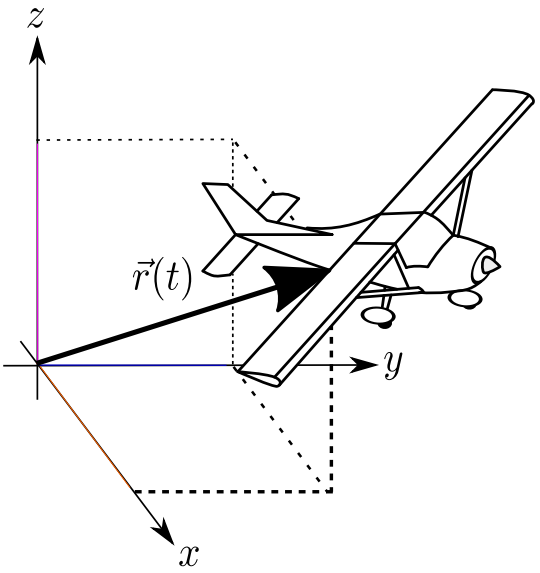
Rechenmethoden für Lehramt Physik

– Skript zur Vorlesung im WS 2019/20 –



Tobias Meng

TECHNISCHE UNIVERSITÄT DRESDEN

Dr. Tobias Meng

Leiter der Forschungsgruppe Quantum Design
Institut für Theoretische Physik
Technische Universität Dresden

Ich freue mich über Ihre Rückfragen und Kommentare zu Skript, Vorlesung, Übung und allem, was sonst noch mit dieser Vorlesung zu tun hat: entweder persönlich während der Vorlesung, in meiner Sprechstunde (Raum BZW/A118) oder per Email an tobias.meng@tu-dresden.de.

Danksagung

Dieses Skript basiert (unter anderem) auf den Skripten von Prof. Dr. S. Rachel und Dr. F. Großmann, die diese Vorlesung in früheren Semestern gehalten haben. Ich danke beiden herzlichst für die Bereitstellung Ihrer Vorlesungsmaterialien. Außerdem danke ich allen, die mich auf Fehler im Skript aufmerksam gemacht haben, insbesondere Benjamin Wolba und Marcus Kundisch.

Dresden, 3. Februar 2020,
Tobias Meng

Vorbemerkungen

Die Vorlesung findet immer Montags in der fünften Doppelstunde (14:50 - 16:20) im Hörsaal REC/C213 statt.

Leider hat die Vorlesung mit 2 **S**emester**w**ochen**s**tunden (SWS) für den gegebenen Inhalt nur sehr wenig Zeit. Um diesem bekannten Problem zu begegnen, soll die Vorlesung in den kommenden Jahren auf 4 SWS verlängert werden. Da die Vorlesung somit leider nur die wichtigsten Rechenmethoden kurz ansprechen kann, gilt:

- Sie sind grundsätzlich selbst dafür verantwortlich, eventuell bestehende Verständnisslücken im **Selbststudium** zu schließen.
- In der Vorlesung wird jede Woche ein **Übungsblatt** mit Hausaufgaben ausgeteilt. Diese dienen dazu, dass Sie die in der Vorlesung erarbeiteten Konzepte üben können und Feedback zu Ihren Rechnungen bekommen.
- Die Übungsblätter werden jede Woche in **Übungsgruppen** besprochen. Nutzen Sie diese auch sehr gerne zum Klären von Verständnisproblemen bei den Hausaufgaben und der Vorlesung allgemein.
- **In der Vorlesung** werden immer wieder **kleine Rechen- oder Verständnisaufgaben** gestellt. Wenn Sie diese bearbeiten, schlagen wir zwei Fliegen mit einer Klappe: Sie üben Rechnen, und ich habe direkt Feedback über das, was schon verstanden ist und das, was nochmal erklärt werden muss.

Generell gilt: machen Sie aktiv mit! Sprechen Sie Verständnisprobleme an, stellen Sie Fragen und kommen Sie gerne zu meiner Sprechstunde. Diese Vorlesung ist ein Angebot an Sie; mein Ziel ist es, Ihnen möglichst viele Rechenkompetenzen zu vermitteln!

Die Webseite der Vorlesung mit allen relevanten Infos, den Übungsblättern und diesem Skript finden Sie unter www.quantum-design.org/teaching.

Literaturvorschläge

Schauen Sie sich am besten verschiedene Lehrbücher an um das für Sie am besten passende Buch zu finden. Und bedenken Sie: das Buch, das am besten zu Ihnen passt, steht vielleicht gar nicht auf der folgenden Liste mit Lehrbuchvorschlägen. Schauen Sie also auch gerne weiter Lehrbücher oder Skripte an.

- M. Otto, *Rechenmethoden für Studierende der Physik im ersten Jahr* (Spektrum, 2011).
- W. Nolting, *Grundkurs Theoretische Physik 1 – Klassische Mechanik* (Springer, 2002), erste ca. 100 Seiten.
- S. Grossmann, *Mathematischer Einführungskurs für die Physik* (Vieweg+Teubner, 2004).
- C. B. Lang und N. Pucker, *Mathematische Methoden in der Physik* (Elsevier/Spektrum, 2005).
- L. Papula, *Mathematik für Ingenieure und Naturwissenschaftler Band 1, Band 2, Klausur- und Übungsaufgaben* (Vieweg+Teubner, 2009/10/12).
- H. Schulz, *Physik mit Bleistift: Einführung in die Rechenmethoden der Naturwissenschaften* (Harri Deutsch, 2006).

Klausur

Termin und Ort der Klausur stehen noch nicht fest, werden aber rechtzeitig bekannt gegeben. Die Klausur wird wahrscheinlich nach Ende der Vorlesungszeit stattfinden.

- Für grundständige Studierende (das "normale" Lehramt Physik) ist die Rechenmethoden-Vorlesung unbenotet. Die Klausur muss nur bestanden werden, dies ist eine Bestehensvoraussetzung für das Modul Physik 1. Wer mindestens 50% der Punkte der Übungsaufgaben als vorrechenbar angekreuzt hat, bekommt 10% Bonuspunkte in der Klausur (mehr dazu in der Übung).
- Im Rahmen der berufsbegleitenden wissenschaftlichen Qualifizierung für Lehrkräfte (also für "Seiteneinsteiger/innen") ist diese Vorlesung Teil des Pflichtmoduls Rechenmethoden (weitere Bestandteile: Übung und Selbststudium). Die Klausur ist in diesem Fall eine benotete Prüfungsleistung (Modulnote = Klausurnote). Prüfungsvorleistung ist das mündliche Lösen von Übungsaufgaben, wobei 1/3 der Aufgaben als vorrechenbar angekreuzt werden müssen (mehr dazu in der Übung). Wer mindestens 50% der Punkte der Übungsaufgaben angekreuzt hat, bekommt 10% Bonuspunkte in der Klausur.

Lernziele

Die Vorlesung hat folgende Lernziele:

- Die Teilnehmenden beherrschen grundlegende Rechenmethoden der Physik. Dies umfasst:
  - Vektoralgebra,
  - lineare Algebra,
  - (Vektor-) Analysis und Analysis von Funktionen mehrerer Variablen (inkl. Koordinatentransformationen und Nabla-Operator),
  - Differentialrechnung und Taylor-Entwicklung,
  - Integralrechnung (inkl. Integralsätzen),
  - Theorie gewöhnlicher Differentialgleichungen,
  - komplexe Zahlen.
- Sie können diese Methoden zur Lösung konkreter Aufgabenstellungen anwenden und ihren Lösungsweg verständlich darstellen.

Hinweis: dieses Skript enthält ganz sicher noch den einen oder anderen Fehler.
Falls Sie also inhaltliche Fragen haben oder einen Fehler finden, melden Sie sich gerne unter tobias.meng@tu-dresden.de!

Inhaltsverzeichnis

Kapitel 1

Vektoralgebra

In diesem ersten Kapitel beschäftigen wir uns mit **Vektoren** und deren **Algebra**, also Recheneigenschaften. Der Stoff dieses Kapitels sollte in großen Teilen in der Schule behandelt worden sein, wir machen hier aber schon erste Verknüpfungen mit darüber hinausgehenden Themen.

1.1 Was ist ein Vektor?

Grob gesagt kann man Vektoren als mathematische Objekte auffassen, die **Richtung und Länge** haben. Wir bezeichnen einen Vektor mit einem **Symbol (der "Name" des Vektors) über dem ein Pfeil steht**, also zum Beispiel $\vec{a}$ (hier ist der Name des Vektors also "a"). Die Länge des Vektors wird alternativ auch als **Betrag** oder **Norm** des Vektors bezeichnet und hat das Symbol $|\vec{a}|$. Der Betrag eines Vektors ist immer eine Zahl ≥ 0.

Einfaches Bild:
Einen Vektor kann man sich als einen Pfeil vorstellen, der zwei Punkte A und B verbindet:

Anwendungen in der Physik:
In der "echten Welt" (und also auch der Physik, die diese ja beschreibt) haben viele Dinge eine Länge und eine Richtung. Der Verbindungsvektor zwischen zwei Punkten ist z. B. die mathematische Version der Antwort auf die Frage "Wo ist der Bahnhof?", die in Worten "200 Meter in diese Richtung!" lautet. Der Vektor wäre in diesem Beispiel also

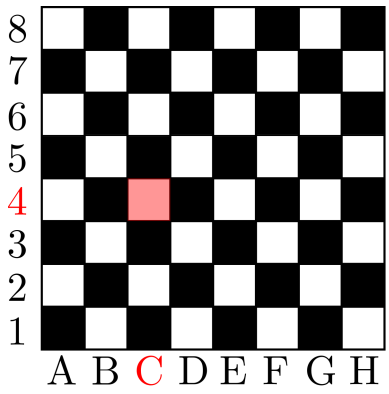
$$\overrightarrow{\text{hier Bahnhof}} \quad \text{mit Betrag} \quad |\overrightarrow{\text{hier Bahnhof}}| = 200\,\text{m}.$$

Andere wichtige vektorielle Größen in der Physik sind unter anderem:

Kraft $\vec{F}$, Ort $\vec{r}$, Impuls $\vec{p}$, Geschwindigkeit $\vec{v}$, Magnetfeld $\vec{B}$.

1.1.1 Koordinatensysteme

Um einen Vektor konkret angeben zu können, benötigen wir ein **Koordinatensystem**. Ein einfaches Beispiel ist die Bezeichnung der Felder eines Schachbretts: jedes Feld ist eindeutig durch die Angabe eines Buchstabens und einer Zahl angegeben, z. B. (C,4).

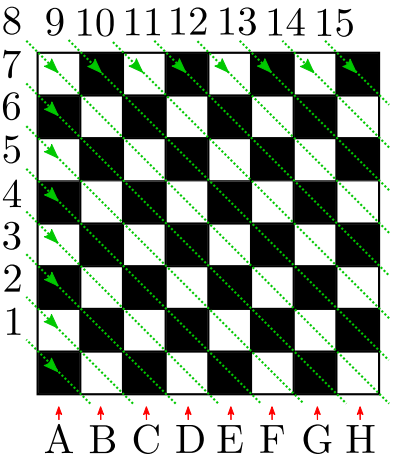


Wichtig: Die Koordinaten des Feldes hängen vom gewählten Koordinatensystem ab!

Rechenaufgabe:

1) Welche Koordinaten hat das selbe Feld im folgenden "schräglinigen" Koordinatensystem?

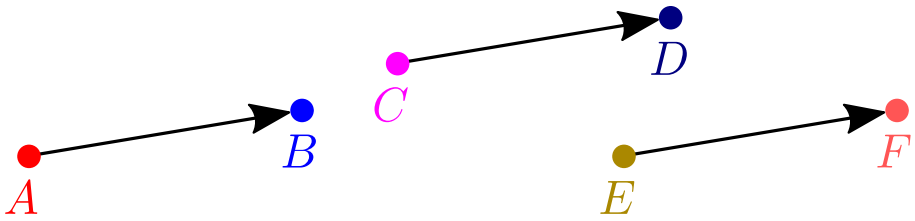
2) Welches Koordinaten hat das Feld, das im schräglinigen System die Koordinaten (C,4) hat, im ursprünglichen Koordinatensystem?



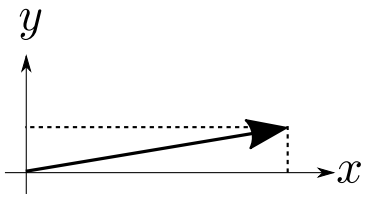
Es gibt also einen Unterschied zwischen einem Feld (das ist, was es ist – unabhängig vom gewählten Koordinatensystem) und seinen Koordinaten (die hängen vom gewählten Koordinatensystem ab). Welches Koordinatensystem man benutzt ist freie Wahl, es gibt aber manchmal Koordinatensysteme, die einfacher sind als andere (beim Schachbrett ist z. B. das erste Koordinatensystem einfacher). Mehr zu Koordinatensystemen und dem Wechsel zwischen verschiedenen Koordinatensystemen gibt es in Kapitel 7.

1.1.2 Standardkoordinaten für Vektoren: kartesische Koordinaten

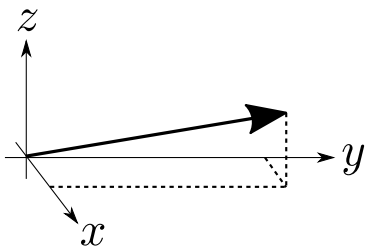
Da ein Vektor laut Definition aus Länge und Richtung besteht, kann der selbe Vektor verschiedene Punktpaare verbinden. In diesem Sinne ist ein Vektor also Repräsentant einer ganzen Klasse von Verbindungspfeilen.



Wie für die Felder des Schachbretts benötigen wir zur konkreten Angabe eines Vektors ein Koordinatensystem. **Wie beim Schachbrett hängen auch beim Vektor die Koordinaten vom gewählten Koordinatensystem ab.** Um einen Vektor mit Zahlen angeben zu können, müssen wir uns zuerst die **Dimension** des Raumes klarmachen, in dem der Vektor liegt. Der Vektor $\overrightarrow{AB}$ von weiter oben liegt in der Ebene der Skript-Seite, die ein zweidimensionales Objekt ist. Das Standard-Koordinatensystem für solche zweidimensionale Vektoren sind die senkrecht zueinander stehenden x und y Koordinaten.

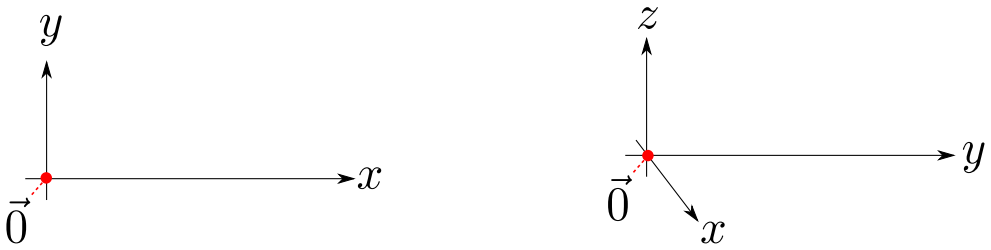


Die meisten Objekte in unserer Welt sind aber dreidimensional. Das Standard-Koordinatensystem für dreidimensionale Vektoren (z.B. der Verbindungsvektor zweier Punkte, die nicht auf der gleichen Höhe liegen) sind die senkrecht zueinander stehenden x, y und z Koordinaten.



Diese Koordinatensysteme nennt man auch "kartesisch".

Was ist jetzt im Koordinatensystem der Anfangspunkt der Vektoren? Da ein Vektor ja nur durch seine Länge und Richtung definiert wird, einigt man sich darauf, den Startpunkt jedes Vektors immer an den selben Punkt zu legen. Diesen nennt man den **Ursprung des Koordinatensystems** und bezeichnet ihn mit $\vec{0}$, 0 oder O. An diesem Punkt schneiden sich die Achsen.



Nachdem man also die Dimension des Vektors (= Anzahl der zu seiner konkreten Angabe nötigen Koordinaten) geklärt hat und ein Koordinatensystem gewählt hat (z. B. das Standard-Koordinatensystem mit x und y Achsen in zwei Dimensionen, bzw. x, y und z Achsen in drei Dimensionen) gibt man den Vektor einfach dadurch an, dass man die Koordinaten (Achsenabschnitte) untereinander schreibt und mit einer großen Klammer umschliesst:

$$\vec{a} = \begin{pmatrix} x\text{-Koordinate} \\ y\text{-Koordinate} \end{pmatrix} \qquad \text{oder} \qquad \vec{b} = \begin{pmatrix} x\text{-Koordinate} \\ y\text{-Koordinate} \\ z\text{-Koordinate} \end{pmatrix}.$$

Diese Vektoren nennt man wegen ihrer Darstellung als Spalte auch **Spaltenvektoren**. Manchmal (siehe später) macht es auch Sinn, einen **Zeilenvektor** zu nutzen, in dem die Einträge in einer Zeile stehen, zum Beispiel $(3, 6, 4)$.

1.1.3 Richtung und Betrag, Einheits- und Basisvektoren

Ein Vektor $\vec{a}$ ist wie gesagt durch seine Richtung und seinen Betrag gekennzeichnet. Der **Betrag** $|\vec{a}|$ (auch **Norm** genannt) ist die Länge des Vektors. Die Richtung des Vektors wird durch den sogenannten **Einheitsvektor in $\vec{a}$-Richtung** angegeben, dem wir das Symbol $\vec{e}_a$ geben. Ein Einheitsvektor ist ein Vektor der Länge 1. Der Einheitsvektor in $\vec{a}$-Richtung ist somit der Vektor, der parallel zu $\vec{a}$ ist und Länge 1 hat. Somit gilt:

$$\vec{a} = |\vec{a}| \cdot \vec{e}_a \qquad \text{“Vektor = Länge mal Richtung”.}$$

Umstellen der Gleichung zeigt, dass die Richtung des Vektors durch folgenden Einheitsvektor beschrieben ist:

$$\vec{e}_a = \frac{\vec{a}}{|\vec{a}|}.$$

Der Betrag eines Vektors berechnet sich als Wurzel aus der Summe der Quadrate der Komponenten:

$$\vec{a} = \begin{pmatrix} a_1 \\ a_2 \end{pmatrix} \quad \Rightarrow \quad |\vec{a}| = \sqrt{a_1^2 + a_2^2}, \tag{1.1a}$$

$$\vec{b} = \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix} \quad \Rightarrow \quad |\vec{b}| = \sqrt{b_1^2 + b_2^2 + b_3^2}. \tag{1.1b}$$

Einen Vektor der Norm 1 (oder auch des Betrages 1 oder auch der Länge 1) nennt man auch **normiert**.

Die Berechnung von $\vec{e}_a$ nennt man auch die **Normierung des Vektors** $\vec{a}$: man berechnet aus dem gegebenen Vektor $\vec{a}$ einen neuen Vektor $\vec{e}_a$, der normiert ist.

In einem gegebenen Koordinatensystem spielen die Einheitsvektoren in Richtung der Koordinatenachsen eine besondere Rolle (siehe später). Diese nennt man **Basisvektoren** (des Koordinatensystems).

Rechenaufgabe:

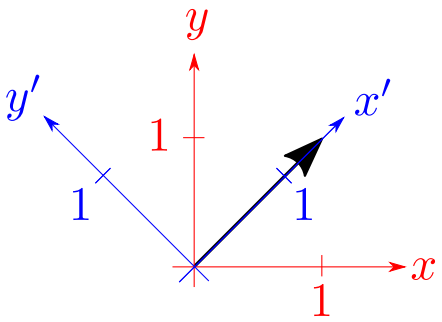
1) Was sind die Basisvektoren des Standard-Koordinatensystems für zweidimensionalen Vektoren?

2) Welchen Betrag hat der Vektor $\vec{c} = \begin{pmatrix} 2 \\ 1 \\ -2 \end{pmatrix}$?

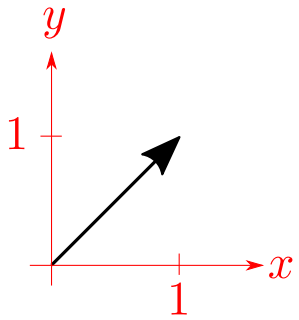
3) Was ist der Einheitsvektor in Richtung des Vektors $\vec{c}$?

1.1.4 Koordinatensystem-Abhängigkeit der Darstellung eines Vektors

Wir können uns nun klarmachen, dass die **Darstellung eines Vektors**, also die Zahlen, die wir in die Spalte schreiben, vom gewählten Koordinatensystem abhängen. Dazu betrachten wir folgendes Beispiel eines gegebenen Vektors $\vec{a}$, den wir in zwei verschiedenen Koordinatensystemen K (rot) und K' (blau) angeben wollen.



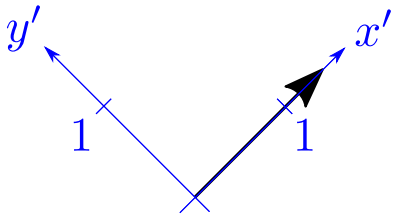
Betrachten wir zunächst das Koordinatensystem K:



Wir können einfach ablesen, dass der Vektor im Koordinatensystem K folgende Darstellung hat:

$$\vec{a}_K = \begin{pmatrix} 1 \\ 1 \end{pmatrix}. \tag{1.2}$$

Der Index K gibt hier das Koordinatensystem an, in dem der Vektor dargestellt ist. Normalerweise gibt man das nicht extra an, hier machen wir es aber zur Klarheit. Das Koordinatensystem K ist zwar das Standard-Koordinatensystem, aber niemand zwingt uns, dieses Koordinatensystem zu verwenden. Alternativ können wir zum Beispiel die gedrehten Koordinaten K' nutzen:



Hier ist der Vektor nur entlang der x'-Achse orientiert, und hat also im Koordinatensystem K' die Darstellung $\vec{a}_{K'} = \begin{pmatrix} a_{x'} \\ a_{y'} \end{pmatrix} = \begin{pmatrix} a_{x'} \\ 0 \end{pmatrix}$. Das Koordinatensystem K' ist im Vergleich zu K einfach nur gedreht. Da sich bei einer Drehung die Länge eines Vektors nicht ändert, und diese durch den Betrag des Vektors gegeben ist, gilt

$$|\vec{a}_K| = |\vec{a}_{K'}| \ , \ |\vec{a}_K| = \sqrt{1^2 + 1^2} = \sqrt{2} \quad \Rightarrow \quad |\vec{a}_{K'}| = \sqrt{{a_{x'}}^2 + 0} = \sqrt{{a_{x'}}^2} = |a_{x'}| = \sqrt{2}. \quad (1.3)$$

Wir finden also, dass $a_{x'} = \sqrt{2}$ oder $a_{x'} = -\sqrt{2}$ sein muss. Da der Vektor $\vec{a}$ in positive x'-Richtung zeigt, finden wir also

$$\vec{a}_{K'} = \begin{pmatrix} \sqrt{2} \\ 0 \end{pmatrix}. \quad (1.4)$$

Wie bei den Feldern des Schachbretts sehen wir also im Vergleich von Gleichungen (1.2) und (1.4), dass die Darstellung eines Vektors vom gewählten Koordinatensystem abhängt. Mehr zu anderen Koordinatensystemen gibt es in Kapitel 4 und 7. Wenn nicht anders angegeben, verwenden wir im folgenden immer das Standard-Koordinatensystem und lassen den Index K bzw. K' weg.

1.2 Ein bisschen echte Mathematik: der Vektorraum

Mathematisch gesehen ist die anschauliche Definition eines Vektors als "Pfeil mit Richtung und Länge" zu eng gefasst. Allgemeiner sind Vektoren die Elemente eines Vektorraums $\mathbb{V}$. Im Vektorraum gibt es zwei Rechenoperationen, die sogenannte Vektoraddition "$\oplus$" und die sogenannte Skalarmultiplikation "$\odot$". Wir betrachten hier den Spezialfall eines "Vektorraums über den reellen Zahlen", der folgende Eigenschaften hat:

1. Sind $\vec{a}$ und $\vec{b}$ Vektoren in $\mathbb{V}$, so ist auch $\vec{a} \oplus \vec{b}$ ein Vektor in $\mathbb{V}$. Für beliebige Vektoren $\vec{a}, \vec{b}, \vec{c} \in \mathbb{V}$ gilt bezüglich der Vektoraddition ferner
   (a) Assoziativgesetz: $\vec{a} \oplus (\vec{b} \oplus \vec{c}) = (\vec{a} \oplus \vec{b}) \oplus \vec{c}$.
   (b) Kommutativgesetz: $\vec{a} \oplus \vec{b} = \vec{b} \oplus \vec{a}$.
   (c) Existenz des neutralen Elements: in $\mathbb{V}$ gibt es einen Vektor $\vec{0}$ so dass $\vec{a} \oplus \vec{0} = \vec{a}$ für alle $\vec{a} \in \mathbb{V}$.
   (d) Existenz des inversen Elements: zu jedem $\vec{a} \in \mathbb{V}$ gibt es einen anderen Vektor $\vec{a}\,'$ in $\mathbb{V}$ so, dass $\vec{a} \oplus \vec{a}\,' = \vec{0}$. (Wir schreiben dieses inverse Element als $\vec{a}\,' = -\vec{a}$).
2. Ist $\vec{a}$ ein Vektor in $\mathbb{V}$ und $\alpha \in \mathbb{R}$ (reelle Zahl) , so ist auch $\alpha \odot \vec{a}$ ein Vektor in $\mathbb{V}$. Für beliebige Vektoren $\vec{a}, \vec{b} \in \mathbb{V}$ und beliebige Zahlen $\alpha, \beta \in \mathbb{R}$ gilt bezüglich der Multiplikation mit Skalaren außerdem:
   (a) Distributivgesetz bzgl. Vektoraddition: $\alpha \odot (\vec{a} \oplus \vec{b}) = \alpha \odot \vec{a} \oplus \alpha \odot \vec{b}$.

(b) Distributivgesetz bzgl. skalarer Addition: $(\alpha + \beta) \odot \vec{a} = \alpha \odot \vec{a} \oplus \beta \odot \vec{a}$.

(c) Kompatibilität der Skalarmultiplikation: $(\alpha\,\beta) \odot \vec{a} = \alpha \odot (\beta \odot \vec{a})$.

(d) Neutralität der Skalarmultiplikation bzgl. der Eins: $1 \odot \vec{a} = \vec{a}$.

Zusatzinformation: Jede beliebige mathematische Struktur, die diese Bedingungen erfüllt, ist ein Vektorraum. Das trifft insbesondere für die oben diskutierten Beispiele zu, beschreibt aber auch noch ganz andere Objekte. In der Quantenmechanik bildet die Gesamtheit aller Zustände, in denen ein quantenmechanisches System sich befinden kann, zum Beispiel einen speziellen Vektorraum, den man "Hilbertraum" nennt. Die Elemente des Hilbertraums sind genau die Zustände, in denen das System sich befinden kann. Diese sind im mathematischen Sinne also auch Vektoren. Man bezeichnet diese Vektoren nicht mit einem Pfeil, sondern mit einem "Ket": der Zustand mit dem Namen Ψ hat das Symbol $|\Psi\rangle$. Ein weiteres Beispiel für einen Vektorraum ist die Menge der Polynome mit der üblichen Addition und Multiplikation (siehe Übung).

1.2.1 Basis und orthonormales Basissystem

Um Vektoren anzugeben, haben wir bisher den Begriff des Koordinatensystems genutzt. Die Einheitsvektoren in Richtung der Koordinatenachsen haben wir Basisvektoren genannt. Die Menge der Basisvektoren bezeichnen wir als "**Basis**". Man kann den Begriff der Basis, bzw. der Basisvektoren, auch allgemeiner fassen: eine Basis ist eine Menge von Vektoren $\{\vec{b}_1, \vec{b}_2, \ldots\}$, die es erlauben, **jeden beliebigen** Vektor $\vec{a}$ als **eindeutige** Linearkombination der Basisvektoren zu schreiben (siehe auch Kapitel 1.5):

$$\vec{a} = \alpha_1(\vec{a})\,\vec{b}_1 + \alpha_2(\vec{a})\,\vec{b}_2 + \ldots = \sum_{i=1,2,\ldots} \alpha_i(\vec{a})\,\vec{b}_i. \tag{1.5}$$

Hierbei sind $\alpha_i(\vec{a})$ Zahlen, die vom konkreten Vektor $\vec{a}$ abhängen. Genau wie es verschiedene mögliche Koordinatensysteme gibt, gibt es auch verschiedene mögliche Basen - welche wir nutzen ist unsere freie Wahl. Alle Basen haben aber die gleiche Anzahl von Basisvektoren. Diese ist gleich der Dimension n, in der der Vektor lebt (= die Dimension des Vektorraumes, den wir betrachten = Anzahl der Koordinatenachsen).

Zur genaueren Definition einer Basis benötigen wir das Konzept der linearen Unabhängigkeit: Vektoren $\vec{a}_i$ heißen **linear unabhängig**, falls

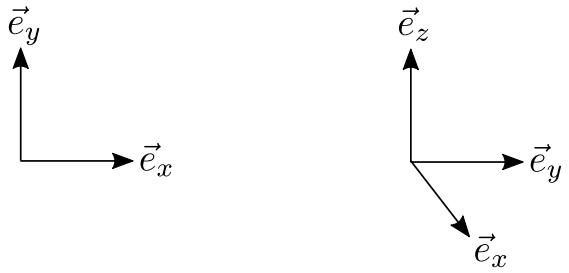
$$\sum_i \alpha_i\,\vec{a}_i = \vec{0} \quad \text{nur erfüllt werden kann, wenn alle Skalare (Zahlen)} \quad \alpha_i = 0. \tag{1.6}$$

Linear unabhängig sind z. B. zwei nicht-parallele Vektoren in der Ebene. Linear abhängig sind z. B. 3 koplanare Vektoren im Raum (drei Vektoren in der selben Ebene). Man kann zeigen, dass die größte Anzahl linear unabhängiger Vektoren gleich der **Dimension n des Vektorraums** ist, in dem die betracheten Vektoren leben. Jede mögliche Kombination von n **linear unabhängigen Vektoren bildet eine Basis**!

Es gibt also viele verschiedene mögliche Basen. Besonders einfach rechnen kann man aber mit einer **Orthonormalbasis**. Diese wird durch n Vektoren $\{\vec{e}_1, \ldots, \vec{e}_n\}$ gebildet, die

1. alle senkrecht aufeinander stehen (also **ortho**gonal zueinander sind – damit sind sie insbesondere auch linear unabhängig) und

2. alle **norm**iert sind, $|\vec{e}_i| = 1$ für alle $i = 1, \ldots, d$.



Ein Beispiel für eine Orthonormalbasis ist die Menge der normierten Basisvektoren in kartesischen Koordinatensystemen.

1.3 Tensoren: mathematische Objekte mit Indizes

Wir haben **Vektoren** durch ihre Koordinaten angegeben, die wir in eine Spalte untereinander geschrieben haben. Für einen gegebenen Vektor $\vec{a}$ bezeichnet den i-ten Eintrag von oben (die i-te Spalte des Vektors) auch als die i-te **Komponente** des Vektors. Diese hat das Symbol a_i. Beispiel:

$$\vec{a} = \begin{pmatrix} 3 \\ 5 \\ 32930 \end{pmatrix} \quad \rightarrow \quad a_1 = 3\,,\ a_2 = 5\,,\ a_3 = 32930.$$

Es gibt in der Physik auch Objekte, die reine Zahlen sind. Diese nennt man auch **Skalare**. Beispiele hierfür sind zum Beispiel die Masse m, die Temperatur T oder die Zeit t.

Andere wichtige physikalische Größen sind hingegen durch **Matrizen** beschrieben, zum Beispiel der Trägheitstensor Θ,

$$\Theta = \begin{pmatrix} \Theta_{xx} & \Theta_{xy} & \Theta_{xz} \\ \Theta_{yx} & \Theta_{yy} & \Theta_{yz} \\ \Theta_{zx} & \Theta_{zy} & \Theta_{zz} \end{pmatrix}.$$

Um eine Komponente einer solchen Matrix zu benennen, muss man sowohl die Zeile als auch die Spalte angeben. Der Eintrag in der i-ten Zeile und j-ten Spalte einer Matrix M wird mit M_{ij} bezeichnet.

Die drei Objektklassen Skalare, Vektoren und Matrizen kann man zum allgemeinen Konzept der **Tensoren** zusammenfassen. Ein Tensor ist ein Objekt mit Komponenten, also Einträgen, in denen irgendetwas steht. Um einen spezifischen Eintrag zu erhalten, muss man bei einem **Tensor der n-ten Stufe** genau n Indizes angeben. Ein Tensor n-ter Stufe hat also Einträge $T_{ij\ldots q}$, wobei T hier eben n Indizes i, j, …, q hat. Somit gilt:

- Ein Skalar ist ein Tensor nullter Stufe.
- Ein Vektor ist ein Tensor erster Stufe.
- Eine Matrix ist ein Tensor zweiter Stufe.

Mehr zu Tensoren und dem Rechnen mit Indizes folgt in Kapitel 3.

1.4 Wichtige Rechenregeln für Vektoren

Im Folgenden besprechen wir noch die wichtigsten Rechenregeln für Vektoren. Diese Regeln sollten Sie am Ende des Semester auswendig können: sie werden in allen anderen Fächern einfach vorausgesetzt und ständig benutzt.

Wichtig: Wir setzen im folgenden ein **orthonormales Koordinatensystem** voraus – in nicht-orthonormalen Koordinatensystemen können manche der folgenden Formeln komplizierter werden.

1.4.1 Transposition

Die **Transposition** macht aus einem Spaltenvektor einen Zeilenvektor. Sie entspricht bildlich dem "auf-die-Seite-legen" eines Vektors, und wird mit einem hochgestellten T bezeichnet.

$$\vec{d} = \begin{pmatrix} d_1 \\ d_2 \\ d_3 \end{pmatrix} \quad \Rightarrow \quad \vec{d}^T = \begin{pmatrix} d_1 & d_2 & d_3 \end{pmatrix} \tag{1.7}$$

Das Anwenden der Transposition auf einen Zeilenvektor ergibt einen Spaltenvektor:

$$\vec{e} = \begin{pmatrix} e_1 & e_2 & e_3 \end{pmatrix} \quad \Rightarrow \quad \vec{e}^T = \begin{pmatrix} e_1 \\ e_2 \\ e_3 \end{pmatrix} \tag{1.8}$$

1.4.2 Addition und Subtraktion von Vektoren

Addition und Subtraktion von Vektoren erfolgt **komponentenweise**. Wir ersetzen hier in der allgemeinen Notation von Kapitel 1.2 das $\oplus$ durch ein einfaches "$+$", und führen auch die Subtraktion "$-$" ein:

$$\vec{a} = \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix} \quad \text{und} \quad \vec{b} = \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix} \quad \Rightarrow \quad \vec{a} + \vec{b} = \begin{pmatrix} a_1 + b_1 \\ a_2 + b_2 \\ a_3 + b_3 \end{pmatrix} \quad \text{und} \quad \vec{a} - \vec{b} = \begin{pmatrix} a_1 - b_1 \\ a_2 - b_2 \\ a_3 - b_3 \end{pmatrix} \tag{1.9}$$

1.4.3 Multiplikation eines Vektors mit einem Skalar

Die Multiplikation eines Vektors mit einem **Skalar** (hier der Einfachheit halber wieder einer rellen Zahl, es könnte je nach Vektorraum aber auch eine komplexe Zahl sein, siehe Kapitel 10) erfolgt auch **komponentenweise**. Wir ersetzen hier in der allgemeinen Notation von Kapitel 1.2 das $\odot$ durch ein Leerzeichen, damit wir es von der Skalarmultiplikation unterscheiden können.

$$\vec{a} = \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix} \quad \text{und} \quad \alpha \text{ Skalar} \quad \Rightarrow \quad \alpha\, \vec{a} = \begin{pmatrix} \alpha\, a_1 \\ \alpha\, a_2 \\ \alpha\, a_3 \end{pmatrix}. \tag{1.10}$$

Es gilt weiterhin für Vektoren $\vec{a}$ und $\vec{b}$ und Skalare α und β:

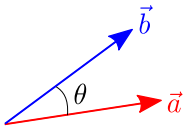
- $\alpha\,(\vec{a} + \vec{b}) = \alpha\,\vec{a} + \alpha\,\vec{b}$.
- $(\alpha + \beta)\,\vec{a} = \alpha\,\vec{a} + \beta\,\vec{a}$

1.4.4 Skalarprodukt von Vektoren

Das Skalarprodukt zweier Vektoren ergibt die Summe der Produkte der Komponenten. Wir bezeichnen es mit einem Punkt zwischen den Vektoren:

$$\vec{a} = \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix} \quad \text{und} \quad \vec{b} = \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix} \quad \Rightarrow \quad \vec{a} \cdot \vec{b} = a_1 b_1 + a_2 b_2 + a_3 b_3. \tag{1.11}$$

Das Skalarprodukt hat eine einfache geometrische Interpretation: bezeichnen wir den Winkel, den die beiden Vektoren $\vec{a}$ und $\vec{b}$ enschliessen, als θ, so gilt



$$\Rightarrow \quad \vec{a} \cdot \vec{b} = |\vec{a}|\,|\vec{b}|\,\cos(\theta) \tag{1.12}$$

Da $-1 \leq \cos(\theta) \leq 1$, folgt aus Gleichung (1.12) die **Schwarzsche Ungleichung**:

$$|\vec{a} \cdot \vec{b}| \leq |\vec{a}|\,|\vec{b}|. \tag{1.13}$$

Es gilt weiterhin für Vektoren $\vec{a}$ und $\vec{b}$ und einen Skalar α:

- $\vec{a} \cdot \vec{b} = \vec{b} \cdot \vec{a}$
- $\alpha\,\vec{a} \cdot \vec{b} = (\alpha\,\vec{a}) \cdot \vec{b} = \vec{a} \cdot \left(\alpha\,\vec{b}\right)$
- $\vec{a} \cdot \vec{a} = |\vec{a}|^2$ und $\vec{a} \cdot \vec{a} = 0 \Leftrightarrow \vec{a} = \vec{0}$ (siehe Gleichung (1.1)) .
- Für $\vec{a}, \vec{b} \neq 0$ ist $\vec{a} \cdot \vec{b} = 0$ genau dann, wenn $\vec{a} \perp \vec{b}$ (siehe Gleichung (1.12)).
- $\vec{a} \cdot \left(\vec{b} + \vec{c}\right) = \vec{a} \cdot \vec{b} + \vec{a} \cdot \vec{c}$

Rechenaufgabe:

Berechnen Sie

$$\left(3\begin{pmatrix} 2 \\ 4 \\ 1 \end{pmatrix} - 2\begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}\right) \cdot \begin{pmatrix} 2 \\ -1 \\ 1 \end{pmatrix} =$$

Wichtig: anders als bei Zahlen, für die es sowohl eine Multiplikation als auch eine Division gibt, ist das Konzept der Division für Vektoren nicht definiert. Es gilt also: **Teile nie durch einen Vektor!**

1.4.5 Kreuzprodukt

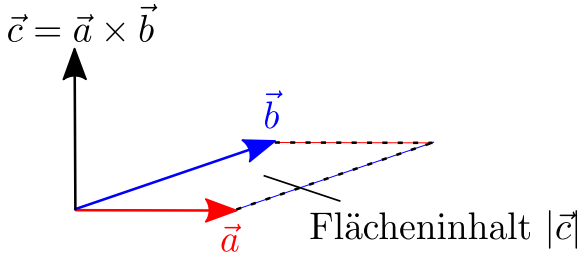
Das Kreuzproukt zweier **dreidimensionaler** Vektoren wird mit $\times$ bezeichnet und ist wie folgt definiert:

$$\vec{a} = \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix} \quad \text{und} \quad \vec{b} = \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix} \quad \Rightarrow \quad \vec{a} \times \vec{b} = \begin{pmatrix} a_2 b_3 - a_3 b_2 \\ a_3 b_1 - a_1 b_3 \\ a_1 b_2 - a_2 b_1 \end{pmatrix}. \tag{1.14}$$

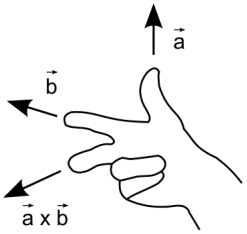
Das Kreuzprodukt ist in der Physik zum Beispiel bei der Lorentz-Kraft wichtig, die ein Teilchen der Ladung q in einem Magnetfeld $\vec{B}$ erfährt, wenn es sich mit Geschwindigkeit $\vec{v}$ bewegt,

$$\vec{F}_{\text{Lorentz}} = q\,\vec{v} \times \vec{B}.$$

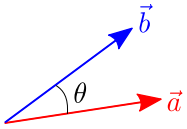
Der Vektor $\vec{c} = \vec{a} \times \vec{b}$ steht senkrecht auf $\vec{a}$ und $\vec{b}$, denn $(\vec{a} \times \vec{b}) \cdot \vec{a} = 0$ und $(\vec{a} \times \vec{b}) \cdot \vec{b} = 0$ (und es gilt Gleichung (1.12)). Die Länge dieses Vektors $\vec{c}$ ist genau der Flächeninhalt des Parallelograms, das von $\vec{a}$ und $\vec{b}$ aufgespannt wird:



Die Richtung (z. B. hoch oder runter) des Vektors $\vec{c}$ folgt aus der "rechte-Hand-Regel" (auch "drei-Finger-Regel"genannt): zeigt $\vec{a}$ entlang des **rechten** Daumens und $\vec{b}$ entlang des **rechten** Zeigefingers, so zeigt $\vec{c} = \vec{a} \times \vec{b}$ in Richtung des **rechten** Mittelfingers (wenn dieser senkrecht zu den anderen beiden steht).



Ähnlich zum Skalarprodukt hat auch das Kreuzprodukt eine geometrische Interpretation: bezeichnen wir den Winkel, den die beiden Vektoren $\vec{a}$ und $\vec{b}$ einschliessen, als θ (mit $0 \leq \theta \leq \pi$), so gilt



$$\Rightarrow \quad |\vec{a} \times \vec{b}| = |\vec{a}|\,|\vec{b}|\,\sin(\theta) \tag{1.15}$$

Es gilt weiterhin für Vektoren $\vec{a}$, $\vec{b}$ und $\vec{c}$ und Skalare α und β:

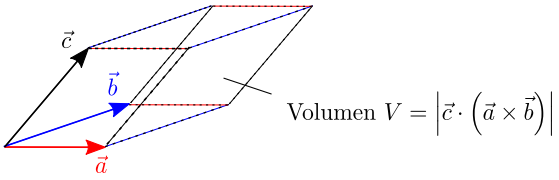
- $\vec{a} \times \vec{b} = -\vec{b} \times \vec{a}$.
- $\vec{a} \parallel \vec{b} \quad \Rightarrow \quad \vec{a} \times \vec{b} = 0$.
- $(\alpha\,\vec{a}) \times (\beta\,\vec{b}) = \alpha\,\beta\,(\vec{a} \times \vec{b})$.
- $\alpha \times (\vec{b} + \vec{c}) = \vec{a} \times \vec{b} + \vec{a} \times \vec{c}$.

1.4.6 Spatprodukt

Als Spatprodukt dreier **dreidimensionaler** Vektoren $\vec{a}$, $\vec{b}$ und $\vec{c}$ bezeichnet man

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

Der Betrag des Spatprodukts der drei Vektoren ergibt das Volumen des von drei Vektoren aufgespannten Parallelepipeds:



1.4.7 Dyadisches Produkt

Als letztes definieren wir noch das "dyadische Produkt" zweier Vektoren $\vec{a}$ und $\vec{b}$, bezeichnet mit $\otimes$, als einen Tensor zweiter Stufe T, für den gilt

$$T = \vec{a}\otimes\vec{b} \quad \Leftrightarrow \quad T_{ij} = a_i\, b_j. \tag{1.16}$$

Für zwei dreidimensionale Vektoren bedeutet das also

$$\vec{a} = \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix} \quad \text{und} \quad \vec{b} = \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix} \quad \Rightarrow \quad \vec{a}\otimes\vec{b} = \begin{pmatrix} a_1\,b_1 & a_1\,b_2 & a_1\,b_3 \\ a_2\,b_1 & a_2\,b_2 & a_2\,b_3 \\ a_3\,b_1 & a_3\,b_2 & a_3\,b_3 \end{pmatrix}.$$

1.5 Zerlegung von Vektoren

Es ist immer wieder hilfreich, einen Vektor $\vec{a}$ in Teile $\vec{a}_i$ zu zerlegen: $\vec{a} = \vec{a}_1 + \vec{a}_2 + \ldots$. Eine wichtige Zerlegungsart ist die in Basisvektoren: hat man eine Basis $\{\vec{b}_1, \vec{b}_2, \ldots, \vec{b}_d\}$, so kann man laut Gleichung (1.5) jeden Vektor schreiben als

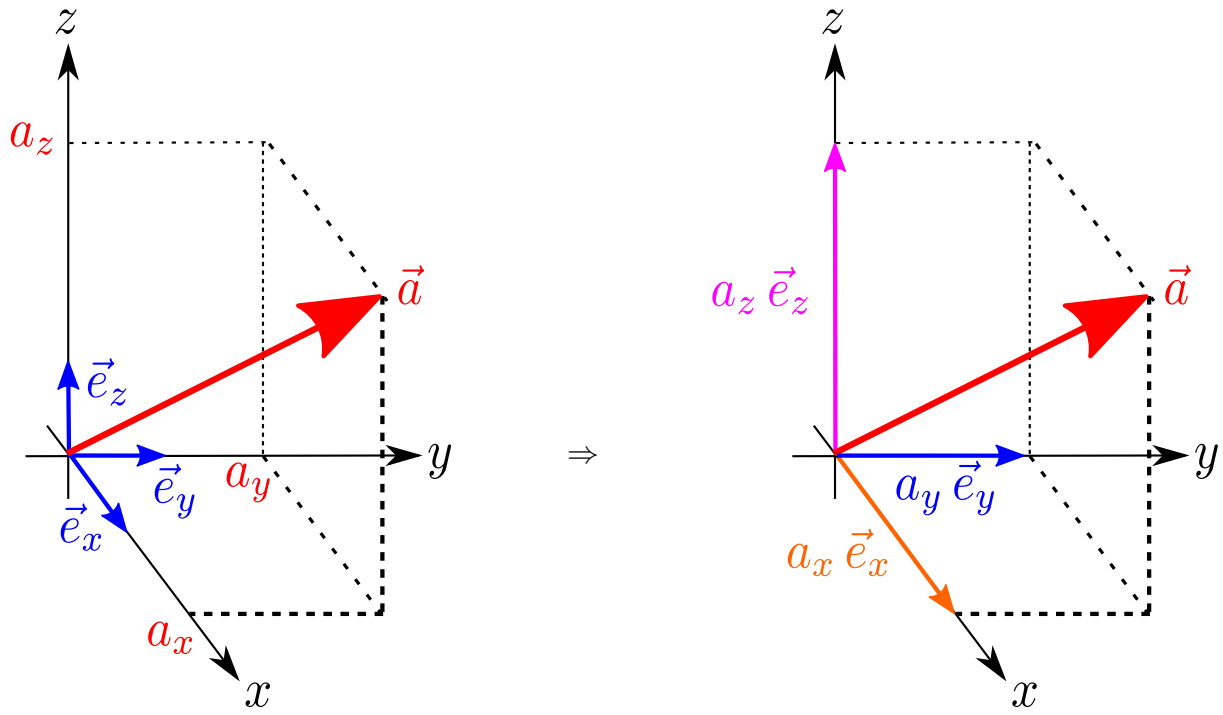
$$\vec{a} = \sum_{i=1}^{d} \alpha_i(\vec{a})\,\vec{b}_i. \tag{1.17}$$

Als Beispiel betrachten wir einen Vektor $\vec{a}$ im dreidimensionalen Vektorraum $\mathbb{R}^3$ (also Vektoren mit drei Zeilen von reellen Zahlen), für den wir die kartesische Standard-Basis $\{\vec{e}_x, \vec{e}_y, \vec{e}_z\}$ nutzen:

$$\vec{a} = \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix}, \quad \vec{e}_x = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \quad \vec{e}_y = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \quad \vec{e}_z = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$$

Wir können den Vektor jetzt in die Basisvektoren zerlegen:

$$\vec{a} = \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix} = \begin{pmatrix} a_1 \\ 0 \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ a_2 \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ 0 \\ a_3 \end{pmatrix} = a_1 \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + a_2 \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} + a_3 \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = a_1\,\vec{e}_x + a_2\,\vec{e}_y + a_3\,\vec{e}_z.$$



Eine andere wichtige Art der Zerlegung eines Vektor $\vec{a}$ ist die in Komponenten parallel und senkrecht zu einem anderen Vektor $\vec{b}$. Hierzu schreiben wir

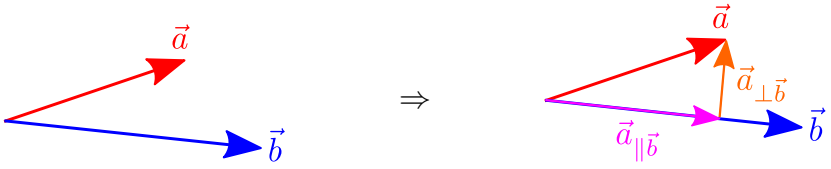
$$\vec{a} = \vec{a}_{\parallel\vec{b}} + \vec{a}_{\perp\vec{b}}.$$

Um den Anteil $\vec{a}_{\parallel\vec{b}}$ parallel zu $\vec{b}$ zu bestimmen, bilden wir einfach das Skalarprodukt mit dem Einheitsvektor $\vec{e}_b$ in $\vec{b}$-Richtung:

$$\vec{a}_{\parallel\vec{b}} = (\vec{a}\cdot\vec{e}_b)\ \vec{e}_b = \frac{\vec{a}\cdot\vec{b}}{|\vec{b}|}\frac{\vec{b}}{|\vec{b}|}. \tag{1.18}$$

Mit Hilfe von Gleichung (1.12) (und ein bisschen Geometrie des Kosinus) kann man sich davon überzeugen, dass dies in der Tat der Anteil von $\vec{a}$ ist, der parallel zu $\vec{b}$ liegt. Der Anteil $\vec{a}_{\perp\vec{b}}$ senkrecht zu $\vec{b}$ kann dann bestimmt werden durch

$$\vec{a}_{\perp\vec{b}} = \vec{a} - \vec{a}_{\parallel\vec{b}}. \tag{1.19}$$



Das wäre im Physikerleben gut zu wissen:

- Definition von "Vektor" kennen und verstehen.
- Konzept des Koordinatensystems kennen.
- Verstehen, dass es viele verschiedene Koordinatensysteme gibt.
- Verstehen, dass die Darstellung des Vektors (also die Zahlen in der Spalte) vom Koordinatensystem abhängt – der Vektor selbst aber nicht.
- Vektoren in kartesischen Koordinaten angeben können.
- Das Konzept eines "Vektorraums" grob beschreiben können.
- Verstehen, was "Basis", "Orthonormalbasis" und "Basisvektoren" sind.
- Vektoren in Basisvektoren und Komponenten parallel/senkrecht zu einem anderen Vektor zerlegen können.
- Einfache Vektor-Rechenoperationen beherschen:
  - Norm berechnen können.
  - Einheitsvektoren berechnen können.
  - Transposition.
  - Multiplikation mit Skalaren.
  - Addition und Subtraktion von Vektoren.
  - Skalarprodukt.
  - Kreuzprodukt.
  - Spatprodukt.

PS: "gut zu wissen" muss nicht heißen, dass das alles in der Klausur abgefragt wird – oder, dass sonst nichts aus diesem Kapitel in der Klausur vorkommt. Wer aber mit den Punkten in diesem Kasten nichts anfangen kann, hat wahrscheinlich in der Klausur und im weiteren Studium Probleme.

Kapitel 2

Abbildungen, Funktionen mehrerer Variablen, Vektorfunktionen, Felder, Folgen und Reihen

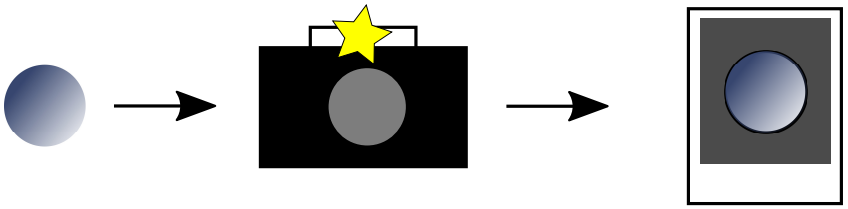
In der Physik geht es oft darum, Zusammenhänge zu beschreiben. Ein Zusammenhang ist oft eine Aussage vom Typ "Wenn A gegeben ist, dann folgt B". Allerdings führen unterschiedliche Ausgangsbedingungen im Allgemeinen zu unterschiedlichen Ergebnissen:

"Wenn ich den Ball fallen lasse, fällt er nach unten. Wenn ich den Ball hochwerfe, fliegt er (zumindest zuerst) nach oben".

In diesem Kapitel werden die Grundlagen zur mathematischen Beschreibung solcher Zusammenhänge gegeben.

2.1 Der Begriff der Abbildung

Ein aus dem Alltag bekanntes Beispiel einer Abbildung ist ein Foto. Dabei erfolgt die Erzeugung des Abbilds dadurch, dass Licht von einem Objekt in eine Kameralinse fällt. Die Kamera macht dann irgendetwas Kompliziertes und erzeugt am Ende ein Bild des Objekts. Man könnte also sagen: "Wenn ein Objekt gegeben ist, dann folgt daraus nach dem Auslösen der Kamera ein bestimmtes Bild (nämlich das des gegebenen Objekts)".



Objekt → Kamera → Bild

Mathematisch ausgedrückt ordnet also die Kamera jedem Objekt ein Bild zu. Allgemeiner definiert die Mathematik zur Beschreibung solcher Zuordnungen den Begriff der "Abbildung" oder "Funktion" wie folgt:

Eine Abbildung oder auch Funktion f ist eine Zuordnung zwischen zwei Mengen $A = \{a_1, a_2, \ldots\}$ und $B = \{b_1, b_2, \ldots\}$, wobei jedem Element a_i der Menge A durch die Funktion f genau ein Element b_j der Menge B zugeordnet wird:

$$a_i \quad \rightarrow \quad f \quad \rightarrow \quad b_j$$

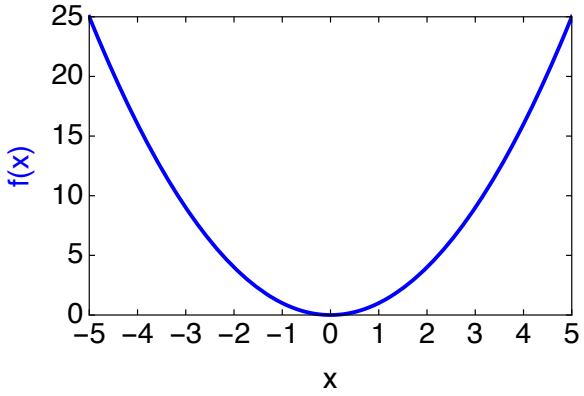
Mathematisch nennt man die Ausgangsmenge A den **Definitionsbereich** und die Endmenge B den **Wertebereich**. Man schreibt die Definition der Funktion wie folgt:

$$f : A \to B, \quad a \mapsto f(a). \tag{2.1}$$

In Worten: f ist eine Funktion von A nach B, die jedem a ein $b = f(a)$ zuordnet. Ein Beispiel für eine Funktion ist $f(x) = x^2$, die für alle reelle Zahlen x definiert ist. Sie ordnet jeder reellen Zahl x die positive reelle Zahl x^2 zu, also z. B. der Zahl -3 die Zahl $(-3)^2 = 9$. Diese schreibt man also mathematisch wie folgt:

$$f : \mathbb{R} \to \mathbb{R}^+, \quad x \mapsto x^2. \tag{2.2}$$

In kurz schreibt man oft auch einfach nur $f(x) = x^2$ ($f(x)$ ist der Funktionswert von f an der Stelle x, in dieser kurzen Form fehlt die Angabe von Definitions- und Wertebereich). Diese Funktion lässt sich wie folgt zeichnen:



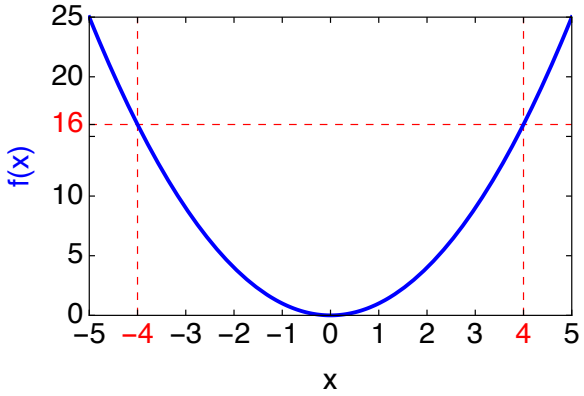
Eines der grundlegendsten Beispiele für die Nutzung von Funktionen in der Physik ist die Angabe des Orts als Funktion der Zeit.

2.2 Umkehrfunktion

Die "Umkehrfunktion" f^{-1} einer Funktion f ordnet einem Funktionswert $y = f(x)$ den Wert x zu:

$$y = f(x) \quad \Rightarrow f^{-1}(y) = x. \tag{2.3}$$

Die Umkehrfunktion ist aber nicht immer eindeutig definiert: oftmals gibt es für einen gebenen Wert y mehrere Werte von x so, dass $y = f(x_1) = f(x_2)$. Ein Beispiel hierfür ist genau die Funktion $f(x) = x^2$, für die sowohl $f(x) = x^2$ ist, als auch $f(-x) = (-x)^2 = x^2$. Es gilt also z. B. $16 = f(4)$ und $16 = f(-4)$.



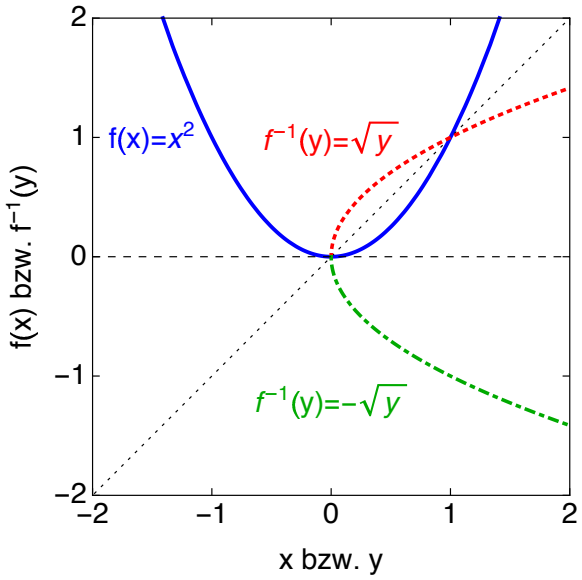
Die Umkehrfunktion "weiß" dann nicht, ob sie dem Wert $y = 4$ den Wert $x = 4$ oder $x = -4$ zuordnen soll - oder, mathematischer gesagt, die Umkehrfunktion kann dann nicht mehr eindeutig definiert werden. Sie zerfällt dann in mehrere "Zweige", sodass die Umkehrfunktion in jedem Zweig eindeutig definiert ist. Für $f(x) = x^2$ gibt es zwei Zweige:

$$f(x) = x^2 \quad \Rightarrow \begin{cases} f^{-1}(y) = \sqrt{y} & \text{auf dem ersten Zweig} \\ f^{-1}(y) = -\sqrt{y} & \text{auf dem zweiten Zweig.} \end{cases} \tag{2.4}$$

Für beide Zweige ist f^{-1} nur für $y \geq 0$ wohldefiniert. Das macht Sinn, denn y soll gleich $f(x)$ sein, und $f(x) = x^2 \geq 0$. Die vollständige Definition von Funktion und Umkehrfunktion sind also:

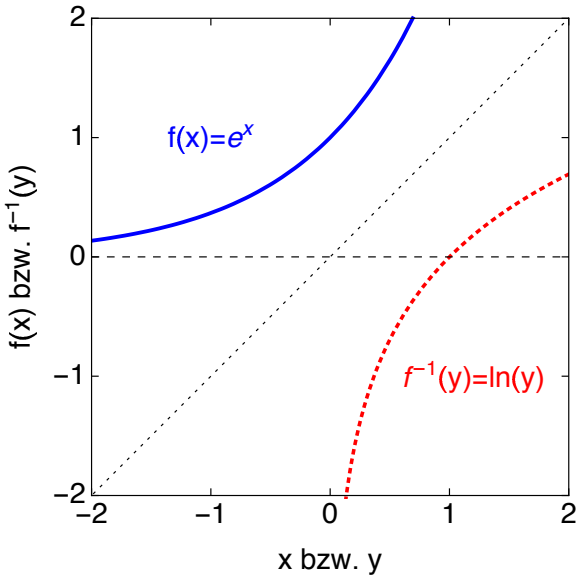
$$f : \mathbb{R} \to \mathbb{R}_0^+, \quad x \mapsto x^2 \quad \Rightarrow \quad \begin{cases} f^{-1} : \mathbb{R}_0^+ \to \mathbb{R}_0^+, & y \mapsto \sqrt{y} \quad \text{auf dem ersten Zweig} \\ f^{-1} : \mathbb{R}_0^+ \to \mathbb{R}_0^-, & y \mapsto -\sqrt{y} \quad \text{auf dem zweiten Zweig.} \end{cases} \tag{2.5}$$

($\mathbb{R}_0^+$ bezeichnet alle positiven reellen Zahlen inklusive der Null, $\mathbb{R}_0^-$ alle negativen reellen Zahlen inklusive der Null.) Durch zeichnen von Funktion und Umkehrfunktion sieht man, dass die Umkehrfunktion sich geometrisch immer genau durch die Spiegelung der Funktion an der Hauptdiagonalen ergibt (das stimmt immer, nicht nur bei $f(x) = x^2$).



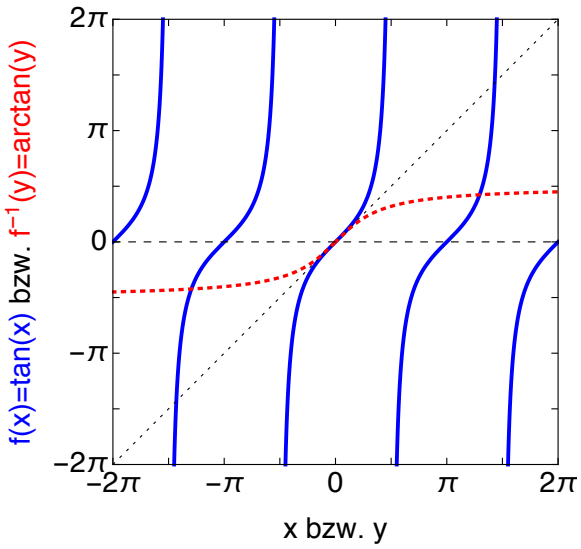
Zu den in der Physik viel-genutzten Funktionen zählt auch die Exponentialfunktion $f(x) = e^x$, die die Umkehrfunktion Logarithmus naturalis $f^{-1}(y) = \ln(y)$ hat:

$$f : \mathbb{R} \to \mathbb{R}_0^+, \quad x \mapsto e^x \quad \Rightarrow f^{-1} : \mathbb{R}_0^+ \to \mathbb{R}, \quad y \mapsto \ln(y). \tag{2.6}$$



Für die trigonometrischen Funktionen $f(x) = \sin(x)$, $f(x) = \cos(x)$ und $(x) = \tan(x)$, die ja periodisch in x sind, haben die Umkehrfunktionen wieder verschiedene Zweige. Bei diesen Umkehrfunktionen nennt man den Zweig, der die Null einschließt, "Arkussinus" $\arcsin(y)$ (für den Sinus), "Arkuskosinus" $\arccos(y)$ (für den Kosinus) und "Arkustangens" $\arctan(y)$ (für den Tangens). Diese sind tabelliert, bzw. in Computern und Taschenrechnern eingespeichert. Für den Arkustangens gilt zum Beispiel:

$$f : \mathbb{R} \to \mathbb{R}, \quad x \mapsto \tan(x) \quad \Rightarrow \quad f^{-1} : \mathbb{R} \to \left[-\frac{\pi}{2}, \frac{\pi}{2}\right], \quad y \mapsto \arctan(y). \tag{2.7}$$

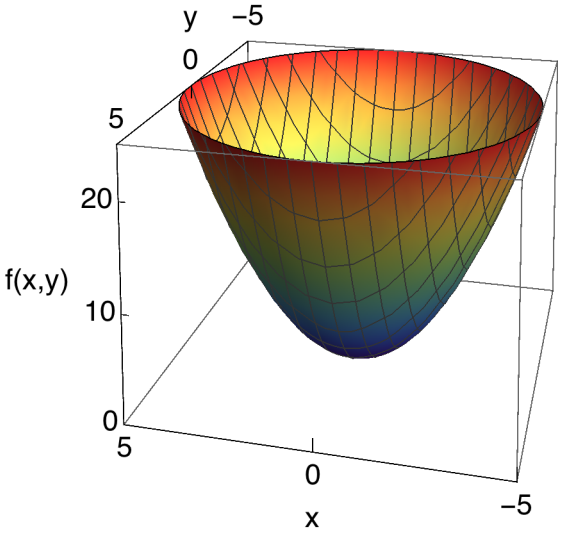


2.3 Skalarwertige Funktionen mehrerer Variablen

Eine Kamera kann auch Bilder von mehreren Objekten gleichzeitig machen. Genauso können auch Funktionen so definiert werden, dass sie mehrere Eingangs-Objekte in ein End-Objekt abbilden – dass sie also mehrere Variablen auf einen Funktionswert abbilden. Wir beschränken uns hier zunächst auf Skalare als Funktionswerte. Ein Beispiel hierfür ist eine Funktion, die zwei Zahlen $x \in \mathbb{R}$ und $y \in \mathbb{R}$ die Summe der Quadrate zuordnet:

$$f : \mathbb{R} \times \mathbb{R} \to \mathbb{R}, \quad (x, y) \mapsto x^2 + y^2. \tag{2.8}$$

In Worten: f ist eine Funktion von zwei reellen Zahlen (das ist die Bedeutung von $\mathbb{R} \times \mathbb{R}$) in die reellen Zahlen, die jedem Paar (x, y) den Wert $x^2 + y^2$. zuordnet. Es gilt also z. B. $f(2, 3) = 2^2 + 3^2 = 4 + 9 = 13$. Diese Funktion lässt sich wie folgt zeichnen:



Man kann die Variablen, hier x und y, auch als Komponenten eines Vektors $\vec{r} = \begin{pmatrix} r_1 \\ r_2 \end{pmatrix} = \begin{pmatrix} x \\ y \end{pmatrix}$ auffassen. Die Funktion f kann man dann als Funktion eines Vektors verstehen. Im konkreten Fall ordnet diese Funktion dem Vektor $\vec{r}$ das Quadrat seines Betrages zu (vgl. Kapitel 1.1.3):

$$f : \mathbb{R}^2 \to \mathbb{R}_0^+, \quad \vec{r} \mapsto |\vec{r}|^2. \tag{2.9}$$

Hierbei haben wir $\mathbb{R} \times \mathbb{R}$ als $\mathbb{R}^2$ geschrieben. Man identifiziert also

$$f(x, y) = f(r_1, r_2) = f(\vec{r}).$$

Dieses Konzept lässt sich direkt auf Funktionen von beliebig vielen Variablen erweitern. Eine solche Funktion weist einem n-Tupel von Eingangswerten $(x_1, \ldots, x_n)$ einen Funktionswert $f(x_1, \ldots, x_n)$ zu (der wieder eine Zahl ist). Das n-Tupel kann man auch als n-dimensionalen Vektor auffassen:

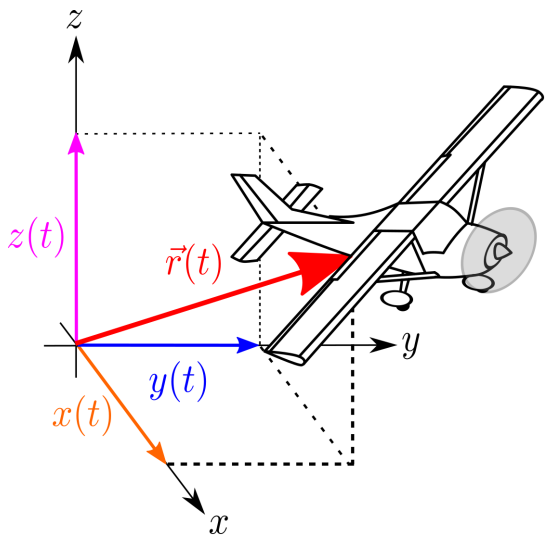
$$\vec{x} = \begin{pmatrix} x_1 \\ x_2 \\ \ldots \\ x_n \end{pmatrix} \quad \Rightarrow \quad f : \mathbb{R}^n \to \mathbb{R}, \quad \vec{x} \mapsto f(\vec{x}) = f(x_1, x_2, \ldots, x_n). \tag{2.10}$$

In der Physik spielen solche Funktionen als (Skalar-)Felder eine große Rolle (siehe ein bisschen weiter unten).

2.4 Vektorwertige Funktionen einer Variablen

Bisher haben wir nur Funktionen betrachtet, die einer oder mehreren Eingangsvariablen eine Zahl zugewiesen haben. Dieses Konzept lässt sich dahingehend verallgemeinern, dass man einer oder mehreren Eingangsvariablen einen **Vektor** zuweist. Eine Situation, in der ein Vektor als Funktionswert

zum Beispiel sinnig ist, ist die Angabe des Ortes eines Flugzeugs als Funktion der Zeit. Während des Fluges ändert sich sowohl die Stelle, über der das Flugzeug fliegt (seine x- und y-Koordinaten), als auch seine Höhe (die z-Koordinate).



Den Ortsvektor $\vec{r}(t)$ des Flugzeugs als Funktion der Zeit muss man also wie folgt angeben

$$\vec{r}: \mathbb{R} \to \mathbb{R}^3, \quad t \mapsto \vec{r}(t) = \begin{pmatrix} x(t) \\ y(t) \\ z(t) \end{pmatrix}. \tag{2.11}$$

Eine vektorwertige Funktion kann man alternativ auch einfach als eine Menge von Funktionen $x(t), y(t), z(t)$ auffassen, die man einfach in einem Vektor untereinander schreibt – hier also die Menge der Koordinaten des Flugzeugs als Funktion der Zeit (davon gibt es drei).

2.5 Vektorwertige Funktionen mehrerer Variablen

Noch allgemeiner kann man Funktionen auch so definieren, dass sie einer Menge von m Eingangsvariablen $x_1, \ldots, x_m$ eine Menge von n skalaren Funktionen $f_1(x_1, \ldots, x_m), \ldots, f_n(x_1, \ldots, x_m)$ zuweisen, wobei n und m beliebige natürliche Zahlen (ohne Null) sind. Die n Funktionen $f_j(x_1, \ldots, x_m)$ fasst man nun wieder als einen n-komponentigen Vektor auf. Genauso kann man die m Variablen als einen m-komponentigen Vektor auffassen. Man erhält also eine Funktion, die einem m-komponentigen Vektor von Variablen einen n-komponentigen Vektor als Funktionswert zuweist:

$$\vec{f}: \mathbb{R}^m \to \mathbb{R}^n, \quad \vec{x} = \begin{pmatrix} x_1 \\ \ldots \\ x_m \end{pmatrix} \mapsto \vec{f}(\vec{x}) = \begin{pmatrix} f_1(\vec{x}) \\ f_2(\vec{x}) \\ \ldots \\ \ldots \\ f_n(\vec{x}) \end{pmatrix} = \begin{pmatrix} f_1(x_1, \ldots, x_m) \\ f_2(x_1, \ldots, x_m) \\ \ldots \\ \ldots \\ f_n(x_1, \ldots, x_m) \end{pmatrix}. \tag{2.12}$$

Eine "normale" Funktion erhält man für $n = 1$ und $m = 1$. Eine skalare Funktion mehrerer Variablen erhält man mit $m > 1$ und $n = 1$. Eine vektorwertige Funktion einer Variablen erhält man mit $m = 1$, $n > 1$.

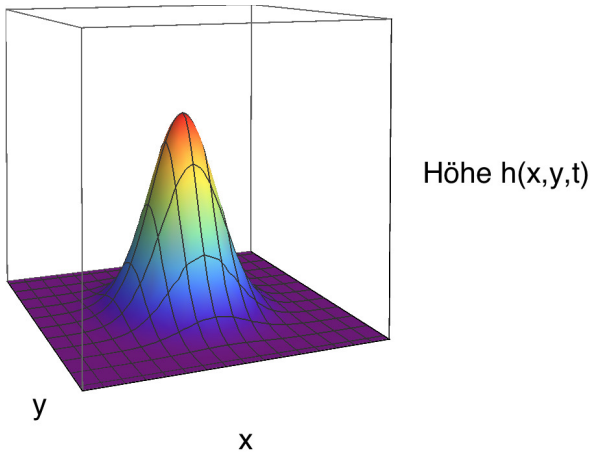
2.6 Felder

In der Physik bezeichnet man mit dem Wort "Feld" eine Größe Φ, die jedem Punkt $\vec{r}$ im Raum und jedem Zeitpunkt t einen Wert $\Phi(\vec{r}, t)$ zuordnet. Manchmal betrachtet man auch sogenannte "statische" Felder $\Phi(\vec{r})$, die nur vom Ort und nicht von der Zeit abhängen. Ein Feld ist also letztlich eine Funktion mehrerer Variablen. Abhängig von der Natur des Funktionswerts unterscheidet man z. B.:

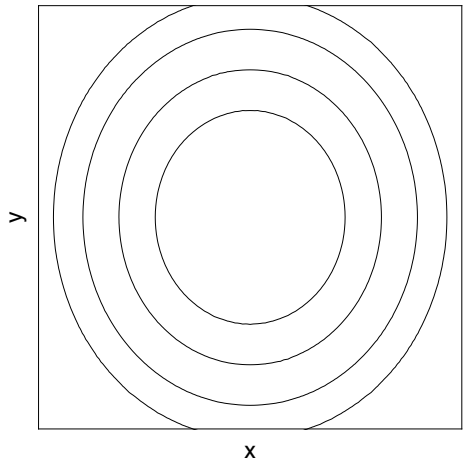
1. **Skalarfelder**: Ein Skalarfeld ordnet jedem Punkt $\vec{r}$ zu jedem Zeitpunkt t eine Zahl (Skalar) zu. Beispiele sind:

   - Temperatur $T(\vec{r}, t)$
   - Druck $p(\vec{r}, t)$
   - Dichte $\rho(\vec{r}, t)$

   Oft interessiert man sich nicht nur für den Wert des Feldes an einem bestimmten Punkt $(\vec{r}, t)$, sondern auch für sogenannte "Äquipotentialflächen" oder "Äquipotentiallinien". Diese sind definiert dadurch, dass das Feld auf Äquipotentialflächen einen konstanten Wert annimmt, $\Phi(\vec{r}, t) = \Phi_0 = \text{const.}$. Ein aus dem Alltag bekanntes Skalarfeld ist die Angabe von Höhe über dem Meerespiegel h, die jedem Punkt der Erdoberfläche (x, y) zum Zeitpunkt t eine Höhe $h(x, y, t)$ zuordnet (wobei man, da Berghöhen sich nur sehr langsam ändern, die Zeit t hier oft weglässt). Die Äquipotentiallinien der Höhe sind Linien konstanter Höhe – die aus Wanderkarten bekannten Höhenlinien.



Höhe $h(x, y, t)$

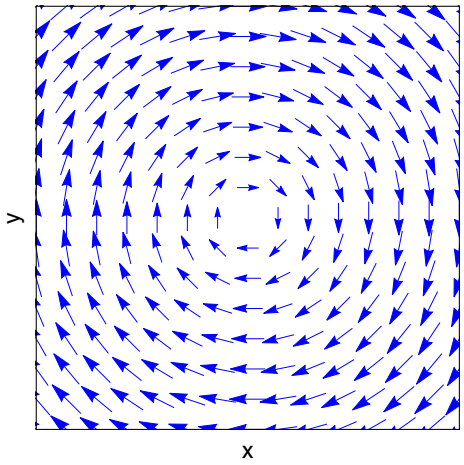


Höhenlinien (Äquipotentiallinien von $h(x, y, t)$)

2. **Vektorfelder**: Ein Vektorfeld ordnet jedem Punkt $\vec{r}$ zu jedem Zeitpunkt t einen Vektor zu. Beispiele sind:

   - elektrisches Feld $\vec{E}(\vec{r}, t)$
   - magnetisches Feld $\vec{B}(\vec{r}, t)$
   - Geschwindigkeit einer Flüssigkeit $\vec{v}(\vec{r}, t)$
   - Kraft (oder auch Kraftfeld) $\vec{F}(\vec{r}, t)$

Ein Vektorfeld wird dadurch dargestellt, dass man für einen gegebenen Zeitpunkt t an jedem Ort $\vec{r}$ den Vektor des Feldes aufträgt. Als Beispiel hier das Feld der Windgeschwindigkeit $\vec{v}(\vec{r}, t)$ um das Auge eines Wirbelsturms herum (50 Meter über der Erdoberfläche) für einen Zeitpunkt t:



3. **Tensorfelder**: der Funktionswert des Feldes ist ein Tensor n-ter Stufe

2.7 Folgen

Eine **Folge** ist Sequenz von durchnummerierten Objekten, also z. B. eine Sequenz von Zahlen. Das i-te Objekt der Folge nennen wir a_i. Als Indizes wählt man normalerweise natürliche Zahlen, $i \in \mathbb{N}$. Formal kann man eine Folge a auffassen als eine Abbildung von der Menge der Indizes i auf die Objekte a_i. Wenn wir die Menge der Objekte a_i mit $\mathbb{X}$ bezeichnen, kann man eine Folge daher als

$$a : \mathbb{N} \to \mathbb{X}, \quad i \mapsto a_i \tag{2.13}$$

schreiben. Von nun an betrachen wir unendliche lange Folgen von reellen Zahlen, d.h. $\mathbb{X} = \mathbb{R}$. Man nennt eine Folge **konvergent** wenn die Zahlen a_i sich mit zunehmend größer werdendem Index i immer mehr einem konstanten Wert a_∞ annähern. Genauer gesagt ist eine Folge konvergent genau dann, wenn

$$\forall\, \epsilon > 0 \quad \exists n : |a_i - a_\infty| < \epsilon \quad \forall i > n. \tag{2.14}$$

In Worten: "Für alle Zahlen $\epsilon > 0$ existiert ein Index n so, dass alle Folgenglieder a_i mit $i > n$ sich von a_∞ um weniger als ϵ unterscheiden". Andernfalls nennt man die Folge **divergent**. Ist die Folge konvergent, so nennt man a_∞ den **Grenzwert** oder **Limes** der Folge, und gibt dies wie folgt an:

$$\lim_{i \to \infty} a_i = a_\infty. \tag{2.15}$$

Selbstcheck:

Welche der folgenden Folgen ist konvergent?

1. Eine sogenannte "alternierende Folge" mit $a_n = (-1)^n$.
2. Eine sogenannte "harmonische Folge" mit $a_n = \frac{1}{n}$.
3. Eine sogenannte "geometrische Folge" mit $a_n = q^n$, wobei q eine reelle Zahl ist.

2.8 Reihen

Eine Reihe bekommt man aus einer Folge indem man die Folgenglieder a_i aufsummiert. Die j-te "Partialsumme" s_j ist dabei definiert als

$$s_j = a_1 + a_2 + \ldots + a_j = \sum_{i=1}^{j} a_i. \tag{2.16}$$

Die Folge der Partialsummen s_j nennt man dann eine **Reihe**. Daher gilt für Reihen, genau wie bei (anderen) Folgen auch, dass Reihen konvergent oder divergent sein können. Eine Reihe ist konvergent, genau dann wenn

$$\forall\, \epsilon > 0 \quad \exists\, n : |s_j - s_\infty| < \epsilon \quad \forall j > n, \tag{2.17}$$

sonst ist sie divergent. Der Grenzwert der Reihe wird dann wie folgt angegeben:

$$\lim_{j\to\infty} s_j = s_\infty. \tag{2.18}$$

Zusatzinformation: geometrische Reihe und vollständige Induktion.

Eine wichtige Reihe ist die sogenannte **geometrische Reihe** mit den Partialsummen

$$s_j = \sum_{i=0}^{j} a_0\, q^i \quad \text{mit} \quad a_0,\, q \in \mathbb{R} \quad \text{und} j \in \mathbb{N}. \tag{2.19}$$

Wann ist die geometrische Reihe konvergent? Hierzu müssen wir erst einmal die Partialsummen s_j berechnen – und zwar für jedes j. Da es unendlich viele natürliche Zahlen gibt, $j = 0, 1, 2, 3, \ldots$, können wir die Partialsummen sicherlich nicht explizit ausrechnen. Wir nutzen stattdessen die Beweistechnik der **vollständigen Induktion**, die es erlaubt, eine Aussage für alle natürlichen Zahlen j zu beweisen. Diese erfolgt nach dem Schema:

1. Induktionsstart: Beweise die Aussage für $j = 0$.
2. Induktionschritt: Beweise, dass wenn die Aussage für j gilt, sie auch für $j + 1$ gilt.

Hiermit beweist man die Aussage für alle j unter den Voraussetzung $|q| \neq 1$. Wir wollen diese Technik also nutzen, um die Partialsummen s_j der geometrischen Reihe für alle natürlichen Zahlen j zu bestimmen. Dazu müssen wir raten, welchen Wert s_j annimmt. Vollständige Induktion erlaubt uns dann, das Geratene zu beweisen.

Um zu raten, was s_j ist, schauen wir uns die ersten paar Glieder der Reihe an. Die können wir einfach ausrechnen, und bekommen:

$$\begin{aligned}
s_0 &= a_0\, q^0 = a_0\, 1 = a_0 \\
s_1 &= a_0\, q^0 + a_0\, q^1 = a_0\, 1 + a_0\, q = a_0\, (1 + q) \\
s_2 &= a_0\, q^0 + a_0\, q^1 + a_0\, q^2 = a_0\, 1 + a_0\, q = a_0\, (1 + q + q^2) \\
s_3 &= a_0\, q^0 + a_0\, q^1 + a_0\, q^2 + a_0\, q^3 = a_0\, 1 + a_0\, q = a_0\, (1 + q + q^2 + q^3)
\end{aligned}$$

Diese können wir jetzt noch geschickt umschreiben (diese Umformung ist offensichtlich eine, die man nicht unbedingt als ersten Ansatz raten würde, aber sie ist hilfreich!):

$$s_0 = a_0 = a_0 \frac{1-q}{1-q}$$
$$s_1 = a_0 (1+q) = a_0 \frac{(1-q)(1+q)}{(1-q)} = a_0 \frac{1-q^2}{1-q}$$
$$s_2 = a_0 (1+q+q^2) = a_0 \frac{(1-q)(1+q+q^2)}{(1-q)} = a_0 \frac{1+q+q^2-q-q^2-q^3}{1-q} = a_0 \frac{1-q^3}{1-q}$$
$$s_3 = a_0 (1+q+q^2+q^3) = a_0 \frac{(1-q)(1+q+q^2+q^3)}{(1-q)} = a_0 \frac{1+q+q^2+q^3-q-q^2-q^3-q^4}{(1-q)}$$
$$= a_0 \frac{1-q^4}{1-q}$$

Wir behaupten jetzt also, dass

$$s_j = a_0 \frac{1-q^{j+1}}{1-q}$$

ist. Dies beweisen wir mittels vollständiger Induktion:

1. Induktionsstart:
$$s_0 = a_0 = a_0 \frac{1-q}{1-q} = a_0 \frac{1-q^{0+1}}{1-q} \quad \checkmark$$

2. Induktionschritt:
$$s_{j+1} = s_j + a_0 q^{j+1} \overset{\text{Voraussetzung}}{=} a_0 \frac{1-q^{j+1}}{1-q} + a_0 q^{j+1} = a_0 \frac{1-q^{j+1}+(1-q)q^{j+1}}{1-q}$$
$$= a_0 \frac{1-q^{j+1}+q^{j+1}-q^{j+2}}{1-q} = a_0 \frac{1-q^{j+2}}{1-q} \quad \checkmark$$

Jetzt, wo wir wissen, welche Werte die Reihenglieder/Partialsummen s_j annehmen, können wir uns dem Grenzwert zuwenden. Laut der Voraussetzung weiter oben haben wir uns auf $q \neq 1$ beschränkt. Wir können uns also auf zwei Fälle beschränken:

- Für $|q| < 1$ gilt, dass $\lim_{j\to\infty} q^j = 0$ ist. Also gilt:
$$\lim_{j\to\infty} s_j = \lim_{j\to\infty} \frac{1-q^{j+1}}{1-q} = \frac{1}{1-q}.$$
Die geometrische Reihe ist also für $|q| < 1$ konvergent und hat den Grenzwert $s_\infty = \frac{1}{1-q}$.
- Für $|q| > 1$ gilt, dass $\lim_{j\to\infty} q^j$ divergiert (für $q > 0$ gilt $\lim_{j\to\infty} q^j = +\infty$, was divergent ist; für $q < 0$ ist s_j eine Folge von Partialsummen mit abwechselndem Vorzeichen und immer größer werdendem Betrag - also auch divergent). Somit ist auch die geometrische Reihe für $|q| > 1$ divergent.

Das wäre im Physikerleben gut zu wissen:

- Definition von Abbildung und Funktion verstehen.
- Funktionen mathematisch genau angeben können (inkl. Definitions- und Wertebereich).
- Verstehen, was eine Umkehrfunktion ist und wissen, wie man sie im Prinzip berechnen kann.
- Verstehen, wann Umkehrfunktionen eindeutig sind, und wann sie verschiedene Zweige haben.
- Einige wichtige Umkehrfunktionen kennen (nicht auswendig die Funktionswerte kennen, aber wissen, wozu diese die Umkehrfunktionen sind):
  - Logarithmus naturalis
  - Arkussinus
  - Arkuskosinus
  - Arkustanges
- Verstehen, was eine skalarwertige Funktion mehrerer Variablen ist.
- Verstehen, was eine vektorwertige Funktion einer oder mehrerer Variablen ist.
- Die Definition von und einige Beispiele für Felder kennen.
- Wissen, was eine Reihe ist.
- Wissen, was eine Folge ist.
- Wissen, was es heißt, wenn eine Reihe oder Folge konvergiert bzw. divergiert.

PS: "gut zu wissen" muss nicht heißen, dass das alles in der Klausur abgefragt wird – oder, dass sonst nichts aus diesem Kapitel in der Klausur vorkommt. Wer aber mit den Punkten in diesem Kasten nichts anfangen kann, hat wahrscheinlich in der Klausur und im weiteren Studium Probleme.

Kapitel 3

Rechnen mit Indizes: Matrizen und Tensoren

In diesem Kapitel werden die grundlegenden Rechenoperationen für Matrizen eingeführt, was uns dann zum Rechnen mit Indizes führt. Eine **Matrix** ist eine Anordnung von Zahlen in einem Tabellenschema. Die Tabelle kann dabei beliebig viele Zeilen und Spalten haben. Anders als eine normale Tabelle, die man wie folgt schreibt,

	Spalte 1	**Spalte 2**	**Spalte 3**	**Spalte 4**
Zeile 1	1	3	4	5
Zeile 2	4	3	1	9
Zeile 3	9	4	3	1

wird eine Matrix durch große Klammern begrenzt:

$$A = \begin{pmatrix} 2 & 1 \\ 1 & 3 \end{pmatrix} \quad \text{oder} \quad B = \begin{pmatrix} -4 & 12 & 4 & 0 \\ 4 & 1 & -4 & 8 \end{pmatrix} \quad \text{oder} \quad C = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad \text{oder} \quad D = \begin{pmatrix} 3 \\ -2 \\ 0 \end{pmatrix}.$$

Eine Matrix mit m Zeilen und n Spalten nennt man eine $(m \times n)$-Matrix. Bei den obigen Beispielen ist also A eine (2×2)-matrix, B eine (2×4)-Matrix und C eine (3×3)-Matrix. Am Beispiel D sieht man, dass man einen (Spalten-)**Vektor** mit n Einträgen also auch als eine $(n \times 1)$-Matrix auffassen kann!

3.1 Matrizen: wichtige Eigenschaften und Rechenregeln

3.1.1 Benennen der Einträge

Um mit Matrizen rechnen zu können müssen wir erst ein mal die einzelnen Einträge der Matrix benennen können. Hier verfahren wir im wesentlichen wie bei Tabellen: wir benennen jeden Eintrag durch die Zeile und Spalte, in der er steht: der Eintrag in der i-ten Zeile und j-ten Spalte einer Matrix M wird M_{ij} (oder auch m_{ij}) genannt. Im Beispiel also

$$A = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix} \quad \text{oder} \quad B = \begin{pmatrix} B_{11} & B_{12} & B_{13} & B_{14} \\ B_{21} & B_{22} & B_{23} & B_{24} \end{pmatrix} \quad \text{oder} \quad C = \begin{pmatrix} C_{11} & C_{12} & C_{13} \\ C_{21} & C_{22} & C_{23} \\ C_{31} & C_{32} & C_{33} \end{pmatrix}.$$

Als Merkregel gilt also:

$M =$ gesamte Matrix , $M_{ij} =$ ein Eintrag der Matrix (jener in der i-ten Zeile und j-ten Spalte).

3.1.2 Addition und Subtraktion von Matrizen

Wenn zwei Matrizen gleich groß sind, sie also beide $(m \times n)$-Matrizen sind (mit gleichem m und n), so können sie addiert und subtrahiert werden. Das passiert dann einfach elementweise:

$$A + B = C \quad \Leftrightarrow \quad C_{ij} = A_{ij} + B_{ij} \quad \text{und} \quad A - B = D \quad \Leftrightarrow \quad D_{ij} = A_{ij} - B_{ij}. \tag{3.1}$$

Im Beispiel:

$$\begin{pmatrix} 1 & 4 \\ 2 & 9 \end{pmatrix} - \begin{pmatrix} 5 & 2 \\ 6 & 0 \end{pmatrix} = \begin{pmatrix} 1-5 & 4-2 \\ 2-6 & 9-0 \end{pmatrix} = \begin{pmatrix} -4 & 2 \\ -4 & 9 \end{pmatrix}.$$

3.1.3 Multiplikation mit Skalaren

Ist A eine Matrix und λ ein Skalar, so ist die Multiplikation von M mit λ elementweise definiert:

$$B = \lambda\, A \quad \Leftrightarrow \quad B_{ij} = \lambda\, A_{ij}. \tag{3.2}$$

Im Beispiel:

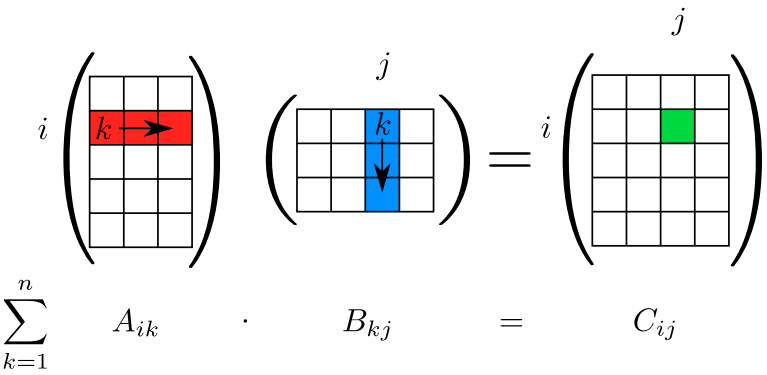
$$3 \begin{pmatrix} 1 & 4 \\ 2 & 9 \end{pmatrix} = \begin{pmatrix} 3 \cdot 1 & 3 \cdot 4 \\ 3 \cdot 2 & 3 \cdot 9 \end{pmatrix} = \begin{pmatrix} 3 & 12 \\ 6 & 27 \end{pmatrix}.$$

3.1.4 Multiplikation von Matrizen

Aus zwei Matrizen A und B kann man genau dann eine neue Matrix $C = A\,B$ durch **Matrixmultiplikation** bilden, wenn die Matrizen "passende" Dimensionen (Größen) haben. Passend bedeutet dabei dass die **Anzahl von Spalten in A gleich der Anzahl der Zeilen in B** sein muss. Mathematisch ausgedrückt bedeutet das:

Sei A eine $(m \times n)$-Matrix. Das Matrixprodukt von $A\,B$ von A mit einer Matrix B ist definiert genau dann, wenn B eine $(n \times p)$-Matrix ist. Es gilt dann:

$$C = A\,B \quad \Leftrightarrow \quad C_{ij} = \sum_{k=1}^{n} A_{ik}\, B_{kj}. \tag{3.3}$$



Da der Index i in der obigen Formel nur bei A_{ik} vorkommt sieht man, dass er die Werte $i = 1, 2, \ldots, m$ annehmen kann. Der Index j kommt nur bei B_{kj} vor, und kann also Werte $j = 1, 2, \ldots, p$ annehmen. Es gilt also:

Das Produkt einer $(m \times n)$-Matrix mit einer $(n \times p)$-Matrix ist eine $(m \times p)$-Matrix.

Achtung: Anders als bei Zahlen ist die Multiplikation von Matrizen nicht kommutativ: $A\,B \neq B\,A$ (außer in ein paar wenigen Ausnahmen).

3.1.5 Transposition

Die "**Transponierte**" einer Matrix erhält man dadurch, dass man Zeilen und Spalten vertauscht (aus der i-ten Zeile wird die i-te Spalte). Man kennzeichnet sie mit einem hochgestellten T. Es gilt also:

$$B = A^T \quad \Leftrightarrow \quad B_{ij} = A_{ji}. \tag{3.4}$$

Im Beispiel:

$$A = \begin{pmatrix} 1 & 4 \\ 2 & 9 \end{pmatrix} \Leftrightarrow B = A^T = \begin{pmatrix} 1 & 2 \\ 4 & 9 \end{pmatrix} \quad \text{oder} \quad C = \begin{pmatrix} 2 & 5 & 1 & 0 \\ 5 & 8 & 4 & 3 \\ 4 & 9 & 7 & 2 \end{pmatrix} \Leftrightarrow D = C^T = \begin{pmatrix} 2 & 5 & 4 \\ 5 & 8 & 9 \\ 1 & 4 & 7 \\ 0 & 3 & 2 \end{pmatrix}.$$

Die Transponierte einer $(m \times n)$-Matrix ist also eine $(n \times m)$-Matrix.

3.1.6 Einige wichtige Arten von Matrizen

- Eine **quadratische Matrix** hat ebensoviele Zeilen wie Spalten, und ist also eine $(n \times n)$-Matrix. Bei quadratischen Matrizen M nennt man die Einträge M_{ii} die Einträge auf der **Hauptdiagonalen**.
- Eine **Einheitsmatrix** ist eine quadratische Matrix, in der alle Einträge auf der Hauptdiagonalen 1 sind und alle weiteren Einträge 0. Einheitsmatrizen werden auch mit $\mathbb{1}$ bezeichnet, manchmal auch noch mit der Angabe der Dimension als Index. Die (3×3)-Einheitsmatrix ist zum Beispiel gegeben durch
$$\mathbb{1}_{3\times3} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$
- Eine **symmetrische Matrix** erfüllt $M = M^T$, wie zum Beispiel
$$M = \begin{pmatrix} 4 & 7 & 5 \\ 7 & 8 & 1 \\ 5 & 1 & 0 \end{pmatrix} = M^T.$$

3.2 Spur

Eine wichtige Kenngröße von quadratischen Matrizen ist ihre **Spur** (mit Symbol "Sp" oder "tr" für das englische "trace"). Diese ist definiert als die Summe der Elemente auf der Hauptdiagonalen:

$$A \text{ ist } (n \times n)\text{-Matrix} \quad \Rightarrow \quad \mathrm{Sp}\,(A) = \sum_{i=1}^{n} A_{ii}. \tag{3.5}$$

Im Beispiel:

$$A = \begin{pmatrix} 1 & 4 & 2 \\ 3 & 4 & 9 \\ 1 & 3 & 9 \end{pmatrix} \quad \Rightarrow \quad \mathrm{Sp}\,(A) = 1 + 4 + 9 = 14.$$

Ein wichtiges Anwendungsgebiet der Spur in der Physik findet sich in der Thermodynamik. Dort wird der Mittelwert einer Messung, die eine Größe $\mathcal{O}$ misst, durch $\bar{\mathcal{O}}$ bezeichnet. Man kann dann zeigen (siehe Vorlesung Thermodynamik!), dass dieser Mittelwert sich als $\bar{\mathcal{O}} = \mathrm{Sp}(\rho\,\mathcal{O})$ berechnet, wobei ρ "statistischer Operator" oder "Dichtematrix" genannt wird. Die Spur hat eine Reihe hilfreicher Eigenschaften (hier seien A, B und C $(n \times n)$-Matrizen):

- Spur von Summen von Matrizen: $\mathrm{Sp}(\lambda_1\,A + \lambda_2\,B) = \lambda_1\,\mathrm{Sp}(A) + \lambda_2\,\mathrm{Sp}(B)$.
- Spur von Transponierten: $\mathrm{Sp}(A^T) = \mathrm{Sp}(A)$.
- Spur von Produkten von Matrizen ist "zyklisch": $\mathrm{Sp}(A\,B\,C) = \mathrm{Sp}(B\,C\,A) = \mathrm{Sp}(C\,A\,B)$.

Rechenaufgabe:

Berechnen Sie folgende Spur:

$$A = \begin{pmatrix} 1 & 2 & 4 \\ 2 & 3 & 0 \end{pmatrix}, \quad B = \begin{pmatrix} 4 & 2 \\ 8 & 7 \\ 1 & 5 \end{pmatrix}, \quad C = \begin{pmatrix} 3 & 7 \\ 3 & 1 \end{pmatrix}$$

$\mathrm{Sp}(A\,B - 3\,C) =$

3.3 Inverse Matrix

Eine weitere in der Physik wichtige Eigenschaft von **quadratischen** Matrizen ist, dass manche davon eine sogenannte "Inverse" haben. Die Inverse einer $(n \times n)$-Matrix A, bezeichnet mit A^{-1}, ist so definiert, dass

$$A\,A^{-1} = \mathbb{1}_{n\times n} = A^{-1}\,A. \tag{3.6}$$

Eine solche Inverse Matrix A^{-1} existiert genau dann, wenn die Determinante der Matrix $\mathrm{Det}(A) \neq 0$ ist (Determinante: siehe Kapitel 3.4).

Die Berechnung der Inversen einer Matrix, falls sie existiert, ist im Allgemeinen eine komplizierte Aufgabe. Für (2×2)-Matrizen gilt folgende halbwegs einfache Formel:

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \quad \Leftrightarrow \quad A^{-1} = \frac{1}{ad - bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}.$$

Im allgemeinen Fall berechnet man die Inverse einer $(n \times n)$-Matrix A mit dem Gauß-Jordan-Verfahren. Dieses hat folgende Schritte:

1. Erweitere die Matrix A mit der Einheitsmatrix $\mathbb{1}_{n \times n}$:

$$A \quad \Rightarrow \quad \left(\begin{array}{ccc|ccc} A_{11} & \dots & A_{1n} & 1 & \dots & 0 \\ \vdots & & \vdots & 0 & \ddots & 0 \\ A_{n1} & \dots & A_{nn} & 0 & \dots & 1 \end{array} \right)$$

2. Führe so lange "elementare Zeilenumformungen" durch, bis die linke Seite der erweiterten Matrix zu einer Einheitsmatrix umgeformt ist. Elementare Zeilenumformungen sind:
   - Multiplizieren einer Zeile (jedes Eintrags der Zeile) mit einem Skalar.
   - Vertauschen von Zeilen.
   - Addition eines Vielfachen einer Zeile zu einer anderen (wieder Eintrag für Eintrag)
3. Die rechte Seite (rechts des vertikalen Trennungsstrichs) ist dann die Inverse A^{-1}.

Als Beispiel berechnen wir die Inverse der Matrix $A = \begin{pmatrix} 2 & 3 \\ 1 & 1 \end{pmatrix}$:

1. Erweitern der Matrix:
$$A \quad \Rightarrow \quad \left(\begin{array}{cc|cc} 2 & 3 & 1 & 0 \\ 1 & 1 & 0 & 1 \end{array} \right)$$
2. Elementare Zeilenumformungen:
   (a) Ziehe das Doppelte von Zeile 2 von Zeile 1 ab:
$$\left(\begin{array}{cc|cc} 2 & 3 & 1 & 0 \\ 1 & 1 & 0 & 1 \end{array} \right) \quad \rightarrow \quad \left(\begin{array}{cc|cc} 0 & 1 & 1 & -2 \\ 1 & 1 & 0 & 1 \end{array} \right)$$
   (b) Ziehe Zeile 1 von Zeile 2 ab:
$$\left(\begin{array}{cc|cc} 0 & 1 & 1 & -2 \\ 1 & 1 & 0 & 1 \end{array} \right) \quad \rightarrow \quad \left(\begin{array}{cc|cc} 0 & 1 & 1 & -2 \\ 1 & 0 & -1 & 3 \end{array} \right)$$
   (c) Vertausche Zeilen 1 und 2:
$$\left(\begin{array}{cc|cc} 0 & 1 & 1 & -2 \\ 1 & 0 & -1 & 3 \end{array} \right) \quad \rightarrow \quad \left(\begin{array}{cc|cc} 1 & 0 & -1 & 3 \\ 0 & 1 & 1 & -2 \end{array} \right)$$
3. Lese ab:
$$A^{-1} = \begin{pmatrix} -1 & 3 \\ 1 & -2 \end{pmatrix}$$

Man kann jetzt in der Tat überprüfen, dass $A\,A^{-1} = \mathbb{1} = A^{-1}\,A$.

3.4 Determinante

Eine weitere wichtige Kenngröße einer quadratischen Matrix M ist ihre **Determinante**, die das Symbol "Det(M)" oder "det(M)" oder "$|M|$" hat. Die Determinante einer $(n \times n)$-Matrix A ist wie folgt definiert

$$A = \begin{pmatrix} A_{11} & \dots & A_{1n} \\ \vdots & & \vdots \\ A_{n1} & \dots & A_{nn} \end{pmatrix} \quad \Rightarrow \quad \mathrm{Det}(A) = \sum_{\sigma \text{ Permutation von 1 bis } n} \mathrm{sgn}(\sigma)\, A_{1,\sigma(1)} \cdot A_{2,\sigma(2)} \cdot \ldots \cdot A_{n,\sigma(n)}. \tag{3.7}$$

Hierbei ist eine "Permutation" der Zahlen 1 bis n ein "Durcheinanderwürfeln" der Folge der Zahlen 1 bis n. Das "Vorzeichen der Permutation" mit dem mathematischen Symbol sgn(σ) ist entweder $+1$ oder -1. Es ist $+1$ wenn die Zahl der Transpositionen (Vertauschung benachbarter Zahlen), die nötig ist, um von der ursprünglichen Folge zur Permutation zu kommen, gerade ist (und -1 sonst). Also im Beispiel: Die Zahlen 1 bis 3 haben folgende Permutationen:

$123\ (\mathrm{sgn} = +1),\ 231\ (\mathrm{sgn} = +1),\ 312\ (\mathrm{sgn} = +1),\ 132\ (\mathrm{sgn} = -1),\ 321\ (\mathrm{sgn} = -1),\ 213\ (\mathrm{sgn} = -1).$

Für eine gegebene Permutation σ ist $\sigma(j)$ die j-te Zahl der Permutation. Also gilt für $\sigma = 312$, dass $\sigma(1) = 3$, $\sigma(2) = 1$ und $\sigma(3) = 2$.

3.4.1 Determinante einer (1×1)-Matrix

Die einfachste quadratische Matrix ist eine (1×1)-Matrix. Es gilt dann:

$$M_1 = (a) \quad \Rightarrow \quad \mathrm{Det}(M_1) = a. \tag{3.8}$$

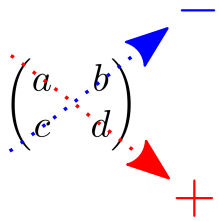
3.4.2 Determinante einer (2×2)-Matrix

Bei einer (2×2)-Matrix folgt aus der allgemeinen Definition in Gleichung (3.7):

$$M_2 = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \quad \Rightarrow \quad \mathrm{Det}(M_2) = a\,d - b\,c. \tag{3.9}$$

Graphisch kann man sich die Determinante von (2×2)-Matrixen so merken:

Determinante = + (Produkt der Diagonalelemente) - (Produkt der nicht-Diagonalelemente).



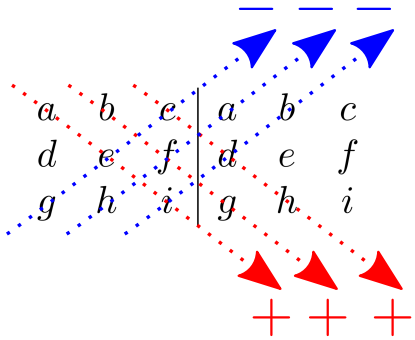
3.4.3 Determinante einer (3×3)-Matrix

Für (3×3)-Matrizen folgt aus der allgemeinen Definition in Gleichung (3.7):

$$M_3 = \begin{pmatrix} a & b & c \\ d & e & f \\ g & h & i \end{pmatrix} \quad \Rightarrow \quad \mathrm{Det}(M_3) = a\,e\,i + b\,f\,g + c\,d\,h - g\,e\,c - h\,f\,a - i\,d\,b. \tag{3.10}$$

Graphisch kann man sich die Determinante von (3×3)-Matrixen wie folgt merken. Man schreibt die Matrix zwei Mal nebeneinander und bildet dann:

Determinante = + (Produkt der "Diagonalelemente") - (Produkt der "nicht-Diagonalelemente").



3.4.4 Determinanten von $(n \times n)$-Matrizen: Entwicklungssatz

Die einfachen Regeln/Bilder zur Berechnung von Determinanten **funktionieren nur für die obigen Fälle**. Bei größeren Matrizen muss die allgemeine Formel (3.7) angewandt werden. Man kann sich das Leben aber ein bisschen leichter machen, indem man den Entwicklungssatz von Laplace nutzt (der auf obiger Formel beruht). Man kann entweder nach Zeilen oder Spalten entwicklen, und wir betrachten hier die Entwicklung nach Zeilen. Für eine $(n \times n)$-Matrix A definieren wir jetzt noch die Matrizen $\tilde{A}(ij)$, die aus der Matrix A durch das Streichen der i-ten Zeile und j-ten Spalte entsteht. Danach wählt man sich eine beliebige Zeile i aus. Der Entwicklungssatz für die Determinante besagt dann:

$$\mathrm{Det}(A) = \sum_{j=1}^{n} (-1)^{i+j} A_{ij} \, \mathrm{Det}(\tilde{A}(ij)). \tag{3.11}$$

Dieses gilt für jede beliebige Wahl der Zeile i. Analog kann man auch nach einer beliebigen Spalte j entwicklen:

$$\mathrm{Det}(A) = \sum_{i=1}^{n} (-1)^{i+j} A_{ij} \, \mathrm{Det}(\tilde{A}(ij)). \tag{3.12}$$

Auch dies gilt wieder für jede Spalte j. Die Entwicklung kann man dann iterativ so lange wiederholen, bis man z.B. nur noch Determinanten von (3×3)-Matrizen berechnen muss. Als Beispiel können wir aber auch die Determinante einer (3×3)-Matrix über den Entwicklungssatz bestimmen (hier Entwicklung nach der ersten Spalte):

$$\begin{aligned} \mathrm{Det} \begin{pmatrix} a & b & c \\ d & e & f \\ g & h & i \end{pmatrix} &= a \, \mathrm{Det} \begin{pmatrix} e & f \\ h & i \end{pmatrix} - d \, \mathrm{Det} \begin{pmatrix} b & c \\ h & i \end{pmatrix} + g \, \mathrm{Det} \begin{pmatrix} b & c \\ e & f \end{pmatrix} \\ &= a\,(e\,i - h\,f) - d\,(b\,i - h\,c) + g\,(b\,f - e\,c). \end{aligned}$$

Dies stimmt in der Tat mit Gleichung (3.10) überein.

3.4.5 Rechenregeln für Determinanten

Auch für Determinanten gibt es eine Reihe hilfreicher Rechenregeln:

- Determinante eines Produkts von Matrizen: $\mathrm{Det}(A\,B) = \mathrm{Det}(A)\,\mathrm{Det}(B)$.
- Determinante des Produkts einer $(n \times n)$-Matrix mit einem Skalar λ: $\mathrm{Det}(\lambda A) = \lambda^n\,\mathrm{Det}(A)$.
- Determinante der Transponierten: $\mathrm{Det}(A) = \mathrm{Det}(A^T)$.
- Determinante der Inversen: $\mathrm{Det}(A^{-1}) = \frac{1}{\mathrm{Det}(A)}$.
- Die Determinante einer Matrix bleibt unverändert, wenn man das Vielfache einer Spalte zu einer anderen Spalte hinzuzählt.
- Die Determinante einer Matrix bleibt unverändert, wenn man das Vielfache einer Zeile zu einer anderen Zeile hinzuzählt.
- Die Determinante einer Matrix, in der zwei Spalten (oder zwei Zeilen) Vielfache voneinander sind, ist Null. Ebenso ist die Determinante einer Matrix, in der alle Einträge in einer Spalte (oder Zeile) Null sind, gleich Null.

3.5 Rechnen mit Indizes: Summenkonvention, Kronecker-Delta und Levi-Civita-Tensor

Beim Rechnen mit Matrizen, Vektoren und Tensoren haben wir zur Definition der Rechenoperationen immer auf die einzelnen Einträge zurückgegriffen (bei einem Vektor $\vec{v}$ auf die Komponenten v_i, bei einer Matrix M auf die Einträge M_{ij}, und ebenso bei Tensoren mit Einträgen $T_{ij\ldots p}$. Obwohl viele Physiker am Anfang ihres Studiums die Tendenz haben, Vektoren und Matrizen voll und ganz auszuschreiben (dann erkennt man einfach besser, wie das komplette Objekt aussieht), ist es mit ein bisschen Übung oft schneller und einfacher, Rechnungen in Komponenten durchzuführen. Als Beispiel erinnern wir uns an die Definition des Skalarprodukts zweier n-komponentiger Vektoren $\vec{a}$ und $\vec{b}$, und dem Matrixprodukt der $(n \times k)$-Matrix A mit der $(k \times m)$-Matrix B:

$$\vec{a} \cdot \vec{b} = \sum_{i=1}^{n} a_i\, b_i \quad \text{und} \quad C_{ij} = \sum_{p=1}^{k} A_{ip}\, B_{pj}.$$

Um beim Rechnen mit Indizes Zeit zu sparen, nutzt man oft die **Einsteinsche Summenkonvention**:

Wenn Indizes auf einer Seite der Gleichung doppelt vorkommen und über alle möglichen Werte des Index summiert wird, dann kann man das Summenzeichen einfach weglassen.

In der Summenkonvention wird das Summenzeichen also einfach weggelassen, durch den doppelt auftretenden Index ist es aber implizit. Die Summation wird – beim Rechnen mit der Summenkonvention – also trotz des weggelassenen Summenzeichens ausgeführt. In unserem Beispiel können wir z. B. annehmen, dass wir wissen, wieviele Komponenten die Vektoren und Matrizen haben. Dann schreiben wir mit der Summenkonvention:

$$\vec{a} \cdot \vec{b} = a_i\, b_i \quad \text{und} \quad C_{ij} = A_{ip}\, B_{pj}.$$

Hier macht es jetzt Sinn, Indizes in zwei Klassen zu unterteilen:

- Ein **Laufindex** ist ein Index, über den summiert wird (bzw. der bei Nutzung der Summenkonvention doppelt auftritt). In unserem Beispiel wäre beim Skalarprodukt $\vec{a} \cdot \vec{b}$ i ein Laufindex, ebenso p beim Matrixprodukt.

- Ein **freier Index** ist ein Index, über den nicht summiert wird. Im Beispiel hat das Skalarprodukt keinen freien Index, bei $C_{ij} = A_{ip}\, B_{pj}$ sind aber i und j freie Indizes.

Beim Rechnen mit Indizes und der Summenkonvention sollte man auf folgendes besonders achten:

- Freie Indizes dürfen nicht umbenannt werden, Laufindizes schon.

- Wenn ein Index in einem Term doppelt auftritt, wird über ihn summiert.

- c_i hingegen impliziert keine Summe, sondern ist einfach die i-te Komponente des Vektors $\vec{c}$. Ebenso impliziert $a_i\, a_i\, a_i$ keine Summe.

Wenn wir die Summenkonvention nutzen, dürfen wir Potenzen von Termen mit Indizes nicht einfach ausführen – sonst wissen wir nicht mehr, wie wir summieren müssen! Beispiele sind:

- $\sqrt{a_k\, a_k} \neq |a_k|$: es gilt $\sqrt{a_k\, a_k} \to \sum_k \sqrt{a_k\, a_k} = \sum_k |a_k|$, wohingegen $|a_k|$ der Betrag der k-ten Komponente des Vektors $\vec{a}$ ist.

- $a_k\, a_k \neq a_k^2$: es gilt $a_k\, a_k \to \sum_k a_k\, a_k = \sum_k a_k^2$, wohingegen a_k^2 das Quadrat der k-ten Komponente des Vektors $\vec{a}$ ist.

3.5.1 Das Kronecker-Delta

Oftmals braucht man in der Physik ein mathematisches Objekt, das besagt: "Wenn genau diese Bedingung erfüllt ist, so folgt jenes". Beim Rechnen mit Indizes braucht man zum Beispiel oft ein Objekt, das einen bestimmten Index auswählt. Hierzu definieren wir das sogenannte "Kronecker-Delta" wie folgt:

$$\delta_{ij} = \begin{cases} 1 & , i = j \\ 0 & , \text{sonst} \end{cases} . \tag{3.13}$$

Man kann sich leicht davon überzeugen, dass δ_{ij} gerade die Komponenten einer Einheitsmatrix sind.

Achtung beim Rechnen mit der Summenkonvention: wenn man den Index i als Laufindex ansieht, der n Werte annehmen kann, und die Summenkovention nutzt, so gilt

$$\delta_{ii} \to \sum_{i=1}^{n} \delta_{ii} = \sum_{i=1}^{n} 1 = n.$$

3.5.2 Der Levi-Civita-Tensor / Epsilon-Tensor

Ein anderes hilfreiches Objekt beim Rechnen mit Indizes ist der sogenannte Levi-Civita-Tensor. Da dieser Tensor immer das Symbol ϵ_{ijk} hat, wird er auch Epsilon-Tensor genannt. Der Levi-Civita-Tensor hat folgende Definition:

$$\epsilon_{ijk} = \begin{cases} 1 & , ijk \text{ Permutation von } 123 \text{ mit Vorzeichen } +1 \text{ (“gerade Permutation”)} \\ -1 & , ijk \text{ Permutation von } 123 \text{ mit Vorzeichen } -1 \text{ (“ungerade Permutation”)} \\ 0 & , \text{sonst.} \end{cases} \tag{3.14}$$

Explizit gilt also

$$\epsilon_{ijk} = \begin{cases} 1 & , ijk \in \{123, 231, 312\} \\ -1 & , ijk \in \{132, 321, 213\} \\ 0 & , \text{sonst.} \end{cases} \tag{3.15}$$

Aus der Definition folgt, dass der Levi-Civita-Tensor sein Vorzeichen ändert, wenn man zwei Indizes vertauscht, also z. B. $\epsilon_{ijk} = -\epsilon_{jik}$ (man sagt auch, dass ϵ_{ijk} “total antisymmetrisch” ist).

Warum ist dieses seltsame Objekt hilfreich? In der Physik tritt es immer wieder auf, weil man es zum Berechnen von Kreuzprodukten in der Indexschreibweise braucht. Kreuzprodukte wiederum sind in der Physik wirklich sehr wichtig, zum Beispiel bei der Lorentz-Kraft, die ein Teilchen der Ladung a in einem Magnetfeld $\vec{B}$ erfährt, wenn es sich mit Geschwindigkeit $\vec{v}$ bewegt,

$$\vec{F}_{\text{Lorentz}} = q\,\vec{v} \times \vec{B}.$$

Wir erinnern uns: bei zwei Vektoren $\vec{a}$ und $\vec{b}$ gilt für das Kreuzprodukt:

$$\vec{a} = \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix} \quad \text{und} \quad \vec{b} = \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix} \quad \Rightarrow \quad \vec{c} = \vec{a} \times \vec{b} = \begin{pmatrix} a_2 b_3 - a_3 b_2 \\ a_3 b_1 - a_1 b_3 \\ a_1 b_2 - a_2 b_1 \end{pmatrix}.$$

In Komponenten und mit Einsteinscher Summenkonvention sieht das Kreuzprodukt wie folgt aus:

$$\vec{c} = \vec{a} \times \vec{b} \quad \Leftrightarrow c_i = \epsilon_{ijk}\, a_j\, b_k. \tag{3.16}$$

So gilt zum Beispiel:

$$c_1 = (+1)\, a_2\, b_3 + (-1)\, a_3\, b_2 = \epsilon_{123}\, a_2\, b_3 + \epsilon_{132}\, a_3\, b_2 = \epsilon_{1ij}\, a_i\, b_j.$$

Im letzten Schritt haben wir benutzt, dass $\epsilon_{1ij} = 0$ ist falls ij nicht die Werte $ij = 23$ oder $ij = 32$ hat. Für das Rechnen mit dem Levi-Civita-Tensor gibt es noch folgende hilfreiche Rechenregeln (alle mit Summenkonvention):

$$\epsilon_{ijk}\, \epsilon_{imn} = \delta_{jm}\, \delta_{kn} - \delta_{jn}\, \delta_{km}. \tag{3.17}$$

$$\epsilon_{ijk}\, \epsilon_{ijn} = 2\, \delta_{kn}. \tag{3.18}$$

$$\epsilon_{ijk}\, \epsilon_{ijk} = 6. \tag{3.19}$$

Das wäre im Physikerleben gut zu wissen:

- Grundlegende Matrizenrechnung beherrschen: Addition, Subtraktion, Mulitplikation mit Skalaren, Transpotition, Matrixmultiplikation.
- Wissen, was eine inverse Matrix ist.
- Die Spur einer Matrix berechnen können.
- Wissen, wie man Determinanten von (2×2)-Matrizen und (3×3)-Matrizen ausrechnen kann.
- Den Entwicklungssatz von Laplace kennen und anwenden können.
- Die Definition von Kronecker-Delta und Levi-Civita-Tensor kennen.
- Mit Indizes rechnen können (braucht Übung, wird irgendwann schön).
- Verstehen, dass Matrizen lineare Abbildungen definieren.

PS: "gut zu wissen" muss nicht heißen, dass das alles in der Klausur abgefragt wird – oder, dass sonst nichts aus diesem Kapitel in der Klausur vorkommt. Wer aber mit den Punkten in diesem Kasten nichts anfangen kann, hat wahrscheinlich in der Klausur und im weiteren Studium Probleme.

Kapitel 4

Basiswechsel, Eigenwerte und Eigenvektoren

In diesem Kapitel wenden wir uns folgendem Problem zu: oftmals ist es in der Physik einfach, die Formel für ein Problem in einem bestimmten Koordinatensystem K aufzustellen. Die Lösung der Gleichung ist aber vielleicht in einem anderen Koordinatensystem K' einfacher, nämlich in genau dem, in dem z. B. die Matrizen, die das physikalische System beschreiben, besonders einfach sind. Wie identifiziert man das "optimale" Koordinatensystem für eine gegebene Rechnung?

4.1 Matrizen als Abbildungen

Eine wichtige Erkenntnis dieses Kapitels ist, dass Matrizen Abbildungen beschreiben/definieren/sind. Um das zu verstehen betrachen wir das (Matrix-)Produkt einer $(m \times n)$-Matrix M mit einem n-komponentigen Vektors $\vec{v}$. Das Produkt $M\,\vec{v}$ ist nach den Rechenregeln des letzten Kapitels eine $(m \times 1)$-Matrix, also ein m-komponentiger Vektor $\vec{w}$:

$$M\,\vec{v} = \vec{w} \quad \text{mit} \quad w_i = M_{ij}\,v_j \tag{4.1}$$

(hier nutzen wir die Summenkonvention, ohne Summenkonvention würden wir für die i-te Komponente von $\vec{w}$ schreiben, dass $w_i = \sum_{j=1}^{n} M_{ij}\,v_j$). Gleichung (4.1) besagt, dass die Matrix M aus dem Vektor $\vec{v}$ den Vektor $\vec{w}$ macht. Anders gesagt:

Die Matrix M bildet den Vektor $\vec{v}$ auf den Vektor $\vec{w}$ ab.

Lebt der Vektor $\vec{v}$ im Vektorraum V und der Vektor $\vec{w}$ im Vektorraum W, so definiert die Matrix M folgende Abbildung $\mathcal{M}$:

$$\mathcal{M} : V \to W, \quad \vec{v} \mapsto \mathcal{M}(\vec{v}) = M\,\vec{v}. \tag{4.2}$$

Die Matrixmultiplikation ist linear: das Endergebnis hängt nur linear von den Komponenten des Eingangsvektors ab. Somit beschreiben Matrizen nur eine spezielle Klasse von Abbildungen: lineare Abbildungen. Eine wichtige Anwendung solcher linearen Abbildungen sind Drehungen eines Vektors. Betrachten wir zum Beispiel hier die Drehung eines zwei-dimensionalen Vektors $\vec{r}$ in der Ebene um einen Winkel ϕ. Wenn wir den gedrehten Vektor $\vec{w}$ nennen, so gilt:

$$\vec{w} = D_\phi\,\vec{v} \quad \text{mit} \quad D_\phi = \begin{pmatrix} \cos(\phi) & -\sin(\phi) \\ \sin(\phi) & \cos(\phi) \end{pmatrix}. \tag{4.3}$$

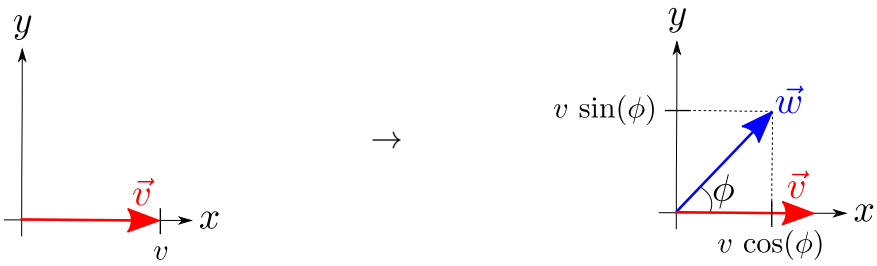
Das können wir mir folgendem Beispiel überprüfen. Wir starten mit einem Vektor $\vec{v}$ entlang der x-Achse, den wir um einen Winkel ϕ drehen:

$$\vec{v} = \begin{pmatrix} v \\ 0 \end{pmatrix} \quad \Rightarrow \quad \vec{w} = D_\phi \, \vec{v} = \begin{pmatrix} \cos(\phi) & -\sin(\phi) \\ \sin(\phi) & \cos(\phi) \end{pmatrix} \begin{pmatrix} v \\ 0 \end{pmatrix} = \begin{pmatrix} v\,\cos(\phi) \\ v\,\sin(\phi) \end{pmatrix}.$$

Da der Vektor durch die Drehung seine Länge nicht geändert hat, gilt $|\vec{v}| = |\vec{w}| = |v|$. Somit gilt (für $|v| = v > 0$)

$$\vec{w} = \begin{pmatrix} v\,\cos(\phi) \\ v\,\sin(\phi) \end{pmatrix}.$$

Mit der geometrischen Definition von Kosinus als "Ankathete durch Hypothenuse" und Sinus als "Gegenkathete durch Hypothenuse" und mit $|\vec{v}| = |\vec{w}|$ entspricht das genau dem Ergebnis der Drehung:



4.2 Eigenvektoren und Eigenwerte

Im folgenden betrachten wir wieder quadratische Matrizen. Das bisher in diesem Kapitel besprochene bedeutet für diese, dass $(n \times n)$-Matrizen Vektoren mit n Komponenten auf andere Vektoren mit n Komponenten abbilden. Im Allgemeinen sind die Eingangs- und Endvektoren $\vec{v}$ und $\vec{w}$ unterschiedlich (wie wir das bei der Drehung gesehen habe). Es kann aber sein, dass es spezielle Vektoren gibt, bei denen die Matrix den Vektor auf einen Vektor abbildet, der in die selbe Richtung zeigt wie der ursprüngliche Vektor (der Endvektor ist also Vielfaches des Anfangsvektors). Diese Vektoren sind für die Abbildung offensichtlich irgendwie speziell (wie genau ist hier noch unklar). Wir fragen also: sei M eine $(n \times n)$-Matrix – gibt es einen Vektor $\vec{v}$, der auf sich selbst oder ein Vielfaches von sich selbst abgebildet wird? Dies entspricht der Gleichung

$$M\,\vec{v} \stackrel{?}{=} \lambda\,\vec{v}, \tag{4.4}$$

wobei λ eine Zahl ist. Wenn es einen solchen Vektor gibt, so nennen wir $\vec{v}$ einen **Eigenvektor der Matrix** M, und λ den **Eigenwert zum Eigenvektor** $\vec{v}$. Je nachdem, welche Matrizen und Vektoren wir uns anschauen, muss es nicht unbedingt Eigenvektoren geben: die Drehmatrix D_ϕ, die eine Drehung der Vektoren in der Ebene um einen Winkel ϕ beschreibt, hat für $\phi \neq 0, 2\pi, 4\pi, 6\pi, \ldots$ keinen Eigenvektor (zumindest keinen, der aus reellen Zahlen besteht). Das macht Sinn, denn D_ϕ dreht jeden Vektor um einen Winkel ϕ. Dabei ändert sich offensichtlich die Richtung jedes Vektors.

4.2.1 Berechnung der Eigenwerte

Um die Eigenwerte zu berechnen, formen wir Gleichung (4.4) wie folgt um:

$$M\,\vec{v} - \lambda\,\vec{v} = (M - \lambda\,\mathbb{1}_{n\times n})\;\vec{v} = 0. \tag{4.5}$$

Man kann zeigen, dass diese Gleichung genau dann von einem Vektor $\vec{v} \neq 0$ gelöst werden kann, wenn Det$(M - \lambda \mathbb{1}_{n\times n})$=0 ist. Um die Eigenwerte einer $(n \times n)$-Matrix M zu bestimmen, betrachtet man also das sogenannte "charakteristische Polynom":

$$\text{Det}\,(M - \lambda\,\mathbb{1}_{n\times n}) = 0. \tag{4.6}$$

Die linke Seite dieser Gleichung ist ein Polynom vom Grad n in λ. Jede Lösung λ der Gleichung ist ein Eigenwert der Matrix M zu einem Eigenvektor.

Bei einer $(n \times n)$-Matrix gibt es im Allgemeinen n Eigenwerte und n Eigenvektoren – zumindest bei den sogenannten "diagonalisierbaren" Matrizen (siehe Kapitel 4.5), auf die wir uns hier beschränken wollen, denn diese sind in der Physik besonders wichtig. Wenn ein bestimmter Wert von λ eine Nullstelle der Vielfachheit m (auch: der Ordnung m, bzw. der Multiplizität m) ist, so gibt es m Eigenvektoren mit dem gleichen Eigenwert λ (auch für dieses Statement beschränken wir uns, ganz genau genommen, auf diagonalisierbare Matrizen). In unserem obigen Rechenbeispiel ist zum Beispiel für $a = b = 3$, $c = 1$ der Wert $\lambda = 3$ eine Nullstelle zweiter Ordnung des charakteristischen Polynoms, und es gibt zwei verschiedene Eigenwerte, die beide den Eigenwert 3 haben.

4.2.2 Berechnung der Eigenvektoren

Ein Eigenvektor $\vec{v}$ zum Eigenwert λ löst laut Defintion die Eigenwert-Gleichung

$$(M - \lambda\,\mathbb{1})\;\vec{v} = 0.$$

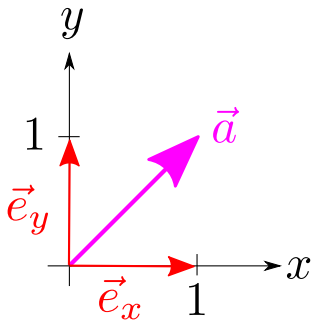
Wenn wir diese Geichung in Komponenten schreiben, erhalten wir ein lineares Gleichungssystem für die Komponenten des Vektors $\vec{v}$: der Vergleich von Gleichungen (B.1) und (4.7a) zeigt, dass wir den Vektor $\vec{v}$ wie in Anhang B beschrieben finden können, wenn wir $\vec{x} = \vec{v}$, $\vec{b} = 0$ und $A = M - \lambda\,\mathbb{1}_{n\times n}$ setzen. Beachte: dieses Gleichungssystem hat unendlich viele Lösungen. Denn erfüllt $\vec{v}$ die Gleichung $M\,\vec{v} = \lambda\,\vec{v}$, so erfüllt auch $\alpha\,\vec{v}$ die Eigenwert-Gleichung zum selben Eigenwert (wobei α irgendeine Zahl ist):

$$M\,\vec{v} = \lambda\,\vec{v} \quad \Rightarrow \quad M\,(\alpha\,\vec{v}) = \alpha\,M\,\vec{v} = \alpha\,\lambda\,\vec{v} = \lambda\,(\alpha\,\vec{v}).$$

In der Praxis versucht man aber oft, die Eigenvektoren zu raten (das geht meistens schneller).

4.3 Basiswechsel und Vektoren

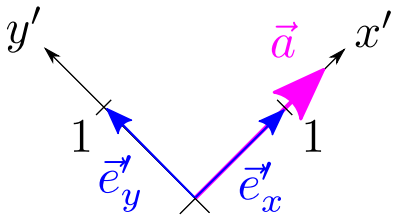
In Kapitel 1.1.4 haben wir gesehen, dass die Darstellung eines Vektors von der gewählten Basis abhängt. In diesem Kapitel haben wir implizit bisher immer eine feste Basis genutzt und konnten so z. B. in Kapitel 4.1 die Vektoren $\vec{v}$ und $\vec{w}$ in Komponenten darstellen. Jetzt wollen wir noch das Konzept des **Basiswechsels** einführen. In Kapitel 1.1.4 haben wir als Beispiel einen Vektor $\vec{a}$ betrachtet, der "schräg" im Standard-Koordinatensystem K mit der Basis $\mathcal{B} = \{\vec{e}_x, \vec{e}_y\}$ liegt ($\vec{e}_{x(y)}$ ist der Einheitsvektor in $x(y)$-Richtung):



Im Koordinatensystem K hat der Vektor die Darstellung

$$\vec{a}_K = \begin{pmatrix} 1 \\ 1 \end{pmatrix}.$$

(Hier haben wir den Index K, der das Koordinatensystem spezifiziert, zur Klarheit wieder explizit angegeben). Wir haben weiterhin ein gedrehtes Koordinatensystem K' mit der Basis $\mathcal{B}' = \{\vec{e}_x', \vec{e}_y'\}$ betrachtet, in dem die x'-Achse genau parallel zu $\vec{a}$ liegt:



Im Koordinatensystem K' hat der selbe Vektor $\vec{a}$ die Darstellung (siehe auch Gleichung (1.4))

$$\vec{a}_{K'} = \begin{pmatrix} \sqrt{2} \\ 0 \end{pmatrix}.$$

Offensichtlich gelangen wir von den Basisvektoren der ursprünglichen Basis $\mathcal{B}$ zu den Basisvektoren der Basis $\mathcal{B}'$ durch eine Drehung, im konkreten Fall durch eine Drehung um einen Winkel $\phi = \pi/4$ $(= 45°)$. Die entsprechende Drehmatrix ist

$$D_{\pi/2} = \begin{pmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix}.$$

Die **Darstellung der neuen Basisvektoren in der alten Basis** lässt sich also aus der **Darstellung der alten Basisvektoren in der alten Basis** wie folgt gewinnen:

$\vec{e}_x{}'$ hat in alter Basis die Darstellung $\quad \vec{e}_{x,K}{}' = D_{\pi/2} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{pmatrix}$, (4.7a)

$\vec{e}_y{}'$ hat in alter Basis die Darstellung $\quad \vec{e}_{y,K}{}' = D_{\pi/2} \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{pmatrix}$. (4.7b)

Das Konzept des Basiswechsels bei einer Drehung lässt sich auf beliebige Basiswechsel erweitern. Im folgenden betrachten wir nicht mehr nur zweidimensionale Vektoren, sondern Vektoren im $\mathbb{R}^n$ (also n-dimensionale Vektoren aus reellen Zahlen). Der $\mathbb{R}^n$ hat die **Orthonormalbasis** $\mathcal{B} = \{\vec{e}_1, \ldots, \vec{e}_n\}$. Die Darstellung des Einheitsvektors $\vec{e}_i$ in der alten Basis sind Spaltenvektoren mit einer 1 in der i-ten Zeile (und Nullen sonst), z. B.

$$\vec{e}_{1,K} = \begin{pmatrix} 1 \\ 0 \\ 0 \\ \vdots \end{pmatrix}.$$

Wir wollen jetzt einen Basiswechsel zu einer neuen **Orthonormalbasis** $\mathcal{B}' = \{\vec{e}_1', \ldots\}$ machen. In der alten Basis soll die **Darstellung der neuen Basisvektoren in der alten Basis** mittels einer Abbildungsmatrix U mit Einträgen (konkreten Zahlen) U_{ij} aus der **Darstellung der alten Basisvektoren in der alten Basis** hervorgehen:

$$\vec{e}_{i,K}{}' = U\,\vec{e}_{i,K} \quad \text{bzw.} \quad \begin{pmatrix} e'_{i,1} \\ e'_{i,2} \\ \vdots \end{pmatrix} = U \begin{pmatrix} e_{i,1} \\ e_{i,2} \\ \vdots \end{pmatrix} \quad \text{für alle Basisvektoren } i. \tag{4.8}$$

Hier ist $e'_{i,j}$ die j-te Komponente der Darstellung von $\vec{e}'_i$ in der **alten** Basis $\mathcal{B}$, wohingegen $e_{i,j} = \delta_{ij}$ die j-te Komponente der Darstellung von $\vec{e}_i$ in der **alten** Basis $\mathcal{B}$ ist. Hieraus folgt, dass die Matrix U einfach durch das spaltenweise nebeneinander-schreiben der Darstellungen $\vec{e}_{i,K}{}'$ der neuen Basisvektoren in der alten Basis entsteht:

$$U = \left(\vec{e}_{1,K}{}' \quad \vec{e}_{2,K}{}' \quad \dots \quad \vec{e}_{n,K}{}'\right). \tag{4.9}$$

Im Fall der Drehung ist die Transformationsmatrix U also nach Gleichungen (4.7) durch

$$U = \left(\vec{e}_{x,K}{}' \quad \vec{e}_{y,K}{}'\right) = \begin{pmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix} \tag{4.10}$$

gegeben. Die Darstellung eines Vektors $\vec{v}$ in der ursprünglichen Basis ist ein Spaltenvektor mit Einträgen v_i, die durch die Zerlegung des Vektors in die Basisvektoren definiert ist:

$$\vec{v} = \sum_j v_j\,\vec{e}_j \quad \rightarrow \quad v_j = \vec{v}\cdot\vec{e}_j. \tag{4.11}$$

In der **neuen** Basis ist die Darstellung des Vektors $\vec{v}$ ein Spaltenvektor mit Einträgen v'_i. Sie sind definiert durch

$$\vec{v} = \sum_i v'_i\,\vec{e}'_i \quad \rightarrow \quad v'_i = \vec{v}\cdot\vec{e}'_i. \tag{4.12}$$

Dargestellt in der alten Basis erhalten wir daher, dass

$$v'_i = \vec{v}\cdot\vec{e}'_i = \sum_j v_j \begin{pmatrix} e_{j,1} \\ e_{j,2} \\ \vdots \end{pmatrix} \cdot U \begin{pmatrix} e_{i,1} \\ e_{i,2} \\ \vdots \end{pmatrix} \tag{4.13}$$

Da nur die i-te Komponente von $\vec{e}_i$ eins ist und alle anderen Komponenten Null sind, $e_{i,j} = \delta_{ij}$, gilt also:

$$v'_i = \sum_j v_j\,U_{ji} = \sum_j v_j\,\left(U^T\right)_{ij} = \sum_j \left(U^T\right)_{ij}\,v_j. \tag{4.14}$$

Wir finden also, dass der Spaltenvektor der Darstellung von $\vec{v}$ im neuen Koordinatensystem K / in der neuen Basis $\mathcal{B}'$ (mit Komponenten v'_i) und der Spaltenvektor der Darstellung von $\vec{v}$ im alten Koordinatensystem K / in der alten Basis $\mathcal{B}$ (mit Komponenten v_i) wie folgt zusammenhängen:

$$\vec{v}_{K'} = U^T\,\vec{v}_K \quad \text{bzw.} \quad \begin{pmatrix} v'_1 \\ v'_2 \\ \vdots \end{pmatrix} = U^T \begin{pmatrix} v_1 \\ v_2 \\ \vdots \end{pmatrix} \tag{4.15}$$

Man kann weiterhin zeigen (siehe Übung), dass $U^T = U^{-1}$ bei Basiswechseln zwischen Orthonormalbasen des $\mathbb{R}^n$ ist.

In Summe ergibt sich also folgendes **Kochrezept**, um von der Darstellung $\vec{v}_K$ eines Vektors $\vec{v} \in \mathbb{R}^n$ in einem Koordinatensystem K mit Orthonormalbasis $\mathcal{B} = \{\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_n\}$ zur Darstellung $\vec{v}_{K'}$ dieses Vektors in einem Koordinatensystem K' mit Orthonormalbasis $\mathcal{B}' = \{\vec{e}'_1, \vec{e}'_2, \ldots, \vec{e}'_n\}$ zu bekommen:

1. Finden Sie die Darstellung $\vec{e}'_{i,K}$ der neuen Basisvektoren in den alten Koordinaten.
2. Stellen Sie die Transformationsmatrix U wie in Gleichung (4.9) gegeben als auf:
$$U = \begin{pmatrix} \vec{e}'_{1,K} & \vec{e}'_{2,K} & \ldots & \vec{e}'_{n,K} \end{pmatrix}.$$
3. Die Darstellung eines Vektors in den neuen Koordinaten folgt dann gemäß Gleichung (4.15):
$$\begin{pmatrix} v'_1 \\ v'_2 \\ \vdots \end{pmatrix} = U^{-1} \begin{pmatrix} v_1 \\ v_2 \\ \vdots \end{pmatrix}$$
wobei bei einem Wechsel zwischen Orthonormalbasen des $\mathbb{R}^n$ gilt, dass $U^T = U^{-1}$.

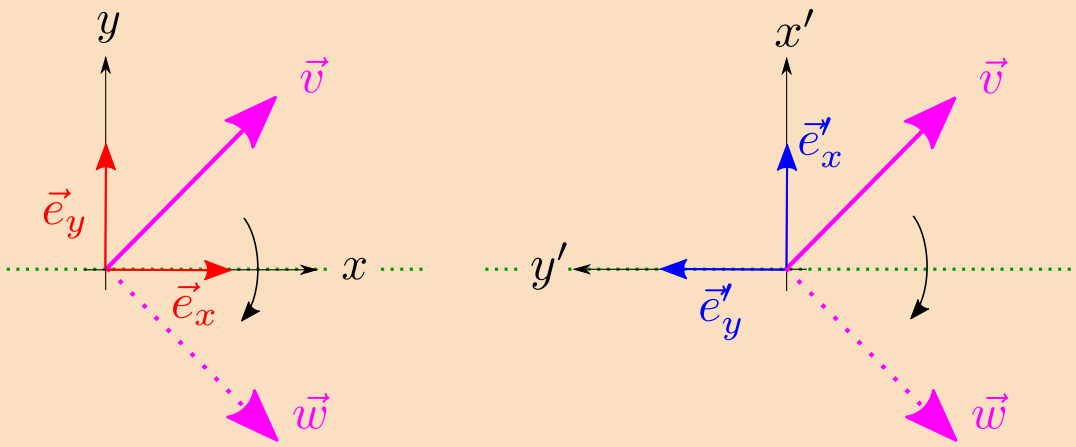
4.4 Basisabhängigkeit der Darstellung von linearen Abbildungen als Matrizen

Wir kommen jetzt wieder zu Matrizen zurück. Eine Matrix M haben wir als Abbildung $\mathcal{M}$ eines n-dimensionalen reellen Vektors auf einen anderen Vektor verstanden:

$$\mathcal{M} : \mathbb{R}^n \to \mathbb{R}^n, \quad \vec{v} \mapsto \mathcal{M}(\vec{v}). \tag{4.16}$$

Bei Vektoren haben wir gesehen, dass die **Darstellung des Vektors**, also die Zahlen im Spaltenvektor, von der gewählten Basis abhängen. Wir zeigen nun, dass das gleiche auch für lineare Abbildungen gilt: **die Darstellung der Abbildung, also die Zahlen in der Matrix, hängen von der gewählten Basis ab**.

Mitmachrechnung:



Bestimmen Sie in den Koordinatensystemen K (linkes Bild) und K' (rechtes Bild) explizit die Matrizen M, die eine Spiegelung an der Horizontalen implementieren, für die also für einen beliebigen Vektor $\vec{v}$ gilt, dass die Darstellungen des Bilds $\vec{w}$ des Vektor sich wie folgt ergibt:

$$\vec{w}_K = M\,\vec{v}_K \quad \text{und} \quad \vec{w}_{K'} = M'\,\vec{v}_{K'}.$$

Um das zu verstehen, betrachten wir zunächst eine **Orthonormalbasis** $\mathcal{B} = \{\vec{e}_1, \ldots, \vec{e}_n\}$ des $\mathbb{R}^n$, die zu einem urpsrünglichen Koordinatensystem K korrespondiert. Wir betrachten nun die lineare Abbildung $\mathcal{M}$. Die Abbildung $\mathcal{M}$ soll aus dem Vektor $\vec{v}$ den Vektor $\mathcal{M}(\vec{v}) = \vec{w}$ machen. In der Basis $\mathcal{B}$ ist die Darstellung des Vektors $\vec{v}$ durch Komponenten v_i gegeben, die Darstellung des Vektors $\vec{w}$ durch Komponenten w_i. In dieser Basis können wir nun die Abbildung durch eine konkrete Matrix M mit Einträgen M_{ij} darstellen, sodass

$$\begin{pmatrix} w_1 \\ w_2 \\ \vdots \end{pmatrix} = M \begin{pmatrix} v_1 \\ v_2 \\ \vdots \end{pmatrix} \quad \Leftrightarrow \quad w_j = \sum_k M_{jk}\, v_k. \tag{4.17}$$

Nun betrachten wir eine **zweite Orthonormalbasis** $\mathcal{B}' = \{\vec{e}'_1, \ldots, \vec{e}'_n\}$. In der alten Basis soll die **Darstellung der neuen Basisvektoren in der alten Basis** mittels eines Abbildungsmatrix U aus der **Darstellung der alten Basisvektoren in der alten Basis** hervorgehen:

$$\begin{pmatrix} e'_{i,1} \\ e'_{i,2} \\ \vdots \end{pmatrix} = U \begin{pmatrix} e_{i,1} \\ e_{i,2} \\ \vdots \end{pmatrix} \quad \text{für alle Basisvektoren } i. \tag{4.18}$$

Diese Matrix ist natürlich, wie schon in Kapitel 4.3 diskutiert, gegeben durch

$$U = \begin{pmatrix} \vec{e}'_{1,K} & \vec{e}'_{2,K} & \ldots & \vec{e}'_{n,K} \end{pmatrix}.$$

Dann gilt laut Gleichung (4.15), dass die Komponenten der Darstellung eines beliebigen Vektors $\vec{v}$ in der neuen und alten Basis wie folgt zusammenhängen:

$$\begin{pmatrix} v'_1 \\ v'_2 \\ \vdots \end{pmatrix} = U^T \begin{pmatrix} v_1 \\ v_2 \\ \vdots \end{pmatrix} \tag{4.19}$$

Für den Vektor $\vec{w}$, den wir aus $\vec{v}$ durch die Abbildung $\mathcal{M}$ erhalten, bedeutet das also folgenden Wechsel der Darstellung

$$w'_i = \sum_j \left(U^T\right)_{ij} w_j = \sum_{j,k} \left(U^T\right)_{ij} M_{jk} v_k \tag{4.20}$$

Zuletzt invertieren wir noch Gleichung (4.19) und nutzen $U^T = U^{-1}$, um die Darstellung von $\vec{v}$ in der neuen Basis zu bestimmen. Wir erhalten:

$$w'_i = \sum_{j,k,l} \left(U^T\right)_{ij} M_{jk} U_{kl} v'_l. \tag{4.21}$$

Wenn wir nun die neue Matrix

$$M' = U^T M U \tag{4.22}$$

definieren, folgt, dass

$$w'_i = \sum_l M'_{il} v'_l. \tag{4.23}$$

Was bedeutet das für die Abbildung $\mathcal{M}$? Wir haben in der ursprünglichen Basis $\mathcal{B}$ gefunden, dass sich die Komponenten der Darstellung eines Vektors wir folgt ändern:

$$w_i = \sum_j M_{ij} v_j \quad \text{beziehungsweise} \quad \begin{pmatrix} w_1 \\ w_2 \\ \vdots \end{pmatrix} = M \begin{pmatrix} v_1 \\ v_2 \\ \vdots \end{pmatrix}, \tag{4.24}$$

wobei M eben die Matrix mit Zahlen M_{ij} ist. In der neuen Basis gilt:

$$w'_i = \sum_j M'_{ij} v'_j \quad \text{beziehungsweise} \quad \begin{pmatrix} w'_1 \\ w'_2 \\ \vdots \end{pmatrix} = M' \begin{pmatrix} v'_1 \\ v'_2 \\ \vdots \end{pmatrix}. \tag{4.25}$$

Die Matrix mit Zahlen M_{ij} ist also die Darstellung der Abbildung $\mathcal{M}$ in der Basis $\mathcal{B}$, die Matrix $M' = U^T M U$ mit Zahlen M'_{ij} ist die Darstellung der selben Abbildung $\mathcal{M}$ in der Basis $\mathcal{B}'$. Zuletzt erinnern wir noch daran, dass bei Wechseln zwischen Orthonormalbasen des $\mathbb{R}^n$ die Transformationsmatrix $U^T = U^{-1}$ erfüllt (siehe Übung).

Als **Kochrezept** gilt also:

Transformieren sich die Darstellungen von Vektoren gemäß

$$\begin{pmatrix} v_1' \\ v_2' \\ \vdots \end{pmatrix} = U^{-1} \begin{pmatrix} v_1 \\ v_2 \\ \vdots \end{pmatrix}$$

so transformiert sich die Darstellung einer Abbildung $\mathcal{M}$ durch Matrizen wie folgt:

$$M' = U^{-1}\, M\, U.$$

4.5 Diagonalisierung von Matrizen

Man kann sich nun folgende Frage stellen: gibt es für eine gegebene Abbildung $\mathcal{M}$ vom $\mathbb{R}^n$ in den $\mathbb{R}^n$ eine Basis, in der die Darstellung von $\mathcal{M}$ besonders einfach ist? Hierfür müssen wir zuerst ein Mal festlegen, was "besonders einfach" heißen soll – offensichtlich soll das Rechnen mit der Abbildung besonders einfach werden. Wir entscheiden uns dafür, dass eine diagonale Matrix eine besonders einfache Darstellung der Abbildung ist (in der Tat ist das Rechnen mit diagonalen Matrizen besonders einfach!).

Wir starten damit, dass wir die Darstellung von $\mathcal{M}$ als eine Matrix M mit konkreten Zahlen in einem Ausgangs-Koordinatensystem K mit Ausgangs-Orthonormalbasis $\mathcal{B}$ kennen. Wir wollen jetzt ein neues Koordinatensystem K' mit der neuen Orthonormalbasis $\mathcal{B}'$ finden, in dem die Darstellung von $\mathcal{M}$ durch eine diagonale Matrix gegeben ist, also einer Matrix die nur Einträge auf der Hauptdiagonalen hat:

$$M' = \begin{pmatrix} M'_{11} & 0 & 0 & \dots \\ 0 & M'_{22} & 0 & \dots \\ \vdots & & \ddots & \\ & & & M'_{nn} \end{pmatrix} \tag{4.26}$$

Laut Gleichung (4.22) gilt (für Matrizen mit reellen Einträgen), dass der Basiswechsel mittels einer Transformationsmatrix U wie folgt implementiert würde (diese werden wir gleich noch genauer spezifizieren):

$$M' = U^{-1}\, M\, U.$$

Falls es eine Matrix U gibt, sodass $M' = U^{-1}\, M\, U$ wirklich diagonal ist, so nennt man die Matrix M **diagonalisierbar**. Das ist leider nicht immer der Fall. Eine $(n \times n)$-Matrix M ist aber zum Beispiel diagonalisierbar, wenn

- sie n unterschiedliche Eigenwerte hat (Achtung: wenn wir in den reellen Matrizen bleiben wollen, müssen diese n unterschiedlichen Eigenwerte alle reell sein!),
- sie nur reelle Einträge hat und symmetrisch ist (also $M = M^T$).

Außerdem ist fast jede Matrix mit komplexen Einträgen diagonalisierbar (in den komplexen Zahlen, siehe Kapitel 10). Betrachten wir im Folgenden nur reelle, symmetrische Matrizen. Man kann dann

zeigen, dass die Transformationsmatrix durch die **Eigenvektoren** definiert wird. Genauer gesagt findet man die Eigenvektoren $\vec{v}_i$ zur Matrix M. Diese erfüllen per Definition die Eigenwertgleichung

$$M\,\vec{v}_i = \lambda_i\,\vec{v}_i. \tag{4.27}$$

Man kann die Eigenvektoren so wählen, dass die $\vec{v}_i$ eine Orthonormalbasis bilden, dass also gilt $\vec{v}_i \cdot \vec{v}_j = \delta_{ij}$ (unter anderem bedeutet das, dass man normierte Eigenvektoren nutzt!). Wir wählen nun die neue Basis als die Basis, die durch die Eigenvektoren aufgespannt wird. Die Spalten der Transformationsmatrix sind dann einfach durch die Darstellung der Basisvektoren in der ursprünglichen Basis $\mathcal{B}$ (im ursprünglichen Koordinatensystem K) gegeben:

$$U = \begin{pmatrix} \vec{v}_{1,K} & \vec{v}_{2,K} & \dots & \vec{v}_{n,K} \end{pmatrix}. \tag{4.28}$$

(Hier gibt der Index das Koordinatensystem / die Basis an, in der der Vektor dargestellt ist.) Diese Basis nennt man auch die **Eigenbasis** der Matrix.

Zur Diagonalisierung kann man dann folgendes **Kochrezept** anwenden:

1. Bestimmen Sie die Eigenwerte.
2. Bestimmen Sie die Eigenvektoren.
3. Bilden Sie die Transformationsmatrix U wie in Gleichung (4.28) angegeben durch

   $$U = \begin{pmatrix} \vec{v}_{1,K} & \vec{v}_{2,K} & \dots & \vec{v}_{n,K} \end{pmatrix},$$

   wobei $\vec{v}_{i,K}$ die Darstellungen der normierten Eigenvektoren (also der normierten neuen Basisvektoren) in der ursprünglichen Basis sind.
4. Diagonalisieren Sie die Matrix wie in Gleichung (4.22) gegeben via

   $$M' = U^{-1}\,M\,U.$$

Man kann zeigen, dass das Ergebnis der Diagonalisierung die Matrix

$$M' = \begin{pmatrix} \lambda_1 & 0 & 0 & \dots \\ 0 & \lambda_2 & 0 & \dots \\ \vdots & & \ddots & 0 \\ 0 & & \dots & \lambda_n \end{pmatrix} \tag{4.29}$$

ist. Da man in der Physik oft mit Matrizen rechnet, ist es in der Praxis extrem hilfreich, Matrizen diagonalisieren zu können.

Das wäre im Physikerleben gut zu wissen:

- Lineare Gleichungssysteme als Matrizen schreiben und wissen, wie man sie lösen kann.
- Eigenwerte einer Matrix berechnen können.
- Eigenvektoren einer Matrix berechnen können (mindestens bei einfachen Matrizen).
- Verstehen, dass die Darstellung von Abbildungen mit Matrizen basisabhängig ist.
- Wissen, wie man einen Basiswechsel für die Darstellung von Vektoren und Matrizen durchführt.
- Wissen, wie man eine Matrix diagonalisiert.

PS: "gut zu wissen" muss nicht heißen, dass das alles in der Klausur abgefragt wird – oder, dass sonst nichts aus diesem Kapitel in der Klausur vorkommt. Wer aber mit den Punkten in diesem Kasten nichts anfangen kann, hat wahrscheinlich in der Klausur und im weiteren Studium Probleme.

Kapitel 5

Differentialrechnung

Im folgenden Kapitel beschäftigen wir uns mit Ableitungen. Das Grundkonzept sollte aus der Schule bekannt sein: die Ableitung einer Funktion an einer Stelle x entspricht der Steigung der Funktion an dieser Stelle. In der Physik spielen Ableitungen eine große Rolle:

- Ableitung als Maß für die Änderung einer Größe: wie ändert sich zum Beispiel der Ort eines Autos als Funktion der Zeit? Wie ändert sich der Druck einen Gases, wenn man seine Temperatur ändert?
- Zur Suche von Extremal-Werten: was muss man tun, damit z. B. die Energie in einem System minimal wird? Anwendung in der analytischen Mechanik: Lagrange-Gleichungen zur Beschreibung der Bewegung eines klassischen Körpers.
- Linearisierung: oft muss man eine gegebene Funktion $f(x)$ nicht vollständig kennen, sondern nur in der Nähe eines Punktes $x = x_0$. In diesem Fall kann man eine sogenannte Taylor-Entwicklung nutzen, welche auf Ableitungen beruht.

5.1 Grundbegriffe der Differentialrechnung

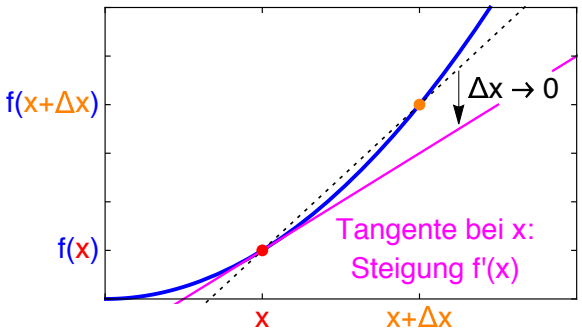
Gegeben sei eine Funktion

$$f : \mathbb{R} \to \mathbb{R}, \quad x \mapsto f(x).$$

Die **Ableitung der Funktion** $f(x)$ **am Punkt** x_0 bezeichnen wir mit Symbol $f'(x)$ oder $\frac{df}{dx}$. Sie ist definiert als

$$f'(x) = \frac{df(x)}{dx} = \lim_{\Delta x \to 0} \frac{f(x + \Delta x) - f(x)}{\Delta x} \tag{5.1}$$

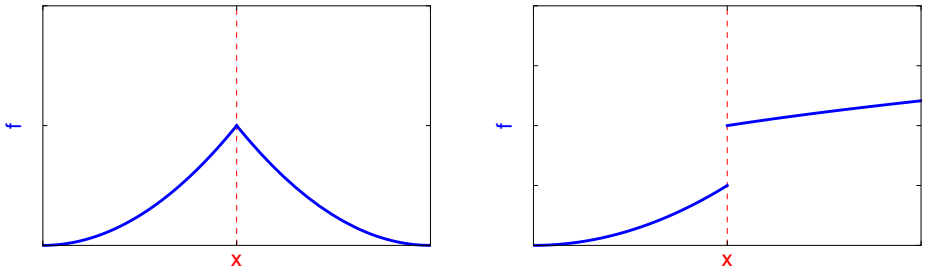
Geometrisch entspricht $f'(x)$ der Steigung der Tangente von $f(x)$ am Punkt x:



Wenn man den Grenzwert von Gleichung (5.1) eindeutig bilden kann (und er kleiner als Unendlich ist), so heißt die Funktion **im Punkt x differenzierbar**. Das muss aber nicht immer der Fall sein: zum Beispiel könnte die Funktion an der Stelle x einen Sprung oder Knick haben. In diesen Fällen gilt, dass die links- und rechtsseitigen Grenzwerte verschieden sind, und sich somit keine eindeutige Tangente definieren lässt,

$$\lim_{\Delta x \to 0^+} \frac{f(x+\Delta x) - f(x)}{\Delta x} \neq \lim_{\Delta x \to 0^-} \frac{f(x+\Delta x) - f(x)}{\Delta x} \tag{5.2}$$

(Hier bezeichnet $\lim_{\Delta x \to 0^+}$ den Grenzwert, in dem Δx von den positiven Zahlen kommend nach Null geht, wobei $\lim_{\Delta x \to 0^-}$ den Grenzwert bezeichnet, bei dem Δx von den negativen Zahlen kommend nach Null geht.)



Die Funktion ist dann nicht differenzierbar im entsprechenden Punkt x.

Man definiert weiterhin das **Differenzial** df einer Funktion als

$$df(x) = f'(x)\,dx, \tag{5.3}$$

wobei $df(x)$ die Änderung des Funktionswerts der Funktion f angibt, die sich aus einer (infinitesimal kleinen) Änderung dx des Variablenwerts ergibt.

Höhere Ableitungen von Funktionen werden dadurch gebildet, dass man die Ableitungen der Ableitungen bildet. Die zweite Ableitung der Funktion f ist zum Beispiel definiert als

$$f''(x) = \frac{df'(x)}{dx} = \frac{d^2 f(x)}{dx^2}. \tag{5.4}$$

Für die n-te Ableitung schreiben wir

$$f^{(n)}(x) = \frac{d^n f(x)}{dx^n}. \tag{5.5}$$

In der Physik spielen **zeitliche Ableitungen** eine besonders große Rolle: sie beschreiben, wie sich Dinge mit der Zeit ändern. Die zeitliche Ableitung des Orts ist zum Beispiel die Geschwindigkeit. Da zeitliche Ableitungen besonders wichtig sind, werden sie oft mit einem besonderen Symbol gekennzeichnet: einem Punkt über der Funktion. Wenn wir die Zeit mit der Variablen t bezeichnen, gilt folgende Notationskonvention:

$$f'(t) = \frac{df(t)}{dt} = \dot{f}(t). \tag{5.6}$$

5.2 Ableitungsregeln

Beim Ableiten gelten folgende wichtige Regeln (im Folgenden sind f und g Funktionen von $\mathbb{R}$ nach $\mathbb{R}$, α und β sind konstante reelle Zahlen):

1. Linearität:
$$\frac{d}{dx}\left(\alpha\, f(x) + \beta\, g(x)\right) = \alpha\, \frac{df(x)}{dx} + \beta\, \frac{dg(x)}{dx}. \tag{5.7}$$

2. Produktregel:
$$\frac{d}{dx}\left(f(x)\, g(x)\right) = g(x)\, \frac{df(x)}{dx} + f(x)\, \frac{dg(x)}{dx}. \tag{5.8}$$

3. Quotientenregel:
$$\frac{d}{dx}\, \frac{f(x)}{g(x)} = \frac{g(x)\, \frac{df(x)}{dx} - f(x)\, \frac{dg(x)}{dx}}{g(x)^2} = \frac{g(x)\, f'(x) - f(x)\, g'(x)}{g(x)^2}. \tag{5.9}$$

4. Kettenregel:
$$\frac{d}{dx}\, f\left(g(x)\right) = \frac{df(g)}{dg}\, \frac{dg(x)}{dx}. \tag{5.10}$$

5. Ableitung der Umkehrfunktion:
$$\frac{df^{-1}(x)}{dx} = \left.\frac{1}{\frac{df(y)}{dy}}\right|_{y=f^{-1}(x)}. \tag{5.11}$$

6. Regel von l'Hôpital (auch Regel von l'Hospital): es kann vorkommen, dass man einen Grenzwert der Form
$$\lim_{x\to x_0} \frac{f(x)}{g(x)}$$
betrachtet, bei dem entweder beide Grenzwerte $\lim_{x\to x_0} f(x) = 0$ und $\lim_{x\to x_0} g(x) = 0$ sind, oder beide Grenzwerte $\lim_{x\to x_0} f(x) = \infty$ und $\lim_{x\to x_0} g(x) = \infty$ sind. Man untersucht dann also Ausdücke der Form $\frac{0}{0}$ oder $\frac{\infty}{\infty}$. Um solche Ausdrücke weiter zu bestimmen, nutzt man folgende Regel:

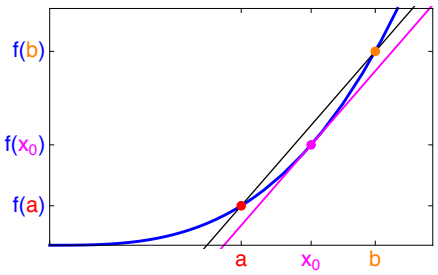
   Falls der Grenzwert $\lim_{x\to x_0} \frac{f'(x)}{g'(x)}$ existiert, so ist er gleich dem Grenzwert $\lim_{x\to x_0} \frac{f(x)}{g(x)}$.

5.3 Mittelwertsatz der Differentialrechnung

Ein wichtiger Satz der Differentialrechung ist der sogenannte Mittelwertsatz. Bildlich gesprochen sagt er folgendes aus:

- Die Ableitung einer Funktion $f'(x_0)$ an der Stelle x_0 ist gleich der Steigung der Tangente in diesem Punkt.

- Wenn man die Tangente in die richtige Richtung verschiebt, so ergeben sich (mindestens) zwei Schnittpunkte mit der ursprünglichen Funktion $f(x)$. (Die Schnittpunkte können wir z. B. mit $x = a$ und $x = b$ bezeichnen.)



In Formeln besagt der Mittelwertsatz:

Sei $f : [a, b] \to \mathbb{R}$ eine Funktion, die auf dem abgeschlossenen Intervall $[a, b]$ (mit $a < b$) definiert und stetig ist. Außerdem sei f im offenen Interval (a, b) differenzierbar. Dann gibt es mindestens ein $x_0 \in (a, b)$, so dass

$$f'(x_0) = \frac{df(x_0)}{dx} = \frac{f(b) - f(a)}{b - a}. \tag{5.12}$$

5.4 Ableitungen einiger wichtiger Funktionen

Folgende Ableitungen von elementaren Funktionen kommen in der Physik häufiger vor:

$\frac{d}{dx} x^y = y\, x^{y-1}$ $(y \in \mathbb{R})$	$\frac{d}{dx} e^x = e^x$	$\frac{d}{dx} \ln(x) = \frac{1}{x}$
$\frac{d}{dx} \sin(x) = \cos(x)$	$\frac{d}{dx} \cos(x) = -\sin(x)$	$\frac{d}{dx} \tan(x) = \frac{1}{\cos(x)^2}$
$\frac{d}{dx} \arcsin(x) = \frac{1}{\sqrt{1-x^2}}$	$\frac{d}{dx} \arccos(x) = -\frac{1}{\sqrt{1-x^2}}$	$\frac{d}{dx} \arctan(x) = \frac{1}{1+x^2}$
$\frac{d}{dx} \sinh(x) = \cosh(x)$	$\frac{d}{dx} \cosh(x) = \sinh(x)$	$\frac{d}{dx} \tanh(x) = \frac{1}{\cosh(x)^2}$
$\frac{d}{dx} \operatorname{arcsinh}(x) = \frac{1}{\sqrt{x^2+1}}$	$\frac{d}{dx} \operatorname{arccosh}(x) = \frac{1}{\sqrt{x^2-1}}$	$\frac{d}{dx} \operatorname{arctanh}(x) = \frac{1}{1-x^2}$

5.5 Ableitungen von skalaren Funktionen mehrerer Variablen

In der Physik betrachten wir oft auch skalare Funktionen von mehreren Variablen, zum Beispiel die Temperatur als Funktion des Ortes und der Zeit, $T(\vec{r}, t) = T(x, y, z, t)$. Um das Konzept der

Ableitung auf solche Funktionen zu verallgemeinern, erinnern wir uns was die Ableitung einer skalaren Funktion einer Variablen uns sagt – nämlich um wie viel sich der Funktionswert ändert, wenn wir die Variable ändern. Bei mehreren Variablen definiert man analog sogenannte **partielle Ableitungen**. Diese beschreiben, um wie viel sich der Funktionswert ändert, wenn wir den Wert von nur einer der Variablen ändern. Alle anderen Variablen werden festgehalten. Man definiert für die Funktion

$$f : \mathbb{R}^n \mapsto \mathbb{R} \ , \ \vec{x} \mapsto f(\vec{x}) = f(x_1, x_2, \ldots, x_n)$$

die **partielle Ableitung nach** x_i als

$$\frac{\partial f(\vec{x})}{\partial x_i} = \lim_{\Delta x_i \to 0} \frac{f(x_1, x_2, \ldots, x_i + \Delta x_i, \ldots, x_n) - f(x_1, x_2, \ldots, x_i, \ldots, x_n)}{\Delta x_i} \tag{5.13}$$

Beachten Sie die neue Notation mit ∂ als Symbol für die partielle Ableitung. Oft schreibt man verkürzt auch

$$\frac{\partial f(\vec{x})}{\partial x_i} = \partial_i f(\vec{x}) = f_{x_i}(\vec{x}). \tag{5.14}$$

5.5.1 Erste Ableitung: der Gradient und der Nabla-Operator

Es ist oft hilfreich, die partiellen (ersten) Ableitungen einer skalaren Funktion mehrerer Variablen in einen Vektor zu packen. Diesen Vektor nennt man dann den **Gradienten** der Funktion $f(\vec{x})$ am Punkt $\vec{x}_0$. Der Gradient hat folgende Symbole:

$$\text{grad}\, f(\vec{x}) = \nabla f(\vec{x}) = \begin{pmatrix} \frac{\partial f(\vec{x})}{\partial x_1} \\ \frac{\partial f(\vec{x})}{\partial x_2} \\ \vdots \\ \frac{\partial f(\vec{x})}{\partial x_n} \end{pmatrix} \tag{5.15}$$

Das Symbol ∇ nennt man "Nabla", beziehungsweise genauer den "Nabla-Operator". (Dieser Operator weist einer Funktion ihren Gradienten zu.) Der Name "Nabla" leitet sich ab von einem harfenähnlichen phönizischen Saiteninstrument, das in etwa die Form dieses Zeichens hatte. Der Nabla-Operator ∇ ist nichts anderes als der Vektor der partiellen Ableitungen,

$$\nabla = \begin{pmatrix} \frac{\partial}{\partial x_1} \\ \frac{\partial}{\partial x_2} \\ \vdots \end{pmatrix} . \tag{5.16}$$

Der Gradient hat eine einfache **geometrische Interpretation**: man kann zeigen, dass er **senkrecht zu den Äquipotentiallinien / Äquipotentialflächen der ursprünglichen Funktion** steht (also den Flächen bzw. Linien, auf denen die Funktion einen gegebenen konstanten Wert annimmt, $f(\vec{x}) = f_0$).

5.5.2 Zweite Ableitungen: Hesse-Matrix und Laplace-Operator

Bei Funktionen mehrerer Variablen gibt es natürlich auch viele verschiedene zweite Ableitungen. Diese bildet man wie bei den höheren Ableitungen einer Funktion von nur einer Variablen dadurch, dass man die Ableitungen der Ableitungen berechnet:

$$\frac{\partial}{\partial x_i}\frac{\partial f(\vec{x})}{\partial x_j} = \frac{\partial^2 f(\vec{x})}{\partial x_i \partial x_j} = \partial_i \partial_j f = \frac{\partial\left(\frac{\partial f(\vec{x})}{\partial x_j}\right)}{\partial x_i}. \tag{5.17}$$

Die zweiten Ableitungen fasst man in der sogenannten **Hesse-Matrix** $H_f(\vec{x})$ zusammen, deren Komponenten/Einträge durch die verschiedenen zweiten Ableitungen gegeben sind:

$$H_f(\vec{x}) \quad \rightarrow \quad H_f(\vec{x})_{ij} = \frac{\partial}{\partial x_i}\frac{\partial f(\vec{x})}{\partial x_j} \tag{5.18}$$

Aus den nicht-gemischten zweiten partiellen Ableitungen kann man weiterhin den sogenannten **Laplace-Operator** Δ konstruieren:

$$\Delta = \sum_i \frac{\partial^2}{\partial x_i^2} \quad \Rightarrow \quad \Delta f(\vec{x}) = \sum_i \frac{\partial^2 f(\vec{x})}{\partial x_i^2} \tag{5.19}$$

(hierbei läuft i über die Komponenten von $\vec{x}$). Der Laplace-Operator kommt in vielen Differentialgleichungen (siehe Kapitel 9) vor, die das Verhalten physikalischer Felder beschreiben. Beispiele sind die Poisson-Gleichung der Elektrostatik, die Navier-Stokes-Gleichungen für Strömungen von Flüssigkeiten oder Gasen und die Diffusionsgleichung für die Wärmeleitung.

5.5.3 Satz von Schwarz

Im Allgemeinen ist die Reihenfolge der höheren Ableitungen wichtig:

$$\frac{\partial}{\partial x_i}\frac{\partial f(\vec{x})}{\partial x_j} \neq \frac{\partial}{\partial x_j}\frac{\partial f(\vec{x})}{\partial x_i} \quad \text{(im Allgemeinen)}. \tag{5.20}$$

Dies vereinfacht sich aber, falls die Funktion f wieder "ausreichend nett" ist. Genauer regelt das der **Satz von Schwarz**, der folgendes aussagt:

> Wenn $f : \mathbb{R}^n \to \mathbb{R}$ mindestens k-mal partiell differenzierbar ist, und alle k-ten partiellen Ableitungen stetig sind, so ist die Reihenfolge aller l-ten Ableitungen mit $l \leq k$ unerheblich.

Es gilt also zum Beispiel: wenn die zweiten Ableitungen von f alle stetig sind, dann ist $\frac{\partial}{\partial x_i}\frac{\partial f(\vec{x})}{\partial x_j} = \frac{\partial}{\partial x_j}\frac{\partial f(\vec{x})}{\partial x_i}$ für alle i und j.

5.5.4 Totales Differenzial

Bei Funktionen mehrerer Variablen verallgemeinert man auch das Konzept des Differenzials zum Konzept des **totalen Differenzials** df:

$$df(\vec{x}) = \sum_i \frac{\partial f(\vec{x})}{\partial x_i}\, dx_i = \frac{\partial f(\vec{x})}{\partial x_1}\, dx_1 + \frac{\partial f(\vec{x})}{\partial x_2}\, dx_2 + \dots. \tag{5.21}$$

$df(\vec{x})$ gibt also die Änderung des Funktionswerts der Funktion f an, die sich aus (infinitesimal kleinen) Änderungen dx_i der Variablenwerte ergibt.

5.5.5 Totale und partielle Ableitung

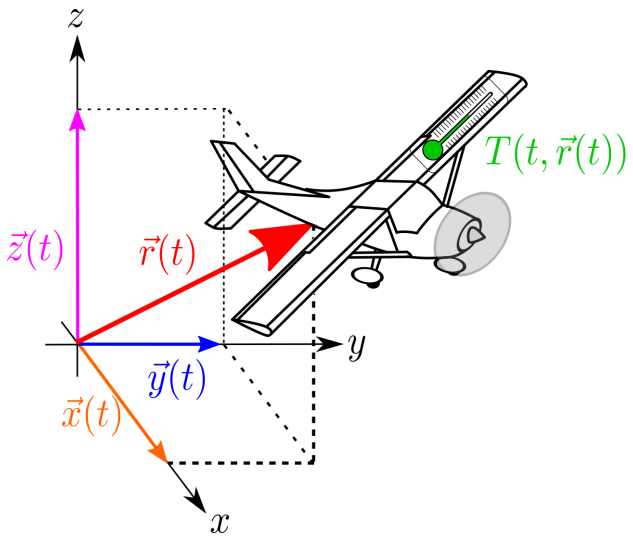
Die Ableitungsregeln verallgemeinern sich auch direkt von den Funktionen einer Variablen auf die Funktionen mehrerer Variablen. Besondere Erwähnung verdient hier noch die Kettenregel: man muss einfach die Kettenregel für jede der Variablen durchführen. So gilt zum Beispiel für zwei Funktionen $f : \mathbb{R} \to \mathbb{R}$ und $g : \mathbb{R} \to \mathbb{R}$ (also Funktionen $f(x)$ und $g(x)$ von nur einer Variablen) und einer Funktion $f : \mathbb{R}^2 \to \mathbb{R}$ (also $h(x, y)$):

$$\frac{dh(f(x), g(x))}{dx} = \frac{\partial h(f, g)}{\partial f}\frac{df(x)}{dx} + \frac{\partial h(f, g)}{\partial g}\frac{dg(x)}{dx}. \tag{5.22}$$

Das ist insbesondere bei Zeitableitungen wichtig. Wir können uns zum Beispiel fragen, welche Temperatur ein Thermometer anzeigt, dass an der Außenhülle eine Flugzeugs angebracht ist. Diese Temperatur T kann sich aus verschiedenen Gründen ändern:

- Die angezeigte Temperatur kann sich ändern, weil sich die Temperatur selbst mit der Zeit t ändert (nachts ist es z. B. kälter als am Tag) – man sagt, dass die Temperatur "explizit zeitabhängig" ist.

- Die angezeigte Temperatur kann sich ändern, weil sich das Flugzeug von einem warmen Ort (Mallorca) zu einem kalten Ort (Stockholm) bewegt hat – man sagt, dass die Temperatur dann auch "implizit zeitabhängig" ist.



Es gilt also:

$$T = T(t, \vec{r}(t)) = T(t, x(t), y(t), z(t).$$

Hier zeigt die erste Variable t die explizite Zeitabhängigkeit der Temperatur an, während die weiteren Variablen $x(t)$, $y(t)$, $z(t)$ die implizite Zeitabhängigkeit beinhalten. Die "gesamte" Ableitung der Temperatur nach der Zeit nennen wir die **totale Ableitung** der Temperatur nach der Zeit. Diese hat das Symbol $\frac{dT}{dt}$ und ist wie folgt definiert:

$$
\begin{aligned}
\underbrace{\frac{dT(t,\vec{r}(t))}{dt}}_{\text{total Ableitung nach der Zeit}} &= \frac{\partial T(t,\vec{r}(t))}{\partial t} + \frac{\partial T(t,\vec{r}(t))}{\partial x}\,\frac{dx(t)}{dt} + \frac{\partial T(t,\vec{r}(t))}{\partial y}\,\frac{dy(t)}{dt} + \frac{\partial T(t,\vec{r}(t))}{\partial z}\,\frac{dz(t)}{dt} \\
&= \underbrace{\frac{\partial T(t,\vec{r}(t))}{\partial t}}_{\text{partielle Ableitung nach der Zeit}} \\
&\quad + \underbrace{\frac{\partial T(t,\vec{r}(t))}{\partial x}}_{\text{Ableitung nach } x}\ \underbrace{\frac{dx(t)}{dt}}_{\text{totale Zeitableitung von } x} \\
&\quad + \underbrace{\frac{\partial T(t,\vec{r}(t))}{\partial y}}_{\text{Ableitung nach } y}\ \underbrace{\frac{dy(t)}{dt}}_{\text{totale Zeitableitung von } y} \\
&\quad + \underbrace{\frac{\partial T(t,\vec{r}(t))}{\partial z}}_{\text{Ableitung nach } z}\ \underbrace{\frac{dz(t)}{dt}}_{\text{totale Zeitableitung von } z}
\end{aligned}
\tag{5.23}
$$

Zusammengefasst gilt:

- Die totale Zeitableitung entspricht der Änderung der auf dem Thermometer angezeigten Temperatur mit der Zeit.
- Diese setzt sich aus mehreren Termen zusammen
  - der partiellen Ableitung nach der Zeit – diese beschreibt die explizite Zeitabhängigkeit der Temperatur,
  - den Kettenregel-Ableitungen nach Ort, und dann Ort nach Zeit – diese beschreiben die Änderung der Temperatur, die von der Änderung des Ortes (von einem warmen zu einem kalten Ort) stammt.

5.6 Ableitungen von vektorwertigen Funktionen

Bei vektorwertigen Funktionen $\vec{f}$ verallgemeinern sich die obigen Formeln im Wesentlichen dadurch, dass man überall f durch $\vec{f}$ ersetzt. Insbesondere sind partielle Ableitungen einer Funktion $\vec{f}: \mathbb{R}^n \to \mathbb{R}^m$, die einem n-komponentigen Vektor $\vec{x}$ den m-komponentigen Funktionswert $\vec{f}(\vec{x})$ zuordnet, wie folgt definiert:

$$
\frac{\partial \vec{f}(\vec{x})}{\partial x_i} = \lim_{\Delta x_i \to 0} \frac{\vec{f}(x_1, x_2, \ldots, x_i + \Delta x_i, \ldots, x_n) - \vec{f}(x_1, x_2, \ldots, x_i, \ldots, x_n)}{\Delta x_i}. \tag{5.24}
$$

Wenn man in kartesischen Koordinaten rechnet, entspricht dies grade der komponentenweisen Ableitung:

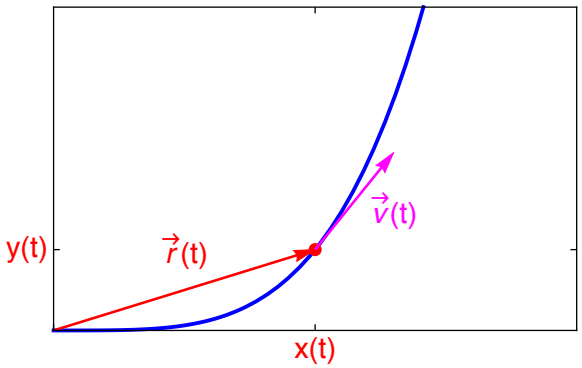
$$
\frac{\partial \vec{f}(\vec{x})}{\partial x_i} = \frac{\partial}{\partial x_i}\left(\sum_{j=1}^{m} f_j(\vec{x})\,\vec{e}_j\right) = \sum_{j=1}^{m} \frac{\partial f_j(\vec{x})}{\partial x_i}\,\vec{e}_j = \begin{pmatrix} \frac{\partial f_1(\vec{x})}{\partial x_i} \\ \frac{\partial f_2(\vec{x})}{\partial x_i} \\ \vdots \\ \frac{\partial f_m(\vec{x})}{\partial x_i} \end{pmatrix}. \tag{5.25}
$$

Die Rechenregeln zur Ableitung verallgemeinern sich damit direkt auf vektorwertige Funktionen und Felder. Es gilt zum Beispiel:

- $\frac{d\left(\vec{f}(x)\times\vec{g}(x)\right)}{dx} = \vec{f}(x) \times \frac{d\vec{g}(x)}{dx} + \vec{g}(x) \times \frac{d\vec{f}(x)}{dx}$.
- $\frac{d}{dx}(\vec{f}(x) \cdot \vec{g}(x)) = \vec{f}(x) \cdot \frac{d\vec{g}(x)}{dx} + \vec{g}(x) \cdot \frac{d\vec{f}(x)}{dx}$.

5.6.1 Ableitungen von Raumkurven

Eine besondere Klasse von vektorwertigen Funktionen sind Raumkurven $\vec{r}(t)$, also der Ort als Funktion eines Parameters (hier der Zeit t). Da diese in der Physik eine besondere Rolle spielen, gehen wir nochmal gesondert auf Raumkurven ein. Wir betrachten Raumkurven im dreidimensionalen Raum (im zweidimensionalen Raum geht das aber genauso). Eine solche Raumkurve ist eine vektorwertige Funktion $\vec{r} : \mathbb{R} \to \mathbb{R}^3$, $t \mapsto \vec{r}(t)$.



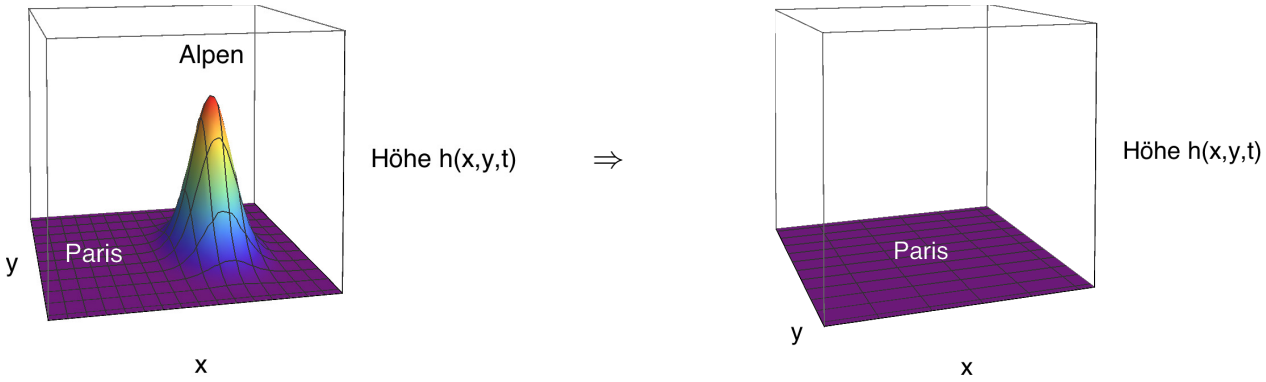
Diese Funktion beinhaltet eine Menge Informationen:

- Ort zu einem Zeitpunkt t: $\vec{r}(t)$.
- Geschwindigkeit zu einem Zeitpunkt t: $\vec{v}(t) = \frac{d\vec{r}}{dt} = \dot{\vec{r}}(t)$.
- Tangenten-Einheitsvektor an die Raumkurve zu einem Zeitpunkt t: $\vec{t}(t) = \frac{\vec{v}(t)}{|\vec{v}(t)|}$
- Die zwischen einer Zeit t_0 und der Zeit t zurückgelegte Strecke ("Bogenlänge"): $s(t, t_0) = \int_{t_0}^{t} dt' \, |\vec{v}(t')|$ (hierzu muss man integrieren, siehe Kapitel 6).

5.7 Taylorreihe

Eines der wichtigsten mathematischen Werkzeuge in der Physik ist die sogenannte Taylorreihe. Die Idee ist hierbei, dass man eine Funktion oft nicht überall kennen muss, sondern nur in der Nähe eines bestimmten Punktes. Zum Beispiel ist bei einer einzelnen Etappe der Tour de France das Höhenprofil $h(x, y)$ nur im Tagesabschnitt relevant, also nur für eine kleine Untermenge von Punkten (x, y).In diesem Unterabschnitt ist es oft sinnvoll, die Funktion des Höhenprofils durch eine einfachere Funktion zu ersetzen, die das Höhenprofil im gegebenen Unterabschnitt gut nähert. Zum Beispiel ist das Höhenprofil der Schlußetappe der Tour de France in Paris mehr oder weniger flach,

$$h(x, y)|_{\text{Paris}} \approx 35\,\text{m}.$$



Höhenprofil der Tour de France

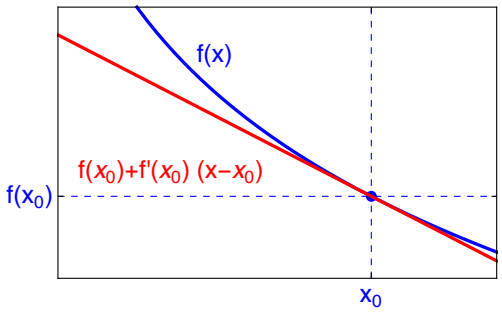
Höhenprofil der letzten Etappe.

Die Taylorreihe ist die mathematisch korrekte Art, aus einer gegebenen Funktion einen Näherung für diese Funktion in der Nähe eines bestimmten Punktes zu bekommen.

5.7.1 Taylorreihe einer skalaren Funktion einer Variablen

Betrachten wir zunächst die Funktion $f : \mathbb{R} \to \mathbb{R}$, $x \mapsto f(x)$. Diese Funktion sei stetig und unendlich oft differenzierbar. Wir wollen diese Funktion in der Nähe des Punkts x_0 nähern. Die Idee ist wie folgt:

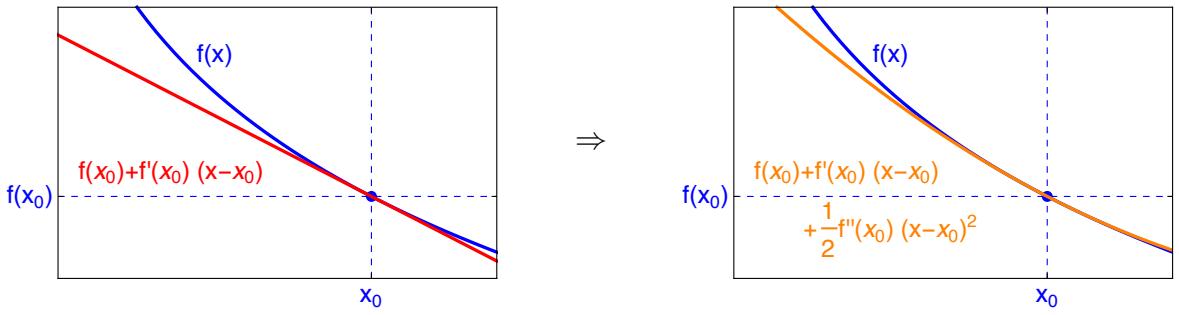
1. Ganz grob gesagt gilt für $x \approx x_0$, dass
$$f(x \approx x_0) \approx f(x_0).$$
2. Wenn wir etwas genauer hinschauen, so gilt in einer Umgebung um x_0, dass wir die Funktion $f(x)$ linearisieren können. Die Steigung der linearisierten Funktion ist dabei die Ableitung von f am Punkt x_0. Es gilt also:
$$f(x \approx x_0) \approx f(x_0) + \frac{df(x_0)}{dx}(x - x_0).$$



Dieses Spiel kann man jetzt immer weiter fortsetzen, und die Funktion immer besser annähern. Dies geschieht formal mit der sogannten "Taylorreihe". Die **Taylorreihe von $f(x)$ um den Punkt $x = x_0$** ist definiert als

$$\begin{aligned} Tf(x, x_0) &= \sum_{n=0}^{\infty} \frac{1}{n!} f^{(n)}(x_0)\,(x - x_0)^n \\ &= f(x_0) + f'(x_0)\,(x - x_0) + \frac{1}{2} f''(x_0)\,(x - x_0)^2 + \frac{1}{6} f'''(x_0)\,(x - x_0)^3 + \dots . \end{aligned} \tag{5.26}$$

Hier haben wir die "nullte" Ableitung der Funktion f als die Funktion selbst definiert, $f^{(0)}(x_0) = f(x_0)$.



Für viele Funktionen (die "ausreichend netten") gilt, dass $Tf(x, x_0) = f(x)$ (und es ist bei der unendlichen Reihe egal, welchen Entwicklungspunkt x_0 man wählt!). **Aber Achtung**: das gilt leider nicht immer!

- Die Reihe $Tf(x, x_0)$ konvergiert nicht immer.
- Selbst wenn $Tf(x, x_0)$ konvergiert, muss der Grenzwert nicht automatisch gleich $f(x)$ sein.

Trotzdem gilt $Tf(x, x_0) = f(x)$ in so vielen Fällen, dass Physiker ständig "taylorn". Das Ziel ist wie gesagt, die Funktion in der Nähe eines Punktes x_0 zu approximieren. Je nach gewünschter Genauigkeit (das hängt dann vom konkreten Problem ab), bricht man bei einer solchen **Taylor-Entwicklung** die Taylorreihe nach der gewünschten Ordnung ab. Will man die Funktion bis zur zweiten Ordnung nähern um den Punkt x_0, so nutzt man also

$$f(x) \approx f(x_0) + \frac{df(x_0)}{dx}(x - x_0) + \frac{d^2 f(x_0)}{dx^2}(x - x_0)^2.$$

5.7.1.1 Einige wichtige Taylorreihen

Im Folgenden listen wir einige wichtige Taylorreihen um $x = 0$ auf (Taylorreihen um diesen Punkt $x = 0$ nennt man auch Maclaurin-Reihen).

1. $e^x = \sum_{n=0}^{\infty} \frac{1}{n!} x^n = 1 + x + \frac{x^2}{2} + \frac{x^3}{6} + \dots$
2. $\log(1 + x) = \sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n} x^n = x - \frac{x^2}{2} + \frac{x^3}{3} + \dots$
3. $\sin(x) = \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n+1)!} x^{2n+1} = x - \frac{x^3}{6} + \frac{x^5}{120} + \dots$
4. $\cos(x) = \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n)!} x^{2n} = 1 - \frac{x^2}{2} + \frac{x^4}{24} + \dots$

Andere Taylorreihen sind in Lehrbüchern oder bei Wikipedia tabelliert (bzw. lassen sich mit den allgemeinen Formeln oben leicht berechnen!).

5.7.2 Taylorreihe einer skalaren Funktion mehrerer Variablen

Das Konzept der Taylorreihe lässt sich auch auf mehrere skalare Funktionen mehrerer Variablen verallgemeinern. Hierzu schreiben wir die allgemein Form der Taylorreihe als Exponentialfunktion von Ableitungen (dies kann man auch als Definition der Exponentialfunktion der Ableitung verstehen):

$$\begin{aligned} Tf(x_0 + \delta x, x_0) &= \sum_{n=0}^{\infty} \frac{1}{n!} f^{(n)}(x_0)\, \delta x^n = \sum_{n=0}^{\infty} \frac{1}{n!} \delta x^n \left(\frac{d}{dx}\right)^n f|_{x=x_0} \\ &= e^{\delta x \frac{d}{dx}} f\Big|_{x=x_0}. \end{aligned} \tag{5.27}$$

Die lässt sich nun einfach auf viele Variablen verallgemeinern:

$$\begin{aligned} Tf(\vec{x}_0 + \delta\vec{x}, \vec{x}_0) &= e^{\delta\vec{x}\cdot\nabla} f\big|_{\vec{x},\vec{x}_0} \\ &= f(\vec{x}_0) + \delta\vec{x}\cdot\nabla f(\vec{x}_0) + (\delta\vec{x}\cdot\nabla)\ (\delta\vec{x}\cdot\nabla)\ f(\vec{x}_0) + \dots \end{aligned} \tag{5.28}$$

Für eine Funktion von zwei Variablen gilt also bis zur zweiten Ordnung (und mit Nutzung des Satz von Schwarz):

$$\begin{aligned} Tf(x+\delta x, y+\delta y, x_0, y_0) \approx f(x_0,y_0) &+ \delta x\,\frac{df(x_0,y_0)}{dx} + \delta y\,\frac{df(x_0,y_0)}{dy} \\ &+ \frac{\delta x^2}{2}\,\frac{d^2f(x_0,y_0)}{dx^2} + \delta x\delta y\,\frac{d^2f(x_0,y_0)}{dxdy} + \frac{\delta y^2}{2}\,\frac{d^2f(x_0,y_0)}{dy^2} \end{aligned} \tag{5.29}$$

5.8 Ableitungen zur Bestimmung von Extremalwerten: Erinnerung an Kurvendiskussion

Oftmals will man besondere Punkte einer Funktion bestimmen, zum Beispiel den Punkt, an dem eine Funktion maximal oder minimal wird: wann wird die Energie eines Systems minimal? Wann wird der Gewinn in einem Unternehmen maximal? Das mathematische Werkzeug hierzu ist die Kurvendiskussion.

5.8.1 Extremalwerte einer skalarwertigen Funktion einer Variablen

Um Kurvendiskussion besser zu verstehen, kann man sich die geometrische Interpretation von Ableitungen einer skalarwertigen Funktion $F : \mathbb{R} \mapsto \mathbb{R}$ von nur einer Variablen in Erinnerung rufen:

- $f'(x) = \frac{df(x)}{dx} > 0$: Funktion steigt bei x an,
- $f'(x) = \frac{df(x)}{dx} = 0$: Funktion ist bei x flach (steigt weder an, noch fällt sie ab).
- $f'(x) = \frac{df(x)}{dx} < 0$: Funktion fällt bei x ab.

Es gilt weiterhin für die zweite Ableitung:

- $f''(x) = \frac{d^2f(x)}{dx^2} > 0$: die Steigung der Tangenten bei x nimmt zu ("Linkskurve"),
- $f''(x) = \frac{d^2f(x)}{dx^2} < 0$: die Steigung der Tangenten bei x nimmt ab ("Rechtskurve").

Notwendige, aber nicht hinreichende Bedingung für das Vorliegen eines lokalen Extrempunkts (Maximum oder Mininum) im Punkt x_0 ist, dass die erste Ableitung verschwindet:

$$f'(x_0) = \frac{df(x_0)}{dx} = 0.$$

Ob an diesem Punkt wirklich ein Extremum vorliegt, und ob es ein Maximum oder Minimum ist, entscheidet sich am Wert der zweiten Ableitung:

$$f'(x_0) = \frac{df(x_0)}{dx} = 0 \quad \Rightarrow \quad \begin{cases} f''(x_0) = \frac{d^2f(x_0)}{fx^2} < 0 & \text{: Funktion hat bei } x_0 \text{ ein lokales Maximum} \\ f''(x_0) = \frac{d^2f(x_0)}{dx^2} = 0 & \text{: Funktion hat bei } x_0 \text{ einen Sattelpunkt} \\ f''(x_0) = \frac{d^2f(x_0)}{dx^2} > 0 & \text{: Funktion hat bei } x_0 \text{ ein lokales Minimum} \end{cases}$$

5.8.2 Extremalwerte einer skalaren Funktion mehrerer Variablen

Für eine Funktion mehrerer Variablen verallgemeinert sich das Konzept von ersten und zweiten Ableitungen zum Gradient und zur Hessematrix. **Notwendige, aber nicht hinreichende Bedingung** für das Vorliegen eines lokalen Extrempunkts (Maximum oder Mininum) im Punkt x_0 ist, dass alle ersten Ableitungen verschwinden:

$$\nabla f(\vec{x}_0) = \begin{pmatrix} \frac{df(\vec{x}_0)}{dx_1} \\ \frac{df(\vec{x}_0)}{dx_2} \\ \cdots \end{pmatrix} = 0.$$

Die Hessematrix $H_f(\vec{x})$ (die Matrix der zweiten Ableitungen) entscheidet darüber, ob dann bei $\vec{x}_0$ ein Extrempunkt vorliegt. Es gilt:

$$\nabla f(\vec{x}_0) = 0 \quad \Rightarrow \quad \begin{cases} H_f(\vec{x}) \text{ hat nur negative Eigenwerte } (H_f(\vec{x}) \text{ ist "negativ definit"}) & : \text{lokales Maximum,} \\ H_f(\vec{x}) \text{ hat nur positive Eigenwerte } (H_f(\vec{x}) \text{ ist "positiv definit"}) & : \text{lokales Minimum.} \end{cases}$$

Das wäre im Physikerleben gut zu wissen:

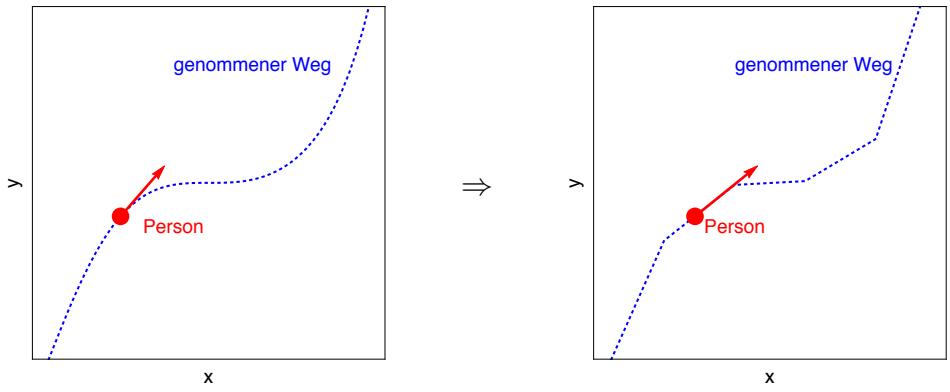
- Wissen, wie eine Ableitung definiert ist und was sie geometrisch bedeutet.
- Ableitungsregeln kennen (Linearität der Ableitung, Produktregel, Quotientenregel, Kettenregel, Ableitung der Umkehrfunktion, Regel von l'Hôpital)
- Die Ableitungen von Polynomen, Exponentialfunktion, Logarithmus, Sinus, Kosinus und Tangens kennen.
- Vektorwertige Funktionen ableiten können.
- Wissen, was ∇ ist und einen Gradienten berechnen können.
- Hessematrix aufstellen können.
- Satz von Schwarz kennen.
- Wissen, was ein totales Differenzial ist.
- Wissen, was eine totale und eine partielle Ableitung ist.
- Wissen, wie die Taylorreihe allgemein definiert ist, und wissen, wozu sie gut ist.
- Die Taylorreihen der Exponentialfunktion, des Logarithmus, von Sinus und Kosinus kennen (zumindest erste zwei oder drei Glieder).
- Extremwerte einer Kurve bestimmen können.

PS: "gut zu wissen" muss nicht heißen, dass das alles in der Klausur abgefragt wird – oder, dass sonst nichts aus diesem Kapitel in der Klausur vorkommt. Wer aber mit den Punkten in diesem Kasten nichts anfangen kann, hat wahrscheinlich in der Klausur und im weiteren Studium Probleme.

Kapitel 6

Integralrechnung

In der Physik werden oft Integrale berechnet. Letztlich hat das etwas damit zu tun, dass viele Dinge nicht in einzelnen diskreten Werten vorliegen, sondern sich kontinuierlich ändern können, man damit aber nicht immer ganz einfach rechnen kann. Ein Beispiel hier ist der Ort einer Person, die über einen Platz läuft:



Wenn man die von der Person zurückgelegte Wegstrecke mit einem Zollstock bestimmen will, hat man das Problem, dass der Weg nicht gerade ist. Näherungsweise kann man die Wegstrecke dann dadurch bestimmen, dass man den zurückgelegten Weg in kleine, gerade Teilstücke zerlegt. Deren Länge kann man mit dem Zollstock bestimmen, und so die gesamte Wegstrecke annähern:

$$\text{zurückgelegter Weg} \approx \sum_i \text{Länge des } i\text{-ten geraden Wegstücks.}$$

Diese Näherung wird offensichtlich besser, wenn man den Weg in immer mehr immer kleinere Teilstücke zerlegt. Im Grenzfall infinitesimal kleiner ("unendlich kleiner") Teilstücke bekommt man dann den exakten zurückgelegten Weg:

$$\text{zurückgelegter Weg} = \lim_{\text{Teilstück-Länge}\to 0} \sum_i \text{Länge des } i\text{-ten geraden Wegstücks.}$$

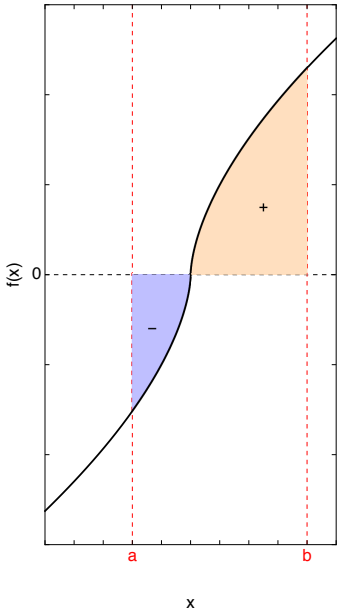
Das aufaddieren unendlich vieler, unendlich kleiner Teilstücke ist genau die Grundidee eines "Integrals". Das Beispiel der Weglänge ist ein Wegintegral, zu dem wir in Kapitel 8.2.1 nochmals zurückkommen werden.

6.1 Integral als Fläche

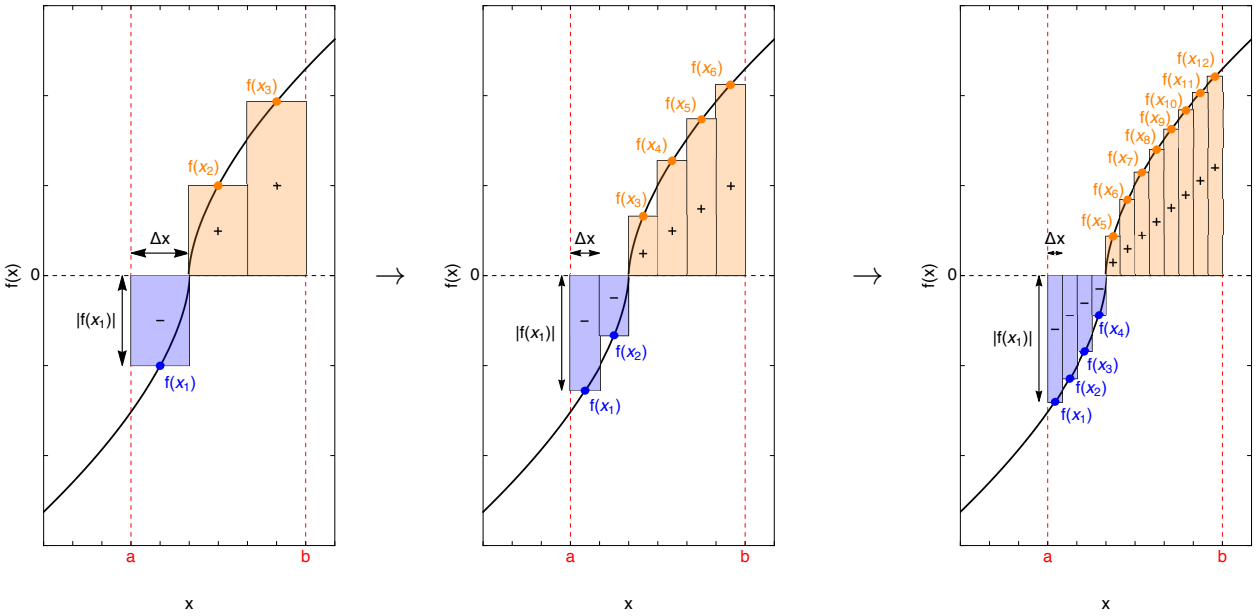
Sie kennen das Integral aus der Schule als eine "vorzeichenbehaftete Fläche": das Integral einer Funktion $f(x)$ (der "Integrand") zwischen den Grenzen a und b, bezeichnet mit dem Symbol

$$\int_a^b dx\, f(x)$$

entspricht der Fläche zwischen der Funktion $f(x)$ und der x-Achse, wobei Flächenanteile oberhalb der x-Achse positiv gezählt werden, während Flächenanteile unterhalb der x-Achse negativ gezählt werden. Das Symbol dx heißt "Differenzial", und ist philosophisch dasselbe Objekt wie in Gleichung (5.3) (für die Funktion $f(x) = x$).



Um die vorzeichenbehaftete Fläche zu bestimmen, muss man sich wieder überlegen, wie man mit der kontinuierlichen, gekrümmten Kurve umgeht. Riemann hat vorgeschlagen, die Fläche unter der Kurve so zu approximieren, dass man sie gut näherungsweise berechnen kann – und zwar durch kleine Rechtecke. Für diese kennen wir den Flächeninhalt natürlich: er ist Höhe mal Breite. Konkret geben wir jedem Rechteck die (kleine) Breite Δx, wobei wir den Grenzwert immer kleinerer Δx nehmen:



Die Höhe des i-ten Rechtecks ist $|f(x_i)|$, die Breite ist Δx. Der Flächeninhalt des i-ten Rechtecks ist also $A_i = \Delta x \cdot |f(x_i)|$. Wenn wir nun die Flächenstücke unterhalb der x-Achse mit einem

Minus-Zeichen versehen, wird aus $|f(x_i)|$ wieder $f(x_i)$. Zuletzt nehmen wir noch den Grenzfall von unendlich vielen unendlich dünnen Rechtecken, $\Delta x \to 0$, und definieren das Integral als den vorzeichenbehafteten Flächeninhalt in diesem Grenzwert:

$$\int_a^b dx\, f(x) = \lim_{\Delta x \to 0} \sum_i \Delta x\, f(x_i) \tag{6.1}$$

In diesem Integral haben wir explizit die **Grenzen** a und b angegeben. Solch ein Integral nennt man auch ein **bestimmtes** Integral. Es ist übrigens egal, ob das dx direkt hinter dem Integralzeichen steht, oder am Ende des Integranden (= der Funktion, die integriert wird). Ebenso ist es egal, wie die Integrationsvariable (hier x) genannt wird:

$$\int_a^b dx\, f(x) = \int_a^b f(x)\, dx = \int_a^b f(t)\, dt = \int_a^b d\xi\, f(\xi). \tag{6.2}$$

Die Konstruktion des Integral über Flächeninhalte von Rechtecken wird nach seinem Erfinder auch "Riemannsches Integral" genannt.

6.2 Praktische Berechnung von Integralen: der Hauptsatz der Integral- und Differentialrechung

Wie genau berechnet man jetzt ein Integral? Offensichtlich ist es etwas unpraktisch, jedes Mal mit dem Geodreieck kleine Rechtecke zu zeichnen, wenn man ein Integral berechnen will. In der Tat macht man das quasi nie, sondern nutzt stattdessen den "Hauptsatz der Differential- und Integralrechnung", auch Fundamentalsatz der Analysis genannt. Dieser verbindet die Integral und Differenzialrechnung. Er besagt im Wesentlichen, dass Integration und Differentiation (Ableitungen) mathematische Gegenstücke sind, und man Integrale daher über sogenannte **Stammfunktionen** berechnen kann. Eine Stammfunktion wiederum ist wie folgt definiert:

Sei $f : \mathbb{R} \to , x \mapsto f(x)$ eine stetige, reelle Funktion. Wir nennen genau dann eine andere Funktion $F(x)$ **Stammfunktion von** $f(x)$, wenn

$$\frac{dF(x)}{dx} = F'(x) = f(x). \tag{6.3}$$

Für eine gegebene Funktion gibt es im Allgemeinen unendlich viele Stammfunktionen: ist $F(x)$ Stammfunktion von $f(x)$, so ist auch $\tilde{F}(x) = F(x) + c$ mit reeller Konstante c Stammfunktion von $f(x)$:

$$\frac{dF(x)}{dx} = F'(x) = f(x) \quad \Rightarrow \quad \frac{d\tilde{F}(x)}{dx} = \frac{d(F(x)+c)}{dx} = \frac{dF(x)}{dx} + \frac{dc}{dx} = F'(x) + 0 = f(x). \tag{6.4}$$

Der Hauptsatz der Integral- und Differentialrechung besagt nun folgendes:

Sei $f : \mathbb{R} \to , x \mapsto f(x)$ eine stetige, reelle Funktion auf dem Intervall $[a, b]$ und $x_0 \in [a, b]$. Dann ist die über das Integral definierte Funktion

$$F(x) = \int_{x_0}^{x} dy\, f(y) \tag{6.5}$$

differenzierbar in $[a, b]$, und sie ist in diesem Intervall eine Stammfunktion von $f(x)$:

$$\frac{d}{dx}F(x) = \frac{d}{dx}\left(\int_{x_0}^{x} dy\, f(y)\right) = f(x). \tag{6.6}$$

Hierbei kann man $x_0 \in [a, b]$ beliebig wählen. Unterschiedliche Werte von x_0 ändern die Stammfunktion genau um eine Konstante. Weiterhin gilt:

$$\int_a^b dx\, f(x) = F(b) - F(a). \tag{6.7}$$

Dieser Satz hat zwei wichtige Aussagen:

- Wenn wir das **Integral** einer Funktion f berechnen wollen, müssen wir nur die **Stammfunktion** F **an den Integralgrenzen** auswerten. Dies kann man etwas kompakter wie folgt schreiben:
$$\int_a^b dx\, f(x) = F(b) - F(a) = [F(x)]_a^b = F(x)\big|_a^b\,.$$
Das ist praktisch, denn Stammfunktionen sind bekannt und tabelliert (bzw. im Computer vorprogrammiert).
- **Integration und Differentiation sind mathematische Gegenstücke**:
$$f(x) \xrightarrow{\int dx} F(x) \xrightarrow{\frac{d}{dx}} f(x).$$

6.3 Integral Know-How

6.3.1 Uneigentliche Integrale

Bisher haben wir nur Integrale betrachtet, die der Fläche unter einer "netten" Funktion in einem abgeschlossenen Intervall entsprechen. Diesen Integralbegriff verallgemeinern wir nun zum uneigentlichen Integral. Ein **uneigentliches Integral** ist ein Integral, bei dem entweder die Integralgrenzen nicht beschränkt sind (also $\pm\infty$), oder der Integrand selbst (die zu integrierende Funktion) an einzelnen Punkten nicht beschränkt ist (divergiert). Wir definieren uneigentliche Integrale über geeignete Grenzwerte:

6.3.1.1 Unendliche Grenzen

Wir definieren für reelle Zahlen a, b und c

$$\int_a^\infty dx\, f(x) = \lim_{b\to\infty}\int_a^b dx\, f(x). \tag{6.8}$$

$$\int_{-\infty}^b dx\, f(x) = \lim_{a\to-\infty}\int_a^b dx\, f(x). \tag{6.9}$$

$$\int_{-\infty}^\infty dx\, f(x) = \lim_{a\to-\infty}\int_a^c dx\, f(x) + \lim_{b\to\infty}\int_c^b dx\, f(x) \text{ (hier ist } c \text{ beliebig)}. \tag{6.10}$$

Im letzten Beispiel ist es wichtig, dass die beiden Grenzwerte für die obere und untere Grenze voneinander unabhängig sind!

6.3.1.2 Nicht-beschränkter Integrand

Die Funktion $f(x)$ habe eine "Polstelle" bei $x = c$ mit $c \in [a, b]$, d.h. sie divergiere an der Stelle c: $\lim_{x\to c} f(x) = \pm\infty$. Dann definieren wir das uneigentliche Integral von f wie folgt:

$$\int_a^b dx\, f(x) = \lim_{\epsilon\to 0} \int_a^{c-\epsilon} dx\, f(x) + \lim_{\delta\to 0} \int_{c+\delta}^b dx\, f(x).$$

Auch hier ist es wieder wichtig, die beiden Grenzwerte unabhängig voneinander zu nehmen.

6.3.2 Ein paar wichtige Stammfunktionen und Integrale

Ohne Beweis und Diskussion sind hier ein paar Stammfunktionen und Integrale gegeben, die in der Physik wichtig sind:

$f(x) = x^n \Rightarrow F(x) = \dfrac{1}{n+1}\, x^{n+1} \quad (n \neq -1)$	$f(x) = \dfrac{1}{x} = x^{-1} \Rightarrow F(x) = \ln(\lvert x\rvert)$
$f(x) = \sin(x) \Rightarrow F(x) = -\cos(x)$	$f(x) = \cos(x) \Rightarrow F(x) = \sin(x)$
$f(x) = e^{\alpha\, x} \Rightarrow F(x) = \dfrac{1}{\alpha} e^{\alpha x}$	$f(x) = \dfrac{1}{\sqrt{x}} \Rightarrow F(x) = 2\sqrt{x}$

Darüber hinaus gibt es noch einige Integrale, deren Wert man zwar bestimmen kann (die also eine wohldefinierte Fläche unter der Kurve haben), bei denen man aber die Stammfunktion nicht als elementare Funktion schreiben kann. In der Physik sind hier besonders wichtig:

- Gamma-Funktion $\Gamma(x)$:
$$\Gamma(x) = \int_0^\infty dt\, t^{x-1}\, e^{-t}.$$
Man kann zeigen, dass $\Gamma(x+1) = x!$ für $x \in \mathbb{Z}$ gilt.
- Fehler-Funktion $\mathrm{erf}(x)$:
$$\mathrm{erf}(x) = \frac{2}{\sqrt{\pi}} \int_0^x dt\, e^{-t^2}.$$

Da sie oft vorkommen sind diese beide Funktionen tabelliert und im Computer vorprogrammiert.

6.3.3 Grundlegende Rechenregeln und Rechentricks für Integrale

Folgende Rechenregeln und -tricks erleichtern das Rechnen mit Integralen:

1. **Integrale sind linear**: für (reelle) Zahlen c_1 und c_2 und Funktionen $f(x)$ und $g(x)$ gilt:
$$\int_a^b dx\ (c_1\, f(x) + c_2 g(x)) = c_1 \int_a^b dx\, f(x) + c_2 \int_a^b dx\, g(x). \tag{6.11}$$
2. **Integrationsintervalle können zerlegt werden**: für (reelle) Zahlen a, b und c gilt
$$\int_a^b dx\, f(x) = \int_a^c dx\, f(x) + \int_c^b dx\, f(x). \tag{6.12}$$

3. **Vertauschen der Grenzen erzeugt ein Minuszeichen**:

$$\int_a^b dx\, f(x) = -\int_b^a dx\, f(x). \tag{6.13}$$

4. **Man kann nach Integralgrenzen ableiten**: unter Nutzung von Gleichung (6.6) folgt:

$$\frac{d}{db}\int_a^b dx\, f(x) = \frac{d}{db}\left(F(b) - F(a)\right) = f(b), \tag{6.14}$$

$$\frac{d}{da}\int_a^b dx\, f(x) = \frac{d}{da}\left(F(b) - F(a)\right) = -f(a). \tag{6.15}$$

   Noch allgemeiner gilt für Grenzen $a(y)$ und $b(y)$, die selbst noch von einem variablen Parameter y abhängen mit der Kettenregel aus Gleichung (5.10):

$$\frac{d}{dy}\int_{a(y)}^{b(y)} dx\, f(x) = \frac{d}{dy}\left(F(b(y)) - F(a(y))\right) = f(b(y))\, b'(y) - f(a(y))\, a'(y) \tag{6.16}$$

5. **Symmetrien erleichtern das integrieren**: Wenn die Funktion $f(x)$ achsensymmetrisch ("gerade") oder punksymmetrisch zum Ursprung ("ungerade") ist, so gilt für die Integration über ein symmetrisches Intervall $[-a, a]$:

$$f(x) = +f(-x) \quad \Rightarrow \quad \int_{-a}^{a} dx\, f(x) = 2\int_0^a dx\, f(x), \tag{6.17}$$

$$f(x) = -f(-x) \quad \Rightarrow \quad \int_{-a}^{a} dx\, f(x) = 0. \tag{6.18}$$

6.3.4 Variablensubstitution

Eine weitere wichtige Integrationstechnik ist die der **Variablensubstitution**, bei der man die Integrationsvariable x durch eine andere Integrationsvariable ersetzt. Diese ist im Allgemeinen eine Funktion der ursprünglichen Variablen. Sei zum Beispiel

$$x = g(u). \tag{6.19}$$

Wir wollen nun aus dem Integral über x ein Integral über u machen. Dabei muss man folgende Punkte beachten:

- **Die Grenzen ändern sich** wie folgt: $x = a \leftrightarrow u = g^{-1}(a)$ und $x = b \leftrightarrow u = g^{-1}(b)$.
- **Das Differenzial ändert sich**: die Definition des Integrals basiert auf Rechtecken der Breite Δx, welche immer kleiner wird. Das übersetzt sich laut Taylor-Entwicklung wie folgt in eine Änderung von u:

$$\Delta x = x_{i+1} - x_i \leftrightarrow g(u_{i+1}) - g(u_i) = g(u_i + \Delta u) - g(u_i) \approx \frac{dg(u_i)}{du}\,\Delta u.$$

  Dies entspricht gerade dem Differential von $x = g(u)$ aus Gleichung (5.3) (wie es sein sollte):

$$x = g(u) \quad \Rightarrow \quad dx = \frac{dg(u)}{du}\, du.$$

Wir finden also, dass sich das Integral wie folgt ändert:

$$x = g(u) \quad \Rightarrow \quad \int_a^b dx\, f(x) = \int_{g^{-1}(a)}^{g^{-1}(b)} du\, \frac{dg(u)}{du}\, f(g(u)). \tag{6.20}$$

Als Beispiel betrachten wir folgendes Integral:

$$\int_2^3 dx\, \frac{1}{5\,x - 7} = ?$$

Wir setzten hier

$$u = 5\,x - 7 = g^{-1}(x) \quad \Rightarrow \quad u + 7 = 5\,x \quad \Rightarrow \quad x = \frac{u+7}{5} \quad \Rightarrow \quad g(u) = \frac{u+7}{5}.$$

Darauf folgt, dass

$$\frac{dg(u)}{du} = \frac{1}{5}.$$

Da $g^{-1}(2) = 3$ und $g^{-1}(3) = 8$ folgt aus Gleichung (6.20):

$$\int_2^3 dx\, \frac{1}{5\,x-7} = \int_3^8 du\, \frac{1}{5}\,\frac{1}{u} = \frac{1}{5}\int_3^8 du\, \frac{1}{u} = \frac{1}{5}\,\left[\ln(|u|)\right]_3^8 = \frac{1}{5}\,(\ln(8) - \ln(3)) = \frac{1}{5}\,\ln\left(\frac{8}{3}\right).$$

Konkret kann man aus der allgemeinen Substitutionsregel zum Beispiel folgende beiden hilfreichen Rechentricks herleiten:

1. **Verschiebetrick**: hiermit kann man additive Konstanten im Integranden loswerden:

$$\int_a^b dx\, f(x + x_0) = \int_{a+x_0}^{b+x_0} du\, f(u) \quad (\text{mit } x = g(u) = u - x_0). \tag{6.21}$$

2. **Reskalierungstrick**: hiermit kann man multiplikative Konstanten im Integranden loswerden:

$$\int_a^b dx\, f(\lambda\, x) = \int_{\lambda\, a}^{\lambda\, b} du\, \frac{1}{\lambda} f(u) \quad (\text{mit } x = g(u) = \frac{u}{\lambda}). \tag{6.22}$$

6.3.5 Partielle Integration

Ein weiterer vielgenutzter Rechentrick ist die partielle Integration, die im Wesentlichen auf der Produktregel basiert (siehe Gleichung (5.8)). Seien $f(x)$ und $g(x)$ zwei differenzierbare Funktionen. Dann gilt laut Produktregel:

$$\frac{d}{dx}\,(f(x)\,g(x)) = \frac{df(x)}{dx}\,g(x) + f(x)\,\frac{dg(x)}{dx}.$$

Wenn man jetzt noch den Hauptsatz der Integral- und Differentialrechnung in Gleichung (6.7) dazunimmt, erhält man die Formel der **partiellen Integration**:

$$\int_a^b dx\, \frac{df(x)}{dx}\, g(x) = \left[f(x)\,g(x)\right]_a^b - \int_a^b dx\, f(x)\, \frac{dg(x)}{dx}. \tag{6.23}$$

Die partielle Integration erlaubt es also, Ableitungen zwischen verschiedenen Funktionen im Integranden hin- und herzuschieben. Das ist manchmal hilfreich, denn Ableitungen und Stammfunktionen von bestimmten Funktionen sind besonders einfach.

Rechenaufgabe:

Bestimmen Sie folgendes Integral mit Hilfe partieller Integration.

$$\int_a^b dx\, x\, e^x =$$

Tipp: der Integrand hat zwei Teile. Wenn Sie den "richtigen" davon als Ableitung verstehen, wird das Integral mit partieller Integration einfacher. Wenn Sie den "falschen" als Ableitung verstehen, wird das Integral schwerer. 50-50-Chance! (Idealerweise nutzen Sie Ihre mathematische Intuition, um die Chance zu erhöhen.)

6.4 Bestimmtes und unbestimmtes Integral am Beispiel von einfachen Integralen von skalaren Funktionen mehrerer Variablen

Bisher haben wir nur Integrale von Funktionen einer Variablen, z. B. $f(x)$, betrachtet. Im folgenden wollen wir auch Integrale von Funktionen mehrerer Variablen betrachten. Dazu betrachten wir die Funktion

$$f : \mathbb{R}^n \to \mathbb{R} \quad , \ \vec{x} \mapsto f(x_1, \ldots, x_n).$$

Für diese Funktion kann man mehrere Arten von "einfachen" Integralen definieren. Beachten Sie, dass hierbei n nicht genauer spezifiziert ist – alles, was wir in diesem Abschnitt sagen, gilt also auch für Integrale von Funktionen nur einer Variablen (also $\int dx\, f(x)$): in beiden Fällen unterscheiden wir bestimmte und unbestimmte Integrale.

6.4.1 Unbestimmtes Integral über eine der Variablen

Das unbestimmte Integral über die Variable x_i (mit $i \in \{1, \ldots, n\}$) ist definiert als

$$\int dx_i\, f(x_1, \ldots, x_n) = F_i(x_1, \ldots, x_n) \quad \text{mit} \quad \frac{d}{dx_i} F_i(x_1, \ldots, x_n) = f(x_1, \ldots, x_n). \tag{6.24}$$

Hierbei ist $F_i(x_1, \ldots, x_n)$ eine Stammfunktion von f bezüglich der Integration nach x_i. Wie die Stammfunktion aus Kapitel 6.2 ist sie nicht eindeutig. Betrachten wir zum Beispiel $\widetilde{F}_i(x_1, \ldots, x_n) = F_i(x_1, \ldots, x_n) + C(x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n)$, wobei $C(x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n)$ eine beliebige Funktion aller Variablen außer x_i ist, so gilt:

$$\frac{d}{dx_i}F_i(x_1,\ldots,x_n) = f(x_1,\ldots,x_n) \quad \Rightarrow \quad \frac{d}{dx_i}\widetilde{F}_i(x_1,\ldots,x_n) = f(x_1,\ldots,x_n).$$

6.4.2 Bestimmtes Integral über eine der Variablen

Hier gilt analog zum bestimmten Integral über nur eine Variable:

$$\int_a^b dx_i\, f(x_1,\ldots,x_n) = F_i(x_1,\ldots,x_{i-1},b,x_{i+1},\ldots,x_n) - F_i(x_1,\ldots,x_{i-1},a,x_{i+1},\ldots,x_n), \tag{6.25}$$

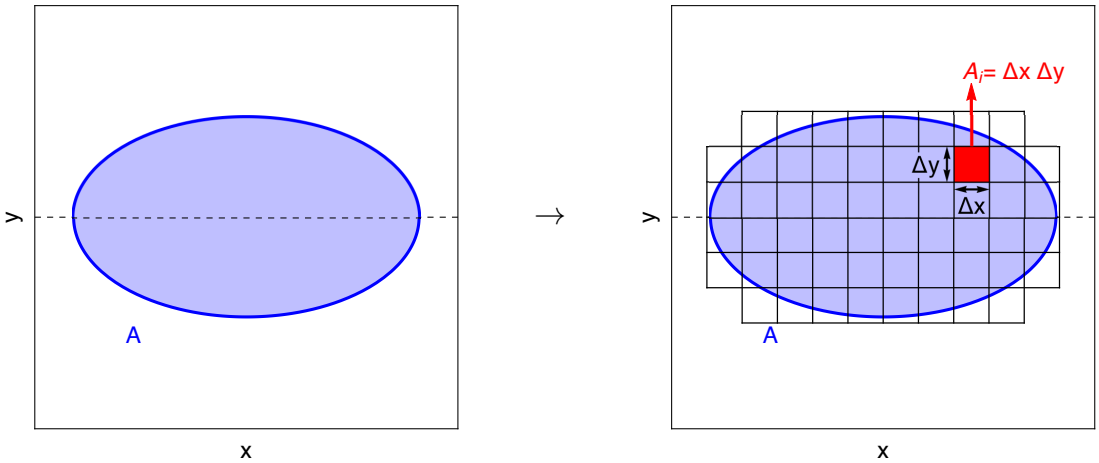
wobei F_i wieder eine Stammfunktion bezüglich x_i ist, $\frac{d}{dx_i}F_i(x_1,\ldots,x_n) = f(x_1,\ldots,x_n)$.

6.5 Mehrfachintegrale

Zuletzt definieren wir noch Mehrfachintegrale.

6.5.1 Flächenintegrale

Der einfachste Fall ist der des **Flächenintegrals**. Die Motivation für das Flächenintegral ist die Berechnung des Flächeninhalts A, der von einer beliebigen Kurve eingeschlossen wird. Auch hier kann man den Flächeninhalt wieder als Summe der Flächeninhalte kleiner Rechtecke, in diesem Fall von kleinen Quadraten, verstehen. Wir wählen hier die Konvention, dass alle Quadrate gezählt werden, die noch ein Stück der zu berechnenenden Fläche einschließen.



$$A \approx \sum_i A_i = \sum_i \Delta x\, \Delta y. \tag{6.26}$$

Der Flächeninhalt innerhalb der Kurve ergibt sich wieder durch den Grenzfall unendlich vieler unendlich kleiner Rechtecke,

$$A = \int_{\text{umrandete Fläche}} dA = \lim_{\Delta x\to 0}\lim_{\Delta y\to 0}\sum_i A_i. \tag{6.27}$$

Flächenintegrale werden mit verschiedenen, äquivalenten Symbolen bezeichnet. Wenn wir die beiden Koordinaten noch in einen Vektor $\vec{r} = \begin{pmatrix} x \\ y \end{pmatrix}$ zusammenfassen, gibt es folgende Notationen:

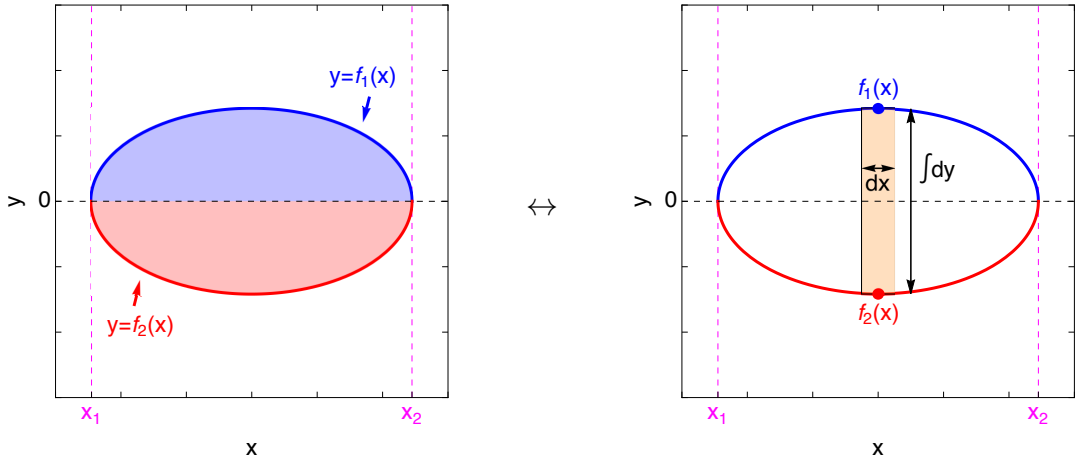
$$\int dA = \int dx \int dy = \int\int dx\, dy = \int d^2r. \tag{6.28}$$

Das Differential d^2r zeigt an, dass man ein zweidimensionales Integral über die Koordinaten, die im Vektor $\vec{r}$ zusammengefasst werden, berechnet.

6.5.1.1 Praktische Berechnung des Flächenintegrals

Die Angabe des Flächenintegrals als Grenzwert von kleinen Quadraten ist zwar mathematisch schön, praktisch aber nicht besonders hilfreich. Wie rechnet man ein Flächenintegral jetzt genau aus? Bei "normalen" Integralen gab es einen ähnlichen Grenzwert von dünnen Rechtecke, den wir in der Praxis aber nie auswerten. Stattdessen nutzt man den Hauptsatz der Integral- und Differentialrechnung und drückt das Integral mit einer Stammfunktion aus.

Flächenintegrale berechnen wir in der Praxis auch nicht über den Grenzwert kleiner Quadrate, sondern führen das Flächenintegral auf "normale" Integrale zurück. Hierzu müssen wir zunächst die Fläche mathematisch beschreiben, deren Flächeninhalt wir berechnen wollen (wir müssen die Fläche "parametrisieren"). Zu diesem Zweck definieren wir zwei Funktionen $f_1(x)$ und $f_2(x)$, so, dass $y = f_1(x)$ die obere Grenze des berandeten Gebietes ist, und $y = f_2(x)$ die untere Grenze des Gebiets:



Die zwei x-Werte bei denen die obere und untere Begrenzung zusammenkommen nennen wir x_1 und x_2. Offensichtlich ist der Flächeninhalt der berandeten Fläche zusammengesetzt aus dem Flächeninhalt zwischen der blauen Funktion $f_1(x)$ und der x-Achse, und dem der Fläche zwischen der roten Funktion $f_2(x)$und der x-Achse. Wir finden also:

$$A = \int_{x_1}^{x_2} dx\, f_1(x) - \int_{x_1}^{x_2} dx\, f_2(x) = \int_{x_1}^{x_2} dx\, (f_1(x) - f_2(x)) = \int_{x_1}^{x_2} dx \int_{f_2(x)}^{f_1(x)} dy. \tag{6.29}$$

Die letzte Gleichung ist eine verschachtelte Form von Integralen, die wir schon kennen:

$$A = \int_{x_1}^{x_2} dx\, h(x) \quad \text{mit} \quad h(x) = \int_{f_2(x)}^{f_1(x)} dy\, g(y) \quad \text{und} \quad g(y) = 1. \tag{6.30}$$

Diese Schreibweise des Integrals besagt, dass wir für jeden x-Wert zwischen x_1 und x_2 den Flächeninhalt eines Rechtecks der Breite dx und Höhe $h(x) = f_1(x) - f_2(x)$ aufsummieren. Diese Höhe wiederum schreiben wir als das Integral von $g(y) = 1$ über die y-Koordinate zwischen den beiden Berandungsfunktionen (das ist eben einfach eine äquivalente Schreibweise!). Gleichung (6.29) ergibt immer den Flächeninhalt – auch dann, wenn die Fläche nicht so schön symmetrisch um $y = 0$ liegt, oder wenn die umrandete Fläche anders geformt ist. Die Definitionen des Flächenintegrals in Gleichungen (6.27) und (6.29) sind äquivalent. Die erste Form ist mathematisch eleganter, aber in der Praxis nutzlos. Die zweite Form erlaubt es, einen Flächeninhalt explizit zu berechnen.

Rechenaufgabe:

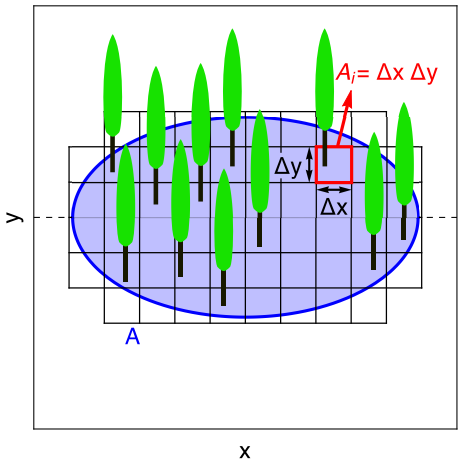
Der Einheitskreis ist durch die Gleichung $x^2 + y^2 = 1$ definiert.

1. Finden Sie die Funktionen $y = f_1(x)$ und $y = f_2(x)$, die den Kreis oben und unten beranden. Was sind die x-Werte, bei denen diese beiden Funktionen zusammenstoßen?

2. Nutzen Sie $\int_0^1 dx\, \sqrt{1 - x^2} = \frac{\pi}{4}$, um den Flächeninhalt des Einheitskreises zu berechnen.

6.5.1.2 Flächenintegrale von Funktionen mehrerer Variablen

Die Berechnung von Flächeninhalten ist tatsächlich nicht der Hauptnutzen von Flächenintegralen. Vielmehr nutzen wir diese in der Physik dazu, um Funktionen über Flächen zu integrieren. Dies erlaubt uns zum Beispiel, aus einer (Flächen-)dichte eine Gesamtzahl zu berechnen. Betrachten wir als Beispiel hierzu ein Stück Wald und fragen uns, wie viele Bäume dort wohl stehen. In einem einfachen Modell könnte man annehmen, dass die Bäume immer genau in 4 Metern Abstand stehen. Das stimmt aber im echten Wald sicher nicht: dort stehen die Bäume an manchen Orten dichter als an anderen.



Um diese Unterschiede im Baumbestand zu berücksichtigen, können wir zunächst eine Baumdichte einführen. Hierzu unterteilen wir unser Waldstück wieder in kleine Quadrate der Fläche $A_i = \Delta x\, \Delta y$. In einem gegebenen Quadrat haben wir dann eine Baumdichte $\rho_{\text{Baum}}(\text{Quadrat } i)$, die wie folgt definiert ist:

$$\rho_{\text{Baum}}(\text{Quadrat } i) = \frac{\text{Anzahl der Bäume im Quadrat } i}{\text{Fläche des Quadrats } i} = \frac{\text{Anzahl der Bäume im Quadrat } i}{\Delta x\, \Delta y}. \tag{6.31}$$

Im konkret abgebildeten Beispiel ist die Baumdichte entweder $\rho_{\text{Baum}} = 0$ (wenn kein Baum im fraglichen Quadrat steht), oder $\rho_{\text{Baum}} = \frac{1\,\text{Baum}}{\Delta x\, \Delta y}$ (wenn ein Baum im fraglichen Quadrat steht). Die Gesamtzahl der Bäume ist dann gegeben durch

$$\begin{aligned}
\text{Anzahl Bäume} &= \sum_{\text{Quadrate } i} \text{Bäume in Quadrat } i \\
&= \sum_{\text{Quadrate } i} \text{Bäume pro Fläche} \cdot \text{Fläche} \\
&= \sum_{\text{Quadrate } i} \text{Baum-Dichte} \cdot \text{Fläche} \\
&= \sum_{\text{Quadrate } i} \rho_{\text{Baum}}(\text{Quadrat } i)\, A_i = \sum_{\text{Quadrate } i} \rho_{\text{Baum}}(\text{Quadrat } i)\, \Delta x\, \Delta y.
\end{aligned}$$

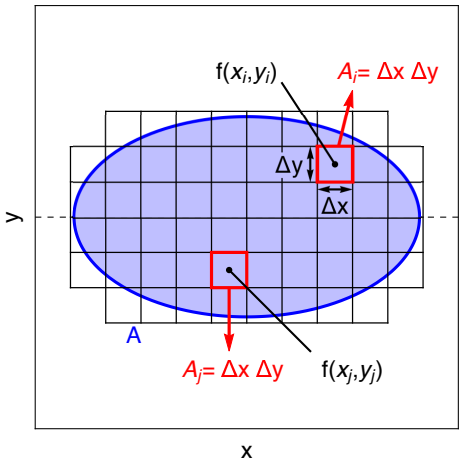
Um jetzt von der durch Quadrate genäherten Fläche auf die echte Waldfläche zu kommen, müssen wir wieder den Grenzwert $\Delta x \to 0$ und $\Delta y \to 0$ nehmen, und gehen von der Summe zum Integral über. Für die Baumdichte heißt das, dass wir von der Funktion $\rho_{\text{Baum}}(\text{Quadrat } i)$ zu einer Funktion der kontinuierlichen Ortskoordinaten $\rho_{\text{Baum}}(x, y)$ übergehen. Diese ist wie folgt definiert

1. $\rho_{\text{Baum}}(x, y) = 0$ falls der Ort mit Koordinaten (x, y) nicht unter einem Baum liegt.
2. $\rho_{\text{Baum}}(x, y) = \frac{1}{\text{Fläche direkt unterhalb des Baums}}$, falls der Ort mit Koordinaten (x, y) unter einem Baum liegt.

Das macht Sinn: $\rho_{\text{Baum}}(x, y)$ ist wieder von der Form $\rho_{\text{Baum}}(x, y) = \frac{\text{Anzahl von Bäumen}}{\text{Fläche}}$ (eben für genau einen Baum, Anzahl von Bäumen $= 1$, und die Fläche unter diesem einen Baum). Diese Funktion misst also den Anteil eines Baumes, der über einem bestimmten Ort liegt. Die Gesamtzahl von Bäumen ist dann durch

$$\text{Anzahl Bäume} = \int dA\, \rho_{\text{Baum}}(x, y) \tag{6.32}$$

gegeben.



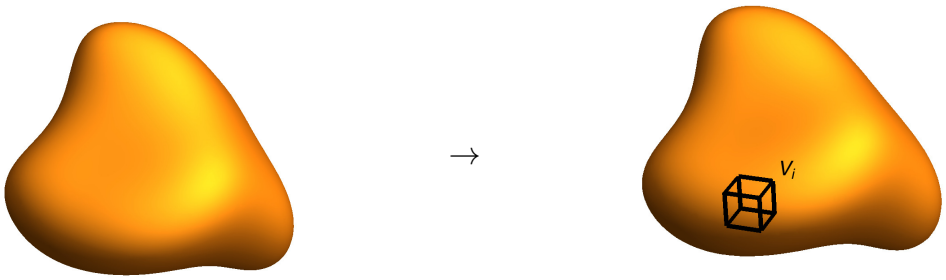
Dieses Prinzip verallgemeinern wir jetzt zum Integral einer beliebigen Funktion $f(x, y)$ über eine Fläche A. Dieses hat folgende äquivalente Notationen:

$$\int\int_A dx\,dy\,f(x,y) = \int\int_A f(x,y)\,dx\,dy = \int dx \int dy\,f(x,y) = \lim_{\Delta x\to 0}\lim_{\Delta y\to 0}\sum_i A_i\,f(x_i,y_i), \tag{6.33}$$

wobei i die kleinen Quadrate durchnummeriert, in die wir die Fläche unterteilt haben (diese haben den Flächeninhalt $A_i = \Delta x\,\Delta y$), und (x_i, y_i) ist ein (beliebiger) Punkt im Quadrat i.

6.5.2 Volumenintegrale von Funktionen mehrerer Variablen

Die Verallgemeinerung vom Flächenintegral zum Volumenintegral ist jetzt ein mathematisch ganz natürlicher Schritt. Ausgangspunkt ist die Frage: Was ist das Volumen eines seltsam geformten Körpers? Ein Quader hat das Volumen Höhe mal Breite mal Tiefe, aber wie berechnet man das Volumen eines "Blobs"?



Die Idee des Volumenintegrals ist wieder ganz einfach: man unterteile den Blob in kleine Würfel mit Kantenlängen Δx in x-Richtung, Δy in y-Richtung und Δz in z-Richtung. Der i-te dieser kleinen Würfel hat dann das Volumen $V_i = \Delta x\,\Delta y\,\Delta z$. Das Gesamtvolumen des Blobs ist dann näherungsweise gegeben durch

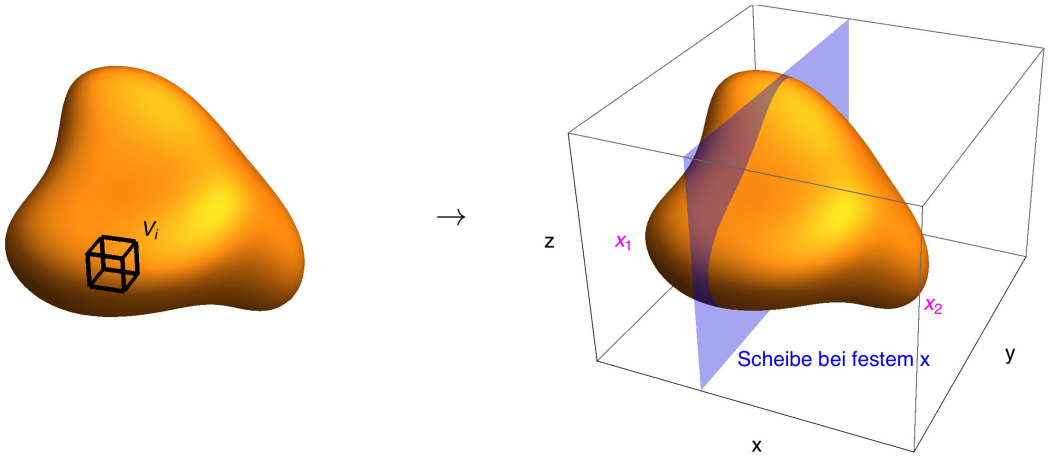
$$V \approx \sum_i V_i = \sum_i \Delta x\,\Delta y\,\Delta z. \tag{6.34}$$

Das echte Volumen des Blobs erhalten wir nun wieder im Grenzwert unendlich vieler, unendlich kleiner Blobs. Dieser Grenzwert definiert das **Volumenintegral**:

$$V = \int_{\text{Volumen des Blobs}} dV = \lim_{\Delta x \to 0} \lim_{\Delta y \to 0} \lim_{\Delta z \to 0} \sum_i V_i. \tag{6.35}$$

Auch Volumenintegrale werden mit verschiedenen, äquivalenten Symbolen bezeichnet. Wenn wir die Koordinaten noch in einen Vektor $\vec{r} = \begin{pmatrix} x \\ y \\ z \end{pmatrix}$ zusammenfassen, gibt es folgende Notationen:

$$\int dV = \int dx \int dy \int dz = \int\int\int dx\, dy\, dz = \int d^3r. \tag{6.36}$$



In der Praxis berechnen wir Volumenintegrale wieder dadurch, dass wir sie auf die uns bekannten eindimensionalen Integrale zurückführen. Dazu schneiden wir den Blob in infinitesimal dünne (y, z)-Scheiben (also Scheiben mit festem x). Diese Scheiben haben eine Dicke dx. Da der Blob ein linkes und rechtes Ende hat, gibt es solche Scheiben nur für x-Werte zwischen x_1 und x_2. Das Volumen des Blobs ist dann das Integral der Volumen $dV = dx\, A(x)$ der Scheiben, wobei $A(x)$ der Flächeninhalt der Scheibe bei x ist:

$$V = \int dV = \int_{x_1}^{x_2} dx\, A(x). \tag{6.37}$$

Zur Berechnung von $A(x)$ nutzen wir das Flächenintegral aus Gleichung (6.29), und erhalten:

$$V = \int_{x_1}^{x_2} dx \int_{f_2(x)}^{f_1(x)} dy \int_{h_2(x,y)}^{h_1(x,y)} dz. \tag{6.38}$$

Hierbei sind $y_1 = f_1(x)$ und $y_2 = f_2(x)$ die minimalen und maximalen y-Werte der Scheibe bei x, und $z_1 = h_1(x, y)$ sowie $z_2 = h_2(x, y)$ die minimalen und maximalen z-Werte der Scheibe bei x am y-Wert y. Diese Integrale berechnen wir dann wiederum über den Hauptsatz der Integral- und Differentialrechung mittels Stammfunktionen. Und wie bei Flächenintegralen liegt der Hauptnutzen von Volumenintegralen nicht im Bestimmen von Volumina, sondern im Integrieren von Funktionen über Volumen. Hier gilt analog zum Integral über Flächen:

$$\int dV\, f(x,y,z) = \int dx \int dy \int dz\, f(x,y,z) = \int\int\int dx\, dy\, dz\, f(x,y,z) = \int d^3r\, f(x,y,z)$$
$$= \lim_{\Delta x\to 0}\lim_{\Delta y\to 0}\lim_{\Delta z\to 0}\sum_i V_i\, f(x_i,y_i,z_i)$$
$$= \lim_{\Delta x\to 0}\lim_{\Delta y\to 0}\lim_{\Delta z\to 0}\sum_i \Delta x\,\Delta y\,\Delta z\, f(x_i,y_i,z_i),$$

wobei $f(x_i, y_i, z_i)$ der Funktionswert an einem (beliebigen) Punkt (x_i, y_i, z_i) im i-ten Würfel ist.

6.5.3 Allgemeine Mehrfachintegrale von Funktionen mehrerer Variablen

Im Allgemeinen betrachtet man natürlich Funktionen von vielen Variablen $x_1, x_2, \ldots, x_n$. Man kann dann auch über alle (oder einen Teil) dieser Variablen integrieren. Formal ist das einfach die Erweiterung des Volumenintegrals;

$$\int d^n r\, f(x_1,\ldots,x_n) = \int dx_1 \ldots \int dx_n\, f(x_1,\ldots,x_n) = \left(\lim_{\Delta x_i\to 0}\Delta x_i\right) f(x_{i,1},\ldots,x_{i,n}), \quad (6.39)$$

wobei $x_{i,j}$ die j-te Komponente eines Vektors $\vec{r}_i$ bezeichnet, der an einen Punkt innerhalb eines Hyperkubus i mit Volumen $\lim_{\Delta x_i\to 0}\Delta x_i$ zeigt.

6.5.4 Reihenfolge und Vertauschung (oder nicht!) von Integralen

Im Allgemeinen darf man die Reihenfolge, in der Integrale ausgeführt werden, nicht verändern. Die Reihenfolge entspricht dabei der Reihenfolge der Integrale **von rechts nach links gelesen**. Beim Flächenintegral

$$I = \int_{x_1}^{x_2} dx \int_{f_2(x)}^{f_1(x)} dy\, f(x,y) \quad (6.40)$$

zum Beispiel führt man zuerst das y-Integral aus, und dann das x-Integral. Betrachten wir folgendes einfaches Beispiel:

$$\int_0^1 dx \int_0^x dy = \int_0^1 dx\,[y]_0^x = \int_0^1 dx\,(x-0) = \int_0^1 dx\, x = \left[\frac{1}{2}x^2\right]_0^1 = \left(\frac{1}{2}1^2 - \frac{1}{2}0^2\right) = \frac{1}{2}.$$

Die Reihenfolge der Integrale ist hier essentiell! Wenn wir hier die Integrale vertauschen, erhalten wir ein anderes Ergebnis:

$$\int_0^x dy \int_0^1 dx = \int_0^x dy\,[x]_0^1 = \int_0^x dy\,(1-0) = \int_0^x dy\, 1 = [y]_0^x = (x-0) = x.$$

Doch nicht nur hier, sondern ganz allgemein gilt: **Integrale dürfen im Allgemeinen nicht vertauscht werden!** Das heißt aber nicht, dass man Integrale nie vertauschen kann. In der "Physiker-Praxis" ist das sogar sehr oft möglich: Integrale können sicher vertauscht werden, wenn sie feste Grenzen haben und der Integrand stetig ist. Dann gilt:

$$\int_a^b dx \int_c^d dy\, f(x,y) = \int_c^d dy \int_a^b dx\, f(x,y). \tag{6.41}$$

Noch mehr Details über die Vertauschbarkeit von Integralen liefert der "Satz von Fubini". Hierzu seien Sie an eines der vielen guten (Analysis-)Lehrbücher verwiesen.

6.5.5 Integrale von vektorwertigen Funktionen

In kartesischen Koordinaten geschieht die **Integration von vektorwertigen Funktionen komponentenweise**. Betrachten wir der Konkretheit halber das Beispiel einer m-komponentigen Funktion $\vec{f}$,

$$\mathbb{R}^n \to \mathbb{R}^m\ , \quad \vec{r} \mapsto \vec{f}(\vec{r}) = \sum_{j=1}^{m} \vec{e}_j\, f_j(\vec{r}). \tag{6.42}$$

Diese wollen wir jetzt über die n Koordinaten integrieren. Dann gilt:

$$\int d^n r\, \vec{f}(\vec{r}) = \sum_{j=1}^{m} \vec{e}_j \int d^n r\, f_j(\vec{r}). \tag{6.43}$$

Das wäre im Physikerleben gut zu wissen:

- Die Bedeutung eines Integrals als Fläche erklären können.
- Wissen, was eine Stammfunktion ist.
- Den Hauptsatz der Differential- und Integralrechnung kennen und seine Bedeutung erklären können.
- Sagen können, was bestimmte und unbestimmte Integrale sind.
- Einige elementare Stammfunktionen kennen (Exponentialfunktion, Polynome, $1/x$, Sinus, Kosinus).
- Wissen, wie uneigentliche Integrale genau definiert sind (mit Grenzwerten!).
- Die Rechentricks aus Kapitel 6.3.3 kennen und anwenden können.
- Variabelsubstitution anwenden können.
- Partielle Integration anwenden können.
- Wissen, wie man Mehrfachintegrale (insbesondere Flächenintegrale und Volumenintegrale) definiert und berechnet.
- Wissen, wann man Integrationsvariablen vertauschen darf.

PS: "gut zu wissen" muss nicht heißen, dass das alles in der Klausur abgefragt wird – oder, dass sonst nichts aus diesem Kapitel in der Klausur vorkommt. Wer aber mit den Punkten in diesem Kasten nichts anfangen kann, hat wahrscheinlich in der Klausur und im weiteren Studium Probleme.

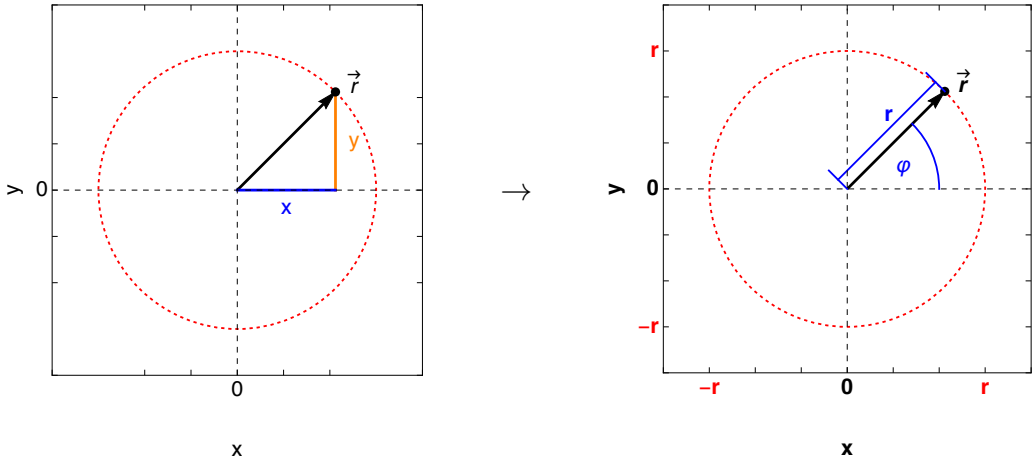
Kapitel 7

Krummlinige Koordinaten und Koordinatenwechsel im Integral

Um die physikalischen Eigenschaften eines gegebenen Systems zu beschreiben, ist es immer hilfreich, die Beschreibung an das System anzupassen. Je besser die "Sprache", in der das System beschrieben wird, zum System passt, desto einfacher ist es in der Praxis, physikalische Fragen für das Problem zu beantworten. Es lohnt sich also auch, das jeweils am besten passende Koordinatensystem zu suchen, und das Problem dann in diesen Koordinaten zu beschreiben. Eine besonders hilfreiche Klasse von Koordinatensystemen sind die sogenannten "krummlinigen Koordinaten". Die wichtigsten Vertreter dieser Klasse sind Polarkoordintaten, Zylinderkoordinaten und Kugelkoordinaten.

7.1 Polarkoordinaten

In der Physik versucht man, Probleme so gut es geht in ein paar grundlegenden Koordinatensystemen zu beschreiben. Wir haben bisher hauptsächlich kartesische Koordinaten genutzt. In zwei Dimensionen waren das die Koordinaten x und y. Oftmals sind physikalische Systeme aber rotationssymmetrisch. In solchen Fällen ist es hilfreich, von kartesischen Koordinaten zu sogenannten **Polarkoordinaten** überzugehen.

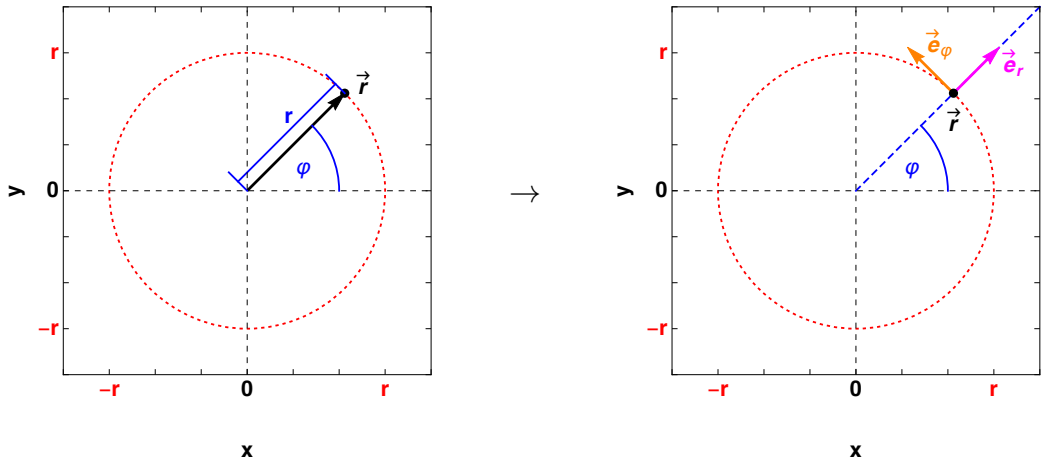


Diese drücken die x- und y-Koordinaten über den Abstand $r = |\vec{r}| = \sqrt{x^2 + y^2}$ aus, den ein Punkts mit dem Ortsvektor $\vec{r}$ zum Ursprung hat, sowie über den Sinus und Kosinus des Winkels φ, den der Ortsvektor $\vec{r}$ mit der x-Achse einschließt.

$$\vec{r} = \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} r\,\cos(\varphi) \\ r\,\sin(\varphi) \end{pmatrix} \tag{7.1}$$

Man kann also jeden Punkt äquivalent durch die Angabe der x- und y-Koordinate angeben, oder durch die Angabe von r und φ. Hierbei ist $r \in [0, \infty[$. Der Winkel φ muss einen beliebigen Winkelbereich der Breite 2π abdecken – eine oft genutzte Konvention ist $\varphi \in [0, 2\pi[$. Für $x = 0 = y$ ist φ unbestimmt.

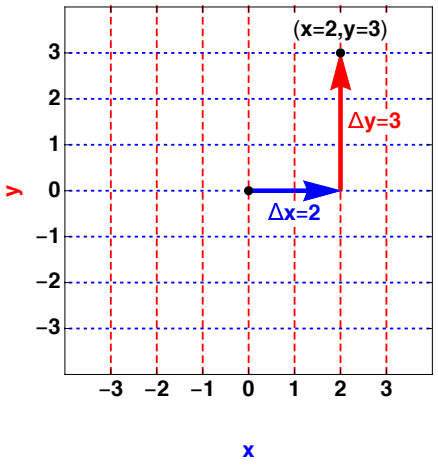
Aber wo liegen die Koordinatenachsen, die zu den neuen Koordinaten (r, φ) gehören? Mathematisch genauer sollten wir fragen; wohin zeigen die neuen Basisvektoren $\vec{e}_r$ und $\vec{e}_\varphi$, die zu unseren neuen Koordinaten gehören?



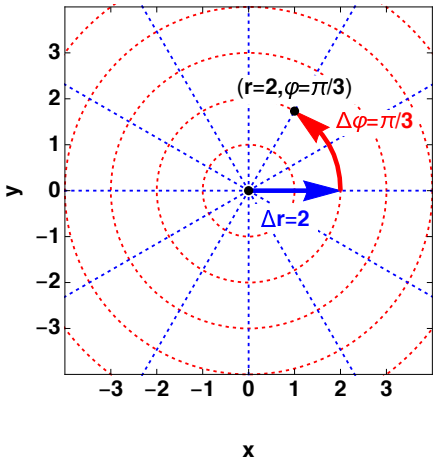
Am einfachsten zu verstehen ist hier $\vec{e}_r$. Dieser Vektor zeigt entlang der Verbindungslinie vom Ursprung zum Punkt $\vec{r}$. **Wichtig:** dies bedeutet, dass der neue Basisvektor $\vec{e}_r$ vom Punkt $\vec{r}$ abhängt, an dem wir ihn anschauen. Der Vektor $\vec{e}_\varphi$ ist ein Vektor, der senkrecht zu $\vec{e}_r$ steht. Somit erhalten wir also wieder ein **orthonormales Koordinatensystem**, wie auch bei kartesischen Koordinaten. Die **Richtung der Basisvektoren (Koordinatenachsen) hängt aber jetzt vom Punkt $\vec{r}$ ab**, an dem wir das Koordinatensystem betrachten! Konkret haben die neuen Basisvektoren die Form

$$\vec{e}_r(r, \varphi) = \begin{pmatrix} \cos(\varphi) \\ \sin(\varphi) \end{pmatrix} \quad \text{und} \quad \vec{e}_\varphi(r, \varphi) = \begin{pmatrix} -\sin(\varphi) \\ \cos(\varphi) \end{pmatrix}. \tag{7.2}$$

Der Einheitsvektor $\vec{e}_r(r, \varphi)$ zeigt in die Richtung, in der die r-Koordinate größer wird, also entlang einer r-Koordinatenlinie (eine Koordinatenlinie ist eine Linie, entlang derer sich nur eine der Koordinaten ändert). Wie der Einheitsvektor selbst muss also auch die Richtung dieser Koordinatenlinie vom Ort abhängen. Das gleiche gilt für die Koordinatenlinien in φ-Richtung. Um uns das bildlich zu veranschaulichen lohnt es sich, sich die Koordinatenlinien im kartesischen System in Erinnerung zu rufen (also mit den "normalen" x- und y-Achsen). Dort finden wir den Punkt mit den Koordinaten $x = 2$ und $y = 3$ dadurch, dass wir vom Ursprung ausgehend zuerst 2 Längeneinheiten ($\Delta x = 2$) entlang einer Koordinatenlinie in x-Richtung gehen, und dann 3 Einheiten entlang der y-Richtung ($\Delta y = 3$):



In Polarkoordinaten ist im Prinzip das nicht anders. Hier sind allerdings die Koordinatenlinien gekrümmt: die Linien der r-Koordinaten gehen radial vom Ursprung nach außen, die Koordinatenlinien in φ-Richtung entsprechen Kreisen (die auf der positiven Hälfte der x-Achse starten und enden). Den Punkt mit Koordinaten $r = 2$ und $\varphi = \pi/3$ finden wir also dadurch, dass wir zuerst entlang der r-Koordinatenlinie zum Winkel $\varphi = 0$ (diese entspricht der positiven Hälfte der x-Achse, die negative Hälfte der x-Achse gehört zum Winkel $\varphi = \pi$ und ist somit die Koodrinatenlinie für die r-Koordinate bei $\varphi = \pi$) eine Strecke $\Delta r = 2$ nach außen gehen, und danach eine Strecke $\Delta\varphi = \pi/3$ entlang der φ-Koordinatenlinie gehen.



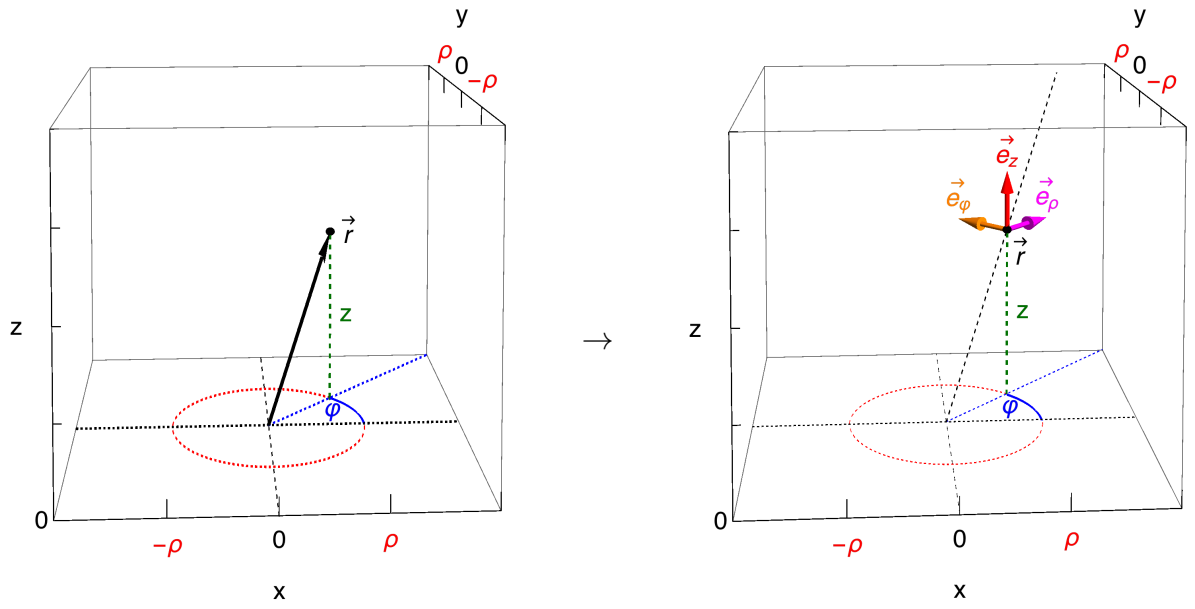
Achtung: da der Winkel φ von der x-Achse "nach oben hin" gemessen wird, müssen wir auf der φ-Koordinatenlinie entgegen dem Uhrzeigersinns gehen.

7.2 Zylinderkoordinaten

Zylinderkoordinaten sind die direkte dreidimensionale Verallgemeinerung von Polarkoordinaten. Sie sind hilfreich, wenn man ein "zylindersymmetrisches" (um eine Achse rotationssymmetrisches) Problem beschreiben will. Beispiele hierfür sind ein stromdurchflossener Draht (der unter anderem ein Magnetfeld erzeugt), oder Wasser, das durch ein Rohr fließt. Die Idee ist einfach, dass man die x und y Komponente des dreidimensionalen Ortstvektors $\vec{r}$ in Polarkoordinaten schreibt, und die z-Komponente so lässt, wie sie ist:

$$\vec{r} = \begin{pmatrix} \rho\,\cos(\varphi) \\ \rho\,\sin(\varphi) \\ z \end{pmatrix}.$$

Zur Beschreibung eines Punktes im Raum geben wir Zylinderkoordinaten also an Stelle der kartesischen Koordinaten (x, y, z) die neuen Koordinaten (ρ, φ, z) an. Hierbei ist $\rho = \sqrt{x^2 + y^2}$ die Länge der Projektion des Vektors $\vec{r}$ in die (x, y)-Ebene, und φ der Winkel, den diese Projektion mit der x-Achse einschließt. Die möglichen Werte dieser Koordinaten sind $\rho \in [0, \infty[$, wohingegen wie bei Polarkoordinaten die übliche Konvention für φ durch $\varphi \in [0, 2\pi[$ gegeben ist.

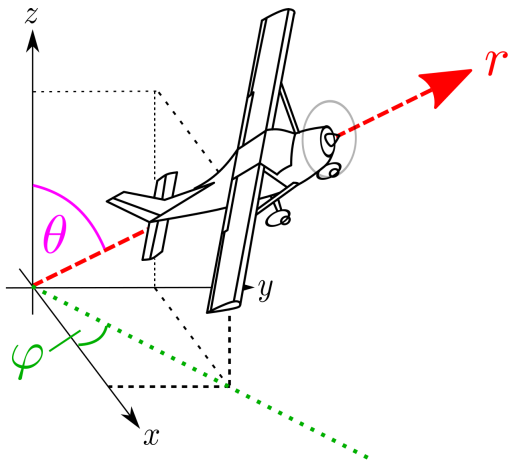


Wie bei Polarkoordinaten hängen die neuen Basisvektoren $\vec{e}_\rho$ und $\vec{e}_\varphi$ vom Punkt ab, an dem wir sie auswerten. Die Basisvektoren bilden ein orthonormales Basissystem und lauten:

$$\vec{e}_\rho(\rho, \varphi, z) = \begin{pmatrix} \cos(\varphi) \\ \sin(\varphi) \\ 0 \end{pmatrix}, \quad \vec{e}_\varphi(\rho, \varphi, z) = \begin{pmatrix} -\sin(\varphi) \\ \cos(\varphi) \\ 0 \end{pmatrix} \quad \text{und} \quad \vec{e}_z(\rho, \varphi, z) = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}. \tag{7.3}$$

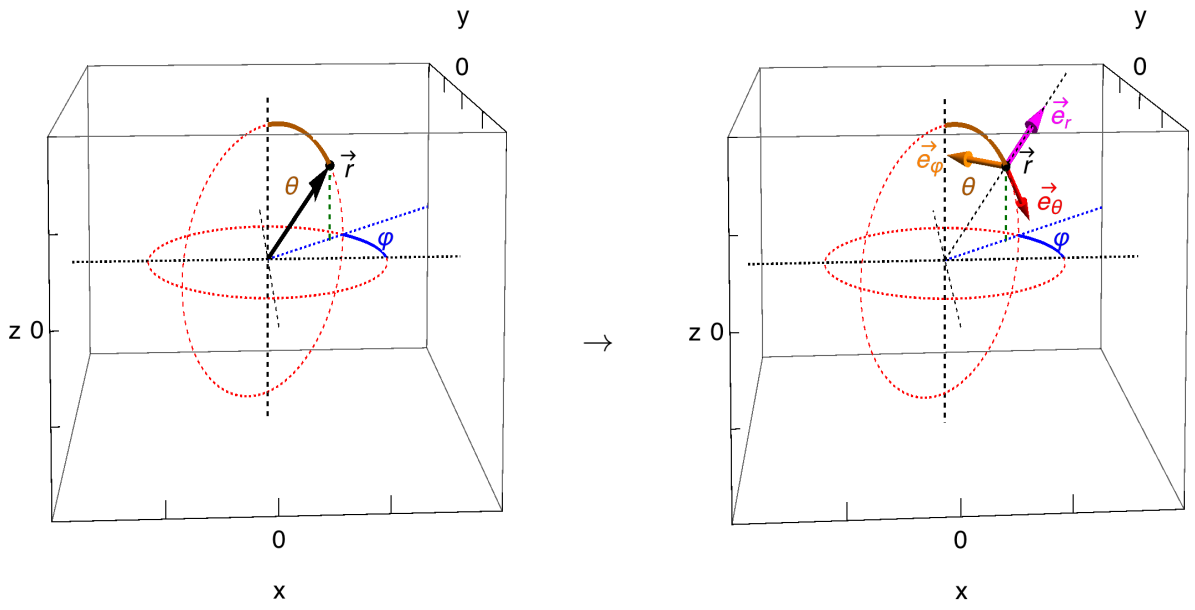
7.3 Kugelkoordinaten

Eine letzte wichtige Sorte von Koordinaten sind die sogenannten Kugelkoordinaten. Diese sind vor allem in der Beschreibung von "kugelsymmetrischen" Problemen hilfreich. Beispiele für solche Probleme sind (in guter Näherung) die Anziehung der Planeten durch die Sonne, oder das elektrische Feld einer Punktladung. Hier gehen wir von den kartesischen Koordinaten (x, y, z) zu Koordinaten (r, θ, φ) über, die wie folgt definiert sind: $r = |\vec{r}| = \sqrt{x^2 + y^2 + z^2}$ ist der Abstand des Punktes mit Orstvektor $\vec{r}$ zum Ursprung, θ ist der Winkel des Ortsvektors mit der z-Achse, und φ der Winkel der Projektion des Ortsvektors in die (x, y)-Ebene (wie bei Zylinderkoordinaten). Betrachten wir diese Koordinaten am Beispiel eines Flugzeugs:



Um in gerader Linie vom Ursprung zu einem beliebigen Punkt zu fliegen, müssen wir den Winkel φ der Flugbahn in der Ebene und den "Anstellwinkel" der Flugbahn θ (aus Konvention gemessen bezüglich der z-Achse) angeben. Diese beiden Winkel bestimmen die Verbindungslinie zwischen dem Ursprung und dem Punkt, zu dem das Flugzeug fliegt. Die letzte fehlende Angabe ist jetzt noch r, die Flugstecke (also der Abstand des Punktes vom Ursprung). Die Angabe der drei Koordinaten (r, θ, φ) legt jeden möglichen Zielpunkt des Flugzeugs eindeutig fest – wir haben also ein weiteres alternatives Koordinatensystem gefunden. Die Wertebereiche dieser Koordinaten sind $r \in [0, \infty[$, $\theta \in [0, \pi]$ und $\varphi \in [0, 2\pi[$ (wieder die üblichen Konventionen, und wieder ist der Winkel φ nicht wohldefiniert wenn $x = y = 0$ ist, wohingegen θ nur für $x = y = z = 0$ nicht wohldefiniert ist). Der Ortsvektor ist dann gegeben durch

$$\vec{r} = \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} r\,\sin(\theta)\,\cos(\varphi) \\ r\,\sin(\theta)\,\sin(\varphi) \\ r\,\cos(\theta) \end{pmatrix} . \tag{7.4}$$



Wie bei Zylinder- und Polarkoordinaten hängen die zugehörigen neuen Basisvektoren $\vec{e}_r$, $\vec{e}_\theta$ und $\vec{e}_\varphi$ vom Punkt ab, an dem wir sie auswerten. Am einfachsten vorzustellen ist dabei der Vektor $\vec{e}_r$, der wie bei Polarkoordinaten entlang der Verbindungslinie vom Ursprung zum Punkt $\vec{r}$ zeigt. Die beiden anderen Vektoren sind senkrecht zu $\vec{e}_r$, und auch senkrecht aufeinander, so dass wir wieder

ein orthonormales Basissystem erhalten. Sie zeigen entlang von Längen- und Breitengrade einer imaginären Kugel vom Radius r und haben die Formeln:

$$\vec{e}_r(r,\theta,\varphi) = \begin{pmatrix} \sin(\theta)\,\cos(\varphi) \\ \sin(\theta)\,\sin(\varphi) \\ \cos(\theta) \end{pmatrix}, \; \vec{e}_\theta(r,\theta,\varphi) = \begin{pmatrix} \cos(\theta)\,\cos(\varphi) \\ \cos(\theta)\,\sin(\varphi) \\ -\sin(\theta) \end{pmatrix} \text{ und } \vec{e}_\varphi(r,\theta,\varphi) = \begin{pmatrix} -\sin(\varphi) \\ \cos(\varphi) \\ 0 \end{pmatrix}. \tag{7.5}$$

7.4 Koordinatenwechsel bei Integralen: die Jacobimatrix

Die in den letzten Abschnitten eingeführten krummlinigen Koordinaten vereinfachen die mathematische Beschreibung von "runden" Problemen erheblich. Das sieht man schon am Beispiel eines Kreises mit Radius R. In kartesischen Koordinaten ist der Kreis die Menge aller Punkte mit $y = \sqrt{R^2 - x^2}$ und der Punkte mit $y = -\sqrt{R^2 - x^2}$, wobei $x \in [-R, R]$. Das ist eine recht komplizierte Definition für ein so einfaches geometrisches Objekt wie einen Kreis. In Polarkoordinaten hingegen ist der Kreis einfach die Menge der Punkte mit $r = R$. Es lohnt sich also, Punkte im Raum in den jeweils passenden krummlinigen Koordinaten (Polar-, Zylinder- oder Kugelkoordinaten) anzugeben.

Es ist sicherlich auch bei Integralen über "runde" Geometrien eine gute Idee, die passenden Koordinaten zu nutzen – in einem kugelsymmetrischen Problem also zum Beispiel Kugelkoordinaten. Wir wollen dann von den kartesischen Integrationsvariablen x, y und z zu den neuen Koordinaten r, θ und φ übergehen. Wie geht das? In Kapitel 6.3.4 haben wir schon das Konzept der Variablensubstitution im Integral kennengelernt. Diese erlaubte es uns, von einer Integrationsvariable x über die Definition $x = g(u)$ zu einer neuen Integrationsvariablen $u = g^{-1}(x)$ überzugehen. Dabei mussten wir nicht nur die Integralgrenzen anpassen, sondern haben auch noch einen Extra-Faktor $g'(u)$ im Integral bekommen:

$$\int_a^b dx\, f(x) = \int_{g^{-1}(a)}^{g^{-1}(b)} du\, \frac{dg(u)}{du}\, f(g(u)).$$

Diese Gleichung wollen wir jetzt zum allgemeinen Koordinatenwechsel im Integral verallgemeinern.

7.4.1 Transformationssatz für Integrale

Ausgangspunkt ist ein Integral über n Koordinaten $x_1, \ldots, x_n$ einer Funktion $f(x_1, \ldots, x_n)$:

$$I = \int dx_1 \ldots \int dx_n\, f(x_1, \ldots, x_n). \tag{7.6}$$

Der Einfachheit halber betrachten wir hier unbestimmte Integrale – wenn man etwas konkret ausrechnen will, muss man noch passende Integrationsgrenzen einsetzen. Wir wollen nun zu neuen Koordinaten $u_1, \ldots, u_n$ übergehen. Als ersten Schritt drücken wir hierzu die ursprünglichen Koordinaten als Funktionen der neuen Koordinaten aus:

$$x_i = x_i(u_1, \ldots, u_n) \quad \forall i = 1, \ldots, n. \tag{7.7}$$

Als Beispiel erinnern wir uns an Polarkoordinaten, in denen der Koordinatenwechsel von x und y zu r und φ erfolgt, wobei

$$x = x(r,\varphi) = r\,\cos(\varphi) \quad \text{und} \quad y = y(r,\varphi) = r\,\sin(\varphi).$$

Der Koordinatenwechsel von den x_i zu den u_j wird durch die sogenannte **Jacobi-Matrix** $\mathcal{J}_{\vec{x}}(\vec{u})$ charakterisiert. Diese ist die Matrix der partiellen Ableitungen der alten Koordinaten nach den neuen:

$$\mathcal{J}_{\vec{x}}(\vec{u})|_{ij} = \frac{\partial x_i(\vec{u})}{\partial u_j} \quad \Rightarrow \quad \mathcal{J}_{\vec{x}}(\vec{u}) = \begin{pmatrix} \frac{\partial x_1(\vec{u})}{\partial u_1} & \cdots & \frac{\partial x_1(\vec{u})}{\partial u_n} \\ \frac{\partial x_2(\vec{u})}{\partial u_1} & \cdots & \frac{\partial x_2(\vec{u})}{\partial u_n} \\ \vdots & & \vdots \\ \frac{\partial x_n(\vec{u})}{\partial u_1} & \cdots & \frac{\partial x_n(\vec{u})}{\partial u_n} \end{pmatrix}. \tag{7.8}$$

Im Beispiel der Polarkoordinaten ergibt dies

$$\mathcal{J}_{x,y}(r,\varphi) = \begin{pmatrix} \frac{\partial x(r,\varphi)}{\partial r} & \frac{\partial x(r,\varphi)}{\partial \varphi} \\ \frac{\partial y(r,\varphi)}{\partial r} & \frac{\partial y(r,\varphi)}{\partial \varphi} \end{pmatrix} = \begin{pmatrix} \cos(\varphi) & -r\,\sin(\varphi) \\ \sin(\varphi) & r\,\cos(\varphi) \end{pmatrix}.$$

Der sogenannte Transformationssatz besagt nun, dass sich Integrale wie folgt transformieren:

$$\int d^n x\, f(x_1,\ldots,x_n) = \int d^n u\ |\mathrm{Det}\,\{\mathcal{J}_{\vec{x}}(\vec{u})\}|\ f(x_1(\vec{u}),\ldots,x_n(\vec{u})). \tag{7.9}$$

Für Polarkoordinaten ergibt sich (wie Sie leicht nachrechnen können):

$$\mathcal{J}_{x,y}(r,\varphi) = \begin{pmatrix} \cos(\varphi) & -r\,\sin(\varphi) \\ \sin(\varphi) & r\,\cos(\varphi) \end{pmatrix}$$

$$\Rightarrow\ \mathrm{Det}\,\{\mathcal{J}_{x,y}(r,\varphi)\} = r\,\cos^2(\varphi) - \left(-r\,\sin^2(\varphi)\right) = r(\cos^2(\varphi)+\sin^2(\varphi)) = r.$$

Hierbei haben wir das allgemeine "Additionstheorem" $\cos^2(\varphi)+\sin^2(\varphi) = 1\ \forall\varphi$ (auch "trigonometrischer Pythagoras" genannt) genutzt. Da $r = |r|$ ist, ergibt sich für Integrale folgende Transformationsregel:

$$\int\int dx\,dy\, f(x,y) = \int dr \int d\varphi\, r\, f(x(r,\varphi),y(r,\varphi)) = \int\int dr\, d\varphi\, r\, f(x(r,\varphi),y(r,\varphi)).$$

Beim Basiswechsel/Koordinatenwechsel in Integralen gilt zusammenfassend folgendes **Kochrezept**:

1. Drücken Sie die ursprünglichen Koordinaten x_i als Funktionen der neuen Koordinaten u_j aus:
$$x_i = x_i(u_1, \ldots, u_n) = x_i(\vec{u}) \quad \forall i = 1, \ldots, n.$$
2. Bestimmen Sie die Jacobimatrix $\mathcal{J}_{\vec{x}}(\vec{u})$ mit Einträgen
$$\mathcal{J}_{\vec{x}}(\vec{u})|_{ij} = \frac{\partial x_i(\vec{u})}{\partial u_j}$$
3. Verändern Sie das Integral gemäss Gleichung (7.9) wie folgt:
$$\int d^n x\, f(x_1, \ldots, x_n) = \int d^n u\ |\mathrm{Det}\,\{\mathcal{J}_{\vec{x}}(\vec{u})\}|\ f(x_1(\vec{u}), \ldots, x_n(\vec{u})).$$

In Kurzschreibweise gibt man auch oft nur die Transformation des infinitesimalen Flächenelements an:

$$dx\, dy \to r\, dr\, d\varphi. \tag{7.10}$$

Analog kann man diese Rechnung auch für Zylinder- und Kugelkoordinaten ausführen. Man findet dann:

- Zylinderkoordinaten: $dx\, dy\, dz \to \rho\, d\rho\, d\varphi\, dz$.
- Kugelkoordinaten: $dx\, dy\, dz \to r^2\, \sin(\theta)\, dr\, d\theta\, d\varphi$.

Eine geometrische Interpretation des Transformationssatzes findet sich übrigens in Anhang C.

Rechenaufgabe:

1. Berechnen Sie die Fläche eines Kreises mit Radius R in den am besten passenden Koordinaten. Nutzen Sie hierfür den Transformationssatz, und überlegen Sie, was die passenden Grenzen sind. Führen Sie dann das Integral aus.

   $$A = \int\int_{\text{Kreis}} d^2r =$$

2. Berechnen Sie analog das Volumen einer Vollkugel mit Radius R in den am besten passenden Koordinaten. Tipp: $\cos(0) = 1$ und $\cos(\pi) = -1$.

   $$V = \int\int\int_{\text{Vollkugel}} d^3r =$$

Das wäre im Physikerleben gut zu wissen:

- Wissen, wie Polar-, Zylinder- und Kugelkoordinaten funktionieren:
  - Definition dieser Koordinaten kennen.
  - Wertebereiche der neuen Variablen kennen.
- Wissen, dass die Basisvektoren in diesen Koordinaten vom Ort abhängen.
- Den letzten Punkt mindestens am Beispiel von $\vec{e}_r$ (Polar- und Kugelkoordinaten) bzw. $\vec{e}_\rho$ (Zylinderkoordinaten) erklären können – und diese also auswendig kennen.
- Wissen, wie die Jacobi-Matrix definiert ist.
- Wissen, wie man in einem Integral einen Koordinatenwechsel durchführt.
- Die Koordinatentransformation für Integrale von kartesischen Koordinaten zu Polar-, Zylinder- und Kugelkoordinaten auswendig kennen.

PS: "gut zu wissen" muss nicht heißen, dass das alles in der Klausur abgefragt wird – oder, dass sonst nichts aus diesem Kapitel in der Klausur vorkommt. Wer aber mit den Punkten in diesem Kasten nichts anfangen kann, hat wahrscheinlich in der Klausur und im weiteren Studium Probleme.

Kapitel 8

Vektoranalysis und Integralsätze von Gauß und Stokes

In diesem Kapitel kommen wir nochmals auf das Konzept der Felder zurück, die schon im Kapitel 2.6 eingeführt wurden. Felder sind mathematisch gesehen einfach Funktionen mehrerer Variablen. Physikalisch sind es Funktionen, die jedem Punkt $\vec{r}$ im Raum zu jeder Zeit t einen Funktionswert zuordnen. Genauer haben wir Felder wie folgt unterschieden:

- Ist der Funktionswert eine Zahl (ein Skalar), so sprechen wir von einem Skalarfeld. Beispiele sind die Temperatur $T(\vec{r}, t)$ oder die Dichte $\rho(\vec{r}, t)$.
- Ist der Funktionswert ein Vektor, so sprechen wir von einem Vektorfeld. Beispiele sind das elektrische Feld $\vec{E}(\vec{r}, t)$ oder das magnetische Feld $\vec{B}(\vec{r}, t)$.
- Ist der Funktionswert ein Tensor n-ter Stufe, so sprechen wir von einem Tensorfeld. Beispiele sind die dielektrische Permittivität $\epsilon(\vec{r}, t)$ (ein Tensor zweiter Stufe, der in der Optik wichtig ist), oder der Riemannsche Krümmungstensor $R^{\alpha}{}_{\beta\gamma\delta}$ (ein Tensor vierter Stufe, der in der Relativitätstheorie die Raumzeit charakterisiert).
- Hängt ein Feld nicht von der Zeit ab, so spricht man von einem statischen Feld.

Felder spielen in der Physik eine zentrale Rolle, denn es geht in der Physik eben meistens nicht um Dinge, die immer den gleichen Wert haben, sondern darum, wie Dinge sich unter bestimmten Bedingungen "verhalten", also ihren Zustand (oder Wert) ändern. In diesem Kapitel gehen wir jetzt nochmal genauer auf das Rechnen mit Feldern ein und betreiben also **Vektoranalysis**.

Oftmals ist der reine Wert eines Feldes nicht das Ende der physikalischen Fragestellung: normalerweise müssen wir aus den Feldern noch weitere Größen berechnen, um am Ende den Ausgang eines Experiments zu beschreiben. Für die "Weiterverarbeitung" von Feldern haben wir zwei Haupt-Rechentechniken: Ableitung (Differentiation) und Integration.

8.1 Differentialoperatoren der Vektoranalysis

Die erste wichtige Rechentechnik bei Feldern ist das Ableiten, also die Differentiation. Die nun folgenden Formeln hängen von der im konkreten Problem betrachteten Raumdimension ab. Diese ist in der Physik normalerweise $d = 3$, manchmal auch $d = 2$ oder $d = 1$, und nur selten $d = 4, 5, \ldots$ (der letzte Fall hat zwar zuerst einmal nichts mit unserer alltäglichen Welt zu tun, kann aber trotzdem manchmal interessant und hilfreich sein). Wir betrachten im folgenden den allgemeinen Fall von d Raumdimensionen. Felder hängen dann vom d-dimensionalen Ortsvektor $\vec{r}$ und der Zeit t ab. Ein

Skalarfeld ϕ und ein n-komponentiges Vektorfeld $\vec{A}$ sind in diesem Fall mathematisch wie folgt definiert:

$$\phi : \mathbb{R}^{d+1} \to \mathbb{R}, \quad (\vec{r}, t) \mapsto \phi(\vec{r}, t) \quad \text{und} \quad \vec{A} : \mathbb{R}^{d+1} \to \mathbb{R}^n, \quad (\vec{r}, t) \mapsto \vec{A}(\vec{r}, t).$$

Diese Felder können wir nun mit verschiedenen Kombinationen von Ableitungen weiterverarbeiten. Diese Kombinationen von Ableitungen basieren alle auf dem Nabla-Operator ∇, den Sie in Kapitel 5.5.1 kennengelernt haben. Im Kontext von Feldern in der Physik ist der Nabla-Operator ist oft der Vektor der **ersten partiellen Ableitungen nach den räumlichen Koordinaten** (und nicht nach der Zeit). Für d Raumdimensionen ist ∇ dann definiert als

$$\nabla = \begin{pmatrix} \frac{\partial}{\partial r_1} \\ \frac{\partial}{\partial r_2} \\ \vdots \\ \frac{\partial}{\partial r_d} \end{pmatrix}, \tag{8.1}$$

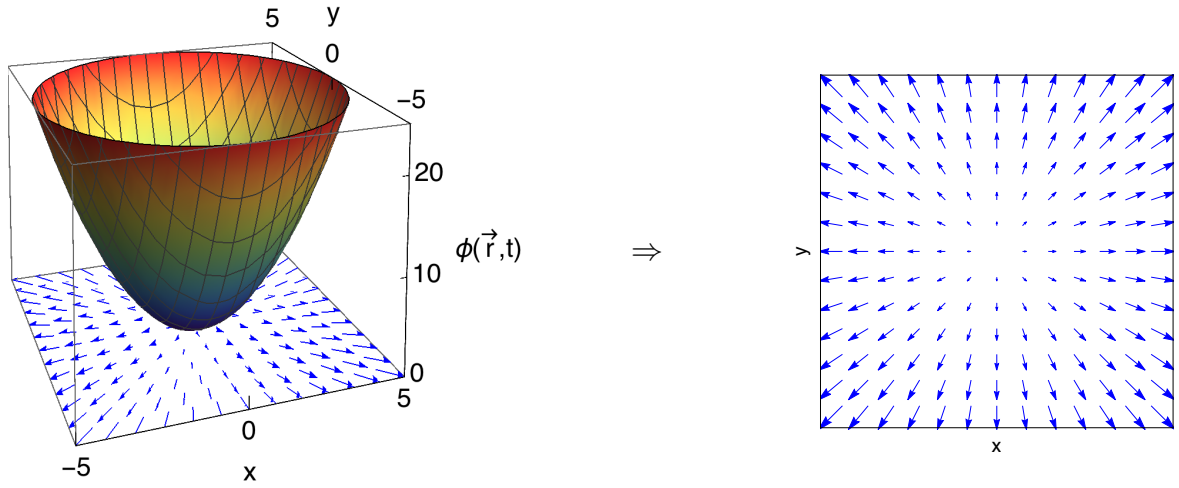
wobei r_i die i-te Komponente des Ortsvektors $\vec{r}$ ist, also die i-te Koordinate. Wir nennen diesen Vektor auch einen **Differentialoperator**: wenn er auf ein Feld losgelassen wird, macht er etwas mit dem Feld (er leitet es ab) – er führt also eine Rechenoperation durch. Je nachdem von welcher Form das Feld ist, können wir mit dem ∇-Operator verschiedene Rechenoperationen definieren.

8.1.1 Skalare Felder und ∇: der Gradient

Wenn wir ein skalares Feld $\phi(\vec{r}, t)$ betrachten, können wir mit dem ∇-Operator den sogenannten Gradienten des Feldes bilden. Wie in Kapitel 5.5.1 bereits behandelt, ist dieser an einem Punkt $\vec{r}_0$ wie folgt definiert:

$$\text{grad}\, \phi(\vec{r}_0, t) = \nabla \phi(\vec{r}_0, t) = \begin{pmatrix} \frac{\partial \phi(\vec{r}_0, t)}{\partial r_1} \\ \frac{\partial \phi(\vec{r}_0, t)}{\partial r_2} \\ \vdots \\ \frac{\partial \phi(\vec{r}_0, t)}{\partial r_d}. \end{pmatrix} \tag{8.2}$$

Der Gradient macht aus einem skalaren Feld ein Vektorfeld. Dieses Vektorfeld hat eine einfache geometrische Interpretation: man kann zeigen, dass der Gradient eines Feldes immer in die **Richtung des steilsten Anstiegs** der Funktionswerte von $\phi(\vec{r}, t)$ zeigt. Der Gradient steht also insbesondere senkrecht zu den Äquipotentiallinien / Äquipotentialflächen von $\phi(\vec{r}, t)$. Zur Illustration zeigen die folgenden beiden Plots das Feld $\phi(\vec{r}, t) = x^2 + y^2$ über seinem Gradientenfeld sowie das Gradientenfeld in der Draufsicht (hier ist $\vec{r} = \begin{pmatrix} x \\ y \end{pmatrix}$ also der zweidimensionale Raum):



Feld $\phi(\vec{r},t) = x^2 + y^2$ Gradient $\nabla\,\phi(\vec{r},t) = \begin{pmatrix} 2x \\ 2y \end{pmatrix}$

8.1.2 d-dimensionale Vektorfelder in d Dimensionen: die Divergenz

Hier betrachtet man ein Vektorfeld, dass ebensoviele Komponenten hat, wie es Raumdimensionen gibt:

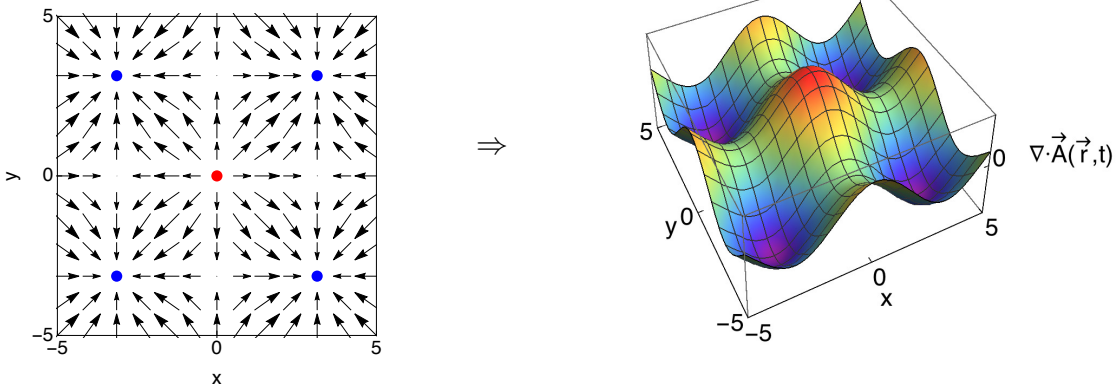
$$\vec{A} : \mathbb{R}^{d+1} \to \mathbb{R}^d, \quad (\vec{r},t) \mapsto \vec{A}(\vec{r},t).$$

Bei einem solchen Feld kann man die sogenannte "Divergenz" bilden, die mit dem Symbol "div" oder durch "$\nabla\cdot$" bezeichnet wird (der Punkt $\cdot$ bezeichnet hierbei das Skalrprodukt mit einem Vektor, der noch rechts an die Divergenz $\nabla\cdot$ "rangepackt" werden muss). Hierzu leitet man jeweils die i-te Komponente des Feldes nach der i-ten Koordinate ab:

$$\text{div}\,\vec{A}(\vec{r}_0,t) = \nabla\cdot\vec{A}(\vec{r}_0,t) = \sum_{i=1}^{d} \frac{\partial A_i(\vec{r}_0,t)}{\partial r_i} = \frac{\partial A_1(\vec{r}_0,t)}{\partial r_1} + \frac{\partial A_2(\vec{r}_0,t)}{\partial r_2} + \ldots + \frac{\partial A_d(\vec{r}_0,t)}{\partial r_d}. \quad (8.3)$$

Beachten Sie, dass Gleichung (8.3) sich in der Tat als Skalarprodukt des ∇-Vektors aus Gleichung (5.16) mit einem Vektorfeld verstehen lässt. Die Divergenz macht aus einem Vektorfeld ein skalares Feld. Die Divergenz eines Vektorfeldes hat anschaulich die Interpretation einer Quellstärke, also der Tendenz des Feldes, von einem gegebenen Punkt wegzufliessen: wenn ein Feld aus einem Punkt "herauskommt", so ist die Divergenz positiv (einen solchen Punkt nennt man auch eine "Quelle" des Vektorfeldes). Geht ein Feld "in einen Punkt hinein", so ist die Divergenz negativ (so ein Punkt ist eine "Senke" des Feldes). Dies zeigt auch das folgende Beispiel des Feldes

$$\vec{A}(\vec{r},t) = \begin{pmatrix} \sin(x) \\ \sin(y) \end{pmatrix} \quad \Rightarrow \quad \nabla\cdot\vec{A}(\vec{r},t) = \cos(x) + \cos(y).$$



Feld $\vec{A}(\vec{r},t)$ Divergenz $\nabla \cdot \vec{A}(\vec{r},t)$

Im linken Plot des Feldes selbst ist die Quelle bei $(0,0)$ rot markiert, während die Senken blau markiert sind. Im rechten Plot der Divergenz des Feldes sieht man gut, dass diese Punkte genau den Maxima und Minima der Divergenz entsprechen. Beachten Sie beim linken Plot, dass die Punkte auf den x- und y-Achsen, bei denen keine Pfeile sind ($x = 0$, $y = \pm\pi$ und $y = 0, x = \pm\pi$) keine Quellen oder Senken sind, da in diesen Punkten das Feld entlang einer Richtung hineinfließt, entlang der anderen Richtung aber herausfließt.

8.1.3 Dreidimensionale Vektorfelder in drei Dimensionen: die Rotation

Die **Rotation** eines Vektorfeldes $\vec{A}(\vec{r},t)$ mit drei Komponenten im dreidimensionalen Raum wird mit "rot" oder "$\nabla\times$" bezeichnet. Sie ist definiert als

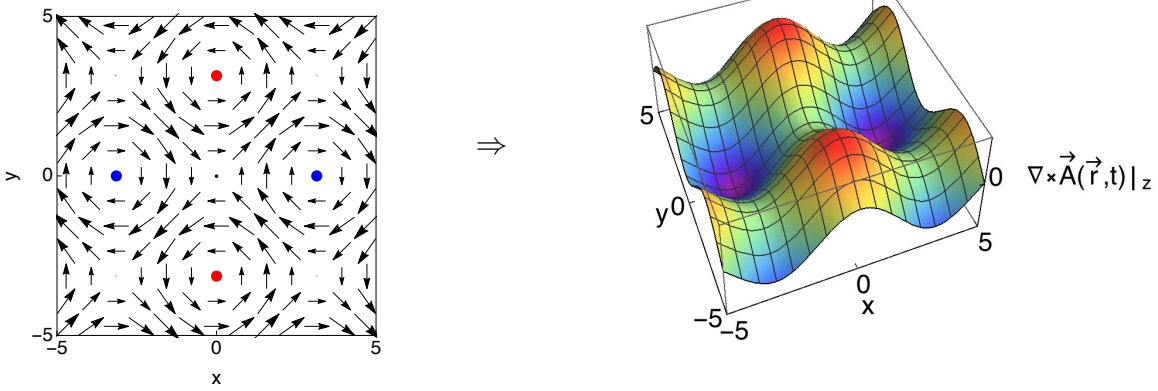
$$\text{rot}\,\vec{A}(\vec{r}_0,t) = \nabla \times \vec{A}(\vec{r}_0,t) \quad \Rightarrow \quad \text{rot}\,\vec{A}(\vec{r}_0,t)\Big|_i = \epsilon_{ijk}\,\partial_j A_k(\vec{r}_0,t) \tag{8.4}$$

(mit Summenkonvention), oder auch explizit

$$\nabla \times \vec{A}(\vec{r}_0,t) = \begin{pmatrix} \frac{\partial A_y(\vec{r}_0,t)}{\partial z} - \frac{\partial A_z(\vec{r}_0,t)}{\partial y} \\ \frac{\partial A_z(\vec{r}_0,t)}{\partial x} - \frac{\partial A_x(\vec{r}_0,t)}{\partial z} \\ \frac{\partial A_x(\vec{r}_0,t)}{\partial y} - \frac{\partial A_y(\vec{r}_0,t)}{\partial x} \end{pmatrix}. \tag{8.5}$$

Beachten Sie, dass Gleichung (8.5) sich in der Tat als Kreuzprodukt des ∇-Vektors aus Gleichung (5.16) mit einem Vektorfeld verstehen lässt. Anschaulich entspricht die Rotation eines Vektorfeldes dem "Grade der Verwirbelung": wenn ein Feld einen Wirbel hat, ist die Rotation des Feldes am Ort des Wirbels groß. Das Vorzeichen der Rotation gibt dabei an, ob der Wirbel im Uhrzeigersinn oder entgegen dem Uhrzeigersinn wirbelt. Dies veranschaulicht das Beispiel des Feldes

$$\vec{A}(\vec{r},t) = \begin{pmatrix} \sin(y) \\ \sin(x) \\ 0 \end{pmatrix} \quad \Rightarrow \quad \nabla \times \vec{A}(\vec{r},t) = \begin{pmatrix} 0 \\ 0 \\ \cos(x) - \cos(y) \end{pmatrix}.$$



Feld $\vec{A}(\vec{r},t)$ in der (x,y)-Ebene z-Komponente der Rotation $\nabla \times \vec{A}(\vec{r},t)$

Links sehen Sie das Vektorfeld $\vec{A}(\vec{r},t)$ in der (x,y)-Ebene in der Draufsicht (das Feld ist in der z-Richtung konstant). Hier sind Wirbel im Uhrzeigersinn blau markiert, Wirbel entgegen des Uhrzeigersinns sind rot markiert. Rechts sehen Sie die Rotation des Feldes. Die nicht markierten Punkte ohne Pfeile im linken Plot (z. B. der Ursprung) kann man auch als "Antiwirbel" bezeichnen, sie haben aber trotzdem verschwindende Rotation.

8.1.4 Zweite Ableitungen: der Laplace-Operator

Zuletzt erinnern wir noch an den Laplace-Operator Δ aus Kapitel 5.5.2. Der Laplace-Operator ist die Summe der zweiten partiellen Ableitungen,

$$\Delta = \sum_{i=1}^{d} \frac{\partial^2}{\partial r_i^2}. \tag{8.6}$$

Der Laplace-Operator kann auch als Skalarprodukt von Divergenz und Gradient aufgefasst werden,

$$\Delta = (\nabla \cdot \nabla) = (\mathrm{div} \cdot \mathrm{grad}). \tag{8.7}$$

Er kann sowohl auf skalare als auch auf Vektorfelder angewandt werden. Im letzteren Fall geschieht die Anwendung in kartesischen Koordinaten komponentenweise:

$$\Delta \vec{A}(\vec{r},t) = \Delta \sum_{i=1}^{d} A_i(\vec{r},t)\, \vec{e}_i = \sum_{i=1}^{d} (\Delta\, A_i(\vec{r},t))\ \vec{e}_i. \tag{8.8}$$

8.1.5 Differentialoperatoren in krummlinigen Koordinaten

Bisher haben wir die Ableitungsoperatoren nur in kartesischen Koordinaten kennengelernt. In Kapitel 7 haben wir aber auch krummlinige Koordinatensysteme eingeführt. Es wäre verlockend, den zweidimensionalen ∇-Operator in Polarkoordinaten einfach als

$$\nabla = \begin{pmatrix} \frac{\partial}{\partial x} \\ \frac{\partial}{\partial y} \end{pmatrix} \overset{?}{\to} \begin{pmatrix} \frac{\partial}{\partial r} \\ \frac{\partial}{\partial \varphi} \end{pmatrix}$$

zu verallgemeinern. **Das ist aber falsch!** Richtig muss man die krummlinigen Koordinaten als Funktion der kartesischen ausdrücken und die Kettenregel beachten. Für den Gradienten ergibt das im zweidimensionalen Fall:

$$\nabla\,\phi(r,\varphi) = \begin{pmatrix} \frac{\partial\phi(r,\varphi)}{\partial x} \\ \frac{\partial\phi(r,\varphi)}{\partial y} \end{pmatrix} = \begin{pmatrix} \frac{\partial\phi(r,\varphi)}{\partial r}\,\frac{\partial r}{\partial x} + \frac{\partial\phi(r,\varphi)}{\partial\varphi}\,\frac{\partial\varphi}{\partial x} \\ \frac{\partial\phi(r,\varphi)}{\partial r}\,\frac{\partial r}{\partial y} + \frac{\partial\phi(r,\varphi)}{\partial\varphi}\,\frac{\partial\varphi}{\partial y} \end{pmatrix}. \tag{8.9}$$

Nachdem man r und φ als Funktionen von x und y ausgedrückt hat,

$$r(x,y) = \sqrt{x^2+y^2},$$
$$\varphi(x,y) = \begin{cases} \arccos\left(\frac{x}{\sqrt{x^2+y^2}}\right) & y \geq 0, \\ 2\pi - \arccos\left(\frac{x}{\sqrt{x^2+y^2}}\right) & y > 0. \end{cases}$$

kann man durch Vergleich mit den Basisvektoren in Gleichung (7.2) zeigen, dass der Gradient in Polarkoordinaten gerade folgende Form hat:

$$\nabla\,\phi(r,\varphi) = \vec{e}_r\,\frac{\partial\phi(r,\varphi)}{\partial r} + \vec{e}_\varphi\,\frac{1}{r}\frac{\partial\phi(r,\varphi)}{\partial\varphi} = \left(\vec{e}_r\,\frac{\partial}{\partial r} + \vec{e}_\varphi\,\frac{1}{r}\frac{\partial}{\partial\varphi}\right)\,\phi(r,\varphi). \tag{8.10}$$

Beachten Sie, dass die Reihenfolge der Ableitungen und der Einheitsvektoren hier wichtig ist, denn die Einheitsvektoren hängen von den Koordinaten ab: bei $\vec{e}_\varphi\,\partial_\varphi\,\phi/r$ wirkt die Ableitung nur auf das Feld ϕ, bei

$$\partial_\varphi\,\vec{e}_\varphi\,\frac{\phi}{r} = \partial_\varphi\,\left(\vec{e}_\varphi\,\frac{\phi}{r}\right) = \frac{\vec{e}_\varphi}{r}\,\partial_\varphi\,\phi + \frac{\phi}{r}\,\partial_\varphi\,\vec{e}_\varphi$$

per Produktregel auch auf den Einheitsvektor! Bei Ableitungen von Vektorfeldern muss man weiterhin beachten, dass auch die Einheitsvektoren von den Koordinaten selbst abhängen. Auf diese Weise kann man zeigen, dass die Ableitungsoperatoren in den verschiedenen Koordinaten folgende Ausdrücke haben (hier nur im Fall von drei Raumdimensionen):

8.1.5.1 Gradient eines skalaren Feldes $\phi(\vec{r},t)$

- kartesische Koordinaten:

$$\nabla\phi(\vec{r},t) = \vec{e}_x\,\frac{\partial\phi(\vec{r},t)}{\partial x} + \vec{e}_y\,\frac{\partial\phi(\vec{r},t)}{\partial y} + \vec{e}_z\,\frac{\partial\phi(\vec{r},t)}{\partial z}.$$

- Zylinderkoordinaten:

$$\nabla\phi(\vec{r},t) = \vec{e}_\rho\,\frac{\partial\phi(\vec{r},t)}{\partial\rho} + \vec{e}_\varphi\,\frac{1}{\rho}\,\frac{\partial\phi(\vec{r},t)}{\partial\varphi} + \vec{e}_z\,\frac{\partial\phi(\vec{r},t)}{\partial z}.$$

- Kugelkoordinaten:

$$\nabla\phi(\vec{r},t) = \vec{e}_r\,\frac{\partial\phi(\vec{r},t)}{\partial r} + \vec{e}_\theta\,\frac{1}{r}\,\frac{\partial\phi(\vec{r},t)}{\partial\theta} + \vec{e}_\varphi\,\frac{1}{r\,\sin(\theta)}\,\frac{\partial\phi(\vec{r},t)}{\partial\varphi}.$$

8.1.5.2 Divergenz eines Vektorfeldes $\vec{A}(\vec{r},t)$

- kartesische Koordinaten:

$$\nabla \cdot \vec{A}(\vec{r},t) = \frac{\partial A_x(\vec{r},t)}{\partial x} + \frac{\partial A_y(\vec{r},t)}{\partial y} + \frac{\partial A_z(\vec{r},t)}{\partial z}.$$

- Zylinderkoordinaten:

$$\nabla \cdot \vec{A}(\vec{r},t) = \frac{1}{\rho}\frac{\partial}{\partial \rho}(\rho\, A_\rho(\vec{r},t)) + \frac{1}{\rho}\frac{\partial A_\varphi(\vec{r},t)}{\partial \varphi} + \frac{\partial A_z(\vec{r},t)}{\partial z}.$$

- Kugelkoordinaten:

$$\nabla \cdot \vec{A}(\vec{r},t) = \frac{1}{r^2}\frac{\partial}{\partial r}(r^2\, A_r(\vec{r},t)) + \frac{1}{r\sin(\theta)}\frac{\partial}{\partial \theta}(\sin(\theta)\, A_\theta(\vec{r},t)) + \frac{1}{r\sin(\theta)}\frac{\partial A_\varphi(\vec{r},t)}{\partial \varphi}.$$

Hier bezeichnen $A_\rho(\vec{r},t)$, $A_r(\vec{r},t)$, $A_\varphi(\vec{r},t)$ und $A_\theta(\vec{r},t)$ die Komponenten von $\vec{A}(\vec{r},t)$ in den krummlinigen Koordinaten, also z. B. $A_\rho(\vec{r},t) = \vec{e}_\rho(\vec{r},t) \cdot \vec{A}(\vec{r},t)$.

8.1.5.3 Rotation eines Vektorfeldes $\vec{A}(\vec{r},t)$

- kartesische Koordinaten:

$$\nabla\times\vec{A}(\vec{r},t) = \vec{e}_x\left(\frac{\partial A_z(\vec{r},t)}{\partial y} - \frac{\partial A_y(\vec{r},t)}{\partial z}\right) + \vec{e}_y\left(\frac{\partial A_x(\vec{r},t)}{\partial z} - \frac{\partial A_z(\vec{r},t)}{\partial x}\right) + \vec{e}_z\left(\frac{\partial A_y(\vec{r},t)}{\partial x} - \frac{\partial A_x(\vec{r},t)}{\partial y}\right).$$

- Zylinderkoordinaten:

$$\begin{aligned}\nabla \times \vec{A}(\vec{r},t) = {} & \vec{e}_\rho\left(\frac{1}{\rho}\frac{\partial A_z(\vec{r},t)}{\partial \varphi} - \frac{\partial A_\varphi(\vec{r},t)}{\partial z}\right) + \vec{e}_\varphi\left(\frac{\partial A_\rho(\vec{r},t)}{\partial z} - \frac{\partial A_z(\vec{r},t)}{\partial \rho}\right) \\ & + \vec{e}_z\,\frac{1}{\rho}\left(\frac{\partial}{\partial \rho}(\rho A_\varphi(\vec{r},t)) - \frac{\partial A_\rho(\vec{r},t)}{\partial \varphi}\right).\end{aligned}$$

- Kugelkoordinaten:

$$\begin{aligned}\nabla \times \vec{A}(\vec{r},t) = {} & \vec{e}_r\,\frac{1}{r\sin(\theta)}\left(\frac{\partial}{\partial \theta}(A_\varphi(\vec{r},t)\sin(\theta)) - \frac{\partial A_\theta(\vec{r},t)}{\partial \varphi}\right) \\ & + \vec{e}_\theta\,\frac{1}{r}\left(\frac{1}{\sin(\theta)}\frac{\partial A_r(\vec{r},t)}{\partial \varphi} - \frac{\partial}{\partial r}(rA_\varphi(\vec{r},t))\right) \\ & + \vec{e}_\varphi\,\frac{1}{r}\left(\frac{\partial}{\partial r}(rA_\theta(\vec{r},t)) - \frac{\partial A_r(\vec{r},t)}{\partial \theta}\right).\end{aligned}$$

8.1.5.4 Laplace-Operator angewandt auf ein skalares Feld $\phi(\vec{r},t)$

- kartesische Koordinaten:

$$\Delta\phi(\vec{r},t) = \frac{\partial^2\phi(\vec{r},t)}{\partial x^2} + \frac{\partial^2\phi(\vec{r},t)}{\partial y^2} + \frac{\partial^2\phi(\vec{r},t)}{\partial z^2}.$$

- Zylinderkoordinaten:

$$\Delta\phi(\vec{r},t) = \frac{1}{\rho}\frac{\partial}{\partial \rho}\left(\rho\,\frac{\partial\phi(\vec{r},t)}{\partial \rho}\right) + \frac{1}{\rho^2}\frac{\partial^2\phi(\vec{r},t)}{\partial \varphi^2} + \frac{\partial^2\phi(\vec{r},t)}{\partial z^2}.$$

- Kugelkoordinaten:

$$\Delta\phi(\vec{r},t) = \frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\,\frac{\partial\phi(\vec{r},t)}{\partial r}\right) + \frac{1}{r^2\sin(\theta)}\frac{\partial}{\partial \theta}\left(\sin(\theta)\frac{\partial\phi(\vec{r},t)}{\partial \theta}\right) + \frac{1}{r^2\sin^2(\theta)}\frac{\partial^2\phi(\vec{r},t)}{\partial \varphi^2}.$$

8.1.6 Ein paar wichtige Rechentricks mit dem ∇-Operator

Wenn man mehrere Differentialoperationen nacheinander ausführt, ergeben sich ein paar Vereinfachungen.

1. **div rot $= \mathbf{0}$**: Für zweifach stetig differenzierbare Vektorfelder $\vec{A}(\vec{r},t)$ gilt:

$$\nabla \cdot \nabla \times \vec{A}(\vec{r},t) = 0, \tag{8.11}$$

   denn mit dem Satz von Schwarz aus Kapitel 5.5.3 und der Antisymmetrie des ϵ-Tensors gilt $\nabla \cdot \nabla \times \vec{A}(\vec{r},t) = \epsilon_{ijk}\partial_i\partial_j A_k(\vec{r},t) = \epsilon_{jik}\partial_j\partial_i A_k(\vec{r},t) = \epsilon_{jik}\partial_i\partial_j A_k(\vec{r},t) = -\epsilon_{ijk}\partial_i\partial_j A_k(\vec{r},t)$.

2. **rot grad $= \mathbf{0}$**: Für zweifach stetig differenzierbare skalare Felder $\phi(\vec{r},t)$ gilt:

$$\nabla \times \nabla\phi(\vec{r},t) = 0, \tag{8.12}$$

   denn $\nabla \times \nabla\phi(\vec{r},t) = \epsilon_{ijk}\partial_j\partial_k\phi(\vec{r},t)$. Der Rest folgt analog zu "div rot $= 0$".

3. **div grad = Laplace**: Für ein skalares Feld $\phi(\vec{r},t)$ gilt:

$$\nabla \cdot \nabla\phi(\vec{r},t) = \Delta\,\phi(\vec{r},t) \tag{8.13}$$

   (dies ist natürlich einfach nur die Definition des Laplace-Operators, siehe Gleichung (8.7)).

4. **rot rot = grad div - div grad = grad div - Laplace**: Für zweifach stetig differenzierbare Vektorfelder $\vec{A}(\vec{r},t)$ gilt:

$$\nabla \times \nabla \times \vec{A}(\vec{r},t) = \nabla(\nabla \cdot \vec{A}(\vec{r},t)) - \Delta\vec{A}(\vec{r},t). \tag{8.14}$$

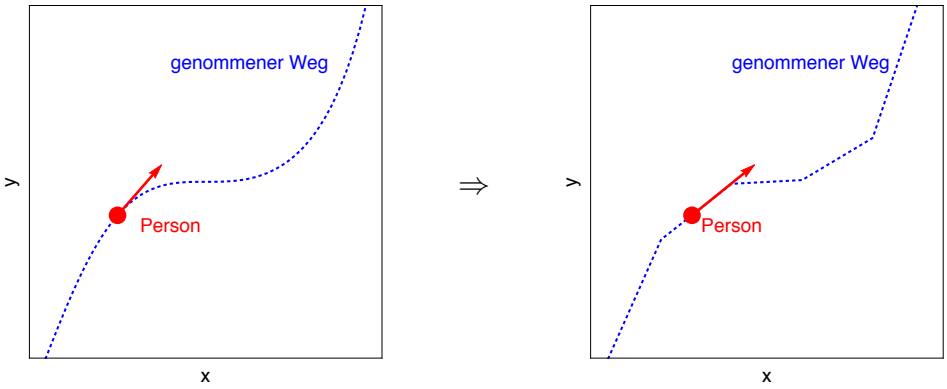
   Auch hier hilft wieder das Rechnen mit dem ϵ-Tensor, siehe 3.5.2, insbesondere Gleichung (3.17): $\nabla \times \nabla \times \vec{A}(\vec{r},t)\Big|_i = \epsilon_{ijk}\epsilon_{klm}\partial_j\partial_l A_m(\vec{r},t) = \partial_i\partial_j A_j(\vec{r},t) - \partial_j\partial_j A_i(\vec{r},t)$.

8.2 Integration von Feldern

Neben Ableitungen kann man mit Feldern auch Integrale bilden. Hier gibt es zum einen die "normalen" Integrale, die schon in Kapitel 6 besprochen wurden (denn letztlich sind Felder nichts anderes als skalarwertige oder vektorwertige Funktionen mehrerer Variablen). Daneben gibt es aber zwei weitere Integraltypen, die im Zusammenhang mit Feldern besonders wichtig sind: Linienintegral und Oberflächenintegrale.

8.2.1 Linienintegral

Das **Linienintegral** wird auch Wegintegral, Kurvenintegral oder Konturintegral genannt. Es bezeichnet die Integration einer Funktion entlang einer Linie, also entlang eines eindimensionalen Teils des Raums. Um dieses Konzept zu verstehen, kommen wir nochmals zum Beispiel einer Person zurück, die über einen Platz läuft:



Wir wollen nun die von der Person zurückgelegte Wegstrecke bestimmen. Das Problem dabei ist, dass der Weg nicht gerade ist. Näherungsweise kann man die Wegstrecke dadurch bestimmen, dass man den zurückgelegten Weg in kleine gerade Teilstücke zerlegt, deren Länge wir mit dem Zollstock bestimmen können. Wenn wir die Länge des i-ten Teilstücks mit Δl_i bezeichnen, so können wir die gesamte Wegstrecke durch die Summe der Teilstücke wir folgt nähern:

$$\text{zurückgelegte Strecke} \approx \sum_i \Delta l_i.$$

Diese Näherung wird offensichtlich besser, wenn man den Weg in immer mehr immer kleinere Teilstücke zerlegt. Im Grenzfall infinitesimal kleiner ("unendlich kleiner") Teilstücke bekommt man dann exakt die zurückgelegte Strecke:

$$\text{zurückgelegte Strecke} = \lim_{\text{alle } \Delta l_i \to 0} \sum_i \Delta l_i.$$

Dies drücken wir als Integral

$$\int_{\text{Weg}} dl = \lim_{\text{alle } \Delta l_i \to 0} \sum_i \Delta l_i \tag{8.15}$$

aus, wobei dl das Differential der Strecke ist. Um solch ein Wegintegral in der Praxis zu berechnen, müssen wir zunächst ein Mal den **Weg parametrisieren**, also eine Funktion $\vec{r}(t)$ finden, die als Funktion eines Parameters t die Orte $\vec{r}(t)$ beschreibt, an denen sich die Person im Laufe des Wegs befindet. Physikalisch ist die Interpretation von t ganz klar die Zeit: als Funktion der Zeit bewegt sich die Person, und zum Zeitpunkt t ist sie eben am Ort $\vec{r}(t)$. Die auf dem i-ten Teilstück zurückgelegte Strecke Δl_i können wir berechnen, indem wir die Geschwindigkeit der Person auf diesem Teilstück, $\vec{v}(t) = \partial_t \vec{r}(t)$ mit der Zeit Δt_i multiplizieren, die die Person für das Wegstück benötigt hat:

$$\Delta l_i = \text{(Weg pro Zeit) mal Zeit} = \frac{\text{Weg}}{\text{Zeit}} \cdot \text{Zeit} = \text{Geschwindigkeit} \cdot \text{Zeit} = |\vec{v}_i|\, \Delta t_i. \tag{8.16}$$

Dann gilt für die zurückgelegte Strecke:

$$\text{zurückgelegte Strecke} = \int_{\text{Weg}} dl = \int |\vec{v}(t)|\, dt = \int \left| \frac{d\vec{r}(t)}{dt} \right| dt = \lim_{\text{alle } \Delta t_i \to 0} \sum_i |\vec{v}_i|\, \Delta t_i. \tag{8.17}$$

Zusammenfassend finden wir also, dass wir zuerst den Weg mit Hilfe einer Hilfsvariablen t als $\vec{r}(t)$ parametrisieren, die anschaulich der Zeit entspricht. Das Wegintegral $\int dl$ drücken wir dann als Integral des Wegparameters t aus: $\int dl = \int dt\, |\partial_t \vec{r}(t)|$.

Die zurückgelegte Wegstrecke im Integral lässt sich also mit folgendem **Kochrezept** berechnen:

1. Parametrisiere Weg: Finde Position als Funktion der Zeit:

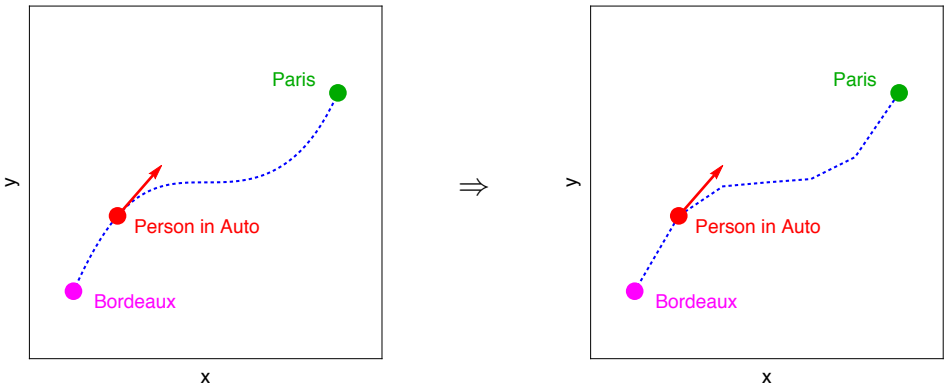
$$\vec{r}(t) = \begin{pmatrix} x(t) \\ y(t) \\ \vdots \end{pmatrix}$$

2. Bestimmen Wegintegral als "Weg = Geschwindigkeit $\times$ Zeit":

$$\text{zurückgelegte Strecke} = \int_{\text{Weg}} dl = \int_{\text{Zeit, die Weg braucht}} |\vec{v}(t)|\, dt = \int_{\text{Zeit, die Weg braucht}} \left|\frac{d\vec{r}(t)}{dt}\right| dt.$$

8.2.2 Linienintegral eines skalaren Feldes

Das Wegintegral kann man nicht nur zum Messen einer Strecke nutzen – besonders interessant wird es, wenn wir eine Funktion entlang eines Weges integrieren. Als Beispiel können wir hier eine Autobahnfahrt in Frankreich betrachten – in unserem Beispielbild sitzt die Person also jetzt im Auto und fährt von Bordeaux nach Paris:



Da es dort die Autobahnmaut gibt, muss die Person auf jedem Teilstück eine Maut bezahlen. Allerdings sind die Mautkosten (Euro pro Kilometer) auf unterschiedlichen Teilstücken unterschiedlich. Die gesamten Kosten sind dann gegeben durch

$$\begin{aligned} \text{Gesamtkosten} &= \sum_i \text{Kosten auf } i\text{-ten Teilstück} \\ &= \sum_i \text{Kosten pro Kilometer auf } i\text{-ten Teilstück } \times \text{ Länge des } i\text{-ten Teilstücks} \end{aligned}$$

Die Kosten pro Kilometer drücken wir nun mathematisch mit einer Funktion $K(\vec{r})$ aus. Diese besagt, dass ein infinitesimal kleines Teilstück der Länge dl, dass um den Ort $\vec{r}$ zentriert ist, die Kosten $K(\vec{r})\, dl$ hat (wir betrachten hier der Einfachheit halber das "Kostenfeld" $K(\vec{r})$ als zeitlich konstant, also als statisches Feld). Zwischen Tours und Orléans zum Beispiel ist $K(\vec{r}) \approx 10\,\text{Cent/km}$. Die Gesamtkosten kann man dann mit Hilfe der Kostenfunktion $K(\vec{r})$ ausdrücken:

$$\text{Gesamtkosten} \approx \sum_i K(\vec{r}_i)\, \Delta l_i = \sum_i K(\vec{r}_i)\, |\vec{v}_i|\, \Delta t_i.$$

Hier ist $\vec{r}_i$ ein Ort auf dem i-ten Teilstück, $\vec{v}_i$ die Geschwindigkeit des Autos auf dem i-ten Teilstück, und Δt_i die Zeit, die das Auto für dieses Teilstück benötigt. Das Wegintegral der Kostenfunktion definieren wir nun als den Grenzwert $\Delta l_i \to 0$ bzw. $\Delta t_i \to 0$. Wenn die Startzeit in Bordeaux t_0 ist und die Endzeit in Paris t_1, so gilt

$$\begin{aligned}\text{Gesamtkosten} &= \int_{\text{Bordeaux}\to\text{Paris}} dl\, K = \int_{t_0}^{t_1} K(\vec{r}(t))\, |\vec{v}(t)|\, dt \\ &= \lim_{\text{alle } \Delta t_i \to 0} \sum_i K(\vec{r}_i)\, |\vec{v}_i|\, \Delta t_i \end{aligned} \tag{8.18}$$

Wenn wir den Weg des Autos mit den Koordinaten $\vec{r}(t)$ beschreiben, so gilt natürlich, dass $\vec{v}(t) = d\vec{r}(t)/dt$ ist. Somit folgt:

$$\text{Gesamtkosten} = \int_{\text{Bordeaux}\to\text{Paris}} dl\, K = \int_{t_0}^{t_1} K(\vec{r}(t)) \left|\frac{d\vec{r}(t)}{dt}\right| dt. \tag{8.19}$$

Diese Definition können wir nun auf ein allgemeines Wegintegral eines skalaren Feldes $\phi(\vec{r}, t)$ entlang eines Weges $\mathcal{C}$ verallgemeinern. Wir parametrisieren den Weg durch eine Kurve $\vec{r}(s)$ so, dass der Startpunkt beim Parameter $s = s_0$ liegt, und der Endpunkt bei $s = s_1$. Dann bezeichnen wir das Wegintegral des Feldes $\varphi(\vec{r}, t)$ entlang des Wegs $\mathcal{C}$ mit

$$\int_{s_0}^{s_1} \left|\frac{d\vec{r}(s)}{ds}\right| \phi(\vec{r}(s), t)\, ds = \int_{\mathcal{C}} |d\vec{r}|\ \phi(\vec{r}, t) \tag{8.20}$$

Beachten Sie, dass der Wegparameter im Allgemeinen nicht gleich der Zeit t sein muss, weshalb wir ihn s genannt haben.

Rechenaufgabe:

Was sind die Fahrtkosten entlang eines Halbkreises mit Radius $1\,\mathrm{km}$, wenn die Kosten pro Kilometer 10 Euro betragen? Nutzen Sie für diese Rechnung das Linienintegral.

8.2.3 Linienintegral eines Vektorfeldes

Bei Vektorfeldern kann man eine weitere Art des Linienintegrals definieren, nämlich eines mit einem Skalarprodukt. Hierzu betrachten wir wieder eine Person, die sich auf einem durch die Kurve $\vec{r}(t)$

parametrisierten Weg bewegt. Zwischen dem Zeitpunkt t und dem infinitesimal späteren Zeitpunkt $t + dt$ hat die Person sich zwischen den Punkten $\vec{r}(t)$ und $\vec{r}(t + dt)$ bewegt. Der Verbindungsvektor dieser beiden Punkte ist

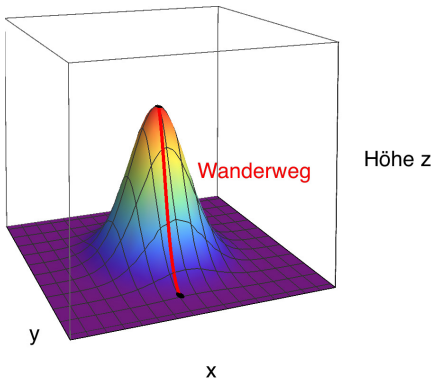
$$d\vec{r}(t) = \vec{r}(t + dt) - \vec{r}(t) = \frac{d\vec{r}(t)}{dt}\, dt. \tag{8.21}$$

Das im letzten Abschnitt berechnete Wegintegral hatte das Differenzial $dl = |d\vec{r}|$. Wir haben also nur den Betrag von $d\vec{r}$ für das Integral gebraucht, und die Richtung des Vektors "weggeschmissen". Manchmal ist das alles, was wir brauchen - manchmal aber ist die Richtung von $d\vec{r}$ auch eine wichtige Information. Betrachten wir als Beispiel ein potentielle Energie, die ein Brötchen hat, wenn es auf einer Höhe z liegt. Wenn wir das Brötchen von der Höhe z zur Höhe $z + \Delta z$ tragen, macht für die potentielle Energie offensichtlich einen großen Unterschied, ob das Brötchen am Ende höher oder tiefer ist, ob also $\Delta z > 0$ oder $\Delta z < 0$ ist. Es ist also im Allgemeinen auch die Richtung des Weges wichtig, nicht nur seine Länge.

Betrachten wir also explizit die potentielle Energie des Brötchens. Diese ist gegeben durch

$$E_{\text{pot}} = m\, g\, z,$$

wobei m die Masse des Brötchens ist und $g \approx 9,81\,\text{m/s}^2$ die Erdbeschleunigung bezeichnet. Wie ändert sich diese potentielle Energie, wenn wir das Brötchen entlang eines Weges $\vec{r}(t)$ während eines Ausflugs auf einen Berg tragen?



Die Strategie der Rechnung ist wieder die gleiche wie bei allen anderen Integralen: wir zerlegen den Weg in kleine Stücke und betrachten die Änderung der potentiellen Energie. Auf dem i-ten Teilstück gilt:

$$\Delta E_{\text{pot,i}} = m\, g\, \Delta z_i,$$

wobei Δz_i die Änderung der Höhe auf dem i-ten Teilstück ist. Wenn das i-te Teilstück zwischen den Zeitpunkten t_i und $t_i + \Delta t_i$ durchlaufen wurde, gilt natürlich:

$$\Delta z_i = z(t_i + \Delta t) - z(t_i) = \frac{z(t_i + \Delta t_i) - z(t_i)}{\Delta t_i}\, \Delta t_i,$$

wobei $z(t)$ die z-Komponente des Orts ist, $z(t) = \vec{e}_z \cdot \vec{r}(t)$. Die Gesamtänderung der potentiellen Energie des Brötchens ist einfach

$$\Delta E_{\text{pot}} = \sum_i \Delta E_{\text{pot,i}} = \sum_i m\, g\, \frac{z(t_i + \Delta t_i) - z(t_i)}{\Delta t_i}\, \Delta t_i = \sum_i m\, g\, \vec{e}_z \cdot \left(\frac{\vec{r}(t_i + \Delta t_i) - \vec{r}(t_i)}{\Delta t_i} \right) \Delta t_i \tag{8.22}$$

Im Grenzfall infinitesimal kleiner Teilstücke geht dieser Ausdruck in folgendes Wegintegral über (wobei t_0 der Startzeitpunkt der Wanderung ist, und wir zum Zeitpunkt t_1 den Gipfel erreicht haben):

$$\Delta E_{\text{pot}} = \int_{t_0}^{t_1} m\, g\, \vec{e}_z \cdot \frac{d\vec{r}(t)}{dt} dt. \tag{8.23}$$

Anstelle des Integrals über den Wegparameter t (den wir immer noch als Zeit interpretieren können) schreibt man das Wegintegral auch oft in folgender Form:

$$\Delta E_{\text{pot}} = \int_{t_0}^{t_1} m\, g\, \vec{e}_z \cdot \frac{d\vec{r}(t)}{dt} dt = \int_{\mathcal{C}} d\vec{r} \cdot \vec{e}_z\, m\, g. \tag{8.24}$$

Hierbei bezeichnet $\mathcal{C}$ den Weg (beim Wandern wären das z. B. der rote Weg in der obigen Skizze). Diesen Ausdruck für die Änderung der potentiellen Energie kann man auch als das klassische "Arbeit ist Kraft mal Weg" verstehen: geleistete Arbeit ist nichts anderes als Energie, und der Weg ist auch offensichtlich im Wegintegral enthalten. Und was die Kraft betrifft: die potentielle Energie, die von der Höhe eines Objekts kommt, ist mit der Erdanziehungskraft verknüpft. Diese zeigt nach unten und hat die Formel

$$\vec{F}_{\text{grav}} = m\, g \begin{pmatrix} 0 \\ 0 \\ -1 \end{pmatrix} = -m\, g\, \vec{e}_z.$$

Hiermit identifizieren wir

$$\Delta E_{\text{pot}} = -\int_{\mathcal{C}} d\vec{r} \cdot \vec{F}_{\text{grav}}, \tag{8.25}$$

also in der Tat "Arbeit ist Kraft mal Weg". Das Minuszeichen sorgt dafür, dass die Änderung der potentiellen Energie positiv ist, wenn das Brötchen den Berg hoch getragen wird (wie es sein sollte).

Diese Definition können wir nun auf ein allgemeines Wegintegral eines Vektorfeldes $\vec{A}(\vec{r}, t)$ entlang eines Weges $\mathcal{C}$ verallgemeinern. Wenn der Weg durch eine Kurve $\vec{r}(s)$ parametrisiert wird, so gilt:

$$\int_{\mathcal{C}} d\vec{r} \cdot \vec{A}(\vec{r}, t) = \int_{s_0}^{s_1} ds\, \frac{d\vec{r}(s)}{ds} \cdot \vec{A}(\vec{r}(s), t) \tag{8.26}$$

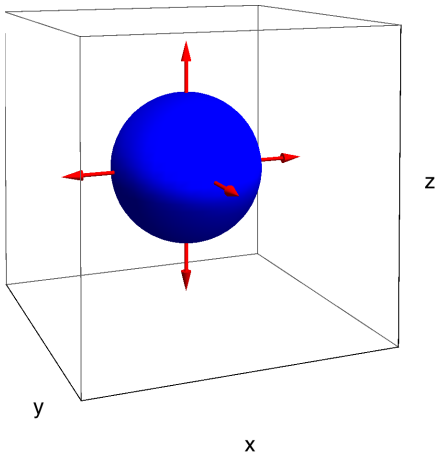
Hier ist wieder s_0 der Wert des Parameters s zu Beginn des Weges und s_1 der Parameterwert am Ende des Weges. Beachten Sie, dass der Wegparameter zwar anschaulich einer "Zeit" entspricht, mathematisch aber nicht das Selbe wie die physikalische Zeit t ist, von der unsere Felder abhängen. Darum haben wir den Wegparamater hier s genannt, während das Feld noch explizit von der physikalischen Zeit t abhängt.

8.2.4 Oberflächenintegral eines Vektorfeldes

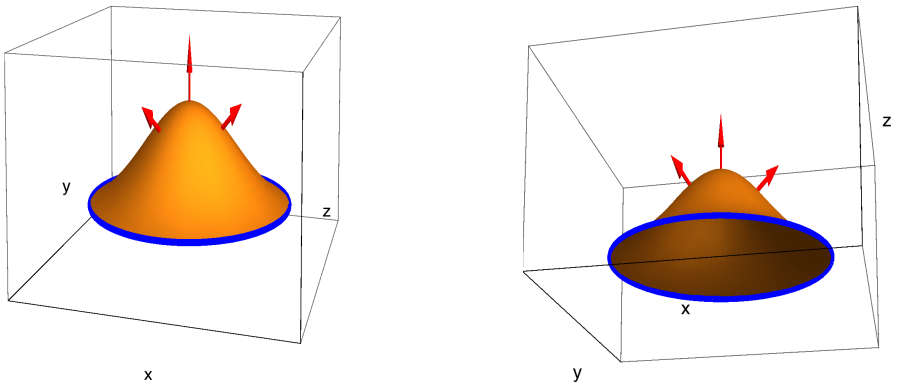
8.2.4.1 Flächen im Raum und ihre Orientierung

Um Oberflächenintegrale von Vektorfeldern zu definieren, müssen wir zunächst das Konzept der Orientierung von Flächen einführen. In einfachen Worten bedeutet dies, das Flächen ein "Innen" von

einem "Außen" trennen. Ein Beispiel hierfür ist die Oberfläche eines Balls. Die **Orientierung** einer Fläche wird mathematisch durch den **Flächennormalenvektor** beschrieben. Dies ist ein **Einheitsvektor**, der in jedem Punkt der Fläche senkrecht auf der Fläche steht und immer **nach außen** zeigt. Beim Beispiel des Balls ist das der Vektor $\vec{e}_r$, der radial nach außen zeigt. Der Normalenvektor hängt also vom Punkt auf dem Ball ab (hier ist der Normalenvektor an verschiedenen Punkten in rot eingezeichnet):

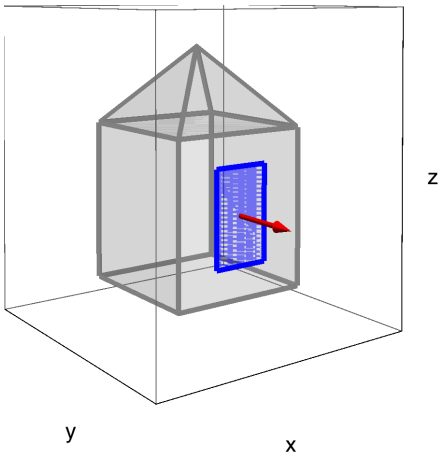


Bei nicht-geschlossenen Flächen wählt man die Richtung der Flächennormalen dadurch aus, dass man der Fläche (willkürlich) eine Oberseite und eine Unterseite zuweist. Der Normalenvektor zeigt dann "eher nach oben" – genauer gesagt senkrecht von der Oberseite weg. Der Normalenvektor ist also im Allgemeinen bei verschiedenen Punkten auf der Fläche unterschiedlich:



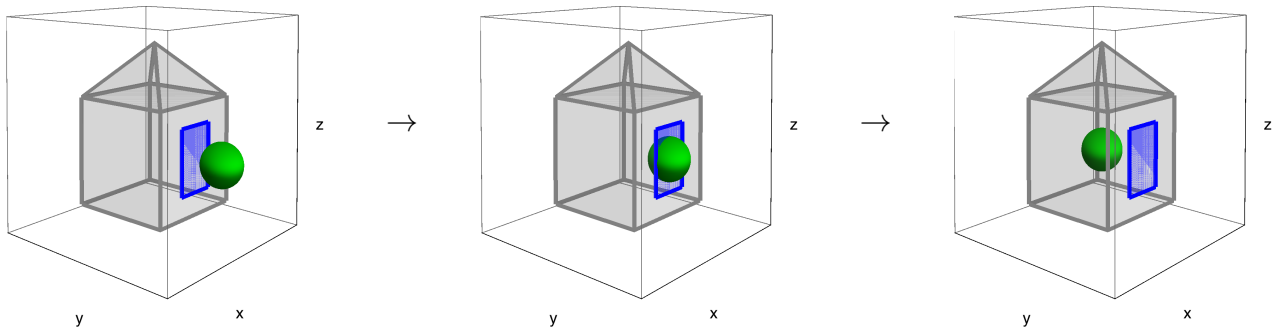
8.2.4.2 Oberflächenintegral

Betrachten wir nun ein Zimmer in einem Haus, das ein Fenster hat. Das Fenster definiert eine Fläche (von z. B. $1\,\mathrm{m} \times 1,5\,\mathrm{m}$). Auch hier gibt es wieder klar ein Innen (das Zimmer) und ein Außen (der Garten). Die Fensterfläche hat also auch wieder einen Normalenvektor (einen Einheitsvektor, der Richtung Garten zeigt). Im folgenden Bild ist der Normalenvektor wieder rot auf der blauen Fensterfläche markiert:



Wir wollen jetzt wissen, wieviele Stechmücken zu einem gegebenen Zeitpunkt bei offenem Fenster in das Zimmer fliegen. Dazu stellen wir uns vor, dass wir eine Lichtschranke am Fenster angebracht haben, die feststellen kann, wenn eine Mücke die Fensterfläche durchfliegt. Allerdings ist hier die Flugrichtung der Stechmücken wichtig: wenn eine Mücke in Richtung des Normalenvektors durch die Fensterfläche fliegt, so verlässt sie ja das Zimmer und erniedrigt die Zahl der Mücken im Zimmer. Fliegt die Mücke entgegen des Normalenvektors, so fliegt sie in das Zimmer, und erhöht so die Zahl der Mücken im Zimmer.

Wir stellen uns jetzt weiter vor, dass das Haus in einem Sumpf steht und Sommer ist – es gibt ziemlich viele Mücken. Wir können also eine Mückendichte $\rho(\vec{r},t)$ (Mücken pro Kubikmeter) definieren. Ein Mückenschwarm (hier in grün illustriert) fliegt nun zum Zeitpunkt $t = t_0$ durch das Fenster ins Haus:



$t < t_0$ $\qquad$ $t = t_0$ $\qquad$ $t > t_0$

Wenn der Mückenschwarm sich mit einer Geschwindigkeit $\vec{v}(\vec{r},t)$ bewegt (die $\vec{r}$-Abhängigkeit der Geschwindigkeit besagt dann z. B., dass die Geschwindigkeit weiter hinten am Schwarm kleiner ist, denn dorthin sind die langsamen Mücken zurückgefallen), so ändert sich die Zahl der Mücken im Zimmer zwischen dem Zeitpunkt $t = t_0$ und $t = t_0 + dt$ durch all die Mücken, die innerhalb dieser Zeitspanne in den Raum geflogen sind. Die Rate der Mückenzahl-Änderung zum Zeitpunkt t ist dabei proportional zur Dichte des Mückenschwarms (wenn die Mücken dichter fliegen, kommen in gegebener Zeit auch mehr ins Zimmer) und zu deren Geschwindigkeit (wenn sie schneller fliegen, ändert sich die Zahl der Mücken im Zimmer schneller). Man kann zeigen, dass die Änderungsrate genau gegeben ist durch

$$\text{Mückenzahl-Änderungsrate bei } t = t_0 : -\int d\vec{S} \cdot \left(\rho(\vec{r},t_0)\,\vec{v}(\vec{r},t_0)\right). \tag{8.27}$$

Hierbei ist

$$d\vec{S} = dS\,\vec{n} = dS \begin{pmatrix} 0 \\ -1 \\ 0 \end{pmatrix} \tag{8.28}$$

das orientierte infinitesimale Flächenelement der Fensterfläche, dS dessen Betrag (also sein Flächeninhalt), und $\vec{n}$ der Normalenvektor der Fläche (der hier in negative y-Richtung zeigt). Das Skalarpodukt $d\vec{S} \cdot \vec{v}$ ist in unserem Beispiel negativ, denn die Flugrichtung der Mücken ist entgegen der Orientierung der Fensterfläche. Wenn die Mücken wieder aus dem Zimmer herausfliegen dreht sich die Richtung ihrer Geschwindigkeit gerade um, $\vec{v} \to -\vec{v}$, so dass sich auch das Vorzeichen des Skalarprodukts (und somit die Mückenzahl-Änderungsrate) umdrehen.

Dieses Konzept können wir nun zum **Oberflächenintegral eines Vektorfeldes** verallgemeinern. Die Fläche sei mit S bezeichnet und habe einen Normalenvektor $\vec{n}(\vec{r})$, der im Allgemeinen noch vom Ort $\vec{r}$ auf der Fläche abhängt. Das Oberflächenintegral eines Vektorfeldes $\vec{A}(\vec{r}, t)$ ist dann wie folgt definiert:

$$\int d\vec{S} \cdot \vec{A}(\vec{r}, t) = \int dS\, \vec{n}(\vec{r}) \cdot \vec{A}(\vec{r}, t). \tag{8.29}$$

Das gerichtete Oberflächenintegral des Vektorfelds $\vec{A}(\vec{r}, t)$ ist also nichts anderes als das normale zweidimensionale Integral des skalaren Feldes $\phi(\vec{r}, t) = \vec{n}(\vec{r}) \cdot \vec{A}(\vec{r}, t)$ über die Fläche S. Dieses können wir wiederum wie in Kapitel 6.5.1 beschrieben berechnen.

8.3 Integralsätze

Bei der praktischen Berechnung von Oberflächenintegralen gibt es zwei besonders hilfreiche "Integralsätze".

8.3.1 Satz von Gauß

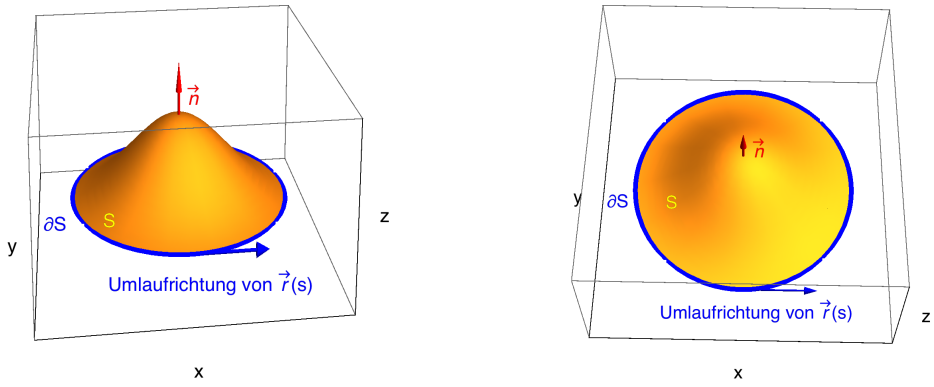
Der Satz von Gauß verbindet Oberflächenintegrale und Volumenintegrale. Gegeben sei ein Volumen V mit einer Oberfläche S, die einen Normalenvektor $\vec{n}(\vec{r})$ hat (der – wenn die Fläche nicht flach ist – wieder vom jeweils betrachteten Punkt auf der Fläche abhängt). Weiter sei $\vec{A}(\vec{r}, t)$ ein stetig differenzierbares Vektorfeld. Dann gilt der **(Integral-) Satz von Gauß**:

$$\int_V dV\, \nabla \cdot \vec{A}(\vec{r}, t) = \int_S d\vec{S} \cdot \vec{A}(\vec{r}, t) = \int_S dS\, \vec{n}(\vec{r}) \cdot \vec{A}(\vec{r}, t). \tag{8.30}$$

In einfachen Worten und bezogen auf unser Mückenbeispiel besagt der Satz von Gauß, dass sich die Mückenzahl im Volumen des Zimmers nur ändert, wenn Mücken durch das Fenster fliegen. In diesem Fall hängt $\vec{A}(\vec{r}, t)$ mit der Mückenbewegung zusammen (genauer gesagt ist es die Mückenstromdichte). Die linke Seite von Gleichung (8.30) enthält somit die "Divergenz der Mückenbewegung", also die Quellen und Senken der Mückenbewegung (vgl. Kapitel 8.1.2) im Volumen des Zimmers. Wir nehmen hier an, dass keine neue Mücken im Zimmer schlüpfen oder alte dort sterben – die Mückenzahl ändert sich also nur, wenn Mücken durch das Fenster fliegen. Die Gleichung besagt also: die Gesamtheit der Quellen und Senken der Mückenbewegung im Zimmer entspricht genau der Gesamtzahl der Mücken, die durch das Fenster fliegen, wobei die Richtung der Mückenbewegung wieder mit der Richtung der Fenster-Flächennormalen verglichen werden muss. Anders gesagt bedeutet Gleichung (8.30) also: die Zahl der Mücken im Zimmer ändert sich (das ist die linke Seite der Gleichung), weil Mücken durch das Fenster fliegen (das ist die rechte Seite der Gleichung).

8.3.2 Satz von Stokes

Der Satz von Stokes verbindet Oberflächenintegrale und Wegintegrale. Gegeben sei eine Oberfläche S, die einen Normalenvektor $\vec{n}(\vec{r})$ hat. Solange diese Fläche nicht geschlossen ist (wie es z. B. die Oberfläche einer Kugel wäre), so hat die Fläche einen Rand. Diesen bezeichnen wir mit ∂S. Der Rand entspricht einer geschlossenen Wegkurve $\vec{r}(s)$. Dabei gibt es ein wichtiges Detail: die **Umlaufrichtung**, mit der der Rand durchlaufen wird. Hier hilft, dass die Fläche gerichtet ist, also einen gegebenen Normalenvektor $\vec{n}$ hat, der in eine spezifische Richtung zeigt. Wenn man von **oben** auf die Fläche schaut, einem der Normalenvektor also entgegenzeigt, so ist der Umlaufsinn der Wegkurve $\vec{r}(s)$ per Konvention als **entgegen dem Uhrzeigersinn** gewählt:



Weiter sei $\vec{A}(\vec{r},t)$ ein stetig differenzierbares Vektorfeld. Dann gilt der **(Integral-) Satz von Stokes**:

$$\int_S d\vec{S}\,\nabla\times\vec{A}(\vec{r},t)=\int_{\partial S} d\vec{r}\cdot\vec{A}(\vec{r},t). \tag{8.31}$$

Wenn wir den Weg $\vec{r}(s)$ um den Rand ∂S wieder so parametrisieren, dass wir bei s_0 starten und bei s_1 die Fläche S ein Mal umrundet haben, können wir den Satz von Stokes auch wie folgt in die aus Kapitel 6 bekannten Integrale umschreiben:

$$\int_S dS\,\vec{n}(\vec{r})\cdot\nabla\times\vec{A}(\vec{r},t)=\int_{s_0}^{s_1} ds\,\frac{d\vec{r}(s)}{ds}\cdot\vec{A}(\vec{r}(s),t).$$

Ist die Fläche S geschlossen, ist sie also zum Beispiel die Oberfläche einer Kugel, so hat sie keinen Rand. Dann gilt:

$$\int_S d\vec{S}\,\nabla\times\vec{A}(\vec{r},t)=0.$$

Das wäre im Physikerleben gut zu wissen:

- Gradient, Divergenz, Rotation und Laplace-Operator kennen und anwenden können.
- Die anschauliche Interpretation von Gradient, Divergenz und Rotation kennen.
- Wissen, dass Differentialoperatoren in krummlinigen Koordinaten andere (komplizierte) Ausdrücke haben.
- Ein Linienintegral eines skalaren Feldes berechnen können.
- Ein Linienintegral eines Vektorfeldes berechnen können.
- Oberflächen orientieren können.
- Oberflächenintegrale berechnen können.
- Den Satz von Gauß kennen.
- Den Satz von Stokes kennen.

PS: "gut zu wissen" muss nicht heißen, dass das alles in der Klausur abgefragt wird – oder, dass sonst nichts aus diesem Kapitel in der Klausur vorkommt. Wer aber mit den Punkten in diesem Kasten nichts anfangen kann, hat wahrscheinlich in der Klausur und im weiteren Studium Probleme.

Kapitel 9

Gewöhnliche Differentialgleichungen

Einer der wichtigsten mathematischen Grundpfeiler der Physik sind Differentialgleichungen. Mit einer Differentialgleichung beschreibt man einen Prozess, bei dem man weiß, wie sich etwas ändert. Ein Beispiel hier ist das Wachstum einer Bakterienkultur. Bakterien vermehren sich durch Zellteilung. Wenn sich in einer Petrischale mehr Bakterien befinden, so teilen sich auch mehr Bakterien. Die Zuwachsrate der Bakterien hängt also davon ab, wieviele Bakterien sich zu einem gegebenen Zeitpunkt in der Petrischale befinden. Die Änderung der Zahl der Bakterien als Funktion der Zeit $dN(t)/dt$ hängt also von N selbst ab. Genauer gesagt ist $dN(t)/dt$ proportional zu $N(t)$: wenn es doppelt so viele Bakterien gibt, teilen sich auch doppelt so viele Bakterien, und die Bakterien-Wachstumsrate ist doppelt so groß:

$$\frac{dN(t)}{dt} = \lambda\, N(t). \tag{9.1}$$

Die Konstante $\lambda \in \mathbb{R}$ bestimmt hier, wie schnell sich die Bakterien konkret vermehren, d. h. ob die Zellteilung also eine Stunde oder einen Tag dauert. Eine solche Gleichung für eine gesuchte Funktion in der auch Ableitungen dieser Funktion vorkommen heißt **Differentialgleichung** (abgekürzt DGL).

9.1 Grundbegriffe von Differentialgleichungen

1. **Gewöhnlich vs. partiell:**
   Zunächst unterscheiden wir zwischen verschiedenen Typen von Differentialgleichungen:

   - Eine **gewöhnliche Differentialgleichung** ist eine Differentialgleichung für eine Funktion, die nur von einer Variablen abhängt. Gleichung (9.1) ist von diesem Typ.
   - Eine **partielle Differentialgleichung** erhält man, wenn man eine Funktion mehrerer Variablen betrachtet (z. B. $f(x, y)$) und in der Differentialgleichung partielle Ableitungen nach mehr als einer Variable auftreten. Ein Beispiel ist die Diffusionsgleichung, die die Änderung eine Dichte $\rho(\vec{r}, t)$ wie folgt beschreibt:

     $$\frac{\partial \rho(\vec{r}, t)}{\partial t} = D \sum_{i=1}^{n} \frac{\partial^2 \rho(\vec{r}, t)}{\partial r_i^2},$$

     wobei $\vec{r}$ der Ort ist, an dem wir die Dichte betrachten, t der Zeitpunkt, D die "Diffusionskonstante" bezeichnet und $\vec{r}$ aus n Komponenten r_i besteht.

   Wir wollen uns hier im weiteren auf die gewöhnlichen Differentialgleichungen beschränken.

2. **Ordnung:**
   Eine weitere Kenngröße einer Differentialgleichung ist ihre **Ordnung**. Tritt in einer Differentialgleichung die n-te Ableitung als höchste Ableitung auf, so hat die Differentialgleichung die Ordnung n. Gleichung (9.1) ist also eine Differentialgleichung der Ordnung 1.

3. **Linearität und Homogenität:**
   Eine lineare Differentialgleichung enthält die unbekannten Funktionen und deren Ableitungen nur in erster Potenz und auch nicht als Produkte. Die allgemeinst-mögliche Form einer linearen gewöhnlichen Differentialgleichung n-ter Ordnung für eine Funktion $f(x)$ ist demnach

$$a_n(x)\,\frac{d^n f(x)}{dx^n} + a_{n-1}(x)\,\frac{d^{n-1} f(x)}{dx^{n-1}} + \ldots + a_1(x)\,\frac{df(x)}{dx} + a_0(x)\,f(x) = b(x). \tag{9.2}$$

   Nichtlineare Differentialgleichungen beinhalten die gesuchte Funktion $f(x)$ und ihre Ableitungen in nichtlinearer Weise, z. B. $\frac{df(x)}{dx} = f(x)^2$.

   In einer homogenen (linearen) Differentialgleichung gibt es nur Terme, welche die unbekannte Funktion oder deren Ableitung enthalten. Eine inhomogene (lineare) Differentialgleichung weist einen solchen Term auf, in dem weder die Funktion selbst, noch eine ihrer Ableitungen vorkommen. Gleichung (9.2) ist also eine homogene Differentialgleichung wenn $b(x) = 0$ ist, und inhomogen falls $b(x) \neq 0$.

9.2 Eindeutigkeit der Lösung einer Differentialgleichung und die Rolle von Randbedingungen

Wir werden uns in der Folge noch genauer mit dem Lösen von Differentialgleichungen befassen. Wir werden sehen, dass eine Differentialgleichung nur dann eindeutig gelöst werden kann, wenn wir weitere Randbedingungen stellen. Beim Bakterienwachstums zum Beispiel kann man durch Einsetzen zeigen, dass die Funktion

$$N(t) = N_0\, e^{\lambda t}$$

die Differentialgleichung (9.1) löst. Die Bakterienzahl steigt also exponentiell an. Allerdings wissen wir immer noch nicht genau, wie viele Bakterien es zu einem gegebenen Zeitpunkt t gibt – dafür müssen wir erst noch N_0 festlegen, die Anzahl der Bakterien zum Zeitpunkt $t = 0$. die Gleichung hat also nur dann eine eindeutige Lösung, wenn wir noch eine Randbedingung angeben. Allgemeiner kann man zeigen:

Eine gewöhnliche Differentialgleichung n-ter Ordnung braucht n Randbedingungen, damit die Lösung eindeutig festgelegt werden kann.

Bezogen auf das Beispiel in Gleichung (9.2) heißt das folgendes: die Gesamtmenge aller Lösungen der Differentialgleichung – die sogenannte **allgemeine Lösung** – hat n Elemente. Es gibt also n unabhängige Funktionen $f_1(x), \ldots, f_n(x)$, die alle jeweils die Differentialgleichung lösen. Die allgemeine Lösung ist dann einfach eine allgemeine Linearkombination all dieser Lösungen:

$$f_{\text{allg.}} = c_1\, f_1(x) + \ldots + c_n\, f_n(x).$$

Die Konstanten c_1 bis c_n werden auch **Integrationskonstanten** genannt. Diese Konstanten können unter Zuhilfenahme von n zusätzlichen Randbedingungen an die Lösung bestimmt werden.

9.2.1 Allgemeine Lösung einer inhomogenen Differentialgleichung

Die allgemeine Lösung der Differentialgleichung (9.2) hängt natürlich nicht nur von den n Randbedingungen ab, sondern auch davon, ob $b(x) = 0$ ist, oder nicht (ob also die Differentialgleichung homogen oder inhomogen ist). Glücklicherweise lässt sich die allgemeine Lösung einer inhomogenen Differentialgleichung durch eine sogenannte **spezielle Lösung** $f_{\text{spez.}}(x)$ und die allgemeine Lösung $f_{\text{allg. hom.}}(x)$ der zugehörigen homogenen Differentialgleichung ausdrücken. Sei also $f_{\text{allg. hom.}}(x)$ die allgemeine Lösung der Gleichung

$$a_n(x)\,\frac{d^n f_{\text{allg. hom.}}(x)}{dx^n} + \ldots + a_0(x)\,f_{\text{allg. hom.}}(x) = 0.$$

Nach unserer obigen Diskussion hat diese die Form

$$f_{\text{allg. hom.}} = c_{1,\text{hom.}}\,f_{1,\text{hom.}}(x) + \ldots + c_{n,\text{hom.}}\,f_{n,\text{hom.}}(x).$$

Als allgemeine Lösung hängt die Funktion$f_{\text{allg. hom.}}(x)$ noch von n Integrationskonstanten ab. Wir suchen jetzt die allgemeine Lösung der Gleichung

$$a_n(x)\,\frac{d^n f_{\text{allg. inhom.}}(x)}{dx^n} + \ldots + a_0(x)\,f_{\text{allg. inhom.}}(x) = b(x).$$

Auch diese hängt wieder von n Integrationskonstanten ab. Die Funktion $f_{\text{allg. inhom.}}(x)$ ist in voller Allgemeinheit aber leider oft schwer direkt zu berechnen. Häufig ist es aber machbar, eine spezielle Lösung zu finden (durch raten oder ähnliches). Dies ist eine spezielle Funktion $f_{\text{spez.}}(x)$, die von keiner Integrationskonstante abhängt und die Gleichung

$$a_n(x)\,\frac{d^n f_{\text{spez.}}(x)}{dx^n} + \ldots + a_0(x)\,f_{\text{spez.}}(x) = b(x)$$

löst. Man kann jetzt zeigen, dass sich die allgemeine Lösung der inhomogenen Differentialgleichung aus der allgemeinen Lösung der homogenen Differentialgleichung und der speziellen Lösung wie folgt bestimmen lässt:

$$f_{\text{allg. inhom.}}(x) = f_{\text{allg. hom.}}(x) + f_{\text{spez.}}(x) \tag{9.3}$$

9.3 Lösungsansätze und -beispiele für gewöhnliche Differentialgleichungen

Wie löst man eine Differentialgleichung jetzt also konkret? Hierzu gibt es verschiedene Strategien. Welche Strategie am besten passt, hängt vom konkreten Problem ab. Die hier besprochenen Ansätze sind bei weitem auch nicht die vollständige Liste von Lösungsstrategien, geben aber einen guten Eindruck davon, wie man Differentialgleichungen in der Praxis löst.

9.3.1 Raten

Eine tatsächlich oft verwendete Strategie beim Lösen von Differentialgleichungen ist das Raten. Das funktioniert deshalb ganz gut, weil sich die Differentialgleichungen der Physik oft ähnlich sehen – und somit auch ihre Lösungen ähnlich sind. Immer wieder vorkommen werden:

- $f'(x) = \alpha\, f(x)$ (mit $\alpha \in \mathbb{R}$) hat die allgemeine Lösung $f(x) = c_0\, e^{\alpha\, x}$.

  Achtung: wir haben hier eine Differentialgleichung erster Ordnung, also brauchen wir für die allgemeine Lösung auch eine Integrationskonstante (das ist c_0).

- $f''(x) = -\alpha\, f(x)$ (mit $\alpha > 0$): hier setzt sich die allgemeine Lösung aus Sinus und Kosinus zusammen: $f(x) = c_1\, \sin(\alpha\, x) + c_2\, \cos(\alpha\, x)$.

  Achtung: wir haben hier eine Differentialgleichung zweiter Ordnung, also brauchen wir für die allgemeine Lösung auch zwei Integrationskonstanten. Diese sind genau c_1 und c_2. Um deren Werte festzulegen, benötigt man zwei Randbedingungen.

9.3.2 Lösung durch Integration für $\frac{d^m f(x)}{dx^m} = g(x)$

Wenn die Differentialgleichung von der Form

$$\frac{d^m f(x)}{dx^m} = g(x) \tag{9.4}$$

ist, und man die Stammfunktionen der Funktion $g(x)$ kennt, so kann man die Differentialgleichung einfach durch m-fache Integration lösen. Bei einer Differentialgleichung erster Ordnung gilt zum Beispiel:

$$\frac{df(x)}{dx} = g(x) \quad \Rightarrow \quad \int_{x_0}^{x} \frac{df(x)}{dx} = f(x) - f(x_0) = \int_{x_0}^{x} g(x) \quad \Rightarrow \quad f(x) = f(x_0) + \int_{x_0}^{x} g(x). \tag{9.5}$$

9.3.3 Trennung der Variablen für $\frac{df(x)}{dx} = g(x)\, h(f(x))$

Falls die Differentialgleichung von der Form

$$\frac{df(x)}{dx} = g(x)\, h(f(x)) \tag{9.6}$$

ist, so kann man sie durch "Trennung der Variablen" lösen. Hierzu bringt man alle x-Terme auf die eine Seite, und alle f-Terme auf die andere Seite:

$$\frac{df(x)}{dx} = g(x)\, h(f(x)) \quad \Rightarrow \quad \frac{df}{h(f)} = g(x)\, dx. \tag{9.7}$$

Hierbei unterdrückt man bei den f-Termen die Variable x. Jetzt löst man die Differentialgleichung durch Integration. Auf der Seite der x-Integration integriert man dabei von einem Startwert $x = y_0$ bis zum Endwert $x = y$. Auf der Seite der f-Integration integriert man von $f(y_0)$ bis $f(y)$:

$$\frac{df(x)}{dx} = g(x)\, h(f(x)) \quad \Rightarrow \quad \frac{df}{h(f)} = g(x)\, dx \quad \Rightarrow \quad \int_{f(y_0)}^{f(y)} \frac{df}{h(f)} = \int_{y_0}^{y} g(x)\, dx. \tag{9.8}$$

Zuletzt löst man nach $f(y)$ auf und ist fertig.

Als Beispiel betrachten wir die Differentialgleichung

$$\frac{df(x)}{dx} = x^2\, f(x).$$

Hier ist $g(x) = x^2$ und $h(f) = f$. Die Trennung der Veränderlichen ergibt

$$\frac{df}{f} = x^2\, dx \quad \Rightarrow \quad \int_{f(y_0)}^{f(y)} \frac{df}{f} = \int_{y_0}^{y} x^2\, dx \quad \Rightarrow \quad \ln\left(f(y)\right) - \ln\left(f(y_0)\right) = \frac{1}{3}(y^3 - y_0^3).$$

Diese Gleichung lösen wir nach $f(y)$ auf, indem wir die rechte und linke Seite in die e-Funktion einsetzen und die Logarithmus-Gesetze $\ln(a) - \ln(b) = \ln(a/b)$ und $e^{\ln(c)} = c$ nutzen:

$$\ln\left(\frac{f(y)}{f(y_0)}\right) = \frac{1}{3}(y^3 - y_0^3) \quad \Rightarrow \quad e^{\ln\left(\frac{f(y)}{f(y_0)}\right)} = e^{\frac{1}{3}(y^3 - y_0^3)} \quad \Rightarrow \quad f(y) = f(y_0)\, e^{(y^3 - y_0^3)/3}.$$

Wenn wir jetzt noch y in x umbenennen, erhalten wir

$$f(x) = f(y_0)\, e^{(x^3 - y_0^3)/3}.$$

Man kann durch Einsetzen testen, dass dies in der Tat eine Lösung der Differentialgleichung ist. Es ist auch die allgemeine Lösung der Differentialgleichung: wenn wir $\tilde{f}_0 = f(y_0)\, e^{-y_0^3/3}$ setzen, bringen wir die Lösung der Differentialgleichung auf die Form

$$f(x) = \tilde{f}_0\, e^{x^3/3}.$$

Hier sehen wir, dass wir eigentlich nur eine unabhängige Integrationskonstante haben, nämlich $\tilde{f}_0$. Wie es sein muss passt das auch genau mit der Ordnung der Differentialgleichung zusammen (diese ist erster Ordnung)!

9.3.4 Exponentialansatz (für homogene lineare Differentialgleichungen mit konstanten Koeffizienten)

Bei homogenen linearen Differentialgleichungen kann man den Exponentialansatz probieren. Hierzu betrachten wir wieder die allgemeine homogene Differentialgleichung

$$a_n \frac{d^n f(x)}{dx^n} + a_{n-1} \frac{d^{n-1} f(x)}{dx^{n-1}} + \ldots + a_1 \frac{df(x)}{dx} + a_0\, f(x) = 0. \tag{9.9}$$

Hierbei ist wichtig, dass die Koeffizienten a_i konstant sind, also nicht von x abhängen. Nun macht man den Ansatz

$$f(x) = e^{\lambda x}. \tag{9.10}$$

Diesen setzt man in die Differentialgeichung ein und erhält:

$$a_n\, \lambda^n\, e^{\lambda x} + a_{n-1}\, \lambda^{n-1}\, e^{\lambda x} + \ldots + a_1\, \lambda\, e^{\lambda x} + a_0\, e^{\lambda x} = 0. \tag{9.11}$$

Wenn wir jetzt durch $e^{\lambda x}$ teilen erhalten wir ein Polynom in λ:

$$a_n\, \lambda^n + a_{n-1}\, \lambda^{n-1} + \ldots + a_1\, \lambda + a_0 = 0. \tag{9.12}$$

Dieses Polynom kann man (aber vielleicht nur in den komplexen Zahlen) lösen. Bei einer Differentialgleichung n-ter Ordnung erhält man im allgemeinen n Lösungen λ_i, so dass die allgemeine Lösung durch

$$f(x) = c_1\, e^{\lambda_1 x} + c_2\, e^{\lambda_2 x} + \ldots + c_n\, e^{\lambda_n x} \tag{9.13}$$

gegeben ist. Hierbei sind die c_i unsere n Integrationskonstanten.

Betrachten wir als Beispiel die Differentialgleichung

$$\frac{d^2 f(x)}{dx^2} - 5\,\frac{df(x)}{dx} + 6\, f(x) = 0.$$

Der Ansatz $f(x) = e^{\lambda x}$ führt uns auf das Polynom

$$\lambda^2 - 5\,\lambda + 6 = 0.$$

Mit der "Mitternachtsformel" folgen die Lösungen

$$\lambda = \frac{5 \pm \sqrt{5^2 - 4 \cdot 1 \cdot 6}}{2 \cdot 1} = \frac{5 \pm \sqrt{25 - 24}}{2} = \frac{5 \pm 1}{2} \quad \Rightarrow \quad \lambda_1 = 3 \text{ und } \lambda_2 = 2.$$

Die allgemeine Lösung der Gleichung ist also

$$f(x) = c_1\, e^{3x} + c_2\, e^{2x},$$

wobei c_1 und c_2 genau die zwei Integrationskonstanten sind, die wir zur Lösung dieser Differentialgleichung zweiter Ordnung benötigen. Man kann durch einsetzen testen, dass dies in der Tat die allgemeine Lösung der Differentialgleichung ist.

9.3.4.1 Entartete Nullstellen

Es kann vorkommen, dass die Anzahl der voneinander verschiedenen Nullstellen λ_i kleiner ist als die Ordnung n der Differentialgleichung. Das passiert dann, wenn eine oder mehrere der Nullstellen λ_i eine Vielfachheit $m > 1$ haben. Bevor wir dies genauer ausführen, erinnern wir noch kurz an die Definition von "Vielfachheit" einer Nullstelle. Betrachten wir hierzu zunächst die Nullstellen der Funktion

$$f(x) = x^2 - x - 6 = (x + 2)\,(x - 3). \tag{9.14}$$

Diese Funktion hat zwei Nullstellen, nämlich $x = -2$ und $x = 3$. Dies sehen wir daran, dass für $x = -2$ die erste Klammer in Gleichung (9.14) gleich Null wird, während für $x = 3$ die zweite Klammer Null wird (und dann nutzen wir noch, das Null mal irgendwas Null ist, hier also z. B. $f(-2) = (-2 + 2) \cdot (-2 - 3) = 0 \cdot (-6) = 0$). Bei der Funktion $h(x) = x^2$ hingegen haben wir nur eine Nullstelle, nämlich $x = 0$. Diese Nullstelle ist aber in einem gewissen Sinn doppelt vorhanden, denn wir können $h(x) = x^2$ schreiben als

$$h(x) = x^2 = x \cdot x. \tag{9.15}$$

Für $x = 0$ wird jetzt sowohl der erste Faktor x als auch der zweite Faktor x in Gleichung (9.15) gleich Null. Wir nennen $x = 0$ daher einer Nullstelle der Vielfachheit 2. Allgemeiner ist die Vielfachheit einer Nullstelle die Potenz, mit der der entsprechende Faktor in der Funktion auftaucht, die man untersucht. Man nennt Nullstellen mit Vielfachheit $m > 1$ auch entartet (oder, genauer, m-fach entartet).

Kommen wir nun zurück zu Differentialgleichungen bei denen der Exponentialansatz zu einer Funktion mit entarteten Nullstellen führt. Ein Beispiel hierfür ist die Differentialgleichung

$$\frac{d^2 f(x)}{dx^2} = 0.$$

Der Ansatz $f(x) = e^{\lambda x}$ führt dann zum Polynom

$$\lambda^2 = 0.$$

Obwohl die Differentialgleichung Ordnung zwei hat, hat diese Gleichung hat nur eine Lösung: $\lambda = 0$. Diese Nullstelle hat die Vielfachheit zwei. Ist jetzt $f(x) = c_1 = c_1\, e^{0\cdot x}$ die allgemeine Lösung der Differentialgleichung? Nein, denn man kann sich leicht davon überzeugen, dass auch

$$f(x) = c_1 + c_2\, x$$

die Differentialgeichung $f''(x) = 0$ löst. Da diese zweite Lösung zwei Integrationskonstanten c_1 und c_2 hat, muss sie auch die allgemeine Lösung der Differentialgleichung sein.

Diese Einsicht kann man wie folgt verallgemeinern. Ist die Anzahl der voneinander verschiedenen λ_i kleiner ist als die Ordnung n der Differentialgleichung, so finden wir die allgemeine Lösung der Differentialgleichung mit folgendem Rezept:

1. Identifiziere die Nullstellen des Polynoms für λ mit Vielfachheit $m > 1$ (im Beispiel $f''(x) = 0$ hat die Nullstelle $\lambda = 0$ die Vielfachheit 2).

2. Ist λ_i Nullstelle der Vielfachheit m, so erhält man die allgemeine Lösung der Differentialgleichung dadurch, dass man in obiger Gleichung (9.13) $c_i\, e^{\lambda_i x}$ wie folgt ersetzt:

$$c_i\, e^{\lambda_i x} \to \sum_{j=1}^{m} c_{i,j}\, x^{j-1}\, e^{\lambda_i x} = (c_{i,1} + c_{i,2}\, x + \ldots + c_{i,m}\, x^{m-1})\, e^{\lambda_i x}.$$

Ist zum Beispiel bei einer Differentialgleichung n-ter Ordnung mit konstanten Koeffizienten die Nullstelle λ_1 eine Nullstelle der Vielfachheit $m = 2$ während alle anderen Nullstellen Vielfachheit 1 haben, so ist die allgemeine Lösung gegeben durch

$$f(x) = (c_{1,1} + c_{1,2}\, x)\, e^{\lambda_1 x} + c_2\, e^{\lambda_2 x} + \ldots + c_{n-1}\, e^{\lambda_{n-1} x}. \tag{9.16}$$

9.3.5 Inhomogene lineare Differentialgleichungen: Variation der Konstanten

Bei linearen Differentialgleichungen haben wir in Kapitel 9.2.1 besprochen, dass sich die allgemeine Lösung einer inhomogenen Differentialgleichung aus der allgemeinen Lösung der homogenen Differentialgleichung und einer speziellen Lösung zusammensetzt. Um die spezielle Lösung zu bestimmen, kann man das Verfahren der **Variation der Konstanten** nutzen. Wir betrachten hier den Spezialfall einer inhomogenen linearen Differentialgleichung **erster Ordung**. Dieses hat folgende Schritte:

1. Löse zuerst die homogene Differentialgleichung.

2. Ersetze in dieser Lösung die Integrationskonstante c_1 durch eine noch unbekannte Funktion der Variable x, also $c_1 \to c_1(x)$.

3. Nutze diese "modifizierte Lösung" der homogenen Differentialgleichung als Ansatz für die inhomogene Differentialgleichung. Einsetzen ergibt dann eine Differentialgleichung für $c_1(x)$, die hoffentlich leichter zu lösen ist.

Betrachten wir als Beispiel die Differentialgleichung

$$x\,\frac{df(x)}{dx} + 2\,f(x) = 1.$$

1. Zuerst lösen wir die homogene Differentialgleichung, zum Beispiel durch Trennung der Variablen:

$$x\,\frac{df(x)}{dx} + 2\,f(x) = 0 \quad \Rightarrow \quad \frac{df}{f} = -2\frac{dx}{x} \quad \Rightarrow \quad \int_{f_0}^{f_1}\frac{df}{f} = -2\int_{x_0}^{x_1}\frac{dx}{x}$$
$$\Rightarrow \quad \ln\left(\frac{f_1}{f_0}\right) = -2\ln\left(\frac{x_1}{x_0}\right) \quad \Rightarrow \quad \ln\left(\frac{f_1}{f_0}\right) = \ln\left(\left(\frac{x_1}{x_0}\right)^{-2}\right) \quad \Rightarrow \quad f_1 = f_0\left(\frac{x_1}{x_0}\right)^{-2}$$

   Wir finden also als allgemeine Lösung der homogenen Differentialgleichung: $f(x) = c_1\,x^{-2}$.

2. Ersetze in dieser Lösung die Integrationskonstante c_1 durch eine noch unbekannte Funktion der Variable x, also $f(x) = c_1\,x^{-2} \to f(x) = c_1(x)\,x^{-2}$.

3. Dies setzen wir nun als Ansatz in die inhomogene Differentialgleichung ein:

$$\begin{aligned}
&x\,\frac{df(x)}{dx} + 2\,f(x) = 1\\
\Rightarrow\quad & x\,\frac{d}{dx}\left[c_1(x)\,x^{-2}\right] + 2\,f(x) = x\left[\frac{dc_1(x)}{dx}\,x^{-2} - 2\,c_1(x)\,x^{-3}\right] + 2\,\left[c_1(x)\,x^{-2}\right] = 1\\
\Rightarrow\quad & \left[\frac{dc_1(x)}{dx}\,x^{-1} - 2\,c_1(x)\,x^{-2}\right] + 2\,\left[c_1(x)\,x^{-2}\right] = \frac{dc_1(x)}{dx}\,x^{-1} = 1\\
\Rightarrow\quad & \frac{dc_1(x)}{dx} = x.
\end{aligned}$$

   Wir haben jetzt also eine neue Differentialgleichung für $c_1(x)$ erhalten, die in der Tat recht einfach ist. Diese können wir direkt durch Integration lösen:

$$\frac{dc_1(x)}{dx} = x \quad \Rightarrow \quad c_1(x_1) - c_1(x_0) = \int_{x_0}^{x_1} dx\,\frac{dc_1(x)}{dx} = \int_{x_0}^{x_1} dx\,x = \frac{1}{2}(x_1^2 - x_0^2).$$

   Wenn wir jetzt das bisher noch allgemeine $x_1 \to x$ setzen und $\tilde{c}_1 = -x_0^2/2$ definieren, erhalten wir

$$c_1(x) = \tilde{c}_1 + \frac{x^2}{2}.$$

   Die allgemeine Lösung der ursprünglichen Differentialgleichung ist also

$$f(x) = \left(\tilde{c}_1 + \frac{x^2}{2}\right)x^{-2} = \frac{1}{2} + \tilde{c}_1\,x^{-2}.$$

   Man kann wieder durch einsetzen testen, dass diese Funktion die Differentialgleichung löst. Da sie immer noch von einer Integrationskonstante $\tilde{c}_1$ abhängt und die Differentialgleichung erster Ordnung ist, ist dies auch die allgemeine Lösung.

   Es ist auch interessant, diese allgemeine Lösung gemäß der Diskussion von Kapitel 9.2.1 als Summe der allgemeinen Lösung der homogenen Differentialgleichung und einer speziellen Lösung zu schreiben. In der Tat ist die allgemeine Lösung der homogenen Differentialgleichung

$$f_{\text{allg. hom.}} = c_1\,x^{-2}.$$

Wir können also sehen, dass anscheinend $f_{\text{spez.}} = 1/2$ eine spezielle Lösung der inhomogenen Differentialgleichung $x\,f'(x) + 2\,f(x) = 1$ sein muss. Und das stimmt auch: die Ableitung einer Konstante ist Null, und $2\,f_{\text{spez.}} = 1$.

Bei einer inhomogenenen linearen Differentialgeichung **höherer Ordnung** kann man das in Kapitel 9.4 beschriebene Verfahren nutzen, um diese Differentialgleichung auf ein System gekoppelter Differentialgleichungen erster Ordnung zurückzuführen. Für dieses kann man die Variation der Konstanten dann analog durchführen.

9.3.6 Reduktion der Ordnung

Wenn in einer Differentialgleichung für die Funktion $f(x)$ nur höhere Ableitungen auftreten (also zum Beispiel nur die zweite und dritte Ableitung, nicht aber die erste Ableitung und der Term $\propto f(x)$ ohne Ableitung), dann kann man sich das Leben dadurch vereinfachen, dass man einfach die erste auftretende Ableitung als eine neue Funktion $g(x)$ definiert. Betrachten wir den Fall, dass nur die m-te und höhere Ableitungen vorkommen:

$$a_n(x)\,\frac{d^n f(x)}{dx^n} + a_{n-1}(x)\,\frac{d^{n-1} f(x)}{dx^{n-1}} + \ldots + a_m(x)\,\frac{d^m f(x)}{dx^m} = b(x). \tag{9.17}$$

Dann definiert man $g(x) = \frac{d^m f(x)}{dx^m}$ und erhält so die einfachere Differentialgleichung

$$a_n(x)\,\frac{d^{n-m} g(x)}{dx^{n-m}} + a_{n-1}(x)\,\frac{d^{n-m-1} g(x)}{dx^{n-m-1}} + \ldots + a_m(x)\,g(x) = b(x). \tag{9.18}$$

9.4 Systeme gekoppelter Differentialgleichungen und Umschreiben von Differentialgleichungen n-ter Ordnung als n Differentialgleichungen erster Ordnung

Zuletzt betrachten wir noch gekoppelte Systeme von Differentialgleichungen. Wie bei "normalen" gekoppelten Gleichungssystemen haben wir hier mehrere Unbekannte, nur eben unbekannte Funktionen und nicht Variablen, die durch gekoppelte Gleichungen spezifiziert werden. Ein Beispiel ist das System

$$\frac{d^2 f(x)}{dx^2} = g(x) \quad \text{und} \quad \frac{dg(x)}{dx} = 2\,f(x) - g(x). \tag{9.19}$$

Wie löst man ein solches System? Die Antwort ist ziemlich ähnlich zum Lösen eines linearen Gleichungssystems: über die Eigenwerte einer Koeffizientenmatrix. Allerdings sind höhere Ableitungen, wie z. B. $\frac{d^2 f(x)}{dx^2}$, hier ein Hindernis. Um solche höheren Ableitungen formal zu umgehen, führt man zuerst für jede höhere Ableitung eine eigene Funktion ein. Bei der hier gegebenen Ableitung zweiter Ordnung führen wir die Funktion $h(x)$ als die Ableitung der Funktion $f(x)$ ein. Dies erlaubt es uns, die gekoppelten Differentialgleichungen als ein (größeres) System von gekoppelten Differentialgleichungen erster Ordnung zu schreiben:

$$\frac{df(x)}{dx} = h(x) \quad \text{und} \quad \frac{dh(x)}{dx} = g(x) \quad \text{und} \quad \frac{dg(x)}{dx} = 2\,f(x) - g(x). \tag{9.20}$$

Jetzt schreiben wir das Gleichungssystem als Matrix:

$$\frac{d}{dx}\begin{pmatrix} f(x) \\ h(x) \\ g(x) \end{pmatrix} = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 2 & 0 & -1 \end{pmatrix} \begin{pmatrix} f(x) \\ h(x) \\ g(x) \end{pmatrix}. \tag{9.21}$$

Allgemeiner gesagt haben wir jetzt also die Ableitung einer Vektorfunktion $\vec{y}(x)$ mit einer Matrix A geschrieben:

$$\frac{d\vec{y}(x)}{dx} = A\,\vec{y}(x) \quad \text{mit} \quad \vec{y}(x) = \begin{pmatrix} f(x) \\ h(x) \\ g(x) \end{pmatrix} \quad \text{und} \quad A = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 2 & 0 & -1 \end{pmatrix}. \tag{9.22}$$

Dieses System lösen wir jetzt dadurch, dass wir die Eigenwerte α_i und zugehörigen Eigenvektoren $\vec{a}_i$ der Matrix A bestimmen (also die Vektoren, für die $A\,\vec{a}_i = \alpha_i\,\vec{a}_i$ gilt). Dann gilt, dass die allgemeine Lösung der Gleichung gegeben ist durch

$$\vec{y}(x) = \sum_i e^{\alpha_i x}\vec{a}_i. \tag{9.23}$$

Dies kann man direkt durch Einsetzen überprüfen. Inhomogene Gleichungssysteme

$$\frac{d\vec{y}(x)}{dx} = A\,\vec{y}(x) + \vec{b}(x) \tag{9.24}$$

kann man z. B. mit der Matrix-Verallgemeinerung der Variation der Konstanten lösen (siehe Übungsblatt 13).

Allgemeiner gilt: durch das Einführen von neuen Symbolen für höhere Ableitungen können wir analog folgende Differentialgleichung n-ter Ordnung als ein System von n gekoppelten Differentialgleichungen erster Ordnung schreiben:

$$\frac{d^5 f(x)}{dx^5} = -7\,f(x) \quad \Rightarrow \quad \text{definiere} \quad h_j(x) = \frac{d^j f(x)}{dx^j} \quad \text{und erhalte:}$$
$$\frac{df(x)}{dx} = h_1(x) \quad \text{und} \quad \frac{dh_1(x)}{dx} = h_2(x) \quad \text{und} \quad \frac{dh_2(x)}{dx} = h_3(x) \quad \text{und}$$
$$\frac{dh_3(x)}{dx} = h_4(x) \quad \text{und} \quad \frac{dh_4(x)}{dx} = -7\,f(x).$$

Diese Differentialgleichungen lassen sich wieder als Gleichungssystem in Matrixform schreiben:

$$\frac{d}{dx}\begin{pmatrix} f(x) \\ h_1(x) \\ h_2(x) \\ h_3(x) \\ h_4(x) \end{pmatrix} = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ -7 & 0 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} f(x) \\ h_1(x) \\ h_2(x) \\ h_3(x) \\ h_4(x) \end{pmatrix}.$$

Zusammenfassend basiert die Lösung eines gekoppelten Systems von linearen Differentialgleichungen mit konstanten Koeffizienten also auf folgenden Schritten:

1. Führen Sie neue Funktionen für die Ableitungen ein, sodass das Gleichungssystem sich als ein gekoppeltes System von Differentialgleichungen **erster** Ordnung schreiben lässt. Achtung: man kann so auch Differentialgleichungen n-ter als gekoppeltes System von n Differentialgleichungen erster Ordnung umschreiben. Beachten Sie außerdem, dass das System von Differentialgleichungen auch inhomogen sein kann.
2. Bringen Sie das Gleichungssystem in eine Matrix-Form.
3. Bestimmen Sie die Eigenwerte und Eigenvektoren der Matrix (wie in Kapitel 4 beschrieben).
4. Konstruieren Sie die allgemeine Lösung wie in Gleichung (9.23) gegeben.

Das wäre im Physikerleben gut zu wissen:

- Wissen, was eine Differentialgleichung ist.
- Wissen, was homogen/inhomogen, gewöhnlich/partiell und linear bedeuten.
- Wissen, was die allgemeine Lösung einer Differentialgleichung ist und wieviele Integrationskonstanten sie bei einer gegebenen Ordnung hat.
- Wissen, dass sich die allgemeine Lösung einer inhomogenen Differentialgleichung aus der allgemeinen Lösung der homogenen Differentialgleichung und einer speziellen Lösung zusammensetzt.
- Ein paar Techniken zum Lösen von Differentialgleichungen nutzen können, insbesondere direkte Integration, Trennung der Variablen, Variation der Konstanten, Exponentialansatz.
- Gekoppelte Systeme von Differentialgleichungen lösen können.
- Eine Differentialgleichung n-ter Ordnung in n gekoppelte Differentialgleichungen erster Ordnung umschreiben können.

PS: "gut zu wissen" muss nicht heißen, dass das alles in der Klausur abgefragt wird – oder, dass sonst nichts aus diesem Kapitel in der Klausur vorkommt. Wer aber mit den Punkten in diesem Kasten nichts anfangen kann, hat wahrscheinlich in der Klausur und im weiteren Studium Probleme.

Kapitel 10

Komplexe Zahlen

In diesem Kapitel führen wir die sogenannten **komplexen** Zahlen ein. Diese sind in der Physik von großer Bedeutung: einerseits erleichtern sie als mathematisch-technisches Hilfsmittel oft das Rechnen, zum Beispiel bei der Berechnung der Eigenschaften von elektrischen Schaltkreisen. Andererseits aber sind komplexe Zahlen auch weit mehr als nur ein mathematischer Trick: insbesondere die Quantenmechanik ist überhaupt nur mit Hilfe von komplexen Zahlen mathematisch formulierbar. Was sind also diese komplexen Zahlen?

10.1 Definition der imaginären Einheit i

Die komplexen Zahlen $\mathbb{C}$ stellen mathematisch eine Erweiterung der reellen Zahlen $\mathbb{R}$ dar, mit Hilfe derer Gleichungen der Form

$$z^2 = -4$$

gelöst werden können. Innerhalb der reellen Zahlen hat diese Gleichung natürlich keine Lösung, denn das Quadrat einer reellen Zahl ist immer ≥ 0. Die komplexen Zahlen beruhen auf der Definition einer neuen mathematische Größe mit dem Symbol i, der sogenannten **imaginären Einheit**. i soll gerade so sein, dass

$$i^2 = -1 \tag{10.1}$$

ist. Für die obige Gleichung gilt also:

$$z^2 = -4 \quad \Rightarrow \quad z = 2\,i \;\text{ oder }\; z = -2\,i.$$

Genauso wie reelle Zahlen oftmals das Symbol x bekommen (z. B. $x = 2$) werden **komplexe Zahlen** häufig mit z bezeichnet. Jede komplexe Zahl z setzt sich aus einem **Realteil** $\mathrm{Re}(z)$ und einem **Imaginärteil** $\mathrm{Im}(z)$zusammen:

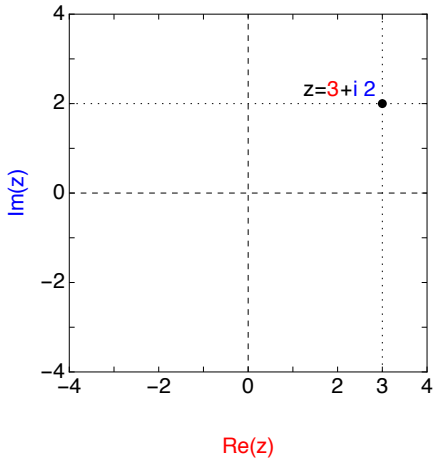
$$z = x + i\,y \quad \text{mit } x \in \mathbb{R} \text{ und } y \in \mathbb{R} \quad \Rightarrow \quad \mathrm{Re}(z) = x \text{ und } \mathrm{Im}(z) = y. \tag{10.2}$$

Der Imaginärteil von z ist also der Teil von z, der mit i multipliziert ist, wohingegen der Realteil von z keinen Faktor von i hat. **Achtung:** sowohl der Real- als auch der Imaginärteil einer komplexen Zahl können Null sein. Die Zahl 5 können wir also wie bisher als eine reelle Zahl verstehen – oder als komplexe Zahl mit verschwindendem Imaginärteil:

$$5 = 5 + i\,0.$$

Eine Zahl mit verschwindendem Realteil, zum Beispiel $z = i\,2$ nennt man auch eine **imaginäre Zahl**.

Graphisch kann man eine komplexe Zahl durch die Angabe von Real- und Imaginärteil in der zweidimensionalen Ebene angeben. Hierzu trägt man auf der x-Achse den Realteil, und auf der y-Achse den Imaginärteil auf. Die Zahl $z = 3 + i\,2$ hat zum Beispiel folgende Darstellung:

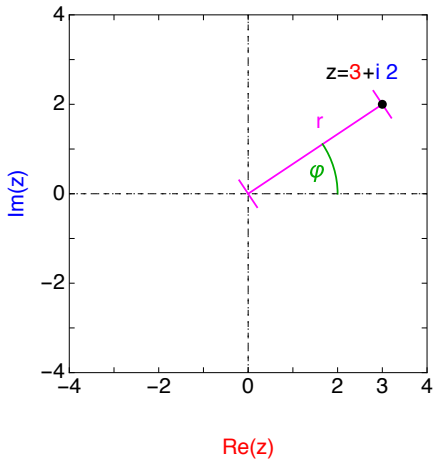


In Anlehnung an die Polarkoordinaten, die wir in Kapitel 7.1 eingeführt haben, kann man eine komplexe Zahl $z = x + i\,y$ statt durch Angabe ihres Real- und Imaginärteils auch durch die Angabe ihres **Betrages** $|z|$ und ihres **Arguments** φ angeben. Diese entsprechen genau dem Radius r und Winkel φ der Polarkoordinaten;

$$z = x + i\,y \quad \Rightarrow \quad x = |z|\,\cos(\varphi) \text{ und } y = |z|\,\sin(\varphi). \tag{10.3}$$

Der Wertebereich des Betrags $|z|$ ist von 0 bis unendlich, $|z| \in [0, \infty[$. Der Wertebereich von φ muss einen Winkel von 2π überdecken. Üblicherweise wählt man die Konvention $\varphi \in]-\pi, \pi]$, man kann aber z. B. auch $\varphi \in [0, 2\pi[$ wählen. Im zweiten Fall gilt

$$z = x + i\,y \quad \Rightarrow \quad |z| = \sqrt{x^2 + y^2} \quad \text{und} \quad \begin{cases} \varphi = \arccos\left(\frac{x}{\sqrt{x^2+y^2}}\right) & , y \geq 0, \\ \varphi = 2\pi - \arccos\left(\frac{x}{\sqrt{x^2+y^2}}\right) & , y < 0, \\ \varphi \text{ ist unbestimmt} & , x = 0 \text{ und } y = 0. \end{cases} \tag{10.4}$$



In dieser sogenannten **Polardarstellung** schreibt man die komplexe Zahl wie folgt:

$$z = x + i\,y = |z|\,e^{i\varphi}. \tag{10.5}$$

Der Vergleich von Gleichung (10.3) und Gleichung (10.5) zeigt, dass die Exponentialfunktion einer imaginären Zahl wie folgt definiert ist:

$$e^{i\varphi} = \cos(\varphi) + i\,\sin(\varphi). \tag{10.6}$$

Diese Gleichung nennt man auch die Eulersche Formel. Aus dieser folgt auch sofort, dass

$$e^{i\pi} = -1. \tag{10.7}$$

Als Seitenbemerkung sei noch gesagt, dass die Eulersche Formel auch deswegen bemerkenswert ist, weil man mit ihr die Grundgrößen 0, 1, e, π und i in einer Formel verbinden kann:

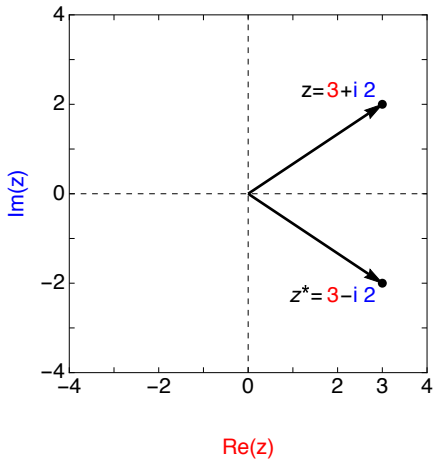
$$e^{i\pi} + 1 = 0. \tag{10.8}$$

10.2 Komplexe Konjugation

Eine wichtige grundlegende Rechenoperation bei komplexen Zahlen ist die sogenannte **komplexe Konjugation**. Die komplexe Konjugation ändert einfach das Vorzeichen des Imaginärteils einer komplexen Zahl. Die komplexe Konjugation wird durch das Symbol * bezeichnet:

$$z = x + i\,y \quad \Rightarrow \quad z^* = x - i\,y \quad (\text{mit } x, y \in \mathbb{R}) \tag{10.9}$$

Die Zahl z^* wird auch als **die zu z komplex konjugierte Zahl** bezeichnet. In der Polardarstellung entspricht die komplexe Konjugation gerade der Spiegelung an der Realteil-Achse (der x-Achse):



Falls z also Betrag $|z|$ und Argument $\varphi(z)$ hat, so hat z^* auch Betrag $|z^*| = |z|$, aber umgedrehtes Argument $\varphi(z^*) = -\varphi(z)$ (zumindest in der Konvention, dass $\varphi \in]-\pi, \pi]$ liegt, sonst muss man $-\varphi(z) \equiv 2\pi - \varphi(z)$ setzen – man sagt dazu auch, dass der Winkel 2π-periodisch ist).

10.3 Rechnen mit komplexen Zahlen

Bei komplexen Zahlen gelten im Wesentlichen die gleichen Rechenregeln wie mit reellen Zahlen – man muss einfach nur auf die imaginäre Einheit i mit $i^2 = -1$ achten. Für zwei komplexe Zahlen $z_1 = x_1 + i\,y_1$ und $z_2 = x_2 + i\,y_2$ (mit $x_1, x_2, y_1, y_2 \in \mathbb{R}$) gilt:

1. Bei **Addition und Subtraktion** werden jeweils die Real- und Imaginärteile addiert und subtrahiert:

$$z_1 \pm z_2 = (x_1 + i\,y_1) \pm (x_2 + i\,y_2) = (x_1 \pm x_2) + i\,(y_1 \pm y_2). \tag{10.10}$$

2. Bei **Multiplikation** muss man alle Terme ausmultiplizieren und $i^2 = -1$ beachten:

$$\begin{aligned} z_1 \cdot z_2 = (x_1 + i\,y_1) \cdot (x_2 + i\,y_2) &= x_1\,x_2 + i\,y_1\,x_2 + i\,x_1\,y_2 + i^2\,y_1\,y_2 \\ &= (x_1\,x_2 - y_1\,y_2) + i(x_1\,y_2 + y_1\,x_2). \end{aligned} \tag{10.11}$$

   Mit $z_1 = |z_1|\,e^{i\varphi_1}$ und $z_2 = |z_2|\,e^{i\varphi_2}$ lässt sich das auch wie folgt schreiben:

$$z_1 \cdot z_2 = \left(|z_1|\,e^{i\varphi_1}\right)\,\left(|z_2|\,e^{i\varphi_2}\right) = |z_1|\,|z_2|\,e^{i(\varphi_1+\varphi_2)}. \tag{10.12}$$

   Bei der Multiplikation zweier komplexer Zahlen multiplizieren sich also ihre Beträge und es addieren sich ihre Argumente.

3. Die **Division** von komplexen Zahlen löst man dadurch, dass man einen gegebenen Bruch mit dem komplex Konjugierten des Nenners erweitert:

$$\frac{z_1}{z_2} = \frac{x_1 + i\,y_1}{x_2 + i\,y_2} = \frac{(x_1 + i\,y_1)\,(x_2 - i\,y_2)}{(x_2 + i\,y_2)\,(x_2 - i\,y_2)} = \frac{x_1\,x_2 + y_1\,y_2}{x_2^2 + y_2^2} + i\,\frac{y_1\,x_2 - x_1\,y_2}{x_2^2 + y_2^2}. \tag{10.13}$$

   Hier haben wir benutzt, dass $i \cdot (-i) = (-1) \cdot i \cdot i = (-1) \cdot (-1) = 1$ ist. Wie die Multiplikation ist auch die Division in der Polardarstellung deutlich einfacher zu schreiben:

$$\frac{z_1}{z_2} = \frac{|z_1|\,e^{i\varphi_1}}{|z_2|\,e^{i\varphi_2}} = \frac{|z_1|}{|z_2|}\,e^{i\varphi_1}\,e^{-i\varphi_2} = \frac{|z_1|}{|z_2|}\,e^{i(\varphi_1-\varphi_2)}. \tag{10.14}$$

   Bei der Division zweier komplexer Zahlen muss man folglich die Beträge dividieren und die Argumente subtrahieren.

4. **Potenzen** von komplexen Zahlen berechnen sich dadurch, dass der Betrag potenziert wird und das Argument mit α multipliziert wird:

$$z^\alpha = \left(|z|\, e^{i\varphi}\right)^\alpha = |z|^\alpha\, e^{i\alpha\varphi} \text{ mit } \alpha \in \mathbb{R}. \tag{10.15}$$

Hierbei ist noch zu beachten, dass das Argument nur bis auf 2π definiert ist – aus der Eulerschen Formel in Gleichung (10.6) folgt zum Beispiel, dass $e^{i\varphi} = e^{i(\varphi+2\pi)}$ ist. Es gilt also:

$$\left(2e^{i2\pi/3}\right)^2 = 2^2\, e^{i2\cdot 2\pi/3} = 4\, e^{i4\pi/3} = 4\, e^{-i2\pi/3}.$$

5. Aus den obigen Rechenregeln folgt für die **komplexe Konjugation**:

   (a) Das Produkt einer komplexen Zahl z mit ihrem komplex Konjugierten z^* ist gleich dem Betragsquadrat der komplexen Zahl:

   $$z\, z^* = |z|^2. \tag{10.16}$$

   (b) Die Summe einer komplexen Zahl z mit Realteil x und Imaginärteil y und ihrer komplex Konjugierten z^* ist gleich dem doppelten Realteil:

   $$z + z^* = (x + i\, y) + (x - i\, y) = 2\, x. \tag{10.17}$$

   (c) Die Differenz einer komplexen Zahl z mit Realteil x und Imaginärteil y und ihrer komplex Konjugierten z^* ist gleich dem doppelten Imaginärteil:

   $$z - z^* = (x + i\, y) - (x - i\, y) = 2\, y. \tag{10.18}$$

10.4 Ein bisschen Funktionentheorie

Die Eulersche Formel $e^{i\varphi} = \cos(\varphi) + i\, \sin(\varphi)$ haben wir in Gleichung (10.6) einfach angegeben - doch wo kommt sie her? Und wie kann man allgemein Funktionen von komplexen Zahle definieren?

10.4.1 Komplexe Funktionen und Cauchy-Riemannsche Differentialgleichungen

Allgemein gesagt ist eine komplexe Funktion eine Abbildung, die eine komplexe Zahl auf eine andere komplexe Zahl abbildet:

$$f : \mathbb{C} \to \mathbb{C}, \quad z = x + i\, y \mapsto u(x, y) + i\, v(x, y), \tag{10.19}$$

wobei $u(x, y)$ und $v(x, y)$ zwei "normale" reelle Funktionen des Realteils x und Imaginärteils y von $z = x + i\, y$ sind. Man könnte also zunächst meinen, dass komplexe Funktionen im Wesentlichen nichts anderes als Funktionen vom $\mathbb{R}^2$ (der zweidimensionalen komplexen Ebene, in der $z = x + i\, y$ lebt) in den $\mathbb{R}^2$ (dort leben der Realteil u und Imaginärteil v von f) sind. Wenn man aber komplex-differenzierbare Funktionen betrachtet, wird deutlich, dass komplexe Funktionen im Vergleich zu Funktionen vom $\mathbb{R}^2$ in den $\mathbb{R}^2$ **sehr viel eingeschränkter** sind.

Um dies einzusehen, definieren wir zunächst die Ableitung einer komplexen Funktion durch den Grenzwert

$$f'(z) = \lim_{h \to 0} \frac{f(z+h) - f(z)}{h}. \tag{10.20}$$

Hierbei sind z und h komplexe Zahlen. Die Funktion f ist genau dann im Punkt z differenzierbar, wenn der Grenzwert in Gleichung (10.20) existiert, und somit insbesondere für alle h gleich ist. Hier kommt nun die Einschränkung komplexer Funktionen im Vergleich zu reellen Funktionen ins Spiel – denn es muss per Definition egal sein, ob h z. B. auf der reellen oder der imaginären Achse nach Null geht. Es muss also für $h = h_x + i\,0 = h_x$ gelten, dass

$$f'(z) = \lim_{h_x \to 0} \frac{u(x+h_x, y) - u(x,y)}{h_x} + \lim_{h_x \to 0} \frac{i\,v(x+h_x, y) - v(x,y)}{h_x} = \partial_x u(x,y) + i\,\partial_x v(x,y). \tag{10.21}$$

Mit $h = 0 + i\,h_y = i\,h_y$ muss weiterhin gelten

$$f'(z) = \lim_{h_y \to 0} \frac{u(x, y+h_y) - u(x,y)}{i\,h_y} + \lim_{h_y \to 0} \frac{i\,v(x, y+h_y) - v(x,y)}{i\,h_y} = -i\,\partial_y u(x,y) + \partial_y v(x,y). \tag{10.22}$$

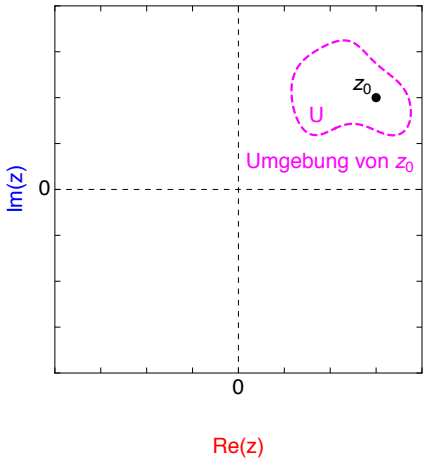
Durch den direkten Vergleich der Real- und Imaginärteile von Gleichungen (10.21) und (10.22) folgt, dass eine komplex-differenzierbare Funktion $f(z) = f(x + i\,y) = u(x,y) + i\,v(x,y)$ folgender Einschränkung unterliegt:

$$\partial_x u(x,y) = \partial_y v(x,y) \quad \text{und} \quad \partial_y u(x,y) = -\partial_x v(x,y). \tag{10.23}$$

Gleichung (10.23) ist das sogenannte **System der Cauchy-Riemannschen Differentialgleichungen**.

10.4.2 Holomorphe Funktionen

Eine besonders wichtige Klasse komplexer Funktionen sind solche, die nicht nur in einzelnen Punkten z_0 differenzierbar sind, sondern mindestens in einer Umgebung U von z_0 (also einer offenen Teilmenge der komplexen Zahlen $\mathbb{C}$, die auch z_0 enthält). Eine solche Funktion heißt **holomorph im Punkt** z_0.



Man kann dann zeigen, dass eine solche holomorphe Funktion in der Umgebung U von z_0, in der sie komplex differenzierbar ist, automatisch nicht nur ein Mal, sondern sogar beliebig oft stetig komplex differenzierbar sein muss. Das klingt vielleicht erst mal überraschend, ist aber eine Konsequenz der Einschränkungen, denen komplex-differenzierbare Funktionen unterliegen. Weiterhin kann man zeigen, dass sich eine solche Funktion in der Umgebung U in einer komplexen Potenzreihe darstellen lässt:

$$f(z) \text{ holomorph in } z_0 \text{ und } z \in U \quad \Rightarrow \quad f(z) = \sum_{n=0}^{\infty} \frac{f^{(n)}(z_0)}{n!}(z - z_0)^n. \tag{10.24}$$

Wichtige Beispiele für holomorphe Funktionen sind:

- Die komplexe Exponentialfunktion $z \mapsto e^z$ – diese ist in ganz $\mathbb{C}$ holomorph.
- Die komplexe Logarithmus-Funktion $z \mapsto \log(z)$ – diese ist in $\mathbb{C}\backslash]-\infty, 0]$ holomorph, also in $\mathbb{C}$ ohne die negative reelle Achse. (Wir definieren den komplexe Logarithmus so, dass er einer komplexen Zahl $z = |z|\, e^{i \arg(z)}$ den Funktionswert $\log(z) = \log(|z|) + i \arg(z)$ zuweist, wobei $\arg(z)$ das Argument von z ist.)

All diese Funktionen können also über ihre Potenzreihe geschrieben werden. Es gibt aber auch wichtige komplexe Funktionen, die **nirgends** holomorph sind:

- Der komplexe Betrag $z \mapsto |z|$.
- Die komplexe Konjugation $z \mapsto z^*$.
- Die Projektion auf den Realteil $z \mapsto \mathrm{Re}(z)$ und die Projektion auf den Imaginärteil $z \mapsto \mathrm{Im}(z)$.

10.5 Grundlagen der Fourier-Transformation

Eine der wichtigsten Anwendungen von komplexen Zahlen in der Physik ist die Fourier-Transformation. Konkret erlaubt die Fourier-Transformation zum Beispiel in der Quantenmechanik den Wechsel von der Orts- zur Impuls-Basis der Wellenfunktionen (und auch wieder zurück).

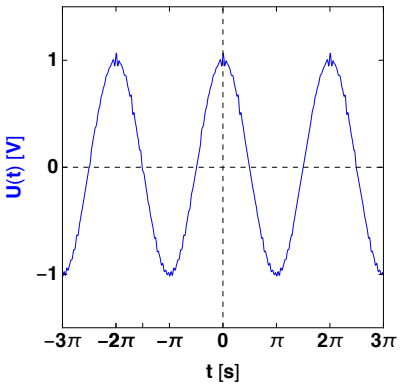
Es gibt verschiedene Blickwinkel auf und Interpretationen der Fourier-Transformation:

- Wechsel zwischen Orts- und Impulsraum in der Quantenmechanik.
- Spektrale Zerlegung von Funksignalen.
- Analyse- und Lösungsmethode für Differenzialgleichungen.

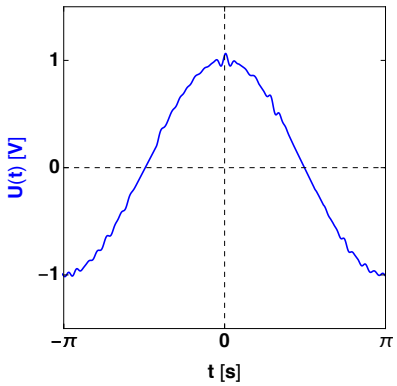
Alle Blickwinkel sind gut und richtig. Der Fakt, dass es diese verschiedenen Sichtweisen auf die Fourier-Transformation gibt, unterstreicht die Bedeutung dieser mathematischen Technik.

10.5.1 Fourier-Transformation und Signalverarbeitung

Ein möglicher Zugang zur Thematik der Fourier-Transformation erfolgt über den Komplex der Signalverarbeitung. Hierbei geht es im Wesentlichen darum, Informationen aus einem gegebenen Signal zu extrahieren oder ein Signal möglichst "effizient" für die Übermittlung vorzubereiten. Betrachten wir als Beispiel die Messung einer Spannung $U(t)$ in Volt als Funktion der Zeit t in Sekunden an einem Stromkreis im Praktikum. Hierbei könnte die Messkurve zum Beispiel so aussehen:

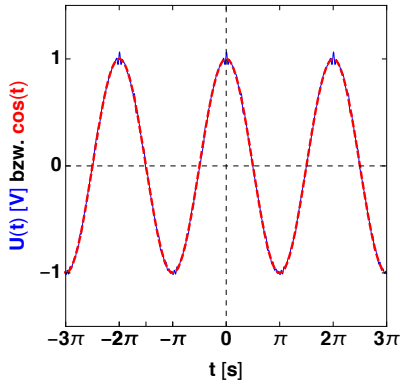


gemessenes Signal



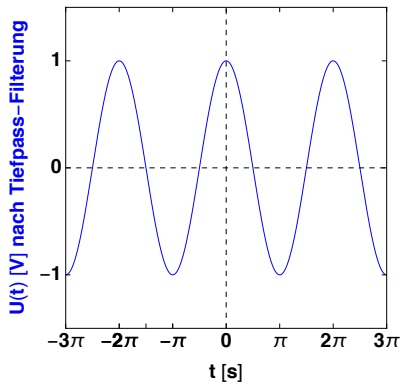
Zoom auf eine Periode

Schauen wir uns dieses Signal genauer an. Es ist tatsächlich perfekt periodisch mit Periode 2π (in einem echten Experiment wäre eine perfekte Periodizität vielleicht nicht gegeben, aber das vernachlässigen wir der Einfachheit halber hier). Ein solches periodisches Signal könnten wir zum Beispiel an einem Schwingkreis messen. Das Signal sieht weiterhin grob wie ein Kosinus aus:



Wir haben also ein Signal, das in ziemlich guter Näherung ein Kosinus ist. Wahrscheinlich auf Grund von nicht ganz perfekten Elementen im Experiment (also nicht ganz perfekten Elementen im Schwingkreis) weicht die gemessene Kurve aber leicht vom Kosinus ab. In den meisten Fällen interessieren uns diese kleinen Abweichungen aber nicht, da sie eben aus unbeabsichtigten, kleinen Fehlern im Experiment stammen.

Um aus dem fast-Kosinus den Kosinus systematisch zu extrahieren, also die zusätzlichen kleinen "Zacken" auf dem Kosinus zu unterdrücken, kann man nun einen Tiefpass-Filter anwenden. Die Idee ist, dass der Kosinus relativ langsam schwingt, wohingegen die Abweichungen vom Kosinus viele kleine, schnelle Zacken hat. Der Tiefpass-Filter unterdrückt jetzt den hochfrequenten Teil des gemessenen Signals (also den Teil, der "schnell zappelt"), und lässt nur den Anteil mit niedrigen Frequenzen passieren. Der Tiefpass-Filter "bügelt" das Signal also gewissermaßen glatt. Angewandt auf das obige Signal erhält man dann in der Tat einen perfekten Kosinus:



Wie im Beispiel des im Praktikum gemessenen Stroms ist es auch sonst oftmals hilfreich, Abweichungen vom "eigentlich gesuchten" Signal aus einem gemessenen Signal entfernen zu können. Beispiele hierfür sind das störende Rauschen beim Fernsehen oder am Telefon, Fehler bei der Datenübertragung im Internet und so weiter. Hierbei ist in der Praxis ein "Signal" meistens eine Welle (Radiowelle, Funkwelle, Lichtwelle im Glasfaserkabel, ...), also eine sich periodisch wiederholende Funktion der Zeit. Die Fourier-Transformation und ihre nahe Verwandte, die Fourier-Reihe, helfen uns dabei, ein besseres Verständnis von Zeit-periodischen Funktionen zu bekommen. Dieses Verständnis können wir dann zum Beispiel nutzen, um ein gegebenes Signal zu verarbeiten.

10.5.2 Die Fourier-Reihe

Die Grundidee der Fourier-Reihe ist, grob gesagt:

Jede periodische Funktion kann als Summe von Sinus- und Kosinus-Funktionen mit unterschiedlichen Argumenten geschrieben werden.

Genauer gesagt gilt: jede "messbare" (das heißt "ausreichend nette", wobei hier nicht wichtig ist, was "ausreichend nett" genau bedeutet – in der Physiker-Praxis ist das quasi jede periodische Funktion) mit $T = \frac{2\pi}{\omega_0}$ periodische Funktion, $f(t) = f(t+T)$, für die das Integral von $|f(t)|^2$ über eine Periode existiert, $\int_0^T |f(t)|^2 < \infty$, ist darstellbar durch eine Summe von Sinus- und Kosinusfunktionen bzw. von komplexen Exponentialfunktionen mit den Frequenzen ω_0, $2\,\omega_0$, $3\,\omega_0$, ... Diese Summe wird Fourier-Reihe genannt und hat die allgemeine Form

$$f(t) = a_0 + \sum_{n=1}^{\infty} a_n \cos(n\,\omega_0\,t) + \sum_{n=1}^{\infty} b_n \sin(n\,\omega_0\,t)\,. \tag{10.25}$$

Hierbei ist $\omega_0 = 2\pi/T$ die sogenannte "Kreisfrequenz" der Schwingung.

Wir betrachten im Folgenden komplexwertige skalare Funktionen f mit Periode T. Die Koeffizienten a_j und b_j der Fourier-Reihe sind dann im Allgemeinen komplexe Zahlen. Mit der Eulerschen Formel in Gleichung (10.6) können wir die Fourier-Reihe wie folgt umschreiben:

$$f(t) = a_0 + \sum_{n=1}^{\infty} \frac{a_n - i\,b_n}{2}\, e^{i\,n\,\omega_0\,t} + \sum_{n=1}^{\infty} \frac{a_n + i\,b_n}{2}\, e^{-i\,n\,\omega_0\,t} = \sum_{n=-\infty}^{\infty} c_n\, e^{i\,n\,\omega_0\,t}, \tag{10.26}$$

mit

$$c_n = \frac{a_n - i\,b_n}{2} \quad \text{für n>0} \quad \text{und} \quad c_n = \frac{a_{|n|} + i\,b_{|n|}}{2} \quad \text{für n<0} \quad \text{und} \quad c_0 = a_0. \tag{10.27}$$

Die Koeffizienten der Fourier-Reihe berechnen sich dabei durch

$$c_n = \frac{1}{T} \int_{-T/2}^{T/2} dt\, f(t)\, e^{-i\,n\,\omega_0\,t} \tag{10.28}$$

Gleichung (10.26) besagt in Worten, dass eine periodische Funktion mit Periode T als Summe von Komponenten mit verschiedenen Frequenzen $\omega_n = n\,\omega_0 = n\,\frac{2\pi}{T}$ dargestellt werden kann. Die Fourier-Reihe ist also die Zerlegung einer periodischen Funktion in Komponenten verschiedener Frequenzen, also der Spektralzerlegung der Funktion.

Zum Beweis der Formel für die Fourierkoeffizienten stellen wir die Funktion durch ihre Fourier-Reihe dar. Dabei müssen wir aufpassen, dass der Summationsindex nicht n heißt – dieser Index ist schon der Index des gesuchten Koeffizienten. Dann gilt:

$$c_n \stackrel{!}{=} \frac{1}{T}\int_{-T/2}^{T/2} dt\, f(t)\, e^{-i\,n\,\omega_0\,t} = \frac{1}{T}\int_{-T/2}^{T/2} dt\, \left(\sum_{m=-\infty}^{\infty} c_m\, e^{i\,m\,\omega_0\,t}\right) e^{-i\,n\,\omega_0\,t}.$$

Wir vertauschen jetzt die Summe und das Integral. Es ist zwar mathematisch nicht direkt offensichtlich, dass wir das dürfen, aber wir machen es trotzdem – Physiker dürfen das meistens, denn die Natur (die wir ja beschreiben wollen) ist meistens nett. Dann erhalten wir

$$\begin{aligned} c_n &\stackrel{!}{=} \sum_{m=-\infty}^{\infty} \frac{1}{T}\int_{-T/2}^{T/2} dt\, c_m\, e^{i\,m\,\omega_0\,t}\, e^{-i\,n\,\omega_0\,t} = \sum_{m=-\infty}^{\infty} \frac{1}{T}\int_{-T/2}^{T/2} dt\, c_m\, e^{i\,(m-n)\,\omega_0\,t} \\ &= \sum_{m=-\infty}^{\infty} \frac{1}{T}\left[\frac{1}{i(m-n)\omega_0}\, c_m\, e^{i\,(m-n)\,\omega_0\,t}\right]_{-T/2}^{T/2} = \sum_{m=-\infty}^{\infty} \frac{1}{T\,\omega_0}\,\frac{c_m}{(m-n)}\,\frac{1}{i}\left(e^{i\,(m-n)\,\omega_0\,T/2} - e^{-i\,(m-n)\,\omega_0\,T/2}\right) \end{aligned}$$

Da $\omega_0 = \frac{2\pi}{T}$ ist, erhalten wir

$$c_n \stackrel{!}{=} \sum_{m=-\infty}^{\infty} \frac{1}{2\pi}\, c_m\, \frac{2}{(m-n)}\, \sin\left(\frac{(m-n)\,\omega_0\,T}{2}\right) = \sum_{m=-\infty}^{\infty} \frac{1}{2\pi}\,\frac{2\,c_m}{(m-n)}\, \sin\left((m-n)\,\pi\right)$$

Da m und n ganze Zahlen sind und $\sin(k\,\pi) = 0$ für ganze Zahlen k ist, ist fast jeder Term der Summe Null. Einzig der Term mit $m = n$ macht Probleme, denn dieser ergibt $\frac{0}{0}$. Hier nutzen wir die Regel von l'Hôpital aus Kapitel 5.2 und erhalten (mit der Kettenregel für die Ableitung des Sinus):

$$c_n \stackrel{!}{=} \lim_{m\to n} \frac{1}{2\pi}\, c_m\, \frac{2}{(m-n)}\, \sin\left((m-n)\,\pi\right) = \frac{1}{2\pi}\, c_n\, \frac{2}{1}\, \cos(0)\, \pi = c_n.$$

Wir finden also, dass die Fourier-Koeffizienten in der Tat wie behauptet berechnet werden können.

Wir finden also folgendes **Kochrezept** zur Fourier-Reihe:

Eine reell- oder komplexwertige periodische Funktion $f(t)$ mit Periode T, es gilt also $f(t) = f(t+T)$, hat die Fourier-Reihe

$$f(t) = \sum_{n=-\infty}^{\infty} c_n\, e^{i\,n\,\omega_0\,t},$$

wobei $\omega_0 = 2\pi/T$ die Kreisfrequenz der oszillierenden Funktion ist. Die Koeffizienten der Fourier-Reihe berechnen sich mittels

$$c_n \stackrel{!}{=} \frac{1}{T}\int_{-T/2}^{T/2} dt\, f(t)\, e^{-i\,n\,\omega_0\,t}.$$

Der Tiefpass-Filter in obigem Beispiel funktioniert übrigens so, dass er die Fourier-Koeffizienten des Signals $U(t)$ berechnet, und dann alle Koeffizienten außer c_{-1}, c_0 und c_1 verwirft (auf Null

setzt). Dann bleibt also noch über:

$$U(t) = \sum_n \sum_{n=-\infty}^{\infty} c_n \, e^{i\,n\,\omega_0\,t} \to c_{-1}\, e^{-i\,\omega_0\,t} + c_0 + c_{-1}\, e^{i\,\omega_0\,t}$$
$$= c_0 + (c_1 + c_{-1})\, \cos(\omega_0 t) + i\,(c_1 - c_{-1})\, \sin(\omega_0 t).$$

Im Beispiel ist $\omega_0 = 1$, $c_0 = 0$ und $c_{-1} = c_1 = 1/2$, so dass nur nur noch $U(t) \to \cos(t)$ übrigbleibt.

10.5.3 Die Fourier-Transformation

Durch die Fourier-Reihe wird eine Funktion mit Periode T als Summe von komplexen Exponentialfunktionen mit Kreisfrequenzen $\omega_n = n\,\omega_0 = n\,2\pi/T$ dargestellt. Die Werte der Frequenzen ω_n sind hier also diskret. Die Fourier-Transformation ist die Verallgemeinerung der Fourier-Reihe zu kontinuierlichen Frequenzen, $\omega_n \to \omega \in \mathbb{R}$. Da sich im Fall der diskreten Fourier-Reihe zwei benachbarte Frequenzen ω_n und ω_{n+1} um $\Delta\omega = 2\pi/T$ unterscheiden, entspricht der Übergang zu kontinuierlichen Frequenzen dem Grenzwert $T \to \infty$, also einer Funktion, die eine unendlich lange Periode hat. Das ist aber nichts anderes als eine nicht-periodische Funktion. In diesem Sinne ist die Fourier-Transformation also die Verallgemeinerung der Fourier-Reihe von periodischen Funktionen zu allen möglichen Funktionen.

Wir definieren die Fourier-Transformierte $\mathcal{F}f(\omega)$ einer im Allgemeinen komplexwertigen Funktion $f(t)$ (die wie gesagt nicht periodisch sein muss!), welche

$$\int_{-\infty}^{\infty} dt\, |f(t)|^2 < \infty \tag{10.29}$$

erfüllt, wie folgt:

$$\mathcal{F}f(\omega) = \int_{-\infty}^{\infty} dt\, e^{i\,\omega\,t}\, f(t). \tag{10.30}$$

Diese entspricht also dem Fourier-Koeffizienten c_n. Achtung: in der Praxis lässt man das $\mathcal{F}$ für die Fourier-Transformation oft weg und schreibt $f(t)$ und $f(\omega)$ – das ist aber ungute (mathematisch falsche) Notation! Die Rücktransformation erfolgt mittels

$$f(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} d\omega\, e^{-i\,\omega\,t}\, \mathcal{F}f(\omega). \tag{10.31}$$

Achtung: Bei der Definition der Fourier-Transformation gibt es leider leicht verschiedene Konventionen. Es ist völlig egal, welche der Definitionen man nutzt, solange man in einer gegebenen Rechnung immer konsistent dieselbe Definition nutzt. In manchen Definitionen wird zum Beispiel an Stelle von Gleichungen (10.30) und (10.31) das Vorzeichen des Arguments der Exponentialfunktionen umgedreht. In anderen Konventionen wird anstelle des Faktors $1/(2\pi)$ in der Rücktransformation sowohl in der Hin- wie in der Rücktransformation ein Faktor von $1/\sqrt{2\pi}$ benutzt.

Die wichtigsten Eigenschaften der Fourier-Transformation sind:

1. **Fourier-Transformation von reellen Funktionen**:

$$f(t) \in \mathcal{R} \quad \Rightarrow \quad (\mathcal{F}f(\omega))^* = \mathcal{F}f(-\omega).$$

2. **Fourier-Transformation der Ableitung**:

$$f(t) = \frac{dg(t)}{dt} \quad \Rightarrow \quad \mathcal{F}f(\omega) = -i\omega\,\mathcal{F}g(\omega).$$

3. **Skalierung der Fourier-Transformation**:

$$g(t) = f(\alpha\,t) \quad \Rightarrow \quad \mathcal{F}g(\omega) = \frac{1}{\alpha}\,\mathcal{F}f\left(\frac{\omega}{\alpha}\right).$$

4. **Fourier-Transformation eines Produkts von Funktionen**:

$$f(t) = g(t)\,h(t) \quad \Rightarrow \quad \mathcal{F}f(\omega) = \frac{1}{2\pi}\int_{-\infty}^{\infty} d\tilde{\omega}\,\mathcal{F}g(\omega-\tilde{\omega})\,\mathcal{F}h(\tilde{\omega}).$$

5. **Satz von Parceval**:

$$\int_{-\infty}^{\infty} dt\,|f(t)|^2 = \int_{-\infty}^{\infty} d\omega\,\frac{1}{2\pi}\,|\mathcal{F}f(\omega)|^2.$$

Das wäre im Physikerleben gut zu wissen:

- Wissen, was i ist.
- Komplexe Zahlen über Real- und Imaginärteil schreiben können.
- Komplexe Zahlen in Polardarstellung können.
- Betrag und Argument einer komplexen Zahl identifizieren können.
- Die Eulersche Formel auswendig kennen.
- Komplexe Zahlen addieren, subtrahieren, multiplizieren, und durch einander teilen können.
- Wissen, was eine holomorphe Funktion ist.
- Wissen, dass holomorphe Funktionen im Vergleich zu Funktionen vom $\mathbb{R}^2$ in den $\mathbb{R}^2$ stark eingeschränkt sind.
- Fourier-Reihen von periodischen Funktionen bilden können.
- Die Interpretation der Fourier-Reihe als Zerlegung einer periodischen Funktion in Komponenten verschiedener Frequenzen kennen.
- Fourier-Transformationen bilden können.
- Die aufgeführten wichtigsten Eigenschaften der Fourier-Transformation kennen.

PS: "gut zu wissen" muss nicht heißen, dass das alles in der Klausur abgefragt wird – oder, dass sonst nichts aus diesem Kapitel in der Klausur vorkommt. Wer aber mit den Punkten in diesem Kasten nichts anfangen kann, hat wahrscheinlich in der Klausur und im weiteren Studium Probleme.

Anhang A

Was sonst noch wichtig wäre

In der Physik gibt es neben den in dieser Vorlesung behandelten Rechentechniken noch eine Reihe weiterer Themen, die wirklich wichtig sind – aber eben ein bisschen weniger wichtig (oder zumindest ein bisschen weniger grundlegend) als das, was wir behandeln konnten. Diese Rechenmethoden werden Sie im Laufe des weitere Studiums noch in anderen Vorlesung sehen. Konkret hätte ich vor allem gerne noch folgende Themen behandelt:

- Weiterführende Diskussionen von Fourier-Reihe und Fourier-Transformation.
- Heavisidefunktion und Diracsche Deltafunktion (und, allgemeiner, Distributionen) – eine kurze Einführung hierzu findet sich in Anhang D.2.
- Greensche Funktionen.
- Statistik.

Sollten Sie also Spaß an Mathematik und/oder Rechenmethoden haben, sind das meine Tipps für das weitere Selbststudium.

Anhang B

Matrizen und lineare Gleichungssysteme

Ein wichtiges Anwendungsgebiet von Matrizenrechnung ist das Lösen von **linearen Gleichungssystemen**. Ein lineares Gleichungssystem ist ein System von n Gleichungen, in denen m Variablen $x_1, \ldots, x_m$ jeweils nur in erster Potenz (also linear) vorkommen. Diese Gleichungen sehen im Allgemeinen wie folgt aus:

$$\begin{aligned} A_{11}\,x_1 + A_{12}\,x_2 + \ldots + A_{1m}\,x_m &= b_1 \\ A_{21}\,x_1 + A_{22}\,x_2 + \ldots + A_{2m}\,x_m &= b_2 \\ &\vdots \quad \vdots \\ A_{n1}\,x_1 + A_{n2}\,x_2 + \ldots + A_{nm}\,x_m &= b_n \end{aligned}$$

Diese Gleichungen kann man auch als ausgeschriebene Matrixgleichung auffassen. Hierzu führen wir die Matrix A ein, deren Einträge A_{ij} genau durch die Koeffizienten A_{ij} in obigem Gleichungssystem gegeben sind. Genauso definieren wir die Vektoren $\vec{x}$ und $\vec{b}$, und schreiben das Gleichungssystem als

$$\underbrace{\begin{pmatrix} A_{11} & A_{12} & \ldots & A_{1m} \\ A_{21} & A_{22} & \ldots & A_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ A_{n1} & A_{n2} & \ldots & A_{nm} \end{pmatrix}}_{=:A} \underbrace{\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_m \end{pmatrix}}_{=:\vec{x}} = \underbrace{\begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix}}_{=:\vec{b}} \quad \Rightarrow \quad A\,\vec{x} = \vec{b}. \tag{B.1}$$

Ist $\vec{b} = 0$, so nennt man das Gleichungssystem "**homogen**". Ist $\vec{b} \neq 0$, so heißt das Gleichungssystem "**inhomogen**". Im Folgenden betrachten wir Gleichungssysteme mit n Gleichungen für n Unbekannte. Für die ist A also eine $(n \times n)$-Matrix. Wann und wie kann man solch ein Gleichungssystem lösen?

1. Falls die Determinante von A nicht verschwindet, $\mathrm{Det}(A) \neq 0$, ist das Gleichungssystem **eindeutig lösbar**. Für die Lösung benötigt man dann die inverse Matrix A^{-1} (siehe Kapitel 3.3). Der Vektor der Variablenwerte, welche die Gleichung lösen, ist dann

$$\vec{x} = A^{-1}\,\vec{b}.$$

2. Falls die Determinante von A Null ist, $\mathrm{Det}(A) = 0$, hat das Gleichungssystem **keine eindeutige Lösung** (und wir können dann auch keine Inverse von A berechnen). Es können wiederum zwei Fälle auftreten:

(a) Das Gleichungssystem hat unendlich viele Lösungen.

(b) Das Gleichungssystem hat keine Lösung.

Um herauszufinden, ob es Lösungen gibt (und von welcher Form diese sind), kann man eine Variante des Gauß-Jordan-Verfahrens anwenden:

1. Erweitere die Matrix A mit dem Vektor $\vec{b}$:

$$A \quad \Rightarrow \quad \left(\begin{array}{ccc|c} A_{11} & \dots & A_{1n} & b_1 \\ \vdots & & \vdots & \vdots \\ A_{n1} & \dots & A_{nn} & b_n \end{array}\right)$$

2. Führe so lange "elementare Zeilenumformungen" durch, bis die linke Seite der erweiterten Matrix auf die sogenannte "Dreiecksform" gebracht ist, also eine Form, bei der alle Einträge unterhalb der Hauptdiagonalen gleich Null sind. Für eine (3×3)-Matrix sieht diese z. B. wie folgt aus:

$$\left(\begin{array}{ccc|c} * & * & * & * \\ 0 & * & * & * \\ 0 & 0 & * & * \end{array}\right)$$

   Hierbei steht $*$ für eine erst mal noch beliebige Zahl. Elementare Zeilenumformungen sind:

   - Multiplizieren einer Zeile (jedes Eintrags der Zeile) mit einem Skalar.
   - Vertauschen von Zeilen.
   - Addition eines Vielfachen einer Zeile zu einer anderen (wieder Eintrag für Eintrag)

3. Wenn das Gleichungssystem unendliche viele Lösungen hat, besteht eine der Zeilen (die letzte Zeile) nur aus Nullen: $(0\ 0\ \dots|0)$. In diesem Fall kann eine der Variablen x_i frei gewählt werden. Man bestimmt die Lösungen dann wie folgt:

   - Definiere eine Matrix $\tilde{A}$ als die linke Seite (links des vertikalen Trennungsstrichs) der erweiterten Dreiecksmatrix und einen Vektor $\tilde{b}$ als die rechte Seite (rechts des vertikalen Trennungsstrichs) der erweiterten Dreiecksmatrix. Die Lösungen des Gleichungssystems sind dann Lösungen des Gleichungssystems

$$\tilde{A}\,\vec{x} = \vec{\tilde{b}},$$

     in dem die letzte Gleichung $0 \cdot x_1 + 0 \cdot x_2 + \ldots + 0 \cdot x_n = 0$ lautet. Diese Gleichung ist offensichtlich richtig. Wenn wir nun zum Beispiel x_n als die frei wählbare Variable aussuchen, können wir die anderen Gleichungen sukzessive nach den anderen x_i auflösen. Diese sind dann Funktionen der Variablen x_n.

4. Ist hingegen eine der Zeilen (die letzte Zeile) von der Form $(0\ 0\ \dots|*)$ mit $* \neq 0$, so hat das Gleichungssystem keine Lösung.

Als Beispiel betrachten wir das lineare Gleichungssystem

$$\begin{aligned} x_1 + x_2 + x_3 &= 1 \\ -x_1 + x_2 - 2\,x_3 &= 4 \\ 2\,x_1 + 2\,x_2 + 2\,x_3 &= 2 \end{aligned}$$

1. Erweitern der Matrix:
$$\left(\begin{array}{ccc|c} 1 & 1 & 1 & 1 \\ -1 & 1 & -2 & 4 \\ 2 & 2 & 2 & 2 \end{array}\right)$$

2. Elementare Zeilenumformungen:

   (a) Ziehe das Doppelte von Zeile 1 von Zeile 3 ab:
$$\left(\begin{array}{ccc|c} 1 & 1 & 1 & 1 \\ -1 & 1 & -2 & 4 \\ 2 & 2 & 2 & 2 \end{array}\right) \quad \rightarrow \quad \left(\begin{array}{ccc|c} 1 & 1 & 1 & 1 \\ -1 & 1 & -2 & 4 \\ 0 & 0 & 0 & 0 \end{array}\right)$$

   (b) Addiere Zeile 1 zu Zeile 2:
$$\left(\begin{array}{ccc|c} 1 & 1 & 1 & 1 \\ -1 & 1 & -2 & 4 \\ 0 & 0 & 0 & 0 \end{array}\right) \quad \rightarrow \quad \left(\begin{array}{ccc|c} 1 & 1 & 1 & 1 \\ 0 & 2 & -1 & 5 \\ 0 & 0 & 0 & 0 \end{array}\right)$$

3. Wir sehen also an der letzten Zeile, dass das Gleichungssystem unendlich viele Lösungen hat. Wir identifzieren jetzt
$$\tilde{A} = \begin{pmatrix} 1 & 1 & 1 \\ -1 & 1 & -2 \\ 0 & 0 & 0 \end{pmatrix} \quad \text{und} \quad \vec{\tilde{b}} = \begin{pmatrix} 1 \\ 5 \\ 0 \end{pmatrix}.$$
   Wir entscheiden uns nun, im Folgenden x_1 und x_2 als Funktion von x_3 auszudrücken. Wir erhalten dann von der zweiten Zeile des linearen Gleichungssystems $\tilde{A}\,\vec{x} = \vec{\tilde{b}}$, dass
$$2\,x_2 - x_3 = 5 \quad \Leftrightarrow \quad x_2 = \frac{5}{2} + \frac{1}{2}\,x_3.$$
   Aus der ersten Zeile von $\tilde{A}\,\vec{x} = \vec{\tilde{b}}$ erhalten wir
$$x_1 + x_2 + x_3 = 1 \quad \Leftrightarrow \quad x_1 = 1 - x_2 - x_3 = -\frac{3}{2} - \frac{3}{2}\,x_3.$$
   (Im letzten Schritt haben wir $x_2 = \frac{5}{2} + \frac{1}{2}\,x_3$ benutzt). Man kann überprüfen, dass das originale Gleichungssystem $A\,\vec{x} = \vec{b}$ durch
$$x_1(x_3) = \frac{5}{2} + \frac{1}{2}\,x_3 \quad \text{und} \quad x_2(x_3) = -\frac{3}{2} - \frac{3}{2}\,x_3 \quad \text{und} \quad x_3$$
   mit beliebigem x_3 gelöst wird.

Anhang C

Geometrische Interpretation des Transformationssatzes

Die Herleitung des Transformationssatzes ist tatsächlich recht kompliziert. Warum tritt zum Beispiel der Betrag der Determinante der Jacobi-Matrix in der Transformation der Integrationsvariablen auf? Um dies ein wenig besser zu verstehen, ist die nun folgende geometrische Interpretation des Integralsatzes hilfreich (zumindest bei zwei- und dreidimensionalen Integralen).

Kommen wir wieder zum Beispiel der Polarkoordinaten zurück. Naiv könnte man ja auch einfach nur die infinitesimalen Längen dx und dy über ihre totalen Differenziale transformieren:

$$dx = \frac{\partial x}{\partial r}\, dr + \frac{\partial x}{\partial \varphi}\, d\varphi = \cos(\varphi)\, dr - r\, \sin(\varphi)\, d\phi, \tag{C.1}$$

$$dy = \frac{\partial y}{\partial r}\, dr + \frac{\partial y}{\partial \varphi}\, d\varphi = \sin(\varphi)\, dr + r\, \cos(\varphi)\, d\phi. \tag{C.2}$$

Allerdings gilt für die Transformation des Integrals **nicht**, dass wir einfach dx und dy mit den totalen Differenzialen ersetzen:

$$\begin{aligned}\int\int dx\, dy\, f(x,y) &= \int\int dr\, d\varphi\, r\, f(x(r,\varphi), y(r,\varphi)\\ &\neq \int\int \left[\cos(\varphi)\, dr - r\, \sin(\varphi)\, d\phi\right]\, \left[\sin(\varphi)\, dr + r\, \cos(\varphi)\, d\phi\right]\; f(x(r\varphi), y(r,\varphi)\end{aligned}$$

Dies liegt daran, dass das $\mathbf{dx\,dy}$ **im Integral nicht das Produkt zweier eindimensionaler Differentiale ist, sondern ein infinitesimales Flächenelement**! Die "Verwirrung" ist letztlich darin begründet, dass diese beiden unterschiedlichen mathematischen Objekte (Produkt von Differentialen vs. Flächenelement) in kartesischen Koordinaten zufällig die gleiche Formel haben – aber eben auch nur in kartesischen Koordinaten. Die Frage ist also: was ist der richtige Ausdruck für ein infinitesimales Flächenelement in Polarkoordinaten? Hierzu wollen wir zunächst nochmal geometrisch verstehen, weshalb ein infinitesimales Flächenelement in kartesischen Koordinaten die Form $dA = dx\, dy$ hat. Betrachten wir also wieder eine Fläche, über die wir integrieren. Diese unterteilen wir in kleine Quadrate der Seitenlängen dx und dy. Es ist nun hilfreich, dieses Quadrat über Vektoren auszudrücken. Einem beliebigen Punkt auf der Fläche mit Koordinaten x und y entspricht (in der zum kartesischen Koordinatensystem passenden kartesischen Orthonormalbasis) der Vektor

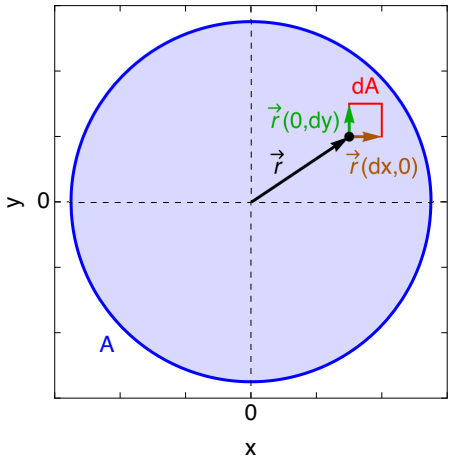
$$\vec{r}(x,y) = \begin{pmatrix} x \\ y \end{pmatrix}. \tag{C.3}$$

Wenn wir der linken unteren Ecke des Quadrats die Koordinaten (x, y) geben, so ist die linke obere Ecke gegeben durch

$$\vec{r}(x, y + dy) = \vec{r}(x, y) + \vec{r}(0, dy). \tag{C.4}$$

Die rechte untere Ecke des Quadrats hat hingegen den Ortsvektor

$$\vec{r}(x + dx, y) = \vec{r}(x, y) + \vec{r}(dx, 0). \tag{C.5}$$



Gemäß Kapitel 1.4.5 gilt, dass wir den Flächeninhalt des Quadrats als Kreuzprodukt der Vektoren $\vec{r}(0, dy)$ und $\vec{r}(dx, 0)$ verstehen können. Da das Kreuzprodukt nur für dreidimensionale Vektoren definiert ist, klappt das natürlich nur, wenn wir die Vektoren dadurch zu Vektoren im dreidimensionalen Raum machen, dass wir ihnen einfach noch eine dritte Komponente $z = 0$ geben:

$$\begin{pmatrix} x \\ y \end{pmatrix} \to \begin{pmatrix} x \\ y \\ 0 \end{pmatrix}.$$

Der Flächeninhalt ist dann gegeben durch

$$dA = |\vec{r}(dx, 0) \times \vec{r}(0, dy)| = \left| \begin{pmatrix} 0 \\ 0 \\ dx\, dy \end{pmatrix} \right| = dx\, dy. \tag{C.6}$$

Diese Kochrezept kann man nun recht einfach auf andere Koordinaten übertragen. Bei Polarkoordinaten geht das wie folgt. Ein Vektor in der Ebene hat den Ausdruck

$$\vec{r}(r, \varphi) = \begin{pmatrix} r\, \cos(\varphi) \\ r\, \sin(\varphi) \\ 0 \end{pmatrix}. \tag{C.7}$$

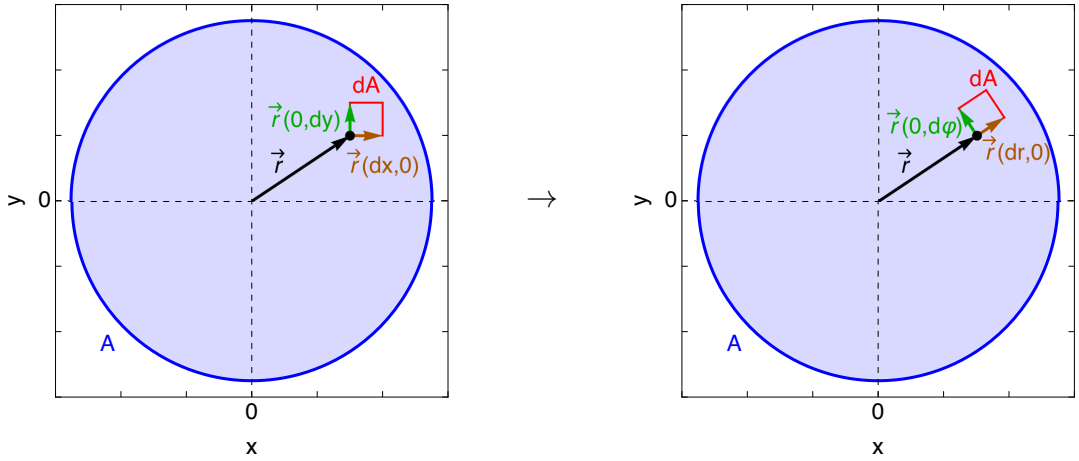
Ein infinitesimales Flächenelement erhalten wir dadurch, dass wir die Fläche des Stücks der Ebene berechnen, dass sich durch infinitesimale Änderungen von r und φ ergibt. Wenn wir r um dr ändern, ändert sich der Ortsvektor wie folgt:

$$\vec{r}(r + dr, \varphi) = \vec{r}(r, \varphi) + \vec{r}(dr, 0) = \vec{r}(r, \varphi) + \begin{pmatrix} dr\, \cos(\varphi) \\ dr\, \sin(\varphi) \\ 0 \end{pmatrix} = \vec{r}(r, \varphi) + \frac{\partial \vec{r}(r, \varphi)}{\partial r}\, dr \tag{C.8}$$

Der Vergleich mit Gleichung (7.2) zeigt, dass der letzte Term im Wesentlichen die dreidimensionale Verallgemeinerung von $\vec{r}(r,\varphi) + \vec{e}_r(r,\varphi)\,dr$ ist. Eine infinitesimale Änderung von φ ändert den Ortsvektor wie folgt (in führender Ordnung in $d\varphi$, hierzu nutzen wir eine Taylor-Entwicklung – das ist OK, denn $d\varphi$ ist infinitesimal klein, $d\varphi \to 0$, so dass die höheren Ordnungen unendlich viel kleiner sind als die erste Ordnung!):

$$\vec{r}(r,\varphi+d\varphi) = \vec{r}(r,\varphi) + \vec{r}(0,d\varphi) = \vec{r}(r,\varphi) + \begin{pmatrix} -r\,\sin(\varphi)\,d\varphi \\ r\,\cos(\varphi)\,d\varphi \\ 0 \end{pmatrix} = \vec{r}(r,\varphi) + \frac{\partial\vec{r}(r,\varphi)}{\partial\varphi}\,d\varphi. \quad \text{(C.9)}$$

Der Vergleich mit Gleichung (7.2) zeigt hier, dass der letzte Term im Wesentlichen die dreidimensionale Verallgemeinerung von $\vec{r}(r,\varphi) + \vec{e}_\varphi\, r\, d\varphi$ ist, wobei $\vec{e}_\varphi$ der Einheitsvektor in φ-Richtung ist.



Das infinitesimale Flächenelement hat also den Flächeninhalt

$$\begin{aligned} dA = |\vec{r}(dr,0) \times \vec{r}(0,d\varphi)| &= \left| \begin{pmatrix} dr\,\cos(\varphi) \\ dr\,\sin(\varphi) \\ 0 \end{pmatrix} \times \begin{pmatrix} -r\,\sin(\varphi)\,d\varphi \\ r\,\cos(\varphi)\,d\varphi \\ 0 \end{pmatrix} \right| \\ = \left| \begin{pmatrix} 0 \\ 0 \\ r\,\cos^2(\varphi)\,dr\,d\varphi - \left(-r\,\sin^2(\varphi)\,dr\,d\varphi\right) \end{pmatrix} \right| &= \left| \begin{pmatrix} 0 \\ 0 \\ r\,dr\,d\varphi \end{pmatrix} \right| = r\,dr\,d\varphi. \end{aligned} \quad \text{(C.10)}$$

Somit gilt also:

$$dA = \begin{cases} dx\,dy & \text{in kartesischen Koordinaten,} \\ r\,dr\,d\varphi & \text{in Polarkoordinaten.} \end{cases} \quad \text{(C.11)}$$

Um nun letztlich zur Determinante der Jacobi-Matrix zu kommen, erinnern wir uns noch an die Definition der Determinante aus Kapitel 3.4, und sehen, dass das Kreuzprodukt genau dem Betrag der Determinante der Jacobi-Matrix entspricht. Dies ist natürlich kein Zufall, denn letztlich haben wir das infinitesimale Flächenelement als

$$dA = \left| \left(\frac{d\vec{r}(r,\varphi)}{dr}\,dr \right) \times \left(\frac{d\vec{r}(r,\varphi)}{d\varphi}\,d\varphi \right) \right| \quad \text{(C.12)}$$

berechnet. Wenn wir das ausschreiben erhalten wir genau den Betrag der Determinante der Jacobi-Matrix. Bei dreidimensionalen Integralen muss man statt eines infinitesimalen Fächenelements ein infinitesimales Volumenelement berechnen. Dessen Volumen ist durch ein Spatprodukt gegeben (siehe Kapitel 1.4.6), was wieder genau dem Betrag der Determinante der entsprechenden (3×3)-Jacobimatrix entspricht.

Anhang D

Heavisidefunktion, Diracsche Deltafunktion, Distributionen

Dieses Kapitel des Anhangs gibt eine kurze Zusammenfassung der Heavisidefunktion und der Dirac Deltafunktion.

D.1 Heavisidefunktion

In der Physik ist es oftmals wichtig, einen bestimmten Zeitpunkt oder Ort zu benennen. Als Beispiel kann man sich vorstellen, dass ab einem Zeitpunkt $t = t_0$ ein Strom der Größe I_0 durch einen Draht fließt weil die Experimentatorin zum Zeitpunkt $t = t_0$ auf den passenden Knopf gedrückt hat. Der Strom $I(t)$ hat dann als Funktion der Zeit die Form

$$I(t) = \begin{cases} 0 & t < t_0 \\ I_0 & t \geq t_0. \end{cases} \tag{D.1}$$

Es gibt aber auch noch viele andere Situationen, die in der Physik eine Aussage der Form "Wenn der Parameter x kleiner als ein Referenzwert x_0 ist, so passiert nichts. Andernfalls passiert etwas." in eine Formel gebracht werden müssen. Man kann dazu jedes mal wieder eine Funktion mit Fallunterscheidung definieren wie wir es in Gleichung (D.1) getan haben. Um sich die Schreibarbeit zu erleichtern, kann man aber auch die nach Oliver Heaviside benannte **Heavisidefunktion** nutzen, die auch Treppen-, Stufen- oder Sprungfunktion genannt wird. Diese hat üblicherweise das Symbol Θ und die Definition

$$\Theta(x) = \begin{cases} 0 & x < 0 \\ 1 & x \geq 0. \end{cases} \tag{D.2}$$

Mit dieser Definition kann man den obigen Strom jetzt in der kompakteren Form

$$I(t) = I_0\,\Theta(t - t_0) \tag{D.3}$$

schreiben.

D.2 Diracsche Deltafunktion und Distributionen

In der Physik ist es oft hilfreich, eine Aussage vom Typ "genau dann, wenn ..." zu machen (und also auch in eine Formel zu packen). Bei Summen haben wir dazu in Kapitel 3.5.1 das Kronecker-Delta δ_{ij} eingeführt, das als

$$\delta_{ij} = \begin{cases} 1 & , i = j \\ 0 & , \text{sonst} \end{cases} \tag{D.4}$$

definiert ist. Es gilt somit, dass

$$\sum_j \delta_{ij}\, a_j = a_i \quad \text{und} \quad \sum_j \delta_{ij} = 1, \tag{D.5}$$

wobei a_i zum Beispiel die Glieder einer Folge oder Reihe sind. Weiterhin haben wir gelernt, dass Integrale als Grenzwerte von Summen zu verstehen sind, in denen die Summanden immer kleiner werden. In der Folge wollen wir nun das Konzept der "Diracschen Delta-Funktion" $\delta(x-x_0)$ einführen. Diese soll das Kronecker-Delta von Folgen/Reihen und Summen zu Funktionen und Integralen wie folgt verallgemeinern:

$$\sum_j \delta_{ij}\, a_j = a_i \;\rightarrow\; \int_{-\infty}^{\infty} dx\, \delta(x - x_0)\, f(x) = f(x_0) \quad \forall\, x_0 \tag{D.6}$$

und

$$\sum_j \delta_{ij} = 1 \;\rightarrow\; \int_{-\infty}^{\infty} dx\, \delta(x - x_0) = 1 \quad \forall\, x_0. \tag{D.7}$$

Während das Kronecker–Delta also für Indizes i und j mit diskreten Werten definiert ist, ist die Dirachsche Delta-Funktion für kontinuierliche Argumente x definiert.

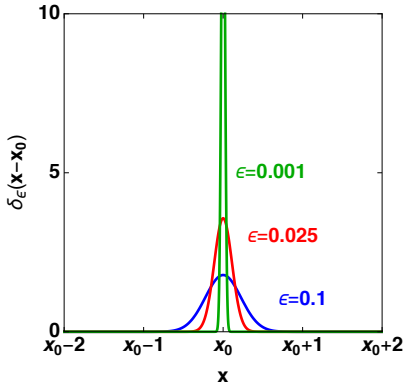
Die Diracsche Delta-"Funktion" $\delta(x)$ ist trotz dieses oft genutzten Namens übrigens mathematisch gesehen nicht wirklich eine Funktion, sondern eine sogenannte "Distribution". Distributionen sind mathematische Objekte, die Funktionen "passend" verallgemeinern – was genau das bedeutet, führt hier aber zu weit. Wir können Diracsche Delta-"Funktion" jedoch durch normale Funktionen $\delta_\epsilon(x)$ so annähern, dass wir $\delta(x)$ als Grenzwert dieser normalen Funktionen erhalten:

$$\delta(x) = \lim_{\epsilon \to 0^+} \delta_\epsilon(x). \tag{D.8}$$

Es gibt mehrere mögliche Wahlmöglichkeiten für die Funktion $\delta_\epsilon(x)$, wir nutzen hier

$$\delta_\epsilon(x) = \frac{1}{\sqrt{\pi\epsilon}}\, e^{-x^2/\epsilon}. \tag{D.9}$$

Die um einen Referenzwert x_0 auf der x-Achse verschobenen Funktionen $\delta_\epsilon(x - x_0)$ sehen wie folgt aus:



Im Grenzwert $\epsilon \to 0$ sehen wir, dass die Diracsche Deltafunktion eine unendlich hohe, unendlich dünne Funktion ist. Dies erklärt auch, warum der normale Funktionenbegriff nicht ausreicht um die Diracsche Deltafunktion zu beschreiben: normale Funktionen dürfen niemals den Funktionswert unendlich haben, es gilt jedoch $\delta(0) = \lim_{\epsilon\to 0} \delta_\epsilon(0) = \infty$. Die Diracsche Deltafunktion und die Heavisidefunktion sind übrigens enge Verwandte, denn man kann zeigen, dass gilt:

$$\frac{d\Theta(x - x_0)}{dx} = \delta(x - x_0). \tag{D.10}$$

D.2.1 Rechnen mit der Deltafunktion

Da die Diracsche Deltafunktion keine normale Funktion ist, muss man beim Rechnen mit ihr extrem vorsichtig sein. Als Faustregel gilt: bekannt sind nur die unten stehenden Rechenregeln. Diese kann man blind anwenden. Alles, was nicht ganz genau zu einer der folgenden Formeln passt, sollte man besser als "nicht bekannt" betrachten. Aus Erfahrung geht alles andere wahrscheinlich schief. Die Rechenregeln sind, für beliebige Funktionen $f(x)$:

1. Integral der Deltafunktion:

$$\int_{-\infty}^{\infty} dx\, f(x)\, \delta(x - x_0) = f(x_0). \tag{D.11}$$

   Mit $f(x) = 1$ folgt daraus auch $\int_{-\infty}^{\infty} dx\, \delta(x - x_0) = 1$.

2. Integral der Ableitung der Deltafunktion:

$$\int_{-\infty}^{\infty} dx\, f(x)\, \frac{d\delta(x - x_0)}{dx} = -\frac{df(x_0)}{dx}. \tag{D.12}$$

3. Für reelle Zahlen $a \neq 0$ gilt:

$$\int_{-\infty}^{\infty} dx\, f(x)\, \delta(a\,(x - x_0)) = \int_{-\infty}^{\infty} dx\, f(x)\, \frac{1}{|a|}\, \delta(x - x_0). \tag{D.13}$$

4. Noch allgemeiner gilt für für eine Funktion $g(x)$ mit N Nullstellen $\tilde{x}_i$, also $g(\tilde{x}_i) = 0\ \forall i$, dass

$$\int_{-\infty}^{\infty} dx\, f(x)\, \delta(g(x)) = \int_{-\infty}^{\infty} dx\, f(x) \left(\sum_{i=1}^{N} \frac{1}{|g'(\tilde{x}_i)|} \delta(x - \tilde{x}_i) \right). \tag{D.14}$$

   Gleichung (D.13) entspricht dem Spezialfall $g(x) = a\,(x - x_0)$ in dem die einzige Nullstelle $\tilde{x}_1 = x_0$ ist (und $g'(x) = a$).