

Theoretische Physik I
Punktmechanik und mathematische
Methoden

Version WS 2018/2019 — Tilman Plehn
Original von Matthias Bartelmann

24. Januar 2019

Zuerst ein paar organisatorische Informationen:

- **stellen Sie Fragen!**
- sollten Sie Werbung für die Fachschaft oder anderes studentisches Engagement machen wollen, dann melden Sie sich gerne bei mir.
- Informationen zum Beispiel zur Organisation der Vorlesung und der Übungen gibt es in der Vorlesung.
- die ersten beiden Übungen in den Vorlesungswochen 2 und 3 sind Präsenzübungen.
- Übungsblätter können gerne in Kleingruppen von bis zu drei Personen bearbeitet werden. Die Lösungen werden einzeln abgegeben, und müssen in der Übungsgruppe vorgerechnet werden können.
- 60% der Übungspunkte sind die Voraussetzung für die Klausurzulassung.
- die beste Vorbereitung für die Klausur sind die Übungsblätter.
- die letzte Vorlesung findet am 24.1. statt, die Woche des 28.1. dient der Klausurvorbereitung (study week).
- erste Klausur am Dienstag, 11.12. 11:00-13:00 (33% der Endnote), zweite Klausur am Dienstag 5.2. 10:00-13:00 (67% der Endnote), beide Teile werden zusammen bewertet.
- zur Klausur können Sie ein doppelseitig beschriebenes oder bedrucktes DIN-A4-Blatt mitnehmen, weitere Hilfsmittel werden nicht benötigt. Wenn es Sie beruhigt, dann können Sie auch gerne einen nicht programmierbaren Taschenrechner mitbringen, Sie werden ihn aber nicht brauchen.
- es gibt eine Nachklausur am Ende der Semesterferien. Sie deckt beide Teile der ersten Klausur ab. Wer zum Beispiel aus Krankheitsgründen nur einen Teil der ersten Klausur(en) schreiben konnte melde sich bitte bei Ramon Winterhalder.

Herzlichen Dank an viele Studentinnen und Studenten, die das Skript kommentiert und korrigiert haben und damit sehr dazu beigetragen haben, es zu verbessern und verständlicher zu machen!

Inhaltsverzeichnis

1	Statistik	1
1.1	Wahrscheinlichkeitsverteilungen	1
1.1.1	Unsicherheiten	1
1.1.2	Daten vs Theorie	4
1.2	Satz von Bayes	5
1.2.1	Bayesische Wahrscheinlichkeit	5
1.2.2	Volumeneffekte	7
1.2.3	Profile-Likelihood	9
2	Die Newtonschen Axiome	12
2.1	Einführung	12
2.2	Newtonsche Axiome	14
2.3	Differentialgleichungen I	15
2.3.1	Definition und Klassifikation	15
2.3.2	Lösung: Trennung der Veränderlichen	16
2.3.3	Lösung: Variation der Konstanten	17
2.4	Fallbewegungen	19
2.4.1	Freier Fall aus geringer Höhe	19
2.4.2	Fall aus geringer Höhe mit Stokesscher Reibung	20
2.4.3	Fall aus geringer Höhe mit Luftwiderstand . .	22
2.4.4	Freier Fall aus großer Höhe	24
3	Mathematische Grundlagen	26
3.1	Differentialgleichungen II	26

3.1.1	Existenz und Eindeutigkeit von Lösungen . . .	26
3.1.2	Lineare Differentialgleichungen n -ter Ordnung	28
3.2	Vektoren I	30
3.2.1	Vektorräume	31
3.2.2	Lineare Abhängigkeit, Basis und Dimension .	33
3.2.3	Beispiele	33
3.3	Taylor-Reihen	34
3.3.1	Taylorsche-Satz	35
3.3.2	Einfache Beispiele	36
4	Impuls, Drehimpuls und Energie	38
4.1	Vektoren II	38
4.1.1	Das Skalarprodukt	38
4.1.2	Summenkonvention, Levi-Civita-Symbol . . .	39
4.1.3	Das Vektorprodukt	40
4.1.4	Koordinaten und Koordinatentransformationen	41
4.1.5	Inertialsysteme	43
4.2	Impuls, Drehimpuls und Energie	43
4.2.1	Impuls	44
4.2.2	Drehmoment und Drehimpuls	44
4.2.3	Energiesatz in einer Dimension	45
5	Bewegung in drei Dimensionen	48
5.1	Kinematik in drei Dimensionen	48
5.1.1	Bahnkurven	48
5.1.2	Bogenlänge, Tangential- und Normalvektoren .	49
5.1.3	Tangential- und Normalkomponenten	51
5.1.4	Kurvenintegrale	52
5.2	Differentialoperatoren	53
5.2.1	Felder, Partielle Ableitungen und der Gradient	53
5.2.2	Divergenz und Rotation	55

5.3	Energiesatz in drei Dimensionen	56
5.3.1	Energieerhaltung bei Potentialkräften	56
5.3.2	Beispiele für Potentialkräfte	57
5.4	Konservative Kraftfelder	58
5.4.1	Unabhängigkeit vom Weg	58
5.4.2	Der Satz von Stokes	60
5.5	Krummlinig-orthogonale Koordinaten	62
5.5.1	Zylinderkoordinaten	63
5.5.2	Sphärische Polarkoordinaten	64
6	Stöße und Streuung	65
6.1	Systeme von Massenpunkten	65
6.1.1	Bewegung des Schwerpunkts	65
6.1.2	Drehimpuls	66
6.1.3	Energie	67
6.1.4	Reduzierte Masse	68
6.2	Stöße und Streuung	68
6.2.1	Elastischer Stoß zwischen zwei Teilchen	68
6.2.2	Transformation der Streuwinkel	69
6.2.3	Energieübertrag bei elastischer Streuung	70
7	Koordinatentransformationen	72
7.1	Darstellung durch Matrizen	72
7.1.1	Drehungen im dreidimensionalen Raum	72
7.1.2	Matrizen und Matrixoperationen	73
7.1.3	Beispiele	74
7.2	Determinanten und Matrixinversion	75
7.2.1	Determinanten	75
7.2.2	Orthonormale Transformationen	76
7.2.3	Inversion von Matrizen	77
7.3	Physikalische Bedeutung	79

7.3.1	Transformation des Drehimpulses	79
8	Kräfte in beschleunigten Bezugssystemen	82
8.1	Zeitabhängige Transformationen	82
8.1.1	Winkelgeschwindigkeit	82
8.1.2	Infinitesimale Transformationen	84
8.2	Bewegung auf der rotierenden Erde	85
8.2.1	Scheinkräfte	85
8.2.2	Corioliskraft	87
9	Bewegung starrer Körper	89
9.1	Tensoren	89
9.1.1	Der Trägheitstensor	89
9.1.2	Tensoren als lineare Abbildungen	90
9.1.3	Transformationseigenschaften	90
9.1.4	Tensorprodukt	91
9.2	Diagonalisierung	92
9.2.1	Diagonalisierung von Matrizen	92
9.2.2	Diagonalisierung des Trägheitstensors	94
10	Trägheitstensor	96
10.1	Volumenintegrale	96
10.1.1	Motivation	96
10.1.2	Jacobi-Determinante	97
10.1.3	Trägheitstensor einer Kugel	99
10.2	Der Satz von Gauß	100
10.2.1	Herleitung	100
10.2.2	Transformation von Divergenz und Rotation	101
11	Harmonische Schwingungen	104
11.1	Harmonischer Oszillator	104
11.1.1	Linearisierte Kraft	104

11.1.2	Freie Schwingungen	105
11.1.3	Komplexe Darstellung	106
11.1.4	Hermitesche Matrizen	106
11.2	Pendel	107
11.2.1	Mathematisches Pendel	107
11.2.2	Physikalisches Pendel	108
11.2.3	Das Foucaultsche Pendel	109
11.3	Erzwungene Schwingungen, Resonanz	110
11.3.1	Periodischem Antrieb	110
11.3.2	Resonanz und Halbwertsbreite	112
11.3.3	Grenzfall schwacher Dämpfung	113
11.3.4	Allgemeine Lösung	113
11.3.5	Beispiel: Schwingung nach einem Kraftstoß	114

Kapitel 1

Statistik

Wahrscheinlichkeitsrechnungen sind in der Physik zentral, sobald wir ein Experiment auswerten, denn zu jedem experimentellen Ergebnis gehört nicht nur der zentrale Messwert, sondern vor allem auch eine Fehlerabschätzung. Wenn man Folgerungen aus einer Messung ziehen möchte, dann sind die Fehlerbalken mindestens so wichtig wie der Zentralwert, denn wir stellen immer dieselbe Frage: *Ist dieses Ergebnis im Rahmen aller Unsicherheiten in Übereinstimmung mit der Theorie-Vorhersage, oder müssen wir die zugrundeliegende Theorie modifizieren?* In diesem Kapitel betrachten wir zwei Aspekte der Statistik physikalischer Messungen. Zuerst widmen wir uns den Wahrscheinlichkeitsverteilungen für verschiedene Fehlerquellen. Im zweiten Schritt beschreiben wir Messungen, die von verschiedenen Fehlerquellen beeinflusst werden. Dabei müssen wir zwischen einem Frequentist-Zugang und einer Bayesischen Wahrscheinlichkeit unterscheiden.

1.1 Wahrscheinlichkeitsverteilungen

1.1.1 Unsicherheiten

- Für jede physikalische Messung kombiniert man drei Arten von Unsicherheiten, wenn man den Fehlerbalken konstruiert. Der wohl bekannteste Beitrag ist die *statistische Unsicherheit*. Wir beginnen mit einem Würfelexperiment, von dem wir wissen, dass das Ergebnis durch eine *Binomialverteilung* beschrieben wird. Dies beruht auf dem Postulat, dass alle Würfelversuche gleichartig und voneinander unabhängig sind. In diesem Fall ist die Wahrscheinlichkeit n Ereignisse in N Versuchen bei einer einzelnen konstanten Wahrscheinlichkeit \bar{n}/N zu beobachten

$$P(n|\bar{n}) = \frac{N!}{n!(N-n)!} \left(\frac{\bar{n}}{N}\right)^n \left(1 - \frac{\bar{n}}{N}\right)^{N-n}. \quad (1.1)$$

In der Physik gehen wir davon aus, dass wir Präzisionsexperimente sehr oft wiederholen können, wobei die positiven Ergebnisse nicht häufig sein müssen, $N \gg n, \bar{n}$. In diesem Grenzfall kann man die Binomialverteilung umformen und erhält

$$\begin{aligned}
P(n|\bar{n}) &= \frac{\bar{n}^n}{n!} \lim_{N \rightarrow \infty} \frac{N!}{(N-n)!} \frac{1}{N^n} \left(1 - \frac{\bar{n}}{N}\right)^{N-n} \\
&= \frac{\bar{n}^n}{n!} \lim_{N \rightarrow \infty} \frac{(N-n+1) \cdots N}{N^n} \left(1 - \frac{\bar{n}}{N}\right)^{-n} \left(1 - \frac{\bar{n}}{N}\right)^N \\
&= \frac{\bar{n}^n}{n!} \lim_{N \rightarrow \infty} \frac{N^n}{N^n} 1^{-n} \left(1 - \frac{\bar{n}}{N}\right)^N \\
&= \frac{\bar{n}^n}{n!} \lim_{N \rightarrow \infty} \left(1 - \frac{\bar{n}}{N}\right)^N \\
&= \frac{\bar{n}^n}{n!} \lim_{N \rightarrow \infty} \exp\left[\ln\left(1 - \frac{\bar{n}}{N}\right)^N\right] \\
&= \frac{\bar{n}^n}{n!} \lim_{N \rightarrow \infty} \exp\left[N \ln\left(1 - \frac{\bar{n}}{N}\right)\right] \\
&= \frac{\bar{n}^n}{n!} \lim_{N \rightarrow \infty} \exp\left[N\left(-\frac{\bar{n}}{N}\right)\right] \quad (\text{Taylor-Reihe, erster Term}) \\
&= \frac{\bar{n}^n}{n!} e^{-\bar{n}} \tag{1.2}
\end{aligned}$$

Dies ist die *Poisson-Verteilung* für die Wahrscheinlichkeit, einen Wert n zu messen, wenn wir \bar{n} erwarten. Sie ist zunächst für alle natürlichen Zahlen $n \geq 0$ definiert. Negative Zahlen ergeben in der Poisson-Verteilung ebenso wenig Sinn wie in einem Zählexperiment. Mathematisch ist für jede Art von Wahrscheinlichkeitsverteilung die Normierung zentral. Sie bedeutet, dass ein System einen der möglichen Messwerte annehmen muss,

$$\sum_{n=0}^{\infty} P(n|\bar{n}) = e^{-\bar{n}} \sum_{n=0}^{\infty} \frac{\bar{n}^n}{n!} = e^{-\bar{n}} e^{\bar{n}} = 1 \tag{1.3}$$

Dazu haben wir lediglich die Reihendarstellung der Exponentialfunktion benötigt.

- In der Praxis ist auch die Definition der Poisson-Verteilung unhandlich, und es ist nicht klar warum wir große Werte von N aber kleine Werte von n und \bar{n} betrachten wollten. Im Grenzwert großer Zahlen von positiven Messungen $n \gg 1$ können wir zunächst die Stirling-Formel

$$n! \approx \sqrt{2\pi n} \left(\frac{n}{e}\right)^n \tag{1.4}$$

benutzen und die Poisson-Verteilung auf positive reelle Zahlen verallgemeinern. Wenn wir Event-Zahlen nahe Null vermeiden und für große Event-Zahlen $n = \bar{n}(1 + \delta)$ oder $\delta = (n - \bar{n})/\bar{n} \lesssim 1$

definieren, dann ergibt sich für die Poisson-Verteilung

$$\begin{aligned}
 P(n|\bar{n}) &= \frac{\bar{n}^{\bar{n}(1+\delta)}}{\sqrt{2\pi\bar{n}(1+\delta)}} \left(\frac{e}{\bar{n}(1+\delta)} \right)^{\bar{n}(1+\delta)} e^{-\bar{n}} \\
 &= \frac{1}{\sqrt{2\pi\bar{n}}} \frac{\bar{n}^{\bar{n}(1+\delta)}}{\sqrt{1+\delta}} \frac{e^{\bar{n}\delta}}{\bar{n}^{\bar{n}(1+\delta)}(1+\delta)^{\bar{n}(1+\delta)}} \\
 &= \frac{1}{\sqrt{2\pi\bar{n}}} \frac{e^{\bar{n}\delta}}{(1+\delta)^{\bar{n}(1+\delta)+1/2}} \\
 &\approx \frac{1}{\sqrt{2\pi\bar{n}}} \frac{e^{\bar{n}\delta}}{(1+\delta)^{\bar{n}(1+\delta)}} \\
 &= \frac{e^{\bar{n}\delta}}{\sqrt{2\pi\bar{n}}} \exp\left[\ln(1+\delta)^{-\bar{n}(1+\delta)}\right] \\
 &= \frac{e^{\bar{n}\delta}}{\sqrt{2\pi\bar{n}}} \exp\left[(-\bar{n}(1+\delta))\ln(1+\delta)\right] \\
 &= \frac{e^{\bar{n}\delta}}{\sqrt{2\pi\bar{n}}} \exp\left[(-\bar{n}-\bar{n}\delta)\left(\delta-\frac{\delta^2}{2}+\dots\right)\right] \\
 &\approx \frac{e^{\bar{n}\delta}}{\sqrt{2\pi\bar{n}}} \exp\left[-\bar{n}\delta-\bar{n}\delta^2+\frac{\bar{n}\delta^2}{2}\right] \\
 &= \frac{1}{\sqrt{2\pi\bar{n}}} e^{-(n-\bar{n})^2/(2\bar{n})} \tag{1.5}
 \end{aligned}$$

Hier haben wir die Taylor-Reihe bis zum zweiten Term entwickelt, eine genaue Definition folgt noch im Laufe der Vorlesung. Im letzten Schritt erhalten wir die *Gauss-Verteilung* oder Normalverteilung im Grenzfall hoher Statistik $N > \bar{n}, n \gg 1$. Wichtig ist auch in diesem Fall die Normierung der Gauss-Verteilung

$$\frac{1}{\sqrt{2\pi\bar{n}}} \int_{-\infty}^{\infty} dn e^{-(n-\bar{n})^2/(2\bar{n})} = 1. \tag{1.6}$$

Die Gauss-Verteilung ist für alle reellen Werte von n definiert und symmetrisch um den Erwartungswert \bar{n} . Ihre Standardabweichung ist $\sigma = \sqrt{\bar{n}}$.

- Neben dem für kleine oder große Werte wohldefiniertem statistischen Fehler gibt es experimentell auch *systematische Fehler*. Sie kommen zum Beispiel aus der Eichung des Versuchsaufbaus oder aus Fehlern auf andere Messgrößen, die für unsere eigentliche Messung benötigt werden. Nur wenn Systematiken mit anderen Messungen im Grenzfall großer Statistik bestimmt werden, dann kann man für sie auch eine Gauss-Verteilung annehmen. Im Idealfall kann man zum Beispiel einen Input-Parameter mit seiner gesamten gemessenen Wahrscheinlichkeitsverteilung angeben.
- Schliesslich gibt es auch für jede theoretische Vorhersage einen *Theorie-Fehler*, weil wir kein komplexes System präzise berechnen können. Beispiele sind in diesem Fall eine Taylorreihe um

eine kleine Störung oder statistische Unsicherheiten in rein numerischen Berechnungen. Das Problem mit theoretischen Fehlern ist, dass wir nicht wissen wie wahrscheinlich zum Beispiel eine Abweichung von 10% im Vergleich zu einer Abweichung um 20% ist. Streng genommen wissen wir zwar, dass bei einer zu großen Diskrepanz der Messung von der Vorhersage die Theorie als widerlegt betrachtet werden, können diese Grenze aber nicht wissenschaftlich präzise festlegen. Mit anderen Worten: Theoretische Fehler können mit unseren statistischen Werkzeugen nicht beschrieben werden. Ein weit verbreiteter Zugang ist es, die Wahrscheinlichkeitsverteilung des Theorie-Fehler für eine Observable zentral als flach anzunehmen und jenseits einer akzeptierten Grenze guten Geschmacks komplett abzuschneiden. Auf diese Weise nehmen wir an, dass anders als bei statistischen Prozessen keine kleine Wahrscheinlichkeit besteht, dass eine Theorie eine beliebig schlechte Vorhersage liefert.

1.1.2 Daten vs Theorie

Nachdem wir die Wahrscheinlichkeitsverteilung der verschiedenen Fehler verstanden haben, stellt sich die Frage, wie man Daten mit Vorhersagen vergleicht. Wir diskutieren kurz zwei verschiedene experimentelle Situationen für ein Zählexperiment, in dem wir die komplette Wahrscheinlichkeitsverteilung $P(n)$ kennen.

- Beginnen wir mit dem Fall, dass Experiment und Theorie gut übereinstimmen. Mit dieser Information können wir zum Beispiel alternative Theorien ausschliessen, die zum Beispiel eine höhere Zählrate vorhersagen. Wir beginnen also mit der *gemessenen Verteilung* $P(n)$ und berechnen den Anteil der Wahrscheinlichkeitsverteilung, die mit der alternativen Theorie und ihrer vorhergesagten Rate n_T verträglich wäre. Dazu integrieren wir die Wahrscheinlichkeit zwischen n_T und noch größeren Zahlen. Wenn zum Beispiel die Bedingung

$$\int_{n_T}^{\infty} dn P(n) < 0.05 \quad \text{mit} \quad \int_{-\infty}^{\infty} dn P(n) = 1 \quad (1.7)$$

an dieses Integral der erlaubten Zählraten erfüllt ist, dann ist die alternative Theorie nur zu 5% erlaubt oder entsprechend zu 95% confidence level ausgeschlossen. Hierbei spielen die Details der widerlegten Zählrate um n_T keine Rolle.

- Alternativ können wir den Fall betrachten, dass eine Vorhersage mit der Messung nicht übereinstimmt. Zum Beispiel beruht die Entdeckung eines neuen Teilchens wie des Higgs-Teilchens

darauf, dass die gängige Theorie ohne dieses neue Teilchen nicht mit der Messung übereinstimmt. Wir betrachten also die *hypothetische Wahrscheinlichkeitsverteilung* im Untergrundmodell und berechnen, wie wahrscheinlich eine gemessene und zu große Ereigniszahl n_E im Rahmen des Untergrundmodells ist, also

$$p_0 = \int_{n_E}^{\infty} dn P(n) . \quad (1.8)$$

In der Teilchenphysik hat man sich darauf geeinigt, dass für eine Entdeckung die Bedingung $p_0 < 3 \times 10^{-7}$ erfüllt ist. Das entspricht im Gauss-Limes fünf Standardabweichungen. In diesem Test spielen die Details des Higgs-Teilchens keine Rolle, es ist lediglich die Vorhersage n_E ohne das Higgs-Teilchen ausgeschlossen.

- Schliesslich können wir mit Hilfe des p_0 -Wertes noch eine weitere Komplikation verstehen: Wenn man zum Beispiel in einer Verteilung einer gemessenen Größe nach einer Anomalie sucht, dann untersucht man die gesamte Kurve. Wenn man eine Kurve in 10 Bereiche unterteilt und diese Bereiche nacheinander untersucht, dann gibt es eine höhere Wahrscheinlichkeit, zum Beispiel eine statistische Fluktuation zu beobachten. In erster Näherung ist diese kombinierte Wahrscheinlichkeit das zehnfache der einfachen Messung. Um eine globale Wahrscheinlichkeit von fünf Standardabweichungen zu erhalten muss man einen lokalen Wert $p_0 < 3 \times 10^{-8}$ unterschreiten. Dieser Effekt heißt *look-elsewhere effect*, wobei er eigentlich look-anywhere effect heißen sollte.

1.2 Satz von Bayes

Im vorigen Kapitel haben wir gezeigt, die man mit Hilfe einer bekannten Wahrscheinlichkeitsverteilung Messungen mit theoretischen Vorhersagen vergleichen kann. Das Problem in der Praxis ist, dass die Wahrscheinlichkeitsverteilung $P(n)$ nicht nur eine der drei statistischen Verteilungen ist. Stattdessen ist sie eine Kombination mehrerer Quellen von statistischen, systematischen und theoretischen Unsicherheiten. Wir müssen also die Frage beantworten, wie wir in einem komplexen System verschiedene Quellen von Unsicherheiten kombinieren können.

1.2.1 Bayesische Wahrscheinlichkeit

- Als Basis für die folgende, formale Diskussion von Wahrscheinlichkeiten beginnen wir mit den drei Kolmogorov-Axiomen: Das erste Axiom besagt, dass jede Wahrscheinlichkeit für alle Ergebnisse eine nicht negative reelle Zahl ist, also $P(A) \in \mathbf{R}$ und

$P(A) \geq 0$. Das zweite Axiom besagt, dass die Wahrscheinlichkeit alles möglichen Ausgänge kombiniert eins ist. Das dritte Axiom besagt, dass sich Wahrscheinlichkeiten für disjunkte Ausgänge addieren, $P(A \cup B) = P(A) + P(B)$.

- Bevor wir uns mit der Frage befassen können, wie man tatsächlich die Wahrscheinlichkeit einer Theorie mit gegebenen Daten berechnen kann, benötigen wir sogenannte bedingte Wahrscheinlichkeiten. Wir betrachten zwei verschiedenen Messungen und definieren die bedingte Wahrscheinlichkeit für ein Ergebnis A unter der Annahme, dass wir auch B beobachten als

$$P(A|B) := \frac{P(A \cap B)}{P(B)} \quad \Leftrightarrow \quad P(A|B) P(B) = P(A \cap B) . \quad (1.9)$$

Weil $A \cap B = B \cap A$ vertauscht ist $P(A \cap B) = P(B \cap A)$ und damit die umgekehrte bedingte Wahrscheinlichkeit

$$P(B|A) = \frac{P(A \cap B)}{P(A)} . \quad (1.10)$$

Diese Formeln sind je nach Geschmack Definitionen oder Axiome, die allerdings aus den elementaren Kolmogorov-Axiomen motiviert. Zwei Messungen A und B sind statistisch unabhängig, wenn für ihre bedingte Wahrscheinlichkeit $P(A \cap B) = P(A)P(B)$ gilt. Diese Relation sollte man allerdings in der Physik nur annehmen, nachdem man sie sorgfältig getestet hat.

- Der Satz von Bayes besagt, dass die bedingte Wahrscheinlichkeit für T gegeben M durch

$$P(T|M) = \frac{P(M|T) P(T)}{P(M)} \quad (1.11)$$

gegeben ist. Dies folgt einfach aus der Definition der bedingten Wahrscheinlichkeit. Wichtig sind für die Anwendungen einige Definitionen: $P(T|M)$ nennt man *posterior probability* oder im folgenden einfach Wahrscheinlichkeit; $P(T)$ und $P(M)$ heißen jeweils *prior probability* oder kurz Prior; wenn man $P(M|T)$ als Funktion vom zweiten Argument T betrachtet, dann heißt es *Likelihood*. Das Kernproblem mit der Likelihood ist, dass sie als Funktion von T nicht normiert ist und sich damit grundlegenden Annahmen der Wahrscheinlichkeitstheorie entzieht.

- Wenn man in (1.11) T mit der Theorie-Vorhersage und M mit den Messungen identifiziert, dann steht auf der linken Seite die Größe, die uns in der Physik eigentlich interessiert: Wie groß ist nach einem Experiment M die Wahrscheinlichkeit, dass eine Theorie T die Messung korrekt beschreibt. Der Satz von Bayes gibt uns

auch die formal korrekte Definition für diese Wahrscheinlichkeit und damit die Größen, die wir nun bestimmen müssen. Die Likelihood $P(M|T)$ kann man zum Beispiel aus einer Simulation eines Experiments erhalten — wir untersuchen alle nach irgendeinem Maß vernünftigen oder relevanten Theorien T und erhalten für jedes T eine Vorhersage für M . Weiterhin können wir uns eventuell behelfen, indem wir den Prior $P(M)$ aus der Normierungsbedingung an unsere gewünschte Wahrscheinlichkeit erhalten. Das Problem bleibt dann der Prior $P(T)$, also ein Maß für Wahrscheinlichkeit im Theorie-Raum. Aus dem Beweis des Satzes von Bayes ist klar, dass man nur totale Wahrscheinlichkeiten $P(T \cap M)$ identifizieren kann, wir uns aber für eine bedingte Wahrscheinlichkeit interessieren und daher die Definition eines Priors nicht vermeiden können.

Streng genommen sind diese Definitionen von Ausschluss und Entdeckung in (1.7) und (1.8) im Frequentist-Zugang definiert, denn die zugrundeliegenden Verteilungen $P(n)$ beschreiben das statistisch reproduzierbare Verhalten einer großen Zahl von Experimenten unter der Annahme einer Theorie, kurz $P(M|T)$.

1.2.2 Volumeneffekte

- Der Satz von Bayes (1.11) lässt uns zwei Möglichkeiten, Theorie und Messungen mathematisch korrekt zu verknüpfen. Im Bayes-Zugang definieren wir die $P(T|M)$ als den offensichtlich *degree of belief*, womit uns nichts übrig bleibt als einen Prior oder Mass über den Wahrscheinlichkeitsraum der Theorie zu definieren. Im Frequentist-Zugang erlauben wir keine Größen, die sich nicht durch eine Verteilung einer großer Zahl experimenteller Ergebnisse definieren lassen. In diesem Fall muss uns die Likelihood $P(M|T) = L(T)$ als *degree of belief* genügen. Sie definiert zum Beispiel den Punkt im Theorie-Raum, der am besten mit der Messung übereinstimmt.
- Den Unterschied zwischen diesen beiden Zugängen kann man anhand eines einfachen Beispiels illustrieren. Betrachten wir einen Becher voll Wasser einerseits und eine Wolke aus Wasserdampf andererseits. Die Wassermenge in der Wolke ist wesentlich größer, aber die Wasserdichte wesentlich geringer.

Die zunächst unspezifische physikalische Frage ist, wo wir wahrscheinlich ein Wassermolekül finden werden. Nach dem Satz von Bayes suchen wir nach dem Objekt, in dem sich ein Wassermolekül wahrscheinlicher befindet. Dies ist offensichtlich die Wolke. Ein Frequentist könnte lediglich fragen, an welchem Ort sich ein Wassermolekül am wahrscheinlichsten befindet, weil die Definition von Wolke oder Wasserbecher eine Frage der Interpretation

und nicht der Messung ist. Die Antwort ist dann der Wasserbecher mit seiner größeren Wasserdichte. Man kann sich nun trefflich streiten, wer die korrekte Antwort gibt, aber man kann sich auch daran erinnern, dass es sich um zwei Antworten auf zwei verschiedene Fragen handelt. In diesem Fall unterscheiden sie sich um einen sogenannten Volumeneffekt, den man in Bayes-Zugang berücksichtigt (oder erzeugt).

- Dieselbe Fragen nach einem Maß, oder besser nach einem Integrationsmaß stellt sich auch in einem anderen Zusammenhang. Nehmen wir an, dass wir zwei Quellen von Unsicherheiten miteinander kombinieren wollen, und dass wir beide individuellen Verteilungen kennen. Uns interessiert aber nur eine der beiden Verteilungen, weil sie zum Beispiel eine Likelihood als Funktion einer physikalisch interessanten Größe ist, der zweite Parameter hingegen intern die Eichung des experimentellen Aufbaus beschreibt. Ein solcher, für das physikalische Ergebnis nicht relevanter Parameter heißt *Nuisance-Parameter* und wird am Ende aus dem experimentellen Ergebnis eliminiert. Wir tun die zum Beispiel durch Integration mit dem als im Bayes-Zugang immer bekannten oder postulierbaren Maß. Wenn beide Verteilungen Gauss-Form haben, dann ergibt sich zunächst ohne Berücksichtigung der Normierung

$$\begin{aligned}
 P(n) &\propto \int_{-\infty}^{\infty} dm e^{-m^2/(2\sigma_1^2)} e^{-(n-m)^2/(2\sigma_2^2)} \\
 &= \int_{-\infty}^{\infty} dm \exp \left[-\frac{m^2(\sigma_1^2 + \sigma_2^2) - 2nm\sigma_1^2 + n^2\sigma_1^2}{2\sigma_1^2\sigma_2^2} \right] \\
 &= \int_{-\infty}^{\infty} dm \exp \left[-\frac{\sigma_1^2 + \sigma_2^2}{2\sigma_1^2\sigma_2^2} \left(m^2 - 2nm\frac{\sigma_1^2}{\sigma_1^2 + \sigma_2^2} + n^2\frac{\sigma_1^2}{\sigma_1^2 + \sigma_2^2} \right) \right] \\
 &= \int_{-\infty}^{\infty} dm \exp \left[-\frac{\sigma^2}{2\sigma_1^2\sigma_2^2} \left(\left(m - n\frac{\sigma_1^2}{\sigma^2} \right)^2 - n^2\frac{\sigma_1^4}{\sigma^4} + n^2\frac{\sigma_1^2}{\sigma^2} \right) \right] \\
 &\hspace{15em} (1.12)
 \end{aligned}$$

An dieser Stelle verschieben wir die Integrationskonstante zu $m' = m - n\sigma_1^2/\sigma^2$ und definieren $\sigma^2 = \sigma_1^2 + \sigma_2^2$. Damit erhal-

ten wir für das Integral

$$\begin{aligned}
 P(n) &\propto \int_{-\infty}^{\infty} dm' \exp \left[-\frac{\sigma^2}{2\sigma_1^2\sigma_2^2} \left(m'^2 - n^2 \frac{\sigma_1^4}{\sigma^4} + n^2 \frac{\sigma_1^2}{\sigma^2} \right) \right] \\
 &= \int_{-\infty}^{\infty} dm' \exp \left[-\frac{\sigma^2}{2\sigma_1^2\sigma_2^2} \left(m'^2 - n^2 \frac{\sigma_1^4 - \sigma_1^2\sigma^2}{\sigma^4} \right) \right] \\
 &= \int_{-\infty}^{\infty} dm' \exp \left[-\frac{\sigma^2}{2\sigma_1^2\sigma_2^2} \left(m'^2 + n^2 \frac{\sigma_1^2\sigma_2^2}{\sigma^4} \right) \right] \\
 &= e^{-n^2/(2\sigma^2)} \int_{-\infty}^{\infty} dm' e^{-m'^2\sigma^2/(2\sigma_1^2\sigma_2^2)} \\
 &= \frac{\sigma_1\sigma_2}{\sqrt{2\pi\sigma}} e^{-n^2/(2\sigma^2)} \tag{1.13}
 \end{aligned}$$

Wenn wir die Normierungskonstanten der beiden originalen Gauss-Verteilungen $1/(\sqrt{2\pi}\sigma_i)$ einsetzen, dann sehen wir, dass die Kombination zweier Gauss-Verteilung wieder eine solche ist. Die Breiten der Gauss-Verteilungen addieren sich im Quadrat.

- In der Herleitung von (1.13) haben wir angenommen, dass wir über m integrieren und bei dieser Integration ein flaches Integrationsmaß haben. Eine Annahme dieser Art müssen wir machen, um die Integration ausführen zu können. Wir könnten stattdessen auch über $d \log m$ integrieren und aus der Substitution ein Integrationsmaß $d \log m / dm = 1/m$ erhalten. Eine solche Variation des Priors oder Integrationsmaßes sollte Teil jeder Analyse von Bayes-Wahrscheinlichkeiten sein, um die Stabilität des Ergebnisses gegenüber veränderten Annahmen zu überprüfen.
- Wenn wir statt mit zwei Gauss-Verteilung mit beliebigen anderen Verteilungen beginnen, dann sagt der zentrale Grenzwertsatz, dass im Grenzfall vieler solcher Verteilungen wieder eine Gauss-Verteilung entsteht. Der einzige Weg, die Addition der Breiten im Quadrat zu entgehen ist, die beiden Parameter m und n nicht als unkorreliert zu betrachten. In diesem Fall können wir bei einer vollständigen Korrelation zum Beispiel eine lineare Addition der beiden Breiten erreichen.

1.2.3 Profile-Likelihood

- Im Frequentist-Zugang existiert kein Integrationsmaß, um einen Nuisance-Parameter auszuintegrieren. Um trotzdem eine Likelihood um eine Dimension zu reduzieren greifen wir auf einen Trick zurück: Wir projizieren den größten Wert der Likelihood als Funktion des Nuisance-Parameters auf den verbleibenden, physikalischen Parameter. Auf diese Weise können wir immer noch verschiedene Werte des physikalischen Parameters miteinander

vergleichen und insbesondere den wahrscheinlichsten Wert ablesen. Nehmen wir wie in (1.13) zunächst an, dass beide individuellen Likelihoods eine Gauss-Verteilung haben,

$$\begin{aligned}
 L(n) &= \max_m e^{-m^2/(2\sigma_1^2)} e^{-(n-m)^2/(2\sigma_2^2)} \\
 &= \max_m \exp \left[-\frac{m^2(\sigma_1^2 + \sigma_2^2) - 2nm\sigma_1^2 + n^2\sigma_1^2}{2\sigma_1^2\sigma_2^2} \right] \\
 &= \max_m \exp \left[-\frac{\sigma_1^2 + \sigma_2^2}{2\sigma_1^2\sigma_2^2} \left(m^2 - 2nm\frac{\sigma_1^2}{\sigma_1^2 + \sigma_2^2} + n^2\frac{\sigma_1^2}{\sigma_1^2 + \sigma_2^2} \right) \right] \\
 &= \max_m \exp \left[-\frac{\sigma^2}{2\sigma_1^2\sigma_2^2} \left(\left(m - n\frac{\sigma_1^2}{\sigma^2} \right)^2 - n^2\frac{\sigma_1^4}{\sigma^4} + n^2\frac{\sigma_1^2}{\sigma^2} \right) \right] \\
 &= \max_{m'} \exp \left[-\frac{\sigma^2}{2\sigma_1^2\sigma_2^2} \left(m'^2 - n^2\frac{\sigma_1^4 - \sigma_1^2\sigma_2^2}{\sigma^4} \right) \right] \\
 &= \max_{m'} \exp \left[-\frac{\sigma^2}{2\sigma_1^2\sigma_2^2} \left(m'^2 + n^2\frac{\sigma_1^2\sigma_2^2}{\sigma^4} \right) \right] \\
 &= e^{-n^2/(2\sigma^2)} \max_{m'} e^{-m'^2\sigma^2/(2\sigma_1^2\sigma_2^2)} = e^{-n^2/(2\sigma^2)}, \tag{1.14}
 \end{aligned}$$

wobei wir wie vorher $m' = m - n\sigma_1^2/\sigma^2$ und $\sigma^2 = \sigma_1^2 + \sigma_2^2$ definieren. Dieses Ergebnis ist identisch mit dem Ergebnis für die Integration im Bayes-Zugang. Allerdings ist in der Profile-Likelihood die Normierung verlorengegangen, die im Bayes-Zugang erhalten wurde.

- Interessant wird es, wenn wir im Extremfall nun zwei Verteilungen mit einer Box-Form kombinieren. Wie nehmen an, dass beide Boxen die Breite $(n_+ - n_-)$ haben. Wenn wir diese Eigenschaft als Θ -Funktion schreiben, dann wird aus der Bedingung $m = n_- \cdots m_+$ mathematisch eine Kombination aus $m < n_+$ und $m > n_-$ oder

$$L \propto \Theta(n_+ - m)\Theta(m - n_-). \tag{1.15}$$

Wenn wir wie vorher beide Gauss-Kurven kombinieren und dann das Maximum über m bilden, dann erhalten wir

$$\begin{aligned}
 L(n) &= \max_m \Theta(n_+ - m) \Theta(m - n_-) \Theta(n_+ - (n - m)) \Theta((n - m) - n_-) \\
 &= \max_{m \in [n_-, n_+]} \Theta((n_+ + m) - n) \Theta(n - (n_- + m)) \\
 &= \Theta(2n_+ - n) \Theta(n - 2n_-). \tag{1.16}
 \end{aligned}$$

Dies ist eine Box der doppelten Breite $2(n_+ - n_-)$, die Profile-Likelihood hat also für zwei unkorrelierte Box-Verteilungen die Breiten linear addiert.

- Aus dieser Diskussion können wir sofort ableiten, in welchem Grenzfall der Bayes-Zugang mit dem Frequentist-Zugang identisch ist: die Verteilungen müssen Gauss-Form haben und, um

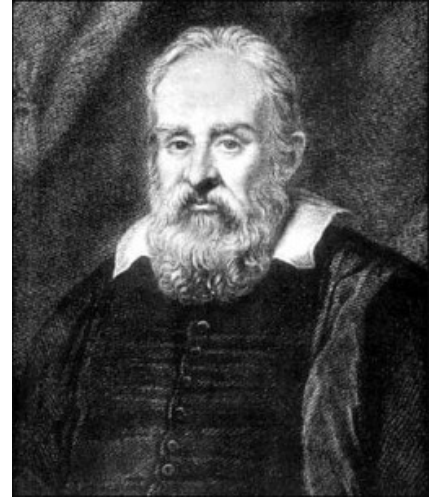
Volumeneffekte zu vermeiden, eine kleine Breite haben. Oder mit anderen Worten, der Bayes-Zugang und der Frequentist-Zugang liefern nur im Grenzfall perfekter Statistik dasselbe Ergebnis.

Kapitel 2

Die Newtonschen Axiome

2.1 Einführung

- Physik ist eine Erfahrungswissenschaft, die Vorgänge in der (meist unbelebten) Natur zu quantifizieren und auf Gesetzmäßigkeiten zurückzuführen sucht. Theoretische Physik sucht die Einheit hinter der Vielfalt, die möglichst fundamentalen Gesetze, die der Vielfalt der Erfahrungstatsachen zugrunde liegen. Diese Gesetze nehmen die Form mathematischer Gleichungen an, in denen die mathematischen Symbole semantisch an die Stelle physikalischer Größen treten. Anders als mathematische Ergebnisse müssen die Gesetze der Physik prüfbare Vorhersagen erlauben.
- Physikalischen Gesetzen liegen notwendigerweise Idealisierungen zugrunde, weil wesentliche von unwesentlichen Eigenschaften physikalischer Systeme unterschieden werden müssen. Erst durch geeignete Näherungen wird theoretische Physik überhaupt möglich, weil erst durch Näherungen Abgrenzungen physikalischer Systeme von ihrer Umwelt eingeführt werden können. Die Beschränkung auf wenige Objekte kann, muss aber nicht die Grundlage solcher Näherungen sein.
- Theoretische Mechanik beschreibt die Gesetze, nach denen sich Körper im Raum unter dem Einfluss von Kräften mit der Zeit bewegen. Sie führt zu Begriffen und Methoden, die sich durch die gesamte theoretische Physik ziehen und vor allem für die Quantenmechanik und Quantenfeldtheorie außerordentlich fruchtbar waren. Anwendungen von Quantenfeldtheorie als mathematischem Modell sind Elementarteilchenphysik oder Festkörperphysik.
- Meilensteine in der geschichtlichen Entwicklung der klassischen Mechanik (im Unterschied zur Quantenmechanik) waren Galileo Galileis Fallversuche, Tycho Brahes Messungen der Marsbahn



Galileo Galilei



Isaac Newton

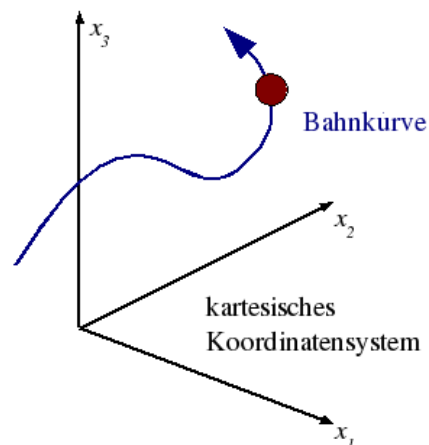
und deren gesetzmäßige Zusammenfassung durch Johannes Kepler sowie Isaac Newtons Axiome und seiner Erklärung der Keplerschen Gesetze durch das Gravitationsgesetz. Weitere für die Entwicklung vor allem der analytischen Mechanik wichtige Personen waren Joseph Lagrange, Leonhard Euler, Jean Baptiste le Rond d'Alembert, William Hamilton und Emmy Noether.

- Die klassische Punktmechanik kennt eine „Vierheit“ von Objekten, nämlich *Körper*, *Kräfte*, *Raum* und *Zeit*.
- *Körper* werden idealisiert als Punkte bestimmter Masse. Ihre Ausdehnung ist sehr klein gegenüber den Dimensionen des gesamten betrachteten Systems, wobei „klein“ oder „groß“ höchst relative Begriffe sind. Zum Beispiel in der Elementarteilchenphysik haben wir es allgemein mit Teilchen zu tun, deren räumliche Ausdehnung wir nicht kennen. Solange wir nur eine obere Grenze zum Beispiel an die Ausdehnung von Elektronen kennen, betrachten wir sie als Punktteilchen. Ein fester Körper wird als System von Massenpunkten aufgefasst, deren Abstände untereinander konstant sind. Ein System aus vielen Massenpunkten bewegt sich im Ganzen so, als würden die äußeren Kräfte an seinem Schwerpunkt angreifen.
- *Kräfte* sind die Ursachen der Änderung einer Bewegung und mathematisch genauer zu definieren. In der klassischen Mechanik wird angenommen, dass Kräfte instantan wirken, d.h. mit unendlicher Ausbreitungsgeschwindigkeit.
- Der *Raum* in der klassischen Mechanik hat drei Raumdimensionen und eine weitere Zeit-Dimension. Die Lage von Körpern in diesem dreidimensionalen Raum hat keinen absoluten Sinn, sondern muss relativ zu anderen Körpern, den Bezugssystemen, angegeben werden. Der physikalische Raum wird als ein reeller, dreidimensionaler Vektorraum aufgefasst. Die momentane Lage eines Massenpunkts im Raum wird durch einen Ortsvektor \vec{x} angegeben. Der Ursprung des Bezugssystems kann beliebig gewählt werden (Homogenität), und seine Achsen können beliebig orientiert werden (Isotropie). Als Bezugssysteme werden wir zunächst kartesische Koordinatensysteme wählen.
- Die *Zeit* spielt in der klassischen Mechanik die Rolle eines unabhängigen Parameters. Ihr Nullpunkt ist im Allgemeinen frei wählbar (Homogenität der Zeit). Raum und Zeit sind in der klassischen Mechanik unabhängig von der Existenz von Körpern und ihrer Bewegung relativ zueinander, sie sind in diesem Sinne absolut. Die Bahnkurve $\vec{x}(t)$ beschreibt, wie sich sein Ort zeitlich ändert. Die Geschwindigkeit \vec{v} ist die Änderung des Ortes mit der

Zeit, dargestellt durch die Zeitableitung des Ortes:

$$\vec{v}(t) = \frac{d\vec{x}}{dt} =: \dot{\vec{x}}(t) . \tag{2.1}$$

Ein Punkt über einem Funktionssymbol bezeichnet die Zeitableitung dieser Funktion.



Bahnkurve in einem kartesischen Koordinatensystem

2.2 Newtonsche Axiome

Newtons Axiome sind die Grundlage aller Rechnungen in der klassischen Mechanik. Sie sind nicht die einzige Art, diese Grundlagen zu formulieren, wie wir später sehen werden, aber sie sind vergleichsweise sehr anschaulich.

1. Trägheitsgesetz, *Lex Prima*: Jeder Körper beharrt in seinem Zustand der Ruhe oder gleichförmigen geradlinigen Bewegung, wenn er nicht durch einwirkende Kräfte gezwungen wird, seinen Zustand zu ändern. ¹

Postuliert wird die Trägheit eines Körpers, sein Beharrungsvermögen. Als „Bewegungsgröße“ wird das Produkt aus Masse m und Geschwindigkeit \vec{v} definiert, d.h. der *Impuls* $\vec{p} = m\vec{v}$. Der Begriff „Masse“ als Proportionalitätskonstante ist an dieser Stelle noch nicht definiert. Damit besagt das Trägheitsgesetz:

$$\vec{p} = m\vec{v} = m\dot{\vec{x}} = \text{konstant} \tag{2.2}$$

in Abwesenheit von Kräften. In diesem Fall ist der Impuls erhalten, wobei auch „Kraft“ noch zu definieren ist.

2. Bewegungsgesetz, *Lex Secunda*: Die Änderung der Bewegung ist der Einwirkung der bewegenden Kraft proportional und geschieht nach der Richtung derjenigen geraden Linie, nach welcher jene Kraft wirkt. ²

Die Änderung der Bewegungsgröße ist die Zeitableitung des Impulses. Sei \vec{F} die Kraft, besagt das Bewegungsgesetz

$$\dot{\vec{p}} = m\ddot{\vec{x}} = \vec{F} \tag{2.3}$$

bei konstanter Masse m . Dieses Axiom besagt, dass Bewegungsgleichungen in der Regel zweite Ableitungen des Ortes nach der Zeit enthalten.

¹Corpus omne perseverare in statu suo quiescendi vel movendi uniformiter in directum nisi quatenus a viribus cogiter statum illum mutare.

²Mutationem motus proportionalem esse vi motrici impressae et fieri secundum lineam rectam, qua vis illa imprimitur.

Kräfte werden definiert, indem man Messvorschriften angibt, die z.B. eine unbekannte Kraft mit der Gravitationskraft vergleichen. Kräfte addieren sich wie Vektoren. Wenn wir die Ortskoordinate kennen, dann ist die Masse m als Proportionalitätskonstante zur Kraft definiert. Sie bezeichnet hier die *träge* Masse, im Gegensatz zur *schweren* Masse, zu der die Gravitationskraft proportional ist.

3. Reaktionsgesetz, *Lex Tertia*: Die Wirkung ist stets der Gegenwirkung gleich, oder die Wirkungen zweier Körper aufeinander sind stets gleich und von entgegengesetzter Richtung.³
4. *Masse und Kräfte*: Offenbar setzen also die Newtonschen Axiome eine geeignete Wahl der Einheiten voraus. Die *Proportionalität* von Impulsänderung und Kraft wird erst durch die geeignete Wahl der Einheit der Masse zu einer Gleichheit. Die *Proportionalität* von schwerer und träger Masse wird zu einer Gleichheit durch die Definition der Gravitationskonstante G . Die träge Masse erweist sich aufgrund der speziellen Relativitätstheorie als geschwindigkeitsabhängig, ist aber in der klassischen Mechanik konstant.

Kräfte hängen im Allgemeinen vom Ort und von der Zeit ab. Einfache Beispiele für Kräfte sind etwa die Gravitations- und die Coulombkraft, die beide indirekt proportional zum Abstandsquadrat sind (das ist eine notwendige Folge der Masselosigkeit der Austauschteilchen). Die Federkraft ist proportional zur Auslenkung einer Feder. Kräfte können aber auch von der Geschwindigkeit abhängen wie etwa die Lorentzkraft auf ein geladenes Teilchen im Magnetfeld oder Reibungskräfte.

2.3 Differentialgleichungen I

2.3.1 Definition und Klassifikation

Gleichungen wie das zweite Newtonsche Bewegungsgesetz verknüpfen Funktionen wie die Bahnkurve $\vec{x}(t)$ mit ihren Ableitungen nach ihrer Variablen. Solche Differentialgleichungen kann man mathematisch beschreiben und oft mit Hilfe von Standardverfahren lösen. Lösungen von Differentialgleichungen sind das Kernproblem der Newtonschen Beschreibung der klassischen Mechanik.

- Differentialgleichungen sind Gleichungen, die Ableitungen einer oder mehrerer Funktionen nach einer oder mehreren Variablen enthalten. *Gewöhnliche* Differentialgleichungen beschrei-

³Actio = reactio.

ben Funktionen einer unabhängigen Variablen, während *partielle* Differentialgleichungen von mehreren unabhängigen Variablen abhängen. Die *Ordnung* einer Differentialgleichung ist der Grad der höchsten auftretenden Ableitung. Wegen Newtons zweitem Axiom führen die Bewegungsgesetze der klassischen Mechanik in der Regel auf gewöhnliche Differentialgleichungen zweiter Ordnung, in denen die unabhängige Variable die Zeit t ist.

- Wenn die unbekannte Funktion $y(x)$ und alle ihre Ableitungen $y'(x)$, $y''(x)$ usw. höchstens linear in einer Differentialgleichung vorkommen, heißt die Gleichung selbst linear. Zum Beispiel ist

$$y''(x) + p(x)y'(x) + q(x)y(x) = 0, \quad (2.4)$$

in der $p(x)$ und $q(x)$ bekannte Funktionen sind, eine gewöhnliche lineare Differentialgleichung zweiter Ordnung in x . In der Newtonschen Mechanik ist der Parameter oft die Zeit t , und Ableitungen werden mit $\dot{y}(t)$ statt mit $y'(x)$ bezeichnet.

- Die Differentialgleichung in (2.4) enthält keine konstanten, von $y(x)$ und seinen Ableitungen unabhängige Terme. Sie heißt daher *homogen*. Inhomogene Differentialgleichungen enthalten solche Konstanten, wie zum Beispiel die Gleichung

$$y''(x) + p(x)y'(x) + q(x)y(x) = r, \quad (2.5)$$

- Das *Richtungsfeld* einer (gewöhnlichen) Differentialgleichung wird dadurch angegeben, dass in jedem Punkt der y - x -Ebene die Steigung $y'(x)$ angegeben wird.

2.3.2 Lösung: Trennung der Veränderlichen

- Beginnen wir mit dem Beispiel des radioaktiven Zerfalls. Die Teilchenzahl $N(t)$ wird durch das Zerfallsgesetz

$$\dot{N}(t) = -\lambda N(t) \quad (2.6)$$

beschrieben, worin $\lambda > 0$ die Zerfallskonstante ist. Diese gewöhnliche, lineare, und homogene Differentialgleichung erster Ordnung lässt sich lösen, indem man die Variablen N und t voneinander trennt,

$$\dot{N} = \frac{dN}{dt} = -\lambda N \quad \Rightarrow \quad \frac{dN}{N} = -\lambda dt. \quad (2.7)$$

Nach der Trennung der Variablen N und t ist es nicht sinnvoll oder hilfreich, N als Funktion der Zeit zu schreiben. Im Endergebnis wird diese Abhängigkeit automatisch wieder auftauchen.

In dieser Form kann die Gleichung auf beiden Seiten integriert werden,

$$\int_{N_0}^{N_1} \frac{dN}{N} = \ln N \Big|_{N_0}^{N_1} = \ln N_1 - \ln N_0$$

$$-\lambda \int_{t_0}^{t_1} dt = -\lambda(t_1 - t_0) \quad (2.8)$$

Wenn man $N_1(t_1)$ als physikalische Observable betrachtet, dann ist die Lösung dieser Gleichung mit bestimmten Integralen

$$\ln N_1(t_1) = -\lambda(t_1 - t_0) + \ln N_0 \quad (2.9)$$

Wenn man die Stammfunktion $N_1(t_1)$ als allgemeine Lösung $N(t)$ betrachtet, dann kann man sie als

$$\ln N(t) = -\lambda t + C \quad \Leftrightarrow \quad N(t) = e^{-\lambda t} e^C, \quad (2.10)$$

schreiben. Die Konstante C ist durch N_0 und t_0 gegeben. Wenn wir statt des bestimmten Integrals zwischen t_0 und t_1 die Stammfunktion des unbestimmten Integrals suchen, dann ist C eine freie Integrationskonstante. In beiden Fällen erlaubt sie uns, eine *Anfangsbedingung* zu stellen: Soll bei $t_0 = 0$ die Anzahl $N = N_0$ sein, so folgt $e^C = N_0$ oder das exponentielle Zerfallsgesetz

$$N(t) = N_0 e^{-\lambda t}. \quad (2.11)$$

Allgemein ist dieses Verfahren bei Gleichungen angebracht, die sich in die Form

$$g(y)y'(x) = f(x) \quad \Leftrightarrow \quad g(y)dy = f(x)dx \quad (2.12)$$

bringen lassen. Im Fall des Zerfallsgesetzes hatten wir $y \rightarrow N$, $g(y) \rightarrow 1/y \rightarrow 1/N$, $x \rightarrow t$, und $f(x) \rightarrow -\lambda$. Dann führt die Separation der Variablen und anschließende Integration über x auf

$$\int g(y)y'(x)dx = \int g(y)dy = \int f(x)dx + C, \quad (2.13)$$

woraus dann implizit die Funktion $y(x)$ bestimmt werden kann.

2.3.3 Lösung: Variation der Konstanten

Das kommende Beispiel ist eine Variation des Zerfallsgesetzes, das mit klassischer Mechanik nichts zu tun hat. Aber es illustriert, wie man mithilfe der „Variation der Konstanten“ inhomogene Differentialgleichungen lösen kann.

- Betrachten wir nun den etwas schwierigeren Fall, in dem Kerne der Sorte A in Kerne der Sorte B und diese dann in stabile Kerne der Sorte C zerfallen. Nach dem einfachen Zerfallsgesetz ist die Anzahl N_A der A -Kerne durch

$$\dot{N}_A(t) = -\lambda_A N_A(t) \quad \Leftrightarrow \quad N_A(t) = N_{A0} e^{-\lambda_A t} \quad (2.14)$$

gegeben. Die Anzahl der Kerne B wird durch ihren Zerfall verringert, aber durch den Zerfall von A -Kernen erhöht,

$$\dot{N}_B(t) = -\lambda_B N_B(t) - \dot{N}_A(t) = -\lambda_B N_B(t) + \lambda_A N_{A0} e^{-\lambda_A t} . \quad (2.15)$$

Wir haben für $N_B(t)$ eine gewöhnliche lineare Differentialgleichung erster Ordnung, die Terme enthält, die nicht von N_B abhängen. Man nennt sie inhomogene Gleichungen und bringt gewöhnlich die homogenen und inhomogenen Teile auf verschiedene Seiten der Gleichung,

$$\dot{N}_B(t) + \lambda_B N_B(t) = \lambda_A N_{A0} e^{-\lambda_A t} . \quad (2.16)$$

- In so einem Fall verschafft man sich zunächst die allgemeine Lösung der homogenen Gleichung, also der Gleichung, die man erhält, wenn man die Inhomogenität auf der rechten Seite gleich Null setzt. Im Fall von (2.16) ist die Lösung das einfache Zerfallsgesetz

$$\dot{N}_B(t) + \lambda_B N_B(t) = 0 \quad \Leftrightarrow \quad N_B(t) = N_{B0} e^{-\lambda_B t} . \quad (2.17)$$

- Eine Lösung der inhomogenen Gleichung kann man sich daraus verschaffen, indem man die Konstante N_{B0} als zeitabhängig auffasst, also $N_{B0} \rightarrow B(t)$ und

$$N_B(t) = B(t) e^{-\lambda_B t} \quad (2.18)$$

$$\dot{N}_B(t) = \dot{B}(t) e^{-\lambda_B t} - B(t)\lambda_B e^{-\lambda_B t} = \dot{B}(t) e^{-\lambda_B t} - \lambda_B N_B(t) . \quad (2.19)$$

Wenn wir annehmen, dass der Ansatz mit $B(t)$ die inhomogene Gleichung löst, dann können wir diese Form in (2.16) einsetzen und eine Differentialgleichung für $B(t)$ erhalten,

$$\left(\dot{B}(t) e^{-\lambda_B t} - \lambda_B N_B(t) \right) + \lambda_B N_B(t) = \lambda_A N_{A0} e^{-\lambda_A t} \quad (2.20)$$

$$\dot{B}(t) = \lambda_A N_{A0} e^{-(\lambda_A - \lambda_B)t} . \quad (2.21)$$

Wir können sie direkt integrieren,

$$B(t) = \frac{\lambda_A}{\lambda_B - \lambda_A} N_{A0} e^{-(\lambda_A - \lambda_B)t} + C , \quad (2.22)$$

wobei wieder die Konstante C durch die Anfangsbedingungen festgelegt werden kann.

- Setzen wir das Ergebnis (2.22) in die Gleichung für N_B in (2.18) ein, so folgt

$$N_B(t) = C e^{-\lambda_B t} + \frac{\lambda_A}{\lambda_B - \lambda_A} N_{A0} e^{-\lambda_A t} . \quad (2.23)$$

Um die Konstante C zu bestimmen, wählen wir die Anzahl der B -Kerne als $N_B(t = 0) = N_{B0}$. Damit erhalten wir

$$N_{B0} = C + \frac{\lambda_A}{\lambda_B - \lambda_A} N_{A0} \quad \Leftrightarrow \quad C = N_{B0} - \frac{\lambda_A}{\lambda_B - \lambda_A} N_{A0} \quad (2.24)$$

und daraus die Lösung für die Differentialgleichung (2.15)

$$N_B(t) = N_{B0} e^{-\lambda_B t} + \frac{\lambda_A}{\lambda_B - \lambda_A} N_{A0} (e^{-\lambda_A t} - e^{-\lambda_B t}) . \quad (2.25)$$

Sie ist, wie alle Lösungen inhomogener Differentialgleichungen, die Summe der allgemeinen homogenen Gleichung (2.17) und einer speziellen Lösung der inhomogenen Gleichung. Man kann sie durch Einsetzen in (2.16) explizit überprüfen.

2.4 Fallbewegungen

2.4.1 Freier Fall aus geringer Höhe

Nach diesem kurzen Ausflug in die Mathematik der Differentialgleichungen können wir in die klassische Mechanik von Kräften und ihren Einfluss auf punktförmige Körper zurückkehren.

- Wir stellen uns eine Punktmasse m vor, die aus einer Höhe r fällt. Wenn die Höhe sehr klein gegenüber dem Erdradius ist, $r \ll R_{\text{Erde}}$, dann ist die Schwerkraft $F_G = -mg$ durch die konstante Erdbeschleunigung $g = 9.81 \text{ m s}^{-2}$ gegeben. Die Bewegungsgleichung lautet einfach

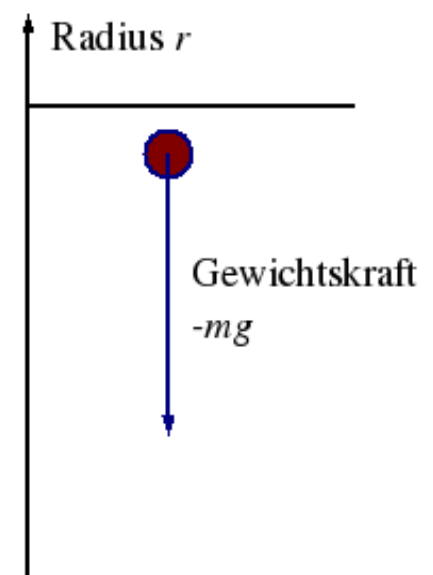
$$m\ddot{r}(t) = m\dot{v}(t) = -mg . \quad (2.26)$$

Das ist eine gewöhnliche, lineare, inhomogene Differentialgleichung zweiter Ordnung, in der allerdings nur die zweite Ableitung vorkommt. Indem wir die Geschwindigkeit $v = \dot{r}$ einführen, erhalten wir eine Differentialgleichung erster Ordnung,

$$\dot{v}(t) = -g , \quad (2.27)$$

deren Lösung sofort durch Integration folgt,

$$v(t) = \dot{r}(t) = -gt + C_1 . \quad (2.28)$$



freier Fall aus geringer Höhe

Die Integrationskonstante C_1 ist die Anfangsgeschwindigkeit $C_1 = v(t = 0) = v_0$. Die Geschwindigkeit, die während des Falls aufgenommen wird, ist negativ, weil unser Koordinatensystem entgegen der Fallbewegung ausgerichtet ist.

- Eine weitere Integration nach der Zeit liefert

$$r(t) = -\frac{g}{2}t^2 + v_0t + C_2, \tag{2.29}$$

wobei die weitere Integrationskonstante C_2 die Anfangshöhe ist, $C_2 = r(t = 0) = r_0$. Da die Bewegungsgleichung von zweiter Ordnung in der Zeit ist, werden zu ihrer vollständigen Lösung zwei Integrationskonstanten benötigt. Diese müssen als Anfangsbedingungen gewählt werden, die in diesem Fall die Bedeutung der Anfangshöhe und der Anfangsgeschwindigkeit haben.

- Wenn der Massenpunkt bei $t = 0$ in der Höhe h losgelassen wird, sind $v_0 = 0$ und $r_0 = h$, und die Lösung lautet

$$v(t) = -gt \quad \text{und} \quad r(t) = h - \frac{g}{2}t^2. \tag{2.30}$$

Die Fallzeit bis $r = 0$ beträgt

$$t_E = \sqrt{\frac{2h}{g}}, \tag{2.31}$$

und die Endgeschwindigkeit ist

$$v_E = -gt_E = -g \sqrt{\frac{2h}{g}} = -\sqrt{2gh}. \tag{2.32}$$

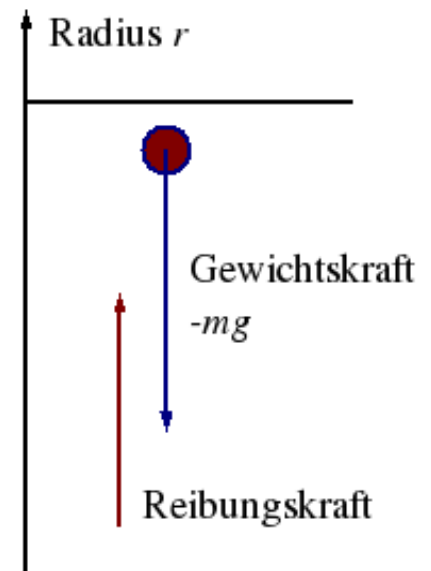
Beispiel: ein Sprung vom Zehnmeterurm dauert $t = 1.4$ s und endet mit einer Geschwindigkeit von $v = -14 \text{ m s}^{-1}$ (etwa -50 km h^{-1}).

2.4.2 Fall aus geringer Höhe mit Stokesscher Reibung

Eine realistischere Beschreibung eines fallenden Körpers beinhaltet mehr als eine Kraft. Zum Beispiel kann man den Luftwiderstand als zweite Kraft berücksichtigen. Wir beginnen erst einmal mit einer Stokesschen Reibungskraft.

- Bei Stokesscher Reibung ist die Reibungskraft proportional zur Geschwindigkeit, $F_R = -K_R v$; sie wirkt der Geschwindigkeit, die wir als positiv gegen die Fallrichtung definiert hatten, entgegen. Die Bewegungsgleichung

$$m\ddot{r}(t) = m\dot{v}(t) = m\frac{dv}{dt} = -mg - K_R v(t) \tag{2.33}$$



gebremster Fall aus geringer Höhe

lässt sich durch Trennung der Variablen v und t lösen:

$$dv = -g \left(1 + \frac{K_R v}{mg} \right) dt$$

$$\int \frac{dv}{1 + \frac{K_R v}{mg}} = \frac{mg}{K_R} \int \frac{dK_R v / (mg)}{1 + K_R v / (mg)} = -g \int dt . \quad (2.34)$$

Beide unbestimmte Integrale lassen sich integrieren,

$$\frac{mg}{K_R} \ln \left(1 + \frac{K_R v}{mg} \right) = -gt + C_1 , \quad (2.35)$$

wobei die beiden Integrationskonstanten gleich zu einer zusammengefasst wurden. Damit lautet die Geschwindigkeit

$$\ln \left(1 + \frac{K_R v(t)}{mg} \right) = \frac{K_R (C_1 - gt)}{mg}$$

$$v(t) = \frac{mg}{K_R} \left(\exp \frac{K_R (C_1 - gt)}{mg} - 1 \right) . \quad (2.36)$$

- Nach sehr langer Zeit, d.h. für $t \rightarrow \infty$, nähert sich v der Endgeschwindigkeit

$$v_E = -\frac{mg}{K_R} , \quad (2.37)$$

an, so dass die Lösung in der Form

$$v(t) = v_E \left(1 - \exp \frac{-C_1 + gt}{v_E} \right) \quad (2.38)$$

geschrieben werden kann. Wählt man weiter die Anfangsgeschwindigkeit, $v(t = 0) = v_0$, dann folgt

$$v_0 = v_E \left(1 - e^{-C_1/v_E} \right) \Rightarrow e^{-C_1/v_E} = 1 - \frac{v_0}{v_E} \quad (2.39)$$

und damit

$$v(t) = v_E \left(1 - \left(1 - \frac{v_0}{v_E} \right) e^{gt/v_E} \right)$$

$$= v_E + (v_0 - v_E) e^{gt/v_E} = (v_0 + |v_E|) e^{-gt/|v_E|} - |v_E| . \quad (2.40)$$

- Diese Lösung ist gleichzeitig eine Differentialgleichung erster Ordnung für die Höhe $r(t)$,

$$\dot{r}(t) = (v_0 + |v_E|) e^{-gt/|v_E|} - |v_E| . \quad (2.41)$$

Eine weitere Integration führt auf

$$r(t) = -(v_0 + |v_E|) \frac{|v_E|}{g} e^{-gt/|v_E|} - |v_E| t + C_2 . \quad (2.42)$$

Setzt man hier $r(t = 0) = r_0$ ein, ergibt sich die Integrationskonstante C_2 zu

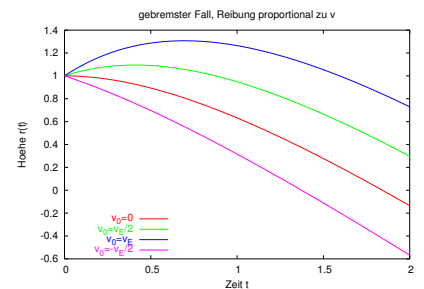
$$C_2 = r_0 + (v_0 + |v_E|) \frac{|v_E|}{g}, \quad (2.43)$$

und damit erhält man die Lösung

$$r(t) = r_0 + (v_0 + |v_E|) \frac{|v_E|}{g} (1 - e^{-gt/|v_E|}) - |v_E|t. \quad (2.44)$$

- Nach sehr langer Zeit, für $t \rightarrow \infty$, verschwindet der Exponentialterm, und r nimmt linear mit der Zeit ab:

$$r(t) \rightarrow r_0 + (v_0 + |v_E|) \frac{|v_E|}{g} - |v_E|t. \quad (2.45)$$



Gebremster Fall mit Stokescher Reibung aus jeweils derselben Höhe, aber mit vier verschiedenen Anfangsgeschwindigkeiten.

2.4.3 Fall aus geringer Höhe mit Luftwiderstand

Im nächsten Schritt ersetzen wir die Stokesche Reibung durch die korrekte Form für den Luftwiderstand.

- Bei Luftwiderstand ist der Betrag der Reibungskraft zum Quadrat der Geschwindigkeit proportional. Die Bewegungsgleichung lautet dann

$$m\ddot{r}(t) = m\dot{v}(t) = m \frac{dv}{dt} = -mg - K_v v(t)|v(t)|. \quad (2.46)$$

Wieder ist die Reibungskraft der Bewegung entgegen gerichtet. Da wir von einem fallenden Körper reden, nehmen wir $v < 0$ oder $|v| = -v$ an. Mit der Definition $v_E^2 = mg/K_v$ folgt nach Trennung der Variablen

$$\begin{aligned} \frac{dv}{dt} &= -g \left(1 - \frac{K_v}{gm} v(t)^2 \right) \\ \int \frac{dv}{1 - \left(\frac{v}{v_E} \right)^2} &= -g \int dt. \end{aligned} \quad (2.47)$$

An dieser Stelle entscheiden wir uns, dass im folgenden $v_E > 0$ und $|v| < v_E$. Bei dem Integral hilft die Partialbruchzerlegung

$$\frac{1}{1 - x^2} = \frac{1}{2} \left(\frac{1}{1 + x} + \frac{1}{1 - x} \right), \quad (2.48)$$

denn sie erlaubt uns zu schreiben

$$\begin{aligned} \int \frac{dv}{2} \left(\frac{1}{1 + \frac{v}{v_E}} + \frac{1}{1 - \frac{v}{v_E}} \right) &= \frac{v_E}{2} \int d\frac{v}{v_E} \left(\frac{1}{\frac{v}{v_E} + 1} - \frac{1}{\frac{v}{v_E} - 1} \right) \\ &= \frac{v_E}{2} \ln \frac{|v + v_E|}{|v - v_E|} \\ &= \frac{v_E}{2} \ln \frac{v_E + v}{v_E - v} = -gt + C_1, \quad (2.49) \end{aligned}$$

Die Integrationskonstante C_1 wird so bestimmt, dass $v = v_0$ bei $t = 0$ ist,

$$C_1 = \frac{v_E}{2} \ln \frac{v_E + v_0}{v_E - v_0}. \quad (2.50)$$

Damit lässt sich die Gleichung für die Geschwindigkeit umformen zu

$$\begin{aligned} \frac{v_E + v}{v_E - v} &= e^{-2gt/v_E} \frac{v_E + v_0}{v_E - v_0} \\ 1 + \frac{v}{v_E} &= e^{-2gt/v_E} \frac{v_E + v_0}{v_E - v_0} \left(1 - \frac{v}{v_E} \right) \\ \frac{v}{v_E} \left(1 + e^{-2gt/v_E} \frac{v_E + v_0}{v_E - v_0} \right) &= e^{-2gt/v_E} \frac{v_E + v_0}{v_E - v_0} - 1 \quad (2.51) \end{aligned}$$

oder

$$v(t) = -v_E \frac{v_E - v_0 - (v_E + v_0) e^{-2gt/v_E}}{v_E - v_0 + (v_E + v_0) e^{-2gt/v_E}}. \quad (2.52)$$

Da wir $v_E > 0$ angenommen haben fallen für $t \rightarrow \infty$ die Exponentialterme weg, und $v \rightarrow -v_E$, d.h. es wird die asymptotische Endgeschwindigkeit im Grenzfall $t \gg v_E/g$ erreicht.

- Eine weitere Integration führt auf die durchfallene Höhe r ,

$$r(t) - r_0 = \int_0^t v(t') dt'. \quad (2.53)$$

Zur Vereinfachung definieren wir

$$x(t) = \frac{2gt}{v_E} \quad a = \frac{v_E + v_0}{v_E - v_0} \quad \Rightarrow \quad \frac{v(x(t))}{v_E} = -\frac{1 - a e^{-x}}{1 + a e^{-x}} \quad (2.54)$$

und erhalten

$$r(t) - r_0 = -\frac{v_E^2}{2g} \int_0^x dx' \frac{1 - a e^{-x'}}{1 + a e^{-x'}}. \quad (2.55)$$

Die Integration führt auf

$$r(t) - r_0 = -\frac{v_E^2}{2g} \left[x' + 2 \ln(1 + ae^{-x'}) \right]_0^x. \quad (2.56)$$

Mit den Definitionen von a und $x(t)$ erhalten wir

$$\begin{aligned} r(t) &= r_0 - \frac{v_E^2}{2g} \left[\frac{2gt}{v_E} + 2 \ln \left(1 + \frac{v_E + v_0}{v_E - v_0} e^{-2gt/v_E} \right) - 2 \ln \left(1 + \frac{v_E + v_0}{v_E - v_0} \right) \right] \\ &= r_0 - v_E t - \frac{v_E^2}{g} \ln \frac{v_E - v_0 + (v_E + v_0) e^{-2gt/v_E}}{v_E - v_0 + v_E + v_0} \\ &= r_0 - v_E t - \frac{v_E^2}{g} \ln \left(\frac{1}{2} - \frac{v_0}{2v_E} + \frac{v_E + v_0}{2v_E} e^{-2gt/v_E} \right) \end{aligned} \quad (2.57)$$

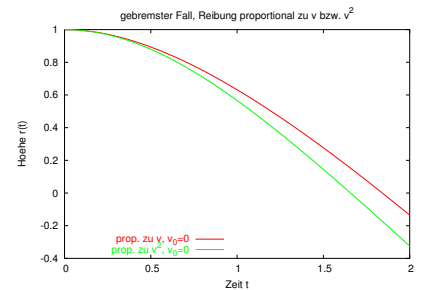
- Wenn wir weiterhin $v_0 = 0$ annehmen, dann erhält man nach genügend langer Zeit, $t \gg v_E/g$, den asymptotischen Verlauf

$$r(t) \rightarrow r_0 - v_E t + \frac{v_E^2}{g} \ln 2, \quad (2.58)$$

Wir können diesen Fall mit dem Ergebnis für Stokesche Reibung nach (2.45) vergleichen, wenn wir annehmen dass wir die Stärke der Reibung so einstellen dass die Endgeschwindigkeit dieselbe ist,

$$r(t) \rightarrow r_0 + \frac{v_E^2}{g} - |v_E|t \quad (2.59)$$

Die beiden Endgeschwindigkeiten hängen von der jeweiligen Stärke der Kraft an. Für einen direkten Vergleich identifizieren wir daher die beiden Endgeschwindigkeiten und vergleichen den Einfluss der Reibung. Wegen $\ln 2 < 1$ liegt der Körper mit Stokescher Reibung in $r(t)$ immer über dem Körper mit Luftwiderstand. Stokesche Reibung ist also effektiver und bremst den freien Fall stärker ab.



Vergleich der durchfallenen Höhen für gebremsten Fall mit Stokescher Reibung (rot) und Luftwiderstand (grün).

2.4.4 Freier Fall aus großer Höhe

Die beiden diskutierten Beispiel nutzen eine Beschreibung der Gravitationskraft, die nur für Körper nahe der Erdoberfläche gilt. Für weiter entfernte Körper hat die Gravitationskraft eine kompliziertere Form als die bisher angenommene lineare Näherung.

- Bei freiem Fall aus großer Höhe, $h \gtrsim R_{\text{Erde}}$, muss berücksichtigt werden, dass sich die Erdbeschleunigung mit der Höhe ändert. Die Gravitationskraft im Abstand r vom Erdmittelpunkt ist

$$F_G = -\frac{GMm}{r^2} \neq -mg \quad (2.60)$$

Wir werden später brauchen dass auch der Erdoberfläche beide Formulierungen der Gravitationskraft dasselbe Ergebnis haben sollten, also $GM/R_{\text{Erde}}^2 = g$. Die allgemeinen Bewegungsgleichung für die allgemeine Form lautet

$$m\ddot{r}(t) = -\frac{GMm}{r(t)^2}. \quad (2.61)$$

Diese lineare Differentialgleichung zweiter Ordnung wandeln wir durch einen *integrierenden Faktor* in eine Differentialgleichung erster Ordnung um. Nach Multiplikation mit \dot{r} erhalten wir

$$\begin{aligned} m\dot{r}(t)\ddot{r}(t) &= -m\frac{GM\dot{r}(t)}{r(t)^2} \\ \frac{m}{2}\frac{d}{dt} \dot{r}(t)^2 &= mGM\frac{d}{dt} \frac{1}{r(t)}. \end{aligned} \quad (2.62)$$

- Wenn wir die beiden Zeitableitungen auf eine Seite bringen, sehen wir, dass die Kombination

$$\frac{m}{2}\dot{r}(t)^2 - \frac{mGM}{r(t)} =: E \quad (2.63)$$

konstant ist, d.h. diese sogenannte „Energie“ E darf sich als Funktion der Zeit nicht ändern. Der erste Ausdruck in Klammern heißt „kinetische“, der zweite „potentielle“ Energie. Die Gleichung (2.63) ist daher die ersten Verwendung des Energiesatzes bei der Lösung eines mechanischen Problems. Die konstante Energie $E = E_0$ ist durch die Anfangsbedingungen gegeben,

$$E_0 = \frac{m}{2}v_0^2 - \frac{mGM}{r_0}. \quad (2.64)$$

Wenn der Massenpunkt im Unendlichen ruht, $r_0 = \infty$ und $v_0 = 0$, ist $E_0 = 0$ und damit für alle Zeiten

$$\frac{1}{2}m\dot{r}(t)^2 = \frac{mGM}{r(t)} = \frac{gR_{\text{Erde}}^2 m}{r(t)}. \quad (2.65)$$

Die Endgeschwindigkeit des freien Falls auf die Erdoberfläche ist durch den Wert für $r(t) = R_{\text{Erde}}$ gegeben, also

$$v_\infty = \sqrt{2gR_{\text{Erde}}} = 11.2 \text{ km s}^{-1}. \quad (2.66)$$

Das ist auch die Fluchtgeschwindigkeit von der Erde, definiert als die Geschwindigkeit, die auf der Erdoberfläche benötigt wird, um das Schwerfeld der Erde zu verlassen.

- Setzt man für v_∞ die Lichtgeschwindigkeit c ein, erhält man den *Schwarzschildradius*

$$r =: R_S = \frac{2GM}{c^2} \quad (2.67)$$

einer Masse M . Er ist für viele Aspekte der relativistischen Astrophysik als „Radius eines schwarzen Lochs“ sehr wichtig.

Kapitel 3

Mathematische Grundlagen

3.1 Differentialgleichungen II

3.1.1 Existenz und Eindeutigkeit von Lösungen

Bevor wir weiter Differentialgleichungen und ihre Lösungen diskutieren, können wir uns kurz der mathematischen Frage nach ihrer Lösbarkeit widmen. In der physikalischen Praxis ist diese Frage typischerweise nicht besonders relevant: Wenn wir ein physikalisches System sinnvoll durch eine Differentialgleichung beschreiben, dann hat diese offenbar eine Lösung, denn ansonsten wäre unser Ansatz unsinnig. Ob diese Lösung mathematisch eindeutig ist, ist in der Newtonschen Mechanik nicht relevant, wenn das entsprechende Experiment einen eindeutigen Ausgang hat. Welche Lösung auch immer das Experiment beschreibt ist die korrekte Lösung. Trotzdem schadet es nicht, wenn wir ein Gefühl für Lösungen von Differentialgleichungen bekommen.

- Zunächst ist die Frage sehr wichtig, unter welchen allgemeinen Bedingungen wir damit rechnen können, Differentialgleichungen überhaupt lösen zu können. Dazu erinnern wir uns zunächst, dass wir Newtons Gleichungen zweiter Ordnung in $r(t)$ in manchen Fällen dadurch gelöst habe, dass wir sie als Differentialgleichung erster Ordnung für $v(t) = \dot{r}(t)$ betrachtet haben. In diesem Sinne lassen sich allgemein Differentialgleichungen n -ter Ordnung als Systeme von Differentialgleichungen erster Ordnung darstellen. Wir fassen dazu die Ableitungen $y^{(i-1)}(x)$ der gesuchten Funktion $y(x)$ als eigene Funktionen $y_i(x)$ auf, also $y(x) = y_1(x)$, $y'(x) = y_2(x)$, $y''(x) = y_3(x)$ usw. Damit wird z.B. die Differentialgleichung zweiter Ordnung aus (2.4), nämlich

$$y''(x) + p(x)y'(x) + q(x)y(x) = 0 ,$$

zum Gleichungssystem erster Ordnung

$$\begin{aligned} y_1'(x) &= y_2(x) \\ y_2'(x) + p(x)y_2(x) + q(x)y_1(x) &= 0. \end{aligned} \quad (3.1)$$

Wir werden uns dementsprechend im Folgenden auf Differentialgleichungen erster Ordnung beschränken können.

- Für Differentialgleichungen erster Ordnung,

$$y'(x) = f(x, y), \quad (3.2)$$

besagt nun die *Lipschitz-Bedingung*, dass eine eindeutige Lösung in einem Intervall I genau dann existiert, wenn für ein $x_0 \in I$ ein Anfangswert

$$y_0 = y(x_0) \quad (3.3)$$

vorgegeben ist, sowie für jedes $x \in I$ und jedes betrachtete Paar von Funktionswerten y, \bar{y} die *Sekantensteigung* durch eine endliche Konstante $k \in \mathbb{R}$ beschränkt ist,

$$\left| \frac{\Delta f}{\Delta y} \right| = \left| \frac{f(x, y) - f(x, \bar{y})}{y - \bar{y}} \right| \leq k. \quad (3.4)$$

Anschaulich bedeutet das, dass man durch endliche Schritte entlang des Richtungsfeldes der Differentialgleichung immer nur endlich weit kommt, also keine unendlichen Sprünge machen kann. Entsprechend wird eine Differentialgleichung n -ter Ordnung erst durch n Anfangsbedingungen eindeutig festgelegt.

Betrachten wir als einfaches Beispiel die Differentialgleichung (2.6), das exponentielle Zerfallsgesetz $\dot{N} = -\lambda N$. Offenbar sind in (3.2) nun $y \rightarrow N$, $x \rightarrow t$ und $f(x, y) \rightarrow -\lambda N(t)$ zu ersetzen. Die Sekantensteigung, deren Beschränktheit überprüft werden muss, ist

$$\left| \frac{\Delta f}{\Delta y} \right| = \left| \frac{-\lambda N + \lambda \bar{N}}{N - \bar{N}} \right| = \lambda. \quad (3.5)$$

Die Lipschitz-Bedingung ist also erfüllt, wenn $\lambda \in \mathbb{R} < \infty$ ist.

Auf der anderen Seite können wir auch die nichtlineare Differentialgleichung erster Ordnung

$$y'(x) = y^2(x) \quad (3.6)$$

untersuchen. Ihre Sekantensteigung

$$\left| \frac{\Delta f}{\Delta y} \right| = \left| \frac{y^2 - \bar{y}^2}{y - \bar{y}} \right| = y + \bar{y} \quad (3.7)$$

ist über den gesamten Raum der reellen Zahlen nicht beschränkt. Wenn $y(x)$ stetig differenzierbar und damit glatt ist, dann kann man durch eine Einschränkung des Parameterbereiches in x und damit in y die Lipschitz-Bedingung und damit eine eindeutige Lösung erzwingen.

- Für physikalische Vorgänge, die fast immer durch Differentialgleichungen zweiter Ordnung in der Zeit beschrieben werden, bedeutet das, dass eine eindeutige Lösung der Bewegungsgleichung in der Nähe einer Zeit t_0 immer angegeben werden kann, wenn die Bewegung im Endlichen und mit endlicher Geschwindigkeit erfolgt und die auftretenden Kräfte stetig sind und sich nicht beliebig schnell ändern. Die Lösung wird durch zwei Anfangsbedingungen an den Ort und die Geschwindigkeit eindeutig bestimmt.

3.1.2 Lineare Differentialgleichungen n -ter Ordnung

Nachdem wir eine Vorstellung davon haben, welche Art von Differentialgleichungen physikalisch relevant sind, widmen wir uns kurz den Lösungen solcher Gleichungen. Insbesondere kommen wir noch einmal auf inhomogene Gleichungen zurück.

- Betrachten wir als Beispiel die lineare, inhomogene Differentialgleichung zweiter Ordnung

$$y''(x) - y(x) = 1 . \tag{3.8}$$

Die dazugehörige homogene Gleichung

$$y''(x) - y(x) = 0 . \tag{3.9}$$

kann zunächst durch den integrierenden Faktor $2y'$ einmal integriert werden,

$$y''(x) = y(x) \tag{3.10}$$

$$2y'y'' = 2yy' \Leftrightarrow (y'^2)' = (y^2)' \Leftrightarrow y' = \pm y \Leftrightarrow y_{1,2}(x) = e^{\pm x} .$$

Solange wir nur an einer Lösung der homogenen Gleichung interessiert sind, können wir die Integrationskonstante vernachlässigen.

Wenn man nicht sofort beiden Lösungen einer linearen homogenen Differentialgleichung zweiter Ordnung sieht, dann kann man sich eine zweite zum Beispiel durch *d'Alembert-Reduktion* verschaffen, wenn man eine Lösung $y_1(x)$ kennt. Man setzt den Ansatz $y_2(x) = y_1(x)f(x)$ mit einer unbekanntem Funktion $f(x)$ in die Gleichung ein und berücksichtigt, dass $y_1(x)$ schon eine Lösung ist, in unserem Fall also

$$y_2(x) = y_1(x)f(x) = e^x f(x) \tag{3.11}$$

mit der Bedingung

$$\begin{aligned} 0 &= y_1'' f + 2y_1' f' + f'' - y_1 f \\ &= (y_1'' - y_1) f + 2y_1' f' + y_1 f'' = 2y_1' f' + y_1 f'' . \end{aligned} \tag{3.12}$$

Es folgt für die unbekannte Funktion $f(x)$

$$\frac{f''}{f'} = -2 \frac{y_1'}{y_1} = -2 \quad \Leftrightarrow \quad (\ln f')' = -2 \quad \Leftrightarrow \quad \ln f' = -2x . \quad (3.13)$$

Man erhält direkt durch Integration, dass wegen $f'(x) = e^{-2x}$ und daher $f(x) = -e^{-2x}/2$ die Funktion $y_2 = y_1 f = -e^{-x}/2$ eine weitere Lösung sein muss. Im Fall einer linearen homogenen Differentialgleichung ist das gleichbedeutend mit der bekannten zweiten Lösung $y_2(x) = e^{-x}$.

- Anhand der Beispiele zum Fallgesetz haben wir gesehen, dass Differentialgleichungen in der Newtonschen Mechanik oft, aber nicht immer linear sind. Zur Erinnerung, linear heißen Differentialgleichungen, deren Koeffizientenfunktionen der Form

$$p_n(x)y^{(n)}(x) + \dots + p_1(x)y'(x) + p_0(x)y(x) = r(x) \quad (3.14)$$

nicht von y abhängen. Die linke Seite wird oft kurz als *linearer Differentialoperator* $L[y]$ bezeichnet, also $L[y] = r(x)$.

Allgemein gilt, dass falls $y_1(x)$ und $y_2(x)$ zwei Lösungen der linearen Differentialgleichung (3.14) sind, so ist auch eine beliebige Linearkombination der beiden eine Lösung, also

$$L[a_1 y_1 + a_2 y_2] = (a_1 + a_2)r(x) \quad \text{mit} \quad a_1, a_2 \in \mathbb{R} . \quad (3.15)$$

Eingesetzt in (3.14) sieht man sofort, dass dies richtig ist.

- Für die allgemeine Lösung einer *homogenen* Differentialgleichung n -ter Ordnung braucht man ein so genanntes *Fundamentalsystem* aus n linear unabhängigen Funktionen $y_i(x)$, $1 \leq i \leq n$. Linear unabhängig ist eine Menge von Funktionen $y_i(x)$ genau dann, wenn jede beliebige Linearkombination aus ihnen nur dann verschwinden kann, wenn alle ihre Koeffizienten verschwinden,

$$\sum_{i=1}^n a_i y_i(x) = 0 \quad \Rightarrow \quad a_i = 0 \quad \text{für alle} \quad 1 \leq i \leq n . \quad (3.16)$$

Es gibt ein einfaches Verfahren, um festzustellen, ob eine Menge von n Lösungen $y_i(x)$ einer linearen homogenen Differentialgleichung linear abhängig ist. Spezialisiert auf $n = 2$ besteht es darin zu prüfen, ob der Ausdruck

$$y_1(x)y_2'(x) - y_2(x)y_1'(x) \quad (3.17)$$

für alle betrachteten $x \in I$ identisch verschwindet. In diesem Fall sind die $y_1(x)$ und $y_2(x)$ linear abhängig.

Eingesetzt in unser Beispiel (3.17) folgt

$$e^x(-e^{-x}) - e^{-x}e^x = -2 \neq 0, \quad (3.18)$$

was belegt, dass die beiden Funktionen $y_1 = e^x$ und $y_2 = e^{-x}$ linear unabhängig sind. Die vollständige Lösung der homogenen Gleichung lautet also

$$y(x) = a_1 e^x + a_2 e^{-x}. \quad (3.19)$$

- Im letzten Schritt gehen wir wieder zurück zur inhomogenen Differentialgleichung. Die allgemeine Lösung einer *inhomogenen* linearen Differentialgleichung erhält man, indem man zur allgemeinen Lösung der dazugehörigen *homogenen* Gleichung eine spezielle Lösung der inhomogenen Gleichung addiert. Lineare inhomogene Differentialgleichungen zweiter Ordnung haben also Lösungen der Form

$$y(x) = a_1 y_1(x) + a_2 y_2(x) + y_{\text{in}}(x), \quad (3.20)$$

wobei $y_1(x)$ und $y_2(x)$ ein Fundamentalsystem bilden und $y_{\text{in}}(x)$ die spezielle Lösung der inhomogenen Gleichung ist. Die beiden reellen Zahlen a_1, a_2 werden durch die beiden Anfangsbedingungen bestimmt, die zur Eindeutigkeit der Lösung notwendig sind.

Während man sich eine Lösung der inhomogenen Gleichung (3.8) durch Variation der Konstanten verschaffen könnte, ist es einfacher zu bemerken, dass $y_{\text{in}} = -1$ eine solche Lösung ist. Damit erhält man die vollständige Lösung der inhomogenen Gleichung

$$y(x) = a_1 e^x + a_2 e^{-x} - 1. \quad (3.21)$$

Als Anfangsbedingungen wählen wir $y(x=0) = 0$ und $y'(x=0) = 0$. Wegen

$$y'(x) = a_1 e^x - a_2 e^{-x} \quad (3.22)$$

erfordert $y'(0) = 0$ die Bedingung $a_1 = a_2$. Demnach ist $y(x) = a_1 e^x + a_1 e^{-x} - 1$, so dass die Bedingung $y(0) = 0$ bedeutet $2a_1 = 1$ oder $a_1 = 1/2$. Also ist die vollständige Lösung der inhomogenen Gleichung mit Anfangsbedingungen

$$y(x) = \frac{e^x + e^{-x}}{2} - 1. \quad (3.23)$$

3.2 Vektoren I

Vektoren und ihre Eigenschaften sind in dieser Vorlesung schon vorgekommen, ohne dass wir sie mathematisch betrachtet haben. Daher ist es an der Zeit, mathematisch ein paar Grundlagen zu Vektoren und ihrer Verknüpfung zu sagen.

3.2.1 Vektorräume

- Zum Beispiel beim freien Fall haben wir uns auf die Analyse von Bewegungen in einer Dimension beschränkt. Wir kommen nun zur Darstellung von Bewegungen in drei Dimensionen. Da Positionen im dreidimensionalen Raum durch Vektoren \vec{x} angegeben werden, besteht die Lösung mechanischer Probleme meistens darin, eine Bahnkurve anzugeben, d.h. eine vektorwertige Funktion $\vec{x}(t)$ der Zeit t , die zeigt, an welchem Ort sich der Massenpunkt zur Zeit t befindet.
- Natürlich ist Ihnen zumindest intuitiv bekannt, was ein Vektor im dreidimensionalen Raum ist. Er hat einen Betrag und eine Richtung und wird z.B. durch drei Komponenten angegeben, bei denen es sich um reelle Zahlen handelt. In der Mathematik und in vielen Bereichen der theoretischen Physik ist ein Vektor aber ein sehr viel allgemeiner definiertes Objekt. Da Vektoren in verschiedensten Bedeutungen und Darstellungen in der Physik vorkommen, lohnt es sich hier, diese allgemeine Definition zu Grunde zu legen.
- In der Mathematik sind Vektoren Objekte, die man addieren, strecken oder stauchen kann, ohne dass sie ihre Vektoreigenschaft verlieren. Präzise definiert wird dies durch die Eigenschaften eines Vektorraums. Ein *Vektorraum* V ist eine Menge von Objekten, für die eine *Addition* definiert ist, die zwei Vektoren v_1, v_2 zu einem neuen Vektor verknüpft,

$$+ : V \times V \rightarrow V, \quad (v_1, v_2) \mapsto v_1 + v_2, \quad (3.24)$$

und für die eine *Multiplikation* mit den Elementen eines *Körpers* K definiert ist, die wiederum einen Vektor ergibt,

$$\cdot : K \times V \rightarrow V, \quad (\lambda, v) \mapsto \lambda \cdot v = \lambda v. \quad (3.25)$$

Wenn man diese beiden Verknüpfungen kombiniert, dann erfüllen die Addition und die Multiplikation in einem Vektorraum die linearen Rechenregeln

$$\begin{aligned} \lambda (v_1 + v_2) &= \lambda v_1 + \lambda v_2 \\ (\lambda_1 + \lambda_2) v &= \lambda_1 v + \lambda_2 v. \end{aligned} \quad (3.26)$$

- In diesen beiden Eigenschaften liegt die große Bedeutung von Vektorräumen in der Physik begründet. Offenbar kann alles durch Vektoren dargestellt werden, was überlagert werden kann, um dabei einen neuen Vektor zu ergeben. Aus (3.15) wissen wir, dass Lösungen linearer Differentialgleichungen durch Überlagerung

konstruiert werden können. Ebenso werden Zustände quantenmechanischer Systeme durch Vektoren beschrieben, weil man experimentell fand, dass zwei Zustände ψ_1 und ψ_2 zu einem neuen Zustand ψ überlagert werden können,

$$\psi = \lambda_1\psi_1 + \lambda_2\psi_2 . \quad (3.27)$$

Der Quantenmechanik liegt demnach auch die mathematische Struktur eines Vektorraums zu Grunde. Alles, was nun über Vektorräume gesagt wird, gilt also gleichermaßen für so verschiedene Dinge wie Orte im Raum oder quantenmechanische Zustände.

- Oben wurde der Begriff eines Körpers eingeführt, ohne ihn genau zu definieren. Dazu müssen wir zunächst daran erinnern, was eine Gruppe ist. *Gruppen* G sind Mengen mathematischer Objekte, zwischen denen eine Verknüpfung \circ definiert ist, die *assoziativ* ist,

$$a_1 \circ (a_2 \circ a_3) = (a_1 \circ a_2) \circ a_3 , \quad (3.28)$$

für die es ein *neutrales Element* e gibt,

$$a \circ e = a , \quad (3.29)$$

und die zu jedem Gruppenelement a ein *inverses Element* a^{-1} enthält,

$$a \circ a^{-1} = e . \quad (3.30)$$

Kommutativ oder *abelsch* heißt eine Gruppe dann, wenn

$$a_1 \circ a_2 = a_2 \circ a_1 \quad (3.31)$$

ist. Ein einfaches Beispiel sind die ganzen Zahlen bezüglich der Addition als Verknüpfung. Die Summe zweier ganzer Zahlen ist wieder eine, das neutrale Element ist die Null und das inverse Element zu einer ganzen Zahl z ist $-z$.

Ein *Körper* K ist nun eine Menge, die bezüglich einer Addition $+$ eine abelsche Gruppe ist und die außerdem bezüglich einer Multiplikation \cdot eine abelsche Gruppe ist. Wir wissen schon dass für das neutrale Element der Addition (in der Regel die Null) ein inverses Element problematisch wird, also schliessen wir es in dieser Definition der Gruppe aus. Außerdem gilt das Distributivgesetz in der Form

$$a \cdot (b + c) = a \cdot b + a \cdot c , \quad (a + b) \cdot c = a \cdot c + b \cdot c . \quad (3.32)$$

Im Unterschied zu den Vektoren heißen die Elemente des Körpers, auf dem ein Vektorraum definiert wird, *Skalare*. In der Regel unterliegt physikalischen Vektorräumen entweder der Körper der reellen Zahlen oder der der komplexen Zahlen.

3.2.2 Lineare Abhängigkeit, Basis und Dimension

Nachdem wir gesehen haben, dass wir formal Lösungen linearer Differentialgleichungen als Vektoren betrachten können, stellt sich die schon bekannte Frage, wie wir wissen dass wir alle Lösungen oder alle Elemente eines Vektorraumes identifiziert haben. In (3.16) haben wir ad hoc einen solchen Test angegeben, für den wir jetzt die notwendigen mathematischen Definitionen nachreichen.

- Mengen von Vektoren v_i mit $1 \leq i \leq n$ aus einem Vektorraum V können zu *Linearkombinationen* v verknüpft werden,

$$v = \lambda_1 v_1 + \lambda_2 v_2 + \dots + \lambda_n v_n . \quad (3.33)$$

Die Menge aller Vektoren v , die auf diese Weise durch die Menge der v_i dargestellt werden können, bildet einen Untervektorraum U von V , $U \subseteq V$. Er wird auch als der von den v_i *aufgespannte* Untervektorraum bezeichnet,

$$U = \text{span}(v_i) . \quad (3.34)$$

- Eine Menge von Vektoren v_i heißt *linear unabhängig*, wenn die Gleichung

$$\lambda_1 v_1 + \lambda_2 v_2 + \dots + \lambda_n v_n = 0 \quad (3.35)$$

nur dann erfüllt werden kann, wenn alle Skalare verschwinden, $\lambda_i = 0$. Insbesondere bedeutet das, dass eine Menge von Vektoren v_i genau dann linear unabhängig ist, wenn sich jeder Vektor $v \in \text{span}(v_i)$ eindeutig aus den Elementen v_i linear kombinieren lässt.

- Eine Menge von Vektoren v_i heißt *Erzeugendensystem* eines Vektorraums V , wenn $V = \text{span}(v_i)$ ist. Ein Erzeugendensystem ist eine *Basis*, wenn seine Elemente linear unabhängig und daher ihre Anzahl minimal sind. Die Anzahl der Elemente einer Basis ist dann die *Dimension* N des Vektorraums. Da sich jeder Vektor v aus V als Linearkombination der Basisvektoren e_i schreiben lässt,

$$v = \sum_{i=1}^N \lambda_i e_i , \quad (3.36)$$

kann jeder Vektor durch das N -Tupel der Zahlen λ_i eindeutig angegeben werden, sobald die Basis bekannt ist.

3.2.3 Beispiele

- Der dreidimensionale Raum \mathbb{R}^3 oder die Menge \mathbb{R}^3 der 3-Tupel (x_1, x_2, x_3) bekommt die Struktur eines Vektorraums über dem

Körper der reellen Zahlen, indem man eine Addition von Vektoren und deren Multiplikation mit Skalaren komponentenweise definiert,

$$v + w = (v_1, v_2, v_3) + (w_1, w_2, w_3) = (v_1 + w_1, v_2 + w_2, v_3 + w_3) \quad (3.37)$$

bzw.

$$\lambda(v_1, v_2, v_3) = (\lambda v_1, \lambda v_2, \lambda v_3). \quad (3.38)$$

Eine geeignete Basis wird durch die Einheitsvektoren $e_1 = (1, 0, 0)$, $e_2 = (0, 1, 0)$ und $e_3 = (0, 0, 1)$ dargestellt. Irgend zwei der Basisvektoren spannen einen Unterraum von \mathbb{R}^3 auf, so lässt sich zum Beispiel die x_1 - x_2 -Ebene durch e_1 und e_2 aufspannen.

- Ein vielleicht weniger nahe liegendes Beispiel sind die Polynome vom Grad $\leq N$ auf dem Intervall $[-1, 1]$,

$$p(x) = \sum_{j=0}^N a_j x^j, \quad a_j \in \mathbb{R}. \quad (3.39)$$

Sie bilden einen Vektorraum über dem Körper der reellen Zahlen, da durch die gewöhnliche Addition zweier solcher Polynome wieder eines entsteht, und weil auch die Multiplikation mit reellen Zahlen offensichtlich definiert ist.

Offenbar bilden verschiedene Potenzen x^j mit $0 \leq j \leq N$ Systeme linear unabhängiger Vektoren aus diesem Vektorraum, da die Gleichung

$$\lambda_0 + \lambda_1 x + \dots + \lambda_n x^n = 0 \quad (3.40)$$

nur dann für alle $x \in [-1, 1]$ erfüllt werden kann, wenn alle λ_j verschwinden. Zudem spannen die x^j mit $0 \leq j \leq N$ den gesamten Vektorraum der Polynome vom Grad $\leq N$ auf, weil jedes Polynom vom Grad $\leq N$ durch Linearkombination der x^j dargestellt werden kann. Also bilden die x^j eine Basis dieses abstrakten Vektorraums, der durch die Werte λ_j dargestellt werden kann.

3.3 Taylor-Reihen

Die Bedeutung von Taylor-Reihen in der Theoretischen Physik lässt sich kaum überschätzen. Der Grund ist, dass sie uns erlauben werden, jede funktionale Abhängigkeit in führende und weniger relevante Beiträge zu unterteilen. Man kann sich also in der mathematischen Beschreibung eines Problems zunächst den numerisch wichtigsten Effekten widmen, und dann Abweichungen vom führenden Verhalten Stück für Stück abarbeiten. Dieser Zugang beschreibt die sogenannte Störungstheorie.

3.3.1 Taylorsche-Satz

- Wir beginnen zunächst mit einer allgemeinen mathematischen Betrachtung, die auf alle Funktionen anwendbar ist, die „genügend glatt“, also genügend oft stetig differenzierbar sind. Sei zunächst ein Intervall der reellen Zahlen $I \in \mathbb{R}$ vorgegeben, und darauf eine Funktion $f : I \rightarrow \mathbb{R}$ definiert, die mindestens $n + 1$ -mal stetig differenzierbar sei. Seien ferner $a, x \in I$ zwei beliebige reelle Zahlen aus dem Intervall. Dann besagt der *Hauptsatz der Differential- und Integralrechnung*

$$f(x) = f(a) + \int_a^x f'(t)dt, \tag{3.41}$$

wobei $f'(x)$ die Ableitung der Funktion $f(x)$ nach x bezeichnet. Wir setzen diesen Satz hier als bekannt voraus.

- Wir behaupten, dass sich $f(x)$ in der folgenden Weise darstellen lasse:

$$f(x) = \sum_{j=0}^{\infty} \frac{(x-a)^j}{j!} f^{(j)}(a) \tag{3.42}$$

$$= f(a) + (x-a)f'(a) + \dots + \frac{(x-a)^n}{n!} f^{(n)}(a) + R_{n+1}(x)$$

wobei $R_{n+1}(x)$ das *Restglied*

$$R_{n+1}(x) = \frac{1}{n!} \int_a^x (x-t)^n f^{(n+1)}(t)dt \tag{3.43}$$

ist. Dies ist die Aussage des *Taylorischen Satzes*, und (3.42) ist die *Taylorische Formel*.

- Der Taylorsche Satz lässt sich am besten durch vollständige Induktion beweisen. Aus dem Hauptsatz der Differential- und Integralrechnung folgt bereits, dass die Aussage (3.42) für $n = 0$ wahr ist, denn es gilt offenbar

$$f(x) = f(a) + R_1(x) \tag{3.44}$$

$$= f(a) + \frac{1}{1!} \int_a^x (x-t)^0 f^{(1)}(t)dt = f(a) + \int_a^x f'(t)dt$$

nach (3.41). Damit ist die Induktion bei $n = 0$ verankert.

Im zweiten Schritt nehmen wir an, die Taylorische Formel gelte für beliebiges $n - 1$ und zeigen, dass sie dann auch für n gilt. Wir setzen also voraus, dass

$$f(x) = f(a) + \sum_{i=1}^{n-1} \frac{f^{(i)}(a)}{i!} (x-a)^i + R_n(x) \tag{3.45}$$

gilt und stellen das Restglied $R_n(x)$ durch partielle Integration anders dar. Wir erhalten

$$\begin{aligned}
 R_n(x) &= \frac{1}{(n-1)!} \int_a^x (x-t)^{n-1} f^{(n)}(t) dt \\
 &= - \int_a^x \frac{d}{dt} \left[\frac{(x-t)^n}{n!} \right] f^{(n)}(t) dt \\
 &= - \frac{(x-t)^n}{n!} f^{(n)}(t) \Big|_a^x + \int_a^x \frac{(x-t)^n}{n!} f^{(n+1)}(t) dt \\
 &= \frac{(x-a)^n}{n!} f^{(n)}(a) + R_{n+1}(x) .
 \end{aligned} \tag{3.46}$$

Setzt man dieses Ergebnis in (3.45) ein, folgt damit

$$\begin{aligned}
 f(x) &= f(a) + \sum_{i=1}^{n-1} \frac{f^{(i)}(a)}{i!} (x-a)^i + \frac{(x-a)^n}{n!} f^{(n)}(a) + R_{n+1}(x) \\
 &= f(a) + \sum_{i=1}^n \frac{f^{(i)}(a)}{i!} (x-a)^i + R_{n+1}(x) .
 \end{aligned} \tag{3.47}$$

und die Gültigkeit des Taylorschen Satzes auch für $n + 1$. Damit ist die Induktion vollständig und der Taylorsche Satz bewiesen.

- Offenbar ist die Taylor-Entwicklung einer Funktion bis zur ersten Ordnung genau dann exakt, wenn $f(x)$ eine lineare Funktion ist, d.h. wenn sie eine Gerade darstellt. Dann ist $f(x) = \alpha + \beta x$ mit zwei beliebigen Konstanten $\alpha, \beta \in \mathbb{R}$, und

$$f(x) = f(a) + \beta(x-a) = \alpha + \beta a + \beta x - \beta a = \alpha + \beta x . \tag{3.48}$$

Deshalb nennt man eine Taylor-Entwicklung bis zur ersten Ordnung auch *Linearisierung* der Funktion in der Umgebung von a . Wie weit diese Umgebung sein kann, hängt natürlich davon ab, bis wohin der Verlauf von $f(x)$ in der Nähe von a durch eine Gerade angenähert werden kann.

3.3.2 Einfache Beispiele

- Meistens wird der Taylorsche Satz dazu verwendet, das Verhalten von Funktionen in der Nähe bestimmter Punkte zu nähern. Man wählt dann einen Punkt a und betrachtet Punkte x in einer kleinen Umgebung von a , für die gewöhnlich $|x-a| \ll 1$ gilt. Auch komplizierte Funktionen lassen sich dann auf einfache Weise darstellen. Betrachten wir einige Beispiele.
- Sei

$$f(x) = \frac{1}{(1 \pm x)^b} \tag{3.49}$$

die Funktion, die in der Nähe von $x = 0$ betrachtet werden soll. Für $x \ll 1$ können wir den Taylorschen Satz bis zur ersten Ordnung verwenden und erhalten

$$f(x) \approx f(0) \mp \frac{b}{(1 \pm 0)^{(b+1)}} x = 1 \mp bx. \quad (3.50)$$

Insbesondere ist also zum Beispiel

$$\frac{1}{\sqrt{1 \pm x}} \approx 1 \mp \frac{x}{2}. \quad (3.51)$$

- Ein weiteres häufig verwendetes Beispiel ist

$$f(x) = \ln(1 \pm x) \quad (3.52)$$

für $x \ll 1$. Dann ist $a = 0$ und damit

$$\ln(1 \pm x) \approx \ln(1 \pm 0) \pm \frac{x - 0}{1 \pm 0} = \pm x. \quad (3.53)$$

- Manchmal reicht eine Linearisierung nicht, um zu physikalisch brauchbaren Aussagen zu kommen. Während die Linearisierung der Sinusfunktion in der Nähe von $x = 0$

$$\sin x \approx \sin 0 + x \cos 0 = x \quad (3.54)$$

ergibt, erhält man durch Linearisierung des Cosinus

$$\cos x = \cos 0 - x \sin 0 = 1, \quad (3.55)$$

also eine Konstante, die den Verlauf der Cosinusfunktion in der Nähe von $x = 0$ nicht erkennen lässt. In diesem Fall muss man bis zur zweiten Ordnung Taylor-entwickeln:

$$\begin{aligned} \cos x &\approx \cos 0 - x \sin 0 - \frac{\cos 0}{2} x^2 = 1 - \frac{x^2}{2} \\ \sin x &\approx \sin 0 + x \cos 0 - \frac{\sin 0}{2} x^2 = x \end{aligned} \quad (3.56)$$

Wenn man beide Funktionen konsistent bis zur zweiten Ordnung entwickelt, dann kann man überprüfen dass

$$\begin{aligned} \sin^2 x + \cos^2 x &= \left(1 - \frac{x^2}{2} + \mathcal{O}(x^3)\right)^2 + \left(x + \mathcal{O}(x^2)\right)^2 \\ &= \left(1 - \frac{x^2}{2}\right)^2 + x^2 + \mathcal{O}(x^3) \\ &= 1 - x^2 + x^2 + \mathcal{O}(x^3) = 1 + \mathcal{O}(x^3), \end{aligned} \quad (3.57)$$

wobei der Term $\mathcal{O}(x^3)$ mathematisch bedeutet, dass in unserer Taylor-Entwicklung Terme dritter und höherer Potenz vernachlässigt werden.

Kapitel 4

Impuls, Drehimpuls und Energie

4.1 Vektoren II

In der linearen Algebra sieht man, dass die Addition und die Streckung von Vektoren nicht die einzigen hilfreichen Konzepte sind. Schon für Vektoren im dreidimensionalen Raum wissen wir, dass zum Beispiel der Öffnungswinkel durch eine weitere Verknüpfung algebraisch berechnet werden kann.

4.1.1 Das Skalarprodukt

- Ein *Skalarprodukt* oder *inneres Produkt* ist eine Abbildung, die zwei Vektoren aus einem Vektorraum eine Zahl aus dem zugrunde liegenden Körper K zuordnet,

$$\langle \cdot, \cdot \rangle : V \times V \rightarrow K, \quad (v, w) \mapsto \langle v, w \rangle, \quad (4.1)$$

und zwar so, dass sie in beiden ihrer Argumente linear ist,

$$\begin{aligned} \langle \lambda_1 v_1 + \lambda_2 v_2, w \rangle &= \lambda_1 \langle v_1, w \rangle + \lambda_2 \langle v_2, w \rangle, \\ \langle v, \mu_1 w_1 + \mu_2 w_2 \rangle &= \mu_1 \langle v, w_1 \rangle + \mu_2 \langle v, w_2 \rangle. \end{aligned} \quad (4.2)$$

Eine solche Abbildung heißt *bilinear*. Skalarprodukte sind positiv definit, d.h. das Skalarprodukt eines Vektors mit sich selbst ist immer positiv, $\langle v, v \rangle > 0$ für alle $v \in V$ mit $v \neq 0$, und $\langle v, v \rangle = 0$ genau dann, wenn $v = 0$. Mithilfe des Skalarprodukts kann einem Vektor $v \in V$ sein Betrag $|v|$ zugeordnet werden,

$$|v| = \sqrt{\langle v, v \rangle}. \quad (4.3)$$

Ist $|v| = 1$, heißt v *Einheitsvektor*.

- Zwei Vektoren $v, w \in V$ heißen *orthogonal*, wenn ihr Skalarprodukt verschwindet, $\langle v, w \rangle = 0$. Besonders bequem sind Basen, deren Elemente e_i paarweise orthogonale Einheitsvektoren sind,

$$\langle e_i, e_j \rangle = \begin{cases} 1 & (i = j) \\ 0 & (i \neq j) \end{cases}. \quad (4.4)$$

Hat man eine solche *Orthonormalbasis*, nimmt das Skalarprodukt eine sehr einfache Form an,

$$\langle a, b \rangle = \left\langle \sum_{i=1}^N \alpha_i e_i, \sum_{j=1}^N \beta_j e_j \right\rangle = \sum_{i=1}^N \sum_{j=1}^N \alpha_i \beta_j \langle e_i, e_j \rangle = \sum_{i=1}^N \alpha_i \beta_i. \quad (4.5)$$

- Im dreidimensionalen Vektorraum \mathbb{R}^3 gibt das Skalarprodukt zwischen zwei Vektoren a, b gerade den Cosinus des Winkels ϕ an, den die beiden Vektoren einschließen,

$$\langle a, b \rangle = |a| |b| \cos \phi. \quad (4.6)$$

Aus der Erweiterung dieser Beziehung auf beliebige Vektorräume leitet sich die Definition orthogonaler Vektoren ab, die oben eingeführt wurde.

- In unserem zweiten Beispiel des Vektorraums der Polynome vom Grad N auf dem Intervall $[-1, 1]$ kann ein Skalarprodukt zwischen zwei Polynomen $p(x)$ und $q(x)$ durch die Definition

$$\langle p(x), q(x) \rangle = \int_{-1}^1 p(x)q(x)dx \quad (4.7)$$

eingeführt werden. Demnach wären z.B. die beiden Polynome $p(x) = x$ und $q(x) = x^2$ zueinander orthogonal, denn

$$\langle x, x^2 \rangle = \int_{-1}^1 x^3 dx = \frac{x^4}{4} \Big|_{-1}^1 = 0. \quad (4.8)$$

Entsprechend lässt sich auf diesem Vektorraum eine orthonormale Basis einführen. Die Darstellung von Polynomen oder allgemein von genügend gutwilligen Funktionen durch Orthonormalbasen hat eine immense Bedeutung in der Physik. Ein Beispiel dafür ist die Fouriertransformation.

4.1.2 Summenkonvention, Levi-Civita-Symbol

- Oft ist es lästig und unnötig, Summen wie etwa in (4.5) auszu-schreiben. Zur Abkürzung verabreden wir die Einsteinsche Summenkonvention, nach der über doppelt auftretende Indizes summiert wird,

$$\sum_{i=1}^N \alpha_i \beta_i =: \alpha_i \beta_i. \quad (4.9)$$

In einer Gleichung kann also derselbe Index einmal auf jeder Seite der Gleichung erscheinen, dann ist es eine Gleichung für die einzelnen Komponenten eines Vektors, oder zweimal auf derselben Seite, dann nutzen wir die Einsteinsche Summenkonvention. Ein Index ohne Summenzeichen kann in einer Gleichung niemals mehr als zweimal erscheinen!

- Ebenso nützlich ist das Kronecker-Symbol δ_{ij} , das durch

$$\delta_{ij} := \begin{cases} 1 & (i = j) \\ 0 & (i \neq j) \end{cases} \quad (4.10)$$

definiert wird. Die Vereinfachung, die aufgrund der Summenkonvention und mithilfe des Kronecker-Symbols erreicht wird, ist offensichtlich, wenn wir (4.4) und (4.5) entsprechend schreiben. (4.4) wird zu

$$\langle e_i, e_j \rangle = \delta_{ij} , \quad (4.11)$$

und aus (4.5) wird

$$\langle \alpha_i e_i, \beta_j e_j \rangle = \alpha_i \beta_j \delta_{ij} = \alpha_i \beta_i . \quad (4.12)$$

- Weiterhin führen wir für dreidimensionale Vektorräume das vollkommen antisymmetrische Levi-Civita-Symbol ϵ_{ijk} ein, das wie folgt definiert ist:

$$\begin{aligned} \epsilon_{123} &= \epsilon_{231} = \epsilon_{312} = 1 \\ \epsilon_{132} &= \epsilon_{321} = \epsilon_{213} = -1 \\ \epsilon_{ijk} &= 0 \quad \text{sonst} , \end{aligned} \quad (4.13)$$

d.h. $\epsilon_{ijk} = 1$ für alle geraden Permutationen von $\{1, 2, 3\}$, $\epsilon_{ijk} = -1$ für alle ungeraden Permutationen davon, und $\epsilon_{ijk} = 0$, wenn mindestens zwei der Indizes i, j, k gleich sind.

Zwischen dem Levi-Civita- und dem Kronecker-Symbol besteht der sehr nützliche Zusammenhang

$$\epsilon_{ijk} \epsilon_{klm} = \delta_{il} \delta_{jm} - \delta_{im} \delta_{jl} , \quad (4.14)$$

mit dem sich viele Rechnungen erheblich abkürzen lassen.

4.1.3 Das Vektorprodukt

- Das *Vektorprodukt* oder äußere Produkt ordnet zwei Vektoren v, w aus V einen Vektor aus V zu,

$$\times : V \times V \rightarrow V , \quad (v, w) \mapsto v \times w , \quad (4.15)$$

und zwar wieder auf bilineare Weise,

$$(\lambda_1 v_1 + \lambda_2 v_2) \times w = \lambda_1 v_1 \times w + \lambda_2 v_2 \times w \quad (4.16)$$

und ebenso für das zweite Argument.

Wegen der Bilinearität reicht es zur Festlegung des Vektorprodukts wieder, die Vektorprodukte der Basisvektoren festzulegen. Mithilfe des Levi-Civita-Symbols definieren wir in drei Dimensionen

$$e_i \times e_j = \epsilon_{ijk} e_k . \quad (4.17)$$

Demzufolge ist das Vektorprodukt *antisymmetrisch*, z.B. $e_1 \times e_2 = e_3$ und $e_2 \times e_1 = -e_3$, und das Vektorprodukt zweier gleicher Basisvektoren verschwindet, $e_i \times e_i = 0$.

Mit dieser Festlegung lautet das Vektorprodukt zweier dreidimensionaler Vektoren a, b

$$a \times b = (\alpha_i e_i) \times (\beta_j e_j) = \alpha_i \beta_j (e_i \times e_j) = \epsilon_{ijk} \alpha_i \beta_j e_k . \quad (4.18)$$

- Aufgrund seiner Definition ist das Vektorprodukt zweier Vektoren a, b orthogonal zu beiden Vektoren,

$$\langle a, a \times b \rangle = \langle \alpha_i e_i, \epsilon_{jkl} \alpha_j \beta_k e_l \rangle = \alpha_i \alpha_j \beta_k \epsilon_{jkl} \delta_{il} = \epsilon_{ijk} \alpha_i \alpha_j \beta_k . \quad (4.19)$$

Da das Produkt $\alpha_i \alpha_j \beta_k$ symmetrisch gegen Vertauschung von i und j ist, ϵ_{ijk} aber antisymmetrisch, verschwindet das Ergebnis, ebenso wie für das Skalarprodukt $\langle b, a \times b \rangle$.

- Ebenso verschwindet das Vektorprodukt zwischen parallelen oder antiparallelen Vektoren a, b , denn dann ist $b = \lambda a$ mit $\lambda \in K$, und

$$a \times b = \lambda \alpha_i \alpha_j \epsilon_{ijk} e_k = 0 . \quad (4.20)$$

Entsprechend kann man dem Vektorprodukt zweier Vektoren den Sinus ihres Zwischenwinkels ϕ zuordnen,

$$|a \times b| = |a| |b| \sin \phi . \quad (4.21)$$

4.1.4 Koordinaten und Koordinatentransformationen

- Die abstrakte Definition von Vektoren und Vektorräumen im dritten Kapitel hat unter Anderem den Vorteil, dass völlig klar wird, dass Vektoren unabhängig von bestimmten Koordinatensystemen existieren. Zwar werden Vektoren im dreidimensionalen Vektorraum \mathbb{R}^3 in der Regel durch reelle Zahlentripel (x_1, x_2, x_3) angegeben, aber diese Zahlentripel dürfen keinesfalls mit den Vektoren selbst verwechselt werden. Zum Beispiel bei den Fallgesetzen

haben wir gelernt, dass die Physik nicht von der Wahl des Koordinatensystems abhängen darf, auch wenn wir für eine konkrete Rechnung zum Beispiel die z -Achse und den Nullpunkt auf dieser Achse sowie in der Zeit festlegen müssen.

- Die Wahl eines Koordinatensystems besteht darin, dass man eine geeignete Basis e_i für einen Vektorraum V festlegt. Dadurch werden Vektoren $v \in V$ eindeutig durch N -Tupel v_i von Zahlen aus dem Körper K darstellbar, $v = v_i e_i$. Die Eindeutigkeit folgt daraus, dass die Basis ein linear unabhängiges Erzeugendensystem des gesamten Vektorraums V ist. Hat man eine Orthonormalbasis gewählt, sind die Koordinaten v_i durch die Skalarprodukte $\langle v, e_i \rangle$ gegeben,

$$\langle v, e_i \rangle = \langle v_j e_j, e_i \rangle = v_j \langle e_j, e_i \rangle = v_j \delta_{ij} = v_i . \quad (4.22)$$

- Ebenso gut können wir durch eine andere Basis e'_i ein anderes Koordinatensystem einführen, in dem derselbe Vektor v nun durch andere Koordinaten v'_i dargestellt wird,

$$v = v_i e_i = v'_j e'_j . \quad (4.23)$$

Die offensichtliche Frage ist, wie sich die neuen Koordinaten v'_i durch die alten v_j ausdrücken lassen. Dazu genügt es zu wissen, wie die neuen Basisvektoren e'_i durch die alten ausgedrückt werden können, d.h. die Beziehungen

$$e'_i = a_{ij} e_j \quad (4.24)$$

müssen durch die Angabe der $N \times N$ Zahlen a_{ij} vollständig festgelegt sein. Um die Zahlen a_{ij} zu berechnen können wir (4.24) zum Beispiel skalar mit e_k multiplizieren,

$$\langle e'_i, e_k \rangle = a_{ij} \langle e_j, e_k \rangle = a_{ij} \delta_{jk} = a_{ik} . \quad (4.25)$$

Die Zahlen a_{ij} sind also die Skalarprodukte der neuen mit den alten Basisvektoren.

Entsprechend können wir auch die neuen Koordinaten v'_i erhalten, indem wir (4.23) skalar mit e'_k multiplizieren,

$$\langle v, e'_k \rangle = v'_k = v_i \langle e_i, e'_k \rangle = a_{ki} v_i . \quad (4.26)$$

Umgekehrt können wir die Koordinaten v_i im alten Bezugssystem durch die neuen v'_j ausdrücken, indem wir (4.23) skalar mit e_k multiplizieren,

$$\langle v, e_k \rangle = v_k = v'_j \langle e'_j, e_k \rangle = v'_j a_{jk} . \quad (4.27)$$

Beachten Sie den Unterschied zwischen (4.26) und (4.27): In (4.26) wird über den zweiten Index von a_{ij} summiert, in (4.27) über den ersten.

4.1.5 Inertialsysteme

- Wenn Koordinatensysteme aus mathematischer Sicht beliebig gewählt werden können, stellt sich die Frage, ob bestimmte Koordinatensysteme gegenüber anderen physikalisch ausgezeichnet sind. Offenbar macht es in der Betrachtung der wirkenden Kräfte einen Unterschied, ob die Bahnkurve eines Körpers in einem Bezugssystem betrachtet wird, das sich mit ihm bewegt, oder in einem Bezugssystem, das sich relativ dazu dreht. Dies führt uns auf den Begriff der *Inertialsysteme*. Ein Inertialsystem ist ein solches Bezugssystem, in dem das erste Newtonsche Axiom gilt, in dem sich also ein kräftefreier Körper geradlinig-gleichförmig bewegt.
- Eine sich drehende Scheibe ist ein Beispiel für ein Bezugssystem, in dem das erste Newtonsche Axiom nicht gilt, denn auf ihr kann ein Körper nur dann in Ruhe bleiben, wenn er durch eine Kraft festgehalten wird.
- Der Begriff des Inertialsystems verdeutlicht, dass die Definitionen der Kraft und des Inertialsystems in einer Weise zyklisch sind, die schon Newton als problematisch empfunden hat: Erst wenn man ein Inertialsystem eingeführt hat, kann man Kräfte sinnvoll definieren. Vorher ist nicht klar, ob eine nicht geradlinig-gleichförmige Bewegung auf eine Kraft zurückgeht oder darauf zurückzuführen ist, dass das Bezugssystem kein Inertialsystem ist. Andererseits kann man Inertialsysteme nicht definieren, ohne auf Kräfte Bezug zu nehmen, weil sie eben als Bezugssysteme definiert sind, in denen sich kräftefreie Körper geradlinig-gleichförmig bewegen.

Erst die Einführung von Näherungen hilft aus dieser zirkulären Situation heraus. Beispielsweise ist die Erde keineswegs ein Inertialsystem, weil sie etwa in Bezug auf ferne Fixsterne rotiert. Effekte, die aufgrund der Erdbewegung auftreten, spielen aber in unseren Labors meist keine Rolle, weil die dort untersuchten Kräfte größer sind. In seiner Allgemeinen Relativitätstheorie ist es Einstein gelungen, Inertialsysteme durch eine sehr viel befriedigendere Klasse von Bezugssystemen zu ersetzen.

4.2 Impuls, Drehimpuls und Energie

In der Einführung zur Newtonschen Mechanik haben wir die Bewegung eines Körpers durch die anliegenden Kräfte beschrieben. Das hat es uns erlaubt, die Bewegung dieses Körpers als Funktion der Zeit zu beschreiben. Bei der Beschreibung des freien Falls aus großer Höhe haben wir aber eine alternative Art der Beschreibung einer Bewegung gesehen,

nämlich unabhängig von der Zeit erhaltene Größen. Solche Größen sehen wir uns nun etwas systematischer an.

4.2.1 Impuls

- Den Impuls \vec{p} haben wir schon als Produkt von Masse m und Geschwindigkeit \vec{v} definiert, als wir das 1. Newtonsche Axiom (2.2) eingeführt haben. Es besagt, dass sich der Impuls eines Körpers dann nicht ändert, wenn keine Kräfte auf ihn wirken,

$$\vec{p} = \text{const.} \quad \text{in Abwesenheit von Kräften} \quad (4.28)$$

Nach den Bemerkungen über allgemeine Bezugs- und Inertialsysteme in Abschnitt 5.1 müssen wir hier präzisieren, dass die Aussage des Trägheitsgesetzes *in Inertialsystemen* gilt.

Das zweite Newtonsche Gesetz (2.3) lautet in Form des Impulses

$$\dot{\vec{p}} = \vec{F}. \quad (4.29)$$

Für zeitliche Veränderungen von Impulsen sind Kräfte verantwortlich. Umgekehrt muss der Impuls also erhalten sein, wenn keine Kräfte wirken.

Die letzte Aussage gilt wiederum in Inertialsystemen. Aus der Sicht von jemanden, der im Karussell fährt, bewegt sich die gesamte Umgebung, ohne dass Kräfte auf sie wirken!

4.2.2 Drehmoment und Drehimpuls

- Das *Moment* einer Kraft \vec{F} bezüglich des Koordinatenursprungs wird durch

$$\vec{M} := \vec{r} \times \vec{F} \quad (4.30)$$

definiert. Wenn ϕ der Winkel zwischen \vec{r} und \vec{F} ist, hat es den Betrag $|\vec{M}| = rF \sin \phi$. Das Moment einer Kraft verschwindet, wenn sie in dieselbe Richtung zeigt wie der Ortsvektor \vec{r} ihres Angriffspunkts.

Das Moment um eine Achse, deren Richtung durch den Einheitsvektor \vec{e} beschrieben wird, ist entsprechend durch

$$M_e = \vec{e} \cdot \vec{M} = \vec{e} \cdot (\vec{r} \times \vec{F}) \quad (4.31)$$

definiert. Es ist offensichtlich ein Skalar.

- Ähnlich wird der *Drehimpuls* um den Koordinatenursprung durch

$$\vec{L} = \vec{r} \times \vec{p} = \vec{r} \times m\vec{v} \quad (4.32)$$

definiert. Seine Zeitableitung ist durch

$$\begin{aligned} \frac{d\vec{L}(t)}{dt} &= \dot{\vec{r}} \times \vec{p} + \vec{r} \times \dot{\vec{p}} \\ &= \frac{\vec{p}}{m} \times \vec{p} + \vec{r} \times \dot{\vec{p}} \\ &= \vec{r} \times \dot{\vec{p}} = \vec{r} \times \vec{F} = \vec{M} \end{aligned} \quad (4.33)$$

gegeben. Demnach gilt der *Drehimpulssatz*

$$\dot{\vec{L}} = \vec{M}. \quad (4.34)$$

In Abwesenheit von Drehmomenten, $\vec{M} = 0$, bleibt der Drehimpuls erhalten, $\dot{\vec{L}} = 0$. Der Vergleich mit (4.29) zeigt, dass Drehmoment und Drehimpuls sich genauso verhalten wie Kraft und Impuls bei einer linearen Bewegung.

4.2.3 Energiesatz in einer Dimension

- Bei der Diskussion der Energie beschränken wir uns zunächst auf eine Dimension. Wie vorher (Abschnitt 1.4.4) bei der Diskussion des freien Falls aus großer Höhe verwenden wir \dot{x} als integrierenden Faktor der Bewegungsgleichung,

$$m\ddot{x} = F(x) \Rightarrow m\dot{x}\ddot{x} = \dot{x}F(x) \quad (4.35)$$

Wir nehmen nun an dass

$$V(x) := - \int_{x_0}^x F(x') dx' \quad \text{oder} \quad F(x) = - \frac{dV(x)}{dx} \quad (4.36)$$

wobei der Anfangspunkt x_0 kann frei gewählt werden kann. Damit ist

$$\frac{d}{dt} V(x) = \frac{dV}{dx} \dot{x} = -F(x)\dot{x} \Rightarrow \frac{m}{2} \frac{d(\dot{x}^2)}{dt} = - \frac{d}{dt} V(x), \quad (4.37)$$

und die Zeitableitung

$$\frac{d}{dt} \left[\frac{m}{2} \dot{x}^2 + V(x) \right] = 0 \quad (4.38)$$

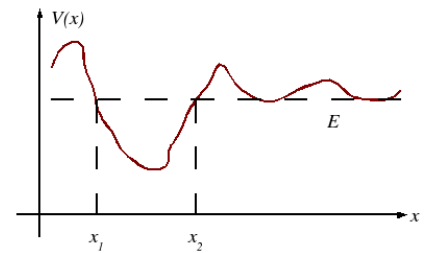
verschwindet.

- Demnach ist die *Energie*

$$E := \frac{m}{2} \dot{x}^2 + V(x) \tag{4.39}$$

konstant. Die Größe $V(x)$ wird als *potentielle Energie* oder als *Potential* bezeichnet.

In einer Dimension kann zu einer Kraft $F(x)$, die nicht von \dot{x} abhängt, immer ein Potential angegeben werden. Wegen der freien Wahl von x_0 ist $V(x)$ nur bis auf eine Konstante bestimmt.



Einschränkung der Bewegung und Umkehrpunkte

- Aus dem Energiesatz in einer Dimension (4.39) lässt sich die Trajektorie des Körpers implizit durch Trennung der Variablen ausdrücken

$$\begin{aligned} \frac{dx}{dt} &= \pm \sqrt{\frac{2}{m}(E - V)} \\ \Rightarrow \pm \int_{x_0}^x \frac{dx'}{\sqrt{\frac{2}{m}(E - V)}} &= t - t_0. \end{aligned} \tag{4.40}$$

Offenbar ist Bewegung nur dort möglich, wo $E - V \geq 0$ ist. Die kinetische Energie $T = (m/2)\dot{x}^2$ ist positiv-semidefinit, $T \geq 0$.

Durch diese Positivität werden *Umkehrpunkte* $x_{1,2}$ definiert, die die Bewegung begrenzen. Bei $x = x_1$ und $x = x_2$ wird die kinetische Energie minimal ($T = 0$) und daher $V = E$. Das ist die Bedeutung des \pm -Zeichens oben: an den Umkehrpunkten ändert sich die Richtung der Bewegung. Dadurch tritt eine (möglicherweise nicht harmonische) Schwingung zwischen x_1 und x_2 auf. Sie hat die Schwingungsperiode

$$\Delta t = 2(t_1 - t_0) = 2 \int_{x_1}^{x_2} \frac{dx}{\sqrt{\frac{2}{m}(E - V)}} = \sqrt{2m} \int_{x_1}^{x_2} \frac{dx}{\sqrt{E - V}} \tag{4.41}$$

- Als Beispiel kann uns der harmonischer Oszillator dienen, für den die Kraft linear von der Auslenkung abhängt,

$$F(x) = -kx \quad \Rightarrow \quad V(x) = - \int_0^x F(x') dx' = \frac{k}{2} x^2, \tag{4.42}$$

wenn man x_0 so wählt, dass $V(x = 0) = 0$. Also ist die Energie

$$E = \frac{m}{2} \dot{x}^2 + \frac{k}{2} x^2 = \frac{m}{2} (\dot{x}^2 + \omega_0^2 x^2), \tag{4.43}$$

mit $\omega_0 := \sqrt{k/m}$. Die Umkehrpunkte werden erreicht, wenn

$$E = \frac{m}{2} \omega_0^2 x^2 \tag{4.44}$$

ist, also bei

$$x_{1,2} = \pm \sqrt{\frac{2E}{m\omega_0^2}}. \tag{4.45}$$

- Die allgemeine Lösung des harmonischen Oszillators ist ebenfalls durch die Bedingung gegeben, dass seine Energie konstant ist. Mit Hilfe der Umkehrpunkte $x_{1,2} = \pm A_0$ schreiben wir dies als

$$E = \frac{m}{2} A_0^2 \omega_0^2 = \text{konst.} \quad (4.46)$$

Aus der allgemeinen Formel kann man die Schwingungsperiode berechnen und erhält

$$\Delta t = \frac{2\pi}{\omega_0} . \quad (4.47)$$

- Die Bewegungsgleichung für den durch einen quadratischen Reibunsterm gedämpften harmonischen Oszillator lassen sich in der Form

$$\frac{d}{dt} \left[\frac{m}{2} \dot{x}^2 + \frac{k}{2} x^2 \right] = -b \dot{x}^2 , \quad (4.48)$$

also mit $V(x) = (k/2)x^2$ schreiben; die Energie nimmt also ab, denn $-b\dot{x}^2 \leq 0$. Es handelt sich um ein *dissipatives* System.

Kapitel 5

Bewegung in drei Dimensionen

5.1 Kinematik in drei Dimensionen

Bislang haben wir für die Bewegung eines Probekörpers mit und ohne Kräfte zwar immer einen dreidimensionalen Raum angenommen, diesen Raum aber in der eigentlichen Berechnung durch eine geeignete Wahl der Koordinaten immer auf eine Dimension reduziert. Mit Hilfe der Rechenregeln in einem Vektorraum können wir diese Rechnungen aber auch allgemein in drei Dimensionen durchführen.

5.1.1 Bahnkurven

- Nach der allgemeinen Einführung in die Grundlagen der Vektorräume verabreden wir, Vektoren im dreidimensionalen Vektorraum \mathbb{R}^3 mit Pfeilen zu kennzeichnen. In der Regel stellen wir Vektoren als *Spaltenvektoren* dar,

$$\vec{x} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}, \quad (5.1)$$

und kennzeichnen die entsprechenden *Zeilenvektoren* als *transponiert*, $\vec{x}^T = (x_1, x_2, x_3)$. Weiterhin notieren wir das Skalarprodukt durch einen Punkt, $\langle \vec{v}, \vec{w} \rangle =: \vec{v} \cdot \vec{w}$.

- Der Ortsvektor \vec{x} eines Massenpunkts ändert sich im Allgemeinen mit der Zeit t , $\vec{x} = \vec{x}(t)$. Die zwischen zwei Zeiten t_1 und $t_2 > t_1$ durchlaufenen Punkte $\vec{x}(t)$ bilden die *Bahnkurve* des Massenpunkts. In drei Dimensionen wird sie dargestellt durch die drei zeitabhängigen Komponenten

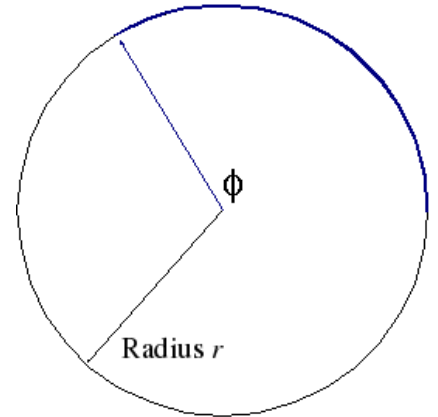
$$\vec{x}(t) = \begin{pmatrix} x_1(t) \\ x_2(t) \\ x_3(t) \end{pmatrix}. \quad (5.2)$$

Beispiele sind eine Kreisbahn in der x_1 - x_2 -Ebene mit Radius R ,

$$\vec{x}_K(t) = R \begin{pmatrix} \cos \phi(t) \\ \sin \phi(t) \\ 0 \end{pmatrix}, \quad (5.3)$$

oder eine Schraubenbahn längs der x_3 -Achse mit Radius R ,

$$\vec{x}_S(t) = \begin{pmatrix} R \cos \phi(t) \\ R \sin \phi(t) \\ x_3(t) \end{pmatrix}. \quad (5.4)$$



Bewegung auf einer Kreisbahn

- Die *Geschwindigkeit* ist wie gewohnt die Ableitung des Ortes nach der Zeit, wobei die Zeitableitung eines Vektors Komponente für Komponente definiert ist

$$\vec{v}(t) = \dot{\vec{x}}(t) = \frac{d\vec{x}(t)}{dt}. \quad (5.5)$$

Die Geschwindigkeiten in den obigen Beispielen sind

$$\vec{v}_K(t) = R \begin{pmatrix} -\dot{\phi} \sin \phi(t) \\ \dot{\phi} \cos \phi(t) \\ 0 \end{pmatrix} \quad \text{und} \quad \vec{v}_S(t) = \begin{pmatrix} -R\dot{\phi} \sin \phi(t) \\ R\dot{\phi} \cos \phi(t) \\ \dot{x}_3(t) \end{pmatrix}. \quad (5.6)$$

- Die *Beschleunigung* ist ebenfalls wie gewohnt die Ableitung der Geschwindigkeit nach der Zeit oder die zweite Ableitung der Ortskoordinate nach der Zeit,

$$\vec{a}(t) = \dot{\vec{v}}(t) = \frac{d\vec{v}(t)}{dt} = \ddot{\vec{x}}(t) = \frac{d^2\vec{x}(t)}{dt^2}. \quad (5.7)$$

Die Beschleunigungen in den obigen Beispielen sind also

$$\vec{a}_K(t) = R \begin{pmatrix} -\ddot{\phi} \sin \phi(t) - \dot{\phi}^2 \cos \phi(t) \\ \ddot{\phi} \cos \phi(t) - \dot{\phi}^2 \sin \phi(t) \\ 0 \end{pmatrix} \quad \text{und} \\ \vec{a}_S(t) = \begin{pmatrix} -R\ddot{\phi} \sin \phi(t) - R\dot{\phi}^2 \cos \phi(t) \\ R\ddot{\phi} \cos \phi(t) - R\dot{\phi}^2 \sin \phi(t) \\ \ddot{x}_3(t) \end{pmatrix}. \quad (5.8)$$

5.1.2 Bogenlänge, Tangential- und Normalvektoren

- Wenn wir die Richtung einer Bahnkurve zu jedem Zeitpunkt t kennen, dann können wir ein beliebig kleines (infinitesimales) Element der *Bogenlänge* längs der Bahnkurve als

$$ds = |d\vec{x}(t)| = \left| \frac{d\vec{x}}{dt} \right| dt = |\vec{v}(t)| dt \quad (5.9)$$

definiere. Die integrierte Bogenlänge

$$s = \int ds = \int_0^t |\vec{v}(t')| dt' . \quad (5.10)$$

entspricht dann der zwischen den Zeiten 0 und t zurückgelegten Strecke.

Oft ist diese Umparametrisierung von der Zeit t zur Bogenlänge s hilfreich. Das ist möglich, weil s mit der Zeit monoton wächst; $\vec{x}(t) \rightarrow \vec{x}(s)$. Die Jacobi-Determinanten dieser Substitution lässt sich aus dem Geschwindigkeitsfeld berechnen.

- Die Richtung der Bewegung ist jetzt durch den *Tangentialvektor* an die Bahnkurve, also die Ableitung der Bahnkurve nach der Bogenlänge, gegeben

$$\vec{\tau} = \frac{d\vec{x}}{ds} \quad \Leftrightarrow \quad d\vec{x} = \vec{\tau} ds . \quad (5.11)$$

Damit kann man einen infinitesimalen Vektor einer Bahnkurve in der Tat in eine Richtung und in die infinitesimale Bogenlänge aufspalten. Aufgrund der Definition der Bogenlänge s ist $\vec{\tau}$ ein Einheitsvektor, $|\vec{\tau}| = 1$. Der Vektor

$$\vec{n}_H = \frac{d\vec{\tau}}{ds} \left| \frac{d\vec{\tau}}{ds} \right|^{-1} \quad (5.12)$$

heißt *Hauptnormalenvektor*. Er steht senkrecht auf $\vec{\tau}$,

$$\vec{\tau}^2 = 1 \Rightarrow \frac{d(\vec{\tau}^2)}{ds} = 2\vec{\tau} \cdot \frac{d\vec{\tau}}{ds} = 0 . \quad (5.13)$$

Der zweite von \vec{n}_H linear unabhängigen Normalenvektor zu $\vec{\tau}$ ist der *Binormalenvektor*

$$\vec{n}_B = \vec{\tau} \times \vec{n}_H . \quad (5.14)$$

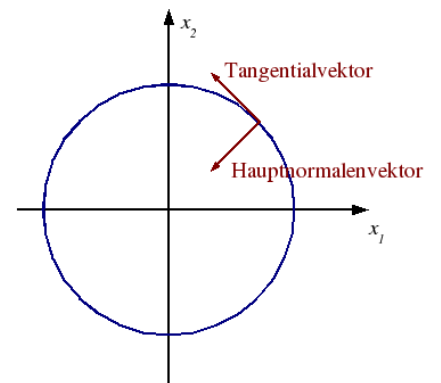
Damit bilden τ , n_H und n_B eine orthogonale Basis unseres dreidimensionalen Raumes.

- Als einfaches Beispiel dient die Kreisbewegung mit konstanter Geschwindigkeit. Hier ergibt sich die Bogenlänge

$$s = \int ds = \int_0^t |\vec{v}(t')| dt' = \int_0^t R\omega dt' = R\omega t . \quad (5.15)$$

Im letzten Schritt haben wir benutzt dass wegen $|\vec{v}(t)| = R\dot{\phi}$ die *Winkelgeschwindigkeit* $\dot{\phi} = \omega$ konstant ist. Die Relation (5.15) erlaubt es uns die bekannte Bahnkurve einer Kreisbewegung durch s auszudrücken,

$$\vec{x}(t) = R \begin{pmatrix} \cos(\omega t) \\ \sin(\omega t) \\ 0 \end{pmatrix} = R \begin{pmatrix} \cos(s/R) \\ \sin(s/R) \\ 0 \end{pmatrix} = \vec{x}(s) . \quad (5.16)$$



Tangential- und Hauptnormalenvektor bei der Kreisbewegung

Aus dieser Form können wir den Tangentialvektor

$$\vec{\tau} = \frac{d\vec{x}(s)}{ds} = \begin{pmatrix} -\sin(s/R) \\ \cos(s/R) \\ 0 \end{pmatrix} = \begin{pmatrix} -\sin \omega t \\ \cos \omega t \\ 0 \end{pmatrix}, \quad (5.17)$$

den Hauptnormalenvektor

$$\vec{n}_H = \frac{1}{R} \begin{pmatrix} -\cos(s/R) \\ -\sin(s/R) \\ 0 \end{pmatrix} \left(\frac{1}{R}\right)^{-1} = - \begin{pmatrix} \cos(s/R) \\ \sin(s/R) \\ 0 \end{pmatrix} \quad (5.18)$$

und den Binormalenvektor

$$\vec{n}_B = \begin{pmatrix} -\sin(s/R) \\ \cos(s/R) \\ 0 \end{pmatrix} \times \begin{pmatrix} -\cos(s/R) \\ -\sin(s/R) \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}, \quad (5.19)$$

ablesen.

5.1.3 Tangential- und Normalkomponenten

- Allgemein definiert man den lokalen Krümmungsradius in Analogie zum Radius des Kreises als die Länge des nicht normierten Hauptnormalenvektors

$$\rho := \left| \frac{d\vec{\tau}}{ds} \right|^{-1} \quad (5.20)$$

Je stärker der Tangentialvektor sich entlang der Bogenlänge ändert, desto größer ist die Krümmung und desto kleiner ist der Radius ρ .

- Die Geschwindigkeit und die Beschleunigung können in ihre tangentialen und normalen Komponenten zerlegt werden, zunächst also

$$\vec{v}(t) = \frac{d\vec{x}(t)}{dt} = \frac{d\vec{x}}{ds} \frac{ds}{dt} = |\vec{v}(t)| \cdot \vec{\tau}. \quad (5.21)$$

Die Geschwindigkeit ist also tangential zur Bahnkurve. Für die Beschleunigung ergibt sich

$$\begin{aligned} \frac{d\vec{v}(t)}{dt} &= \frac{d(v\vec{\tau})}{dt} = \dot{v}\vec{\tau} + v \frac{d\vec{\tau}}{dt} = \dot{v}\vec{\tau} + v \frac{d\vec{\tau}}{ds} \frac{ds}{dt} \\ &= \dot{v}\vec{\tau} + v \frac{\vec{n}_H}{\rho} |\vec{v}| = \dot{v}\vec{\tau} + \frac{v^2}{\rho} \vec{n}_H. \end{aligned} \quad (5.22)$$

mit $|\vec{v}| = v$. Sie hat also eine tangentiale Komponente und eine Komponente in Richtung der Hauptnormalen, die mit zunehmendem Krümmungsradius abnimmt.

5.1.4 Kurvenintegrale

- Die Einführung des Energiesatzes durch einen integrierenden Faktor hat ergeben, dass die Energie in einer Dimension erhalten ist, wenn die Kraft als negative Ableitung eines Potentials nach dem Ort dargestellt werden kann,

$$F(x) = -\frac{dV(x)}{dx} \Leftrightarrow V(x) = -\int_{x_0}^x F(x')dx'. \quad (5.23)$$

Die Frage ist, welche der beiden Formulierungen wir am besten in drei Dimensionen verallgemeinern können. Das Integral können wir einfach auf ein Integral entlang einer Bahnkurve in drei Dimensionen verallgemeinern. Der Anfangspunkt x_0 der Integration war beliebig, aber in drei Dimensionen müssen wir den Integrationsweg festlegen. Die offensichtliche Frage ist dann, wann man ein Potential $V(\vec{x})$ einführen kann, das vom Integrationsweg unabhängig wird. Hierzu beginnen wir mit einer Diskussion der Kurvenintegrale.

- Eine endlich lange Kurve C lässt sich durch eine Funktion darstellen, die einem Kurvenparameter t einen Punkt $\vec{x}(t)$ im Raum zuordnet, wobei der Kurvenparameter einem bestimmten endlichen oder unendlichen Intervall $I = [t_a, t_b]$ entnommen ist,

$$C : I \rightarrow \mathbb{R}^3, \quad t \mapsto \vec{x}(t). \quad (5.24)$$

Die Endpunkte der Kurve sind $\vec{x}_a = \vec{x}(t_a)$ und $\vec{x}_b = \vec{x}(t_b)$. Betrachten wir nun eine vektorwertige Funktion $\vec{A}(\vec{x})$ in \mathbb{R}^3 , die wir entlang dieser Kurve integrieren wollen. Einer kleinen Verschiebung von \vec{x} zu $\vec{x} + \delta\vec{x}$ können wir den Skalar

$$\delta\Phi = \vec{A}(\vec{x}) \cdot \delta\vec{x} \quad (5.25)$$

zuordnen. Um nun das *Kurvenintegral* über $\vec{A}(\vec{x})$ zu berechnen, beginnen wir bei \vec{x}_a und gehen in infinitesimalen Schritten $d\vec{x}$ längs der Kurve nach \vec{x}_b . In jedem Schritt berechnen wir $\vec{A}(\vec{x}) \cdot d\vec{x}$ und summieren alle Ergebnisse auf. Damit erhalten wir das Kurvenintegral

$$\Phi = \int_{\vec{x}_a}^{\vec{x}_b} \vec{A}(\vec{x}) \cdot d\vec{x}. \quad (5.26)$$

Das Kurvenintegral heißt geschlossen, wenn Anfangs- und Endpunkt übereinstimmen, $\vec{x}_a = \vec{x}_b$. Es wird dann durch

$$\int_{\vec{x}_a}^{\vec{x}_b} \vec{A}(\vec{x}) \cdot d\vec{x} = \oint_C \vec{A}(\vec{x}) \cdot d\vec{x} \quad (5.27)$$

gekennzeichnet.

- Bei der praktischen Berechnung geht man entsprechend vor: Man stellt das gerichtete Wegelement $d\vec{x}$ mithilfe des Kurvenparameters t dar,

$$d\vec{x} = \frac{d\vec{x}}{dt} dt, \quad (5.28)$$

so dass das Kurvenintegral in das gewöhnliche, eindimensionale Integral

$$\int_{\vec{x}_a}^{\vec{x}_b} \vec{A}(\vec{x}) \cdot d\vec{x} = \int_{t_a}^{t_b} \vec{A}(\vec{x}) \cdot \frac{d\vec{x}}{dt} dt \quad (5.29)$$

übergeht, das dann wie gewohnt ausgeführt werden kann.

5.2 Differentialoperatoren

Nach dieser kurzen Diskussion der Integration müssen wir nun die Ableitung vektorwertige Funktionen im n -dimensionalen Raum betrachten.

5.2.1 Felder, Partielle Ableitungen und der Gradient

- Ein *Feld* f ist eine Funktion des Raumes, d.h. eine Funktion, die jedem Raumpunkt \vec{x} einen Wert $f(\vec{x})$ zuordnet. Wenn das Feld jedem Raumpunkt einen Vektor zuordnet $\vec{f}(\vec{x})$ heißt es *Vektorfeld*. Solange wir ein solches Vektorfeld nur als N -Tupel-Feld betrachten ist es nicht anderes als eine Anordnung verschiedener Skalarfelder. Später werden wir sehen, dass diese Skalarfelder auf eine spezielle Art miteinander verknüpft sind, aber im Moment können wir diesen Aspekt problemlos vernachlässigen.
- Was wir hingegen verstehen müssen ist wie wir Felder behandeln, deren Argument ein Vektor ist. Wir erinnern uns zuert daran, dass für Funktionen $f(x)$ die Ableitung nach x eindeutig als

$$f'(x) = \frac{df(x)}{dx} := \lim_{\epsilon \rightarrow 0} \frac{f(x + \epsilon) - f(x)}{\epsilon} \quad (5.30)$$

definiert ist. Für Funktionen mehrerer Variabler definiert man eine *partielle Ableitung*, bei der nur die Änderung längs einer Koordinatenrichtung untersucht wird,

$$\frac{\partial f(x_1, \dots, x_N)}{\partial x_i} := \lim_{\epsilon \rightarrow 0} \frac{f(x_1, \dots, x_i + \epsilon, \dots, x_N) - f(x_1, \dots, x_i, \dots, x_N)}{\epsilon}, \quad (5.31)$$

während jeweils alle anderen Koordinaten konstant gehalten werden. Die partielle Ableitung nach der Koordinate x_i wird oft durch andere Schreibweisen abgekürzt, von denen

$$\frac{\partial f(x_1, \dots, x_N)}{\partial x_i} =: \partial_i f(x_1, \dots, x_N) =: f_{,i}(x_1, \dots, x_N) \quad (5.32)$$

die gebräuchlichsten sind.

- Zum Beispiel hat die Funktion

$$f(x_1, x_2, x_3) = \frac{\sin x_1 \cos x_2}{x_3} \quad (5.33)$$

die partiellen Ableitungen

$$\begin{aligned} \partial_1 f(x_1, x_2, x_3) &= \frac{\cos x_1 \cos x_2}{x_3} \\ \partial_2 f(x_1, x_2, x_3) &= -\frac{\sin x_1 \sin x_2}{x_3} \\ \partial_3 f(x_1, x_2, x_3) &= -\frac{\sin x_1 \cos x_2}{x_3^2} . \end{aligned} \quad (5.34)$$

- Für die partielle Ableitung gelten offensichtlich dieselben Regeln wie für die gewöhnliche Ableitung, nämlich die Produktregel

$$\partial_i(fg) = (\partial_i f)g + f(\partial_i g) \quad (5.35)$$

und die Kettenregel

$$\partial_i(f(g)) = \frac{df}{dg} \partial_i g , \quad (5.36)$$

wobei f eine Funktion einer skalaren Größe und g eine Funktion der Koordinaten \vec{x}_i ist.

- Mit Hilfe der Kettenregel definieren wir die *totale Ableitung* einer Feldfunktion $f(\vec{x})$ als

$$\frac{df(\vec{x}(t))}{dt} := \frac{\partial f(\vec{x}(t))}{\partial x_i} \frac{dx_i(t)}{dt} = \partial_i f(\vec{x}(t)) v_i(t) . \quad (5.37)$$

Wenn wir die Kurve $\vec{x}(t)$ und damit $\vec{v}(t)$ kennen, dann beschreibt dies die Ableitung des Feldes f entlang dieser Kurve. Beispielsweise ist die totale Ableitung der Funktion $f(\vec{x})$ aus (5.33) längs der Kurve $\vec{x}(t) = (v_1 t, v_2 t, 1)$ durch

$$\frac{df}{dt} = v_1 \partial_1 f + v_2 \partial_2 f = v_1 \cos(v_1 t) \cos(v_2 t) - \sin(v_1 t) \sin(v_2 t) \quad (5.38)$$

gegeben.

- Der *Gradient* $\vec{\nabla}f$ einer skalaren Funktion $f(\vec{x})$ ist als der N -dimensionale Vektor definiert, dessen Komponenten

$$\vec{\nabla}f(\vec{x}) = \begin{pmatrix} \partial_1 f \\ \vdots \\ \partial_N f \end{pmatrix} \quad \text{bzw.} \quad (\vec{\nabla}f)_i = \partial_i f \quad (5.39)$$

betragen. Das Zeichen ∇ wird „Nabla“ ausgesprochen. Einem stetig differenzierbaren, skalaren Feld $f(\vec{x})$ ordnet der Gradient in jedem Punkt einen Vektor $\vec{\nabla}f$ zu. Der vektorwertige Operator

$$\vec{\nabla} = \vec{e}_1 \frac{\partial}{\partial x_1} + \dots + \vec{e}_N \frac{\partial}{\partial x_N} = \vec{e}_i \partial_i \quad (5.40)$$

heißt Nabla-Operator. Ausgedrückt durch den Nabla-Operator lautet die totale Ableitung (5.37) einer Funktion $f(\vec{x})$

$$\frac{df(\vec{x}(t))}{dt} = \vec{v} \cdot \vec{\nabla}f(\vec{x}(t)). \quad (5.41)$$

Die *Richtungsableitung* eines skalaren Feldes längs einer Richtung, die durch den Einheitsvektor \vec{e} gegeben ist, ist das Skalarprodukt aus dem Gradienten und dem Vektor \vec{e} ,

$$\vec{e} \cdot \vec{\nabla}f(\vec{x}). \quad (5.42)$$

Liegt \vec{e} in einer Fläche, in der die Funktion f konstant ist, muss die Richtungsableitung verschwinden. Für $\vec{\nabla}f \cdot \vec{e} = 0$ steht der Gradient also senkrecht auf Flächen konstanter Funktionswerte. Der Gradient zeigt demnach die Richtung des steilsten Anstiegs der Funktion $f(\vec{x})$ an.

5.2.2 Divergenz und Rotation

- Wenn wir den Nabla-Operator unabhängig von seiner Herleitung betrachten, dann kann man ihn offensichtlich auch auf ein Vektorfeld anwenden. Die *Divergenz* $\vec{\nabla} \cdot \vec{f}(\vec{x})$ eines Vektorfeldes $\vec{f}(\vec{x})$ kann als das Skalarprodukt des Nabla-Operators mit dem Vektorfeld $\vec{f}(\vec{x})$ aufgefasst werden. Sie ist durch die Summe der partiellen Ableitungen

$$\vec{\nabla} \cdot \vec{f}(\vec{x}) = \sum_{i=0}^N \frac{\partial f_i}{\partial x_i} = \partial_i f_i \quad (5.43)$$

definiert. Wenn \vec{f} überall differenzierbar ist, ordnet die Divergenz jedem Punkt des Raumes einen Skalar zu.

- Alternativ kann man den Nabla-Operator auf ein Vektorfeld durch das Vektorprodukt anwenden. Die *Rotation* $\vec{\nabla} \times \vec{f}(\vec{x})$ eines Vektorfeldes $\vec{f}(\vec{x})$ ist

$$\vec{\nabla} \times \vec{f}(\vec{x}) = \epsilon_{ijk} \partial_i f_j(\vec{x}) \vec{e}_k . \quad (5.44)$$

Sie ordnet einem differenzierbaren Vektorfeld ein Vektorfeld zu.

- Zahlreiche wichtige Rechenregeln für den Umgang mit dem Nabla-Operator lassen sich mithilfe der Produkt- und der Kettenregel sowie der Beziehung (4.14) herleiten. Seien f und g skalare Felder sowie \vec{v} und \vec{w} Vektorfelder, dann gelten zum Beispiel die Formeln:

$$\begin{aligned} \vec{\nabla}(fg) &= f\vec{\nabla}g + g\vec{\nabla}f \\ \vec{\nabla} \cdot (f\vec{v}) &= \vec{\nabla}f \cdot \vec{v} + f\vec{\nabla} \cdot \vec{v} \\ \vec{\nabla} \times (f\vec{v}) &= \vec{\nabla}f \times \vec{v} + f\vec{\nabla} \times \vec{v} \\ \vec{\nabla} \times (\vec{v} \times \vec{w}) &= (\vec{w} \cdot \vec{\nabla})\vec{v} + \vec{v}(\vec{\nabla} \cdot \vec{w}) - (\vec{v} \cdot \vec{\nabla})\vec{w} + \vec{w}(\vec{\nabla} \cdot \vec{v}) \\ \vec{\nabla} \times (\vec{\nabla} \times \vec{v}) &= \vec{\nabla}(\vec{\nabla} \cdot \vec{v}) - (\vec{\nabla} \cdot \vec{\nabla})\vec{v} . \end{aligned} \quad (5.45)$$

Der Operator

$$\vec{\nabla}^2 := \vec{\nabla} \cdot \vec{\nabla} = \partial_i \partial_i =: \Delta \quad (5.46)$$

heißt auch Laplace-Operator.

5.3 Energiesatz in drei Dimensionen

Nachdem wir nun in mehreren Dimensionen differenzieren und integrieren können, sollten wir unseren eindimensionalen Energiesatz verallgemeinern können.

5.3.1 Energieerhaltung bei Potentialkräften

- Aus der nun vektorwertigen Bewegungsgleichung folgt nach Multiplikation mit $\dot{\vec{x}}$ als vektorwertigem integrierendem Faktor

$$m\ddot{\vec{x}} = \vec{F} \quad \Rightarrow \quad \frac{d}{dt} \left(\frac{m}{2} \dot{\vec{x}}^2 \right) = \vec{F} \cdot \frac{d\vec{x}}{dt} . \quad (5.47)$$

Analog zum eindimensionalen Fall ist

$$T = \frac{m}{2} \dot{\vec{x}}^2 = \frac{\vec{p}^2}{2m} \quad (5.48)$$

die *kinetische Energie*. Wenn wir (5.47) als Differential betrachten, dann ist die infinitesimale Änderung der kinetischen Energie

gleich der von der Kraft \vec{F} längs des Wegelements $d\vec{x}$ verrichtete Arbeit,

$$d\left(\frac{m}{2}\dot{\vec{x}}^2\right) = \vec{F} \cdot d\vec{x}. \quad (5.49)$$

Die *Leistung* ist die pro Zeiteinheit verrichtete Arbeit,

$$\text{Leistung} = \frac{\text{Arbeit}}{\text{Zeiteinheit}} = \vec{F} \cdot \frac{d\vec{x}}{dt}. \quad (5.50)$$

- Um aus der Bewegungsgleichung auf eine Erhaltungsgröße zu schließen, müsste sich die rechte Seite in (5.47) als Zeitableitung einer anderen Funktion schreiben lassen. Das ist allgemein nicht der Fall, gilt aber für die Klasse der Potentialkräfte mit

$$\vec{F} = -\vec{\nabla}V(\vec{x}). \quad (5.51)$$

Für solche Kräfte gilt nach der Definition der totalen Ableitung

$$\vec{F} \cdot \frac{d\vec{x}}{dt} = -\vec{\nabla}V \cdot \frac{d\vec{x}}{dt} = -\frac{dV(\vec{x})}{dt}. \quad (5.52)$$

Damit ist dann wieder

$$\frac{m}{2}\dot{\vec{v}}^2 + V(\vec{x}) = T + V = E = \text{konst.}, \quad (5.53)$$

und die Energie ist erhalten. Potentialkräfte $\vec{F} = -\vec{\nabla}V$ heißen daher *konservativ*, andere *dissipativ*.

5.3.2 Beispiele für Potentialkräfte

- Ein einfaches Beispiel liefert der harmonische Oszillator in drei Dimensionen mit dem Potential

$$V = \frac{k}{2}\vec{x}^2 \quad \Rightarrow \quad \vec{F} = -\vec{\nabla}V = -k\vec{x}. \quad (5.54)$$

- Ein zweites Beispiel sind Zentralkräfte, die zu einem festen Zentrum hin oder davon weg zeigen. Der Betrag der Kraft hängt nur vom Abstand von diesem Zentrum ab. Wenn man das Zentrum in den Ursprung legt, ist

$$\vec{F} = F(r)\vec{e}_r = F(r)\frac{\vec{x}}{r}. \quad (5.55)$$

Wir zeigen dass solche Kräfte immer ein Potential haben, indem wir das Potential ausrechnen. Mit dem Ansatz

$$V(r) = -\int_{r_0}^r F(r')dr' \quad (5.56)$$

erhalten wir für die Kraft in drei Dimensionen

$$\begin{aligned}
 F_i &= -\frac{\partial}{\partial x_i} V(r) \\
 &= \frac{\partial}{\partial x_i} \int_{r_0}^r F(r') dr' = \frac{\partial r}{\partial x_i} \frac{d}{dr} \int_{r_0}^r F(r') dr' = F(r) \frac{x_i}{r}. \quad (5.57)
 \end{aligned}$$

Das ist genau die Kraft in (5.55).

5.4 Konservative Kraftfelder

5.4.1 Unabhängigkeit vom Weg

Schon bei der Herleitung des dreidimensionalen Energiesatzes haben wir zwei Dinge gesehen: Erstens ergibt die Integraldarstellung des Potentials in (5.23) nur einen Sinn, wenn das Kurvenintegral unabhängig vom Weg ist. Zweitens gibt es Klassen von Kräften, die sich mit Hilfe eines solchen Potentials schreiben lassen. Eine mathematische Verknüpfung würde es uns ersparen, konservative Kräfte immer wieder durch das Berechnen des Potentials zu erkennen.

- Nehmen wir also zunächst an, das Kurvenintegral

$$\Phi(\vec{x}) = \int_{\vec{x}_a}^{\vec{x}} \vec{A}(\vec{x}') \cdot d\vec{x}' \quad (5.58)$$

sei unabhängig vom Weg. Wir betrachten es hier als Funktion des Endpunkts der Kurve. Wir verlängern sie um ein kleines Stück $\delta\vec{x}$, betrachten die Differenz für beliebig kleine Verschiebungen $\delta\vec{x}$, und vergleichen sie mit der totalen Ableitung

$$\begin{aligned}
 \Phi(\vec{x} + \delta\vec{x}) - \Phi(\vec{x}) &= \vec{\nabla}\Phi(\vec{x}) \cdot d\vec{x} \\
 &= \int_{\vec{x}}^{\vec{x}+\delta\vec{x}} \vec{A}(\vec{x}') \cdot d\vec{x}' = \vec{A}(\vec{x}) \cdot d\vec{x} \\
 \Rightarrow \vec{A} &= \vec{\nabla}\Phi. \quad (5.59)
 \end{aligned}$$

Der letzte Schritt ist einfach der Hauptsatz der Integralrechnung. Das bedeutet dass für ein vom Weg unabhängiges Kurvenintegral der Integrand ein Gradientenfeld sein muss.

- Wenn wir umgekehrt annehmen dass der Integrand ein Gradientenfeld ist, dann ist das Kurvenintegral

$$\int_{\vec{x}_a}^{\vec{x}_b} \vec{A} \cdot d\vec{x} = \int_{\vec{x}_a}^{\vec{x}_b} \vec{\nabla}\Phi \cdot d\vec{x} = \int_{\Phi(\vec{x}_a)}^{\Phi(\vec{x}_b)} d\Phi = \Phi(\vec{x}_b) - \Phi(\vec{x}_a) \quad (5.60)$$

vom Weg unabhängig. Kombiniert haben wir also gezeigt, dass ein Kurvenintegral über ein Vektorfeld \vec{A} dann und nur dann vom Weg unabhängig ist, wenn es ein Skalarfeld Φ gibt, dessen Gradient \vec{A} ist, $\vec{A} = \vec{\nabla}\Phi$.

- Wie können wir einfach entscheiden, ob \vec{A} ein Gradientenfeld ist? Wir können direkt zeigen, dass für jedes Gradientenfeld die Rotation verschwindet,

$$\vec{\nabla} \times \vec{A} = \vec{\nabla} \times \vec{\nabla}\Phi = \epsilon_{ijk} \partial_i \partial_j \Phi \vec{e}_k \equiv 0. \quad (5.61)$$

- In die andere Richtung nehmen wir $\vec{\nabla} \times \vec{A} = 0$ an. Wir schreiben das Kurvenintegral über \vec{A} längs eines speziellen Weges, der stückweise parallel zu den drei Koordinatenachsen führt und definieren uns dadurch ein mögliches Potential Φ ,

$$\Phi(\vec{x}) = \int_{x_a}^x A_x(x', y_a, z_a) dx' + \int_{y_a}^y A_y(x, y', z_a) dy' + \int_{z_a}^z A_z(x, y, z') dz', \quad (5.62)$$

von dem uns lediglich die Existenz interessiert. Die drei partiellen Ableitungen von Φ sind dann

$$\begin{aligned} \partial_z \Phi &= A_z(x, y, z) \\ \partial_y \Phi &= A_y(x, y, z_a) + \int_{z_a}^z \partial_y A_z(x, y, z') dz' \\ \partial_x \Phi &= A_x(x, y_a, z_a) + \int_{y_a}^y \partial_x A_y(x, y', z_a) dy' + \int_{z_a}^z \partial_x A_z(x, y, z') dz'. \end{aligned} \quad (5.63)$$

In der z -Komponente bedeutet das schon $\vec{\nabla}\Phi = \vec{A}$. Unsere Annahme $\vec{\nabla} \times \vec{A} = 0$ gibt uns die drei Bedingungen $\partial_y A_z = \partial_z A_y$, $\partial_x A_y = \partial_y A_x$ und $\partial_x A_z = \partial_z A_x$ und damit

$$\partial_y \Phi = A_y(x, y, z_a) + \int_{z_a}^z \partial_z A_y(x, y, z') dz' = A_y(x, y, z) \quad (5.64)$$

und entsprechend $\partial_x \Phi = A_x(x, y, z)$. Damit haben wir $\vec{\nabla}\Phi = \vec{A}$ für alle drei Komponenten bewiesen und in zwei Schritten gezeigt, dass $\vec{\nabla} \times \vec{A} = 0$ äquivalent dazu ist, dass \vec{A} ein Gradientenfeld ist.

- Diese Schlussfolgerung gilt aber nur dann, wenn \vec{A} überall längs des Integrationsweges definiert ist. Deswegen muss man die Einschränkung anbringen, dass die obige Aussage nur in einem Raumbereich gilt, der „keine Löcher hat“, so dass kein möglicher Integrationsweg durch einen Bereich läuft, in dem \vec{A} nicht definiert ist. Dies wird präzise durch die Definition eines *einfach zusammenhängenden Gebiets* formuliert: Ein Gebiet G heißt einfach zusammenhängend, wenn jede geschlossene Kurve in G stetig zu einem Punkt zusammengezogen werden kann, ohne das Gebiet zu verlassen.

- Wir haben also gezeigt, dass Kraftfelder \vec{F} , die auf einem einfach zusammenhängenden Gebiet G definiert sind, dann und nur dann ein Potential haben, wenn sie in G wirbelfrei sind, $\vec{\nabla} \times \vec{F} = 0$. Und genau unter dieser Bedingung ist die Darstellung des Potentials als Kurvenintegral auch unabhängig vom Weg.

5.4.2 Der Satz von Stokes

Es bietet sich an dieser Stelle an, einen der beiden Integralsätze zu besprechen, die in der Physik sehr wichtig sind. Der Satz von Stokes ersetzt Integrale über Vektorfelder durch Integrale über Differentialoperatoren dieser Vektorfelder.

- Wir betrachten zunächst ein Vektorfeld \vec{A} und ein geschlossenes Kurvenintegral über \vec{A} längs eines infinitesimal kleinen Weges in der x - y -Ebene. Ausgehend vom Punkt (x_0, y_0) gehen wir jeweils parallel zu den Achsen zu den Punkten $(x_0 + dx, y_0)$, $(x_0 + dx, y_0 + dy)$, $(x_0, y_0 + dy)$ und zurück zu (x_0, y_0) . Dann lässt sich das geschlossene Kurvenintegral in der Form

$$\oint_C \vec{A} \cdot d\vec{x} = \int_1 A_x(x_0, y_0)dx + \int_2 A_y(x_0 + dx, y_0)dy \quad (5.65)$$

$$- \int_3 A_x(x_0 + dx, y_0 + dy)dx - \int_4 A_y(x_0, y_0 + dy)dy$$

schreiben, worin die Minuszeichen daher kommen, dass man längs des 3. und des 4. Wegstücks gegen die Achsenrichtung läuft. Mit Hilfe von Taylor-Näherungen der Art

$$A_y(x_0 + dx, y_0) = A_y(x_0, y_0) + \partial_x A_y(x_0, y_0)dx, \quad (5.66)$$

bestimmen wir die Komponenten des Feldes \vec{A} an den Eckpunkten des Weges. Das Integral in (5.65) nähern wir auf diese Weise durch Flächeninhalte kleiner Rechtecke an,

$$\oint_C \vec{A} \cdot d\vec{x} = \int_1 A_x(x_0, y_0)dx + \int_2 [A_y(x_0, y_0) + \partial_x A_y(x_0, y_0)dx] dy$$

$$- \int_3 [A_x(x_0, y_0) + \partial_y A_x(x_0, y_0)dy] dx - \int_4 A_y(x_0, y_0)dy$$

$$= \int (\partial_x A_y - \partial_y A_x) dx dy = \int (\vec{\nabla} \times \vec{A})_z dx dy. \quad (5.67)$$

Das Kurvenintegral über die infinitesimal kleine, geschlossene Kurve erweist sich also identisch zur Rotation des Feldes \vec{A} in Richtung der Senkrechten zur Fläche, multipliziert mit dem Flächeninhalt.

- Wir können dieses Ergebnis auf beliebige Flächen S und ihre (geschlossenen) Randkurven ∂S erweitern, indem wir Flächen in infinitesimale, benachbarte Zellen ds zerlegen und die Zirkulation um ihre Randkurven aufsummieren. Dabei fallen die Beiträge aller Wege heraus, die nicht Teile der Randkurve sind, und wir müssen berücksichtigen, dass wir jeder Zelle eine Richtung geben müssen, die auf ihr senkrecht steht und so orientiert ist, dass sie mit der Orientierung ihrer Randkurve ein Rechtssystem bildet, also in (5.67) infinitesimal

$$\vec{e}_z dx dy =: d\vec{s}. \quad (5.68)$$

Das Ergebnis ist der *Stokes'sche Satz*

$$\oint_{\partial S} \vec{A} \cdot d\vec{x} = \int_S (\vec{\nabla} \times \vec{A}) \cdot d\vec{s}, \quad (5.69)$$

der besagt, dass das Flächenintegral über die Rotation eines Vektorfeldes gleich seiner Zirkulation über die Randkurve der Fläche ist. Wieder gilt die Einschränkung, dass die Fläche S ein einfach zusammenhängendes Gebiet darstellt, weil sich anderenfalls nicht alle Beiträge der Wegelemente im Inneren der Fläche herausheben.

Der Stokes'sche Satz zeigt dann, dass das Kurvenintegral eines Vektorfeldes längs eines geschlossenen Weges genau dann verschwindet, wenn das Integral über seine Rotation innerhalb der eingeschlossenen Fläche verschwindet.

- Dass die Einschränkung auf zusammenhängende Gebiete sehr wesentlich ist, zeigt das folgende Beispiel. Gegeben sei das Kraftfeld

$$\vec{F}(\vec{x}) = \frac{1}{x_1^2 + x_2^2} \begin{pmatrix} -x_2 \\ x_1 \\ 0 \end{pmatrix}, \quad (5.70)$$

dessen Rotation verschwindet,

$$\vec{\nabla} \times \vec{F}(\vec{x}) = \begin{pmatrix} \partial_2 F_3 - \partial_3 F_2 \\ \partial_3 F_1 - \partial_1 F_3 \\ \partial_1 F_2 - \partial_2 F_1 \end{pmatrix} = 0. \quad (5.71)$$

Das Integral über \vec{F} längs der geschlossenen Kreiskurve

$$\vec{x}(\phi) = R \begin{pmatrix} \cos \phi \\ \sin \phi \\ 0 \end{pmatrix}, \quad 0 \leq \phi < 2\pi \quad (5.72)$$

beträgt aber

$$\oint \vec{F} \cdot d\vec{x} = \int_0^{2\pi} \vec{F} \cdot \frac{d\vec{x}}{d\phi} d\phi = \int_0^{2\pi} (\sin^2 \phi + \cos^2 \phi) d\phi = 2\pi \neq 0. \quad (5.73)$$

Die Kraft ist *nicht* konservativ, denn wegen der Unstetigkeit des Kraftfeldes für $x_1 = 0 = x_2$ ist das von der Kurve eingeschlossene Gebiet nicht einfach zusammenhängend.

5.5 Krummlinig-orthogonale Koordinaten

- Wir haben bisher vor allem kartesische Koordinaten verwendet. Man kann sie dadurch kennzeichnen, dass man drei raumfeste Koordinatenebenen einführt, in denen jeweils eine Koordinate konstant ist, nämlich x_1 in der x_2 - x_3 -, x_2 in der x_1 - x_3 - und x_3 in der x_1 - x_2 -Ebene. Wir haben aber auch schon gesehen, dass sie für die Beschreibung physikalischer Systeme manchmal ungeschickt sind, weil die physikalischen Größen auf Flächen konstant sind, die nicht mit den Koordinatenflächen zusammenfallen. Ein Beispiel sind kugelsymmetrische Systeme, deren physikalische Eigenschaften auf Flächen mit konstantem Radius, also Kugeln, konstant sind.
- Zur Verallgemeinerung denken wir uns im dreidimensionalen Raum drei Koordinatenflächen, die dadurch definiert sind, dass auf ihnen jeweils ein Parameter q_i konstant ist. Punkte im Raum kann man dann immer durch drei Koordinatenflächen beschreiben, die sich dort schneiden. Die zu ihnen gehörenden Parameter (q_1, q_2, q_3) verwenden wir als Koordinaten. Die Einheitsnormalenvektoren der Koordinatenflächen werden als neue Basisvektoren \vec{e}_i verwendet.
- Abstände im Raum dürfen nicht von den Koordinaten abhängen. Wegen des Satzes von Pythagoras ist der Abstand ds zweier infinitesimal benachbarter Punkte im euklidischen Raum durch die Summe der quadratischen Abstände dx_i^2 gegeben. Durch Substitution in den q -Raum erhalten wir, wenn wir die beiden beteiligten Summe ausnahmsweise ausschreiben

$$\begin{aligned}
 ds^2 &= dx_1^2 + dx_2^2 + dx_3^2 \\
 &= \sum_i dx_i^2 = \sum_{i,j} \left(\frac{\partial x_i}{\partial q_j} dq_j \right)^2 = \sum_i \left(\frac{\partial x_i}{\partial q_j} \right)^2 dq_j^2 \\
 &= \sum_j dq_j^2 \sum_i \left(\frac{\partial x_i}{\partial q_j} \right)^2 =: \sum_j dq_j^2 h_j^2, \quad (5.74)
 \end{aligned}$$

wobei wir h_j durch

$$h_j^2 = \left(\frac{\partial x_1}{\partial q_j} \right)^2 + \left(\frac{\partial x_2}{\partial q_j} \right)^2 + \left(\frac{\partial x_3}{\partial q_j} \right)^2 \quad (5.75)$$

definieren. In dieser Form hat der quadratische Abstand ds^2 die Form

$$ds^2 = (h_1 dq_1)^2 + (h_2 dq_2)^2 + (h_3 dq_3)^2 . \quad (5.76)$$

- Die Abstände längs der neuen Koordinatenrichtungen sind also statt dx_i nun $h_i dq_i$. Wenn wir annehmen, dass die neue Koordinatenbasis ebenso wie die alte orthonormal ist, $\vec{e}_i \cdot \vec{e}_j = \delta_{ij}$, dann können wir durch einfache Analogie schreiben

$$\vec{\nabla} f = \frac{\partial f}{\partial q_i} \left(\frac{\vec{e}}{h} \right)_i \quad \text{oder} \quad (\vec{\nabla} f)_i = \frac{1}{h_i} \frac{\partial f}{\partial q_i} . \quad (5.77)$$

wobei die zweite Gleichung komponentenweise gilt und nicht die Summenkonvention nutzt.

5.5.1 Zylinderkoordinaten

- Betrachten wir zunächst ein System mit Zylindersymmetrie. Wir bezeichnen die Zylinderachse willkürlich als \vec{e}_3 . Eine zweite Schar von Koordinatenflächen erhalten wir, indem wir die x_1 - x_3 -Ebene um den Winkel ϕ um die \vec{e}_3 -Achse drehen. Als dritte Schar von Koordinatenflächen führen wir Flächen ein, die den Zylinder senkrecht im Abstand z von der x_1 - x_2 -Ebene schneiden. Diese drei Koordinatenflächen sind jeweils durch die Konstanz der drei Parameter $\rho = (x_1^2 + x_2^2)^{1/2} \in [0, \infty)$, $\phi \in [0, 2\pi)$ und $z \in (-\infty, \infty)$ definiert.
- Ihre Einheitsnormalenvektoren bilden die Basis \vec{e}_ρ , \vec{e}_ϕ und \vec{e}_z . Aufgrund ihrer Konstruktion hängen sie mit den kartesischen Einheitsvektoren durch

$$\vec{e}_\rho = \cos \phi \vec{e}_1 + \sin \phi \vec{e}_2 , \quad \vec{e}_\phi = \cos \phi \vec{e}_2 - \sin \phi \vec{e}_1 , \quad \vec{e}_z = \vec{e}_3 \quad (5.78)$$

zusammen. Wegen $x_1 = \rho \cos \phi$, $x_2 = \rho \sin \phi$ und $x_3 = z$ gilt nach der Definition in (5.75)

$$\begin{aligned} h_\rho^2 &= \cos^2 \phi + \sin^2 \phi = 1 , & h_\phi^2 &= \rho^2 \sin^2 \phi + \rho^2 \cos^2 \phi = \rho^2 , \\ h_z^2 &= 1 . \end{aligned} \quad (5.79)$$

und für den infinitesimalen quadratische Abstand

$$ds^2 = d\rho^2 + \rho^2 d\phi^2 + dz^2 . \quad (5.80)$$

Nach (5.77) ist der Gradient einer Funktion $f(\rho, \phi, z)$ in Zylinderkoordinaten

$$\begin{aligned} \vec{\nabla} f &= \frac{\vec{e}_\rho}{h_\rho} \partial_\rho f + \frac{\vec{e}_\phi}{h_\phi} \partial_\phi f + \frac{\vec{e}_z}{h_z} \partial_z f \\ &= \vec{e}_\rho \partial_\rho f + \frac{\vec{e}_\phi}{\rho} \partial_\phi f + \vec{e}_z \partial_z f . \end{aligned} \quad (5.81)$$

5.5.2 Sphärische Polarkoordinaten

- Sphärische Polarkoordinaten sind dann angemessen, wenn das betrachtete System Kugelsymmetrie hat. Entsprechend führt man zunächst Kugelflächen als eine Schar von Koordinatenflächen ein. Dann wählt man \vec{e}_3 als (beliebige) Polachse und führt als zweite Schar von Koordinatenflächen solche ein, die durch Drehung der x_1 - x_3 -Ebene um den Winkel ϕ um die \vec{e}_3 -Achse entstehen. Als dritte Schar von Koordinatenflächen wählt man Kegelflächen um die \vec{e}_3 -Achse mit Öffnungswinkel θ . Auf diesen drei Koordinatenflächen sind also jeweils $r = (x_1^2 + x_2^2 + x_3^2)^{1/2} \in [0, \infty)$, $\phi \in [0, 2\pi)$ und $\theta \in [0, \pi]$ konstant.
- Wegen

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = r \begin{pmatrix} \cos \phi \sin \theta \\ \sin \phi \sin \theta \\ \cos \theta \end{pmatrix} \quad (5.82)$$

sind

$$\begin{aligned} h_r^2 &= \cos^2 \phi \sin^2 \theta + \sin^2 \phi \sin^2 \theta + \cos^2 \theta = 1, \\ h_\phi^2 &= r^2 \sin^2 \theta (\sin^2 \phi + \cos^2 \phi) = r^2 \sin^2 \theta, \\ h_\theta^2 &= r^2 \cos^2 \theta (\cos^2 \phi + \sin^2 \phi) + r^2 \sin^2 \theta = r^2. \end{aligned} \quad (5.83)$$

Aus dem infinitesimalen quadratischen Wegelement wird

$$ds^2 = dr^2 + r^2 \sin^2 \theta d\phi + r^2 d\theta^2, \quad (5.84)$$

und der Gradient in sphärischen Polarkoordinaten lässt sich in der Form

$$\vec{\nabla} f = \vec{e}_r \partial_r f + \frac{\vec{e}_\phi}{r \sin \theta} \partial_\phi f + \frac{\vec{e}_\theta}{r} \partial_\theta f \quad (5.85)$$

darstellen. Für eine radialsymmetrische Funktion $f(r)$, die also nur vom Radius r abhängt, aber nicht von den Winkeln θ und ϕ , erhalten wir daraus das bekannte Ergebnis

$$\vec{\nabla} f(r) = \frac{df(r)}{dr} \vec{e}_r = f'(r) \vec{e}_r \quad (5.86)$$

auf sehr einfache Weise.

Kapitel 6

Stöße und Streuung

6.1 Systeme von Massenpunkten

Stöße oder Streuung von Teilchen sind vermutlich die am weitesten verbreitete experimentelle Methode, um ein physikalisches System zu untersuchen. Sie wird noch wichtiger, wenn wir in der Quantenmechanik lernen werden, dass auch die Bestrahlung eines Objektes mit Licht als Streuung mit Licht-Teilchen betrachtet werden kann. Bevor wir miteinander wechselwirkende Teilchen mit Hilfe von Streuprozessen betrachten, müssen wir zunächst die bekannten Erhaltungssätze für eine allgemeine Anzahl von Teilchen formulieren.

6.1.1 Bewegung des Schwerpunkts

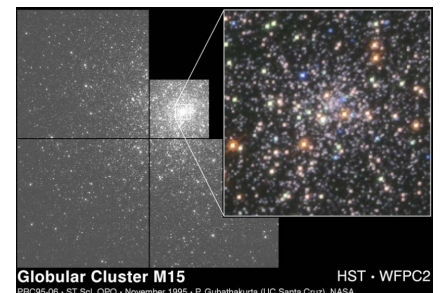
- Gegeben seien N Massenpunkte mit den Massen m_i an den Ortsvektoren \vec{x}_i , $1 \leq i \leq N$. Die Kraft des i -ten Massenpunkts auf den j -ten sei \vec{F}_{ij} , und zusätzlich wirke auf den i -ten Massenpunkt die äußere Kraft $\vec{F}_i^{(e)}$. Die Bewegungsgleichungen für das Teilchen i lauten also

$$m_i \ddot{\vec{x}}_i = \sum_{j \neq i} \vec{F}_{ji} + \vec{F}_i^{(e)}. \quad (6.1)$$

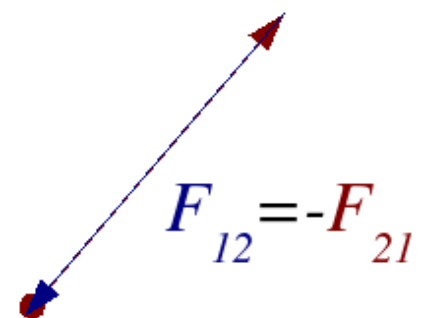
- Mithilfe der Gesamtmasse $M = \sum_i m_i$ definieren wir die *Schwerpunktcoordinate*

$$M \vec{X} = \sum_i m_i \vec{x}_i. \quad (6.2)$$

Summiert man die Bewegungsgleichungen für alle Teilchen, so



Beispiel für ein System von Massenpunkten: der Kugelsternhaufen M15



Kräfte zwischen zwei Teilchen heben sich paarweise auf.

erhält man wegen des dritten Newtonschen Axioms $\vec{F}_{ji} = -\vec{F}_{ij}$

$$\begin{aligned} M\ddot{\vec{X}} &= \sum_i m_i \ddot{\vec{x}}_i = \sum_{i,j \neq i} \vec{F}_{ji} + \sum_i \vec{F}_i^{(e)} \\ &= \sum_{j>i} \vec{F}_{ij} + \sum_{j<i} \vec{F}_{ij} + \sum_i \vec{F}_i^{(e)} \\ &= \sum_{j>i} (\vec{F}_{ij} + \vec{F}_{ji}) + \sum_i \vec{F}_i^{(e)} = \sum_i \vec{F}_i^{(e)}. \end{aligned} \quad (6.3)$$

Das Gesamtsystem bewegt sich also so, als wäre seine gesamte Masse in seinem Schwerpunkt vereinigt und bewege sich aufgrund der äußeren Kräfte.

- Wenn keine äußeren Kräfte wirken, ist der Gesamtimpuls erhalten,

$$M\dot{\vec{X}} = \sum_i m_i \dot{\vec{x}}_i = \sum_i \vec{p}_i =: \vec{P} = \text{konstant}. \quad (6.4)$$

- Wenn wir neue Koordinaten $\vec{x}_i^* = \vec{x}_i - \vec{X}$ relativ zum Schwerpunkt einführen, dann finden wir

$$\begin{aligned} \sum_i m_i \vec{x}_i^* &= \sum_i m_i \vec{x}_i - M\vec{X} = \vec{0} \\ \sum_i \vec{p}_i^* &= \sum_i m_i \dot{\vec{x}}_i^* = \sum_i m_i \dot{\vec{x}}_i - M\dot{\vec{X}} = \vec{0}. \end{aligned} \quad (6.5)$$

6.1.2 Drehimpuls

- Wir nehmen nun an, dass die inneren Kräfte zwischen zwei Massenpunkten längs der Verbindungslinie zwischen diesen Punkten wirken,

$$\vec{F}_{ij} \parallel (\vec{x}_i - \vec{x}_j) \quad \text{oder} \quad \vec{F}_{ij} \times (\vec{x}_i - \vec{x}_j) = \vec{0}. \quad (6.6)$$

Der Gesamtdrehimpuls bezüglich des Koordinatenursprungs ist

$$\vec{L} = \sum_i \vec{L}_i = \sum_i (\vec{x}_i \times \vec{p}_i). \quad (6.7)$$

Er hat die Zeitableitung oder das Gesamtdrehmoment

$$\begin{aligned} \vec{M} &= \frac{d\vec{L}}{dt} = \sum_i (\dot{\vec{x}}_i \times \vec{p}_i) = \sum_i m_i (\dot{\vec{x}}_i \times \ddot{\vec{x}}_i) \\ &= \sum_i \left(\dot{\vec{x}}_i \times \sum_{j \neq i} \vec{F}_{ji} \right) + \sum_i \dot{\vec{x}}_i \times \vec{F}_i^{(e)} \\ &= \sum_{j>i} (\dot{\vec{x}}_i \times \vec{F}_{ji} + \dot{\vec{x}}_j \times \vec{F}_{ij}) + \sum_i \dot{\vec{x}}_i \times \vec{F}_i^{(e)} \\ &= \sum_{j>i} [(\dot{\vec{x}}_i - \dot{\vec{x}}_j) \times \vec{F}_{ji}] + \sum_i \dot{\vec{x}}_i \times \vec{F}_i^{(e)} = \sum_i \dot{\vec{x}}_i \times \vec{F}_i^{(e)}. \end{aligned} \quad (6.8)$$

Innere Kräfte tragen auch zum Gesamtdrehimpuls eines Systems von Massenpunkten *nicht* bei.

- Der Gesamtdrehimpuls im Schwerpunktsystem hat die Form

$$\begin{aligned}
 \vec{L} &= \sum_i m_i (\vec{x}_i \times \dot{\vec{x}}_i) = \sum_i m_i (\vec{X} + \vec{x}_i^*) \times (\dot{\vec{X}} + \dot{\vec{x}}_i^*) \\
 &= \vec{X} \times M\dot{\vec{X}} + \left(\vec{X} \times \sum_i \vec{p}_i^* \right) + \left(\sum_i m_i \vec{x}_i^* \times \dot{\vec{X}} \right) + \sum_i (\vec{x}_i^* \times \vec{p}_i^*) \\
 &= \vec{X} \times M\dot{\vec{X}} + \sum_i (\vec{x}_i^* \times \vec{p}_i^*) . \tag{6.9}
 \end{aligned}$$

Der erste Term ist der Drehimpuls der Bewegung des Schwerpunkts um den Ursprung, der letzte ist der gesamte innere Drehimpuls.

6.1.3 Energie

- Wir nehmen nun zusätzlich an, dass die Kräfte zwischen den Massenpunkten Potentialkräfte seien

$$\vec{F}_{ji} = -\vec{\nabla}_i V_{ji}(|\vec{x}_i - \vec{x}_j|) . \tag{6.10}$$

wobei ∇_i die Ableitung nach der Koordinate x_i ist. Wenn sie nur vom skalaren Abstand abhängen, dann erfüllen die Potentiale $V_{ij} = V_{ji}$. Ebenso seien die äußeren Kräfte Potentialkräfte,

$$\vec{F}_i^{(e)} = -\vec{\nabla}_i V_i^{(e)}(\vec{x}_i) . \tag{6.11}$$

- Die Bewegungsgleichungen nach (6.1) können summandenweise mit dem integrierenden Faktor $\dot{\vec{x}}_i$ multipliziert werden,

$$\begin{aligned}
 m_i \ddot{\vec{x}}_i &= - \sum_i \vec{\nabla}_i \sum_{j \neq i} V_{ji}(|\vec{x}_i - \vec{x}_j|) - \sum_i \vec{\nabla}_i V_i^{(e)}(\vec{x}_i) \tag{6.12} \\
 \sum_i m_i \dot{\vec{x}}_i \cdot \ddot{\vec{x}}_i &= - \sum_i \dot{\vec{x}}_i \cdot \vec{\nabla}_i \sum_{j \neq i} V_{ji}(|\vec{x}_i - \vec{x}_j|) - \sum_i \dot{\vec{x}}_i \cdot \vec{\nabla}_i V_i^{(e)}(\vec{x}_i) .
 \end{aligned}$$

Im zweiten Term müssen wir beachten, dass V_{ij} von zwei Argumenten x_i und x_j abhängt. Wenn wir die Ortsableitung in eine Zeitableitung umformen,

$$\frac{d}{dt} \left[\sum_i \frac{m_i}{2} \dot{\vec{x}}_i^2 + \frac{1}{2} \sum_{j \neq i} V_{ji}(|\vec{x}_i - \vec{x}_j|) + \sum_i V_i^{(e)}(\vec{x}_i) \right] = 0 , \tag{6.13}$$

dann sorgt der Faktor 1/2 vor dem zweiten Term mit der Bedingung $V_{ij} = V_{ji}$ dafür, dass in der Kettenregel nur die eine Abhängigkeit in (6.12) beiträgt. Physikalisch bedeutet das, dass die potentielle Energie des Teilchens i bezüglich des Teilchens j nur einmal gezählt wird, weil sie identisch mit der potentiellen Energie des Teilchens j bezüglich des Teilchens i ist. Das ist der Energiesatz eines Systems von N Massenpunkten.

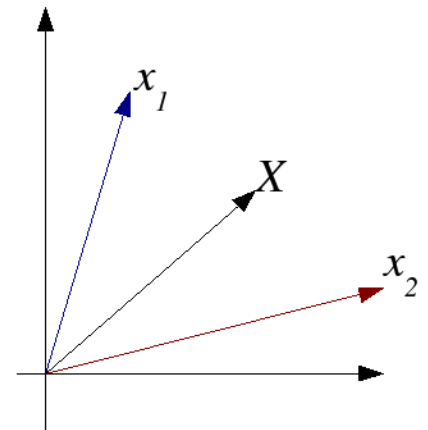
6.1.4 Reduzierte Masse

- Für Systeme mit zwei Teilchen führen wir den Verbindungsvektor $\vec{r} = \vec{x}_1 - \vec{x}_2$ ein. Für ihn bekommen wir also Bewegungsgleichung

$$\begin{aligned}
 m_1 m_2 \ddot{\vec{r}} &= m_2 m_1 \ddot{\vec{x}}_1 - m_1 m_2 \ddot{\vec{x}}_2 \\
 &= -m_2 \vec{\nabla}_1 V(|\vec{r}|) + m_1 \vec{\nabla}_2 V(|\vec{r}|) \\
 &= -m_2 \frac{dV(|\vec{r}|)}{d|\vec{r}|} \vec{\nabla}_1 \sqrt{(\vec{x}_1 - \vec{x}_2)^2} + m_1 \frac{dV(|\vec{r}|)}{d|\vec{r}|} \vec{\nabla}_2 \sqrt{(\vec{x}_2 - \vec{x}_1)^2} \\
 &= -m_2 \frac{dV(|\vec{r}|)}{d|\vec{r}|} \frac{\vec{r}}{|\vec{r}|} - m_1 \frac{dV(|\vec{r}|)}{d|\vec{r}|} \frac{\vec{r}}{|\vec{r}|} \\
 &= -(m_1 + m_2) \frac{dV(|\vec{r}|)}{d|\vec{r}|} \vec{\nabla}_r |\vec{r}|, \tag{6.14}
 \end{aligned}$$

Das ist die Bewegungsgleichung eines Massenpunkts mit der reduzierten Masse μ in einem äußeren Potential V ,

$$\mu = \frac{m_1 m_2}{m_1 + m_2} \quad \Rightarrow \quad \mu \ddot{\vec{r}} = -\vec{\nabla}_r V(|\vec{r}|). \tag{6.15}$$



Schwerpunkt zweier Massenpunkte gleicher Masse

- Bei $m_1 = m_2 =: m$ ist die reduzierte Masse $\mu = m/2$, und der Schwerpunkt liegt in der Mitte zwischen den beiden Massen. Bei $m_1 \gg m_2$ ist die reduzierte Masse $\mu \approx m_2$. Das entspricht der Bewegung von m_2 um den „festen“ Massenpunkt m_1 . Der Schwerpunkt liegt annähernd am Ort von m_1 , $\vec{X} \approx \vec{x}_1$, was die nachträgliche Rechtfertigung dafür liefert, die Bewegung eines Planeten um die Sonne durch die Bewegung eines Massenpunktes um die ortsfeste Sonne anzunähern.

6.2 Stöße und Streuung

6.2.1 Elastischer Stoß zwischen zwei Teilchen

- Elastisch* heißt ein Stoß zwischen zwei Teilchen, wenn in seinem Verlauf keine mechanische Energie in andere Energieformen (z.B. durch Verformung) umgewandelt wird. Wir können also bei der Beschreibung eines elastischen Stoßes von den Erhaltungssätzen des Impulses und der Energie Gebrauch machen.
- Gegeben seien wieder zwei Teilchen mit den Massen m_1 und m_2 . Ihre Geschwindigkeiten vor dem Stoß seien $\vec{v}_{1,2}$, und nach dem Stoß $\vec{v}'_{1,2}$. Die Impulserhaltung fordert

$$m_1 \vec{v}_1 + m_2 \vec{v}_2 = m_1 \vec{v}'_1 + m_2 \vec{v}'_2, \tag{6.16}$$

während die Energieerhaltung

$$\frac{m_1}{2} \vec{v}_1^2 + \frac{m_2}{2} \vec{v}_2^2 = \frac{m_1}{2} \vec{v}_1^{*2} + \frac{m_2}{2} \vec{v}_2^{*2} \quad (6.17)$$

verlangt. Welche Aussagen sind allein aufgrund der Erhaltungssätze möglich?

- Im Schwerpunktsystem lauten der Impulssatz

$$m_1 \vec{v}_1^* = -m_2 \vec{v}_2^* \quad \text{und} \quad m_1 \vec{v}_1^{*'} = -m_2 \vec{v}_2^{*'} \quad (6.18)$$

und der Energiesatz

$$\frac{m_1}{2} \vec{v}_1^{*2} + \frac{m_2}{2} \vec{v}_2^{*2} = \frac{m_1}{2} \vec{v}_1^{*2'} + \frac{m_2}{2} \vec{v}_2^{*2'} \quad (6.19)$$

Der Winkel zwischen der Einfalls- und der Ausfallsrichtung heißt Streuwinkel ϑ^* . Man kann ihn als Teil von Zylinderkoordinaten definieren und dabei annehmen dass der Azimutalwinkel ϕ durch die Rotationssymmetrie keine Rolle spielt. Im Schwerpunktsystem ist der Streuwinkel für beide Teilchen gleich.

- Setzt man den Impulssatz (6.18) in den Energiesatz (6.19) ein, um \vec{v}_2^* bzw. $\vec{v}_2^{*'}$ zu eliminieren, folgt

$$m_1 \vec{v}_1^{*2} + m_2 \left(\frac{m_1}{m_2} \vec{v}_1^* \right)^2 = m_1 \vec{v}_1^{*2'} + m_2 \left(\frac{m_1}{m_2} \vec{v}_1^{*'} \right)^2 \Rightarrow |\vec{v}_1^*| = |\vec{v}_1^{*'}| \Rightarrow |\vec{v}_2^*| = |\vec{v}_2^{*'}|. \quad (6.20)$$

Die Beträge der Geschwindigkeiten vor und nach dem Stoß sind also für beide Massenpunkte gleich. Damit sind die Erhaltungssätze erschöpft, sie machen also insbesondere über den Streuwinkel keine Aussage. Er hängt vom wirksamen Kraftgesetz ab (z.B. bei der Coulombstreuung). Wir wissen bisher nur, dass $0 \leq \vartheta^* \leq \pi$ ist.

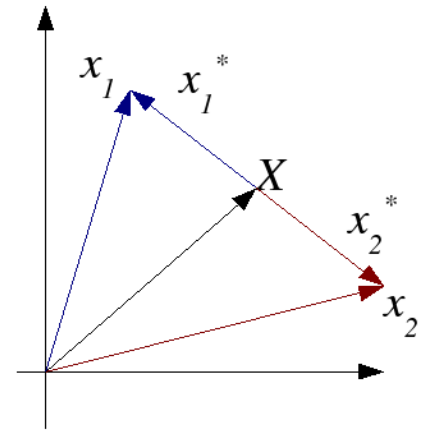
6.2.2 Transformation der Streuwinkel

- Wenn der Massenpunkt 2 anfänglich im Laborsystem ruht, $\vec{v}_2 = 0$, dann bewegt sich der Schwerpunkt im Laborsystem als $M\dot{\vec{X}} = m_1 \vec{v}_1$. Für die Bewegung im Schwerpunktsystem gilt also

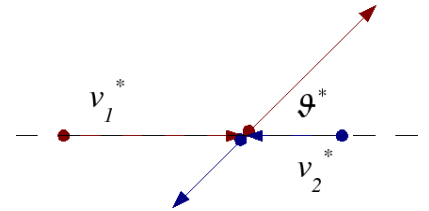
$$\vec{v}_1 = \vec{v}_1^* + \dot{\vec{X}} = \vec{v}_1^* + \frac{m_1}{M} \vec{v}_1 \Leftrightarrow \vec{v}_1^* = \frac{M - m_1}{M} \vec{v}_1 = \frac{m_2}{M} \vec{v}_1 \quad (6.21)$$

Bei einer Bewegung des Schwerpunkts in Richtung von \vec{v}_1 ändern die beiden Komponenten zur Berechnung des Streuwinkels wie im Bild beschrieben

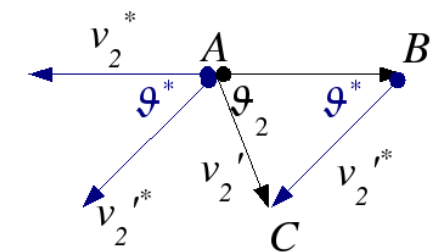
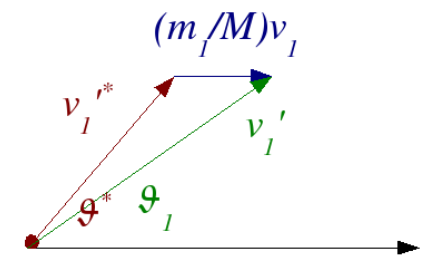
$$\begin{aligned} v_1' \sin \vartheta_1 &= v_1^* \sin \vartheta^* \\ v_1' \cos \vartheta_1 &= v_1^* \cos \vartheta^* + \frac{m_1}{M} v_1 \end{aligned} \quad (6.22)$$



Transformation auf Schwerpunktsystem



Streuung im Schwerpunktsystem



zur Transformation des Streuwinkels

und daraus folgt für den Streuwinkels ϑ_1 im Laborsystem

$$\tan \vartheta_1 = \frac{v_1^* \sin \vartheta^*}{v_1^* \cos \vartheta^* + (m_1/M)v_1} = \frac{\sin \vartheta^*}{\cos \vartheta^* + (m_1/m_2)}, \quad (6.23)$$

denn $v_1 = (M/m_2)v_1^*$ nach (6.21).

- Betrachten wir nun die Bewegung des zweiten Massepunktes. Wegen $\vec{v}_2 = 0$ ist

$$\vec{v}_2^* = -\dot{\vec{X}} = -\frac{m_1}{M}\vec{v}_1. \quad (6.24)$$

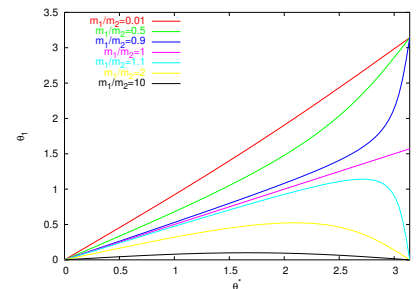
Da außerdem noch $|\vec{v}_2^*| = |\vec{v}_1^*|$ ist, muss das Dreieck, das aus $\dot{\vec{X}}$, \vec{v}_2^* und \vec{v}_1^* gebildet wird (Dreieck ABC in der Skizze am Rand), gleichschenkelig sein. Die Winkel BAC und ACB sind daher gleich ϑ_2 , und weil der Winkel ABC gleich ϑ^* ist, folgt

$$\vartheta_2 = \frac{\pi - \vartheta^*}{2}. \quad (6.25)$$

- Wenn wir noch einmal (6.23) betrachten, dann sind im Grenzfalle $m_1 \ll m_2$ die beiden Winkel gleich, $\vartheta_1 \approx \vartheta^*$. Für $m_1 < m_2$ kann $\tan \vartheta_1$ alle Werte zwischen null und unendlich annehmen, also sind alle Streuwinkel $0 \leq \vartheta_1 \leq \pi$ möglich. Für $m_1 \geq m_2$ sind kleine Werte von $\tan \vartheta_1$ weiterhin möglich, aber der Nenner kann nicht verschwinden. Damit können wir einen maximalen Streuwinkel berechnen,

$$\sin \vartheta_1^{\max} = \frac{m_2}{m_1}. \quad (6.26)$$

Schließlich sind für $m_1 = m_2$ auch $\vartheta_1 = \vartheta^*/2$ und $\vartheta_2 = \pi/2 - \vartheta_1$.



Streuwinkel ϑ_1 im Laborsystem als Funktion des Streuwinkels ϑ^* im Schwerpunktsystem für verschiedene Massenverhältnisse

6.2.3 Energieübertrag bei elastischer Streuung

- Betrachten wir als Beispiel ein Neutron der Masse m , das an einem Kern der Masse Am streut. Wie in der vorigen Betrachtung nehmen wir den Kern im Laborsystem als im Ursprung ruhend an ($\vec{v}_2 = 0$).
- Nach (6.21) ist

$$\begin{aligned} \vec{v}_1^2 &= \left(\vec{v}_1^* + \dot{\vec{X}} \right)^2 = \left(\vec{v}_1^* + \frac{m}{m + Am} \vec{v}_1 \right)^2 \\ &= \vec{v}_1^{*2} + \frac{2}{1 + A} |\vec{v}_1^*| |\vec{v}_1| \cos \vartheta^* + \frac{\vec{v}_1^2}{(1 + A)^2}. \end{aligned} \quad (6.27)$$

Ferner gilt wegen der Energieerhaltung (6.20)

$$\vec{v}_1^{*2} = \vec{v}_1^2 = \left(\vec{v}_1 - \dot{\vec{X}} \right)^2 = \left(\vec{v}_1 - \frac{\vec{v}_1}{1 + A} \right)^2 = \vec{v}_1^2 \left(\frac{A}{1 + A} \right)^2. \quad (6.28)$$

Eingesetzt in (6.27) bedeutet dies

$$\begin{aligned}\vec{v}_1'^2 &= \vec{v}_1^2 \frac{A^2}{(1+A)^2} + \frac{2}{1+A} \frac{A}{1+A} \vec{v}_1^2 \cos \vartheta^* + \frac{\vec{v}_1^2}{(1+A)^2} \\ &= \vec{v}_1^2 \left[1 + \frac{2A}{(1+A)^2} (\cos \vartheta^* - 1) \right].\end{aligned}\quad (6.29)$$

- Für die *relative* Änderung der kinetischen Energie ergibt sich damit

$$\frac{T - T'}{T} = 1 - \frac{\vec{v}_1'^2}{\vec{v}_1^2} = \frac{2A}{(1+A)^2} (1 - \cos \vartheta^*). \quad (6.30)$$

Eine Mittelung über ϑ^* ergibt unter der Annahme, dass die Streuwinkel im Schwerpunktsystem gleichverteilt sind

$$\begin{aligned}\left\langle \frac{T - T'}{T} \right\rangle &= \frac{1}{4\pi} \int_0^{2\pi} d\varphi \int_0^\pi \frac{2A}{(1+A)^2} (1 - \cos \vartheta^*) \sin \vartheta^* d\vartheta^* \\ &= \frac{A}{(1+A)^2} \int_{-1}^1 (1 - \mu) d\mu = \frac{2A}{(1+A)^2};\end{aligned}\quad (6.31)$$

das ist der mittlere, relative Energieverlust des Neutrons. Er wird maximal, wenn $A = 1$ gilt, d.h. die Streuung von Neutronen an Atomkernen bremst die Neutronen dann am effektivsten ab, wenn die Atomkerne möglichst leicht sind. Ein Beispiel dafür ist die Neutronenmoderation in Kernreaktoren z.B. mit Wasser.

Kapitel 7

Koordinatentransformationen

7.1 Darstellung durch Matrizen

Mathematisch haben wir Koordinatentransformationen schon formal anhand von Basen von Vektorräumen eingeführt. Die Umrechnung von kartesischen Koordinaten in Zylinder-Koordinaten und Kugel-Koordinaten haben wir explizit angegeben. Im letzten Kapitel haben wir eine konkrete Transformation im dreidimensionalen Raum ausgenutzt, nämlich die Umrechnung zwischen Laborsystem und Schwerpunktsystem. In diesem Kapitel werden wir uns systematisch der Transformation zwischen Koordinatensystemen mit Hilfe von Matrizen widmen.

7.1.1 Drehungen im dreidimensionalen Raum

- Stellen wir uns zum Beispiel vor, dass wir ein vorher eingeführtes Koordinatensystem um seine z -Achse um einen Winkel ϕ drehen. Dadurch gehen die Basisvektoren \vec{e}_1 und \vec{e}_2 in neue Basisvektoren \vec{e}'_1 und \vec{e}'_2 über. Nach (4.24) sind die Einträge in die Matrix, die eine Basis in eine andere transformiert, durch die Skalarprodukte der alten mit den neuen Basisvektoren gegeben.

In unserem Beispiel schließen \vec{e}_1 und \vec{e}'_1 sowie \vec{e}_2 und \vec{e}'_2 jeweils den Winkel ϕ miteinander ein, also

$$\vec{e}'_1 \cdot \vec{e}_1 = \cos \phi = \vec{e}'_2 \cdot \vec{e}_2, \quad (7.1)$$

während \vec{e}_1 mit \vec{e}'_2 den Winkel $90^\circ + \phi$ und \vec{e}_2 mit \vec{e}'_1 den Winkel $90^\circ - \phi$ einschließt, so dass

$$\vec{e}'_2 \cdot \vec{e}_1 = \cos(90^\circ + \phi) = -\sin \phi, \quad \vec{e}'_1 \cdot \vec{e}_2 = \cos(90^\circ - \phi) = \sin \phi \quad (7.2)$$

sind. Da $\vec{e}'_3 = \vec{e}_3$ ist, bleiben die Skalarprodukte mit \vec{e}'_3 unverändert,

$$\vec{e}'_1 \cdot \vec{e}_3 = 0 = \vec{e}'_2 \cdot \vec{e}_3, \quad \vec{e}'_3 \cdot \vec{e}_3 = 1. \quad (7.3)$$

Damit lauten die 3×3 Zahlen, die diese Transformation sicherlich komplett beschreiben $a_{ij} = \vec{e}'_i \cdot \vec{e}_k$:

$$\begin{array}{lll} a_{11} = \cos \phi & a_{12} = \sin \phi & a_{13} = 0 \\ a_{21} = -\sin \phi & a_{22} = \cos \phi & a_{23} = 0 \\ a_{31} = 0 & a_{32} = 0 & a_{33} = 1. \end{array} \quad (7.4)$$

7.1.2 Matrizen und Matrixoperationen

- Die 3×3 Zahlen a_{ij} , die oben bei der Darstellung von Koordinatentransformationen aufgetreten sind, lassen sich durch das rechteckige Zahlenschema einer *Matrix* darstellen, die in diesem Fall quadratisch ist,

$$A = \begin{pmatrix} \cos \phi & \sin \phi & 0 \\ -\sin \phi & \cos \phi & 0 \\ 0 & 0 & 1 \end{pmatrix}. \quad (7.5)$$

- Allgemein sind Matrizen Zahlenschemata aus $M \times N$ Zahlen eines Körpers K , die in M Zeilen und N Spalten angeordnet sind,

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1N} \\ a_{21} & a_{22} & \cdots & a_{2N} \\ \vdots & \vdots & \cdots & \vdots \\ a_{M1} & a_{M2} & \cdots & a_{MN} \end{pmatrix} =: (a_{ij}). \quad (7.6)$$

Die Zahlen a_{ij} heißen *Matrixelemente*. Die $M \times N$ -Matrizen bilden einen Vektorraum \mathcal{M}_{MN} über K , wenn man die Addition und die Multiplikation mit Skalaren elementweise definiert,

$$\begin{array}{l} + : \quad \mathcal{M}_{MN} \times \mathcal{M}_{MN} \rightarrow \mathcal{M}_{MN}, \quad (A, B) \mapsto (a_{ij} + b_{ij}), \\ \cdot : \quad K \times \mathcal{M}_{MN} \rightarrow \mathcal{M}_{MN}, \quad (\lambda, A) \mapsto (\lambda a_{ij}). \end{array} \quad (7.7)$$

- Außerdem führt man eine Matrixmultiplikation ein, durch die $L \times M$ -Matrizen mit $M \times N$ -Matrizen multipliziert werden, um $L \times N$ -Matrizen zu ergeben,

$$\mathcal{M}_{LM} \times \mathcal{M}_{MN} \rightarrow \mathcal{M}_{LN}, \quad (A, B) \mapsto AB = (a_{ij}b_{jk}) \quad (7.8)$$

mit $1 \leq i \leq L$, $1 \leq j \leq M$ und $1 \leq k \leq N$. Indem man die Koordinatendarstellungen von (Spalten-)Vektoren aus N -dimensionalen Vektorräumen V_N als $N \times 1$ -Matrizen auffasst, ist damit auch eine

Multiplikation von $M \times N$ -Matrizen mit N -dimensionalen Spaltenvektoren definiert, deren Ergebnis ein Vektor aus einem M -dimensionalen Vektorraum V_M ist,

$$\mathcal{M}_{MN} \times V_N \rightarrow V_M, \quad (A, v) \mapsto (a_{ij}v_j), \quad (7.9)$$

wobei $1 \leq i \leq M$ und $1 \leq j \leq N$ sind. Ebenso kann man Zeilenvektoren aus V_N als $1 \times N$ -Matrizen auffassen und sie mit $N \times M$ -Matrizen multiplizieren, um M -dimensionale Zeilenvektoren zu erhalten,

$$V_N \times \mathcal{M}_{MN} \rightarrow V_M, \quad (v, A) \mapsto (v_i a_{ij}). \quad (7.10)$$

Man schreibt die Multiplikationen (7.9) und (7.10) auch kurz in der Form

$$A \cdot v \quad \text{und} \quad v^T \cdot A. \quad (7.11)$$

Ebenso kann man in (7.10) zuerst die Zeilen und Spalten der Matrix A vertauschen, wodurch die Matrix in ihre transponierte A^T übergeht und die Indizes vertauscht werden,

$$A = (a_{ij}), \quad A^T = (a_{ji}^T) = (a_{ji}). \quad (7.12)$$

Offensichtlich gilt dann

$$(v^T \cdot A)_j = v_i a_{ij} = (a_{ji}^T)v_i = (A^T \cdot v)_j. \quad (7.13)$$

- Ganz allgemein gilt offenbar für beliebige Matrizen $A \in \mathcal{M}_{LM}$ und $B \in \mathcal{M}_{MN}$

$$(A \cdot B)^T = (a_{ij}b_{jk})^T = (a_{kj}b_{ji}) = (a_{jk}^T b_{ij}^T) = B^T \cdot A^T. \quad (7.14)$$

7.1.3 Beispiele

- Ein physikalisch relevantes Beispiel für die Multiplikation zweier Matrizen stellt die Kombination zweier Drehungen im dreidimensionalen Raum dar, von denen die eine um die z -Achse um einen Winkel ϕ dreht, die andere dann um die y -Achse um einen Winkel ψ . Die kombinierte Drehung wird durch das Produkt der beiden Drehmatrizen

$$R_1 = \begin{pmatrix} \cos \phi & \sin \phi & 0 \\ -\sin \phi & \cos \phi & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad R_2 = \begin{pmatrix} \cos \psi & 0 & \sin \psi \\ 0 & 1 & 0 \\ -\sin \psi & 0 & \cos \psi \end{pmatrix} \quad (7.15)$$

dargestellt. Das Produkt lautet nach der Multiplikationsregel (7.8)

$$R_2 R_1 = \begin{pmatrix} \cos \phi \cos \psi & \sin \phi \cos \psi & \sin \psi \\ -\sin \phi & \cos \phi & 0 \\ -\cos \phi \sin \psi & -\sin \phi \sin \psi & \cos \psi \end{pmatrix}. \quad (7.16)$$

Indem man die Reihenfolge der Faktoren vertauscht, sieht man, dass die Matrixmultiplikation nicht kommutativ ist,

$$R_1 R_2 = \begin{pmatrix} \cos \phi \cos \psi & \sin \phi & \cos \phi \sin \psi \\ -\sin \phi \cos \phi & \cos \phi & -\sin \phi \sin \psi \\ -\sin \psi & 0 & \cos \psi \end{pmatrix}, \quad (7.17)$$

was die alltägliche Erfahrung bekräftigt, dass das Ergebnis zweier nacheinander ausgeführter räumlicher Drehungen im Allgemeinen von der Reihenfolge der Drehungen abhängt.

- Der Vektor \vec{x} mit den alten Koordinaten

$$\vec{x} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \quad (7.18)$$

vor der ersten Drehung hat im neuen Koordinatensystem nach der ersten Drehung R_1 die Koordinaten

$$\vec{x}' = R_1 \cdot \vec{x} = \begin{pmatrix} \cos \phi + \sin \phi \\ \cos \phi - \sin \phi \\ 1 \end{pmatrix}. \quad (7.19)$$

7.2 Determinanten und Matrixinversion

7.2.1 Determinanten

- Quadratischen Matrizen können Zahlen zugeordnet werden, die als ihre Determinanten bezeichnet werden und die für viele auch physikalische Betrachtungen unverzichtbar sind,

$$\det : \mathcal{M}_{NN} \rightarrow K, \quad A \mapsto \det A. \quad (7.20)$$

- In ihrer Definition tauchen *Permutationen* von Indizes auf. Permutationen waren uns schon kurz bei der Einführung des Levi-Civita-Symbols begegnet. Eine Permutation π bringt eine Menge von Indizes $\{i_1, i_2, \dots, i_N\}$ in eine andere Reihenfolge $\{i_{\pi(1)}, i_{\pi(2)}, \dots, i_{\pi(N)}\}$, indem sie ein Paar oder mehrere Paare von Indizes vertauscht. Sie heißt *gerade* oder *ungerade*, wenn eine gerade oder ungerade Anzahl von Vertauschungen zweier Elemente durchgeführt wird. Geraden Permutationen wird ein positives, ungeraden ein negatives Vorzeichen $\text{sign}(\pi)$ zugeordnet.
- Die N Indizes können auf $N! = 1 \cdot 2 \cdot \dots \cdot N$ Weisen angeordnet werden, also gibt es für N Zahlen $N!$ Permutationen. Zum Beispiel haben die drei Zahlen 1, 2, 3 $3! = 6$ Permutationen, von denen

$$\begin{aligned} (1, 2, 3), \quad (2, 3, 1), \quad (3, 1, 2) & \text{ gerade und} \\ (2, 1, 3), \quad (1, 3, 2), \quad (3, 2, 1) & \text{ ungerade} \end{aligned} \quad (7.21)$$

sind. Auf diese Weise haben wir das Levi-Civita-Symbol in drei Dimensionen definiert.

- Die Determinante einer Matrix $A \in \mathcal{M}_{NN}$ ist durch

$$\det A = \sum_{\pi} \text{sign}(\pi) a_{1\pi(1)} a_{2\pi(2)} \dots a_{N\pi(N)} \quad (7.22)$$

definiert. In drei Dimensionen verwenden wir die sechs Permutationen aus (7.21) und (7.22) und erhalten

$$\begin{aligned} \det A &= (a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32}) \\ &\quad - (a_{12}a_{21}a_{33} + a_{11}a_{23}a_{32} + a_{13}a_{22}a_{31}) \\ &= a_{11}(a_{22}a_{33} - a_{23}a_{32}) - a_{12}(a_{21}a_{33} - a_{23}a_{31}) \\ &\quad + a_{13}(a_{21}a_{32} - a_{22}a_{31}). \end{aligned} \quad (7.23)$$

Ebenfalls in drei Dimensionen kann die Determinante mithilfe des Levi-Civita-Symbols durch

$$\det A = \epsilon_{ijk} a_{1i} a_{2j} a_{3k} \quad (7.24)$$

dargestellt werden.

- Determinanten haben eine Reihe interessanter Eigenschaften, von denen einige hier ohne Begründung zusammengestellt werden. Zunächst gelten aufgrund der Definition der Determinante

$$\det(\lambda A) = \lambda^N \det A \quad \text{und} \quad \det(A^T) = \det A. \quad (7.25)$$

Weitere wichtige Aussagen sind das Multiplikationstheorem

$$\det(A \cdot B) = \det A \det B \quad (7.26)$$

für $N \times N$ -Matrizen A, B, C und dass die Determinante einer Matrix verschwindet, in der die Zeilen- oder Spaltenvektoren linear abhängig sind.

- Determinanten werden oft auch durch Betragsstriche gekennzeichnet, $\det A \equiv |A|$.

7.2.2 Orthonormale Transformationen

- Wir kehren noch einmal zu den $N \times N$ -Matrizen zurück, durch die Koordinatentransformationen dargestellt werden können. Wir haben sie in (7.6) als quadratische Schemata der Zahlen a_{ij} eingeführt, die die neue Basis e'_i durch die alte e_j ausdrücken, $e'_i = a_{ij}e_j$. Für Transformationen, die Orthonormalbasen in Orthonormalbasen überführen, gilt

$$\langle e'_i, e'_j \rangle = \langle a_{ik}e_k, a_{jl}e_l \rangle = a_{ik}a_{jl} \delta_{kl} = a_{ik}a_{jk} = a_{ik}a_{kj}^T = \delta_{ij} \quad (7.27)$$

gelten. Mit anderen Worten, die Transformationsmatrix A muss dann die *Orthonormalitätsrelation*

$$AA^T = I = A^T A \quad (7.28)$$

erfüllen, wobei I die *Einheitsmatrix* mit den Elementen $I_{ij} = \delta_{ij}$ ist. Die zweite Gleichung in (7.28) gilt wegen der Transpositionsregel (7.14). Eine solche Transformation heißt selbst orthonormal.

- Offenbar ist $\det I = 1$, und deswegen gilt wegen (7.25) und (7.26)

$$1 = \det(AA^T) = \det A \det A^T = (\det A)^2, \quad (7.29)$$

woraus folgt, dass die Determinante einer orthonormalen Transformation $\det A = \pm 1$ ist. Die Transformation heißt *eigentlich*, wenn $\det A = 1$ ist, anderenfalls *uneigentlich*.

- Die Bedeutung uneigentlicher orthonormaler Transformationen sieht man an folgendem Beispiel. Eine Spiegelung im dreidimensionalen Raum an der y - z -Ebene kann durch die Transformationsmatrix

$$A = \begin{pmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad (7.30)$$

dargestellt werden. Offensichtlich ist $\det A = -1$, weshalb diese Spiegelung eine uneigentliche orthonormale Transformation ist. Sie bildet die Einheitsvektoren \vec{e}_i entsprechend

$$\vec{e}_1 \rightarrow -\vec{e}_1, \quad \vec{e}_2 \rightarrow \vec{e}_2, \quad \vec{e}_3 \rightarrow \vec{e}_3 \quad (7.31)$$

ab. Wenn vor der Spiegelung die Beziehung $\vec{e}_i \times \vec{e}_j = \epsilon_{ijk} \vec{e}_k$ zwischen den Basisvektoren galt, wie sie in (4.17) eingeführt wurde, dann gilt nach der Spiegelung

$$\vec{e}_i \times \vec{e}_j = -\epsilon_{ijk} \vec{e}_k. \quad (7.32)$$

Man sagt, das Koordinatensystem wurde von einem Rechts- in ein Linkssystem transformiert und hat damit seine *Orientierung* oder *Helizität* vertauscht. Eigentliche orthonormale Transformationen überführen also Rechts- in Rechtssysteme, während uneigentliche Transformationen die Helizität vertauschen.

7.2.3 Inversion von Matrizen

- Eine Matrix A^{-1} heißt *invers* zu einer $N \times N$ -Matrix A , wenn sie die Bedingung

$$A^{-1}A = I \quad (7.33)$$

erfüllt. Wenn A^{-1} existiert, lassen sich ihre Elemente x_{ij} aus (7.33) eindeutig bestimmen. Sie müssen offenbar die $N \times N$ Gleichungen

$$x_{ij}a_{jk} = \delta_{ik} \quad (7.34)$$

erfüllen. Die *Cramersche Regel* besagt, dass die x_{ij} durch

$$x_{ji} = \frac{(-1)^{i-j}A_{ij}}{\det A} \quad (7.35)$$

gegeben sind, wobei die A_{ij} die *Unterdeterminanten* von A sind. Das sind die Determinanten der $(N - 1) \times (N - 1)$ -Matrizen, die man erhält, indem man aus A die i -te Zeile und die j -te Spalte streicht.

Insbesondere zeigt sie, dass die Inverse einer Matrix nur dann existiert, wenn $\det A \neq 0$ ist. Man nennt die Matrix dann *regulär*, anderenfalls *singulär*.

- Für praktische Rechnungen ist die Cramersche Regel beinahe ungeeignet, weil sie enormen Rechenaufwand erfordert. Wir geben hier die Formeln für reguläre 2×2 - und 3×3 -Matrizen an:

$$\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}^{-1} = \frac{1}{a_{11}a_{22} - a_{12}a_{21}} \begin{pmatrix} a_{22} & -a_{12} \\ -a_{21} & a_{11} \end{pmatrix} \quad (7.36)$$

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}^{-1} = \frac{1}{\det A} \begin{pmatrix} a_{22}a_{33} - a_{23}a_{32} & a_{13}a_{32} - a_{12}a_{33} & a_{12}a_{23} - a_{13}a_{22} \\ a_{23}a_{31} - a_{21}a_{33} & a_{11}a_{33} - a_{13}a_{31} & a_{13}a_{21} - a_{11}a_{23} \\ a_{21}a_{32} - a_{22}a_{31} & a_{12}a_{31} - a_{11}a_{32} & a_{11}a_{22} - a_{12}a_{21} \end{pmatrix}.$$

Überzeugen Sie sich, dass diese Formeln stimmen!

- Wenden wir die Multiplikationsregel für Determinanten auf (7.33) an, folgt

$$\det(A^{-1}) = \frac{1}{\det A} = (\det A)^{-1}. \quad (7.37)$$

Multiplikation der Gleichung

$$A^{-1}(AA^{-1}) = A^{-1} \quad (7.38)$$

mit $(A^{-1})^{-1}$ von links ergibt

$$I = (A^{-1})^{-1}A^{-1}(AA^{-1}) = AA^{-1}, \quad (7.39)$$

d.h. Rechts- und Linksinverse sind gleich und $(A^{-1})^{-1} = A$.

- Die Inverse $(AB)^{-1}$ eines Produkts AB lässt sich ähnlich einfach bestimmen. Multiplizieren wir die Bestimmungsgleichung

$$(AB)^{-1}(AB) = I \quad (7.40)$$

von rechts zuerst mit B^{-1} und dann mit A^{-1} , folgt

$$(AB)^{-1} = B^{-1}A^{-1}. \quad (7.41)$$

Außerdem folgt direkt aus der Cramerschen Regel, dass Transposition und Inversion vertauscht werden können,

$$(A^T)^{-1} = (A^{-1})^T. \quad (7.42)$$

- Für orthonormale Matrizen gilt offenbar schon aufgrund ihrer Definition, dass Inversion und Transposition identisch sind,

$$A^T = A^{-1}. \quad (7.43)$$

Das wird wiederum durch Drehmatrizen veranschaulicht. Die Inverse einer Drehmatrix, die eine Koordinatendrehung um eine beliebige Achse um einen Winkel ϕ beschreibt, muss eine Drehmatrix sein, die um dieselbe Achse, aber um den Winkel $-\phi$ dreht. Der Vergleich mit den Drehmatrizen R_1 und R_2 aus (7.15) zeigt, dass der Vorzeichenwechsel in ϕ identisch mit der Transposition der Matrizen ist.

7.3 Physikalische Bedeutung

Für die Physik ist es enorm wichtig, die Auswirkungen von Koordinatentransformationen auf physikalische Systeme zu kennen. Wenn ein physikalisches System nach einer Koordinatentransformation durch dieselbe Gleichung beschrieben wird wie vorher, heißt es symmetrisch unter dieser Transformation. Symmetrieüberlegungen spielen in der Physik eine herausragende, fundamentale Rolle, insbesondere deswegen, weil bestimmte Symmetrien zur Folge haben, dass bestimmte physikalische Größen erhalten sind.

7.3.1 Transformation des Drehimpulses

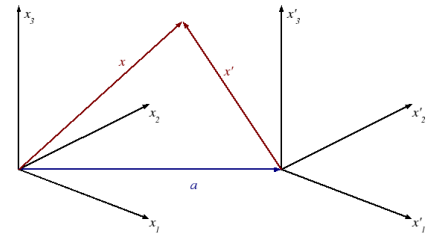
- Wir untersuchen nun, wie sich die Bewegungsgleichungen ändern, wenn wir zu neuen Koordinaten übergehen. Das Ausgangssystem K am Ursprung O habe die Basis \vec{e}_i , und das neue System K' sei am Ursprung O' und habe die Basis \vec{e}'_i .

- Für die Verschiebung des Ursprungs um einen konstanten Vektor \vec{a} gilt

$$\vec{x}' = \vec{x} - \vec{a}, \quad \dot{\vec{x}}' = \dot{\vec{x}}, \quad \ddot{\vec{x}}' = \ddot{\vec{x}}. \quad (7.44)$$

Damit ändert sich die Bewegungsgleichung zu

$$m\ddot{\vec{x}} = \vec{F}(\vec{x}, \dot{\vec{x}}, t) \rightarrow m\ddot{\vec{x}}' = \vec{F}'(\vec{x}', \dot{\vec{x}}', t) = \vec{F}(\vec{x}' + \vec{a}, \dot{\vec{x}}', t). \quad (7.45)$$



Übergang zu neuem Bezugssystem

- Wenn die Bezugssysteme K und K' denselben Ursprung haben, aber gegeneinander verdreht sind, gilt

$$\vec{x}' = R \cdot \vec{x} \quad \Leftrightarrow \quad \vec{x} = R^{-1} \cdot \vec{x}', \quad (7.46)$$

wobei R eine *orthogonalen* Drehmatrix ist, also $RR^T = I$. Wenn R zeitlich konstant ist, folgt

$$\begin{aligned} \dot{\vec{x}}' &= \frac{d}{dt}(R\vec{x}) = R\dot{\vec{x}} \\ \ddot{\vec{x}}' &= R\ddot{\vec{x}} \\ m\ddot{\vec{x}}' &= \vec{F}'(\vec{x}', \dot{\vec{x}}', t) = R\vec{F}(R^{-1}\vec{x}', R^{-1}\dot{\vec{x}}', t) \end{aligned} \quad (7.47)$$

Hier nehmen wir an, dass die Kraft wie die Geschwindigkeit transformiert.

- Die Komponenten des Drehimpulses im neuen Koordinatensystem sind

$$L'_i = \epsilon_{ijk} x'_j p'_k = \epsilon_{ijk} (R_{jl} x_l) (R_{km} p_m) = (\epsilon_{ijk} R_{jl} R_{km}) x_k p_m \quad (7.48)$$

Da R orthogonal ist, $R_{pq} R_{iq} = \delta_{pi}$, lässt sich der Ausdruck in Klammern umschreiben als

$$\begin{aligned} \epsilon_{ijk} R_{jl} R_{km} &= \epsilon_{pjk} \delta_{pi} R_{jl} R_{km} = \epsilon_{pjk} (R_{iq} R_{pq}) R_{jl} R_{km} \\ &= (\det R) R_{iq} \epsilon_{qlm}. \end{aligned} \quad (7.49)$$

Im letzten Schritt benutzen wir die Definition der Determinante, (7.24). Daraus ergibt sich für den transformierten Drehimpuls

$$L'_i = (\det R) R_{iq} (\epsilon_{qlm} x_l p_m) = (\det R) R_{iq} L_q = \pm R_{iq} L_q. \quad (7.50)$$

Anders als die Kraft transformiert der Drehimpuls nur für eigentliche Rotationen wie die Geschwindigkeit.

- Über die mathematische Definition hinaus wird einer physikalischen Größe Vektorcharakter zugeschrieben, wenn sie sich bei Koordinatentransformationen wie eine Geschwindigkeit transformiert, also entsprechend

$$v'_i = R_{ij} v_j \quad \text{für} \quad \det R = 1. \quad (7.51)$$

Man unterscheidet

– *polare Vektoren*, für die

$$v'_i = R_{ij}v_j \quad \text{für} \quad \det R = \pm 1 \quad (7.52)$$

gilt, und

– *axiale Vektoren*, die bei uneigentlichen Transformationen gespiegelt werden,

$$a'_i = (\det R)R_{ij}a_j . \quad (7.53)$$

Also ist der Drehimpuls ein *axialer* Vektor.

- Dass wir als Standard-Vektor die Geschwindigkeit $\dot{\vec{x}}$ und nicht den Ortsvektor \vec{x} definieren hat einen Grund: Sei \vec{a}' der Ursprung des alten, ungestrichenen Systems im neuen, gestrichenen. Dann hängt der Ortsvektor im gestrichenen System mit dem im ungestrichenen nach

$$x'_i = a'_i + R_{ij}x_j \quad (7.54)$$

zusammen. Für einen weiteren Punkt mit Ortsvektor \vec{y} ist $y'_i = a'_i + y_j R_{ij}$, d.h. der Verbindungsvektor der beiden Punkte transformiert sich wie

$$x'_i - y'_i = R_{ij}(x_i - y_i) . \quad (7.55)$$

Als Vektoren werden in der Physik allgemein solche Größen definiert, die sich wie Koordinatendifferenzen transformieren. Der Ortsvektor selbst ist demnach kein Vektor und wird als *gebundener Vektor* bezeichnet, die Geschwindigkeit hingegen ist ein Vektor.

Kapitel 8

Kräfte in beschleunigten Bezugssystemen

8.1 Zeitabhängige Transformationen

Im letzten Kapitel haben wir Transformationen zwischen Koordinatensystemen betrachtet, die zunächst die betrachtete Physik nicht ändern. Mit anderen Worten, die Änderung der Koordinaten bewirkt keine Änderung der beteiligten Kräfte und damit der Bewegungsgleichungen. Bei der Rotation des Koordinatensystems war das schon weniger klar. In diesem Kapitel betrachten wir Transformationen, die das Kräftegleichgewicht im Bezugssystem tatsächlich ändern.

8.1.1 Winkelgeschwindigkeit

- Wenn sich zwei Koordinatensysteme relativ zueinander beliebig bewegen, sind Ortsvektoren in ihnen durch

$$\vec{x}'(t) = \vec{a}'(t) + R(t) \cdot \vec{x}(t) \quad (8.1)$$

miteinander verknüpft. Dabei sind $\vec{a}'(t)$ und $R(t)$ jetzt vorgegebene Funktionen der Zeit. Daraus erhalten wir die Zeitableitung

$$\dot{\vec{x}}' = \dot{\vec{a}}' + R \cdot \dot{\vec{x}} + \dot{R} \cdot \vec{x} = \dot{\vec{a}}' + R \left[\dot{\vec{x}} + R^{-1} \dot{R} \cdot \vec{x} \right]. \quad (8.2)$$

Der erste Term in den eckigen Klammern ist erwartet, für den zweiten Term brauchen wir eine physikalische Interpretation. Aus der Orthonormalität der Drehmatrix folgt

$$\dot{R}^T R + R^T \dot{R} = 0 \quad \Leftrightarrow \quad (R^T \dot{R})^T = -R^T \dot{R}. \quad (8.3)$$

Die Matrix $R^T \dot{R}$ muss also, wenn man sie transponiert, ihr Vorzeichen wechseln,

$$a_{ii} = 0 \quad \text{und} \quad a_{ij} = -a_{ji}. \quad (8.4)$$

Solche Matrizen heißen schiefsymmetrisch und haben in drei Dimensionen drei unabhängige Elemente. Es muss also möglich sein, die schiefsymmetrische Matrix $R^T \dot{R}$ in der Form

$$R^T \dot{R} = \begin{pmatrix} 0 & -\omega_3 & +\omega_2 \\ \omega_3 & 0 & -\omega_1 \\ -\omega_2 & +\omega_1 & 0 \end{pmatrix} \quad (8.5)$$

zu schreiben. Mithilfe des Levi-Civita-Symbols kann (8.5) als

$$(R^T \dot{R})_{ij} = -\epsilon_{ijk} \omega_k \quad (8.6)$$

dargestellt werden, was sich als sehr nützlich erweisen wird.

- Als Beispiel betrachten wir eine Drehung um die z -Achse um den Winkel $-\omega t$,

$$R = \begin{pmatrix} \cos \omega t & -\sin \omega t & 0 \\ \sin \omega t & \cos \omega t & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad (8.7)$$

Damit ist

$$\begin{aligned} R^T \dot{R} &= \begin{pmatrix} \cos \omega t & \sin \omega t & 0 \\ -\sin \omega t & \cos \omega t & 0 \\ 0 & 0 & 1 \end{pmatrix} \omega \begin{pmatrix} -\sin \omega t & -\cos \omega t & 0 \\ \cos \omega t & -\sin \omega t & 0 \\ 0 & 0 & 0 \end{pmatrix} \\ &= \omega \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \Rightarrow \vec{\omega} = \begin{pmatrix} 0 \\ 0 \\ \omega \end{pmatrix}, \end{aligned} \quad (8.8)$$

- Um die Rolle von ω_i zu verstehen untersuchen wir seine Transformationseigenschaften. Dazu nehmen wir an dass das ungestrichene System in (8.1) einem dritten, gestirnten System durch

$$\vec{x} = \vec{c} + S \vec{x}^* \quad (8.9)$$

hervorgeht. Dabei seien die Verschiebung \vec{c} und die Drehung S nun konstant. Für gestrichene Vektoren \vec{x}' gilt nach (8.1)

$$\begin{aligned} \vec{x}' &= \vec{a}' + R(\vec{c} + S \vec{x}^*) = (\vec{a}' + R\vec{c}) + (RS) \vec{x}^* \\ &= \vec{a}'' + R^* \vec{x}^*, \end{aligned} \quad (8.10)$$

wobei wir in der zweiten Reihe gestirnte Koordinaten analog zu (8.1) definieren. Also ist $R^* = RS$, und damit

$$\begin{aligned} -\epsilon_{ijk} \omega_k^* &= (R^{*T} \dot{R}^*)_{ij} = (S^T R^T \dot{R} S)_{ij} \\ &= S_{ai} R_{ba} \dot{R}_{bc} S_{cj} = -S_{ai} \epsilon_{acl} \omega_l S_{cj}, \end{aligned} \quad (8.11)$$

Wegen $SS^T = I$ folgt $S_{lk} S_{nk} \omega_n = \delta_{ln} \omega_n = \omega_l$, also gilt

$$\epsilon_{ijk} \omega_k^* = \epsilon_{acl} S_{ai} S_{cj} S_{lk} S_{nk} \omega_n = \det S \epsilon_{ijk} S_{nk} \omega_n. \quad (8.12)$$

Demnach transformiert sich $\vec{\omega}$ unter Koordinatendrehungen nach

$$\vec{\omega}^* = \det S S^T \vec{\omega} \quad \Leftrightarrow \quad \vec{\omega} = \det S S \vec{\omega}^* \quad (8.13)$$

was zeigt, dass der Vektor $\vec{\omega}$ ein *axialer* Vektor ist, der nach unserer vorigen Betrachtung in die Richtung der Drehachse zeigt.

- Setzen wir jetzt noch (8.6) in (8.2) ein, dann folgt

$$\dot{\vec{x}}' = \dot{\vec{a}}' + R(\dot{\vec{x}} + \vec{\omega} \times \vec{x}) . \quad (8.14)$$

und analog für einen beliebigen Vektor \vec{v} mit $\vec{v}' = R\vec{v}$

$$\dot{\vec{v}}' = R(\dot{\vec{v}} + \vec{\omega} \times \vec{v}) . \quad (8.15)$$

8.1.2 Infinitesimale Transformationen

- Betrachten wir statt der Drehung um einen endlichen Winkel die Drehung zwischen einem Zeitpunkt t und einem infinitesimal späteren Zeitpunkt $t + dt$, dann erhalten wir für die zeitabhängige Drehmatrix die Taylorreihe

$$R(t + dt) = R(t) + \dot{R}(t)dt = R(t) \left[I + R^T(t) \dot{R}(t) dt \right] . \quad (8.16)$$

Daraus wird mit (8.6)

$$R_{ik}(t + dt) = R_{ij}(t) \left[\delta_{jk} - \epsilon_{jkl} \omega_l(t) dt \right] . \quad (8.17)$$

Seien die Koordinatenachsen zunächst so orientiert, dass wie vorhin beschrieben $\vec{\omega}$ zur Zeit t in \vec{e}_3 -Richtung zeigt. Weiterhin sollen die gestrichenen Achsen bei t mit den ungestrichenen übereinstimmen, dann gilt

$$R(t) = I , \quad (R^T \dot{R})(t) = \begin{pmatrix} 0 & -|\omega| & 0 \\ |\omega| & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \quad (8.18)$$

und damit nach (8.16)

$$R(t + dt) = \begin{pmatrix} 1 & -|\omega|dt & 0 \\ |\omega|dt & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} . \quad (8.19)$$

Dies ist eine infinitesimalen Drehung um die \vec{e}_3 -Achse um den Winkel $d\varphi = |\omega|dt$ im positiven Drehsinn und im Zeitintervall dt . Deswegen wird $\vec{\omega}$ als der Vektor der momentanen Winkelgeschwindigkeit bezeichnet. Seine Richtung gibt die Lage der momentanen Drehachse an.

- Der implizit oben benutzte Begriff der infinitesimalen Transformationen ist in sehr vielen Bereichen der Physik sehr wichtig, z.B. in der Quantenmechanik oder der Feldtheorie. Deshalb betrachten wir sie noch etwas näher. Für eine komplett infinitesimale Transformation gilt

$$\vec{x}' = \vec{a}' + R\vec{x} = d\vec{a}' + (I + dR)\vec{x}. \quad (8.20)$$

Wegen der Orthonormalität ist $R^T R = I$, und daher

$$(I + dR)^T (I + dR) = I \quad \Rightarrow \quad dR^T + dR = 0. \quad (8.21)$$

Auch dR ist schiefsymmetrisch und kann analog zu (8.6) als

$$dR_{ij} = -\epsilon_{ijk} d\varphi_k \quad (8.22)$$

dargestellt werden, wobei die $d\varphi_k$ infinitesimale Drehwinkel sind. Für infinitesimale Koordinatentransformationen gilt demnach allgemein

$$\vec{x}' = \vec{x} + d\vec{a}' + d\vec{\varphi} \times \vec{x}. \quad (8.23)$$

- Wie erwähnt, heißt eine Transformation *aktiv*, bei der das physikalische System sich in einem festen Koordinatensystem bewegt, und *passiv*, wenn das physikalische System fest bleibt, aber das Koordinatensystem bewegt wird. Mathematisch sind diese Arten der Transformation äquivalent, aber sie müssen physikalisch streng unterschieden werden.

8.2 Bewegung auf der rotierenden Erde

Als ein Beispiel für Bewegung in einem nichtinertialen System betrachten wir nun Kräfte, die auf der rotierenden Erde auftreten. Wir führen dazu zwei Koordinatensysteme ein, nämlich eines, das fest mit einem Punkt auf der Erdoberfläche verbunden ist, und ein Inertialsystem, in das wir die rotierende Erde einbetten.

8.2.1 Scheinkräfte

- Gegeben seien also ein (gestrichenes) Inertialsystem mit dem Ursprung im Erdmittelpunkt und ein (ungestrichenes) System, das mit der Erdoberfläche verbunden ist und seinen Ursprung in \vec{a}' auf der Erdoberfläche hat. Die \vec{e}_3 -Achse zeige in \vec{a}' senkrecht von der Erdoberfläche weg.

- Ein Vektor im ungestrichenen Oberflächensystem geht durch eine zeitabhängige Rotation und eine Verschiebung aus dem gestrichenen Inertialsystem hervor

$$\vec{x}' = \vec{a}' + R\vec{x}. \quad (8.24)$$

Aus Sicht des Oberflächensystems ist der Ortsvektor des Erdmittelpunkts \vec{a} konstant, so dass die Transformation die einfachere Form

$$\vec{x}' = R(\vec{a} + \vec{x}), \quad (8.25)$$

bekommt. Die Geschwindigkeit im gestrichenen Inertialsystem ist nach (8.15)

$$\dot{\vec{x}}' = R \left[\dot{\vec{x}} + \vec{\omega} \times (\vec{a} + \vec{x}) \right], \quad (8.26)$$

und die Beschleunigung entsprechend

$$\ddot{\vec{x}}' = \dot{R} \left[\dot{\vec{x}} + \vec{\omega} \times (\vec{a} + \vec{x}) \right] + R \left[\ddot{\vec{x}} + \dot{\vec{\omega}} \times (\vec{a} + \vec{x}) + \vec{\omega} \times \dot{\vec{x}} \right]. \quad (8.27)$$

- Für eine Bewegungsgleichung im ungestrichenen Oberflächensystem multiplizieren wir von links mit $R^T = R^{-1}$ und erhalten damit die Beschleunigung im erdfesten System,

$$R^T \ddot{\vec{x}}' = R^T \dot{R} \left[\dot{\vec{x}} + \vec{\omega} \times (\vec{a} + \vec{x}) \right] + \left[\ddot{\vec{x}} + \dot{\vec{\omega}} \times (\vec{a} + \vec{x}) + \vec{\omega} \times \dot{\vec{x}} \right]. \quad (8.28)$$

Den ersten Term schreiben wir in Komponenten, verwenden dabei das Ergebnis (8.6) und erhalten

$$\begin{aligned} (R^T \dot{R})_{ij} \left[\dot{\vec{x}} + \vec{\omega} \times (\vec{a} + \vec{x}) \right]_j &= -\epsilon_{ijk} \omega_k \left[\dot{x}_j + (\vec{\omega} \times (\vec{a} + \vec{x}))_j \right] \\ &= (\vec{\omega} \times \dot{\vec{x}})_i + [\vec{\omega} \times (\vec{\omega} \times (\vec{a} + \vec{x}))]_i. \end{aligned} \quad (8.29)$$

Indem wir dieses Ergebnis in (8.28) einsetzen, folgt

$$\begin{aligned} R^T \ddot{\vec{x}}' &= \vec{\omega} \times \dot{\vec{x}} + \vec{\omega} \times (\vec{\omega} \times (\vec{a} + \vec{x})) + \ddot{\vec{x}} + \dot{\vec{\omega}} \times (\vec{a} + \vec{x}) + \vec{\omega} \times \dot{\vec{x}} \\ &= \ddot{\vec{x}} + \dot{\vec{\omega}} \times (\vec{a} + \vec{x}) + 2\vec{\omega} \times \dot{\vec{x}} + \vec{\omega} \times (\vec{\omega} \times (\vec{a} + \vec{x})). \end{aligned} \quad (8.30)$$

- Setzen wir diese Beschleunigung im Sinne des zweiten Newtonschen Axioms einer externen Kraft $\vec{F}_{\text{ext}} = -\vec{\nabla}V$ gleich, erhalten wir eine Bewegungsgleichung mit einer Reihe Kraftterme,

$$\begin{aligned} m\ddot{\vec{x}} &= -m \left[\dot{\vec{\omega}} \times (\vec{a} + \vec{x}) \right] - 2m\vec{\omega} \times \dot{\vec{x}} \\ &\quad - m \left[\vec{\omega} \times (\vec{\omega} \times \vec{x}) \right] - m \left[\vec{\omega} \times (\vec{\omega} \times \vec{a}) \right] - \vec{\nabla}V(\vec{x}). \end{aligned} \quad (8.31)$$

Man kann diese vier weiteren Terme als *Scheinkräfte* auffassen. Die einzelnen Scheinkräfte haben die Bedeutung:

1. Die Kraft $-m [\dot{\vec{\omega}} \times (\vec{a} + \vec{x})]$ tritt nur auf, wenn die Drehachse sich ändert. Für die Erde ist $\vec{\omega} \approx \text{const.}$ und damit $\dot{\vec{\omega}} \approx 0$, abgesehen von Präzession und Polschwankungen.
2. Die Corioliskraft $-2m(\vec{\omega} \times \dot{\vec{x}}) =: \vec{F}_C$ steht senkrecht auf $\vec{\omega}$ und $\dot{\vec{x}}$ und verschwindet für Bewegungen längs der Drehachse $\vec{\omega}$.
3. Die Zentrifugalkraft $-m [\vec{\omega} \times (\vec{\omega} \times \vec{x})] =: \vec{F}_Z$ kann noch vereinfacht werden. Dabei benutzen wir

$$\begin{aligned} \vec{\omega} \times \vec{\omega} \times \vec{x} &= \epsilon_{ijk} \omega_j (\epsilon_{klm} \omega_l x_m) \\ &= (\delta_{il} \delta_{jm} - \delta_{im} \delta_{jl}) \omega_j \omega_l x_m = \omega_i \omega_j x_j - x_i \omega_j \omega_j, \end{aligned} \quad (8.32)$$

und erhalten

$$\vec{F}_Z = m [\vec{\omega}^2 \vec{x} - \vec{\omega}(\vec{\omega} \cdot \vec{x})] = m\omega^2 \left[\vec{x} - \frac{\vec{\omega}(\vec{x} \cdot \vec{\omega})}{\omega^2} \right] = m\omega^2 \vec{y} \quad (8.33)$$

wobei \vec{y} der senkrechte Abstand des Massenpunkts von der Drehachse ist.

4. Der Term $-m [\vec{\omega} \times (\vec{\omega} \times \vec{a})]$ ist schließlich unabhängig von \vec{x} , aber proportional zu m . Er liefert einen Beitrag zur Erdbeschleunigung, $\vec{g} \rightarrow \vec{g} - \vec{\omega} \times (\vec{\omega} \times \vec{a})$.
- Allen Scheinkräften ist gemein, dass sie ihren Ursprung in der Trägheitskraft $-m\ddot{\vec{x}}$ haben. Darüber hinaus haben sie die Proportionalität zu m mit der Schwerkraft gemein, was die Allgemeine Relativitätstheorie im Äquivalenzprinzip benutzt. Scheinkräfte zeigen an, dass das Bezugssystem kein Inertialsystem ist.

8.2.2 Corioliskraft

- Seien \vec{e}_1 und \vec{e}_2 nun nach Süden bzw. nach Osten orientiert, und \vec{e}_3 zeige weiterhin senkrecht nach oben. Der Ursprung des erdfesten Systems befinde sich bei der geografischen Breite ϑ , so dass der Winkel zur Nordpolrichtung \vec{e}_3' und damit der Winkelgeschwindigkeit der Erdrotation durch $90^\circ - \vartheta$ gegeben ist. Im Oberflächensystem ist sie dann

$$\vec{\omega} = \omega \begin{pmatrix} -\cos \vartheta \\ 0 \\ \sin \vartheta \end{pmatrix} = \omega \begin{pmatrix} 0 \\ 0 \\ \sin \vartheta \end{pmatrix} + \omega \begin{pmatrix} -\cos \vartheta \\ 0 \\ 0 \end{pmatrix} = \vec{\omega}_\perp + \vec{\omega}_\parallel \quad (8.34)$$

Die Winkelgeschwindigkeit $\vec{\omega}$ hat also orthogonale Komponenten senkrecht zur Erdoberfläche und tangential in Nord-Süd-Richtung.

- Entsprechend hat die Corioliskraft zwei Komponenten

$$\vec{F}_C = 2m(\dot{\vec{x}} \times \vec{\omega}_\perp) + 2m(\dot{\vec{x}} \times \vec{\omega}_\parallel) . \quad (8.35)$$

Weil für Bewegung auf der Erdoberfläche $\dot{\vec{x}} \perp \vec{e}_3$ ist, ist der erste Term tangential zur Erdoberfläche. Er zeigt auf Nordhalbkugel ($\vartheta > 0$) nach rechts, auf der Südhalbkugel ($\vartheta < 0$) nach links relativ zur Richtung von $\dot{\vec{x}}$. Dies ist von großer Bedeutung für Stürme, Meeresströmungen, Flussläufe, Ballistik usw. Der zweite Term steht senkrecht auf der Erdoberfläche.

Kapitel 9

Bewegung starrer Körper

9.1 Tensoren

In den vergangenen Kapiteln haben wir verschiedene Definitionen von Vektoren gegeben, die mehr oder weniger geometrisch oder algebraisch waren, und die wir je nach Fragestellung ausgenutzt haben. Als nächstes müssen wir uns mit Objekten befassen, die ähnlich wie Vektoren über ihre Transformationseigenschaften definiert werden, aber eben keine Skalare oder Vektoren sind. Sie heißen Tensoren.

9.1.1 Der Trägheitstensor

- Bevor wir uns der mathematischen Definition von Tensoren zuwenden, beginnen wir mit einem Beispiel. Die kinetische Energie eines starren Körpers bestehend aus den Massenpunkte m_i an den Orten \vec{x}_i im Inertialsystem ist

$$T = \frac{1}{2} \sum_{i=1}^N m_i \dot{\vec{x}}_i^2, \quad (9.1)$$

- Aus dem körperfesten System \vec{x} , definiert durch $\dot{\vec{x}}_i = 0$, erhalten wir das Inertialsystem nach (8.15) als $\dot{\vec{x}}_i' = R(t)(\dot{\vec{x}}_i + \vec{\omega} \times \vec{x}_i)$ und damit

$$\begin{aligned} T &= \frac{1}{2} \sum_{i=1}^N m_i (\vec{\omega} \times \vec{x}_i)^2 \\ &= \frac{1}{2} \sum_{i=1}^N m_i (\omega_j \omega_k \delta_{jk} \vec{x}_i^2 - \omega_j \omega_k x_{i,j} x_{i,k}) \\ &= \frac{1}{2} \omega_j \left[\sum_{i=1}^N m_i (\vec{x}_i^2 \delta_{jk} - x_{i,j} x_{i,k}) \right] \omega_k =: \frac{1}{2} \omega_j \Theta_{jk} \omega_k. \end{aligned} \quad (9.2)$$

Das Objekt Θ ist offensichtlich eine Matrix, deren genaue Eigenschaften wir im folgenden untersuchen werden.

9.1.2 Tensoren als lineare Abbildungen

- Wir hatten das Skalarprodukt als eine Abbildung eingeführt, die zwei Vektoren (v_1, v_2) aus einem Vektorraum V eine Zahl aus dem Körper K zuordnet,

$$V \times V \rightarrow K, \quad (v_1, v_2) \mapsto \langle v_1, v_2 \rangle \in K. \quad (9.3)$$

Weiterhin hatten wir vom Skalarprodukt verlangt, dass es bilinear sei, d.h. linear in jedem seiner beiden Argumente.

- Alternativ können wir sagen, dass durch das Skalarprodukts ein Satz Vektoren v_1 andere Vektoren v_2 linear in den Körper K abzubilden. Damit definiert jeder Vektor $v_1 \in V$ eine lineare Abbildung des Vektorraums in den Körper K ,

$$V : V \rightarrow K, \quad v_1(v_2) \mapsto \langle v_1, v_2 \rangle. \quad (9.4)$$

- Entsprechend führen wir nun *Tensoren* als multilineare Abbildungen von V nach K ein, d.h. als Abbildungen, die einer beliebigen Menge von n Vektoren $v_i \in V$ eine Zahl aus dem Körper K zuordnen,

$$T : \underbrace{V \times \dots \times V}_{n \text{ mal}} \rightarrow K, \quad (v_1, \dots, v_n) \mapsto T(v_1, \dots, v_n). \quad (9.5)$$

Die Anzahl n der Vektoren, die der Tensor nach K abbilden kann, heißt sein *Rang*, oder man nennt ihn Tensor n -ter *Stufe*. Multilinearität bedeutet, dass der Tensor in jedem seiner Argumente linear ist,

$$T(\dots, v_i + \lambda w_i, \dots) = T(\dots, v_i, \dots) + \lambda T(\dots, w_i, \dots). \quad (9.6)$$

Vektoren selbst sind demnach Tensoren erster Stufe. Wir beschränken uns im Folgenden der Einfachheit halber auf Tensoren zweiter Stufe, aber die folgenden Aussagen gelten entsprechend auch für Tensoren beliebiger Stufe.

9.1.3 Transformationseigenschaften

- Zunächst stellen wir fest, dass Tensoren vom selben Rang einen K -Vektorraum \mathcal{T} bilden, wenn wir die offensichtliche Addition und Multiplikation mit Skalaren definieren,

$$(T_1 + \lambda T_2)(v_1, v_2) = T_1(v_1, v_2) + \lambda T_2(v_1, v_2). \quad (9.7)$$

- Wie bei Vektoren reicht es wegen der Multilinearität der Tensoren aus, zu wissen, welche Zahlen der Tensor T den Basisvektoren e_i zuordnet. Wir bezeichnen sie als *Komponenten* T_{ij} des Tensors,

$$\begin{aligned} T(e_i, e_j) &= T_{ij}, \quad 1 \leq i, j \leq n \\ \Rightarrow T(v, w) &= T(v_i e_i, w_j e_j) = v_i w_j T(e_i, e_j) = v_i T_{ij} w_j. \end{aligned} \quad (9.8)$$

Tensoren zweiter Stufe lassen sich also durch Matrizen darstellen, die ihre Komponenten enthalten, sobald man eine Basis für den Vektorraum festgelegt hat.

- Jetzt können wir die Transformation eines Tensors $T \rightarrow T'$ unter einer orthogonalen Transformation $e'_i = R_{ik} e_k$ ableiten,

$$\begin{aligned} v_i T_{ij} w_j &= T(v, w) = T(v'_i e'_i, w'_j e'_j) \\ &= T(v'_i R_{ik} e_k, w'_j R_{jl} e_l) \\ &= v'_i R_{ik} T_{kl} R_{jl} w'_j \quad \Leftrightarrow \quad T'_{ij} = R_{ik} T_{kl} R_{lj} \end{aligned} \quad (9.9)$$

In Matrixschreibweise heisst das

$$T' = RTR^T = RTR^{-1}. \quad (9.10)$$

- Um festzustellen, ob das Objekt Θ in (9.2) tatsächlich ein Tensor ist, untersuchen wir sein Transformationsverhalten. Wir definieren ein neues orthonormales Koordinatensystem \vec{e}_i^* mit gleichem Ursprung, also $x_j = R_{jk} x_k^*$. Wegen der Orthonormalität von R ist das Skalarprodukt invariant,

$$\begin{aligned} \vec{x}^2 &= R_{jl} x_l^* R_{jm} x_m^* \\ &= R_{lj}^T R_{jm} x_l^* x_m^* = \delta_{lm} x_l^* x_m^* = \vec{x}^{*2} \\ \Rightarrow \Theta_{jk} &= \sum_{i=1}^N m_i \left(\vec{x}_i^{*2} \delta_{jk} - R_{jl} x_{i,l}^* R_{km} x_{i,m}^* \right) \\ &= \sum_{i=1}^N m_i \left(\vec{x}_i^{*2} R_{jl} R_{km} \delta_{lm} - R_{jl} R_{km} x_{i,l}^* x_{i,m}^* \right) \\ &= R_{jl} \Theta_{lm}^* R_{mk}^T. \end{aligned} \quad (9.11)$$

In der Tat ist Θ als ein *Tensor zweiter Stufe*. Er wird *Trägheitstensor* des starren Körpers genannt.

9.1.4 Tensorprodukt

- Aus zwei Vektoren (v_1, v_2) lässt sich ein Tensor zweiter Stufe aufbauen, indem man ein sogenanntes *Tensorprodukt* oder *äußere Produkt* der beiden bildet,

$$\otimes : V \times V \rightarrow \mathcal{T}, \quad (v_1, v_2) \mapsto v_1 \otimes v_2. \quad (9.12)$$

Der so entstandene Tensor ordnet einem Paar von Vektoren w_1, w_2 das Produkt der Abbildungen $v_1(w_1)$ und $v_2(w_2)$ zu,

$$(v_1 \otimes v_2)(w_1, w_2) = v_1(w_1) v_2(w_2) . \quad (9.13)$$

- Mit Hilfe der Basisvektoren e_i kann man durch das Tensorprodukt eine Basis der Tensoren bilden, also für Tensoren zweiter Stufe die Elemente $(e_i \otimes e_j)$. Mit deren Hilfe schreiben wir

$$\begin{aligned} T(v, w) &= T_{ij} (e_i \otimes e_j)(v_k e_k, w_l e_l) = T_{ij} v_k w_l e_i(e_k) e_j(e_l) \\ &= T_{ij} v_k w_l \delta_{ik} \delta_{jl} = v_i w_j T_{ij} . \end{aligned} \quad (9.14)$$

Der Vektorraum der Tensoren zweiter Stufe im n -dimensionalen Raum hat also die Dimension n^2 hat, weil er durch die n^2 Basis-tensoren $e_i \otimes e_j$ aufgespannt wird.

- Wendet man einen Tensor T auf zwei gleiche Basisvektoren an, erhält man seine *Spur*

$$T(e_k, e_k) = T_{ij} (e_i \otimes e_j)(e_k, e_k) = T_{ij} \delta_{ik} \delta_{jk} = T_{kk} . \quad (9.15)$$

- Auch hier zeigt (9.14), dass eine orthonormale Koordinatentransformation $e'_i = R_{ij} e_j$ die Tensorkomponenten T_{ij} wie folgt transformiert:

$$\begin{aligned} T &= T_{ij} (e_i \otimes e_j) = T'_{kl} (e'_k \otimes e'_l) \\ &= T'_{kl} R_{ki} R_{lj} (e_i \otimes e_j) = R^T_{ik} T'_{kl} R_{lj} (e_i \otimes e_j) . \end{aligned} \quad (9.16)$$

9.2 Diagonalisierung

9.2.1 Diagonalisierung von Matrizen

- Wir haben gesehen, dass sich Tensoren bei orthonormalen Koordinatendrehungen wie $T' = RTR^T$ transformieren. Dies gilt allgemein auch für Matrizen A , wenn wir verlangen, dass Gleichungen der Form $y = Ax$ durch Koordinatentransformationen R ihre Form nicht verändern, also nach der Transformation $y' = A'x'$ gilt,

$$y' = Ry = R Ax = R A R^T x' = A' x' \quad \Leftrightarrow \quad A' = R A R^T \quad (9.17)$$

Können wir diese Transformation benutzen, um eine Matrix A' in die besonders einfache *Diagonalform*

$$A' = \begin{pmatrix} A_1 & 0 & 0 \\ 0 & A_2 & 0 \\ 0 & 0 & A_3 \end{pmatrix} \quad (9.18)$$

zu bringen? Wenn wir (9.17) von links mit R^T multiplizieren, dann erhalten wir

$$\sum_k R_{ik}^T A'_{kj} = A_j R_{ij}^T = \sum_k A_{ik} R_{kj}^T \quad (9.19)$$

Für jeden festen Wert von j ist dies ein lineares Gleichungssystem, in dem wir R_{ij}^T als vertikaler Vektor r_i auffassen können. Für die Matrix R sind sie die Zeilenvektoren $\vec{r} = (r_{j1}, r_{j2}, r_{j3})$. Wir können dann (9.19) als

$$A\vec{r} = A_j\vec{r}, \quad (9.20)$$

schreiben. Diese sogenannten *Eigenvektoren* \vec{r}_i sind ein ausgezeichnetes Koordinatensystem, in dem die Matrix A Diagonalförm annimmt. Es heißt *Hauptachsensystem* von A .

- Allgemein betrachten wir also ein lineares Gleichungssysteme der Form

$$Av = \lambda v \quad \Leftrightarrow \quad (A - \lambda I)v = 0. \quad (9.21)$$

Vektoren v , für die wir diese einfache Form erhalten, heißen *Eigenvektoren*, die zugehörigen Zahlen λ heißen *Eigenwerte*. Wenn ein Inverses existiert, dann folgt aus dieser Gleichung

$$v = (A - \lambda I)^{-1} 0 = 0. \quad (9.22)$$

Wenn (9.21) für nicht verschwindende v gelten soll, dann darf es also kein solches Inverses geben, und damit muss

$$\det(A - \lambda I) = 0 \quad (9.23)$$

sein. Diese Gleichung heißt *charakteristisches Polynom* von A .

- Als Lösungen des charakteristischen Polynoms gibt es nach dem Fundamentalsatz der Algebra immer n Eigenwerte, die komplex sein können. Da wir komplexe Zahlen erst später besprechen, behaupten wir hier ohne Beweis, dass die Eigenwerte dann reell sind, wenn die Matrix symmetrisch ist, $A^T = A$.
- Betrachten wir als Beispiel die symmetrische Matrix

$$A = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}. \quad (9.24)$$

Ihr charakteristisches Polynom lautet

$$\det(A - \lambda I) = \det \begin{pmatrix} -\lambda & 1 & 0 \\ 1 & -\lambda & 0 \\ 0 & 0 & -\lambda \end{pmatrix} = -\lambda(\lambda^2 - 1) = 0 \quad (9.25)$$

und hat die Lösungen $\lambda = -1, 0, 1$. Die Matrix A hat also die Diagonalform

$$A' = \begin{pmatrix} -1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} =: \text{diag}(-1, 0, 1). \quad (9.26)$$

- Die zugehörigen Eigenvektoren \vec{r} sind die Lösungen von

$$(A - \lambda I)\vec{r} = \begin{pmatrix} -\lambda & 1 & 0 \\ 1 & -\lambda & 0 \\ 0 & 0 & -\lambda \end{pmatrix} \begin{pmatrix} r_1 \\ r_2 \\ r_3 \end{pmatrix} = \begin{pmatrix} -\lambda r_1 + r_2 \\ r_1 - \lambda r_2 \\ -\lambda r_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}. \quad (9.27)$$

Für $\lambda = -1$ muss also $r_2 = -r_1$ und $r_3 = 0$ sein. Wir können eine der beiden Komponenten als $r_1 = 1$ wählen und erhalten den Eigenvektor zum Eigenwert $\lambda = -1$

$$\vec{r}_1 = \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}. \quad (9.28)$$

Analoge Rechnungen führen auf die beiden anderen Eigenvektoren

$$\vec{r}_2 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}, \quad \vec{r}_3 = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}. \quad (9.29)$$

9.2.2 Diagonalisierung des Trägheitstensors

- Der Trägheitstensor Θ in (9.2) lässt sich durch eine reelle, symmetrische Matrix darstellen. Sie hat drei reelle Eigenwerte Θ_1 , Θ_2 und Θ_3 , die *Hauptträgheitsmomente*. Die dazugehörigen Eigenvektoren bezeichnen wir als *Hauptträgheitsachsen* oder *Hauptachsen*.
- Betrachten wir als Beispiel zwei Massenpunkte m_1 und m_2 , die durch eine masselose Stange miteinander verbunden sind. m_2 befinde sich im Ursprung, m_1 laufe in der x_1 - x_2 -Ebene um. Dann sind $x_1 = l \cos \varphi$, $x_2 = l \sin \varphi$, $x_3 = 0$, und

$$(\Theta_{jk}) = m_1 l^2 \begin{pmatrix} \sin^2 \varphi & -\cos \varphi \sin \varphi & 0 \\ -\cos \varphi \sin \varphi & \cos^2 \varphi & 0 \\ 0 & 0 & 1 \end{pmatrix}. \quad (9.30)$$

Seine Eigenwerte sind durch das charakteristische Polynom

$$\begin{aligned} \det(\Theta - \lambda) &= \begin{vmatrix} m_1 l^2 \sin^2 \varphi - \lambda & -m_1 l^2 \cos \varphi \sin \varphi & 0 \\ -m_1 l^2 \cos \varphi \sin \varphi & m_1 l^2 \cos^2 \varphi - \lambda & 0 \\ 0 & 0 & m_1 l^2 - \lambda \end{vmatrix} \\ &= (m_1 l^2 - \lambda) [(m_1 l^2 \sin^2 \varphi - \lambda)(m_1 l^2 \cos^2 \varphi - \lambda) \\ &\quad - m_1^2 l^4 \sin^2 \varphi \cos^2 \varphi] \\ &= (m_1 l^2 - \lambda) \lambda (\lambda - m_1 l^2) \end{aligned} \quad (9.31)$$

bestimmt. Die Lösungen sind $\lambda = m_1 l^2, 0$. Die Eigenvektoren zu diesen Eigenwerten sind

$$\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}, \quad \begin{pmatrix} -\sin \varphi \\ \cos \varphi \\ 0 \end{pmatrix}, \quad \begin{pmatrix} \cos \varphi \\ \sin \varphi \\ 0 \end{pmatrix}. \quad (9.32)$$

Kapitel 10

Trägheitstensor

10.1 Volumenintegrale

10.1.1 Motivation

- Wir hatten im letzten Kapitel gesehen, dass der Trägheitstensor für einen starren Körper durch

$$\Theta_{jk} = \sum_{i=1}^N m_i \left[\vec{x}_i^2 \delta_{jk} - x_{i,j} x_{i,k} \right] \quad (10.1)$$

gegeben ist. Wenn der Körper kontinuierlich ist führt man einen Grenzübergang zu beliebig kleinen Massenpunkten durch, die sich an den kontinuierlichen Orten \vec{x} innerhalb des Körpers befinden, das infinitesimal kleine Volumen $dV = dx_1 dx_2 dx_3$ einnehmen und die Masse

$$dm = \rho(\vec{x}) dV \quad (10.2)$$

besitzen, wobei $\rho(\vec{x})$ die Dichte am Ort \vec{x} ist. Dann geht (10.1) in das *Volumenintegral*

$$\int \int \int dx_1 dx_2 dx_3 \rho(\vec{x}) \left[\vec{x}^2 \delta_{jk} - x_j x_k \right] \quad (10.3)$$

über, wobei die Integrationsgrenzen durch den Rand des Körpers vorgegeben sind.

- Allgemein haben 3-dimensionale Volumenintegrale die Form

$$I = \int \int \int dx_1 dx_2 dx_3 f(x_1, x_2, x_3) . \quad (10.4)$$

Hier ist f eine skalare Funktion, vektorwertige Funktionen würden einfach komponentenweise integriert werden.

- Oft ist die größte Schwierigkeit bei Volumenintegralen die Parametrisierung der Integrationsgrenzen. Um das zu illustrieren, berechnen wir das Volumen einer Kugel mit Radius R aus dem Volumenintegral über die Funktion

$$f(\vec{x}) = \begin{cases} 1 & x_1^2 + x_2^2 + x_3^2 \leq R^2 \\ 0 & \text{sonst} \end{cases} . \quad (10.5)$$

Das Integral ist also

$$V = \int_{x_1^2+x_2^2+x_3^2 \leq R^2} dx_1 dx_2 dx_3 \quad (10.6)$$

Wir betrachten die Integrationen eine nach der anderen. Wenn x_1 und x_2 gegeben sind, dann muss x_3 die Bedingung $x_3^2 = R^2 - x_1^2 - x_2^2$ erfüllen, also

$$\int_{-\sqrt{R^2-x_1^2-x_2^2}}^{\sqrt{R^2-x_1^2-x_2^2}} dx_3 = 2 \sqrt{R^2 - x_1^2 - x_2^2} . \quad (10.7)$$

Wenn nur noch x_1 vorgegeben ist, muss x_2 aus dem Intervall $x_1^2 + x_2^2 = R^2$ stammen. Das folgende Integral über x_2 ist also

$$\begin{aligned} & 2 \int_{-\sqrt{R^2-x_1^2}}^{\sqrt{R^2-x_1^2}} \sqrt{R^2 - x_1^2 - x_2^2} dx_2 \\ &= \left[x_2 \sqrt{R^2 - x_1^2 - x_2^2} + (R^2 - x_1^2) \arcsin \frac{x_2}{\sqrt{R^2 - x_1^2}} \right]_{-\sqrt{R^2-x_1^2}}^{\sqrt{R^2-x_1^2}} \\ &= \pi(R^2 - x_1^2) . \end{aligned} \quad (10.8)$$

Schließlich ist das verbleibende Integral über x_1

$$\pi \int_{-R}^R (R^2 - x_1^2) dx_1 = \pi \left(R^2 x_1 - \frac{x_1^3}{3} \right)_{-R}^R = \frac{4\pi}{3} R^3 . \quad (10.9)$$

Diese Rechnung ist unbefriedigend, weil die kartesischen Koordinaten der Symmetrie der Kugel nicht angepasst sind. Deswegen behandeln wir zunächst die Frage, wie Volumenintegrale in anderen, der Situation besser angepassten Koordinaten ausgeführt werden können.

10.1.2 Jacobi-Determinante

- Volumenintegrale werden erheblich vereinfacht, wenn man Koordinaten verwendet, die der Symmetrie des Integrationsvolumens

angepasst sind, z.B. Zylinder- oder Kugelkoordinaten. Nehmen wir an, solche Koordinaten seien durch $u_i(x_j)$ gegeben, dann sind ihre totalen Ableitungen

$$du_i = (\partial_{x_j} u_i) dx_j \quad \text{oder} \quad d\vec{u} = \bar{J} d\vec{x}, \quad (10.10)$$

mit $J_{ij} = \partial_{x_j} u_i$. Wenn die Matrix \bar{J} invertierbar oder $\det \bar{J} \neq 0$ ist folgt daraus

$$d\vec{x} = J d\vec{u} \quad \text{mit} \quad J_{ij} = \partial_{u_j} x_i \quad (10.11)$$

Die Matrix $J = \bar{J}^{-1}$ heißt *Jacobimatrix* der Koordinatentransformation.

- Für das Volumenintegral brauchen wir das dreidimensionale Volumenelement $dV = dx_1 dx_2 dx_3$. Bevor wir es durch das Produkt $du_1 du_2 du_3$ ausdrücken finden wir zuerst eine orthogonale Transformation R mit der J diagonal wird,

$$J' = \text{diag}(J_1, J_2, J_3) \\ \text{mit} \quad d\vec{x}' = R d\vec{x} = R J d\vec{u} = (RJR^T) R d\vec{u} =: J' d\vec{u}' \quad (10.12)$$

In der Diagonalform folgt für das Volumenintegral

$$dx'_1 dx'_2 dx'_3 = J_1 J_2 J_3 du'_1 du'_2 du'_3 = |\det J'| du'_1 du'_2 du'_3 \\ = |\det J| du'_1 du'_2 du'_3, \quad (10.13)$$

wobei die Betragsstriche gewährleisten, dass das Volumenelement positiv bleibt. Weil ein Volumenintegral durch eine Koordinatendrehung nicht verändert wird, erhalten wir

$$dx_1 dx_2 dx_3 = |\det J| du_1 du_2 du_3 \\ \Rightarrow \int f(\vec{x}) dx_1 dx_2 dx_3 = \int f[\vec{x}(\vec{u})] |\det J| du_1 du_2 du_3 \quad (10.14)$$

Die Determinante $\det J$ heißt *Jacobideterminante* oder *Funktionaldeterminante*. Sie wird oft in der Form

$$\det J = \det \left(\frac{\partial x_i}{\partial u_j} \right) = \frac{\partial(x_1, x_2, x_3)}{\partial(u_1, u_2, u_3)} \quad (10.15)$$

geschrieben.

- Als Beispiel berechnen wir wieder das Volumen der Kugel, aber diesmal in Kugelkoordinaten (r, θ, ϕ) . Zunächst gilt

$$x_1 = r \sin \theta \cos \phi, \quad x_2 = r \sin \theta \sin \phi, \quad x_3 = r \cos \theta, \quad (10.16)$$

und daher ist die Jacobimatrix

$$J = \begin{pmatrix} \sin \theta \cos \phi & r \cos \theta \cos \phi & -r \sin \theta \sin \phi \\ \sin \theta \sin \phi & r \cos \theta \sin \phi & r \sin \theta \cos \phi \\ \cos \theta & -r \sin \theta & 0 \end{pmatrix}. \quad (10.17)$$

Ihre Determinante kann man berechnen und erhält

$$\det J = r^2 \sin \theta . \quad (10.18)$$

Das Volumenintegral über die Kugel lautet daher

$$\begin{aligned} \int_0^R dr \int_0^\pi d\theta \int_0^{2\pi} d\phi r^2 \sin \theta &= \int_0^R dr r^2 \int_0^\pi d\theta \sin \theta \int_0^{2\pi} d\phi \\ &= \frac{R^3}{3} \times 2 \times 2\pi = \frac{4\pi}{3} R^3 . \end{aligned} \quad (10.19)$$

- Der Vollständigkeit halber geben wir noch die Jacobideterminante für die Transformation von kartesischen in Zylinderkoordinaten (ρ, ϕ, z) an. Hier ist

$$x_1 = \rho \cos \phi , \quad x_2 = \rho \sin \phi , \quad x_3 = z , \quad (10.20)$$

daher sind die Jacobimatrix und ihre Determinante

$$J = \begin{pmatrix} \cos \phi & -\rho \sin \phi & 0 \\ \sin \phi & \rho \cos \phi & 0 \\ 0 & 0 & 1 \end{pmatrix} \Rightarrow \det J = \rho . \quad (10.21)$$

10.1.3 Trägheitstensor einer Kugel

- Als weiteres Beispiel berechnen wir den Trägheitstensors einer homogenen Kugel mit Radius R und Gesamtmasse M oder Dichte ρ . Nach (10.3) und mit (10.18) ist

$$\Theta_{ij} = \int_0^R dr \int_0^\pi d\theta \int_0^{2\pi} d\phi \rho \left[r^2 \delta_{ij} - x_i x_j \right] r^2 \sin \theta . \quad (10.22)$$

Wir können uns durch eine physikalische Überlegung viel Rechenarbeit sparen. Zunächst müssen die Nebendiagonalelemente von Θ_{ij} alle verschwinden, weil gegenüber einer Kugel keine Orientierung der drei Koordinatenachsen bevorzugt sein kann, weshalb die drei Koordinatenachsen bei beliebiger Orientierung Hauptachsen von Θ sein müssen. Weiterhin müssen alle Diagonalelemente gleich sein, weil die Kugel keine der drei Koordinatenrichtungen auszeichnet. Wir schließen also, dass aus Symmetriegründen

$$\Theta_{11} = \Theta_{22} = \Theta_{33} \quad \text{und} \quad \Theta_{12} = \Theta_{13} = \Theta_{23} = 0 \quad (10.23)$$

gelten müssen.

- Also beschränken wir uns auf die Berechnung von Θ_{11} oder besser $\Theta_{11} + \Theta_{22} + \Theta_{33} = 3\Theta_{11}$,

$$\begin{aligned} 3\Theta_{11} &= \sum_{i=1}^3 \rho \int_0^R dr \int_0^\pi d\theta \int_0^{2\pi} d\phi (r^2 - x_i^2) r^2 \sin \theta \\ &= \rho \int_0^R dr \int_0^\pi d\theta \int_0^{2\pi} d\phi (3r^2 - r^2) r^2 \sin \theta \\ &= 4\pi\rho \int_0^R dr r^4 \int_0^\pi d\theta \sin \theta = 8\pi\rho \frac{R^5}{5}. \end{aligned} \quad (10.24)$$

Wegen $M = 4\pi R^3 \rho/3$ folgt daraus

$$\Theta_{11} = \Theta_{22} = \Theta_{33} = \frac{2}{5}MR^2. \quad (10.25)$$

10.2 Der Satz von Gauß

10.2.1 Herleitung

- Mit Hilfe dieser Volumenintegrale können wir einen Integralsatz ähnlich dem Satz von Stokes herleiten. Der Gaußsche Satz besagt dass Flächenintegrale über Vektorfelder \vec{A} durch Volumenintegrale über deren Divergenz $\vec{\nabla} \cdot \vec{A}$ ersetzt werden können,

$$\int_{\partial V} \vec{A} \cdot d\vec{\sigma} = \int_V \vec{\nabla} \cdot \vec{A} dV, \quad (10.26)$$

wobei ∂V der Rand des Volumens V ist, also seine gesamte Oberfläche, und die Flächenelemente $d\vec{\sigma}$ entlang der Flächennormalen nach außen gerichtet sind.

- Für den Beweis betrachten wir ein kleines, würfelförmiges Volumen mit den Abmessungen dx_1 , dx_2 und dx_3 . Seine Oberfläche besteht aus sechs Quadraten, die sich paarweise gegenüberliegen und parallel zu einer der drei Koordinatenachsen sind. Die äußeren Flächennormalen sind also entweder parallel oder antiparallel zu den Koordinatenachsen. Der Beitrag der beiden zur x_2 - x_3 -Ebene parallelen Quadrate zum Flächenintegral auf der linken Seite von (10.26) ist zum Beispiel

$$\begin{aligned} & -A_1(x_1, x_2, x_3) dx_2 dx_3 + A_1(x_1 + dx_1, x_2, x_3) dx_2 dx_3 \\ &= -A_1(x_1, x_2, x_3) dx_2 dx_3 + A_1(x_1, x_2, x_3) dx_2 dx_3 \\ & \quad + \partial_{x_1} A_1(x_1, x_2, x_3) dx_1 dx_2 dx_3 \\ &= \partial_{x_1} A_1(x_1, x_2, x_3) dV. \end{aligned} \quad (10.27)$$

Entsprechende Beiträge kommen von den beiden anderen Flächenpaaren, so dass wir für das infinitesimale Volumen dV

$$\begin{aligned} \lim_{dV \rightarrow 0} \int_{\partial V} \vec{A} \cdot d\vec{\sigma} &= \int (\partial_{x_1} A_1 + \partial_{x_2} A_2 + \partial_{x_3} A_3) dV \\ &= \int (\vec{\nabla} \cdot \vec{A}) dV \end{aligned} \quad (10.28)$$

erhalten. Nun denken wir uns wie beim Satz von Stokes ein endlich großes Volumen aus infinitesimal kleinen Würfeln aufgebaut. Dann heben sich die inneren Flächenintegrale heraus, weil an der Trennfläche zwischen zwei benachbarten Würfeln die Beiträge betragsgleich aber antiparallelen sind. Es bleibt also nur das Flächenintegral über den Rand des Volumens übrig, während die Divergenz über das gesamte Volumen integriert wird.

10.2.2 Transformation von Divergenz und Rotation

- Die Integralsätze von Gauß und Stokes geben uns eine einfache Methode, die Differentialoperatoren Divergenz und Rotation in krummlinig-orthogonale Koordinatensysteme zu übersetzen. Wir hatten in (5.76) gesehen, dass nach einer Transformation in solche Koordinaten das Längenelement ds durch

$$ds^2 = (h_1 dq_1)^2 + (h_2 dq_2)^2 + (h_3 dq_3)^2 \quad (10.29)$$

ausgedrückt werden kann, wobei die h_i durch

$$h_j^2 = \left(\frac{\partial x_1}{\partial q_j} \right)^2 + \left(\frac{\partial x_2}{\partial q_j} \right)^2 + \left(\frac{\partial x_3}{\partial q_j} \right)^2 \quad (10.30)$$

wie in (5.75) gegeben sind. Analog ist das Volumenintegral in unserer üblichen Diagonalform

$$dV = h_1 h_2 h_3 dq_1 dq_2 dq_3 . \quad (10.31)$$

- Führen wir die Betrachtung aus der Herleitung des Satzes von Gauß nun in krummlinig-orthogonalen Koordinaten durch, erhalten wir zunächst für die Beiträge zweier längs der q_1 -Achse gegenüberliegender Begrenzungsflächen eines infinitesimalen Volumenelements zum Flächenintegral

$$\begin{aligned} &-(A_{q_1} h_2 h_3)(q_1, q_2, q_3) dq_2 dq_3 + (A_{q_1} h_2 h_3)(q_1 + dq_1, q_2, q_3) dq_2 dq_3 \\ &= \frac{\partial(A_{q_1} h_2 h_3)}{\partial q_1} dq_1 dq_2 dq_3 . \end{aligned} \quad (10.32)$$

Allgemein hängen auch die h_i von den Koordinaten q_j ab. Das gesamte Flächenintegral wird

$$\left(\frac{\partial(A_{q_1} h_2 h_3)}{\partial q_1} + \frac{\partial(A_{q_2} h_1 h_3)}{\partial q_2} + \frac{\partial(A_{q_3} h_1 h_2)}{\partial q_3} \right) dq_1 dq_2 dq_3 . \quad (10.33)$$

Nach dem Satz von Gauß gilt also

$$\vec{\nabla} \cdot \vec{A} = \frac{1}{h_1 h_2 h_3} \left(\frac{\partial(A_{q_1} h_2 h_3)}{\partial q_1} + \frac{\partial(A_{q_2} h_1 h_3)}{\partial q_2} + \frac{\partial(A_{q_3} h_1 h_2)}{\partial q_3} \right). \quad (10.34)$$

- Durch eine völlig analoge Betrachtung, ausgehend von (5.67) für die dritte Komponente der Rotation, erhält man die Darstellung der Rotation in krummlinig-orthogonalen Koordinaten,

$$\begin{aligned} (\vec{\nabla} \times \vec{A})_{q_1} &= \frac{1}{h_2 h_3} \left(\frac{\partial(A_{q_3} h_3)}{\partial q_2} - \frac{\partial(A_{q_2} h_2)}{\partial q_3} \right), \\ (\vec{\nabla} \times \vec{A})_{q_2} &= \frac{1}{h_1 h_3} \left(\frac{\partial(A_{q_1} h_1)}{\partial q_3} - \frac{\partial(A_{q_3} h_3)}{\partial q_1} \right), \\ (\vec{\nabla} \times \vec{A})_{q_3} &= \frac{1}{h_1 h_2} \left(\frac{\partial(A_{q_2} h_2)}{\partial q_1} - \frac{\partial(A_{q_1} h_1)}{\partial q_2} \right). \end{aligned} \quad (10.35)$$

- Für Kugelkoordinaten hatten wir $(q_1, q_2, q_3) = (r, \theta, \phi)$ und

$$h_r = 1, \quad h_\theta = r, \quad h_\phi = r \sin \theta. \quad (10.36)$$

Daher ist $h_r h_\theta h_\phi = r^2 \sin \theta$ und

$$\vec{\nabla} \cdot \vec{A} = \frac{1}{r^2} \frac{\partial(r^2 A_r)}{\partial r} + \frac{1}{r \sin \theta} \frac{\partial(\sin \theta A_\theta)}{\partial \theta} + \frac{1}{r \sin \theta} \frac{\partial A_\phi}{\partial \phi}. \quad (10.37)$$

Für die Komponenten der Rotation erhalten wir

$$\begin{aligned} (\vec{\nabla} \times \vec{A})_r &= \frac{1}{r \sin \theta} \frac{\partial(\sin \theta A_\phi)}{\partial \theta} - \frac{1}{r \sin \theta} \frac{\partial A_\theta}{\partial \phi}, \\ (\vec{\nabla} \times \vec{A})_\theta &= \frac{1}{r \sin \theta} \frac{\partial A_r}{\partial \phi} - \frac{1}{r} \frac{\partial(r A_\phi)}{\partial r}, \\ (\vec{\nabla} \times \vec{A})_\phi &= \frac{1}{r} \frac{\partial(r A_\theta)}{\partial r} - \frac{1}{r} \frac{\partial A_r}{\partial \theta}. \end{aligned} \quad (10.38)$$

- Betrachten wir als Beispiel das Radialfeld

$$\vec{F} = e^{-r} \vec{e}_r \quad (10.39)$$

und integrieren es über eine Kugelfläche mit Radius R um den Ursprung. Das gerichtete Flächenelement ist

$$d\vec{\sigma} = r^2 \sin \theta d\theta d\phi \vec{e}_r, \quad (10.40)$$

so dass das Flächenintegral der Kraft über die Kugelfläche bei $r = R$

$$R^2 \int_0^\pi \sin \theta d\theta \int_0^{2\pi} d\phi \vec{F}(r, \theta, \phi) \cdot \vec{e}_r = 4\pi R^2 e^{-R} \quad (10.41)$$

ist. Andererseits ist die Divergenz von \vec{F} wegen (10.37) durch

$$\vec{\nabla} \cdot \vec{F} = \frac{1}{r^2} \frac{\partial(r^2 F_r)}{\partial r} = \frac{1}{r^2} (2r - r^2) e^{-r} \quad (10.42)$$

gegeben, so dass das Volumenintegral darüber

$$\begin{aligned} 4\pi \int_0^R r^2 dr \vec{\nabla} \cdot \vec{F} &= 4\pi \int_0^R (2r - r^2) e^{-r} dr \\ &= 4\pi r^2 e^{-r} \Big|_0^R = 4\pi R^2 e^{-R} \end{aligned} \quad (10.43)$$

ist, wie es nach dem Gaußschen Satz sein muss.

Kapitel 11

Harmonische Schwingungen

11.1 Harmonischer Oszillator

11.1.1 Linearisierte Kraft

- Auf einen Körper im mechanischen *Gleichgewicht* wirkt insgesamt keine Kraft, $F = 0$. Bei genügend kleiner Auslenkung aus der Gleichgewichtslage x_0 kann die Kraft in eine Taylorreihe entwickelt werden,

$$F = F(x_0) + F'(x_0)(x - x_0) + \dots \approx F'(x_0)(x - x_0) \quad (11.1)$$

Der Einfachheit halber verschieben nehmen wir $x_0 = 0$ an. Weiter nehmen wir an, dass $F'(x_0) \neq 0$ ist und vernachlässigen Terme höherer Ordnung in der Taylor-Reihe, d.h. wir *linearisieren* die Kraft um die Gleichgewichtslage und erhalten

$$F = F'(0) x \quad (11.2)$$

- Wenn $F'(0) > 0$ ist, entfernt sich das System aus der Gleichgewichtslage, und das Gleichgewicht ist *labil*. Das Gleichgewicht ist für $F'(0) < 0$ stabil. Entsprechend schreiben wir für die *linearisierte Kraft*

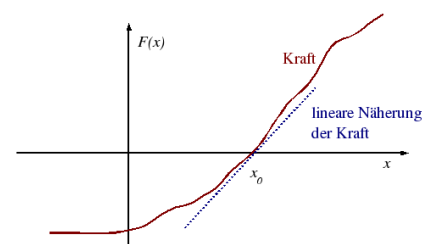
$$F'(0) := -k \quad \text{oder} \quad F(x) = -kx, \quad k > 0. \quad (11.3)$$

Die Bewegungsgleichung ist dann

$$m\ddot{x} = -kx \quad (11.4)$$

und ergibt

$$\ddot{x} + \omega_0^2 x = 0 \quad \text{mit} \quad \omega_0 := \sqrt{\frac{k}{m}}. \quad (11.5)$$



Entwicklung der Kraft um die Gleichgewichtslage

11.1.2 Freie Schwingungen

- Die Bewegungsgleichung (11.5) ist offenbar eine gewöhnliche, lineare, homogene Differentialgleichung zweiter Ordnung mit konstanten Koeffizienten. Zu ihrer Lösung brauchen wir also ein Fundamentalsystem aus zwei linear unabhängigen Lösungen. Wir sehen sofort das $x_1(t) = \sin \omega_0 t$ und $x_2(t) = \cos \omega_0 t$ die Gleichung lösen. Dass sie linear unabhängig voneinander sind, stellt sich durch Anwendung von (3.17) heraus

$$x_1 \dot{x}_2 - x_2 \dot{x}_1 = -\omega_0 \sin^2 \omega_0 t - \omega_0 \cos^2 \omega_0 t = -\omega_0 \neq 0. \quad (11.6)$$

Die allgemeine Lösung unserer Gleichung (11.5) lautet also

$$x(t) = C_1 \sin \omega_0 t + C_2 \cos \omega_0 t \quad (11.7)$$

Die Konstanten C_1, C_2 müssen durch die beiden Anfangsbedingungen bestimmt werden. Wenn wir den Anfangsort x_0 und die Anfangsgeschwindigkeit v_0 festlegen, folgt

$$\begin{aligned} x(t=0) = x_0 &\Rightarrow C_2 = x_0 \\ \dot{x}(t=0) = v_0 &\Rightarrow C_1 = \frac{v_0}{\omega_0} \end{aligned} \quad (11.8)$$

und damit

$$x(t) = \frac{v_0}{\omega_0} \sin \omega_0 t + x_0 \cos \omega_0 t. \quad (11.9)$$

- Wegen der trigonometrischen Beziehung $\cos(x-y) = \cos x \cos y + \sin x \sin y$ lässt sich die allgemeine Lösung auch in die Form

$$x(t) = A_0 \cos(\omega_0 t - \delta_0) \quad (11.10)$$

bringen. Dazu setzt man $x = \omega_0$ sowie $A_0 \cos y = x_0$ und $A_0 \sin y = v_0/\omega_0$ und erhält

$$A_0 = \sqrt{x_0^2 + \frac{v_0^2}{\omega_0^2}} \quad (11.11)$$

und

$$\tan \delta_0 = \frac{v_0}{x_0 \omega_0}. \quad (11.12)$$

$A_0 \geq 0$ heißt Amplitude, $\omega_0 t - \delta_0$ Phase der Schwingung. Dabei ist $0 \leq \delta_0 < 2\pi$.

- Die durch (11.10) bzw. (11.11) beschriebene Bewegung heißt *harmonische Schwingung*. Sie hat die *Kreisfrequenz* ω_0 , die *Frequenz* ν , und die *Schwingungsperiode* T ,

$$\nu = \frac{\omega_0}{2\pi} \Rightarrow T = \frac{1}{\nu} = \frac{2\pi}{\omega_0}. \quad (11.13)$$

11.1.3 Komplexe Darstellung

- Es lohnt sich zur Beschreibung des harmonischen Oszillators komplexe Zahlen einzuführen, weil sich durch sie vieles enorm vereinfacht. Wir multiplizieren an die Bewegungsgleichung (11.5) $2\ddot{x}$ als integrierendem Faktor,

$$2\dot{z}\ddot{z} + 2\omega_0^2\dot{z}z = 0 \quad \Rightarrow \quad \frac{d(\dot{z}^2)}{dt} = -\omega_0^2 \frac{d(z^2)}{dt} . \quad (11.14)$$

Sie kann integriert werden und ergibt modulo einer Konstanten, die wir zunächst vernachlässigen

$$\dot{z}^2 = -\omega_0^2 z^2 \quad \Leftrightarrow \quad \frac{\dot{z}}{z} = \pm i\omega_0 \quad \Leftrightarrow \quad z = e^{\pm i\omega_0 t} . \quad (11.15)$$

- Diese beiden Lösungen sind linear unabhängig, denn

$$\left| -i\omega_0 e^{i\omega_0 t} e^{-i\omega_0 t} - i\omega_0 e^{-i\omega_0 t} e^{i\omega_0 t} \right| = 2\omega_0 \neq 0 . \quad (11.16)$$

Natürlich hat eine komplexe Lösung an sich noch keinen physikalischen Sinn, sodass man den Real- oder Imaginärteil der komplexen Lösung wählt, sobald man eine physikalische Aussage treffen will.

11.1.4 Hermitesche Matrizen

- Zum Abschluss dieser einführenden Diskussion komplexer Zahlen beschäftigen wir uns noch mit einer komplexen Erweiterung der symmetrischen Matrizen. Sei A eine Matrix mit Elementen $a_{ij} \in \mathbb{C}$. Wenn man alle ihre Elemente komplex konjugiert, heißt die Matrix selbst komplex konjugiert, $A^* = (a_{ij}^*)$. Wenn man sie zusätzlich transponiert, heißt sie adjungiert,

$$A^\dagger = (A^*)^T = (a_{ji}^*) . \quad (11.17)$$

Aus Gründen, die gleich offensichtlich werden, spielen solche Matrizen vor allem in der Quantenmechanik eine tragende Rolle, die gleich ihren Adjungierten sind,

$$A = A^\dagger . \quad (11.18)$$

Sie heißen *selbstadjungiert* oder *hermitesch*.

- Seien nun A eine hermitesche Matrix, λ_i und λ_j irgend zwei ihrer Eigenwerte und v_i und v_j die zugehörigen Eigenvektoren. Dann gilt nach Definition

$$Av_i = \lambda_i v_i , \quad Av_j = \lambda_j v_j . \quad (11.19)$$

Durch Multiplikation der ersten Gleichung mit v_j^\dagger und der zweiten mit v_i^\dagger erhalten wir

$$v_j^\dagger A v_i = \lambda_i \langle v_i, v_j \rangle, \quad v_i^\dagger A v_j = \lambda_j \langle v_i, v_j \rangle, \quad (11.20)$$

denn das Skalarprodukt zweier komplexwertiger Vektoren v und w muss $\langle v, w \rangle = v^\dagger w$ lauten, um zu garantieren, dass $\langle v, v \rangle \geq 0$ ist. Die adjungierte der zweiten Gleichung ist

$$v_j^\dagger A^\dagger v_i = v_j^\dagger A v_i = \lambda_j^* \langle v_i, v_j \rangle, \quad (11.21)$$

wobei wir $A^\dagger = A$ ausgenutzt haben. Subtrahieren wir dieses Ergebnis von der ersten Gleichung (11.20), folgt

$$0 = (\lambda_i - \lambda_j^*) \langle v_i, v_j \rangle. \quad (11.22)$$

Für $i = j$ folgt daraus wegen $\langle v_i, v_i \rangle > 0$, dass $\lambda_i = \lambda_i^*$, die Eigenwerte also *reell* sind. Für $i \neq j$ sieht man dass Eigenvektoren zu verschiedenen Eigenwerten hermitescher Matrizen *orthogonal* sein müssen.

- Schließlich sei noch erwähnt, dass unitäre Matrizen U solche sind, deren Adjungierte gleich ihren Inversen sind,

$$U^\dagger = U^{-1}. \quad (11.23)$$

Sie verallgemeinern die Beziehung $R^T = R^{-1}$ orthogonaler Matrizen und treten ebenfalls in der Quantenmechanik häufig auf.

11.2 Pendel

In den folgenden Beispielen zeigen wir, wie sich auch kompliziertere Schwingungsvorgänge leicht mit Hilfe der komplexen Zahlen lösen lassen. Das bedeutet nicht, dass man nicht auch mit den üblichen trigonometrischen Funktionen zum Erfolg kommt, aber dieser Ansatz wäre wesentlich komplizierter.

11.2.1 Mathematisches Pendel

- Ein Massenpunkt der Masse M sei im homogenen Gravitationsfeld an einem masselosen Faden der Länge l aufgehängt. Das Pendel bewege sich in der x - y -Ebene. Für einen kleinen Auslenkwinkel ϕ aus der Ruhelage zeigt der Faden in die Richtung

$$\begin{aligned} \vec{x} &= l(\sin \phi \vec{e}_x - \cos \phi \vec{e}_y) \approx -l \vec{e}_y \\ \Rightarrow \dot{\vec{x}} &= l \dot{\phi} (\cos \phi \vec{e}_x + \sin \phi \vec{e}_y) \approx l \dot{\phi} \vec{e}_x \\ \Rightarrow \ddot{\vec{x}} &= l \ddot{\phi} (\cos \phi \vec{e}_x + \sin \phi \vec{e}_y) - l \dot{\phi}^2 (\sin \phi \vec{e}_x - \cos \phi \vec{e}_y) \\ &\approx l \ddot{\phi} \vec{e}_x + l \dot{\phi}^2 \vec{e}_y. \end{aligned} \quad (11.24)$$

Die Gravitationskraft zeigt immer nach unten,

$$\vec{F} = -Mg\vec{e}_y, \quad (11.25)$$

wobei g die Schwerebeschleunigung ist. Senkrecht zum Pendel wirkt also die Rückstellkraft

$$\begin{aligned} \vec{F}_r &= \vec{F} - \left(\vec{F} \cdot \frac{\vec{x}}{|\vec{x}|} \right) \frac{\vec{x}}{|\vec{x}|} \\ &= -Mg\vec{e}_y - Mg \cos \phi (\sin \phi \vec{e}_x - \cos \phi \vec{e}_y) \\ &= -\sin^2 \phi Mg\vec{e}_y - Mg \cos \phi \sin \phi \vec{e}_x \approx -Mg\phi \vec{e}_x. \end{aligned} \quad (11.26)$$

Damit ist der führende Term in der Bewegungsgleichung eindimensional,

$$-Mg\phi = Ml\ddot{\phi} \quad \Leftrightarrow \quad \ddot{\phi} + \frac{g}{l}\phi = 0. \quad (11.27)$$

Das ist offensichtlich eine harmonische Oszillatorgleichung für ϕ . Die Schwingungsfrequenz ist unabhängig von der Masse M des Pendelkörpers

$$\omega = \sqrt{\frac{g}{l}}, \quad (11.28)$$

11.2.2 Physikalisches Pendel

- Nun ersetzen wir den Massenpunkt M durch eine homogene Scheibe derselben Masse, der Dicke d und mit Radius R , die in ihrem Schwerpunkt an der Pendelstange fixiert wird. Ändert sich dadurch das Verhalten des Pendels? Wir müssen jetzt berücksichtigen, dass sich der Pendelkörper dreht und sein Drehimpuls sich der Pendelschwingung fortwährend ändert.
- Bestimmen wir zunächst den Trägheitstensor. Dazu führen wir Zylinderkoordinaten so ein, dass die z -Richtung senkrecht zur Scheibe steht und der Ursprung im Scheibenmittelpunkt liegt. Wieder müssen aus Symmetriegründen die beiden Hauptträgheitsmomente Θ_{11} und Θ_{22} gleich sein. Wir brauchen das dritte Hauptträgheitsmoment Θ_{33} , da die Scheibe um die z -Achse gedreht werden wird. Also berechnen wir

$$\begin{aligned} \Theta_{33} &= \int_0^R r dr \int_0^{2\pi} d\phi \int_{-d/2}^{d/2} dz \rho (\vec{x}^2 \delta_{33} - x_3^2) \\ &= \rho \int_0^R r dr \int_0^{2\pi} d\phi \int_{-d/2}^{d/2} dz ((r^2 + z^2) - z^2) \\ &= \rho \int_0^R r^3 dr \int_0^{2\pi} d\phi \int_{-d/2}^{d/2} dz \\ &= \rho \frac{R^4}{4} 2\pi d = \frac{MR^2}{2}, \end{aligned} \quad (11.29)$$

mit der Masse $M = \pi\rho R^2 d$. Da die Drehachse um die Pendellänge l entfernt ist, ist das gesamte Trägheitsmoment die Summe des Pendels

$$\Theta = \frac{MR^2}{2} + Ml^2. \quad (11.30)$$

Die Drehimpulsänderung in die z -Richtung ist dann

$$\dot{L} = \Theta \dot{\omega} = \Theta \ddot{\phi} \vec{e}_z = M \left(\frac{R^2}{2} + l^2 \right) \ddot{\phi} \vec{e}_z. \quad (11.31)$$

- Das Drehmoment der einzigen Kraft, der Gravitationskraft, ist

$$\vec{M} = \vec{F} \times l \frac{\vec{x}}{|\vec{x}|} = -Mg \vec{e}_y \times l \sin \phi \vec{e}_x \approx Mgl \phi \vec{e}_z. \quad (11.32)$$

woraus für die Bewegungsgleichung $\vec{M} = \dot{L}$ in einer Dimension

$$\left(\frac{R^2}{2} + l^2 \right) \ddot{\phi} + gl\phi = 0 \quad (11.33)$$

folgt. Das relative Vorzeichen können wir aus der Bedingung bestimmen, dass unser Gleichgewicht stabil ist. Die Schwingungsfrequenz

$$\omega = \sqrt{\frac{2gl}{R^2 + 2l^2}}. \quad (11.34)$$

hängt nach wie vor nicht von der Pendelmasse ab, wird aber mit zunehmendem Scheibenradius immer kleiner. Für $R \rightarrow 0$ bekommen wir das Ergebnis (11.28) für das mathematische Pendel zurück.

11.2.3 Das Foucaultsche Pendel

- Auf ein Pendel, das auf der Erde schwingt, wirkt die Corioliskraft $\vec{F}_C = 2m\dot{\vec{x}} \times \vec{\Omega}$ als weitere Scheinkraft. Es folgt also aus der Bewegungsgleichung

$$\ddot{\vec{x}} - 2\dot{\vec{x}} \times \vec{\Omega} + \omega^2 \vec{x} = 0, \quad (11.35)$$

wobei $\omega = \sqrt{l/g}$ die übliche Pendelfrequenz und $\vec{\Omega}$ die Winkelgeschwindigkeit der Erde sind. Solange die Corioliskraft klein gegenüber der Rückstellkraft ist können wir die Pendelbewegung als Schwingung in einer sich langsam drehenden Ebene nähern. Wir wählen dafür die x - y -Ebene und führen die z -Achse als lokale

Senkrechte ein. Für kleine Auslenkungen ist also \dot{z} vernachlässigbar klein und

$$\dot{\vec{x}} \times \vec{\Omega} = \begin{pmatrix} \dot{y}\Omega_z - \dot{z}\Omega_y \\ -\dot{x}\Omega_z + \dot{z}\Omega_x \\ \dot{x}\Omega_y - \dot{y}\Omega_x \end{pmatrix} \approx \begin{pmatrix} \dot{y}\Omega_z \\ -\dot{x}\Omega_z \\ \dot{x}\Omega_y - \dot{y}\Omega_x \end{pmatrix}. \quad (11.36)$$

Die Bewegungsgleichungen in der x - y -Ebene lauten daher

$$\begin{aligned} \ddot{x} - 2\dot{y}\Omega_z + \omega^2 x &= 0 \\ \dot{y} + 2\dot{x}\Omega_z + \omega^2 y &= 0. \end{aligned} \quad (11.37)$$

Mit Hilfe der komplexe Größe

$$\xi := x + iy \quad (11.38)$$

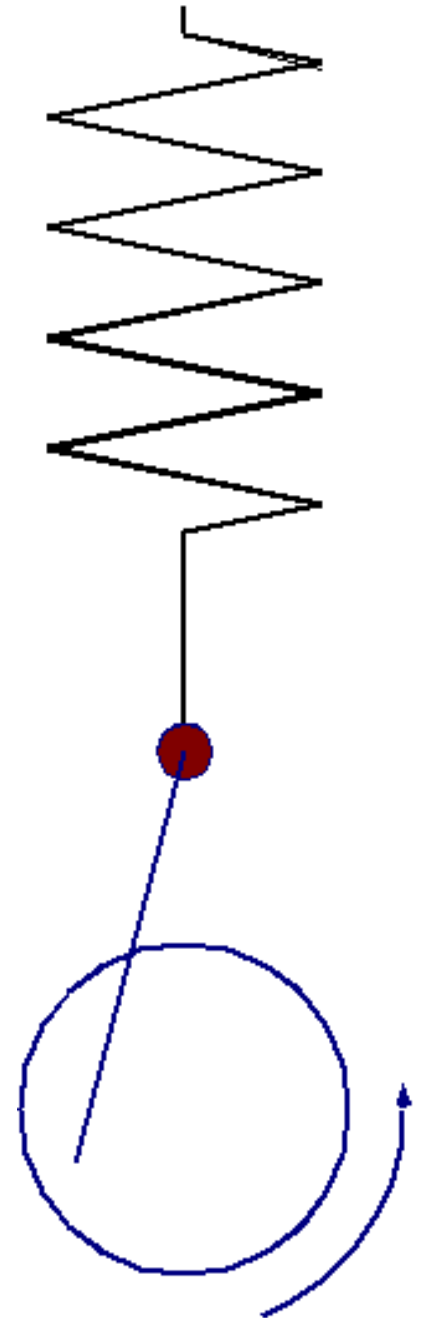
schreiben wir sie als eine Gleichung

$$\ddot{\xi} + 2i\Omega_z \dot{\xi} + \omega^2 \xi = 0. \quad (11.39)$$

Dies ist mathematisch eine Schwingung mit imaginärer Dämpfung $i\Omega_z$. Da die Rotationsperiode der Erde sehr viel größer als die Schwingungsperiode des Pendels ist, ist die Dämpfung schwach, $\Omega_z \ll \omega$. Die Lösung ist dann nach dem Übungsblatt 10

$$\begin{aligned} \xi(t) &= e^{-i\Omega_z t} e^{\pm i\omega t} \\ &= [\cos(\Omega_z t) - i \sin(\Omega_z t)] [\cos(\omega t) \pm i \sin(\omega t)] \\ &= \cos(\Omega_z t) \cos(\omega t) \pm \sin(\Omega_z t) \sin(\omega t) \\ &\quad + i [\pm \cos(\Omega_z t) \sin(\omega t) - \sin(\Omega_z t) \cos(\omega t)] \\ \begin{pmatrix} x(t) \\ y(t) \end{pmatrix} &= \begin{pmatrix} \cos(\Omega_z t) & \sin(\Omega_z t) \\ -\sin(\Omega_z t) & \cos(\Omega_z t) \end{pmatrix} \begin{pmatrix} \cos(\omega t) \\ \pm \sin(\omega t) \end{pmatrix}. \end{aligned} \quad (11.40)$$

Die Pendelebene der üblichen Schwingung dreht sich mit der Frequenz Ω_z .



erzwungene Schwingung: Federpendel mit periodischem Antrieb

11.3 Erzwungene Schwingungen, Resonanz

11.3.1 Periodischem Antrieb

- Auf den Massenpunkt mit einer allgemeinen Bewegungsgleichung wirke in diesem Beispiel von außen eine periodische Kraft $F_e = c \cos(\omega t)$, mit der die Bewegungsgleichung

$$m\ddot{x} + b\dot{x} + kx = c \cos(\omega t). \quad (11.41)$$

wird. Wir lösen die Gleichung wieder in der komplexen Ebene,

$$m\ddot{z} + b\dot{z} + kz = c e^{i\omega t} \quad (11.42)$$

mit $z(t) \in \mathbb{C}$, wobei die physikalische Lösung der Realteil der komplexen Lösung sein wird.

- Wir brauchen jetzt eine spezielle Lösung der inhomogenen Gleichung plus die allgemeine Lösung der homogenen Gleichung. Letztere ist wieder die Lösung des gedämpften harmonischen Oszillators,

$$\begin{aligned} z(t) &= e^{-\lambda t} e^{\pm i\bar{\omega}t} \\ x(t) &= \Re z(t) = e^{-\lambda t} \cos(\pm\bar{\omega}t) \end{aligned} \quad (11.43)$$

mit

$$\lambda = \frac{b}{2m} \quad \omega_0^2 = \frac{k}{m} \quad \bar{\omega} = \sqrt{\omega_0^2 - \lambda^2}, \quad (11.44)$$

falls $\lambda < \omega_0$ ist.

- Die spezielle Lösung gewinnen wir über den Ansatz

$$\begin{aligned} z(t) = z_0 e^{i\omega t} &\Rightarrow (-m\omega^2 + ib\omega + k)z_0 = c \\ &\Leftrightarrow (-\omega^2 + 2i\lambda\omega + \omega_0^2)z_0 = \frac{c}{m}. \end{aligned} \quad (11.45)$$

Wie erhalten

$$z_0 = \frac{c/m}{(\omega_0^2 - \omega^2) + 2i\omega\lambda} = \frac{c/m [(\omega_0^2 - \omega^2) - 2i\omega\lambda]}{(\omega_0^2 - \omega^2)^2 + 4\omega^2\lambda^2}. \quad (11.46)$$

und damit für die Amplitude und Phase von z_0

$$A = \frac{c/m}{\sqrt{(\omega_0^2 - \omega^2)^2 + 4\omega^2\lambda^2}} > 0 \quad \tan \delta = \frac{2\omega\lambda}{\omega_0^2 - \omega^2}. \quad (11.47)$$

Die spezielle Lösung ist also für $0 \leq \delta < \pi$

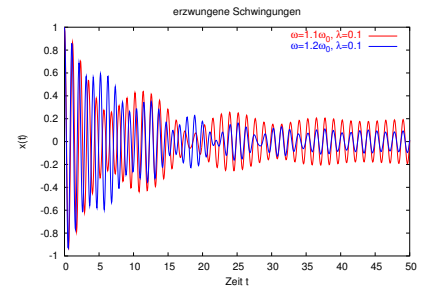
$$\begin{aligned} z(t) &= z_0 e^{i\omega t} =: A e^{-i\delta} e^{i\omega t} = A e^{i(\omega t - \delta)} \\ x(t) &= \Re z(t) = A \cos(\omega t - \delta), \end{aligned} \quad (11.48)$$

und die allgemeine Lösung für die erzwungene Schwingung

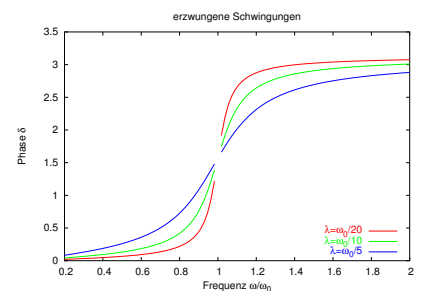
$$x(t) = A \cos(\omega t - \delta) + \bar{A} e^{-\lambda t} \cos(\bar{\omega}t - \bar{\delta}) \quad (11.49)$$

führt. A und δ sind bereits festgelegt, wogegen \bar{A} und $\bar{\delta}$ noch zur Erfüllung der Anfangsbedingungen zur Verfügung stehen.

- Für lange Zeiten $t \gg 1/\lambda$ verschwindet der zweite Term mit seiner exponentiellen Dämpfung. Nach einer *Einschwingung* schwingt der Oszillator dann mit derselben Frequenz wie die äußere Kraft, aber um die Phase δ verschoben.



Erzwungene Schwingungen mit schwacher Dämpfung; deutlich sichtbar sind der Einschwingvorgang und Schwebungen



Phase δ der erzwungenen Schwingung nach dem Einschwingen für verschiedene Werte der Dämpfung λ

11.3.2 Resonanz und Halbwertsbreite

- Die Amplitude A aus (11.47) zeigt das typische *Resonanzverhalten*. Nach der Kettenregel erreicht sie ein Maximum bei

$$\begin{aligned} \frac{d}{d\omega} \left((\omega_0^2 - \omega^2)^2 + 4\omega^2\lambda^2 \right) &= -4(\omega_0^2 - \omega^2)\omega + 8\omega\lambda^2 \\ &= 4\omega \left(-\omega_0^2 + \omega^2 + 2\lambda^2 \right) = 0 \end{aligned} \quad (11.50)$$

$$\text{oder } \omega_{\max} = \sqrt{\omega_0^2 - 2\lambda^2} \approx \omega_0 \left(1 - \frac{\lambda^2}{\omega_0^2} + \mathcal{O}\left(\frac{\lambda^3}{\omega_0^3}\right) \right),$$

wobei wir in der Taylorreihe eine schwache Dämpfung $\lambda \ll \omega_0$ annehmen. Ohne Dämpfung ist die Resonanzfrequenz identisch mit der Frequenz des freien Pendels, bei kleiner Dämpfung beginnen sich die beiden zu unterscheiden. Das Maximum erreicht die Höhe

$$A_{\max} = \frac{c/m}{\sqrt{(2\lambda^2)^2 + 4(\omega_0^2 - 2\lambda^2)\lambda^2}} = \frac{c/m}{2\lambda\sqrt{\omega_0^2 - \lambda^2}} \quad (11.51)$$

erreicht.

- Die Halbwertsbreite der Resonanz ist durch die Bedingung

$$A^2(\omega_{1,2}) = \frac{1}{2}A_{\max}^2 \quad (11.52)$$

definiert. Streng genommen ist dies die Halbwertsbreite für eine Intensität, die quadratisch von der Amplitude abhängt. Sie führt zur Bedingung

$$\begin{aligned} A^2(\omega_{1,2}) &= \frac{c^2/m^2}{(\omega_0^2 - \omega_{1,2}^2)^2 + 4\omega_{1,2}^2\lambda^2} = \frac{c^2/m^2}{8\lambda^2(\omega_0^2 - \lambda^2)} \\ \Leftrightarrow \omega_0^4 - 2\omega_0^2\omega_{1,2}^2 + \omega_{1,2}^4 + 4\omega_{1,2}^2\lambda^2 &= 8\lambda^2\omega_0^2 - 8\lambda^4 \\ \Leftrightarrow \omega_{1,2}^4 - 2\omega_{1,2}^2\omega_{\max}^2 + \omega_0^4 - 8\lambda^2\omega_0^2 + 8\lambda^4 &= 0, \end{aligned} \quad (11.53)$$

mit der Lösung

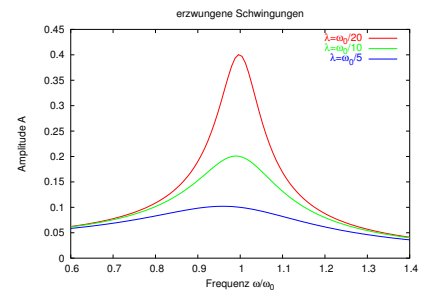
$$\begin{aligned} \omega_{1,2}^2 &= \omega_{\max}^2 \pm \sqrt{\omega_{\max}^4 - \omega_0^4 + 8\lambda^2\omega_0^2 - 8\lambda^4} \\ &= \omega_{\max}^2 \pm \sqrt{6\lambda^2\omega_0^2 - 4\lambda^4} \\ &\approx \omega_0^2 \left(1 \pm \frac{\sqrt{6}\lambda}{\omega_0} \right) \left(1 + \mathcal{O}\left(\frac{\lambda^2}{\omega_0^2}\right) \right) \end{aligned} \quad (11.54)$$

oder

$$\omega_{1,2} \approx \omega_0 \left(1 \pm \frac{\sqrt{3}\lambda}{\sqrt{2}\omega_0} \right) = \omega_0 \pm \frac{\sqrt{3}\lambda}{\sqrt{2}}. \quad (11.55)$$

Die komplette Halbwertsbreite für schwache Dämpfung ist damit

$$\Gamma = \sqrt{6}\lambda. \quad (11.56)$$



Resonanzverhalten der Amplitude erzwungener Schwingungen

11.3.3 Grenzfall schwacher Dämpfung

- Um den Fall schwacher Dämpfung zu betrachten schreiben wir die allgemeine Lösung (11.49) in der Form

$$x(t) = A \cos(\omega t - \delta) + e^{-\lambda t}(C_1 \cos \bar{\omega} t + C_2 \sin \bar{\omega} t). \quad (11.57)$$

Die Anfangsbedingungen legen die beiden Konstanten fest,

$$\begin{aligned} x(t=0) &= A \cos \delta + C_1 = x_0 \\ &\Rightarrow C_1 = x_0 - A \cos \delta \\ \dot{x}(t=0) &= A\omega \sin \delta - \lambda C_1 + C_2 \bar{\omega} = v_0 \\ &\Rightarrow C_2 = \frac{1}{\bar{\omega}}(v_0 + \lambda x_0 - A\lambda \cos \delta - A\omega \sin \delta). \end{aligned}$$

Im Grenzfall $\lambda \rightarrow 0$ gilt nach (11.47) $\delta \rightarrow 0$, und damit

$$C_1 \rightarrow x_0 - A \quad C_2 \rightarrow \frac{v_0}{\omega_0} \quad A \rightarrow \frac{c/m}{\omega_0^2 - \omega^2}. \quad (11.58)$$

Die allgemeine Lösung nimmt dann die Gestalt

$$x(t) = \frac{c/m}{\omega_0^2 - \omega^2}(\cos \omega t - \cos \omega_0 t) + x_0 \cos \bar{\omega} t + \frac{v_0}{\omega_0} \sin \omega_0 t \quad (11.59)$$

an.

- Für die Resonanzbedingung $\omega \rightarrow \omega_0$ geht der erste Term gegen den Grenzwert

$$\lim_{\omega \rightarrow \omega_0} \frac{\cos \omega t - \cos \omega_0 t}{\omega_0^2 - \omega^2} = \lim_{\omega \rightarrow \omega_0} \frac{t \sin \omega t}{2\omega} = \frac{t \sin \omega_0 t}{2\omega_0}, \quad (11.60)$$

was sich anhand der Regeln von l'Hospital feststellen lässt. Im Spezialfall $x_0 = 0$ erhält man dann die Lösung

$$x(t) = \left(\frac{v_0}{\omega_0} + \frac{ct}{2m\omega_0} \right) \sin \omega_0 t. \quad (11.61)$$

Der zweite Term in Klammern wächst linear mit der Zeit, und man spricht von einem *säkularen Anwachsen* der Amplitude.

11.3.4 Allgemeine Lösung

- Für harmonisch getriebene Schwingungen konnten wir die spezielle Lösung der inhomogenen Gleichung durch den Ansatz (11.45) erraten. Es gibt aber ein allgemeines Konstruktionsverfahren. Wir gehen dabei von der ungedämpften Gleichung

$$\ddot{x} + \omega_0^2 x = \frac{F(t)}{m} \quad (11.62)$$

aus und schreiben sie in der Form

$$\frac{d}{dt} (\dot{x} + i\omega_0 x) - i\omega_0 (\dot{x} + i\omega_0 x) = \frac{F(t)}{m} . \quad (11.63)$$

Mit der komplexen Funktion

$$\xi(t) = \dot{x}(t) + i\omega_0 x(t) \quad (11.64)$$

reduziert sich die Bewegungsgleichung auf eine Differentialgleichung 1. Ordnung reduziert,

$$\dot{\xi} - i\omega_0 \xi = \frac{F(t)}{m} . \quad (11.65)$$

- Die dazugehörige homogene Gleichung $\dot{\xi} = i\omega_0 \xi$ hat die Lösung $\xi(t) = Ae^{i\omega_0 t}$. Die Lösung der inhomogenen Gleichung erhalten wir durch Variation der Konstanten $A \rightarrow A(t)$,

$$\begin{aligned} (\dot{A} + iA\omega_0) e^{i\omega_0 t} - i\omega_0 A e^{i\omega_0 t} &= \dot{A} e^{i\omega_0 t} = \frac{F(t)}{m} \\ \Rightarrow A(t) &= \int_0^t \frac{F(t')}{m} e^{-i\omega_0 t'} dt' . \end{aligned} \quad (11.66)$$

Die allgemeine Lösung für beliebige äußere Kraft ist dann

$$\xi(t) = \xi_0 + e^{i\omega_0 t} \left(\int_0^t \frac{F(t')}{m} e^{-i\omega_0 t'} dt' \right) . \quad (11.67)$$

Da $x(t)$ reell sein muss, folgt aus (11.64), dass $x(t) = \Im(\xi(t))/\omega_0$.

11.3.5 Beispiel: Schwingung nach einem Kraftstoß

- Als Beispiel betrachten wir eine Kraft, die zwischen den Zeiten $t = 0$ und $t = \tau$ den konstanten Betrag F_0 hat und sonst verschwindet. Nach (11.67) ist dann

$$\xi(t) = \xi_0 + e^{i\omega t} \begin{cases} 0 & t < 0 \\ \frac{iF_0}{m\omega} (e^{-i\omega t} - 1) & t = 0 \dots \tau \\ \frac{iF_0}{m\omega} (e^{-i\omega \tau} - 1) & t > \tau \end{cases} . \quad (11.68)$$

- Wenn wir verlangen, dass der Oszillator anfangs in Ruhe ist, also $\dot{x} = 0 = x$ bis $t = 0$, muss $\xi_0 = 0$ sein. Dann ist

$$\xi(t) = \begin{cases} 0 & t < 0 \\ \frac{iF_0}{m\omega} (1 - e^{i\omega t}) & t = 0 \dots \tau \\ \frac{iF_0}{m\omega} (e^{i\omega(t-\tau)} - e^{i\omega t}) & t \geq \tau \end{cases} , \quad (11.69)$$

woraus wir

$$x(t) = \frac{\Im \xi(t)}{\omega} = \begin{cases} 0 & t < 0 \\ \frac{F_0}{m\omega^2} (1 - \cos \omega t) & t = 0 \dots \tau \\ \frac{F_0}{m\omega^2} (\cos \omega(t - \tau) - \cos \omega t) & t > \tau \end{cases} \quad (11.70)$$

erhalten.