

Theoretische Physik I
Punktmechanik und mathematische
Methoden

Version WS 2023/2024 — Tilman Plehn
Original von Matthias Bartelmann

14. Mai 2024

Inhaltsverzeichnis

1	Statistik	1
1.1	Wahrscheinlichkeitsverteilungen	1
1.1.1	Zählexperimente	1
1.1.2	Unsicherheiten	4
1.1.3	Daten vs Theorie	6
1.2	Satz von Bayes	7
1.2.1	Bayesische Wahrscheinlichkeit	7
1.2.2	Volumeneffekte und Profile-Likelihood	10
1.3	Taylor-Reihen	13
1.3.1	Taylorscher Satz	13
1.3.2	Einfache Beispiele	14
2	Die Newtonschen Axiome	16
2.1	Einführung	16
2.2	Newtonsche Axiome	18
2.3	Differentialgleichungen I	19
2.3.1	Definition und Klassifikation	19
2.3.2	Lösung: Trennung der Veränderlichen	20
2.3.3	Lösung: Variation der Konstanten	21
2.4	Fallbewegungen	23
2.4.1	Freier Fall aus geringer Höhe	23
2.4.2	Fall aus geringer Höhe mit Stokesscher Reibung	24
2.4.3	Fall aus geringer Höhe mit Luftwiderstand . . .	26

2.4.4	Freier Fall aus großer Höhe	28
3	Mathematische Grundlagen	30
3.1	Differentialgleichungen II	30
3.1.1	Existenz und Eindeutigkeit von Lösungen	30
3.1.2	Lineare Differentialgleichungen n -ter Ordnung	32
3.2	Vektoren I	34
3.2.1	Vektorräume	35
3.2.2	Lineare Abhängigkeit, Basis und Dimension	37
3.2.3	Beispiele	37
4	Impuls, Drehimpuls und Energie	39
4.1	Vektoren II	39
4.1.1	Das Skalarprodukt	39
4.1.2	Summenkonvention, Levi-Civita-Symbol	40
4.1.3	Das Vektorprodukt	41
4.1.4	Koordinaten und Koordinatentransformationen	42
4.1.5	Inertialsysteme	43
4.2	Impuls, Drehimpuls und Energie	44
4.2.1	Impuls	44
4.2.2	Drehmoment und Drehimpuls	45
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 Gradient	53

5.2.2	Divergenz und Rotation	56
5.3	Energiesatz in drei Dimensionen	56
5.3.1	Energieerhaltung bei Potentialkräften	57
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	61
5.5	Krummlinig-orthogonale Koordinaten	62
5.5.1	Zylinderkoordinaten	63
5.5.2	Sphärische Polarkoordinaten	64
6	Stöße und Streuung	66
6.1	Systeme von Massenpunkten	66
6.1.1	Impuls	66
6.1.2	Drehimpuls	67
6.1.3	Energie	68
6.1.4	Reduzierte Masse	69
6.2	Stöße und Streuung	69
6.2.1	Elastischer Stoß im Schwerpunktsystem	69
6.2.2	Laborsystem	70
6.2.3	Energieübertrag	71
7	Koordinatentransformationen	73
7.1	Darstellung durch Matrizen	73
7.1.1	Drehungen im dreidimensionalen Raum	73
7.1.2	Matrizen und Matrixoperationen	74
7.1.3	Beispiele	75
7.2	Determinanten und Matrixinversion	76
7.2.1	Determinanten	76
7.2.2	Orthonormale Transformationen	77
7.2.3	Inversion von Matrizen	78

7.3	Transformation des Drehimpulses	80
8	Beschleunigte Bezugssysteme	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	Die Euler-Winkel	89
9.2	Tensoren	92
9.2.1	Der Trägheitstensor	92
9.2.2	Tensoren als lineare Abbildungen	92
9.2.3	Transformationseigenschaften	93
9.3	Diagonalisierung	94
9.3.1	Diagonalisierung von Matrizen	94
9.3.2	Diagonalisierung des Trägheitstensors	96
10	Trägheitstensor	98
10.1	Volumenintegrale	98
10.1.1	Motivation	98
10.1.2	Jacobi-Determinante	99
10.1.3	Trägheitstensor einer Kugel	101
10.2	Drehimpuls und Trägheitsmomente	101
10.2.1	Drehimpuls	101
10.2.2	Trägheitsmomente	102
10.3	Der Satz von Gauß	103
10.3.1	Herleitung	103
10.3.2	Transformation von Divergenz und Rotation	104

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. Statistik ist keiner der klassischen Aspekte von Theorie 1, aber sie ist sehr mathematisch und irgendwo müssen Sie diese Grundlagen ja lernen.

1.1 Wahrscheinlichkeitsverteilungen

1.1.1 Zählexperimente

- Die einfachsten Messungen in der Physik sind Zählexperimente, also die Analyse von einer Anzahl von Signalen in irgendeinem Detektor. Beispiele sind Ausschläge eines Pendels, radioaktive Zerfälle, oder LHC-Events mit zwei Myonen. Wir beginnen mit einem Würfelexperiment, dessen zentrale Annahme ist, dass alle Würfelversuche gleichartig und voneinander unabhängig sind. Wir nehmen an, dass wir N Würfe haben, mit einer konstanten Wahrscheinlichkeit \bar{n}/N für ein bestimmtes Ergebnis. Diese Art von Messung heißt auch Bernoulli-Prozess. In diesem Fall ist die

Wahrscheinlichkeit, n Ereignisse in N Versuchen zu beobachten

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

Der erste Term rührt daher, dass wir die n Ereignisse auf N Versuche aufteilen können. Diese Verteilung heisst *Binomialverteilung*. Die wohl wichtigste mathematische Eigenschaft von Wahrscheinlichkeitsverteilungen ist die Normierung, also die Bedingung dass ein System einen der möglichen Messwerte annehmen muss

$$\begin{aligned} \sum_{n=0}^N P(n|\bar{n}) &= \sum_{n=0}^N \binom{N}{n} \left(\frac{\bar{n}}{N}\right)^n \left(1 - \frac{\bar{n}}{N}\right)^{N-n} \\ &= \left[\frac{\bar{n}}{N} + \left(1 - \frac{\bar{n}}{N}\right)\right]^N = 1^N = 1. \end{aligned} \quad (1.2)$$

Die zweite Zeile ist einfach die binomische Formel.

- In der Physik gehen wir davon aus, dass wir Präzisionsexperimente sehr oft wiederholen können, wobei positive Ergebnisse selten sind, also

$$N \gg n, \bar{n}. \quad (1.3)$$

In diesem Grenzfall kann man die Binomialverteilung umformen und erhält

$$\begin{aligned} P(n|\bar{n}) &= \frac{\bar{n}^n}{n!} \lim_{N \rightarrow \infty} \frac{N!}{(N-n)!} \frac{1}{N^n} \left(1 - \frac{\bar{n}}{N}\right)^{N-n} \\ &= \frac{\bar{n}^n}{n!} \lim_{N \rightarrow \infty} \frac{(N-n+1) \cdots N}{N^n} \left(1 - \frac{\bar{n}}{N}\right)^{-n} \left(1 - \frac{\bar{n}}{N}\right)^N \\ &= \frac{\bar{n}^n}{n!} \lim_{N \rightarrow \infty} \frac{N^n}{N^n} 1^{-n} \left(1 - \frac{\bar{n}}{N}\right)^N \\ &= \frac{\bar{n}^n}{n!} \lim_{N \rightarrow \infty} \left(1 - \frac{\bar{n}}{N}\right)^N \quad (N \gg n) \\ &= \frac{\bar{n}^n}{n!} \lim_{N \rightarrow \infty} \exp \left[\ln \left(1 - \frac{\bar{n}}{N}\right)^N \right] \\ &= \frac{\bar{n}^n}{n!} \lim_{N \rightarrow \infty} \exp \left[N \ln \left(1 - \frac{\bar{n}}{N}\right) \right] \\ &= \frac{\bar{n}^n}{n!} \lim_{N \rightarrow \infty} \exp \left[N \left(-\frac{\bar{n}}{N}\right) \right] \quad (N \gg \bar{n}) \\ &= \frac{\bar{n}^n}{n!} e^{-\bar{n}} \end{aligned} \quad (1.4)$$

Dies ist die *Poisson-Verteilung* für die Wahrscheinlichkeit, einen Wert n zu messen, wenn wir \bar{n} erwarten. Im vorletzten Schritt haben wir eine Näherungsformel für den Logarithmus benutzt, die wir im Anschluss an das Statistik-Kapitel herleiten werden, nämlich

$$\ln(1 + \delta) \approx \delta - \frac{\delta^2}{2} + \cdots \quad \text{falls } |\delta| \ll 1. \quad (1.5)$$

Die Poisson-Verteilung ist zunächst für alle natürlichen Zahlen $n \geq 0$ definiert. Negative Zahlen ergeben in der Poisson-Verteilung ebenso wenig Sinn wie in einem Zählexperiment. Auch hier stimmt die Normierung der Wahrscheinlichkeitsverteilung

$$\sum_{n=0}^{\infty} P(n|\bar{n}) = e^{-\bar{n}} \sum_{n=0}^{\infty} \frac{\bar{n}^n}{n!} = e^{-\bar{n}} e^{\bar{n}} = 1 \quad (1.6)$$

Dazu haben wir lediglich die Reihendarstellung der Exponentialfunktion benötigt.

- In der Praxis ist auch die Definition der Poisson-Verteilung unhandlich. Im Grenzwert einer hinreichend großen Zahl von positiven Messungen $n \gg 1$ können wir zunächst die Stirling-Formel

$$n! \approx \sqrt{2\pi n} \left(\frac{n}{e}\right)^n \quad (1.7)$$

benutzen und die Poisson-Verteilung auf positive reelle Zahlen verallgemeinern,

$$P(n|\bar{n}) = \frac{\bar{n}^n}{\sqrt{2\pi n}} \left(\frac{e}{n}\right)^n e^{-\bar{n}}. \quad (1.8)$$

Wenn wir als weitere Forderung an (1.3) annehmen dass

$$N \gg n, \bar{n} \gg 1, \quad (1.9)$$

wir also Zählraten nahe Null vermeiden und unsere Variablen als

$$n = \bar{n}(1 + \delta) \quad \Leftrightarrow \quad \delta = \frac{n - \bar{n}}{\bar{n}} \quad (1.10)$$

umschreiben, dann ergibt sich für die Poisson-Verteilung

$$\begin{aligned} P(n|\bar{n}) &= \frac{\bar{n}^{\bar{n}(1+\delta)}}{\sqrt{2\pi \bar{n}(1+\delta)}} \left(\frac{e}{\bar{n}(1+\delta)}\right)^{\bar{n}(1+\delta)} e^{-\bar{n}} \\ &= \frac{1}{\sqrt{2\pi \bar{n}}} \frac{\bar{n}^{\bar{n}(1+\delta)}}{\sqrt{1+\delta}} \frac{e^{\bar{n}\delta}}{\bar{n}^{\bar{n}(1+\delta)}(1+\delta)^{\bar{n}(1+\delta)}} \\ &= \frac{1}{\sqrt{2\pi \bar{n}}} \frac{e^{\bar{n}\delta}}{(1+\delta)^{\bar{n}(1+\delta)+1/2}}. \end{aligned} \quad (1.11)$$

Im Nenner können wir den letzten Term im Exponenten sicherlich vernachlässigen und erhalten mit demselben Trick in (1.5), aber

jetzt mit dem quadratischen Term

$$\begin{aligned}
P(n|\bar{n}) &\approx \frac{1}{\sqrt{2\pi\bar{n}}} \frac{e^{\bar{n}\delta}}{(1+\delta)^{\bar{n}(1+\delta)}} \\
&= \frac{e^{\bar{n}\delta}}{\sqrt{2\pi\bar{n}}} \exp\left[\ln(1+\delta)^{-\bar{n}(1+\delta)}\right] \\
&= \frac{e^{\bar{n}\delta}}{\sqrt{2\pi\bar{n}}} \exp\left[(-\bar{n}(1+\delta))\ln(1+\delta)\right] \\
&= \frac{e^{\bar{n}\delta}}{\sqrt{2\pi\bar{n}}} \exp\left[(-\bar{n}-\bar{n}\delta)\left(\delta - \frac{\delta^2}{2} + \dots\right)\right] \\
&\approx \frac{e^{\bar{n}\delta}}{\sqrt{2\pi\bar{n}}} \exp\left[-\bar{n}\delta - \bar{n}\delta^2 + \frac{\bar{n}\delta^2}{2}\right] \\
&= \frac{1}{\sqrt{2\pi\bar{n}}} \exp\left[-\frac{\bar{n}\delta^2}{2}\right] \\
&= \frac{1}{\sqrt{2\pi\bar{n}}} e^{-(n-\bar{n})^2/(2\bar{n})} \tag{1.12}
\end{aligned}$$

Im letzten Schritt erhalten wir die *Gauss-Verteilung* oder Normalverteilung. Bei der Anwendung von (1.3) auf (1.10) waren wir ein bisschen schlampig, weil wir nie wirklich $|n - \bar{n}| \ll \bar{n}$ gezeigt haben. Im Nachhinein können wir das aber überprüfen. Die Breite der Gauss-Verteilung ist nämlich gegeben durch

$$\frac{(n - \bar{n})^2}{2\bar{n}} \approx \frac{1}{2} \quad \Leftrightarrow \quad |n - \bar{n}| \approx \sqrt{\bar{n}} \ll \bar{n} . \tag{1.13}$$

Dasselbe gilt für die Poisson-Verteilung, man kann es nur nicht so schön ablesen. Die Gauss-Verteilung ist für alle reellen Werte von n definiert, und auch hier stimmt die Normierung

$$\frac{1}{\sqrt{2\pi\bar{n}}} \int_{-\infty}^{\infty} dn e^{-(n-\bar{n})^2/(2\bar{n})} = 1 . \tag{1.14}$$

1.1.2 Unsicherheiten

- Wir haben drei verschiedene Formeln für die Wahrscheinlichkeit, n Signale in N Versuchen, mit einer individuellen Wahrscheinlichkeit \bar{n}/N zu beobachten. Alle drei Verteilung beschreiben die möglichen Ausgänge eines Zähl-experiments, mit einem bevorzugten Ergebnis \bar{n} , aber endlichen Wahrscheinlichkeiten für einen ganzen Bereich in n . Zum Beispiel die Gauss-Verteilung ist symmetrisch um den Erwartungswert \bar{n} , mit einer Breite oder Standardabweichung $\sigma = \sqrt{\bar{n}}$. Wir interpretieren eine solche Breite als Unsicherheit unseres Zähl-experiments, einen Beitrag zum Fehlerbalken einer physikalische Messung. Diese sogenannte *statistische Unsicherheit* hat die wichtige Eigenschaft, dass sie für

hohe Statistik nach der Annahme von (1.9) verschwindet,

$$\lim_{\bar{n} \rightarrow \infty} \frac{\sigma}{\bar{n}} = \lim_{\bar{n} \rightarrow \infty} \frac{1}{\sqrt{\bar{n}}} = 0. \quad (1.15)$$

- 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. Offensichtlich verschwinden solche systematischen Fehler nicht, wenn wir sie Signal-Messung oft wiederholen, die Eichung des Versuchsaufbaus oder Hilfsmessungen aber weiterhin nutzen.

Nur wenn Systematiken durch anderen Messungen bestimmt werden, kann man für sie auch eine Gauss-Verteilung annehmen. Im Idealfall kann man zum Beispiel einen Input-Parameter mit seiner gesamten gemessenen Wahrscheinlichkeitsverteilung angeben. Ein Beispiel ist die Messung einer Rate, die aus einem Signal und einem Untergrund besteht,

$$n = b + s \quad \text{und} \quad \bar{n} = \bar{b} + \bar{s}, \quad (1.16)$$

wobei uns nur das Signal s interessiert und wir den Untergrund b aus einer Kontrollmessung erhalten. Wenn wir uns für einen Parameter wie s interessieren, dann nennen wir ihn parameter of interest, wenn ein Parameter wie b lediglich intern in einer Analyse auftaucht, dann nennen wir ihn nuisance parameter.

Nur als Nebenbemerkung — die Annahme, dass sich Untergrund und Signal einfach addieren ist klassisch und in der Quantenphysik nicht richtig, ein Aspekt der auch Selim Jochim in der Experimentalphysik großen Spaß bereitet.

- 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 Näherungs-Formel oder statistische Unsicherheiten in 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.

1.1.3 Daten vs Theorie

Nachdem wir Wahrscheinlichkeitsverteilung von Signalaraten berechnen können 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 $P(n)$ gemessen haben.

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

$$\int_{s+b}^{\infty} dn P(n) < 0.05 \quad \text{mit} \quad \int_{-\infty}^{\infty} dn P(n) = 1 \quad (1.17)$$

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 $s + b$ 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 Theorie ohne dieses neue Teilchen nicht mit der Messung übereinstimmt. Wir betrachten also die *hypothetische Wahrscheinlichkeitsverteilung* im Untergrundmodell $P(b)$ wie in (1.16) und berechnen, wie wahrscheinlich eine gemessene und zu große Ereigniszahl $n > b$ im Rahmen des Untergrundmodells ist, also

$$p_0 = \int_n^{\infty} db P(b) . \quad (1.18)$$

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 ebensowenig eine

Rolle wie die Breite der gemessenen Verteilung keine Rolle, es ist lediglich die Vorhersage b 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 eine statistische Fluktuation zu beobachten. In erster Näherung ist diese kombinierte Wahrscheinlichkeit einfach 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.
- Streng genommen sind diese Definitionen von Ausschluss und Entdeckung in (1.17) und (1.18) in einem Frequentist-Zugang definiert, denn die zugrundeliegenden Verteilungen $P(n)$ beschreiben das statistisch reproduzierbare Verhalten einer großen Zahl von Experimenten, entweder als gemessene Verteilung $P(n)$ für den Ausschluss von Modellen oder als erwartete Untergrund-Verteilung zur Etablierung einer besseren Theorie.

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

$$P(A) \in \mathbf{R} \quad \text{und} \quad P(A) \geq 0 . \quad (1.19)$$

Das zweite Axiom besagt, dass die Wahrscheinlichkeit aller möglichen Ausgänge kombiniert eins ist,

$$\sum_i P(A_i) = 1 . \quad (1.20)$$

Das ist die Normierungsbedingung, die wir schon mehrfach überprüft hatten. Das dritte Axiom besagt, dass sich Wahrscheinlichkeiten für disjunkte Ausgänge addieren,

$$P(A \cup B) = P(A) + P(B) . \quad (1.21)$$

- In der disjunkten Wahrscheinlichkeit verknüpfen wir zwei Messungen logisch mit einem ‘oder’. Alternativ können wir auch die die Frage stellen, was die Wahrscheinlichkeit für zwei Messungen A und B ist. Zwei Messungen A und B sind unabhängig, wenn für diese *gemeinsame Wahrscheinlichkeit*

$$P(A \cap B) = P(A)P(B) \quad (1.22)$$

gilt. Diese Relation sollte man in der Physik nur annehmen, nachdem man sie sorgfältig getestet hat.

- Da zwei Messungen im Allgemeinen nicht unabhängig sind, definieren 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

$$P(A|B) := \frac{P(A \cap B)}{P(B)} \quad \Leftrightarrow \quad P(A|B) P(B) = P(A \cap B) . \quad (1.23)$$

Wegen $A \cap B = B \cap A$ gilt für die umgekehrte bedingte Wahrscheinlichkeit

$$P(A \cap B) = P(B \cap A) \quad \Rightarrow \quad P(B|A) = \frac{P(A \cap B)}{P(A)} . \quad (1.24)$$

Diese Formeln sind je nach Geschmack Definitionen oder Axiome, die allerdings aus den Kolmogorov-Axiomen motiviert sind.

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

$$P(T|M) = \frac{P(M|T) P(T)}{P(M)} \quad (1.25)$$

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; 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. Den Term $P(M)$ nennen wir Evidenz, sie wird typischerweise durch die Normierungsbedingung im Raum der Theorie-Parameter ersetzt,

$$1 = \int dT P(T|M) = \frac{1}{P(M)} \int dT P(M|T) P(T)$$

$$\Leftrightarrow P(M) = \int dT P(M|T) P(T). \quad (1.26)$$

Weiterhin heisst $P(T)$ *prior probability* oder kurz Prior. In (1.26) sehen wir, dass der Prior bei der Integration über Theorie-Parameter eine Art Integrationsmass definiert.

- Auf der linken Seite von (1.25) steht die Größe, die uns in der Physik interessiert: Wie groß ist nach einem Experiment M die Wahrscheinlichkeit, dass eine Theorie T die Messung korrekt beschreibt. Der Satz von Bayes gibt uns 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 simulieren die Messung für alle möglichen Theorie-Parameter T und erhalten eine Vorhersage für M . Das Problem ist der Prior $P(T)$, das Maß für Wahrscheinlichkeit im Theorie-Raum. Aus dem Beweis des Satzes von Bayes ist klar, dass wir einen Prior nicht vermeiden können, wenn wir uns für die bedingte Wahrscheinlichkeit $P(T|M)$ interessieren.
- Der Satz von Bayes sagt uns auch, wie wir zwei unabhängige Messungen miteinander kombinieren. Für die kombinierte Wahrscheinlichkeit machen wir erst eine Messung, berechnen deren Wahrscheinlichkeitsverteilung, nutzen diese als Prior für die zweite Messung, und berechnen dann die kombinierte Wahrscheinlichkeit. Das ist logisch klar, aber technisch kompliziert. Für die Likelihood ist diese Kombination einfacher. Nehmen wir an, in unserem Zählexperiment sind nicht alle Ereignisse gleich, so dass eine Funktion $p(x)$ für jedes Ereignis zum Beispiel die Wahrscheinlichkeit beschreibt, dass das Ereignis aus einem Signal und nicht aus dem Untergrund kommt. In dem Fall ist die Likelihood einer Menge von Einzelmessungen eine Kombination von Zählstatistik und der Wahrscheinlichkeit der einzelnen Ereignisse, also

$$P(\{x_i\}|T) = \frac{\bar{s}^n}{n!} e^{-\bar{s}} \prod_{i=1}^n p(x_i) = \frac{1}{n!} e^{-\bar{s}} \prod_{i=1}^n (\bar{s} p(x_i)). \quad (1.27)$$

Diese Form hat den Vorteil, dass wir zwei Likelihoods für zwei unabhängige Datensätze einfach kombinieren können

$$P(\{x_i, x_j\}|T) = \frac{1}{n_1!n_2!} e^{-\bar{s}_1 - \bar{s}_2} \prod_{i=1}^{n_1+n_2} (\bar{s}p(x_i)) . \quad (1.28)$$

Wenn wir hier noch die Likelihood durch das Verhältnis zweier Likelihoods für verschiedene Theorie-Hypothesen ersetzen, dann fällt auch noch die Komplikation im Vorfaktor weg. Likelihood-Verhältnisse sind außerdem mathematisch die optimale Methode um anhand von Daten zwei Hypothesen zu unterscheiden.

1.2.2 Volumeneffekte und Profile-Likelihood

- Der Satz von Bayes (1.25) lässt uns zwei Möglichkeiten, Theorie und Messungen mathematisch korrekt zu verknüpfen. Im Bayes-Zugang definieren wir die $P(T|M)$ als den *degree of belief*. Dazu müssen wir einen Prior im Wahrscheinlichkeitsraum der Theorie zu definieren. Der Frequentist-Zugang erlaubt diese Annahme nicht, also muss uns die Likelihood $P(M|T)$ 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 mit der größeren Wahrscheinlichkeit befindet, also offensichtlich die Wolke. Ein Frequentist fragt hingegen, an welchem Ort sich ein Wassermolekül am wahrscheinlichsten befindet. Das ist 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 Integrationsmaß stellt sich auch in Zusammenhang mit der Signalmessung in Gegenwart eines Untergrundes, (1.16). Nehmen wir der Einfachheit halber an, dass unser Untergrund b um $\bar{b} = 0$ fluktuiert und durch eine Gauss-Verteilung der Breite σ_b beschrieben ist. Die Frage ist nun, wie

sich die Verteilung des Signals s und die Verteilung der beobachteten Kombination $n = s + b$ zueinander verhalten. Wenn auch die Messung von n eine Gauss-Form hat, dann ergibt sich für das Signal s zunächst ohne Berücksichtigung der Normierung

$$\begin{aligned}
 P(s) &\propto \int_{-\infty}^{\infty} db e^{-b^2/(2\sigma_b^2)} e^{-n^2/(2\sigma_n^2)} \\
 &= \int_{-\infty}^{\infty} db \exp\left[-\frac{b^2}{2\sigma_b^2}\right] \exp\left[-\frac{(s+b)^2}{2\sigma_n^2}\right] \\
 &= \int_{-\infty}^{\infty} db \exp\left[-\frac{b^2(\sigma_n^2 + \sigma_b^2) + 2sb\sigma_b^2 + s^2\sigma_b^2}{2\sigma_b^2\sigma_n^2}\right] \\
 &= \int_{-\infty}^{\infty} db \exp\left[-\frac{\sigma_b^2 + \sigma_n^2}{2\sigma_b^2\sigma_n^2} \left(b^2 + 2sb\frac{\sigma_b^2}{\sigma_b^2 + \sigma_n^2} + s^2\frac{\sigma_b^2}{\sigma_b^2 + \sigma_n^2}\right)\right] \\
 &= \int_{-\infty}^{\infty} db \exp\left[-\frac{\sigma_s^2}{2\sigma_b^2\sigma_n^2} \left(\left(b + s\frac{\sigma_b^2}{\sigma_s^2}\right)^2 - s^2\frac{\sigma_b^4}{\sigma_s^4} + s^2\frac{\sigma_b^2}{\sigma_s^2}\right)\right],
 \end{aligned} \tag{1.29}$$

mit

$$\sigma_s^2 := \sigma_b^2 + \sigma_n^2. \tag{1.30}$$

An dieser Stelle verschieben wir $b \rightarrow b' = b + s\sigma_b^2/\sigma_s^2$,

$$\begin{aligned}
 P(s) &\propto \int_{-\infty}^{\infty} db' \exp\left[-\frac{\sigma_s^2}{2\sigma_b^2\sigma_n^2} \left(b'^2 - s^2\frac{\sigma_b^4}{\sigma_s^4} + s^2\frac{\sigma_b^2}{\sigma_s^2}\right)\right] \\
 &= \int_{-\infty}^{\infty} db' \exp\left[-\frac{\sigma_s^2}{2\sigma_b^2\sigma_n^2} \left(b'^2 + s^2\frac{\sigma_b^2\sigma_s^2 - \sigma_b^4}{\sigma_s^4}\right)\right] \\
 &= \int_{-\infty}^{\infty} db' \exp\left[-\frac{\sigma_s^2}{2\sigma_b^2\sigma_n^2} \left(b'^2 + s^2\frac{\sigma_b^2\sigma_n^2}{\sigma_s^4}\right)\right] \\
 &= e^{-s^2/(2\sigma_s^2)} \int_{-\infty}^{\infty} db' e^{-b'^2\sigma_s^2/(2\sigma_b^2\sigma_n^2)} \\
 &= \frac{\sqrt{2\pi}\sigma_s}{\sigma_b\sigma_n} e^{-s^2/(2\sigma_s^2)}.
 \end{aligned} \tag{1.31}$$

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 (1.31) haben wir angenommen, dass wir über b mit einem konstanten Maß integrieren. Eine Annahme dieser Art müssen wir machen, um die Integration ausführen zu können. Wir könnten stattdessen auch über $d \log b$ integrieren und aus der Substitution ein Integrationsmaß $d \log b/db = 1/b$ erhalten.

- 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 b 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.
- Im Frequentist-Zugang existiert kein Integrationsmaß, um einen Nuisance-Parameter wie b auszuintegrieren. Um trotzdem eine Likelihood für s zu erhalten projizieren wir den größten Wert der Likelihood als Funktion von b 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 (1.31) zunächst an, dass beide individuellen Likelihoods eine Gauss-Verteilung haben,

$$\begin{aligned}
 L(s) &= \max_b e^{-b^2/(2\sigma_b^2)} e^{-(s+b)^2/(2\sigma_n^2)} \\
 &= \max_b \exp \left[-\frac{b^2(\sigma_n^2 + \sigma_b^2) + 2sb\sigma_b^2 + s^2\sigma_b^2}{2\sigma_b^2\sigma_n^2} \right] \\
 &= \max_b \exp \left[-\frac{\sigma_b^2 + \sigma_n^2}{2\sigma_b^2\sigma_n^2} \left(b^2 + 2sb\frac{\sigma_b^2}{\sigma_b^2 + \sigma_n^2} + s^2\frac{\sigma_b^2}{\sigma_b^2 + \sigma_n^2} \right) \right] \\
 &= \max_b \exp \left[-\frac{\sigma_s^2}{2\sigma_b^2\sigma_n^2} \left(\left(b + s\frac{\sigma_b^2}{\sigma_s^2} \right)^2 - s^2\frac{\sigma_b^4}{\sigma_s^4} + s^2\frac{\sigma_b^2}{\sigma_s^2} \right) \right] \\
 &= \max_{b'} \exp \left[-\frac{\sigma_s^2}{2\sigma_b^2\sigma_n^2} \left(b'^2 + s^2\frac{\sigma_b^2\sigma_s^2 - \sigma_b^4}{\sigma_s^4} \right) \right] \\
 &= \max_{b'} \exp \left[-\frac{\sigma_s^2}{2\sigma_b^2\sigma_n^2} \left(b'^2 + s^2\frac{\sigma_b^2\sigma_n^2}{\sigma_s^4} \right) \right] \\
 &= e^{-s^2/(2\sigma_s^2)} \max_{b'} e^{-b'^2\sigma_s^2/(2\sigma_b^2\sigma_n^2)} \\
 &= e^{-s^2/(2\sigma_s^2)}, \tag{1.32}
 \end{aligned}$$

wobei wir wie vorher $b' = b + s\sigma_b^2/\sigma_s^2$ und $\sigma_s^2 = \sigma_b^2 + \sigma_n^2$ definieren. Dieses Ergebnis ist identisch mit dem Ergebnis für die Integration im Bayes-Zugang. Allerdings ist in der Profile-Likelihood die Normierung verlorengegangen.

- 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.

1.3 Taylor-Reihen

In (1.5) haben wir $\ln(1 + \delta)$ für $|\delta| \ll 1$ durch $\delta + \dots$ genähert. Eine solche Näherung heisst Taylor-Reihe, und ihre Bedeutung 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.

1.3.1 Taylorscher Satz

- Wir beginnen zunächst mit einer allgemeinen mathematischen Betrachtung, die auf alle Funktionen anwendbar ist, die genügend oft stetig differenzierbar sind. Wir nehmen an es gebe ein Intervall $I \in \mathbb{R}$ und darauf eine Funktion $f : I \rightarrow \mathbb{R}$, 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*

$$f(x) - f(a) = \int_a^x f'(t) dt, \quad (1.33)$$

wobei $f'(x)$ die Ableitung der Funktion $f(x)$ nach x bezeichnet.

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

$$f(x) = \sum_{j=0}^{\infty} \frac{(x-a)^j}{j!} f^{(j)}(a) \quad (1.34)$$

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

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

$$R_{n+1}(x) = \frac{1}{n!} \int_a^x (x-t)^n f^{(n+1)}(t) dt \quad (1.35)$$

ist. Dies ist die Aussage des *Taylorschen Satzes*, und (1.34) ist die *Taylorsche Formel*.

- Der Taylorsche Satz lässt sich am besten durch vollständige Induktion beweisen. Aus dem Hauptsatz der Differential- und Integralrechnung folgt bereits, dass die Aussage (1.34) für $n = 0$ wahr ist, denn es gilt offenbar

$$f(x) = f(a) + R_1(x) \quad (1.36)$$

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

nach (1.33). Damit ist die Induktion bei $n = 0$ verankert.

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

$$f(x) = f(a) + \sum_{i=1}^{n-1} \frac{f^{(i)}(a)}{i!} (x-a)^i + R_n(x) \quad (1.37)$$

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

$$\begin{aligned} R_n(x) &= \frac{1}{(n-1)!} \int_a^x (x-t)^{n-1} f^{(n)}(t) dt \\ &= - \int_a^x \frac{d}{dt} \left[\frac{(x-t)^n}{n!} \right] f^{(n)}(t) dt \\ &= - \frac{(x-t)^n}{n!} f^{(n)}(t) \Big|_a^x + \int_a^x \frac{(x-t)^n}{n!} f^{(n+1)}(t) dt \\ &= \frac{(x-a)^n}{n!} f^{(n)}(a) + R_{n+1}(x). \end{aligned} \quad (1.38)$$

Setzt man dieses Ergebnis in (1.37) ein, folgt damit

$$\begin{aligned} f(x) &= f(a) + \sum_{i=1}^{n-1} \frac{f^{(i)}(a)}{i!} (x-a)^i + \frac{(x-a)^n}{n!} f^{(n)}(a) + R_{n+1}(x) \\ &= f(a) + \sum_{i=1}^n \frac{f^{(i)}(a)}{i!} (x-a)^i + R_{n+1}(x). \end{aligned} \quad (1.39)$$

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

$$f(x) = f(a) + \beta(x-a) = \alpha + \beta a + \beta x - \beta a = \alpha + \beta x. \quad (1.40)$$

Deshalb nennt man eine Taylor-Entwicklung bis zur ersten Ordnung auch *Linearisierung* der Funktion in einer Umgebung von a .

1.3.2 Einfache Beispiele

- Der Taylorsche Satz gibt uns eine Näherung für das Verhalten einer Funktion in der Nähe bestimmter Punkte. Man wählt einen

Punkt a und betrachtet Punkte x in einer kleinen Umgebung von a , also $|x - a| \ll 1$. Betrachten wir ein Beispiel,

$$f(x) = \frac{1}{(1 \pm x)^b} \quad (1.41)$$

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

$$f(x) \approx f(0) \mp \frac{b}{(1 \pm 0)^{(b+1)}} x = 1 \mp bx. \quad (1.42)$$

Insbesondere ist also

$$\frac{1}{\sqrt{1 \pm x}} \approx 1 \mp \frac{x}{2}. \quad (1.43)$$

- Ein weiteres Beispiel haben wir in (1.5) benutzt,

$$f(x) = \ln(1 \pm x) \quad (1.44)$$

ebenfalls für $x \ll 1$. Dann ist

$$\ln(1 \pm x) \approx \ln(1 \pm 0) \pm \frac{1}{1 \pm 0} x = \pm x. \quad (1.45)$$

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

$$\sin x \approx \sin 0 + x \cos 0 = x \quad (1.46)$$

ergibt, erhält man durch Linearisierung des Cosinus

$$\cos x = \cos 0 - x \sin 0 = 1, \quad (1.47)$$

also eine Konstante. Wenn man den Verlauf der Cosinusfunktion in der Nähe von $x = 0$ berücksichtigen will, dann muss man bis zur zweiten Ordnung taylorln

$$\begin{aligned} \cos x &\approx \cos 0 - x \sin 0 - \frac{x^2}{2} \cos 0 = 1 - \frac{x^2}{2} \\ \sin x &\approx \sin 0 + x \cos 0 - \frac{x^2}{2} \sin 0 = x \end{aligned} \quad (1.48)$$

Wenn wir beide Funktionen konsistent bis zur zweiten Ordnung entwickeln, dann finden wir

$$\begin{aligned} \sin^2 x + \cos^2 x &= \left(x + \mathcal{O}(x^2)\right)^2 + \left(1 - \frac{x^2}{2} + \mathcal{O}(x^3)\right)^2 \\ &= x^2 + \left(1 - \frac{x^2}{2}\right)^2 + \mathcal{O}(x^3) \\ &= x^2 + 1 - x^2 + \mathcal{O}(x^3) = 1 + \mathcal{O}(x^3), \end{aligned} \quad (1.49)$$

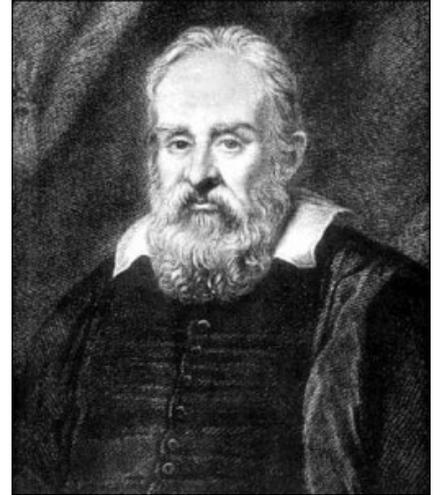
wobei der Term $\mathcal{O}(x^3)$ mathematisch bedeutet, dass in unserer Taylor-Entwicklung Terme dritter und höherer Potenz vernachlässigt werden.

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



Galileo Galilei



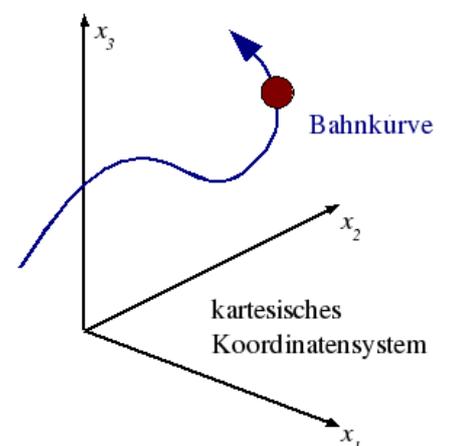
Isaac Newton

Galileis Fallversuche, die Analyse der Marsbahn durch Johannes Kepler, basierend auf Tycho Brahes Messungen, 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. 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.
- *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. Die 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:

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

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 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“ definieren wir das Produkt aus Masse m und Geschwindigkeit \vec{v} , 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:

$$\vec{p} = m\vec{v} = m\dot{\vec{x}} = \text{konstant} \quad (2.2)$$

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, dann ist das Bewegungsgesetz

$$\dot{\vec{p}} = m\ddot{\vec{x}} = \vec{F} \quad (2.3)$$

bei konstanter Masse m . Nach diesem Axiom enthalten Bewegungsgleichungen in der Regel zweite Ableitungen des Ortes nach der Zeit.

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

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.

³Actio = reactio.

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 beschreiben 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 illustriert, wie man inhomogene Differentialgleichungen lösen kann.

- Betrachten wir den komplexeren 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. Für solche inhomogene Gleichungen bringt man die homogenen und inhomogenen Teile auf verschiedene Seiten,

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

- Unser Rezept beginnt damit, erst einmal die allgemeine Lösung der homogenen Gleichung, also der Gleichung, die man erhält, wenn man die Inhomogenität auf der rechten Seite gleich Null setzt, zu berechnen. In unserem Fall ist das

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

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

$$\begin{aligned} N_B(t) &= B(t) e^{-\lambda_B t} \\ \dot{N}_B(t) &= \dot{B}(t) e^{-\lambda_B t} - B(t) \lambda_B e^{-\lambda_B t} = \dot{B}(t) e^{-\lambda_B t} - \lambda_B N_B(t) . \end{aligned} \quad (2.18)$$

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

$$\begin{aligned} (\dot{B}(t) e^{-\lambda_B t} - \lambda_B N_B(t)) + \lambda_B N_B(t) &= \lambda_A N_{A0} e^{-\lambda_A t} \\ \dot{B}(t) &= \lambda_A N_{A0} e^{-(\lambda_A - \lambda_B)t} \\ \Rightarrow \quad B(t) &= \frac{\lambda_A}{\lambda_B - \lambda_A} N_{A0} e^{-(\lambda_A - \lambda_B)t} + C , \end{aligned} \quad (2.19)$$

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

- Setzen wir (2.19) in (2.18) ein, so folgt

$$N_B(t) = C e^{-\lambda_B t} + \frac{\lambda_A}{\lambda_B - \lambda_A} N_{A0} e^{-\lambda_A t} . \quad (2.20)$$

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

$$N_{B0} = C + \frac{\lambda_A}{\lambda_B - \lambda_A} N_{A0} \quad \Leftrightarrow \quad C = N_{B0} - \frac{\lambda_A}{\lambda_B - \lambda_A} N_{A0} \quad (2.21)$$

und daraus die Lösung für die Differentialgleichung (2.15)

$$N_B(t) = N_{B0} e^{-\lambda_B t} + \frac{\lambda_A}{\lambda_B - \lambda_A} N_{A0} (e^{-\lambda_A t} - e^{-\lambda_B t}) . \quad (2.22)$$

- Man kann allgemein zeigen, dass alle Lösungen inhomogener Differentialgleichungen, die Summe der allgemeinen homogenen Gleichung und einer speziellen Lösung der inhomogenen Gleichung sind. Wir kommen in der Vorlesung noch einige Male auf diese Eigenschaft zurück.

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 nach dem Trägheitsgesetz

$$m\ddot{r}(t) = m\dot{v}(t) = F_G = -mg. \quad (2.23)$$

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

$$\begin{aligned} \dot{v}(t) &= -g, \\ v(t) = \dot{r}(t) &= -gt + C_1. \end{aligned} \quad (2.24)$$

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

$$r(t) = -\frac{g}{2}t^2 + v_0t + C_2, \quad (2.25)$$

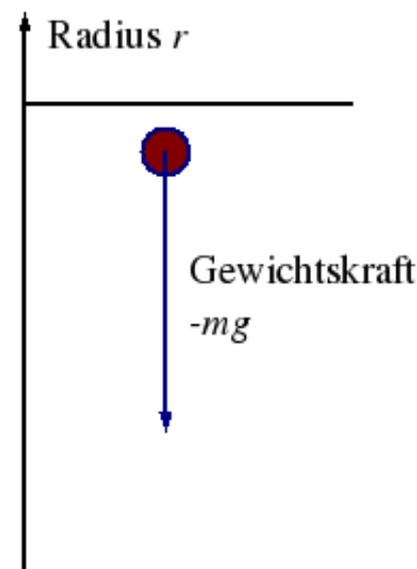
wobei die weitere Integrationskonstante C_2 die Anfangshöhe ist, $C_2 = r(t = 0) = r_0$. Da die Bewegungsgleichung von zweiter Ordnung in der Zeit ist, werden zu ihrer vollständigen Lösung zwei Integrationskonstanten benötigt. Diese werden durch die Anfangsbedingungen, in diesem Fall die Anfangshöhe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

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

$$v(t) = -gt \quad \text{und} \quad r(t) = h - \frac{g}{2}t^2. \quad (2.26)$$

Die Fallzeit bis $r(t_E) = 0$ beträgt

$$t_E = \sqrt{\frac{2h}{g}}, \quad (2.27)$$



freier Fall aus geringer Höhe

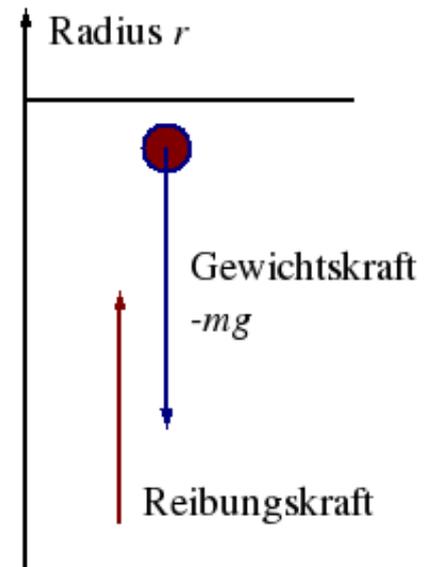
und die Endgeschwindigkeit ist

$$v_E = -gt_E = -g \sqrt{\frac{2h}{g}} = -\sqrt{2gh}. \quad (2.28)$$

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.



gebremster Fall aus geringer Höhe

- 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

$$m\ddot{r}(t) = m\dot{v}(t) = m \frac{dv}{dt} = -mg - K_R v(t) \quad (2.29)$$

lässt sich durch Trennung der Variablen v und t lösen:

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

$$\int \frac{dv}{1 + \frac{K_R v}{mg}} = \frac{mg}{K_R} \int \frac{dK_R v / (mg)}{1 + K_R v / (mg)} = -g \int dt. \quad (2.30)$$

Beide unbestimmte Integrale lassen sich integrieren,

$$\frac{mg}{K_R} \ln \left(1 + \frac{K_R v}{mg} \right) = -gt + C_1, \quad (2.31)$$

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

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

$$v(t) = \frac{mg}{K_R} \left(\exp \frac{K_R (C_1 - gt)}{mg} - 1 \right). \quad (2.32)$$

- Nach sehr langer Zeit, d.h. für $t \rightarrow \infty$, balancieren sich die beiden Kräfte und v nähert sich einer Endgeschwindigkeit

$$\dot{v}_E = 0 \quad \Rightarrow \quad v_E = -\frac{mg}{K_R}, \quad (2.33)$$

so dass die Lösung (2.32) in der Form

$$v(t) = v_E \left(1 - \exp \frac{-C_1 + gt}{v_E} \right) \quad (2.34)$$

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

$$v_0 = v_E \left(1 - e^{-C_1/v_E} \right) \Rightarrow e^{-C_1/v_E} = 1 - \frac{v_0}{v_E} \quad (2.35)$$

und damit

$$v(t) = v_E \left(1 - \left(1 - \frac{v_0}{v_E} \right) e^{gt/v_E} \right) = v_E + (v_0 - v_E) e^{gt/v_E} . \quad (2.36)$$

Um die Formel einfacher lesen zu können nutzen wir $v_E < 0$ für die Fallbewegung

$$v(t) = (v_0 + |v_E|) e^{-gt/|v_E|} - |v_E| = \begin{cases} v_0 & t = 0 \\ v_E & t \gg \frac{|v_E|}{g} = \frac{m}{K_R} \end{cases} . \quad (2.37)$$

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

$$\dot{r}(t) = (v_0 + |v_E|) e^{-gt/|v_E|} - |v_E| . \quad (2.38)$$

Eine weitere Integration führt auf

$$r(t) = -(v_0 + |v_E|) \frac{|v_E|}{g} e^{-gt/|v_E|} - |v_E|t + C_2 . \quad (2.39)$$

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

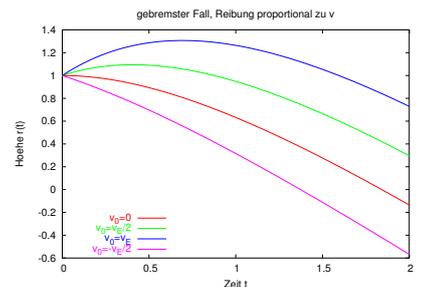
$$C_2 = r_0 + (v_0 + |v_E|) \frac{|v_E|}{g} , \quad (2.40)$$

und damit erhält man

$$r(t) = r_0 + (v_0 + |v_E|) \frac{|v_E|}{g} \left(1 - e^{-gt/|v_E|} \right) - |v_E|t . \quad (2.41)$$

- Nach hinreichend langer Zeit balancieren sich die beiden Kräfte, und r nimmt linear mit der Zeit ab,

$$r(t) \rightarrow r_0 + (v_0 + |v_E|) \frac{|v_E|}{g} - |v_E|t . \quad (2.42)$$



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

2.4.3 Fall aus geringer Höhe mit Luftwiderstand

Das Problem mit dem Luftwiderstand ist, dass er durch die Stokesche Reibung nicht korrekt beschrieben wird. Stattdessen ist beim Luftwiderstand der Betrag der Reibungskraft proportional zum Quadrat der Geschwindigkeit.

- Die Bewegungsgleichung lautet dann

$$m\ddot{r}(t) = m\dot{v}(t) = m\frac{dv}{dt} = -mg + K_v v(t)^2. \quad (2.43)$$

Wieder führt die Gewichtskraft zu einer Beschleunigung, und die Reibungskraft ist der Bewegung entgegen gerichtet. Für einen fallenden Körper ist $v < 0$ und zumindest vor dem Einfluss der Reibung auch $\dot{v} = -g < 0$. Wie vorher folgt nach Trennung der Variablen

$$\begin{aligned} \frac{dv}{dt} &= -g \left(1 - \frac{K_v}{gm} v(t)^2 \right) = -g \left(1 - \frac{v(t)^2}{v_E^2} \right) & v_E^2 &= \frac{mg}{K_v} \\ \int \frac{dv}{1 - \left(\frac{v}{v_E}\right)^2} &= -g \int dt. \end{aligned} \quad (2.44)$$

Wie vorher schreiben wir der Einfachheit halber die Formel als Funktion von $|v_E| > 0$. Bei dem Integral hilft die Partialbruchzerlegung

$$\frac{1}{1-x^2} = \frac{1}{2} \left(\frac{1}{1+x} + \frac{1}{1-x} \right), \quad (2.45)$$

denn sie erlaubt uns zu schreiben

$$\begin{aligned} \int \frac{dv}{2} \left(\frac{1}{1 + \frac{v}{|v_E|}} + \frac{1}{1 - \frac{v}{|v_E|}} \right) &= \frac{|v_E|}{2} \int d\frac{v}{v_E} \left(\frac{1}{\frac{v}{|v_E|} + 1} - \frac{1}{\frac{v}{|v_E|} - 1} \right) \\ &= \frac{|v_E|}{2} \ln \frac{|v + |v_E||}{|v - |v_E||} \\ &= \frac{|v_E|}{2} \ln \frac{|v_E| + v}{|v_E| - v} = -gt + C_1. \end{aligned} \quad (2.46)$$

Im letzten Schritt nutzen wir $v < 0$ und $|v| < |v_E|$. Die Integrationskonstante C_1 wird wie immer durch $v(t=0) = v_0$ bestimmt,

$$\begin{aligned} C_1 &= \frac{|v_E|}{2} \ln \frac{|v_E| + v_0}{|v_E| - v_0} \\ \Rightarrow \frac{|v_E|}{2} \ln \frac{|v_E| + v}{|v_E| - v} &= -gt + \frac{|v_E|}{2} \ln \frac{|v_E| + v_0}{|v_E| - v_0}. \end{aligned} \quad (2.47)$$

Diese Gleichung lässt sich umformen zu

$$\begin{aligned} \frac{|v_E| + v}{|v_E| - v} &= e^{-2gt/|v_E|} \frac{|v_E| + v_0}{|v_E| - v_0} \\ 1 + \frac{v}{|v_E|} &= e^{-2gt/|v_E|} \frac{|v_E| + v_0}{|v_E| - v_0} \left(1 - \frac{v}{|v_E|}\right) \\ \frac{v}{|v_E|} \left(1 + e^{-2gt/|v_E|} \frac{|v_E| + v_0}{|v_E| - v_0}\right) &= e^{-2gt/|v_E|} \frac{|v_E| + v_0}{|v_E| - v_0} - 1 \\ v(t) &= -|v_E| \frac{|v_E| - v_0 - (|v_E| + v_0) e^{-2gt/|v_E|}}{|v_E| - v_0 + (|v_E| + v_0) e^{-2gt/|v_E|}}. \end{aligned} \quad (2.48)$$

Wieder wird die asymptotische Endgeschwindigkeit im Grenzfall $t \gg |v_E|/g$ erreicht.

- Eine weitere Integration führt auf die durchfallene Höhe r ,

$$r(t) - r_0 = \int_0^t v(t') dt'. \quad (2.49)$$

Zur Vereinfachung definieren wir

$$x(t) = \frac{2gt}{|v_E|} \quad a = \frac{|v_E| + v_0}{|v_E| - v_0} \quad \Rightarrow \quad \frac{v(t)}{|v_E|} = -\frac{1 - a e^{-x}}{1 + a e^{-x}} \quad (2.50)$$

und erhalten

$$r(t) - r_0 = -\frac{v_E^2}{2g} \int_0^x dx' \frac{1 - a e^{-x'}}{1 + a e^{-x'}}. \quad (2.51)$$

Die Integration führt auf

$$r(t) - r_0 = -\frac{v_E^2}{2g} \left[x' + 2 \ln(1 + a e^{-x'}) \right]_0^x. \quad (2.52)$$

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

$$\begin{aligned} r(t) &= r_0 - \frac{v_E^2}{2g} \left[\frac{2gt}{|v_E|} + 2 \ln \left(1 + \frac{|v_E| + v_0}{|v_E| - v_0} e^{-2gt/|v_E|} \right) - 2 \ln \left(1 + \frac{|v_E| + v_0}{|v_E| - v_0} \right) \right] \\ &= r_0 - |v_E|t - \frac{v_E^2}{g} \ln \frac{|v_E| - v_0 + (|v_E| + v_0) e^{-2gt/|v_E|}}{|v_E| - v_0 + |v_E| + v_0} \\ &= r_0 - |v_E|t - \frac{v_E^2}{g} \ln \left(\frac{1}{2} - \frac{v_0}{2|v_E|} + \frac{|v_E| + v_0}{2|v_E|} e^{-2gt/|v_E|} \right) \end{aligned} \quad (2.53)$$

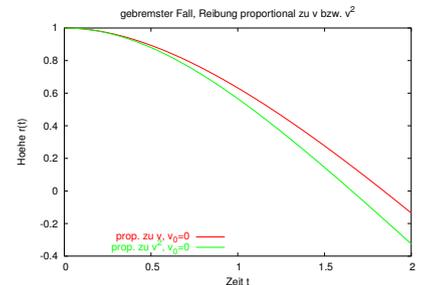
- Wenn wir weiterhin $v_0 = 0$ annehmen, dann erhält man nach genügend langer Zeit, $t \gg |v_E|/g$, den asymptotischen Verlauf

$$r(t) \rightarrow r_0 + \frac{v_E^2}{g} \ln 2 - |v_E|t, \quad (2.54)$$

Wir können diesen Fall mit dem Ergebnis für Stokessche Reibung nach (2.42) vergleichen, wenn wir annehmen dass wir die Stärke der Reibung so einstellen dass die Endgeschwindigkeit dieselbe ist,

$$r(t) \rightarrow r_0 + \frac{v_E^2}{g} - |v_E|t . \tag{2.55}$$

Wegen $\ln 2 < 1$ liegt der Körper mit Stokesscher Reibung in $r(t)$ immer über dem Körper mit Luftwiderstand. Stokessche Reibung ist also effektiver und bremst den freien Fall stärker ab.



Vergleich der durchfallenen Höhen für gebremsten Fall mit Stokesscher 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 Graviationskraft 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

$$F_G = -\frac{GMm}{r^2} \neq -mg \tag{2.56}$$

Auf der Erdoberfläche müssen beide Formulierungen der Gravitationskraft dasselbe Ergebnis haben, also $GM/R_{\text{Erde}}^2 = g$. Die allgemeinen Bewegungsgleichung für einen freien Fall ohne Reibung lautet

$$m\ddot{r}(t) = -\frac{GMm}{r(t)^2} . \tag{2.57}$$

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

$$\begin{aligned} m\dot{r}(t) \ddot{r}(t) &= -m \frac{GM\dot{r}(t)}{r(t)^2} \\ \frac{m}{2} \frac{d}{dt} \dot{r}(t)^2 &= mGM \frac{d}{dt} \frac{1}{r(t)} . \end{aligned} \tag{2.58}$$

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

$$\frac{m}{2} \dot{r}(t)^2 - \frac{mGM}{r(t)} =: E \tag{2.59}$$

konstant ist. 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 (2.59) entspricht also einer Energie-Erhaltung. Die konstante Energie $E = E_0$ ist durch die Anfangsbedingungen gegeben,

$$E_0 = \frac{m}{2}v_0^2 - \frac{mGM}{r_0} . \quad (2.60)$$

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

$$\frac{m}{2}v(t)^2 = \frac{mGM}{r(t)} = \frac{gR_{\text{Erde}}^2 m}{r(t)} . \quad (2.61)$$

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

$$v_\infty = \sqrt{2gR_{\text{Erde}}} = 11.2 \text{ km s}^{-1} . \quad (2.62)$$

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

- Setzt man in (2.61) für v_∞ die Lichtgeschwindigkeit c ein, erhält man den *Schwarzschildradius* einer Masse M

$$R_S = \frac{2GM}{c^2} , \quad (2.63)$$

der salopp sagt, wie viel Masse ein Objekt haben muss, dass ein Lichtstrahl es nicht verlassen kann. 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. Ob diese Lösung eindeutig ist, ist in der Newtonschen Mechanik nicht relevant, wenn das entsprechende Experiment einen eindeutigen Ausgang hat. Das Experiment beschreibt immer die korrekte Lösung. Trotzdem schadet es nicht, wenn wir ein Gefühl für Lösungen von Differentialgleichungen bekommen.

- Zunächst ist es 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 haben, 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 der gesuchten Funktion als eigene Funktionen auf, also $y(x) = y_1(x)$, $y'(x) = y_2(x)$, $y''(x) = y_3(x)$ usw. Damit wird eine Differentialgleichung zweiter Ordnung,

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

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 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 $\dot{N} = -\lambda N$. Wir ersetzen in (3.2) nun $y \rightarrow N$, $x \rightarrow t$ und $f(x, y) \rightarrow -\lambda N(t)$. Die Sekantensteigung ist dann

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

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

Wir können 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 Raum der reellen Zahlen nicht beschränkt. Wenn in dieser Situation aber $y(x)$ stetig differenzierbar und damit glatt ist, dann kann man durch eine Einschränkung des Parameterbereiches in x auch den Bereich von y einschränken und damit die Lipschitz-Bedingung 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. Wir haben schon mehrfach gesehen, dass die Lösung dann durch zwei Anfangsbedingungen an den Ort und die Geschwindigkeit eindeutig bestimmt wird.

3.1.2 Lineare Differentialgleichungen n -ter Ordnung

Von den Fallgesetzen haben wir eine Vorstellung, welche Art von Differentialgleichungen physikalisch relevant sind. Insbesondere müssen wir noch einmal auf inhomogene Gleichungen zurück kommen.

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

$$y''(x) - y(x) = 1. \quad (3.8)$$

Die dazugehörige homogene Gleichung

$$y''(x) - y(x) = 0 \quad \Leftrightarrow \quad y''(x) = y(x) \quad (3.9)$$

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

$$\begin{aligned} 2y'y'' = 2yy' &\Leftrightarrow (y^2)' = (y^2)' \\ &\Leftrightarrow y' = \pm y \quad \Leftrightarrow y_{1,2}(x) = e^{\pm x}. \end{aligned} \quad (3.10)$$

Solange wir nur an einer Lösung der homogenen Gleichung interessiert sind, können wir die Integrationskonstante vernachlässigen und haben damit die beiden erlaubten Lösungen.

- Wenn man nur eine Lösung $y_1(x)$ einer linearen homogenen Differentialgleichung zweiter Ordnung hat, dann kann man sich eine zweite zum Beispiel durch *d'Alembert-Reduktion* verschaffen. 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,

$$\begin{aligned} 0 &= y_1''f + 2y_1'y_1'f' + y_1f'' - y_1f \\ &= (y_1'' - y_1)f + 2y_1'y_1'f' + y_1f'' = 2y_1'y_1'f' + y_1f''. \end{aligned} \quad (3.11)$$

Daraus folgt mit $y_1 = e^x = y_1'$

$$\begin{aligned} \frac{f''}{f'} &= -2\frac{y_1'}{y_1} = -2 &\Leftrightarrow (\ln f')' &= -2 \\ \Leftrightarrow \ln f' &= -2x &\Leftrightarrow f'(x) &= e^{-2x}. \end{aligned} \quad (3.12)$$

Man erhält dann durch Integration

$$f = -\frac{1}{2}e^{-2x} \Leftrightarrow y_2 = y_1 f = -\frac{1}{2}e^{-x} \quad (3.13)$$

als weitere Lösung. Das ist die bekannte zweiten Lösung $y_2(x) = e^{-x}$.

- 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,

$$\sum_{i=1}^n a_i y_i(x) = 0 \Rightarrow a_i = 0 \quad \text{für alle } 1 \leq i \leq n. \quad (3.14)$$

Wenn wir einen Satz Lösungen kennen, dann wollen wir wissen ob sie in der Tat unabhängig sind. Es gibt ein einfaches Verfahren, die festzustellen. Spezialisiert auf $n = 2$ besteht es darin zu prüfen, ob der Ausdruck

$$y_1(x)y_2'(x) - y_2(x)y_1'(x) \quad (3.15)$$

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 (3.15) folgt

$$e^x(-e^{-x}) - e^{-x}e^x = -2 \neq 0, \quad (3.16)$$

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

$$y(x) = a_1 e^x + a_2 e^{-x}. \quad (3.17)$$

- Jetzt betrachten wir wieder die inhomogenen Differentialgleichung. Wir hatten schon gesagt, dass man die allgemeine Lösung einer *inhomogenen* linearen Differentialgleichung erhält, 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

$$y(x) = a_1 y_1(x) + a_2 y_2(x) + y_{\text{in}}(x), \quad (3.18)$$

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

$$y(x) = a_1 e^x + a_2 e^{-x} - 1. \quad (3.19)$$

Um die Konstanten $a_{1,2}$ festzulegen nutzen wir wie immer Anfangsbedingungen, zum Beispiel $y'(x=0) = 0$,

$$y'(x) = a_1 e^x - a_2 e^{-x} = 0 \quad \text{für } x = 0, \quad (3.20)$$

und damit $a_1 = a_2$. Als zweite Bedingung nehmen wir an dass $y(x=0) = 0$, also

$$y(x) = a_1 e^x + a_1 e^{-x} - 1 = 0 \quad \text{für } x = 0, \quad (3.21)$$

oder $a_1 = 1/2$. Damit ist die vollständige Lösung der inhomogenen Gleichung mit Anfangsbedingungen

$$y(x) = \frac{e^x + e^{-x}}{2} - 1. \quad (3.22)$$

- Formaler gesehen haben wir anhand unserer Beispiele gesehen, dass Differentialgleichungen in der Newtonschen Mechanik oft linear. Das heisst, ihre Koeffizientenfunktionen in der Form

$$p_n(x)y^{(n)}(x) + \dots + p_1(x)y'(x) + p_0(x)y(x) = r(x) \quad (3.23)$$

hängen nicht von y ab. 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 (3.23) sind, so ist auch eine beliebige Linearkombination der beiden eine Lösung, also

$$L[a_1 y_1 + a_2 y_2] = (a_1 + a_2)r(x) \quad \text{mit } a_1, a_2 \in \mathbb{R}. \quad (3.24)$$

Eingesetzt in (3.23) sieht man sofort, dass dies richtig ist. Die nächste Frage ist, was wir über diese Art von Struktur sagen können.

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.
- Wir wissen intuitiv was ein Vektor im dreidimensionalen Raum ist. Er hat einen Betrag und eine Richtung und wird 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,

$$+ : V \times V \rightarrow V, \quad (v_1, v_2) \mapsto v_1 + v_2, \quad (3.25)$$

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

$$\cdot : K \times V \rightarrow V, \quad (\lambda, v) \mapsto \lambda \cdot v = \lambda v. \quad (3.26)$$

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

$$\begin{aligned} \lambda (v_1 + v_2) &= \lambda v_1 + \lambda v_2 \\ (\lambda_1 + \lambda_2) v &= \lambda_1 v + \lambda_2 v. \end{aligned} \quad (3.27)$$

- Hier 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 (3.24) 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,

$$\psi = \lambda_1\psi_1 + \lambda_2\psi_2 . \quad (3.28)$$

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.

- In (3.26) wurde der Begriff eines Körpers eingeführt, ohne ihn 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,

$$a_1 \circ (a_2 \circ a_3) = (a_1 \circ a_2) \circ a_3 , \quad (3.29)$$

für die es ein *neutrales Element* e gibt,

$$a \circ e = a , \quad (3.30)$$

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

$$a \circ a^{-1} = e . \quad (3.31)$$

Kommutativ oder *abelsch* heißt eine Gruppe dann, wenn

$$a_1 \circ a_2 = a_2 \circ a_1 \quad (3.32)$$

ist. Ein einfaches Beispiel sind die ganzen Zahlen bezüglich der Addition. Die Summe zweier ganzer Zahlen ist wieder eine, das neutrale Element ist die 0 und das inverse Element zu einer ganzen Zahl n ist $-n$. Ein weiteres Beispiel sind die reellen Zahlen bezüglich der Multiplikation, mit dem neutralen Element 1 und dem Inversen $1/n$.

Ein *Körper* K ist eine abelsche Gruppe bezüglich einer Addition $+$, die ausserdem auch bezüglich einer Multiplikation \cdot eine abelsche Gruppe ist. Da wir schon wissen dass für das neutrale Element der Addition, die Null, ein inverses Element der Multiplikation problematisch wird, erlauben wir diese Ausnahme in der Definition des Körpers. Außerdem gilt das Distributivgesetz in der Form

$$a \cdot (b + c) = a \cdot b + a \cdot c , \quad (a + b) \cdot c = a \cdot c + b \cdot c . \quad (3.33)$$

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 (3.14) 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,

$$v = \lambda_1 v_1 + \lambda_2 v_2 + \dots + \lambda_n v_n . \quad (3.34)$$

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,

$$U = \text{span}(v_i) . \quad (3.35)$$

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

$$\lambda_1 v_1 + \lambda_2 v_2 + \dots + \lambda_n v_n = 0 \quad (3.36)$$

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,

$$v = \sum_{i=1}^N \lambda_i e_i , \quad (3.37)$$

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,

$$\begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix} + \begin{pmatrix} w_1 \\ w_2 \\ w_3 \end{pmatrix} = \begin{pmatrix} v_1 + w_1 \\ v_2 + w_2 \\ v_3 + w_3 \end{pmatrix} \quad \text{und} \quad \lambda \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix} = \begin{pmatrix} \lambda v_1 \\ \lambda v_2 \\ \lambda v_3 \end{pmatrix}. \quad (3.38)$$

Eine geeignete Basis sind die Einheitsvektoren

$$e_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \quad e_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \quad e_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}. \quad (3.39)$$

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 $[-1, 1]$,

$$p(x) = \sum_{j=0}^N a_j x^j, \quad a_j \in \mathbb{R}. \quad (3.40)$$

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

$$\lambda_0 + \lambda_1 x + \dots + \lambda_n x^n = 0 \quad (3.41)$$

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.

Kapitel 4

Impuls, Drehimpuls und Energie

4.1 Vektoren II

Um unsere eher zufällige Definition von der Energie als Erhaltungsgröße systematisch zu verstehen benötigen wir weitere Eigenschaften von Vektoren im 3-dimensionalen Raum. Wir beginnen damit, dass die Addition und die Streckung von Vektoren nicht die einzigen hilfreichen Konzepte sind, denken wir zum Beispiel an die Definition eines Öffnungswinkels.

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 und damit *bilinear* ist,

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

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 $v = 0$. Damit können wir einem Vektor $v \in V$ sein Betrag $|v|$ zuordnen,

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

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) auszusprechen. 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. Dabei müssen wir uns erinnern, dass ein Index in einer Gleichung niemals mehr als zweimal erscheinen kann!

- Ebenso nützlich ist das Kronecker-Symbol

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

Damit wird (4.4) zu

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

und (4.5) entsprechend

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

- Weiterhin ist für dreidimensionale Vektorräume das vollkommen antisymmetrische Levi-Civita-Symbol ϵ_{ijk} nützlich,

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

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 Zusammenhang

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

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,

$$\begin{aligned} \langle a, a \times b \rangle &= \langle \alpha_i e_i, \epsilon_{jkl} \alpha_j \beta_k e_l \rangle \\ &= \alpha_i \alpha_j \beta_k \epsilon_{jkl} \delta_{il} = \alpha_i \alpha_j \beta_k \epsilon_{jki} = 0 . \end{aligned} \quad (4.19)$$

Im letzten Schritt haben wir verwendet, dass das Produkt $\alpha_i \alpha_j \beta_k$ symmetrisch gegen Vertauschung von i und j ist, ϵ_{ijk} aber antisymmetrisch.

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

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

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

- Unsere abstrakte Definition von Vektoren und Vektorräumen macht klar, 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 nicht mit den Vektoren selbst verwechselt werden.
- 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. Für eine Orthonormalbasis 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. Wir wissen dass die neuen Basisvektoren e'_i durch die alten ausgedrückt werden können

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

Diese $N \times N$ Gleichungen legen alle a_{ij} vollständig fest. Um die 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 aus den alten v_i erhalten, und umgekehrt,

$$\begin{aligned} v'_k &= \langle v, e'_k \rangle = v_i \langle e_i, e'_k \rangle = a_{ki} v_i \\ v_k &= \langle v, e_k \rangle = v'_j \langle e'_j, e_k \rangle = v'_j a_{jk} . \end{aligned} \quad (4.26)$$

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 geradliniggleichförmige 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 geradliniggleichförmig 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,

$$\vec{p} = \text{const.} \quad \text{in Abwesenheit von Kräften } \vec{F} \quad (4.27)$$

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

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

$$\dot{\vec{p}} = \vec{F}. \quad (4.28)$$

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

4.2.2 Drehmoment und Drehimpuls

- Es stellt sich heraus, dass es sinnvoll ist den *Drehimpuls* um den Koordinatenursprung durch

$$\vec{L} = \vec{x} \times \vec{p} = \vec{x} \times m\vec{v} \quad (4.29)$$

zu definieren. Im Fall $\vec{r} \parallel \vec{p}$ gibt es keine Drehbewegung und der Drehimpuls ist null. Seine Zeitableitung ist

$$\begin{aligned} \frac{d\vec{L}(t)}{dt} &= \dot{\vec{x}} \times \vec{p} + \vec{x} \times \dot{\vec{p}} \\ &= \frac{\vec{p}}{m} \times \vec{p} + \vec{x} \times \dot{\vec{p}} \\ &= \vec{x} \times \dot{\vec{p}} = \vec{x} \times \vec{F}. \end{aligned} \quad (4.30)$$

- Ganz allgemein wird das *Moment* einer Kraft \vec{F} bezüglich des Koordinatenursprungs durch

$$\vec{M} := \vec{x} \times \vec{F} \quad (4.31)$$

definiert. Wenn ϕ der Winkel zwischen \vec{x} und \vec{F} ist, hat es den Betrag $|\vec{M}| = xF \sin \phi$. Das Moment einer Kraft verschwindet also, wenn sie in dieselbe Richtung zeigt wie der Ortsvektor \vec{x} ihres Angriffspunkts. Eingesetzt in (4.30) ergibt dies den *Drehimpulssatz*

$$\dot{\vec{L}} = \vec{M}. \quad (4.32)$$

In Abwesenheit von Drehmomenten, $\vec{M} = 0$, bleibt der Drehimpuls erhalten, $\dot{\vec{L}} = 0$. Der Vergleich mit (4.28) 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,

$$m\ddot{x} = F(x) \quad \Rightarrow \quad m\dot{x}\ddot{x} = \dot{x}F(x) \quad (4.33)$$

Wir definieren nun das Integral

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

$$\frac{d}{dt} V(x) = \frac{dV}{dx} \dot{x} = -F(x) \dot{x} \quad (4.34)$$

wobei der Anfangspunkt x_0 kann frei gewählt werden kann. Damit wird (4.33)

$$\frac{m}{2} \frac{d(\dot{x}^2)}{dt} = - \frac{d}{dt} V(x) \quad \Leftrightarrow \quad \frac{d}{dt} \left[\frac{m}{2} \dot{x}^2 + V(x) \right] = 0 \quad (4.35)$$

- Demnach ist die *Energie*

$$E := \frac{m}{2} \dot{x}^2 + V(x) \quad (4.36)$$

erhalten. 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.

- Nach (4.36) lässt sich die Trajektorie des Körpers implizit durch Trennung der Variablen ausdrücken

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

$$\Rightarrow \pm \int_{x_0}^x \frac{dx'}{\sqrt{\frac{2}{m}(E - V)}} = t - t_0. \quad (4.37)$$

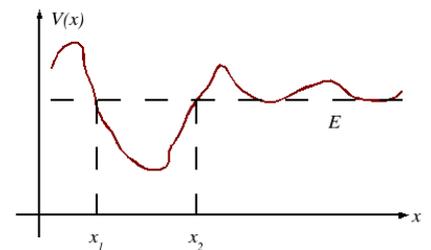
Offenbar ist Bewegung nur dort möglich, wo $E - V \geq 0$ ist. Die kinetische Energie $T = (m/2)\dot{x}^2$ ist positiv-semidefinit, $T \geq 0$.

An den sogenannten *Umkehrpunkten* $x_{1,2}$ wird die kinetische Energie minimal ($T = 0$) und daher $V = E$. Wegen des \pm -Zeichens ändert sich dort die Richtung der Bewegung. Dadurch tritt eine Art von Schwingung zwischen x_1 und x_2 auf. Sie hat die (halbe) Schwingungsperiode

$$t_2 - t_1 = \int_{x_1}^{x_2} \frac{dx}{\sqrt{\frac{2}{m}(E - V)}} = \sqrt{\frac{m}{2}} \int_{x_1}^{x_2} \frac{dx}{\sqrt{E - V}} \quad (4.38)$$

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

$$F(x) = -kx \quad \Rightarrow \quad V(x) = - \int_0^x F(x') dx' = \frac{k}{2} x^2, \quad (4.39)$$



Einschränkung der Bewegung und Umkehrpunkte

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

$$E = \frac{m}{2}\dot{x}^2 + \frac{k}{2}x^2 = \frac{m}{2}(\dot{x}^2 + \omega_0^2 x^2) \quad \text{mit} \quad \omega_0 := \sqrt{\frac{k}{m}}. \quad (4.40)$$

Die Umkehrpunkte liegen bei

$$T = 0 \quad \Leftrightarrow \quad E = \frac{m}{2}\omega_0^2 x_{1,2}^2 \quad \Leftrightarrow \quad x_{1,2} = \pm \sqrt{\frac{2E}{m\omega_0^2}}. \quad (4.41)$$

- Die allgemeine Lösung des harmonischen Oszillators ist ebenfalls durch die Bedingung gegeben, dass seine Energie konstant ist. Mit Hilfe der Umkehrpunkte schreiben wir dies als

$$E = \frac{m}{2}\omega_0^2 A_0^2 = \text{konst.} \quad \text{mit} \quad x_{1,2} = \pm A_0 = \pm \sqrt{\frac{2E}{m\omega_0^2}}. \quad (4.42)$$

Aus der allgemeinen Formel (4.37) kann man dann die Zeit zwischen den Umkehrpunkten berechnen,

$$\begin{aligned} t_2 - t_1 &= \int_{-A_0}^{A_0} \frac{dx}{\omega_0 \sqrt{A_0^2 - x^2}} \\ &= \frac{1}{\omega_0} \arcsin \frac{x}{A_0} \Big|_{-A_0}^{A_0} \\ &= \frac{\arcsin 1 - \arcsin(-1)}{\omega_0} = \frac{\pi}{\omega_0}. \end{aligned} \quad (4.43)$$

Für einen kompletten Durchlauf müssen wir von t_1 über t_2 wieder zu t_1 , also ist die Schwingungsperiode $2(t_2 - t_1) = 2\pi/\omega_0$.

- Durch einen (quadratischen) Reibungsterm kommt eine dritte Form der Energie oder ein Energieverlust des Systems hinzu. Die Bewegungsgleichung für einen gedämpften harmonischen Oszillator (4.39) lässt sich dann in der Form

$$\frac{dE}{dt} = \frac{d}{dt} \left[\frac{m}{2}\dot{x}^2 + \frac{k}{2}x^2 \right] = -b\dot{x}^2 \quad (4.44)$$

schreiben. Die Energie nimmt ab wenn $-b\dot{x}^2 \leq 0$. Es handelt sich um ein sogenanntes *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 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 zu Vektorräumen bezeichnen wir Vektoren im dreidimensionalen Vektorraum \mathbb{R}^3 mit Pfeilen. In der Regel stellen wir diese Vektoren als *Spaltenvektoren* mit transponierten Zeilenvektoren dar,

$$\vec{x} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \quad \text{und} \quad \vec{x}^T = (x_1, x_2, x_3). \quad (5.1)$$

Weiterhin notieren wir das Skalarprodukt durch einen Punkt,

$$\langle \vec{x}, \vec{y} \rangle \equiv \vec{x} \cdot \vec{y}. \quad (5.2)$$

- 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,

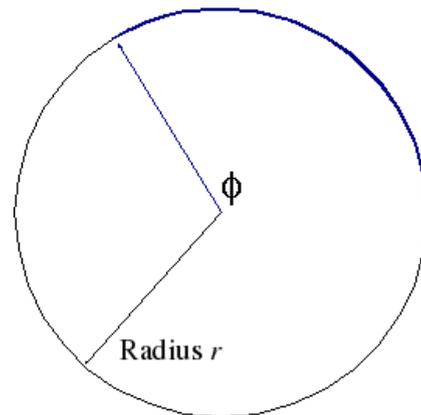
$$\vec{x}(t) = \begin{pmatrix} x_1(t) \\ x_2(t) \\ x_3(t) \end{pmatrix}. \quad (5.3)$$

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

$$\vec{x}_K(t) = R \begin{pmatrix} \cos \phi(t) \\ \sin \phi(t) \\ 0 \end{pmatrix}, \quad (5.4)$$

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

$$\vec{x}_S(t) = \begin{pmatrix} R \cos \phi(t) \\ R \sin \phi(t) \\ x_3(t) \end{pmatrix}. \quad (5.5)$$



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

$$\vec{v}(t) = \dot{\vec{x}}(t) = \frac{d\vec{x}(t)}{dt}. \quad (5.6)$$

Die Geschwindigkeiten in den obigen Beispielen sind

$$\vec{v}_K(t) = R \begin{pmatrix} -\dot{\phi} \sin \phi(t) \\ \dot{\phi} \cos \phi(t) \\ 0 \end{pmatrix} \quad \text{und} \quad \vec{v}_S(t) = \begin{pmatrix} -R\dot{\phi} \sin \phi(t) \\ R\dot{\phi} \cos \phi(t) \\ \dot{x}_3(t) \end{pmatrix}. \quad (5.7)$$

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

$$\vec{a}(t) = \dot{\vec{v}}(t) = \frac{d\vec{v}(t)}{dt} = \ddot{\vec{x}}(t) = \frac{d^2\vec{x}(t)}{dt^2}. \quad (5.8)$$

Die Beschleunigungen in den obigen Beispielen sind also

$$\vec{a}_K(t) = R \begin{pmatrix} -\ddot{\phi} \sin \phi(t) - \dot{\phi}^2 \cos \phi(t) \\ \ddot{\phi} \cos \phi(t) - \dot{\phi}^2 \sin \phi(t) \\ 0 \end{pmatrix} \quad \text{und} \\ \vec{a}_S(t) = \begin{pmatrix} -R\ddot{\phi} \sin \phi(t) - R\dot{\phi}^2 \cos \phi(t) \\ R\ddot{\phi} \cos \phi(t) - R\dot{\phi}^2 \sin \phi(t) \\ \ddot{x}_3(t) \end{pmatrix}. \quad (5.9)$$

5.1.2 Bogenlänge, Tangential- und Normalvektoren

- Die Bewegung an einem Punkt der Bahnkurve ist durch die *Bogenlänge* ds und den *Tangentialvektor* τ an die Bahnkurve gegeben,

$$d\vec{x} = \vec{\tau} ds. \quad (5.10)$$

Hier spalten wir einen infinitesimalen Vektor einer Bahnkurve in eine Richtung und in die infinitesimale Bogenlänge auf. Ein beliebig kleines (infinitesimales) Element der *Bogenlänge* längs der Bahnkurve ist

$$ds = |d\vec{x}(t)| = \left| \frac{d\vec{x}}{dt} \right| dt = |\vec{v}(t)| dt \quad (5.11)$$

Die integrierte Bogenlänge

$$s = \int ds = \int_0^t |\vec{v}(t')| dt' . \quad (5.12)$$

ergibt die 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,

$$\vec{x}(t) \rightarrow \vec{x}(s) \quad \text{mit} \quad \frac{ds}{dt} = |\vec{v}(t)| . \quad (5.13)$$

Die Jacobi-Determinante lässt sich also aus dem Geschwindigkeitsfeld berechnen.

- Nach der Aufspaltung in (5.10) ist der Tangentialvektor ein Einheitsvektor,

$$\vec{\tau} = \frac{d\vec{x}}{ds} \quad \text{mit} \quad |\vec{\tau}| = 1 . \quad (5.14)$$

Als Teil der entsprechenden Basis von \mathbb{R}^3 steht *Hauptnormalenvektor*

$$\vec{n}_H = \frac{d\vec{\tau}}{ds} \left| \frac{d\vec{\tau}}{ds} \right|^{-1} \quad (5.15)$$

senkrecht auf $\vec{\tau}$,

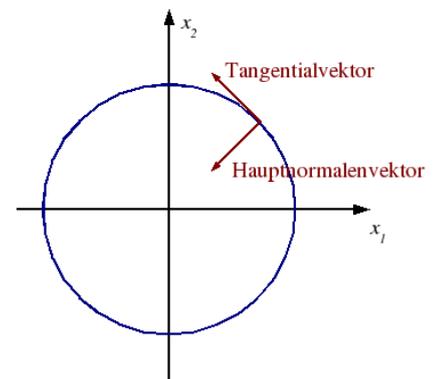
$$\vec{\tau}^2 = 1 \quad \Rightarrow \quad 0 = \frac{d(\vec{\tau}^2)}{ds} = 2\vec{\tau} \cdot \frac{d\vec{\tau}}{ds} . \quad (5.16)$$

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

$$\vec{n}_B = \vec{\tau} \times \vec{n}_H . \quad (5.17)$$

- Als einfaches Beispiel dient die Kreisbewegung mit konstanter Geschwindigkeit. Hier ergibt sich die Bogenlänge nach (5.12)

$$s = \int ds = \int_0^t |\vec{v}(t')| dt' = \int_0^t R\omega dt' = R\omega t . \quad (5.18)$$



Tangential- und Hauptnormalenvektor bei der Kreisbewegung

Im letzten Schritt haben wir benutzt dass wegen $|\vec{v}(t)| = R\dot{\phi}$ auch die *Winkelgeschwindigkeit* $\dot{\phi} = \omega$ konstant ist. Die Relation (5.18) erlaubt es uns die bekannte Bahnkurve einer Kreisbewegung durch s auszudrücken,

$$\vec{x}(t) = R \begin{pmatrix} \cos(\omega t) \\ \sin(\omega t) \\ 0 \end{pmatrix} = R \begin{pmatrix} \cos(s/R) \\ \sin(s/R) \\ 0 \end{pmatrix} = \vec{x}(s). \quad (5.19)$$

Aus dieser Form können wir den Tangentialvektor

$$\vec{\tau} = \frac{d\vec{x}(s)}{ds} = \begin{pmatrix} -\sin(s/R) \\ \cos(s/R) \\ 0 \end{pmatrix} = \begin{pmatrix} -\sin \omega t \\ \cos \omega t \\ 0 \end{pmatrix}, \quad (5.20)$$

den Hauptnormalenvektor

$$\vec{n}_H = \frac{1}{R} \begin{pmatrix} -\cos(s/R) \\ -\sin(s/R) \\ 0 \end{pmatrix} \left(\frac{1}{R}\right)^{-1} = - \begin{pmatrix} \cos(s/R) \\ \sin(s/R) \\ 0 \end{pmatrix} \quad (5.21)$$

und den Binormalenvektor

$$\vec{n}_B = \begin{pmatrix} -\sin(s/R) \\ \cos(s/R) \\ 0 \end{pmatrix} \times \begin{pmatrix} -\cos(s/R) \\ -\sin(s/R) \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}, \quad (5.22)$$

ablesen.

5.1.3 Tangential- und Normalkomponenten

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

$$\vec{v}(t) = \frac{d\vec{x}(t)}{dt} = \frac{d\vec{x}}{ds} \frac{ds}{dt} = |\vec{v}(t)| \cdot \vec{\tau}. \quad (5.23)$$

Die Geschwindigkeit ist also tangential zur Bahnkurve.

- Für die Beschleunigung ergibt sich mit $v = |\vec{v}|$

$$\frac{d\vec{v}(t)}{dt} = \frac{d(v\vec{\tau})}{dt} = \dot{v}\vec{\tau} + v \frac{d\vec{\tau}}{dt} = \dot{v}\vec{\tau} + v \frac{d\vec{\tau}}{ds} \frac{ds}{dt} = \dot{v}\vec{\tau} + v^2 \frac{d\vec{\tau}}{ds}, \quad (5.24)$$

Sie hat eine tangentielle Komponente und eine Komponente in Richtung der Hauptnormalen. Für konstantes $|\vec{v}|$ steht die Beschleunigung senkrecht auf der Geschwindigkeit oder der Bahnkurve. Offenbar spielt bei der Berechnung der Beschleunigung der nicht normierte Hauptnormalenvektor $d\vec{\tau}/ds$ eine Rolle.

- Wir können diesen nicht normierten Hauptnormalenvektor mit dem lokalen *Krümmungsradius* verknüpfen. Wir legen an einen Punkt der Bahnkurve nicht eine Tangente, sondern einen Kreis an, der lokal dieselbe Krümmung hat wie die Kurve. Der Radius dieses Kreises ist dann durch die Länge des nicht normierten Hauptnormalenvektors gegeben,

$$\frac{1}{\rho} := \left| \frac{d\vec{\tau}}{ds} \right|. \quad (5.25)$$

gegeben. Je stärker der Tangentialvektor sich entlang der Bogenlänge ändert, desto größer ist die Krümmung und desto kleiner ist der Radius ρ . Für die Beschleunigung heisst das

$$\frac{d\vec{v}(t)}{dt} = \dot{v}\vec{\tau} + \frac{v^2}{\rho}\vec{n}_H. \quad (5.26)$$

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,

$$F(x) = -\frac{dV(x)}{dx} \Leftrightarrow V(x) = -\int_{x_0}^x F(x')dx'. \quad (5.27)$$

Das Integral können wir entlang einer Bahnkurve $s(t)$ in drei Dimensionen auswerten. Der Anfangspunkt x_0 der Integration war beliebig, aber in drei Dimensionen müssen wir den Integrationsweg festlegen.

- Zwei offensichtliche Fragen aus der Definition dieses Potentials sind, ersten, wie wir sie auf die Kraft als Vektor $\vec{F}(\vec{x})$ verallgemeinern, und wann ein Potential $V(\vec{x})$ vom Integrationsweg unabhängig ist. 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,

$$C : I \rightarrow \mathbb{R}^3, \quad t \mapsto \vec{x}(t). \quad (5.28)$$

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 der Projektion oder dem Skalarprodukt

$$\delta\Phi = \vec{A}(\vec{x}) \cdot \delta\vec{x} \quad (5.29)$$

zuordnen. Um 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,

$$\Phi = \int_{\vec{x}_a}^{\vec{x}_b} \vec{A}(\vec{x}) \cdot d\vec{x}. \quad (5.30)$$

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

$$\int_{\vec{x}_a}^{\vec{x}_b} \vec{A}(\vec{x}) \cdot d\vec{x} = \oint_C \vec{A}(\vec{x}) \cdot d\vec{x} \quad (5.31)$$

gekennzeichnet.

- Bei der praktischen Berechnung geht man analog zu (5.11) und (5.12) vor: Man stellt das gerichtete Wegelement $d\vec{x}$ mithilfe des Kurvenparameters t dar,

$$d\vec{x} = \frac{d\vec{x}}{dt} dt, \quad (5.32)$$

so dass das Kurvenintegral in das eindimensionale Integral

$$\int_{\vec{x}_a}^{\vec{x}_b} \vec{A}(\vec{x}) \cdot d\vec{x} = \int_{t_a}^{t_b} \vec{A}(\vec{x}) \cdot \frac{d\vec{x}}{dt} dt = \int_{t_a}^{t_b} \vec{A}(\vec{x}) \cdot \vec{v}(t) dt \quad (5.33)$$

übergeht.

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 Gradient

- Ein *Feld* f ist 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 vernachlässigen.

- Was wir hingegen benötigen ist die Ableitung von Feldern, deren Argument ein Vektor ist. Wir erinnern uns zuert daran, dass für Funktionen $f(x)$ die Ableitung nach x als

$$f'(x) = \frac{df(x)}{dx} := \lim_{\epsilon \rightarrow 0} \frac{f(x + \epsilon) - f(x)}{\epsilon} \quad (5.34)$$

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

$$\frac{\partial f(x_1, \dots, x_N)}{\partial x_i} := \lim_{\epsilon \rightarrow 0} \frac{f(x_1, \dots, x_i + \epsilon, \dots, x_N) - f(x_1, \dots, x_i, \dots, x_N)}{\epsilon}, \quad (5.35)$$

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

$$\frac{\partial f(x_1, \dots, x_N)}{\partial x_i} =: \partial_i f(x_1, \dots, x_N) =: f_{,i}(x_1, \dots, x_N) \quad (5.36)$$

die gebräuchlichsten sind.

- Zum Beispiel hat die Funktion

$$f(x_1, x_2, x_3) = \frac{\sin x_1 \cos x_2}{x_3} \quad (5.37)$$

die partiellen Ableitungen

$$\begin{aligned} \partial_1 f(x_1, x_2, x_3) &= \frac{\cos x_1 \cos x_2}{x_3} \\ \partial_2 f(x_1, x_2, x_3) &= -\frac{\sin x_1 \sin x_2}{x_3} \\ \partial_3 f(x_1, x_2, x_3) &= -\frac{\sin x_1 \cos x_2}{x_3^2}. \end{aligned} \quad (5.38)$$

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

$$\partial_i(fg) = (\partial_i f)g + f(\partial_i g) \quad (5.39)$$

und die Kettenregel

$$\partial_i(f(g)) = \frac{df}{dg} \partial_i g, \quad (5.40)$$

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* eines Feldes $f(\vec{x})$ als

$$\frac{df(\vec{x}(t))}{dt} := \frac{\partial f(\vec{x}(t))}{\partial x_i} \frac{dx_i(t)}{dt} = \partial_i f(\vec{x}(t)) v_i(t). \quad (5.41)$$

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 (5.37) längs der Kurve $\vec{x}(t) = (v_1 t, v_2 t, 1)$

$$\frac{df}{dt} = v_1 \partial_1 f + v_2 \partial_2 f = v_1 \cos(v_1 t) \cos(v_2 t) - \sin(v_1 t) \sin(v_2 t). \quad (5.42)$$

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

$$\vec{\nabla} f(\vec{x}) = \begin{pmatrix} \partial_1 f \\ \vdots \\ \partial_N f \end{pmatrix} \quad \text{bzw.} \quad (\vec{\nabla} f)_i = \partial_i f. \quad (5.43)$$

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

$$\vec{\nabla} = \vec{e}_1 \frac{\partial}{\partial x_1} + \dots + \vec{e}_N \frac{\partial}{\partial x_N} = \vec{e}_i \partial_i \quad (5.44)$$

heißt Nabla-Operator. Ausgedrückt durch den Nabla-Operator lautet die totale Ableitung (5.41) einer Funktion $f(\vec{x})$

$$\frac{df(\vec{x}(t))}{dt} = \vec{v} \cdot \vec{\nabla} f(\vec{x}(t)). \quad (5.45)$$

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} ,

$$\vec{e} \cdot \vec{\nabla} f(\vec{x}). \quad (5.46)$$

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

$$\vec{\nabla} \cdot \vec{f}(\vec{x}) = \partial_i f_i \tag{5.47}$$

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

$$\vec{\nabla} \times \vec{f}(\vec{x}) = \epsilon_{ijk} \partial_i f_j(\vec{x}) \vec{e}_k . \tag{5.48}$$

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:

$$\begin{aligned} \vec{\nabla}(fg) &= f\vec{\nabla}g + g\vec{\nabla}f \\ \vec{\nabla} \cdot (f\vec{v}) &= \vec{\nabla}f \cdot \vec{v} + f\vec{\nabla} \cdot \vec{v} \\ \vec{\nabla} \times (f\vec{v}) &= \vec{\nabla}f \times \vec{v} + f\vec{\nabla} \times \vec{v} \\ \vec{\nabla} \times (\vec{v} \times \vec{w}) &= (\vec{w} \cdot \vec{\nabla})\vec{v} + \vec{v}(\vec{\nabla} \cdot \vec{w}) - (\vec{v} \cdot \vec{\nabla})\vec{w} + \vec{w}(\vec{\nabla} \cdot \vec{v}) \\ \vec{\nabla} \times (\vec{\nabla} \times \vec{v}) &= \vec{\nabla}(\vec{\nabla} \cdot \vec{v}) - (\vec{\nabla} \cdot \vec{\nabla})\vec{v} . \end{aligned} \tag{5.49}$$

Der Operator

$$\vec{\nabla}^2 := \vec{\nabla} \cdot \vec{\nabla} = \partial_i \partial_i =: \Delta \tag{5.50}$$

heißt Laplace-Operator.

5.3 Energiesatz in drei Dimensionen

Nachdem wir nun in mehreren Dimensionen differenzieren und integrieren können, werden wir auch den eindimensionalen Energiesatz verallgemeinern.

5.3.1 Energieerhaltung bei Potentialkräften

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

$$m\ddot{\vec{x}} = \vec{F}$$

$$\Rightarrow m\dot{x}_i\dot{x}_i = m \frac{d}{dt} \frac{\dot{x}_i\dot{x}_i}{2} \equiv \frac{d}{dt} \left(\frac{m}{2} \dot{\vec{x}}^2 \right) = \vec{F} \cdot \frac{d\vec{x}}{dt} . \quad (5.51)$$

Analog zum eindimensionalen Fall ist

$$T = \frac{m}{2} \dot{\vec{x}}^2 = \frac{\vec{p}^2}{2m} \quad (5.52)$$

die *kinetische Energie*. Wenn wir (5.51) 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*,

$$d\left(\frac{m}{2} \dot{\vec{x}}^2\right) = \vec{F} \cdot d\vec{x} . \quad (5.53)$$

Die *Leistung* ist die pro Zeiteinheit verrichtete Arbeit,

$$\text{Leistung} = \frac{\text{Arbeit}}{\text{Zeiteinheit}} = \vec{F} \cdot \frac{d\vec{x}}{dt} . \quad (5.54)$$

- Um aus der Bewegungsgleichung auf eine Erhaltungsgröße zu schließen, müsste sich die rechte Seite in (5.51) als Zeitableitung einer anderen Funktion schreiben lassen. Das ist allgemein nicht der Fall, gilt aber für die Klasse der Potentialkräfte mit

$$\vec{F} = -\vec{\nabla}V(\vec{x}) . \quad (5.55)$$

Für solche Kräfte gilt nach der Definition der totalen Ableitung

$$\vec{F} \cdot \frac{d\vec{x}}{dt} = -\vec{\nabla}V \cdot \frac{d\vec{x}}{dt} = -\frac{dV(\vec{x})}{dt} . \quad (5.56)$$

Damit ist dann wieder

$$\frac{m}{2} \dot{\vec{v}}^2 + V(\vec{x}) = T + V = E = \text{konst.} , \quad (5.57)$$

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

$$V = \frac{k}{2} \vec{x}^2 \quad \Rightarrow \quad \vec{F} = -\vec{\nabla}V = -k\vec{x} . \quad (5.58)$$

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

$$\vec{F} = F(r) \vec{e}_r = F(r) \frac{\vec{x}}{r}. \quad (5.59)$$

Wir zeigen dass solche Kräfte immer ein Potential haben, indem wir den Ansatz

$$V(r) = - \int_{r_0}^r F(r') dr' \quad (5.60)$$

bestätigen. Wir erhalten damit für die Kraft in drei Dimensionen

$$\begin{aligned} F_i &= - \frac{\partial}{\partial x_i} V(r) \\ &= \frac{\partial}{\partial x_i} \int_{r_0}^r F(r') dr' = \frac{\partial r}{\partial x_i} \frac{d}{dr} \int_{r_0}^r F(r') dr' = F(r) \frac{x_i}{r}. \end{aligned} \quad (5.61)$$

Das ist genau die Kraft in (5.59).

5.4 Konservative Kraftfelder

5.4.1 Unabhängigkeit vom Weg

Bei der Herleitung des Energiesatzes haben wir zwei Dinge gesehen. Erstens ergibt die Integraldarstellung des Potentials in (5.27) 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, solche sogenannten konservative Kräfte immer wieder durch das Berechnen des Potentials zu erkennen.

- Wir betrachten sogenannte konservative Kräfte und ihre Potentiale in zwei Schritten. Zuerst zeigen wir, dass Wegunabhängigkeit äquivalent zur Existenz eines Potentials ist. Dann zeigen wir, dass ein Feld genau dann ein Potential hat, wenn seine Rotation verschwindet.
- Nehmen wir also zunächst an, das Kurvenintegral in (5.30)

$$\Phi(\vec{x}) = \int_{\vec{x}_a}^{\vec{x}} \vec{A}(\vec{x}') \cdot d\vec{x}' \quad (5.62)$$

sei unabhängig vom Weg. Wir betrachten es nun als Funktion des Endpunkts der Kurve und verlängern sie um ein infinitesimal kleines Stück $\delta\vec{x} \rightarrow d\vec{x}$ und erhalten wegen der angenommenen Unabhängigkeit vom Weg lediglich die totale Ableitung

$$\begin{aligned} \Phi(\vec{x} + \delta\vec{x}) - \Phi(\vec{x}) &= \vec{\nabla}\Phi(\vec{x}) \cdot d\vec{x} \\ &= \int_{\vec{x}}^{\vec{x}+\delta\vec{x}} \vec{A}(\vec{x}') \cdot d\vec{x}' = \vec{A}(\vec{x}) \cdot d\vec{x} \\ \Rightarrow \vec{A} &= \vec{\nabla}\Phi. \end{aligned} \tag{5.63}$$

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

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

$$\int_{\vec{x}_a}^{\vec{x}_b} \vec{A} \cdot d\vec{x} = \int_{\vec{x}_a}^{\vec{x}_b} \vec{\nabla}\Phi \cdot d\vec{x} = \int_{\Phi(\vec{x}_a)}^{\Phi(\vec{x}_b)} d\Phi = \Phi(\vec{x}_b) - \Phi(\vec{x}_a) \tag{5.64}$$

Es ist 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,

$$\vec{\nabla} \times \vec{A} = \vec{\nabla} \times \vec{\nabla}\Phi = \epsilon_{ijk} \partial_i \partial_j \Phi \vec{e}_k \equiv 0. \tag{5.65}$$

- In die andere Richtung definieren wir uns ein Potential Φ als Kurvenintegral über \vec{A} längs eines speziellen Weges, der stückweise parallel zu den drei Koordinatenachsen führt,

$$\Phi(\vec{x}) = \int_{x_a}^x A_x(x', y_a, z_a) dx' + \int_{y_a}^y A_y(x, y', z_a) dy' + \int_{z_a}^z A_z(x, y, z') dz', \tag{5.66}$$

und beweisen, dass sein Gradient \vec{A} ist, wenn $\vec{\nabla} \times \vec{A} = 0$ gilt. Die drei partiellen Ableitungen von Φ sind

$$\begin{aligned} \partial_z \Phi &= A_z(x, y, z) \\ \partial_y \Phi &= A_y(x, y, z_a) + \int_{z_a}^z \partial_y A_z(x, y, z') dz' \\ \partial_x \Phi &= A_x(x, y_a, z_a) + \int_{y_a}^y \partial_x A_y(x, y', z_a) dy' + \int_{z_a}^z \partial_x A_z(x, y, z') dz'. \end{aligned} \tag{5.67}$$

Unsere Annahme $\vec{\nabla} \times \vec{A} = 0$ gibt uns die drei Bedingungen

$$\partial_y A_z = \partial_z A_y \quad \partial_x A_y = \partial_y A_x \quad \partial_x A_z = \partial_z A_x, \quad (5.68)$$

und damit

$$\partial_y \Phi = A_y(x, y, z_a) + \int_{z_a}^z \partial_{z'} A_y(x, y, z') dz' = A_y(x, y, z) \quad (5.69)$$

und entsprechend $\partial_x \Phi = A_x(x, y, z)$. Damit haben wir $\vec{\nabla} \Phi = \vec{A}$ komponentenweise bewiesen und 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.
- Die Einschränkung auf zusammenhängende Gebiete illustrieren wir mit dem Feld

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

dessen Rotation verschwindet,

$$\vec{\nabla} \times \vec{A}(\vec{x}) = \begin{pmatrix} \partial_2 A_3 - \partial_3 A_2 \\ \partial_3 A_1 - \partial_1 A_3 \\ \partial_1 A_2 - \partial_2 A_1 \end{pmatrix} = 0. \quad (5.71)$$

Das Integral über \vec{A} 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

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

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.

- Wir haben also gezeigt, dass Kraftfelder \vec{A} , die auf einem einfach zusammenhängenden Gebiet G definiert sind, dann und nur dann ein Potential haben, wenn sie in G rotationsfrei sind, $\vec{\nabla} \times \vec{A} = 0$. 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 jetzt an, analog 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 wieder ein Vektorfeld \vec{A} und ein jetzt geschlossenes Kurvenintegral 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 analog zu (5.66) in der Form

$$\oint_C \vec{A} \cdot d\vec{x} = \int A_x(x_0, y_0) dx + \int A_y(x_0 + dx, y_0) dy \quad (5.74)$$

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

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

$$A_y(x_0 + dx, y_0) = A_y(x_0, y_0) + \partial_x A_y(x_0, y_0) dx, \quad (5.75)$$

bestimmen wir die Komponenten des Feldes \vec{A} an den Eckpunkten des Weges, vernachlässigen Terme der Ordnung dx^2 und dy^2 , und erhalten

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

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

Das Integral in (5.74) nähern wir auf diese Weise durch Flächeninhalte kleiner Rechtecke an,

$$\oint_C \vec{A} \cdot d\vec{x} = \int (\partial_x A_y - \partial_y A_x) dx dy = \int (\vec{\nabla} \times \vec{A})_z dx dy. \quad (5.76)$$

Das Kurvenintegral über die infinitesimal kleine, geschlossene Kurve ist 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 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. Wir geben jeder Zelle eine Richtung, die auf ihr senkrecht steht und so orientiert ist, dass sie mit der Orientierung ihrer Randkurve ein Rechtssystem bildet, also in (5.76) infinitesimal

$$\vec{e}_z dx dy =: d\vec{s}. \quad (5.77)$$

Das Ergebnis ist der *Stokes'sche Satz*

$$\oint_{\partial S} \vec{A} \cdot d\vec{x} = \int_S (\vec{\nabla} \times \vec{A}) \cdot d\vec{s}, \quad (5.78)$$

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 auch, dass das Kurvenintegral eines Vektorfeldes längs eines geschlossenen Weges genau dann verschwindet, wenn das Integral über seine Rotation innerhalb der eingeschlossenen Fläche verschwindet.

5.5 Krummlinig-orthogonale Koordinaten

- Wir haben bisher vor allem kartesische Koordinaten verwendet. Diese kann man auch dadurch kennzeichnen, dass man drei raumfeste Ebenen 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

$$\begin{aligned} ds^2 &= \sum_i dx_i^2 = \sum_{i,j} \left(\frac{\partial x_i}{\partial q_j} dq_j \right)^2 = \sum_{i,j} \left(\frac{\partial x_i}{\partial q_j} \right)^2 dq_j^2 \\ &= \sum_j dq_j^2 \sum_i \left(\frac{\partial x_i}{\partial q_j} \right)^2 =: \sum_j dq_j^2 h_j^2, \end{aligned} \quad (5.79)$$

wobei wir h_j durch

$$h_j^2 = \left(\frac{\partial x_1}{\partial q_j} \right)^2 + \left(\frac{\partial x_2}{\partial q_j} \right)^2 + \left(\frac{\partial x_3}{\partial q_j} \right)^2 \quad (5.80)$$

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

$$ds^2 = (h_1 dq_1)^2 + (h_2 dq_2)^2 + (h_3 dq_3)^2. \quad (5.81)$$

- 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, dann können wir durch einfache Analogie schreiben

$$(\vec{\nabla} f)_i = \frac{\partial f}{\partial (h_i dq_i)} = \frac{1}{h_i} \frac{\partial f}{\partial q_i}. \quad (5.82)$$

wobei diese Gleichung komponentenweise gilt und nicht die Summenkonvention nutzt.

5.5.1 Zylinderkoordinaten

- Betrachten wir konkret ein System mit Zylindersymmetrie. Wir bezeichnen die Zylinderachse willkürlich als \vec{e}_3 . Drei Koordinatenflächen sind durch konstante Parameter $\rho = (x_1^2 + x_2^2)^{1/2} \in [0, \infty)$, $\phi \in [0, 2\pi)$ und $z \in (-\infty, \infty)$ definiert. Wir wissen, dass man die kartesischen Koordinaten als

$$x_1 = \rho \cos \phi \quad x_2 = \rho \sin \phi \quad x_3 = z \quad (5.83)$$

schreiben kann.

- Die neue Basis von Einheitsnormalenvektoren \vec{e}_ρ , \vec{e}_ϕ und \vec{e}_z hängt mit den kartesischen Einheitsvektoren durch

$$\begin{aligned}\vec{e}_\rho &= \cos \phi \vec{e}_1 + \sin \phi \vec{e}_2 \\ \vec{e}_\phi &= \cos \phi \vec{e}_2 - \sin \phi \vec{e}_1 \\ \vec{e}_z &= \vec{e}_3\end{aligned}\quad (5.84)$$

zusammen. Nach (5.80) gilt

$$\begin{aligned}h_\rho^2 &= \cos^2 \phi + \sin^2 \phi = 1 \\ h_\phi^2 &= \rho^2 \sin^2 \phi + \rho^2 \cos^2 \phi = \rho^2 \\ h_z^2 &= 1 \\ \Rightarrow \quad ds^2 &= d\rho^2 + \rho^2 d\phi^2 + dz^2.\end{aligned}\quad (5.85)$$

Nach (5.82) ist der Gradient einer Funktion $f(\rho, \phi, z)$ in Zylinderkoordinaten

$$\begin{aligned}\vec{\nabla} f &= \frac{\vec{e}_\rho}{h_\rho} \partial_\rho f + \frac{\vec{e}_\phi}{h_\phi} \partial_\phi f + \frac{\vec{e}_z}{h_z} \partial_z f \\ &= \vec{e}_\rho \partial_\rho f + \frac{\vec{e}_\phi}{\rho} \partial_\phi f + \vec{e}_z \partial_z f.\end{aligned}\quad (5.86)$$

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

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = r \begin{pmatrix} \cos \phi \sin \theta \\ \sin \phi \sin \theta \\ \cos \theta \end{pmatrix}\quad (5.87)$$

sind

$$\begin{aligned}h_r^2 &= \cos^2 \phi \sin^2 \theta + \sin^2 \phi \sin^2 \theta + \cos^2 \theta = 1 \\ h_\phi^2 &= r^2 \sin^2 \theta (\sin^2 \phi + \cos^2 \phi) = r^2 \sin^2 \theta \\ h_\theta^2 &= r^2 \cos^2 \theta (\cos^2 \phi + \sin^2 \phi) + r^2 \sin^2 \theta = r^2 \\ \Rightarrow \quad ds^2 &= dr^2 + r^2 \sin^2 \theta d\phi^2 + r^2 d\theta^2,\end{aligned}\quad (5.88)$$

und der Gradient in sphärischen Polarkoordinaten lässt sich als

$$\vec{\nabla} f = \vec{e}_r \partial_r f + \frac{\vec{e}_\phi}{r \sin \theta} \partial_\phi f + \frac{\vec{e}_\theta}{r} \partial_\theta f \quad (5.89)$$

darstellen. Für eine radialsymmetrische Funktion $f(r)$, die nur vom Radius r abhängt, erhalten wir daraus das bekannte Ergebnis

$$\vec{\nabla} f(r) = \frac{df(r)}{dr} \vec{e}_r = f'(r) \vec{e}_r \quad (5.90)$$

auf sehr einfache Weise.

Kapitel 6

Stöße und Streuung

6.1 Systeme von Massenpunkten

Stöße oder Streuung von Teilchen sind die am weitesten verbreitete experimentelle Methode, um ein physikalisches System zu untersuchen, vor allen wenn wir in der Quantenmechanik lernen, dass auch die Bestrahlung eines Objektes mit Licht als Streuung mit Licht-Teilchen betrachtet werden kann. Um wechselwirkende Teilchen in Streuprozessen zu behandeln, beginnen wir mit den bekannten Erhaltungssätze für eine allgemeine Anzahl von Teilchen.

6.1.1 Impuls

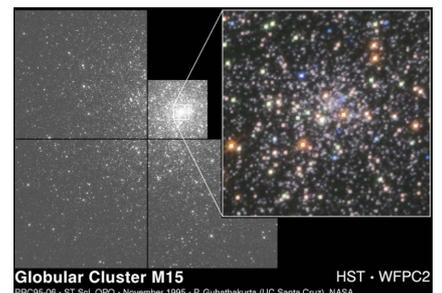
- Gegeben seien N Massenpunkte mit den Massen m_i an den Ortsvektoren \vec{x}_i , $1 \leq i \leq N$. Die Kraft des i -ten Massenpunkts auf den j -ten sei \vec{F}_{ij} , und zusätzlich wirke auf den i -ten Massenpunkt die äußere Kraft $\vec{F}_i^{(e)}$. Die Bewegungsgleichungen für das Teilchen i lauten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

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

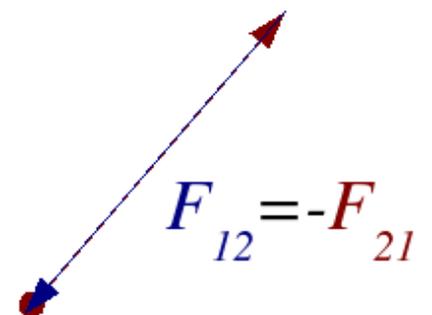
- Mithilfe der Gesamtmasse $M = \sum_i m_i$ definieren wir die *Schwerpunktkoordinate*

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

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 ($\vec{F}_{ji} = -\vec{F}_{ij}$)

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

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,

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

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

$$\sum_i m_i \vec{x}_i^* = \sum_i m_i \vec{x}_i - \sum_i m_i \vec{X} = \sum_i m_i \vec{x}_i - M\vec{X} = \vec{0}, \quad (6.5)$$

dann wird aus der Impulserhaltung

$$\sum_i \vec{p}_i^* = \sum_i m_i \dot{\vec{x}}_i^* = \sum_i m_i \dot{\vec{x}}_i - M\dot{\vec{X}} = \vec{0}. \quad (6.6)$$

6.1.2 Drehimpuls

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

$$\vec{F}_{ij} \parallel (\vec{x}_i - \vec{x}_j) \quad \text{oder} \quad \vec{F}_{ij} \times (\vec{x}_i - \vec{x}_j) = \vec{0}. \quad (6.7)$$

Der Gesamtdrehimpuls bezüglich des Koordinatenursprungs ist

$$\vec{L} = \sum_i \vec{L}_i = \sum_i (\vec{x}_i \times \vec{p}_i). \quad (6.8)$$

Er hat die Zeitableitung oder das Gesamtdrehmoment

$$\begin{aligned} \vec{M} &= \frac{d\vec{L}}{dt} = \sum_i (\dot{\vec{x}}_i \times \dot{\vec{p}}_i) = \sum_i (\dot{\vec{x}}_i \times m_i \ddot{\vec{x}}_i) \\ &= \sum_i \left(\dot{\vec{x}}_i \times \sum_{j \neq i} \vec{F}_{ji} \right) + \sum_i \dot{\vec{x}}_i \times \vec{F}_i^{(e)} \\ &= \sum_{j>i} (\dot{\vec{x}}_i \times \vec{F}_{ji} + \dot{\vec{x}}_j \times \vec{F}_{ij}) + \sum_i \dot{\vec{x}}_i \times \vec{F}_i^{(e)} \\ &= \sum_{j>i} [(\dot{\vec{x}}_i - \dot{\vec{x}}_j) \times \vec{F}_{ji}] + \sum_i \dot{\vec{x}}_i \times \vec{F}_i^{(e)} = \sum_i \dot{\vec{x}}_i \times \vec{F}_i^{(e)}. \end{aligned} \quad (6.9)$$

Unsere inneren Kräfte tragen auch zum Gesamtdrehmoment eines Systems von Massenpunkten *nicht* bei.

- Der Gesamtdrehimpuls im Schwerpunktsystem hat die Form

$$\begin{aligned}
 \vec{L} &= \sum_i m_i (\vec{x}_i \times \dot{\vec{x}}_i) = \sum_i m_i (\vec{X} + \vec{x}_i^*) \times (\dot{\vec{X}} + \dot{\vec{x}}_i^*) \\
 &= \vec{X} \times M\dot{\vec{X}} + \left(\vec{X} \times \sum_i \vec{p}_i^* \right) + \left(\sum_i m_i \vec{x}_i^* \times \dot{\vec{X}} \right) + \sum_i (\vec{x}_i^* \times \vec{p}_i^*) \\
 &= \vec{X} \times M\dot{\vec{X}} + \sum_i (\vec{x}_i^* \times \vec{p}_i^*) . \tag{6.10}
 \end{aligned}$$

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

6.1.3 Energie

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

$$\vec{F}_{ji} = -\vec{\nabla}_i V_{ji}(|\vec{x}_i - \vec{x}_j|) . \tag{6.11}$$

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,

$$\vec{F}_i^{(e)} = -\vec{\nabla}_i V_i^{(e)}(\vec{x}_i) . \tag{6.12}$$

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

$$\begin{aligned}
 m_i \ddot{\vec{x}}_i &= - \sum_i \vec{\nabla}_i \sum_{j \neq i} V_{ji}(|\vec{x}_i - \vec{x}_j|) - \sum_i \vec{\nabla}_i V_i^{(e)}(\vec{x}_i) \tag{6.13} \\
 \sum_i m_i \dot{\vec{x}}_i \cdot \ddot{\vec{x}}_i &= - \sum_i \dot{\vec{x}}_i \cdot \vec{\nabla}_i \sum_{j \neq i} V_{ji}(|\vec{x}_i - \vec{x}_j|) - \sum_i \dot{\vec{x}}_i \cdot \vec{\nabla}_i V_i^{(e)}(\vec{x}_i) .
 \end{aligned}$$

Wenn wir die Ortsableitung in eine Zeitableitung umformen,

$$\frac{d}{dt} \left[\sum_i \frac{m_i}{2} \dot{\vec{x}}_i^2 + \frac{1}{2} \sum_{j \neq i} V_{ji}(|\vec{x}_i - \vec{x}_j|) + \sum_i V_i^{(e)}(\vec{x}_i) \right] = 0 , \tag{6.14}$$

dann sorgt der Faktor 1/2 vor dem zweiten Term mit der Bedingung $V_{ij} = V_{ji}$ dafür, dass in der Kettenregel nur die Abhängigkeit von \vec{x}_i 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.

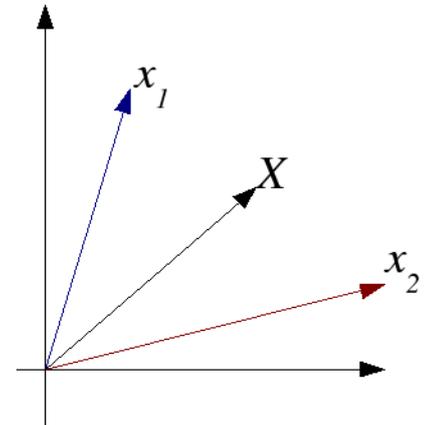
6.1.4 Reduzierte Masse

- Für Systeme mit zwei Teilchen führen wir zusätzlich zum Schwerpunkt die Relativkoordinate $\vec{r} = \vec{x}_1 - \vec{x}_2$ ein und erhalten

$$\begin{aligned}
 m_1 m_2 \ddot{\vec{r}} &= m_2 m_1 \ddot{\vec{x}}_1 - m_1 m_2 \ddot{\vec{x}}_2 \\
 &= -m_2 \vec{\nabla}_1 V(|\vec{r}|) + m_1 \vec{\nabla}_2 V(|\vec{r}|) \\
 &= -m_2 \frac{dV(|\vec{r}|)}{d|\vec{r}|} \frac{\vec{r}}{|\vec{r}|} + m_1 \frac{dV(|\vec{r}|)}{d|\vec{r}|} \frac{\vec{r}}{|\vec{r}|} \\
 &= -m_2 \frac{dV(|\vec{r}|)}{d|\vec{r}|} \frac{\vec{r}}{|\vec{r}|} - m_1 \frac{dV(|\vec{r}|)}{d|\vec{r}|} \frac{\vec{r}}{|\vec{r}|} \\
 &= -(m_1 + m_2) \frac{dV(|\vec{r}|)}{d|\vec{r}|} \frac{\vec{r}}{|\vec{r}|}, \tag{6.15}
 \end{aligned}$$

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

$$\mu = \frac{m_1 m_2}{m_1 + m_2} \Rightarrow \mu \ddot{\vec{r}} = -\vec{\nabla}_r V(|\vec{r}|). \tag{6.16}$$



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 im Schwerpunktsystem

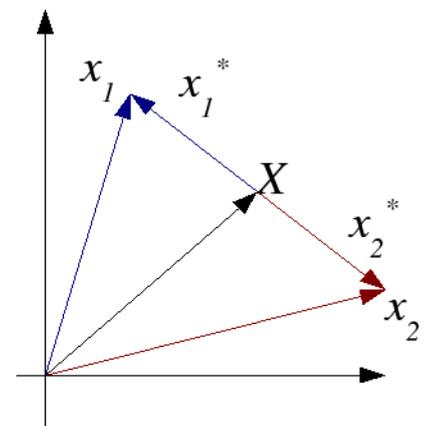
- Elastisch heißt ein Stoß zwischen zwei Teilchen, wenn in seinem Verlauf keine mechanische Energie in andere Energieformen (z.B. durch Verformung) umgewandelt wird. Wir können also von Impuls- und Energie-Erhaltung Gebrauch machen.
- Gegeben seien wieder zwei Teilchen mit den Massen m_1 und m_2 . Ihre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

$$m_1 \vec{v}_1 + m_2 \vec{v}_2 = m_1 \vec{v}'_1 + m_2 \vec{v}'_2, \tag{6.17}$$

während die Energieerhaltung

$$\frac{m_1}{2} \vec{v}_1^2 + \frac{m_2}{2} \vec{v}_2^2 = \frac{m_1}{2} \vec{v}'_1^2 + \frac{m_2}{2} \vec{v}'_2^2 \tag{6.18}$$

verlangt.



Transformation auf Schwerpunktkoordinaten

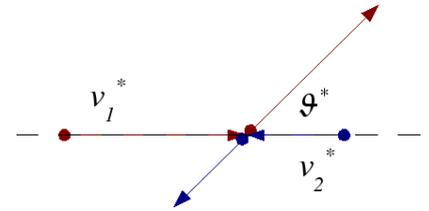
- Im Schwerpunktsystem lauten der Impulssatz

$$m_1 \vec{v}_1^* = -m_2 \vec{v}_2^* \quad \text{und} \quad m_1 \vec{v}_1'^* = -m_2 \vec{v}_2'^* , \quad (6.19)$$

und der Energiesatz

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

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.



Streuung im Schwerpunktsystem

- Setzt man den Impulssatz (6.19) in den Energiesatz (6.20) ein, um \vec{v}_2^* bzw. $\vec{v}_2'^*$ zu eliminieren, folgt

$$\begin{aligned} m_1 \vec{v}_1^{*2} + m_2 \left(\frac{m_1}{m_2} \vec{v}_1^* \right)^2 &= m_1 \vec{v}_1'^{*2} + m_2 \left(\frac{m_1}{m_2} \vec{v}_1'^* \right)^2 \\ \Rightarrow |\vec{v}_1^*| &= |\vec{v}_1'^*| \quad \Rightarrow |\vec{v}_2^*| = |\vec{v}_2'^*| . \end{aligned} \quad (6.21)$$

Die Beträge der Geschwindigkeiten vor und nach dem Stoß sind also für beide Massenpunkte gleich. Die Erhaltungssätze machen also über den Streuwinkel keine Aussage. Er hängt vom Kraftgesetz ab.

6.2.2 Laborsystem

- Wenn im *Laborsystem* der Massenpunkt 2 anfänglich ruht, $\vec{v}_2 = 0$, dann bewegt sich der Schwerpunkt als $M \dot{\vec{X}} = m_1 \vec{v}_1$. Die beiden Systeme sind also verknüpft durch

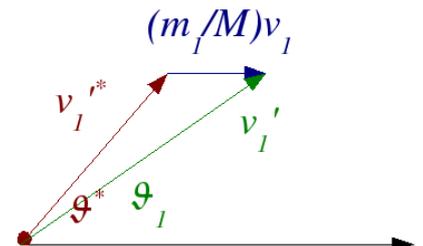
$$\vec{v}_1 = \vec{v}_1^* + \dot{\vec{X}} = \vec{v}_1^* + \frac{m_1}{M} \vec{v}_1 \quad \Leftrightarrow \quad \vec{v}_1^* = \frac{M - m_1}{M} \vec{v}_1 = \frac{m_2}{M} \vec{v}_1 . \quad (6.22)$$

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

$$\begin{aligned} v_1' \sin \vartheta_1 &= v_1'^* \sin \vartheta^* \\ v_1' \cos \vartheta_1 &= v_1'^* \cos \vartheta^* + \frac{m_1}{M} v_1 \\ \Rightarrow \tan \vartheta_1 &= \frac{v_1'^* \sin \vartheta^*}{v_1'^* \cos \vartheta^* + (m_1/M) v_1} . \end{aligned} \quad (6.23)$$

Mit (6.22) folgt für den Streuwinkels ϑ_1 im Laborsystem

$$\tan \vartheta_1 = \frac{\sin \vartheta^*}{\cos \vartheta^* + (m_1/m_2)} . \quad (6.24)$$



- Nach (6.24) sind im Grenzfall $m_1 \ll m_2$ die Winkel in beiden Systemen gleich, $\vartheta_1 \approx \vartheta^*$. Für $m_1 < m_2$ kann $\tan \vartheta_1$ alle Werte zwischen null und unendlich annehmen, also $0 \leq \vartheta_1 \leq \pi$. Für $m_1 > m_2$ kann der Nenner von $\tan \vartheta_1$ aber nicht mehr verschwinden. Wir können den maximalen Streuwinkel als Maximum von $\tan \vartheta_1$ als Funktion von ϑ^* berechnen,

$$\tan \vartheta_1 < \frac{m_2}{m_1} \Leftrightarrow \tan \vartheta_1^{\max} = \frac{m_2}{m_1}. \quad (6.25)$$

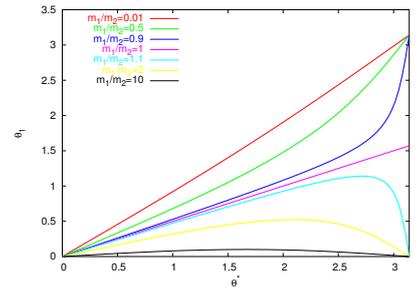
Schließlich sind für $m_1 = m_2$ auch $\vartheta_1 = \vartheta^*/2$ und $\vartheta_2 = \pi/2 - \vartheta_1$.

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

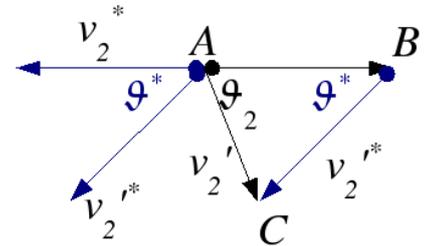
$$\vec{v}_2^* = -\dot{\vec{X}} = -\frac{m_1}{M} \vec{v}_1. \quad (6.26)$$

Da außerdem noch $|\vec{v}_2^*| = |\vec{v}_2|$ ist, muss das Dreieck, das aus $\dot{\vec{X}}$, \vec{v}_2^* und \vec{v}_2 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

$$\vartheta_2 = \frac{\pi - \vartheta^*}{2}. \quad (6.27)$$



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



zur Transformation des Streuwinkels

6.2.3 Energieübertrag

- Betrachten wir als Beispiel ein Neutron der Masse m , das an einem Kern der Masse Am streut, also $M = m(1 + A)$. Wie in der vorigen Betrachtung nehmen wir den Kern im Laborsystem als im Ursprung ruhend an ($\vec{v}_2 = 0$).
- Nach (6.22) ist $\dot{\vec{X}} = \vec{v}_1/(1 + A)$ und

$$\begin{aligned} \vec{v}_1^2 &= \left(\vec{v}_1^* + \dot{\vec{X}} \right)^2 = \left(\vec{v}_1^* + \frac{1}{1 + A} \vec{v}_1 \right)^2 \\ &= \vec{v}_1^{*2} + \frac{2}{1 + A} |\vec{v}_1^*| |\vec{v}_1| \cos \vartheta^* + \frac{\vec{v}_1^2}{(1 + A)^2}. \end{aligned} \quad (6.28)$$

Ferner gilt wegen der Energieerhaltung (6.21)

$$\vec{v}_1^{*2} = \vec{v}_1^2 = \left(\vec{v}_1 - \dot{\vec{X}} \right)^2 = \left(\vec{v}_1 - \frac{\vec{v}_1}{1 + A} \right)^2 = \vec{v}_1^2 \left(\frac{A}{1 + A} \right)^2. \quad (6.29)$$

Eingesetzt in (6.28) bedeutet dies

$$\begin{aligned} \vec{v}_1^2 &= \vec{v}_1^2 \frac{A^2}{(1 + A)^2} + \frac{2}{1 + A} \frac{A}{1 + A} \vec{v}_1^2 \cos \vartheta^* + \frac{\vec{v}_1^2}{(1 + A)^2} \\ &= \vec{v}_1^2 \frac{1 + A^2}{(1 + A)^2} + \frac{2}{1 + A} \frac{A}{1 + A} \vec{v}_1^2 \cos \vartheta^* \\ &= \vec{v}_1^2 \left[1 + \frac{2A}{(1 + A)^2} (\cos \vartheta^* - 1) \right]. \end{aligned} \quad (6.30)$$

- Für die relative Änderung der kinetischen Energie ergibt sich damit

$$\frac{T - T'}{T} = 1 - \frac{v_1'^2}{v_1^2} = \frac{2A}{(1+A)^2}(1 - \cos \vartheta^*). \quad (6.31)$$

Eine Mittelung über ϑ^* ergibt unter der Annahme, dass die Streuwinkel im Schwerpunktsystem gleichverteilt sind und mit (5.88)

$$\begin{aligned} \left\langle \frac{T - T'}{T} \right\rangle &= \frac{1}{4\pi} \int_0^{2\pi} d\varphi \int_0^\pi \frac{2A}{(1+A)^2} (1 - \cos \vartheta^*) \sin \vartheta^* d\vartheta^* \\ &= \frac{A}{(1+A)^2} \int_{-1}^1 (1 - c) dc = \frac{2A}{(1+A)^2}. \end{aligned} \quad (6.32)$$

Das ist der mittlere, relative Energieverlust des Neutrons. Er wird maximal, wenn $A = 1$ gilt. Ein Beispiel für diese Wahl des passenden Materials ist die Neutronenmoderation in Kernreaktoren mit Wasser oder Xenon-Detektoren für die Suche nach dunkler Materie.

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, $a_{ij} = \vec{e}'_i \cdot \vec{e}_j$.

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

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

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 $\vec{e}'_3 = \vec{e}_3$, 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 Skalarprodukte dieser Transformation

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

7.1.2 Matrizen und Matrixoperationen

- Die 3×3 Skalarprodukte a_{ij} , die bei der Darstellung von Koordinatentransformationen auftreten, 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,

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

- Außerdem führt man eine Matrixmultiplikation ein, durch die $L \times M$ -Matrizen mit $M \times N$ -Matrizen multipliziert werden und $L \times N$ -Matrizen 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 $i \leq L$, $j \leq M$ und $k \leq N$. Indem man N -dimensionale (Spalten-)Vektoren als $N \times 1$ -Matrizen auffasst, ist damit auch eine Multiplikation von $M \times N$ -Matrizen mit N -dimensionalen Spaltenvektoren definiert. Deren Ergebnis ist ein Vektor aus einem M -dimensionalen Vektorraum V_M ,

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

mit $i \leq M$ und $j \leq N$. 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

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

Analog kann man die Zeilen und Spalten der Matrix A vertauschen, wodurch die Matrix in ihre Transponierte A^T übergeht,

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

- Offensichtlich gilt

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

Ganz allgemein gilt für 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 ist 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 . Sie 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,

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

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 als Zahlen sogenannte Determinanten zugeordnet werden, die für viele auch physikalische Betrachtungen unverzichtbar sind. In ihrer Definition tauchen *Permutationen* von Indizes auf. Permutationen waren uns schon beim Levi-Civita-Symbol begegnet. Eine Permutation π bringt eine Reihe 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 Paare von Indizes vertauscht. Sie heißt *gerade* oder *ungerade*, wenn eine gerade oder ungerade Anzahl von Vertauschungen durchgeführt wird. Geraden Permutationen wird ein positives, ungeraden ein negatives Vorzeichen $\text{sign}(\pi)$ zugeordnet.
- N Indizes können auf $N! = 1 \cdot 2 \cdot \dots \cdot N$ Weisen angeordnet werden, es gibt also $N!$ Permutationen. Zum Beispiel haben die drei Zahlen 1, 2, 3 $3! = 6$ Permutationen, von denen

$$\begin{array}{ll} (1, 2, 3), & (2, 3, 1), & (3, 1, 2) & \text{gerade und} \\ (2, 1, 3), & (1, 3, 2), & (3, 2, 1) & \text{ungerade} \end{array} \quad (7.20)$$

sind. Damit hatten wir das Levi-Civita-Symbol definiert.

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

$$\det : \mathcal{M}_{NN} \rightarrow K, \quad A \mapsto \det A. \quad (7.21)$$

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 (7.20) und (7.22). Am einfachsten kann die Determinante in drei Dimensionen mithilfe des Levi-Civita-Symbols durch

$$\det A = \epsilon_{ijk} a_{1i} a_{2j} a_{3k} \quad (7.23)$$

dargestellt werden.

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

$$\det(\lambda A) = \lambda^N \det A \quad \text{und} \quad \det(A^T) = \det A . \quad (7.24)$$

Weiterhin gilt das Multiplikationstheorem

$$\det(A \cdot B) = \det A \det B \quad (7.25)$$

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

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

7.2.2 Orthonormale Transformationen

- Wir kehren noch einmal zu $N \times N$ -Matrizen für Koordinatentransformationen zurück. 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 ineinander überführen, gilt

$$\langle e'_i, e'_j \rangle = \langle a_{ik} e_k, a_{jl} e_l \rangle = a_{ik} a_{jl} \delta_{kl} = a_{ik} a_{jk} = a_{ik} a_{kj}^T = \delta_{ij} . \quad (7.26)$$

Die Transformationsmatrix A muss die *Orthonormalitätsrelation*

$$AA^T = I = A^T A \quad (7.27)$$

erfüllen, wobei I die *Einheitsmatrix* mit den Elementen $I_{ij} = \delta_{ij}$ ist. Die zweite Gleichung in (7.27) gilt wegen der Transpositionsregel (7.14).

- Offenbar ist $\det I = 1$, und deswegen gilt wegen (7.24) und (7.25)

$$\begin{aligned} 1 &= \det(AA^T) = \det A \det A^T = (\det A)^2 \\ \Rightarrow \det A &= \pm 1 . \end{aligned} \quad (7.28)$$

Die orthonormale Transformation heißt *eigentlich*, wenn $\det A = 1$ ist, anderenfalls *uneigentlich*.

- Ein Beispiel für eine uneigentliche orthonormale Transformationen ist eine Spiegelung im dreidimensionalen Raum an der y - z -Ebene,

$$A = \begin{pmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \Rightarrow \det A = -1. \quad (7.29)$$

Diese Spiegelung bildet die Einheitsvektoren \vec{e}_i entsprechend

$$\vec{e}_1 \rightarrow -\vec{e}_1, \quad \vec{e}_2 \rightarrow \vec{e}_2, \quad \vec{e}_3 \rightarrow \vec{e}_3 \quad (7.30)$$

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

$$\vec{e}_i \times \vec{e}_j = -\epsilon_{ijk} \vec{e}_k. \quad (7.31)$$

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

$$A^{-1}A = I \quad (7.32)$$

erfüllt. Wenn A^{-1} existiert, lassen sich ihre Elemente x_{ij} aus (7.32) eindeutig aus den $N \times N$ Gleichungen

$$x_{ij}a_{jk} = \delta_{ik} \quad (7.33)$$

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

$$x_{ji} = \frac{(-1)^{i-j} A_{ij}}{\det A} \quad (7.34)$$

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 (7.34), 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 ungeeignet, weil sie enormen Rechenaufwand erfordert. Wir geben hier die Formeln für reguläre 2×2 - und 3×3 -Matrizen an:

$$\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}^{-1} = \frac{1}{a_{11}a_{22} - a_{12}a_{21}} \begin{pmatrix} a_{22} & -a_{12} \\ -a_{21} & a_{11} \end{pmatrix} \quad (7.35)$$

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

Überzeugen Sie sich, dass diese Formeln stimmen!

- Wenden wir die Multiplikationsregel für Determinanten auf (7.32) an, folgt

$$\det(A^{-1}) = \frac{1}{\det A} = (\det A)^{-1}. \quad (7.36)$$

- Die Manipulation der Gleichung

$$\begin{aligned} AA^{-1} = I &\Leftrightarrow A^{-1}[AA^{-1}] = [A^{-1}A]A^{-1} = A^{-1} \\ &\Leftrightarrow [(A^{-1})^{-1}A^{-1}]AA^{-1} = (A^{-1})^{-1}[A^{-1}A]A^{-1} \\ &\Leftrightarrow AA^{-1} = (A^{-1})^{-1}A^{-1} \\ &\Leftrightarrow (A^{-1})^{-1} = A. \end{aligned} \quad (7.37)$$

d.h. Rechts- und Linksinverse sind gleich.

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

$$(AB)^{-1}(AB) = I \quad (7.38)$$

von rechts zuerst mit B^{-1} und dann mit A^{-1} , folgt

$$(AB)^{-1} = B^{-1}A^{-1}. \quad (7.39)$$

Außerdem folgt direkt aus der Cramerschen Regel, dass Transposition und Inversion vertauscht werden können,

$$(A^T)^{-1} = (A^{-1})^T. \quad (7.40)$$

- Für orthonormale Matrizen gilt aufgrund ihrer Definition (7.27)

$$A^T = A^{-1}. \quad (7.41)$$

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 Transformation des Drehimpulses

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 wir sehen werden dass bestimmte Symmetrien zur Folge haben, dass bestimmte physikalische Größen erhalten sind.

- 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 $\vec{a} = -\vec{a}'$ gilt

$$\vec{x}' = \vec{x} - \vec{a} \equiv \vec{x} + \vec{a}', \quad \dot{\vec{x}}' = \dot{\vec{x}}, \quad \ddot{\vec{x}}' = \ddot{\vec{x}}. \quad (7.42)$$

Damit wird die Bewegungsgleichung

$$\begin{aligned} m\ddot{\vec{x}} &= \vec{F}(\vec{x}, \dot{\vec{x}}, t) \\ \Leftrightarrow m\ddot{\vec{x}}' &= \vec{F}(\vec{x}' + \vec{a}, \dot{\vec{x}}', t) \equiv \vec{F}'(\vec{x}', \dot{\vec{x}}', t). \end{aligned} \quad (7.43)$$

Die Kraft transformiert sich nicht wie der Ort, sondern wie die Geschwindigkeit.

- Wenn die Bezugssysteme K und K' denselben Ursprung haben, aber gegeneinander verdreht sind, gilt

$$\vec{x}' = R \cdot \vec{x} \quad \Leftrightarrow \quad \vec{x} = R^{-1} \cdot \vec{x}', \quad (7.44)$$

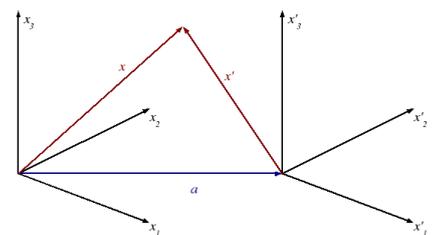
wobei R eine *orthogonalen* Drehmatrix ist, also $RR^T = I$. Wenn R zeitlich konstant ist, folgt

$$\begin{aligned} \dot{\vec{x}}' &= \frac{d}{dt}(R\vec{x}) = R\dot{\vec{x}} \\ \ddot{\vec{x}}' &= R\ddot{\vec{x}} \\ \Rightarrow m\ddot{\vec{x}}' &= R\vec{F}(R^{-1}\vec{x}', R^{-1}\dot{\vec{x}}', t) = \vec{F}'(\vec{x}', \dot{\vec{x}}', t). \end{aligned} \quad (7.45)$$

Die Kraft transformiert sich wieder wie die Beschleunigung oder die Geschwindigkeit.

- Der Drehimpuls im rotierten Koordinatensystem ist

$$L'_i = \epsilon_{ijk} x'_j p'_k = \epsilon_{ijk} (R_{jl} x_l) (R_{km} p_m) = (\epsilon_{ijk} R_{jl} R_{km}) x_l p_m \quad (7.46)$$



Übergang zu neuem Bezugssystem

Mit $RR^T = I$ lässt sich dies umschreiben,

$$\begin{aligned} \epsilon_{ijk}R_{jl}R_{km} &= \epsilon_{pjk}\delta_{pi}R_{jl}R_{km} = \epsilon_{pjk}(RR^T)_{pi}R_{jl}R_{km} \\ &= \epsilon_{pjk}R_{pq}R_{iq}R_{jl}R_{km} \\ &= R_{iq}\epsilon_{qlm} \det R . \end{aligned} \quad (7.47)$$

Im letzten Schritt benutzen wir eine Folgerung aus der Definition der Determinante, (7.23), nämlich

$$\epsilon_{pjk}R_{pq}R_{jl}R_{km} = \epsilon_{qlm} \det R = \epsilon_{qlm}\epsilon_{ijk}R_{1i}R_{2j}R_{3k} \quad (7.48)$$

für alle q, l, m . Für die Kombination 1, 2, 3 ist diese Identität offensichtlich, für die anderen Permutation kann man sie explizit bestätigen. Daraus ergibt sich für den transformierten Drehimpuls

$$L'_i = (\det R) R_{iq}(\epsilon_{qlm}x_l p_m) = (\det R) R_{iq}L_q = \pm R_{iq}L_q . \quad (7.49)$$

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 $v'_i = R_{ij}v_j$. Man unterscheidet allgemein

- *polare Vektoren*, für die trotz $\det R = \pm 1$ immer gilt

$$v'_i = R_{ij}v_j . \quad (7.50)$$

- *axiale Vektoren*, die bei uneigentlichen Transformationen gespiegelt werden,

$$\tilde{v}'_i = (\det R)R_{ij}\tilde{v}_j . \quad (7.51)$$

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. Für eine kombinierte Verschiebung und Drehung gilt

$$x'_i = a'_i + R_{ij}x_j \quad (7.52)$$

Für einen weiteren Punkt ist $y'_i = a'_i + R_{ij}y_j$, der Verbindungsvektor der beiden Punkte transformiert sich also wie

$$x'_i - y'_i = R_{ij}(x_i - y_i) . \quad (7.53)$$

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

Beschleunigte Bezugssysteme

8.1 Zeitabhängige Transformationen

Im letzten Kapitel haben wir konstante 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 ändern.

8.1.1 Winkelgeschwindigkeit

- Wenn sich zwei Koordinatensysteme relativ zueinander beliebig bewegen, sind Ortsvektoren in ihnen durch die zeitabhängige Version von (7.52)

$$\vec{x}'(t) = \vec{a}'(t) + R(t) \cdot \vec{x}(t) \quad (8.1)$$

miteinander verknüpft. 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 $R^T \dot{R}$ wechselt also unter Transposition ihr Vorzeichen,

$$(R^T \dot{R})_{ii} = 0 \quad \text{und} \quad (R^T \dot{R})_{ij} = -(R^T \dot{R})_{ji}. \quad (8.4)$$

Solche Matrizen heißen schiefsymmetrisch und haben in drei Dimensionen drei unabhängige Elemente,

$$R^T \dot{R} = \begin{pmatrix} 0 & -\omega_3 & +\omega_2 \\ \omega_3 & 0 & -\omega_1 \\ -\omega_2 & +\omega_1 & 0 \end{pmatrix} \quad (8.5)$$

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 $\vec{\omega}$ 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 \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 (8.6) gibt uns

$$\begin{aligned} -\epsilon_{ijk} \omega_k^* &= (R^{*T} \dot{R}^*)_{ij} = (S^T R^T \dot{R} S)_{ij} \\ &= S_{ia}^T (R^T \dot{R})_{ab} S_{bj} \\ &= -\omega_l \epsilon_{abl} S_{ai} S_{bj}. \end{aligned} \quad (8.11)$$

Analog zu (7.46) schreiben wir

$$\begin{aligned}\epsilon_{abl}S_{ai}S_{bj} &= \epsilon_{abc}\delta_{cl}S_{ai}S_{bj} = \epsilon_{abc}(SS^T)_{cl}S_{ai}S_{bj} \\ &= \epsilon_{abc}S_{ck}S_{lk}S_{ai}S_{bj} \\ &= S_{lk}\epsilon_{ijk} \det S ,\end{aligned}\quad (8.12)$$

und damit

$$\begin{aligned}\epsilon_{ijk}\omega_k^* &= \det S \epsilon_{ijk}S_{lk} \omega_l & \Leftrightarrow & \vec{\omega}^* = \det S S^T \vec{\omega} \\ & & \Leftrightarrow & \vec{\omega} = \det S S \vec{\omega}^* .\end{aligned}\quad (8.13)$$

Dies zeigt, dass der Vektor $\vec{\omega}$ ein *axialer* Vektor ist, der nach unserer vorigen Betrachtung in die Richtung der Drehachse zeigt.

- Setzen wir zum Schluss noch (8.6) in (8.2) ein, dann folgt

$$\dot{\vec{x}}' = \dot{\vec{a}}' + R(\dot{\vec{x}} + \vec{\omega} \times \vec{x}) . \quad (8.14)$$

oder für die Zeitableitung eines beliebigen Vektors \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 nur eine Drehung zwischen einem Zeitpunkt t und einem infinitesimal späteren Zeitpunkt $t + dt$, dann erhalten wir für die zeitabhängige Drehmatrix die linearisierte 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 vorher 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 Zeitintervall dt . Damit ist $\vec{\omega}$ die momentanen Winkelgeschwindigkeit, seine Richtung gibt die momentanen Drehachse an.

- Diese infinitesimalen Transformationen sind in der Physik sehr wichtig, z.B. in der Quantenmechanik oder der Feldtheorie. Deshalb betrachten wir sie noch etwas näher. Die infinitesimale Version der kompletten Transformation (8.1) ist

$$\vec{x}' = \vec{a}' + R\vec{x} = d\vec{a}' + (I + dR)\vec{x}. \tag{8.20}$$

Wegen der Orthonormalität ist

$$\begin{aligned} R^T R = I &\Leftrightarrow (I + dR)^T (I + dR) = I \\ &\Leftrightarrow dR^T + dR = 0. \end{aligned} \tag{8.21}$$

Wenn dR schiefsymmetrisch ist, kann es analog zu (8.6) als

$$dR_{ij} = -\epsilon_{ijk} d\varphi_k \tag{8.22}$$

dargestellt werden, wobei die $d\varphi_k$ infinitesimale Drehwinkel sind. Für infinitesimale Koordinatentransformationen gilt demnach allgemein

$$\vec{x}' = \vec{x} + d\vec{a}' + d\vec{\varphi} \times \vec{x}. \tag{8.23}$$

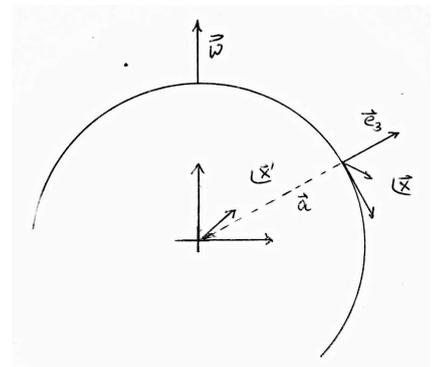
- Eine Transformation heisst *aktiv*, wenn das physikalische System sich in einem festen Koordinatensystem bewegt, und *passiv*, wenn das physikalische System fest bleibt, aber das Koordinatensystem bewegt wird. Mathematisch sind diese Arten der Transformation äquivalent, aber sie müssen physikalisch 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

$$\vec{x}'(t) = \vec{a}'(t) + R(t)\vec{x}(t) . \quad (8.24)$$

Aus Sicht des Oberflächensystems ist der Ortsvektor des Erdmittelpunkts \vec{a} konstant. Die Rotation um die Erdachse gibt dann $\vec{a}'(t) = R(t)\vec{a}$, und die einfachere Form

$$\vec{x}'(t) = R(t)\vec{a} + R(t)\vec{x}(t) = R(t)(\vec{a} + \vec{x}(t)) , \quad (8.25)$$

bekommt. Die Geschwindigkeit und die Beschleunigung im gestrichenen Inertialsystem sind nach (8.14) und (8.15)

$$\begin{aligned} \dot{\vec{x}}' &= R \left[\dot{\vec{x}} + \vec{\omega} \times (\vec{a} + \vec{x}) \right] \\ \ddot{\vec{x}}' &= \dot{R} \left[\dot{\vec{x}} + \vec{\omega} \times (\vec{a} + \vec{x}) \right] + R \left[\ddot{\vec{x}} + \dot{\vec{\omega}} \times (\vec{a} + \vec{x}) + \vec{\omega} \times \dot{\vec{x}} \right] . \end{aligned} \quad (8.26)$$

- Für Bewegungsgleichungen in ungestrichenen Oberflächensystem müssen wir nach $\ddot{\vec{x}}$ auflösen und multiplizieren zuerst von links mit $R^T = R^{-1}$,

$$R^T \ddot{\vec{x}}' = R^T \dot{R} \left[\dot{\vec{x}} + \vec{\omega} \times (\vec{a} + \vec{x}) \right] + \left[\ddot{\vec{x}} + \dot{\vec{\omega}} \times (\vec{a} + \vec{x}) + \vec{\omega} \times \dot{\vec{x}} \right] . \quad (8.27)$$

Den ersten Term schreiben wir in Komponenten, verwenden dabei das Ergebnis (8.6) und erhalten

$$\begin{aligned} (R^T \dot{R})_{ij} \left[\dot{\vec{x}} + \vec{\omega} \times (\vec{a} + \vec{x}) \right]_j &= -\epsilon_{ijk} \omega_k \left[\dot{x}_j + (\vec{\omega} \times (\vec{a} + \vec{x}))_j \right] \\ &= (\vec{\omega} \times \dot{\vec{x}})_i + [\vec{\omega} \times (\vec{\omega} \times (\vec{a} + \vec{x}))]_i . \end{aligned} \quad (8.28)$$

Der gesamte Ausdruck (8.27) wird dann

$$\begin{aligned} R^T \ddot{\vec{x}}' &= \vec{\omega} \times \dot{\vec{x}} + \vec{\omega} \times (\vec{\omega} \times (\vec{a} + \vec{x})) + \ddot{\vec{x}} + \dot{\vec{\omega}} \times (\vec{a} + \vec{x}) + \vec{\omega} \times \dot{\vec{x}} \\ &= \ddot{\vec{x}} + \dot{\vec{\omega}} \times (\vec{a} + \vec{x}) + 2\vec{\omega} \times \dot{\vec{x}} + \vec{\omega} \times (\vec{\omega} \times (\vec{a} + \vec{x})) . \end{aligned} \quad (8.29)$$

- Auf der linken Seite steht $\ddot{\vec{x}}'$ im Inertialsystem. Es gilt also das zweiten Newtonschen Axiom mit einer möglichen Potentialkraft, die unser eigentliches Physik-Problem definiert,

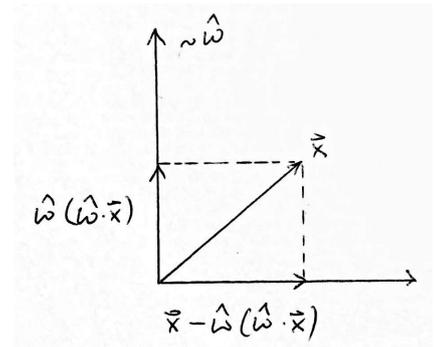
$$mR^T \ddot{\vec{x}}' = \vec{F}_{\text{ext}} = -\vec{\nabla}V . \quad (8.30)$$

Damit erhalten wir die Bewegungsgleichung im ungestrichenen Oberflächensystem

$$\begin{aligned} m\ddot{\vec{x}} &= -m \left[\dot{\vec{\omega}} \times (\vec{a} + \vec{x}) \right] - 2m\vec{\omega} \times \dot{\vec{x}} \\ &\quad - m \left[\vec{\omega} \times (\vec{\omega} \times \vec{x}) \right] - m \left[\vec{\omega} \times (\vec{\omega} \times \vec{a}) \right] - \vec{\nabla}V(\vec{x}) . \end{aligned} \quad (8.31)$$

Man kann diese vier weiteren Terme als *Scheinkräfte* auffassen. Sie 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



$$\begin{aligned} \vec{\omega} \times \vec{\omega} \times \vec{x} &= \epsilon_{ijk} \omega_j (\epsilon_{klm} \omega_l x_m) & (8.32) \\ &= (\delta_{il} \delta_{jm} - \delta_{im} \delta_{jl}) \omega_j \omega_l x_m = \omega_i \omega_j x_j - x_i \omega_j \omega_j, \end{aligned}$$

und erhalten

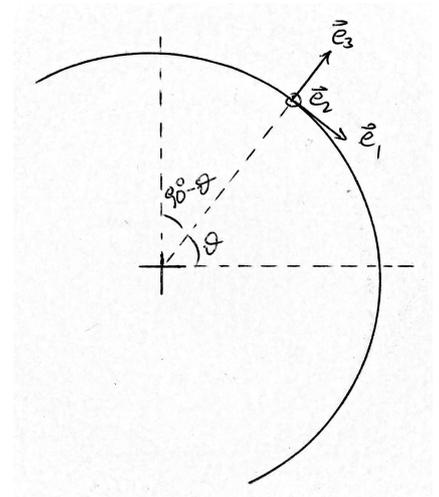
$$\vec{F}_Z = m [\vec{\omega}^2 \vec{x} - \vec{\omega}(\vec{\omega} \cdot \vec{x})] = m \omega^2 [\vec{x} - \hat{\omega}(\vec{x} \cdot \hat{\omega})] \quad (8.33)$$

wobei $\hat{\omega}$ der Einheitsvektor der Winkelgeschwindigkeit und damit der Term in Klammern der senkrechte Abstand des Massenpunkts von der Drehachse ist.

4. Der Term $-m [\vec{\omega} \times (\vec{\omega} \times \vec{a})]$ ist unabhängig von \vec{x} , aber proportional zu m . Er liefert einen Beitrag zur Erdbeschleunigung, $\vec{g} \rightarrow \vec{g} - \vec{\omega} \times (\vec{\omega} \times \vec{a})$. Diese beiden letzten Kräfte sind verwandt und heben sich für $\vec{x} = -\vec{a}$ auf.
- 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

$$\vec{F}_C = 2m(\dot{\vec{x}} \times \vec{\omega}_\perp) + 2m(\dot{\vec{x}} \times \vec{\omega}_\parallel) . \quad (8.35)$$

Für Bewegung auf der Erdoberfläche, $\dot{\vec{x}} \perp \vec{e}_3$, 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 usw. Der zweite Term steht senkrecht auf der Erdoberfläche.

Kapitel 9

Bewegung starrer Körper

9.1 Die Euler-Winkel

- Ein starrer Körper ist ein ausgedehntes Objekt, dessen Teile in fester räumlicher Beziehung zueinander stehen. Man kann sich starre Körper als aus vielen Massenpunkten aufgebaut denken, deren Abstände untereinander zeitlich konstant bleiben.
- Nehmen wir an, wir bräuchten eine große Zahl N von Massenpunkten, um einen starren Körper aufzubauen. Jeder dieser N Massenpunkte konnte sich vorher frei in jeder Richtung des Raums bewegen, weshalb wir drei Zahlen brauchten, um seine Lage im Raum festzulegen. Allgemein bezeichnet man die Anzahl unabhängiger Parameter, die zur vollständigen Charakterisierung der Lage des Systems notwendig sind, als die Anzahl der *Freiheitsgrade*. Zum Beispiel hat ein System von N Massenpunkten ohne weitere Nebenbedingungen $f = 3N$ Freiheitsgrade, nämlich die N Ortsvektoren \vec{x}_i mit ihren jeweils drei Komponenten.
- Ein *starrer Körper* hat nun $f = 6$ Freiheitsgrade, nämlich die Lage seines Schwerpunkts und drei Winkel, die seine Orientierung im Raum angeben. Wir beginnen unsere Untersuchung der Bewegung starrer Körper damit, diese drei Winkel eindeutig festzulegen.
- Gegeben sei dazu ein Inertialsystem \vec{e}'_i und ein körperfestes System \vec{e}_i , deren Ursprünge zusammen fallen. Zwischen den Koordinaten \vec{x}' im Inertialsystem und den Koordinaten \vec{x} im körperfesten System besteht die Verbindung

$$\vec{x}' = R \cdot \vec{x}, \quad (9.1)$$

wobei R eine (zeitabhängige) Drehmatrix ist. Da sie symmetrisch ist, sind von ihren neun Elementen sechs unabhängig. Diese wer-

den durch die drei *Euler-Winkel* $(\vartheta, \varphi, \psi)$ parametrisiert, die den Intervallen

$$0 \leq \vartheta \leq \pi, \quad 0 \leq \varphi \leq 2\pi, \quad 0 \leq \psi \leq 2\pi \quad (9.2)$$

entnommen werden. Damit haben wir für den starren Körper sechs *verallgemeinerte Koordinaten* gefunden, in denen sich seine Bewegung ausdrücken lässt, nämlich die drei Koordinaten des Schwerpunkts und die drei Euler-Winkel. Wir geben nun eine Konstruktion der Euler-Winkel an.

- Nach der Drehung R schneidet die \vec{e}_1 - \vec{e}_2 -Ebene die \vec{e}'_1 - \vec{e}'_2 -Ebene in der *Knotenlinie*, deren Richtung durch $\vec{e}'_3 \times \vec{e}_3$ gegeben ist. Die Drehmatrix R , ausgedrückt durch die Eulerwinkel, wird nun wie folgt konstruiert:

1. Zunächst wird um \vec{e}'_3 um den Winkel φ gedreht. Dadurch

$$\vec{x}' = D_3(\varphi) \cdot \vec{x}^*, \quad D_3(\varphi) = \begin{pmatrix} \cos \varphi & -\sin \varphi & 0 \\ \sin \varphi & \cos \varphi & 0 \\ 0 & 0 & 1 \end{pmatrix}. \quad (9.3)$$

2. Nun wird um \vec{e}_1^* um den Winkel ϑ gedreht. Dadurch entsteht ein Koordinatensystem \vec{e}_i^{**} mit

$$\vec{x}^* = D_1(\vartheta) \cdot \vec{x}^{**}, \quad D_1(\vartheta) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos \vartheta & -\sin \vartheta \\ 0 & \sin \vartheta & \cos \vartheta \end{pmatrix}. \quad (9.4)$$

3. Schließlich wird um \vec{e}_3^{**} um den Winkel ψ gedreht. Dadurch gelangt man in das körperfeste Koordinatensystem \vec{e}_i , und

$$\vec{x}^{**} = D_3(\psi) \cdot \vec{x}, \quad D_3(\psi) = \begin{pmatrix} \cos \psi & -\sin \psi & 0 \\ \sin \psi & \cos \psi & 0 \\ 0 & 0 & 1 \end{pmatrix}. \quad (9.5)$$

- Insgesamt ergibt sich daraus die Drehung

$$\vec{x}' = D_3(\varphi) \{D_1(\vartheta) [D_3(\psi) \cdot \vec{x}]\} =: R(\varphi, \vartheta, \psi) \cdot \vec{x} \quad (9.6)$$

mit der dazugehörigen gesamten Drehmatrix

$$R(\varphi, \vartheta, \psi) = D_3(\varphi)D_1(\vartheta)D_3(\psi). \quad (9.7)$$

- Die Drehmatrix $D_1(\vartheta)D_3(\psi)$ lautet zunächst

$$D_1(\vartheta)D_3(\psi) = \begin{pmatrix} \cos \psi & -\sin \psi & 0 \\ \cos \vartheta \sin \psi & \cos \vartheta \cos \psi & -\sin \vartheta \\ \sin \vartheta \sin \psi & \sin \vartheta \cos \psi & \cos \vartheta \end{pmatrix}, \quad (9.8)$$

und die volle Drehmatrix $R(\varphi, \vartheta, \psi)$ ist schließlich durch

$$R(\varphi, \vartheta, \psi) = \begin{pmatrix} \cos \varphi \cos \psi - \sin \varphi \cos \vartheta \sin \psi \\ \sin \varphi \cos \psi + \cos \varphi \cos \vartheta \sin \psi \\ \sin \vartheta \sin \psi \\ -\cos \varphi \sin \psi - \sin \varphi \cos \vartheta \cos \psi & \sin \varphi \sin \vartheta \\ -\sin \varphi \sin \psi + \cos \varphi \cos \vartheta \cos \psi & -\cos \varphi \sin \vartheta \\ \sin \vartheta \cos \psi & \cos \vartheta \end{pmatrix} \quad (9.9)$$

gegeben.

- Manchmal werden andere Konventionen für die Eulerwinkel benutzt, etwa indem man im zweiten Schritt nicht um die \vec{e}_1^* -, sondern um die \vec{e}_2^* -Achse dreht. Die Drehmatrix $R(\varphi, \vartheta, \psi)$ lautet dann natürlich anders, aber das Konstruktionsprinzip bleibt dasselbe.
- Der Konstruktion der Drehmatrix entsprechend kann der Vektor der momentanen Winkelgeschwindigkeit wie folgt durch die drei Eulerwinkel ausgedrückt werden:

$$\vec{\omega} = \dot{\varphi} \vec{e}_3^* + \dot{\vartheta} \vec{e}_1^{**} + \dot{\psi} \vec{e}_3. \quad (9.10)$$

Dabei sind die Einheitsvektoren längs der Koordinatenachsen bestimmt durch $\vec{e}_i' \cdot \vec{e}_j = R_{ij}$, also

$$\vec{e}_1^{**} \cdot \vec{e}_j = [D_3(\psi)]_{1j} = \begin{pmatrix} \cos \psi \\ -\sin \psi \\ 0 \end{pmatrix} \quad (9.11)$$

sowie

$$\vec{e}_3^* \cdot \vec{e}_j = [D_1(\vartheta)D_3(\psi)]_{3j} = \begin{pmatrix} \sin \vartheta \sin \psi \\ \sin \vartheta \cos \psi \\ \cos \vartheta \end{pmatrix}. \quad (9.12)$$

Die vektorielle Winkelgeschwindigkeit im körperfesten System lässt sich demnach auf

$$\vec{\omega} = \dot{\vartheta} \begin{pmatrix} \cos \psi \\ -\sin \psi \\ 0 \end{pmatrix} + \dot{\varphi} \begin{pmatrix} \sin \vartheta \sin \psi \\ \sin \vartheta \cos \psi \\ \cos \vartheta \end{pmatrix} + \dot{\psi} \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}. \quad (9.13)$$

- Ergänzend stellen wir noch fest, dass die Winkelgeschwindigkeit $\vec{\omega}$ nicht eine zeitliche Ableitung eines Vektors \vec{c} ist,

$$\vec{\omega} \neq \frac{d\vec{c}}{dt}. \quad (9.14)$$

Wenn das so wäre, folgte nämlich aus

$$\dot{\vec{c}} = \frac{\partial \vec{c}}{\partial \varphi} \dot{\varphi} + \frac{\partial \vec{c}}{\partial \vartheta} \dot{\vartheta} + \frac{\partial \vec{c}}{\partial \psi} \dot{\psi} \quad (9.15)$$

zusammen mit (9.10), dass

$$\frac{\partial c_1}{\partial \vartheta} = \cos \psi, \quad \frac{\partial c_1}{\partial \psi} = 0 \quad \Rightarrow \quad \frac{\partial^2 c_1}{\partial \psi \partial \vartheta} \neq \frac{\partial^2 c_1}{\partial \vartheta \partial \psi}, \quad (9.16)$$

gelten müsste, was auf einen Widerspruch führt.

9.2 Tensoren

In den vergangenen Kapiteln haben wir verschiedene Definitionen von Vektoren gegeben, die mehr oder weniger geometrisch oder algebraisch waren. Wir sollten uns also systematisch mit Objekten befassen, die wie Skalare oder Vektoren über ihre Transformationseigenschaften definiert werden.

9.2.1 Der Trägheitstensor

- Wir beginnen mit einem Beispiel. Die kinetische Energie eines starren Körpers bestehend aus den Massenpunkte m_i an den Orten \vec{x}'_i im Inertialsystem ist

$$T = \frac{1}{2} \sum_{i=1}^N m_i \dot{\vec{x}}_i'^2, \quad (9.17)$$

- Aus dem körperfesten System \vec{x} , definiert durch $\dot{\vec{x}}_i = 0$, erhalten wir das Inertialsystem nach (8.15) als $\dot{\vec{x}}_i' = R(t)(\dot{\vec{x}}_i + \vec{\omega} \times \vec{x}_i)$, also

$$\begin{aligned} T &= \frac{1}{2} \sum_{i=1}^N m_i (\vec{\omega} \times \vec{x}_i)^2 \\ &= \frac{1}{2} \sum_{i=1}^N m_i \epsilon_{ijk} \epsilon_{lab} \omega_j x_{i,k} \omega_a x_{i,b} \\ &= \frac{1}{2} \sum_{i=1}^N m_i (\omega_j \omega_k \delta_{jk} \vec{x}_i^2 - \omega_j \omega_k x_{i,j} x_{i,k}) \\ &= \frac{1}{2} \omega_j \left[\sum_{i=1}^N m_i (\vec{x}_i^2 \delta_{jk} - x_{i,j} x_{i,k}) \right] \omega_k =: \frac{1}{2} \omega_j \Theta_{jk} \omega_k. \end{aligned} \quad (9.18)$$

Θ ist eine Matrix, deren Eigenschaften wir untersuchen wollen.

9.2.2 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,

$$V \times V \rightarrow K, \quad (v_1, v_2) \mapsto \langle v_1, v_2 \rangle \in K. \quad (9.19)$$

Weiterhin hatten wir vom Skalarprodukt verlangt, dass es bilinear sei, d.h. linear in jedem seiner beiden Argumente.

- Entsprechend führen wir nun *Tensoren* als multilineare Abbildungen von V nach K ein, die n beliebigen Vektoren $v_i \in V$ eine Zahl aus K zuordnen,

$$T : \underbrace{V \times \dots \times V}_{n \text{ mal}} \rightarrow K, \quad (v_1, \dots, v_n) \mapsto T(v_1, \dots, v_n). \quad (9.20)$$

Die Anzahl n heißt *Rang*, man nennt den Tensor n -ter *Stufe*. Multilinearität bedeutet wieder

$$T(\dots, v_i + \lambda w_i, \dots) = T(\dots, v_i, \dots) + \lambda T(\dots, w_i, \dots). \quad (9.21)$$

Vektoren sind demnach Tensoren erster Stufe. Wir beschränken uns im Folgenden auf Tensoren zweiter Stufe, aber die folgenden Aussagen gelten entsprechend für Tensoren beliebiger Stufe.

9.2.3 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,

$$(T_1 + \lambda T_2)(v_1, v_2) = T_1(v_1, v_2) + \lambda T_2(v_1, v_2). \quad (9.22)$$

- Wie bei Vektoren reicht es wegen der Multilinearität aus, zu wissen, welche Zahlen der Tensor T den Basisvektoren e_i zuordnet. Wir bezeichnen sie als *Komponenten* T_{ij} des Tensors,

$$\begin{aligned} T(e_i, e_j) &=: T_{ij} \quad (1 \leq i, j \leq n) \\ \Rightarrow T(v, w) &= T(v_i e_i, w_j e_j) = v_i w_j T(e_i, e_j) =: v_i T_{ij} w_j. \end{aligned} \quad (9.23)$$

Tensoren zweiter Stufe lassen sich also durch Matrizen darstellen, sobald man eine Basis für den Vektorraum festgelegt hat.

- Damit können wir die Transformation eines Tensors $T \rightarrow T'$ unter einer orthogonalen Transformation $e'_i = R_{ik} e_k$ ableiten,

$$\begin{aligned} v_i T_{ij} w_j &= T(v, w) = T(v'_i e'_i, w'_j e'_j) \\ &= T(v'_i R_{ik} e_k, w'_j R_{jl} e_l) \\ &= v'_i R_{ik} T_{kl} R_{jl} w'_j \Leftrightarrow T'_{ij} = R_{ik} T_{kl} R_{lj}^T \end{aligned} \quad (9.24)$$

In Matrixschreibweise heisst das

$$T' = R T R^T = R T R^{-1}. \quad (9.25)$$

- Um festzustellen, ob das Objekt Θ in (9.18) 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,

$$\begin{aligned} \vec{x}^2 &= R_{jl}x_l^* R_{jm}x_m^* = R_{lj}^T R_{jm} x_l^* x_m^* = \delta_{lm}x_l^* x_m^* = \vec{x}^{*2} \\ \Rightarrow \Theta_{jk} &= \sum_{i=1}^N m_i \left(\vec{x}_i^{*2} \delta_{jk} - R_{jl}x_{i,l}^* R_{km}x_{i,m}^* \right) \\ &= \sum_{i=1}^N m_i \left(\vec{x}_i^{*2} R_{jl}R_{km}\delta_{lm} - R_{jl}R_{km} x_{i,l}^* x_{i,m}^* \right) \\ &= R_{jl}\Theta_{lm}^* R_{mk}^T. \end{aligned} \quad (9.26)$$

In der Tat ist Θ als ein *Tensor zweiter Stufe*. Er wird *Trägheitstensor* des starren Körpers genannt.

9.3 Diagonalisierung

9.3.1 Diagonalisierung von Matrizen

- Tensoren transformieren unter orthonormalen Koordinatendrehungen also wie $T' = RTR^T$. Dies gilt allgemein 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,

$$A'x' = y' = Ry = RAx = RAR^T x' \quad \Leftrightarrow \quad A' = RAR^T \quad (9.27)$$

Können wir diese Transformation benutzen, um eine Matrix A' in die besonders einfache *Diagonalform*

$$A' = \begin{pmatrix} A_1 & 0 & 0 \\ 0 & A_2 & 0 \\ 0 & 0 & A_3 \end{pmatrix} \quad (9.28)$$

zu bringen? Wenn wir (9.27) von links mit R^T multiplizieren, dann erhalten wir für diese Diagonalform in Komponenten

$$\sum_k R_{ik}^T A'_{kj} = \sum_k A_{ik} R_{kj}^T = A_i R_{ij}^T \quad (9.29)$$

Für jeden festen Wert von j ist dies ein lineares Gleichungssystem, in dem wir R_{ij}^T als Spaltenvektor r_i auffassen können, oder als Zeilenvektoren für die Matrix R . Wir können also (9.29) als

$$A\vec{r} = A_j\vec{r}, \quad (9.30)$$

schreiben. Diese *Eigenvektoren* \vec{r}_i definieren das ausgezeichnetes Koordinatensystem, in dem A Diagonalform annimmt. Es heißt *Hauptachsensystem* von A .

- Allgemein betrachten wir also ein lineares Gleichungssystem

$$Av = \lambda v \quad \Leftrightarrow \quad (A - \lambda I)v = \vec{0}. \quad (9.31)$$

Vektoren v , für die wir diese einfache Form erhalten, heißen *Eigenvektoren*, die zugehörigen Zahlen λ heißen *Eigenwerte*. Wenn ein Inverses existiert, dann folgt aus (9.31)

$$v = (A - \lambda I)^{-1} \vec{0} = \vec{0}. \quad (9.32)$$

Wenn (9.31) für nicht verschwindende v gelten soll, dann darf es also kein solches Inverses geben, und damit muss

$$\det(A - \lambda I) = 0 \quad (9.33)$$

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

$$A = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}. \quad (9.34)$$

Ihr charakteristisches Polynom lautet

$$\det(A - \lambda I) = \det \begin{pmatrix} -\lambda & 1 & 0 \\ 1 & -\lambda & 0 \\ 0 & 0 & -\lambda \end{pmatrix} = -\lambda(\lambda^2 - 1) = 0 \quad (9.35)$$

und hat die Lösungen $\lambda = -1, 0, 1$. Die Matrix A hat also die Diagonalform

$$A' = \begin{pmatrix} -1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} =: \text{diag}(-1, 0, 1). \quad (9.36)$$

- Die zugehörigen Eigenvektoren \vec{r} sind die Lösungen von

$$(A - \lambda I)\vec{r} = \begin{pmatrix} -\lambda & 1 & 0 \\ 1 & -\lambda & 0 \\ 0 & 0 & -\lambda \end{pmatrix} \begin{pmatrix} r_1 \\ r_2 \\ r_3 \end{pmatrix} = \begin{pmatrix} -\lambda r_1 + r_2 \\ r_1 - \lambda r_2 \\ -\lambda r_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}. \quad (9.37)$$

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 $\lambda = -1$

$$\vec{r}_1 = \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}. \quad (9.38)$$

Analoge Rechnungen führen auf die beiden anderen Eigenvektoren

$$\vec{r}_2 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}, \quad \vec{r}_3 = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}. \quad (9.39)$$

9.3.2 Diagonalisierung des Trägheitstensors

- Der Trägheitstensor Θ in (9.18) 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 der Länge l miteinander verbunden sind; m_2 befinde sich im Ursprung, trägt also nicht zum Trägheitstensor bei; m_1 laufe in der x_1 - x_2 -Ebene um. Dann sind die m_1 -Koordinaten $x_1 = l \cos \varphi$, $x_2 = l \sin \varphi$, $x_3 = 0$, und

$$\begin{aligned} \Theta &= m_1 \begin{pmatrix} l^2 - l^2 \cos^2 \varphi & -l^2 \sin \varphi \cos \varphi & 0 \\ -l^2 \sin \varphi \cos \varphi & l^2 - l^2 \sin^2 \varphi & 0 \\ 0 & 0 & l^2 \end{pmatrix} \\ &= m_1 l^2 \begin{pmatrix} \sin^2 \varphi & -\cos \varphi \sin \varphi & 0 \\ -\cos \varphi \sin \varphi & \cos^2 \varphi & 0 \\ 0 & 0 & 1 \end{pmatrix}. \end{aligned} \quad (9.40)$$

Seine Eigenwerte sind durch das charakteristische Polynom

$$\begin{aligned} \det(\Theta - \lambda) &= \begin{vmatrix} m_1 l^2 \sin^2 \varphi - \lambda & -m_1 l^2 \cos \varphi \sin \varphi & 0 \\ -m_1 l^2 \cos \varphi \sin \varphi & m_1 l^2 \cos^2 \varphi - \lambda & 0 \\ 0 & 0 & m_1 l^2 - \lambda \end{vmatrix} \\ &= (m_1 l^2 - \lambda) [(m_1 l^2 \sin^2 \varphi - \lambda)(m_1 l^2 \cos^2 \varphi - \lambda) \\ &\quad - m_1^2 l^4 \sin^2 \varphi \cos^2 \varphi] \\ &= (m_1 l^2 - \lambda) \lambda (\lambda - m_1 l^2) \end{aligned} \quad (9.41)$$

bestimmt. Die Lösungen sind $\lambda = m_1 l^2, 0$. Die Eigenvektoren zu diesen Eigenwerten sind

$$\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \quad \begin{pmatrix} -\sin \varphi \\ \cos \varphi \\ 0 \end{pmatrix} \quad \begin{pmatrix} \cos \varphi \\ \sin \varphi \\ 0 \end{pmatrix}. \quad (9.42)$$

Kapitel 10

Trägheitstensor

10.1 Volumenintegrale

10.1.1 Motivation

- Der Trägheitstensor eines starren Körpers ist nach (9.18) durch

$$\Theta_{jk} = \sum_{i=1}^N m_i [\vec{x}_i^2 \delta_{jk} - x_{i,j} x_{i,k}] \quad (10.1)$$

gegeben. 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

$$dm = \rho(\vec{x}) dV = \rho(\vec{x}) dx_1 dx_2 dx_3 \quad (10.2)$$

besitzen, wobei $\rho(\vec{x})$ die Dichte am Ort \vec{x} ist. Dann geht (10.1) in das *Volumenintegral*

$$\int dx_1 dx_2 dx_3 \rho(\vec{x}) [\vec{x}^2 \delta_{jk} - x_j x_k] \quad (10.3)$$

über, wobei die Integrationsgrenzen durch den Rand des Körpers vorgegeben sind.

- Allgemein haben 3-dimensionale Volumenintegrale die Form

$$I = \int dx_1 dx_2 dx_3 f(x_1, x_2, x_3) . \quad (10.4)$$

Hier ist f eine skalare Funktion, vektorwertige Funktionen werden einfach komponentenweise integriert.

- Oft ist die größte Schwierigkeit bei Volumenintegralen die Parametrisierung der Integrationsgrenzen. Um das zu illustrieren, berechnen wir das Volumen einer Kugel mit Radius R ,

$$f(\vec{x}) = \begin{cases} 1 & x_1^2 + x_2^2 + x_3^2 \leq R^2 \\ 0 & \text{sonst} \end{cases}$$

$$\Rightarrow V = \int_{x_1^2+x_2^2+x_3^2 \leq R^2} dx_1 dx_2 dx_3 . \quad (10.5)$$

Wenn x_1 und x_2 gegeben sind, dann gilt für x_3

$$V = \int dx_1 dx_2 \int_{-\sqrt{R^2-x_1^2-x_2^2}}^{\sqrt{R^2-x_1^2-x_2^2}} dx_3$$

$$= 2 \int dx_1 dx_2 \sqrt{R^2 - x_1^2 - x_2^2} . \quad (10.6)$$

Wenn nur noch x_1 vorgegeben ist, dann können wir x_2 -Integral mit Hilfe einer Formelsammlung lösen

$$V = 2 \int dx_1 \int_{-\sqrt{R^2-x_1^2}}^{\sqrt{R^2-x_1^2}} dx_2 \sqrt{R^2 - x_1^2 - x_2^2}$$

$$= \dots$$

$$= \pi \int dx_1 (R^2 - x_1^2) . \quad (10.7)$$

Schließlich ist das verbleibende Integral über x_1

$$V = \pi \int_{-R}^R dx_1 (R^2 - x_1^2) = \pi \left(R^2 x_1 - \frac{x_1^3}{3} \right)_{-R}^R = \frac{4\pi}{3} R^3 . \quad (10.8)$$

In dieser Rechnung sind offenbar die kartesischen Koordinaten der Symmetrie der Kugel nicht angepasst sind. Die Frage ist, wie Volumenintegrale in besser angepassten Koordinaten ausgeführt werden können.

10.1.2 Jacobi-Determinante

- Koordinaten, die der Symmetrie des Integrationsvolumens angepasst sind, sind zum Beispiel Zylinder- oder Kugelkoordinaten. Nehmen wir an, solche Koordinaten seien durch $u_i(x_j)$ gegeben, dann sind ihre totalen Ableitungen

$$du_i = (\partial_{x_j} u_i) dx_j \quad \text{oder} \quad d\vec{u} = \bar{J} d\vec{x} , \quad (10.9)$$

mit $\bar{J}_{ij} = \partial_{x_j} u_i$. Wenn die Matrix \bar{J} invertierbar ist, dann folgt

$$d\vec{x} = J d\vec{u} \quad \text{mit} \quad J_{ij} = \partial_{u_j} x_i \quad (10.10)$$

Die Matrix $J = \bar{J}^{-1}$ heißt *Jacobimatrix* der Koordinatentransformation.

- Da eine Rotation das Volumen erhält, definieren wir nun eine orthogonale Transformation R , die J diagonalisiert,

$$J' = \text{diag}(J_1, J_2, J_3)$$

$$\text{mit } d\vec{x}' = R d\vec{x} = R J d\vec{u} = (RJR^T) R d\vec{u} =: J' d\vec{u}' . \quad (10.11)$$

In der Diagonalform wird das Volumenintegral

$$\begin{aligned} dx'_1 dx'_2 dx'_3 &= |J_1 J_2 J_3| du'_1 du'_2 du'_3 = |\det J'| du'_1 du'_2 du'_3 \\ &= |\det J| du'_1 du'_2 du'_3 , \end{aligned} \quad (10.12)$$

wobei der Betrag gewährleistet, dass das Volumenelement positiv bleibt. Weil das Volumen erhalten ist, gilt in den originalen Koordinaten für (10.4)

$$\begin{aligned} dx_1 dx_2 dx_3 &= |\det J| du_1 du_2 du_3 \\ \Rightarrow \int f(\vec{x}) dx_1 dx_2 dx_3 &= \int f(\vec{x}(\vec{u})) |\det J| du_1 du_2 du_3 . \end{aligned} \quad (10.13)$$

Die Determinante $\det J$ heißt *Jacobideterminante* oder *Funktionaldeterminante*,

$$\det J = \det \begin{pmatrix} \frac{\partial x_i}{\partial u_j} \end{pmatrix} = \left| \frac{\partial x_i}{\partial u_j} \right| = \frac{\partial(x_1, x_2, x_3)}{\partial(u_1, u_2, u_3)} . \quad (10.14)$$

- Als Beispiel berechnen wir wieder das Volumen der Kugel in den Kugelkoordinaten (r, θ, ϕ) aus (5.87)

$$x_1 = r \sin \theta \cos \phi , \quad x_2 = r \sin \theta \sin \phi , \quad x_3 = r \cos \theta , \quad (10.15)$$

und mit der Jacobimatrix

$$J = \begin{pmatrix} \sin \theta \cos \phi & r \cos \theta \cos \phi & -r \sin \theta \sin \phi \\ \sin \theta \sin \phi & r \cos \theta \sin \phi & r \sin \theta \cos \phi \\ \cos \theta & -r \sin \theta & 0 \end{pmatrix} . \quad (10.16)$$

Ihre Determinante kann man berechnen und erhält

$$\det J = r^2 \sin \theta . \quad (10.17)$$

Das Volumenintegral über die Kugel mit Radius R ist dann

$$\begin{aligned} V &= \int_0^R dr r^2 \int_0^\pi d\theta \sin \theta \int_0^{2\pi} d\phi \\ &= \frac{R^3}{3} \times 2 \times 2\pi = \frac{4\pi}{3} R^3 . \end{aligned} \quad (10.18)$$

- Der Vollständigkeit halber geben wir auch die Jacobideterminante für die Transformation in Zylinderkoordinaten (ρ, ϕ, z) an,

$$x_1 = \rho \cos \phi , \quad x_2 = \rho \sin \phi , \quad x_3 = z . \quad (10.19)$$

Die Jacobimatrix und ihre Determinante sind

$$J = \begin{pmatrix} \cos \phi & -\rho \sin \phi & 0 \\ \sin \phi & \rho \cos \phi & 0 \\ 0 & 0 & 1 \end{pmatrix} \Rightarrow \det J = \rho . \quad (10.20)$$

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 (10.17) ist

$$\Theta_{ij} = \int_0^R dr \int_0^\pi d\theta \int_0^{2\pi} d\phi \rho [r^2 \delta_{ij} - x_i x_j] r^2 \sin \theta . \quad (10.21)$$

Wir können uns durch eine physikalische Überlegung viel Rechenarbeit sparen. Nach (9.26) ist Θ_{ij} ein Tensor und kann rotiert werden. Für eine Kugel kann keine Orientierung der drei Koordinatenachsen bevorzugt sein, weshalb die drei Koordinatenachsen bei beliebiger Orientierung Hauptachsen von Θ sein müssen. Alle Nicht-Diagonalelemente von Θ_{ij} müssen also verschwinden und die Diagonalelemente müssen gleich sein,

$$\Theta_{11} = \Theta_{22} = \Theta_{33} \quad \text{und} \quad \Theta_{12} = \Theta_{13} = \Theta_{23} = 0 . \quad (10.22)$$

- Wir berechnen am einfachsten

$$\begin{aligned} 3\Theta_{11} &= \Theta_{11} + \Theta_{22} + \Theta_{33} \\ &= \sum_{i=1}^3 \rho \int_0^R dr \int_0^\pi d\theta \int_0^{2\pi} d\phi (r^2 - x_i^2) r^2 \sin \theta \\ &= \rho \int_0^R dr \int_0^\pi d\theta \int_0^{2\pi} d\phi (3r^2 - r^2) r^2 \sin \theta \\ &= 2 \times 2\pi\rho \int_0^R dr r^4 \int_0^\pi d\theta \sin \theta = 8\pi\rho \frac{R^5}{5} . \end{aligned} \quad (10.23)$$

Mit $M = 4\pi/3 \times R^3 \times \rho$ folgt

$$\Theta_{11} = \Theta_{22} = \Theta_{33} = \frac{2}{5}MR^2 . \quad (10.24)$$

10.2 Drehimpuls und Trägheitsmomente

10.2.1 Drehimpuls

- Um den Drehimpuls des starren Körpers zu berechnen, setzen wir den Körper in ein gestrichenes Inertialsystem, das momentan mit dem körperfesten System zusammenfällt. Zu diesem Zeitpunkt ist die Drehmatrix zwischen beiden die Einheitsmatrix, $R = I$, und daher $\vec{x}_i = \vec{x}'_i$. Die Geschwindigkeiten der Massenpunkte sind hingegen nicht identisch, sondern wegen (8.15) $\vec{x}'_i = \vec{\omega} \times \vec{x}_i$. Also

ist der Drehimpuls im Inertialsystem

$$\begin{aligned}
 \vec{L}' &= \sum_{i=1}^N m_i (\vec{x}'_i \times \dot{\vec{x}}'_i) \\
 &= \sum_{i=1}^N m_i [\vec{x}_i \times (\vec{\omega} \times \vec{x}_i)] \\
 &= \sum_{i=1}^N m_i [\vec{x}_i^2 \vec{\omega} - \vec{x}_i (\vec{x}_i \cdot \vec{\omega})] \equiv \Theta \vec{\omega} .
 \end{aligned} \tag{10.25}$$

Der Trägheitstensor im Inertialsystem ist zeitabhängig und muss allgemein aus dem körperfesten System ins Inertialsystem gedreht werden, $\Theta' = R\Theta R^T$.

10.2.2 Trägheitsmomente

- Das *Trägheitsmoment* um eine Achse \hat{n} ist durch

$$\begin{aligned}
 \hat{n}^T \Theta \hat{n} &= \hat{n}_j \hat{n}_k \Theta_{jk} = \hat{n}_j \hat{n}_k \sum_{i=1}^N m_i [\vec{x}_i^2 \delta_{jk} - x_{i,j} x_{i,k}] \\
 &= \sum_{i=1}^N m_i [\vec{x}_i^2 - (\vec{x}_i \cdot \hat{n})^2]
 \end{aligned} \tag{10.26}$$

definiert, wobei der Term in Klammern nach Pythagoras der senkrechte Abstand des Massenpunktes i von der Drehachse ist.

- Die Diagonalelemente Θ_{ii} sind die Trägheitsmomente um die Koordinatenachsen. Die Θ_{jk} mit $j \neq k$ heißen *Deviationsmomente*. Die Eigenwerte von Θ heißen *Hauptträgheitsmomente*, ihre Eigenvektoren sind die *Hauptträgheitsachsen*.
- Die Bedingung $\vec{y}^T \Theta \vec{y} = y_j y_k \Theta_{jk} = 1$ definiert das *Trägheitsellipsoid*. Im System der Hauptträgheitsachsen ist offensichtlich, dass es sich um ein Ellipsoid handelt, denn

$$\Theta_1 y_1^2 + \Theta_2 y_2^2 + \Theta_3 y_3^2 = 1 . \tag{10.27}$$

- Für einen starren Körpers der Masse $M = \sum m_i$ können wir den Trägheitstensor um eine beliebige Achse in den Ortsvektor des

Schwerpunkts \vec{X} und die Schwerpunktkoordinaten \vec{x}_i^* aufspalten,

$$\begin{aligned}
 \Theta_{jk} &= \sum_{i=1}^N m_i \left(\vec{x}_i^2 \delta_{jk} - x_{i,j} x_{i,k} \right) \\
 &= \sum_{i=1}^N m_i \left[(\vec{x}_i^* + \vec{X})^2 \delta_{jk} - (x_{i,j}^* + X_j)(x_{i,k}^* + X_k) \right] \\
 &= \sum_{i=1}^N m_i \left[(\vec{x}_i^*)^2 \delta_{jk} - x_{i,j}^* x_{i,k}^* \right] + M \left(\vec{X}^2 \delta_{jk} - X_j X_k \right) \\
 &\quad + 2\delta_{jk} \vec{X} \cdot \sum_{i=1}^N m_i \vec{x}_i^* - X_j \sum_{i=1}^N m_i x_{i,k}^* - X_k \sum_{i=1}^N m_i x_{i,j}^* \\
 &= \sum_{i=1}^N m_i \left[(\vec{x}_i^*)^2 \delta_{jk} - x_{i,j}^* x_{i,k}^* \right] + M \left(\vec{X}^2 \delta_{jk} - X_j X_k \right) \\
 &\equiv \Theta_{jk}^* + M \left(\vec{X}^2 \delta_{jk} - X_j X_k \right) . \tag{10.28}
 \end{aligned}$$

Die drei Terme in der zweiten Zeile verschwinden aufgrund der Definition des Schwerpunkts. Damit ist das gesamte Trägheitsmoment um eine Achse \hat{n}

$$\hat{n}^T \Theta \hat{n} = \hat{n}^T \Theta^* \hat{n} + M \left[\vec{X}^2 - (\vec{X} \cdot \hat{n})^2 \right] . \tag{10.29}$$

Wie schon so oft ist der Term in Klammern der senkrechte Abstand des Schwerpunkts von der Drehachse.

10.3 Der Satz von Gauß

10.3.1 Herleitung

- Für 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,

$$\int_{\partial V} \vec{A} \cdot d\vec{\sigma} = \int_V (\vec{\nabla} \cdot \vec{A}) dV , \tag{10.30}$$

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.

- Wie beim Satz von Stokes bauen wir ein endlich großes Volumen aus infinitesimal kleinen Würfeln auf. 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.

- Für das infinitesimale würfelförmige Volumen $dV = dx_1 dx_2 dx_3$ besteht die Oberfläche aus sechs Quadraten, die sich paarweise gegenüberliegen. Die äußeren Flächennormalen sind 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 (10.30) ist

$$\begin{aligned} A_1 d\sigma_1 &= -A_1(x_1, x_2, x_3) dx_2 dx_3 + A_1(x_1 + dx_1, x_2, x_3) dx_2 dx_3 \\ &= -A_1(x_1, x_2, x_3) dx_2 dx_3 + A_1(x_1, x_2, x_3) dx_2 dx_3 \\ &\quad + \partial_{x_1} A_1(x_1, x_2, x_3) dx_1 dx_2 dx_3 \\ &= \partial_{x_1} A_1(x_1, x_2, x_3) dV . \end{aligned} \tag{10.31}$$

Alle drei Flächenpaaren ergeben zusammen

$$\begin{aligned} \lim_{dV \rightarrow 0} \int_{\partial V} \vec{A} \cdot d\vec{\sigma} &= \int (\partial_{x_1} A_1 + \partial_{x_2} A_2 + \partial_{x_3} A_3) dV \\ &= \int (\vec{\nabla} \cdot \vec{A}) dV . \end{aligned} \tag{10.32}$$

Um (10.30) zu erhalten heben sich alle internen Flächen auf der linken Seite auf, während die Divergenz auf der rechten Seite über das gesamte Volumen integriert wird.

10.3.2 Transformation von Divergenz und Rotation

- Die Integralsätze von Gauß und Stokes zeigen uns eine einfache Methode, um Divergenz und Rotation in krummlinig-orthogonale Koordinatensysteme zu übersetzen. Nach (5.81) ist bei einer solchen Transformation das Längenelement ds durch

$$ds^2 = (h_1 dq_1)^2 + (h_2 dq_2)^2 + (h_3 dq_3)^2 \tag{10.33}$$

gegeben ist, mit wie in (5.80)

$$h_j^2 = \left(\frac{\partial x_1}{\partial q_j} \right)^2 + \left(\frac{\partial x_2}{\partial q_j} \right)^2 + \left(\frac{\partial x_3}{\partial q_j} \right)^2 . \tag{10.34}$$

Analog ist das Volumenintegral in unserer üblichen Diagonalform

$$dV = h_1 h_2 h_3 dq_1 dq_2 dq_3 . \tag{10.35}$$

- Betrachten wir die Herleitung des Satzes von Gauß in krummlinig-orthogonalen Koordinaten, dann erhalten wir analog zu (10.31) für die Beiträge zweier längs der q_1 -Achse gegenüberliegender Begrenzungsflächen

$$\begin{aligned} &- (A_{q_1} h_2 h_3)(q_1, q_2, q_3) dq_2 dq_3 + (A_{q_1} h_2 h_3)(q_1 + dq_1, q_2, q_3) dq_2 dq_3 \\ &= \frac{\partial(A_{q_1} h_2 h_3)}{\partial q_1} dq_1 dq_2 dq_3 , \end{aligned} \tag{10.36}$$

wobei auch die h_i von den Koordinaten q_j abhängen. Das gesamte Flächenintegral wird

$$\left(\frac{\partial(A_{q_1} h_2 h_3)}{\partial q_1} + \frac{\partial(A_{q_2} h_1 h_3)}{\partial q_2} + \frac{\partial(A_{q_3} h_1 h_2)}{\partial q_3} \right) dq_1 dq_2 dq_3 . \quad (10.37)$$

Im Satz von Gauß muss dann in den neuen Koordinaten gelten gilt dann

$$\vec{\nabla} \cdot \vec{A} = \frac{1}{h_1 h_2 h_3} \left(\frac{\partial(A_{q_1} h_2 h_3)}{\partial q_1} + \frac{\partial(A_{q_2} h_1 h_3)}{\partial q_2} + \frac{\partial(A_{q_3} h_1 h_2)}{\partial q_3} \right) . \quad (10.38)$$

- Durch eine völlig analoge Betrachtung, ausgehend von (5.76) für die dritte Komponente der Rotation, erhält man die Darstellung der Rotation in krummlinig-orthogonalen Koordinaten,

$$\begin{aligned} (\vec{\nabla} \times \vec{A})_{q_1} &= \frac{1}{h_2 h_3} \left(\frac{\partial(A_{q_3} h_3)}{\partial q_2} - \frac{\partial(A_{q_2} h_2)}{\partial q_3} \right) , \\ (\vec{\nabla} \times \vec{A})_{q_2} &= \frac{1}{h_1 h_3} \left(\frac{\partial(A_{q_1} h_1)}{\partial q_3} - \frac{\partial(A_{q_3} h_3)}{\partial q_1} \right) , \\ (\vec{\nabla} \times \vec{A})_{q_3} &= \frac{1}{h_1 h_2} \left(\frac{\partial(A_{q_2} h_2)}{\partial q_1} - \frac{\partial(A_{q_1} h_1)}{\partial q_2} \right) . \end{aligned} \quad (10.39)$$

- Für Kugelkoordinaten hatten wir $(q_1, q_2, q_3) = (r, \theta, \phi)$ und

$$h_r = 1 , \quad h_\theta = r , \quad h_\phi = r \sin \theta . \quad (10.40)$$

Daher ist $h_r h_\theta h_\phi = r^2 \sin \theta$ und

$$\vec{\nabla} \cdot \vec{A} = \frac{1}{r^2} \frac{\partial(r^2 A_r)}{\partial r} + \frac{1}{r \sin \theta} \frac{\partial(\sin \theta A_\theta)}{\partial \theta} + \frac{1}{r \sin \theta} \frac{\partial A_\phi}{\partial \phi} . \quad (10.41)$$

Für die Komponenten der Rotation erhalten wir

$$\begin{aligned} (\vec{\nabla} \times \vec{A})_r &= \frac{1}{r \sin \theta} \frac{\partial(\sin \theta A_\phi)}{\partial \theta} - \frac{1}{r \sin \theta} \frac{\partial A_\theta}{\partial \phi} , \\ (\vec{\nabla} \times \vec{A})_\theta &= \frac{1}{r \sin \theta} \frac{\partial A_r}{\partial \phi} - \frac{1}{r} \frac{\partial(r A_\phi)}{\partial r} , \\ (\vec{\nabla} \times \vec{A})_\phi &= \frac{1}{r} \frac{\partial(r A_\theta)}{\partial r} - \frac{1}{r} \frac{\partial A_r}{\partial \theta} . \end{aligned} \quad (10.42)$$

- Betrachten wir den Satz von Gauß für das Radialfeld

$$\vec{F} = e^{-r} \vec{e}_r \quad (10.43)$$

und integrieren es über eine Kugelfläche mit Radius R um den Ursprung. Das gerichtete Flächenelement ist

$$d\vec{\sigma} = r^2 \sin \theta d\theta d\phi \vec{e}_r , \quad (10.44)$$

und damit das Flächenintegral über die Kugel­fläche bei $r = R$

$$R^2 \int_0^\pi \sin \theta d\theta \int_0^{2\pi} d\phi \vec{F}(r, \theta, \phi) \cdot \vec{e}_r = 4\pi R^2 e^{-R} \quad (10.45)$$

ist. Andererseits ist die Divergenz von \vec{F} wegen (10.41)

$$\begin{aligned} \vec{\nabla} \cdot \vec{F} &= \frac{1}{r^2} \frac{\partial(r^2 F_r)}{\partial r} = \frac{1}{r^2} \frac{\partial(r^2 e^{-r})}{\partial r} = \frac{2r - r^2}{r^2} e^{-r} \\ \Rightarrow 4\pi \int_0^R r^2 dr \vec{\nabla} \cdot \vec{F} &= 4\pi \int_0^R (2r - r^2) e^{-r} dr \\ &= 4\pi r^2 e^{-r} \Big|_0^R = 4\pi R^2 e^{-R}, \end{aligned} \quad (10.46)$$

genau wie es nach dem Satz von Gauß sein muss.