

Mathematische Methoden

Johannes Berg
Institut für Theoretische Physik
Universität zu Köln

24. Februar 2016

Inhaltsverzeichnis

Inhaltsverzeichnis	1
0 Grundlagen	5
0.1 Der physikalische Raum	5
0.2 Grundlegende Begriffe	7
1 Vektoren und Vektorräume	13
1.1 Verschiebungen in der Ebene	13
1.2 Raumpunkte und Verschiebungen	16
1.3 Axiome des Vektorraums	18
1.4 Basissysteme	20
1.5 Das Skalarprodukt und die Norm	22
1.6 Das Kreuzprodukt	25
2 Analysis	29
2.1 Folgen	30
2.2 Reihen	31
2.3 Grenzwert von Funktionen	33

2.4	Differentiation	36
2.5	Taylorreihe	41
2.6	Integration	43
2.7	Mehrfachintegrale	47
3	Komplexe Zahlen	53
3.1	Komplexe Zahlen und \mathbb{R}^2	53
3.2	Realteil und Imaginärteil, Betrag und Argument komplexer Zahlen	54
3.3	Die Euler-Formel	55
3.4	Wurzeln komplexer Zahlen und der komplexe Logarithmus	56
4	Differentialgleichungen	59
4.1	Differentialgleichungen in der Physik	59
4.2	Terminologie	60
4.3	Gewöhnliche DGL erster Ordnung	61
4.4	Gewöhnliche lineare DGL	67
4.5	Partielle DGL (PDG)	77
5	Vektoranalysis	83
5.1	Integration von skalaren und vektoriellen Feldern	85
5.2	Linienintegrale	85
5.3	Differentialoperatoren I: Der Gradient	89
5.4	Das Oberflächenintegral	92
5.5	Das Volumenintegral	97
5.6	Die Rotation (engl. “curl”)	99
5.7	Die Divergenz	100
5.8	Integraltheoreme	101

“Let me end on a more cheerful note. The miracle of the appropriateness of the language of mathematics for the formulation of the laws of physics is a wonderful gift which we neither understand nor deserve. We should be grateful for it and hope that it will remain valid in future research and that it will extend, for better or for worse, to our pleasure, even though perhaps also to our bafflement, to wide branches of learning.” (Eugene Wigner)

Diese Vorlesung gibt eine Einführung in mathematische Methoden, derer sich die Physik (und viele weitere Wissenschaften) bedient. Dabei nutzt die Physik mathematische Methoden durchaus auch als Handwerkszeug, also um Rechenmethoden für ein konkretes Problem zu finden. Wichtiger noch ist aber die Suche nach dem passenden mathematischen Formalismus um bestimmte physikalische Sachverhalte zu beschreiben. Ein Beispiel ist das zweite Newtonsche Gesetz $\vec{F} = m\vec{a}$. Implizit in dieser Formulierung ist die Annahme, dass die Kraft \vec{F} und die Beschleunigung $\vec{a} \equiv d^2\vec{x}/dt^2$ Vektoren (im Ortsraum) sind¹. Gemeinsam bilden sie eine Differentialgleichung $\vec{F} = m\frac{d^2\vec{x}}{dt^2}$, ihre Lösung $\vec{x}(t)$ sagt die Bahnkurve voraus, die ein Teilchen unter dieser Kraft nimmt. Ohne Kenntnis von Differentialgleichungen und Vektoren lässt sich diese grundlegende Gleichung der Mechanik nicht verstehen.

Dabei treten oft dieselben mathematischen Konzepte in unterschiedlichen physikalischen Zusammenhängen auf. Hier ein kurzer Überblick verschiedener Teilgebiete der Physik und verwandter Gebiete, jeweils mit den mathematischen Methoden, die dort Verwendung finden.

1. Mechanik: *Differentialgleichungen (DGL), Vektorrechnung*
2. Elektromagnetismus: *DGL, Vektoranalysis*
3. Quantenmechanik: *DGL, lineare Algebra*

¹Das Gleichheitszeichen \equiv bedeutet Gleichheit im Sinne einer Definition. In einer anderen Welt könnte die z.B. Kraft auch ungleich der Masse mal Beschleunigung sein, die Beschleunigung ist aber als zweite Ableitung des Ortsvektors *definiert*.

4. Relativitätstheorie und Kosmologie: *Vektoren und Tensoren, Differentialgeometrie*
5. Teilchenphysik und Quantenfeldtheorie: Gruppentheorie, *lineare Algebra, Funktionentheorie, ...*
6. Meteorologie: *DGL, lineare Algebra*
7. Geologie: Gruppentheorie, *DGL*, Statistik

Die kursiv gesetzten Themen bilden ein Grundgerüst mathematischer Methoden für die ersten Semester des Physikstudiums, und damit die Themen dieser Vorlesungen. Der erste Teil dieses Skriptes basiert auf einer Vorlesungsmitschrift von Sebastian Breustedt und Nikolaus Rademacher, bei denen ich mich herzlich bedanke. Vielen Dank auch an Dennis Finck, Donald Köster und Steffen Karalus für engagiertes Korrekturlesen und Anmerkungen. Fehler (bestimmt noch viele) und Anregungen bitte direkt an berg@thp.uni-koeln.de. Die Lektüre dieses Skriptes ersetzt weder den Vorlesungsbesuch, noch das eigenständige Nacharbeiten des Stoffes anhand von Textbüchern, noch das selbstständige Bearbeiten der Übungen.

Wir beginnen mit einigen Vorstellungen zum Thema Raum und Koordinaten, der Bühne auf der sich physikalische Prozesse abspielen. Um das dabei benutzte Vokabular zu definieren, entwickeln wir die mathematischen Begriffe der Menge, Abbildung, und Gruppe. Diese Überlegungen dienen als Vorbereitung auf das erste Kapitel, das sich mit Vektorräumen beschäftigt.

0.1 Der physikalische Raum

Alles Geschehen in unserem Universum läuft im Raum ab, Grund genug sich mit dem dreidimensionalen Raum unserer Wahrnehmung und Erfahrung zu beschäftigen. Mit der Definition eines mathematischen Modells dieses Raumes müssen wir bis Abschnitt 1.2 warten; die Überlegungen hier dienen als Motivation, zunächst die dazu nötigen Begriffe zu definieren. Der Inhalt dieses Abschnittes ist ihnen (vermutlich) aus der analytischen Geometrie bekannt.

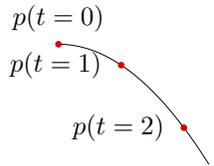
Um die Dinge einfach zu halten, betrachten wir uns zunächst einen eindimensionalen Raum, konkret eine gerade Linie _____. Dieser Raum umfasst eine (unendlich große) Menge von Punkten sowie Beziehungen zwischen diesen Punkten, d.h. Angaben wie man von einem Punkt zu einem anderen Punkt gelangt. Diese Beziehung zwischen Punkten sind Verschiebungen.

Analoges gilt für den zweidimensionalen Raum, zum Beispiel dieses Blatt Papier bzw. die Oberfläche des Bildschirms auf dem Sie diesen Text lesen. Der Hörsaal ist ein Beispiel für einen dreidimensionalen Raum.

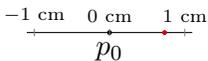
Ein Punkt p in einem solchen Raum könnte nun die Position einer Punktmasse beschreiben. Bewegt sich die Punktmasse, so hängt p von der Zeit ab; die Funktion $p(t)$ gibt zu jedem Zeitpunkt t den Punkt $p = p(t)$ an, an dem sich die Masse befindet. $p(t)$ beschreibt also die Bahnkurve an, auf der sich die Punktmasse bewegt. Die Berechnung solcher Bahnkurven ist

Ziel der Mechanik, die dazu verwendeten Differentialgleichungen werden wir in Kapitel 4 kennenlernen.

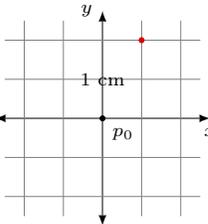
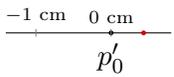
Die Abbildung links zeigt ein Beispiel für eine solche Bahnkurve in 2 Dimensionen. Eine Punktmasse bewege sich zum Zeitpunkt $t = 0$ mit einer nicht-verschwindenden Geschwindigkeit v_0 horizontal, die Schwerkraft beschleunigt die Punktmasse nach unten. Wir legen eine zweidimensionale Ebene so, dass die gesamte Bewegung innerhalb dieser Ebene verläuft (dazu muss die Ebene den Anfangspunkt $p(t = 0)$, die horizontale Achse der Anfangsgeschwindigkeit, und die vertikale Achse einschließen). Die Bahnkurve lässt sich dann in zwei Dimensionen darstellen, links erkennen sie die Form der Wurfparabel.



0.1.1 Koordinaten



Zur Beschreibung von Punkten in solchen Räumen sind **Koordinaten** nützlich. Im Beispiel des eindimensionalen Raumes genügt ein ausgezeichnete Punkt (der **Koordinatenursprung**), eine Längeneinheit (1 Meter, 1 Zoll, oder 1 Smoot), und eine Richtung um jedem Punkt auf der Linie eine Zahl, seine **Koordinate** zuzuordnen. Ursprung, Längeneinheit und Richtung definieren ein **Koordinatensystem** für den eindimensionalen Raum. Der rote Punkt im Beispiel links hat zum Beispiel die Koordinate $x = 0.732$ cm. Die Koordinaten eines Punktes hängen natürlich von der Wahl des Koordinatensystems ab. Die zweite Abbildung links zeigt denselben Raum mit demselben Punkt. In diesem Beispiel ist allerdings der Ursprung p'_0 anders gewählt, entsprechend ist die Koordinate des Punktes eine andere, hier $x' = x - 0.3$.



Analog lässt sich jeder Punkt der zweidimensionalen Ebene mit zwei Koordinaten beschreiben. Ein besonders einfaches Koordinatensystem besteht aus einem Punkt p_0 als Ursprung und zwei rechtwinkligen Achsen; solche Koordinaten werden als **kartesische Koordinate** bezeichnet. In dem Beispiel links hat der rote Punkt die Koordinaten $x = 1$ cm und $y = 2$ cm. Weitere Koordinatensysteme werden wir später kennenlernen. Analog ist ein Punkt im dreidimensionalen Raum durch 3 Koordinaten charakterisiert.

Eine Bahnkurve wie das Beispiel oben kann nun durch eine Funktion angegeben werden, die zu jedem Zeitpunkt t die Koordinaten des Punktes angibt, an dem sich eine Masse befindet. Im Beispiel mit der Wurfparabel sind diese Funktionen $x(t) = v_0 t$ und $y(t) = -\frac{1}{2} g t^2$. Zweifaches Ableiten nach der Zeit führt auf $\frac{d^2 x}{dt^2} = 0$ und $\frac{d^2 y}{dt^2} = -g$, die Beschleunigung in x -Richtung ist also Null, in y -Richtung ist sie die Fallbeschleunigung $-g$. Die Koordinaten des Massepunktes gehorchen also Differentialgleichungen (hier dem zweiten Newtonschen Gesetz, Beschleunigung gleich Kraft durch Masse). Im Kapitel 4 werden wir solche Gleichungen systematisch formulieren und lösen.

0.2 Grundlegende Begriffe

Mit diesen einfachen Betrachtungen haben wir bereits recht weit vorgegriffen und Begriffe wie Menge, Funktion, Abbildung verwendet. Im Folgenden werden wir diese Begriffe definieren und eine einheitliche Notation entwickeln.

0.2.1 Mengen

Die **Menge** ist ein grundlegender Begriff der Mathematik, er bezeichnet eine Zusammenfassung von wohlunterscheidbaren Objekten. Ein Beispiel ist die Menge der natürlichen Zahlen, die wir als $\mathbb{N} = \{1, 2, 3, \dots\}$ bezeichnen. Die Zahl 13 ist ein Element dieser Menge, wir schreiben $13 \in \mathbb{N}$. Analog gilt "Donnerstag" $\notin \mathbb{N}$. Die **Teilmenge** (oder auch **Untermenge**) einer Menge M enthält Elemente dieser Menge, z.B. sind die natürlichen Zahlen eine Teilmenge der reellen Zahlen \mathbb{R} , wir schreiben $\mathbb{N} \subset \mathbb{R}$. Die Teilmenge von M , deren Elemente die Eigenschaft E haben, schreiben wir als $\{x|x \in M, E(x)\}$. Zum Beispiel schreiben wir die Menge der geraden Zahlen als $\{x|x \in \mathbb{N}, x/2 \in \mathbb{N}\}$.

Die **Vereinigung** zweier Mengen sind alle Elemente, die in der einen oder der anderen Menge vertreten sind, $M \cup N = \{x|x \in M \text{ oder } x \in N\}$. Der **Durchschnitt** zweier Mengen sind alle Elemente die in der einen und der anderen Menge vertreten sind, $M \cap N = \{x|x \in M \text{ und } x \in N\}$.

Die Menge, die kein Element enthält, wird als leere Menge bezeichnet und als \emptyset notiert.

Als nächstes definieren wir das **Produkt** $A \times B$ zweier Mengen A, B , als die Menge aller geordneten Paare der Elemente aus A und B . Für $A = \{\text{Peter, Ida}\}$ und $B = \{\text{Müller, Maier, Schmidt}\}$ ist $A \times B$ gleich der Menge $\{(\text{Peter, Müller}), (\text{Peter, Maier}), (\text{Peter, Schmidt}), (\text{Ida, Müller}), (\text{Ida, Maier}), (\text{Ida, Schmidt})\}$. Geordnet bedeutet, dass das erste Element eines solchen Paares aus der ersten Menge A stammt, und das zweite aus der zweiten Menge B . Die Menge $B \times A$ enthält damit $\{(\text{Müller, Peter}), \dots\}$.

0.2.2 Abbildungen

Eine **Abbildung** f zwischen zwei Mengen ist eine Zuordnung, die jedem Element einer Menge X (der **Definitionsmenge**) jeweils ein Element einer zweiten Menge Y (**Zielmenge**) zuweist. Hierfür verwenden wir die folgende Schreibweise $f : X \rightarrow Y; x \mapsto f(x)$, wobei $x \in X$ und $f(x) \in Y$. Im täglichen Gebrauch sind wir meist etwas lax und sprechen von der Funktion $f(x)$, meinen aber eine solche Abbildung. Sogenannte **reelle Funktionen** sind ein häufig auftretender Spezialfall, sie haben als Zielmenge die reellen Zahlen \mathbb{R} . Oft ist auch die Definitionsmenge die Menge der reellen Zahlen, die Abbildung $f : \mathbb{R} \rightarrow \mathbb{R}; x \mapsto f(x)$ ist das was Sie bereits als "Funktion einer Veränderlichen" kennengelernt haben, und wir werden die Sprechweise auch im Alltag beibehalten. Funktionen dieser Art mit $f(-x) = f(x)$

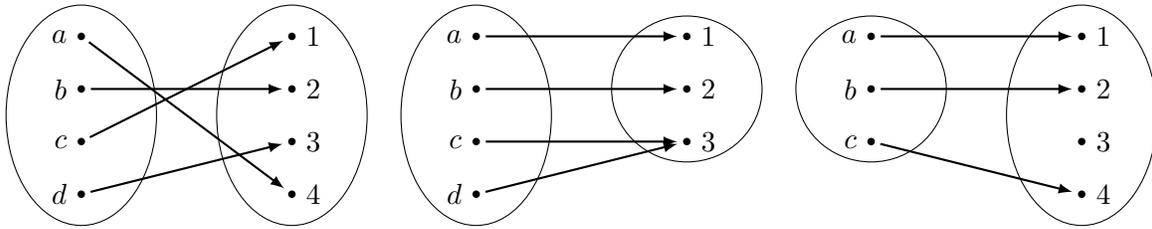


Abbildung 0.1: Je ein Beispiel für bijektive, eine surjektive, und eine injektive Abbildung (von links nach rechts).

für alle x heißen **gerade Funktionen**, solche mit $f(-x) = -f(x)$ heißen **ungerade Funktionen**.

Beispiel

Ein Paar reeller Zahlen ist ein Element der Menge $\mathbb{R} \times \mathbb{R}$. Die Addition von zwei reellen Zahlen a, b ist eine Abbildung aus der Menge der Paare reeller Zahlen auf die Menge der reellen Zahlen, also $+$: $\mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R}, (a, b) \mapsto a + b$. Machen Sie sich klar, dass sich hier $\mathbb{R} \times \mathbb{R}$ nicht auf die Operation der Multiplikation zweier Zahlen bezieht, sondern auf die Bildung der Paare reeller Zahlen (die Paare von Zahlen, die dann addiert werden).

Ist jedes Element der Zielmenge genau ein Abbild eines Elementes der Definitionsmenge, heißt die Abbildung **bijektiv**. Insbesondere existiert dann ein Inverses dieser Abbildung $f^{-1} : Y \rightarrow X, y \mapsto f^{-1}(y) = x$. Gibt es eine bijektive Abbildung zwischen zwei Mengen, so heißen die Mengen **gleichmächtig**. Mit dieser Definition kann man auch die Anzahl von Elementen in Mengen mit unendlich vielen Elementen vergleichen.

Tritt jedes Element der Zielmenge mindestens einmal als Abbildung der Definitionsmenge auf, heißt die Abbildung **surjektiv**, d.h. $f(X)=Y$. Eine Abbildung heißt **injektiv** wenn jedem Element der Zielmenge höchstens ein Element der Definitionsmenge zugeordnet ist. Damit ist eine Abbildung dann und nur dann bijektiv wenn sie sowohl surjektiv als auch injektiv ist.

Abbildungen lassen sich hintereinander ausführen. Folgt z.B. auf die Abbildung f von der Menge X auf die Menge Y noch eine weitere Abbildung g von Y auf die Menge Z so schreiben wir die kombinierte Abbildung $X \rightarrow Z$ als $g \circ f : X \rightarrow Z, x \mapsto z = g(f(x))$.

0.2.3 Gruppen

Die minimale Struktur, die man einer Menge sinnvollerweise geben kann, führt auf das Konzept der Gruppe. Eine **Gruppe** ist eine Menge M sowie eine Verknüpfung der Elemente dieser Menge mit bestimmten Eigenschaften. Diese Verknüpfung nimmt zwei Elemente

von M und generiert aus ihnen ein neues Element von M . Die Verknüpfung ist also eine Abbildung $M \times M \rightarrow M$, $(x, y) \mapsto z = x * y$ und sie gehorcht den folgenden Forderungen (den Gruppenaxiomen ¹):

1. dem Assoziativgesetz $x * (y * z) = (x * y) * z$
2. der Existenz eines neutralen Elements e mit $x * e = e * x = x$ für alle $x \in M$.
3. der Existenz von inversen Elementen: für alle $x \in M$ existiert ein x^{-1} so dass $x * x^{-1} = x^{-1} * x = e$.

Dabei ist bereits die Eigenschaft der Verknüpfung $M \times M \rightarrow M$ nichttrivial, sie bedeutet dass die Verknüpfung stets auf Elemente von M führt, und nicht auf andere Objekte. Man sagt in diesem Zusammenhang, dass M unter der Verknüpfung **abgeschlossen** ist.

Ein einfaches Beispiel für eine Gruppe sind die reellen Zahlen, verknüpft mit der Operation der Addition. Die natürlichen Zahlen bis 9 und die Null, $M = \{0, 1, \dots, 9\}$ bilden allerdings keine Gruppe unter der Addition, da diese Menge nicht abgeschlossen ist: $3 + 8$ ist kein Element von M . Auch die natürlichen Zahlen mit Null bilden keine Gruppe unter Addition (fehlendes Inverses).

Aus diesen Axiomen folgen bereits eine Reihe interessanter Eigenschaften, z.B. kann es nur ein neutrales Element geben, denn gäbe es zwei neutrale Elemente e und e' dann gälte $e = e * e' = e' * e = e'$. Analog gibt es zu jedem Element x nur ein inverses Element x^{-1} . Angenommen es gäbe ein Element a mit zwei Inversen b und c , dann gilt unter Nutzung des ersten Axioms $b = b * e = b * (a * c) = (b * a) * c = e * c = c$.

¹(Für Puristen:) Diese Forderungen lassen sich noch weiter reduzieren. Zum Beispiel genügt in 2. die Forderung, dass $x * e = x \forall x$ (e wird dann als rechtsneutrales Element bezeichnet). Daraus kann man dann leicht folgern, dass auch $e * x = x$; beginnen wir mit $x * e = x$ und multiplizieren links mit x^{-1} und rechts mit x .

Beispiel

Gruppen spielen eine zentrale Rolle in der modernen Physik bei der Beschreibungen von Symmetrien. Zum Beispiel bildet die Menge der Rotationen um eine feste Achse um Vielfache von 60° eine Gruppe: Ihre Elemente sind $0^\circ, 60^\circ, 120^\circ, 180^\circ, 240^\circ, 300^\circ$. Die Verknüpfung zweier Rotationen ist definiert als ihre Hintereinanderausführung. Das neutrale Element ist damit die Rotation um 0° , die Inverse von z.B. 60° ist das Element 300° .

In diesem Beispiel ist die Verknüpfung **kommutativ**, d.h. $x * y = y * x$ für alle $x, y \in M$. Gruppen mit einer kommutativen Verknüpfung heißen Abelsche Gruppen. Im Allgemeinen ist allerdings $x*y \neq y*x$! Ein Beispiel sind Rotationen um 90° und ihre Vielfache um die x -Achse oder um die y -Achse. Nehmen sie ein Buch zur Hand (ein Smartphone geht notfalls auch), legen Sie es auf den Tisch, rotieren Sie es um 90° um die Achse die durch den Tisch geht. Dann rotieren Sie es um 90° um eine Achse senkrecht dazu. Beginnen Sie noch einmal von vorn, und führen Sie nun diese Operationen in umgekehrter Reihenfolge aus. Die Endposition des Buches wird in den beiden Fällen nicht die Gleiche sein!

Beispiel

Ein weiteres Beispiel einer Gruppe ist die Menge der Permutationen. Eine **Permutation** von Objekten ist ihre Anordnung in einer bestimmten Reihenfolge, beziehungsweise eine Transformation, die die Reihenfolge in der Objekte angeordnet sind, ändert. Beispiel für eine solche Transformation ist die Transformation, die A, B, C in C, A, B überführt. Dabei wurde das Objekt an erster Stelle an die 2. Stelle umgesetzt, das zweite an die 3. Stelle, und das letzte Element an die 1. Stelle. Eine Möglichkeit diese Transformation zu notieren ist $(2,3,1)$. (Vorsicht, diese Transformation überführt die Reihenfolge $1,2,3$ in $3,1,2$.)^a

Für die $n = 3$ Zahlen $1, 2, 3$ gibt die Permutationen $(1, 2, 3), (1, 3, 2), (2, 1, 3), (2, 3, 1), (3, 1, 2), (3, 2, 1)$.

In jeder Permutation tritt jede der 3 Zahlen genau einmal auf, aber jeweils auf unterschiedlichen Positionen. Platzieren wir die Zahl eins zuerst, gibt es für sie 3 mögliche Positionen (an erster, zweiter, oder dritter Stelle), platzieren wir als nächstes die Zahl zwei sind noch zwei Möglichkeiten geblieben, und die 3 kommt in die letzte verbliebene Position. Insgesamt gibt es also $1 \times 2 \times 3 = 6$ Permutationen von 3 Objekten. Für n Elementen gibt es $n! \equiv 1 \times 2 \times 3 \dots n$ Permutationen.

Man kann Permutationen verknüpfen, indem man sie hintereinander ausführt. Nummerieren wir (willkürlich) für $n = 3$ die 6 Permutationen in der obigen Reihenfolge mit $\sigma_1, \sigma_2, \sigma_3, \sigma_4, \sigma_5, \sigma_6$. Wenn wir jetzt zuerst $\sigma_2 = (1, 3, 2)$ anwenden und daraufhin $\sigma_3 = (2, 1, 3)$, dann wird z.B. das erste Element zunächst an der 1. Stelle belassen, und dann an die 2. Stelle verschoben. Das Ergebnis schreiben wir als $\sigma_3 * \sigma_2$ (wie bei der Verknüpfung von Abbildungen von rechts nach links gelesen) und erhalten $\sigma_3 * \sigma_2 = (2, 3, 1) = \sigma_4$.

Das erste Element $\sigma_1 = (1, 2, 3)$ entspricht dem neutralen Element, denn es lässt die Reihenfolge unverändert. Permutationen kommutieren nicht, z.B. ist der erste Eintrag von $\sigma_2 * \sigma_3$ gleich 3. Jede Permutation läßt sich durch eine Serie paarweiser Vertauschungen generieren, benötigt man eine gerade (ungerade) Zahl von Vertauschungen spricht man von einer geraden

(ungeraden) Permutation.

^aZur Charakterisierung einer Permutation kann man alternativ auch das Ergebnis der Permutation auf 1,2,3 nutzen, und in der Vorlesung haben wir das auch gemacht. Die beiden Darstellungen enthalten die gleiche Information. Die oben genannte Schreibweise ist für konkrete Rechnungen praktischer.

Schließlich noch etwas Notation. In unseren Definitionen haben wir häufig die Wendungen “für alle x ” oder “es existiert” benutzt. Wir kürzen sie ab mit \forall als “für alle” und \exists für “es existiert”.

Info: Mit ein wenig mehr Struktur als eine Gruppe kommen wir bereits bei Objekten wie Zahlen an: ein **Körper** ist eine Menge mit einer als ‘additiv’ bezeichneten Verknüpfung, die den Gruppenaxiomen gehorcht und kommutativ ist, sowie einer als ‘multiplikativ’ bezeichneten Verknüpfung, die (bis auf das additive Nullelement) ebenso eine Abelsche Gruppe bildet. Die Menge rationalen Zahlen oder der reellen Zahlen bilden einen solchen Körper.

Vektoren spielen eine zentrale Rolle bei der Beschreibung von physikalischen Vorgängen im dreidimensionalen Raum oder der vierdimensionalen Raumzeit. Wir beginnen mit einem einfachen Beispiel, der Verschiebung in der Ebene, bevor wir uns dem Begriff des Vektorraums auf allgemeinerem (und etwas abstrakterem) Niveau nähern.

1.1 Verschiebungen in der Ebene

Wir betrachten Verschiebungen in der Ebene.

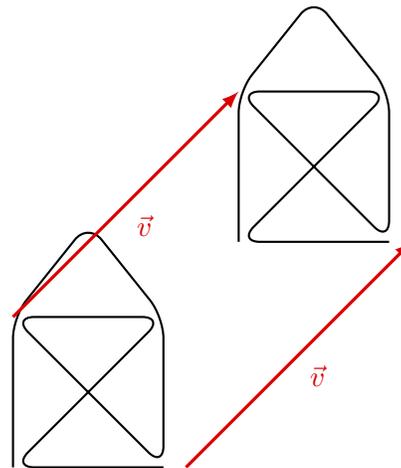
Eine Verschiebung \vec{v} bildet jeden Punkt p der Ebene auf einen anderen Punkt p' ab. Abstände der Punkte untereinander und ihre Orientierung zueinander bleiben erhalten. Alle Verschiebungspfeile der Abbildung sind also einander äquivalent, und sollen als eine einzelne Verschiebung \vec{v} betrachtet werden.

Zwei Verschiebungen \vec{v} und \vec{w} lassen sich kombinieren ("verknüpfen"), indem man erst \vec{v} und dann \vec{w} ausführt. Das Ergebnis ist eine neue Verschiebung, die wir \vec{u} nennen. Wir schreiben

$$\vec{u} = \vec{w} + \vec{v} . \quad (1.1)$$

Anstelle des Pluszeichens $+$ hätten wir auch ein beliebiges anderes Symbol nutzen können für eine Verknüpfung, die aus zwei Verschiebungen eine dritte generiert. Die Schreibweise mit dem Pluszeichen nutzt aus, dass diese Verknüpfung zweier

Verschiebungen Eigenschaften hat, die uns aus der Verknüpfung zweier Zahlen durch Addition geläufig sind. Wir bezeichnen sie daher auch als die 'Addition' zweier Verschiebungen.

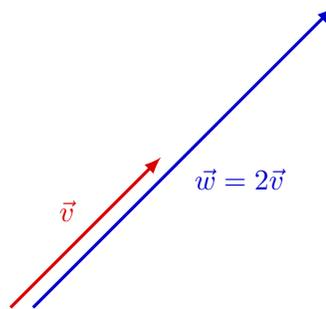
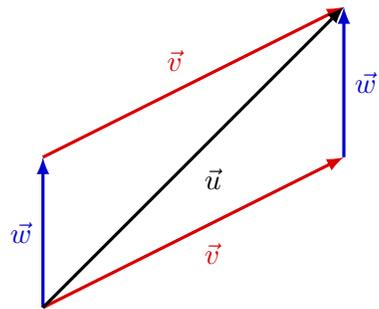


Es gilt nämlich $\vec{u} = \vec{v} + \vec{w}$, also $\vec{v} + \vec{w} = \vec{w} + \vec{v}$ und ebenso $(\vec{v} + \vec{w}) + \vec{x} = \vec{v} + (\vec{w} + \vec{x})$, die Hintereinanderausführung von Verschiebungen ist also kommutativ und assoziativ. Es existiert auch eine Verschiebung, die jeden Punkt auf sich selbst abbildet (das sogenannte Nullelement). Diese Eigenschaften hat die Hintereinanderausführung von Verschiebungen gemeinsam mit der Addition von Zahlen: Die Verschiebungen bilden eine Gruppe, wenn man sie durch Hintereinanderausführung verknüpft (wir sagen auch sie bilden eine Gruppe unter Hintereinanderausführung).

Diese Eigenschaften von Verschiebungen haben wir übrigen nicht bewiesen. Diese folgen aus unserer Alltagserfahrung; streng genommen handelt es sich um experimentelle Ergebnisse, aus denen wir fordern die Verknüpfung von Verschiebungen in der Ebene möge kommutativ und assoziativ sein.

Zusätzlich zur Addition von zwei Verschiebungen lässt sich die Multiplikation einer Verschiebung \vec{v} mit einer Zahl λ als Streckung von \vec{v} um einen Faktor λ definieren. Dabei ist dann $\vec{v} + \vec{v} = 2\vec{v}$. Bei $\lambda < 0$ ist die Streckung in Gegenrichtung gemeint. Es gelten $\lambda\vec{v} + \nu\vec{v} = (\lambda + \nu)\vec{v}$ und $\lambda(\nu\vec{v}) = (\lambda\nu)\vec{v}$ (wieder experimentelles Ergebnis, bzw. Forderung).

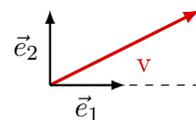
Durch Addition von Verschiebungen entsteht wieder eine Verschiebung (und nicht ein anderes Objekt), ebenso wie durch die Multiplikation einer Verschiebung mit einer Zahl wieder eine Verschiebung entsteht. Kurz: Verschiebungen kann man addieren (definiert durch Hintereinanderausführung) und mit Zahlen multiplizieren (definiert durch Streckung). Weitere physikalische und mathematische Objekte mit diesen beiden Eigenschaften werden wir im Abschnitt 1.3 finden. Solche Objekte werden wir dort als Vektoren bezeichnen. So bildet die Menge der Verschiebungen mit den Operationen Hintereinanderausführung und Streckung einen sogenannten Vektorraum.



1.1.1 Basissysteme von Verschiebungen

Basissysteme erlauben, Verschiebungen durch Zahlen zu beschreiben, was konkrete Rechnungen stark vereinfacht.

Wir beginnen damit, zwei ausgezeichnete Richtungen in der Ebene ('rechts' und 'oben') und eine Längeneinheit ('1 m') zu wählen. Die Verschiebung \vec{e}_1 verschiebe alle Punkte um eine Längeneinheit nach rechts, \vec{e}_2 verschiebe alle Punkte um eine Längeneinheit nach oben.



Alle Verschiebungen lassen sich dann durch Streckungen von \vec{e}_1 und \vec{e}_2 und Hintereinanderausführung der resultierenden Verschiebungen generieren, z.B. $\vec{v} = 2\vec{e}_1 + \vec{e}_2$. Ein Basissystem aus Verschiebungen, die senkrecht aufeinanderstehen (dazu später mehr) bezeichnet man als **kartesische Basis**. Haben die Verschiebungen auch noch alle die gleiche Länge spricht man von einem **orthonormalen** Basissystem. Eine Basis muss allerdings nicht aus solchen orthonormalen Verschiebungen bestehen, in der Praxis hat eine solche Basis jedoch klare Vorteile (einen werden wir in Abschnitt 1.5.1 kennenlernen).

Jede Verschiebung \vec{v} in der Ebene ist damit durch 2 Zahlen charakterisiert, $\vec{v} = v_1\vec{e}_1 + v_2\vec{e}_2$. Diese Zahlen v_1 und v_2 heißen **Komponenten** von \vec{v} in der Basis \vec{e}_1, \vec{e}_2 . Die Komponenten der Verschiebung \vec{v} lassen sich in Spaltenschreibweise zusammenfassen,

$$\mathbf{v} = \begin{pmatrix} v_1 \\ v_2 \end{pmatrix},$$

wobei \mathbf{v} aus den Komponenten der Verschiebung in der gewählten Basis besteht. \mathbf{v} wird auch als **Darstellung** der Verschiebung \vec{v} in der gewählten Basis bezeichnet, oder als **Koordinatenvektor**. Die Komponenten werden dann als Koordinaten des Vektors bezeichnet (nicht zu verwechseln mit den Koordinaten eines Punktes im Raum aus Abschnitt 0.1.1). Gegeben die Basis, bezeichnet jedes \mathbf{v} eindeutig eine Verschiebung \vec{v} . Trotzdem handelt es sich um unterschiedliche Objekte, \vec{v} bezeichnet eine Verschiebung (ist also unabhängig von der Wahl einer Basis), \mathbf{v} die Komponenten von \vec{v} in einer bestimmten Basis¹. Basisvektoren haben in ihrer Basis stets die Darstellung $(1, 0, 0, \dots), (0, 1, 0, \dots), \dots$. Die Aussage $\mathbf{e}_1 = (1, 0, \dots)$ ist korrekt, aber recht inhaltsleer, wenn nicht gesagt wird was \vec{e}_1 ist (z.B. horizontale Verschiebung Richtung Osten um 1 m).

Für die Addition zweier Verschiebungen gilt nun

$$\vec{v} + \vec{w} = (v_1\vec{e}_1 + v_2\vec{e}_2) + (w_1\vec{e}_1 + w_2\vec{e}_2) = (v_1 + w_1)\vec{e}_1 + (v_2 + w_2)\vec{e}_2 \quad (1.2)$$

(der zweite Schritt folgt aus der Assoziativität der Addition von Verschiebungen und der Forderung $\lambda\vec{v} + \nu\vec{v} = (\lambda + \nu)\vec{v}$ an die Streckung). In Komponentenschreibweise

$$\mathbf{v} + \mathbf{w} = \begin{pmatrix} v_1 + w_1 \\ v_2 + w_2 \end{pmatrix}. \quad (1.3)$$

Dieses Ergebnis ist nicht zu unterschätzen! Wir haben nun eine Rechenregel, die die Addition (Hintereinanderausführung) von Verschiebungen durch die Addition von Zahlen

¹In Textbüchern wird oft kein Unterschied zwischen diesen Objekten gemacht, der Sinn erschließt sich allerdings meist aus dem Kontext.

beschreibt: Gegeben zwei Verschiebungen, addieren wir einfach ihre Komponenten und erhalten die Komponenten der Summe der Verschiebungen.

Aufgabe

Zeigen Sie analog, dass die Verschiebung $\vec{w} = \lambda\vec{v}$ die Komponenten $w_1 = \lambda v_1$ und $w_2 = \lambda v_2$ hat.

Info: Beim Rechnen mit Vektoren und ihren Komponenten treten häufig Summen auf, zum Beispiel $v_1\vec{e}_1 + v_2\vec{e}_2 + v_3\vec{e}_3$. Um unnötige Schreibarbeit zu vermeiden, nutzen wir das Summenzeichen \sum . Zum Beispiel ist $1 + 2 + 3$ die Summe der natürlichen Zahlen i von 1 bis 3, was wir kompakt als $\sum_{i=1}^3 i$ schreiben. Statt i hätten wir natürlich auch eine andere Variable nutzen können. Analog ist $v_1\vec{e}_1 + v_2\vec{e}_2 + v_3\vec{e}_3 = \sum_{i=1}^3 v_i\vec{e}_i$. Die Grenzen der Summe schreiben wir oft nicht mit, wenn sie aus dem Kontext folgen (z.B. in 3 Dimensionen wenn i von 1 bis 3 läuft). Doppelte Summen wie $\sum_{i=1}^3 \sum_{j=1}^3 ij$ schreiben wir noch kompakter als $\sum_{i,j=1}^3 ij$.

1.2 Raumpunkte und Verschiebungen

Es mag verwunderlich erscheinen, dass wir als Beispiele für Vektoren Verschiebungen genutzt haben, aber noch nicht über ‘Ortsvektoren’ gesprochen haben. Tatsächlich sind die Punkte einer Ebene oder des dreidimensionalen Raumes keine Elemente eines Vektorraumes, denn man kann Punkte im Raum nicht sinnvoll addieren. Mit der Wahl eines ausgezeichneten Punktes, des sogenannten **Ursprungs** können wir jedoch jeden Punkt im Raum durch eine Verschiebung beschreiben.

Beispiel

Die Universität befindet sich in Köln-Sülz am Punkt p . Anstatt der Beschreibung dieses Punktes im Raum, z.B. durch ein Koordinatensystem, kann man aber genauso gut angeben, welche Verschiebung \vec{r} uns vom Koordinatenursprung p_0 (der in diesem Beispiel natürlich am Dom liegt) zur Universität bringt, wir schreiben $p = p_0 + \vec{r}$ und haben damit jedem Punkt einen Vektor (und nur einen) zugeordnet.

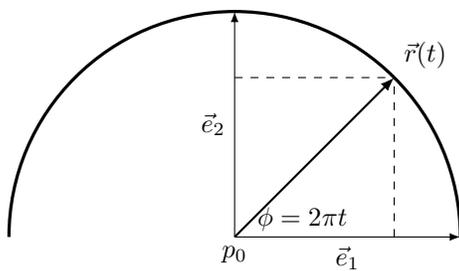
Nach Wahl eines Ursprungs lassen sich Punkte im Raum als Verschiebungen vom Ursprung, und damit als Vektoren beschreiben. Etwas präziser: jeder Punkt entspricht genau einer Verschiebung, und jede Verschiebung genau einem Punkt. Die Verschiebung \vec{r} wird als **Ortsvektor** bezeichnet. Sie beschreibt einen Raumpunkt allerdings nur bei gegebenem Ursprung.

Eine wichtige Anwendung des Konzepts des Ortsvektors sind sogenannte **Bahnkurven**. Denken wir an eine Punktmasse, die sich durch den Raum bewegt. Zu jedem Zeitpunkt t

befindet sie sich an einem anderen Punkt p im Raum. Die Bahnkurve ist also durch eine Funktion der Zeit $p(t)$, oder äquivalent bei Wahl eines Ursprungs durch den zeitabhängigen Ortsvektor $\vec{r}(t)$ beschrieben. Wählen wir nun noch eine (feste, zeitlich unveränderte) Basis für die Ortsvektoren, beschreiben die Komponenten $\mathbf{r}(t)$ die Bewegung der Punktmasse. Verschiebungen in der Ebene sind durch ein Zahlenpaar charakterisiert, kompakt schreiben wir $\mathbf{v} \in \mathbb{R} \times \mathbb{R} \equiv \mathbb{R}^2$. Als Beispiel für Abbildungen nutzen wir noch einmal die obige Raumkurve $\mathbf{r}(t)$. Formal betrachtet ist eine Raumkurve in zwei Dimensionen eine Abbildung aus den reellen Zahlen \mathbb{R} auf die Verschiebungen in der Ebene \mathbb{R}^2 ,

$$\begin{aligned} \mathbf{r}(t) : \mathbb{R} &\rightarrow \mathbb{R}^2 \\ t &\mapsto \mathbf{r} = \mathbf{r}(t) = \begin{pmatrix} r_1(t) \\ r_2(t) \end{pmatrix}. \end{aligned} \quad (1.4)$$

Beispiel



Betrachten wir einen Punkt der sich mit gleichförmiger Geschwindigkeit auf einem Halbkreis mit Radius 1 bewegt. Zur Zeit $t = 0$ beginne die Bewegung, zur Zeit $t = 1/2$ sei das Ziel erreicht. Legen wir p_0 in das Zentrum des Halbkreises und wählen die Basis entlang der Achsen wie in der Abbildung, ist die Bahnkurve

$$\mathbf{r}(t) = \begin{pmatrix} \cos(2\pi t) \\ \sin(2\pi t) \end{pmatrix}. \quad (1.5)$$

Aufgabe

Berechnen Sie die Komponenten einer dreidimensionalen spiralförmigen Bahnkurve um die z -Achse (Wendeltreppe, Korkenzieher). Vergleichen Sie Ihr Ergebnis mit

$$\mathbf{r}(t) = \begin{pmatrix} \cos(2\pi t) \\ \sin(2\pi t) \\ v_z t \end{pmatrix}. \quad (1.6)$$

Info: Für den physikalischen Raum (ohne Wahl eines Ursprungs) ist ein sogenannter affiner Raum die adäquate Beschreibung. Ein **affiner Raum** $(M, V, +)$ ist eine Menge M von Punkten, ein Vektorraum V und eine Verknüpfung $+$

$$\begin{aligned} M \times V &\rightarrow M \\ (p, \vec{v}) &\mapsto p + \vec{v} \in M \end{aligned}$$

mit der Eigenschaft $p + (\vec{u} + \vec{v}) = (p + \vec{u}) + \vec{v}$. Zudem existiert zu jedem Paar $p, q \in M$ genau ein Vektor $\vec{v} \in V, p = q + \vec{v}$.

Das heißt, Punkte $p, q \in M$ des affinen Raumes können subtrahiert werden, $p - q = \vec{v}$ (das Ergebnis ist die Verschiebung von q nach p). Punkte des affinen Raumes können aber nicht addiert werden, und bilden daher auch keinen Vektorraum.

Als Alternative zu den Koordinaten aus Abschnitt 0.1.1 kann als Koordinatensystem eines affinen Raumes auch ein ausgezeichneter Punkt p_0 (der Koordinatenursprung) und eine Basis $\{\vec{e}_i\} = \vec{e}_1, \vec{e}_2, \dots$ von V genutzt werden. Jeder Punkt $p \in M$ kann dann als

$$p = p_0 + \sum_{i=1}^k v_i \vec{e}_i \quad (1.7)$$

geschrieben werden, und wird mit dem Verschiebungsvektor \vec{v} von p_0 nach p identifiziert. \vec{v} kann dann als Ortsvektor bezeichnet werden. Diese Zuordnung ist allerdings nur möglich, nachdem wir p_0 (willkürlich) gewählt haben, während die Punkte im Raum unabhängig von einer solchen Wahl existieren; es handelt sich beim Ortsvektor \vec{v} also um eine Darstellung von p bezüglich p_0 .

1.3 Axiome des Vektorraums

Mengen, für deren Elemente eine Additionsoperation und eine Multiplikation mit Zahlen existiert treten in den unterschiedlichsten Zusammenhängen auf. Als konkretes Beispiel werden wir zwar weiter die Menge der Verschiebungen benutzen, es lassen sich allerdings auch für ganz andere Objekte die Operationen Addition und Multiplikation mit Zahlen definieren; sie bilden also auch Vektorräume. Daher ist eine formale Definition des Vektorraums sinnvoll; sie erlaubt eine einheitliche Beschreibung dieser Mengen.

Ausgerüstet mit den Begriffen zu Mengen und Gruppen aus Abschnitt 0.2.1 und 0.2.3 können wir den Begriff des Vektorraumes präzise fassen. Ein Vektorraum über den reellen Zahlen \mathbb{R} ist eine Menge V von Elementen (genannt Vektoren) und zwei Verknüpfungen. Die erste Verknüpfung ist die Addition von zwei Vektoren. Diese Verknüpfung bildet zwei Vektoren (ein Element von $V \times V$) auf ein Element von V ab. Die zweite Verknüpfung ist die Multiplikation einer reellen Zahl mit einem Vektor, sie bildet also ein Element von $\mathbb{R} \times V$ auf V ab.

Ein **Vektorraum** über den reellen Zahlen \mathbb{R} ist ein Tripel $(V, +, \times)$ bestehend aus einer Menge V und zwei Verknüpfungen $+$ und \times .

$$V \times V \longrightarrow V, (\vec{v}, \vec{w}) \longmapsto \vec{v} + \vec{w} \quad (\vec{v}, \vec{w} \in V)$$

$$\mathbb{R} \times V \longrightarrow V, (\lambda, \vec{v}) \longmapsto \lambda \times \vec{v} \quad (\lambda \in \mathbb{R}, \vec{v} \in V)$$

mit den folgenden Eigenschaften:

1. $(V, +)$ bildet eine Gruppe
2. $\vec{v} + \vec{w} = \vec{w} + \vec{v} \forall \vec{w}, \vec{v} \in V$ (Kommutativität)
3. $(\lambda + \mu) \times \vec{v} = \lambda \times \vec{v} + \mu \times \vec{v}$
4. $\lambda \times (\mu \times \vec{v}) = (\lambda \times \mu) \vec{v}$
5. $\lambda \times (\vec{v} + \vec{w}) = \lambda \times \vec{v} + \lambda \times \vec{w}$
6. $1 \times \vec{v} = \vec{v}$

Das Multiplikationszeichen wird häufig weggelassen, $\lambda \vec{v} \equiv \lambda \times \vec{v}$.

Aufgabe

Prüfen Sie, ob die Verschiebungen in der Ebene, mit den Operationen der Hintereinanderausführung und der Streckung, die Axiome des Vektorraumes erfüllen.

Beispiel

- Verschiebungen (mit Hintereinanderausführung als Addition, Streckung als Multiplikation)
- Geschwindigkeiten; als Verschiebungen pro Zeit lassen sich Geschwindigkeiten genauso addieren und mit Zahlen multiplizieren wie Verschiebungen. Ein Massepunkt der sich mit konstanter Geschwindigkeit bewegt und pro Zeitintervall τ um \vec{s} verschoben wird hat die Geschwindigkeit \vec{s}/τ .
- n -Tupel von reellen Zahlen (1-Tupel: einzelne Zahl, 2-Tupel: ein Paar von Zahlen, ...) mit Addition definiert als Addition der Komponenten der Tupel, und analog der Multiplikation. Zum Beispiel gilt für 3-Tupel

$$\mathbf{x} = \begin{pmatrix} x \\ y \\ z \end{pmatrix}, \mathbf{x}_1 + \mathbf{x}_2 = \begin{pmatrix} x_1 + x_2 \\ y_1 + y_2 \\ z_1 + z_2 \end{pmatrix}.$$

Damit bilden auch die Vektorkomponenten einen Vektorraum.

- Funktionen $f(x)$; durch Addition $f(x)+g(x)$ oder Multiplikation $\lambda f(x)$ für alle x entstehen neue Funktionen. Folgende Überlegung mag hilfreich sein: Approximieren wir eine Funktion $f(x)$ durch diskrete Funktionswerte bei x_1, x_2, x_3, \dots . Diese Funktionswerte

$$\begin{pmatrix} f(x_1) \\ f(x_2) \\ f(x_3) \\ \dots \end{pmatrix} \tag{1.8}$$

bilden unter der Addition von Funktionen und der Multiplikation von Funktionen mit Zahlen einen Vektorraum (er ist gleich dem Raum der n -Tupel im obigen Beispiel). Legt man die diskreten Funktionswerte x_1, x_2, x_3, \dots beliebig dicht, so approximiert man die Funktion $f(x)$ beliebig gut. Daraus folgt auch, dass der von Funktionen gebildete Vektorraum unendlichdimensional ist.

- Und hier noch ein Beispiel für etwas, das kein Vektorraum ist: einzelne Kugeln Eis verschiedener Sorten. Die Menge { Vanille, Erdbeere, Schokolade } bildet *keinen* Vektorraum, weil keine (sinnvolle) Additionsregel für Eiskugeln existiert, bei der die Addition von zwei Kugeln wieder eine Eiskugel ergibt.
- Permutationen bilden ebenso keinen Vektorraum, sie lassen sich nicht mit Zahlen multiplizieren.
- Ein weiteres Beispiel einer Menge die keinen Vektorraum (über den reellen Zahlen) bildet sind die natürlichen Zahlen $\mathbb{N} = \{1, 2, 3, \dots\}$. Die Zahlen 3 und 5 sind Elemente dieser Menge, aber $3 + (-1)5$ ist es nicht.

Die Elemente der Menge V werden als Vektoren bezeichnet, verkürzend spricht man oft vom Vektorraum V und meint das Tripel $(V, +, \times)$. Schreibt man $\vec{u} - \vec{v}$ für $u + (-1\vec{v})$, und $\frac{\vec{u}}{\lambda}$ für $(\frac{1}{\lambda})\vec{u}$, so sind auch Subtraktion und Division definiert. Gelten diese Axiome auch für eine Untermenge (Teilmenge) $U \subset V$, spricht man von einem **Unterraum** $(U, +, \times)$. Dann generieren Addition und Multiplikation von Vektoren aus U nur Vektoren aus U . Zum Beispiel bilden Verschiebungen entlang einer Geraden bilden einen Unterraum der Verschiebungen in der Ebene. Begriffe wie „Länge“, „Richtung“ tauchen in diesen Axiomen nicht auf, sie lassen sich auch nicht für alle Vektorräume definieren.

1.4 Basissysteme

Vektoren $\vec{e}_1, \vec{e}_2, \vec{e}_3, \dots, \vec{e}_n \in V$ nennt man ein **Erzeugendensystem**, wenn man jeden Vektor $\vec{v} \in V$ schreiben kann als

$$\vec{v} = v_1\vec{e}_1 + v_2\vec{e}_2 + v_3\vec{e}_3 + \dots + v_n\vec{e}_n, v_1, \dots, v_n \in \mathbb{R} \quad (1.9)$$

Durch Multiplikation dieser Vektoren mit Zahlen v_1, v_2, \dots, v_n und Addition der Vektoren lässt sich also jedes Element des Vektorraumes erzeugen. Man bezeichnet (1.9) als **Linearkombination** der Vektoren $\vec{e}_1, \vec{e}_2, \vec{e}_3, \dots, \vec{e}_n \in V$.

Die Beispiele i)-iii) sind alle Erzeugendensysteme der Verschiebungen in der Ebene, iv) allerdings nicht.

Vektoren $\vec{e}_1, \vec{e}_2 \dots \vec{e}_n$ heißen **linear unabhängig**, wenn gilt

$$\text{sei } v_1\vec{e}_1 + v_2\vec{e}_2 + v_3\vec{e}_3 + \dots + v_n\vec{e}_n = \vec{0}, \text{ dann gilt } v_1 = v_2 = \dots = v_n = 0. \quad (1.10)$$

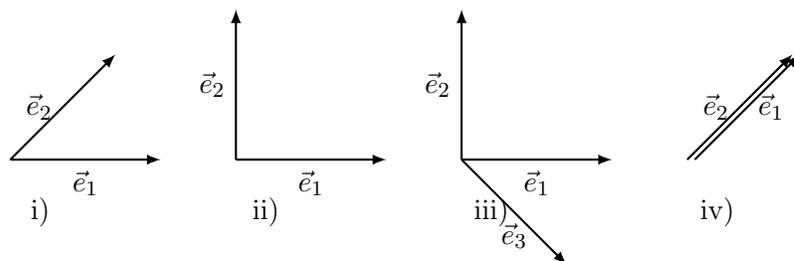
Anders ausgedrückt kann aus einer Menge linear unabhängiger Vektoren kein Element als Linearkombination der anderen Vektoren geschrieben werden. Beispiele i) und ii) zeigen linear unabhängige Vektoren, nicht aber iii) oder iv).

Vektoren $\vec{e}_1, \vec{e}_2, \dots, \vec{e}_k$ bilden eine sogenannte **Basis** von V , wenn sie sowohl ein Erzeugendensystem sind als auch linear unabhängig sind. Denn dann kann jeder Vektor \vec{v} auf eine einzige Art als eine Summe $\vec{v} = v_1\vec{e}_1 + v_2\vec{e}_2 + \dots + v_k\vec{e}_k$ ausgedrückt werden: Erstens kann jeder Vektor als Linearkombination dieser Basisvektoren ausgedrückt werden, denn die Basisvektoren bilden ein Erzeugendensystem. Und zweitens, wäre die Darstellung nicht eindeutig dann hätten wir $\vec{v} = \sum_{i=1}^k v_i\vec{e}_i = \sum_{i=1}^k v'_i\vec{e}_i$ mit $v_i \neq v'_i$ für mindestens ein i . Damit hätten wir $\sum_{i=1}^k (v_i - v'_i)\vec{e}_i = \vec{0}$ mit $v_i - v'_i \neq 0$ für mindestens ein i , dann wären die Vektoren $\vec{e}_1, \vec{e}_2, \dots, \vec{e}_k$ aber (im Widerspruch zur Annahme) nicht linear unabhängig.

Gegeben eine Basis $\vec{e}_1, \vec{e}_2, \dots, \vec{e}_k$ ist also jeder Vektor \vec{v} eindeutig durch das Tupel $\mathbf{v} = \begin{pmatrix} v_1 \\ v_2 \\ \dots \\ v_k \end{pmatrix}$ bestimmt. Jedem \vec{v} ist genau ein \mathbf{v} zugeordnet ist, und umgekehrt jedem \mathbf{v}

genau ein \vec{v} , die Abbildung zwischen Vektoren und ihrer Darstellung ist also bijektiv. Die Zahl der Basisvektoren k heißt die Dimension des Vektorraumes. Interessanterweise ist sie unabhängig von der Wahl der Basis (was wir aber nicht beweisen werden, siehe z.B. Fischer und Kaul Abschnitt 6.4, oder Kerner und v. Wahl 7.3.10).

Gegeben eine Basis $\vec{e}_1, \vec{e}_2, \dots$ ergeben sich aus den Axiomen des Vektorraumes Rechenregeln für die Komponenten von Vektoren. Damit läßt sich (zumindest in endlichdimensionalen Vektorräumen) mit Vektoren wie mit Zahlen rechnen. Zum Beispiel sind zwei Vektoren \vec{x} und \vec{y} dann und nur dann gleich, wenn ihre Komponenten gleich sind, denn die Darstellung eines Vektors \vec{x} in einer Basis als $\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$ ist eindeutig. Weitere Rechenregeln



sind

$$\mathbf{0} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \quad (1.11)$$

$$\mathbf{x} + \mathbf{y} = \begin{pmatrix} x_1 + y_1 \\ x_2 + y_2 \end{pmatrix}$$

$$\lambda \mathbf{x} = \begin{pmatrix} \lambda x_1 \\ \lambda x_2 \end{pmatrix}$$

Aufgabe

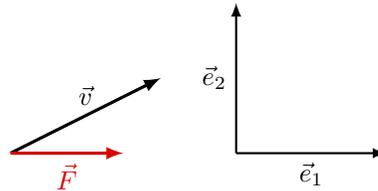
Leiten Sie den Rest der Rechenregeln aus den Axiomen des Vektorraumes her.

1.5 Das Skalarprodukt und die Norm

Beispiel

Betrachten wir die Verschiebung \vec{v} eines Massepunktes gegen eine konstante Kraft^a. In *einer* Dimension ist die Arbeit, die die Kraft am Massepunkt verrichtet, definiert als vF . v und F sind die Komponenten der Verschiebung und Kraft. In mehr als einer Dimension trägt nur die Komponente der Verschiebung parallel zur Kraft zur Arbeit bei. Zerlegt man Kraft und Verschiebung in die Komponenten einer orthonormalen Basis (aus Vektoren der Länge eins, die senkrecht aufeinander stehen, siehe unten), $\vec{F} = F_1\vec{e}_1 + F_2\vec{e}_2 + F_3\vec{e}_3$ und $\vec{v} = v_1\vec{e}_1 + v_2\vec{e}_2 + v_3\vec{e}_3$ kann man die Punktmasse zunächst um die Distanz v_1 entlang von \vec{e}_1 verschieben, dann um v_2 entlang von \vec{e}_2 , und schließlich um v_3 entlang \vec{e}_3 . Die dabei insgesamt an der Punktmasse verrichtete Arbeit ist

$$v_1F_1 + v_2F_2 + v_3F_3 . \quad (1.12)$$



^aWarum ist die Kraft ein Vektor? Die Beschleunigung ist ein Vektor, implizit aus dem zweiten Newtonschen Gesetz ist damit auch die Kraft gleich Masse (Skalar) mal Beschleunigung ein Vektor. Siehe 2.4.1.

Formal handelt es sich bei der Arbeit um eine Abbildung von zwei Vektoren (Verschiebung, Kraft) auf eine reelle Zahl (Arbeit, eine sogenannte **skalare** Größe). Wir führen eine neue Verknüpfung von zwei Vektoren auf die reellen Zahlen ein, das sogenannte Skalarprodukt oder innere Produkt $V \times V \rightarrow \mathbb{R}$, $(\vec{v}, \vec{w}) \mapsto \vec{v} \cdot \vec{w}$.

Wir beginnen im dreidimensionalen (affinen) Raum und wählen einen Ursprung und ein Basis aus drei rechtwinklig zueinander stehenden Vektoren der Länge eins. Eine solche

Orthonormalbasis lässt sich aus einem kartesischen Koordinatensystem konstruieren (siehe 0.1.1), indem man vom Ursprung aus jeweils eine Verschiebung eine Längeneinheit in Richtung der x -, y -, und z -Achsen betrachtet².

In orthonormaler Basis definieren wir das **Skalarprodukt** von zwei Vektoren \mathbf{v} und \mathbf{w} , als

$$\vec{v} \cdot \vec{w} \equiv \mathbf{v} \cdot \mathbf{w} = v_1 w_1 + v_2 w_2 + v_3 w_3 + \dots v_n w_n \quad (1.13)$$

Damit ist Arbeit definiert als das Skalarprodukt von Kraft und Verschiebung. Räume, in denen das obige Skalarprodukt definiert ist, heißen **euklidisch**, (1.13) wird auch als das euklidische oder kanonische oder Standard-Skalarprodukt bezeichnet.

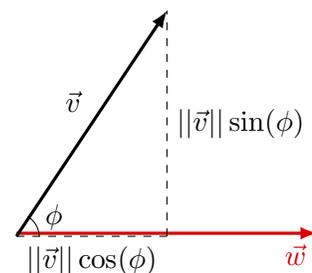
Aus dem Skalarprodukt eines Vektors mit sich selbst lässt sich die sogenannte **Norm** $\|\vec{v}\|$ definieren, die Länge des Vektors \vec{v}

$$\|\vec{v}\| = \sqrt{\vec{v} \cdot \vec{v}} = \sqrt{v_1^2 + v_2^2 + v_3^2 + \dots} \quad (1.14)$$

Geometrische Interpretation

Das Skalarprodukt hat eine geometrische Interpretation: Wählen wir die (orthonormalen) Basisvektoren \vec{e}_1, \vec{e}_2 in der Ebene, die von \vec{v}, \vec{w} aufgespannt wird, und \vec{e}_1 parallel zu \vec{w} , dann gilt $\vec{v} \cdot \vec{w} = v_1 w_1 + v_2 w_2 + v_3 w_3 + \dots = v_1 w_1 = (\|\vec{v}\| \cos(\phi))(\|\vec{w}\|) = \|\vec{v}\| \|\vec{w}\| \cos(\phi)$. Damit gibt das Skalarprodukt an, wie stark sich zwei Vektoren ‘ähnlich’ sind: $\vec{v} \cdot \vec{w} / (\|\vec{v}\| \|\vec{w}\|)$ ist 1, wenn $\vec{v} = \vec{w}$, es ist 0 wenn die beiden Vektoren senkrecht aufeinander stehen, und -1 wenn $\vec{v} = -\vec{w}$. In euklidischen Räumen gilt damit die **Cauchy-Schwarzsche Ungleichung**

$$|\vec{v} \cdot \vec{w}| \leq \|\vec{v}\| \|\vec{w}\| .$$



1.5.1 Skalarprodukträume

Skalarprodukte treten in einem breiteren Kontext auf als Kraft und Arbeit; in nicht-euklidischen Räumen kann oft ein Skalarprodukt definiert werden, das festlegt wann zwei Vektoren ‘orthogonal’ sind oder welchen ‘Winkel’ sie einschliessen. Mit Hilfe der Norm wird dann definiert was mit der ‘Länge’ eines Vektors gemeint ist.

Ein Vektorraum V über \mathbb{R} heißt **Skalarproduktraum**, wenn für alle Vektoren $\vec{v}, \vec{w}, \vec{w}_1, \vec{w}_2 \in V$ ein Skalarprodukt $\langle \vec{v}, \vec{w} \rangle$ definiert ist mit

- $\langle \vec{v}, \alpha \vec{w}_1 + \beta \vec{w}_2 \rangle = \alpha \langle \vec{v}, \vec{w}_1 \rangle + \beta \langle \vec{v}, \vec{w}_2 \rangle$

²Zur Konstruktion der kartesischen Basis benötigen wir Achsen, die senkrecht aufeinander stehen; um eine Definition dieses Begriffes kommen wir also nicht herum. Wenn wir Längen messen können, ist das leicht zu erreichen: Zwei Achsen stehen senkrecht aufeinander wenn die beiden Punkte die jeweils eine Längeneinheit vom Ursprung entfernt sind die Distanz $\sqrt{2}$ haben.

- $\langle \vec{v}, \vec{w} \rangle = \langle \vec{w}, \vec{v} \rangle$
- $\langle \vec{v}, \vec{v} \rangle \geq 0$ und wenn $\langle \vec{v}, \vec{v} \rangle = 0 \leadsto \vec{v} = \vec{0}$

Beispiel

Wir definieren das Skalarprodukt zwischen zwei Funktionen $f(x)$ und $g(x)$ als $\langle f, g \rangle = \int dx f(x)g(x)$. Überzeugen Sie sich, dass diese Definition die obigen Forderungen erfüllt.

Aufgabe

Zeigen sie, dass das euklidische Skalarprodukt (1.13) diese Forderungen erfüllt.

Diese allgemeinen Eigenschaften helfen z.B. auch das Skalarprodukt von zwei Vektoren in einer nicht-kartesischen Basis $\{\vec{e}_i\}$ auszurechnen, wenn die Skalarprodukte der Basisvektoren untereinander bekannt sind. Das Skalarprodukt von zwei Vektoren $\vec{u} = \sum_i u_i \vec{e}_i$ und $\vec{v} = \sum_j v_j \vec{e}_j$ ist

$$\begin{aligned}
 \left\langle \sum_i u_i \vec{e}_i, \sum_j v_j \vec{e}_j \right\rangle &\stackrel{i)}{=} \sum_j v_j \left\langle \sum_i u_i \vec{e}_i, \vec{e}_j \right\rangle \\
 &\stackrel{ii)}{=} \sum_j v_j \left\langle \vec{e}_j, \sum_i u_i \vec{e}_i \right\rangle \\
 &= \sum_{ij} u_i v_j \langle \vec{e}_j, \vec{e}_i \rangle \stackrel{ii)}{=} \sum_{ij} u_i v_j \langle \vec{e}_i, \vec{e}_j \rangle .
 \end{aligned} \tag{1.15}$$

Sind also die Skalarprodukte zwischen allen Paaren von Basisvektoren \vec{e}_i, \vec{e}_j bekannt, kann mit diesem Ausdruck das Skalarprodukt zwischen \vec{u} und \vec{v} aus ihren Komponenten in der Basis $\{\vec{e}_i\}$ berechnet werden. Dabei ist das Skalarprodukt zwischen diesen Vektoren unabhängig von der Wahl der Basis, denn Norm und Winkel ändern sich ja nicht unter durch eine Änderung der Basis. Allerdings ändert sich durch die Wahl der Basis ggf. der Zusammenhang (1.15) zwischen dem Skalarprodukt und den Komponenten in einer gegebenen Basis. Wenn die Basis $\{\vec{e}_i\}$ orthonormal ist, erhalten wir aus (1.15) mit $\langle \vec{e}_i, \vec{e}_j \rangle = \begin{cases} 1, & \text{wenn } i = j. \\ 0, & \text{wenn } i \neq j \end{cases} \equiv \delta_{ij}$ wieder das euklidische Skalarprodukt (1.13), $\sum_{ij} u_i v_j \langle \vec{e}_i, \vec{e}_j \rangle = \sum_i u_i v_i$.

In einer orthonormalen Basis können mit Hilfe des Skalarproduktes auch leicht die Komponenten eines Vektors berechnet werden

$$\vec{e}_i \cdot \vec{v} = \vec{e}_i \cdot \left(\sum_j v_j \vec{e}_j \right) = \sum_j v_j \vec{e}_i \cdot \vec{e}_j = v_i . \tag{1.16}$$

1.6 Das Kreuzprodukt

Ein weiteres Produkt von zwei Vektoren kann im dreidimensionalen Raum definiert werden. Als Ergebnis dieser neuen Multiplikation von zwei Vektoren erhalten wir einen weiteren Vektor. Am leichtesten lässt sich das sogenannte Kreuzprodukt in einer kartesischen Basis definieren. Wir betrachten eine sogenannte **rechtshändiges** kartesisches Basissystem im dreidimensionalen Raum, bei der die Basisvektoren $\vec{e}_1, \vec{e}_2, \vec{e}_3$ orthonormal sind und wie Daumen, Zeigefinger und Mittelfinger Ihrer rechten Hand angeordnet sind. Das **Kreuzprodukt** zwischen $\mathbf{v}, \mathbf{w} \in \mathbb{R}^3$ ist in einer solchen Basis definiert als

$$\mathbf{v} \times \mathbf{w} = \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix} \times \begin{pmatrix} w_1 \\ w_2 \\ w_3 \end{pmatrix} = \begin{pmatrix} v_2 w_3 - v_3 w_2 \\ v_3 w_1 - v_1 w_3 \\ v_1 w_2 - v_2 w_1 \end{pmatrix} \quad (1.17)$$

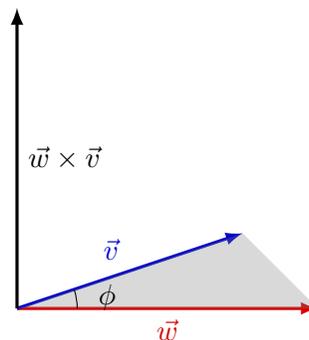
Geometrische Interpretation

1. $\mathbf{v} \times \mathbf{w}$ ist ein Vektor, der senkrecht auf \mathbf{v} und \mathbf{w} steht.

$$\begin{aligned} \mathbf{v} \cdot (\mathbf{v} \times \mathbf{w}) &= v_1 v_2 w_3 - v_1 v_3 w_2 \\ &\quad + v_2 v_3 w_1 - v_2 v_1 w_3 \\ &\quad + v_3 v_1 w_2 - v_3 v_2 w_1 \\ &= 0, \end{aligned} \quad (1.18)$$

und analog für $\mathbf{w} \cdot (\mathbf{v} \times \mathbf{w})$.

2. $\|\mathbf{v} \times \mathbf{w}\| = \|\mathbf{v}\| \|\mathbf{w}\| \sin \phi$, wobei ϕ der Winkel zwischen \mathbf{v} und \mathbf{w} ist.



$$\begin{aligned} \|\mathbf{v} \times \mathbf{w}\|^2 + (\mathbf{v} \cdot \mathbf{w})^2 &= v_2^2 w_3^2 - 2v_2 w_3 v_3 w_2 + v_3^2 w_2^2 \\ &\quad + v_3^2 w_1^2 - 2v_3 w_1 v_1 w_3 + v_1^2 w_3^2 \\ &\quad + v_1^2 w_2^2 - 2v_1 w_2 v_2 w_1 + v_2^2 w_1^2 \\ &\quad + (v_1 w_1 + v_2 w_2 + v_3 w_3)^2 \\ &= (v_1^2 + v_2^2 + v_3^2)(w_1^2 + w_2^2 + w_3^2) \\ &= \|\mathbf{v}\|^2 \|\mathbf{w}\|^2 \\ \hookrightarrow \|\mathbf{v} \times \mathbf{w}\|^2 &= \|\mathbf{v}\|^2 \|\mathbf{w}\|^2 - (\mathbf{v} \cdot \mathbf{w})^2 = (1 - \cos^2 \phi) \|\mathbf{v}\|^2 \|\mathbf{w}\|^2 \\ &= \sin^2 \phi \|\mathbf{v}\|^2 \|\mathbf{w}\|^2 \end{aligned} \quad (1.19)$$

Damit ist das Kreuzprodukt von zwei Vektoren \mathbf{v} und \mathbf{w} ein Vektor, der senkrecht auf der von \mathbf{v} und \mathbf{w} aufgespannten Ebene steht.

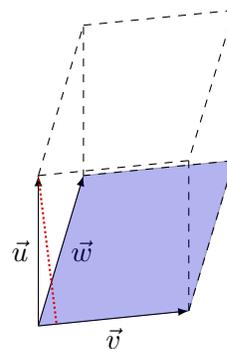
3. Diese geometrische Interpretation lässt allerdings noch die Richtung des Kreuzproduktes frei. Nach der Definition (1.17) folgt diese der "Rechte-Hand-Regel": Daumen = \mathbf{v} ; Zeigefinger = \mathbf{w} ; Mittelfinger = $\mathbf{v} \times \mathbf{w}$. Daraus folgt auch die wichtigste algebraische Eigenschaft des Kreuzproduktes, es ist **antikommutativ**, $\mathbf{v} \times \mathbf{w} = -\mathbf{w} \times \mathbf{v}$.

Das Kreuzprodukt zweier Verschiebungen $\mathbf{v} \times \mathbf{w}$ ist selbst keine Verschiebung, denn die Einheiten sind $[\text{Länge}]^2$ nicht $[\text{Länge}]$. Es beschreibt vielmehr eine Fläche, nämlich das von seinen beiden Argumenten aufgespannte Parallelogramm. Dabei beschreibt die Richtung von $\mathbf{v} \times \mathbf{w}$ die Orientierung des Parallelogramms im Raum, die Norm von $\mathbf{v} \times \mathbf{w}$ seinen Flächeninhalt.

Info: Das Kreuzprodukt hängt in zwei Punkten von der Wahl der Basis ab: in seiner Orientierung senkrecht zur von \vec{v}, \vec{w} , aufgespannten Oberfläche und in seiner Norm durch die Wahl der Längeneinheit. Es ist nur in drei Dimensionen definiert. Trotz dieser klaren Mängel wird ihnen das Kreuzprodukt an vielen Stellen begegnen. Wenn diese Beschränkungen irritieren: eine allgemeine Definition findet sich im Rahmen des Cartan-Kalküls (Vorlesungen "Differentialgeometrie" oder "Geometrie in der Physik").

Das Kreuzprodukt und Flächen und Volumina

Das Kreuzprodukt $\mathbf{v} \times \mathbf{w}$ ist nützlich um die Fläche des von \vec{v} und von \vec{w} aufgespannten Parallelogramms zu berechnen (durch die Norm von $\mathbf{v} \times \mathbf{w}$). Eine weitere wichtige Anwendung aus der Geometrie betrifft den von drei Vektoren $\vec{u}, \vec{v}, \vec{w}$ aufgespannten Körper, der sogenannte **Spats** oder das **Parallelepiped**. Sein Volumen ist die Grundfläche (das von \vec{v}, \vec{w} aufgespannte Parallelogramm), mal die Höhe des Parallelepipeds (senkrecht zur Grundfläche, gepunktete Linie in rot). Da $\mathbf{v} \times \mathbf{w}$ senkrecht zur Grundfläche steht ist das orientierte Volumen damit $\mathbf{u} \cdot (\mathbf{v} \times \mathbf{w})$, da das Skalarprodukt von \mathbf{u} und $\mathbf{v} \times \mathbf{w}$ gleich der Grundfläche mal die Komponente von \vec{u} senkrecht zur Grundfläche ist³. Welches Paar der Vektoren $\vec{u}, \vec{v}, \vec{w}$ man zur Definition der Grundfläche nimmt kann dabei keine Rolle spielen, wir erwarten also $\mathbf{u} \cdot (\mathbf{v} \times \mathbf{w}) = \mathbf{v} \cdot (\mathbf{w} \times \mathbf{u}) = \mathbf{w} \cdot (\mathbf{u} \times \mathbf{v})$ ⁴.



³Das orientierte Volumen gibt über das Vorzeichen außer dem Volumen noch die Orientierung der Elemente des Spats an: das orientierte Volumen hat einen positiven Wert, wenn die 3 Elemente des Spats ein rechtshändiges System bilden, einen negativen Wert für ein linkshändiges System. Das Volumen des Spats ist dann einfach der Betrag des orientierten Volumens.

⁴Bei diesen **zyklischen Vertauschungen** bleibt die Reihenfolge der drei Vektoren $\vec{u}, \vec{v}, \vec{w}$ erhalten. Damit bleibt die Orientierung des Spats erhalten, z.B. bilden $\vec{v}, \vec{w}, \vec{u}$ ein rechtshändiges System wenn $\vec{u}, \vec{v}, \vec{w}$

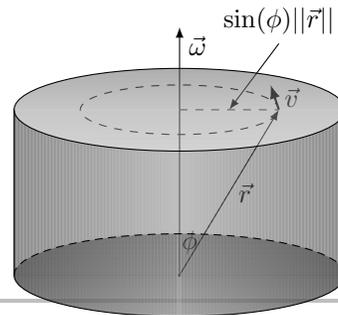
Das Kreuzprodukt und Rotationen

Die wichtigste Anwendung des Kreuzproduktes in der Mechanik liegt in Verbindung mit Rotationen. Betrachten wir eine sich drehende Scheibe und konstruieren einen Vektor ω , der senkrecht zur Scheibe steht und als Norm die Winkelgeschwindigkeit hat ($2\pi/\text{Drehperiode}$). Jeder Punkt auf der Scheibe bewegt sich nun mit einer anderen Geschwindigkeit \mathbf{v} . Legen wir den Ursprung in den Scheibenmittelpunkt, bewegt sich jeder Punkt auf der Scheibe senkrecht zu seinem Ortsvektor *und* senkrecht zu dem Winkelgeschwindigkeitsvektor ω , die ihrerseits senkrecht aufeinanderstehen. Betrag der Geschwindigkeit ist $|\mathbf{r}||\omega|$ und somit

$$\mathbf{v} = \omega \times \mathbf{r} . \quad (1.20)$$

Aufgabe

Hier haben wir uns auf eine Scheibe beschränkt, so dass \vec{r} und $\vec{\omega}$ senkrecht aufeinander stehen. Betrachten Sie einen rotierenden Zylinder, um zu zeigen, dass dieses Ergebnis allgemein gilt. (Hinweis: Nutzen Sie dass der Abstand eines Punktes von der Rotationsachse $\sin(\phi)|\vec{r}|$ ist.) Zeigen Sie, dass dieses Ergebnis für beliebiger Körper bei Rotation mit fester Drehachse gilt.



ein rechtshändiges System bilden.

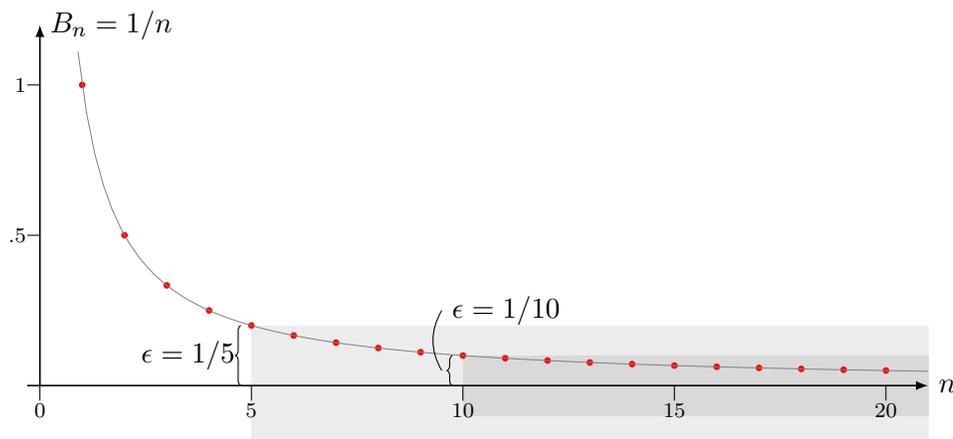
Zwei

Analysis

Analysis beinhaltet die Themen Grenzwerte, Differentiation und Integration. In der Physik spielen diese Themen eine zentrale Rolle bei der Beschreibung der Veränderung physikalischer Größen in Raum und Zeit: die Position eines Teilchens als Funktion der Zeit in der Mechanik, oder die räumliche und zeitliche Änderung elektromagnetischer Felder. Wir diskutieren die Differentiation von Funktionen einer und mehrerer Veränderlicher sowie Integrale über eine und mehrere Variablen.

Einiges dieser Themen ist Ihnen aus der Schule geläufig, wir beginnen daher nur mit einem kurzen Abriss der Themen Folgen und Reihen, Grenzwert, Differentiation und Integration von Funktionen einer Veränderlichen. Eine tiefere Darstellung bleibt den Analysis-Vorlesungen vorbehalten, die die meisten von Ihnen hören werden. Zentraler Begriff ist der Grenzwert einer Folge oder einer Funktion. Mit Hilfe des Grenzwerts einer unendlichen Folge von Zahlen lernen wir Summen mit einer unendlichen Zahl von Termen zu berechnen. Praktische Anwendungen liegen in der Entwicklung von Funktionen als Potenzenreihen, sogenannte Taylorreihen; weite Teile der theoretischen Physik fußen auf kunstvollen Reihenentwicklungen.

2.1 Folgen



Wir betrachten eine (unendliche) **Folge** von Zahlen B_1, B_2, B_3, \dots . Ein Beispiel ist die Folge $1, 1/2, 1/3, \dots$ mit $B_n = 1/n$. Die Abbildung zeigt, wie die Terme B_n für wachsendes n immer näher an Null heranrücken, ohne die Null für endliche n je zu erreichen. Dennoch rückt $1/n$ für wachsende n beliebig nahe an die Null heran. Dieses “beliebig nahe” lässt sich präzise fassen: egal wie klein man $\epsilon > 0$ wählt, gibt es immer ein N , so dass für alle $n > N$ die Differenz zwischen B_n und einer Zahl B kleiner ist als ϵ . (Hier ist $B = 0$.) Die Abbildung zeigt Beispiele für $\epsilon = 1/5$ und $1/10$. Existiert für eine Folge für jedes $\epsilon > 0$ ein $N \in \mathbb{N}$, so dass für alle $n > N$ gilt $|B_n - B| < \epsilon$, so nennen wir die Folge **konvergent** mit **Grenzwert** oder **Limes** B und schreiben $\lim_{n \rightarrow \infty} B_n = B$. Diese Definition beschreibt das Verhalten der Folge B_n für unendliche n , kommt aber selbst eleganterweise völlig ohne den Begriff “unendlich” aus.

Beispiel

Wir zeigen, dass der Grenzwert $\lim_{n \rightarrow \infty} 1/n$ gleich Null ist. Für ein gegebenes $\epsilon > 0$ wählen wir $N \in \mathbb{N}$ größer $1/\epsilon$. Also gilt für alle $n > N$ dass $|B_n - B| = |1/n - 0| = 1/n < 1/N < \epsilon$, wobei die erste Ungleichung aus $n > N$, die zweite aus $N > 1/\epsilon$ folgt. (Überzeugen Sie sich, dass dies für andere Werte von B als Null nicht gilt.)

2.2 Reihen

Wir betrachten Folgen, deren einzelne Terme durch Summen gebildet werden. Eine Folge von Zahlen, die durch die Teilsummen $b_1, b_1 + b_2, b_1 + b_2 + b_3, \dots$ gebildet wird,

$$B_n = \sum_{k=1}^n b_k, \quad (2.1)$$

heißt **Reihe**. Ist die Folge dieser Teilsummen konvergent, sprechen wir von einer **konvergenten Reihe**. Damit läßt sich auch einer Summe mit unendlich vielen Termen ein Wert zuweisen, nämlich der Grenzwert der Folge B_1, B_2, B_3, \dots . Informell bezeichnet man $b_1 + b_2 + b_3 + \dots$ als Reihe und meint die Folge der Teilsummen (denn nur dann kann man Konvergenz definieren).

Beispiel

Als Beispiel betrachten wir die sogenannte **geometrische Reihe**

$$\begin{aligned} \sum_{k=0}^{\infty} x^k &= 1 + 0.1 + 0.01 + 0.001 + \dots = 1.1111\dots && \text{für } x = 0.1 && (2.2) \\ &= 1 + 2 + 4 + 8 + \dots && \text{für } x = 2. \end{aligned}$$

Ob die unendliche Reihe $\sum_{k=0}^{\infty} x^k$ konvergent ist oder nicht hängt also von x ab. Die Summe B_n der ersten Terme bis x^n dieser Reihe läßt sich leicht berechnen

$$\begin{aligned} (1-x)B_n &= (1-x) \sum_{k=0}^n x^k = \sum_{k=0}^n x^k - \sum_{k=0}^n x^{k+1} && (2.3) \\ &= (1 + x + x^2 + \dots + x^n) - (x + x^2 + \dots + x^n + x^{n+1}) = 1 - x^{n+1} \\ B_n &= \frac{1 - x^{n+1}}{1 - x} \\ \curvearrowright B &= \lim_{n \rightarrow \infty} B_n = \frac{1}{1 - x} \text{ für } |x| < 1 \end{aligned}$$

Das letzte Gleichheitszeichen der zweiten Zeile ergibt sich, weil sich alle Terme in $\sum_{k=0}^n x^k$ und $-\sum_{k=0}^n x^{k+1}$ gegenseitig wegheben, bis auf den ersten und den letzten Term. Für $n \rightarrow \infty$ geht x^k gegen Null, wenn $|x| < 1$. Die geometrische Reihe konvergiert also für $|x| < 1$ und divergiert für $|x| > 1$.

2.2.1 Potenzreihen

Reihen $\sum_{k=0}^{\infty} a_k x^k = a_0 + a_1 x + a_2 x^2 + \dots$, $a_k \in \mathbb{R}$ aus unterschiedlichen Potenzen einer Variablen heißen **Potenzreihen**. Die geometrische Reihe (2.2) ist ein Beispiel für eine

Potenzreihe, mit $a_k = 1 \forall k$. Bricht man die Reihe nach einer n Termen ab, so erhält man ein Polynom $n - 1$ -ter Ordnung.

Grund sich mit Potenzreihen zu beschäftigen ist, dass sich eine Vielzahl von Funktionen durch unendliche Potenzreihen darstellen lässt. Bricht man die Potenzreihe nach einer endlichen Zahl von Termen ab erhält man oft eine nützliche Näherung, wir werden diesem Gedanken noch in Abschnitt 2.5 nachgehen.

Der Konvergenzradius

Hängen die Terme einer Reihe von einem Parameter x ab, kann es also sein, daß die Reihe für bestimmte Werte von x konvergiert, für andere aber nicht. Mit der geometrischen Reihe (2.2) haben wir ein Beispiel kennengelernt.

Nehmen wir an, eine Potenzreihe in $\sum_{k=0}^{\infty} a_k x^k$ konvergiere für einen bestimmten Wert von $x = x_1 \neq 0$. Also ist $|a_k x_1^k| \leq C \forall k$, wobei $C \in \mathbb{R}$ eine endliche Konstante ist, denn $a_k x_1^k$ kann nicht beliebig mit k wachsen. Aus $\sum_{k=0}^{\infty} a_k x^k = \sum_{k=0}^n a_k x_1^k \left(\frac{x}{x_1}\right)^k$ folgt ausserdem, dass jeder Term in $\sum a_k x^k$ ist betragsmäßig kleiner als die Terme in

$$\sum_{k=0}^{\infty} C \left| \frac{x}{x_1} \right|^k .$$

Für $|x| < |x_1|$ konvergiert diese geometrische Reihe, und da Reihe $\sum_k a_k x^k$ nur betragsmäßig kleinere Terme enthält konvergiert auch sie. D. h. wenn eine Reihe für einen Wert von $x = x_1$ konvergiert, muss sie das für alle betragsmäßig kleineren Werte auch tun.

Zu jeder Potenzreihe $\sum_{k=0}^{\infty} a_k x^k$ gibt es also ein $\rho \in [0, \infty]$ mit der Eigenschaft, dass die Reihe für $|x| < \rho$ konvergiert, und für $|x| > \rho$ divergiert. ρ heißt **Konvergenzradius**. Über das Verhalten der Potenzreihe bei $x = \rho$ macht dieser Satz übrigens keine Aussage, dort gibt es auch kein allgemeine gültiges Verhalten: es gibt Reihen die am Konvergenzradius konvergieren, Reihen die divergieren, oder sogar an einem Punkt konvergiere, an einem anderen divergieren.

Ausblick (ohne Beweise)

- Der Konvergenzradius lässt sich durch einen Grenzwert $\rho = \lim_{n \rightarrow \infty} \left| \frac{a_n}{a_{n+1}} \right|$ bestimmen, wenn ab einem bestimmten n alle a_n von Null verschieden sind, und wenn der Grenzwert existiert. (Für den Fall, dass er nicht existiert, s. Satz von Cauchy-Hadamard.)
- Eine Reihe heißt **absolut konvergent**, wenn auch die Reihe ihrer Absolutbeträge konvergiert. $\sum_{k=0}^{\infty} b_k$ ist absolut konvergent, wenn $\sum_{k=0}^{\infty} |b_k|$ konvergiert. Innerhalb ihres Konvergenzradius ist jede Reihe absolut konvergent.

- Absolut konvergente Reihen können Term für Term multipliziert werden

$$\left(\sum_{k=0}^{\infty} a_k \right) \left(\sum_{k'=0}^{\infty} b_{k'} \right) = \sum_k \sum_{k^1} a_k b_{k^1}$$

- eine Potenzreihe $\sum_{k=0}^{\infty} a_k x^k$ heißt **gleichmäßig konvergent** in einem Bereich $D \subset \mathbb{R}$ wenn

$$\forall \epsilon > 0 \exists N \in \mathbb{N} : \forall n > N \forall x \in D \left| \sum_{k=0}^n a_k x^k - \sum_{k=0}^{\infty} a_k x^k \right| < \epsilon$$

(Ein strengeres Kriterium als Konvergenz an jedem einzelnen Punkt x ; eine Beziehung zwischen N und ϵ muss jetzt statt an einem einzelnen Punkt im gesamten Definitionsbereich der Funktion gelten.) Innerhalb ihres Konvergenzradius konvergieren Potenzreihen gleichmäßig.

- Gleichmäßig konvergente Potenzreihen können Term für Term differenziert und integriert werden

$$\lim_{n \rightarrow \infty} \int dx \left(\sum_{k=0}^{\infty} a_k x^k \right) = \lim_{n \rightarrow \infty} \sum_{k=0}^n \left(\int dx a_k x^k \right)$$

und Limites (Grenzwerte) vertauschen bei gleichmäßiger Konvergenz der Reihe

$$\lim_{x \rightarrow A} \lim_{y \rightarrow B} f(x, y) = \lim_{y \rightarrow B} \lim_{x \rightarrow A} f(x, y) .$$

Der Beweis dieser Aussagen ist ein wichtiger Teil von Vorlesungen in Analysis.

2.3 Grenzwert von Funktionen

2.3.1 Grenzwert einer Funktion bei unendlichem Argument

Die Definition der Konvergenz einer Folge lässt sich leicht auf Funktionen verallgemeinern. Damit können wir dann das Verhalten einer Funktion (z.B. $1/x$) für große Werte des Arguments charakterisieren. Eine Funktion $f(x)$ hat den Grenzwert F bei unendlich, wenn man für jedes $\epsilon > 0$ ein Wert von x (nennen wir ihn δ) existiert, so dass sich rechts von diesem Wert die Funktion weniger als ϵ von F unterscheidet. In kompakter Form hat die Funktion f für $x \rightarrow \infty$ den **Grenzwert** F , wenn $\forall \epsilon > 0 \exists \delta : \forall x > \delta |f(x) - F| < \epsilon$. Man sagt auch die Funktion konvergiert für $x \rightarrow \infty$ gegen F .

Beispiel

Wir suchen den Grenzwert von $1/x$ bei unendlich und prüfen ob dieser gleich Null ist. Für jedes $\epsilon > 0$ können wir ein δ angeben, so dass für $x > \delta$ $1/x$ weniger als ϵ von null abweicht, nämlich $\delta = 1/\epsilon$.

Es gibt Funktionen, die ohne obere Schranke im Unendlichen anwachsen, z.B. $f(x) = x$. Für jedes T gibt es also ein Zahl (nenne wir sie wieder δ), so dass rechts davon die Funktion stets grösser als T ist. Man sagt der Grenzwert von f bei $x \rightarrow \infty$ ist unendlich, wenn $\forall T \exists \delta : \forall x > \delta f(x) > T$. Analog kann mit mit $f(x) < T$ der Grenzwert $-\infty$ definiert werden.

Der Grenzwert einer Funktion bei $-\infty$ ist analog definiert, nur dass nun für $x < \delta$, $|f(x) - F| < \epsilon$ gefordert ist, bzw. $x < \delta$, $f(x) > T$.

2.3.2 Grenzwert einer Funktion bei endlichem Argument

Oft sind auch Grenzwerte bei endlichem Argument interessant, z.B. wenn die Funktion an einem bestimmten Punkt nicht definiert ist. Ein Beispiel werden wir im nächsten Abschnitt kennenlernen.

Betrachten wir als Beispiel $g(x) = (x - x_0)^2$. (Dort ist tatsächlich die Funktion bei x_0 definiert.) Nähert sich x dem Wert x_0 , nähert sich $g(x)$ der Null an. Was bedeutet das "nähert sich der Null an"? Eine Abweichung von 10^{-10} von Null? Oder 10^{-100} ?

Der Sachverhalt lässt sich wieder präziser formulieren: Fordern wir, dass der Abstand zwischen $g(x)$ und 0 einen Wert $\epsilon > 0$ nicht übersteige. Dabei können wir ϵ so klein wählen wie wir wollen. Dann gibt es in unserem Beispiel einen Wert δ , so dass sobald x näher als δ an x_0 liegt, diese Forderung erfüllt ist, d.h. $|g(x)| < \epsilon$. Geometrisch betrachtet liegt die Funktion innerhalb eines Kastens mit Seitenlänge 2δ und Höhe 2ϵ um x_0 und den Grenzwert 0.

Diese Denkweise wird zur Definition des Grenzwertes einer Funktion genutzt: Sei X eine Teilmenge von \mathbb{R} und $x_0 \in \mathbb{R}$ ein Punkt, bei dem in jeder Umgebung unendlich viele Elemente von X liegen¹. Gibt es für jedes $\epsilon > 0$ ein δ , so dass für alle $x \in X$ mit $0 < |x - x_0| < \delta$ gilt $|g(x) - G| < \epsilon$, dann bezeichnen wir G als **Grenzwert** von $g(x)$ für x gegen x_0 , oder $\lim_{x \rightarrow x_0} g(x) = G$. Der Unterschied zwischen dem Funktionswert $g(x)$ und dem Grenzwert wird also beliebig klein, wenn man x genügend nahe bei x_0 wählt. Dazu muss $g(x)$ bei x_0 nicht einmal definiert sein.

¹Hier fehlt noch der Begriff der Umgebung eines Punktes. Für reelle Zahlen ist das einfach, gegeben ein $\epsilon > 0$ ist die ϵ -Umgebung von x_0 die Menge aller Punkte mit $|x - x_0| < \epsilon$.

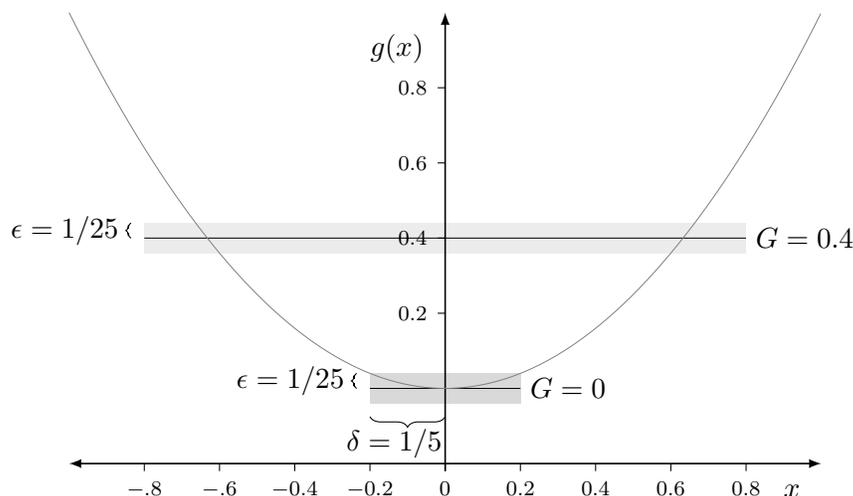


Abbildung 2.1: Grenzwert von $g(x) = x^2$ für $x \rightarrow 0$. Für $G = 0$ und $\epsilon = 1/25$ existiert ein δ , so dass für $-\delta < x < \delta$ gilt $|g(x) - G| < \epsilon$ (dunkelgraue Fläche (ein solches δ lässt sich für jedes $\epsilon < 0$ finden). Für $G = 0.4$ allerdings lässt sich für $\epsilon = 1/25$ kein solches δ finden (hellgraue Fläche).

Beispiel

In unserem Beispiel $g(x) = (x - x_0)^2$ ist $g(x) - 0 < \epsilon$ für alle Werte von x mit $|x - x_0| < \delta = \sqrt{\epsilon}$. $g(x) - G < \epsilon$ für alle Werte von x im Intervall $x_0 - \delta < x < x_0 + \delta$ lässt sich nur mit $G = 0$ für alle Werte von ϵ erfüllen. $G = 0$ ist daher als Grenzwert eindeutig bestimmt, siehe Abbildung.

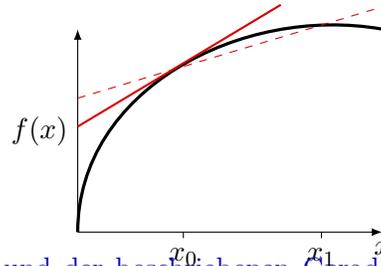
Aus diesem Beispiel entnehmen wir, daß sich Grenzwerte nicht einfach “ausrechnen” lassen, wir haben zunächst G geraten und verifiziert. (Wir hätten aber auch nicht geraten, dass $(x - x_0)^2$ nahe bei x_0 andere Werte als solche nahe bei Null annimmt.)

An Punkten, bei denen die Funktion einen Sprung hat, gibt es nach dieser Definition keinen Grenzwert. Ein Beispiel ist die Funktion $\text{sgn}(x)$, die das Vorzeichen von x angibt und bei $x = 0$ von -1 nach 1 springt. Hier ist die Definition eines rechtsseitigen und ein linksseitigen Grenzwertes sinnvoll. Wir ersetzen $0 < |x - x_0| < \delta$ durch $0 < x - x_0 < \delta$ zur Definition des rechtsseitigen Grenzwertes und durch $0 > x - x_0 > -\delta$ zur Definition des linksseitigen Grenzwertes. Geometrisch ist die damit einfach die “Kiste” mit Breite 2δ um x_0 herum ersetzt durch ihre rechte bzw. linke Hälfte. Damit hat $\text{sgn}(x)$ bei $x = 0$ den rechtsseitigen Grenzwert 1 aber den linksseitigen Grenzwert -1 , wir schreiben $\lim_{x \rightarrow 0^+} \text{sgn}(x) = 1$ und $\lim_{x \rightarrow 0^-} \text{sgn}(x) = -1$.

2.4 Differentiation

2.4.1 Differentiation von Funktionen einer Veränderlichen

Die Ableitung $f'(x)$ einer Funktion $f(x)$ beschreibt die Steigung $s = f'(x = x_0)$ von $f(x)$ am Punkt x_0 . Diese Formulierung verdeckt allerdings, wie nützlich das Konzept der Ableitung in der Praxis ist: Für eine hinreichend "glatte" Funktion schmiegt sich $f(x)$ bei x_0 an die Gerade durch den Punkt mit Koordinaten $(x_0, f(x_0))$ mit Steigung s , $f(x)$ kann also nahe x_0 durch diese Gerade approximiert werden.

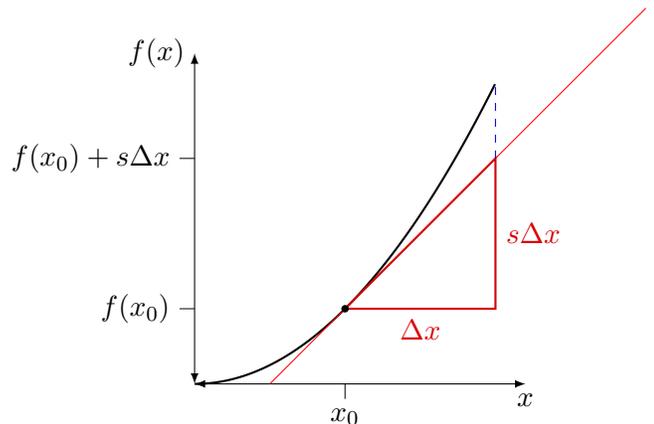


Info: Der Unterschied zwischen der Funktion $f(x)$ und der beschriebenen Geraden ist wieder eine Funktion von $x - x_0$, die bei $x = x_0$ verschwindet. Schreiben wir diesen Unterschied als $(x - x_0)\psi(x)$ und approximieren $\psi(x)$ wiederum durch eine Gerade, und führen dieses Prinzip iterativ fort, kommen wir auf eine Näherung für $f(x)$, die im Prinzip beliebig genau sein kann. Diese Idee führt in Abschnitt 2.5 auf die sogenannte **Taylor-Reihe**. Wie können wir diese Ableitung (Steigung der Tangente zu $f(x)$ an einem Punkt x_0) bestimmen? Eine Gerade durch die Punkte mit Koordinaten $(x_0, f(x_0))$ und $(x_1, f(x_1))$ (genannt **Sekante**) hat Steigung $S = \frac{f(x_1) - f(x_0)}{x_1 - x_0}$. Je näher x_1 bei x_0 liegt, desto besser approximiert S die Steigung von $f(x)$ bei x_0 . Setzt man aber einfach $x_1 = x_0$ ist allerdings der Nenner dieses Bruchs Null! Wir nutzen daher den Grenzwert: **Der Grenzwert $f'(x) = \lim_{\delta x \rightarrow 0^+} \frac{f(x + \delta x) - f(x)}{\delta x}$ heißt (rechte) Ableitung einer Funktion $f(x)$ ($f : \mathbb{R} \rightarrow \mathbb{R}$) am Punkt x (wenn der Grenzwert existiert). Analog definiert $\lim_{\delta x \rightarrow 0^-} \frac{f(x - \delta x) - f(x)}{\delta x}$ die linke Ableitung, die sich von der rechten aber nur unterscheidet wenn $f(x)$ bei x_0 einen Knick hat.**

In der Nähe von x_0 lässt sich dann die Funktion durch eine Gerade $f(x) \approx f(x_0) + (x - x_0)s$ mit Steigung $s = f'(x_0)$ nähern. Zur Herleitung genügt es ein Dreieck mit Kantenlängen $\Delta x \equiv x - x_0$ und $s\Delta x$ zu betrachten (siehe Abbildung rechts). Die gestrichelte blaue Linie gibt den Fehler an, den wir machen, wenn wir die Funktion $f(x)$ bei x_0 durch eine Gerade nähern. Schreiben wir diesen Fehler als $\phi(x)$ erhalten wir

$$f(x) = f(x_0) + (x - x_0)f'(x_0) + \phi(x), \quad (2.4)$$

Umstellen der einzelnen Terme führt auf $f'(x_0) = \frac{f(x) - f(x_0)}{x - x_0} - \frac{\phi(x)}{x - x_0}$. Wir betrachten den Limes $x \rightarrow x_0$



dieser Gleichung, zusammen mit der Definition der Ableitung erhalten wir $\lim_{x \rightarrow x_0} \frac{\phi(x)}{x-x_0} = 0$. Damit ist nicht nur $\phi(x_0) = 0$, sondern der Fehler der linearen Näherung $\phi(x)$ verschwindet bei x_0 schneller als $x - x_0$. Bei x_0 ist $\phi(x)$ von der Ordnung $(x - x_0)^2$ oder kleiner, wir schreiben $f(x) = f(x_0) + (x - x_0)f'(x_0) + \mathcal{O}((x - x_0)^2)$. $\mathcal{O}(x^n)$ bedeutet "von der selben Ordnung wie x^n im Grenzfall $x \rightarrow 0$ ", präziser $f(x) = \mathcal{O}(g(x))$ wenn $0 < \lim_{x \rightarrow 0} f(x)/g(x) < \infty$. (Dieselbe Notation wird auch verwendet um die Asymptotik einer Funktion bei $x \rightarrow \infty$ zu beschreiben, welche Asymptotik gemeint ist folgt aus dem Kontext.)

Beispiel

Die Funktion $f(x) = 3x^2 + 15x^3 + 17x^4$ ist für kleine x von der Ordnung x^2 , denn $\lim_{x \rightarrow 0} \frac{f(x)}{x^2} = \lim_{x \rightarrow 0} (3 + 15x + 17x^2) = 3$ liegt zwischen 0 und ∞ , also ist $f(x) = \mathcal{O}(x^2)$ (für $x \rightarrow 0$). Analog ist auch $f(x) = 3x^2 + \mathcal{O}(x^3)$, womit wir notieren, dass wir bei kleinen Werten von x den Beitrag $15x^3 + 17x^4$ zu $f(x)$ gegenüber dem Beitrag $3x^2$ vernachlässigen können.

Beispiel

Betrachten wir die Funktion $f(x) = x^2$. Mit $(x + \delta x)^2 = x^2 + 2x\delta x + \delta x^2$ ist

$$f'(x) = \lim_{\delta x \rightarrow 0} \frac{f(x + \delta x) - f(x)}{\delta x} = \lim_{\delta x \rightarrow 0} \frac{(x + \delta x)^2 - x^2}{\delta x} = \lim_{\delta x \rightarrow 0} (2x + \delta x) = 2x . \quad (2.5)$$

Die Ableitung höherer Polynome lässt sich analog herleiten. Bei der Ableitung von Exponentialfunktionen $f(x) = a^x$ erhalten wir

$$f'(x) = \lim_{\delta x \rightarrow 0} \frac{a^{x+\delta x} - a^x}{\delta x} = a^x \lim_{\delta x \rightarrow 0} \frac{a^{\delta x} - 1}{\delta x} . \quad (2.6)$$

Ableitung der Exponentialfunktion ist also die Exponentialfunktion selbst, mal eine Konstante. $e = 2.7182818284590\dots$ ist definiert als die einzige positive Zahl für die diese Konstante den Wert eins hat; $\lim_{h \rightarrow 0} \frac{e^h - 1}{h} = 1$, damit ist die Exponentialfunktion e^x gleich ihrer Ableitung.

Für die Ableitung von Summen, Produkten und Quotienten gelten Regeln, die sie aus der Schule kennen, und sich mit (2.4) herleiten lassen. Als Beispiel betrachten wir die Kettenregel $(g \circ f)'(x) = g'(f(x)) f'(x)$, oder $\frac{d}{dx}g(f(x)) = g'(f(x))f'(x)$. Wir schreiben $\phi(x) = (x - x_0)\psi(x)$, aus $\lim_{x \rightarrow x_0} \frac{\phi(x)}{x-x_0} = 0$ folgt $\psi(x_0) = 0$. Analog zu (2.4) schreiben wir

$$g(y) = g(y_0) + (y - y_0)g'(y_0) + (y - y_0)\psi_2(y) \quad (2.7)$$

und setzen dann $y = f(x)$. Aus (2.4) erhalten wir $y - y_0 = (x - x_0)f'(x_0) + (x - x_0)\psi(x)$, setzen wir diese Ergebnis in (2.7) ein, erhalten wir

$$g(f(x)) = g(f(x_0)) + [(x - x_0)f'(x_0) + (x - x_0)\psi(x)] g'(f(x_0)) + [(x - x_0)f'(x_0) + (x - x_0)\psi(x)] \psi_2(f(x)) . \quad (2.8)$$

Bilden wir daraus die Ableitung von $g(f(x))$ erhalten wir die **Kettenregel**

$$\begin{aligned} \frac{d}{dx}g(f(x)) &\equiv \lim_{x \rightarrow x_0} \frac{g(f(x)) - g(f(x_0))}{(x - x_0)} \\ &= g'(f(x_0))f'(x_0) + \lim_{x \rightarrow x_0} (\psi(x)) g'(f(x_0)) + f'(x_0)\psi_2(f(x)) + \psi(x)\psi_2(f(x)) \\ &= g'(f(x_0))f'(x_0) . \end{aligned} \quad (2.9)$$

Etwas lockerer schreiben wir auch

$$g(f(x + \delta x)) = g(f(x) + \delta x f'(x) + \mathcal{O}(\delta x^2)) = g(f(x)) + g'(f(x))\delta x f'(x) + \mathcal{O}(\delta x^2) \quad (2.10)$$

woraus wieder die Kettenregel folgt.

Ableitung eines Vektors

Diese Definition der Ableitung funktioniert ohne Modifikation auch für die Ableitung eines Vektors, der von einer Veränderlichen abhängt. Beispiel ist eine Bahnkurve, die die Bewegung eines Massepunktes beschreibt, siehe Abschnitt 1.2. Mathematisch beschreiben wir diese Bahnkurve durch die (vektorwertige) Funktion $\vec{r}(t)$. Der Vektor $\vec{r}(t + \delta t) - \vec{r}(t)$ (Differenz zweier Vektoren) gibt die Verschiebung des Massepunktes zwischen den Zeitpunkten t und $t + \delta t$ an. Teilen wir durch δt erhalten wir im Grenzfall $\delta t \rightarrow 0$ die **Vektorgeschwindigkeit**

$$\vec{v}(t) = \frac{d\vec{r}}{dt} \equiv \lim_{\delta t \rightarrow 0} \frac{\vec{r}(t + \delta t) - \vec{r}(t)}{\delta t} . \quad (2.11)$$

Bei zeitlich konstanten Basisvektoren sind die Komponenten der Vektorgeschwindigkeit leicht zu bestimmen,

$$\vec{v}(t) = \sum_{i=1}^n v_i \vec{e}_i = \frac{d}{dt} \sum_{i=1}^n r_i(t) \vec{e}_i = \sum_{i=1}^n \frac{dr_i}{dt} \vec{e}_i , \quad (2.12)$$

und somit sind die Komponenten der Vektorgeschwindigkeit $v_i(t) = \frac{dr_i}{dt}$. Ihre Einheiten sind (wegen des Faktors δt) in (2.11) [Länge/Zeit], nicht wie bei dem Ortsvektor [Länge]. Trotzdem ist die Vektorgeschwindigkeit wiederum ein Vektor: Es handelt es sich um die Differenz zweier Ortsvektoren (Differenz zweier Vektoren ist ein Vektor), multipliziert mit $1/\delta t$, einer skalaren Größe. Würden wir also eine neue Basis nutzen, so daß sich die Komponenten von \vec{r} ändern, dann würden sich die Komponenten von \vec{v} auf die gleiche Art und Weise ändern wie die von \vec{r} .

Beispiel

Wir kehren zurück zu der Bahnkurve aus Abschnitt 1.2, die die gleichförmige Bewegung eines Massepunktes entlang eines Halbkreises beschreibt

$$\mathbf{r}(t) = \begin{pmatrix} \cos(2\pi t) \\ \sin(2\pi t) \end{pmatrix}. \quad (2.13)$$

Die Geschwindigkeit des Massepunktes ist dann

$$\mathbf{v}(t) = \begin{pmatrix} -(2\pi) \sin(2\pi t) \\ (2\pi) \cos(2\pi t) \end{pmatrix} \quad (2.14)$$

und die Beschleunigung $\mathbf{a}(t) \equiv \frac{d\mathbf{v}}{dt}$

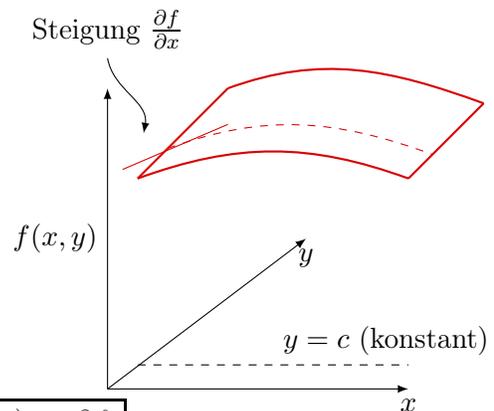
$$\mathbf{a}(t) = \begin{pmatrix} -(2\pi)^2 \cos(2\pi t) \\ -(2\pi)^2 \sin(2\pi t) \end{pmatrix}. \quad (2.15)$$

Damit stehen Geschwindigkeit und Ortsvektor bei der kreisförmigen Bewegung senkrecht zueinander, und die Beschleunigung ist parallel und in Gegenrichtung zum Ortsvektor.

Das 2te Newtonsche Gesetz, $\vec{F} = m\vec{a}$ ist also nicht nur eine Verknüpfung der Größen Kraft, Masse und Beschleunigung, sondern enthält auch implizit die Forderung, dass die Kraft ein Vektor ist. Bei einer Basisänderung müssen sich die Komponenten der Kraft also genauso ändern wie die Komponenten eines Ortsvektors.

2.4.2 Die partielle Ableitung

Auch für Funktionen mehrere Veränderlicher können Ableitungen definiert werden. Betrachten wir zunächst eine Funktion $f(x, y)$, die von zwei Variablen abhängt. Diese Funktion kann als gekrümmte Oberfläche über der x, y -Ebene visualisiert werden. Bei konstantem $y = c$ hängt $f(x, y = c) =: f_c(x)$ nur von x ab, so dass wir $f'_c(x)$ definieren können. Eine zweite Ableitung kann gebildet werden, wenn statt y die Variable $x = c$ konstant gehalten wird. Ist $f(x_1, x_2, x_3, \dots, x_n)$ eine Funktion auf einem (offenen) Quader $\subset \mathbb{R}^n$, heißt der Grenzwert



$$\lim_{\delta x_i \rightarrow 0} \frac{f(x_1, x_2, \dots, x_i + \delta x_i, \dots) - f(x_1, x_2, \dots, x_i, \dots)}{\delta x_i} \equiv \frac{\partial f}{\partial x_i} \quad (2.16)$$

partielle Ableitung $\frac{\partial f}{\partial x_i}$ von f nach x_i (wenn der Grenzwert existiert).

Beispiel

$$\frac{\partial}{\partial x_1}(x_1^3 + x_1x_2x_3 + x_1^2x_2) = 3x_1^2 + x_2x_3 + 2x_1x_2$$

$$\frac{\partial}{\partial x_2}(\sin(x_1x_2)) = x_1 \cos(x_1x_2)$$

Die partiellen Ableitung gibt an, wie sich der Funktionswert von f unter einer kleinen Veränderung von x_i ändert; $f(\dots, x_i + \delta x_i, \dots) = f(\dots, x_i, \dots) + \frac{\partial f}{\partial x_i}(\dots, x_i, \dots)\delta x_i + \mathcal{O}(\delta x_i^2)$. Geometrisch gibt $\frac{\partial f}{\partial x_i}$ die Steigung der durch f definierten Oberfläche in Richtung der x_i -Koordinate an.

Auch wenn sich *zwei* Variablen x_i und x_j verändern ist die Änderung des Funktionswertes f in linearer Näherung durch die beiden partiellen Ableitungen definiert (da alle anderen Variablen konstant gehalten werden, schreiben wir sie im Folgenden gar nicht mit). Wir ändern erst x_i um δx_i (bei konstantem x_j), und dann x_j (bei konstantem $x_i + \delta x_i$). Die Reihenfolge spielt dabei keine Rolle, denn der Funktionswert hängt nur von den Variablen x_i, x_j ab, nicht davon wie wir von einem Punkt zu einem anderen gekommen sind. Bei dem ersten Schritt ändert sich der Funktionswert um $\delta x_i \frac{\partial f}{\partial x_i}(x_i, x_j) + \mathcal{O}(\delta x_i^2)$, bei dem zweiten Schritt um $\delta x_j \frac{\partial f}{\partial x_j}(x_i + \delta x_i, x_j) + \mathcal{O}(\delta x_j^2) = \delta x_j \frac{\partial f}{\partial x_j}(x_i, x_j) + \mathcal{O}(\delta x_j^2) + \mathcal{O}(\delta x_i \delta x_j)$. Im letzten Schritt haben wir $\frac{\partial f}{\partial x_j}(x_i + \delta x_i, x_j)$ wiederum durch eine lineare Näherung approximiert. Nehmen wir die Änderungen in den beiden Schritten zusammen erhalten wir

$$f(x_i + \delta x_i, x_j + \delta x_j) = f(x_i, x_j) + \delta x_i \frac{\partial f}{\partial x_i}(x_i, x_j) + \delta x_j \frac{\partial f}{\partial x_j}(x_i, x_j) + \mathcal{O}(\delta x_{i/j}^2). \quad (2.17)$$

Auch dieses Ergebnis lässt sich geometrisch interpretieren, die Steigungen in den beiden Richtungen definieren eine Ebene, die am Punkt x_i, x_j tangential zu f ist, in erster Ordnung in $\delta x_i, \delta x_j$ wird f durch diese Ebene beschrieben. Dieses Argument lässt sich leicht auf Änderungen in mehr als zwei Variablen verallgemeinern.

Ableitungen höherer Ordnung sind durch Hintereinander-Ausführung von $\frac{\partial}{\partial x_i}$ und $\frac{\partial}{\partial x_j}$ definiert. Dabei gilt die **Schwarzsche Regel**

$$\frac{\partial}{\partial x_i} \left(\frac{\partial}{\partial x_j} f \right) = \frac{\partial}{\partial x_j} \left(\frac{\partial}{\partial x_i} f \right) \equiv \frac{\partial^2}{\partial x_i \partial x_j} f, \quad (2.18)$$

die sich wie folgt herleiten lässt

$$\begin{aligned} \frac{\partial}{\partial x_i} \left(\frac{\partial f}{\partial x_j} \right) &= \lim_{\delta x_i \rightarrow 0} \frac{1}{\delta x_i} \left(\frac{\partial f(x_1, \dots, x_i + \delta x_i, \dots, x_j)}{\partial x_j} - \frac{\partial f(x_1 \dots x_i \dots x_j)}{\partial x_j} \right) \\ &= \lim_{\delta x_i \rightarrow 0} \lim_{\delta x_j \rightarrow 0} \frac{1}{\delta x_i \delta x_j} (f(x_i + \delta x_i, x_j + \delta x_j) - f(x_i + \delta x_i, x_j) - f(x_i, x_j + \delta x_j) + f(x_i, x_j)) \end{aligned}$$

wenn die Grenzwerte vertauschen:

$$= \lim_{\delta x_j \rightarrow 0} \lim_{\delta x_i \rightarrow 0} \frac{1}{\delta x_i \delta x_j} (\dots) = \frac{\partial}{\partial x_j} \left(\frac{\partial f}{\partial x_i} \right)$$

Einen Beweis (der nicht auf Vertauschung der Grenzwerte beruht) gibt Kerner und v. Wahl in 9.2.8. Dort wird gezeigt dass dieser Satz gilt, wenn die zweite Ableitung von f stetig ist.

2.5 Taylorreihe

Unser Ziel bei der Einführung von Potenzreihen war es gewesen, möglichst allgemeine Funktionen als Potenzreihen darzustellen (innerhalb des Konvergenzradius). Wir betrachten eine Reihe aus den Potenzen von x ,

$$f(x) = \sum_{k=0}^{\infty} a_k x^k = a_0 + a_1 x + a_2 x^2 + a_3 x^3 + \dots \quad (2.19)$$

Dazu sind zunächst die Koeffizienten der Potenzreihe a_0, a_1, a_2 zu bestimmen. Für $k = 0$ betrachten wir die linke und rechte Seite von (2.19) bei $x = 0$.

$$f(x) = \sum_{k=0}^{\infty} a_k x^k = a_0 + a_1 x + a_2 x^2 + \dots$$

$$k = 0 : \quad f(0) = a_0 + a_1 x|_{x=0} + a_2 x^2|_{x=0} + \dots = a_0 \quad \curvearrowright a_0 = f(0)$$

$$k = 1 : \quad f'(0) = a_1 + 2 a_2 x|_{x=0} + 3 a_3 x^2|_{x=0} + \dots = a_1 \quad \curvearrowright a_1 = f'(0)$$

$$k = 2 : \quad f^{(2)}(0) = 1 \times 2 a_2 + 2 \times 3 a_3 x|_{x=0} + \dots = 2 a_2 \quad \curvearrowright a_2 = \frac{1}{2!} f^{(2)}(0)$$

$$k = 3 : \quad f^{(3)}(0) = 1 \times 2 \times 3 a_3 + 4 \times 3 \times 2 a_4 x|_{x=0} + \dots = 3! a_3 \quad \curvearrowright a_3 = \frac{1}{3!} f^{(3)}(0)$$

$$\text{allgemeines } k : \quad \curvearrowright a_k = \frac{1}{k!} f^{(k)}(0)$$

Analog können wir auch die Funktion $f(x)$ als Potenzreihe in $x - x_0$ schreiben, d.h. $f(x) = \sum_{k=0}^{\infty} a_k (x - x_0)^k$, mit neuen Koeffizienten $a_k = \frac{1}{k!} f^{(k)}(x_0)$. Die Potenzreihen in Potenzen

von x und $x - x_0$ sind natürlich eng miteinander verknüpft: Die Terme der Potenzreihe $(x - x_0)^n$ lässt sich ausmultiplizieren und die Potenzreihe in Potenzen von x schreiben, wir erhalten dann wieder (2.19) mit den entsprechenden Koeffizienten.

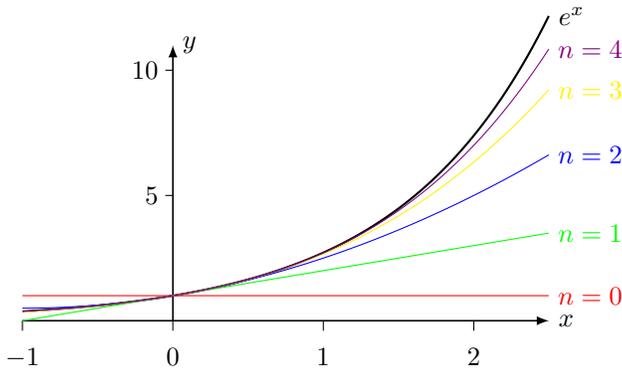
Ist die Funktion $f : (a, b) \Rightarrow \mathbb{R}$ beliebig oft differenzierbar und $x_0 \in (a, b)$, heißt

$$\sum_{k=0}^{\infty} \frac{f^{(k)}(x_0)}{k!} (x - x_0)^k$$

die **Taylorreihe** bei x_0 , ihre Koeffizienten die **Taylorkoeffizienten**. $R_n(x) = \sum_{k=n+1}^{\infty} \frac{f^{(k)}(x_0)}{k!} (x - x_0)^k$ heißt **Restglied** der Reihe (n -ter Ordnung).

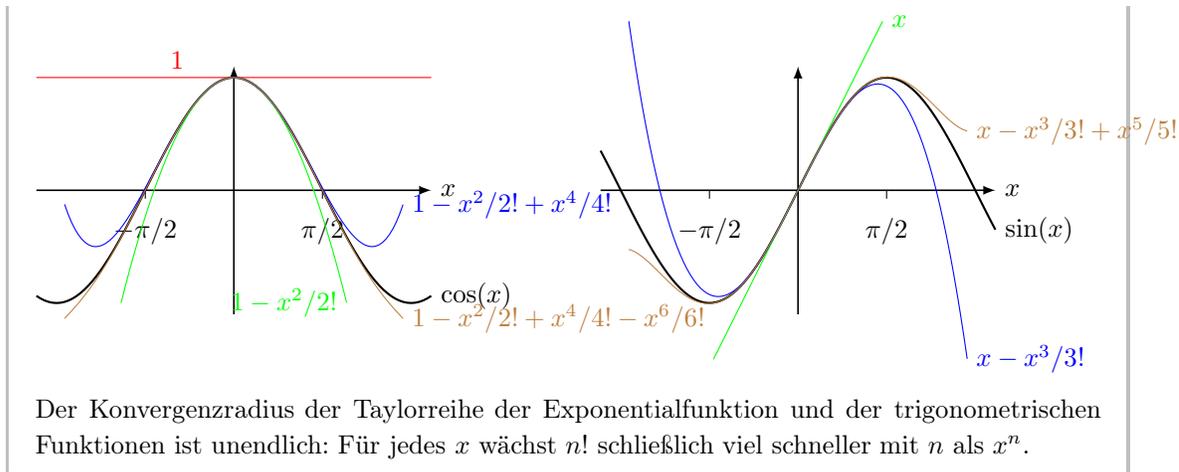
— *Beispiel* —

Die Exponentialfunktion und trigonometrische Funktionen. Die Exponentialfunktion e^x hat eine besonders einfache Taylorreihe, da ihre erste Ableitung (sowie alle weiteren Ableitungen) e^x sind, damit ist die Taylorreihe bei $x = 0$ gleich $1 + x + \frac{1}{2!}x^2 + \frac{x^3}{3!} + \frac{x^4}{4!} + \dots$. Für kleine Werte von x gibt die Taylorreihe oft bereits nach einer kleinen Zahl von Termen eine brauchbare Näherung; $e^x = 1 + x + \mathcal{O}(x^2)$ (Taylorreihe bis zur ersten Ordnung in x), $e^x = 1 + x + x^2/2 + \mathcal{O}(x^3)$ (Taylorreihe bis zur zweiten Ordnung in x , etc).



Die Abbildung zeigt die Exponentialfunktion und ihre Taylorentwicklung bei $x = 0$ bis zur n -ten Potenz von x für $n = 0, 1, 2, 3, 4$.

Als weiteres Beispiel noch die Taylorreihen für die trigonometrischen Funktionen bei $x_0 = 0$. Mit $f(x) = \cos(x)$, $\frac{df}{dx} = -\sin(x)$, $\frac{d^2f}{dx^2} = -\cos(x)$ erhalten wir die Taylor-Reihe der Cosinus-Funktion $1 - \frac{x^2}{2} + \frac{x^4}{4!} - \frac{x^6}{6!} + \dots$, und analog die Taylor-Reihe der Sinus-Funktion $x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + \dots$



Info: Wir haben nicht gefragt, welche Funktionen sich prinzipiell durch eine konvergente Taylorreihe (2.19) darstellen lassen. Funktionen die (lokal) gleich ihrer Taylorreihe sind heißen **analytisch**. Die Frage, welche Funktionen analytisch sind, findet erst im Rahmen der Funktionentheorie (Theorie von Funktionen komplexer Veränderlicher) eine zufriedenstellende Antwort.

2.6 Integration

Die Idee der Zerlegung einer Größe in eine große Zahl (kleiner) Summanden tritt in den unterschiedlichsten Zusammenhängen auf, und führt stets auf eine Art von Integral. Der Ihnen bekannte Fall, das Problems die Fläche unter einer Kurve zu bestimmen, ist das prominenteste Beispiel. Dabei wird die Fläche unter einer Kurve in eine (große) Anzahl von Streifen zerlegt, die Fläche der (kleinen, daher fast rechteckigen) Streifen berechnet, und die Fläche der Teilstücke aufaddiert.

Betrachten wir z.B. eine Ladung Q , die sich auf einem Draht befindet. Gegeben die Ladungsdichte $\rho(x)$ auf dem Draht pro Länge Draht (ein kleines Teilstück des Drahtes der Länge Δx hat also Ladung $\rho(x)\Delta x$) ist die Gesamtladung ein Integral der Ladungsdichte über die Länge des Drahtes. Ist der Draht gerade, kommen wir auf ein Integral $\int dx\rho(x)$ vom Typ "Fläche unter der Kurve". Die Idee des Zerlegens und Aufsummierens kleiner Terme funktioniert aber auch für einen gebogenen Draht, oder eine zweidimensionale Oberfläche oder ein dreidimensionales Volumen. Solche Probleme werden wir in Kapitel 5 behandeln. Oberflächen- und Volumenintegrale führen auf Mehrfachintegrale, die wir hier diskutieren.

2.6.1 Integrale von Funktionen einer Veränderlichen

Wir beginnen natürlich mit einem einfachen Fall, und greifen auf die Idee der “Fläche unter einer Kurve” zurück. Zur physikalischen Motivation können wir uns die Funktion $f(x)$ als eine (lineare) Dichte in einer Dimension denken, die Masse oder Ladung eines kleinen Teilstücks der Länge Δx ist dann $f(x)\Delta x$, oder, geometrisch gedacht, die Fläche eines schmalen Rechtecks mit Breite Δx und Höhe $f(x)$. *Klein* bedeutet hier zunächst so klein, dass die Veränderung von $f(x)$ im Intervall Δx vernachlässigt werden kann. Letztendlich werden wir den Grenzfall $\Delta x \rightarrow 0$ betrachten.

Zur Berechnung der Fläche unter der Funktion $f(x)$ zwischen x_a und x_b betrachten die Zerlegung des Intervalls $[x_a, x_b]$ in eine große Anzahl N von Teilstücken; $x_0 \equiv x_a$, $x_N \equiv x_b$ und $x_1 < x_2 < x_3 \dots$ liegen zwischen x_a und x_b . Die Werte x_0, x_1, x_2, \dots werden auch als Stützstellen bezeichnet. Die Summe über alle Teilstücke der Länge $x_1 - x_0$ (erstes Teilstück), $x_2 - x_1$ (zweites Teilstück) \dots , jeweils multipliziert mit $f(x_0)$, $f(x_1)$, \dots ergibt die Summe²

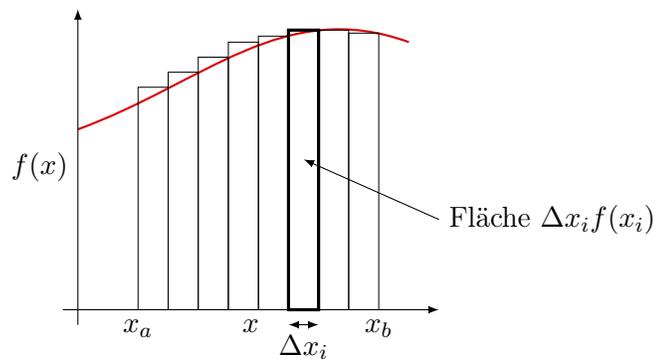
$$\sum_{i=0}^{N-1} (x_{i+1} - x_i) f(x_i) \equiv \sum_{i=0}^{N-1} \Delta x_i f(x_i) \quad (2.20)$$

mit $\Delta x_i \equiv x_{i+1} - x_i$. Diese Summe wird als Riemann-Summe bezeichnet. Im nächsten Schritt machen wir die Längen der Teilstücke sehr klein; die Rechtecke der Abbildung approximieren die Fläche unter der Kurve also immer besser. Den Grenzfall bei dem die Längen der Teilstücke Δx_i gegen Null gehen, bezeichnen wir als das Riemannintegral

$$\boxed{\int_a^b dx f(x) = \lim_{\Delta x_i \rightarrow 0} \sum_{i=0}^{N-1} \Delta x_i f(x_i)} \quad (2.21)$$

In diesem Grenzfall geht natürlich auch die Zahl der Teilstücke gegen unendlich. Wie genau der Grenzfall $\Delta x_i \rightarrow 0$ für alle Δx_i durchzuführen ist, ist damit noch nicht gesagt. Es stellt sich als ausreichend heraus, die Länge des längsten Teilstückes $\max_i(\Delta x_i)$ gegen Null zu führen (siehe z.B. Kerner und von Wahl). Funktionen, für die das Integral (2.21) existiert, heißen Riemann-integrierbar. Nach dieser Definition ist das Integral einer Funktion mit nur negativen Funktionswerten negativ, sind die Funktionswerte gehäuft positiv

² Hier kann man fragen ob nicht $f(x_{i+1})$ statt $f(x_i)$ der korrekte Ausdruck wäre. Kernpunkt des Beweises der Existenz des sogenannten Riemann-Integrals ist, dass der Unterschied zwischen diesen beiden Fällen im Grenzfall $N \rightarrow \infty$ verschwindet.



wie negativ ist das Integral null: Bei negativen Funktionswerten trägt die Fläche über der Kurve negativ zum Integral bei. Unter Vertauschung der Integrationsgrenzen ändert das Integral sein Vorzeichen, da die Δx_i alle ihr Vorzeichen ändern.

Das Integral und die Stammfunktion

Betrachten wir die Fläche unter der Kurve im Intervall $[0, x]$ als Funktion von x , und nennen diese Funktion $F(x)$. Vergrößern wir nun das Intervall durch Hinzufügen eines (beliebig kleinen) Teilstücks Δx , vergrößert sich die Fläche um $\Delta x f(x)$, also $F(x + \Delta x) = F(x) + \Delta x f(x) + \mathcal{O}(\Delta x^2)$ und somit

$$\frac{dF}{dx} = f(x) . \tag{2.22}$$

Aus dieser Gleichung ist die sogenannten **Stammfunktion** $F(x)$ von $f(x)$ nur bis auf eine additive Konstante bestimmt. Bei der Subtraktion von Stammfunktionen verschwindet diese Konstante allerdings und für allgemeine Integralgrenzen gilt dann

$$\int_a^b dx f(x) = F(b) - F(a) . \tag{2.23}$$

Diese Aussage wird als Fundamentalsatz der Analysis bezeichnet. Integration und Differentiation sind also inverse Operationen; um das Integral einer Funktion $f(x)$ zu bestimmen, suchen wir die Funktion $F(x)$, deren Ableitung nach x gleich $f(x)$ ist. Wir schreiben auch das sogenannte **unbestimmte Integral** ohne die Integrationsgrenzen anzugeben als $\int dx f(x) = F(x)$ (Gleichheit verstanden als “bis auf eine unbestimmte Konstante”).

2.6.2 Techniken zur Berechnung von Integralen

Wichtigste Technik ist die Ableitungen vieler Funktionen zu kennen, und so die Stammfunktion zu raten und zu verifizieren: Eine allgemeine Methode um systematisch zu einer beliebigen Funktion die Stammfunktion anzugeben gibt es nicht. Wir diskutieren einige nützliche Strategien.

Partielle Integration

Partielle Integration ist eine nützliche Technik, wenn der Integrand $f(x)$ sich als Produkt zweier Funktionen schreiben lässt. Sei $f(x) = u'(x)v(x)$. Aus der Produktregel der Ableitung erhalten wir $\frac{d}{dx}(u(x)v(x)) = u'(x)v(x) + u(x)v'(x)$. Umstellen der Terme und Integration auf beiden Seiten ergibt

$$\int_a^b dx u'(x)v(x) = \int_a^b dx \frac{d}{dx}(u(x)v(x)) - \int_a^b dx u(x)v'(x) = [u(x)v(x)]_a^b - \int_a^b dx u(x)v'(x) . \tag{2.24}$$

Bei geschickter Wahl von $u(x)$ und $v(x)$ ist das resultierende Integral oft einfach zu bestimmen.

Integration durch Substitution

Dieses Integrationsprinzip folgt aus der Kettenregel der Differentialrechnung. Gegeben sei eine Funktion $f(x)$ mit Stammfunktion $F(x)$. Wir betrachten nun eine Funktion $x(y)$ mit Umkehrfunktion $y(x)$ und definieren eine weitere Stammfunktion $G(y) = F(x(y))$ mit Ableitung $g(y)$. Nach der Kettenregel ist $\frac{dG}{dy} = \frac{dF}{dx} \frac{dx}{dy}$, bzw. $g(y) = f(x(y)) \frac{dx}{dy}$. Integration über x ergibt

$$\int_{y(a)}^{y(b)} dy \frac{dx}{dy} f(x(y)) = \int_{y(a)}^{y(b)} dy g(y) = [G(y)]_{y(a)}^{y(b)} = [F(x(y))]_{y(a)}^{y(b)} = [F(x)]_a^b = \int_a^b dx f(x). \quad (2.25)$$

Ist $f(x)$ schwer zu integrieren, kann eine geschickt gewählte Substitution auf eine Funktion $\frac{dx}{dy} f(x(y))$ führen, die leichter zu integrieren ist.

Beispiel

Gesucht sei das Integral der Funktion $\sqrt{1-x^2}$ über das Intervall $[0, 1]$. Da $\sqrt{1-\sin(t)^2} = \cos(t)$ erwarten wir, dass eine trigonometrische Variablensubstitution nützlich könnte. Wir führen eine neue Variable $\sin(y) = x$, bzw. $y = \arcsin(x)$ ein und erhalten mit $\frac{dx}{dy} = \cos(y)$

$$\int_0^1 dx \sqrt{1-x^2} = \int_{\arcsin(0)}^{\arcsin(1)} dy \frac{dx}{dy} \sqrt{1-\sin^2(y)} = \int_0^{\pi/2} dy \cos^2(y), \quad (2.26)$$

das mit Hilfe partieller Integration berechnet werden kann. Oder mit der trigonometrischen Formel $\cos^2(y) = \frac{1+\cos(2y)}{2}$.

Im obigen Beispiel und in (2.25) muss $x(y)$ bijektiv sein, damit wir die Integralgrenzen nach der Substitution $y(a)$ und (b) angeben können (durch Aufteilen in Intervalle auf denen $x(y)$ dann bijektiv ist lässt sich das leicht beheben). Umkehrbarkeit ist allerdings nicht bei jeder Variablensubstitution nötig. Wenn ein Integrand in der Form $\int_a^b dx \frac{dy}{dx} f(y(x))$ geschrieben werden kann führt Variablensubstitution auf $\int_{y(a)}^{y(b)} dy f(y)$ ohne dass eine Umkehrfunktion gebraucht wurde.

Die Variablensubstitution kann geometrisch interpretiert werden. Ein $\int_a^b dx f(x) \approx \sum_i \Delta x f(x_i)$ unterteilen wir das Integrationintervall in Teilintervalle konstanter Länge Δx mit "Stützstellen" bei $0, \Delta x, 2\Delta x, \dots$. Dasselbe Intervall lässt sich mit neuen Stützstellen bei $y(0), y(0) + \Delta y, y(0) + 2\Delta y, \dots$ beschreiben, die Intervalllängen in beiden Fällen sind durch $\Delta x \approx \Delta y \frac{dx}{dy}$ verknüpft. Eine Merkregel für die Variablensubstitution (2.25) ist $dy = dx \frac{dy}{dx}$ (eine Aussage die hier nur unter dem Integral Sinn macht).

Differenzieren nach einem Parameter

Wenn der Integrand durch Ableitung nach einem Parameter aus einer Funktion hervorgeht, können wir das ausnutzen.

Beispiel

Gesucht sei das Integral $\int_0^\infty dx x e^{-x}$, das wir als $\int_0^\infty dx x e^{-\alpha x}$ mit $\alpha = 1$ schreiben.

$$\begin{aligned}\int_0^\infty dx x e^{-\alpha x} &= \int_0^\infty dx \left(-\frac{d}{d\alpha}\right) e^{-\alpha x} = \left(-\frac{d}{d\alpha}\right) \int_0^\infty dx e^{-\alpha x} \\ &= -\frac{d}{d\alpha} \left[-\frac{1}{\alpha} e^{-\alpha x}\right]_0^\infty = -\frac{d}{d\alpha} \alpha^{-1} = \frac{1}{\alpha^2}\end{aligned}\quad (2.27)$$

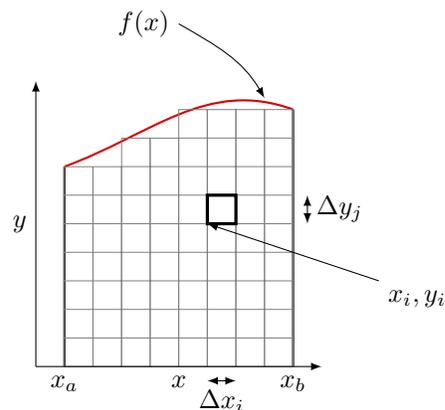
Im zweiten Schritt haben wir Integration und Ableitung vertauscht. Für die endliche Riemannsumme (2.20) ist dies natürlich zulässig. Für das Riemannintegral ist es das mit Einschränkung auch (der Integrand muss stetig sein und seine Ableitung ebenso, siehe z.B. Kerner und von Wahl).

2.7 Mehrfachintegrale

Ziel ist es, das Integral als Riemannsumme über eine Variable zu verallgemeinern. Anstatt (kleine) Intervalle einer Geraden (x -Achse) aufzusummieren, können wir dann kleine Flächenelemente oder Volumenelemente addieren. Damit können wir dann auch über Oberflächen und Volumen integrieren.

Beispiel

In einem sehr beschränkten Sinne haben wir bereits ein Oberflächenintegral berechnet. Die Fläche unter der Kurve $f(x)$ im Intervall $[x_a, x_b]$ lässt sich nicht nur in schmale Rechtecke mit Fläche $\Delta x f(x)$ zerlegen, sondern auch in kleine Flächenelemente mit Fläche $\Delta x \Delta y$. Diese Flächenelemente haben ihre linke untere Ecke bei x_i, y_i und eine Fläche $\Delta x \Delta y$. Als Riemannsumme ist dann die Fläche unter der Kurve $f(x)$ die Summe der Flächen aller Elemente. Die Zahl der Flächenelemente entlang der y -Achse hängt hier von $f(x)$ ab, $N_y(i) \approx f(x_i)/\Delta y$ für konstante Δy .



Schreiben wir die Funktion $f(x) = \int_0^{f(x)} dy$, dann ist die Fläche unter der Kurve

$$\int_{x_a}^{x_b} dx f(x) = \int_{x_a}^{x_b} dx \int_0^{f(x)} dy \equiv \int_{x_a}^{x_b} \int_0^{f(x)} dx dy \quad (2.28)$$

Bei dieser Berechnung der Fläche wurden einzelne Flächenelemente aufsummiert. Das ist natürlich ein sehr eingeschränktes Problem. Auf einer 2D Oberfläche S könnte sich aber zum Beispiel elektrische Ladung mit einer Ladungsdichte $\rho(x, y)$ pro kleiner Fläche befinden. Die Riemannsumme der Gesamtladung ist dann $\sum_{i,j}^{N_x, N_y} \Delta x_i \Delta y_j \rho(x_i, y_j)$ und führt im Grenzfall beliebig kleiner Flächenelemente auf das Mehrfachintegral

$$\lim_{\Delta x_i \rightarrow 0, \Delta y_j \rightarrow 0} \sum_{i=0, j=0}^{N_x-1, N_y-1} \Delta x_i \Delta y_j \rho(x_i, y_j) = \iint_S dx dy \rho(x, y), \quad (2.29)$$

(wenn der Grenzfall existiert)³. Zur konkreten Berechnung der Riemannsumme könnte man zunächst die Summe auf Spalten mit konstantem x beschränken, und dann die Beiträge der Spalten addieren, oder umgekehrt zunächst die Beiträge von Zeilen mit konstantem y addieren und dann die Beiträge aller Zeilen zusammennehmen. Im Grenzfall $\Delta x_i \rightarrow 0, \Delta y_j \rightarrow 0$ erhalten wir im ersten Fall zunächst ein Integral $\int dy f(x, y)$ das dann noch über x integriert werden muss, und umgekehrt im zweiten Fall. Für Summen mit unendlich vielen Termen ist die Gleichheit nicht selbstverständlich. Der **Satz von Fubini**

$$\iint dx dy \rho(x, y) = \int dx \left(\int dy \rho(x, y) \right) = \int dy \left(\int dx \rho(x, y) \right). \quad (2.30)$$

(s. Kerner und von Wahl) gilt für stetige Funktionen $\rho(x, y)$ und besagt, dass es keine Rolle spielt, ob das Mehrfachintegral zunächst über die x -Variable oder über die y Variable ausgewertet wird.

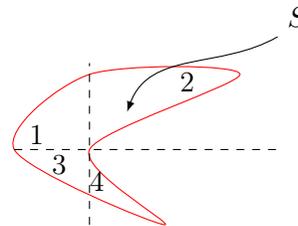
Beispiel

Wir berechnen die Masse einer rechteckigen Platte S mit Begrenzung durch die Linien $x = 0, x = 3, y = 0, y = 2$ mit Dichte pro Fläche $\rho(x, y) = xy$. Die Masse eines kleinen rechteckigen Teilstücks mit Kantenlängen $\Delta x, \Delta y$ ist $\approx xy \Delta x \Delta y$, die Gesamtmasse

$$M = \iint_S dx dy (xy) = \int_0^3 \int_0^2 dx dy (xy) = \left(\int_0^3 dx x \right) \left(\int_0^2 dy y \right) = 9/2 \times 4/2 = 9. \quad (2.31)$$

³ Da klar ist wie viele Variablen hier integriert werden schreibt man auch $\int_S dx dy \rho(x, y)$, S bezeichnet die Fläche über die integriert wird.

Diese Vertauschbarkeit der Integrationen führt uns auf das Thema der Integrationsgrenzen, die bei Mehrfachintegralen komplizierter sein können als bei Integralen einer Veränderlichen. In (2.29) haben wir die Integrationsgrenzen nicht spezifiziert, und einfach ein Integral über die Fläche S angegeben. Die expliziten Integrationsgrenzen eines Mehrfachintegrals hängen davon ab, in welcher Reihenfolge wir die beiden Integrationen auswerten (das Ergebnis selbst hängt nicht von der Reihenfolge ab), also ob wir die Flächenelemente in horizontale oder vertikale Streifen zusammensetzen. Werten wir im Beispiel oben zunächst das Integral (2.28) über y (bei konstantem x) aus, sind die Integrationsgrenzen einfach $\int_{x_a}^{x_b} \int_0^{f(x)}$. Werten wir zunächst das Integral über x (bei konstantem y) aus sind die Integrationsgrenzen bei diesem Beispiel komplizierter. Einige Oberflächen zerlegt man sinnvollerweise in Teilstücke, die separat ausgewertet werden (s. rechts).



Das Volumenintegral

Die gleiche Denkweise wie beim Integral über eine Oberfläche lässt sich auch auf Volumina anwenden. Das Volumen eines Quaders mit Kantenlängen $\Delta x_i, \Delta y_j, \Delta z_k$ ist $\Delta x_i \Delta y_j \Delta z_k$. Gibt $\rho(x, y, z)$ zum Beispiel eine Dichte pro kleinem Volumenelement an, dann ist die in einem Volumen V enthaltene Masse durch die Riemannsumme

$$\sum_{i,j,k} \Delta x_i \Delta y_j \Delta z_k \rho(x_i, y_j, z_k) \quad (2.32)$$

approximiert, wobei die Summe über die Indizes i, j, k alle (kleinen) Volumenelemente addiert. Im Grenzfalle unendlich kleiner Volumenelemente $\lim_{\Delta x \rightarrow 0, \Delta y \rightarrow 0, \Delta z \rightarrow 0}$ erhalten wir das Volumenintegral

$$\int_V dx dy dz \rho(x, y, z) = \int dx \int dy \int dz \rho(x, y, z) . \quad (2.33)$$

Wieder können die Integrationen in beliebiger Reihenfolge vorgenommen werden (unter milden Forderungen an den Integranden).

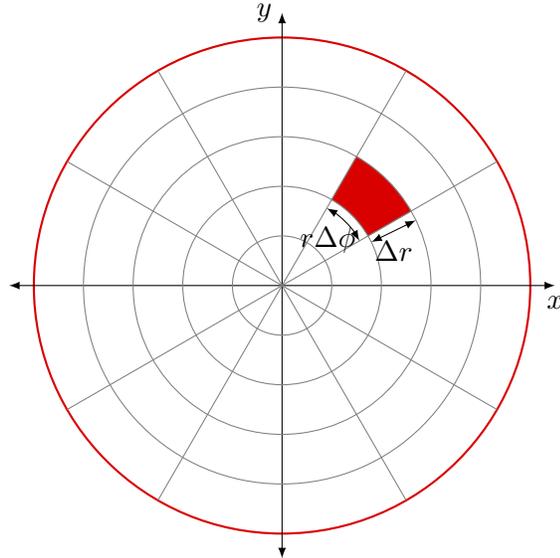
2.7.1 Substitution der Variablen in Mehrfachintegralen

Mehrfachintegrale sind also prinzipiell nicht schwieriger zu berechnen als Integrale von Funktionen einer Veränderlichen. Allerdings können die Integrationsgrenzen Schwierigkeiten machen. Geschickte Änderung der Integrationsvariablen kann hier helfen. Mit einer Variablensubstitution können Symmetrien eines Problems ausgenutzt werden, z.B. wenn Integrationsgrenzen oder der Integrand eine bestimmte Symmetrie aufweist. Als konkretes

Beispiel betrachten wir die Variablentransformation von kartesischen Koordinaten (x, y) zu Polarkoordinaten (r, θ) in zwei Dimensionen. Bei dem Flächenintegral über die Kreisfläche mit Radius 1

$$\int_S dx dy \rho(x, y) = \int_{-1}^1 dx \int_{-\sqrt{1-x^2}}^{\sqrt{1-x^2}} dy \rho(x, y) \quad (2.34)$$

ist vor allem der Fall interessant, bei dem auch der Integrand nur von der Entfernung r eines Punktes p von der Kreismitte abhängt, $\rho(x, y) = \rho(r)$. Dann können sowohl der Integrand als auch die Integrationsgrenzen durch eine einzelne Variable ausgedrückt werden. Wir definieren 2 neue Variablen, die sogenannten **Polarkoordinaten**. Die erste Koordinate ist die **Radialkoordinate** $r(x, y) = \sqrt{x^2 + y^2}$, die den Abstand vom Ursprung angibt. Die zweite Koordinate, die **Winkelkoordinate** ϕ , ist definiert durch den Winkel, den die Linie vom Ursprung zum Punkt p mit der positiven x -Achse einschliesst, $\tan \phi = y/x$. Mit der Umkehrfunktion des Tangens und Kenntnis in welchen der vier Quadranten p liegt kann die Winkelkoordinate leicht bestimmt werden.



Unser Ziel ist es, durch eine Variablensubstitution das Integral über die Kreisfläche als Integral über r und ϕ zu schreiben. Eine Vielzahl von Problemen weist eine Art von Symmetrie auf (hier: Symmetrie des Integranden und der Integrationsgrenzen unter Rotation), nutzt man neue Variablen erhält man oft einfacher zu lösende Integrale.

Jeder Punkt in der Ebene ist eindeutig durch ein Paar (x, y) und ein Paar (r, ϕ) , $\phi \in [0, 2\pi]$ beschrieben (bis auf einen einzelnen Punkt, den Ursprung, dort ist ϕ nicht eindeutig definiert); die Abbildung von (x, y) auf (r, ϕ) ist also invertierbar. Die Linien mit konstanter Variable r bilden Ringe, Linien mit konstantem ϕ bilden radial nach außen laufende Geraden mit Steigung $\tan(\phi)$. Eine kleine Änderung der Koordinaten (x, y) zu $(x + \Delta x, y + \Delta y)$ definiert ein kleines Rechteck der Fläche $\Delta x \Delta y$. Eine kleine Änderung der Koordinaten r, ϕ definiert eine Intersektion aus Ring und Kuchenstück, siehe Abbildung. Die Fläche dieses Objekts ist (für kleine Änderungen $\Delta r, \Delta \phi$, wenn dieses Objekt fast ein Rechteck ist) $\Delta r (r \Delta \phi)$. Damit ist das Integral (2.34)

$$\int_S dx dy \rho(x, y) = \int_0^1 \int_0^{2\pi} dr d\phi r \rho(r) = 2\pi \int_0^1 dr r \rho(r) . \quad (2.35)$$

Der zusätzliche Faktor r ergibt sich aus der Substitution der Variablen, man kann ihn als höherdimensionales Equivalent des Faktors $\frac{dx}{dy}$ in (2.25) auffassen. Weitere Variablentransformationen dieser Art werden in den Übungen diskutiert. Den Transformationssatz, der allgemeine Variablentransformationen behandelt werden wir im nächsten Semester kennenlernen.

Thema dieses Kapitels ist die Erweiterung der reellen Zahlen auf die komplexen Zahlen.

Die natürlichen Zahlen bilden die Basis, auf der weitere Zahlkörper aufbauen. Zum Beispiel wurden ganze Zahlen $\dots, -3, -2, -1, 0, 1, 2, 3, \dots$ eingeführt um Gleichungen wie $x + 3 = 2$ lösen zu können. Analoge Erweiterungen der natürlichen Zahlen führten auf die rationalen Zahlen und die irrationalen Zahlen. Nur Gleichungen wie $x^2 = -1$ blieben bislang ohne Lösung. Zur Lösung dieser Gleichungen führen wir einen Zahlkörper mit einem neuen Element i ein, das die Eigenschaft $i^2 = -1$ hat.

3.1 Komplexe Zahlen und \mathbb{R}^2

Wie betrachten den Vektorraum $\mathbb{R}^2 = \left\{ \begin{pmatrix} a \\ b \end{pmatrix} : a, b \in \mathbb{R} \right\}$ mit der Addition von Vektoren und Multiplikation mit reellen Zahlen definiert durch

$$\begin{pmatrix} a \\ b \end{pmatrix} + \begin{pmatrix} a' \\ b' \end{pmatrix} = \begin{pmatrix} a + a' \\ b + b' \end{pmatrix} \quad (3.1)$$

$$\lambda \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} \lambda a \\ \lambda b \end{pmatrix} \quad \lambda \in \mathbb{R} . \quad (3.2)$$

Zusätzlich definieren wir noch eine Multiplikation von zwei Vektoren

$$\begin{pmatrix} a \\ b \end{pmatrix} \cdot \begin{pmatrix} a' \\ b' \end{pmatrix} = \begin{pmatrix} a a' - b b' \\ a b' + b a' \end{pmatrix} . \quad (3.3)$$

Diese Multiplikation

ist kommutativ:
$$\begin{pmatrix} a' \\ b' \end{pmatrix} \cdot \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} a'a - b'b \\ a'b + b'a \end{pmatrix} = \begin{pmatrix} a'a' - b'b' \\ a'b' + b'a' \end{pmatrix} = \begin{pmatrix} a \\ b \end{pmatrix} \cdot \begin{pmatrix} a' \\ b' \end{pmatrix}$$

hat ein Einselement:
$$\begin{pmatrix} 1 \\ 0 \end{pmatrix} \cdot \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} a \\ b \end{pmatrix}$$

hat ein Inverses:
$$\begin{pmatrix} a \\ b \end{pmatrix}^{-1} \equiv \frac{1}{a^2 + b^2} \begin{pmatrix} a \\ -b \end{pmatrix} \quad \begin{pmatrix} a \\ b \end{pmatrix}^{-1} \cdot \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

und erlaubt eine Lösung von $x^2 = -1$

$$\begin{pmatrix} 0 \\ 1 \end{pmatrix} \cdot \begin{pmatrix} 0 \\ 1 \end{pmatrix} = - \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

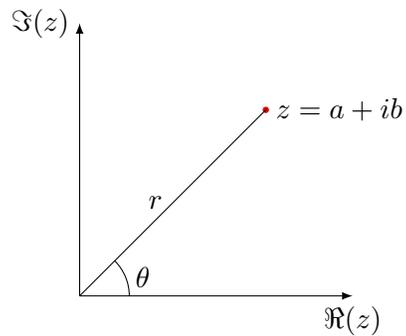
Wir schreiben $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$ als 1 und $\begin{pmatrix} 0 \\ 1 \end{pmatrix}$ als i mit $i \cdot i = -1$. $z = a + ib$ mit $a, b \in \mathbb{R}$ entspricht dann

$\begin{pmatrix} a \\ b \end{pmatrix}$ und wird als **komplexe Zahl** bezeichnet. Jede polynome (algebraische) Gleichung n -ten Grades (mit komplexen Koeffizienten) hat n Lösungen in den komplexen Zahlen (Fundamentalsatz der Algebra). Mit der Erweiterung auf die komplexen Zahlen sind wir also am Endpunkt einer Entwicklung angekommen: alle algebraischen Gleichungen sind nun lösbar¹.

3.2 Realteil und Imaginärteil, Betrag und Argument komplexer Zahlen

a wird als **Realteil**, b als **Imaginärteil** einer komplexen Zahl $z = a + ib$ bezeichnet. Wir schreiben $a = \Re(a + ib)$ und $b = \Im(a + ib)$. Es sind also zwei reelle Zahlen nötig um eine einzelne komplexe Zahl zu spezifizieren. Analog zur Zahlengeraden der reellen Zahlen, werden komplexe Zahlen graphisch in der Ebene dargestellt (**komplexe Ebene** oder **Gaußsche Ebene**). Reelle Zahlen liegen dabei auf der x -Achse, rein imaginäre Zahlen auf der y -Achse.

Eine komplexe Zahl $z = a + ib$ kann auch in Polarkoordinaten $r = \sqrt{a^2 + b^2}$, $\tan \theta = \frac{b}{a}$ dargestellt werden. $r = |z| \equiv \sqrt{a^2 + b^2}$ wird als **Betrag** einer komplexen Zahl z bezeichnet, θ als ihr **Argument**. $z^* = a - ib$ definiert die **komplex Konjugierte** von $z = a + ib$. Dann ist $z z^* = (a + ib)(a - ib) = a^2 - i ab + i ba - (i)^2 b^2 = a^2 + b^2 = |z|^2$



¹Das heißt natürlich nicht, dass nicht noch weitere Zahlkörper definiert werden können oder keine interessanten Eigenschaften haben.

Multiplikation von komplexen Zahlen kann einfach durch Ausmultiplizieren erfolgen

$$zz' = (a + ib)(a' + ib') = aa' + iab' + iba' + (i \cdot i)bb' = (aa' - bb') + i(ab' + ba') , \quad (3.4)$$

das Ergebnis stimmt mit (3.3) überein. Für die Division gilt nutzen wir $\frac{z_1}{z_2} = \frac{z_1}{z_2} \frac{z_2^*}{z_2^*} = \frac{z_1 z_2^*}{|z_2|^2}$ (in Übereinstimmung mit dem inversen Element der Multiplikation).

3.3 Die Euler-Formel

Wir nutzen die Taylor-Reihe der Exponentialfunktion, um die Exponentialfunktion auch für komplexe Zahlen zu definieren.

$$\begin{aligned} e^{ix} &\equiv \sum_{k=0}^{\infty} \frac{(ix)^k}{k!} = 1 + ix - \frac{x^2}{2} + \dots \\ &= \sum_{k=0,2,4,6,\dots} \frac{(ix)^k}{k!} + \sum_{k=1,3,5,7,\dots} \frac{(ix)^k}{k!} \\ &= \sum_{k=0}^{\infty} (-1)^k \frac{x^{2k}}{(2k)!} + i \sum_{k=0}^{\infty} (-1)^k \frac{x^{2k+1}}{(2k+1)!} \\ &= \cos x + i \sin x \end{aligned}$$



$$\boxed{e^{ix} = \cos x + i \sin x} \quad (3.5)$$

Die sogenannte **Eulerformel** verbindet trigonometrische Formeln mit der Exponentialfunktion. Eine komplexe Zahl z mit Betrag r und Argument θ ist damit $z = re^{i\theta}$. Das Produkt zweier komplexer Zahlen z_1 und z_2 ist dann $z_1 z_2 = r_1 r_2 e^{i\theta_1} e^{i\theta_2} = r_1 r_2 e^{i(\theta_1 + \theta_2)}$, d.h. der Betrag des Produktes ist $r_1 r_2$, sein Argument ist $\theta_1 + \theta_2$. Multiplikation komplexer Zahlen wird also durch Drehungen in der komplexen Ebene beschrieben.

Diese Eigenschaft komplexer Zahlen führt auf ihre wichtigste praktische Anwendung: Bei Schwingungen und Oszillationen erlaubt die Eulerformel (komplizierte) trigonometrische Berechnungen mit der (einfachen) Exponentialfunktion zu führen.

Beispiel

Eine einfache Anwendung der Eulerschen Formel sind die trigonometrischen Additionstheoreme

$$\begin{aligned} \cos(x+y) + i \sin(x+y) &= e^{i(x+y)} = e^{ix} e^{iy} \\ &= (\cos x + i \sin x)(\cos y + i \sin y) \\ &= \cos x \cos y - \sin x \sin y + i(\sin x \cos y + \sin y \cos x) . \end{aligned} \tag{3.6}$$

Aus dem Realteil dieser Gleichung erhalten wir $\cos(x+y) = \cos x \cos y - \sin x \sin y$, aus dem Imaginärteil $\sin(x+y) = \sin x \cos y + \sin y \cos x$.

3.4 Wurzeln komplexer Zahlen und der komplexe Logarithmus

Die Eulerformel erlaubt auch Wurzeln und Logarithmen komplexer Zahlen zu definieren. Für die n -te Potenz einer komplexen Zahl $z = re^{i\theta}$ gibt die Eulerformel (3.5)

$$z^n = r^n e^{in\theta} , \tag{3.7}$$

ein Ergebnis das auch als **Formel von de Moivre** bezeichnet wird. Diese Beziehung lässt sich nutzen, um auch die n -te Wurzel einer komplexen Zahlen zu definieren. Dabei muss allerdings beachtet werden, dass man auf der rechten Seite von (3.7) noch $1 = e^{2\pi ki}$ mit $k \in \{0, 1, \dots, n-1\}$ hinzufügen kann. Entsprechend ist die Wurzel einer komplexen Zahl nicht eindeutig

$$z^{1/n} = r^{1/n} \exp\left\{i \frac{\theta + 2k\pi}{n}\right\} , \tag{3.8}$$

sondern eine komplexe Zahl hat n komplexe n -te Wurzeln.

Aufgabe

Malen sie die 2-ten, 3-ten und 4-ten Wurzeln von 1 in der komplexen Ebene auf.

Eine komplexe Zahl w mit $e^w = z$ heißt der komplexe (natürliche) Logarithmus von z . Damit ist auch eine Zahl $w + 2k\pi i$ mit ganzzahligem k ein Logarithmus von z , d.h. es gibt eine unendliche Zahl von komplexen Logarithmen. Diese Mehrdeutigkeit spiegelt die Periodizität der komplexen Exponentialfunktion wider. Oft ist es sinnvoll, sich auf einen dieser Logarithmen zu beschränken. Konvention ist es, den Imaginärteil von w auf das Intervall $] -\pi, \pi]$ zu beschränken. Der so definierte Logarithmus heißt Hauptwert von w . “Vergisst” man die Mehrdeutigkeit der komplexen Wurzel oder des Logarithmus, kann dies allerdings absurde Ergebnisse nach sich ziehen. Kurz: $1 = e^{2\pi ki}$, aber daraus folgt nicht $\log(1) = 0 = 2\pi ki \ \forall k \in \mathbb{Z}$.

Info: Die komplexen Zahlen erscheinen zunächst abstrakt und schwer zugänglich. Das galt für die irrationalen Zahlen oder die negativen Zahlen aber sicher auch einmal. In der Mathematik spielen die komplexen Zahlen wegen der sogenannten algebraischen Abgeschlossenheit eine wichtige Rolle: Jede algebraische Gleichung von Grad größer Null hat eine Lösung in den komplexen Zahlen. In der Physik ist der Zusammenhang zwischen Trigonometrie und der Exponentialfunktion zentral bei der Beschreibung von Schwingungen und Oszillationen. In Kapitel 4 werden wir davon extensiv Gebrauch machen. Und schließlich kann man auch einen Vektorraum über den komplexen Zahlen definieren, also einen Vektorraum dessen Elemente (statt mit reellen Zahlen) mit komplexen Zahlen multipliziert werden können. Ein solcher Vektorraum spielt die zentrale Rolle in der Quantenmechanik.

Vier

Differentialgleichungen

Poincare map <https://www.youtube.com/watch?v=CgFVRVB7jM0>

Physik ist in weiten Teilen Naturbeschreibung mit Differentialgleichungen: viele Naturgesetze können mit Gleichungen formuliert werden, die eine Funktion sowie Ableitungen dieser Funktion enthalten. Wir lernen unterschiedliche Arten von Differentialgleichungen kennen und entwickeln Strategien zur Lösung von Differentialgleichungen.

4.1 Differentialgleichungen in der Physik

Einzelne Messungen können durch Zahlen beschrieben werden, die ihnen zugrundeliegenden Prozesse sind jedoch (meist) durch Funktionen beschrieben, z.B. die Bahnkurve $\mathbf{r}(t)$ eines Teilchens oder ein Potential $V(x, y, z)$.

Oft gehorchen diese Funktionen Differentialgleichungen

$$m \frac{d^2 \vec{r}}{dt^2} = \vec{F}(\vec{r}, t) \quad \text{Newton II}$$

$$\nabla \cdot \vec{E} = \rho(\vec{r})/\epsilon_0 \quad \text{Maxwell I}$$

Eine Differentialgleichung (DGL) ist eine Gleichung für eine unbekannte Funktion, in der Ableitungen dieser Funktion auftreten. Die klassische Mechanik, Elektrodynamik, Quantenmechanik basieren alle auf Differentialgleichungen. Bevor wir unterschiedliche Arten von DGL zu klassifizieren und lösen lernen, diskutieren wir mehrere Beispiele.

Das zweite Newtonsche Gesetz in einer Dimension, beschrieben durch die Variable y

$$F(y(t), t) = ma(t) = m \frac{d^2 y(t)}{dt^2} ,$$

verbindet die Beschleunigung einer Punktmasse mit der Kraft die auf die Punktmasse wirkt. Letztere kann sowohl von der Zeit, als auch vom Ort abhängen. Wir betrachten mehrere Spezialfälle.

1. Konstante Kraft $F = mg$. Dieser Fall beschreibt die Bewegung eines Teilchens in einem konstanten Gravitationsfeld

$$\frac{d^2y}{dt^2} = g \quad (4.1)$$

Diese DGL ist eine Gleichung für eine Funktion $y(t)$ der Zeit, d.h. rechte und linke Seite sind Funktionen der Zeit (die rechte Seite ist in diesem Beispiel eine konstante Funktion). Die gesuchte Lösung, eingesetzt in die linke Seite der Differentialgleichung, muss für alle Zeiten t gleich der rechten Seite sein. Gesucht ist eine also Funktion $y(t)$, deren zweite Ableitung nach t für alle Zeiten gleich g ist.

$y(t) = \frac{1}{2}gt^2 + bt + a$ löst diese Gleichung, da $\frac{d^2}{dt^2}(gt^2/2) = g$ und $\frac{d^2}{dt^2}(at + b) = 0$. Die Koeffizienten a und b sind aus den Anfangsbedingungen $y(t=0) = y_0$, $\frac{dy}{dt}(0) = v_0$ zu bestimmen: Einsetzen von $t = 0$ gibt $a = y_0$ und $b = v_0$.

2. Zeitabhängige Kraft $F = Am \cos(\Omega t)$. Dieser Fall beschreibt ein Teilchen, das einer oszillierenden Kraft unterworfen ist

$$\frac{d^2y}{dt^2} = A \cos(\Omega t) \quad (4.2)$$

$y(t) = -\frac{A}{\Omega^2} \cos(\Omega t) + at + b$ löst die Gleichung, da $\frac{d^2}{dt^2} \cos(\Omega t) = -\Omega^2 \cos(\Omega t)$ und $\frac{d^2}{dt^2}(at + b) = 0$.

3. Ortsabhängige Kraft $F(y) = -ky$. Dieser Fall beschreibt eine Punktmasse an einer linearen Feder.

$$\frac{d^2y}{dt^2} = -\frac{k}{m} y(t) \equiv -\omega_0^2 y(t), \quad \omega_0^2 \equiv \frac{k}{m} \quad (4.3)$$

Diese DGL unterscheidet sich qualitativ von den ersten beiden Beispielen. Gesucht ist eine Funktion $y(t)$, deren zweite Ableitung nicht gleich einer bekannten Funktion ist, sondern proportional zur gesuchten Funktion selbst. Wir werden effektive Strategien zur Lösung solcher DGL kennenlernen, hier erraten wir die Lösung. Wir erwarten eine Oszillation der Masse an der Feder mit zeitlich konstanter Amplitude; $y(t) = A \cos(\omega_0 t - \phi)$. Einsetzen in die DGL zeigt, dass dieser Ansatz tatsächlich die DGL löst, da zweifaches Ableiten einer Kosinus oder Sinusfunktion wieder eine Kosinus/Sinusfunktion ist.

4.2 Terminologie

Eine Gleichung für eine Funktion, die eine oder mehrere Ableitung enthält, heißt **Differentialgleichung**, die höchste Ordnung dieser Ableitungen heißt **Ordnung der DGL**.

Die Variable, nach der abgeleitet wird, heißt **unabhängige Variable**, die Variable, die abgeleitet wird, heißt **abhängige Variable**. DGL, die nur Ableitungen nach einer einzelnen Variablen enthalten, heißen **gewöhnliche DGL**. DGL, die (partielle) Ableitungen nach mehreren Variablen enthalten, heißen **partielle DGL**.

4.3 Gewöhnliche DGL erster Ordnung

Beispiel

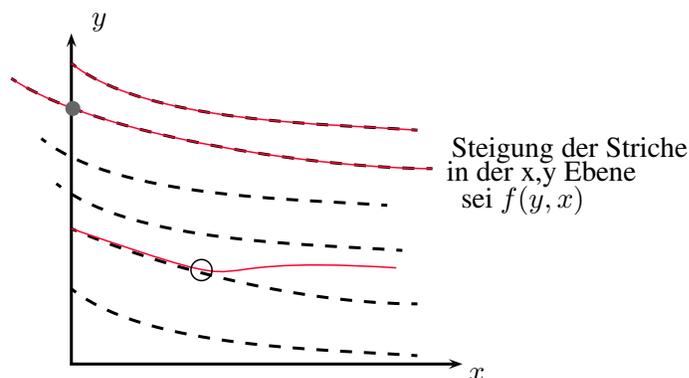
Radioaktiver Zerfall. Die Zahl der pro kleinem Zeitintervall zerfallenden Kerne ist proportional zur Zahl der vorhandenen Kerne y

$$\frac{dy}{dt} = -\lambda y(t) \quad (4.4)$$

Die Zahl radioaktiver Kerne nimmt exponentiell mit der Zeit ab: tatsächlich wird $\frac{dy}{dx} + \lambda y(t) = 0$ gelöst durch $y(t) = y_0 e^{-\lambda t}$. Die Konstante y_0 kann aus der Anfangsbedingung $y(x = 0) = y_0 e^{-\lambda 0} = y_0$ bestimmt werden.

An dieser Stelle ist eine konzeptionelle Hürde zu überwinden: die DGL $\frac{dy}{dx} = -\lambda y(t)$ ist auf der rechten Seite eine Funktion von y (hier die Funktion y selbst): Die Zahl der pro Zeiteinheit zerfallenden Kerne ist gegeben durch λy , proportional zur Zahl der Kerne y . Die DGL setzt sie gleich der negativen Ableitung von y nach der Zeit. Finden wir allerdings die Lösung $y(t) = Ae^{-\lambda t}$ und setzen sie in die DGL ein, steht auf der rechten Seite eine Funktion von t allein, nämlich $-\lambda Ae^{-\lambda t}$. Dennoch ist die DGL selbst eine Gleichung, die explizit y und t enthalten kann, erst ihre Lösung etabliert eine Beziehung $y = y(t)$ ¹. Die rechte Seite einer DGL wie (4.4) kann also sowohl von der unabhängigen Variablen abhängen, hier t , als auch von der abhängigen Variable, hier y .

¹Nicht anders ist es bei einer algebraischen Gleichung wie $x - 3 = 4$. Prinzipiell kann x jeden reellen Wert annehmen. Aber nur $x = 7$ löst die Gleichung. In dem obigen Beispiel kann die Zahl der Kerne $y(t)$ zum Zeitpunkt t jeden Wert annehmen, aber nur $y(t) = Ae^{-\lambda t}$ löst die DGL $\frac{dy}{dt} = -\lambda y(t)$.



Eine allgemeine DGL kann nach $\frac{dy}{dx}$ aufgelöst werden (linke Seite), auf der rechten Seite stehen dann Terme, die von der abhängigen Variable y und vielleicht noch von der unabhängigen Variable x abhängen. In dem Beispiel des radioaktiven Zerfalls ist die DGL (4.4) bereits nach $\frac{dy}{dx}$ aufgelöst, auf der rechten Seite steht ein Term, der nur von y abhängt und als Funktion von y geschrieben werden kann $\frac{dy}{dx} = f(y(x))$ mit $f(y) = -\lambda y$. Den letzten Ausdruck können wir auch als $f(y(x)) = y(x)$ schreiben, $f(y) = y$ betont, dass wir die rechte Seite dieser DGL kennen wenn wir y kennen (wenn 1000 Atome vorhanden sind, werden pro kleiner Zeiteinheit 1000λ Atome zerfallen, dazu müssen wir nicht die Uhrzeit (abhängige Variable) wissen). Die DGL erster Ordnung zum Beispiel

$$\left(\frac{dy}{dx}\right)^3 = \sin(y(x)) + x^3 \quad (4.5)$$

kann aufgelöst werden zu $\frac{dy}{dx} = \sqrt[3]{\sin(y(x)) + x^3}$. Die rechte Seite ist bekannt wenn y und x bekannt sind, $\frac{dy}{dx} = f(y(x), x)$ mit $f(y, x) = \sqrt[3]{\sin(y) + x^3}$. Allgemeine DGL erster Ordnung können also geschrieben werden als

$$\frac{dy}{dx} = f(y(x), x) , \quad (4.6)$$

indem man die DGL nach der Ableitung erster Ordnung auflöst. Im Beispiel des radioaktiven Zerfalls (4.4) ist $f(y, x) = -\lambda y$. Die Lösung dieser DGL lässt sich als eine Kurve $y(x)$ visualisieren, die überall die durch die Funktion $f(y, x)$ vorgegebene Steigung f hat. Eine solche Kurve kann man geometrisch konstruieren, indem man die xy -Ebene zunächst mit kleinen Strichen füllt, die jeweils die Steigung $f(y, x)$ haben. Gegeben eine Anfangsbedingung $y(x) = y_x$ bei einem gegebenen x , können wir uns vorstellen den "Strichen zu folgen". Die entstehende Kurve ist überall tangential zu den kleinen Strichen, hat also überall die Steigung $f(y(x), x)$ und löst damit die DGL.

Info: Diese graphische Darstellung legt auch eine einfache numerische Methode nahe, um DGL mit Hilfe eines Computers zu lösen. Ausgehend von $y(0)$, das durch die Anfangsbedingung festgelegt sei, suchen wir nach Näherungen der Lösung auf einem diskreten Gitter $x = 0, \Delta x, 2\Delta x, \dots$. Dazu machen wir diskrete Schritte entlang des Gitters. Aus $(y(\Delta x) - y(0))/\Delta x \approx dy/dx = f(y, x)$ am ersten Gitterpunkt folgt der Wert $y(\Delta x) \approx y(0) + f(y(0), x)\Delta x$. Dieser Schritt lässt sich nun wiederholen um iterativ $y(2\Delta x), y(3\Delta x), \dots$ zu bestimmen. Die Schrittweite Δx bestimmt die Genauigkeit des Ergebnisses. Mehr dazu in der Computerphysikvorlesung!

4.3.1 Eindeutigkeit

Sei eine DGL $\frac{dy}{dx} = f(y(x), x)$ auf einem Rechteck $(a, b) \times (c, d)$ definiert, und ein endliches $L > 0$ existiert mit

$$|f(y, x) - f(\tilde{y}, x)| \leq (y - \tilde{y})L \quad \forall x, y, \tilde{y} \text{ auf dem Rechteck} \quad (4.7)$$

so sind zwei Lösungen, die bei einem $x_0 \in [a, b]$ denselben Wert haben, im gesamten Intervall $[a, b]$ gleich.

Info: Dieser Eindeutigkeit der Lösung kann recht leicht bewiesen werden: Integration auf beiden Seiten der DGL (4.6) ergibt die Integralgleichung $y(x) - y(x_0) = \int_{x_0}^x dx' f(x', y(x'))$. Nehmen wir an, daß wir zwei Funktionen $y(x)$ und $Y(x)$ mit dieser Eigenschaft gefunden haben. Wir betrachten nun ein x mit $x - x_0 < \frac{1}{2L}$ und bezeichnen den größten Unterschied zwischen $y(x)$ und $Y(x)$ im Intervall $[x, x_0]$ mit S . Dann gilt

$$|Y(x) - y(x)| \leq \int dx' |f(x', Y(x')) - f(x', y(x'))| \leq L \int dx' |Y(x') - y(x')| \leq LS(x - x_0) \leq S/2. \quad (4.8)$$

Durch Auswertung bei dem Maximum von $|Y(x) - y(x)|$ erhalten wir $S \leq S/2$, ein Widerspruch, der sich nur durch $S = 0$ auflösen lässt. Die beiden Funktionen sind also im Intervall $[x_0, x]$ gleich.

Die Existenz der Lösung wird mit dem **Satz von Picard-Lindelöf** bewiesen. Man beginnt mit $y_0(x) = y(x_0)$, d.h. einer konstanten Funktion, die die Anfangsbedingung erfüllt. Dann nutzt man die Iterationsvorschrift $y_{n+1}(x) = y(x_0) + \int_{x_0}^x dx' f(x', y_n(x'))$ und zeigt, dass die Folge von Funktionen $y_n(x)$ konvergiert, der Grenzfall $n \rightarrow \infty$ löst damit die DGL (s. Kerner und von Wahl).

Die sogenannte **Lipschitz-Bedingung** (4.7) ist in der Praxis fast immer erfüllt. Damit legt eine Anfangsbedingung die Lösung einer DGL erster Ordnung eindeutig fest. Von einer Gleichung, die die Natur zu beschreiben soll, erwarten wir diese Eindeutigkeit: Gegeben die Anfangsbedingung und ein Naturgesetz wie das zweite Newtonsche Gesetz, ist der Zustand des Systems für alle weiteren Zeiten eindeutig festgelegt. Gäbe es mehrere Lösungen wäre diese Theorie keine (vollständige) Beschreibung des Systems.

Ein Nebeneffekt ist Arbeitersparnis: Schaffen wir es - egal wie - eine Lösung der DGL zu finden, die die Anfangsbedingung erfüllt, sind wir am Ziel, weitere Lösungen gibt es nicht.

4.3.2 Gekoppelte Differentialgleichungen mehrerer Veränderlicher

Die graphische Konstruktion der Lösung einer DGL und die Eindeutigkeit der Lösung scheint zunächst auf DGL erster Ordnung beschränkt. Dieser Zugang lässt sich allerdings leicht auf DGL höherer Ordnung verallgemeinern. Dazu definieren wir DGL in mehreren Veränderlichen $y_1(x), y_2(x), \dots$

$$\begin{aligned} \frac{dy_1}{dx} &= f_1(y_1, y_2, y_3, x) && \text{gekoppelte DGL erster Ordnung} \\ \frac{dy_2}{dx} &= f_2(y_1, y_2, y_3, x) \\ \frac{dy_3}{dx} &= f_3(y_1, y_2, y_3, x) \end{aligned}$$

In Vektorschreibweise können wir dieses Gleichungssystem kompakt als $\frac{d\mathbf{y}}{dx} = \mathbf{f}(\mathbf{y}, x)$ schreiben. Man bezeichnet diese Differentialgleichungen als **gekoppelt**, wenn sie sich nicht als mehrere DGL in nur jeweils einer einzelnen Variablen schreiben lassen, d.h. wenn in f_1 ausser y_1 noch z.B. y_2 etc. auftritt. Lösungen solcher Gleichungssysteme sind wieder eindeutig, wenn eine Verallgemeinerung der Lipschitz-Bedingung (4.7) erfüllt ist.

Differentialgleichungen höherer Ordnung lassen sich als gekoppelte DGL erster Ordnung schreiben.

Beispiel

Wir betrachten das zweite Newtonsche Gesetz, eine gewöhnliche DGL zweiter Ordnung. Indem wir eine neue Variable $v(t) = \frac{dy}{dt}$ einführen erhalten wir zwei gekoppelte DGL erster Ordnung

$$\begin{aligned} m \frac{dv}{dt} &= F(y, t) && \text{zwei gekoppelte DGL} \\ \frac{dy}{dt} &= v(t) . \end{aligned}$$

Wir werden uns in Abschnitt 4.4.5 noch näher mit gekoppelten DGL beschäftigen.

4.3.3 Separation der Variablen

Eine einfache Lösungsmethode existiert für einen Spezialfall von DGL erster Ordnung, nämlich sogenannte **separable** Gleichungen

$$\frac{dy}{dx} = h(y)g(x) , \tag{4.9}$$

bei denen die rechte Seite ein Produkt von zwei (bekannten) Funktionen von x und von y ist.

Beispiel

Hier zwei einfache Beispiele. $\frac{dy}{dx} = \sin(x) \cos(y)$ ist separabel, denn mit $h(y) = \cos(y)$ und $g(x) = \sin(x)$ lässt sich diese DGL in die Form (4.9) bringen. Dagegen ist $\frac{dy}{dx} = \sin(x) + \cos(y)$ nicht separabel.

Die gesuchte Lösung ist eine Funktion $y(x)$, die (4.9) erfüllt. Mit dieser (noch unbekannt) Funktion $y(x)$ können beide Seiten als Funktion von x verstanden werden. Umformung und Integration über ein Intervall $[x_b, x_e]$ ergibt dann

$$\int_{x_b}^{x_e} dx g(x) = \int_{x_b}^{x_e} dx \frac{dy}{dx} h^{-1}(y(x)) = \int_{y(x_b)}^{y(x_e)} dy h^{-1}(y). \quad (4.10)$$

Im letzten Schritt wurde durch Variablensubstitution integriert. Für gegebenes x_b ergibt die Auswertung des Integrals links eine Funktion von x_e , rechts eine Funktion von y_e , die nach $y_e = y(x_e)$ aufgelöst die Lösung der DGL (4.9) ergibt.

Beispiel

Wir wenden die Methode der Separation der Variablen auf die Differentialgleichung (4.4) an, die den radioaktiven Zerfall beschreibt. Anfangsbedingung sei $y(t = 0) = y_0$. Mit $f(y, t) = -\lambda y = h(y)g(t)$ schreiben wir $h(y) = y$ und $g(t) = -\lambda$. Aus (4.10) erhalten wir mit $t_b = 0$

$$\begin{aligned} \int_{t_b}^{t_e} dt g(t) &= -\lambda \int_0^{t_e} dt = -\lambda [t]_0^{t_e} = -\lambda t_e & (4.11) \\ &= \int_{y(t_b)}^{y(t_e)} dy h^{-1}(y) = \int_{y_0}^{y(t_e)} dy \frac{1}{y} = [\ln(y)]_{y_0}^{y(t_e)} = \ln(y(t_e)) - \ln(y_0) = \ln\left(\frac{y(t_e)}{y_0}\right). \end{aligned}$$

Auflösen nach $y(t_e)$ ergibt das erwartete Ergebnis $y(t) = y_0 e^{-\lambda t}$.

4.3.4 Gewöhnliche lineare DGL erster Ordnung

Ein weiterer wichtiger Spezialfall sind lineare DGL. Eine gewöhnliche lineare DGL erster Ordnung ist von der Form

$$\frac{dy}{dx} + a(x) y(x) = f(x). \quad (4.12)$$

Die Terme $y(x)$ und $\frac{dy}{dx}$ treten in erster Potenz auf, die Gleichung ist also linear in $y(x)$ und seiner Ableitung. Der Term $f(x)$ auf der rechten Seite wird als **Inhomogenität** der DGL bezeichnet. Ist die rechte Seite gleich null spricht man von einer **homogenen** DGL.

Lösung der homogenen Gleichung

Wir suchen die Lösung von

$$\frac{dy}{dx} + a(x)y(x) = 0. \quad (4.13)$$

In Analogie mit der DGL $\frac{dy}{dx} + \lambda y(x) = 0$, die durch die Exponentialfunktion $y(x) = y(0)e^{-\lambda x}$ gelöst wird, versuchen wir zuerst den Ansatz $y(x) = y(0)e^{-a(x)x}$. Mit

$$\frac{dy}{dx} = -y(0) \frac{d}{dx} (a(x)x) e^{-a(x)x} \quad (4.14)$$

ergibt er allerdings keine Lösung der DGL. Der zweite Versuch mit $y(x) = y(0) \exp \left\{ - \int_0^x dx' a(x') \right\} \equiv y(0) e^{-A(x)}$ ist allerdings erfolgreich

$$\begin{aligned} y(x) &= y(0) \exp \left\{ - \int_0^x dx' a(x') \right\} \equiv y(0) e^{-A(x)} \\ \frac{dy}{dx} &= y(0) \frac{d}{dx} \left(- \int_0^x dx' a(x') \right) e^{-A(x)} \\ &= -y(0) a(x) e^{-A(x)} = -a(x)y(x). \end{aligned} \quad (4.15)$$

Lösung der inhomogenen Gleichung

Als Ansatz zur Lösung der inhomogenen DGL

$$\frac{dy}{dx} + a(x)y(x) = f(x) \quad (4.16)$$

wählen wir $y(x) = c(x)e^{-A(x)}$ mit einer noch zu bestimmenden Funktion $c(x)$. Bei der Ableitung dieses Ansatzes nach x erhalten wir nach der Produktregel 2 Terme: $\frac{dy}{dx} = -a(x)c(x)e^{-A(x)} + c'(x)e^{-A(x)}$. Der erste Term und $a(x)y(x)$ heben sich gegenseitig weg. Der zweite Term muss dann $c'(x)e^{-A(x)}$ gleich der rechten Seite von (4.16) sein:

$$\begin{aligned} \curvearrowright \quad \frac{dy}{dx} + a(x)y(x) &= c'(x)e^{-A(x)} \stackrel{\text{DGL}}{=} f(x) \\ \curvearrowright \quad c'(x) &= e^{A(x)} f(x) \\ c(x) &= \int_0^x dx' e^{A(x')} f(x') \\ y(x) &= c(x)e^{-A(x)} = \int_0^x dx' e^{-A(x)+A(x')} f(x') \\ &= \int_0^x dx' G(x, x') f(x') \quad \text{mit } G(x, x') = e^{-A(x)+A(x')} \end{aligned} \quad (4.17)$$

$G(x, x')$ heißt die **Greensche Funktion** dieser DGL. Die Werte von $f(x')$ im Intervall $0 \leq x' \leq x$ bestimmen also $y(x)$ bei einem gegebenen Wert von x . Die Greenfunktion $G(x, x')$ gibt an, wie $y(x)$ von $f(x')$ abhängt. Ist x die Zeit t , dann beschreibt die Greensche Funktion $G(t, t')$ den Einfluss von $f(t')$ auf die Lösung zu einem späteren Zeitpunkt t .

Bei dieser Lösung fällt auf, daß es keinen freien Parameter zur Wahl einer Anfangsbedingung gibt, bei $x = 0$ ist $y = 0$. Addieren wir allerdings zur Lösung der inhomogenen DGL die Lösung der homogenen Gleichung, erhalten wir weitere Lösung der inhomogenen DGL: Ist $y_1(x)$ eine Lösung der inhomogenen DGL und $y_2(x)$ eine Lösung der homogenen DGL, dann ist auch $y(x) = y_1(x) + by_2(x)$ eine Lösung, da

$$\begin{aligned} \frac{dy}{dx} + a(x)y(x) &= \frac{d}{dx}(y_1(x) + by_2(x)) + a(x)(y_1(x) + by_2(x)) \\ &= \left[\frac{dy_1}{dx} + a(x)y_1(x)\right] + b\left[\frac{dy_2}{dx} + a(x)y_2(x)\right] = f(x). \end{aligned} \quad (4.18)$$

Im letzten Schritt haben wir ausgenutzt, dass y_1 die inhomogene Gleichung $\frac{dy_1}{dx} + a(x)y_1(x) = f(x)$ löst und y_2 die die homogene Gleichung $\frac{dy_2}{dx} + a(x)y_2(x) = 0$. Die Lösung mit der Anfangsbedingung $y(x=0) = b$ ist dann

$$y(x) = \int_0^x dx' e^{-A(x)+A(x')} f(x') + b e^{-A(x)}. \quad (4.19)$$

Der zweite Term in diesem Ausdruck wird als **allgemeine Lösung** bezeichnet (da er nicht von der speziellen Form der Inhomogenität abhängt), der erste Term als **spezielle Lösung**. Bei $x = 0$ ist der erste Term gleich null und die Exponentialfunktion im zweiten Term gleich eins, $b = y(x=0)$ ist damit gleich der frei gewählten Anfangsbedingung. Analog, wenn die Anfangsbedingung nicht bei $x = 0$, sondern bei einem anderen Wert von $x = a$ vorliegt, wählen wir als untere Integralgrenzen den entsprechenden Wert von x und den Vorfaktor des zweiten Term als $y(a)$.

Diese Methode zur Lösung von Differentialgleichungen wird als **Variation der Konstanten** bezeichnet². Den Vorfaktor $c(x)$ der Exponentialfunktion, der in der homogenen Lösung eine Konstante ist, lässt man nun variieren, um eine Lösung einer inhomogenen Gleichung zu finden.

4.4 Gewöhnliche lineare DGL

Wir betrachten die gewöhnliche DGL n -ter Ordnung für $y(x)$ mit einer reellen Veränderlichen x

$$a_n(x) \frac{d^n y}{dx^n} + a_{n-1}(x) \frac{d^{n-1} y}{dx^{n-1}} + \dots + a_1(x) \frac{dy}{dx} + a_0(x) y(x) = f(x). \quad (4.20)$$

²Leider ein reichlich widersinniger Ausdruck.

Diese DGL heißt **lineare DGL**, da $y(x)$ und seine Ableitungen in jedem Term in erster Potenz auftreten. Viele fundamentale Gleichungen der Physik sind lineare DGL (Maxwell-Gleichungen der Elektrodynamik, Schrödingergleichung der Quantenmechanik). Die rechte Seite $f(x)$ wird wieder als **Inhomogenität** der DGL bezeichnet. Die Gleichung lässt sich in der Form

$$\left[a_n(x) \frac{d^n}{dx^n} + a_{n-1}(x) \frac{d^{n-1}}{dx^{n-1}} + \dots + a_1(x) \frac{d}{dx} + a_0(x) \right] y(x) = f(x) \quad (4.21)$$

schreiben, wobei wir den Term in eckigen Klammern als eine Abbildung

$\mathcal{L} \equiv \left[a_n(x) \frac{d^n}{dx^n} + a_{n-1}(x) \frac{d^{n-1}}{dx^{n-1}} + \dots + a_1(x) \frac{d}{dx} + a_0(x) \right]$ (einen Differentialoperator) von Funktionen auf Funktionen auffassen können, mit

$$y(x) \xrightarrow{\mathcal{L}} a_n(x) \frac{d^n y}{dx^n} + a_{n-1}(x) \frac{d^{n-1} y}{dx^{n-1}} + \dots + a_1(x) \frac{dy}{dx} + a_0(x) y(x). \quad (4.22)$$

Der Differentialoperator \mathcal{L} einer linearen DGL ist eine lineare Abbildung (im Raum der Funktionen), $\mathcal{L}(a y_1(x) + b y_2(x)) = a \mathcal{L} y_1(x) + b \mathcal{L} y_2(x)$.

Dieser Zugang erlaubt die kompakte Schreibweise der DGL $\mathcal{L} y = f$. Die Lösung y folgt dann aus der Inversen des Operators \mathcal{L} : Gesucht ist $y(t)$, so daß der Differentialoperator angewandt auf $y(t)$ die Funktion $f(t)$ ergibt.

Aus der Linearität von \mathcal{L} ergibt sich auch eine wichtige mathematische Eigenschaft, die alle homogenen ($f(t) = 0$) DGL teilen. Seien $y_1(x)$ und $y_2(x)$ Lösungen einer linearen, homogenen DGL, d.h. $\mathcal{L} y_1 = 0$, $\mathcal{L} y_2 = 0$. Dann ist auch die Linearkombination $a y_1 + b y_2$ Lösungen der DGL, denn

$$\mathcal{L}(a y_1 + b y_2) = a \mathcal{L} y_1 + b \mathcal{L} y_2 = 0. \quad (4.23)$$

Ebenso ist λy_1 eine Lösung. Die Lösungen einer linearen homogenen DGL bilden also einen Vektorraum, man kann sie addieren und mit Zahlen multiplizieren und erhält wieder eine Lösung der DGL.

Info: Dieses Ergebnis folgt aus der Linearität von \mathcal{L} : die Lösungen linearer homogener Gleichungen bilden einen Vektorraum. Diese Eigenschaft ist also nicht auf Differentialgleichungen beschränkt.

Beispiel

Als Beispiel betrachten wir die lineare DGL $\frac{d^2 y}{dt^2} + \omega_0^2 y(t) = 0$, $\mathcal{L} = \frac{d^2}{dt^2} + \omega_0^2$ ist der dazugehörige (lineare) Differentialoperator, mit ihm kann die Gleichung als $\mathcal{L} y = 0$ geschrieben werden. Diese homogene lineare gewöhnliche DGL beschreibt zum Beispiel eine Punktmasse an einer linearen Feder, $m \frac{d^2 y}{dt^2} = -k y$. Sowohl $y(t) = \cos(\omega_0 t)$ als auch $y(t) = \sin(\omega_0 t)$ lösen diese DGL. Auch $y(t) = a \cos(\omega_0 t) + b \sin(\omega_0 t)$ ist eine Lösung für beliebige a, b , denn $\frac{d^2 y}{dt^2} = -a \omega_0^2 \cos(\omega_0 t) - b \omega_0^2 \sin(\omega_0 t) = -\omega_0^2 y(t)$. Diese Linearkombination von Sinus- und Kosinusfunktionen gleicher Kreisfrequenz wiederum

eine oszillierende Funktion, da

$$A \cos(\omega_0 t - \phi) = A \cos(\omega_0 t) \cos(-\phi) - A \sin(\omega_0 t) \sin(-\phi) = a \cos(\omega_0 t) + b \sin(\omega_0 t) \quad (4.24)$$

mit $A \cos(\phi) \equiv a$ und $A \sin(\phi) \equiv b$ und damit $A = \sqrt{a^2 + b^2}$ und $\tan(\phi) = b/a$ (s. Additionstheorem (3.6)).

Für inhomogene Gleichungen gilt etwas ähnliches: Sei $y_1(x)$ eine Lösung einer inhomogenen Gleichung $\mathcal{L} y_1 = f$ und $y_2(x)$ eine Lösung der dazugehörigen *homogenen* Gleichung $\mathcal{L} y_2 = 0$, dann ist $y_1(x) + a y_2(x)$ eine weitere Lösung der inhomogenen Gleichung:

$$\mathcal{L}(y_1 + a y_2) = \mathcal{L} y_1 + a \mathcal{L} y_2 = f . \quad (4.25)$$

Beispiel

Diese Eigenschaft linearer DGL haben wir bereits bei den Beispielen (4.1) und (4.2) kennengelernt. Die Lösung der inhomogenen Gleichung $\frac{d^2 y}{dt^2} = g$ ist $y(t) = gt^2/2$, Lösungen der dazugehörigen homogenen $\frac{d^2 y}{dt^2} = 0$ Gleichung sind $y(t) = 1$ und $y(t) = t$, ihre Linearkombination $at + b$ löst ebenso die homogene Gleichung. Erst die Summe der Lösung der inhomogenen Gleichung und die der homogene Gleichung, $y(t) = gt^2/2 + at + b$, ist dann die vollständige Lösung der DGL. Die Koeffizienten a und b werden aus der Anfangsbedingung (Ort und Geschwindigkeit) bestimmt.

Komplexe Lösungen linearer DGL Für lineare DGL sind Realteil und Imaginärteil einer komplexen Lösung wiederum Lösungen der DGL. Diese Eigenschaft wird ihren enormen Wert zeigen, wenn wir in Abschnitt 4.4.1 die lineare Bewegungsgleichung des harmonischen Oszillators lösen.

Sei \mathcal{L} ein reeller Operator (die Koeffizienten aller Ableitungen sind reellwertig), $z(x) = z_R(x) + i z_I(x)$ aber eine komplexe Lösung der (homogenen) DGL. Aus $\mathcal{L} z = 0 = 0 + i 0$ und der Linearität des Differentialoperators \mathcal{L} folgt durch Vergleich von Real- und Imaginärteil

$$\begin{aligned} 0 &= \mathcal{L}(z_R + i z_I) = \mathcal{L} z_R + i \mathcal{L} z_I \\ \Leftrightarrow &\mathcal{L} z_R = 0, \mathcal{L} z_I = 0 . \end{aligned} \quad (4.26)$$

Beispiel

Als Beispiel betrachten wir wieder die lineare DGL $\frac{d^2 y}{dt^2} + \omega_0^2 y(t) = 0$. Sie wird gelöst durch die komplexe Funktion $z(t) = e^{i\omega_0 t}$, da $\frac{d^2 z}{dt^2} = -\omega_0^2 e^{i\omega_0 t} = -\omega_0^2 z(t)$. Real- und Imaginärteil von $z(t) = \cos(\omega_0 t) + i \sin(\omega_0 t)$ sind ebenso Lösungen der DGL.

4.4.1 Gewöhnliche lineare DGL zweiter Ordnung mit konstanten Koeffizienten

Gewöhnliche lineare DGL zweiter Ordnung können in die Form

$$\frac{d^2y}{dx^2} + a_1(x) \frac{dy}{dx} + a_0(x) y(x) = f(x) \quad (4.27)$$

gebracht werden (indem wir beide Seiten durch einen möglichen Koeffizienten $a_2(x)$ teilen). Wir beschränken uns hier auf konstante Koeffizienten (obwohl es, wie bei der DGL erster Ordnung, wieder Lösungen zu allgemeinen Koeffizienten gibt, die Greensche Funktion für diesen Fall ist aber komplizierter). DGL mit konstanten Koeffizienten sind sowohl von den physikalischen Anwendungen, als auch vom Lösungsansatz her interessant.

Beispiel

Eine Masse m an einer Feder mit Federstärke k bewege sich in einer Dimension durch eine viskose Flüssigkeit (Reibungskraft proportional zur Geschwindigkeit) und wird von einer externen Kraft $F_{\text{ext}}(t)$ angetrieben. Aus der Newtonschen Bewegungsgleichung $m \frac{d^2x}{dt^2} = -kx - \gamma \frac{dx}{dt} + F_{\text{ext}}(t)$ erhalten wir mit $b = \gamma/m$, $c = k/m$, $f(t) = F_{\text{ext}}/m$

$$\frac{d^2x}{dt^2} + b \frac{dx}{dt} + cx = f(t) . \quad (4.28)$$

4.4.2 Gedämpfter harmonischer Oszillator (homogene Gleichung)

Wir beginnen mit der Analyse des Falles ohne externe Antriebskraft, $f(t) = 0$.

$$\frac{d^2x}{dt^2} + b \frac{dx}{dt} + cx(t) = 0 . \quad (4.29)$$

Wir erwarten generell als Lösung eine sinusförmige Schwingung mit exponentiell abfallender Amplitude und wählen als Ansatz $x(t) = Ae^{-\xi t} \cos(\omega t - \phi)$. Einsetzen dieses Ansatzes in die DGL (4.29) führt allerdings auf eine aufwändige trigonometrische Rechnung (ausprobieren!).

Wir erinnern uns: bei linearen DGL sind Realteile und Imaginärteile von komplexen Lösungen ebenso Lösungen der DGL. **Der komplexer Ansatz $x(t) = Ae^{i\omega t}$ $A, \omega \in \mathbb{C}$ beschreibt sowohl die Oszillation, als auch den exponentiellen Abfall von x mit t .** Mit der Zerlegung eines komplexen ω in Realteil und Imaginärteil $\omega = \omega_R + i\omega_I$ ($\omega_R, \omega_I \in \mathbb{R}$) ist nämlich

$$e^{i\omega t} = e^{i\omega_R t} e^{-\omega_I t} = (\cos(\omega_R t) + i \sin(\omega_R t)) e^{-\omega_I t} , \quad (4.30)$$

der Realteil von $e^{i\omega t}$ ist also eine exponentiell abfallende Kosinusfunktion. Auch eine Phasenverschiebung dieser Funktion lässt sich mit einer komplexen Amplitude $A = \tilde{A}e^{-i\phi}$

($\tilde{A}, \phi \in \mathbb{R}$) beschreiben, $Ae^{i\omega_R t} = \tilde{A}e^{-i\phi+i\omega_R t} = \tilde{A}[\cos(\omega_R t - \phi) + i \sin(\omega_R t - \phi)]$ mit Realteil $\tilde{A} \cos(\omega_R t - \phi)$.

Der komplexe Exponentialansatz ist die wichtigste Anwendung der Eulerformel (3.5) in der Physik. Die Ableitungen von $Ae^{i\omega t}$ nach t sind jetzt nämlich einfach

$$\begin{aligned}x(t) &= Ae^{i\omega t} \\ \frac{dx}{dt} &= i\omega Ae^{i\omega t} \\ \frac{d^2x}{dt^2} &= -\omega^2 Ae^{i\omega t} .\end{aligned}$$

Für die DGL (4.29) ergibt der Ansatz $x(t) = Ae^{i\omega t}$ dann

$$\begin{aligned}(-\omega^2 + bi\omega + c) Ae^{i\omega t} &= 0 & \forall t \\ A = 0 & \text{ (triviale Lösung)} \\ -\omega^2 + ib\omega + c &= 0\end{aligned}$$

Die erste Lösung $A = 0$ ist uninteressant und entspricht einem ruhenden Pendel. Die zweite Lösung führt auf eine quadratische Gleichung in ω mit den beiden Lösungen

$$\omega_{+/-} = \frac{-ib \pm \sqrt{-b^2 + 4c}}{-2} = \frac{i}{2} b \pm \frac{1}{2} \sqrt{4c - b^2} . \quad (4.31)$$

Mithilfe des Exponentialansatzes $Ae^{i\omega t}$ ist es also gelungen, die lineare Differentialgleichung (4.29) auf eine algebraische Gleichung zu reduzieren. Die Lösung der homogenen DGL (4.29) ist der Realteil von $A_+ e^{i\omega_+ t} + A_- e^{i\omega_- t}$. Dieser Ansatz löst lineare DGL mit konstanten Koeffizienten, unabhängig von der Ordnung der DGL.

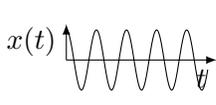
Wir betrachten nun die Lösung des harmonischen Oszillators in unterschiedlichen Regimes. Für den Spezialfall ohne Reibung $b = 0$ erhalten wir die beiden Lösungen

$$\omega_{+/-} = \pm\omega_0 \equiv \pm\sqrt{c} = \pm\sqrt{k/m},$$

und somit $z(t) = A_+ e^{i\omega_0 t} + A_- e^{-i\omega_0 t}$ ($A_+, A_- \in \mathbb{C}$) mit einem Realteil der Form $x(t) = \tilde{A} \cos(\omega_0 t - \phi)$. Diese Lösung hat zwei freie Parameter, \tilde{A} und ϕ , die noch aus den Anfangsbedingungen (Position und Geschwindigkeit) zu bestimmen sind. Die Auslenkung des Oszillators zum Zeitpunkt $t = 0$ sei x_0 , seine Geschwindigkeit sei v_0 . Setzen wir in der Lösung $x(t)$ und ihrer Ableitung nach der Zeit diese Anfangsbedingungen ein, erhalten wir

$$\tilde{A} \cos(-\phi) = x_0 \quad (4.32)$$

$$-\omega_0 \tilde{A} \sin(-\phi) = v_0 . \quad (4.33)$$



Teilen wir die zweite Gleichung durch ω_0 und nutzen $\cos^2(-\phi) + \sin^2(-\phi) = 1$ erhalten wir $\tilde{A} = \sqrt{x_0^2 + v_0^2/\omega_0^2}$. Teilen wir die zweite Gleichung durch die erste, erhalten wir $\tan(-\phi) = v_0/(x_0\omega_0)$, und damit die Phasenverschiebung ϕ .

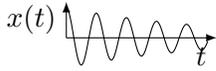
Wir führen nun eine Fallentscheidung für steigendes b (steigende Dämpfung) durch

1. **Gedämpfte harmonische Schwingung** $4c - b^2 > 0$

$$\omega_{+/-} = \pm\omega_R + i\omega_I \quad \omega_R = \frac{1}{2}\sqrt{4c - b^2} \quad , \quad \omega_I = b/2$$

$$z(t) = A_+e^{i\omega_+t} + A_-e^{i\omega_-t}$$

mit Realteil $Ae^{-\omega_I t} \cos(\omega_R t - \phi)$,



die Kreisfrequenz $\omega_R = \frac{1}{2}\sqrt{4c - b^2}$ nimmt mit zunehmender Dämpfung b ab, die Abklingrate $\omega_I = b/2$ wächst mit b .

2. **Kritische Dämpfung** $4c - b^2 = 0$

$$\omega_{\pm} = \frac{i}{2}b$$

ergibt $x(t) = A_{\pm}e^{i\omega_{\pm}t}$



Da ω rein imaginär ist fällt $x_R(t)$ ohne Oszillation exponentiell ab. Überraschend ist, dass der Exponentialansatz nur eine Lösung findet. Wir benötigen aber 2 freie Parameter, um alle Anfangsbedingungen umsetzen zu können. Die zweite Lösung findet sich durch Betrachtung des Grenzfalles $b \rightarrow b_c \equiv 2\sqrt{c}$. Wir betrachten eine Linearkombination der beiden Lösungen $e^{i\omega_+t}$ und $e^{i\omega_-t}$ aus (4.31) und führen den Grenzfalle $\omega_+ \rightarrow \omega_-$ aus

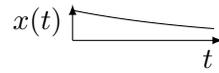
$$\lim_{\omega_+ \rightarrow \omega_-} \frac{e^{i\omega_+t} - e^{i\omega_-t}}{i\omega_+ - i\omega_-} = \lim_{\omega_+ \rightarrow \omega_-} e^{i\omega_+t} \frac{1 - e^{i(\omega_- - \omega_+)t}}{i\omega_+ - i\omega_-}$$

$$= \lim_{\omega_+ \rightarrow \omega_-} e^{i\omega_+t} \frac{1 - (1 + (i\omega_- - i\omega_+)t)}{i\omega_+ - i\omega_-} = te^{i\omega_-t}$$

und erhalten eine weitere Lösung, die sich nicht einfach aus dem Exponentialansatz ergibt. Die allgemeine Lösung ist dann die Linearkombination der beiden Lösungen

$$x(t) = A e^{i\omega t} + B t e^{i\omega t} \quad i\omega = -b/2$$

mit zwei freien Parametern (Realteile von A, B).



3. Überdämpfter Oszillator $4c - b^2 < 0$

$$\omega_{+/-} = \frac{i}{2}b \pm \frac{i}{2}\sqrt{|4c - b^2|} = \frac{i}{2} \left(b \pm \sqrt{b^2 - 4c} \right)$$

Die allgemeine Lösung ist dann gegeben durch

$$x(t) = A_+ e^{i\omega_+ t} + A_- e^{i\omega_- t} = A_+ e^{-\omega_+ t} + A_- e^{-\omega_- t} .$$

Wegen $\omega_+ > \omega_-$ entsprechen die beiden Terme einem (relativ) schnellen und einem langsamen Abklingprozess, beide ohne Oszillation. Für lange Zeiten ist diese Lösung durch den zweiten (langsamen) Term dominiert.

Zum Schluss vergleichen wir noch das Abklingverhalten unter überkritischer Dämpfung und kritischer Dämpfung. Wegen

$$\omega_{-I} = \frac{b - \sqrt{b^2 - 4c}}{2} < \frac{b}{2} = \omega_{cI} \quad (\text{kritisches Abklingen})$$

führt kritische Dämpfung mit $b = b_c$ zum schnellsten Abklingen, bei der nach einer Auslenkung des Oszillators $x(t)$ so schnell wie möglich zur Ruhelage $x = 0$ zurückkehrt.

4.4.3 Gedämpfter harmonischer Oszillator (inhomogene Gleichung)

$$\mathcal{L}x = f \quad \mathcal{L} = \frac{d^2}{dt^2} + b\frac{d}{dt} + \omega_0^2$$

Physikalisch beschreibt die Inhomogenität $f(t)$ eine externe Kraft, die den Oszillator antreibt. Wir beschränken uns auf zwei Spezialfälle. Wie wir im Zusammenhang mit der Fourieranalyse sehen werden, ist deren Lösung bereits ausreichend, um die Lösung des harmonischen Oszillators unter einer beliebigen periodischen Antriebskraft zu berechnen. Der erste Spezialfall ist eine konstante externe Kraft $f(t) = f$. Als Ansatz für die spezielle Lösung wählen wir $x(t) = A$, einsetzen in die DGL ergibt $A = f/c$.

4.4.4 Sinusförmige externe Kraft

Der zweite Spezialfall ist eine sinusförmige externe Kraft. Dieser Fall ist wichtig für eine Vielzahl von Anwendungen (regelmäßiges Anschubsen einer Schaukel, regelmäßige Windstöße gegen eine Brücke). Im nächsten Kapitel werden wir sehen, wie sich auch nicht-sinusförmige Kräfte aus Sinusschwingungen unterschiedlicher Frequenz zusammensetzen

lassen. Die Kreisfrequenz Ω und Antriebsamplitude F sind vorgegeben, sie hängen auch nicht von den Eigenschaften des Oszillators ab. Gesucht ist also die Lösung der DGL

$$\frac{d^2}{dt^2} x + b \frac{dx}{dt} + \omega_0^2 x(t) = F \cos \Omega t \quad (4.34)$$

Wir nutzen wieder die Linearität der DGL um eine komplexe Lösung zu finden. Hierbei lassen wir auch eine komplexe Inhomogenität zu und schreiben die Antriebskraft als $F e^{i\Omega t}$ (mit reellem Ω und F). Denn für einen reellen linearen Differentialoperator \mathcal{L} und die DGL $\mathcal{L}z = f_z = f_R + i f_I$ ist der Realteil einer Lösung von $\mathcal{L}z = f_z$ eine Lösung von $\mathcal{L}y = f_R$

$$f_R + i f_I = \mathcal{L}(z_R + i z_I) = \mathcal{L}z_R + i \mathcal{L}z_I \quad \leadsto \mathcal{L}z_R = f_R . \quad (4.35)$$

Wir suchen also die Lösung von $\frac{d^2}{dt^2} x + b \frac{dx}{dt} + c x(t) = F e^{i\Omega t}$ und nutzen wieder einen komplexen Ansatz $z(t) = A e^{i\omega t}$. Wir erhalten die Gleichung

$$(-\omega^2 + i b \omega + \omega_0^2) A e^{i\omega t} = F e^{i\Omega t} , \quad (4.36)$$

die für alle Zeiten t erfüllt sein muss. Dazu muss $\omega = \Omega$ gelten, d.h. der Oszillator folgt der Kreisfrequenz der externen Kraft. Für die (komplexe) Amplitude A erhalten wir

$$A = |A| e^{-i\phi} = \frac{F}{\omega_0^2 - \Omega^2 + i\Omega b} \quad (4.37)$$

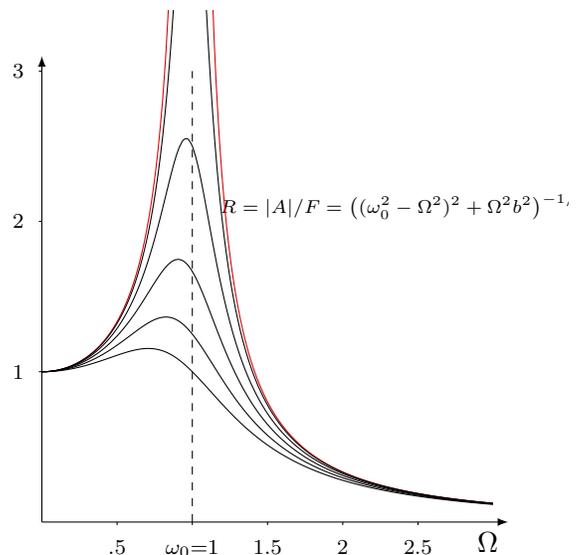
mit Betrag $|A| = F((\omega_0^2 - \Omega^2)^2 + \Omega^2 b^2)^{-1/2}$ und Phase $\phi = \arctan(\frac{\Omega b}{\omega_0^2 - \Omega^2})$.

Der Realteil dieser komplexen Lösung

$$\Re(A e^{i\Omega t}) = \Re(|A| e^{i\Omega t - i\phi}) = |A| \cos(\Omega t - \phi) \quad (4.38)$$

ist die Lösung der inhomogenen Gleichung.

$R = |A|/F$ gibt das Verhältnis der Antriebsamplitude F und der Amplitude der inhomogenen Lösung an und hängt ab von der Kreisfrequenz der Antriebskraft Ω , der Kreisfrequenz des ungedämpften harmonischen Oszillators $\omega_0 = \sqrt{k/m}$, und dem Dämpfungsparameter b . Bei kleinen Werten der Dämpfung kann R für $\Omega \approx \omega_0$ sehr groß werden. Dieses Phänomen wird als **Resonanz** bezeichnet. Die Abbildung zeigt R für $\omega_0 = 1$ und $b = 0$ (rot) und $b = 0.2, 0.4, \dots$. Der Effekt lässt sich am einfachsten mit dem Kind auf der Schaukel illustrieren: Aufgabe ist mit kleiner Antriebsamplitude A eine große



Schwingung der Schaukel zu erreichen. Dies erzielt ein externer Schaukelanschubser, wenn er der Schaukel immer einen Schubs gibt, wenn sich die Schaukel gerade ihm weg bewegt. Die Kreisfrequenz der Schaukel muss also gleich der des Antriebs sein, der Antrieb muss in Phase mit der Bewegung der Schaukel sein. Ist die Kreisfrequenz des Antriebs Ω gleich der Kreisfrequenz des gedämpften harmonischen Oszillators $\sqrt{\omega_0^2 - b^2/4}$ erreicht R sein Maximum. (Ohne Dämpfung liegt dieses Maximum bei Unendlich. In einem realistischen System würde der Oszillator allerdings bei großen Auslenkungen Effekte der Nichtlinearität zeigen.)

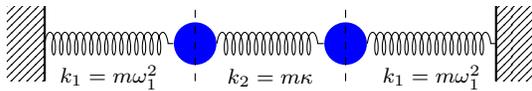
Um die vollständige Lösung zu bestimmen, müssen wir zu dieser Lösung der inhomogenen Gleichung noch die allgemeine Lösung der homogenen Gleichung addieren. Die beiden freien Koeffizienten der allgemeinen Lösung sind dann aus den Anfangsbedingungen zu bestimmen. Für $b \neq 0$ fällt die Amplitude der allgemeinen Lösung allerdings exponentiell ab, für große Zeiten ist $x(t)$ also asymptotisch durch die Lösung der inhomogenen Gleichung gegeben. Die Amplitude von $x(t)$ ist dann konstant $|A|$, über eine ganze Periode ist die Energie die durch die externe Kraft dem Pendel zugeführt wird genau gleich der Energiedissipation durch Reibung.

4.4.5 Gekoppelte lineare DGL und Normalmoden

Wir betrachten als Beispiel ein ungedämpftes gekoppeltes Federpendel, bei dem zwei Punktmassen an linearen Federn noch durch eine weitere Feder gekoppelt sind. Die Auslenkungen der beiden Massen von der Ruheposition sind durch die beiden Variablen x_1 und x_2 beschrieben. Die Bewegungsgleichungen sind

$$\begin{aligned} m \frac{d^2 x_1}{dt^2} &= -k_1 x_1 + k_2 (x_2 - x_1) \\ m \frac{d^2 x_2}{dt^2} &= -k_1 x_2 + k_2 (x_1 - x_2) . \end{aligned} \tag{4.39}$$

Bei der Lösung einer solchen gekoppelten DGL mit mehreren abhängigen Variablen stehen wir vor der Schwierigkeit, dass wir die Lösung einer Variablen kennen müssten um die DGL in der anderen Variablen zu lösen.



In diesem Fall erhalten wir jedoch durch eine einfache Variablentransformation eine DGL ohne Kopplung zwischen den Variablen! Wir betrachten neue Variablen $y_1 = (x_1 + x_2)/\sqrt{2}$

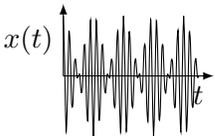
und $y_2 = (x_1 - x_2)/\sqrt{2}$. Durch Addition der Gleichungen (4.39) und durch Subtraktion der Gleichungen erhalten wir

$$\frac{d^2 y_1}{dt^2} = -\omega_1^2 y_1 \quad (4.40)$$

$$\frac{d^2 y_2}{dt^2} = -\omega_2^2 y_2 \quad (4.41)$$

mit $\omega_1^2 = k_1/m$ und $\omega_2^2 = k_1/m + \kappa$. Diese beiden Gleichungen sind entkoppelt und lassen sich unabhängig voneinander lösen. Wir erhalten $y_1 = Ae^{i\omega_1 t}$ und $y_2 = Be^{i\omega_2 t}$, wobei A und B aus den Anfangsbedingungen bestimmt werden müssen (siehe harmonischer Oszillator ohne Reibung). Danach transformieren wir zurück zu den ursprünglichen Variablen $x_1 = (y_1 + y_2)\sqrt{2}$ und $x_2 = (y_1 - y_2)\sqrt{2}$. Wir spielen diese Vorgehensweise für drei unterschiedlichen Anfangsbedingungen durch: 1. $x_1(0) = x_2(0)$ und $\frac{dx_1}{dt}(0) = \frac{dx_2}{dt}(0)$; beide Massen haben zu Beginn die gleiche Auslenkung und Geschwindigkeit. Damit ist $y_2(0) = 0$ und $\frac{dy_2}{dt}(0) = 0$ und $y_2(t)$ für alle Zeiten null. Wir erhalten $x_1(t) = x_2(t) = A/\sqrt{2}e^{i\omega_1 t}$. 2. $x_1(0) = -x_2(0)$ und $\frac{dx_1}{dt}(0) = -\frac{dx_2}{dt}(0)$, beide Massen sind zu Beginn in unterschiedliche Richtungen gleich stark ausgelenkt und haben entgegengesetzte Geschwindigkeiten. Damit ist $y_1(0) = 0$ und $\frac{dy_1}{dt}(0) = 0$, $y_1(t)$ für alle Zeiten null und $x_1(t) = -x_2(t) = B/\sqrt{2}e^{i\omega_2 t}$. 3. $x_1(0) = a, x_2(0) = 0$ und $\frac{dx_1}{dt}(0) = \frac{dx_2}{dt}(0) = 0$. Wir erhalten $y_1(t) = a/\sqrt{2}e^{i\omega_1 t}$ und $y_2(t) = a/\sqrt{2}e^{i\omega_2 t}$ und damit

$$\begin{aligned} x_1(t) &= (y_1(t) + y_2(t))/\sqrt{2} = \frac{a}{2}(e^{i\omega_1 t} + e^{i\omega_2 t}) \\ &= \frac{a}{2}e^{i\frac{\omega_1 + \omega_2}{2}t}(e^{i\frac{\omega_1 - \omega_2}{2}t} + e^{-i\frac{\omega_1 - \omega_2}{2}t}) \\ &= ae^{i\bar{\omega}t} \cos(\Delta\omega t) \end{aligned} \quad (4.42)$$



mit $\bar{\omega} \equiv \frac{\omega_1 + \omega_2}{2}$ und $\Delta\omega \equiv \frac{\omega_1 - \omega_2}{2}$. Die Amplitude der Schwingung $e^{i\bar{\omega}t}$ oszilliert durch den Term $\cos(\Delta\omega t)$ mit Kreisfrequenz $\Delta\omega$ ³

Im ersten Beispiel bewegen sich beide Massen mit Kreisfrequenz ω_1 und mit gleicher Phase, im zweiten Beispiel mit Kreisfrequenz ω_2 und mit einer Phasenverschiebung von π (Gegenphase). Eine solche Dynamik, bei der alle Komponenten eines Systems mit gleicher Frequenz schwingen, mit konstanter Amplitude und relativer Phase, bezeichnet man als **Normalmode** des Systems. Mit den Anfangsbedingungen 1. und 2. wurden also jeweils zwei unterschiedliche Normalmoden angeregt, die Anfangsbedingung 3. regt eine Kombination der beiden Normalmoden an.

³Das akustische Gegenstück zu dieser Bewegung wird als Schwebung bezeichnet, gut beobachten kann man es bei zwei Tönen sehr ähnlicher Frequenz, wie sie beim Stimmen von Instrumenten auftreten. Ein Beispiel ist hier <https://www.youtube.com/watch?v=5hxQDAmNWE>.

Später werden wir einfache Techniken der linearen Algebra nutzen, um die Normalmoden von allgemeinen System zu finden.

4.5 Partielle DGL (PDG)

DGL mit Ableitungen nach unterschiedlichen Variablen (partielle Ableitungen) werden als **partielle DGL (PDG)** bezeichnet. Die Klassifizierung nach Ordnung der DGL, linear/nichtlinear ist analog zu gewöhnlichen DGL. PDG treten bei der räumlichen oder raumzeitlichen Beschreibung von Feldern auf.

Diffusionsgleichung

$$\frac{\partial \rho}{\partial t} = D \frac{\partial^2 \rho}{\partial x^2} \quad \text{für die Teilchendichte } \rho(x, t)$$

(s. auch Skizze unten)

Laplacegleichung der Elektrostatik

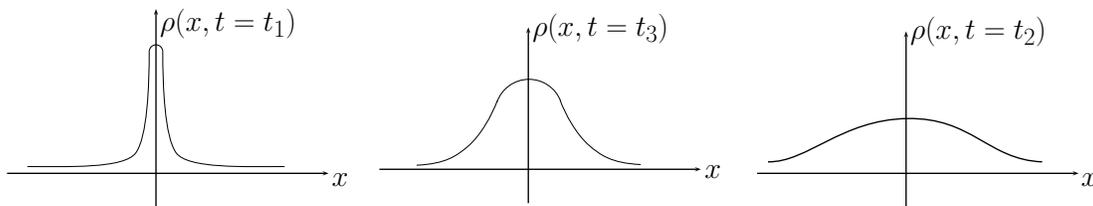
$$\frac{\partial^2 \phi}{\partial x^2} + \frac{\partial^2 \phi}{\partial y^2} + \frac{\partial^2 \phi}{\partial z^2} = 0 \quad \text{für das elektr. Potential } \phi(x, y, z)$$

Wellengleichung

$$\frac{\partial^2 \phi}{\partial x^2} = \frac{1}{c^2} \frac{\partial^2 \phi}{\partial t^2}$$

Schrödingergleichung der Quantenmechanik

$$i\hbar \frac{\partial}{\partial t} \psi = -\frac{\hbar^2}{2m} \frac{\partial^2}{\partial x^2} \psi(x, t) + V(x) \psi(x, t) \quad \text{für Wellenfunktion } \psi(x, t) \text{ eines Te}$$



PDG sind generell viel schwerer zu lösen als gewöhnliche Differentialgleichungen, außerdem ist die Theorie zur Existenz und Eindeutigkeit von Lösungen sehr viel komplexer als bei gewöhnlichen Differentialgleichungen (auch wenn wir bei PDG die die Natur beschreiben davon ausgehen, dass Lösungen dieser Gleichungen existieren und eindeutig sind).

4.5.1 Randbedingungen

Zusätzlich zu den Anfangsbedingungen schränken bei PDF oft Randbedingungen die Lösung ein. Als Beispiel betrachten wir die Wellengleichung in einer Dimension $\frac{\partial^2 \phi}{\partial x^2} = \frac{1}{c^2} \frac{\partial^2 \phi}{\partial t^2}$. Sie beschreibt zum Beispiel die (kleine) Auslenkung eines Gummibandes, das zwischen zwei Punkten aufgespannt ist. An diesen Punkten, z.B. $x = 0$ und $x = L$ ist dann $\phi(x = 0, t) = 0, \phi(x = L, t) = 0$ für alle Zeiten t . Analog in 2D: Eine über eine Öffnung gespannte Membran (Trommel) mit Auslenkung $\phi(x, y, t)$ gehorcht der Wellengleichung $\frac{\partial^2 \phi}{\partial x^2} + \frac{\partial^2 \phi}{\partial y^2} = \frac{1}{c^2} \frac{\partial^2 \phi}{\partial t^2}$. Dann ist auf der Öffnung (auf der wir die Membran festgezurrut haben) $\phi = 0$ für alle Zeiten t .

Unsere Strategie zur Lösung von linearen homogenen PDG (wie die obigen Beispiele) ist analog zu gewöhnlichen DGL. Wir suchen zunächst unterschiedliche Lösungen (ggf. mit Einschränkungen durch Randbedingungen), und suchen dann die Linearkombination dieser Lösungen, die die Anfangsbedingungen erfüllt.

4.5.2 Separationsansatz

Der Separationsansatz erlaubt eine Vielzahl (aber bei weitem nicht alle) PDG auf gewöhnliche DGL zurückzuführen, die dann verhältnismäßig leicht gelöst werden können. Sämtliche oben aufgeführte PDG lassen sich durch diesen Ansatz lösen. Vor allem in der Quantenmechanik stellt sich der Separationsansatz als sehr mächtig heraus.

Als konkretes Beispiel betrachten wir die Wellengleichung in 1D

$$\frac{\partial^2 \phi}{\partial x^2} = \frac{1}{c^2} \frac{\partial^2 \phi}{\partial t^2} \tag{4.43}$$

mit Randbedingung $\phi(x = 0, t) = \phi(x = L, t) = 0 \forall t$ (z.B. ein Seil, das zwischen zwei Punkten bei $x = 0$ und $x = L$ aufgespannt ist, mit kleiner Auslenkung $\phi(x, t)$ in einer Richtung rechtwinklig zur x -Achse).

Als Ansatz für die Lösung nutzen wir das Produkt einer (noch zu bestimmenden) Funktion von x und einer (ebenso noch zu bestimmenden) Funktion von t ; $\phi(x, t) = f(x) g(t)$. Diesen Ansatz setzen wir in die Wellengleichung ein und teilen auf beiden Seiten durch $f(x) g(t)$

$$\begin{aligned} \frac{\partial^2}{\partial x^2} (f(x) g(t)) &= \frac{1}{c^2} \frac{\partial^2}{\partial t^2} (f(x) g(t)) \\ g(t) \frac{d^2}{dx^2} f(x) &= \frac{1}{c^2} f(x) \frac{d^2}{dt^2} g(t) && \text{(Linearität der PDG)} \\ \underbrace{\frac{1}{f(x)} \frac{d^2}{dx^2} f(x)}_{\text{Funktion von } x} &= \underbrace{\frac{1}{c^2} \frac{1}{g(t)} \frac{d^2}{dt^2} g(t)}_{\text{Funktion von } t} && \text{(durch } \phi(x, t) \text{ geteilt)} \end{aligned}$$

Diese Gleichung muss für alle x, t gelten. Die linke Seite ist eine Funktion von x , die rechte Seite eine Funktion von t . x und t sind unabhängige Variablen. Die einzige Möglichkeit ist, dass beide Seiten konstante Funktionen sind

$$\begin{aligned} \frac{1}{f(x)} \frac{d^2}{dx^2} f(x) &= \alpha & \leadsto & \frac{d^2}{dx^2} f(x) = \alpha f(x) \\ \frac{1}{c^2} \frac{1}{g(t)} \frac{d^2}{dt^2} g(t) &= \alpha & \leadsto & \frac{d^2}{dt^2} g(t) = c^2 \alpha g(t) \end{aligned}$$

Die Konstante α ist noch zu bestimmen. Aus der PDF in zwei Variablen hat der Separationsansatz zwei *gewöhnliche* DGL gemacht, die jetzt noch zu lösen sind:

$$\begin{aligned} \frac{d^2 f}{dx^2} &= \alpha f(x) & f(x) &= A \sin kx & kL &= \pi n & , & n \in \mathbb{N} \\ \frac{d^2 g}{dt^2} &= c^2 \alpha g(t) & g(t) &= B e^{ikct} \\ & & &= B e^{i\omega t} \\ & & \text{mit } \omega &= kc & \leadsto & \alpha &= -k^2 . \end{aligned} \quad (4.44)$$

Jedes $k = \frac{\pi n}{L}$ mit $n \in \mathbb{N}$ löst die DGL mit der obigen Randbedingung und ergibt eine Lösung

$$\phi(x, t) = C_n \sin\left(\frac{\pi x n}{L}\right) e^{i \frac{\pi n}{L} ct} . \quad (4.45)$$

Die Konstante α in (4.44) ist dann durch $\alpha = -k^2$ gegeben. Prinzipiell könnte α auch positiv sein, dann wäre $f(x)$ aber eine Exponentialfunktion $e^{\pm kx}$ und die Randbedingung $f(0) = 0, f(L) = 0$ ließe sich nicht implementieren.

$k = \pi n/L = 2\pi/\lambda$ wird als **Kreiswellenzahl** der Lösung bezeichnet. n gibt also an, wieviele halbe Perioden in das Intervall $[0, L]$ passen. Lösungen zu hohen Werten von n haben eine kürzere Wellenlänge λ . Die Kreisfrequenz $\omega = ck$ hängt ebenso von n ab, $e^{i \frac{\pi n}{L} ct} = e^{i\omega t}$ mit $\omega = \pi cn/L$. Eine Lösung der Form (4.45) zeichnet sich dadurch aus, dass $\phi(x, t)$ für alle Punkte x mit der gleichen Frequenz oszillieren, bei konstanter Amplitude und relativer Phase zueinander. Eine solche Bewegung ist also wieder eine Normalmode des Systems.

Da die Wellengleichung (4.43) eine lineare homogene Gleichung ist, lösen auch Linearkombinationen von Lösungen mit unterschiedlichem n die Wellengleichung, die allgemeine Lösung der Wellengleichung ist also

$$\phi(x, t) = \sum_{n=1}^{\infty} C_n \sin\left(\frac{\pi x n}{L}\right) e^{i \frac{\pi n}{L} ct} \quad (4.46)$$

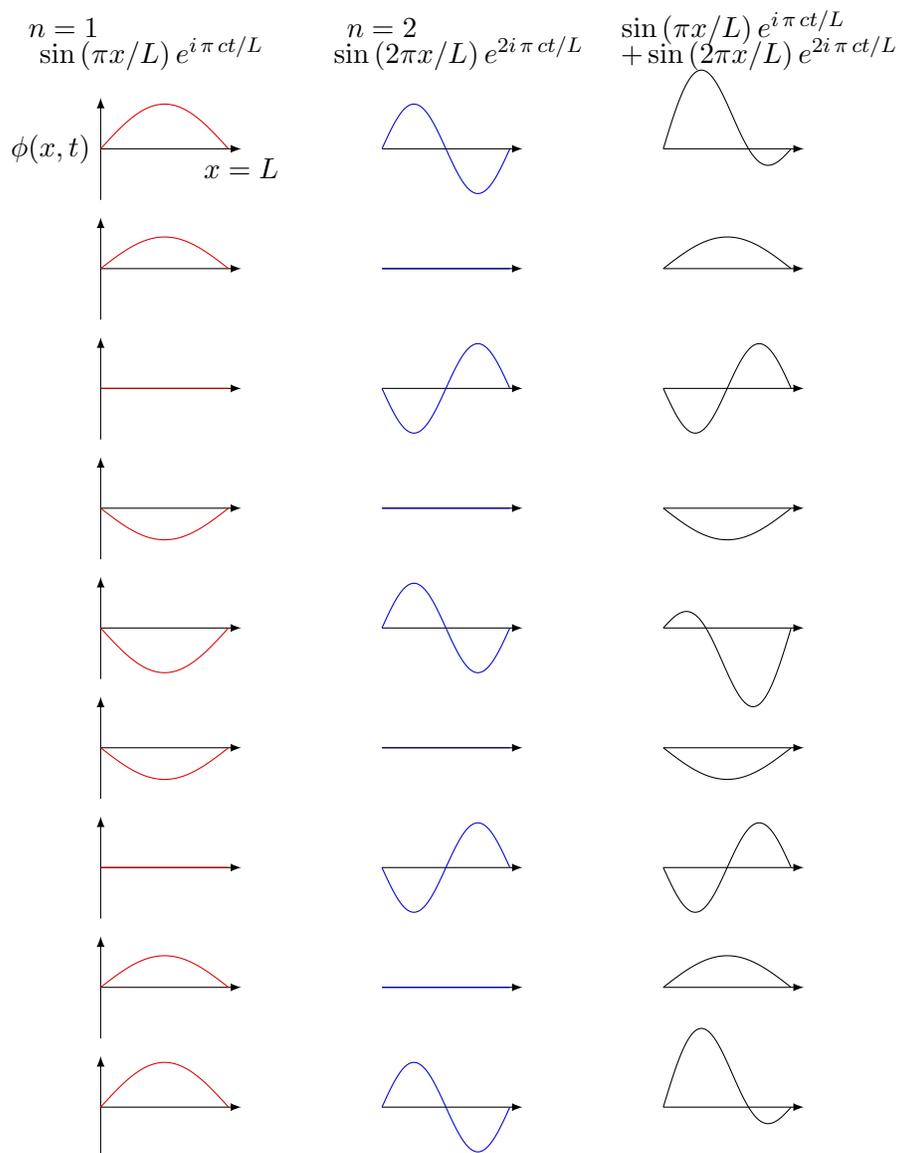
Diese allgemeine Lösung ist nicht mehr von der Form $f(x)g(t)$, unterliegt also nicht der Einschränkung des Separationsansatzes. Die einzelnen Terme in (4.45) sind die Normalmoden der Wellengleichung. Die Koeffizienten $C_n \equiv |C_n|e^{i\phi}$ geben die Amplitude und Phase der Normalmoden an.

Die Abbildung zeigt links eine Periode der ersten Mode (Zeit läuft von oben nach unten), $n = 1$, mit $C_1 = 1$ und $C_n = 0$ sonst. Die Kreiswellenzahl k ist gleich π/L , so daß genau eine halbe Sinusfunktion in das Intervall $[0, L]$ passt. Die Kreisfrequenz $\omega = ck$ ist $\pi c/L$, bis zur letzten Zeile erste Mode genau eine Periode durchlaufen, dort ist also $\pi ct/L = 2\pi$ oder $t = 2L/c$. Die zweite Mode mit doppelt so hoher Kreiswellenzahl und Kreisfrequenz (Mitte) hat $C_2 = 1$ und $C_n = 0$ sonst. Eine sogenannte **Überlagerung** oder **Superposition** der beiden ersten Moden mit $C_1 = C_2 = 1$ und $C_n = 0$ sonst ist rechts gezeigt.

Grundsätzlich müssen die Koeffizienten C_n der Normalmoden aus den Anfangsbedingungen bestimmt werden. In unserem Beispiel der Wellengleichung ist dies die Auslenkung zum Zeitpunkt $t = 0$, $\phi(x, t = 0) \equiv \phi_0(x)$, und die Geschwindigkeit jedes Punktes des Gummibandes zum Zeitpunkt $t = 0$, $\partial\phi/\partial t(x, t = 0) = v_0(x)$. Die Koeffizienten C_n müssen also so gewählt werden, dass

$$\sum_{n=1}^{\infty} C_n \sin\left(\frac{\pi xn}{L}\right) = \phi_0(x) . \quad (4.47)$$

Eine zweite Gleichung erhalten wir, wenn wir die Ableitung von (4.46) nach der Zeit gleich der Geschwindigkeit $v_0(x)$ setzen. Aber läßt sich überhaupt jede Funktion $\phi_0(x)$, die die Randbedingung $\phi_0(0) = \phi_0(L) = 0$ erfüllt, als eine Summe von (prinzipiell unendliche vielen) Sinustermen schreiben? Dieser Frage wird uns später auf das Thema der Fourieranalyse führen.



Fünf

Vektoranalysis

Wir führen Vektorfelder ein, die jedem Punkt im Raum einen Vektor zuordnen. Anwendungen sind Kraftfelder, wie z.B. das elektrische oder magnetische Feld. Wir diskutieren die Integration von solchen Feldern über Linien, Oberflächen und Volumina. Zur Beschreibung dieser Felder verallgemeinern wir das Konzept der Ableitung einer Funktion und führen eine Reihe von Differentialoperatoren ein.

Bisher haben wir uns mit den Elementen von Vektorräumen beschäftigt um Verschiebungen, Geschwindigkeiten, Beschleunigungen (\propto Kräften) eines Massepunktes zu beschreiben. Mehrere Massepunkte könnten dann durch mehrere Vektoren beschrieben werden. Viele physikalische Situationen benötigen eine Erweiterung, die über den Vektorraum hinausgeht: ein Vektorfeld ordnet *jedem Punkt* im Raum einen Vektor zu.

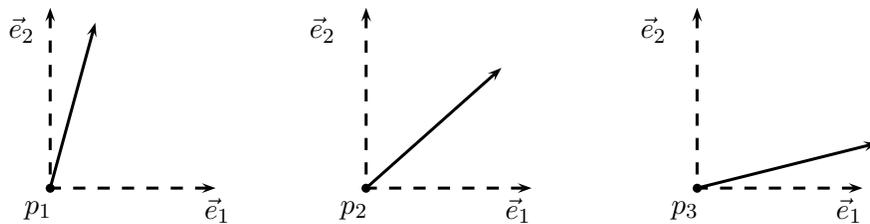
Beispiel

- Zur Beschreibung eines strömenden Gases müssen wir berücksichtigen, dass Richtung und Geschwindigkeit der Gasteilchen vom Ort abhängen. Jedem Punkt im Raum ist also ein Vektor zugeordnet, der angibt in welche Richtung und mit welcher Geschwindigkeit sich Teilchen des Gases an diesem Punkt bewegen. Die Geschwindigkeitsvektoren an benachbarten Punkten sind nicht unabhängig (wenn z.B. Gasteilchen erhalten bleiben).
- Das elektrische Feld $\vec{E} = \vec{F}/q$ hat an unterschiedlichen Punkten im Raum unterschiedliche Richtungen und Stärken. Die Vektoren \vec{E} an unterschiedlichen Punkten sind nicht unabhängig, sie gehorchen den Maxwellgleichungen. Der Begriff des Vektorfeldes ist zentral für den Elektromagnetismus.
- Analoges gilt für das magnetische Feld und für das Gravitationsfeld.

Sei M ein (affiner) Raum (s. 1.3), dann ist ein **Vektorfeld** $\vec{v}(p)$ auf M eine Abbildung

$$\vec{v} : M \longrightarrow V, p \longmapsto \vec{v}(p) , \quad (5.1)$$

die jedem Punkt in M ein Element \vec{v} eines Vektorraumes V zuordnet. Damit “sitzt” auf jedem Punkt des affinen Raumes ein Vektorraum. Die Wahl eines Vektors für jeden Raumpunkt definiert ein konkretes Vektorfeld. Häufig haben der Raum M und der Raum V die gleiche Dimension. Ein Vektorfeld lässt sich in einer Basis darstellen, indem man für jeden Punkt im Raum eine Basis wählt.



Hier ist exemplarisch für alle Punkte p des Raumes M an drei Punkten p_1, p_2, p_3 eine kartesische Basis definiert. Zusätzlich haben diese Basisvektoren jeweils dieselbe Orientierung relativ zu einem kartesischen Koordinatensystem. Man spricht daher auch von einer **raumfesten Basis**. Aus nichtkartesischen Koordinaten kann man Basisysteme definieren, die nicht raumfest sind.

Mit einer Basis an jedem Punkt p kann nun der Vektor \vec{v} an jedem Punkt dargestellt werden durch

$$\mathbf{v}(p) = \begin{pmatrix} v_1(p) \\ v_2(p) \end{pmatrix} \quad \text{für } \vec{v}(p) = v_1(p)\vec{e}_1(p) + v_2(p)\vec{e}_2(p) \quad (5.2)$$

und analog in höherdimensionalen Räumen.

Wählen wir für den affinen Raum M einen Ursprung p_0 und schreiben $p = p_0 + \vec{r}$, dann ist das Vektorfeld in zwei Dimensionen charakterisiert durch

$$\mathbf{v}(\mathbf{r}) = \begin{pmatrix} v_1(\mathbf{r}) \\ v_2(\mathbf{r}) \end{pmatrix} = \begin{pmatrix} v_1(x, y) \\ v_2(x, y) \end{pmatrix} , \quad (5.3)$$

in drei Dimensionen entsprechend

$$\mathbf{v}(\mathbf{r}) = \begin{pmatrix} v_1(x, y, z) \\ v_2(x, y, z) \\ v_3(x, y, z) \end{pmatrix} .$$

In zwei Dimensionen ist ein Vektorfeld also durch zwei Funktionen von zwei Variablen charakterisiert, in drei Dimensionen durch drei Funktionen von drei Variablen.

Analog zum Vektorfeld definieren wir ein **skalares Feld**, das jedem Punkt im Raum eine skalare Größe zuordnet, d.h. eine Abbildung

$$f : M \longrightarrow \mathbb{R}, p \longmapsto f(p) . \quad (5.4)$$

Ein skalares Feld ist durch eine Funktion der Koordinaten charakterisiert, z.B in 3D

$$f(\mathbf{r}) = f(x, y, z) . \quad (5.5)$$

Ein Beispiel für ein skalares Feld ist die Dichte eines Gases, dessen Konzentration an unterschiedlichen Stellen unterschiedlich hoch sein mag, oder die Temperatur, die im Raum variiert.

5.1 Integration von skalaren und vektoriellen Feldern

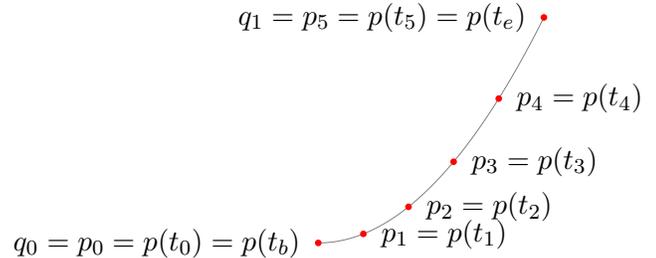
Das Integral $\int_{x_a}^{x_b} dx f(x)$ gibt z.B. das Gewicht eines geraden Stabes, wenn dessen lineare Dichte $f(x)$ ist. x beschreibt die Position entlang des Stabes, $f(x)\Delta x$ ist dann das Gewicht eines kleinen Teilstücks der Länge Δx des Stabes bei x , und das Integral ergibt sich aus dem Limes bei dem die Länge aller Teilstücke gegen null geht. Wie könnte man nun sein Gewicht berechnen, wenn der Stab gebogen wäre? Die Zerlegung des Stabes in eine große Zahl kleiner Teilstücke führt auf das sogenannte Linienintegral. Eine wichtige Anwendung dieses Integrals liegt in der Berechnung der Arbeit, die ein Kraftfeld entlang einer gegebenen Bahnkurve verrichtet. Die Verallgemeinerung auf Integrale von Feldern über Oberflächen und Volumina führt auf Flächen und Volumenintegrale. Sie spielen eine wichtige Rolle in der Elektrodynamik, wo z.B. elektrische Felder über Oberflächen oder Ladungsdichten über Volumina integriert werden. Diese Integrale werden wir in Abschnitt 5.4 und 5.5 definieren.

5.2 Linienintegrale

Bleiben wir zunächst beim Beispiel des eindimensionalen gebogenen Drahtes mit räumlich veränderlicher Dichte. Der Draht bilde eine Kurve C im Raum. Die Dichte f pro Länge eines kurzen Teilstücks hängt nun vom Ort p ab, an dem sich das Teilstück befindet. Zerlegen wir die Kurve C in kleine Intervalle zwischen den Raumpunkten $p_0, p_1, p_2, \dots, p_N$ auf der Kurve C . $\Delta \vec{r}_i \equiv p_{i+1} - p_i$ gibt die Verschiebung von p_i nach p_{i+1} an. Die Masse des i -ten Teilstücks ist dann die Länge des Teilstücks mal die Dichte, $\|\Delta \vec{r}_i\| f(p_i)$. Die Gesamtmasse als Riemannsumme (vgl. die Riemannsumme (2.20)) ist damit

$$\sum_{i=0}^{N-1} \|(p_{i+1} - p_i)\| f(p_i) = \sum_{i=0}^{N-1} \|\Delta \vec{r}_i\| f(p_i) . \quad (5.6)$$

Zur Auswertung dieser Riemann-Summe im Grenzfalle bei dem die einzelnen Schritte klein werden, $\|\Delta\vec{r}_i\| \rightarrow 0$, definieren wir zunächst eine Parametrisierung der Kurve C . Eine Parametrisierung einer Raumkurve C ist eine Funktion $p(t)$, die jedem $t \in [t_b, t_e]$ einen Punkt auf C zuordnet (b für Beginn, e für Ende). Etwas formaler ist eine **Parametrisierung** $p(t)$ einer Kurve C (als Punktmenge) eine differenzierbare Abbildung mit $p([t_b, t_e]) = C$, $p(t_b) = q_0$ (Anfangspunkt), $p(t_e) = q_1$ (Endpunkt) und $\frac{dp}{dt} \neq 0 \quad \forall t \in [t_b, t_e]$. Bei gegebenem Ursprung p_u , mit $p = p_u + \vec{r}$ nutzen wir auch $\vec{r}(t)$ als Parametrisierung der Kurve.



Eine Parametrisierung einer Raumkurve können sie durch Ameise visualisieren, die im Zeitintervall $[t_b, t_e]$ die Raumkurve C (gebogener Draht) entlang krabbelt. Im kleinem Zeitintervall $\Delta t_i = t_{i+1} - t_i$ legt sie die Strecke $\approx \Delta t_i \frac{d\vec{r}(t_i)}{dt}$ zurück, dieses Teilstück trägt dann $\approx \Delta t_i \frac{d\vec{r}(t_i)}{dt} f(p(t_i))$ zur Gesamtmasse des Drahtes bei. Dabei haben wir kleine Teilstücke durch Geraden approximiert (lineare Approximation), $p(t_{i+1}) - p(t_i) = \vec{r}(t_{i+1}) - \vec{r}(t_i) = \Delta t_i \frac{d\vec{r}(t_i)}{dt} + \mathcal{O}(\Delta t_i^2)$.

Im Grenzfalle $\Delta t_i \rightarrow 0 \forall i$ geht die Zahl dieser Teilstücke gegen unendlich und wir definieren das **Linienintegral** als

$$\int_C dr f(\vec{r}) \equiv \lim_{\|\Delta\vec{r}_i\| \rightarrow 0} \sum_{i=0}^{N-1} \|\Delta\vec{r}_i\| f(p_i) = \lim_{\Delta t_i \rightarrow 0} \sum_{i=0}^{N-1} \Delta t_i \left\| \frac{d\vec{r}(t_i)}{dt} \right\| f(p_i) = \int_{t_b}^{t_e} dt \left\| \frac{d\vec{r}(t)}{dt} \right\| f(\vec{r}(t)) \quad (5.7)$$

Bei der Notation $\int_C dr f(\vec{r})$ kann man sich dr als die Länge eines sehr kurzen Teilstücks denken.

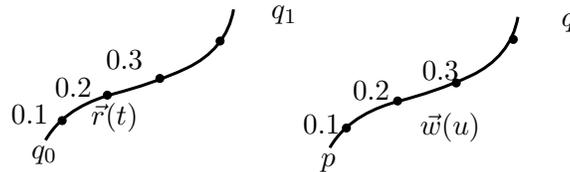
Beispiel

Wir setzen $f(p) = 1$ um die Bogenlänge L des Halbkreises mit Radius 1 zu berechnen,

$L = \int_C dr$ (s. Beispiel in 1.2)

$$\begin{aligned}
 \vec{r}(t) &= \cos(2\pi t)\vec{e}_1 + \sin(2\pi t)\vec{e}_2 \\
 \mathbf{r}(t) &= \begin{pmatrix} \cos(2\pi t) \\ \sin(2\pi t) \end{pmatrix} \\
 \int_C dr &= \int_0^{1/2} dt \left\| \frac{d\vec{r}}{dt} \right\| \\
 \frac{d\mathbf{r}}{dt} &= 2\pi \begin{pmatrix} -\sin(2\pi t) \\ \cos(2\pi t) \end{pmatrix}, \quad \left\| \frac{d\mathbf{r}}{dt} \right\| = \sqrt{(2\pi \sin(2\pi t))^2 + (2\pi \cos(2\pi t))^2} = 2\pi \\
 &= \int_0^{1/2} dt \left\| \frac{d\mathbf{r}}{dt} \right\| = 2\pi \int_0^{1/2} dt = \pi \tag{5.8}
 \end{aligned}$$

Die Parametrisierung der Raumkurve C erlaubt also die Auswertung eines Linienintegrals als (Riemann-) Integral über eine Variable, den Parameter der Raumkurve. Eine wichtige Frage ist allerdings, ob die Definition des Linienintegrals (5.7) von der Wahl der Parametrisierung $\vec{r}(t)$ abhängt. Wir betrachten dazu eine zweite Parametrisierung einer Kurve C , $\vec{w}(u)$.



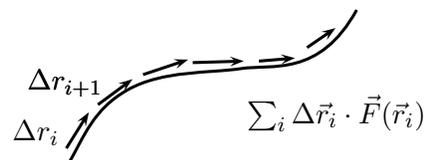
Dann existiert eine differenzierbare Funktion $t(u)$, die jedem u ein t zuordnet, so dass $\vec{w}(u) = \vec{r}(t(u))$. Jeder Schritt $\Delta t > 0$ führt $\vec{r}(t)$ weiter auf der Kurve C von q_0 nach q_1 , und ebenso jeder Schritt Δu . Daher ist $dt/du > 0$ entlang der Raumkurve und mit der Kettenregel $\frac{d\vec{w}}{du} = \frac{d}{du}\vec{r}(t(u)) = \frac{d\vec{r}}{dt} \frac{dt}{du}$ erhalten wir

$$\int_{u_b}^{u_e} du \left\| \frac{d\vec{w}}{du} \right\| f(\vec{w}(u)) = \int_{u_b}^{u_e} du \left\| \frac{d\vec{r}}{dt} \right\| \frac{dt}{du} f(\vec{r}(t(u))) = \int_{t_b}^{t_e} dt \left\| \frac{d\vec{r}}{dt} \right\| f(\vec{r}(t)) \tag{5.9}$$

Ein Wechsel der Parametrisierung lässt sich also als Variablensubstitution der Parametrisierungsvariablen interpretieren.

Das Linienintegral eines Vektorfeldes

Gesucht sei die Arbeit, die ein Kraftfeld $\vec{F}(\vec{r})$ an einer Punktmasse verrichtet, die sich entlang einer Raumkurve C bewegt. Bei der Parametrisierung der



Raumkurve können Sie wieder an eine Ameise denken, die im Zeitintervall Δt_i die Strecke $\Delta \vec{r}_i \approx \Delta t_i \frac{d\vec{r}(t_i)}{dt}$ zurücklegt. Dabei verrichtet das Kraftfeld die Arbeit $\approx \Delta \vec{r}_i \cdot \vec{F}(\vec{r}(t_i)) \approx \Delta t_i \frac{d\vec{r}(t_i)}{dt} \cdot \vec{F}(\vec{r}(t_i))$. Wir definieren das Linienintegral eines Vektorfeldes

$$\int_C d\mathbf{r} \cdot \mathbf{F}(\mathbf{r}) \equiv \int_{t_b}^{t_e} dt \frac{d\mathbf{r}}{dt} \cdot \mathbf{F}(\mathbf{r}(t)) \quad (5.10)$$

$$\stackrel{\text{kart. Basis}}{=} \int_{t_b}^{t_e} dt \frac{dx}{dt} F_x(\mathbf{r}(t)) + \int_{t_b}^{t_e} dt \frac{dy}{dt} F_y(\mathbf{r}(t)) + \int_{t_b}^{t_e} dt \frac{dz}{dt} F_z(\mathbf{r}(t))$$

Bei der Notation $\int_C d\mathbf{r} \cdot \mathbf{F}(\mathbf{r})$ können sie sich $d\mathbf{r}$ als kleinen Vektor entlang eines Teilstücks der Kurve C denken.

Beispiel

Wir betrachten ein Kraftfeld in 2 Dimensionen, $\mathbf{F}(\mathbf{r}) = \begin{pmatrix} xy \\ -y^2 \end{pmatrix}$. Ein Massepunkt bewegt sich entlang einer Parabel $C : y = x^2$ vom Punkt $(0,0)$ zu $(1,1)$. Wir berechnen die am Massepunkt verrichtete Arbeit W mit der Parametrisierung $x(t) = t, y(t) = t^2$ und damit $\frac{d\mathbf{r}}{dt} = \begin{pmatrix} \frac{dx}{dt} \\ \frac{dy}{dt} \end{pmatrix} = \begin{pmatrix} 1 \\ 2t \end{pmatrix}$. Wir erhalten

$$\begin{aligned} W &= \int_C d\mathbf{r} \cdot \mathbf{F}(\mathbf{r}) = \int_0^1 dt \frac{d\mathbf{r}}{dt} \cdot \mathbf{F}(\mathbf{r}(t)) = \int_0^1 dt \begin{pmatrix} 1 \\ 2t \end{pmatrix} \cdot \begin{pmatrix} t^3 \\ -t^4 \end{pmatrix} \\ &= \int_0^1 dt (t^3 - 2t^5) = \left[\frac{1}{4}t^4 - \frac{2}{6}t^6 \right]_0^1 = \frac{1}{4} - \frac{1}{3} = -\frac{1}{12}. \end{aligned} \quad (5.11)$$

Aus dem Spezialfall eines Feldes mit nur einer Komponente $F_x(\vec{r}) = f(\vec{r})$ lässt sich das Linienintegral über eine ausgezeichnete Koordinate definieren

$$\int_C dx f(\vec{r}) \equiv \int_{t_b}^{t_e} dt \frac{dx}{dt} f(\vec{r}(t)). \quad (5.12)$$

Hierbei werden bei der Zerlegung der Raumkurve C die Änderungen der x -Komponente entlang der Teilstücke aufsummiert.

Beispiel

$$\int_C dx = \int_{t_b}^{t_e} dt \frac{dx(t)}{dt} = \int_{x_b}^{x_e} dx = [x]_{x_b}^{x_e} = x_e - x_b \quad (5.13)$$

Im zweiten Schritt wurde das Integral durch die Substitution $x = x(t)$ ausgewertet.

Und schließlich lässt sich analog zum Linienintegral mit dem Skalarprodukt auch ein Integral mit Kreuzprodukt definieren.



$$\int_c \mathbf{dr} \times \mathbf{F}(\mathbf{r}) \equiv \int_{t_b}^{t_e} dt \frac{d\mathbf{r}}{dt} \times \mathbf{F}(\mathbf{r}(t)) \quad (5.14)$$

Dieses Integral tritt zum Beispiel bei der Berechnung des magnetischen Feldes eines stromdurchflossenen Drahtes auf.

5.3 Differentialoperatoren I: Der Gradient

Wie verändert sich die Komponenten \mathbf{v} eines Vektorfeldes $\mathbf{v}(\mathbf{r})$ unter einer (kleinen) Änderung von \mathbf{r} ? Wie ändert sich der Wert eines skalaren Feldes $f(\mathbf{r})$ unter einer solchen kleinen Änderung? Vektorfelder in drei Dimensionen werden durch drei Funktionen die von drei Veränderlichen abhängen beschrieben, und analog wird ein skalares Feld von einer Funktion von drei Veränderlichen beschrieben. Entsprechend ist das Konzept der Ableitung von Feldern mathematisch reicher als das der Ableitung einer einzelnen Funktion einer Variablen.

Wir werden drei sogenannte Differentialoperatoren auf Feldern diskutieren, die jeweils unterschiedliche Aspekte von Feldern charakterisieren; den Gradienten eines skalaren Feldes (ergibt ein Vektorfeld), die Rotation eines Vektorfeldes (ergibt ein Vektorfeld, siehe 5.6), und die Divergenz eines Vektorfeldes (ergibt ein skalares Feld, siehe 5.7).

Für eine Funktion einer Veränderlichen hatten wir in Abschnitt 2.4.1 die Ableitung als lineare Näherung einer Funktion $f(x)$ an einem Punkt verstanden, $f(x + \Delta x) = f(x) + \frac{df}{dx} \Delta x + \mathcal{O}(\Delta x^2)$. Der Gradient verallgemeinert die Ableitung und gibt an, wie sich ein skalares Feld $f(\mathbf{r})$ unter einer (kleinen) Änderung $\Delta \mathbf{r}$ ändert.

Sei $f(x_1, x_2, x_3, \dots)$ eine Funktion $\mathbb{R}^n \rightarrow \mathbb{R}$, $\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ \dots \end{pmatrix} \mapsto f(\mathbf{x})$, und x_1, x_2, \dots, x_n kartesische Koordinaten des \mathbb{R}^n des Punktes $p_0 + \mathbf{x}$, dann heißt der Spaltenvektor

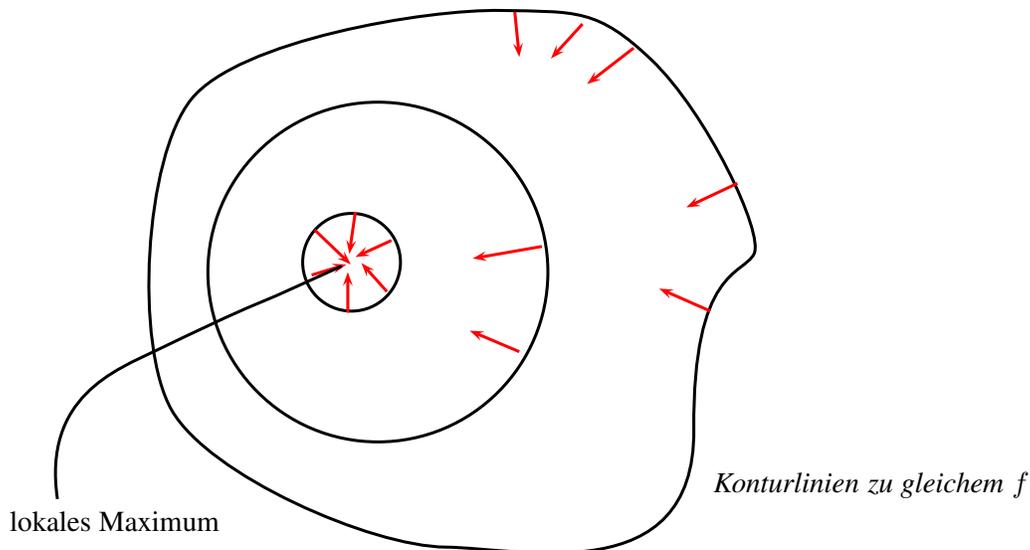
$$\nabla f = \begin{pmatrix} \partial f / \partial x_1 \\ \partial f / \partial x_2 \\ \partial f / \partial x_3 \\ \vdots \end{pmatrix} \quad (5.15)$$

der **Gradient** der Funktion $f(x_1, x_2, x_3, \dots) \equiv f(\mathbf{r})$ am Punkt $p = p_0 + \mathbf{x}$.

Betrachten wir die Funktion am Punkt p und an einem zweiten Punkt $p + \Delta \mathbf{x}$. Das Skalarprodukt von ∇f mit einem Verschiebungsvektor $\Delta \mathbf{x}$ gibt die Änderung von $f(\mathbf{x})$ zu $f(\mathbf{x} + \Delta \mathbf{x})$ in linearer Näherung an

$$\Delta f \equiv f(\mathbf{x} + \Delta \mathbf{x}) - f(\mathbf{x}) = \sum_{i=1}^n \frac{\partial f}{\partial x_i} \Delta x_i + \mathcal{O}(\|\Delta \mathbf{x}\|^2) = \nabla f \cdot \Delta \mathbf{x} + \mathcal{O}(\|\Delta \mathbf{x}\|^2). \quad (5.16)$$

Der Gradient einer Funktion mehrerer Veränderlicher $f(x_1, x_2, \dots)$, bzw. $f(\mathbf{x})$ kann also als Verallgemeinerung der Ableitung einer Funktion einer Variablen betrachtet werden. Zur Visualisierung des Gradienten stellen wir eine Funktion $f(p)$ durch Konturlinien dar, die wie Höhenlinien einer Landkarte Punkte mit gleichem Wert der Funktion f verbinden.



Dann steht der Vektor ∇f senkrecht auf den Konturlinien, denn für ein $\Delta \mathbf{x}$ entlang einer Kontour ist $\nabla f \cdot \Delta \mathbf{x} = 0$. Betrachten wir alle $\Delta \mathbf{x}$ mit konstanter (kleiner) Norm. Da das Skalarprodukt zweier Vektoren konstanter Länge dann maximal ist, wenn beide Vektoren

in dieselbe Richtung zeigen, zeigt ∇f in die Richtung maximaler Veränderung von f (in linearer Näherung, also für kleine $\|\Delta \mathbf{x}\|$). Der Gradient zeigt damit in die Richtung der größten Veränderung von f .

Beispiel

$$f(x, y, z) = \sqrt{x^2 + y^2 + z^2}$$

$$\nabla f = \begin{pmatrix} \partial f / \partial x \\ \partial f / \partial y \\ \partial f / \partial z \end{pmatrix} = \frac{1}{2\sqrt{x^2 + y^2 + z^2}} \begin{pmatrix} 2x \\ 2y \\ 2z \end{pmatrix} = \frac{\mathbf{r}}{\|\mathbf{r}\|}$$

Hier bilden die Punkte mit gleichem Wert der Funktion f kugelförmige Schalen, entsprechend zeigt ∇f radial nach außen. In zwei Dimensionen ist auch noch eine alternative Visualisierung möglich: Betrachtet man $f(x, y) = \sqrt{x^2 + y^2}$ als Höhenprofil über der xy -Ebene erhält man einen Trichter mit Mittelpunkt beim Ursprung.

5.3.1 Gradientenfelder

Ein Vektorfeld $\mathbf{v}(\mathbf{r})$ heißt **Gradientenfeld**, wenn eine Funktion $f(\mathbf{r})$ existiert, so dass

$$\mathbf{v}(\mathbf{r}) = \nabla f(\mathbf{r}) \tag{5.17}$$

Ein allgemeines Vektorfeld aus d -dimensionalen Vektoren wird durch d Funktionen beschrieben, ein Gradientenfeld ist also ein sehr spezielles Vektorfeld, denn es ist durch die *eine* Funktion $f(\mathbf{r})$ vollständig beschrieben. Jedes Gradientenfeld weist eine Beziehung zwischen den partiellen Ableitungen seiner Komponenten auf: Sei $\mathbf{v}(\mathbf{x}) = \nabla f$ ein Gradientenfeld. Aus der Schwarzschen Regel für partielle Ableitungen $\frac{\partial^2 f}{\partial x_i \partial x_j} = \frac{\partial^2 f}{\partial x_j \partial x_i}$ gilt für die Komponenten $v_i = \frac{\partial f}{\partial x_i}$ und $v_j = \frac{\partial f}{\partial x_j}$ auch $\frac{\partial v_j}{\partial x_i} = \frac{\partial v_i}{\partial x_j}$.

Ein Gradientenfeld im \mathbb{R}^3 hat damit

$$\begin{aligned} \frac{\partial v_z}{\partial y} - \frac{\partial v_y}{\partial z} &= 0 \\ \frac{\partial v_x}{\partial z} - \frac{\partial v_z}{\partial x} &= 0 \\ \frac{\partial v_y}{\partial x} - \frac{\partial v_x}{\partial y} &= 0. \end{aligned} \tag{5.18}$$

In einfach zusammenhängenden Räumen (grob ausgedrückt: ohne Löcher) stellt sich heraus, dass diese Bedingung auch hinreichend für ein Gradientenfeld ist. Zwei skalare Funktionen $f(\mathbf{r})$ und $f(\mathbf{r}) + C$ generieren dasselbe Gradientenfeld.

Sei $\mathbf{F}(\mathbf{x})$ ein Gradientenfeld, d.h. $\mathbf{F}(\mathbf{r}) = \nabla f(\mathbf{x})$. Dann gilt für sein Linienintegral entlang einer Linie C

$$\begin{aligned} \int_C d\mathbf{x} \cdot \mathbf{F}(\mathbf{x}) &= \int_{t_b}^{t_e} dt \frac{d\mathbf{x}}{dt} \cdot \mathbf{F}(\mathbf{x}(t)) & \frac{d}{dt}(f(\mathbf{x}(t))) &= \nabla f \cdot \frac{d\mathbf{x}}{dt} \\ &= \int_{t_b}^{t_e} dt \frac{d}{dt}(f(\mathbf{x}(t))) & &= [f(\mathbf{x}(t))]_{t_b}^{t_e} \end{aligned}$$

Damit ist das Linienintegral eines Gradientenfeldes wegunabhängig, es hängt nur von Anfangs- und Endpunkt ab. Das Gravitationsfeld ist ein Gradientenfeld, ebenso das von statischen Ladungen erzeugte elektrische Feld. Die von einem Gradientenfeld verrichtete Arbeit um eine Punktmasse von Punkt q_0 zu q_q zu bringen ist also unabhängig vom konkret gewählten Weg. Die skalare Funktion f wird in diesem Zusammenhang als Potential bezeichnet.

Beispiel

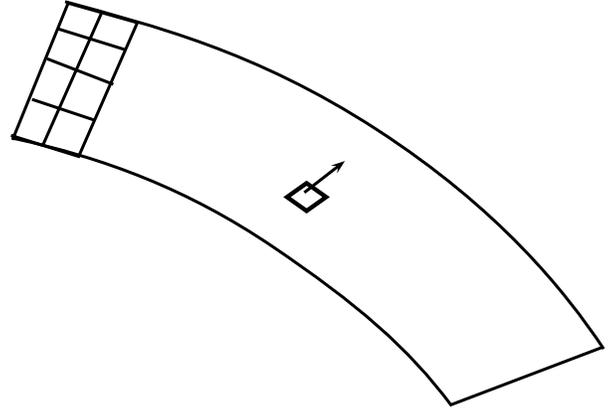
Ein Beispiel für ein Gradientenfeld ist das Kraftfeld der Schwerkraft. Im Zusammenhang von Kraftfeldern spricht man auch von einem konservativen Feld. Erstes Beispiel ist ein homogenes Kraftfeld in Richtung der z -Achse, $\mathbf{F} = -mg\mathbf{e}_z$. Dieses Vektorfeld ist ein Gradientenfeld, da $\mathbf{F} = -\nabla V(\mathbf{r})$ mit $V(\mathbf{r}) = V(x, y, z) = mgz$. V wird in diesem Zusammenhang als Gravitationspotential bezeichnet. Die Arbeit, die das Kraftfeld an einer Punktmasse verrichtet, die entlang einer Kurve C mit Anfangspunkt p und Endpunkt q bewegt wird ist $W = \int_C d\mathbf{r} \cdot \mathbf{F}(\mathbf{r}) = -mg(z_q - z_p)$. Der Wert des Graviationspotentials an einem Punkt wird auch als potentielle Energie bezeichnet. Ist die an der Punktmasse verrichtete Arbeit positiv (z.B. wenn die Masse im Gravitationspotential fällt), wird potentielle Energie in kinetische Energie umgewandelt. Ist die Arbeit negativ, wird kinetische Energie in potentielle Energie umgewandelt – die Masse wird abgebremst.

Ein zweites Beispiel ist das Kraftfeld einer zweiten Punktmasse mit Masse M am Ursprung, $\mathbf{F} = -\frac{mMG}{r^2} \frac{\mathbf{r}}{|\mathbf{r}|}$. Die Vektoren dieses Feldes zeigen radial nach innen, ihre Norm wird mit zunehmendem Abstand von Ursprung immer kleiner. Das entsprechende Potential ist $V(\mathbf{r}) = mMG/|\mathbf{r}|$.

5.4 Das Oberflächenintegral

Das Oberflächenintegral erlaubt Vektor- und Skalarfelder auf Oberflächen zu integrieren. Wir benötigen dazu das Konzept der Mehrfachintegrale aus Abschnitt 2.7. Ein Beispiel ist die Berechnung der Gesamtladung auf einer gekrümmten Oberfläche mit gegebener Dichte pro kleiner Fläche. Weitere wichtige Anwendungen in der Elektrostatik sind Oberflächenintegrale des elektrischen Feldes im Zusammenhang mit dem Gaußschen Integralsatz, den wir in Sektion 5.8 diskutieren werden.

Wir zerlegen wieder eine Oberfläche in kleine Teilflächen, über die wir dann summieren werden. Bei einer einzelnen kleinen Teilfläche nutzen wir, dass ihr Beitrag zum Integral lediglich von ihrem Flächeninhalt und von ihrer Orientierung im Raum abhängt: ist die Teilfläche hinreichend klein, dann ist der Integrand (z.B. eine Dichte oder ein Vektorfeld) konstant über die Teilfläche. Die genaue Form der Teilfläche (dreieckig, rund, elliptisch, ...) ist dann irrelevant. Zur Charakterisierung des Flächeninhalts und der Orientierung einer Teilfläche ΔS definieren wir einen Vektor $\Delta \vec{v}ecS$, der senkrecht zur Teilfläche ΔS steht, und dessen Länge $\|\Delta \vec{v}ecS\|$ gleich ihrem Flächeninhalt ist. Dieser Vektor wird als **vektorielles Oberflächenelement** oder **Oberflächenelement**



bezeichnet. Die Orientierung von $\Delta \vec{v}ecS$ wird von Fall zu Fall unterschiedlich gewählt¹. Bei geschlossenen Oberflächen wählt man die Richtung von $\Delta \vec{v}ecS$ meist so, dass die Oberflächenelemente nach außen zeigen.

Die Summe über N solcher kleinen Flächenelemente führt uns im Grenzfall $N \rightarrow \infty$ auf die Flächenintegrale

$$\int_S d\vec{v}ecS \cdot \vec{v}ecF(\vec{v}ecr) \equiv \lim_{N \rightarrow \infty} \sum_{\text{Teilflächen } i} \Delta \vec{v}ecS_i \cdot \vec{v}ecF(\vec{v}ecr_i) \quad \text{eines Vektorfeldes } \vec{v}ecF(\vec{v}ecr) \quad (5.19)$$

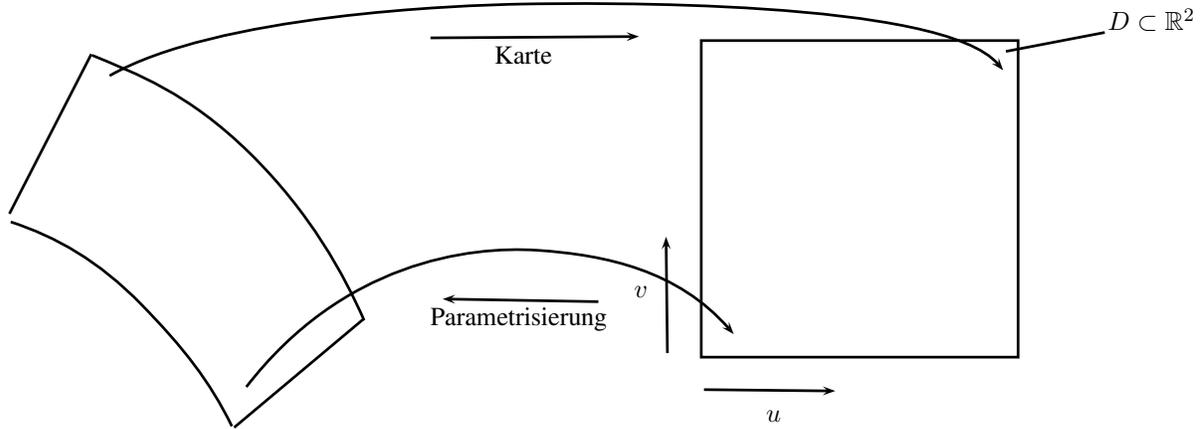
$$\int_S dS f(\vec{v}ecr) \equiv \lim_{N \rightarrow \infty} \sum_{\text{Teilflächen } i} \|\Delta \vec{v}ecS_i\| f(\vec{v}ecr_i) \quad \text{eines skalaren Feldes } f(\vec{v}ecr) \quad (5.20)$$

Bei einer Flüssigkeit konstanter Dichte mit Fließgeschwindigkeit $\vec{v}ecF(\vec{v}ecr)$ gibt das Flächenintegral über $\vec{v}ecF(\vec{v}ecr)$ den Gesamtfluss pro Zeit über S an. Bei einer Dichte pro kleiner Fläche eines dünnen Bleches, gibt das Flächenintegral über $f(\vec{v}ecr)$ die Masse eines Bleches mit Oberfläche S an.

Zur konkreten Definition und Berechnung solcher Integrale benötigen wir wieder eine Parametrisierung. Zweidimensionale Oberflächen benötigen zwei Variablen (u, v) um einen Punkt $p = p_0 + \vec{r}$ auf der Oberfläche eindeutig festzulegen. Eine **Parametrisierung**

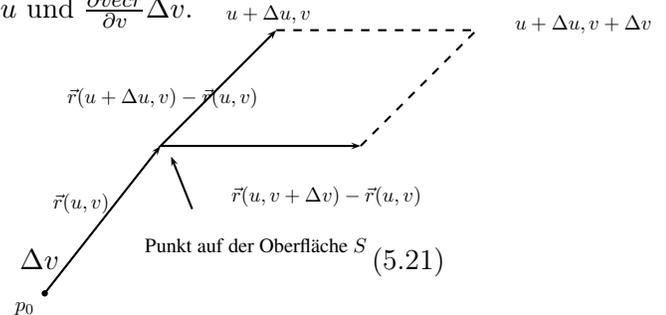
¹Es gibt auch Oberflächen, für die keine solche Orientierung möglich ist, wir also nicht eine Seite als 'außen' und die andere als 'innen' bezeichnen können, genausowenig wie man eine Seite blau und die andere rot malen könnte. Der berühmte Möbiusstreifen ist ein Beispiel. Für solche Oberflächen können wir kein Oberflächenintegral definieren.

einer Oberfläche S (als Punktmenge) ist eine bijektive Abbildung von $D \subset \mathbb{R}^2$ auf S , $(u, v) \in D \mapsto \text{vecr}(u, v) : p_0 + \text{vecr}(u, v) \in S$, die stetig differenzierbar ist und deren Inverses ebenso stetig differenzierbar ist.



Wir zerlegen nun D in kleine rechteckige Teilstücke mit Kantenlängen $\Delta u_i, \Delta v_i$, wie bei der Definition des Mehrfachintegrals in 2.7. Diese Zerlegung von D definiert auch eine Zerlegung der Oberfläche S in kleine Teilstücke. Näherungsweise (kleine $\Delta u_i, \Delta v_i$) handelt es sich dabei um Parallelogramme mit Kanten $\frac{\partial \text{vecr}}{\partial u} \Delta u$ und $\frac{\partial \text{vecr}}{\partial v} \Delta v$. Zur Berechnung des Vektors des entsprechenden Flächenelementes $\Delta \text{vec} S$ nutzen wir das Kreuzprodukt

$$\begin{aligned} \Delta \mathbf{S} &= \frac{\partial \mathbf{r}}{\partial u} \Delta u \times \frac{\partial \mathbf{r}}{\partial v} \Delta v \\ &= \left(\frac{\partial \mathbf{r}}{\partial u} \times \frac{\partial \mathbf{r}}{\partial v} \right) \Delta u \Delta v \end{aligned} \quad (5.21)$$



Gegeben eine Parametrisierung $\mathbf{r}(u, v)$ erhalten wir damit für die Oberflächenintegrale (5.19)

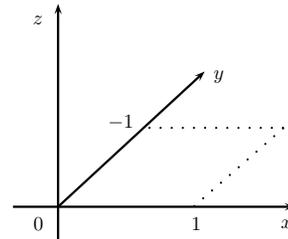
$$\begin{aligned} \int_S d\mathbf{S} \cdot \mathbf{F}(\mathbf{r}) &\equiv \int_D dudv \left(\frac{\partial \mathbf{r}}{\partial u} \times \frac{\partial \mathbf{r}}{\partial v} \right) \cdot \mathbf{F}(\mathbf{r}(u, v)) \\ \int_S dS f(\mathbf{r}) &\equiv \int_D dudv \left\| \frac{\partial \mathbf{r}}{\partial u} \times \frac{\partial \mathbf{r}}{\partial v} \right\| f(\mathbf{r}(u, v)) \end{aligned} \quad (5.22)$$

Beispiel

Wir betrachten das Integral der Funktion $f(\mathbf{r}) = 1$ über eine ebene quadratische Oberfläche in der z -Ebene. $S = \{(x, y, z) \in \mathbb{R}^3 | 0 \leq x \leq 1, 0 \leq y \leq 1, z = 0\}$. Die einfachste Parametrisierung dieser Oberfläche ist $x(u, v) = u, y(u, v) = v, z(u, v) = 0$, also $\mathbf{r}(u, v) = \begin{pmatrix} u \\ v \\ 0 \end{pmatrix}$.

$$\frac{\partial \mathbf{r}}{\partial u} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \quad \frac{\partial \mathbf{r}}{\partial v} = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$$
$$\frac{\partial \mathbf{r}}{\partial u} \times \frac{\partial \mathbf{r}}{\partial v} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$$

$$\int_S dS = \int_D du dv = \int_0^1 du \int_0^1 dv = 1$$



Beispiel

Ein weniger einfaches Beispiel ist die Oberfläche einer Kugel mit Radius R . Zur Parametrisierung nutzen wir die **sphärischen Koordinaten** oder **Kugelkoordinaten** in 3D. Die beiden Winkelparameter θ, ϕ (statt u, v) sind so definiert, dass der Ortsvektor \mathbf{r} mit der z -Achse den **Polarwinkel** θ einschließt, und seine Projektion auf die xy -Ebene mit der x -Achse den **Azimutwinkel** ϕ einschließt.

$$\mathbf{r}(\theta, \phi) = \begin{pmatrix} x(\theta, \phi) \\ y(\theta, \phi) \\ z(\theta, \phi) \end{pmatrix} = \begin{pmatrix} R \sin \theta \cos \phi \\ R \sin \theta \sin \phi \\ R \cos \theta \end{pmatrix} \quad (5.23)$$

Die Werte des Polarwinkels θ liegen zwischen 0 (Nordpol) und π (Südpol), die des Azimutwinkels ϕ zwischen 0 und 2π . Die Ableitung des Ortsvektors nach diesen Parametern ergibt

$$\frac{\partial \mathbf{r}}{\partial \theta} = R \begin{pmatrix} \cos \theta \cos \phi \\ \cos \theta \sin \phi \\ -\sin \theta \end{pmatrix} \quad \frac{\partial \mathbf{r}}{\partial \phi} = R \begin{pmatrix} -\sin \theta \sin \phi \\ \sin \theta \cos \phi \\ 0 \end{pmatrix} \quad (5.24)$$

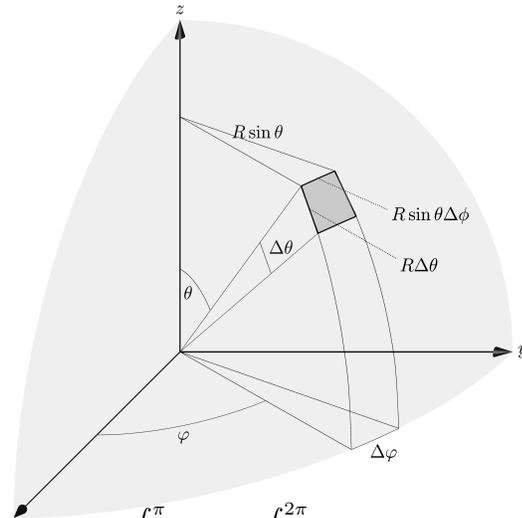
und damit das Kreuzprodukt

$$\frac{\partial \mathbf{r}}{\partial \theta} \times \frac{\partial \mathbf{r}}{\partial \phi} = \begin{pmatrix} R^2 \sin^2 \theta \cos \phi \\ R^2 \sin^2 \theta \sin \phi \\ R^2 \cos \theta \sin \theta (\cos^2 \phi + \sin^2 \phi) \end{pmatrix} = R^2 \sin \theta \begin{pmatrix} \sin \theta \cos \phi \\ \sin \theta \sin \phi \\ \cos \theta \end{pmatrix}. \quad (5.25)$$

Das Resultat läßt sich auch in der Form $R \sin \theta \mathbf{r}(\theta, \phi)$ schreiben (s. Definition der Kugelkoordinaten oben). Das vektorielle Flächenelement $\mathbf{dS} = \left(\frac{\partial \mathbf{r}}{\partial \theta} \times \frac{\partial \mathbf{r}}{\partial \phi} \right) \Delta \theta \Delta \phi$ zeigt also radial nach außen, wie auf einer Kugel zu erwarten. Auch sein Flächeninhalt $R^2 \sin \theta \Delta \theta \Delta \phi$ läßt sich geometrisch verstehen. Ein Kreis mit konstantem Polarwinkel θ hat Radius $R \sin \theta$. Das kleine Rechteck, das durch kleine Änderungen $\Delta \theta$ und $\Delta \phi$ definiert ist hat also Seitenlängen $R \sin \theta \Delta \phi$ und $R \Delta \theta$ und damit den Flächeninhalt $R^2 \sin \theta \Delta \theta \Delta \phi$.

Mit diesem Ergebnis können wir das Flächenintegral leicht berechnen und erhalten

$$\begin{aligned} \int_S dS &= \int_D d\theta d\phi \left\| \frac{\partial \mathbf{r}}{\partial \theta} \times \frac{\partial \mathbf{r}}{\partial \phi} \right\| = \int_D d\theta d\phi R^2 \sin \theta = R^2 \int_0^\pi d\theta \sin \theta \int_0^{2\pi} d\phi \\ &= R^2 [-\cos \theta]_0^\pi [\phi]_0^{2\pi} = 4\pi R^2. \end{aligned} \quad (5.26)$$



5.5 Das Volumenintegral

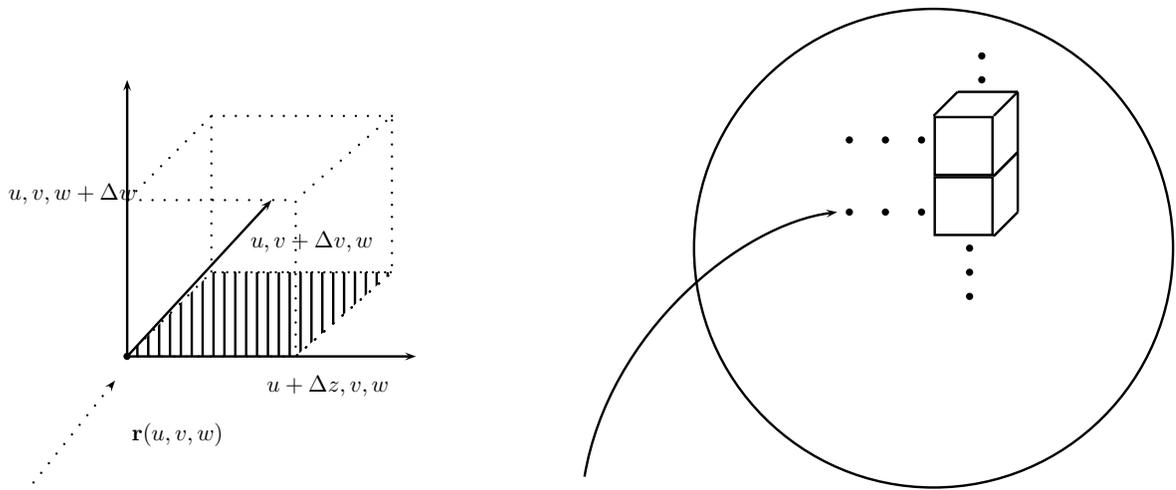
Das Volumenintegral ermöglicht ein Skalarfeld $\phi(\mathbf{r})$ oder ein Vektorfeld $\mathbf{F}(\mathbf{r})$ über ein Volumen zu integrieren. Ein Beispiel ist das Integral einer Ladungsdichte über ein Volumen. Zur Definition des Volumenintegrals gehen wir vor wie bei der Definition des Linien- und Oberflächenintegrals und unterteilen das Volumen in eine große Zahl kleiner Volumenelemente mit Volumen ΔV . Innerhalb eines Volumenelements ist der Integrand $\phi(\mathbf{r})$ oder $\mathbf{F}(\mathbf{r})$ näherungsweise konstant, so dass wir von jedem Volumenelement ΔV mit dem Integrand multiplizieren, und diese Beiträge aller Volumenelemente aufsummieren. Im Grenzfall bei dem ΔV aller Volumenelemente gegen null gehen erhalten wir das Volumenintegral.

Die Parametrisierung eines Volumens erfordert drei Variablen (u, v, w) , ein Raumpunkt ist also durch eine Funktion $\vec{r}(u, v, w)$ festgelegt mit $(u, v, w) \in D \subset \mathbb{R}^3$. Die Kurven $\vec{r}(u, v, w)$ bei denen jeweils alle Parameter bis auf einen festgehalten werden kann man sich (näherungsweise) als Netz aus lauter kleinen Parallelepipeden visualisieren. Diese Parallelepipede nutzen wir als Volumenelemente.

Das Volumen eines kleinen Parallelepipeds, das durch die Vektoren $\frac{\partial \vec{r}}{\partial u} \Delta u$, $\frac{\partial \vec{r}}{\partial v} \Delta v$ und $\frac{\partial \vec{r}}{\partial w} \Delta w$ aufgespannt wird ist

$$\Delta V = \left| \frac{\partial \mathbf{r}}{\partial u} \cdot \left(\frac{\partial \mathbf{r}}{\partial v} \times \frac{\partial \mathbf{r}}{\partial w} \right) \right| \Delta u \Delta v \Delta w . \quad (5.27)$$

Das Spatprodukt hat eine Orientierung (ist positiv für einen rechtshändigen Spat, negativ für einen linkshändigen, s. 1.6); daher ist das Volumen des von 3 Vektoren aufgespannten Parallelepipeds der Betrag des Spatproduktes.



ΔV gibt das Volumen eines kleinen Volumenelementes an

Führen wir nun die Zahl der Volumenelemente gegen unendlich und die $\Delta u, \Delta v, \Delta w$ für jedes einzelne Element gegen null, erhalten wir als Definition des **Volumenintegrals**

$$\int_V dV \phi(\mathbf{r}) \equiv \int_D du dv dw \left| \frac{\partial \mathbf{r}}{\partial u} \cdot \left(\frac{\partial \mathbf{r}}{\partial v} \times \frac{\partial \mathbf{r}}{\partial w} \right) \right| \phi(\mathbf{r}(u, v, w)) \quad (5.28)$$

$$\int_V dV \mathbf{F}(\mathbf{r}) \equiv \int_D du dv dw \left| \frac{\partial \mathbf{r}}{\partial u} \cdot \left(\frac{\partial \mathbf{r}}{\partial v} \times \frac{\partial \mathbf{r}}{\partial w} \right) \right| \mathbf{F}(\mathbf{r}(u, v, w)) \quad (5.29)$$

Für kartesische Koordinaten ist das Spatprodukt eins, $\int_V dV \phi(\mathbf{r}) = \int_V dx dy dz \phi(x, y, z)$. Eine Änderung der Parametrisierung, z.B. von kartesischen Koordinaten zu Kugelkoordinaten, kann auch wieder als Variablensubstitution der Parametrisierungsvariablen verstanden werden.

Beispiel

Als Beispiel berechnen wir noch das Volumen einer Kugel mit Radius R . Zur Parametrisierung bieten sich wieder Kugelkoordinaten an

$$\mathbf{r}(r, \theta, \phi) = \begin{pmatrix} x(r, \theta, \phi) \\ y(r, \theta, \phi) \\ z(r, \theta, \phi) \end{pmatrix} = \begin{pmatrix} r \sin \theta \cos \phi \\ r \sin \theta \sin \phi \\ r \cos \theta \end{pmatrix} \quad (5.30)$$

mit einer **Radialkoordinate** $r = \sqrt{x^2 + y^2 + z^2}$, die den Abstand eines Punktes vom Ursprung im Kugelmittelpunkt angibt. Durch Ableitung nach der Radialkoordinate erhalten wir $\frac{\partial \mathbf{r}}{\partial r} = \begin{pmatrix} \sin \theta \cos \phi \\ \sin \theta \sin \phi \\ \cos \theta \end{pmatrix}$, und zusammen mit dem Ergebnis aus dem Beispiel zum Flächenintegral

$$\frac{\partial \mathbf{r}}{\partial \theta} \times \frac{\partial \mathbf{r}}{\partial \phi} = r^2 \sin \theta \begin{pmatrix} \sin \theta \cos \phi \\ \sin \theta \sin \phi \\ \cos \theta \end{pmatrix} \quad (5.31)$$

erhalten wir das Spatprodukt

$$\frac{\partial \mathbf{r}}{\partial r} \cdot \left(\frac{\partial \mathbf{r}}{\partial \theta} \times \frac{\partial \mathbf{r}}{\partial \phi} \right) = r^2 \sin \theta . \quad (5.32)$$

Das Volumenelement in Kugelkoordinaten ist also $r^2 \sin \theta \Delta r \Delta \theta \Delta \phi$. Auch dieses Ergebnis lässt sich aus geometrischen Überlegungen herleiten; die Kanten des kleinen Quaders, der durch Koordinatenänderungen $\Delta r, \Delta \theta$ und $\Delta \phi$ generiert wird sind $\Delta r, r \Delta \theta$ und $r \sin \theta \Delta \phi$.

Das Volumen der Kugel ist dann

$$\begin{aligned} \int_V dV &= \int_D dr \, d\theta \, d\phi \left| \frac{\partial \mathbf{r}}{\partial r} \cdot \left(\frac{\partial \mathbf{r}}{\partial \theta} \times \frac{\partial \mathbf{r}}{\partial \phi} \right) \right| \\ &= \int_D dr \, d\theta \, d\phi r^2 \sin \theta = \int_0^R dr r^2 \int_0^\pi d\theta \sin \theta \int_0^{2\pi} d\phi = \frac{4}{3} \pi R^3 . \end{aligned} \quad (5.33)$$

5.6 Die Rotation (engl. "curl")

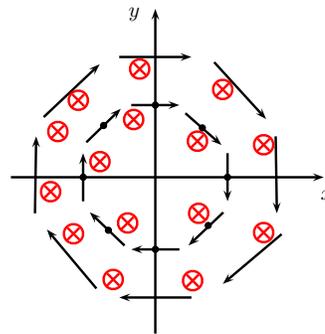
In 3 Dimensionen und kartesischen Koordinaten lässt sich für ein Vektorfeld $\mathbf{v}(\mathbf{x})$ ein Feld von Spaltenvektoren definieren, die sogenannte **Rotation** von $\mathbf{v}(\mathbf{x})$, mit Komponenten

$$\nabla \times \mathbf{v} = \begin{pmatrix} \frac{\partial v_z}{\partial y} - \frac{\partial v_y}{\partial z} \\ \frac{\partial v_x}{\partial z} - \frac{\partial v_z}{\partial x} \\ \frac{\partial v_y}{\partial x} - \frac{\partial v_x}{\partial y} \end{pmatrix} \quad (5.34)$$

$\nabla \times$ soll an das Kreuzprodukt $\begin{pmatrix} \partial/\partial x \\ \partial/\partial y \\ \partial/\partial z \end{pmatrix} \times \begin{pmatrix} v_x \\ v_y \\ v_z \end{pmatrix}$ erinnern. Die Komponenten der Rotation sind uns bereits begegnet, aus (5.18) hat ein Gradientenfeld damit die Rotation Null, $\nabla \times \nabla f = 0 \, \forall f(\mathbf{x})$, oder kurz $\nabla \times \nabla = 0$. Ein Vektorfeld, dessen Rotation nicht verschwindet, kann also kein Gradientenfeld sein, Linienintegrale über ein solches Vektorfeld sind also vom Weg abhängig.

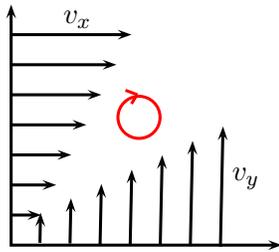
Beispiel

$$\begin{aligned} \mathbf{v}(\mathbf{r}) &= \begin{pmatrix} y \\ -x \\ 0 \end{pmatrix} \\ \nabla \times \mathbf{v} &= \begin{pmatrix} \frac{\partial 0}{\partial y} - \frac{\partial}{\partial z}(-x) \\ \frac{\partial}{\partial z} y - \frac{\partial}{\partial x} 0 \\ \frac{\partial}{\partial x}(-x) - \frac{\partial}{\partial y}(y) \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ -2 \end{pmatrix} \end{aligned}$$

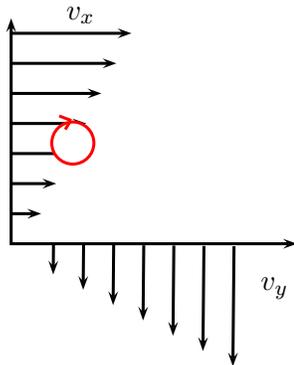


Die Rotation gibt an ob und wo "Wirbel" im Vektorfeld auftreten. Wirbel wie in dem obigen Beispiel führen dazu, dass Linienintegrale über ein Vektorfeld wegabängig sind.

In dem Beispiel oben ist das Linienintegral entlang C von einem Punkt zurück zu sich selbst abhängig davon, wie oft und in welcher Richtung C den Ursprung umläuft. Die Rotation eines Vektorfeldes lässt sich visualisieren; $\mathbf{v}(\mathbf{x})$ sei das Geschwindigkeitsfeld einer strömenden Flüssigkeit. $\nabla \times \mathbf{v}(\mathbf{x})$ beschreibt, wie ein kleiner Ball mit festem Mittelpunkt \mathbf{x} rotiert. $\|\nabla \times \mathbf{v}\|$ ist proportional zur Winkelgeschwindigkeit, die Richtung von $\nabla \times \mathbf{v}$ gibt die Drehachse an.



$$\frac{\partial v_y}{\partial x} = \frac{\partial v_x}{\partial y}$$



$\frac{\partial v_y}{\partial x}$ und $\frac{\partial v_x}{\partial y}$ haben unterschiedliche Vorzeichen:

$$\Rightarrow \frac{\partial v_x}{\partial x} - \frac{\partial v_x}{\partial y} \neq 0$$

Die Rotation spielt eine wichtige Rolle in der Magnetostatik, eines der Maxwell-Gesetze verbindet die Rotation des Magnetfeldes mit der Flussdichte des elektrischen Stroms.

5.7 Die Divergenz

ist eine Abbildung von einem Vektorfeld $\mathbf{v}(\mathbf{x})$ auf ein skalares Feld, definiert in kartesischen Koordinaten als

$$\nabla \cdot \mathbf{v} \equiv \frac{\partial v_x}{\partial x} + \frac{\partial v_y}{\partial y} + \frac{\partial v_z}{\partial z} \quad (5.35)$$

(und analog in n Dimensionen). Die Notation $\nabla \cdot$ soll an das Skalarprodukt erinnern:

$$\nabla \cdot \mathbf{v} = \begin{pmatrix} \partial/\partial x \\ \partial/\partial y \\ \partial/\partial z \end{pmatrix} \cdot \begin{pmatrix} v_x \\ v_y \\ v_z \end{pmatrix} \quad (5.36)$$

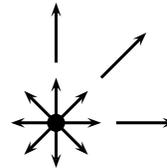
Wie die Rotation spielt auch die Divergenz eine wichtige Rolle im Elektromagnetismus. Eine der Maxwellgleichungen verbindet die Divergenz des elektrischen Feldes mit der elektrischen Ladungsdichte.

Beispiel

$$\mathbf{v}(\mathbf{x}) = \begin{pmatrix} x \\ y \\ z \end{pmatrix}$$

$$\nabla \cdot \mathbf{v} = \frac{\partial v_x}{\partial x} + \frac{\partial v_y}{\partial y} + \frac{\partial v_z}{\partial z} = \frac{\partial x}{\partial x} + \frac{\partial y}{\partial y} + \frac{\partial z}{\partial z} = 3$$

Damit ist die Divergenz des Vektorfeldes $\begin{pmatrix} x \\ y \\ z \end{pmatrix}$ konstant.



Auch die Divergenz hat eine intuitive Interpretation: $\mathbf{v}(\mathbf{x})$ beschreibe die Bewegung einer Flüssigkeit oder eines Gases und sei gleich dem Geschwindigkeitsfeld. Die Zahl der Teilchen, die eine kleine Fläche pro Zeit und Flächeninhalt durchströmt ist dann $\mathbf{v}(\mathbf{r}) \cdot \mathbf{S}$ (wobei \mathbf{S} ein Einheitsvektor senkrecht zur Oberfläche ist). $\nabla \cdot \mathbf{v}$ gibt an, wieviele Teilchen netto aus einem kleinen Volumen hinausfließen (pro Volumen). Dieses Bild werden wir beim Gaußschen Integralsatz weiter ausbauen.

5.8 Integraltheoreme

Der Fundamentalsatz der Analysis

$$\int_a^b dx \left(\frac{d}{dx} F(x) \right) = F(b) - F(a) \quad (5.37)$$

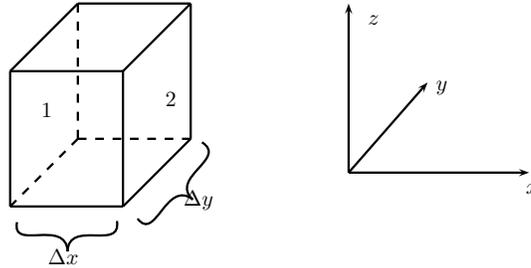
verbindet die Werte von $F(x)$ an den Rändern des Intervalls $[a, b]$ mit dem Integral von $\frac{dF}{dx}$ über das Intervall $[a, b]$ (s. 2.6.1). In höheren Dimensionen existieren Verallgemeinerungen dieses Satzes, die z.B. Integrale über ein Volumen einer Ableitung mit einem Integral über den Rand dieses Volumens (eine Oberfläche) verbinden. Dabei spielen Differentialoperatoren wie die Rotation und die Divergenz die Rolle der Ableitung auf der linken Seite. Wichtige Anwendungen dieser Integralsätze liegen in der Elektro- und Magnetostatik.

5.8.1 Der Gaußsche Integralsatz

Wir betrachten ein Vektorfeld in 3 Dimensionen in kartesischer Basis und kartesischen Koordinaten

$$\mathbf{v}(\mathbf{r}) = \begin{pmatrix} v_x(x, y, z) \\ v_y(x, y, z) \\ v_z(x, y, z) \end{pmatrix} \quad (5.38)$$

$\mathbf{v}(\mathbf{r})$ könnte zum Beispiel ein Strömungsfeld sein, wobei $\mathbf{v} \cdot \Delta \mathbf{S}$ angibt, wie viele Teilchen pro Zeit durch ein kleines Flächenelement $\Delta \mathbf{S}$ strömen. Wir betrachten nun einen (kleinen) Quader mit Kantenlänge $\Delta x, \Delta y, \Delta z$.



Der Netto-Fluss in x -Richtung aus dem Quader heraus ist

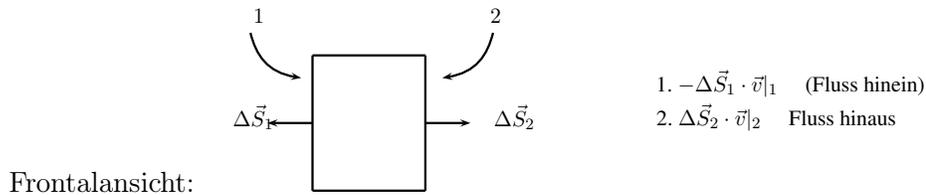
$$(v_x|_2 - v_x|_1) \Delta y \Delta z \simeq \frac{\partial v_x}{\partial x} \Delta x \Delta y \Delta z$$

In y - und z -Richtung ist der Netto-Fluss:

$$\begin{aligned} \frac{\partial v_y}{\partial y} \Delta x \Delta y \Delta z & \quad \text{durch die hintere und vordere Wand} \\ \frac{\partial v_z}{\partial z} \Delta x \Delta y \Delta z & \quad \text{durch obere und untere Wand} \end{aligned}$$

$$\left(\frac{\partial v_x}{\partial x} + \frac{\partial v_y}{\partial y} + \frac{\partial v_z}{\partial z} \right) \Delta x \Delta y \Delta z = \nabla \cdot \mathbf{v} \Delta x \Delta y \Delta z \quad (5.39)$$

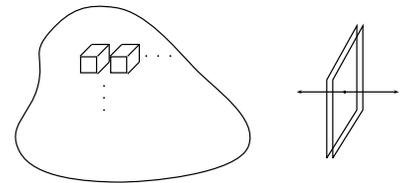
ist der gesamte Netto-Fluss aus dem Quader heraus. Dieser Netto-Fluss lässt sich auch als Oberflächenintegral schreiben.



Der Netto Fluss in x -Richtung aus dem Quader hinaus ist also $\Delta \mathbf{S}_1 \cdot \mathbf{v}|_1 + \Delta \mathbf{S}_2 \cdot \mathbf{v}|_2$. Summiert man die Beiträge von den 6 Seiten des kleinen Quaders erhält man den Netto Fluss aus dem Quader hinaus $\int d\mathbf{S} \cdot \mathbf{v}$ und damit

$$\nabla \cdot \mathbf{v} \Delta x \Delta y \Delta z = \int_S d\mathbf{S} \cdot \mathbf{v}(\mathbf{r}) \quad \text{für kleine Quader} \quad (5.40)$$

Für ein Volumen V , das sich in viele (kleine) Quader zerlegen lässt, summieren wir dieses Ergebnis für einen Quader über die vielen Teilquader. Im Inneren des Volumens heben sich die Oberflächenintegrale auf der rechten Seite gegenseitig weg und so erhalten wir den **Gaußschen Integralsatz**



$$\boxed{\int_V dV \nabla \cdot \mathbf{v}(\mathbf{r}) = \int_{\partial V} d\mathbf{S} \cdot \mathbf{v}}, \quad (5.41)$$

wobei $S = \partial V$ die Oberfläche beschreibt, die das Volumens V begrenzt. Dieser Satz lässt hat eine einfache geometrische Interpretation. Hat ein Vektorfeld \mathbf{v} eine Quelle innerhalb des Volumens V , dann führt diese Quelle zu einem Netto-Fluss aus dem Volumen heraus. Liegen mehrere Quellen innerhalb des Volumens, addieren sich ihre Beiträge zum Netto-Fluss. Liegt die Quelle aber außerhalb von V , führt der Fluss nach V hinein und wieder hinaus, trägt also nicht zum Nettofluss bei.

Beispiel

Der Gaußsche Integralsatz findet wichtige Anwendungen in der klassischen Feldtheorie, zum Beispiel in der Elektrostatik und in der Gravitationstheorie. Das Gravitationsfeld $\mathbf{g}(\mathbf{r}) \equiv \mathbf{F}(\mathbf{r})/m$ hat keine Quellen, sondern sogenannte **Senken**, die durch Massen entstehen: $\nabla \cdot \mathbf{g}$ ist also proportional zur Massendichte $\rho(\mathbf{r})$. Das Gaußsche Gravitationsgesetz postuliert $\nabla \cdot \mathbf{g} = 4\pi G\rho(\mathbf{r})$. Die Schwerkraft im Abstand r zu einer Punktmasse am Ursprung lässt sich leicht herleiten, indem wir ein kugelförmiges Volumen V mit Radius r um den Ursprung betrachten. Das Volumenintegral über die Massendichte $\rho(\mathbf{r})$ ist

$$\begin{aligned} M &= \int_V dV \rho(\mathbf{r}) = -\frac{1}{4\pi G} \int_V dV \nabla \cdot \mathbf{g}(\mathbf{r}) \\ &= -\frac{1}{4\pi G} \int_{\partial V} d\mathbf{S} \cdot \mathbf{g}(\mathbf{r}) = -\frac{1}{4\pi G} \int_{\partial V} dS g(r) = -\frac{1}{4\pi G} 4\pi g(r) r^2. \end{aligned} \quad (5.42)$$

Im zweiten Schritt haben wir das Gaußsche Gravitationsgesetz genutzt, im dritten den Gaußschen Integralsatz. Zur Auswertung des Oberflächenintegrals haben wir genutzt, dass das Gravitationsfeld einer Punktmasse radial nach innen zeigt (Kugelsymmetrie des Problems) und seine Norm auf ∂V konstant ist, d.h. $d\mathbf{S} \cdot \mathbf{g} = dS g(r)$. Damit erhalten wir die bekannte Beziehung $g(r) = -GM/r^2$.

Diese Methode Felder auszurechnen stellt sich als sehr mächtig heraus, wenn die den Feldern zugrundeliegende Masseverteilung eine Symmetrie aufweist. Erstes Beispiel ist eine kugelsymmetrische Masseverteilung $\rho(\mathbf{r}) = \rho(r)$, wir erwarten ein kugelsymmetrisches Feld $g(\mathbf{r}) = g(r)\hat{\mathbf{r}}$. Wir betrachten wieder ein kugelförmiges Volumen V mit Radius r um den Ursprung und berechnen das Volumenintegral über die Massendichte

$$M_V = \int_V dV \rho(r) = -\frac{1}{4\pi G} \int_{\partial V} d\mathbf{S} \cdot \mathbf{g}(\mathbf{r}) = -\frac{1}{4\pi G} 4\pi g(r) r^2. \quad (5.43)$$

mit denselben Schritten wie bei dem Beispiel mit der Punktmasse. Wir erhalten $g(r) = -GM_V/r^2$, wobei M_V nun die von V eingeschlossene Masse ist! Das Gravitationsfeld $g(r)$ einer kugelsymmetrischen Masseverteilung ist also gleich dem Feld einer Punktmasse, die so gross ist wie die innerhalb einer Kugel mit Radius r eingeschlossene Masse. Die Masse ausserhalb r trägt dahingegen nicht zum Gravitationsfeld bei. Wenn wir das Gravitationsfeld einer Kugel berechnen, können wir also die Kugel als Punktmasse behandeln ohne einen Fehler zu machen, selbst wenn der Abstand von der Kugeloberfläche nicht gross ist (im Vergleich zum Radius der Kugel).

Ebenso lässt sich das Gravitationsfeld *im Inneren* einer massiven Kugel mit Radius R und konstanter Dichte berechnen. Wir betrachten nun ein Volumen V mit Radius $r < R$ und erhalten $M_V = \frac{4}{3}\pi r^3$ und mit (5.43) $g(r) = -\frac{4}{3}\pi G r$, das Feld steigt also innerhalb der Kugel linear an. Auch das Gravitationsfeld einer Hohlkugel lässt sich auf diese Weise schnell bestimmen: ein Volumen V im Inneren der Hohlkugel schliesst keine Masse ein, das Feld ist also null.

Diese Methode findet auch in der Elektrostatik breite Anwendung; die erste Maxwellglei-

chung $\nabla \cdot \mathbf{E} = \rho_E(r)/\epsilon_0$ besagt, dass elektrische Ladungen Quellen für das elektrische Feld $\mathbf{E}(\mathbf{r})$ sind. Diese Gleichung ist von derselben Form wie Gaußsche Gravitationsgesetz. Bei Ladungsverteilungen oder Masseverteilungen, die einer bestimmten Symmetrie unterliegen, lassen sich oft Volumina/Oberflächen finden, so dass sich das Feld durch den Gaußschen Integralsatz leicht berechnen lässt. Man bezeichnet diese Oberflächen als **Gaußsche Volumina/Oberflächen**.

Info: Der Vollständigkeit halber noch eine Herleitung des Gaußschen Gravitationsgesetzes. Wir beginnen mit einem Massepunkt M im Ursprung und einem Volumen V , das den Ursprung enthält und auf dessen Oberfläche jeder Punkt mit einer geraden Linie vom Ursprung erreicht werden kann, ohne dass die Linie vorher die Oberfläche schneidet (ein sogenanntes sternartiges Volumen). Das Oberflächenintegral

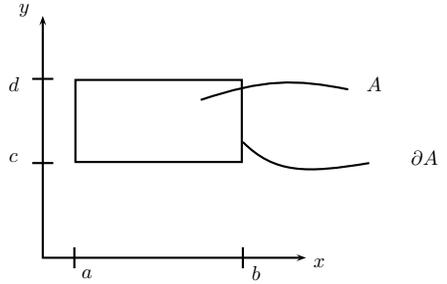
$$\int_{\partial V} d\mathbf{S} \cdot \mathbf{g}(\mathbf{r}) = -GM \int_{\partial V} d\mathbf{S} \cdot \hat{\mathbf{r}}/r^2 \quad (5.44)$$

erscheint nun sehr viel schwieriger zu berechnen, da das Feld $\mathbf{g}(\mathbf{r})$ nicht mehr auf der Oberfläche ∂V konstant ist. Ausserdem zeigen die vektoriellen Flächenelemente $\Delta\mathbf{S}$ nicht mehr radial nach aussen. Doch auch dieses Integral lässt sich leicht berechnen. Wir nutzen Kugelkoordinaten: die Oberfläche ∂V ist dann durch eine Funktion $r(\theta, \phi)$ beschrieben, die angibt, wie weit Punkte auf der Oberfläche in einer bestimmten Richtung vom Ursprung entfernt sind. Bei gegebenen Koordinaten θ, ϕ beschreibt dann $\Delta\theta$ und $\Delta\phi$ ein Flächenelement, dessen Projektion in radialer Richtung $\Delta\mathbf{S} \cdot \hat{\mathbf{r}}$ gleich dem Flächeninhalt $r^2 \sin(\theta)\Delta\theta\Delta\phi$ ist. Integration über θ und ϕ gibt für das Oberflächenintegral $-4\pi GM$. Addieren wir weitere Punktmassen trägt jede Punktmasse $-4\pi GM$ zum Oberflächenintegral bei, da das Volumen beliebig gewählt war muss die Masse auch nicht im Ursprung liegen. Für eine beliebige Masseverteilung erhalten wir $\int_{\partial V} d\mathbf{S} \cdot \mathbf{g}(\mathbf{r}) = -4\pi G \int_V dV \rho(\mathbf{r})$. Aus dem Gaußschen Integralsatz folgt $\int_V dV \nabla \cdot \mathbf{g} = -4\pi G \int_V dV \rho(\mathbf{r})$, da dieses Ergebnis für beliebige (sternförmige) Volumina gilt müssen die Integranden gleich sein, also $\nabla \cdot \mathbf{g} = -4\pi G \rho(\mathbf{r})$.

Beschreibt man das Feld durch ein Potential, $\mathbf{g} = -\nabla\phi$, so erhält man die sogenannte **Poissongleichung** $\Delta\phi \equiv \nabla \cdot \nabla\phi = -4\pi G\rho$. Dabei haben wir aus den beiden Differentialoperatoren ∇ und $\nabla \cdot$ einen neuen Differentialoperator definiert. Der sogenannte **Laplace-Operator** Δ bildet ein skalares Feld auf ein skalares Feld ab, in dem er zunächst den Gradienten eines skalaren Feldes generiert und dann die Divergenz dieses Gradientenfeldes bildet. In kartesischen Koordinaten und drei Dimensionen ist $\Delta\phi = (\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2})\phi(x, y, z)$. Der Laplace-Operator tritt in der klassischen Feldtheorie auf, aber auch als Diffusionsoperator in der Wärmeleitung. Er spielt auch eine zentrale Rolle in der Quantenmechanik.

5.8.2 Das Greensche Theorem in der Ebene

Wir beschränken uns zunächst auf 2 Dimensionen, um einen einfachen Integralsatz herzuleiten, der Oberflächenintegrale mit Linienintegralen um den Rand der Oberfläche verbindet. Wir betrachten zwei Funktionen von x und y , $P(x, y)$ und $Q(x, y)$, mit kontinuierlichen partiellen Ableitungen. Diese Funktionen sind frei gewählt und ändern ihren Wert als Funktion von x, y . Auf einer besonders einfachen Fläche, einem Rechteck A , werden wir nun ein konkretes Oberflächenintegral über A mit einem konkreten Linienintegral über den Rand dieses Rechtecks ∂A vergleichen.



Oberflächenintegral:

$$\iint_A dx dy \frac{\partial Q}{\partial x} = \int_a^b dx \int_c^d dy \frac{\partial Q}{\partial x} = \int_c^d dy [Q(b, y) - Q(a, y)] \quad (5.45)$$

Wegintegral:

$$\oint_{\partial A} dy Q(x, y) = \int_c^d dy Q(b, y) + \int_d^c dy Q(a, y) = \int_c^d dy [Q(b, y) - Q(a, y)]$$

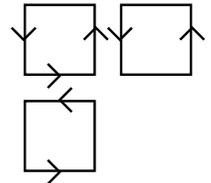
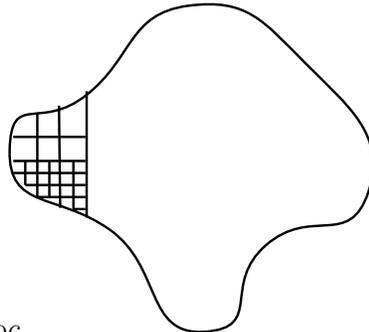
($\oint_{\partial A}$ bezeichnet das Integral über die geschlossene Kurve ∂A .)

$$\leadsto \iint_A dx dy \frac{\partial Q}{\partial x} = \oint_{\partial A} dy Q(x, y) \quad \text{für das gewählte Rechteck} \quad (5.46)$$

$$\text{Analog}^2 : - \iint_A dx dy \frac{\partial P}{\partial y} = \oint_{\partial A} dx P(x, y) \quad (5.47)$$

$$\leadsto \iint_A dx dy \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y} \right) = \oint_{\partial A} (dx P(x, y) + dy Q(x, y)) \quad (5.48)$$

Dieses Ergebnis gilt natürlich zunächst nur für rechteckige Flächen. Betrachten wir aber eine ebene Fläche A beliebiger Form (ohne Löcher), die in eine Menge (beliebig kleiner) Rechtecke zerlegt werden kann. Summieren wir (5.46) über all diese Rechtecke, so heben sich Beiträge der Linienintegrale *aus dem Inneren* von A gegenseitig weg. Die Summe von (1) über alle Rechtecke, in die A zerlegt wurde ergibt dann



auf der linken Seite ein Oberflächenintegral, auf der rechten Seite erhalten wir ein Linienintegral über den Rand von A . Unter milden Bedingungen (die Kurve die den Rande von A beschreibt, darf sich nicht schneiden) erhalten wir

$$\boxed{\int_A dx dy \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y} \right) = \oint_{\partial A} (dx P(x, y) + dy Q(x, y))} \quad (5.49)$$

Dieses Ergebnis heißt Greensches Theorem in der Ebene.

Beispiel

Betrachten wir als konkretes Beispiel die Funktionen $Q(x, y) = x$ und $P(x, y) = 0$.

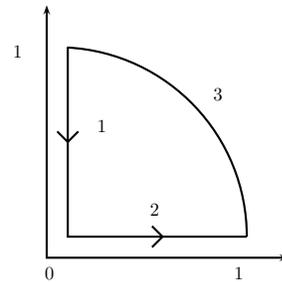
$$\int_A dx dy \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y} \right) = \iint_A dx dy$$

$$\oint_{\partial A} (dx P(x, y) + dy Q(x, y)) = \oint_{\partial A} dy x$$

Dieses Ergebnis kann genutzt werden, um die Berechnung des Flächeninhaltes von A zu ersetzen durch die Berechnung eines Linienintegrals. Auf diesem Prinzip basieren alte mechanische Apparate zur Bestimmung von Flächeninhalten (<http://de.wikipedia.org/wiki/Planimeter>). Als Beispiel berechnen wir das entsprechende Linienintegral um den Viertelkreis mit Radius eins.

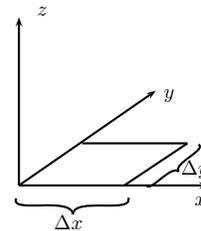
- Teilstück 1: $x = 0 \curvearrowright$ kein Beitrag zum Integral
- Teilstück 2: $\frac{dy}{dt} = 0 \curvearrowright$ kein Beitrag zum Integral

- Teilstück 3: $y(\theta) = \sin \theta$
 $x(\theta) = \cos \theta$
 $\int_0^{\pi/2} d\theta \frac{dy(\theta)}{d\theta} x(\theta) = \int_0^{\pi/2} d\theta \cos^2 \theta = \pi/4$



5.8.3 Satz von Stokes

Betrachten wir wieder ein rechteckiges Flächenelement und legen die z -Achse senkrecht zu seiner Oberfläche. Wir betrachten ein Vektorfeld $\mathbf{v}(\mathbf{r})$ und nutzen das Greensche Theorem mit $Q(x, y) = v_y(x, y, z = 0)$ und $P(x, y) = v_x(x, y, z = 0)$. ($z = 0$ ist konstant auf dem Rechteck.)



$$\iint_A dx dy \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y} \right) = \iint_A dx dy \left(\frac{\partial v_y}{\partial x} - \frac{\partial v_x}{\partial y} \right) = \int_A d\mathbf{S} \cdot \nabla \times \mathbf{v} \quad (5.50)$$

z -Komponente von $\nabla \times \mathbf{v}$.

Für das Linienintegral über den Rand von A erhalten wir

$$\begin{aligned} \oint_{\partial A} (dx P(x, y) + dy Q(x, y)) &= \oint_{\partial A} dx v_x(x, y, z) + dy v_y(x, y, z) + dz v_z(x, y, z) \\ &= \oint_{\partial A} d\mathbf{r} \cdot \mathbf{v}(\mathbf{r}). \end{aligned} \quad (5.51)$$

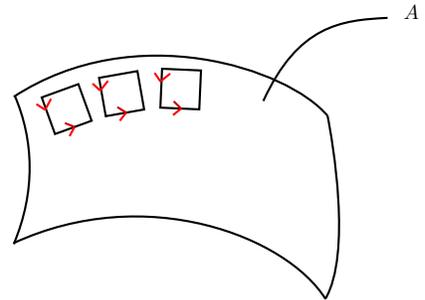
Der letzte Schritt folgt aus $dz/dt = 0$ für jede Parametrisierung des Rechtecks. Aus dem Greenschen Theorem folgt für dieses Rechteck

$$\int_A d\mathbf{S} \cdot (\nabla \times \mathbf{V}) = \oint_{\partial A} d\mathbf{r} \cdot \mathbf{v}(\mathbf{r}). \quad (5.52)$$

Als Beziehung zwischen Vektoren gilt dieser Ausdruck unabhängig von der Wahl der kartesischen Basis (auch wenn wir zur Herleitung eine konkrete Basis genutzt haben, in der das Rechteck in der xy -Ebene liegt).

Für eine Oberfläche, die man in (viele) kleine Rechtecke zerlegen kann, heben sich die Beiträge kleiner Rechtecke im Inneren von A gegenseitig weg. Aus der Summe von (5.52) über diese Rechtecke erhalten wir damit den **Satz von Stokes** (math.: Stokes-Kelvin-Theorem)

$$\boxed{\int_A d\mathbf{S} \cdot (\nabla \times \mathbf{v}) = \oint_{\partial A} d\mathbf{r} \cdot \mathbf{v}(\mathbf{r})}. \quad (5.53)$$



Analog zum Gaußschen Integralsatz findet der Satz von Stokes Anwendung in der Magnetostatik.

Info: Mit dem Satz von Stokes können wir auch den Zusammenhang zwischen Vektorfeldern $\mathbf{v}(\mathbf{r})$ mit Rotation $\nabla \times \mathbf{v} = 0$ und Gradientenfeldern zeigen. Für ein Vektorfeld mit verschwindender Rotation sind nach (5.53) Linienintegrale um eine geschlossenen Kurve null. Damit sind auch Linienintegrale vom Weg unabhängig: Wir teilen eine geschlossene Kurve C in ein Teilstück C_1 von Punkt p_1 zu p_2 und ein Teilstück C_2 von p_2 zu p_1 und schreiben das Linienintegral von \mathbf{v} entlang einer Linie C als I_C . Aus $I_{C_1} + I_{C_2} = I_C = 0$ folgt dann $I_{C_1} = -I_{C_2}$. Drehen

wir nun die Integrationsrichtung in einem der Teilstücke um, erhalten wir aus zwei unterschiedlichen Integrationswegen von p_1 nach p_2 dasselbe Ergebnis, Linienintegrale sind also wegunabhängig. Mit diesem Ergebnis lässt sich auch $f(\mathbf{r})$ konstruieren, so dass $\mathbf{v} = \nabla f$: $f(\mathbf{r}) = \int_{C(p_0, p_0 + \mathbf{r})} d\mathbf{r} \cdot \mathbf{v}(\mathbf{r})$. Dabei beschreibt $C(p_0, p_0 + \mathbf{r})$ eine beliebige Kurve vom Ursprung zum Punkt $p_0 + \mathbf{r}$. Nach dem obigen Ergebnis hängt das so definierte skalare Feld nur vom Anfangs- und Endpunkt dieser Kurve ab. $f(\mathbf{r})$ ist nur bis auf eine additive Konstante bestimmt, die durch die Wahl des Ursprungs festgelegt wird.

Randnotiz: Dass Vektorfelder mit Rotation null Gradientenfelder sind gilt für **einfach zusammenhängenden Räume** (vereinfacht: Räume ohne Löcher wie z.B. \mathbb{R}^3 ; genauer Räume in denen für alle Punktpaare ein Weg existiert der sie verbindet, und jeder geschlossene Weg auf einen Punkt zusammengezogen werden kann). In nicht einfach zusammenhängenden Räumen kann nämlich noch im Inneren einer berandeten Oberfläche ein weiterer Rand existieren, auf dem das Linienintegral ggf. nicht verschwindet.