB}) = \sum_{i=1}^n (\mathbf{AB})_{ii} = \sum_{i,j=1}^n \mathbf{A}_{ij} \mathbf{B}_{ji} = \sum_{i,j=1}^n \mathbf{B}_{ij} \mathbf{A}_{ji} = \sum_{j=1}^n (\mathbf{AB})_{jj} = \text{Tr}(\mathbf{BA}) \quad (4.51)$$

Daraus folgt dann automatisch  $\text{Tr}(\mathbf{ABC}) = \text{Tr}(\mathbf{BCA})$  etc. Damit ist für einen Basiswechsel

$$\text{Tr}(\mathbf{TAT}^{-1}) = \text{Tr}(\mathbf{AT}^{-1}\mathbf{T}) = \text{Tr}(\mathbf{A}). \quad (4.52)$$

Die Spur ist also basisunabhängig.

### 4.5 Determinante

Eine weitere, sehr wichtige Eigenschaft einer Matrix / linearen Abbildung ist die *Determinante*. Auch diese ist basisunabhängig.

#### 4.5.1 Permutationen und ihr Signum

Wir haben bereits anlässlich des Kreuzproduktes das (dreidimensionale) Levi-Civita-Symbol  $\epsilon_{ijk}$ ,  $i, j, k \in \{1, 2, 3\}$  kennengelernt. Es verschwindet wenn mindestens zwei Indizes den gleichen Wert haben. Es bleibt die Situation, in dem alle drei möglichen Indizes auftauchen und vertauscht sein. Eine solche Reihung aller Elemente von  $\{1, 2, \dots, n\}$  nennt man eine *Permutation*. Jede dieser Permutationen können wir aus der Ausgangsanordnung durch paarweises Vertauschen erhalten (wir tauschen z.B. erst das erste Element an seinen Platz, dann das zweite etc.) Die Zahl dieser Vertauschungen ist nicht eindeutig aber - da sich zwei am Ende überflüssige Vertauschungen immer wieder aufheben ist es zumindest eindeutig, ob die Zahl der Permutationen gerade oder ungerade ist - dies bezeichnet man als das Signum der Permutation (gerade / ungerade) und man weist ihm den numerischen Wert nach der Zahl der Vertauschungen  $n_p$  zu

$$\text{sgn}(P) = (-1)^{n_p} = \begin{cases} 1 & \text{P gerade} \\ -1 & \text{P ungerade} \end{cases} \quad (4.53)$$

Für  $n = 3$  waren die zyklischen Permutationen gerade und alle anderen ungerade. In anderen Fällen ist das deutlich strukturierter. Die Menge aller Permutationen nennt man die Symmetrische Gruppe  $S_n$ .

### 4.5.2 Die Determinante

Die Determinante einer quadratischen Matrix  $\mathbf{A} \in \mathbb{R}^{n \times n}$  ist dann definiert über die Leibniz-Regel

$$\det \mathbf{A} = \sum_{P \in S_n} \operatorname{sgn}(P) \prod_{i=1}^n A_{iP_i} \quad (4.54)$$

Dieser Ausdruck bedeutet, dass wir i) eine Permutation ziehen ii) durch die Zeilen durchgehen und aus jeder Zeile das durch die Permutation angegebene Spalte Element herausgreifen iii) wir diese Elemente miteinander multiplizieren, dann das Signum der Permutation drannmultiplizieren. Am Ende iv) werden alle diese Beiträge zusammengezählt.

Nähern wir uns dieser Formel schrittweise: Im Fall  $n = 1$  ist  $\det \mathbf{A} = A_{11}$ . Für  $n = 2$  gibt es nur zwei Permutationen

$$\det \mathbf{A} = A_{11}A_{22} - A_{12}A_{21} \quad (4.55)$$

Dies ist der Nenner aus der Formel für die Matrixinverse der  $2 \times 2$  Matrix. Für  $n = 3$  geht das analog zur Jägerzaunregel beim Kreuzprodukt

$$\det \mathbf{A} = A_{11}A_{22}A_{33} - A_{11}A_{23}A_{32} - A_{12}A_{21}A_{33} + A_{12}A_{23}A_{31} - A_{13}A_{22}A_{31} + A_{13}A_{21}A_{32} \quad (4.56)$$

Für größere  $n$  brauchen wir eine Strategie, um die Permutationen alle systematisch zu erfassen. Hierzu hilft uns die *Laplace-Regel*:

1. Wir greifen uns eine beliebige Zeile  $i$  einer Matrix heraus - idealerweise eine mit vielen Nullen.
2. Wir berechnen die Minoren  $M_{ij}$ ,  $j = 1 \dots n$ . Dies sind die Matrizen, die entstehen, wenn wir Reihe  $i$  und Spalte  $j$  herauslassen.
3. Dann berechnen wir  $\det A = \sum_j A_{ij}(-1)^{i+j} \det M_{ij}$ .

Dies funktioniert auch, wenn wir Spalte  $j$  festhalten analog. Die Formel in Schritt 3. numeriert einfach die Permutationen sinnvoll durch. Dies ist ein Beispiel für eine *Rekursionsformel*: Im dritten Schritt müssen wir die Determinanten der Minoren ggf. wieder mittels Laplaceregeln ausrechnen, die Regel also wieder in sich selbst einsetzen.

Bevor wir den Sinn dieser Konstruktion zur Lösung praktischer Aufgaben ergründen, rechnen wir ein paar Beispiele.

### 4.5.3 Spezialfälle

Dies sind Matrizen  $\mathbf{A}$  mit  $A_{ij} = 0$  für  $i > j$ . Diese Matrizen spielen in der Physik von Zerfallsprozessen eine Rolle. Wenn wir die Leibniz-Regel hinschreiben sehen wir, dass nur die Permutation  $1, \dots, n$  eine Rolle spielt - jedes Tauschen liefert einen null: Für obere (und analog für untere) Dreiecksmatrizen ist

$$\det \mathbf{A} = \prod_{i=1}^n A_{ii}. \quad (4.57)$$

Dies gilt insbesondere auch für Diagonalmatrizen.

### 4.5.4 Eigenschaften

Die Nützlichkeit der Determinante erklärt sich aus ihren Eigenschaften:

1. Invarianz unter Transposition  $\det \mathbf{A} = \det \mathbf{A}^T$ . Der Beweis kann per Induktion und Laplace-Formel erfolgen, wobei wir einmal nach Zeilen und einmal nach Spalten entwickeln
2. Invarianz unter Zeilen- bzw. Spaltentausch: Laplace-Formel - der Tausch kann durch eine Vertauschung von Indizes in der Permutation rückgängig gemacht werden, das kehrt aber jedes Signum um. Korollar: Sind zwei Spalten (oder zwei Zeilen) gleich, dann verschwindet die Determinante.
3. Multilinearität: Die Determinante ist in jeder Spalte bzw. Zeile linear: Distributivgesetz in jedem Beitrag der Leibniz-Formel
4. Skalierung:  $\det(r\mathbf{A}) = r^n \det \mathbf{A}$  : Taucht in jedem Faktor der Leibniz-Formel auf
5. Wenn Zeilen (oder Spalten) linear abhängig sind, ist  $\det \mathbf{A} = 0$ . Dies werden wir am meisten benutzen. Wenn die Spalten linear abhängig sind, dann können wir eine Spalte der Matrix als Linearkombination der anderen Spalten schreiben, z.B.  $\mathbf{A}_1 = \sum_{i=2}^n \lambda_i \mathbf{A}_i$ . Dann können wir per Multilinearität die Determinante als Summe von Determinanten von Matrizen mit identischen Spalten schreiben (die erste und die  $i$ te) - die verschwinden alle.
6. Korollar: Die Determinante verschwindet genau dann, wenn die Matrix nicht invertierbar ist. Damit können wir die Invertierbarkeit testen.
7. Determinantenmultiplikationssatz: ein aufwändiger Standardbeweis zeigt  $\det \mathbf{AB} = \det \mathbf{A} \det \mathbf{B}$
8. Daraus folgt automatisch dass  $\det \mathbf{A}^{-1} = (\det \mathbf{A})^{-1}$
9. Aus diesen beiden letzten Eigenschaften können wir die Basisunabhängigkeit schließen

$$\det(\mathbf{T}^{-1} \mathbf{A} \mathbf{T}) = \det \mathbf{T}^{-1} \det \mathbf{A} \det \mathbf{T} = \det \mathbf{A} \quad (4.58)$$

Die Determinante hat also zwei interessante Eigenschaften: Sie charakterisiert die lineare Abbildung unabhängig von der Basis, und sie erlaubt es, auf Invertierbarkeit zu testen - mit einer recht aufwändigen Rechnung, aber weniger aufwändig als vollständiger Gaußalgorithmus.

## 4.6 Diagonalisierung und Eigenwerte

Diagonalmatrizen sind besonders einfach - sie strecken die Basisvektoren um einen festen Wert und mischen sie nicht miteinander. Es ist darum eine lohnende Frage, ob es für eine gegebene Matrix  $\mathbf{A} \in \mathbb{R}^{n \times n}$  eine Basis gibt, in der sie diagonal ist, und was diese Basis sowie die Diagonalelemente sind. Dieses Ziel verfolgen wir mit der Untersuchung von Eigenwerten und -vektoren.

### 4.6.1 Eigenvektoren und Eigenwerte

Eine Diagonalmatrix  $\mathbf{D} = \text{diag}(\lambda_1, \dots, \lambda_n)$  operiert auf einen Vektor der Standardbasis durch Multiplikation  $\mathbf{D}\hat{e}_i = \lambda_i\hat{e}_i$ . Wenn wir jetzt einen Basiswechsel mit einer Matrix  $\mathbf{T}$  durchführen wird daraus

$$\mathbf{T}\mathbf{D}\mathbf{T}^{-1}\mathbf{T}\hat{e}_i = \lambda_i\mathbf{T}\hat{e}_i \quad (4.59)$$

D.h. die transformierte Matrix  $\mathbf{A} = \mathbf{T}\mathbf{D}\mathbf{T}^{-1}$  multipliziert den transformierten Vektor  $\vec{v}_i = \mathbf{T}\hat{e}_i$  mit dem gleichen Wert  $\lambda_i$ . In der Praxis haben wir die transformierte Matrix  $\mathbf{A}$  vorgegeben und suchen Vektor  $\vec{v}_i \neq \vec{0}$  und  $\lambda_i$  so, dass

$$\mathbf{A}\vec{v}_i = \lambda_i\vec{v}_i \quad (4.60)$$

ist. Dies nennt man ein *Eigenwertproblem*,  $\lambda_i$  einen *Eigenwert* und  $\vec{v}_i$  den zugehörigen *Eigenvektor*.

### 4.6.2 Das charakteristische Polynom

Wir können das Eigenwertproblem umschreiben zu

$$(\mathbf{A} - \lambda\mathbf{1})\vec{v} = \vec{0}. \quad (4.61)$$

Wir suchen also nichtverschwindende Vektoren so, dass für passendes  $\lambda$  die Anwendung der Matrix  $\mathbf{A} - \lambda\mathbf{1}$  den Nullvektor liefert. Dazu darf diese Matrix nicht invertierbar sein - sonst wäre der Nullvektor die einzige Lösung. Zur Überprüfung der Invertierbarkeit nehmen wir die Determinante

$$Z(\lambda) = \det(\mathbf{A} - \lambda\mathbf{1}) \quad (4.62)$$

$Z(\lambda)$  heißt *charakteristisches Polynom* der Matrix  $\mathbf{A}$  und seine Nullstellen sind die Eigenwerte. Wenn wir einen Eigenwert gefunden haben, sind die Lösungen des oben angegebenen LGS die Eigenvektoren. Das charakteristische Polynom hat nach Leibniz-Formel den Grad  $n$ . Wir nennen die Menge aller Eigenwerte

einer Matrix das *Spektrum* der Matrix. Als Determinante ist das charakteristische Polynom basisunabhängig, d.h.

Bevor wir uns mit seinen Eigenschaften beschäftigen, wollen wir ein kleines Beispiel explizit ausrechnen. Wir betrachten

$$\mathbf{A} = \begin{pmatrix} 1 & -1 \\ -\frac{1}{2} & \frac{3}{2} \end{pmatrix} \quad (4.63)$$

Das Charakteristische Polynom ist

$$\begin{aligned} Z(\lambda) &= \det \begin{pmatrix} 1-\lambda & -1 \\ -\frac{1}{2} & \frac{3}{2}-\lambda \end{pmatrix} = (1-\lambda) \left( \frac{3}{2}-\lambda \right) - \frac{1}{2} \\ &= \lambda^2 - \frac{5}{2}\lambda + 1 = (\lambda-2) \left( \lambda + \frac{1}{2} \right) \end{aligned}$$

Die Eigenwerte von  $\mathbf{A}$  sind damit  $\lambda_1 = 2$  und  $\lambda_2 = -\frac{1}{2}$ . Zur Bestimmung der zugehörigen Eigenvektoren finden wir Lösungen des LGS

$$(\mathbf{A} - 2\mathbf{1}) \vec{v}_1 = \begin{pmatrix} -1 & -1 \\ -\frac{1}{2} & -\frac{1}{2} \end{pmatrix} \vec{v}_1 = \vec{0}. \quad (4.64)$$

Diese sind

$$\vec{v}_1 = \mu \begin{pmatrix} 1 \\ -1 \end{pmatrix}. \quad (4.65)$$

Für den anderen Eigenwert betrachten wir analog

$$\left( \mathbf{A} + \frac{1}{2}\mathbf{1} \right) \vec{v}_2 = \begin{pmatrix} \frac{1}{2} & -1 \\ -\frac{1}{2} & 1 \end{pmatrix} \vec{v}_2 = \vec{0} \quad \vec{v}_2 = \mu \begin{pmatrix} 2 \\ 1 \end{pmatrix}. \quad (4.66)$$

### 4.6.3 Eigenschaften von Eigenwerten und -vektoren

#### 4.6.3.1 Nullstellen des Charakteristischen Polynoms

Das charakteristische Polynom vom Grad  $n$  hat nach dem schon zu Beginn des Semesters erwähnten *Fundamentalsatz der Algebra*  $n$  komplexe Nullstellen, wobei einige von ihnen zusammenfallen können. Im rein reellen können es deutlich weniger Nullstellen sein. Das heißt, eine rein reelle Matrix kann durchaus komplexe Eigenwerte haben. Ein in der Quantenmechanik sehr wichtiges Beispiel ist

$$\mathbf{A} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \quad (4.67)$$

mit Eigenwerten  $\pm i$  und Eigenvektoren  $\begin{pmatrix} 1 \\ \pm i \end{pmatrix}$ . Darum hat das charakteristische Polynom auch die Form

$$Z(\lambda) = (-1)^n \prod_{i=1}^n (\lambda - \lambda_i) \quad (4.68)$$

Insbesondere ist  $Z(0) = \det \mathbf{A} = \prod_{i=1}^n \lambda_i$  das Produkt der Eigenwerte (einschließlich deren Multiplizität).



### 4.6.3.2 Determinante und Invertierbarkeit

Die Bedingung für die Existenz des Eigenwertes null sind nicht verschwindende Lösungen von

$$(\mathbf{A} - 0\mathbf{1}) \vec{v} = \mathbf{A}\vec{v} = \vec{0} \quad (4.69)$$

also die gleiche Bedingung wie für nicht-Invertierbarkeit. Eine Matrix ist also genau dann invertierbar, wenn sie den Eigenwert null nicht besitzt. Dies passt dazu, dass das Verschwinden der Determinante die Nicht-Invertierbarkeit anzeigt.

### 4.6.3.3 Lineare Unabhängigkeit verschiedener Eigenvektoren

Wir zeigen, dass Eigenvektoren zu verschiedenen Eigenwerten linear unabhängig sind. Sei  $\mathcal{M} = \{\vec{v}_i \neq \vec{0}, i = 1 \dots m, \mathbf{A}\vec{v}_i = \lambda_i \vec{v}_i, \lambda_i \neq \lambda_j \text{ für } i \neq j\}$  eine solche Menge an Eigenvektoren. Wir gehen per Induktion über  $i$  vor. Der Induktionsanfang ist einfach: Eine Menge mit einem einzigen nichtverschwindenden Vektor ist immer linear unabhängig. Die Induktionshypothese gelte bei  $j < m$ , die Teilmenge bis  $j$  sei bereits linear unabhängig. Wenn jetzt der neu hinzukommende Vektor  $\vec{v}_{j+1}$  die Menge linear abhängig macht, dann existieren Koeffizienten  $c_k$  so, dass gilt

$$\vec{0} = \mathbf{A}\vec{0} = \mathbf{A} \sum_{k=0}^{j+1} c_k \vec{v}_k = \sum_{k=0}^{j+1} c_k \lambda_k \vec{v}_k \quad (4.70)$$

sowie analog

$$\vec{0} = \lambda_{j+1} \vec{0} = \sum_{k=0}^{j+1} c_k \lambda_{j+1} \vec{v}_k. \quad (4.71)$$

Wenn wir diese beiden Gleichungen subtrahieren erhalten wir

$$\vec{0} = \sum_{k=0}^j c_k (\lambda_k - \lambda_{j+1}) \vec{v}_k \quad (4.72)$$

Da hier der letzte Summand herausgefallen ist und die Vektoren bis  $\vec{v}_j$  linear unabhängig sind, und da  $\lambda_k \neq \lambda_{j+1}$  nach Voraussetzung (alle Eigenvektoren zu verschiedenen Eigenwerten) folgt daraus, dass alle  $c_k = 0$  sind. Dann muss auch  $c_{j+1} = 0$  sein und die Vektoren sind in der Tat linear unabhängig.

## 4.6.4 Diagonalisierbarkeit und Diagonalisierung

Ausgangspunkt unserer Überlegung war es, dass wir gerne eine Basis finden möchten, in der eine Matrix / die von ihr dargestellte lineare Abbildung Diagonalform hat. Dies hat uns zum Konzept von Eigenwert/Eigenvektor geführt. Unser ursprüngliches Ziel ist dann erfüllt, wenn wir eine Basis aus Eigenvektoren haben - genau dann ist die Matrix *diagonalisierbar*.

#### 4.6.4.1 Nichtentartetes Spektrum

In dem Fall, dass alle  $n$  Eigenwerte verschieden sind, also keine Nullstellen des charakteristischen Polynoms zusammenfallen, ist die Sache einfach. Man nennt alle einfachen Eigenwerte dieser Art *nichtentartet*. Zu jedem nichtentarteten Eigenwert haben wir einen eindimensionalen Untervektorraum als *Eigenraum* und da die Eigenvektoren linear unabhängig sind und es  $n$  Eigenwerte gibt, existiert eine Basis des zugrundeliegenden Vektorraums aus Eigenvektoren.

#### 4.6.4.2 Entartetes Spektrum

Wenn mehrere Eigenwerte zusammenfallen, ist die Sache ambivalent. Wenn ein Eigenwert  $k$  fach entartet ist, kann ein bis zu  $k$  dimensionaler Eigenvektorraum zu diesem Eigenwert vorliegen, muss aber nicht - die Eigenraumdimension kann kleiner sein. Wenn dies passiert, ist die Matrix *nicht diagonalisierbar*. Dies illustrieren wir an Beispielen

Wir haben oben schon die Matrix hingeschrieben, die die Ableitung in der Basis von Monomen beschreibt.

$$\mathbf{D} = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \\ 0 & 0 & 0 & 0 \end{pmatrix} \quad (4.73)$$

Das charakteristische Polynom ist  $Z(\lambda) = \lambda^4$  d.h. es gibt nur den vierfach entarteten Eigenwert  $\lambda = 0$ . Dennoch ist der einzige Eigenraum eindimensional

$$\text{span} \left\{ \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} \right\} \quad (4.74)$$

In der Quanteninformation ist folgende Matrix wichtig

$$\mathbf{C} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix}. \quad (4.75)$$

Das charakteristische Polynom ist  $Z(\lambda) = (1 - \lambda)^2 (\lambda^2 - 1) = (\lambda - 1)^3 (\lambda + 1)$ . Es gibt einen dreifach entarteten Eigenwert mit  $\lambda = 1$ . Der zugehörige Eigenraum ist tatsächlich

$$\text{span} \left\{ \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \\ 1 \end{pmatrix} \right\} \quad (4.76)$$

### 4.6.4.3 Spektralsatz

Auch hier haben wir bisher keinen a priori-test für Diagonalisierbarkeit. Eine Hilfe ist der *Spektralsatz*, der in der Mathematik bewiesen wird: Eine Matrix ist genau dann diagonalisierbar, wenn

$$\mathbf{A}\mathbf{A}^T = \mathbf{A}^T\mathbf{A}. \quad (4.77)$$

Eine solche Matrix nennt man auch *normale Matrix*.

### 4.6.4.4 Diagonalisierung

Wie oben ausgeführt ist Diagonalisierbarkeit äquivalent mit der Existenz einer Basis aus Eigenvektoren. Sind diese berechnet, dann können wir die Basiswechsellmatrix  $\mathbf{T}$  aufstellen, indem wir die Eigenvektoren nebeneinander schreiben

$$\mathbf{T} = \left( \vec{v}_1 \quad \vec{v}_2 \quad \cdots \quad \vec{v}_n \right) \quad (4.78)$$

Dann ist

$$\mathbf{A} = \mathbf{T} \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_n) \mathbf{T}^{-1} \quad (4.79)$$

und somit die Matrix diagonalisiert.

## 4.6.5 Beispiele und spezielle Matrizen

### 4.6.5.1 Drehmatrizen

Wir haben oben schon die Drehmatrix

$$\mathbf{R}_\phi = \begin{pmatrix} \cos \phi & -\sin \phi \\ \sin \phi & \cos \phi \end{pmatrix} \quad \mathbf{R}_\phi^T = \begin{pmatrix} \cos \phi & \sin \phi \\ -\sin \phi & \cos \phi \end{pmatrix} \quad \mathbf{R}_\phi \mathbf{R}_\phi^T = \mathbf{1} = \mathbf{R}_\phi^T \mathbf{R}_\phi \quad (4.80)$$

eingeführt. Sie ist normal, also diagonalisierbar. Das charakteristische Polynom ist  $Z(\lambda) = (\cos \phi - \lambda)^2 + \sin^2 \phi = \lambda^2 - 2\lambda \cos \phi + 1$ . Seine Nullstellen sind

$$\lambda_{1/2} = \cos \phi \pm \sqrt{\cos^2 \phi - 1} = e^{\mp i\phi} \quad (4.81)$$

Auch hier haben wir also eine reelle Matrix mit komplexen Eigenwerten. Wir berechnen die Eigenvektoren durch Gaußelimination in der Matrix

$$\mathbf{R}_\phi - e^{-i\phi} \mathbf{1} = \begin{pmatrix} i \sin \phi & \sin \phi \\ -\sin \phi & i \sin \phi \end{pmatrix} \quad \vec{v}_1 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -i \end{pmatrix} \quad (4.82)$$

und analog haben wir für den anderen Eigenwert

$$\vec{v}_2 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ i \end{pmatrix} \quad (4.83)$$

Damit ist

$$\mathbf{T} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ -i & i \end{pmatrix} \quad \mathbf{T}^{-1} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & i \\ 1 & -i \end{pmatrix} \quad (4.84)$$

Und somit

$$\mathbf{R}_\phi = \frac{1}{2} \begin{pmatrix} 1 & 1 \\ -i & i \end{pmatrix} \begin{pmatrix} e^{-i\phi} & 0 \\ 0 & e^{i\phi} \end{pmatrix} \begin{pmatrix} 1 & i \\ 1 & -i \end{pmatrix} \quad (4.85)$$

### 4.6.5.2 Orthogonale Matrizen

Matrizen, die das Skalarprodukt zwischen Vektoren erhalten

$$\langle \mathbf{A}\vec{v}, \mathbf{A}\vec{v} \rangle = \langle \vec{v}, \vec{v} \rangle \quad (4.86)$$

heißen *orthogonal*. Im Anschauungsraum sind dies die längenerhaltenden Abbildungen, in fortgeschrittenen Anwendungen lassen sich auch andere Erhaltungsgrößen oft durch solche Matrizen ausdrücken. Orthogonale Matrizen sind invertierbar, da

$$\langle \mathbf{A}\vec{v}, \mathbf{A}\vec{v} \rangle = 0 \Rightarrow \langle \vec{v}, \vec{v} \rangle = 0 \Rightarrow \vec{v} = \vec{0}. \quad (4.87)$$

Das Produkt von zwei orthogonalen Matrizen ist orthogonal

$$\langle \mathbf{A}\mathbf{B}\vec{v}, \mathbf{A}\mathbf{B}\vec{v} \rangle = \langle \mathbf{B}\vec{v}, \mathbf{B}\vec{v} \rangle = \langle \vec{v}, \vec{v} \rangle \quad (4.88)$$

und ähnlich kann man zeigen, dass die Inverse einer orthogonalen Matrix wieder orthogonal ist.

Wie können wir diese Eigenschaft an einer Matrix ablesen? Wir schreiben die definierende Eigenschaft im  $\mathbb{R}^n$  aus

$$\begin{aligned} \langle \mathbf{A}\vec{v}, \mathbf{A}\vec{v} \rangle &= \sum_{i=1}^n (\mathbf{A}\vec{v})_i (\mathbf{A}\vec{v})_i = \sum_{i,j,k} A_{ij} v_j A_{ik} v_k \\ &= \sum_{j,k} v_j v_k \sum_i A_{ji}^T A_{ik} = \sum_j v_j v_j. \end{aligned}$$

Das wir  $\mathbf{A}$  so einschränken müssen, dass dies unabhängig von der Wahl der  $v_j$  und  $v_k$  gilt, erzwingt dies

$$\sum_i A_{ji}^T A_{ik} = \delta_{jk} \quad (4.89)$$

Diese Gleichungen können wir kompakt und im Ergebnis transparenter in Matrixform schreiben

$$\langle \mathbf{A}\vec{v}, \mathbf{A}\vec{v} \rangle = (\mathbf{A}\vec{v})^T (\mathbf{A}\vec{v}) = \vec{v}^T \mathbf{A}^T \mathbf{A} \vec{v} = \vec{v}^T \vec{v} \quad (4.90)$$

was wenn wir eine Basis von  $\vec{v}$  s abarbeiten erzwingt, dass  $\mathbf{A}^T \mathbf{A} = \mathbf{1}$  ist, also  $\mathbf{A}^{-1} = \mathbf{A}^T$ . Kompakt kann man also sagen, dass die Spalten von  $\mathbf{A}$  orthonormiert sind - orthogonale Matrizen bilden eine ONB auf eine ONB ab. Laut Konstruktion sind diese Bedingungen notwendig und hinreichend.

Schon bei der Diskussion der inversen Matrix haben wir gezeigt, dass Rechtsinverse = Linksinverse ist. Damit ist auch  $\mathbf{A}\mathbf{A}^T = \mathbf{A}^T \mathbf{A}$  und somit sind orthogonale Matrizen normal, also sind sie diagonalisierbar. Um auch die Länge der Eigenvektoren zu erhalten, müssen die Eigenwerte alle vom Betrag eins sein, haben also die Form  $\lambda = e^{i\phi}$

### 4.6.5.3 Symmetrische Matrizen

Eine Matrix heißt symmetrisch wenn  $\mathbf{A} = \mathbf{A}^T$  ist. Diese sind dann automatisch normal und somit diagonalisierbar. Für allgemeine lineare Abbildungen können wir diese Eigenschaft über das Skalarprodukt definieren

$$\langle \vec{v}, \mathbf{A}\vec{w} \rangle = \langle \mathbf{A}\vec{v}, \vec{w} \rangle. \quad (4.91)$$

Zum Zeigen der Äquivalenz beider Definitionen betrachten wir diesen Ausdruck in der Standardbasis

$$\hat{e}_j^T (\mathbf{A}\hat{e}_i) = A_{ji} = (\mathbf{A}\hat{e}_j)^T \hat{e}_i = \hat{e}_j \mathbf{A}^T \hat{e}_i = A_{ij}. \quad (4.92)$$

Außerdem sind Eigenvektoren zu verschiedenen Eigenwerten automatisch orthogonal bzgl. des Standardskalarproduktes: Seien  $\vec{v}_i/\vec{v}_j$  zwei verschiedene Eigenvektoren zu verschiedenen Eigenwerten, dann ist

$$\lambda_i \langle \vec{v}_i, \vec{v}_j \rangle = \langle \mathbf{A}\vec{v}_i, \vec{v}_j \rangle = \langle \vec{v}_i \mathbf{A}\vec{v}_j \rangle = \lambda_j \langle \vec{v}_i, \vec{v}_j \rangle. \quad (4.93)$$

Dies ist bei  $\lambda_i \neq \lambda_j$  nur erfüllbar, wenn  $\langle \vec{v}_i, \vec{v}_j \rangle = 0$  ist. Damit können wir immer eine ONB aus Eigenvektoren bilden: In entarteten Eigenräumen können wir mit dem Gram-Schmidt-Verfahren immer eine orthogonale Eigenbasis finden. Damit wird eine symmetrische Basis immer durch eine orthogonale Transformation diagonalisiert.

## 4.7 Matrizen und lineare Abbildungen in komplexen Vektorräumen

Wir haben uns bisher auf reelle Vektorräume konzentriert. Es spricht aber nichts dagegen, diese gesamte Maschinerie auch in komplexen Vektorräumen anzuschauen, die sich dann im  $\mathbb{C}^n$  darstellen lassen. Die wichtigste Änderung ergibt sich bereits beim Skalarprodukt: Um zu garantieren, dass auch komplexe Vektoren wie z.B.  $(i \ 0)^T$  eine reelle, nichtnegative Länge haben ist das komplexe Skalarprodukt

$$\langle \vec{w}, \vec{v} \rangle = \sum_i w_i^* v_i = \vec{w}^{*,T} \vec{v} \quad (4.94)$$

Generell nennt man eine Matrix die wie hier  $\vec{w}$  als  $n \times 1$  Matrix komplex konjugiert und transponiert wird auch die hermitesch konjugierte Matrix und schreibt  $\mathbf{A}^\dagger = \mathbf{A}^{*,T}$ . Man sagt eigentlich „A-Kreuz“, der Anglizismus ist aber „A-Dagger“ (A-Dolch). Die Eigenschaften, die wir transponierten Matrizen zugeschrieben haben gelten analog für konjugierte Matrizen.

Das Gegenstück zur orthogonalen Matrix heißt *unitär* und hat  $\mathbf{U}^\dagger = \mathbf{U}^{-1}$  als definierenden Eigenschaft. Das Gegenstück zur symmetrischen Matrix heißt *hermitesch* und ist charakterisiert durch  $\mathbf{A}^\dagger = \mathbf{A}$ .

### 4.7.1 Eigenwerte hermitescher Matrizen

Damit ausgerüstet können wir noch eine wichtige Eigenschaft hermitescher / symmetrischer Matrizen betrachten: Reelle Eigenwerte. Dies benötigt komplexe Vektorräume weil die Eigenvektoren diagonalisierbarer symmetrischer Matrizen durchaus komplex sein können. Dies ist aber bei hermiteschen Matrizen nicht der Fall. Wir betrachten

$$\lambda_i \langle \vec{v}_i, \vec{v}_i \rangle = \langle \vec{v}_i, \mathbf{A}\vec{v}_i \rangle = \langle \mathbf{A}\vec{v}_i, \vec{v}_i \rangle = \lambda_i^* \langle \vec{v}_i, \vec{v}_i \rangle \quad (4.95)$$

was  $\lambda_i = \lambda_i^*$  erzwingt.

# Kapitel 5

## Gewöhnliche Differenzialgleichungen

### 5.1 Einführung

Wir haben bereits gelernt, algebraische Gleichungen aufzulösen. Differenzialgleichungen sind Gleichungen, in denen Funktionen und ihre Ableitungen miteinander verknüpft sind. Die allgemeine Form der Newtonschen Bewegungsgleichung in einer Dimension

$$m\ddot{x} = F(\dot{x}, x, t) \quad (5.1)$$

ist eine solche Differenzialgleichung. Wir haben schon gelernt, wie man die im Spezialfall, dass  $x$  und  $\dot{x}$  auf der rechten Seite nicht vorkommen löst, nämlich durch zweifaches Integrieren

$$\ddot{x} = \frac{F(t)}{m} \quad \dot{x} = \frac{1}{m} \int_{t_0}^t d\tau F(\tau) + v_0 \quad x(t) = \frac{1}{m} \int_{t_0}^t d\bar{\tau} \int_{t_0}^{\bar{\tau}} d\tau F(\tau) + v_0(t-t_0) + x_0 \quad (5.2)$$

Dieses sehr einfache Beispiel einer „unechten“ Differenzialgleichung enthält schon Eigenschaften, die uns immer wieder begegnen werden: Es wird integriert, und die Lösung der Differenzialgleichung allein ist nicht eindeutig - eindeutig wird sie erst dann, wenn die Anfangsbedingungen, hier für Ort und Geschwindigkeit, angegeben werden. Die Kombination aus Differenzialgleichung und diesen Werten nennt man auch *Anfangswertproblem*.

Wir behandeln hier nur Differenzialgleichungen von Funktionen einer Variablen, gewöhnliche Differenzialgleichungen. Die Differenzialgleichungen der Vektoranalysis heißen partielle Differenzialgleichungen. Die Ordnung einer Differenzialgleichung ist gegeben durch die höchste vorkommende Ableitung - das Beispiel ist also eine DGI zweiter Ordnung. Bei der Behandlung von Differenzialgleichungen in einer Vorlesung wie unserer hat man bei Erstkontakt manchmal den Eindruck, dass es hier hauptsächlich um ad-Hoc Lösungsverfahren für bestimmte Gleichungstypen geht. Das ist richtig - aber das ist in der „echten“

Mathematikvorlesung auch nicht anders. Dort werden die Sätze über die Existenz und Eindeutigkeit von Lösungen sorgfältiger bewiesen als bei uns.

## 5.2 Differenzialgleichungen mit getrennten Veränderlichen

Ein wichtiger erster Fall ist die Situation, dass die Differenzialgleichung getrennte Variable hat, d.h. sich so auflösen lässt, dass auf einer Seite nur die Funktion und auf der anderen Seite nur die Variable steht. Wir suchen

$$\frac{df}{dx} = g(x)h(f) \quad (5.3)$$

dies können wir auflösen zu

$$\frac{df}{h(f)} = g(x)dx \quad (5.4)$$

Wenn der Anfangswert  $f_0 = f(x_0)$  vorgegeben ist, integriert sich dies zu

$$\int_{f_0}^f \frac{d\bar{f}}{h(\bar{f})} = \int_{x_0}^x g(\bar{x})d\bar{x}. \quad (5.5)$$

Dies ist eine algebraische Gleichung für  $f(x)$ , deren Lösung uns die gewünschte Funktion gibt.

Diese simpel hingeschriebene Methode illustrieren wir an Beispielen. Wir starten mit

$$\frac{df}{dx} = s \frac{f}{x} \quad (5.6)$$

Wir lösen auf zu

$$\frac{df}{f} = s \frac{dx}{x} \quad (5.7)$$

integriert zu

$$\log \frac{f}{f_0} = s \log \frac{x}{x_0}. \quad (5.8)$$

Beide Seiten potenzieren liefert

$$f = f_0 \left( \frac{x}{x_0} \right)^s = g_0 x^s \quad (5.9)$$

wobei wir die Integrationskonstanten zusammengefasst haben. Das hätten wir natürlich auch geraten: Die Ableitung eines Monoms dividiert durch die Variable und multipliziert mit dem Exponenten.

Betrachten wir den etwas allgemeineren Fall (mit  $\alpha \neq -1$ )

$$\frac{df}{dx} = sx^\alpha f \quad (5.10)$$



Wir lösen wiederum auf zu

$$\frac{df}{f} = sx^\alpha dx \quad (5.11)$$

bzw.

$$\log \frac{f}{f_0} = s \int_{x_0}^x d\bar{x} x^\alpha = \frac{s}{\alpha + 1} (x^{\alpha+1} - x_0^{\alpha+1}) \quad (5.12)$$

Also erhalten wir

$$f = g_0 \exp\left(\frac{sx^{\alpha+1}}{\alpha + 1}\right). \quad (5.13)$$

Eine sehr praktische Anwendung dieses Ergebnisses findet sich in der eindimensionalen Punktmechanik. Wenn wir dort nur eine konservative Kraft haben, dann ist die Energie erhalten

$$E = \frac{m}{2} \left(\frac{dx}{dt}\right)^2 + V(x). \quad (5.14)$$

Dies können wir bis auf ein Vorzeichen nach der Geschwindigkeit auflösen

$$\frac{dx}{dt} = \pm \sqrt{\frac{2}{m} (E - V(x))}. \quad (5.15)$$

Das Vorzeichen hängt von der Bewegungsrichtung ab - zu größeren oder zu kleineren  $x$ . Wir lösen auf zu

$$t - t_0 = \int_{x_0}^x \frac{d\bar{x}}{\sqrt{\frac{2}{m} (E - V(\bar{x}))}}. \quad (5.16)$$

Wenn wir das Integral auf der rechten Seite lösen können, dann haben wir  $t(x)$  und wenn diese Funktion umkehrbar ist haben wir  $x(t)$ . Dies ist also eigentlich eine Strategie zur exakten Lösung der Newtonschen Bewegungsgleichung - allerdings ist in der Praxis das Integral oft schwer zu knacken.

Ein interessanter Spezialfall ist die Periode einer gebundenen Bewegung. Umkehrpunkte der Bewegung haben Sie schon oft ausgerechnet, es sind die Lösungen der Gleichung

$$V(x_i) = E \quad (5.17)$$

Zwischen diesen bewegt sich ein Teilchen periodisch hin und her. Die Periode ist dann zwei mal die Zeit zwischen zwei Umkehrpunkten

$$T = \sqrt{2m} \int_{x_n}^{x_{n+1}} \frac{d\bar{x}}{\sqrt{E - V(\bar{x})}}. \quad (5.18)$$

Als Beispiel betrachten wir das Potenzgesetzpotenzial  $V(x) = ax^s$  mit  $a, s > 0$ . Diese hat zwei Umkehrpunkte

$$x_{1/2} = \pm \left(\frac{E}{a}\right)^{1/s}. \quad (5.19)$$

Alternativ können wir die Energie ausdrücken durch  $E = ax_1^s$ . Die Periode dieser Bewegung ist

$$T = \sqrt{\frac{2m}{a}} \int_{-x_1}^{x_1} \frac{d\bar{x}}{\sqrt{x_1^s - x^s}} = \sqrt{\frac{8m}{a}} \int_0^{x_1} \frac{d\bar{x}}{\sqrt{x_1^s - \bar{x}^s}}. \quad (5.20)$$

Im letzten Schritt haben wir ausgenutzt, dass der Integrand gerade ist. Das Integral können wir dimensionslos machen, indem wir  $\xi = \bar{x}/x_1$  substituieren und erhalten

$$T = \sqrt{\frac{8m}{a}} x_1 \int_0^1 \frac{d\xi}{\sqrt{x_1^s - x_1^s \xi^s}} = \sqrt{\frac{8m}{a}} x_1^{1-s/2} \int_0^1 \frac{d\xi}{\sqrt{1 - \xi^s}}. \quad (5.21)$$

Auch ohne Berechnung des verbleibenden Integrals (das ja nur noch von  $s$  abhängt, also nur eine Zahl ist) sehen wir etwas erstaunliches: Für den Fall  $s = 2$  hängt das Integral nicht von der Bewegungsamplitude  $x_1$  ab (und damit nicht von der Periode). Hier stehen zwei Faktoren im Wettbewerb: Wenn die Energie steigt, dann wird das Teilchen schneller, muss aber auch einen längeren Weg zurücklegen. Bei  $s = 2$  heben sich diese beiden Phänomene weg. Bei weicheren Potenzialen,  $0 < s < 2$  wächst die Periode mit der Energie, sonst fällt sie. Das Integral am Ende ist

$$\int_0^1 \frac{d\xi}{\sqrt{1 - \xi^s}} = \sqrt{\pi} \frac{\Gamma(1 + \frac{1}{s})}{\Gamma(\frac{1}{2} + \frac{1}{s})}. \quad (5.22)$$

Die hier eingeführte Funktion ist eine spezielle transzendente Funktion, die Eulersche Gammafunktion. Diese ist eine Verallgemeinerung der Fakultät, insbesondere ist

$$\Gamma(n+1) = \Gamma(n) \quad (5.23)$$

### 5.2.1 Autonome Differenzialgleichungen

Ein Spezialfall der Differenzialgleichung mit getrennten Veränderlichen ist die *autonome Differenzialgleichung*

$$\frac{df}{dx} = g(f). \quad (5.24)$$

Die Lösung folgt dem gerade gerechneten Beispiel, es ist

$$\int \frac{df}{g(f)} = x - x_0 \quad (5.25)$$

Ein Beispiel für eine autonome DGL ist die *logistische Gleichung*. Sie beschreibt die Dynamik einer Population mit einer natürlichen Wachstumsrate  $\Gamma$  mit einer Kapazitätsbegrenzung wenn zu viele Individuen da sind

$$\frac{df}{dt} = cf(\Gamma - f). \quad (5.26)$$

Wir starten von einer Anfangspopulation  $f(0) = f_0$ . Variablentrennung liefert

$$\begin{aligned} t &= \frac{1}{c} \int_{f_0}^f d\bar{f} \frac{1}{\bar{f}(\Gamma - \bar{f})} = \frac{1}{\Gamma c} \int_{f_0}^f d\bar{f} \left[ \frac{1}{\bar{f}} + \frac{1}{\Gamma - \bar{f}} \right] \\ &= \frac{1}{\Gamma c} \log \left( \frac{\bar{f}}{\Gamma - \bar{f}} \right) \Big|_{f_0}^f = \frac{1}{\Gamma c} \log \left( \frac{f}{\Gamma - f} \frac{\Gamma - f_0}{f_0} \right). \end{aligned}$$

Wir exponenzieren

$$\frac{1}{\Gamma/f - 1} \frac{\Gamma - f_0}{f_0} = e^{\Gamma ct} \quad (5.27)$$

isolieren die Funktion

$$\frac{\Gamma}{f} - 1 = \frac{\Gamma - f_0}{f_0} e^{-\Gamma ct} \quad (5.28)$$

und lösen auf zu

$$f = \Gamma \left( 1 + \frac{\Gamma - f_0}{f_0} e^{-\Gamma ct} \right)^{-1} = \frac{\Gamma f_0}{f_0 + (\Gamma - f_0) e^{-\Gamma ct}}. \quad (5.29)$$

Diese Lösungen haben eine qualitativ interessante Struktur: Im Fall langer Zeiten erreichen wir  $\lim_{t \rightarrow \infty} f(t) = \frac{\Gamma}{f_0}$  unabhängig von der Anfangsbedingung.

### 5.3 Lineare Differenzialgleichungen

Lineare Differenzialgleichungen sind solche, in denen die gesuchte Funktion und Ableitung nur in linearer Potenz vorkommen, d.h.

$$\frac{df}{dx} = g(x)f + h(x). \quad (5.30)$$

Den Fall  $h(x) = 0$  nennen wir eine *homogene* lineare Differenzialgleichung, sonst sprechen wir von einer *inhomogenen* Gleichung.

#### 5.3.1 Homogene lineare Differenzialgleichungen

In diesem Fall haben wir nur

$$\frac{df}{dx} = g(x)f \quad (5.31)$$

In diesem Fall gilt das *Superpositionsprinzip*. Sind  $f_1$  und  $f_2$  Lösungen der genannten Gleichung (z.B. zu verschiedenen Anfangswerten) dann ist aufgrund der Linearität auch  $\lambda f_1 + \mu f_2$  eine Lösung. Da diese Gleichung ein Spezialfall der Variablentrennung ist, können wir die Lösung auch gleich angeben als

$$f(x) = a \exp \left( \int_{x_0}^x d\bar{x} g(\bar{x}) \right). \quad (5.32)$$

Die ersten Beispiele oben waren alle linear.

### 5.3.2 Inhomogene lineare Differenzialgleichungen

Wenn wir jetzt die vollständige lineare Differenzialgleichung anschauen, dann gilt zunächst einmal nicht das ursprüngliche Superpositionsprinzip, und die Variablen lassen sich auch nicht trennen. Was tun? Wir beobachten, dass wenn  $f_0(x)$  eine spezielle Lösung der inhomogenen Gleichung und  $f_1(x)$  eine Lösung der entsprechenden homogenen Gleichung ist, dann haben wir auch wieder eine Lösung der inhomogenen DGL

$$\frac{d}{dx} (f_0 + f_1) = g(x)f_0(x) + h(x) + g(x)f_1(x) = g(x)f(x) + h(x). \quad (5.33)$$

Das bedeutet, als allgemeine Strategie, dass wir uns ein  $f_0$  beschaffen müssen, das die Differenzialgleichung für irgendeine Anfangsbedingung löst) und wir dann das richtige  $f_1$  suchen müssen, so dass wir die richtige Anfangsbedingung erhalten. Dazu benutzen wir die Methode der *Variation der Konstanten*. Wir machen für die Lösung der inhomogenen DGL den Ansatz

$$f_0(x) = a(x) \exp\left(\int_{x_0}^x d\bar{x} g(\bar{x})\right) \quad (5.34)$$

d.h. wir behandeln die Integrationskonstante der homogenen Lösung wie eine Variable. Diesen Ansatz setzen wir in die DGL ein und erhalten

$$\begin{aligned} \frac{d}{dx} \left( a(x) \exp\left(\int_{x_0}^x d\bar{x} g(\bar{x})\right) \right) &= \frac{da}{dx} \exp\left(\int_{x_0}^x d\bar{x} g(\bar{x})\right) + g(x)a(x) \exp\left(\int_{x_0}^x d\bar{x} g(\bar{x})\right) \\ &= \frac{da}{dx} \exp\left(\int_{x_0}^x d\bar{x} g(\bar{x})\right) + g(x)f_0(x) = g(x)f_0(x) + h(x). \end{aligned}$$

Damit können wir auflösen zu

$$\frac{da}{dx} = h(x) \exp\left(-\int_{x_0}^x d\bar{x} g(\bar{x})\right). \quad (5.35)$$

Da  $a(x)$  auf der rechten Seite gar nicht mehr vorkommt, müssen wir nur noch einmal integrieren und erhalten

$$a(x) = a_0 + \int_{x_0}^x dx_1 h(x_1) \exp\left(-\int_{x_0}^{x_1} d\bar{x} g(\bar{x})\right) \quad (5.36)$$

und damit ist die allgemeine Lösung

$$\begin{aligned} f_0(x) &= a_0 \exp\left(\int_{x_0}^x d\bar{x} g(\bar{x})\right) + \int dx_1 h(x_1) \exp\left(-\int_{x_0}^{x_1} d\bar{x} g(\bar{x})\right) \exp\left(\int_{x_0}^x d\bar{x} g(\bar{x})\right) \\ &= a_0 \exp\left(\int_{x_0}^x d\bar{x} g(\bar{x})\right) + \int_{x_0}^x dx_1 h(x_1) \exp\left(-\int_x^{x_1} d\bar{x} g(\bar{x})\right) \end{aligned}$$

Wir können entweder in diese Gleichung einsetzen, oder die Schritte ihrer Herleitung wiederholen. Als Beispiel betrachten wir die Bewegungsgleichung eines Teilchen mit Reibung unter einer äußeren Kraft

$$\dot{v}(t) = -\gamma v(t) + F(t). \quad (5.37)$$

Wir wollen dies mit einer Anfangsgeschwindigkeit  $v(0)$  lösen. Die Lösung der homogenen DGL ist

$$v(t) = v_0 \exp\left(-\gamma \int_0^t d\bar{t}\right) = v_0 e^{-\gamma t}. \quad (5.38)$$

Für die spezielle inhomogene Lösung setzen wir  $v_0(t)$  an und erhalten

$$\frac{dv}{dt} = \frac{dv_0}{dt} e^{-\gamma t} - \gamma v_0(t) e^{-\gamma t} = -\gamma v(t) + \frac{dv_0}{dt} e^{-\gamma t} = -\gamma v(t) + F(t). \quad (5.39)$$

Damit haben wir für den Vorfaktor

$$\frac{dv_0}{dt} = F(t) e^{\gamma t} \quad (5.40)$$

und somit

$$v_0(t) = \int_0^t d\tau F(\tau) e^{\gamma \tau} \quad (5.41)$$

und somit die volle Lösung

$$v(t) = v_0 e^{-\gamma t} + \int_0^t d\tau F(\tau) e^{-\gamma(t-\tau)} = v_0 e^{-\gamma t} + \int_0^t d\tau F(t-\tau) e^{-\gamma \tau}. \quad (5.42)$$

Diese Lösung zeigt einige Charakteristika von solchen angetriebenen Systemen: Nach einer Zeit von etwa  $1/\gamma$  spielt die Anfangsbedingung keine große Rolle mehr und auch an Kräfte erinnert sich das System über die etwa gleiche Dauer - der Rest der Vorgeschichte landet in der Reibung. Im Fall einer konstanten Kraft haben wir

$$v(t) = v_0 e^{-\gamma t} + F \int_0^t d\tau e^{-\gamma \tau} = v_0 e^{-\gamma t} + \frac{F}{\gamma} (1 - e^{-\gamma t}). \quad (5.43)$$

Das bedeutet, dass Teilchen startet mit  $v = v_0$  und endet mit  $v = F/\gamma$ . Bei dieser Endgeschwindigkeit geht die Energie, die durch die Kraft zugeführt wird, in die Reibung über. Dazwischen interpoliert die Geschwindigkeit mit einer Exponentialfunktion. Im Fall einer harmonischen Kraft  $F \cos \omega t$  haben wir für  $v_0 = 0$

$$\begin{aligned} v(t) &= F e^{-\gamma t} \int_0^t d\tau \cos \omega \tau e^{\gamma \tau} = F e^{-\gamma t} \operatorname{Re} \left[ \int_0^t d\tau e^{(i\omega + \gamma)\tau} \right] \\ &= F e^{-\gamma t} \operatorname{Re} \left[ \frac{e^{(i\omega + \gamma)t} - 1}{i\omega + \gamma} \right] = \frac{F}{\gamma^2 + \omega^2} \operatorname{Re} [(-i\omega + \gamma)(e^{i\omega t} - e^{-\gamma t})] \\ &= \frac{F}{\gamma^2 + \omega^2} [\gamma \cos \omega t + \omega \sin \omega t + \gamma e^{-\gamma t}]. \end{aligned}$$

Das Teilchen folgt dem Antrieb also mit einer gewissen Phasenverschiebung.

### 5.3.3 Lineare Differenzialgleichungen mit konstanten Koeffizienten

Besonders einfach ist, was wir schon als Beispiel behandelt haben, die lineare DGL mit konstanten Koeffizienten

$$\frac{df}{dx} = cf \quad (5.44)$$

mit der Lösung  $f = f_0 e^{-cx}$ .

## 5.4 Systeme von Differenzialgleichungen

### 5.4.1 Definitionen

Ein System von Differenzialgleichungen ist, analog zum System von Gleichungen. Es enthält im Allgemeinen  $n$  Funktionen und  $m$  Gleichungen

$$\frac{df_i}{dx} = g_i(f_1, \dots, f_n, x) \quad x = 1 \dots m \quad (5.45)$$

wobei der Fall  $m = n$  der interessanteste ist. Dieser Fall schließt Differenzialgleichungen höherer Ordnung für eine einzelne Funktion ein. Eine einzige Differenzialgleichung der Ordnung  $n$

$$\frac{d^n f}{dx^n} = g(f, f', \dots, f^{(n-1)}, x) \quad (5.46)$$

kann umgeschrieben werden als System von  $n$  Differenzialgleichungen für die Funktionen  $f, f' \dots f^{(n-1)}$  als

$$\begin{aligned} \frac{df}{dx} &= f' \\ \frac{df'}{dx} &= f'' \\ &\vdots \\ \frac{df^{(n-2)}}{dx^{n-2}} &= f^{(n-1)} \\ \frac{d^n f}{dx^n} &= g(f, f', \dots, f^{(n-1)}, x) \end{aligned}$$

Ein einfaches Beispiel ist die eindimensionale Newtonsche Bewegungsgleichung

$$m \frac{d^2 x}{dt^2} = F(x, \dot{x}, t) \quad (5.47)$$

die wir aufteilen können als

$$\begin{aligned} \frac{dx}{dt} &= \dot{x} \\ \frac{d\dot{x}}{dt} &= \frac{1}{m} F(x, \dot{x}, t) \end{aligned}$$

Die Newtonschen Bewegungsgleichungen für ein Teilchen in drei Dimensionen sind schon sechs gekoppelte Gleichungen.

$$m \frac{d^2 r_i}{dt^2} = F_i(\vec{r}, t). \quad (5.48)$$

Für  $N$  Punktmassen haben wir sogar  $3N$  gekoppelte Differenzialgleichungen.

Ähnlich wie eine einzelne DGL erster Ordnung nur durch Angabe einer Integrationskonstante eindeutig lösbar ist, die typischerweise als Anfangswert angegeben wird, muss hier eine Integrationskonstante pro Gleichung angegeben werden. Dies kann wiederum als Anfangswert geschehen - für die Newtonschen Bewegungsgleichungen z.B. als Anfangswerte von Ort und Geschwindigkeit. Manchmal gibt man auch Randwerte an, d.h. man teilt die Angabe der  $N$  Integrationskonstanten auf Anfang und Ende auf - dies kann aber die Lösbarkeit einschränken.

### 5.4.2 Lineare Systeme mit konstanten Koeffizienten

Die Theorie der Differenzialgleichungssysteme bzw. Differenzialgleichungen mit konstanten Koeffizienten ist sehr reichhaltig. Wir behandeln hier den wichtigsten Spezialfall: Ein System mit konstanten Koeffizienten. Wir starten mit dem *homogenen System*: Für  $n$  Funktionen  $f_1, \dots, f_n$  sind diese

$$\begin{aligned} \frac{df_1}{dx} &= A_{11}f_1 + A_{12}f_2 + \dots + A_{1n}f_n \\ \frac{df_2}{dx} &= A_{21}f_1 + A_{22}f_2 + \dots + A_{2n}f_n \\ &\vdots \\ \frac{df_n}{dx} &= A_{n1}f_1 + A_{n2}f_2 + \dots + A_{nn}f_n \end{aligned}$$

Mit anderen Worten, wir können die Koeffizienten  $A_{ij}$  in einer Matrix versammeln und die Funktionen in einen Spaltenvektor und erhalten

$$\frac{d}{dx} \begin{pmatrix} f_1 \\ f_2 \\ \vdots \\ f_n \end{pmatrix} = \mathbf{A} \begin{pmatrix} f_1 \\ f_2 \\ \vdots \\ f_n \end{pmatrix}. \quad (5.49)$$

### 5.4.3 Matrixexponentiale

Zur (formalen) Lösung dieses Systems definieren wir die Exponentialfunktion einer Matrix  $\mathbf{M}$  über die Potenzreihenentwicklung

$$\exp(\mathbf{M}) = \sum_{n=0}^{\infty} \frac{\mathbf{M}^n}{n!}. \quad (5.50)$$

Diese Reihenentwicklung konvergiert - die Matrixpotenzen wachsen sicherlich nicht schneller als die Fakultät im Nenner. Die praktische Berechnung dieses Ausdrucks kann mehrere Wege nehmen<sup>1</sup>. In manchen Fällen werden die Matrixpotenzen sehr schnell auf die Null führen, z.B. bei

$$\mathbf{M} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \quad \mathbf{M}^2 = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} \dots \quad (5.51)$$

In diesem Beispiel ist

$$\exp(\mathbf{M}) = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \quad (5.52)$$

Wenn die zugrundeliegende Matrix diagonalisierbar ist, haben wir einen vereinfachten Zugang. Es ist dann

$$\mathbf{M} = \mathbf{TDT}^{-1} \quad \mathbf{D} = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_n). \quad (5.53)$$

Wir haben dann nämlich

$$\mathbf{M}^n = (\mathbf{TDT}^{-1})^n = \mathbf{TDT}^{-1}\mathbf{TDT}^{-1}\dots\mathbf{DT}^{-1} = \mathbf{TD}^n\mathbf{T}^{-1} \quad (5.54)$$

Damit ist

$$\exp(\mathbf{M}) = \sum_{n=0}^{\infty} \frac{1}{n!} \mathbf{TD}^n\mathbf{T}^{-1} = \mathbf{T} \exp(\mathbf{D}) \mathbf{T}^{-1} \quad (5.55)$$

Die Exponentialfunktion der Diagonalmatrix ist einfach berechnet: Es ist

$$\mathbf{D}^k = \text{diag}(\lambda_1^k, \dots, \lambda_n^k) \quad (5.56)$$

also ist

$$\exp(\mathbf{D}) = \text{diag}(e^{\lambda_1}, \dots, e^{\lambda_n}) \quad (5.57)$$

und somit

$$\exp \mathbf{M} = \mathbf{T} \text{diag}(e^{\lambda_1}, \dots, e^{\lambda_n}) \mathbf{T}^{-1} \quad (5.58)$$

#### 5.4.4 Lösung des Systems mit konstanten Koeffizienten

Damit können wir die Lösung des Systems Gl. (5.49) angehen. Diese ist

$$\vec{f}(t) = \exp(\mathbf{A}t) \vec{f}(0). \quad (5.59)$$

Wir können leicht überprüfen dass der Anfangswert stimmt:  $\exp(\mathbf{A}0) = \mathbf{1}$ . Für die Differenzialgleichung realisieren wir, dass

$$(\mathbf{A}t)^n = t^n \mathbf{A}^n \quad (5.60)$$

<sup>1</sup>genauer entnehmen Sie dem Paper „21 dubious ways to exponentiate a Matrix“ von Clive Moler



ist und somit

$$\begin{aligned} \frac{d}{dt} \exp(\mathbf{A}t) &= \frac{d}{dt} \sum_{k=0}^{\infty} \frac{t^k}{k!} \mathbf{A}^k = \sum_{k=1}^{\infty} \frac{kt^{k-1}}{k!} \mathbf{A}^k \\ &= \sum_{s=0}^{\infty} \frac{t^s}{s!} \mathbf{A} \mathbf{A}^s = \mathbf{A} \exp(\mathbf{A}t) \end{aligned}$$

wobei wir in der Summation  $s = k - 1$  gesetzt haben. Damit ist

$$\frac{d\vec{f}}{dt} = \mathbf{A} \exp(\mathbf{A}t) \vec{f}(0) = \mathbf{A} \vec{f}(t) \quad (5.61)$$

wie gewünscht.

### 5.4.5 Beispiel: Der gedämpfte harmonische Oszillator

Ein gedämpfter freier harmonischer Oszillator wird beschrieben durch die Differenzialgleichung

$$\ddot{x}(t) + 2\gamma\dot{x}(t) + \omega_0^2 x(t) = 0 \quad (5.62)$$

Hier haben wir bereits die Kraftkonstante der Rückstellkraft  $F = -Kx$  umgeschrieben mit  $\omega_0^2 = K/m$  und ähnlich die Reibungskonstante angepasst. Alle Parameter dieser Gleichung haben die Dimension  $(\text{Zeit})^{-1}$ . Wir erkennen dies als System mit konstanten Koeffizienten der Form

$$\frac{d}{dt} \begin{pmatrix} x \\ \dot{x} \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ -\omega_0^2 & -2\gamma \end{pmatrix} \begin{pmatrix} x \\ \dot{x} \end{pmatrix} \equiv \mathbf{M} \begin{pmatrix} x \\ \dot{x} \end{pmatrix}. \quad (5.63)$$

Wir berechnen das charakteristische Polynom

$$Z(\lambda) = \det(\mathbf{M} - \lambda\mathbf{1}) = \det \begin{pmatrix} -\lambda & 1 \\ -\omega_0^2 & -2\gamma - \lambda \end{pmatrix} = \lambda(\lambda + 2\gamma) + \omega_0^2. \quad (5.64)$$

Seine Nullstellen erhalten wir mit der Mitternachtsformel

$$\lambda_{1/2} = -\gamma \pm \sqrt{\gamma^2 - \omega_0^2} \quad (5.65)$$

An dieser Stelle werden wir nachher eine wichtige physikalische Fallunterscheidung durchführen. Im Augenblick dient uns dieses Beispiel aber der Illustration unseres Formalismus und wir ziehen diesen weiter durch. Wir ermitteln die Eigenvektoren als nichttriviale Lösungen von

$$\begin{pmatrix} \gamma \mp \sqrt{\gamma^2 - \omega_0^2} & 1 \\ -\omega_0^2 & -\gamma \mp \sqrt{\gamma^2 - \omega_0^2} \end{pmatrix} \vec{v}_{1/2} = \vec{0}. \quad (5.66)$$

Bevor wir starten entdimensionalisieren wir die Matrix indem wir  $Q = \omega_0/\gamma$  einführen zu

$$\begin{pmatrix} \gamma(1 \mp \sqrt{1 - Q^2}) & 1 \\ -\gamma^2 Q^2 & \gamma(-1 \mp \sqrt{1 - Q^2}) \end{pmatrix} \quad (5.67)$$

Wir tauschen die Zeilen und erzeugen eine eins in der linken oberen Ecke

$$\begin{pmatrix} 1 & \frac{1 \pm \sqrt{1-Q^2}}{\gamma Q^2} \\ \gamma(1 \mp \sqrt{1-Q^2}) & 1 \end{pmatrix} \quad (5.68)$$

Wir multiplizieren die obere Zeile mit  $\gamma(1 \mp \sqrt{1-Q^2})$  und subtrahieren das Ergebnis von der unteren Zeile. Dabei nutzen wir dass nach dem dritten Binom

$$\gamma \frac{1 \pm \sqrt{1-Q^2}}{\gamma Q^2} (1 \mp \sqrt{1-Q^2}) = \frac{1}{Q^2} (1 - (1-Q^2)) = 1(1) \quad (5.69)$$

ist und erhalten eine in der unteren Zeile ausgeäumte Matrix. Daraus bestimmen wir die Eigenvektoren

$$\vec{v}_{1/2} = \begin{pmatrix} 1 \pm \sqrt{1-Q^2} \\ \gamma Q^2 \end{pmatrix} \quad (5.70)$$

Die Basiswechselmatrix ist also

$$\mathbf{T} = \begin{pmatrix} 1 + \sqrt{1-Q^2} & 1 - \sqrt{1-Q^2} \\ \gamma Q^2 & \gamma Q^2 \end{pmatrix} \quad (5.71)$$

Die Determinante dieser Matrix ist  $D = 2\gamma Q^2 \sqrt{1-Q^2}$  und damit ist die Inverse

$$\mathbf{T}^{-1} = \frac{1}{2\gamma Q^2 \sqrt{1-Q^2}} \begin{pmatrix} \gamma Q^2 & -1 + \sqrt{1-Q^2} \\ -\gamma Q^2 & 1 + \sqrt{1-Q^2} \end{pmatrix} \quad (5.72)$$

Die Zeitentwicklung wird damit beschrieben durch

$$\begin{aligned} \exp(\mathbf{M}t) &= \frac{e^{-\gamma t}}{2\gamma Q^2 \sqrt{1-Q^2}} \begin{pmatrix} 1 + \sqrt{1-Q^2} & 1 - \sqrt{1-Q^2} \\ \gamma Q^2 & \gamma Q^2 \end{pmatrix} \\ &\begin{pmatrix} \gamma Q^2 e^{\gamma \sqrt{1-Q^2} t} & (-1 + \sqrt{1-Q^2}) e^{\gamma \sqrt{1-Q^2} t} \\ -\gamma Q^2 e^{-\gamma \sqrt{1-Q^2} t} & (1 + \sqrt{1-Q^2}) e^{-\gamma \sqrt{1-Q^2} t} \end{pmatrix} \\ &= \frac{e^{-\gamma t}}{\sqrt{1-Q^2}} \begin{pmatrix} \sqrt{1-Q^2} \cosh \bar{\gamma} t + \sinh \bar{\gamma} t & -\frac{1}{\gamma} \sinh \bar{\gamma} t \\ \gamma Q^2 \sinh \bar{\gamma} t & \sqrt{1-Q^2} \cosh \bar{\gamma} t - \sinh \bar{\gamma} t \end{pmatrix} \end{aligned}$$

mit  $\bar{\gamma} = \gamma \sqrt{1-Q^2}$ . Dies ist exakt und reichlich unübersichtlich.

Wir können etwas Licht ins Dunkel bringen, indem wir spezielle Anfangsbedingungen ansetzen. Wenn der Harmonische Oszillator ausgelenkt aus der Ruhe startet, also  $v(0) = 0$  ist, dann haben wir

$$x(t) = x(0) \left( \cosh \bar{\gamma} t + \frac{1}{\sqrt{1-Q^2}} \sinh \bar{\gamma} t \right) e^{-\gamma t} \quad (5.73)$$

und analog beim Start ohne Auslenkung,  $x(0) = 0$  haben wir

$$x(t) = -\frac{v(0)}{\gamma \sqrt{1-Q^2}} e^{-\gamma t} \sinh \bar{\gamma} t. \quad (5.74)$$

### 5.4.6 Physikalische Grenzfälle

Wesentlich einfacher wird die Interpretation der physikalischen Grenzfälle. Wir starten von den Eigenwerten

$$\lambda_{1/2} = -\gamma \pm \sqrt{\gamma^2 - \omega_0^2} \quad (5.75)$$

#### 5.4.6.1 Ungedämpft

Im ungedämpften Fall haben wir  $\gamma = 0$ . Damit wird  $\lambda_{1/2} = \pm i\omega_0$ . Wir lösen diesen Fall zur Übung noch einmal separat. Die Eigenvektoren berechnen sich aus

$$\begin{pmatrix} \mp i\omega_0 & 1 \\ -\omega_0^2 & \mp i\omega_0 \end{pmatrix} \vec{v}_{1/2} = 0. \quad (5.76)$$

zu

$$\vec{v}_{1/2} = \begin{pmatrix} 1 \\ \pm i\omega_0 \end{pmatrix} \quad (5.77)$$

Die Basiswechselmatrizen sind

$$\mathbf{T} = \begin{pmatrix} 1 & 1 \\ i\omega_0 & -i\omega_0 \end{pmatrix} \quad \mathbf{T}^{-1} = \frac{i}{2\omega_0} \begin{pmatrix} -i\omega_0 & -1 \\ -i\omega_0 & 1 \end{pmatrix} \quad (5.78)$$

und somit als allgemeine Lösung

$$\begin{aligned} \begin{pmatrix} x(t) \\ \dot{x}(t) \end{pmatrix} &= \frac{i}{2\omega_0} \mathbf{T} \begin{pmatrix} e^{i\omega_0 t} & 0 \\ 0 & e^{-i\omega_0 t} \end{pmatrix} \begin{pmatrix} -i\omega_0 & -1 \\ -i\omega_0 & 1 \end{pmatrix} \begin{pmatrix} x(0) \\ v(0) \end{pmatrix} \\ &= \frac{i}{2\omega_0} \mathbf{T} \begin{pmatrix} e^{i\omega_0 t} & 0 \\ 0 & e^{-i\omega_0 t} \end{pmatrix} \begin{pmatrix} -i\omega_0 x(0) - v(0) \\ -i\omega_0 x(0) + v(0) \end{pmatrix} \\ &= \frac{i}{2\omega_0} \begin{pmatrix} 1 & 1 \\ i\omega_0 & -i\omega_0 \end{pmatrix} \begin{pmatrix} e^{i\omega_0 t} (-i\omega_0 x(0) - v(0)) \\ e^{-i\omega_0 t} (-i\omega_0 x(0) + v(0)) \end{pmatrix} \\ &= \begin{pmatrix} x(0) \cos \omega_0 t + \frac{v(0)}{\omega_0} \sin \omega_0 t \\ v(0) \cos \omega_0 t - \omega_0 x(0) \sin \omega_0 t \end{pmatrix}. \end{aligned}$$

Der harmonische Oszillator schwingt also mit der Frequenz  $\omega_0$ . Wir können die Koordinate umschreiben mit dem komplexen Ansatz

$$x(t) = x(0) \cos \omega_0 t + \frac{v(0)}{\omega_0} \sin \omega_0 t = \operatorname{Re} \left[ \left( x(0) + \frac{v(0)}{i\omega_0} \right) e^{i\omega_0 t} \right] = A \cos(\omega_0 t + \phi) \quad (5.79)$$

mit  $A = \sqrt{x_0^2 + \left(\frac{v(0)}{\omega_0}\right)^2}$  und  $\phi = -\arctan\left(\frac{v(0)}{\omega_0 x(0)}\right)$ .  $A$  ist die Amplitude der Schwingung und  $\phi$  beschreibt die Phasenlage.

### 5.4.6.2 Unterdämpft

Jetzt betrachten wir den Fall  $\gamma < \omega_0$ . Auch hier sind die Eigenwerte komplex und können geschrieben werden als

$$\lambda_{1/2} = -\gamma \pm i\bar{\omega} \quad \bar{\omega} = \sqrt{\omega_0^2 - \gamma^2} \quad (5.80)$$

Diesen Fall diskutieren wir durch Umschreiben der allgemeinen Lösung. Dazu nutzen wir  $\sqrt{1 - Q^2} = \sqrt{1 - \left(\frac{\omega_0}{\gamma}\right)^2} = \frac{i\bar{\omega}}{\gamma}$  und  $\bar{\gamma} = i\bar{\omega}$ . Wir nutzen auch

$$\cosh ix = \frac{1}{2}(e^{ix} + e^{-ix}) = \cos x \quad \sinh ix = \frac{1}{2}(e^{ix} - e^{-ix}) = i \sin x \quad (5.81)$$

Wir erhalten aus der allgemeinen Lösung

$$\begin{aligned} x(t) &= \frac{e^{-\gamma t}}{\sqrt{1 - Q^2}} \left[ x(0) \left( \sqrt{1 - Q^2} \cosh \bar{\gamma} t + \sinh \bar{\gamma} t \right) + \frac{v(0)}{\gamma} \sinh \bar{\gamma} t \right] \\ &= e^{-\gamma t} \left[ x(0) \left( \cos \bar{\omega} t + \frac{\gamma}{\bar{\omega}} \sin \bar{\omega} t \right) + \frac{v(0)}{\bar{\omega}} \sin \bar{\omega} t \right] \end{aligned}$$

Es handelt sich also um eine zerfallende Schwingung mit verlangsamter Frequenz  $\bar{\omega}$  die leicht phasenverschoben ist durch die Dämpfung. Instruktiv ist der Grenzfall schwacher Dämpfung,  $\gamma \ll \omega_0$ . Dann ist  $\bar{\omega} = \omega_0 + O(Q^{-2})$  und

$$x(t) = e^{-\gamma t} \left( x(0) \cos \bar{\omega} t + \frac{v(0)}{\omega_0} \sin \bar{\omega} t \right) \quad (5.82)$$

also eine freie Schwingung mit langsam zerfallender Amplitude.

### 5.4.6.3 Überdämpft

Im Fall  $\omega_0 > \gamma$  sind beide Eigenwerte rein reell. Beide sind negativ und geordnet nach  $\lambda_2 < \lambda_1 < 0$ . Hier ist unsere ursprüngliche Formel schon das Beste, was wir hinschreiben können. Da  $\gamma > \bar{\gamma}$  ist, ist die Zeitentwicklung eine Überlagerung von einem schnellerem und einem langsamerem exponentiellen Zerfall.

### 5.4.6.4 Kritisch gedämpft

Interessant wird die Situation bei  $\gamma = \omega$ , dem Übergang zwischen unter- und überdämpft. Hier fallen beide Eigenwerte zusammen,  $\lambda_1 = \lambda_2$  und es ist  $Q = 1$ . Die Matrix  $\mathbf{M}$  ist dann nicht diagonalisierbar. Interessanterweise existiert aber der entsprechende Grenzwert der allgemeinen Lösung

$$\begin{aligned} x(t) &= \lim_{Q \rightarrow 1} \frac{e^{-\gamma t}}{\sqrt{1 - Q^2}} \left[ x(0) \left( \sqrt{1 - Q^2} \cosh \bar{\gamma} t + \sinh \bar{\gamma} t \right) + \frac{v(0)}{\gamma} \sinh \bar{\gamma} t \right] \\ &= e^{-\gamma t} [x(0) (1 + \gamma t) + v(0)t] \end{aligned}$$

Hier sieht man (Zeichung in der Vorlesung) noch den Ansatz einer deformierten Oszillation. Man nennt dieses Ergebnis den *aperiodischen Grenzfall*.

### 5.4.7 Der angetriebene harmonische Oszillator

Wir betrachten einen harmonischen Oszillator mit einem zeitlich harmonischen Antrieb

$$\ddot{x}(t) + 2\gamma\dot{x}(t) + \omega_0^2 x(t) = f_{\omega,0} \cos(\omega t + \phi) \quad (5.83)$$

Da dies eine lineare Differenzialgleichung ist gilt wie üblich bei linearen Objekten dass wenn  $x_\omega(t)$  eine Lösung dieser inhomogenen DGI ist und  $x_0(t)$  eine Lösung der homogenen DGI Gl. (5.62), dann ist  $x_\omega + \lambda x_0$  wieder eine Lösung der inhomogenen DGI. Wir müssen also lediglich eine spezielle Lösung des getriebenen HO finden ohne auf die Anfangsbedingungen Rücksicht zu nehmen und können dann mittels freier Lösungen auf die Anfangsbedingungen anpassen.

#### 5.4.7.1 Spezielle Lösung, komplexer Ansatz

Zum finden der speziellen Lösung nutzen wir einen *komplexen Ansatz*. Wir schreiben die rechte Seite um als

$$f_{\omega,0} \cos(\omega t + \phi) = \operatorname{Re} [f_\omega e^{i\omega t}] \quad f_\omega = f_{\omega,0} e^{i\phi} \quad (5.84)$$

Auf der linken Seite machen wir den Ansatz

$$x(t) = \operatorname{Re} [x_\omega e^{i\omega t}] \quad (5.85)$$

Für den Ansatz der Lösung als komplexe Exponentialfunktion heiligt der Zweck die Mittel - wir sehen gleich, dass dies eine spezielle Lösung produziert. Der komplexe Ansatz gilt, solange die durchzuführende Rechenoperation linear ist und damit mit der Realteilbildung kompatibel: Für die Berechnung nichtlinearer Größen wie z.B. der kinetischen Energie  $m\dot{x}^2/2$  muss *erst* der Realteil genommen und dann eingesetzt werden.

Dieser Ansatz macht aus unserer DGI

$$f_\omega e^{i\omega t} = [-\omega^2 + 2i\gamma\omega + \omega_0^2] x_\omega e^{i\omega t} \quad (5.86)$$

wobei wir nach den Differenziationen bereits ausgeklammert haben. Wir dividieren die Exponentialfunktion ab und lösen auf

$$x_\omega = \frac{f_\omega}{\omega_0^2 - \omega^2 + 2i\gamma\omega} = \frac{e^{i\alpha_\omega}}{\sqrt{(\omega_0^2 - \omega^2)^2 + 4\gamma^2\omega^2}} f_\omega \quad \tan \alpha_\omega = \frac{2\gamma\omega}{\omega^2 - \omega_0^2}. \quad (5.87)$$

Wir sehen also, dass eine nicht-abklingende Oszillation mit der Antriebsfrequenz eine spezielle Lösung der DGI bildet.

#### 5.4.7.2 Gesamtlösung

Wenn wir jetzt noch eine freie Lösung dazuaddieren, um auf die Anfangsbedingung einzugehen, so ist dies wie oben gesehen bestenfalls eine Oszillation mit einer Frequenz von  $\omega_0$  oder kleiner, die mit einer Rate von ungefähr  $\gamma$  abklingt. Nach einer Zeit  $t \gg \gamma^{-1}$  tritt also nur noch die spezielle Lösung auf und das System hat seine Anfangsbedingung vergessen. Wir konzentrieren uns damit ab jetzt auf die Diskussion der speziellen Lösung.

### 5.4.7.3 Resonanz

Die Lösung Gl. (5.87) zeigt das Phänomen der *Resonanz*. Dazu betrachten wir die Amplitude  $A_\omega = \left( (\omega_0^2 - \omega^2)^2 + 4\gamma^2\omega^2 \right)^{-1/2}$  und die Phasenverschiebung  $\alpha_\omega$  zwischen Antrieb und Bewegung als Funktion der Antriebsfrequenz  $\omega$  relativ zur *Eigenfrequenz*  $\omega_0$ . Die entsprechenden Kurven skizzieren wir in der Vorlesung. Wir konzentrieren uns auf den unterdämpften Fall

Die Amplitude hat bei niedrigen Frequenzen,  $\omega \ll \omega_0$  die Form

$$\begin{aligned} A_\omega &= \frac{f_\omega}{\omega_0^2} \left( 1 - 2\omega^2/\omega_0^2 + \omega^4/\omega_0^4 + 4\gamma^2\omega^2/\omega_0^4 \right)^{-1/2} \\ &= \frac{f_\omega}{\omega_0^2} \left( 1 + \frac{\omega^2}{\omega_0^2} \right) + O(\omega^4) \end{aligned}$$

sie beginnt also bei einem konstanten Wert und wächst dann quadratisch an, zunächst unabhängig von der Dämpfung. Die Phasenverschiebung in diesem Regime ist  $\alpha_\omega \simeq 0$ . Dies kann man so verstehen, dass der intrinsisch schnelle Oszillator von einer langsamen Kraft mitgeschleppt wird und dieser instantan folgt - dies nennt man auch *adiabatische* Zeitentwicklung (der Begriff *adiabatisch* erhält dann in der Thermodynamik und statistischen Physik eine erweiterte Bedeutung).

Bei hohen Frequenzen tritt der umgekehrte Fall auf. Wir entwickeln dort

$$\begin{aligned} A_\omega &= \frac{f_\omega}{\omega^2} \left( 1 - 2\omega_0^2/\omega^2 + \omega_0^4/\omega^4 + 4\gamma^2/\omega^2 \right)^{-1/2} \\ &= \frac{f_\omega}{\omega^2} \left( 1 + \frac{\omega_0^2}{\omega^2} \right) + O(\omega_0^4) \end{aligned}$$

Die Amplitude klingt also mit immer größer werdenden Frequenzen wie  $1/\omega^2$  ab und nähert sich dieser Asymptote von oben. Die Phasenverschiebung ist  $\alpha_\omega \simeq \pi$ . Hier setzt der Oszillator der jetzt verglichen mit dem Antrieb langsam ist seine Trägheit dem Antrieb entgegen.

Dazwischen geht, bei  $\omega \simeq \omega_0$  die Amplitude durch ein Maximum von

$$A_{\omega=\omega_0} = \frac{f_\omega}{2\gamma\omega_0} \quad (5.88)$$

mit  $\alpha_\omega = \pi/2$ . Hier tritt *Resonanz* auf: Die Antriebsfrequenz ist auf die Eigenfrequenz abgestimmt und der Energieübertrag vom Antrieb auf den Oszillator ist optimal. Das Aufschaukeln des Oszillators ist nur dadurch limitiert, dass Reibung auftritt und je stärker das System oszilliert desto mehr Energie geht in den Reibungsmechanismus - bei langen Zeiten stellt sich ein *dynamisches Gleichgewicht* ein, bei dem genau so viel Energie in die Dämpfung übergeht wie der Antrieb dem Oszillator zuführen kann.

Im Resonanzfall lohnt es sich, auch das entsprechende Anfangswertproblem zu lösen. Speziell betrachten wir  $x(0), \dot{x}(0) = 0$  und  $\phi = 0$ . Die spezielle Lösung ist dann

$$x_\omega = \frac{f_\omega}{2\gamma\omega_0} \sin \omega_0 t. \quad (5.89)$$

Diese Lösung startet schon mal aus der Ruhe, allerdings ist  $\dot{x}_\omega(0) = \frac{f_\omega}{2\gamma}$ . Wir nutzen unser Ergebnis des freien Oszillators, Gl. (5.82) und erhalten

$$x(t) = \frac{f_\omega}{2\gamma\omega_0} (1 - e^{-\gamma t}) \sin \omega_0 t \quad (5.90)$$

Es entsteht also das Bild dass wir etwa eine Zeit von  $\gamma^{-1}$  brauchen, um das Resonanzverhalten zu etablieren.

## 5.5 Fortgeschrittene Techniken zur Lösung von linearen Differenzialgleichungen

### 5.5.1 Überlagerung von Antrieben beim harmonischen Oszillator

Im letzten Kapitel haben wir spezielle Lösungen des getriebenen harmonischen Oszillators

$$\ddot{x}(t) + 2\gamma\dot{x}(t) + \omega_0^2 x(t) = f(t) \quad (5.91)$$

angeschaut in dem Fall, dass  $f(t) = f_{\omega,0} \cos(\omega_0 t + \phi)$  und die Lösung  $x_\omega(t)$  erhalten. Tatsächlich können wir daraus bereits sehr viel mehr lernen. Die linke Seite schreiben wir schematisch als  $Dx$  mit

$$D = \frac{d^2}{dt^2} + 2\gamma \frac{d}{dt} + \omega_0^2 \quad (5.92)$$

wir nennen  $D$  einen *Differenzialoperator* (Operator kann hier mit Abbildung gleichgesetzt werden). Wie schon oben ausgeführt, ist dieser linear, also

$$D(\lambda x(t) + \mu y(t)) = \lambda D x(t) + \mu D y(t). \quad (5.93)$$

Darum ist für den Fall, dass die Inhomogenität eine Summe über Kosinusfunktionen ist

$$f(t) = \sum_{\omega_i} f_{\omega_i} \cos(\omega_i t + \phi_i) \quad (5.94)$$

die Lösung

$$x(t) = \sum_{\omega_i} x_{\omega_i}(t) = \operatorname{Re} \left[ \sum_{\omega_i} \frac{f_{\omega_i}}{\omega_0^2 - \omega_i^2 + 2i\gamma\omega_i} e^{i(\omega_i t + \phi_i)} \right] \quad (5.95)$$

Wir sehen also, dass für die verschiedenen Frequenzen des Antriebs der Oszillator wie ein Filter wirkt - er unterdrückt Frequenzen die mehrere  $\gamma$  von  $\omega_0$  verstimmt sind.

Wir scheinen uns das Leben hier besonders einfach gemacht zu haben - der Vektorraum der von den harmonischen Antrieben aufgespannt wird kann gelöst werden. Wie groß ist der? Welcher Teil des Raums aller möglichen Antriebsfunktionen kann das sein? Die Beantwortung dieser Frage benötigt etwas Anlauf, denn wir können hier nicht einfach Dimensionen abzählen - die Summendarstellung in diesem Kapitel können unendlich sein, aber es gibt auch sehr viele Funktionen.

### 5.5.2 Das Dirac-Delta

Das Dirac-Delta ist ein Grenzwert einer Funktionenfolge  $\delta(x) = \lim_{\epsilon \rightarrow 0^+} \delta_\epsilon(x)$  mit den Eigenschaften

$$\lim_{\epsilon \rightarrow 0^+} \delta(x \neq 0) = 0 \quad \lim_{\epsilon \rightarrow 0^+} \int_{-\infty}^{\infty} dx \delta_\epsilon(x) = 1 \quad (5.96)$$

Wir suchen also eine Kurve, die im Grenzfall überall außerhalb von  $x = 0$  verschwindet, aber in  $x = 0$  so unendlich unendlich ist, dass das Integral darüber eins ergibt. Dieses Objekt ist keine Funktion im engeren Sinn und Interpretationen seiner Bedeutung machen dann Sinn, wenn über das Dirac-Delta integriert wird. Die für uns wichtigste Eigenschaft des Dirac-Delta ist seine Anwendung zum Ausschneiden von Funktionswerten. Es ist

$$\int_{-\infty}^{\infty} dx f(x) \delta(x) = \int_{-\infty}^{\infty} dx f(0) \delta(x) = f(0). \quad (5.97)$$

Hier haben wir Grenzübergänge vertauscht. Damit das funktioniert, muss  $f(x)$  mindestens *quadratintegrabel sein*, es muss also folgendes Integral existieren

$$\int_{-\infty}^{\infty} dx |f(x)|^2 < \infty. \quad (5.98)$$

Die Menge dieser Funktionen heißt  $L_2$ . Damit trägt der Integrationsbereich außerhalb der Null nichts zum Integral bei, d.h. es ist für  $a < 0 < b$  auch

$$\int_a^b dx f(x) \delta(x) = f(0). \quad (5.99)$$

Folgende weitere Eigenschaften sind interessant und ergeben sich aus Integrationsregeln

$$\delta(f(x)) = \frac{1}{|f'(f^{-1})|} \delta(f).$$

$$\delta(cx) = \frac{1}{|c|} \delta(x)$$

$$\int_{-\infty}^{\infty} dx f(x) \delta(x - x') = f(x')$$

$$\int_{-\infty}^x dx' \delta(x') = H(x) = \begin{cases} 1 & x > 0 \\ \frac{1}{2} & x = 0 \\ 0 & x < 0 \end{cases}$$

### 5.5.3 Annäherung durch Funktionenfolgen

Wir betrachten Beispiele der Annäherung dieser Funktion durch Folgen. Nahelegend (aber unpraktisch) sind Rechtecke

$$\delta_\epsilon(x) = \frac{1}{2\epsilon} \begin{cases} 1 & |x| < \epsilon \\ 0 & \text{sonst} \end{cases} \quad (5.100)$$



Besser sind Gaußfunktionen. Dazu überlegen wir zuerst mal

$$J^2 = \left( \int_{-\infty}^{\infty} dx e^{-x^2} \right)^2 = \int_{-\infty}^{\infty} dx \int_{-\infty}^{\infty} dy e^{-(x^2+y^2)} \quad (5.101)$$

Wir substituieren in Polarkoordinaten und finden

$$J^2 = \int_0^{\infty} \rho d\rho \int_0^{2\pi} d\phi e^{-\rho^2} = 2\pi \int_0^{\infty} d\rho \rho e^{-\rho^2} = \pi \quad (5.102)$$

also  $J = \sqrt{\pi}$ . Wenn wir jetzt noch eine Breite erlauben in der Form

$$\int_{-\infty}^{\infty} dx e^{-a^2 x^2} = \int_{-\infty}^{\infty} \frac{d\xi}{|a|} e^{-\xi^2} = \frac{\sqrt{\pi}}{a} \quad (5.103)$$

Damit ist folgendes eine geeignete Näherung für das Diracdelta

$$\delta_{\epsilon}(x) = \frac{1}{\sqrt{\pi\epsilon}} e^{-x^2/\epsilon} \quad (5.104)$$

Dies eignet sich für numerische Rechnungen sehr gut - die Funktion ist überall differenzierbar und konvergiert für  $x \gg \sqrt{\epsilon}$  sehr schnell.

Ein drittes (und im nächsten Kapitel entscheidendes) Beispiel ist die *Lorentzkurve*

$$\delta_{\epsilon}(x) = \frac{1}{\pi} \frac{\epsilon}{\epsilon^2 + x^2}. \quad (5.105)$$

Die Integraleigenschaft rechnen wir nach

$$\begin{aligned} J_{\epsilon} &= \int_{-\infty}^{\infty} \frac{dx}{\pi} \frac{\epsilon}{\epsilon^2 + x^2} = \frac{1}{\pi} \int_{-\infty}^{\infty} \frac{dx}{\epsilon} \frac{1}{1 + (x/\epsilon)^2} \\ &= \frac{1}{\pi} \int_{-\infty}^{\infty} \frac{ds}{1 + s^2} = \frac{1}{\pi} \arctan s \Big|_{-\infty}^{\infty} = 1 \end{aligned}$$

#### 5.5.4 Fourierzerlegung

Die Fourierzerlegung ist die Zerlegung einer Funktion (in der Zeit) nach Frequenzen. Wir kennen es z.B., dass das Radiospektrum nach Frequenzen zerlegt wird oder auch Töne nach Frequenzen, hoch und tief, dargestellt werden. Wir wollen jetzt die mathematische Grundlage dazu legen. Dazu betrachten wir die Quadratintegrale Funktion

$$f_k(x) = \frac{1}{2\pi} e^{ikx} e^{-\epsilon|x|} \quad \epsilon > 0. \quad (5.106)$$

Wir wollen jetzt zeigen, dass diese Funktionen im Limes  $\epsilon \rightarrow 0^+$  eine Basis des  $L_2$  bilden. Die lineare Unabhängigkeit ist offensichtlich. Zur Basisdarstellung betrachten wir zunächst

$$D_{\epsilon}(x) = \int_{-\infty}^{\infty} dk f_k(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} dk e^{ikx} e^{-\epsilon|x|}. \quad (5.107)$$

Der Realteil des Integranden ist gerade, der Imaginärteil ist ungerade. Damit ist das Integral reell und wir können schreiben

$$D_\epsilon(x) = \frac{1}{\pi} \operatorname{Re} \left[ \int_0^\infty dk e^{i(k+i\epsilon)x} \right] = \frac{1}{\pi} \operatorname{Re} \left[ \frac{e^{i(k+i\epsilon)x}}{i(k+i\epsilon)} \right]_0^\infty = \frac{1}{\pi} \operatorname{Re} \left[ \frac{-1}{i(k+i\epsilon)} \right] = \frac{1}{\pi} \frac{\epsilon}{k^2 + \epsilon^2} \quad (5.108)$$

Also haben wir gerade eine Lorentzkurve und die wird zum Diracdelta. Damit können wir für eine beliebige quadratintegrale Form schreiben

$$\begin{aligned} f(x) &= \int_{-\infty}^\infty dx' f(x') \delta(x-x') \\ &= \lim_{\epsilon \rightarrow 0^+} \int_{-\infty}^\infty \frac{dx' dk}{2\pi} f(x') e^{ik(x-x')} e^{-\epsilon|x-x'|} \\ &= \int_{-\infty}^\infty dk \frac{e^{ikx}}{2\pi} \tilde{f}(k) \\ \tilde{f}(k) &= \int_{-\infty}^\infty dx' f(x') e^{-ikx'}. \end{aligned}$$

Wir nennen  $\tilde{f}(k)$  die Fouriertransformierte von  $f(x)$ , manchmal auch geschrieben als  $\tilde{f} = \mathcal{F}(f)$  und  $f = \mathcal{F}^{-1}(\tilde{f})$ . Wir sehen, dass die Transformation zwischen  $f$  und  $\tilde{f}$  sehr symmetrisch ist (identisch bis auf ein Vorzeichen und einen Normierungsfaktor). Sie werden in Büchern verschiedene Konventionen sehen, wo das  $2\pi$  zu stehen hat. Wir haben im Ergebnis gezeigt, dass die  $f_k(x)$  eine Basis des  $L_2$  bilden d.h. das wir jede Funktion als (kontinuierliche) Linearkombination dieser Funktionen darstellen können.

Folgende Eigenschaften der Fouriertransformation sind wichtig:

1. Linearität: Aus der Linearität der Integration folgt auch  $\mathcal{F}(\lambda f(x) + \mu g(x)) = \lambda \mathcal{F}(f) + \mu \mathcal{F}(g)$
2. Differentiationsatz: Wir betrachten die Ableitung in Fourierdarstellung

$$\frac{df(x)}{dx} = \frac{d}{dx} \int \frac{dk}{2\pi} e^{ikx} \tilde{f}(k) = \int \frac{dk}{2\pi} ik e^{ikx} \tilde{f}(k) \quad (5.109)$$

Daraus folgt  $\mathcal{F}\left(\frac{df}{dx}\right) = ik\mathcal{F}(f)$ . Aus der Ableitung wird also eine algebraische Operation. Wir werden später sehen, wie sich damit Differenzialgleichungen kompakt lösen lassen.

3. Faltungssatz: Wir berechnen die Fouriertransformierte eines Produkts

$$\begin{aligned} \mathcal{F}(fg) &= \int dx f(x)g(x) e^{-ikx} = \int \frac{dx dk' dk''}{(2\pi)^2} \tilde{f}(k') e^{ik'x} \tilde{g}(k'') e^{ik''x} e^{-ikx} \\ &= \int \frac{dk' dk''}{2\pi} \tilde{f}(k') \tilde{g}(k'') \delta(k' + k'' - k) \\ &= \int \frac{dk'}{2\pi} \tilde{f}(k') \tilde{g}(k - k') \end{aligned}$$

Das Integral auf der rechten Seite heißt *Faltung von f an g*,  $f \circ g$ . Der Faltungssatz lautet dann  $\mathcal{F}(fg) = \mathcal{F}(f) \circ \mathcal{F}(g)$  und aus Symmetriegründen auch  $\mathcal{F}^{-1}(\mathcal{F}(f)\mathcal{F}(g)) = f \circ g$ .

4. Satz von Parseval über die Normierung

$$\begin{aligned} \int_{-\infty}^{\infty} dx |f(x)|^2 &= \frac{1}{4\pi^2} \int dx dk dk' f(k) e^{ikx} \bar{f}(k')^{-ik'x} = \\ &= \frac{1}{2\pi} \int dk dk' f(k) \bar{f}(k') \delta(k - k') = \\ &= \frac{1}{2\pi} \int dk |f(k)|^2 \end{aligned}$$

5. Reelle Funktion: Wenn  $f(x) = \bar{f}(x)$  ist, dann ist

$$\int dk \tilde{f}(k) e^{ikx} = \int dk \bar{\tilde{f}}(k) e^{-ikx} = \int dk \bar{\tilde{f}}(-k) e^{ikx} \quad (5.110)$$

also  $f(k) = \bar{f}(-k)$ . Für eine rein imaginäre Funktion ist  $f(k) = -\bar{f}(-k)$ .

6. Gerade Funktion: Wenn  $f(x) = f(-x)$  ist

$$\int dk f(x) e^{ikx} = \int dk f(x) \cos kx \quad (5.111)$$

und damit ist  $\tilde{f}$  auch gerade (analoges gilt für ungerade Funktionen)

### 5.5.5 Anwendung auf den getriebenen Oszillator

Wir können jetzt für einen beliebigen Antrieb  $f(t)$  mit Fouriertransformierter  $\tilde{f}(\omega)$  einfach schreiben

$$x(t) = \text{Re} \left[ \int_{-\infty}^{\infty} \frac{d\omega}{2\pi} \frac{\tilde{f}(\omega)}{\omega_0^2 - \omega_i^2 + 2i\gamma\omega_i} e^{i\omega t} \right] \quad (5.112)$$

Dieses Integral ist im Allgemeinen recht knifflig, wir lernen aber ganz am Ende einen guten Trick, es auszurechnen.

## 5.6 Greensche Funktionen

Eine andere Methode, lineare Differenzialgleichungen mit beliebiger Inhomogenität zu lösen ist die Methode der Greenschen Funktionen. Die beruht auf dem gleichen Prinzip wie oben: Wir finden eine Basis für den Raum aller möglichen Antriebe, lösen die DGl für diese Basisfunktion, und schreiben den Antrieb dann als Linearkombination.

### 5.6.1 Definition und Lösungsformel

Wir betrachten eine allgemeine lineare Differenzialgleichung

$$\mathbf{D}f(x) = g(x) \quad (5.113)$$

wobei  $\mathbf{D}$  ein linearer Differenzialoperator ist, der nach  $x$  differenziert. Er kann selber von  $x$  abhängen und es kann sich um ein System bzw. eine DGL höherer Ordnung handeln. Die Greensche Funktion zu  $\mathbf{D}$  ist dann eine Funktion  $G(x, x')$  ist dann eine Funktion von zwei Argumenten mit der Eigenschaft

$$\mathbf{D}G(x, x') = \delta(x - x'). \quad (5.114)$$

Weiterhin wird nur nach  $x$  differenziert - die Variable  $x'$  ist ein Index. Im Spezialfall von konstanten Koeffizienten - also keiner expliziten  $x$ - Abhängigkeit von  $G$  reicht der Ansatz  $G(x - x')$ . Wenn wir die Greensche Funktion haben, dann ist die Lösung der ursprünglichen DGL einfach

$$f(x) = \int_{-\infty}^{\infty} dx G(x, x') g(x'). \quad (5.115)$$

Diese Formel können wir einfach testen

$$\begin{aligned} \mathbf{D}f(x) &= \mathbf{D} \int_{-\infty}^{\infty} dx' G(x, x') g(x') = \int_{-\infty}^{\infty} dx' \mathbf{D}[G(x, x')] g(x') \\ &= \int_{-\infty}^{\infty} dx' \delta(x - x') g(x') = g(x). \end{aligned}$$

Hier haben wir genutzt, dass  $\mathbf{D}$  nach  $x$  differenziert, dass auf der rechten Seite nur die Rolle eines Index hat - hier wird über  $x'$  integriert. Die Definitionseigenschaft von  $G$  führt dazu, dass diese Funktion manchmal auch *Impulsantwort* genannt wird.

Die Herausforderung besteht jetzt in der Bestimmung von  $G$ .

### 5.6.2 Greensche Funktion des getriebenen harmonischen Oszillators

Wir können und wollen hier nicht auf die reichhaltigen Methoden zur Bestimmung von Greenschen Funktionen eingehen, sondern gehen gleich auf den harmonischen Oszillator ein. Die Greensche Funktion muss erfüllen

$$\left[ \frac{d^2}{dt^2} + 2\gamma \frac{d}{dt} + \omega_0^2 \right] G(t - t') = \delta(t - t'). \quad (5.116)$$

Außerdem wollen wir, dass die Ursache der Wirkung folgt und das Teilchen aus der Ruhe startet (dies ist strenggenommen nicht wichtig - es gibt viele Greensfunktionen - aber liefert natürlich eine physikalisch sinnvolle spezielle Lösung. Dies erreichen wir mit der Forderung von  $G(\tau < 0) = 0$ . Damit ist die Greensche Funktion des harmonischen Oszillators bei  $\tau > 0$  einfach eine Lösung

des freien harmonischen Oszillators mit der entsprechenden Anfangsbedingung. Zu deren Bestimmung müssen wir uns um eine Umgebung von  $\tau = 0$  kümmern. Wir integrieren für  $\epsilon > 0$

$$\int_{-\epsilon}^{\epsilon} d\tau \left[ \frac{d^2 G(\tau)}{d\tau^2} + 2\gamma \frac{dG(\tau)}{d\tau} + \omega_0^2 G(\tau) \right] = 1. \quad (5.117)$$

Im Limes  $\epsilon \rightarrow 0^+$  erhalten wir, weil  $G$  stetig sein muss,

$$1 = \lim_{\epsilon \rightarrow 0^+} [G'(\epsilon) + 2\gamma G(\epsilon)]. \quad (5.118)$$

Da der rechte Grenzwert eine endliche Zahl ist, muss die Funktion bei 0 stetig aber nur einseitig differenzierbar sein, sprich  $G(0^+) = 0$  aber  $G'(0^+) = 1$  - alles andere führt zu Unendlichkeiten. Damit ist  $G(\tau > 0)$  eine Lösung des Anfangswertproblems mit  $G(0) = 0$  und  $G'(0) = 1$ . Im unterdämpften Fall ist das

$$G(\tau) = \frac{\sin \omega_0 \tau}{\omega_0} e^{-\gamma \tau} H(\tau) \quad (5.119)$$

wobei  $H$  die schon mehrfach eingeführte Heaviside-Stufenfunktion ist.

### 5.6.3 Anwendung

Wenn wir als Antrieb selber eine Heavisidefunktion nehmen,  $f(t) = aH(t)$  dann erhalten wir

$$\begin{aligned} x(t > 0) &= \frac{a}{\omega_0} \int_0^t d\tau e^{-\gamma(t-\tau)} \sin \omega_0(t-\tau) = \\ &= \frac{a}{\omega_0} \operatorname{Im} \int_0^t d\tau e^{i(\omega_0+i\gamma)\tau} = \frac{a}{\omega_0} \operatorname{Im} \left[ \frac{e^{i(\omega_0+i\gamma)\tau}}{i(\omega_0+i\gamma)} \right]_0^t \\ &= -\frac{a}{\omega_0} \operatorname{Re} \left[ \frac{e^{(i\omega_0+i\gamma)t} - 1}{\omega_0 + i\gamma} \right] = \frac{a}{\omega_0^2} (1 - e^{-\gamma t} \cos \omega_0 t). \end{aligned}$$

Wie können wir das verstehen? Eine konstante Kraft dieser Art entspricht einem Potenzialbeitrag von  $-amx$ . Damit können wir das Potenzial quadratisch ergänzen gemäß

$$\frac{1}{2} m \omega_0^2 x^2 - amx = \frac{1}{2} m \omega_0^2 \left( x^2 - 2 \frac{a}{\omega_0^2} x \right) = \frac{m}{2} \omega_0^2 \left( x - \frac{a}{\omega_0^2} \right)^2 - \frac{ma^2}{2\omega_0^2} \quad (5.120)$$

Damit handelt es sich hier um einen harmonischen Oszillator mit Gleichgewichtslage bei  $x_0 = \frac{a}{\omega_0^2}$  und einem verschobenen Referenzpunkt. Wenn wir das Potenzial also abrupt umschalten, dann führt der Oszillator eine gedämpfte Schwingung hin zum neuen Nullpunkt aus.

## 5.7 Komplexe Differenzierbarkeit

Wir sehen hier, dass komplexe Funktionen auch für die Integration große Vorteile haben.

### 5.7.1 Analytische Funktionen

Wir betrachten komplexwertige Funktionen komplexer Variablen  $f : \mathbb{C} \rightarrow \mathbb{C}$ . Wir könnten die ähnlich zweidimensionaler Vektorfelder behalten, aber wir wollen echte komplexe Differenzierbarkeit behandeln. Schreiben wir im reellen  $z = x + iy$  und  $f = u + iv$  dann ist eine Funktion genau dann nach  $z$  differenzierbar, wenn es egal ist, wie man den Real- und Imaginärteil der Ableitung findet. Dies ist beschrieben in den Cauchy-Riemannschen Differenzialgleichungen

$$\frac{\partial u}{\partial x} = \frac{\partial v}{\partial y} \quad \frac{\partial u}{\partial y} = -\frac{\partial v}{\partial x} \quad (5.121)$$

erfüllt sind. Das bedeutet im Wesentlichen, dass die Funktion von  $z$  und nicht von  $z^*$  abhängt (wir können uns z.B. leicht davon überzeugen, dass  $f = z$  die DGLen erfüllt und  $g = z^*$  sie verletzt). Tatsächlich sind alle Funktionen, die durch eine Potenzreihe in  $z$  geschrieben werden können, im Konvergenzkreis differenzierbar (und das sogar unendlich oft). Solche Funktionen heißen analytisch oder holomorph. Bei genauerer Betrachtung von Gl. (5.121) sehen wir, dass die komplex konjugierte  $f^*$  quellenfrei ist (erste Gleichung) und wirbelfrei (zweite Gleichung).

Wir wollen jetzt auch Funktionen zulassen, die nicht überall holomorph sind, sondern Singularitäten in isolierten Punkten besitzt (d.h., die Menge der Singularitäten darf keinen Häufungspunkt besitzen). Solche Funktionen nennen wir *meromorph*. Ähnlich zur Taylorreihe können um eine Singularität in  $z_0$  die Funktion in eine Laurentreihe entwickeln

$$f(z) = \sum_{s=-m}^{\infty} r_s (z - z_0)^s \quad r_{-m} \neq 0. \quad (5.122)$$

Wir nennen dann  $z_0$  einen  $m$ -fachen Pol. Bei  $m = \infty$  reden wir von einer wesentlichen Singularität. Nicht-isolierte Singularitäten treten z.B. bei mehrdeutigen Funktionen wie der Wurzel und dem Logarithmus auf.

### 5.7.2 Integrale über geschlossene Wege

Wir betrachten zunächst eine Funktion auf einem Gebiet auf dem sie analytisch ist und einen Weg  $C$  in diesem Gebiet. Wir betrachten ein geschlossenes Wegintegral

$$\oint_C dz f(z) = \int dt \dot{z} f(z(t)) \quad (5.123)$$

wobei wir davon ausgehen, dass der Weg als  $z(t)$  parameterisiert ist. Wenn wir wieder zur reellen Schreibweise übergehen sehen wir, dass aus der Wirbelfreiheit von  $f^*$  folgt, dass dieses Integral verschwindet. Geschlossene Wegintegrale über analytische Funktionen verschwinden also. Jetzt betrachten wir meromorphe Funktionen. Speziell betrachten wir  $(z - z_0)^m$  und als Integrationsweg  $z(t) = z_0 + re^{it}$  mit  $0 \leq t < 2\pi$ . Damit ist

$$I_m = \oint_C dz (z - z_0)^m = ir \int_0^{2\pi} e^{it} dt r^m e^{imt} = ir^{m+1} \int_0^{2\pi} dt e^{i(m+1)t} \quad (5.124)$$

Für  $m \neq -1$  haben wir damit

$$I_m = \frac{r^{m+1}}{m+1} e^{i(m+1)t} \Big|_0^{2\pi} = 0. \quad (5.125)$$

Für  $m = -1$  ist hingegen

$$I_{-1} = i \int_0^{2\pi} dt = 2\pi i. \quad (5.126)$$

Da  $z_0$  nach Konstruktion kein Häufungspunkt der Menge der Singularitäten ist existieren weitere geschlossene Wege um  $z_0$ . Wir können diese in den bereits betrachteten Kreis umbauen, indem wir Wege durch ein Gebiet ohne Singularitäten addieren und subtrahieren. Damit sehen wir, dass für jeden Weg, der nur die Singularität  $z_0$  einschließt und dabei die Windungszahl  $n(z_0, C)$  hat gilt

$$\oint_C dz f(z) = 2\pi i n(z_0, C) r_{-1}. \quad (5.127)$$

Den Entwicklungskoeffizienten  $r_{-1}$  nennen wir das *Residuum* von  $f$  in  $z_0$ . Wir können jetzt weiter gehen und einen Weg wählen, der mehrere Singularitäten  $z_i$  mit Umlaufzahlen  $n(z_i, C)$  einschließt. Dann ist

$$\oint_C dz f(z) = 2\pi i \sum_i n(z_i, C) \operatorname{Res}(f(z), z_i). \quad (5.128)$$

Dies ist der Residuensatz.

### 5.7.3 Der Residuensatz

An vielen Stellen müssen wir definite bzw. uneigentliche Integrale berechnen, bei denen es sich lohnt, die Funktion ins Komplexe erweitert zu betrachten - also komplexwertige Funktionen komplexer Variablen zu betrachten,  $f: \mathbb{C} \rightarrow \mathbb{C}$ , also in die komplexe Zahlenebene geeignet analytisch fortzusetzen. Die Theorie analytischer Funktionen (bzw. meromorpher Funktionen, das sind Funktionen die überall bis auf isolierte Punkte analytisch sind) liefert dann mächtige Werkzeuge zur Integralberechnung, insbesondere den Residuensatz. Wir skizzieren hier den Gedankengang und die Methodik, die genauere Theorie sollte in der Funktionentheorie behandelt werden.

Eine komplexe Funktion  $f(z) = f(x + iy) = u(x, y) + iv(x, y)$  mit reellen  $x, y$  sowie  $u, v$  heißt komplex differenzierbar, wenn sie die Cauchy-Riemanschen Differenzialgleichungen

$$u_x = v_y \quad u_y = -v_x \quad (5.129)$$

erfüllt. Genau dann ist es nämlich egal, aus welcher Richtung in der komplexen Zahlenebene wir den Differenzialquotienten bilden - wir haben eine Ableitung die wieder eine Funktion ist, keine Matrix. Es stellt sich heraus, dass dies äquivalent ist zum Verschwinden der Rotation bei einer ebenen Kurve, die dann ja nur

eine  $z$ -Komponente hat. Ist somit  $C$  ein geschlossener Integrationsweg und  $f$  analytisch, dann ist

$$\oint_C dz f(z) = 0. \quad (5.130)$$

Jetzt betrachten wir einen Integrationsweg  $C_\epsilon = \epsilon e^{it}$   $0 \leq t < 2\pi$  und  $\epsilon > 0$  klein. Wir integrieren die neue Funktion  $g(z) = f(z)/z$

$$J = \lim_{\epsilon \rightarrow 0} \oint_{C_\epsilon} g(z) = \lim_{\epsilon \rightarrow 0} \int_0^{2\pi} dt \left( \frac{dz}{dt} \right) \frac{f(\epsilon e^{it})}{\epsilon e^{it}}. \quad (5.131)$$

Jetzt werten wir die Ableitung der Kurve aus. Da  $f$  analytisch ist können wir um  $z = 0$  Taylorentwickeln und insbesondere existiert der Grenzwert  $f(0)$  und kann entsprechend erreicht werden. Damit haben wir

$$J = \lim_{\epsilon \rightarrow 0} \int_0^{2\pi} dt i\epsilon e^{it} \frac{f(0)}{\epsilon e^{it}} = if(0) \int_0^{2\pi} dt = 2\pi if(0). \quad (5.132)$$

Dieses Ergebnis lässt sich mächtig verallgemeinern. Nehmen wir einen Integrationspfad  $C$  der den Ursprung  $N(C, 0)$  mal im Uhrzeigersinn umkreist (bei Umkreisen im Gegenuhrzeigersinn ist diese Windungszahl  $N$  negativ). Dieser lässt sich dann durch das Anstückeln eines geschlossenen Weges  $D = C - C_\epsilon^N$  der den Ursprung nicht einschließt durch  $N$  faches infinitesimales Umkreisen des Ursprungs ausdrücken. Da  $g$  in einem Gebiet das  $D$  enthält analytisch ist, haben wir  $\oint_D dz g(z) = 0$  und damit  $\oint_C \frac{f(z)}{z} = Nf(0)$ . Natürlich war es nicht zwingend, den Pol der Funktion gerade bei 0 anzusetzen - wir können das alles auch um  $z_0$  verschieben. Dies liefert uns den Integralsatz von Cauchy

$$\oint_C \frac{f(z)}{z - z_0} = N(C, z_0) f(z_0). \quad (5.133)$$

Das Integral ist also nur durch den Funktionswert bei  $z_0$  bestimmt. Wir konnten dies aus der Taylorentwicklung von  $f$  und dem Integral über  $(z - z_0)^{-1}$  schließen. Wie sieht das bei höheren Potenzen aus? Betrachten wir  $(z - z_0)^m$  und als Integrationsweg  $z(t) = z_0 + r e^{it}$  mit  $0 \leq t < 2\pi$ . Damit ist

$$I_m = \oint_C dz (z - z_0)^m = ir \int_0^{2\pi} e^{it} dt r^m e^{imt} = ir^{m+1} \int_0^{2\pi} dt e^{i(m+1)t} \quad (5.134)$$

Für  $m \neq -1$  haben wir damit

$$I_m = \frac{r^{m+1}}{m+1} e^{i(m+1)t} \Big|_0^{2\pi} = 0. \quad (5.135)$$

Für  $m = -1$  ist hingegen

$$I_{-1} = i \int_0^{2\pi} dt = 2\pi i. \quad (5.136)$$



Und wenn wir den Integranden noch mit  $f(z)$  multiplizieren bleibt das Argument bestehen.

Wir wollen jetzt auch Funktionen zulassen, die nicht überall holomorph sind, sondern Singularitäten in isolierten Punkten besitzt (d.h., die Menge der Singularitäten darf keinen Häufungspunkt besitzen). Solche Funktionen nennen wir *meromorph*. Ähnlich zur Taylorreihe können um eine Singularität in  $z_0$  die Funktion in eine Laurentreihe entwickeln

$$f(z) = \sum_{s=-m}^{\infty} r_s (z - z_0)^s \quad r_{-m} \neq 0. \quad (5.137)$$

Wir nennen dann  $z_0$  einen  $m$ -fachen Pol. Bei  $m = \infty$  reden wir von einer wesentlichen Singularität. Nicht-isolierte Singularitäten treten z.B. bei mehrdeutigen Funktionen wie der Wurzel und dem Logarithmus auf. Wir haben gerade gelernt, dass beim Umkreisen von  $z_0$  nur der Term  $r_{-1}$  zum Integral beiträgt. Diesen nennen wir das Residuum von  $f$  in  $z_0$ ,  $\text{Res}(f, z_0)$ . Da alle Pole isoliert sind, können wir wiederum kleine Integrationswege definieren, die nur einen Pol umkreisen und daraus einen großen Integrationsweg zusammenstückeln und erhalten daraus den Residuensatz:  $f(z)$  sei meromorph mit isolierten Singularitäten  $\{z_j\}$ . Dann ist

$$\oint_C dz f(z) = 2\pi i \sum_j N(C, z_j) \text{Res}(f(z), z_j). \quad (5.138)$$

Jetzt benötigen wir eine Formel zur Berechnung des Residuums. Im Fall  $m = 1$ , wenn wir also einen Pol erster Ordnung haben, ist das einfach:  $\text{Res}(f(z), z_j) = \lim_{z \rightarrow z_j} (z - z_j) f(z)$ . Insbesondere ist bei Quotienten aus analytischen Funktionen

$$\text{Res}\left(\frac{p}{q}, z_j\right) = \frac{p(z_j)}{q'(z_j)} \quad (5.139)$$

Bei Polen höherer Ordnung würde dies divergieren (d.h., wenn Sie die Ordnung des Pols falsch haben und nicht wissen, dann merken Sie das hier). Für allgemeines  $m$  verwandeln wir die Laurentreihe erst in eine Taylorreihe und isolieren dann den richtigen Term durch mehrfaches Differenzieren

$$\text{Res}(f(z), z_0) = \frac{1}{(m-1)!} \frac{d^{m-1}}{dz^{m-1}} [(z - z_0)^m f(z)]|_{z=z_0} \quad (5.140)$$

Die wichtigste Anwendung ist in der Physik die Berechnung uneigentlicher Integrale. Diese müssen wir zuerst in geschlossene Wegintegrale umwandeln, die dann mit dem Residuensatz ausgewertet werden können. Ein wichtiger Fall ist

$$I_R = \int_{-\infty}^{\infty} dx R(x) e^{ikx} \quad k > 0 \quad (5.141)$$

wobei  $R(x)$  eine rationale Funktion ist, deren Nennergrad den Zählergrad um mindestens eins übersteigt. Wir schließen den Integrationsweg mit einem Halbkreis in der oberen Hälfte der Zahlenebene. Wir müssen uns als erstes davon

überzeugen, dass das ergänzte Teilstück  $A$  keinen Beitrag liefert. Dazu setzen wir zunächst die Integrationsgrenzen auf ein großes  $S > 0$

$$I_R(S) = \int_{-\infty}^{\infty} dx R(x)e^{ikx} + \int_{A(S)} dz R(z)e^{ikz} \quad (5.142)$$

wobei  $A(S)$  die Kurve zu einem Rechteck in der oberen Halbebene schließt. Die Komponenten parallel zur imaginären Achse verschwinden, wenn sie nur weit genug draußen sind. Die Komponente parallel zur imaginären Achse wird exponentiell unterdrückt. Unter Beachtung der Reihenfolge der Grenzwerte erhält man dass

$$I_R = 2\pi i \sum_{\text{Im} z_i > 0} \text{Res}(R(z)e^{ikz}, z_i) \quad (5.143)$$

ist.