
Mathematische Ergänzungen
zur Theoretischen Physik 2

H. van Hees

17. August 2017



Inhaltsverzeichnis

Inhaltsverzeichnis	3
1 Variationsrechnung und kanonische Mechanik	7
1.1 Das Brachistochronenproblem	7
1.2 Die Euler-Lagrange-Gleichungen	8
1.2.1 Lösung des Brachistochronenproblems	9
1.3 Das Fundamentallemma der Variationsrechnung	10
1.4 Isoperimetrische Probleme	11
1.5 Das Hamiltonsche Prinzip	13
1.5.1 Die Lagrange-Funktion	13
1.5.2 Äquivalente Lagrange-Funktionen	16
1.5.3 Beispiel: Freier Fall bzw. schiefer Wurf	17
1.5.4 Beispiel: Harmonischer Oszillator	18
1.6 Das Noether-Theorem (Lagrange-Form)	20
1.6.1 Symmetrien	20
1.6.2 Raum-Zeit-Symmetrien	22
1.6.3 Zweiteilchensystem mit Wechselwirkungs-Zentralpotential	24
1.7 Beispiel: Kepler-Bewegung	26
1.8 Die Hamiltonsche kanonische Mechanik	28
1.8.1 Beispiel: Der harmonische Oszillator	30
1.9 Kanonische Transformationen	30
1.10 Die symplektische Struktur des Phasenraums und Poisson-Klammern	33
1.11 Beispiele zu kanonischen Transformationen	35
1.11.1 Beliebige Diffeomorphismen im Konfigurationsraum	36
1.11.2 Freier Fall	38
1.11.3 Eindimensionaler harmonischer Oszillator	39
1.12 Lie-Gruppen und -Algebren	40
1.13 Das Noether-Theorem (Hamilton-Form)	42
1.13.1 Die Galilei-Gruppe (Hamilton-Formulierung)	43
1.13.2 Beispiel: Hamilton-Funktion für freies Teilchen	46
1.13.3 Beispiel: Hamilton-Funktion für ein Zweiteilchensystem	47
1.13.4 Beispiel für dynamische Symmetrie: Isotroper harmonischer Oszillator	50
1.14 Die Symmetrien des Keplerproblems	57
1.15 Ausblick auf die Quantentheorie	60

2	Spezielle Relativitätstheorie	61
2.1	Das speziell-relativistische Raum-Zeit-Modell	61
2.2	Das Zwillingsparadoxon	68
2.3	Allgemeine Lorentz-Transformationen	69
2.4	Addition von Geschwindigkeiten	72
2.5	Relativgeschwindigkeit zweier Teilchen	73
2.6	Die Lorentz-Gruppe als Lie-Gruppe	74
2.7	Speziell relativistische Mechanik	78
2.8	Zerfall und Stöße von Teilchen	81
2.8.1	Zweiteilchenzerfall	81
2.8.2	Zweiteilchenstöße	83
2.8.3	Kinematik im Schwerpunktssystem	83
2.8.4	Kinematik im Laborsystem	84
2.9	Äquivalenz zwischen Energie und Masse	85
2.9.1	Voll inelastischer Stoß	85
2.9.2	Zerfall eines Pions in zwei Photonen	86
2.10	Geladene Teilchen im elektromagnetischen Feld	87
2.10.1	Teilchen im homogenen elektrischen Feld	88
2.10.2	Teilchen in einem homogenen Magnetfeld	90
2.10.3	Teilchen in parallelen homogenen elektrischen und magnetischen Feldern	91
2.10.4	Teilchen im gekreuzten homogenen elektromagnetischen Feld	93
2.11	Relativistische geladene Teilchen und das Wirkungsprinzip	94
2.11.1	(3+1)-Formalismus	94
2.11.2	Manifest kovariantes Hamilton-Prinzip	95
2.12	Tensoranalysis im Minkowski-Raum	97
2.13	Lorentz-Transformationen des elektromagnetischen Feldes	98
A	Matrix-Lie-Gruppen und -Lie-Algebren	103
A.1	Unitäre Vektorräume	103
A.2	Unitäre und selbstadjungierte Abbildungen	107
A.3	Spektralsatz für selbstadjungierte Operatoren	109
A.4	Frobenius-Skalarprodukt für lineare Abbildungen	111
A.5	Die Matrixexponentialfunktion	114
A.6	Grundlagen der Gruppentheorie	118
A.7	Die Gruppen SU(n) und ihre Lie-Algebren	119
	Literaturverzeichnis	121

Einleitung

Allgemeine Hinweise zur Vorlesung

Dies ist das Manuskript zur Vorlesung „Mathematische Ergänzungen zur Theoretischen Physik 2“. Ziel dieser Vorlesung ist es, die in der Vorlesung „Theoretische Physik 2“ benötigten mathematischen Methoden zu vertiefen und vor allem auf konkrete physikalische Probleme anzuwenden. Der Schwerpunkt liegt entsprechend weniger auf formalen Beweisen als vielmehr auf der Vermittlung der Rechentechnik, die sehr wichtig für das Verständnis der theoretischen Physik ist.

Inhaltlich ist die Vorlesung durch die physikalischen Anwendungen in der klassischen analytischen Mechanik festgelegt. Aus mathematischer Sicht betrachtet kommt im zweiten Semester als wesentliche Neuerung die **Variationsrechnung** und die Behandlung von **Lie-Symmetrien** mit ihren Anwendungen auf **Differentialgleichungen** hinzu.

Die Variationsrechnung beschäftigt sich mit dem Auffinden von Extrema von sogenannten Funktionalen, d.h. von Abbildungen von Funktionen auf reelle Zahlen. Ein klassisches Beispiel anhand dessen die Variationsrechnung vor allem von Leonard Euler im frühen 18. Jh. entwickelt wurde, ist das **Brachistochronenproblem**. Die Frage dabei ist, auf welche Kurve man ein Partikelchen, das sich im (als homogen genäherten) Schwerfeld der Erde reibungsfrei entlang dieser Kurve bewegt, führen muß, daß es in minimaler Zeit von einem Punkt zu einem anderen niedriger aber nicht senkrecht gelegenen Punkt gelangt. Eine andere Problemklasse sind die sog. **isoperimetrischen Probleme**. Dabei sucht man z.B. diejenige geschlossene ebene Kurve vorgegebener Länge, die eine möglichst große Fläche umspannt (**Problem der Dido**) oder die geschlossene Fläche vorgegebenen Oberflächeninhalts, die ein möglichst großes Volumen im dreidimensionalen Raum umschließt. Dies sind **Variationsprobleme mit Nebenbedingungen**.

Für die gesamte Physik ist das **Hamiltonsche Prinzip der kleinsten Wirkung** von zentraler Bedeutung, und es wird daher den Hauptaugenmerk in dieser Vorlesung einnehmen. Es handelt sich auf den ersten Blick einfach um eine Umformulierung der Bewegungsgleichungen (in unserem Fall vor allem die Newtonschen Bewegungsgleichungen) in Form eines Variationsprinzips. Es zeigt sich aber, daß es oft eine wesentlich einfachere Formulierung der Probleme der Mechanik erlaubt, indem man die Bewegungsgleichung sehr einfach in Form beliebiger **generalisierter Koordinaten** aufstellen und dann einfacher lösen kann als in den ursprünglichen kartesischen Koordinaten der Newtonschen Mechanik. Wir werden das Hamiltonsche Prinzip der kleinsten Wirkung mit und ohne Nebenbedingungen sowohl in der Lagrange- als auch der Hamilton-Formulierung besprechen.

Die letztere Form ist von besonderer Wichtigkeit, da sie eine besonders elegante Formulierung von **Symmetrieprinzipien** erlaubt. Dabei verstehen wir unter einer Symmetrie eine Transformation der Variablen (und evtl. der Zeit), unter der die Bewegungsgleichungen invariant sind. Z.B. ergibt sich oft eine Symmetrie aus den geometrischen Gegebenheiten eines mechanischen Problems. Ein wichtiges Beispiel ist die **Bewegung in einem Zentralpotential**, wo die auf einen Massenpunkt wirkende Kraft als Gradient eines Potentials darstellbar ist, das nur vom Abstand zu einem vorgegebenen Zentrum abhängt. Es ist schon anschaulich klar, daß die Bewegungsgleichungen unabhängig unter Drehungen der Koordinaten um dieses Zentrum sein werden und sich eine Behandlung des Problems in Kugelkoordinaten anbietet.

Eine weitere wichtige Symmetrie sind die **Symmetrien des Raum-Zeit-Modells**. Im Fall der Newtonschen Mechanik ist z.B. der Raum ein Euklidischer homogener isotroper Punktraum, und es gilt der Trägheitssatz, d.h. es gibt eine ausgezeichnete Klasse von Bezugssystemen, die **Inertialsysteme**, und die Bewegungsgleichungen sind unabhängig von der Wahl des Inertialsystems, d.h. man kann anhand physikalischer Beobachtungen nicht entscheiden, welche „absolute Geschwindigkeit“ das Inertialsystem besitzt, denn die Bewegungsgleichungen sehen in allen Inertialsystemen gleich aus. Die Menge aller Transformationen der Zeit und Raumkoordinaten, die diese Eigenschaften besitzen, bilden mathematisch eine algebraische Struktur, die man **Gruppe** nennt, die **Galilei-Gruppe**. Jede solche Transformation läßt sich aus **zeitlichen Translationen, räumlichen Translationen, Drehungen und Galilei-Boosts** zusammensetzen. Insgesamt hängt eine solche allgemeinste Transformation von 10 Parametern ab, und zu jeder dieser speziellen Transformationen, die jeweils nur von einem kontinuierlichen Parameter abhängen, läßt sich eine **Erhaltungsgröße** herleiten, d.h. eine Größe, die entlang der Trajektorie des mechanischen Systems zeitlich konstant ist. Umgekehrt entspricht jeder solchen Erhaltungsgröße ein Einparametersymmetriegruppe (**Noether-Theorem**). Im Fall der Galilei-Gruppe sind diese Erhaltungsgrößen **Energie (1 Erhaltungsgröße), Impuls (ein Vektor, also 3 Erhaltungsgrößen), Drehimpuls (ein Axialvektor, also 3 Erhaltungsgrößen) und die Schwerpunktschwindigkeit (ein Vektor, also 3 Erhaltungsgrößen)**.

Es zeigt sich nun, daß das Newtonsche Raumzeit-Modell die Natur nur näherungsweise beschreibt und nur auf Bewegungen anwendbar ist, in der die Geschwindigkeiten gegenüber einem gegebenen Inertialsystem klein gegen die **Lichtgeschwindigkeit** sind. Eine weit bessere Beschreibung von Raum und Zeit wird durch die **Spezielle Relativitätstheorie** gegeben. Wir werden die mathematischen Grundlagen dieser Theorie (**Minkowski-Raum**) genau besprechen und die allgemeinen analytischen Methoden der auch auf die speziell-relativistische Mechanik anwenden.

Falls es die Zeit erlaubt, werden wir auch noch die Anfangsgründe der **Allgemeinen Relativitätstheorie** besprechen, soweit sie die relativistische Bewegung von Teilchen in vorgegebenen Gravitationsfeldern betrifft.

Literaturempfehlungen: Die Literatur zum Thema „Mathematik für Physiker“ ist nahezu unerschöpflich. Einen sehr guten Überblick über die Standardmathematik, die alle Physiker beherrschen sollten, ist [AHK⁺15]. Von denselben Autoren ist daraus auch ein Lehrbuch für Mathematiker hervorgegangen, das sich zum Nachlesen einiger Beweise empfiehlt, die im vorgenannten Buch für Physiker und Ingenieure fehlen [ABH⁺13].

Der Klassiker zur Variationsrechnung ist [CH24]. Ansonsten findet sich die Variationsrechnung in jedem modernen Lehrbuch der theoretischen Physik, z.B. [BFK⁺15, LL97, Sch07, Nol13, Gre03, Joo89].

Zusätzlich zu diesem Skript, das sich entsprechend des Charakters der „mathematischen Ergänzungen“ eher auf die Rechentechniken konzentriert sei auch auf mein Manuskript zur klassischen Mechanik [Hee08] hingewiesen.

Kapitel 1

Variationsrechnung und kanonische Mechanik

In diesem Kapitel betrachten wir **Extremalprobleme**, bei denen man Funktionen sucht, die eine bestimmte reelle Größe minimal oder maximal machen. Wir gehen dabei vom **Brachistochronenproblem** als typischem Beispiel aus und entwickeln die allgemeinen mathematischen Methoden zur Lösung solcher Probleme. Nach einer kurzen Betrachtung der **isoperimetrischen Probleme** als weiterem Beispiel wenden wir uns der Anwendung in der **Mechanik**, also dem **Hamiltonschen Prinzip der kleinsten Wirkung**, das man mit einigem Recht als das umfassendste mathematische Prinzip für die gesamte moderne Physik bezeichnen kann.

1.1 Das Brachistochronenproblem

Das **Brachistochronenproblem** steht am Anfang der Entwicklung der Variationsrechnung. Der schwierige Name leitet sich aus dem Griechischen her (brachistos=kürzest, chronos=Zeit). Es handelt sich um das Problem, diejenige Kurve zu finden, die zwei vorgegebene nicht vertikal übereinander gelegene Punkte verbindet, entlang derer man ein Partikelchen im homogenen Schwerfeld der Ebene reibungsfrei gleiten lassen muß, so daß es in möglichst kurzer Zeit von einem zum anderen Punkt gelangt.

Dazu müssen wir uns zunächst überlegen, wie wir die Bewegung entlang einer beliebigen Kurve $y = f(x)$ in der (x, y) -Ebene beschreiben können. Wir wählen die y -Achse nach unten, so daß die Schwerkraft auf das Teilchen durch $\vec{F} = mg\vec{e}_y$ gegeben ist. Diese Kraft läßt sich offenbar als Gradient des Potentials

$$V(\vec{x}) = -mgy \Rightarrow \vec{F} = -\vec{\nabla}V = mg\vec{e}_y \quad (1.1.1)$$

schreiben. Demnach wird der Energiesatz gelten

$$E = T + V = \frac{m}{2}\dot{\vec{x}}^2 + V(\vec{x}) = \text{const.} \quad (1.1.2)$$

Nehmen wir nun an, daß sich der Anfangspunkt bei $(x, y) = (0, 0)$ befindet, und das Teilchen aus der Ruhe losläuft, folgt

$$\frac{m}{2}\dot{\vec{x}}^2 - mgy = 0. \quad (1.1.3)$$

Die Geschwindigkeit ergibt sich aus

$$\begin{pmatrix} \dot{x} \\ \dot{y} \end{pmatrix} = \begin{pmatrix} \dot{x} \\ f'(x)\dot{x} \end{pmatrix} \Rightarrow \dot{\vec{x}} = \dot{x} \begin{pmatrix} 1 \\ f'(x) \end{pmatrix}, \quad (1.1.4)$$

wobei $f'(x)$ die Ableitung der Funktion f nach x bedeutet. Ein Punkt über einem Symbol bedeutet stets die Ableitung nach der Zeit. Setzen wir dies in (1.1.3) ein folgt

$$\frac{m}{2}\dot{x}^2[1 + f'^2(x)] + mgf(x) = 0. \quad (1.1.5)$$

Dies können wir nun nach $dt/dx = 1/\dot{x}$ auflösen:

$$\frac{dt}{dx} = \sqrt{\frac{1+f'^2(x)}{2gf(x)}}. \quad (1.1.6)$$

Bewegt sich also das Teilchen von $(x, y) = 0$ zu $(x_0, y_0 = f(x_0))$, ist die Gesamtdauer der Bewegung

$$A[f] = \int_0^{x_0} dx \sqrt{\frac{1+f'^2(x)}{2gf(x)}}. \quad (1.1.7)$$

Es muß natürlich $y = f(x) > 0$ sein, damit das Integral reell ist.

Dies ist ein sog. **Funktional**, d.h. eine Abbildung, die einer Funktion f eine reelle Zahl zuordnet. In unserem Beispiel ordnet sie jeder Kurve, die die vorgegebenen Punkte $P_1 = (0, 0)$ und $P_2 = (x_0, y_0 = f(x_0))$ verbindet, die von einem reibungsfrei auf der Kurve im Schwerfeld der Erde gleitenden Teilchen benötigte Zeit, um vom Punkt P_1 zum Punkt P_2 zu gelangen.

Die Aufgabe ist es nun, diejenige Kurve zu finden, die (1.1.7) möglichst klein macht.

1.2 Die Euler-Lagrange-Gleichungen

Wir schreiben das Funktional zunächst in allgemeinerer Form hin. Es sei $L(f, f')$ eine Funktion von zwei reellen Parametern, die **Lagrange-Funktion** des Variationsproblems. Dann sei das Funktional

$$A[f] = \int_{x_1}^{x_2} dx L[f(x), f'(x)] \quad (1.2.1)$$

gegeben, wobei $f'(x)$ hier die Ableitung von f bedeutet.

Wir suchen nun diejenige Funktion f , die das Funktional $A[f]$ extremal macht, wobei $y_1 = f(x_1)$ und $y_2 = f(x_2)$ fest vorgegebene Werte sein sollen.

Dazu erinnern wir uns, daß eine differenzierbare Funktion $g(\eta)$ bei $\eta = 0$ nur dann extremal werden kann (also ein lokales Minimum oder Maximum besitzen kann), wenn $g'(0) = 0$ ist. Wir definieren diese Funktion als

$$g(\eta) = A[f + \eta \delta f]. \quad (1.2.2)$$

Dabei ist δf eine beliebige Funktion mit $\delta f(x_1) = \delta f(x_2) = 0$. Wir variieren also f um die gesuchte Funktion f , wobei wir die Endpunkte festhalten (vgl. die nebenstehende Abbildung). Wir sagen auch, wir **variieren** die Funktion f um die das Funktional $A[f]$ extremierende Funktion. Daher rührt der Name **Variationsrechnung**.

Die notwendige Bedingung für das Vorliegen eines Extremums ist nun bei $\eta = 0$

$$g'(\eta)|_{\eta=0} = \int_{x_1}^{x_2} dx \frac{d}{d\eta} L[f(x) + \eta \delta f(x), f'(x) + \eta \delta f'(x)] \Big|_{\eta=0} \stackrel{!}{=} 0. \quad (1.2.3)$$

Nach der Kettenregel für die Ableitung von Funktionen mehrerer Veränderlicher gilt nun

$$g'(\eta)|_{\eta=0} = \int_{x_1}^{x_2} dx \left[\delta f(x) \frac{\partial L[f(x), f'(x)]}{\partial f} + \delta f'(x) \frac{\partial L[f(x), f'(x)]}{\partial f'} \right] \stackrel{!}{=} 0. \quad (1.2.4)$$

Wir wenden nun auf den zweiten Term die partielle Integration an. Wegen $\delta f(x_1) = \delta f(x_2) = 0$ ergibt sich dann

$$g'(\eta)|_{\eta=0} = \int_{x_1}^{x_2} dx \delta f(x) \left[\frac{\partial L[f(x), f'(x)]}{\partial f} - \frac{d}{dx} \left(\frac{\partial L[f(x), f'(x)]}{\partial f'} \right) \right] \stackrel{!}{=} 0. \quad (1.2.5)$$

Diese Bedingung muß nun für beliebige Funktionen $\delta f(x)$ mit $\delta f(x_1) = \delta f(x_2) = 0$ gelten. Offenbar ist das der Fall, wenn die eckige Klammer unter dem Integral verschwindet. Diese notieren wir in der etwas bequemeren abgekürzten Schreibweise

$$\frac{\partial L}{\partial f} - \frac{d}{dx} \frac{\partial L}{\partial f'} = 0. \quad (1.2.6)$$

Dies sind die **Euler-Lagrange-Gleichungen** des Variationsproblems.

Im nächsten Abschnitt zeigen wir, daß tatsächlich diese Schlußweise sogar streng gilt, d.h. (1.2.5) kann nur dann für alle $\delta f(x)$ gelten, wenn die Euler-Lagrange-Gleichungen erfüllt sind.

Damit haben wir das Variationsproblem auf die Aufgabe, die Euler-Lagrange-Gleichungen zu lösen, zurückgeführt. Diese sind offensichtlich gewöhnliche Differentialgleichungen 2. Ordnung, und dies ist auch genau die Form der Bewegungsgleichungen in der Newtonschen Mechanik (mit der Zeit als unabhängige und die Koordinaten der Trajektorien der Teilchen als abhängige Variablen).

Wir bemerken nur gleich noch eine weitere Folgerung aus (1.2.6). Dazu berechnen wir die totale Ableitung von $L[f(x), f'(x)]$ nach x . Wir lassen der Bequemlichkeit halber wieder die Argumente von L weg und erhalten dann aus der Kettenregel

$$\frac{d}{dx} L = f'(x) \frac{\partial L}{\partial f} + f''(x) \frac{\partial L}{\partial f'}. \quad (1.2.7)$$

Für die Lösung des Variationsproblems gilt nun die Euler-Lagrange-Gleichung, d.h. wir können $\partial L / \partial f$ in (1.2.7) Hilfe dieser Gleichung eliminieren. Wir erhalten dann

$$\frac{d}{dx} L = f'(x) \frac{d}{dx} \frac{\partial L}{\partial f'} + f''(x) \frac{\partial L}{\partial f'} = \frac{d}{dx} \left[f'(x) \frac{\partial L}{\partial f'} \right]. \quad (1.2.8)$$

Dabei haben wir im letzten Schritt die Produktregel verwendet. Indem wir die beiden Ableitungen auf eine Seite der Gleichung bringen, erhalten wir somit

$$\frac{d}{dx} H = 0 \quad \text{mit} \quad H := f' \frac{\partial L}{\partial f'} - L, \quad (1.2.9)$$

d.h. die **Hamilton-Funktion** des Variationsproblems ist eine Erhaltungsgröße. Wir bemerken, daß dies gilt, weil wir angenommen haben, daß L nicht explizit von x sondern nur implizit über die Abhängigkeit von $f(x)$ und $f'(x)$ von x abhängt. Wir können (1.2.9) sofort integrieren zu

$$H = E = \text{const} \quad (1.2.10)$$

für die Lösung der Euler-Lagrange-Gleichung. Man nennt H ein **erstes Integral** dieser Differentialgleichung, denn sie ist nurmehr nur noch eine Differentialgleichung 1. Ordnung, die i.a., einfacher zu lösen ist als eine Differentialgleichung 2. Ordnung.

1.2.1 Lösung des Brachistochronenproblems

Jetzt können wir uns mit der Lösung des Brachistochronenproblems beschäftigen. Hier ist gemäß (1.1.7)

$$L(f, f') = \sqrt{\frac{1 + f'^2(x)}{2gf(x)}}, \quad (1.2.11)$$

und wir können direkt (1.2.10) verwenden. Dazu berechnen wir zunächst die Hamilton-Funktion mit (1.2.9). Es gilt

$$\frac{\partial L}{\partial f'} = \frac{f'}{\sqrt{2gf(1+f'^2)}} \quad (1.2.12)$$

und damit

$$H = f' \frac{\partial L}{\partial f'} - L = -\frac{1}{\sqrt{2gf(1+f'^2)}}. \quad (1.2.13)$$

Für die Lösung des Extremalproblems ist gemäß (1.2.10) diese Größe konstant. Bequemer läßt sich die Kurve also durch

$$f(1+f'^2) = C = \text{const} \quad (1.2.14)$$

beschreiben.

Um eine Lösung zu finden, machen wir den Ansatz

$$y = f(x) = A \sin^2\left(\frac{\lambda}{2}\right) = \frac{A}{2}(1 - \cos \lambda). \quad (1.2.15)$$

Dann ist die Kurve als Parameterdarstellung $\vec{x} = \vec{x}(\lambda)$ gegeben. Es gilt mit der Kettenregel

$$f' = \frac{df}{dx} = \frac{df}{d\lambda} \frac{d\lambda}{dx} = \frac{df}{d\lambda} \left(\frac{dx}{d\lambda}\right)^{-1} = A \sin\left(\frac{\lambda}{2}\right) \cos\left(\frac{\lambda}{2}\right) \left(\frac{dx}{d\lambda}\right)^{-1}. \quad (1.2.16)$$

Setzen wir dies in (1.2.14) ein, folgt durch einfache Umformungen

$$\frac{C}{A \sin^2(\lambda/2)} = 1 + A^2 \sin^2\left(\frac{\lambda}{2}\right) \cos^2\left(\frac{\lambda}{2}\right) \left(\frac{dx}{d\lambda}\right)^{-2} \quad (1.2.17)$$

Setzen wir dann

$$\frac{dx}{d\lambda} = A \sin^2\left(\frac{\lambda}{2}\right), \quad (1.2.18)$$

folgt

$$\frac{C}{A \sin^2(\lambda/2)} = 1 + \frac{\cos^2(\lambda/2)}{\sin^2(\lambda/2)} = \frac{\sin^2(\lambda/2) + \cos^2(\lambda/2)}{\sin^2(\lambda/2)} = \frac{1}{\sin^2(\lambda/2)} \Rightarrow A = C. \quad (1.2.19)$$

Also ist

$$\begin{aligned} \frac{dx}{d\lambda} &= C \sin^2\left(\frac{\lambda}{2}\right) = \frac{C}{2}(1 - \cos \lambda) \Rightarrow x(\lambda) = \frac{C}{2}(\lambda - \sin \lambda) + C_2; \quad C_2 = \text{const.} \\ y(\lambda) &= C \sin^2\left(\frac{\lambda}{2}\right) = \frac{C}{2}(1 - \cos \lambda). \end{aligned} \quad (1.2.20)$$

Da der Punkt $(0,0)$ auf der Kurve liegen soll, können wir $C_2 = 0$ setzen. Dann ist C durch den zweiten Punkt $P_2 = (x_2, y_2)$ bestimmt. Bei der Kurve handelt es sich um eine **Zykloide**.

1.3 Das Fundamentallemma der Variationsrechnung

Wir holen nun den Beweis des sog. **Fundamentallemmas der Variationsrechnung** nach. Sei $f : [t_a, t_b] \rightarrow \mathbb{R}$ stetig. Gilt dann für jede auf dem Intervall für jede beliebig oft stetig differenzierbare Funktion δ

$$\int_{t_a}^{t_b} dt f(t) \delta(t) = 0, \quad (1.3.1)$$

so ist notwendig $f(t) = 0$ für alle $t \in [t_a, t_b]$.

Beweis: Als ersten Schritt beweisen wir, daß die Funktion

$$m : \mathbb{R} \rightarrow \mathbb{R} : \quad m(x) = \begin{cases} \exp\left(-\frac{1}{1-x^2}\right) & \text{für } |x| < 1 \\ 0 & \text{für } |x| \geq 1 \end{cases} \quad (1.3.2)$$

eine überall beliebig oft stetig differenzierbare Funktion ist. Sie erfüllt dies mit Sicherheit in den offenen Mengen $(-1, 1)$, $(-\infty, -1)$ und $(1, \infty)$. An den Stellen $x = \pm 1$ besitzt die für das Intervall $(-1, 1)$ zur Definition benutzte Funktion Ableitungen, die das Produkt einer rationalen Funktion mit Polen bei $x = \pm 1$ multipliziert mit dem Differentialausdruck darstellen. Der Exponentialausdruck konvergiert allerdings für $x \rightarrow \pm 1$ stärker gegen 0 als jedes Polynom, so daß der Grenzwert jeweils 0 ergibt. Damit sind aber alle Ableitungen von m wie m selbst stetig bei $x = \pm 1$.

Zum Beweis des Fundamentallemmas nehmen wir nun an, es gäbe eine Stelle $t_0 \in (t_1, t_2)$ mit $f(t_0) \neq 0$. Da f nach Voraussetzung stetig ist, existiert eine Umgebung $U_a(t_0) = [t - a, t + a] \cap (t_1, t_2)$, in der f beständig dasselbe Vorzeichen besitzt wie an der Stelle t_0 .

Die Funktion

$$\delta(t) = m\left(\frac{t - t_0}{a}\right) \quad (1.3.3)$$

ist nun in der besagten Umgebung $U_a(t_0)$ positiv und beliebig oft differenzierbar. Im Widerspruch zur Voraussetzung gilt also offenbar

$$\int_{t_1}^{t_2} dt f(t) \delta(t) = \int_{U_a(t_0)} dt f(t) \delta(t) \neq 0, \quad (1.3.4)$$

denn der Integrand besitzt im Integrationsbereich gemäß unserer Konstruktion beständig dasselbe Vorzeichen und ist nicht beständig 0. Da aber das Integral nach Voraussetzung verschwindet, muß die Annahme, daß für irgendein $t_0 \in (t_1, t_2)$ gilt $f(t_0) \neq 0$, falsch sein, also ist $f \equiv 0$ in (t_1, t_2) . q.e.d.

1.4 Isoperimetrische Probleme

In diesem Abschnitt beschreiben wir eine geringfügig andere Art von Problemen. Ein klassisches Beispiel ist das **Problem der Dido**, bei dem verlangt wird, diejenige geschlossene Kurve in der Ebene zu bestimmen, die bei vorgegebener Länge L die größtmögliche Fläche umschließt.

In diesem Fall ist es am einfachsten, die Kurve mit der vom Anfangspunkt aus gezählten Bogenlänge s zu parametrisieren: $\vec{x} = \vec{x}(s) = (x(s), y(s))$. Die umschlossene Fläche berechnen wir, indem wir infinitesimale Dreiecke (vgl. Skizze) aufsummieren. Die Flächennormalenvektoren sind

$$d\vec{F} = \frac{1}{2} \text{vec } x \times d\vec{x} = \frac{1}{2} ds \vec{x} \times \frac{d\vec{x}}{ds} = \frac{1}{2} ds \begin{pmatrix} 0 \\ 0 \\ xy - yx \end{pmatrix}. \quad (1.4.1)$$

Dabei soll der Punkt die Ableitung nach s bedeuten. Wir nehmen an, die Kurve werde im Gegenuhrzeigersinn durchlaufen. Dann ist die z -Komponente des obigen Flächennormalenvektors stets positiv, und wir erhalten für die Fläche

$$F = \frac{1}{2} \int_0^\ell ds (xy - yx). \quad (1.4.2)$$

Beim Variieren der Kurve zur Bestimmung des Extremums müssen wir aber noch die **Nebenbedingung** beachten, daß die Bogenlänge stets ℓ sein muß. Da wir die Bogenlänge s selbst als Parameter gewählt haben, gilt $|\dot{\vec{x}}| = 1$, und die Nebenbedingung ist trivialerweise

$$L = \int_0^\ell ds |\dot{\vec{x}}| = \int_0^\ell ds. \quad (1.4.3)$$

1. Variationsrechnung und kanonische Mechanik

Wie schon bei Extremwertproblemen für Funktionen mehrerer Variabler mit Nebenbedingungen können wir die Nebenbedingungen mit **Lagrangeparametern** berücksichtigen, d.h. wir variieren die Kurve ohne die Nebenbedingungen zu berücksichtigen und führen eine zusätzliche Funktion λ ein und verlangen, daß das Funktional

$$A[\vec{x}, \lambda] = \frac{1}{2} \int_0^\ell ds [x\dot{y} - y\dot{x} + \lambda(|\dot{\vec{x}}| - 1)] \quad (1.4.4)$$

unter Variation der Kurve \vec{x} und λ extremal wird. Die Lagrangefunktion ist in diesem Fall also

$$L(\vec{x}, \dot{\vec{x}}, \lambda) = \frac{1}{2} [x\dot{y} - y\dot{x} + \lambda(|\dot{\vec{x}}| - 1)] \quad (1.4.5)$$

und die Variation nach \vec{x} und λ liefert wieder die Euler-Lagrange-Gleichungen (*nachrechnen!*)

$$\begin{aligned} \frac{\partial L}{\partial x} - \frac{d}{ds} \frac{\partial L}{\partial \dot{x}} &= 0, \\ \frac{\partial L}{\partial y} - \frac{d}{ds} \frac{\partial L}{\partial \dot{y}} &= 0, \\ \frac{\partial L}{\partial \lambda} &= 0. \end{aligned} \quad (1.4.6)$$

Für die Lagrangefunktion (1.4.5) ergibt sich

$$\begin{aligned} \dot{y} - \frac{d}{ds} \left(-y + \lambda \frac{\dot{x}}{|\dot{\vec{x}}|} \right) &= 0, \\ -\dot{x} - \frac{d}{ds} \left(x + \lambda \frac{\dot{y}}{|\dot{\vec{x}}|} \right) &= 0, \\ |\dot{\vec{x}}| &= 1. \end{aligned} \quad (1.4.7)$$

Die beiden oberen Gleichungen können wir sofort aufintegrieren und die dritte Gleichung zur Vereinfachung verwenden. Es folgt

$$2y - \lambda\dot{x} = C_1, \quad 2x + \lambda\dot{y} = C_2. \quad (1.4.8)$$

Da die Lagrangefunktion nicht explizit von s abhängt, ist die dazugehörige Hamilton-Funktion entlang der Trajektorie (*nachrechnen!*)

$$H = \dot{x} \frac{\partial L}{\partial \dot{x}} + \dot{y} \frac{\partial L}{\partial \dot{y}} - L = \lambda = \text{const.} \quad (1.4.9)$$

Wegen $\lambda = \text{const}$ folgt durch Ableiten der ersten Gleichung (1.4.8)

$$\dot{y} = \frac{\lambda}{2} \ddot{x} \quad (1.4.10)$$

und durch Einsetzen in (1.4.8)

$$\frac{\lambda^2}{2} \ddot{x} + 2x = C_2. \quad (1.4.11)$$

Dies ist eine lineare Differentialgleichung mit konstanten Koeffizienten. Die allgemeine Lösung ergibt sich als Summe der allgemeinen Lösung der homogenen Gleichung (also für $C_2 = 0$) und einer speziellen Lösung der inhomogenen Gleichung. Eine spezielle Lösung ergibt sich offenbar durch den Ansatz $x = A = \text{const}$. Dann ergibt sich $2A = C_2 \Rightarrow A = C_2/2$. Setzen wir also $x = C_2/2 + \xi$ folgt

$$\ddot{\xi} = -\frac{4}{\lambda^2} \xi, \quad (1.4.12)$$

und diese Gleichung können wir sofort lösen durch

$$\xi(s) = x(s) - \frac{C_1}{2} = k_1 \cos\left(\frac{2}{\lambda}s\right) + k_2 \sin\left(\frac{2}{\lambda}s\right) = k \cos\left(\frac{2}{\lambda}s + \varphi\right) \quad (1.4.13)$$

mit $k_1, k_2 = \text{const}$ bzw. $k, \varphi = \text{const}$. Aus der ersten Gleichung von (1.4.9) folgt dann

$$y(s) = \frac{1}{2}(C_2 + \lambda\dot{x}) = \frac{1}{2}(C_2 + \lambda\xi) = \frac{C_2}{2} - k \sin\left(\frac{2}{\lambda}s + \varphi\right). \quad (1.4.14)$$

Es handelt sich also um einen Kreis. Es ist klar, daß die absolute Lage durch das Problem unbestimmt ist, d.h. wir können ohne Beschränkung der Allgemeinheit $C_1 = C_2 = 0$ setzen. Die beiden weiteren Konstanten λ und k bestimmen sich dann aus der Nebenbedingung, daß die Bogenlänge des Kreises $\ell = 2\pi k$ sein muß. Der Kreis soll für das Intervall $s \in [0, \ell]$ genau einmal durchlaufen werden, d.h. es muß $2\ell/\lambda = 2\pi$ sein, d.h. $\lambda = \frac{\pi}{\ell}$. Die Wahl von φ ist unerheblich. Diese Phase bestimmt nur, ab wo die Bogenlänge $s \in [0, \ell]$ gezählt wird. Wir können also auch $\varphi = 0$ setzen und die Lösung des Problems lautet damit schließlich

$$x(s) = \frac{\ell}{2\pi} \cos\left(\frac{2\pi}{\ell}s\right), \quad y(s) = \frac{\ell}{2\pi} \sin\left(\frac{2\pi}{\ell}s\right). \quad (1.4.15)$$

1.5 Das Hamiltonsche Prinzip

Das **Hamiltonsche Prinzip** stellt eines der wichtigsten mathematischen Beschreibungsweisen für die Dynamik physikalischer Systeme dar. Die klassische Mechanik ist dafür das einfachste Beispiel. Seine Anwendungen umfassen jedoch weit größere Teile der Physik, insbesondere die klassische Feldtheorie (Elektrodynamik, Allgemeine Relativitätstheorie), und es spielt eine entscheidende Rolle bei der Formulierung der Quantentheorie.

1.5.1 Die Lagrange-Funktion

Die Idee besteht dabei darin, die Bewegungsgleichungen für dynamische Systeme als **Euler-Lagrange-Gleichungen** geeigneter **Variationsprobleme** zu beschreiben. Wir betrachten zuerst den einfachsten Fall eines Teilchens, das sich in einem vorgegebenen äußeren **konservativen Kraftfeld** bewegt. Dabei heißt ein Kraftfeld konservativ, wenn man es als Gradient eines von der Zeit unabhängigen skalaren Potentials V beschreiben kann. Am einfachsten läßt sich dieses Problem in kartesischen Koordinaten x^k bzgl. eines Inertialsystems beschreiben¹ Dann gilt die Newtonsche Bewegungsgleichung

$$m\ddot{x}^k = m\ddot{x}_k = -\nabla_k V = -\frac{\partial V}{\partial x^k}. \quad (1.5.1)$$

Dabei ist zu beachten, daß für kartesische Koordinaten $x^k = x_k$ und $\nabla_k = \nabla^k$ ist.

Diese Bewegungsgleichung läßt sich nun leicht als Variationsproblem beschreiben. Wir benötigen dazu nur eine Lagrange-Funktion $L(x^k, \dot{x}^k)$, deren Euler-Lagrange-Gleichungen

$$\frac{d}{dt} \frac{\partial L}{\partial \dot{x}^k} = \frac{\partial L}{\partial x^k} \quad (1.5.2)$$

der Newtonschen Bewegungsgleichung (1.5.1) entsprechen.

¹Wir verwenden in dieser Vorlesung für alle Koordinaten den Ricci-Kalkül mit hoch- und tiefgestellten Indizes, entsprechend kontra- und kovariantem Transformationsverhalten bzgl. Basiswechseln. Wir kommen darauf noch ausführlicher zu sprechen. Die entsprechende Theorie wird ausführlich auch im Skript zu „Mathematische Ergänzungen zur Theoretischen Physik 1“ [Hee14] behandelt.

Dann ergibt sich die Trajektorie des Teilchens als die Lösung des Variationsproblems, das sog. **Wirkungsfunktional**

$$S[x^k] = \int_{t_1}^{t_2} dt L(x^k, \dot{x}^k) \quad (1.5.3)$$

extremal zu machen, wobei die Endpunkte der Bahnen $x^k(t)$ festgehalten werden, d.h. für die Variation um die Trajektorie gilt $\delta x^k(t_1) = \delta x^k(t_2) = 0$. Dies ist das **Hamilton-Prinzip der kleinsten Wirkung**.

Um nun die Lagrange-Funktion zu bestimmen, verlangen wir, daß

$$m\ddot{x}_k = \frac{d}{dt} \frac{\partial L}{\partial \dot{x}^k}, \quad \frac{\partial L}{\partial x^k} = -\frac{\partial V}{\partial x^k} \quad (1.5.4)$$

gilt, denn dann ist die Newtonsche Bewegungsgleichung (1.5.1) offenbar durch die Euler-Lagrange-Gleichung (1.5.2) beschrieben. Die erste Gleichung (1.5.4) können wir einmal nach er Zeit integrieren, wobei für unsere Zwecke zunächst eine spezielle Lösung der entsprechenden Differentialgleichungen ausreicht, denn jede Lagrange-Funktion, deren Euler-Lagrange-Gleichung die Newtonsche Bewegungsgleichung reproduziert, ist physikalisch gleichberechtigt. Wir können also einfach

$$m\dot{x}_k = \frac{\partial L}{\partial \dot{x}^k} \quad (1.5.5)$$

verlangen, und diese Gleichungen werden offensichtlich durch

$$L(x^k, \dot{x}^k) = \frac{m}{2} \dot{\vec{x}}^2 + \tilde{L}(x^k) \quad (1.5.6)$$

gelöst. Dabei ist \tilde{L} nur noch von den x^k abhängig, also bzgl. der \dot{x}^k konstant. Diese verbliebene Funktion muß gemäß der zweiten Gleichung in (1.5.4)

$$\frac{\partial \tilde{L}}{\partial x^k} = -\frac{\partial V}{\partial x^k} \quad (1.5.7)$$

erfüllen. Dies wird offenbar durch $\tilde{L} = -V$ erfüllt, und eine mögliche Lagrange-Funktion ist damit durch

$$L(x^k, \dot{x}^k) = \frac{m}{2} \dot{\vec{x}}^2 - V(\vec{x}) \quad (1.5.8)$$

gegeben.

Wie bei der Herleitung von (1.2.9) gezeigt, ergibt sich, daß die **Hamilton-Funktion**

$$H = \sum_{k=1}^3 \dot{x}^k \cdot \frac{\partial L}{\partial \dot{x}^k} - L \quad (1.5.9)$$

entlang der Trajektorie erhalten ist. Setzt man (1.5.8) ein, erhält man den Energiesatz, der stets gilt, wenn die Kräfte konservativ sind und demzufolge die Lagrange-Funktion nicht explizit von der Zeit abhängt, d.h. die Größe

$$H = \frac{m}{2} \dot{\vec{x}}^2 + V(\vec{x}) = E \quad (1.5.10)$$

ist für die Lösung der Euler-Lagrange-Gleichung, also die Trajektorie des Teilchens unter Einfluß der konservativen Kraft, erhalten. Dies ist die von der Newtonschen Mechanik her bekannte **Energieerhaltung**. Man bezeichnet den ersten Term $T = m\dot{\vec{x}}^2/2$ als **kinetische** und das Potential $V(\vec{x})$ der Kräfte als **potentielle Energie**.

Der erste Vorteil des Hamilton-Prinzips gegenüber der Newtonschen Bewegungsgleichung ist, daß es sich leicht auf **verallgemeinerte Koordinaten** umschreiben läßt. Dies ist bequem, wenn bei einer Anwendung

bestimmte **Symmetrien** vorliegen, z.B. wenn das Potential $V(\vec{x}) = V(r)$ mit $r = |\vec{x}|$ gilt, d.h. wenn ein **Zentralpotential** vorliegt. Dann ist es offensichtlich, daß statt der oben verwendeten kartesischen Kugelkoordinaten (r, ϑ, φ) zu benutzen.

Seien also q^j ($j \in \{1, 2, 3\}$) beliebige verallgemeinerte Koordinaten, die zumindest einen Teil des Raumes umkehrbar eindeutig beschreiben, d.h. die kartesischen Koordinaten x^k sind umkehrbar eindeutige Funktionen der q^j :

$$x^k = x^k(q^j). \quad (1.5.11)$$

Die Trajektorien in diesem Raumbereich werden dann eindeutig als Funktionen $q^k = q^k(t)$ der generalisierten Koordinaten von der Zeit beschrieben. Wir müssen nun nur die Lagrange-Funktion durch diese generalisierten Koordinaten und deren Zeitableitungen ausdrücken. Zunächst gilt wegen der Kettenregel.

$$T = \frac{m}{2} \sum_{i,j,k,l=1}^3 \frac{\partial x^k}{\partial q^i} \frac{\partial x^l}{\partial q^j} \dot{q}^i \dot{q}^j \delta_{kl}. \quad (1.5.12)$$

Hier und im folgenden bietet es sich an, die **Einsteinsche Summationskonvention** zu verwenden, wonach wir das Summenzeichen einfach weglassen und über Indizes, die doppelt in einer Gleichung auftreten, summieren. In der hier verwendeten Schreibweise des sog. **Ricci-Kalküls** muß dabei stets ein Index dieses Paares oben und einer unten stehen. Ableitungen nach Größen mit einem oberen Index sind dabei Größen mit einem unteren Index. Wir kommen darauf gleich nochmals zurück.

Definieren wir nun

$$g_{ij}(q) = \delta_{kl} \frac{\partial x^k}{\partial q^i} \frac{\partial x^l}{\partial q^j}, \quad (1.5.13)$$

ergibt sich

$$T = \frac{m}{2} g_{ij}(q) \dot{q}^i \dot{q}^j. \quad (1.5.14)$$

Dabei sind die über die generalisierten Koordinaten q^k vom Ort abhängigen Koeffizienten g_{ij} die Komponenten der **Euklidischen Metrik** bzgl. der lokalen Basisvektoren

$$\vec{g}_j = \frac{\partial \vec{x}}{\partial q^j}, \quad g_{ij} = \vec{g}_i \cdot \vec{g}_j. \quad (1.5.15)$$

Das Potential können wir über die Parametrisierung (1.5.11) direkt durch die generalisierten Koordinaten ausdrücken, wobei wir dem etwas laxen Physikerbrauch in der Notation folgendend, das gleiche Symbol V für die Funktion verwenden, d.h. wir schreiben einfach

$$V(q) = V[\vec{x}(q)]. \quad (1.5.16)$$

Die Wirkung wird nun einfach durch

$$S[q] = \int_{t_1}^{t_2} dt L(q, \dot{q}) \quad \text{mit} \quad L(q, \dot{q}) = T(q, \dot{q}) - V(q) = \frac{m}{2} g_{ij}(q) \dot{q}^i \dot{q}^j - V(q) \quad (1.5.17)$$

gegeben.

Die Bewegungsgleichungen ergeben sich wieder nach dem Hamilton-Prinzip der kleinsten Wirkung, nur daß wir jetzt die Trajektorie vermöge der verallgemeinerten Koordinaten $q^k = q^k(t)$ ausdrücken. Entsprechend gelten die Euler-Lagrange-Gleichungen für beliebige generalisierten Koordinaten, d.h. sie sind **forminvariant unter allgemeinen Koordinatentransformationen**:

$$\frac{d}{dt} \frac{\partial L}{\partial \dot{q}^k} = \frac{\partial L}{\partial q^k}. \quad (1.5.18)$$

Wichtig ist noch der Begriff der **kanonisch konjugierten Impulse**. Sie werden definiert als

$$p_k = \frac{\partial L}{\partial \dot{q}^k}. \quad (1.5.19)$$

Wir können dann die Euler-Lagrange-Gleichungen (1.5.18) in der Form

$$\dot{p}_k = \frac{\partial L}{\partial q^k} \quad (1.5.20)$$

schreiben. Daraus ist ersichtlich, daß ein kanonisch konjugierter Impuls p_k **entlang der Trajektorie erhalten** ist, wenn die Lagrange-Funktion nicht von der generalisierten Koordinate q^k abhängt. Man nennt eine solche generalisierte Koordinate **zyklisch**.

Es ist also für die Integration der Bewegungsgleichungen eines Problems wünschenswert, solche generalisierten Koordinaten einzuführen, für die möglichst viele Koordinaten zyklisch sind. Dies erreicht man i.a. dadurch, daß man **Symmetrien** eines Problems ausnutzt. Zu einer systematischen Betrachtung dazu kommen wir in Abschnitt 1.6.

1.5.2 Äquivalente Lagrange-Funktionen

Es ist klar, daß die soeben hergeleitete Lagrange-Funktion nicht die einzige Funktion ist, die auf die Bewegungsgleichungen führt. Wir bezeichnen zwei Lagrange-Funktionen als **äquivalent**, wenn sie zu denselben Bewegungsgleichungen führen. Dies ist sicher immer dann der Fall, wenn die Variation der Wirkungen, die aus diesen beiden Lagrange-Funktionen vermöge (1.5.17) definiert werden, übereinstimmen.

Sei also

$$L'(q^k, \dot{q}^k, t) = L(q^k, \dot{q}^k, t) + \Lambda(q^k, \dot{q}^k, t). \quad (1.5.21)$$

Dann sind die Lagrange-Funktionen L' und L offenbar äquivalent, wenn die mit der Funktion Λ gebildete Euler-Lagrange-Gleichungen identisch erfüllt werden, also *für alle Trajektorien* ($q^k(t)$)

$$\frac{d}{dt} \frac{\partial \Lambda}{\partial \dot{q}^k} = \frac{\partial \Lambda}{\partial q^k} \quad (1.5.22)$$

gilt. Führen wir also die Zeitableitung auf der rechten Seite aus, erhalten wir²

$$\frac{\partial^2 \Lambda}{\partial \dot{q}^j \partial \dot{q}^k} \dot{q}^j + \frac{\partial^2 \Lambda}{\partial \dot{q}^j \partial q^k} \ddot{q}^k + \frac{\partial}{\partial t} \frac{\partial \Lambda}{\partial \dot{q}^k} = \frac{\partial \Lambda}{\partial q^k}. \quad (1.5.23)$$

Dabei bedeutet die partielle Zeitableitung $\partial/\partial t$ die Ableitung nach der expliziten Zeitabhängigkeit. Da nun voraussetzungsgemäß die rechte Seite dieser Gleichung nicht von \ddot{q}^k abhängt, muß notwendig

$$\frac{\partial^2 \Lambda}{\partial \dot{q}^j \partial \dot{q}^k} = 0 \quad (1.5.24)$$

gelten. Dies ist aber nur möglich, wenn Λ eine lineare Funktion in die \dot{q}^k ist, wobei die Koeffizienten von q und t abhängen dürfen. Damit wird

$$\Lambda(q^k, \dot{q}^k, t) = \tilde{\Lambda}_j(q^k, t) \dot{q}^k + \tilde{\Lambda}(q^k, t). \quad (1.5.25)$$

Setzen wir dies in (1.5.23) ein, erhalten wir nach einer einfachen Rechnung

$$\partial_t \tilde{\Lambda}_k = \frac{\partial \tilde{\Lambda}}{\partial q^k}. \quad (1.5.26)$$

²Es gilt wieder die Einsteinsche Summenkonvention!

Das bedeutet, daß

$$\tilde{\Lambda}_k = \frac{\partial}{\partial q^k} \Omega(q^k, t) \quad (1.5.27)$$

mit einer beliebigen Funktion Ω , die nur von den q^k und t abhängt, setzen müssen. Aus (1.5.26) folgt dann, daß

$$\tilde{\Lambda} = \partial_t \Omega \quad (1.5.28)$$

ist. Setzen wir nun (1.5.27) und (1.5.28) in (1.5.25) ein, folgt

$$\Lambda = \frac{\partial \Omega}{\partial q^k} \dot{q}^k + \partial_t \Omega = \frac{d}{dt} \Omega. \quad (1.5.29)$$

Das bedeutet, daß die beiden Lagrangefunktionen L' und L genau dann äquivalent sind, wenn es eine Funktion $\Omega(q^k, t)$ gibt, so daß

$$L'(q^k, \dot{q}^k, t) = L(q^k, \dot{q}^k, t) + \frac{d}{dt} \Omega(q^k, t) \quad (1.5.30)$$

gilt.

Betrachten wir nun die Wirkungen. Offenbar ist

$$S'[q] = \int_{t_1}^{t_2} dt L' = S[q] + \Omega[q^k(t_2), t_2] - \Omega[q^k(t_1), t_1]. \quad (1.5.31)$$

Die Wirkungen unterscheiden sich also nur durch **Randterme**, also Werte der Trajektorien an den Randpunkten des betrachteten Zeitintervalls $[t_1, t_2]$, und zwar nur von den $q^k(t_1)$ und $q^k(t_2)$ nicht von \dot{q}^k . Da beim Hamiltonschen Prinzip definitionsgemäß die Randpunkte festgehalten werden, stimmen die Variationen der beiden Wirkungsfunktionale überein.

1.5.3 Beispiel: Freier Fall bzw. schiefer Wurf

Wir betrachten in folgenden einige sehr einfache Beispiele, um die grundlegende Rechentechnik des Lagrange-Formalismus zu erläutern. Beginnen wir mit dem freien Fall in Erdnähe. Dazu nähern wir die Gravitationskraft auf ein Punktteilchen als konstant:

$$\vec{F} = -m g \vec{e}_3, \quad (1.5.32)$$

wobei \vec{e}_3 ein Basisvektor der kartesischen Basis (\vec{e}_j) ($j \in \{1, 2, 3\}$) ist. Natürlich läßt sich dieses Beispiel am einfachsten durch direkte Anwendung der Newtonschen Bewegungsgleichungen lösen, aber wir wollen den Lagrange-Formalismus verwenden, um an diesem einfachen Beispiel die grundlegenden Rechenschritte zu erläutern.

Zunächst müssen wir uns überlegen, welche generalisierten Koordinaten wir wählen. Das entscheidet sich am einfachsten, wenn wir zunächst das Potential der Kraft bestimmen³, d.h. wir suchen eine skalare Funktion V , so daß

$$\vec{\nabla} \Phi = -\vec{F} = -m g \vec{e}_3 \quad (1.5.33)$$

gilt. Offensichtlich sind dazu kartesische Koordinaten besonders bequem, denn es gilt

$$\vec{\nabla} \Phi = \sum_{j=1}^3 \vec{e}_j \frac{\partial \Phi}{\partial x^j} \Rightarrow \frac{\partial \Phi}{\partial x^1} = \frac{\partial \Phi}{\partial x^2} = 0, \quad \frac{\partial \Phi}{\partial x^3} = +m g. \quad (1.5.34)$$

³Um den Lagrange-Formalismus anwenden zu können, müssen wir annehmen, daß die Kraft ein Potential besitzt.

Die ersten beiden Gleichungen besagen, daß $\Phi = \Phi(x^3)$ ist, d.h. nicht von x^1 und x^2 abhängt, und die letzte Gleichung läßt dich sofort integrieren zu

$$\Phi(x^3) = m g x^3. \quad (1.5.35)$$

Man könnte noch eine beliebige Konstante addieren, aber diese spielt im folgenden keine Rolle für die Bewegungsgleichungen, weil in den Euler-Lagrange-Gleichungen nur Ableitungen vorkommen, so daß eine additive Konstante in der Lagrange-Funktion irrelevant für die Bewegungsgleichungen ist. Die Lagrangefunktion lautet also in unseren kartesischen Koordinaten

$$L = T - V = \frac{m}{2} \dot{x}^2 - m g x^3. \quad (1.5.36)$$

Die Lagrangefunktion hängt nicht von x^1 und x^2 ab, d.h. diese Koordinaten sind zyklisch und folglich sind die dazugehörigen generalisierten Impulse wegen (1.5.20) Erhaltungsgrößen, d.h. es gilt

$$\begin{aligned} p_1 &= \frac{\partial L}{\partial \dot{x}^1} = m \dot{x}^1 = \text{const}, \\ p_2 &= \frac{\partial L}{\partial \dot{x}^2} = m \dot{x}^2 = \text{const}. \end{aligned} \quad (1.5.37)$$

Damit lassen sich die Bewegungsgleichungen für x^1 und x^2 sofort lösen:

$$\begin{aligned} x^1(t) &= v_0^1 t + x_0^1, & v_0^1 &= \dot{x}^1(0) = \text{const}, & x_0^1 &= x^1(0) = \text{const}, \\ x^2(t) &= v_0^2 t + x_0^2, & v_0^2 &= \dot{x}^2(0) = \text{const}, & x_0^2 &= x^2(0) = \text{const}. \end{aligned} \quad (1.5.38)$$

Auch die Bewegungsgleichung für x^3 läßt sich in diesem Fall am einfachsten direkt integrieren. Um sie aus dem Lagrange-Formalismus herzuleiten, betrachten wir die entsprechende Euler-Lagrange-Gleichung

$$\frac{d}{dt} \frac{\partial L}{\partial \dot{x}^3} = \dot{p}_3 = m \ddot{x}^3 \stackrel{!}{=} \frac{\partial L}{\partial x^3} = -m g, \quad (1.5.39)$$

und dies kann man in der Tat einfach zweimal nach der Zeit integrieren. Unter Berücksichtigung der Anfangsbedingungen, folgt

$$x^3(t) = -\frac{g}{2} t^2 + v_0^3 t + x_0^3. \quad (1.5.40)$$

Wir bemerken noch, daß die Lagrangefunktion nicht explizit von der Zeit abhängt und somit wegen (1.5.10) die Energie erhalten ist. Es ist also

$$H = \frac{m}{2} \dot{x}^2 + m g x^3 = E = \text{const}. \quad (1.5.41)$$

Das bestätigt man unmittelbar durch Einsetzen der Lösungen (1.5.38) und (1.5.40) (*Übung!*).

1.5.4 Beispiel: Harmonischer Oszillator

Als nächstes betrachten wir den **eindimensionalen harmonischen Oszillator**. Die Lagrange-Funktion lautet in diesem Fall

$$L = \frac{m}{2} \dot{x}^2 - \frac{k}{2} x^2. \quad (1.5.42)$$

In der Tat lautet dann die Bewegungsgleichung

$$\frac{d}{dt} \frac{\partial L}{\partial \dot{x}} = m \ddot{x} \stackrel{!}{=} \frac{\partial L}{\partial x} = -k x. \quad (1.5.43)$$

Auch diese **lineare Differentialgleichung 2. Ordnung** mit konstanten Koeffizienten läßt sich leicht direkt integrieren (vgl. [Hee14], Abschnitt 5.2). Wir wollen hier aber den alternativen Weg über den Energieerhaltungssatz verwenden, um diese generell wichtige Methode zu erläutern. Da nämlich L nicht explizit zeitabhängig ist, gilt der Energie-Erhaltungssatz, d.h.

$$H = \dot{x} \frac{\partial L}{\partial \dot{x}} - L = \frac{m}{2} \dot{x}^2 + \frac{k}{2} x^2 = E = \text{const.} \quad (1.5.44)$$

Dies ist eine Differentialgleichung 1. Ordnung und insofern einfacher als die Bewegungsgleichung. Sie ist zwar nicht linear, läßt sich aber leicht durch „Trennung der Variablen“ lösen:

$$\dot{x} = \pm \sqrt{\frac{1}{m}(2E - kx^2)}, \quad (1.5.45)$$

d.h. es gilt

$$\int_0^t dt' = t = \pm \int_{x_0}^x dx' \sqrt{\frac{m}{2E - kx'^2}}. \quad (1.5.46)$$

Das Vorzeichen bestimmt sich dabei wie folgt. Wegen (1.5.44) ist $|x_0| \leq \sqrt{2E/k}$, so daß das Argument unter der Wurzel positiv ist. Das Vorzeichen der Anfangsgeschwindigkeit bestimmt dann das Vorzeichen in (1.5.46) und ob $x < x_0$ oder $x > x_0$ wird, denn die Zeit muß immer monoton wachsen, d.h. $t > 0$ sein. Erreicht dann x die Werte $\pm \sqrt{2E/k}$, muß sich das Vorzeichen der Wurzel und die Änderungsrichtung von x wieder umkehren. Das Teilchen bewegt sich also stets zwischen diesen beiden Endpunkten hin- und her. Das wird auch an der nebenstehenden Skizze deutlich, wo das Potential $V(x) = kx^2/2$ aufgetragen und der Wert der Gesamtenergie also horizontale Linie eingetragen ist. Da die kinetische Energie $T = m\dot{x}^2/2 \geq 0$ ist, können nur die Bereich von x durch den Massenpunkt erreicht werden, wo $V(x) \leq E$ ist, und damit bestimmt die Gleichung $V(x) = E$ die beiden Umkehrpunkte der Bewegung.

Wir können nun das Integral (1.5.46) ausführen:

$$\begin{aligned} \int_{x_0}^x dx' \sqrt{\frac{m}{2E - kx'^2}} &= \sqrt{\frac{m}{2E}} \int_{x_0}^x dx' \sqrt{\frac{1}{1 - k/(2E)x'^2}} \\ &= \sqrt{\frac{m}{k}} \left[\arccos \left(\sqrt{\frac{k}{2E}} x \right) - \arccos \left(\sqrt{\frac{k}{2E}} x_0 \right) \right]. \end{aligned} \quad (1.5.47)$$

Setzen wir $\omega = \sqrt{k/m}$, liefert dies gemäß (1.5.46)

$$x(t) = \pm \sqrt{\frac{2E}{k}} \cos(\omega t - \varphi_0), \quad \varphi_0 = \mp \arccos \left(\sqrt{\frac{k}{2E}} x_0 \right). \quad (1.5.48)$$

Wir können die allgemeine Lösung einfacher in der Form

$$\begin{aligned} x(t) &= A \cos(\omega t) + B \sin(\omega t), \quad x(0) = A = x_0, \quad \dot{x}(0) = B\omega = v_0, \\ x(t) &= x_0 \cos(\omega t) + \frac{v_0}{\omega} \sin(\omega t). \end{aligned} \quad (1.5.49)$$

angeben. Um dies in die Form (1.5.48) zu bringen, bemerken wir, daß wegen der Additionstheoreme für die Winkelfunktionen

$$x(t) = C \cos(\omega t - \varphi_0) = C [\cos(\omega t) \cos \varphi_0 + \sin(\omega t) \sin \varphi_0] \quad (1.5.50)$$

gilt. Vergleich wir dies mit (1.5.49), ergibt sich

$$C \cos \varphi_0 = x_0, \quad C \sin \varphi_0 = \frac{v_0}{\omega}. \quad (1.5.51)$$

Zunächst folgt

$$C^2 = C^2(\cos^2 \varphi_0 + \sin^2 \varphi_0) = x_0^2 + \left(\frac{v_0}{\omega}\right)^2. \quad (1.5.52)$$

Wir können dann willkürlich $C > 0$ annehmen. Dann ist

$$\cos \varphi_0 = \frac{x_0}{C} \Rightarrow \varphi_0 = \pm \arccos\left(\frac{x_0}{C}\right). \quad (1.5.53)$$

Da $\arccos \alpha \in [0, \pi]$ für $\alpha \in [0, 1]$, bestimmt sich das Vorzeichen aus dem Vorzeichen von v_0 , d.h.

$$\varphi_0 = \text{sign } v_0 \arccos\left(\frac{x_0}{C}\right). \quad (1.5.54)$$

Nun ist wegen $\omega = \sqrt{k/m}$

$$C^2 = x_0^2 + \frac{m}{k} v_0^2 = \frac{1}{k}(kx_0^2 + mv_0^2) = \frac{2E}{k} \Rightarrow C = +\sqrt{\frac{2E}{k}}, \quad (1.5.55)$$

was mit der Definition (1.5.53) für φ_0 auf die Form (1.5.48) aufgrund der spezifischen Wahl für das Vorzeichen der **Amplitude** $C > 0$. Den Parameter φ_0 nennt man die **Phase** der Schwingung.

1.6 Das Noether-Theorem (Lagrange-Form)

Das Noether-Theorem stellt eines der wichtigsten Resultate der mathematischen Physik für die gesamte moderne Theorie-Bildung dar. Es stellt einen Zusammenhang zwischen **Symmetrien** der Wirkung (bzw. der Variation der Wirkung) und **Erhaltungsgrößen** her, d.h. Größen, die als Funktionen der generalisierten Koordinaten, ihrer Zeitableitungen und evtl. explizit der Zeit entlang der Trajektorien des Systems, also für die Lösungen der Bewegungsgleichungen zeitlich konstant bleiben.

1.6.1 Symmetrien

Dabei verstehen wir unter einer Symmetrie im Sinne des Lagrange-Formalismus eine Transformation der Zeit- und verallgemeinerten Koordinaten der Art

$$t' = t'(t, q^k, \dot{q}^k), \quad q'^j = q'^j(t, q^k, \dot{q}^k), \quad (1.6.1)$$

für die die Lagrange-Funktion $L(q'^k, \dot{q}'^k, t')$, geschrieben in den ursprünglichen Koordinaten und der ursprünglichen Zeitvariablen äquivalent zur ursprünglichen Lagrange-Funktion $L(q^k, \dot{q}^k, t)$ ist, d.h. es existiert eine Funktion $\Omega(q^k, t)$, so daß⁴

$$L(q'^k, \dot{q}'^k, t') \frac{dt'}{dt} = L(q^k, \dot{q}^k, t) + \frac{d}{dt} \Omega(q^k, t). \quad (1.6.2)$$

Dabei haben wir den Faktor dt'/dt eingeführt, damit sich die Integration in der transformierten Wirkung

$$S'[q'] = \int_{t'_1}^{t'_2} dt' L(q'^k, \dot{q}'^k, t') = \int_{t_1}^{t_2} dt \frac{dt'}{dt} L(q'^k, \dot{q}'^k, t') \quad (1.6.3)$$

⁴Man beachte, daß hier auf der linken und der rechten Seite der Gleichung *dieselbe* Lagrange-Funktion stehen!

auf beiden Seiten auf t bezieht.

Dabei ist definitionsgemäß

$$\dot{q}'^k = \frac{d}{dt'} q'^k. \quad (1.6.4)$$

Im folgenden genügen uns „infinitesimale Transformationen“, d.h. wir schreiben

$$t' = t + \epsilon T(t, q^k, \dot{q}^k), \quad q'^j = q^j + \epsilon Q^j(t, q^k, \dot{q}^k) \quad (1.6.5)$$

und betrachten die Symmetriebedingung (1.6.2) als gültig zur Ordnung $\mathcal{O}(\epsilon)$. Wir benötigen zunächst gemäß (1.6.4)

$$\dot{q}'^j = \frac{d}{dt'} q'^j \stackrel{(1.6.5)}{=} \frac{dt}{dt'} \frac{d}{dt} [q^j + \epsilon Q^j(t, q^k, \dot{q}^k)]. \quad (1.6.6)$$

Nun gilt

$$\frac{dt}{dt'} = \left(\frac{dt'}{dt} \right)^{-1} \stackrel{(1.6.5)}{=} \frac{1}{1 + \epsilon \frac{d}{dt} T} = 1 - \epsilon \frac{d}{dt} T + \mathcal{O}(\epsilon^2) \quad (1.6.7)$$

und damit

$$\dot{q}'^j = \left(1 - \epsilon \frac{d}{dt} T + \mathcal{O}(\epsilon^2) \right) \left(\dot{q}^j + \epsilon \frac{d}{dt} Q^j \right) = \dot{q}^j + \epsilon \left(\frac{d}{dt} Q^j - \dot{q}^j \frac{d}{dt} T \right) + \mathcal{O}(\epsilon^2). \quad (1.6.8)$$

Entwickeln wir nun die linke Seite von (1.6.2) bis zur Ordnung $\mathcal{O}(\epsilon)$, erhalten wir mit (1.6.8)

$$L(q'^k, \dot{q}'^k, t') = L(q^k, \dot{q}^k, t) + \epsilon \left[Q^k \frac{\partial L}{\partial q^k} + \left(\frac{d}{dt} Q^k - \dot{q}^k \frac{d}{dt} T \right) \frac{\partial L}{\partial \dot{q}^k} + T \frac{\partial L}{\partial t} \right] + \mathcal{O}(\epsilon^2). \quad (1.6.9)$$

In (1.6.2) müssen wir nur fordern, daß $\Omega = -\epsilon \tilde{\Omega} + \mathcal{O}(\epsilon^2)$ ist⁵, denn wir fordern die Äquivalenz der Lagrange-Funktionen auf der linken und rechten Seite von (1.6.2) nur bis auf Größen der Ordnung $\mathcal{O}(\epsilon^2)$, und Ω muß von der Ordnung $\mathcal{O}(\epsilon)$ sein. Damit folgt durch Einsetzen von (1.6.9) in (1.6.2) unter Berücksichtigung des Faktors $dt'/dt = 1 + \epsilon dT/dt$, daß eine Symmetrie vorliegt, wenn

$$Q^k \frac{\partial L}{\partial q^k} + \left(\frac{d}{dt} Q^k - \dot{q}^k \frac{d}{dt} T \right) \frac{\partial L}{\partial \dot{q}^k} + L \frac{d}{dt} T + T \frac{\partial L}{\partial t} + \frac{d}{dt} \tilde{\Omega} = 0 \quad (1.6.10)$$

gilt. Betrachten wir nun, was dies für die **Trajektorien** des Systems bedeutet, also für diejenigen $q^k(t)$, die die Bewegungsgleichungen

$$\frac{d}{dt} \frac{\partial L}{\partial \dot{q}^k} = \frac{\partial L}{\partial q^k} \quad (1.6.11)$$

erfüllen. Zunächst ist

$$Q^k \frac{\partial L}{\partial q^k} + \left(\frac{d}{dt} Q^k \right) \frac{\partial L}{\partial \dot{q}^k} = Q^k \frac{d}{dt} \frac{\partial L}{\partial \dot{q}^k} + \left(\frac{d}{dt} Q^k \right) \frac{\partial L}{\partial \dot{q}^k} = \frac{d}{dt} \left(Q^k \frac{\partial L}{\partial \dot{q}^k} \right) \quad (1.6.12)$$

und weiter

$$\frac{d}{dt} L = \dot{q}^k \frac{\partial L}{\partial q^k} + \ddot{q}^k \frac{\partial L}{\partial \dot{q}^k} + \frac{\partial L}{\partial t} = \dot{q}^k \frac{d}{dt} \frac{\partial L}{\partial \dot{q}^k} + \ddot{q}^k \frac{\partial L}{\partial \dot{q}^k} + \frac{\partial L}{\partial t} = \frac{d}{dt} \left(\dot{q}^k \frac{\partial L}{\partial \dot{q}^k} \right) + \frac{\partial L}{\partial t}. \quad (1.6.13)$$

Setzen wir dies in (1.6.10) für $\partial L/\partial t$ sowie (1.6.12) ein, folgt nach einigen Umformungen (*Übung!*)

$$\frac{d}{dt} \left(Q^k p_k - TH + \tilde{\Omega} \right) = 0, \quad (1.6.14)$$

⁵Die Einführung des zusätzlichen Vorzeichens ist Konvention.

wobei wir die generalisierten Impulse und die Hamilton-Funktion

$$p_k = \frac{\partial L}{\partial \dot{q}^k}, \quad H = \dot{q}^k \frac{\partial L}{\partial \dot{q}^k} - L \quad (1.6.15)$$

verwendet haben. Dies ist ein **Noether-Theorem**: Liegt eine Symmetrie unter der „infinitesimalen Transformation“ (1.6.5) vor, d.h. erfüllt die Lagrange-Funktion die Bedingung (1.6.10), so ergibt sich eine **Erhaltungsgröße** entlang der Trajektorien des Systems gemäß (1.6.14).

1.6.2 Raum-Zeit-Symmetrien

Das Newtonsche Raum-Zeit-Modell weist viele Symmetrien auf, und diese Symmetrien sollten von den fundamentalen Naturgesetzen für abgeschlossene Systeme erfüllt sein. Wir wollen nun für die Spezialfälle eines Systems aus einem bzw. zwei Teilchen die Bedingungen an die Lagrange-Funktion suchen, die sich aus dieser Einschränkung ergeben. Dies geschieht am einfachsten in kartesischen Koordinaten.

Betrachten wir zunchst den Fall für ein Teilchen. Dann wählen wir als die q^k die kartesischen Koordinaten x^k des Ortsvektors dieses Teilchens. Die Raum-Zeit-Symmetrien umfassen nun die **Homogenität von Raum und Zeit**, die **Isotropie des Raums** und die Symmetrie unter **Galilei-Boosts**, d.h. die Unabhängigkeit der Naturgesetze in Inertialsystemen, d.h. wenn man von einem Inertialsystem zu einem dazu gleichförmig geradlinig bewegten Bezugssystem übergeht.

Betrachten wir zunächst die Homogenität in der Zeit, d.h. daß beliebige Translationen in der Zeit eine Symmetrietransformation sein muß. Dies wird in unserem oben entwickelten Formalismus durch

$$t' = t + \epsilon, \quad q'^k = x^k \Rightarrow T = 1, \quad Q^k = 0 \quad (1.6.16)$$

erfüllt. Setzen wir dies in die Symmetriebedingung (1.6.10) ein, erhalten wir

$$\frac{\partial L}{\partial t} + \frac{d}{dt} \tilde{\Omega} = 0. \quad (1.6.17)$$

Offenbar läßt sich dies für $\tilde{\Omega} = 0$ und die Bedingung

$$\frac{\partial L}{\partial t} = 0 \quad (1.6.18)$$

erfüllen, d.h. die Lagrange-Funktion darf nicht explizit von der Zeit abhängen. Gemäß (1.6.14) ist die **Hamilton-Funktion**, also die Energie die der Zeittranslationssymmetrie entsprechende Erhaltungsgröße. Im folgenden gehen wir davon aus, daß (1.6.18) erfüllt ist.

Der **Homogenität des Raumes** entspricht die Translationsinvarianz in einer beliebigen Richtung \vec{n}

$$q'^k = x^k + \epsilon n^k \quad t' = t \Rightarrow Q^k = n^k = \text{const}, \quad T = 0. \quad (1.6.19)$$

Die Symmetriebedingung (1.6.10) verlangt, daß

$$n^k \frac{\partial L}{\partial x^k} + \frac{d}{dt} \tilde{\Omega} = 0 \quad (1.6.20)$$

ist. Dies läßt sich wieder mit $\tilde{\Omega} = 0$ erfüllen, und die Symmetrie verlangt dann, daß L nicht von den x^k abhängt, denn wir können $\vec{n} = \vec{e}_k$ für $k \in \{1, 2, 3\}$ setzen, d.h. alle drei kartesischen Koordinaten müssen zyklisch sein, und folglich sind gemäß (1.6.14) alle drei **Komponenten des kanonisch konjugierten Impulses**

$$L = L(\dot{x}^k), \quad p_k = \frac{\partial L}{\partial \dot{x}^k} \quad (1.6.21)$$

erhalten.

Der **Isotropie des Raumes** entspricht die Symmetrie unter Drehungen um beliebige Drehachsen. Eine infinitesimale Drehung ist durch

$$t' = t, \quad q'^k = x^k - \epsilon \delta^{ik} n^j \epsilon_{ijl} x^l \Rightarrow T = 0, \quad Q^k = -\delta^{ik} n^j \epsilon_{ijl} x^l, \quad n^j = \text{const} \quad (1.6.22)$$

definiert. Wegen $\partial L / \partial x^k = 0$ verlangt die Symmetriebedingung (1.6.10), wieder mit $\tilde{\Omega} = 0$ nur noch

$$-\delta^{ik} n^j \epsilon_{ijl} \dot{x}^l \frac{\partial L}{\partial \dot{x}^k} = 0 \quad (1.6.23)$$

Wir können wieder $\vec{n} = \vec{e}_j$ für die drei kartesischen Basisvektoren \vec{e}_j setzen. Demnach muß $\partial L / \partial \dot{x}^k \propto \dot{x}^k$ sein, d.h. es gilt

$$L = L(\dot{\vec{x}}^2). \quad (1.6.24)$$

Gemäß (1.6.14) sind die entsprechenden Erhaltungsgrößen die drei Komponenten des **Drehimpulses**, denn

$$\vec{\ell} \cdot \vec{n} := Q^k p_k = -\delta^{ik} n^j \epsilon_{ijl} x^l p_k = \vec{n} \cdot (\vec{x} \times \vec{p}). \quad (1.6.25)$$

Da wir für \vec{n} wieder die drei kartesischen Basisvektoren setzen dürfen, sind alle drei Drehimpulskomponenten

$$\vec{\ell} = \vec{x} \times \vec{p} \quad (1.6.26)$$

erhalten.

Die Invarianz unter **Galilei-Boosts** wird durch

$$t' = t, \quad q'^k = x^k - \epsilon n^k t \Rightarrow T = 0, \quad Q^k = -n^k t, \quad n^k = \text{const} \quad (1.6.27)$$

beschrieben. Mit (1.6.24) verlangt die Symmetriebedingung (1.6.10)

$$-n^k \frac{\partial L}{\partial \dot{x}^k} + \frac{d}{dt} \tilde{\Omega} = -2n^k \dot{x}_k L'(\dot{\vec{x}}^2) + \dot{x}^k \frac{\partial \tilde{\Omega}}{\partial x^k} + \frac{\partial \tilde{\Omega}}{\partial t} = 0. \quad (1.6.28)$$

Dies kann für eine Funktion $\tilde{\Omega} = \tilde{\Omega}(x^k, t)$ nur für

$$\tilde{\Omega} = n_k x^k = m \vec{n} \cdot \vec{x}, \quad m = \text{const} \quad (1.6.29)$$

erfüllt werden. Setzen wir dies in (1.6.28) ein, erhalten wir für L die Differentialgleichung

$$\frac{d}{d(\dot{\vec{x}}^2)} L = \frac{m}{2} \Rightarrow L = \frac{m}{2} \dot{\vec{x}}^2. \quad (1.6.30)$$

Gemäß (1.6.14) ist also

$$m \vec{n} \cdot \vec{x}_s = \vec{n} \cdot (m \vec{x} - \vec{p} t) \quad (1.6.31)$$

und, da wir wieder $\vec{n} = \vec{e}_j$ für alle drei kartesischen Basisvektoren \vec{e}_j setzen dürfen, der **Schwerpunktvektor**

$$\vec{K} = m \vec{X}_s = m \vec{x} - \vec{p} t \quad (1.6.32)$$

erhalten. In der Tat beschreibt (1.6.30) offensichtlich ein freies Teilchen, das sich mit der Geschwindigkeit $\dot{\vec{x}} = \vec{p}/m$ und dem Impuls $\vec{p} = m \dot{\vec{x}} = \text{const}$ geradlinig gleichförmig bewegt.

1.6.3 Zweiteilchensystem mit Wechselwirkungs-Zentralpotential

Wir können uns nun auch leicht ein abgeschlossenes Zweiteilchensystem definieren. Offensichtlich erfüllt nämlich die Lagrange-Funktion

$$L = \frac{m_1}{2} \dot{\vec{x}}_1^2 + \frac{m_2}{2} \dot{\vec{x}}_2^2 - V(r) \quad \text{mit} \quad r = |\vec{x}_1 - \vec{x}_2| \quad (1.6.33)$$

die volle Galilei-Symmetrie. Man rechnet mit Hilfe von (1.6.10) und (1.6.14) leicht nach (*Übung!*), daß die Erhaltungsgrößen durch

$$H = \frac{m_1}{2} \dot{\vec{x}}_1^2 + \frac{m_2}{2} \dot{\vec{x}}_2^2 + V(r) \quad (\text{Zeittranslationsinvarianz}), \quad (1.6.34)$$

$$\vec{P} = \vec{p}_1 + \vec{p}_2 = m_1 \dot{\vec{x}}_1 + m_2 \dot{\vec{x}}_2 \quad (\text{räumliche Translationsinvarianz}), \quad (1.6.35)$$

$$\vec{l} = m_1 \vec{x}_1 \times \dot{\vec{x}}_1 + m_2 \vec{x}_2 \times \dot{\vec{x}}_2 \quad (\text{räumliche Drehungen}), \quad (1.6.36)$$

$$\vec{x}_s = M \vec{X} - \vec{P} t, \quad M = m_1 + m_2, \quad \vec{X} = \frac{m_1 \vec{x}_1 + m_2 \vec{x}_2}{M} \quad (\text{Galilei-Boostinvarianz}). \quad (1.6.37)$$

Es liegt nun nahe, diese Symmetrien möglichst dadurch auszunutzen, daß man geeignete neue Koordinaten einführt. Offenbar eignen sich dafür als erstes, die **Schwerpunkts- und Relativkoordinaten**

$$\vec{X} = \frac{m_1 \vec{x}_1 + m_2 \vec{x}_2}{M}, \quad \vec{r} = \vec{x}_1 - \vec{x}_2. \quad (1.6.38)$$

Zur Umrechnung der Lagrange-Funktion in die neuen Koordinaten benötigen wir die Umkehrtransformation (*nachrechnen!*)

$$\vec{x}_1 = \frac{M \vec{X} + m_2 \vec{r}}{M}, \quad \vec{x}_2 = \frac{M \vec{X} - m_1 \vec{r}}{M}. \quad (1.6.39)$$

Damit ergibt sich die kinetische Energie durch einfaches Einsetzen und einfache Umformungen mit Hilfe der neuen Koordinaten (*nachrechnen!*)

$$T = \frac{m_1}{2} \dot{\vec{x}}_1^2 + \frac{m_2}{2} \dot{\vec{x}}_2^2 = \frac{M}{2} \dot{\vec{X}}^2 + \frac{\mu}{2} \dot{\vec{r}}^2 \quad \text{mit} \quad \mu = \frac{m_1 m_2}{M}. \quad (1.6.40)$$

Die kinetische Energie ist schon durch die Relativkoordinaten gegeben, d.h.

$$L = T - V = \frac{M}{2} \dot{\vec{X}}^2 + \frac{\mu}{2} \dot{\vec{r}}^2 - V(r). \quad (1.6.41)$$

Damit haben wir die **Translationssymmetrie** und **Galilei-Boostsymmetrie** bereits vollständig ausgenutzt, denn die Schwerpunktskoordinaten \vec{X} sind zyklisch und folglich der dazugehörige kanonisch konjugierte Impuls, der in diesem Fall einfach der gewöhnliche Gesamtimpuls des Systems

$$\vec{P} = \frac{\partial L}{\partial \dot{\vec{X}}} = M \dot{\vec{X}} = \text{const} \quad (1.6.42)$$

ist. Damit bewegt sich der Schwerpunkt geradlinig gleichförmig, denn wir können diese Gleichung wegen $\vec{P} = \text{const}$ sehr einfach nach der Zeit integrieren, um

$$\vec{X}(t) = \frac{\vec{P}}{M} t + \vec{X}_0, \quad (1.6.43)$$

wobei die Integrationskonstanten $\vec{X}_0 = \text{const}$ durch die Anfangsbedingungen $\vec{x}_j(t=0) = \vec{x}_j^{(0)}$ und $\dot{\vec{x}}_j(t=0) = \vec{v}_j^{(0)}$, ebenso wie \vec{P} , bei $t=0$ bestimmt sind:

$$\vec{P} = m_1 \vec{v}_1^{(0)} + m_2 \vec{v}_2^{(0)}, \quad \vec{X}_0 = \frac{m_1 \vec{x}_1^{(0)} + m_2 \vec{x}_2^{(0)}}{M}. \quad (1.6.44)$$

Die Bewegungsgleichungen für die Relativkoordinate \vec{r} entkoppeln vollständig von der nunmehr gelösten Schwerpunktsbewegung. Wir haben also das Problem der Bewegung zweier Körper mit einer Zentralkraftwechselwirkung auf ein effektives Einteilchenproblem zurückgeführt. Wir brauchen also nur noch die Lagrange-Funktion

$$L_{\text{rel}} = \frac{\mu}{2} \dot{\vec{r}}^2 - V(|\vec{r}|) \quad (1.6.45)$$

zu betrachten. Wie oben bei der Anwendung des Noether-Theorems auf den Fall der Bewegung eines Teilchens ist diese Lagrange-Funktion unter beliebigen Drehungen des Vektors \vec{r} invariant. Dies führt auf die Erhaltung des **Gesamtdrehimpulses**, den wir als

$$\vec{l} = \vec{r} \times \vec{p} \quad \text{mit} \quad \vec{p} = \frac{\partial L}{\partial \dot{\vec{r}}} = \mu \dot{\vec{r}} \quad (1.6.46)$$

schreiben können. Nun ist aber

$$\vec{r} \cdot \vec{l} = 0, \quad (1.6.47)$$

d.h. die Bewegung verläuft in der zeitlich konstanten Ebene senkrecht zum Drehimpuls \vec{l} . Geometrisch ist $dA = \frac{1}{2} dt |\vec{r} \times \dot{\vec{r}}| = \frac{l}{2\mu} dt$ die in einer kleinen Zeit dt überstrichene Fläche des Radiusvektors \vec{r} . Dies ist das **2. Keplersche Gesetz**, das demnach offensichtlich für beliebige Zentralkräfte gilt.

Legen wir nun die r^3 -Achse des Koordinatensystems in Richtung von \vec{l} und den Koordinatenursprung so, daß $r_0^3 = r^3(t=0) = 0$ ist, können wir Polarkoordinaten in der r^1 - r^2 -Ebene einführen, d.h.

$$\begin{pmatrix} r^1 \\ r^2 \\ r^3 \end{pmatrix} = \begin{pmatrix} r \cos \varphi \\ r \sin \varphi \\ r_0^3 \end{pmatrix}, \quad r_0^3 = 0 = \text{const.} \quad (1.6.48)$$

Dann ist

$$\dot{\vec{r}} = \dot{r} \begin{pmatrix} \cos \varphi \\ \sin \varphi \\ 0 \end{pmatrix} + r \dot{\varphi} \begin{pmatrix} -\sin \varphi \\ \cos \varphi \\ 0 \end{pmatrix} \Rightarrow \dot{\vec{r}}^2 = \dot{r}^2 + r^2 \dot{\varphi}^2. \quad (1.6.49)$$

Dies in (1.6.45) eingesetzt liefert

$$L_{\text{rel}} = \frac{\mu}{2} (\dot{r}^2 + r^2 \dot{\varphi}^2) - V(r). \quad (1.6.50)$$

Es ist also in diesen Koordinaten φ eine zyklische Koordinate. Dies ist zu erwarten, weil sie die Drehung des Radiusvektors in der r^1 - r^2 -Ebene beschreibt und das System unter diesen Drehungen invariant ist. Der dazugehörige kanonisch konjugierte Impuls ist

$$p_\varphi = \frac{\partial L_{\text{rel}}}{\partial \dot{\varphi}} = \mu r^2 \dot{\varphi} = \text{const.} \quad (1.6.51)$$

Dies ist aber wegen

$$\vec{l} = \mu \vec{r} \times \dot{\vec{r}} = \mu r \dot{\varphi} \vec{e}_3 \quad (1.6.52)$$

der Betrag des Drehimpulses: $p_\varphi = l$.

Nun hängt L_{rel} auch nicht explizit von der Zeit ab. Folglich ist die Hamilton-Funktion

$$H_{\text{rel}} = p_j q^j - L_{\text{rel}} = T + V = \frac{\mu}{2} (\dot{r}^2 + r^2 \dot{\varphi}^2) + V(r) = \frac{\mu}{2} \dot{r}^2 + \frac{l^2}{2\mu r^2} + V(r) = E = \text{const.} \quad (1.6.53)$$

1.7 Beispiel: Kepler-Bewegung

Als konkretes Beispiel für ein exakt lösbares Zentralkraftproblem betrachten wir die **Kepler-Bewegung**, das für die Himmelsmechanik eine herausragende Bedeutung besitzt. In unserem Sonnensystem dominiert die Masse der Sonne nämlich über die Massen aller übrigen Himmelskörper, und man kann in einer ersten Näherung zur Beschreibung der Planetenbewegung annehmen, daß man es mit einem Zweikörperproblem der im vorigen Abschnitt beschriebenen Art zu tun hat. Das Potential ist das **Newtonsche Gravitationspotential**

$$V(r) = -\frac{Gm_1m_2}{r} = -\frac{\alpha}{r} \quad (1.7.1)$$

mit der **Newtonschen Gravitationskonstanten** $G = 6,67408(31) \cdot 10^{-11} \text{ m}^3 \text{ kg}^{-1} \text{ s}^{-2}$. Wir setzen dies in (1.6.53) ein:

$$\frac{\mu}{2} \dot{r}^2 + \frac{l^2}{2\mu r^2} - \frac{\alpha}{r} = E. \quad (1.7.2)$$

Diese Differentialgleichung läßt sich zwar formal integrieren. Die Lösung ist aber nicht ganz einfach zu interpretieren.

In diesem Fall können wir aber die **Bahnkurve** bestimmen, indem wir $r(\varphi)$ berechnen. Wir nehmen dabei an, daß $l > 0$ ist. Dann ist

$$\dot{\varphi} = \frac{l}{\mu r^2} \Rightarrow \dot{r} = \frac{dr}{d\varphi} \dot{\varphi} = r'(\varphi) \frac{l}{\mu r^2}. \quad (1.7.3)$$

Dabei bedeutet der Strich die Ableitung nach φ . Setzen wir dann $r = 1/s$, folgt

$$s' = -\frac{r'}{r^2} \stackrel{(1.7.3)}{\Rightarrow} \dot{r} = \frac{l}{\mu} s'. \quad (1.7.4)$$

Einsetzen (1.7.2) ergibt dann

$$(s')^2 + s^2 - \frac{2\mu\alpha}{l^2} s = \frac{2\mu E}{l^2}. \quad (1.7.5)$$

Die Lösung vereinfacht sich, indem wir diese Gleichung nochmals nach φ ableiten:

$$s' \left(s'' + s - \frac{\mu\alpha}{l^2} \right) = 0. \quad (1.7.6)$$

Wir haben also entweder $s' = 0$, oder die Klammer muß verschwinden.

Der erste Fall $s' = 0$ bedeutet, daß $r = \text{const}$ ist, d.h. \vec{r} beschreibt eine Kreisbahn. Im zweiten Fall gilt

$$s'' + s = \frac{\mu\alpha}{l^2}. \quad (1.7.7)$$

Dies ist eine **lineare Differentialgleichung 2. Ordnung mit konstanten Koeffizienten**. Wie in [Hee14] in Kapitel 5 ausführlich gezeigt, ergibt sich die allgemeine Lösung dieser Gleichung durch die Summe aus der allgemeinen Lösung der homogenen Gleichung

$$s''_{\text{hom}} + s_{\text{hom}} = 0 \quad (1.7.8)$$

und irgendeiner beliebigen Lösung der inhomogenen Gleichung. Die allgemeine Lösung der homogenen Gleichung ist offenbar durch

$$s_{\text{hom}}(\varphi) = A \cos(\varphi + \varphi_0) \quad (1.7.9)$$

gegeben. Offenbar ist weiter

$$s_{\text{inh}}(\varphi) = \frac{\mu\alpha}{l^2} \quad (1.7.10)$$

eine spezielle Lösung der inhomogenen Gleichung. Wir können die allgemeine Lösung von (1.7.7) also in der Form

$$s(\varphi) = \frac{\mu\alpha}{l^2} [1 + \epsilon \cos(\varphi + \varphi_0)] \Rightarrow r(\varphi) = \frac{1}{s(\varphi)} = \frac{l^2}{\mu\alpha} \frac{1}{1 + \epsilon \cos(\varphi + \varphi_0)}. \quad (1.7.11)$$

Dabei sind ϵ und φ_0 Integrationskonstanten. Wir können offenbar durch Drehen der Basis in der r^1 - r^2 -Ebene erreichen, daß die Integrationskonstante $\varphi_0 = 0$ und $\epsilon > 0$ sind. Dann kennzeichnet $\varphi = 0$ den maximalen Wert von s , d.h. $\varphi = 0$ kennzeichnet den minimalen Wert von r entlang der Bahn der beiden Körper. Für die Planetenbewegung um die Sonne ist das der Punkt, an dem sich Planet und Sonne am nächsten kommen, der **Perihel**. Dann wird

$$r(\varphi) = \frac{p}{1 + \epsilon \cos \varphi} \quad \text{mit} \quad p = \frac{l^2}{\mu\alpha}. \quad (1.7.12)$$

Das ist aber gemäß der Betrachtung in [Hee14] (Abschnitt 3.3) ein **Kegelschnitt** mit dem Koordinatenursprung in einem Brennpunkt, und zwar für $\epsilon = 0$ ein **Kreis**, für $0 < \epsilon < 1$ eine **Ellipse**, für $\epsilon = 1$ eine Parabel und für $\epsilon > 1$ eine Hyperbel.

Wir können leicht verstehen, warum es sowohl gebundene Kreis- bzw. Ellipsenbahnen als auch ungebundene Parabel- bzw. Hyperbelbahnen gibt, indem wir auf (1.7.5) zurückgehen. Dazu setzen wir die Lösung (1.7.11) in (1.7.12) ein. Dies ergibt nach einfachen Umformungen (*Übung!*)

$$\epsilon^2 = 1 + \frac{2l^2 E}{\mu\alpha^2}. \quad (1.7.13)$$

Wir haben also für $E < 0$ eine Ellipsen- oder Kreisbahn, für $E = 0$ eine Parabelbahn und für $E > 0$ eine Hyperbelbahn. Dies wird anhand der folgenden Skizze klar, in der das **effektive Potential**

$$V_{\text{eff}}(r) = -\frac{\alpha}{r} + \frac{l^2}{2\mu r^2} \quad (1.7.14)$$

geplottet ist. Gemäß (1.7.13) muß offenbar $E \geq -\mu\alpha^2/(2l^2)$ sein, damit $\epsilon^2 \geq 0$ sein kann. Für den Minimalwert berührt E das Potential nur in seinem Minimum, und demzufolge muß $r = \text{const}$ sein, denn Werte für r , wo $V_{\text{eff}}(r) > E$ wird, kann das System entlang seiner Bahn nicht erreichen, weil $T_{\text{eff}} = \mu\dot{r}^2/2 \geq 0$ ist und folglich (1.7.2) nicht erfüllt sein kann. Für größere Werte schneidet die Linie $E = \text{const}$ den Graphen von V_{eff} entweder in zwei Punkten (für $-\mu\alpha^2/(2l^2) < E < 0$) oder in genau einem Punkt (für $E \geq 0$). Im ersteren Fall, bleibt r auf das Intervall zwischen diesen beiden Schnittpunkten des Graphen von V_{eff} beschränkt. An diesen Punkten wird $T_{\text{eff}} = 0$, und die Bewegung kehrt sich um. In diesem Fall ist die Bewegung also periodisch entsprechend der Bewegung der Körper auf Ellipsen- bzw. Kreisbahnen um den gemeinsamen Schwerpunkt, der ein Brennpunkt der Ellipse bzw. der Mittelpunkt des Kreises ist. Das ist das **1. Keplersche Gesetz**. Für $E \geq 0$ erreichen die Körper den Punkt größter Annäherung, worauf sich die Bewegung umkehrt und sie sich wieder voneinander entfernen. Die beiden Körper beschreiben dann beide Parabel- bzw. Hyperbelbahnen mit dem gemeinsamen Schwerpunkt als Brennpunkt.

Wir können schließlich auch noch das **3. Keplersche Gesetz** für einen Planeten herleiten. In der Umlaufzeit T überstreicht \vec{r} die Fläche der Ellipsenbahn $A = \pi a b$ andererseits ist

$$A = \frac{lT}{2\mu}. \quad (1.7.15)$$

Wir müssen also nur noch

$$T = \frac{2\pi\mu a b}{l} \quad (1.7.16)$$

mit Hilfe von a und Konstanten ausdrücken. Dazu verwenden wir die Geometrie der Ellipse aus [Hee14] (Abschnitt 3.3). Demnach ist

$$T = \frac{2\pi\mu}{l} a \sqrt{a^2 - e^2} = \frac{2\pi\mu}{l} a^2 (1 - \epsilon^2). \quad (1.7.17)$$

Andererseits ist

$$p = \frac{b^2}{a} = a(1 - \epsilon^2) \Rightarrow a = \frac{p}{1 - \epsilon^2} \stackrel{((1.7.12-1.7.13))}{=} \frac{\alpha}{2E} \Rightarrow E = -\frac{\alpha}{2a}. \quad (1.7.18)$$

bzw.

$$1 - \epsilon^2 = -\frac{2l^2 E}{\mu \alpha^2} = \frac{l^2}{\mu \alpha a}. \quad (1.7.19)$$

Quadrieren wir (1.7.17) und verwenden (1.7.19), erhalten wir schließlich

$$T^2 = \frac{4\pi^2 \mu a^3}{\alpha} \stackrel{(1.6.40, 1.7.1)}{=} \frac{4\pi^2}{G(m_1 + m_2)} a^3. \quad (1.7.20)$$

Falls nun, wie in unserem Sonnensystem $m_1 \gg m_2$ (wobei m_1 die Masse der Sonne und m_2 die Masse des Planeten ist), gilt

$$T^2 = \frac{4\pi^2}{G(m_1 + m_2)} a^3 \approx \frac{4\pi^2}{G m_1} a^3, \quad (1.7.21)$$

d.h. die Quadrate der Umlaufzeiten verhalten sich wie die Kuben der großen Halbachsen der Kepler-Ellipsen, und dies ist das **3. Keplersche Gesetz**.

1.8 Die Hamiltonsche kanonische Mechanik

In diesem Abschnitt kommen wir nun auf die im Hinblick auf die Erweiterung der Physik im Sinne der Quantentheorie wichtigste Formulierung der Mechanik, den **Hamilton-Formalismus**. Auch innerhalb der klassischen Mechanik ermöglicht diese Formulierung die tiefgründigste Behandlung der Symmetrieprinzipien und elegante Methoden zur Lösung der Bewegungsgleichung.

Die Grundidee besteht darin, das Hamiltonsche Prinzip der kleinsten Wirkung im Konfigurationsraum, parametrisiert durch verallgemeinerte Koordinaten q^k , zu einem Hamiltonschen Prinzip im sog. Phasenraum zu erweitern. Der Phasenraum beschreibt dabei die Bewegung des Teilchens durch die verallgemeinerten Koordinaten q^k und die dazugehörigen **kanonisch konjugierten Impulse**

$$p_j = \frac{\partial L}{\partial \dot{q}^j}. \quad (1.8.1)$$

Wir gehen dabei davon aus, daß sich die \dot{q}^k vermöge (1.8.1) eindeutig als Funktionen der q^k und p_j ausdrücken lassen. Es gibt zwar Fälle, sog. **singuläre Probleme**, bei denen dies nicht der Fall ist, aber wir behandeln diesen recht komplizierten Fall in dieser Vorlesung nicht. Der interessierte Leser sei auf [Dir58] verwiesen.

Im regulären Fall, also wenn sich die \dot{q}^j vermöge (1.8.1) als Funktionen der q^k und p_j schreiben lassen (und evtl. auch explizit von der Zeit, falls die Lagrange-Funktion L explizit zeitabhängig ist), folgt nun für das totale Differential der **Hamilton-Funktion**

$$H = p_k \dot{q}^k - L \Rightarrow dH = \left(p_k - \frac{\partial L}{\partial \dot{q}^k} \right) d\dot{q}^k + \dot{q}^k dp_k - \frac{\partial L}{\partial q^k} dq^k - \frac{\partial L}{\partial t} dt. \quad (1.8.2)$$

Wegen (1.8.1) verschwindet die Klammer vor den $d\dot{q}^k$, und es folgt

$$dH = \dot{q}^k dp_k - \frac{\partial L}{\partial q^k} dq^k - \frac{\partial L}{\partial t} dt. \quad (1.8.3)$$

Da wir voraussetzungsgemäß die \dot{q}^k durch die p_j und q^k ausdrücken können, können wir also $H = H(q^k, p_j, t)$ annehmen, d.h. die Hamilton-Funktion als Funktion der generalisierten Koordinaten q^k ,

der dazugehörigen kanonisch konjugierten Impulse p_j und evtl. explizit der Zeit t betrachten. Der Vergleich mit (1.8.4) ergibt dann

$$\dot{q}^k = \frac{\partial H}{\partial p_k}, \quad \frac{\partial L}{\partial q^k} = -\frac{\partial H}{\partial q^k}, \quad \frac{\partial L}{\partial t} = -\frac{\partial H}{\partial t}. \quad (1.8.4)$$

Betrachten wir nun die **Bewegungsgleichungen**. Aus dem Hamiltonschen Prinzip, wie wir es oben in Abschnitt 1.5 hergeleitet haben, gelten für die Lösungen der Bewegungsgleichungen die **Euler-Lagrange-Gleichungen**

$$\dot{p}_k = \frac{d}{dt} \frac{\partial L}{\partial \dot{q}^k} = \frac{\partial L}{\partial q^k}. \quad (1.8.5)$$

Wegen (1.8.4) gilt also

$$\dot{q}^k = \frac{\partial H}{\partial p_k}, \quad \dot{p}_j = -\frac{\partial H}{\partial q^j}. \quad (1.8.6)$$

Dies sind die **Hamiltonschen kanonischen Gleichungen**. Sie sind eine weitere äquivalente Formulierung der Bewegungsgleichungen der Newtonschen Mechanik, soweit es sich um ein dynamisches System handelt, das sich mit dem Lagrange-Formalismus beschreiben läßt für das sich die generalisierten Geschwindigkeiten \dot{q}^k als Funktionen der q^k und p_j ausdrücken lassen (**reguläre Hamiltonsche Systeme**). Es handelt sich nun um ein System gewöhnlicher Differentialgleichungen für eine **Trajektorie im Phasenraum**, der durch (q^k, p_j) aufgespannt wird. Hat man also f voneinander unabhängige generalisierte Koordinaten q^k , ist der Phasenraum $2f$ -dimensional. Die Bewegungsgleichungen für die q^k aus dem Hamiltonschen Prinzip waren zweiter Ordnung, und ihre Lösung wird eindeutig durch Vorgabe von Anfangsbedingungen q^k und \dot{q}^j zu einer Zeit t_0 bestimmt werden. Entsprechend sind die kanonischen Gleichungen (1.8.6) Differentialgleichungen 1. Ordnung für die q^k und p_j , deren Lösungen eindeutig durch Vorgabe von Anfangsbedingungen q^k und p_j zur Zeit $t = t_0$ eindeutig bestimmt werden. Da wir die p_j durch die q^k und \dot{q}_j ausdrücken können, bestimmen die Anfangsbedingungen für die Bewegungsgleichungen in der Lagrange-Form auch eindeutig die Anfangsbedingungen im Phasenraum, wie sie für die Lösung der kanonischen Gleichung vorgegeben werden müssen. Da wir weiter angenommen haben, daß das Hamiltonsche System regulär ist, bestimmt auch umgekehrt die Vorgabe der Anfangsbedingung im Phasenraum eindeutig die Anfangsbedingung im Raum der q^k und \dot{q}^k . Die Bewegungsgleichungen in Lagrangescher Formulierung und die Hamiltonschen kanonischen Gleichungen sind also in der Tat vollständig zueinander äquivalent.

Als nächstes folgt aber der entscheidende Vorteil der Hamiltonschen Formulierung, der darin besteht, daß wir das Hamilton-Prinzip der kleinsten Wirkung bzgl. Variationen im Konfigurationsraum erweitern können zu einem Wirkungsprinzip bzgl. *Variationen im Phasenraum*.

Wir bemerken nun, daß wir die Hamiltonschen kanonischen Gleichungen (1.8.6) auch aus einem **erweiterten Prinzip der kleinsten Wirkung** erhalten können. Dazu schreiben wir die Wirkung als Wirkung im Phasenraum, indem wir (1.8.2) verwenden, um die Lagrange-Funktion durch die Hamilton-Funktion auszudrücken. Entscheidend ist dabei nun aber, daß wir die Variationsmöglichkeiten dahingehend erweitern, daß wir die generalisierten Koordinaten q^k und die kanonisch konjugierten Impulse p_j *unabhängig* voneinander variieren dürfen. Dabei sollen die Variationen der q^k den üblichen festen Randbedingungen der Trajektorien im Konfigurationsraum wie beim ursprünglichen Hamilton-Prinzip erfüllen, also $\delta q^k(t_1) = \delta q^k(t_2) \equiv 0$, während wir den Variationen der p_j keinerlei Randbedingungen auferlegen. Dann lautet das **Phasenraum-Wirkungsfunktional**

$$S[q^k, p_j] = \int_{t_1}^{t_2} dt (\dot{q}^k p_k - H). \quad (1.8.7)$$

Die Variation lautet dann

$$\delta S = \int_{t_1}^{t_2} dt \left[\delta \dot{q}^k p_k + \left(\dot{q}^k - \frac{\partial H}{\partial p_k} \right) \delta p_k - \frac{\partial H}{\partial q^k} \delta q^k \right]. \quad (1.8.8)$$

Da beim Hamiltonschen Variationsprinzip die Zeit nicht mitvariiert wird, gilt $\delta \dot{q}^k = d(\delta q^k)/dt$, und wir können im ersten Term partiell integrieren und die Randbedingungen $\delta q^k(t_1) = \delta q^k(t_2)$ ausnutzen. Dann folgt

$$\delta S = \int_{t_1}^{t_2} dt \left[\left(\dot{q}^k - \frac{\partial H}{\partial p_k} \right) \delta p_k - \left(\dot{p}_k + \frac{\partial H}{\partial q_k} \right) \delta q^k \right]. \quad (1.8.9)$$

Dann wird δS für alle nun als *unabhängig* voneinander angenommen Variationen δq^k und δp_k genau dann stationär, wenn die Hamiltonschen kanonischen Gleichungen (1.8.6) erfüllt sind, d.h. sie folgen aus dem eben betrachteten **verallgemeinerten Hamiltonschen Prinzip der kleinsten Wirkung für Phasenraumtrajektorien**.

1.8.1 Beispiel: Der harmonische Oszillator

Als einfaches Beispiel betrachten wir den harmonischen Oszillator in der Hamilton-Formulierung des Wirkungsprinzips. Die Lagrange-Funktion lautet

$$L = \frac{m}{2} \dot{x}^2 - \frac{m\omega^2}{2} x^2. \quad (1.8.10)$$

Der kanonisch konjugierte Impuls ist

$$p = \frac{\partial L}{\partial \dot{x}} = m\dot{x} \Rightarrow \dot{x} = \frac{p}{m}. \quad (1.8.11)$$

Die Geschwindigkeit \dot{x} läßt sich also eindeutig durch p ausdrücken, und es liegt demzufolge ein reguläres Hamiltonsches System vor. Daraus folgt eindeutig die Hamiltonfunktion als Funktion von x und p , also

$$H = \dot{x}p - L = m\dot{x}^2 - L = \frac{m}{2} \dot{x}^2 + \frac{m\omega^2}{2} x^2 = \frac{p^2}{2m} + \frac{m\omega^2}{2} x^2. \quad (1.8.12)$$

Die kanonischen Bewegungsgleichungen lauten demzufolge gemäß (1.8.6)

$$\dot{x} = \frac{\partial H}{\partial p} = \frac{p}{m}, \quad \dot{p} = -\frac{\partial H}{\partial x} = -m\omega^2 x. \quad (1.8.13)$$

Setzt man nun die erste in die zweite Gleichung ein, folgt

$$\dot{p} = m\ddot{x} = -m\omega^2 x \Rightarrow \ddot{x} = -\omega^2 x, \quad (1.8.14)$$

also wie zu erwarten die korrekte Bewegungsgleichung für den harmonischen Oszillator.

1.9 Kanonische Transformationen

Die Bedeutung des erweiterten Hamiltonschen Prinzips die Bewegungsgleichung nicht nur forminvariant unter allgemeinem Wechsel der Koordinaten q^j des Konfigurationsraums (in Gestalt der Euler-Lagrange-Gleichungen) sind, sondern auch unter allgemeineren die generalisierten Impulse einschließenden Transformationen, die die Variation des im Phasenraum formulierten Wirkungsfunktional (1.8.7) invariant lassen. Solche Transformationen bezeichnen wir mit Hamilton als **kanonische Transformationen**.

Wir gehen nun wie bei der analogen Frage bei der Lagrangeschen Formulierung in Abschnitt 1.5.2 vor, wobei wir die am Ende gegebene Betrachtung anhand des Wirkungsfunktional ausnutzen.

Um die Bedingungen an eine beliebige Transformation

$$q^k = q^k(Q^k, P_k, t), \quad p_k = p_k(Q^k, P_k, t) \quad (1.9.1)$$

dafür zu finden, daß sie die Hamiltonschen kanonischen Gleichungen forminvariant lassen, müssen wir nur verlangen, daß die Variation des Wirkungsfunktional (1.8.7) in beiden kanonischen Koordinatensystemen gleich ist, wobei wir zulassen wollen, daß für die neuen kanonischen Koordinaten (Q^k, P_k) auch eine neue Hamilton-Funktion $H'(Q^k, P_k, t)$ eingeführt werden kann. Die Variation der Wirkung bleibt gemäß dieser Betrachtung invariant unter dieser kanonischen Transformation, wenn

$$\Delta I = \int_{t_1}^{t_2} dt [p_k \dot{q}^k - H - P_k \dot{Q}^k + H'], \quad (1.9.2)$$

gelesen als Wegintegral im sogenannten erweiterten von (t, q, p) parametrisierten Phasenraum, ein totales Differential

$$dt(H' - H) + dq^k p_k - dQ^k P_k = df \quad (1.9.3)$$

sein muß, wobei f gemäß der auf der linken Seite auftretenden Differentiale als eine Funktion von q, Q und t aufzufassen ist. Der Vergleich zwischen der linken und rechten Seite zeigt weiterhin, daß

$$H' - H = \partial_t f, \quad p_k = \frac{\partial f}{\partial q^k}, \quad P_k = -\frac{\partial f}{\partial Q^k} \quad (1.9.4)$$

gilt. Die Transformation (1.9.1) ist also genau dann eine kanonische Transformation, wenn es eine Funktion f der alten und der neuen generalisierten Koordinaten q^k und Q^k gibt, so daß die Beziehungen (1.9.4) gelten. Nach dem Lemma von Poincaré ist das wenigstens lokal nur dann der Fall, wenn

$$\frac{\partial p_k}{\partial Q^l} = -\frac{\partial P_l}{\partial q^k} \quad (1.9.5)$$

ist.

Es ist klar, daß es viel einfacher ist, wenn wir die Funktion f willkürlich vorgeben und die „alten Koordinaten“ (q, p) mit Hilfe der Bedingung (1.9.4) durch die „neuen Koordinaten“ (Q, P) ausdrücken. Wir nennen daher f auch **Erzeugende der kanonischen Transformation**. Ist f explizit zeitabhängig, dürfen wir dabei nicht vergessen, auch die Hamiltonfunktion gemäß (1.9.4) zu transformieren.

Es ist manchmal allerdings bequemer, die erzeugende Funktion mit Hilfe anderer Paare alter und neuer Phasenraumkoordinaten auszudrücken. Hier bewährt sich das schon bei der Herleitung der Hamiltonschen kanonischen Gleichungen angewandte Prinzip der **Legendre-Transformation**. Als Beispiel leiten wir den für das folgende wichtigsten Fall her, daß wir die Erzeugende als Funktion g der alten Konfigurationsraumkoordinaten q und der neuen kanonischen Impulse P vorgeben. Dann schreiben wir

$$f(q, Q, t) = g_1(q, P, t) - Q^k P_k \Rightarrow df = dq^k \frac{\partial g_1}{\partial q^k} - dQ^k P_k + \left(\frac{\partial g_1}{\partial P_k} - Q^k \right) dP_k + \frac{\partial g_1}{\partial t} dt. \quad (1.9.6)$$

Das bedeutet, daß

$$Q^k = \frac{\partial g_1}{\partial P_k} \quad (1.9.7)$$

sein muß, damit f die geforderte Abhängigkeit von q und Q hat. Setzen wir (1.9.6) in (1.9.4) ein, finden wir, daß die kanonische Transformation durch g gemäß

$$H' = H + \partial_t g_1, \quad p_k = \frac{\partial g_1}{\partial q^k}, \quad Q^k = \frac{\partial g_1}{\partial P_k} \quad (1.9.8)$$

erzeugt wird. Aus der Vertauschbarkeit der zweiten Ableitungen von g_1 nach den q und P folgt daraus die Beziehung

$$\frac{\partial p_k}{\partial P_l} = \frac{\partial Q^l}{\partial q^k}. \quad (1.9.9)$$

Als nächstes betrachten wir die Legendre-Transformation

$$f(q, Q, t) = g_2(p, Q, t) + q^k p_k \Rightarrow df = dp_k \left(\frac{\partial g_2}{\partial p_k} + q^k \right) + dQ^k \frac{\partial g_2}{\partial Q^k} + dq^k p_k + \frac{\partial g_2}{\partial t} dt. \quad (1.9.10)$$

Daraus folgt wieder wie im bei (1.9.6)

$$H' = H + \partial_t g_2, \quad q^k = -\frac{\partial g_2}{\partial p_k}, \quad P_k = -\frac{\partial g_2}{\partial Q^k}. \quad (1.9.11)$$

Die Vertauschbarkeit der zweiten Ableitung von g_2 nach p und Q liefert damit die Bedingung

$$\frac{\partial q^k}{\partial Q^l} = \frac{\partial P_l}{\partial p_k}. \quad (1.9.12)$$

Schließlich kombinieren wir beide Legendretransformationen (1.9.6) und (1.9.10) indem wir von g_2 ausgehen:

$$g_2(p, Q, t) = g_3(p, P, t) - Q^k P_k \Rightarrow dg_2 = dp_k \frac{\partial g_3}{\partial p_k} + dP_k \left(\frac{\partial g_3}{\partial P_k} - Q^k \right) - dQ^k P_k + \frac{\partial g_3}{\partial t} dt. \quad (1.9.13)$$

Mit (1.9.11) folgt daraus

$$H' = H + \partial_t g_3, \quad q^k = -\frac{\partial g_3}{\partial p_k}, \quad Q^k = \frac{\partial g_3}{\partial P_k}. \quad (1.9.14)$$

Die Vertauschbarkeit der zweiten Ableitungen von g_3 liefert schließlich die Bedingung

$$\frac{\partial q^k}{\partial P_l} = -\frac{\partial Q^l}{\partial p_k}. \quad (1.9.15)$$

Aus den Bedingungen (1.9.5, 1.9.9, 1.9.12, 1.9.15) können wir nun die Unabhängigkeit der sog. **Poisson-Klammer** von der Wahl der kanonisch konjugierten Phasenraumvariablen nachweisen, also die **Kovarianz der Poisson-Klammer** bzgl. kanonischer Transformationen. Dabei ist die Poisson-Klammer für zwei beliebige Funktion $A(q, p)$ und $B(q, p)$ durch

$$\{A, B\} = \frac{\partial A}{\partial q^k} \frac{\partial B}{\partial p_k} - \frac{\partial B}{\partial q^k} \frac{\partial A}{\partial p_k}. \quad (1.9.16)$$

Dazu schreiben wir für zwei beliebige Phasenraumfunktionen A und B die Poisson-Klammern mit den neuen Phasenraumkoordinaten (Q, P) an,

$$\begin{aligned} \{A, B\}^{(Q, P)} &= \frac{\partial A}{\partial Q^k} \frac{\partial B}{\partial P_k} - (A, B) \\ &= \left(\frac{\partial A}{\partial q^j} \frac{\partial q^j}{\partial Q^k} + \frac{\partial A}{\partial p_j} \frac{\partial p_j}{\partial Q^k} \right) \left(\frac{\partial B}{\partial q^m} \frac{\partial q^m}{\partial P_k} + \frac{\partial B}{\partial p_m} \frac{\partial p_m}{\partial P_k} \right) - (A, B), \end{aligned} \quad (1.9.17)$$

wobei (A, B) für den Ausdruck steht, der durch den voranstehenden Term durch Vertauschen von A mit B hervorgeht. Ausmultiplizieren der Klammern und Berücksichtigung der Antisymmetrisierung bzgl. Vertauschen von A und B ergibt unter Zuhilfenahme der Bedingungen (1.9.5, 1.9.9, 1.9.12, 1.9.15) in der Tat die Poissonklammer geschrieben in den „alten Variablen“ (q, p) . Als Beispiel für diese Rechnung betrachten wir den antisymmetrisierten ersten Term

$$\frac{\partial A}{\partial q^j} \frac{\partial B}{\partial q^m} \left(\frac{\partial q^j}{\partial Q^k} \frac{\partial q^m}{\partial P_k} - \frac{\partial q^m}{\partial Q^k} \frac{\partial q^j}{\partial P_k} \right) = -\frac{\partial A}{\partial q^j} \frac{\partial B}{\partial q^m} \underbrace{\left(\frac{\partial q^j}{\partial Q^k} \frac{\partial q^m}{\partial P_k} + \frac{\partial q^j}{\partial P_k} \frac{\partial q^m}{\partial Q^k} \right)}_{\partial q^j / \partial p_m} = 0. \quad (1.9.18)$$

Zusammenfassen der übrigen Terme ergibt dann in der Tat

$$\{A, B\}^{(Q, P)} = \{A, B\}^{(q, p)} := \{A, B\}. \quad (1.9.19)$$

1.10 Die symplektische Struktur des Phasenraums und Poisson-Klammern

In diesem Abschnitt verwenden wir das erweiterte Hamiltonsche Prinzip um einen Einblick in die lokale geometrische Struktur des Phasenraums zu gewinnen. Dies liefert auch noch weitere Einsichten zur Poisson-Klammer und der eben hergeleiteten Forminvarianz unter kanonischen Transformationen.

Dazu gehen wir von der oben hergeleiteten Kovarianzbedingung (1.9.3) aus, die wir jedoch in der Abhängigkeit von den neuen Koordinaten (Q, P) und in seiner Gestalt in Form von Zeitableitungen anschreiben, d.h. wir setzen $F(Q, P, t) = f[q(Q, P), Q, t]$:

$$p_k \dot{q}^k - H(q^k, p_k, t) = P^k \dot{Q}_k - H'(Q^k, P_k, t) + \frac{d}{dt} F(Q, P, t). \quad (1.10.1)$$

Integration dieser Gleichung bzgl. t von t_1 bis t_2 und Substitution der (q, p) zugunsten der (Q, P) auf der linken Seite ergibt schließlich

$$\int_{t_1}^{t_2} dt \left[\dot{Q}^j \left(P_j - p_k \frac{\partial q^k}{\partial Q^j} \right) - p_k \dot{P}_j \frac{\partial q^k}{\partial P_j} - (H' - H) \right] = -F(Q, P, t) \Big|_{t_1}^{t_2}. \quad (1.10.2)$$

Jetzt fassen wir das linksstehende Integral als Kurvenintegral im sogenannten erweiterten Phasenraum, der durch die $(2f + 1)$ Parameter (Q, P, t) parametrisiert wird, auf. Da die rechte Seite nur von Anfangs- und Endpunkt der Kurve abhängt, muß nach dem Poincaréschen Lemma der Integrand ein totales Differential sein. Dabei ist $-F$ das Potential des über die Kurve integrierten Feldes im erweiterten Phasenraum:

$$\frac{\partial(-F)}{\partial Q^j} = P_j - p_k \frac{\partial q^k}{\partial Q^j}, \quad \frac{\partial(-F)}{\partial P_l} = -p_k \frac{\partial q^k}{\partial P_l}, \quad \frac{\partial(-F)}{\partial t} = -(H' - H). \quad (1.10.3)$$

Das ist aber nur der Fall, wenn die Integrabilitätsbedingungen, die aus der Vertauschbarkeit der partiellen Ableitungen von F nach den Q und P folgen, erfüllt sind:

$$(Q^j, P_l)_{\text{lb}} := \frac{\partial q^k}{\partial Q^j} \frac{\partial p_k}{\partial P_l} - \frac{\partial q^k}{\partial P_l} \frac{\partial p_k}{\partial Q^j} = \delta_{jl}, \quad (Q^j, Q^l)_{\text{lb}} = (P_j, P_l)_{\text{lb}} = 0. \quad (1.10.4)$$

Das für beliebige Phasenraumtransformationen $q(Q, P, t)$, $p(Q, P, t)$ definierte Objekt $(X, Y)_{\text{lb}}$, wobei für X und Y beliebige neue Phasenraumkoordinaten (Q, P) eingesetzt werden können, heißt **Lagrange-Klammer**. Eine **kanonische Transformation** liegt also genau dann vor, wenn die Lagrangeklammern invariant unter dieser Transformation sind. Es ist nämlich klar, daß für die identische Transformation $Q = q$, $P = p$ die Lagrangeklammerbeziehungen erfüllt sind.

Die Lagrangeklammer stellt nun aber keine besonders bequem zu berechnende Konstruktion dar, wenn es darum geht, für gegebene neue Koordinaten (Q, P) zu entscheiden, ob sie eine kanonische Transformation der (q, p) darstellen. Wir wollen vielmehr die Bedingungen (1.10.4) durch Ableitungen der neuen Phasenraumkoordinaten nach den alten ausdrücken. Dazu betrachten wir die Jacobimatrix der inversen kanonischen Transformation

$$J = \begin{pmatrix} \partial Q^i / \partial q^j & \partial Q^i / \partial p_j \\ \partial P_i / \partial q^j & \partial P_i / \partial p_j \end{pmatrix}. \quad (1.10.5)$$

Dabei sollen die partiellen Ableitungen jeweils als $f \times f$ -Blockmatrix verstanden werden. Damit ist J insgesamt eine $(2f) \times (2f)$ -Matrix. Die Bedingungen (1.10.4) lassen sich nun durch die inverse Matrix

$$J^{-1} = \begin{pmatrix} \partial q^j / \partial Q^i & \partial p_j / \partial Q^i \\ \partial q^j / \partial P_i & \partial p_j / \partial P_i \end{pmatrix} \quad (1.10.6)$$

ausdrücken. Dazu führen wir die sogenannte **symplektische Fundamentalmatrix**

$$S = \begin{pmatrix} 0 & \delta_i^j \\ -\delta_i^j & 0 \end{pmatrix} \quad (1.10.7)$$

ein. Offenbar gilt $S^{-1} = -S = S^t$. Man bezeichnet einen $(2f)$ -dimensionalen Vektorraum mit einer antisymmetrischen Bilinearform dieser Art als **symplektischen Vektorraum**. Analog dazu, wie im euklidischen Raum die Einheitsmatrix das kanonische Skalarprodukt darstellt, gilt dies für die antisymmetrische Bilinearform eines symplektischen Vektorraums mit der symplektischen Fundamentalmatrix S .

Das Analogon zu den orthogonalen Transformationen im euklidischen Raum entsprechen bilden im symplektischen Raum die **symplektischen Abbildungen**, die entsprechend als diejenigen umkehrbar eindeutigen linearen Abbildungen im \mathbb{R}^{2f} definiert sind, die das durch S definierte symplektische Produkt $(x, y)_s = x^t S y$ zweier beliebiger Vektoren x und y invariant lassen. Ist T die darstellende Matrix einer symplektischen Abbildung, dann muß gelten $T^t S T = S$.

Man kann nun die Bedingungen (1.10.4) kurz wie folgt schreiben:

$$(J^{-1})^t S^t J^{-1} = S^t \Rightarrow J S J^t = S. \quad (1.10.8)$$

Das bedeutet aber, daß J in jedem Punkt des Phasenraums eine symplektische Transformation ist. Demnach ist eine Transformation $(q, p) \leftrightarrow (Q, P)$ genau dann kanonisch, wenn sie **lokal symplektisch** ist. Man bezeichnet solche lokalen symplektischen Transformationen einer Mannigfaltigkeit auch kurz als **Symplektomorphismen**.

Die in (1.10.8) letztgenannte Form der Bedingungen (1.10.4) lassen sich in Komponenten zerlegt kurz wie folgt schreiben

$$\{Q^j, P_l\} := \frac{\partial Q^j}{\partial q^k} \frac{\partial P_l}{\partial p_k} - \frac{\partial Q^j}{\partial p_k} \frac{\partial P_l}{\partial q^k} = \delta_l^j, \quad \{Q^j, Q^l\} = \{P_j, P_l\} = 0. \quad (1.10.9)$$

Damit ist aber eine **manifest kovariante Charakterisierung** der lokal symplektischen Struktur des Phasenraums durch die bereits im vorigen Abschnitt eingeführten **Poisson-Klammern**

$$\{A, B\} = \frac{\partial A}{\partial q^k} \frac{\partial B}{\partial p_k} - \frac{\partial A}{\partial p_k} \frac{\partial B}{\partial q^k} \quad (1.10.10)$$

gegeben, wobei A und B beliebige im Phasenraum definierte Funktionen sind.

Auch die Bewegungsgleichungen lassen sich leicht mit Hilfe von Poissonklammern formulieren. Die totale Zeitableitung einer im Phasenraum definierten Funktion entlang der durch die Bewegung des Systems gegebenen Trajektorie, ist nämlich offenbar durch

$$\frac{dA}{dt} = \frac{\partial A}{\partial q^k} \dot{q}^k + \frac{\partial A}{\partial p_k} \dot{p}_k + \frac{\partial A}{\partial t} = \{A, H\} + \frac{\partial A}{\partial t} \quad (1.10.11)$$

gegeben. Dabei haben wir die Hamiltonschen kanonischen Gleichungen (1.8.6) benutzt und einer eventuellen expliziten (d.h. nicht durch die Trajektorie $[q(t), p(t)]$ eingepprägten) Zeitabhängigkeit, von A Rechnung getragen.

Wir betrachten nun noch einige wichtige Eigenschaften der Poisson-Klammer, die zeigen, daß sie auf dem Vektorraum der glatten Phasenraumfunktionen (d.h. beliebig oft partiell stetig differenzierbaren Funktionen $A(q, p)$) ein sog. **Lie-Produkt** bilden, d.h. mit der Poisson-Klammer als Produkt wird dieser Funktionenvektorraum V zu einer **Lie-Algebra**. Das Lie-Produkt ist zunächst eine bilineare Abbildung $V \times V \rightarrow V$, d.h.

sie bildet zwei glatte Phasenraumfunktionen in eine neue glatte Phasenraumfunktion ab, und für beliebige Phasenraumfunktionen A, B und C und beliebige $\lambda_1, \lambda_2 \in \mathbb{R}$ gilt (*nachprüfen!*)

$$\{\lambda_1 A + \lambda_2 B, C\} = \lambda_1 \{A, C\} + \lambda_2 \{B, C\} \quad (1.10.12)$$

und die analoge Beziehung für das zweite Argument. Weiter rechnet man nach, daß die Beziehungen

$$\{A, B\} = -\{B, A\}, \quad (1.10.13)$$

$$\{A, \{B, C\}\} + \{B, \{C, A\}\} + \{C, \{A, B\}\} = 0 \quad (1.10.14)$$

gelten (*nachrechnen!*). Dies sind die definierenden Eigenschaften eines **Lie-Produkts**. Die zweite Zeile heißt **Jacobi-Identität**. Die so definierte Algebra der Poissonklammern macht also den Funktionenraum V zu einer **Lie-Algebra**.

Die Poissonklammern genügen zusätzlich zur Lie-Produktstruktur auch der sog. **Derivationseigenschaft**

$$\{AB, C\} = A\{B, C\} + \{A, C\}B, \quad (1.10.15)$$

wobei das Produkt zweier Funktionen punktweise aufgefaßt wird. Die Bezeichnung „Derivationseigenschaften“ rührt daher, daß (1.10.15) die Form der Produktregel für Ableitungen besitzt. Wie wir im nächsten Abschnitt sehen werden, können wir für eine fest vorgegebene Phasenraumfunktion C mittels

$$\mathfrak{L}_C A = \{A, C\} \quad (1.10.16)$$

den Operator \mathfrak{L}_C , die durch C generierte **Lie-Ableitung**, definieren. Offensichtlich ist dieser Operator eine lineare Abbildung im Vektorraum der Phasenraumfunktionen, der wegen (1.10.15) die Produktregel

$$\mathfrak{L}_C AB = B\mathfrak{L}_C A + A\mathfrak{L}_C B. \quad (1.10.17)$$

Weiter gilt die Produktregel auch für Poisson-Klammern:

$$\mathfrak{L}_C [\{A, B\}] = \{\mathfrak{L}_C A, B\} + \{A, \mathfrak{L}_C B\}. \quad (1.10.18)$$

Zum *Beweis* rechnen wir die rechte Seite aus, indem wir die Definition der Lie-Ableitung (1.10.16) verwenden und mittels (1.10.13) und (1.10.14) umformen:

$$\begin{aligned} \{\mathfrak{L}_C A, B\} + \{A, \mathfrak{L}_C B\} &= \{\{A, C\}, B\} + \{A, \{B, C\}\} \\ &\stackrel{(1.10.13)}{=} \{A, \{B, C\}\} - \{B, \{A, C\}\} \\ &\stackrel{(1.10.13)}{=} \{A, \{B, C\}\} + \{B, \{C, A\}\} \\ &= \{A, \{B, C\}\} + \{B, \{C, A\}\} + \{C, \{A, B\}\} - \{C, \{A, B\}\} \\ &\stackrel{(1.10.14)}{=} -\{C, \{A, B\}\} \stackrel{(1.10.13)}{=} \{\{A, B\}, C\} \stackrel{(1.10.17)}{=} \mathfrak{L}_C [\{A, B\}]. \end{aligned} \quad (1.10.19)$$

1.11 Beispiele zu kanonischen Transformationen

Hier wollen wir einige Beispiele zu kanonischen Transformationen, die wir in den vorigen Abschnitten auf verschiedene Arten definiert haben. Zum einen kann man kanonische Transformationen definieren, indem man **erzeugende Funktionen** willkürlich festlegt. Sie sind stets als Funktionen eines Satzes alter und neuer kanonischer Koordinaten definiert und dürfen auch explizit zeitabhängig sein. Man kann alle vier

Kombinationen von alten und neuen Koordinaten verwenden, je nach Bequemlichkeit für die jeweilige Anwendung. Wir stellen die vier Fälle nochmals übersichtlich zusammen. Sie wurden hergeleitet in den Gln. (1.9.4,1.9.8,1.9.11,1.9.14)

$$g = g(q, Q, t): \quad p_k = \frac{\partial g}{\partial q^k}, \quad P_k = -\frac{\partial g}{\partial Q^k}, \quad H' = H + \frac{\partial g}{\partial t}, \quad (1.11.1)$$

$$g = g(q, P, t): \quad p_k = \frac{\partial g}{\partial q^k}, \quad Q^k = \frac{\partial g}{\partial P_k}, \quad H' = H + \frac{\partial g}{\partial t}, \quad (1.11.2)$$

$$g = g(p, Q, t): \quad q^k = -\frac{\partial g}{\partial p_k}, \quad P_k = -\frac{\partial g}{\partial Q^k}, \quad H' = H + \frac{\partial g}{\partial t}, \quad (1.11.3)$$

$$g = g(p, P, t): \quad q^k = -\frac{\partial g}{\partial p_k}, \quad Q^k = \frac{\partial g}{\partial P_k}, \quad H' = H + \frac{\partial g}{\partial t}. \quad (1.11.4)$$

Andererseits haben wir auch gezeigt, daß eine Transformation $(q, p) \leftrightarrow (Q, P)$ genau dann kanonisch ist, wenn die Poisson-Klammern ungeändert bleiben:

$$\{q^k, q^j\} = \{Q^k, Q^j\} = \{p_k, p_j\} = \{P_k, P_j\} = 0, \quad \{q^k, p_j\} = \{Q^k, P_j\} = \delta_j^k. \quad (1.11.5)$$

Dabei ist die Poisson-Klammer zwischen Phasenraumfunktionen ihrerseits forminvariant unter den kanonischen Transformationen, kann also sowohl mittels der alten als auch der neuen Koordinaten ausgerechnet werden gemäß

$$\{A, B\} = \frac{\partial A}{\partial q^k} \frac{\partial B}{\partial p_k} - \frac{\partial B}{\partial q^k} \frac{\partial A}{\partial p_k} = \frac{\partial A}{\partial Q^k} \frac{\partial B}{\partial P_k} - \frac{\partial B}{\partial Q^k} \frac{\partial A}{\partial P_k}. \quad (1.11.6)$$

1.11.1 Beliebige Diffeomorphismen im Konfigurationsraum

In der Lagrangeformulierung der analytischen Mechanik konnten wir von beliebigen generalisierten Koordinaten q zu beliebigen neuen generalisierten Koordinaten Q übergehen. Die Abbildung mußte nur (zumindest lokal) umkehrbar und differenzierbar (also ein **Diffeomorphismus**) sein. Es ist klar, daß demnach eine solche Transformation stets zu einer kanonischen Transformation erweiterbar sein muß. Wir geben also lediglich

$$q^k = f^k(Q, t) \quad (1.11.7)$$

vor und fragen, ob wir diese Vorgabe durch eine kanonische Transformation im Hamiltonformalismus realisieren können. In diesem Fall ist das sehr einfach, denn wir brauchen z.B. nur auf den Generator vom Typ (1.11.3), $g = g(p, Q, t)$ zurückzugreifen. Zunächst ist

$$q^k = -\frac{\partial g}{\partial p_k} = f^k(Q, t) \Rightarrow g(p, Q, t) = -p_j f^j(Q, t). \quad (1.11.8)$$

Man beachte, daß wir nicht den allgemeinst möglichen Fall betrachten. Wir brauchen nur eine mögliche kanonische Transformation, d.h. wir verzichten auf das Ausschreiben der jeweiligen „Integrationskonstanten“ beim Hochintegrieren partieller Ableitungen. Gemäß (1.11.3) müssen wir dann die neuen kanonischen Impulse gemäß

$$P_k = -\frac{\partial g}{\partial Q^k} = p_j \frac{\partial f^k}{\partial Q^j}. \quad (1.11.9)$$

Voraussetzungsgemäß ist (1.11.7) ein Diffeomorphismus und damit diese Gleichung nach den p_j auflösbar, weil die Jacobi-Matrix $\partial f^k / \partial Q^j$ eines Diffeomorphismus invertierbar ist. Damit haben wir aber die gewünschte kanonische Transformation gefunden.

Wir bemerken, daß sich die p_j kontravariant zu dq^j transformieren, was die Wahl der Indexstellung der kanonischen Variablen inspiriert. In der Tat gilt (bei festgehaltener Zeit t !)

$$dq^k = dQ^j \frac{\partial q^k}{\partial Q^j}, \quad p_k = P_j \frac{\partial Q^j}{\partial q^k}. \quad (1.11.10)$$

Als **Beispiel** betrachten wir die Transformation von kartesischen Koordinaten $(q^k) = (x, y, z)$ zu Kugelkoordinaten $(Q^k) = (r, \vartheta, \varphi)$:

$$(q^k) = \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} r \cos \varphi \sin \vartheta \\ r \sin \varphi \sin \vartheta \\ r \cos \vartheta \end{pmatrix}. \quad (1.11.11)$$

Dann gilt wegen (1.11.8)

$$g(p, Q) = -p_k q^k(Q) = -(p_x r \cos \varphi \sin \vartheta + p_y r \sin \varphi \sin \vartheta + p_z r \cos \vartheta). \quad (1.11.12)$$

Die neuen zu den Q kanonisch konjugierten Impulse sind durch

$$\begin{pmatrix} P_r \\ P_\vartheta \\ P_\varphi \end{pmatrix} = - \begin{pmatrix} \partial_r g \\ \partial_\vartheta g \\ \partial_\varphi g \end{pmatrix} = \begin{pmatrix} p_x \cos \varphi \sin \vartheta + p_y \sin \varphi \sin \vartheta + p_z \cos \vartheta \\ p_x r \cos \varphi \cos \vartheta + p_y r \sin \varphi \cos \vartheta - p_z r \sin \vartheta \\ -p_x r \sin \varphi \sin \vartheta + p_y r \cos \varphi \sin \vartheta \end{pmatrix}. \quad (1.11.13)$$

Da $\partial_t g = 0$ ist nun einfach $H' = H$, aber wir müssen H' durch die Q und P ausdrücken, d.h. wir müssen (1.11.13) nach \vec{p} umstellen. Dazu schreiben wir das lineare Gleichungssystem in Matrixschreibweise

$$\begin{pmatrix} P_r \\ P_\vartheta \\ P_\varphi \end{pmatrix} = \begin{pmatrix} \cos \varphi \sin \vartheta & \sin \varphi \sin \vartheta & \cos \vartheta \\ r \cos \varphi \cos \vartheta & r \sin \varphi \cos \vartheta & -r \sin \vartheta \\ -r \sin \varphi \sin \vartheta & r \cos \varphi \sin \vartheta & 0 \end{pmatrix} \begin{pmatrix} p_x \\ p_y \\ p_z \end{pmatrix} = \hat{T} \begin{pmatrix} p_x \\ p_y \\ p_z \end{pmatrix}. \quad (1.11.14)$$

Die inverse Matrix ergibt sich zu (*nachrechnen!*)

$$\hat{T}^{-1} = \begin{pmatrix} \cos \varphi \cos \vartheta & \cos \varphi \cos \vartheta / r & -\sin \varphi / (r \sin \vartheta) \\ \sin \varphi \sin \vartheta & \sin \varphi \cos \vartheta / r & \cos \varphi / (r \sin \vartheta) \\ \cos \vartheta & -\sin \vartheta / r & 0 \end{pmatrix} \quad (1.11.15)$$

und damit (*nachrechnen!*)

$$H' = H = \frac{1}{2m} \vec{p}^2 = \frac{1}{2m} (P_r, P_\vartheta, P_\varphi) \hat{T}^{-1T} \hat{T}^{-1} \begin{pmatrix} P_r \\ P_\vartheta \\ P_\varphi \end{pmatrix} = \frac{1}{2m} \left(P_r^2 + \frac{P_\vartheta^2}{r^2} + \frac{P_\varphi^2}{r^2 \sin^2 \vartheta} \right). \quad (1.11.16)$$

Freilich erhält man das gleiche Ergebnis auch, indem man die Berechnung der generalisierten Impulse mit der Lagrange-Funktion vornimmt. Zunächst ergibt sich

$$\dot{\vec{x}} = \dot{r} \vec{e}_r + r \dot{\vartheta} \vec{e}_\vartheta + r \sin \vartheta \dot{\varphi} \vec{e}_\varphi, \quad (1.11.17)$$

wobei $\vec{e}_r = \partial_r \vec{x}$, $\vec{e}_\vartheta = 1/r \partial_\vartheta \vec{x}$ und $\vec{e}_\varphi = 1/(r \sin \vartheta) \partial_\varphi \vec{x}$ die orthonormierten Basisvektoren der Kugelkoordinaten sind. Für das freie Teilchen ist

$$L = \frac{m}{2} \dot{\vec{x}}^2 = \frac{m}{2} (\dot{r}^2 + r^2 \dot{\vartheta}^2 + r^2 \sin^2 \vartheta \dot{\varphi}^2). \quad (1.11.18)$$

Die kanonisch konjugierten Impulse sind

$$P_r = \frac{\partial L}{\partial \dot{r}} = m \dot{r}, \quad P_\vartheta = \frac{\partial L}{\partial \dot{\vartheta}} = m r^2 \dot{\vartheta}, \quad P_\varphi = \frac{\partial L}{\partial \dot{\varphi}} = m r^2 \sin^2 \vartheta \dot{\varphi}, \quad (1.11.19)$$

und die Hamiltonfunktion ergibt sich zu

$$H = \dot{r}P_r + \dot{\vartheta}P_{\vartheta} + \dot{\varphi}P_{\varphi} - L = L = \frac{1}{2m} \left(P_r^2 + \frac{P_{\vartheta}^2}{r^2} + \frac{P_{\varphi}^2}{r^2 \sin^2 \vartheta} \right), \quad (1.11.20)$$

was in der Tat mit (1.11.17) übereinstimmt. In diesem Fall erweist sich die Umrechnung von kartesischen zu Kugelkoordinaten über den Lagrangeformalismus als bequemer als mittels der kanonischen Transformation.

1.11.2 Freier Fall

Wir betrachten den freien Fall eines Massenpunktes in Erdnähe. Die z -Achse sei senkrecht nach oben gerichtet, und wir betrachten das Problem in kartesischen Koordinaten $(q^k) = (x, y, z)$. Dann lautet die Hamiltonfunktion

$$H = \frac{\vec{p}^2}{2m} + mgz. \quad (1.11.21)$$

Es liegt nahe, neben den selbstverständlich zyklischen Koordinaten x und y auch noch eine neue Koordinate Z dadurch zyklisch zu machen, daß wir in ein beschleunigtes Bezugssystem übergehen mit

$$z = Z - \frac{g}{2}t^2, \quad p_3 = P_3 - mgt. \quad (1.11.22)$$

Um dies mit einer zeitabhängigen kanonischen Transformation zu bewerkstelligen, liegt es nahe, die erzeugende Funktion in der Form (1.11.3) zu verwenden. Der Ansatz (1.11.22) ergibt nach einfachen Integrationen

$$g(\vec{p}, \vec{X}, t) = -p_1X - p_2Y - p_3 \left(Z - \frac{g}{2}t^2 \right) - mgtZ + \tilde{g}(t). \quad (1.11.23)$$

Dann erhalten wir gemäß (1.11.3)

$$\vec{x} = -\vec{\nabla}_{\vec{p}} g = \begin{pmatrix} X \\ Y \\ Z - gt^2/2 \end{pmatrix}, \quad \vec{P} = -\vec{\nabla}_{\vec{X}} g = \begin{pmatrix} p_1 \\ p_2 \\ p_3 + mgt \end{pmatrix}. \quad (1.11.24)$$

Nun ist (*nachrechnen*)

$$H' = H + \partial_t g = \frac{\vec{p}^2}{2m} + mg^2t^2 + \frac{d\tilde{g}}{dt}. \quad (1.11.25)$$

Setzen wir nun noch

$$\tilde{g}(t) = -\frac{mg^2t^3}{3}, \quad (1.11.26)$$

wird die neue Hamilton-Funktion tatsächlich zu der für ein freies Teilchen

$$H' = \frac{\vec{p}^2}{2m}. \quad (1.11.27)$$

Die Lösungen der Bewegungsgleichungen sind demnach in den neuen Koordinaten

$$\vec{X}(t) = \vec{X}_0 + \frac{1}{m}\vec{P}_0t, \quad \vec{P}(t) = \vec{P}_0 = \text{const.} \quad (1.11.28)$$

Die allgemeine Lösung der Bewegungsgleichung für die alten Koordinaten liefert dann gemäß (1.11.24) wie zu erwarten das korrekte Resultat

$$\vec{x}(t) = \vec{X}_0 + \frac{1}{m} - \frac{g}{2}t^2\vec{e}_3, \quad \vec{p}(t) = \vec{P}_0 - mgt\vec{e}_3. \quad (1.11.29)$$

1.11.3 Eindimensionaler harmonischer Oszillator

Der eindimensionale harmonische Oszillator ist durch die Hamiltonfunktion

$$H = \frac{p^2}{2m} + \frac{m\omega^2}{2}q^2 \quad (1.11.30)$$

definiert. Wir wollen nun neue kanonische Variablen (Q, P) finden, so daß Q zyklisch ist, d.h. die neue Hamiltonfunktion nicht von Q abhängt. Das erreichen wir offenbar für eine beliebige Funktion f , indem wir

$$p = \sqrt{2m}f(P)\cos Q, \quad q = \frac{\sqrt{2}f(P)}{\omega\sqrt{m}}\sin Q \quad (1.11.31)$$

setzen. Die kanonische Transformation enthält dabei die Zeit nicht explizit, und es ist folglich

$$H'(Q, P) = H[q(Q, P), p(Q, P)] = f^2(P). \quad (1.11.32)$$

Wir müssen nur sicherstellen, daß wir f so bestimmen können, daß (1.11.31) tatsächlich eine kanonische Transformation ist. Dazu berechnen wir die einzige nichttriviale kanonische Poisson-Klammer

$$\{q, p\} = \frac{\partial q}{\partial Q} \frac{\partial p}{\partial P} - \frac{\partial p}{\partial Q} \frac{\partial q}{\partial P} = \frac{2}{\omega} f(P) f'(P) \stackrel{!}{=} 1. \quad (1.11.33)$$

Die Differentialgleichung für f ist leicht lösbar:

$$f(P) f'(P) = \frac{1}{2} \frac{d}{dP} f^2 = \frac{\omega}{2} \Rightarrow f = \sqrt{\omega P}. \quad (1.11.34)$$

Damit sind aber alle Poisson-Klammern die kanonischen, und mit (1.11.34) ist (1.11.31) kanonisch:

$$q = \frac{2\sqrt{P}}{\sqrt{m\omega}} \sin Q, \quad p = 2\sqrt{m\omega P} \cos Q. \quad (1.11.35)$$

Die neue Hamiltonfunktion ist also gemäß (1.11.33)

$$H'(Q, P) = \omega P \Rightarrow \dot{Q} = \frac{\partial H'}{\partial P} = \omega, \quad \dot{P} = -\frac{\partial H'}{\partial Q} = 0. \quad (1.11.36)$$

Dabei haben wir die Hamiltonschen kanonischen Gleichungen für die neuen Koordinaten benutzt. Die Bewegungsgleichungen sind sofort lösbar

$$Q(t) = \omega t + Q_0, \quad P(t) = P_0 = \text{const.} \quad (1.11.37)$$

Dabei haben wir die Anfangsbedingungen $Q(0) = Q_0, P(0) = P_0$ gleich eingearbeitet. Mittels der kanonischen Transformation (1.11.35) erhalten wir dann die Lösungen für die ursprünglichen Phasenraumkoordinaten einfach durch Einsetzen

$$q(t) = \sqrt{\frac{4P_0}{m\omega}} \sin(\omega t + Q_0), \quad p(t) = \sqrt{4m\omega P_0} \cos(\omega t + Q_0). \quad (1.11.38)$$

Dies ist offenbar die allgemeine Lösung der Bewegungsgleichungen für die ursprünglichen Phasenraumkoordinaten, wenn auch in nicht ganz expliziter Form (*nachrechnen!*). Offenbar gilt

$$q(0) = q_0 = \sqrt{\frac{4P_0}{m\omega}} \sin Q_0, \quad p(0) = p_0 = \sqrt{4m\omega P_0} \cos Q_0. \quad (1.11.39)$$

Damit folgt aber in der Tat die Lösung der Bewegungsgleichungen für den harmonischen Oszillator in der gewohnten Form. Dazu müssen wir nur die Additionstheoreme für sin und cos verwenden und (1.11.39) anwenden

$$\begin{aligned} q(t) &= \sqrt{\frac{4P_0}{m\omega}} [\cos(\omega t) \sin Q_0 + \sin(\omega t) \cos Q_0] = q_0 \cos(\omega t) + \frac{p_0}{m\omega} \sin(\omega t), \\ p(t) &= \sqrt{4m\omega P_0} [\cos(\omega t) \cos Q_0 - \sin(\omega t) \sin Q_0] = p_0 \cos(\omega t) - m\omega x_0 \sin(\omega t). \end{aligned} \quad (1.11.40)$$

1.12 Lie-Gruppen und -Algebren

Wir können nun mit Hilfe der Lie-Ableitung (1.10.17) kanonische Transformationen definieren, die von einem reellen Parameter abhängen und eine sog. **Einparameter-Lie-Gruppe** bilden. Diese Transformationen wirken auf beliebige Phasenraumfunktionen. Für eine beliebig vorgegebene Phasenraumfunktion G , den **Generator** (oder **infinitesimale Erzeugende**) definiert die Differential-Gleichung

$$\frac{dO_\alpha(q, p)}{d\alpha} = \mathfrak{L}_G O_\alpha(q, p), \quad O_{\alpha=0}(q, p) = O(q, p) \quad (1.12.1)$$

eine Transformation $O_{\alpha=0} \rightarrow O_\alpha$. Offenbar gilt

$$O_\alpha(q, p) = O(q_\alpha, p_\alpha), \quad (1.12.2)$$

denn es ist

$$\begin{aligned} \frac{d}{d\alpha} O(q_\alpha, p_\alpha) &= \frac{dq_\alpha^k}{d\alpha} \frac{\partial}{\partial q^k} O(q_\alpha, p_\alpha) + \frac{dp_{\alpha,k}}{d\alpha} \frac{\partial}{\partial p_k} O(q_\alpha, p_\alpha) \\ &= \{q^k, G\} \frac{\partial}{\partial q^k} O(q_\alpha, p_\alpha) + \{p_k, G\} \frac{\partial}{\partial p_k} O(q_\alpha, p_\alpha) \\ &= \{O_\alpha, G\}. \end{aligned} \quad (1.12.3)$$

Daß (1.12.1) darüber hinaus auch eine Schar endlicher kanonischer Transformationen, wobei $\alpha \in \mathbb{R}$ der Scharparameter ist, definiert, erkennen wir wie folgt. Formal können wir nämlich (1.12.1) mit Hilfe der **Operatorexponentialfunktion** integrieren:

$$\begin{aligned} O_\alpha(q, p) &= O(q_\alpha, p_\alpha) = \exp(\alpha \mathfrak{L}_G \cdot) O(q, p) \\ &:= O(q, p) + \alpha \{O(q, p), G\} + \frac{\alpha^2}{2!} \{\{O(q, p), G\}, G\} + \dots \\ &:= \sum_{k=0}^{\infty} \frac{\alpha^k}{k!} \{O, G\}_k \end{aligned} \quad (1.12.4)$$

Wenden wir dies nun auf $O(q, p) = \{q^k, q^j\} = 0$, $O(q, p) = \{q^k, p_j\} = \delta_j^k$ und $O(q, p) = \{p^j, p^k\} = 0$ an, folgt wegen $\{\text{const}, G\} = 0$ sofort, daß

$$\{q_\alpha^j, q_\alpha^k\} = \{q^j, q^k\} = 0, \quad \{p_{\alpha,j}, q_\alpha^k\} = p_j q^k = \delta_j^k, \quad \{p_{\alpha,j}, p_{\alpha,k}\} = \{p_j, p_k\} = 0, \quad (1.12.5)$$

d.h. es gelten für jedes $\alpha \in \mathbb{R}$ für die $(Q, P) = (q_\alpha, p_\alpha)$ die kanonischen Poisson-Klammerrelationen, und folglich ist die Transformation $(q, p) \mapsto (Q, P)$ für jedes $\alpha \in \mathbb{R}$ eine kanonische Transformation. Wir bemerken noch, daß sich der Generator unter so definierten Transformationen trivialerweise nicht ändert, denn es ist wegen $\{G, G\} = 0$

$$G_\alpha = \exp(\alpha \mathfrak{L}_G) G = G. \quad (1.12.6)$$

Wir können also die definierende Differentialgleichung (1.12.1) auch in der Form

$$\frac{d}{d\alpha} O_\alpha = \mathfrak{L}_{G_\alpha} O_\alpha = \{O_\alpha, G_\alpha\} \quad (1.12.7)$$

schreiben.

Wir gehen auf die Diskussion der Konvergenz der Reihe (1.12.4) nicht näher ein. Wir werden unten an einigen Beispielen diese Konvergenz explizit nachweisen. Wenn sie aber konvergiert, löst sie auch die Differentialgleichung (1.12.1):

$$\begin{aligned} \frac{d}{d\alpha} \exp(\alpha \mathfrak{L}_G) O(q, p) &= \sum_{k=1}^{\infty} \frac{\alpha^{k-1}}{k!} \{O, G\}_k \\ &= \sum_{k=0}^{\infty} \frac{\alpha^k}{k!} \{O, G\}_{k+1} = \{\exp(\alpha \mathfrak{L}_G) O, G\} = \\ &= \{O_\alpha, G\}. \end{aligned} \quad (1.12.8)$$

Damit ist gezeigt, daß die Einparameter-Lie-Gruppen durch ihre „infinitesimalen Erzeugenden“ eindeutig definiert sind.

Das ist nun mathematisch insofern interessant, als die infinitesimalen Generatoren zusammen mit den Poissonklammern eine **Liealgebra** bilden, wie wir oben bereits gezeigt haben, d.h. jeder Liegruppe ist eindeutig eine Lie-Algebra zugeordnet, und wenigstens lokal für die Einparameteruntergruppen der Liegruppe kann man durch Integration der Differentialgleichung (1.12.1) die Liegruppe aus ihrer Liealgebra rekonstruieren.

Die Einparameter-Lie-Gruppen sind **homomorph zur additiven Gruppe der reellen Zahlen**. Um dies zu sehen, sei G der Generator der Einparametergruppe. Dann definieren wir das Gruppenelement als den in (1.12.4) definierten Operator

$$g_\alpha = \exp(\alpha \mathfrak{L}_G), \quad \alpha \in \mathbb{R}. \quad (1.12.9)$$

Wir wollen nun zeigen, daß für $\alpha, \beta \in \mathbb{R}$ gilt $g_\alpha g_\beta = g_{\alpha+\beta} = g_\beta g_\alpha$. Wegen der algebraischen Produkteigenschaften der Poissonklammer gilt

$$\begin{aligned} g_\alpha g_\beta &= \sum_{k=0}^{\infty} \frac{(\alpha \mathfrak{L}_G)^k}{k!} \sum_{l=0}^{\infty} \frac{(\beta \mathfrak{L}_G)^l}{l!} = \sum_{n=0}^{\infty} \sum_{m=0}^n \frac{(\alpha \mathfrak{L}_G)^m (\beta \mathfrak{L}_G)^{n-m}}{m! (n-m)!} = \\ &= \sum_{n=0}^{\infty} \frac{1}{n!} \sum_{m=0}^n \binom{n}{m} (\alpha \mathfrak{L}_G)^m (\beta \mathfrak{L}_G)^{n-m} = \sum_{n=0}^{\infty} \frac{[(\alpha + \beta) \mathfrak{L}_G]^n}{n!} = \exp[(\alpha + \beta) \mathfrak{L}_G]. \end{aligned} \quad (1.12.10)$$

Dies zeigt, daß in der Tat die Abbildung $\alpha \mapsto g_\alpha$ einen Homomorphismus von der additiven Gruppe der reellen Zahlen in die von G erzeugte Einparametergruppe darstellt.

Wir betrachten im folgenden auch kompliziertere Gruppen, die von endlich vielen Generatoren und Parametern abhängig sind. Wir behandeln auch hier wieder nur solche Gruppen, die als kanonische Transformationen von Phasenraumfunktionen erzeugt werden.

Seien G_a mit $a \in \{1, 2, \dots, d\}$ linear unabhängige Phasenraumfunktionen. Ist dann der von ihnen aufgespannte Untervektorraum $U = \text{span}(G_a)_{a \in \{1, \dots, d\}}$ bzgl. der Poissonklammeralgebra abgeschlossen, so heißt U eine **Unter-Liealgebra** der Liealgebra der Phasenraumfunktionen.

Zur algebraischen Abgeschlossenheit ist es offenbar notwendig und hinreichend, daß es für je zwei Basisfunktionen von U reelle Zahlen f^c_{ab} gibt, so daß

$$\{G_a, G_b\} = f^c_{ab} G_c = -f^c_{ba} G_c. \quad (1.12.11)$$

Die f^c_{ab} werden als die **antisymmetrischen Strukturkoeffizienten der Liealgebra** bzw. der von ihnen erzeugten Liegruppe bezeichnet. Diese Liegruppe wird dabei definitionsgemäß durch die Elemente

$$g(\alpha_1, \dots, \alpha_d) = \exp\left(\sum_a \alpha_a \mathfrak{L}_{G_a}\right) \quad (1.12.12)$$

algebraisch erzeugt.

Zu beachten ist, daß i.a. diese Gruppe **nicht** kommutativ ist; das ist nur der Fall, wenn die G_a wechselseitig konstante Poissonklammern besitzen. Allgemein gilt nämlich für eine beliebige Phasenraumfunktion O

$$\begin{aligned} [\mathfrak{L}_A, \mathfrak{L}_B]O &:= (\mathfrak{L}_A \mathfrak{L}_B - \mathfrak{L}_B \mathfrak{L}_A)O \\ &= \{\{O, B\}, A\} - \{\{O, A\}, B\} = \{O, \{B, A\}\} \\ &= \mathfrak{L}_{\{B, A\}}O, \end{aligned} \quad (1.12.13)$$

wobei wir von den Lie-Algebraeigenschaften (1.10.13) und (1.10.14) Gebrauch gemacht haben (*nachrechnen!*). Dies zeigt, daß die \mathfrak{L}_{G_a} genau dann vertauschen, wenn die Poissonklammern der G_a untereinander jeweils konstant sind. Dann und nur dann ist die erzeugte Liealgebra isomorph zur additiven Gruppe des \mathbb{R}^d .

Im allgemeinen Fall ist übrigens die Exponentialabbildung noch nicht einmal surjektiv auf die von den Elementen (1.12.12) algebraisch erzeugte Gruppe.

1.13 Das Noether-Theorem (Hamilton-Form)

Wir haben im vorigen Abschnitt gesehen, daß wir kanonische Transformationen mittels Generatoren vermöge der Lie-Ableitung gemäß (1.12.3) definieren können, und dies Einparameter-Lie-Gruppen vermöge der Operatorexponentialfunktion (1.12.4) definiert. Wir betrachten im folgenden auch Generatoren, die explizit von der Zeit abhängen können.

Um zu sehen, wie wir in einem solchen Fall die kanonische Transformation der Phasenraumvariablen und der Hamilton-Funktion zu definieren haben, benötigen wir zunächst nur eine kanonische Transformation, die sich nur „infinitesimal“ von der Identität unterscheidet. Für einen infinitesimalen Parameter $\delta\alpha$ ist nun

$$\exp(\delta\alpha \mathfrak{L}_G) = 1 + \delta\alpha \mathfrak{L}_G. \quad (1.13.1)$$

Wir verwenden dazu den in Abschnitt 1.9 entwickelten Formalismus mittels Generatoren der Art (1.9.8), also für eine erzeugende Funktion der Art $g_1 = g_1(q, P, t)$. Für eine nur infinitesimal von der Identität abweichende Transformation gilt

$$g_1(q, P, t) = q_j P^j + G(q, P, t) \delta\alpha = q_j P^j + G(q, p, t) \delta\alpha + \mathcal{O}(\delta\alpha^2), \quad (1.13.2)$$

und mit (1.9.8) folgt dann in der Tat, daß

$$\begin{aligned} H'(Q, P, t) &= H(q, p, t) + \delta\alpha \partial_t G(q, p, t) + \mathcal{O}(\delta\alpha^2), \\ p_k &= P_k + \delta\alpha \frac{\partial}{\partial q^k} G(q, p, t) + \mathcal{O}(\delta\alpha^2), \\ Q^k &= q^k + \delta\alpha \frac{\partial}{\partial p_k} G(q, p, t) + \mathcal{O}(\delta\alpha^2). \end{aligned} \quad (1.13.3)$$

Folglich ist

$$H'(Q, P, t) - H(q, p, t) = \delta\alpha \partial_t G, \quad Q^k - q^k = \delta\alpha \mathfrak{L}_G q^k, \quad P_k - p_k = \delta\alpha \mathfrak{L}_G p_k. \quad (1.13.4)$$

D.h. wir definieren die kanonische Transformation der Hamilton-Funktion vermöge

$$\frac{d}{d\alpha} H'_\alpha = \partial_t G_\alpha = \partial_t G, \quad \frac{d}{d\alpha} q_\alpha^k = \mathfrak{L}_G q_\alpha^k, \quad \frac{d}{d\alpha} p_{\alpha,k} = \mathfrak{L}_G p_{\alpha,k}. \quad (1.13.5)$$

Für jedes $\alpha \in \mathbb{R}$ haben wir dann eine kanonische Transformation über die Lösung der letzten beiden Differentialgleichungen von (1.13.5) gegeben. Dies definiert aber gemäß den Ausführungen im vorigen Abschnitt eine **Einparameter-Lie-Gruppe**.

Diese Einparameter-Lie-Gruppe bildet nun eine **Einparameter-Symmetriegruppe** definitionsgemäß genau dann, wenn für jedes $\alpha \in \mathbb{R}$

$$H'_\alpha - H_\alpha = 0 \Rightarrow \frac{d}{d\alpha} (H'_\alpha - H_\alpha) \stackrel{(1.13.5)}{=} \partial_t G - \mathfrak{L}_G H = \partial_t G + \{G, H\} = \mathfrak{L}_H G + \partial_t G = 0 \quad (1.13.6)$$

ist. Dann gilt aber gemäß (1.13.5) für die Lösungen der Hamiltonschen kanonischen Gleichungen (1.8.6), also für die Phasenraumtrajektorien

$$\mathfrak{L}_H G + \partial_t G = \frac{d}{dt} G, \quad (1.13.7)$$

denn für diese Phasenraumtrajektorien ist

$$\begin{aligned} \frac{d}{dt} G(q, p, t) &= \dot{q}^k \frac{\partial G}{\partial q^k} + \dot{p}_k \frac{\partial G}{\partial p_k} = \frac{\partial G}{\partial q^k} \frac{\partial H}{\partial p_k} - \frac{\partial G}{\partial p_k} \frac{\partial H}{\partial q^k} + \partial_t G \\ &= \{G, H\} + \partial_t G = \mathfrak{L}_H G + \partial_t G. \end{aligned} \quad (1.13.8)$$

Aus der obigen Herleitung folgt, daß auch umgekehrt jede Phasenraumfunktion, die entlang der Phasenraumtrajektorien (also der Lösungen der Hamiltonschen kanonischen Gleichungen (1.8.6) erhalten ist, Generator einer entsprechenden Einparameter-Symmetriegruppe des Systems ist. Dies ist die vollständige Formulierung des **Noether-Theorems**:

Jede Einparameter-Lie-Symmetriegruppe eines Hamiltonischen Systems definiert eine Erhaltungsgröße und umgekehrt jede Erhaltungsgröße eine Einparameter-Lie-Symmetriegruppe.

Es ergibt sich daraus sofort, daß die Generatoren der Symmetriegruppe des Hamiltonischen Systems eine **Unter-Lie-Algebra** der Lie-Algebra der Phasenraumfunktionen bilden, wobei die Poisson-Klammern gemäß (1.12.13) umkehrbar eindeutig (modulo Konstanten, denn $\mathfrak{L}_G \equiv \mathfrak{L}_{G+\text{const}}$, wobei hier const für eine beliebige von den Phasenraumkoordinaten unabhängige aber evtl. zeitabhängige Funktion steht) über die Kommutatoren der entsprechenden Lie-Ableitungen verknüpft sind.

1.13.1 Die Galilei-Gruppe (Hamilton-Formulierung)

Wir können nun die Analyse aus Abschnitt 1.6.2 wesentlich eleganter mit dem nun ausformulierten Apparat der Lie-Gruppen- und -Algebrentheorie formulieren. Dabei können wir die Erhaltungsgrößen als Generatoren der entsprechenden Lie-Symmetrien über den eleganten Formalismus mit Hilfe der Lie-Ableitung bzw. Poisson-Klammern verwenden.

Zeittranslationsinvarianz: Die Zeitentwicklung des dynamischen Systems wird durch die **Hamilton-Funktion** im Sinne dieses Formalismus erzeugt, denn die Hamiltonschen kanonischen Gleichungen lassen sich mit Hilfe der Poisson-Klammern und damit der Lie-Ableitung ausdrücken (*nachrechnen!*)

$$\dot{q}^k = \frac{\partial H}{\partial p_k} = \{q^k, H\} = \mathfrak{L}_H q^k, \quad \dot{p}_k = -\frac{\partial H}{\partial q^k} = \{p_k, H\} = \mathfrak{L}_H p_k. \quad (1.13.9)$$

Die zur Zeittranslation gehörigen Erhaltungsgröße ist demnach gemäß der Symmetrieanalyse im Lagrange-Formalismus die **Hamilton-Funktion**. Die Bedingung für Zeittranslationsinvarianz ist gemäß (1.13.6) $\partial_t H =$

0, d.h. die Hamilton-Funktion darf nicht explizit von der Zeit abhängen. Dann gilt aber in der Tat entlang der Trajektorien

$$\frac{dH}{dt} = \mathfrak{L}_H H + \partial_t H = \mathfrak{L}_H H = 0. \quad (1.13.10)$$

Räumliche Translationsinvarianz: Eine infinitesimale Translation wird für die kartesischen Ortsvariablen durch

$$\delta x^k = \delta a^j \stackrel{!}{=} \delta a^j \mathfrak{L}_{G_j} x^k = \delta a^j \{x^k, G_j\} = \delta a^j \frac{\partial G_j}{\partial p_k} \quad (1.13.11)$$

definiert. Für den Generator ergibt sich also die partielle Differentialgleichung

$$\frac{\partial G_j}{\partial p_k} = \delta_j^k \Rightarrow G_j = p_j + \tilde{G}_j(\vec{x}). \quad (1.13.12)$$

Die einfachste Wahl für den Generator ist aber $\tilde{G}_j(\vec{x}) = 0$. Dann folgt

$$G_j = p_j \Rightarrow \delta p_k = \delta a^j \mathfrak{L}_{p_j} p_k = 0. \quad (1.13.13)$$

In der Tat generiert $p_n = \vec{n} \cdot \vec{p}$ Translationen in eine beliebige Richtung Richtung \vec{n} . Das können wir in diesem Fall auch mit der Operatorexponentialfunktion (1.12.4) verifizieren, denn es ist

$$x_\alpha^k = \exp(\alpha \mathfrak{L}_{\vec{n} \cdot \vec{p}}) x^k = x^k + \alpha n^j \{x^k, p_j\} = x^k + \alpha n^k, \quad (1.13.14)$$

wobei alle höheren Terme verschwinden, weil $\mathfrak{L}_{\vec{n} \cdot \vec{p}} x^k = n^k = \text{const}$ ist. Ebenso folgt wegen $\mathfrak{L}_{\vec{n} \cdot \vec{p}} p^k = 0$

$$p_{\alpha,k} = p_k = \text{const}. \quad (1.13.15)$$

Die Translationsgruppe ist abelsch, und in der Tat vertauschen die Lie-Ableitungen der Generatoren und auch die Poisson-Klammern verschwinden, denn es gelten die kanonischen Vertauschungsregeln $\{p_k, p_j\} = 0$. Aus der Hamiltonschen kanonischen Gleichung folgt sofort, daß der Impuls p_k genau dann erhalten ist, wenn $\partial H / \partial x^k = 0$ ist, also die Hamilton-Funktion nicht von der dazugehörigen Ortskoordinate x^k abhängt (also x^k zyklisch ist).

Isotropie des Raumes: Eine infinitesimale Drehung um eine Drehachse in Richtung eines Einheitsvektors \vec{n} wird für die kartesischen Orts- und Impulskoordinaten durch

$$\delta \vec{x} = \delta \varphi \vec{n} \times \vec{x}, \quad \delta \vec{p} = \delta \varphi \vec{n} \times \vec{p} \quad (1.13.16)$$

beschrieben (vgl. [Hee14]). Wir geben nun der Einfachheit halber für die kartesischen Koordinaten der Ortsvektoren und Impulse die Unterscheidung in ko- und kontravariante Indizes auf, wobei aber die Einsteinsche Summationskonvention weiter gelten soll. In Komponentenschreibweise lautet (1.13.16)

$$\delta x_k = \delta \varphi \epsilon_{klm} n_l x_m, \quad \delta p_m = \delta \varphi \epsilon_{klm} n_l p_m. \quad (1.13.17)$$

Drückt man dies mit Hilfe eines Generators G aus, folgt

$$\mathfrak{L}_G x_k = \{x_k, G\} = \frac{\partial G}{\partial p_k} \stackrel{!}{=} \epsilon_{klm} n_l x_m \Rightarrow G = \epsilon_{klm} n_l x_m p_k + \tilde{G}(\vec{x}) = \vec{n} \cdot \vec{\ell} + \tilde{G}(\vec{p}). \quad (1.13.18)$$

Dabei ist $\vec{\ell} = \vec{x} \times \vec{p}$ der Drehimpuls. Für den Impuls folgt daraus

$$\mathfrak{L}_G p_k = \{p_k, G\} = -\frac{\partial G}{\partial x_k} = -\epsilon_{mlk} n_l p_m - \frac{\partial \tilde{G}}{\partial x_k} = +\epsilon_{klm} n_l p_m - \frac{\partial \tilde{G}}{\partial x_k} \stackrel{!}{=} \epsilon_{klm} n_k p_m \Rightarrow \frac{\partial \tilde{G}}{\partial x_k} = 0. \quad (1.13.19)$$

Wir können also ohne Beschränkung der Allgemeinheit $\vec{G} = 0$ setzen. Die zu einer Drehung um \vec{n} gehörigen Erhaltungsgrößen sind also die Komponenten des Drehimpulses in Richtung von \vec{n} . Auch hier läßt sich die endliche Drehung wieder mit Hilfe der Operatorexponentialfunktion gewinnen. Betrachten wir der Einfachheit halber eine Drehung um die 3-Achse, setzen also $\vec{n} = \vec{e}_3$, so folgt

$$\begin{aligned}\mathcal{L}_{\ell_3} \vec{x} &= \vec{e}_3 \times \vec{x}, \\ \mathcal{L}_{\ell_3}^2 \vec{x} &= \vec{e}_3 \times (\vec{e}_3 \times \vec{x}) = \vec{e}_3 x^3 - \vec{x}, \\ \mathcal{L}_{\ell_3}^3 \vec{x} &= \vec{e}_3 \times (\vec{e}_3 x^3 - \vec{x}) = -\vec{e}_3 \times \vec{x}.\end{aligned}\tag{1.13.20}$$

Durch Rekursion gelangen wir damit für alle $k \in \mathbb{N}$

$$\mathcal{L}_{\ell_3}^{2k} \vec{x} = (-1)^k (\vec{x} - \vec{e}_3 x^3) = (-1)^k \begin{pmatrix} x^1 \\ x^2 \\ 0 \end{pmatrix}, \quad \mathcal{L}_{\ell_3}^{2k-1} \vec{x} = (-1)^{k+1} \vec{e}_3 \times \vec{x} = (-1)^{k+1} \begin{pmatrix} -x^2 \\ x^1 \\ 0 \end{pmatrix}.\tag{1.13.21}$$

Damit folgt

$$\begin{aligned}\vec{x}_\varphi &= \exp(\varphi \mathcal{L}_{\ell_3}) \vec{x} \\ &= \begin{pmatrix} 0 \\ 0 \\ x^3 \end{pmatrix} + \left(1 - \frac{\varphi^2}{2!} + \frac{\varphi^4}{4!} - \dots\right) \begin{pmatrix} x^1 \\ x^2 \\ 0 \end{pmatrix} + \left(\varphi - \frac{\varphi^3}{3!} + \frac{\varphi^5}{5!} - \dots\right) \begin{pmatrix} -x^2 \\ x^1 \\ 0 \end{pmatrix} \\ &= \begin{pmatrix} x^1 \cos \varphi - x^2 \sin \varphi \\ x^1 \sin \varphi + x^2 \cos \varphi \\ x^3 \end{pmatrix} = \hat{D}_3(-\varphi) \vec{x}.\end{aligned}\tag{1.13.22}$$

Weiter ist für die Impulse

$$\mathcal{L}_{\ell_3} p_j = \{x^1 p_2 - x^2 p_1, p_j\} = \{x^1, p_j\} p_2 - \{x^2, p_j\} p_1 = \delta_j^1 p_2 - \delta_j^2 p_1\tag{1.13.23}$$

bzw.

$$\mathcal{L}_{\ell_3} \vec{p} = \vec{e}_3 \times \vec{p},\tag{1.13.24}$$

und dieselbe Rechnung wie für \vec{x} ergibt

$$\vec{p}_\varphi = \hat{D}_3(-\varphi) \vec{p},\tag{1.13.25}$$

d.h. \vec{x} und \vec{p} verhalten sich unter Drehungen wie die Komponenten von Vektoren, wie es sein muß.

Drehungen sind nicht abelsch, bilden aber eine Untergruppe der Galileigruppe. Die entsprechende **Lie-Algebra** der Drehgruppe ergibt sich aus den Poisson-Klammern für die entsprechenden Erzeugenden, also die Drehimpulskomponenten zu (*nachrechnen!*)⁶

$$\begin{aligned}\{\ell_a, \ell_b\} &= \epsilon_{aij} \epsilon_{bkl} \{x_i p_j, x_k p_l\} \\ &= \epsilon_{aij} \epsilon_{bkl} (\{x_i, x_k p_l\} p_j + x_i \{p_j, x_k p_l\}) \\ &= \epsilon_{aij} \epsilon_{bkl} (\delta_{il} x_k p_j - \delta_{jk} x_i p_l) \\ &= -\epsilon_{aji} \epsilon_{bki} x_k p_j + \epsilon_{aij} \epsilon_{blj} x_i p_l \\ &= -(\delta_{ab} \delta_{jk} - \delta_{ak} \delta_{jb}) x_k p_j + (\delta_{ab} \delta_{il} - \delta_{al} \delta_{ib}) x_i p_l \\ &= x_a p_b - x_b p_a = \epsilon_{cab} \ell_c.\end{aligned}\tag{1.13.26}$$

⁶Wir schreiben für die hier verwendeten kartesischen Komponenten der Einfachheit halber alle Indizes als untere Indizes, denn die Komponenten der euklidischen Metrik sind $\vec{e}_j \cdot \vec{e}_k = \delta_{jk}$ und somit für alle Vektor- bzw. Tensor-komponenten die ko- und kontravarianten Komponenten gleich, denn es ist $V_j = \delta_{jk} V^k = V^j$.

Dabei haben wir die in [Hee14] ausführlich behandelte Formel

$$\epsilon_{abc}\epsilon_{dec} = \delta_{ad}\delta_{be} - \delta_{ae}\delta_{bd} \quad (1.13.27)$$

verwendet. Die Drehimpulsalgebra (also die Lie-Algebra der Drehgruppe $SO(3)$) ist also durch

$$\{\ell_a, \ell_b\} = \epsilon_{cab}\ell_c \quad (1.13.28)$$

gegeben, und die Strukturkonstanten der Algebra sind gerade die Levi-Civita-Symbole. Man beachte, daß in diesem Fall diese Strukturkonstanten nicht nur in den beiden hinteren Indizes sondern vollständig antisymmetrisch sind. Dies ist der Besonderheit zu verdanken, daß wir eine kartesische Basis für die Komponenten von Vektoren und insbesondere des Drehimpulses verwendet haben.

Galilei-Boosts. Ein infinitesimaler Galilei-Boost in der Richtung \vec{n} ist durch

$$\delta\vec{x} = -\delta v\vec{n}t, \quad \delta\vec{p} = -m\delta v\vec{n} \quad (1.13.29)$$

definiert. Für den entsprechenden Generator gilt also

$$\mathfrak{L}_G x_k = \{x_k, G\} = \frac{\partial G}{\partial p_k} \stackrel{!}{=} -n_k t \Rightarrow G = -\vec{n} \cdot \vec{p}t + \tilde{G}(\vec{x}) \quad (1.13.30)$$

und weiter

$$\mathfrak{L}_G p_k = \{p_k, G\} = -\frac{\partial G}{\partial x_k} = -\frac{\partial \tilde{G}}{\partial x_k} \stackrel{!}{=} -mn_k \Rightarrow \tilde{G} = m\vec{n} \cdot \vec{x}. \quad (1.13.31)$$

Damit ist $G = \vec{n} \cdot \vec{k}$ mit

$$\vec{K} = m\vec{x} - \vec{p}t. \quad (1.13.32)$$

Die endlichen Transformationen ergeben sich hier am einfachsten durch Lösung der Differentialgleichung. Für einen Boost in x^3 -Richtung haben wir z.B.

$$\begin{aligned} \frac{d}{dv} x_{v,k} &= \mathfrak{L}_{K_{v,3}} x_{v,k} = \{x_{v,k}, mx_{v,3} - p_3 t\} = -t\delta^{k3} \\ \Rightarrow x_{v,1} &= x_1, \quad x_{v,2} = x_2, \quad x_{v,3} = x_3 - vt, \end{aligned} \quad (1.13.33)$$

Für die Impulse folgt

$$\begin{aligned} \frac{d}{dv} p_{v,k} &= \mathfrak{L}_{K_{v,3}} p_{v,k} = \{p_{v,k}, mx_{v,3} - p_{v,3}\} = -m\delta_{3k} \\ \Rightarrow p_{v,1} &= p_1, \quad p_{v,2} = p_2, \quad p_{v,3} = p_3 - mv = p_3 - mv. \end{aligned} \quad (1.13.34)$$

Wie vom Noether-Theorem zu erwarten, generieren also die durch (1.13.32) definierten Komponenten von \vec{K} tatsächlich Galilei-Boosts. Da die Galilei-Boosts eine abelsche Untergruppe der Galilei-Gruppe bilden, kommutieren die Schwerpunktskoordinaten.

1.13.2 Beispiel: Hamilton-Funktion für freies Teilchen

Wir können die bekannte Hamilton-Funktion $H = \vec{p}^2/(2m)$ für ein freies Teilchen nun aus der Forderung herleiten, daß die Bewegungsgleichung Galilei-invariant sein muß. Dazu müssen alle 10 Generatoren der Galilei-Gruppe die Symmetriebedingung (1.13.6) erfüllen. Die zeitliche Translationsinvarianz, für die $G = H$ ist, verlangt

$$\{H, H\} + \partial_t H = \partial_t H = 0 \Rightarrow H = H(\vec{x}, \vec{p}), \quad (1.13.35)$$

d.h. H darf nicht explizit zeitabhängig sein. Für die räumliche Translationsinvarianz ist $G = p_j$ für $j \in \{1, 2, 3\}$ und damit

$$\{p_j, H\} = -\frac{\partial H}{\partial x^j} = 0 \Rightarrow H = H(\vec{p}). \quad (1.13.36)$$

Die Drehungen werden durch die drei Komponenten des Drehimpulses $\ell_j = \epsilon_{jkl} x^k p_l$ erzeugt. Daraus folgt

$$\{\ell_j, H\} = \epsilon_{jkl} \{x^k p_l, H\} = \epsilon_{jkl} \{x^k, H\} p_l = \epsilon_{jkl} \frac{\partial H}{\partial p_k} p_l \stackrel{!}{=} 0. \quad (1.13.37)$$

Das bedeutet aber, daß

$$\frac{\partial H}{\partial p_k} = f(\vec{p}^2) p_k \Rightarrow H(\vec{p}) = \frac{1}{2} f(\vec{p}^2). \quad (1.13.38)$$

Nun müssen wir noch die Boost-Invarianz erfüllen. Die Generatoren sind durch (1.13.32) gegeben. Die Symmetriebedingung (1.13.6) verlangt dann

$$\begin{aligned} \{K^j, H\} + \partial_t K^j &= \frac{m}{2} \{x^j, f(\vec{p}^2)\} - p_j = m f'(\vec{p}^2) p_j - p_j \stackrel{!}{=} 0 \\ \Rightarrow f'(\vec{p}^2) &= \frac{1}{m} \Rightarrow f(\vec{p}^2) = \frac{\vec{p}^2}{m} \Rightarrow H = \frac{\vec{p}^2}{2m}. \end{aligned} \quad (1.13.39)$$

1.13.3 Beispiel: Hamilton-Funktion für ein Zweiteilchensystem

Wir betrachten nun ein abgeschlossenes Zweiteilchensystem und leiten die Hamilton-Funktion her, der vollständig symmetrisch unter Galilei-Transformationen ist. Wir verwenden wieder kartesische Koordinaten (\vec{x}_j, \vec{p}_j) ($j \in \{1, 2\}$) als Phasenraumvariablen und werten wieder die Symmetriebedingung (1.13.6) für alle 10 Generatoren der Galilei-Gruppe aus. Die Zeittranslationsinvarianz (Generator H) verlangt wieder, daß

$$\partial_t H = 0 \Rightarrow H = H(\vec{x}_1, \vec{x}_2, \vec{p}_1, \vec{p}_2). \quad (1.13.40)$$

Die dazugehörige Erhaltungsgröße, also der Generator für Zeittranslationen ist die **Hamilton-Funktion** (also die **Gesamtenergie**). Die räumlichen Translationen werden durch den **Gesamtimpuls**

$$\vec{P} = \vec{p}_1 + \vec{p}_2 \quad (1.13.41)$$

erzeugt. Die Symmetriebedingung verlangt demnach

$$\{P_k, H\} = -\sum_{j=1}^2 \frac{\partial H}{\partial x_j^k} \stackrel{!}{=} 0 \quad \text{bzw.} \quad \vec{\nabla}_{x_1} H = -\vec{\nabla}_{x_2} H. \quad (1.13.42)$$

Das bedeutet, daß

$$H = H(\vec{x}_1 - \vec{x}_2, \vec{p}_1, \vec{p}_2) \quad (1.13.43)$$

sein muß. Die Drehungen werden durch den **Gesamtdrehimpuls**

$$\vec{l} = \sum_{j=1}^2 \vec{x}_j \times \vec{p}_j \quad (1.13.44)$$

erzeugt. Die Auswertung der Symmetriebedingungen ergibt dann analog wie bei der Betrachtung im Falle eines Teilchens oben, daß

$$H = H(|\vec{x}_1 - \vec{x}_2|, |\vec{p}_1|, |\vec{p}_2|) \quad (1.13.45)$$

ist. Die Galilei-Boosts werden schließlich durch die **Schwerpunktskoordinaten**

$$\vec{K} = \sum_{j=1}^2 (m_j \vec{x}_j - t \vec{p}_j) = \sum_{j=1}^2 m_j \vec{x}_j - t \vec{P} \quad (1.13.46)$$

erzeugt. Die Symmetriebedingung verlangt unter Berücksichtigung von (1.13.43)

$$\sum_{j=1}^2 \{m_j \vec{x}_j, H\} - \vec{P} = m_1 \frac{\partial H}{\partial \vec{p}_1} + m_2 \frac{\partial H}{\partial \vec{p}_2} - \vec{p}_1 - \vec{p}_2 = 0. \quad (1.13.47)$$

Integriert man dies, findet man schließlich

$$H = \sum_{j=1}^2 \frac{1}{2m_j} \vec{p}_j^2 + V(|\vec{x}_1 - \vec{x}_2|), \quad (1.13.48)$$

d.h. die Bewegung zweier Teilchen, die über ein Zentralpotential wechselwirken. Die kanonischen Bewegungsgleichungen lauten

$$\dot{\vec{x}}_j = \frac{\partial H}{\partial \vec{p}_j} = \frac{1}{m_j} \vec{p}_j, \quad \dot{\vec{p}}_1 = -\dot{\vec{p}}_2 = -\frac{\partial H}{\partial \vec{x}_1} = -\frac{\vec{x}_1 - \vec{x}_2}{|\vec{x}_1 - \vec{x}_2|} V'(|\vec{x}_1 - \vec{x}_2|). \quad (1.13.49)$$

Wir können dieses Gleichungssystem allerdings erheblich vereinfachen, indem wir zu Schwerpunkts- und Relativkoordinaten übergehen, d.h.

$$\vec{X} = \frac{m_1 \vec{x}_1 + m_2 \vec{x}_2}{m_1 + m_2}, \quad \vec{r} = \vec{x}_1 - \vec{x}_2. \quad (1.13.50)$$

Dies können wir mittels der Betrachtungen in Abschnitt 1.11.1 zu einer kanonischen Transformation ergänzen. Dazu müssen wir (1.13.51) zunächst nach den \vec{x}_j auflösen (*nachrechnen!*):

$$\vec{x}_1 = \vec{X} + \frac{m_2}{m_1 + m_2} \vec{r}, \quad \vec{x}_2 = \vec{X} - \frac{m_1}{m_1 + m_2} \vec{r}. \quad (1.13.51)$$

Gemäß (1.11.8) ist die erzeugende Funktion der kanonischen Transformation

$$g(\vec{p}_1, \vec{p}_2, \vec{X}, \vec{r}) = -\vec{p}_1 \cdot \left(\vec{X} + \frac{m_2}{m_1 + m_2} \vec{r} \right) - \vec{p}_2 \cdot \left(\vec{X} - \frac{m_1}{m_1 + m_2} \vec{r} \right). \quad (1.13.52)$$

Damit ergibt sich für die kanonisch konjugierten Impulse zu \vec{X} und \vec{r}

$$\vec{P}_X = -\partial_{\vec{X}} g = \vec{p}_1 + \vec{p}_2, \quad \vec{P}_r = -\partial_{\vec{r}} g = \frac{1}{m_1 + m_2} (m_2 \vec{p}_1 - m_1 \vec{p}_2). \quad (1.13.53)$$

Umgekehrt ergibt sich (*nachrechnen!*)

$$\vec{p}_1 = \vec{P}_r + \frac{m_1}{m_1 + m_2} \vec{P}_X, \quad \vec{p}_2 = \vec{P}_r - \frac{m_2}{m_1 + m_2} \vec{P}_X \quad (1.13.54)$$

und schließlich (*nachrechnen!*)

$$H' = H = \frac{1}{2M} \vec{P}_X^2 + \frac{1}{2\mu} \vec{P}_r^2 + V(r), \quad (1.13.55)$$

wobei $M = m_1 + m_2$, $\mu = m_1 m_2 / M$ und $r = |\vec{r}|$ ist.

Wie aus dem Noether-Theorem zu erwarten, sind die Schwerpunktskoordinaten zyklisch und somit die Schwerpunktsimpulse \vec{P}_X Erhaltungsgrößen. Die Bewegungsgleichungen separieren folglich in die freie Bewegung des Schwerpunkts und die Relativbewegung, die durch die **effektive Einteilchen-Hamilton-Funktion**

$$H_{\text{eff}} = \frac{1}{2\mu} \vec{P}_r^2 + V(r) \quad (1.13.56)$$

beschrieben wird. Diese Hamilton-Funktion weist nun noch die Rotationssymmetrie um den Ursprung der Relativkoordinaten auf. Nach dem Noether-Theorem ist der entsprechende **Relativdrehimpuls**

$$\vec{\ell}_r = \vec{r} \times \vec{P}_r = \text{const} \quad (1.13.57)$$

erhalten. Wegen

$$\dot{\vec{r}} = \frac{\partial H_{\text{eff}}}{\partial \vec{P}_r} = \frac{1}{\mu} \vec{P}_r \quad (1.13.58)$$

folgt

$$\vec{\ell}_r = \mu \vec{r} \times \dot{\vec{r}} = \text{const}, \quad (1.13.59)$$

und das bedeutet, daß die Bewegung in einer Ebene senkrecht zu $\vec{\ell}_r$ stattfindet. Setzen wir $\vec{l}_r = l_r \vec{e}_3$ und führen Kugelkoordinaten ein, ist daher $\vartheta = \pi/2 = \text{const}$ und folglich der kanonisch konjugierte Impuls $P_\vartheta \propto \dot{\vartheta} = 0$. Der kanonisch konjugierte Impuls zu φ ist $p_\varphi = l_r$ (*nachrechnen*). Daraus ergibt sich schließlich gemäß (1.11.20)

$$H_{\text{eff}} = \frac{1}{2\mu} \left(P_r^2 + \frac{P_\varphi^2}{r^2} \right) + V(r). \quad (1.13.60)$$

Da also φ zyklisch ist, ist in der Tat $P_\varphi = \text{const}$ und damit das Problem auf die eindimensionale Radialbewegung mit der entsprechenden Hamilton-Funktion

$$H_{\text{eff}} = \frac{1}{2\mu} P_r^2 + V_{\text{eff}}(r), \quad V_{\text{eff}}(r) = \frac{1}{2\mu r^2} \ell_r^2 + V(r) \quad (1.13.61)$$

mit $\ell_r = p_\varphi = \text{const}$ reduziert.

Wir bemerken noch, daß sich diese Betrachtungen sehr leicht auf **Vielteilchensysteme** erweitern lassen. Ein Beispiel für ein entsprechendes Modell eines **abgeschlossenen Systems**, das entsprechend die volle Galilei-Symmetrie aufweist, ist durch ein über **Paarzentralwechselwirkungen** interagierende Punktteilchensystem gegeben. In kartesischen Koordinaten ergibt sich

$$H = \sum_{j=1}^N \frac{\vec{p}_j^2}{2m_j} + \frac{1}{2} \sum_{i \neq j} V_{ij}(|\vec{x}_i - \vec{x}_j|). \quad (1.13.62)$$

Der Faktor 1/2 rührt daher, daß die Wechselwirkungspotentiale für jedes Teilchenpaar nur einmal in die Gesamtenergie, also die Hamilton-Funktion, eingeht, aber in der Summe jeweils zweimal vorkommt (z.B. die Wechselwirkungsenergie zwischen Teilchen 1 und 2 zum einen für $i = 1, j = 2$ und zum anderen für $j = 1, i = 2$).

Die Bewegungsgleichungen lauten (*nachrechnen!*)

$$\dot{\vec{x}}_k = \vec{\nabla}_{\vec{p}_k} H = \frac{1}{m_j} \vec{p}_j, \quad \dot{\vec{p}}_k = -\vec{\nabla}_{\vec{x}_k} H = -\frac{\vec{x}_k - \vec{x}_j}{|\vec{x}_k - \vec{x}_j|} V'_{ij}(|\vec{x}_k - \vec{x}_j|). \quad (1.13.63)$$

Man kann leicht zeigen (*Übung!*), daß die Erzeugenden der Galilei-Transformationen durch die Hamilton-Funktion (Zeittranslationen), den Gesamtimpuls (Raumtranslationen)

$$\vec{P} = \sum_{j=1}^N \vec{p}_j, \quad (1.13.64)$$

Gesamtdrehimpuls (räumliche Drehungen)

$$\vec{\ell} = \sum_{j=1}^N \vec{x}_j \times \vec{p}_j \quad (1.13.65)$$

und für die Galilei-Boosts durch

$$\vec{K} = \sum_{j=1}^N m_j \vec{x}_j - \vec{P}t, \quad M = \sum_{j=1}^N m_j \quad (1.13.66)$$

gegeben ist.

1.13.4 Beispiel für dynamische Symmetrie: Isotroper harmonischer Oszillator

Wir betrachten nochmals das Beispiel des **harmonischen Oszillators** in einer Dimension. Die Hamilton-Funktion ist

$$H = \frac{p^2}{2m} + \frac{m\omega^2}{2}x^2. \quad (1.13.67)$$

Wir vereinfachen diese Hamiltonfunktion, indem wir zu neuen Phasenraumkoordinaten

$$X = \sqrt{m\omega}x, \quad P = \frac{p}{\sqrt{m\omega}} \quad (1.13.68)$$

übergehen. Dies ist offenbar eine kanonische Transformation, denn sie erfüllen offenbar die kanonischen Poisson-Klammerrelationen

$$\{X, X\} = m\omega \{x, x\}, \quad \{P, P\} = \frac{1}{m\omega} \{p, p\}, \quad \{X, P\} = \{x, p\} = 1. \quad (1.13.69)$$

Die Hamiltonfunktion ist nun in den neuen Variablen ausgedrückt.

$$H = \frac{\omega}{2}(P^2 + X^2). \quad (1.13.70)$$

Jetzt ist offensichtlich, daß diese Hamilton-Funktion unter ebenen Drehungen, also SO(2)-Transformationen der Art

$$\begin{pmatrix} X_\alpha \\ P_\alpha \end{pmatrix} = \begin{pmatrix} \cos \alpha & -\sin \alpha \\ \sin \alpha & \cos \alpha \end{pmatrix} \begin{pmatrix} X \\ P \end{pmatrix} \quad (1.13.71)$$

invariant ist. Um zu sehen, ob wir dies als kanonische Transformation darstellen können, bilden wir die Ableitung nach α bei $\alpha = 0$:

$$\left. \frac{d}{d\alpha} \begin{pmatrix} X_\alpha \\ P_\alpha \end{pmatrix} \right|_{\alpha=0} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} X \\ P \end{pmatrix} = \begin{pmatrix} -P \\ X \end{pmatrix}. \quad (1.13.72)$$

Wir suchen also eine Erzeugende $G = G(X, P)$ mit

$$\mathfrak{L}_G X = \{X, G\} = \frac{\partial G}{\partial P} = -P, \quad \mathfrak{L}_G P = \{P, G\} = -\frac{\partial G}{\partial X} = X. \quad (1.13.73)$$

Dies läßt sich sofort zu

$$G(X, P) = -\frac{1}{2}(P^2 + X^2) = -\frac{1}{\omega}H(X, P) \quad (1.13.74)$$

integrieren.

Diese Symmetrie führt also auf nichts neues, denn da H nicht explizit zeitabhängig ist, ist H eine Erhaltungsgröße (bis auf konstante Faktoren die einzige). Es ist nun allerdings klar, daß wir mit (1.13.72) bereits die Lösung der Bewegungsgleichungen gefunden haben, denn es ist ja wegen (1.13.71)

$$\begin{aligned} \begin{pmatrix} X(t) \\ Y(t) \end{pmatrix} &= \exp(t\mathfrak{L}_{H_0}) \begin{pmatrix} X_0 \\ P_0 \end{pmatrix} = \exp(-\omega t \mathfrak{L}_{G_0}) \begin{pmatrix} X_0 \\ P_0 \end{pmatrix} = \begin{pmatrix} X_{\alpha=-\omega t} \\ P_{\alpha=-\omega t} \end{pmatrix} \\ &= \begin{pmatrix} \cos(\omega t) & \sin(\omega t) \\ -\sin(\omega t) & \cos(\omega t) \end{pmatrix} \begin{pmatrix} X_0 \\ P_0 \end{pmatrix} \\ &= \begin{pmatrix} X_0 \cos(\omega t) + P_0 \sin(\omega t) \\ P_0 \cos(\omega t) - X_0 \sin(\omega t) \end{pmatrix}. \end{aligned} \quad (1.13.75)$$

Dies löst offenbar in der Tat die kanonischen Gleichungen

$$\dot{X} = \frac{\partial H}{\partial P} = \omega P, \quad \dot{P} = -\frac{\partial H}{\partial X} = -\omega X \quad (1.13.76)$$

mit beliebig vorgegebenen Anfangsbedingungen $X(t=0) = X_0$ und $P(t=0) = P_0$, d.h. (1.13.75) ist die vollständige Lösung der Bewegungsgleichungen.

Weniger trivial wird diese Symmetrie für den isotropen harmonischen Oszillator. Dann ist die Hamilton-Funktion in kartesischen Koordinaten durch

$$H = \frac{\vec{p}^2}{2m} + \frac{m\omega^2}{2}\vec{x}^2 \quad (1.13.77)$$

gegeben. Genau wie beim obigen Beispiel für den eindimensionalen Oszillator ist wieder

$$\vec{X} = \sqrt{m\omega}\vec{x}, \quad \vec{P} = \frac{1}{\sqrt{m\omega}}\vec{p} \quad (1.13.78)$$

eine kanonische Transformation, und in diesen Koordinaten ist

$$H = \frac{\omega}{2}(\vec{X}^2 + \vec{P}^2). \quad (1.13.79)$$

Offenbar ist diese Hamilton-Funktion nicht nur unter räumlichen Drehungen $\vec{X}' = \hat{D}\vec{X}$, $\vec{P}' = \hat{D}\vec{P}$ mit $\hat{D} \in \text{SO}(3)$ invariant sondern unter beliebigen Drehungen der sechs Phasenraumvariablen. Schreiben wir

$$\vec{\Xi} = \begin{pmatrix} \vec{X} \\ \vec{P} \end{pmatrix} \quad (1.13.80)$$

und definieren das „kanonische Skalarprodukt“ in diesem sechsdimensionalen Raum

$$\vec{\Xi}_1 \cdot \vec{\Xi}_2 = \vec{X}_1 \cdot \vec{X}_2 + \vec{P}_1 \cdot \vec{P}_2, \quad (1.13.81)$$

ist klar, daß $H = \omega\vec{\Xi}^2/2$ invariant unter der Transformation

$$\vec{\Xi}' = \hat{O}\vec{\Xi}, \quad \hat{O} \in \text{SO}(6). \quad (1.13.82)$$

1. Variationsrechnung und kanonische Mechanik

Dabei ist die $SO(6)$ wieder die Gruppe der $\mathbb{R}^{6 \times 6}$ -Matrizen, die durch (1.13.81) Skalarprodukte beliebiger Vektoren $\vec{\Xi}_1$ und $\vec{\Xi}_2$ invariant läßt, d.h. für die

$$\hat{O}^T \hat{O} = \hat{O} \hat{O}^T = \mathbb{1}_6 \quad (1.13.83)$$

gilt.

Wir betrachten nun wieder infinitesimale Transformationen. Da weiter das Skalarprodukt (1.13.81) wieder durch kartesische Koordinaten ausgedrückt ist, unterscheiden wir auch hier ko- und kontravariante Indizes nicht und notieren alle Indizes als untere Indizes. Für eine „infinitesimale“ $SO(6)$ -Transformation $O_{ij} = \delta_{ij} + \delta \omega_{ij}$ folgt wieder aus der Invarianz der Skalarprodukte (*nachrechnen!*) $\delta \omega_{ij} = -\delta \omega_{ji}$.

Allerdings wollen wir nur solche Symmetrien betrachten, die wir als kanonische Transformationen realisieren können. Da die oben betrachteten $SO(6)$ -Transformationen linear in $\vec{\Xi} = (\vec{X}, \vec{p})$ sind müssen die entsprechende Erzeugende bilinear in diesen Variablen sein, d.h. wir machen den Ansatz

$$g = \alpha_{ij} X_i X_j + \beta_{ij} X_i P_j + \gamma_{ij} P_i P_j. \quad (1.13.84)$$

Damit dies die Erzeugende einer Symmetrietransformation ist, muß sie nach dem Noether-Theorem erhalten sein. Verwenden wir die kanonischen Gleichungen

$$\dot{\vec{X}} = \frac{\partial H}{\partial \vec{P}} = \omega \vec{P}, \quad \dot{\vec{P}} = -\frac{\partial H}{\partial \vec{X}} = -\omega \vec{P}, \quad (1.13.85)$$

ergibt sich für die Zeitableitung von (1.13.86)

$$\dot{g} = \omega [(\alpha_{ij} - \gamma_{ij})(X_i P_j + X_j P_i) + \beta_{ij}(P_i P_j - X_i X_j)]. \quad (1.13.86)$$

Damit dies für alle möglichen Anfangsbedingungen an (\vec{X}, \vec{P}) verschwindet, muß offenbar $\alpha_{ij} = \gamma_{ij}$ und $\beta_{ij} = -\beta_{ji}$ sein, d.h. die Generatoren werden durch alle symmetrischen Matrizen $\alpha_{ij} = \gamma_{ij}$ und antisymmetrischen Matrizen $\beta_{ij} = -\beta_{ji}$ charakterisiert. Damit haben wir die folgenden 9 unabhängigen Generatoren

$$A_{ij} = \frac{1}{2}(X_i X_j + P_i P_j), \quad B_{ij} = X_i P_j - X_j P_i = \epsilon_{ijk} \ell_k, \quad \vec{\ell} = \vec{X} \times \vec{P}. \quad (1.13.87)$$

Wir haben die willkürlichen Koeffizienten geeignet gewählt. Es ist klar, daß wir die gewöhnlichen Drehungen, die von den Drehimpulsen $\vec{\ell}$ generiert werden, als Symmetriegruppe erhalten. Dies sind die Drehungen, die gleichermaßen auf \vec{X} und \vec{P} wirken. Betrachten wir nun die A_{ij} . Die entsprechende kanonische Transformation wird wieder durch die Differentialgleichung (1.12.1) bestimmt

$$\begin{aligned} \frac{dX_k}{d\alpha} &= \mathfrak{L}_{A_{ij}} X_k = \frac{1}{2} \{X_k, P_i P_j\} = \frac{1}{2} (\{X_k, P_i\} P_j + P_i \{X_k, P_j\}) = \frac{1}{2} (\delta_{ik} P_j + P_i \delta_{kj}), \\ \frac{dP_k}{d\alpha} &= \mathfrak{L}_{A_{ij}} P_k = \frac{1}{2} \{P_k, X_i X_j\} P_k = \frac{1}{2} (X_i \{P_k, X_j\} + \{P_k, X_i\} X_j) = -\frac{1}{2} (\delta_{jk} X_i + \delta_{ik} X_j). \end{aligned} \quad (1.13.88)$$

Betrachten wir das Beispiel $i = 1, j = 2$, ergibt sich

$$\frac{d\vec{X}}{d\alpha} = \frac{1}{2} \begin{pmatrix} P_2 \\ P_1 \\ 0 \end{pmatrix}, \quad \frac{d\vec{P}}{d\alpha} = -\frac{1}{2} \begin{pmatrix} X_2 \\ X_1 \\ 0 \end{pmatrix}. \quad (1.13.89)$$

Die Lösung dieses Differentialgleichungssystems (*Übung*) ergibt, in bequemer Kombination von Variablen Paaren zu zweidimensionalen Vektoren ausgedrückt

$$\begin{aligned} \begin{pmatrix} X_1(\alpha) \\ P_2(\alpha) \end{pmatrix} &= \begin{pmatrix} \cos(\alpha/2) & \sin(\alpha/2) \\ -\sin(\alpha/2) & \cos(\alpha/2) \end{pmatrix} \begin{pmatrix} X_1 \\ P_2 \end{pmatrix}, \\ \begin{pmatrix} X_2(\alpha) \\ P_1(\alpha) \end{pmatrix} &= \begin{pmatrix} \cos(\alpha/2) & \sin(\alpha/2) \\ -\sin(\alpha/2) & \cos(\alpha/2) \end{pmatrix} \begin{pmatrix} X_2 \\ P_1 \end{pmatrix}. \end{aligned} \quad (1.13.90)$$

Ein Generator A_{ij} mit $i \neq j$ generiert also eine simultane Drehung in der (X_i, P_j) und (X_j, P_i) um denselben Winkel $\alpha/2$. Dies entspricht auch sicher einer speziellen $SO(6)$ -Transformation. Für $i = j$ erhalten wir offenbar eine Drehung in der (X_i, P_i) -Ebene um den Winkel α (wie im eindimensionalen Fall oben).

Natürlich können nicht alle 9 Erhaltungsgrößen unabhängig voneinander sein. Offenbar entspricht z.B.

$$\sum_{i=1}^3 A_{ii} = \frac{1}{2}(\vec{X}^2 + \vec{P}^2) = \frac{H}{\omega}. \quad (1.13.91)$$

Insgesamt kann es nur maximal 6 unabhängige Erhaltungsgrößen geben, denn es gibt auch nur 6 linear unabhängige Anfangsbedingungen, d.h. wir können nur $\vec{X}(t=0) = \vec{X}_0$ und $\vec{P}(t=0) = \vec{P}_0$ unabhängig vorgeben. Dann ist die Lösung der Bewegungsgleichungen eindeutig bestimmt und folglich auch die Erhaltungsgrößen. Um nun die Lösungen der Gleichung qualitativ zu verstehen, können wir die Erhaltungsgrößen ausnutzen, ohne die Bewegungsgleichungen explizit lösen zu müssen. Zunächst bedeutet die Erhaltung des Drehimpulses $\vec{\ell} = \vec{x} \times \vec{p} = m\vec{x} \times \dot{\vec{x}}$, daß die Bewegung in einer Ebene senkrecht zum Drehimpuls erfolgt, wobei wir annehmen, daß $\vec{\ell} \neq 0$. Wir können durch eine geeignete Drehung der Koordinatenachsen offenbar stets erreichen, daß $\vec{\ell} = \ell \vec{e}_3$ ist, d.h. die Bewegung in der 12-Ebene verläuft. Dann ist effektiv $H = \omega(A_{11} + A_{22})$. Da $\{A_{11}, A_{22}\}$ vertauschen, ist folglich entsprechend der Analogie zum eindimensionalen Fall [vgl. (1.13.75)]

$$\begin{pmatrix} X_1(t) \\ P_1(t) \end{pmatrix} = \begin{pmatrix} X_{10} \cos(\omega t) + P_{10} \sin(\omega t) \\ P_{10} \cos(\omega t) - X_{10} \sin(\omega t) \end{pmatrix}, \quad \begin{pmatrix} X_2(t) \\ P_2(t) \end{pmatrix} = \begin{pmatrix} X_{20} \cos(\omega t) + P_{20} \sin(\omega t) \\ P_{20} \cos(\omega t) - X_{20} \sin(\omega t) \end{pmatrix}. \quad (1.13.92)$$

Die Bahnkurve können wir übersichtlicher in der Form

$$\vec{X} = \vec{A} \cos(\omega t) + \vec{B} \sin(\omega t), \quad \vec{A} = \vec{X}_0, \quad \vec{B} = \vec{P}_0/\omega. \quad (1.13.93)$$

schreiben.

Um die Bahnkurve einfacher zu identifizieren, können wir zusammen mit einer Phase φ neue Vektoren \vec{A}' und \vec{B}' einführen gemäß

$$\vec{X} = \vec{A}' \cos(\omega t + \varphi) + \vec{B}' \sin(\omega t + \varphi) \quad (1.13.94)$$

Mit den Additionstheoremen der Winkelfunktionen folgt

$$\vec{A} = \vec{A}' \cos \varphi + \vec{B}' \sin \varphi, \quad \vec{B} = -\vec{A}' \sin \varphi + \vec{B}' \cos \varphi \quad (1.13.95)$$

bzw. umgekehrt

$$\vec{A}' = \vec{A} \cos \varphi - \vec{B} \sin \varphi, \quad \vec{B}' = \vec{A} \sin \varphi + \vec{B} \cos \varphi. \quad (1.13.96)$$

Nun wollen wir verlangen, daß \vec{A}' und \vec{B}' aufeinander senkrecht stehen, d.h.

$$\vec{A}' \cdot \vec{B}' = 0 \Rightarrow (\vec{A}^2 - \vec{B}^2) \sin \varphi \cos \varphi + \vec{A} \cdot \vec{B} (\cos^2 \varphi - \sin^2 \varphi) = 0. \quad (1.13.97)$$

Mit den Doppelwinkeltheoremen folgt

$$\frac{1}{2}(\vec{A}^2 - \vec{B}^2) \sin(2\varphi) = -\vec{A} \cdot \vec{B} \cos(2\varphi). \quad (1.13.98)$$

Falls $\vec{A} \cdot \vec{B} = 0$, können wir natürlich $\varphi = 0$ setzen, um diese Gleichung zu erfüllen. Andernfalls brauchen wir nur

$$\varphi = \frac{1}{2} \operatorname{arccot} \left(\frac{\vec{B}^2 - \vec{A}^2}{2\vec{A} \cdot \vec{B}} \right) \quad (1.13.99)$$

zu setzen. Dann ist klar, daß (1.13.94) eine Ellipse in der Ebene senkrecht zum Drehimpuls mit Halbachsen \vec{A}' bzw. \vec{B}' .

Wir kommen noch einmal zur durch die Generatoren (1.13.87) definierten Symmetriegruppe des isotropen harmonischen Oszillators zurück. Wir können die Gruppe noch durch etwas andere Generatoren charakterisieren. Dazu definieren wir zunächst die komplexen Vektoren

$$\vec{a} = \frac{1}{\sqrt{2}}(\vec{X} + i\vec{P}). \quad (1.13.100)$$

Dann gilt

$$H = \omega \vec{a}^* \cdot \vec{a} =: \omega \langle \vec{a} | \vec{a} \rangle. \quad (1.13.101)$$

Für komplexe Vektoren $\vec{a}, \vec{b} \in \mathbb{C}^3$ definiert die **Sesquilinearform**

$$\langle \vec{a} | \vec{b} \rangle := \vec{a}^* \cdot \vec{b} = a_1^* b_1 + a_2^* b_2 + a_3^* b_3 \quad (1.13.102)$$

ein Skalarprodukt und macht \mathbb{C}^3 zu einem **unitären komplexen Vektorraum**. Es gelten die üblichen Rechenregeln für das Skalarprodukt, nur daß (1.13.103) bzgl. des ersten Argumente **antilinear** ist, d.h. es gilt für beliebige $\lambda_1, \lambda_2 \in \mathbb{C}$ und Vektoren $\vec{a}_1, \vec{a}_2 \in \mathbb{C}^3$

$$\langle \lambda_1 \vec{a}_1 | \vec{b} \rangle = \lambda_1^* \langle \vec{a}_1 | \vec{b} \rangle + \lambda_2^* \langle \vec{a}_2 | \vec{b} \rangle \quad (1.13.103)$$

und beim Vertauschen der Argumente haben wir

$$\langle \vec{b} | \vec{a} \rangle = \vec{b}^* \cdot \vec{a} = (\vec{a}^* \cdot \vec{b})^* = \langle \vec{a} | \vec{b} \rangle^*. \quad (1.13.104)$$

Lesen wir \vec{a} als Spaltenvektor und definieren den **hermitesch adjungierten Vektor** als den Zeilenvektor $\vec{a}^\dagger = (a_1^*, a_2^*, a_3^*)$ Für irgendwelche Matrizen $\hat{M} \in \mathbb{C}^{3 \times 3}$ gilt

$$\langle \vec{a} | \hat{M} \vec{b} \rangle = \vec{a}^\dagger \hat{M} \vec{b} = (\hat{M}^\dagger \vec{a})^\dagger \vec{b} = \langle \hat{M}^\dagger \vec{a} | \vec{b} \rangle, \quad (1.13.105)$$

wobei die **hermitesch adjungierte Matrix** durch

$$(\hat{M}^\dagger)_{jk} = M_{kj}^* \quad (1.13.106)$$

definiert ist, d.h. die transponierte und komplex konjugierte Matrix.

Es ist nun natürlich wie beim reellen Skalarprodukt nach solchen Matrizen zu fragen, die das unitäre Skalarprodukt im \mathbb{C}^3 invariant lassen. Solche Matrizen nennt man **unitäre Matrizen**. Für sie muß für alle Vektoren $\vec{a}, \vec{b} \in \mathbb{C}^3$

$$\langle \hat{U} \vec{a} | \hat{U} \vec{b} \rangle = \langle \vec{a} | \vec{b} \rangle \quad (1.13.107)$$

gelten. Nun ist

$$\langle \hat{U} \vec{a} | \hat{U} \vec{b} \rangle = \langle \hat{U}^\dagger \hat{U} \vec{a} | \vec{b} \rangle. \quad (1.13.108)$$

Damit demnach (1.13.107) für alle $\vec{a}, \vec{b} \in \mathbb{C}^3$ erfüllt sein kann muß

$$\hat{U}^\dagger \hat{U} = \mathbb{1}_3 \quad (1.13.109)$$

sein. Daraus folgt aber bekanntlich auch $\hat{U} \hat{U}^\dagger = \mathbb{1}$, d.h. die Matrix \hat{U} ist genau dann unitär, wenn $\hat{U}^{-1} = \hat{U}$ ist. Offensichtlich bilden diese Matrizen eine Gruppe (*warum?*). Diese nennt man die **Gruppe der unitären**

Transformationen im dreidimensionalen komplexen unitären Vektorraum und bezeichnet sie mit $U(3)$. Es gilt weiter

$$\det(\hat{U}^\dagger \hat{U}) = \det \mathbb{1}_3 = 1 \Rightarrow |\det \hat{U}| = 1. \quad (1.13.110)$$

Die unitären Matrizen mit der Determinante 1 bilden eine Untergruppe der $U(3)$, die **spezielle unitäre Gruppe** $SU(3)$ (*warum?*).

Wir betrachten nun infinitesimal von der Identität abweichende unitäre Transformationen $\hat{U} = \mathbb{1}_3 + i\delta\alpha \hat{T}$ ⁷. Die Unitaritätsbedingung (1.13.109) verlangt dann für $\delta\alpha \in \mathbb{R}$

$$\hat{U}^\dagger \hat{U} = (\mathbb{1}_3 - i\delta\alpha \hat{T}^\dagger)(\mathbb{1}_3 + i\delta\alpha \hat{T}) = \mathbb{1}_3 + i\delta\alpha(\hat{T} - \hat{T}^\dagger) + \mathcal{O}(\delta\alpha^2) \stackrel{!}{=} \mathbb{1}_3 \Rightarrow \hat{T}^\dagger = \hat{T}. \quad (1.13.111)$$

Man nennt die Matrix \hat{T} **selbstadjungiert**.

Eine beliebige Matrix $\hat{T} \in \mathbb{C}^{n \times n}$ besitzt $2n^2$ reelle Parameter. Ist sie selbstadjungiert, gilt $T_{ij}^* = T_{ji}$. Demnach sind die Diagonalelemente reell und die Elemente unterhalb der Diagonalen sind komplex konjugiert zu denen oberhalb der Diagonalen. Folglich besitzt eine selbstadjungierte $\mathbb{C}^{n \times n}$ -Matrix $n + (n^2 - n) = n^2$ unabhängige reelle Parameter. Für $n = 3$ sind dies gerade 9 freie reelle Parameter.

Wir zeigen nun, daß die 9 infinitesimalen Transformationen, die durch die Generatoren (1.13.84) definiert sind gerade die $U(3)$ -Transformationen der Vektoren \vec{a} (1.13.100) entsprechen. Es ist klar, daß wegen (1.13.101) die Hamilton-Funktion des harmonischen Oszillators unter beliebigen $U(3)$ -Transformationen invariant ist. Dazu betrachten wir nun die Poisson-Klammern für die Komponenten der Vektoren \vec{a} und \vec{a}^* . Offenbar gilt

$$\begin{aligned} \{a_j, a_k^*\} &= \frac{1}{2} \{X_j + iP_j, X_k - iP_k\} = -\frac{i}{2} (\{X_j, P_k\} - \{P_j, X_k\}) = -i\delta_{jk}, \\ \{a_j, a_k\} &= \frac{1}{2} \{X_j + iP_j, X_k + iP_k\} = \frac{i}{2} (\{X_j, P_k\} + \{P_j, X_k\}) = 0, \\ \{a_j^*, a_k^*\} &= \frac{1}{2} \{X_j - iP_j, X_k - iP_k\} = -\frac{i}{2} (\{X_j, P_k\} + \{P_j, X_k\}) = 0. \end{aligned} \quad (1.13.112)$$

Weiter folgt aus den Bewegungsgleichungen (1.13.103)

$$\dot{\vec{a}} = \mathfrak{L}_H \vec{a} = \{\vec{a}, H\} = \vec{X} + i\vec{P} = \omega(\vec{P} - i\vec{X}) = -i\omega\vec{a} \Rightarrow \dot{\vec{a}}^* = +i\omega\vec{a}^*. \quad (1.13.113)$$

Nun gilt

$$\tilde{A}_{ij} = (a_i^* a_j + a_j^* a_i) = (X_i X_j + P_i P_j) = 2A_{ij}, \quad \tilde{B}_{ij} = i(a_i a_j^* - a_j a_i^*) = X_i P_j - X_j P_i = \epsilon_{ijk} \ell_k = B_{ij}. \quad (1.13.114)$$

Die von den Generatoren (1.13.84) aufgespannte Lie-Algebra läßt sich also auch durch \tilde{A} und \tilde{B} aufspannen. Die entsprechenden Lie-Ableitungen von \vec{a} ergeben sich also zu

$$\mathfrak{L}_{\tilde{A}_{ij}} \vec{a} = \vec{e}_k \{a_k, a_i^* a_j + a_i a_j^*\} = -i\vec{e}_k (\delta_{ki} a_j + \delta_{kj} a_i) = i(\vec{e}_i a_j + \vec{e}_j a_i) = -i\hat{A}_{ij} \vec{a}. \quad (1.13.115)$$

Dabei ist \hat{A}_{ij} für $i \neq j$ eine symmetrische rein reelle (also auch selbstadjungierte) Matrix. Dies sind offenbar drei voneinander linear unabhängige symmetrische selbstadjungierte Matrizen, die eine Basis für alle diese antisymmetrischen selbstadjungierten Matrizen bilden. Angewandt auf den Vektor \vec{a} vertauschen sie die beiden Komponenten mit indizes i und j und annullieren alle übrigen Komponenten. Die Linearkombination von Generatoren

$$\frac{1}{\omega} H = \frac{1}{2} \sum_{i=1}^3 A_{ii} \quad (1.13.116)$$

⁷Der Faktor i erweist sich gleich noch als bequem, ist aber nur Konvention.

entspricht der Matrix $-i/2 \sum_i \hat{A}_{ii} = -i\mathbb{1}_3$. Offenbar generiert sie diejenigen Transformationen, die \vec{a} mit einem gemeinsamen Phasenfaktor $\exp(-i\alpha)$ mit $\alpha \in \mathbb{R}$ multiplizieren.

Weiter gilt

$$\mathfrak{L}_{\hat{B}_{ij}} \vec{a} = i\vec{e}_k \{a_k, a_i a_j^* - a_i^* a_j\} = -i\vec{e}_k (i\delta_{jk} a_i - i\delta_{ik} a_j) = -i(i\vec{e}_j a_i - i\vec{e}_i a_j) = -i\hat{B}_{ij} \vec{a}. \quad (1.13.117)$$

Dabei sind die \hat{B}_{ij} 3 linear unabhängige antisymmetrische rein imaginäre (und damit auch selbstadjungierte) 3×3 -Matrizen. Offenbar lassen sich mit den \hat{A}_{ij} und \hat{B}_{ij} alle selbstadjungierten $\mathbb{C}^{3 \times 3}$ -Matrizen durch Linearkombination aufbauen. Wir können also mit den Lie-Ableitungen $\mathfrak{L}_{\hat{A}_{ij}}$ und $\mathfrak{L}_{\hat{B}_{ij}}$ alle U(3)-Transformationen generieren, denn man kann zeigen, daß sich alle U(3)-Matrizen in der Form $\exp(i\hat{T})$ mit $\hat{T} = \hat{T}^\dagger$ schreiben lassen. Folglich ist die vollständige Symmetriegruppe des harmonischen Oszillators im Sinne eines Hamiltonschen Systems die U(3).

Diese Betrachtungen entfalten ihre volle Schönheit allerdings erst in ihrer Anwendung auf den harmonischen Oszillator in der Quantentheorie. Die übliche Wahl der 9 selbst-adjungierten Matrizen der U(3) ist neben der Einheitsmatrix $\mathbb{1}_3$ (die wie oben gesehen die Multiplikation des Vektors $\vec{a} \in \mathbb{C}^3$ mit einem Phasenfaktor generiert) sind die folgenden 8 **Gell-Mann-Matrizen**

$$\begin{aligned} \hat{\lambda}_1 = \hat{A}_{12} &= \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, & \hat{\lambda}_2 = \hat{B}_{12} &= \begin{pmatrix} 0 & -i & 0 \\ -i & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, & \hat{\lambda}_3 = \frac{1}{2}(\hat{A}_{11} - \hat{A}_{22}) &= \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \\ \hat{\lambda}_4 = \hat{A}_{13} &= \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}, & \hat{\lambda}_5 = \hat{B}_{13} &= \begin{pmatrix} 0 & 0 & -i \\ 0 & 0 & 0 \\ i & 0 & 0 \end{pmatrix}, & \hat{\lambda}_6 = \hat{A}_{23} &= \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}, \\ \hat{\lambda}_7 = \hat{B}_{23} &= \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -i \\ 0 & i & 0 \end{pmatrix}, & \hat{\lambda}_8 = \frac{1}{\sqrt{3}}(\hat{A}_{11} + \hat{A}_{22} - 2\hat{A}_{33}) &= \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -2 \end{pmatrix}. \end{aligned} \quad (1.13.118)$$

Diese selbstadjungierten Matrizen sind zum einen alle **spurlos**, d.h.

$$\text{tr } \hat{\lambda}_j = \sum_{k=1}^3 (\lambda_j)_{kk} = 0 \quad (1.13.119)$$

und **orthogonal** im Sinne des auf Matrizen definierten **Frobenius-Skalarprodukts** (*nachrechnen*)⁸

$$\text{tr}(\hat{\lambda}_j^\dagger \hat{\lambda}_k) = \text{tr}(\hat{\lambda}_j \hat{\lambda}_k) = 2\delta_{jk}. \quad (1.13.120)$$

Sie bilden eine Lie-Algebra mit dem Kommutator als Lie-Produkt. Die über die Exponentialfunktion gegebenen Matrizen

$$\hat{U} = \exp\left(-\frac{i}{2} \sum_{j=1}^8 \alpha_j \hat{\lambda}_j\right) \quad (1.13.121)$$

sind wegen

$$\hat{U}^\dagger = \exp\left(+\frac{i}{2} \sum_{j=1}^8 \alpha_j \hat{\lambda}_j\right) = \hat{U}^{-1} \quad \text{und} \quad \det \exp(\hat{A}) = \exp(\text{tr } \hat{A}) \quad (1.13.122)$$

allesamt speziell unitäre Matrizen⁹. Die $\hat{\lambda}_j$ sind eine Basis für die zur Gruppe SU(3) gehörige Lie-Algebra $\mathfrak{su}(3)$. Die Matrizen λ_1 , λ_2 und λ_3 besitzen nur von 0 verschiedene Matrix-Elemente für die ersten beiden

⁸Details zur Matrizenrechnung s. Anhang A.

⁹Die hier verwendeten Sätze über Matrizen, die Matrixexponentialfunktion und Determinanten finden sich ebenfalls in Anhang A.

Komponenten der Vektoren $\vec{a} \in \mathbb{C}^3$. Daher bilden sie eine Unter-Lie-Algebra und die entsprechenden Matrix-exponentialabbildungen liefern SU(2)-Matrizen, d.h. die Unter-Lie-Algebra ist die entsprechende Lie-Algebra $\mathfrak{su}(2)$.

1.14 Die Symmetrien des Keplerproblems

Ein interessantes Beispiel für die Anwendung der Symmetrieprinzipien ist auch das Keplerproblems, das wir ja schon in Abschnitt 1.7 im Rahmen der Lagrangeschen Formulierung behandelt haben. Wir können uns auf das restringierte Problem beschränken, d.h. wir haben bereits die Schwerpunktsbewegung absepariert, und wir können das System so behandeln, daß wir ein festes Zentrum („Sonne“), welches wir in den Koordinatenursprung eines kartesischen Koordinatensystems setzen, vor uns haben, um das sich der Körper („Planet“) der reduzierten Masse μ bewegt. Der Hamiltonoperator dieses restringierten Problems ist in kartesischen Koordinaten durch

$$H(\vec{x}, \vec{p}) = \frac{\vec{p}^2}{2\mu} - \frac{K}{|\vec{x}|}, \quad \text{mit } K = Gm_1m_2 \quad (1.14.1)$$

gegeben, wobei G die Newtonsche Gravitationskonstante ist.

Unmittelbar ersichtlich ist zunächst die verbliebene **Rotationssymmetrie**, die freilich aus der Rotationssymmetrie des Gravitationspotentials resultiert und **jedem Zentralkraftproblem** zukommt. Wie wir wissen, sind die dazugehörigen Noetherschen Erhaltungsgrößen die drei Komponenten des Drehimpulses:

$$\vec{l} = \vec{x} \times \vec{p}. \quad (1.14.2)$$

Im Sinne der obigen Liegruppentheorie sind diese gemäß dem Noetherschen Theorem die Erzeugenden der Drehungen, d.h.

$$D(\vec{\alpha}) = \exp(\mathfrak{L}_{\vec{\alpha} \cdot \vec{l}}) \quad (1.14.3)$$

ist der Operator einer Drehung um die Drehachse in Richtung $\vec{\alpha}/|\alpha|$ um den Drehwinkel α .

Durch direktes Nachrechnen, was ein wenig Ableitungsarbeit erfordert, weist man nach, daß außer dem Drehimpuls und selbstverständlich der Hamiltonfunktion selbst auch der **Runge-Lenz-Vektor** erhalten ist, also die Poissonklammern seiner Komponenten mit der Hamiltonfunktion, verschwinden. Dieser ist durch

$$\vec{A} = \frac{1}{K\mu} \left(\vec{p} \times \vec{l} - \mu \vec{x} \frac{K}{|\vec{x}|} \right) \quad (1.14.4)$$

gegeben. Einfacher als die Berechnung der Poissonklammern ist die Bildung der Zeitableitung, die aufgrund der Bewegungsgleichungen verschwindet (*Übung*).

Zunächst ist es nunmehr trivial, die Bahnkurve anzugeben. Dazu multiplizieren wir (1.14.4) mit \vec{x}

$$\vec{x} \cdot \vec{A} = r|\vec{A}| \cos \phi = \frac{\vec{l}^2}{K\mu} - r \Rightarrow r = \frac{\vec{l}^2}{K\mu} \frac{1}{1 + |\vec{A}| \cos \phi}. \quad (1.14.5)$$

Dabei ist ϕ der Winkel zwischen dem zeitlich konstanten Runge-Lenz-Vektor \vec{A} und dem Ortsvektor \vec{x} und $r = |\vec{x}|$. Der Vergleich mit (1.7.12) zeigt, daß

$$p = \frac{\vec{l}^2}{K\mu}, \quad \epsilon = |\vec{A}| \quad (1.14.6)$$

Parameter und Exzentrizität des Kegelschnittes sind.

Es ist klar, daß mit \vec{A} auch \vec{A}^2 eine Erhaltungsgröße ist, die aber nicht unabhängig von den übrigen Erhaltungsgrößen sein kann, denn \vec{l} und \vec{A} bilden ja zusammen schon sechs Erhaltungsgrößen. In der Tat liefert die direkte Rechnung

$$\epsilon^2 = \vec{A}^2 = 1 + \frac{2E\vec{l}^2}{K^2\mu}. \quad (1.14.7)$$

Dabei ist E die Gesamtenergie des Systems, welche natürlich auch erhalten ist. Der Vergleich von (1.14.6) zeigt vollständige Übereinstimmung mit (1.7.13), wo wir durch, verglichen mit der quasi algebraischen Methode, derer wir uns hier bedient haben, mühsame Integrationen zu diesem Resultat gelangt sind.

Mit (1.14.5) ist auch erwiesen, daß \vec{A} richtungsmäßig vom Zentrum zum Perihel des Kegelschnitts zeigt, und das entspricht auch der Zählung des Winkels ϕ , der hier freilich genau wie in Abschnitt 1.7 gewählt wurde. Wir sehen wieder, daß für $\vec{l} \neq 0$ die konkrete Bahnform durch das Vorzeichen der Gesamtenergie bestimmt ist:

- (1) $E < 0$: Dann ist nach (1.14.7) $0 \leq \epsilon < 1$, und der Kegelschnitt ist eine Ellipse, die sich für $\epsilon = 0$ zum Kreis spezialisiert. Kehrt man in diesem Falle das Vorzeichen von \vec{A} um, erhielte man einen Vektor, der vom Zentrum zum Aphel weist, was freilich an den Resultaten und der Bedeutung von \vec{A} als zusätzlichen Symmetrieparametern nichts ändern würde.
- (2) $E = 0$: Dann ist nach (1.14.7) $\epsilon = 0$, und die Bahn eine Parabel.
- (3) $E > 0$: Dann ist nach (1.14.7) $\epsilon > 1$, und die Bahn eine Hyperbel.

Wir haben bis jetzt letztlich durch viel einfachere Rechnung lediglich die gleichen Resultate wie in Abschnitt 1.7 erhalten. Mathematisch interessant ist im jetzigen Zusammenhang die Identifikation der Symmetriegruppe, die durch die 6 unabhängigen Erhaltungsgrößen \vec{l} und \vec{a} erzeugt wird. Dazu bilden wir die Poissonklammern der Komponenten von \vec{A} untereinander sowie mit \vec{l} , denn von \vec{l} kennen wir diese bereits:

$$\{l_j, l_k\} = \epsilon_{jkl} l_l. \quad (1.14.8)$$

Die Poissonklammern der Komponenten von \vec{A} untereinander und mit \vec{l} erhält man durch direkte Rechnung (am besten unter Zuhilfenahme eines Computer-Algebrasystems):

$$\begin{aligned} \{A_j, A_k\} &= -\frac{2E}{K^2\mu} \epsilon_{jkl} l_l, \\ \{l_j, A_k\} &= \epsilon_{jkl} A_l. \end{aligned} \quad (1.14.9)$$

Letzteres wissen wir allerdings auch ohne Rechnung, denn es ist klar, daß \vec{A} sich unter Drehungen wie ein Vektor verhält. Wie wir aus (1.14.8-1.14.9) ersehen, bilden \vec{l} und \vec{A} eine **Unterliealgebra der allgemeinen Poissonklammer-Liealgebra**. Durch Exponentiation und Bildung des algebraischen Abschlusses ergibt sich also insgesamt eine durch diese sechs Generatoren erzeugte Lie-Gruppe.

Wir wollen nunmehr die Gruppen identifizieren, welche mit jedem der drei Fälle $E < 0$, $E = 0$ und $E > 0$ verbunden sind. Wir werden auch gleich sehen, warum wir diese Fallunterscheidung durchführen müssen.

Beginnen wir mit $E < 0$. Dann können wir eine neue Basis der Liealgebra wählen, indem wir

$$\vec{a} = \sqrt{-\frac{K^2\mu}{2E}} \vec{A} \quad (1.14.10)$$

eingeführen. Das ergibt dann die einfachere Liealgebra

$$\{l_j, l_k\} = \epsilon_{jkl} l_l, \quad \{a_j, a_k\} = \epsilon_{jkl} l_l, \quad \{l_j, a_k\} = \epsilon_{jkl} a_l. \quad (1.14.11)$$

Wir wollen nun zeigen, daß dies die Algebra der $SO(4)$, also der Gruppe der Rotationen im \mathbb{R}^4 ist. Das sind diejenigen linearen Abbildungen, die das euklidische Skalarprodukt in vier Dimensionen ungeändert lassen und die Determinante 1 besitzen.

Eine reelle 4×4 -Matrix M ist offenbar genau dann eine $SO(4)$ -Matrix, wenn $M^t M = \mathbb{1}_4$ ist. Daraus leitet man wie im Fall der Drehgruppe her, daß eine „infinitesimal“ von der Einheitsmatrix verschiedene Matrix $M_{jk} = \delta_{jk} + \delta \alpha_{jk}$ genau dann eine $SO(4)$ -Matrix ist, wenn $\delta \alpha_{jk} = -\delta \alpha_{kj}$ ist. In der Tat besitzt also die $SO(4)$ insgesamt $(4 \cdot 4 - 4)/2 = 6$ linear unabhängige Erzeugende.

Die Drehgruppe $SO(3)$ in drei Dimensionen ist offenbar eine Untergruppe, die aus denjenigen $SO(4)$ -Matrizen besteht, die nur die ersten drei Komponenten eines jeden \mathbb{R}^4 -Vektors ändern. Dies entspricht den drei Erzeugenden $\delta \alpha_{jk} = -\epsilon_{ljk}$ mit $l \in \{1, 2, 3\}$. Identifiziert man also $(\hat{l}_l)_{jk} = -\epsilon_{ljk}$, wobei wir jedes \hat{l}_j als 3×3 -Matrix auffassen, entspricht dies der Drehimpulsalgebra für unsere Drehimpulskomponenten \vec{l} . In der Tat rechnet man nach, daß diese Matrizen die Kommutatorrelationen der Drehimpulsalgebra erfüllen:

$$[\hat{l}_j, \hat{l}_k] := \hat{l}_j \hat{l}_k - \hat{l}_k \hat{l}_j = \epsilon_{jkl} \hat{l}_l. \quad (1.14.12)$$

Wir erweitern diese 3×3 -Matrizen einfach in unserem Sinne zu 4×4 -Matrizen, daß wir die unterste Zeile und die vierte Spalte durch Nullen auffüllen. Im \mathbb{R}^4 agieren diese Matrizen dann für die ersten drei Komponenten als „infinitesimale Drehungen“ und lassen die letzte Komponente ungeändert.

Jetzt fehlen offensichtlich noch die Erzeugenden, die eine Drehung in der Ebene, gebildet aus dem vierten Einheitsvektor und einem der drei verbliebenen Einheitsvektoren generieren. Man rechnet nach, daß diese Erzeugenden \hat{a}_j genau die den Poisson-Klammern (1.14.11) entsprechenden Kommutatorrelationen ergeben. Dabei sind die Poisson-Klammern durch Kommutatoren und vice versa zu ersetzen. Wir haben also einen Isomorphismus von Liealgebren, wobei das Produkt der Lie-Algebra zum einen durch die Poisson-Klammern und zum anderen durch die Kommutatoren von Matrizen gegeben sind. Entsprechend sind die Basiselemente der Liealgebra die Drehimpulskomponenten und Runge-Lenz-Vektorkomponenten einerseits und die entsprechenden Matrizen \hat{l}_j und \hat{a}_k andererseits. Wir können also die Lie-Algebra, unabhängig von ihrer Wirkung auf vierdimensionale Vektoren mit der Lie-Algebra der Gruppe $SO(4)$ identifizieren. Man spricht daher auch allgemein von der $SO(4)$ -Symmetrie des Keplerproblems, obwohl dies nur die Symmetriegruppe für $E < 0$, also die gebundene Bewegung (Ellipsen- bzw. Kreisbahnen) ist.

Der Vollständigkeit halber wollen wir zeigen, daß die Liealgebren von $SO(4)$ und $SO(3) \otimes SO(3)$ isomorph sind. Dazu gehen wir erneut zu einer neuen Basis der Lie-Algebra über:

$$\alpha_j = \frac{l_j + a_j}{2}, \quad \beta_j = \frac{l_j - a_j}{2}. \quad (1.14.13)$$

Für diese gelten die Poissonklammerrelationen

$$\{\alpha_j, \alpha_k\} = \epsilon_{jkl} \alpha_l, \quad \{\alpha_j, \beta_k\} = 0, \quad \{\beta_j, \beta_k\} = \epsilon_{jkl} \beta_l. \quad (1.14.14)$$

Die Liealgebra zerfällt also in zwei $SO(3)$ -Teilalgebren, deren Elemente untereinander verschwindende Lieklammern besitzen, d.h. es handelt sich um eine direkte Summe der beiden $SO(3)$ -Algebren: $\mathfrak{so}(4) \cong \mathfrak{so}(3) \oplus \mathfrak{so}(3)$. Durch Exponentiation zeigt man, daß dies einem direkten Produkt der Gruppen entspricht.

Kommen wir nun zum Fall $E = 0$, also dem Fall der Parabelbahnen. Dann ist die Liealgebra wegen (1.14.8-1.14.9) durch

$$\{l_j, l_k\} = \epsilon_{jkl} l_l, \quad \{A_j, A_k\} = 0, \quad \{l_j, A_k\} = \epsilon_{jkl} A_l \quad (1.14.15)$$

gegeben. Das ist aber die gleiche Algebra wie zwischen Impuls und Drehimpuls, so daß wir sofort die Liealgebra der Isomorphiegruppe $ISO(3)$ des \mathbb{R}^3 , aufgefaßt als euklidischer affiner Punktraum, erkennen. Freilich handelt es sich hierbei nicht um den physikalischen Raum.

Schließlich betrachten wir noch $E > 0$, also den Fall der Hyperbelbahnen. Hier gehen wir genauso vor wie bei $E < 0$ mit dem Unterschied, daß wir jetzt

$$\vec{a} = \sqrt{\frac{K^2 \mu}{2E}} \vec{A} \quad (1.14.16)$$

definieren müssen. Die Liealgebra ist dann

$$\{l_j, l_k\} = \epsilon_{jkl} l_l, \quad \{a_j, a_k\} = -\epsilon_{jkl} l_l, \quad \{l_j, a_k\} = \epsilon_{jkl} a_l. \quad (1.14.17)$$

Dies ist die Liealgebra zur Gruppe $SO(1,3)^\dagger$, der Gruppe, die die Bilinearform

$$x, y \in \mathbb{R}^4 \mapsto \eta_{\mu\nu} x^\mu y^\nu \quad \text{mit} \quad (\eta_{\mu\nu}) = \text{diag}(1, -1, -1, -1) \quad (1.14.18)$$

invariant lassen. Die Indizes μ, ν laufen unserer Konvention gemäß von 0 bis 3. Die sich durch Exponentiation ergebende Gruppe läßt auch das Vorzeichen der 0-Komponenten von Vektoren ungeändert. Die entsprechende Gruppe ist die **eigentlich orthochrone Lorentz-Gruppe**. Diese Struktur auf dem \mathbb{R}^4 hat eine sehr wichtige Bedeutung, legt sie doch die Raumzeit-Struktur in der **Speziellen Relativitätstheorie** fest, worauf wir im nächsten Kapitel ausführlich eingehen werden. Der Beweis dieser Behauptung erfolgt genau analog zum Fall $E < 0$ und sei dem Leser zur Übung überlassen. Die Drehungen, erzeugt durch die Drehimpulskomponenten \vec{l} , läßt man dabei auf die unteren drei Komponenten wirken. Klarerweise bilden diese eine Untergruppe der $SO(1,3)^\dagger$. Die drei Operationen, die die 0-Komponente und je eine der drei unteren Komponenten der Vierervektoren ändern, bilden dann die übrigen Basiselemente der $so(1,3)$ -Algebra. Von dieser Matrixalgebra kann man wie im Fall der $so(4)$ zeigen, daß sie im Sinne von Liealgebren isomorph zu der Poissonklammeralgebra (1.14.17) ist.

Es sei noch betont, daß der wahre Nutzen dieser algebraischen Betrachtungen erst in der Quantentheorie des Wasserstoffatoms voll zutage tritt, weil sie die Entartung der Energieniveaus zu analysieren helfen.

1.15 Ausblick auf die Quantentheorie

Kapitel 2

Spezielle Relativitätstheorie

In der Physik um 1900 herum war eine der interessantesten ungeklärten Fragen die nach dem Nachweis des sogenannten **Äthers**. Dies war eine „Substanz“, dessen Schwingungen die elektromagnetischen Wellen, die von James Clerk Maxwell im Rahmen seiner Theorie der **Elektrodynamik** vorhergesagt (1865) und durch Heinrich Hertz entdeckt worden waren (1886), fortleiten sollte, ähnlich wie die Schwingungen der Luft Schallwellen entsprechen. Allerdings hatte diese Substanz sehr ungewöhnlich theoretische Eigenschaften und konnte zum anderen auch nicht experimentell nachgewiesen werden.

So unternahmen Michelson und Morley mit einem Interferometer ein Experiment zur Messung der Bewegung der Erde durch den Äther, demzufolge sich aufgrund der unterschiedlichen Lichtgeschwindigkeit in Richtung der Bewegung und senkrecht dazu ein Effekt ergeben sollte, wenn das Interferometer um 90° gedreht wird. Ein solcher Effekt konnte aber nicht nachgewiesen werden.

In den Maxwell-Gleichungen des elektromagnetischen Feldes ergab sich in einem Inertialsystem, in dem der Äther ruhen sollte, eine bestimmte Phasengeschwindigkeit für die elektromagnetischen Wellen im Vakuum, die Lichtgeschwindigkeit c . Da aber die Maxwell-Gleichungen nicht invariant unter Galilei-Boosts sind, sollte sich eine Änderung der Lichtgeschwindigkeit für ein Bezugssystem ergeben, das sich gegenüber dem Ruhssystem des Äthers bewegt.

Die Lösung dieses Widerspruchs, die endgültig Einstein 1905 gelang, ist, daß die der Newtonschen Mechanik zugrundeliegende mathematische Beschreibung von Raum und Zeit nur näherungsweise Gültigkeit besitzen und daher die Transformationsformeln zwischen Koordinaten (und der Zeit!) unterschiedlicher inertialer Bezugssysteme modifiziert werden müssen. Während dies in der Hauptvorlesung ausführlich physikalisch motiviert wird, beschränken wir uns hier im wesentlichen auf die Mathematik des der speziellen Relativitätstheorie zugrundeliegenden Raum-Zeit-Modells, des sog. **Minkowski|-Raumes**.

2.1 Das speziell-relativistische Raum-Zeit-Modell

In seiner berühmten Arbeit „Zur Elektrodynamik bewegter Körper“ [Ein05] ging Einstein von zwei Postulaten aus:

1. Es existiert ein Inertialsystem, in dem ein Körper in Ruhe verharrt oder sich gleichförmig geradlinig (also mit konstanter Geschwindigkeit) bewegt, solange keine Kräfte auf ihn einwirken.
2. Die Ausbreitungsgeschwindigkeit **elektromagnetischer Wellen** (insbesondere des **Lichtes**) im Vakuum ist unabhängig von der Geschwindigkeit der Lichtquelle und des Empfängers (Beobachters) dieser Wellen.

Das erste Postulat entspricht genau dem **Galileischen Trägheitsgesetz**, während das zweite der üblichen Realisierung dieses Postulats in der Newtonschen Mechanik widerspricht, denn es sollte unmöglich sein, die

2. Spezielle Relativitätstheorie

Absolutgeschwindigkeit eines Bezugssystems durch irgendwelche physikalischen Vorgänge zu bestimmen, da die physikalischen Grundgesetze in jedem Inertialsystem gleich aussehen, und jedes gegen ein Inertialsystem gleichförmig geradlinig bewegte Bezugssystem ist notwendig wieder ein Inertialsystem, denn gemäß dem Trägheitsgesetz sind ruhende oder gleichförmig geradlinig bewegte Körper gleichberechtigt. Anders ausgedrückt kann man Geschwindigkeiten nur in Bezug auf einen Referenzkörper definieren.

Wir können nun mit diesen Postulaten die speziell-relativistische Raum-Zeit-Struktur finden. Dazu denken wir uns eine punktförmige Lichtquelle, die in einem Inertialsystem A ruht, von der aus ein kurzer Lichtimpuls zur Zeit $t = 0$ ausgesandt wird. Da die Lichtgeschwindigkeit c ist, befindet sich das Lichtsignal zur Zeit t auf der Kugeloberfläche mit dem Radius ct , d.h. für diese Kugeloberfläche \vec{x} gilt

$$c^2 t^2 - \vec{x}^2 = 0. \quad (2.1.1)$$

Wir betrachten nun einen Beobachter B , der sich relativ zu diesem Bezugssystem mit der konstanten Geschwindigkeit $\vec{v} = c\vec{\beta}$ relativ zu A bewegt. Dies definiert ein anderes Inertialsystem. Dessen Zeit \bar{t} und räumliche Koordinaten $\vec{\bar{x}}$ seien so gewählt, daß zur Zeit $t = 0$ auch $\bar{t} = 0$ ist und daß die Koordinatenursprünge zu dieser Zeit zusammenfallen. Gemäß dem zweiten Postulat muß nun auch für B bzgl. seiner Raum-Zeit-Koordinaten (2.1.1) gelten, d.h. auch er beobachtet eine Kugelwelle, die von seinem Koordinatenursprung zur Zeit $\bar{t} = 0$ ausgesandt wurde, d.h. für die Position des Wellenpakets gilt auch für ihn

$$c^2 \bar{t}^2 - \vec{\bar{x}}^2 = 0. \quad (2.1.2)$$

Wegen des ersten Postulats muß nun die Transformation zwischen den Zeiten und der räumlichen Koordinaten bzgl. zweier Inertialsystem linear sein, denn jede beliebige gleichförmig geradlinige Bewegung in dem einen Inertialsystem muß ebenso eine gleichförmig geradlinige Bewegung im anderen Inertialsystem sein.

Um die entsprechenden linearen Transformationen zu finden, ist es am bequemsten, Minkowski [Min09] zu folgen zunächst **vierdimensionale** Vektoren $(x^\mu) = (ct, x^1, x^2, x^3)$ einzuführen (mit $x^0 = ct$). Die Basisvektoren bezeichnen wir mit e_μ (man beachte hierbei die Indexstellung: obere Indizes für die Komponenten und untere Indizes für die Basisvektoren). Dann führen wir eine Bilinearform, das **Minkowski-Produkt**

$$\mathbf{x} \cdot \mathbf{y} = x^0 y^0 - \vec{x} \cdot \vec{y} \quad (2.1.3)$$

ein, so daß dann die Gleichung (2.1.1) in der Form

$$\mathbf{x} \cdot \mathbf{x} =: \mathbf{x}^2 = 0 \quad (2.1.4)$$

geschrieben werden kann. Offenbar gilt für die hierbei eingeführten „pseudo-kartesischen Basisvektoren“

$$\mathbf{e}_\mu \cdot \mathbf{e}_\nu = \eta_{\mu\nu} := \begin{cases} 1 & \text{for } \mu = \nu = 0, \\ -1 & \text{for } \mu = \nu, \quad \mu, \nu \in \{1, 2, 3\}, \\ 0 & \text{for } \mu \neq \nu, \quad \mu, \nu \in \{0, 1, 2, 3\}. \end{cases} \quad (2.1.5)$$

Damit können wir nämlich das **Minkowski-Produkt** in der Form

$$\mathbf{x} \cdot \mathbf{y} = x^\mu y^\nu \mathbf{e}_\mu \cdot \mathbf{e}_\nu = \eta_{\mu\nu} x^\mu y^\nu \quad (2.1.6)$$

schreiben. Dann folgt (2.1.2) aus (2.1.1), wenn wir annehmen, daß auch B solche pseudo-kartesischen Basisvektoren $\bar{\mathbf{e}}_\mu$ benutzt wie A . Das bedeutet dann, daß die lineare Transformation zwischen den Raumzeit-Koordinaten der beiden Systeme einfach einer Basistransformation mit einer Matrix $(\Lambda^\mu_\nu) \in \mathbb{R}^{4 \times 4}$ entspricht, so daß

$$\mathbf{e}_\nu = \bar{\mathbf{e}}_\mu \Lambda^\mu_\nu \quad (2.1.7)$$

gilt. Da nun sowohl \bar{e}_μ als auch e_ν pseudo-kartesisch im Sinne von (2.1.5) sein sollen, folgt für die Transformationsmatrix

$$\eta_{\rho\sigma} = e_\rho \cdot e_\sigma = \Lambda^\mu{}_\rho \bar{e}_\mu \cdot \Lambda^\nu{}_\sigma \bar{e}_\nu = \Lambda^\mu{}_\rho \Lambda^\nu{}_\sigma \eta_{\mu\nu}. \quad (2.1.8)$$

Dies ist sehr ähnlich zur Definition einer orthogonalen Matrix in einem vierdimensionalen Euklidischen Vektorraum. Allerdings weist das Minkowski-Produkt den entscheidenden Unterschied auf, daß die räumlichen Koordinaten mit einem $-$ -Zeichen in das Produkt eingehen. Das Minkowskiprodukt ist also im Gegensatz zu dem üblichen **Skalarprodukt** eines Euklidischen Vektorraums nicht positiv definit, sondern $x \cdot x = x^2$ kann sowohl positiv als auch negativ sein, und aus $x \cdot x = 0$ folgt nicht notwendig, daß $x = 0$ ist.

Andererseits ist aber die Matrix $(\eta_{\mu\nu}) = \text{diag}(1, -1, -1, -1)$ invertierbar, und wir können wie in einem Euklidischen Raum mit Hilfe dieser Matrix und ihrer Inversen $(\eta^{\mu\nu}) = \text{diag}(1, -1, -1, -1)$ (*nachrechnen!*) statt der kovarianten auch kontravariante Basisvektoren und Komponenten von Vektoren einführen, z.B. $e^\mu = \eta^{\mu\nu} e_\nu$ bzw. $x_\mu = g_{\mu\nu} x^\nu$, so daß $x = x_\mu e^\mu = x^\mu e_\mu = \eta_{\mu\nu} x^\mu e^\nu$ sowie für zwei Vektoren x und y das Minkowski-Produkt $x \cdot y = \eta_{\mu\nu} x^\mu y^\nu = x^\mu y_\mu = x_\nu y^\nu = \eta^{\mu\nu} x_\mu y_\nu$ schreiben kann.

Wir betrachten nun die Transformation zwischen zwei Inertialsystem. Es möge sich das System des Beobachters B mit den kovarianten Basisvektoren \bar{e}_μ mit der Geschwindigkeit $\vec{v} = \beta c \vec{e}_1$ relativ zum System des Beobachters A mit den kovarianten Basisvektoren \vec{e}_ν bewegen, und der Zeitnullpunkt sei so gewählt, daß für $t = \bar{t} = 0$ die Ursprünge der Bezugssystem zusammenfallen. Es ist klar, daß in diesem Fall $\bar{x}^2 = x^2$ und $\bar{x}^3 = x^3$ ist und wir daher nur die Transformation in der (12)-Ebene zu betrachten brauchen.

Wir betrachten nun die Bewegung der inertialen Beobachter A und B vom Bezugssystem von A aus. Die Beobachter denken wir uns dazu im Ursprung der jeweiligen räumlichen Koordinatensysteme lokalisiert. Wir tragen nun die $x^0 = ct$ -Achse von A senkrecht nach oben und die x^1 -Achse nach rechts auf (vgl. Abb. 2.1). Wir können dann Bewegungen punktförmiger Objekte als Linien in diesem **Minkowski-Diagramm** betrachten, die sog. **Weltlinien**. Die Weltlinie des Beobachters A ist offensichtlich durch $x_A(t) = ct e_0$ gegeben, denn die räumlichen Koordinaten ändern sich nicht, da der Beobachter A in seinem Bezugssystem ruht. Der Beobachter B bewegt sich hingegen relativ zu A , und A wird in seinen Raum-Zeit-Koordinaten dessen Weltlinie durch

$$x_B = ct(e_0 + \beta e_1) \quad (2.1.9)$$

beschreiben. Andererseits muß B seine eigene Weltlinie durch

$$x_B = c\bar{t}\bar{e}_0 \quad (2.1.10)$$

beschreiben. Die Zeitachse \bar{e}_0 muß nun aber $\bar{e}_0 \cdot \bar{e}_0 = 1$ erfüllen. Wir müssen also den Tangentenvektor $dx_B/dt = c(e_0 + \beta e_1)$ im Sinne des Minkowskiprodukts normieren können. Wegen $dx_B/dt \cdot dx_B/dt = c^2(1 - \beta^2)$, ist dies nur möglich, wenn $|\beta| < 1$ ist, d.h. B kann nur dann in einem Inertialsystem sein, wenn seine Geschwindigkeit $v = \beta c$ die Bedingung $|v| < c$ erfüllt, d.h. c stellt eine absolute **Grenzgeschwindigkeit** für die Relativgeschwindigkeit zwischen inertialen Beobachtern dar¹ Es ergibt sich dann also

$$\bar{e}_0 = \gamma(e_0 + \beta e_1) \quad \text{mit} \quad \gamma = \frac{1}{\sqrt{1 - \beta^2}}. \quad (2.1.11)$$

Wir können nun auch den räumlichen Basisvektor in der (12)-Ebene definieren. Für ihn muß gelten

$$\bar{e}_1 \cdot \bar{e}_0 = 0, \quad \bar{e}_1 \cdot \bar{e}_1 = -1 \Rightarrow \bar{e}_1 = \gamma(\beta e_0 + e_1). \quad (2.1.12)$$

Es ist klar, daß dies gerade der an der Winkelhalbierenden zwischen Raum- und Zeitachse von A gespiegelte Vektor \bar{e}_0 . Die Winkelhalbierende wird in den Koordinaten von A aber durch $x_L = ct(e_0 + e_1)$ beschrieben.

¹Dabei ist die Relativgeschwindigkeit als die Geschwindigkeit des einen inertialen Beobachters im Ruhssystem des anderen inertialen Beobachters zu verstehen.

2. Spezielle Relativitätstheorie

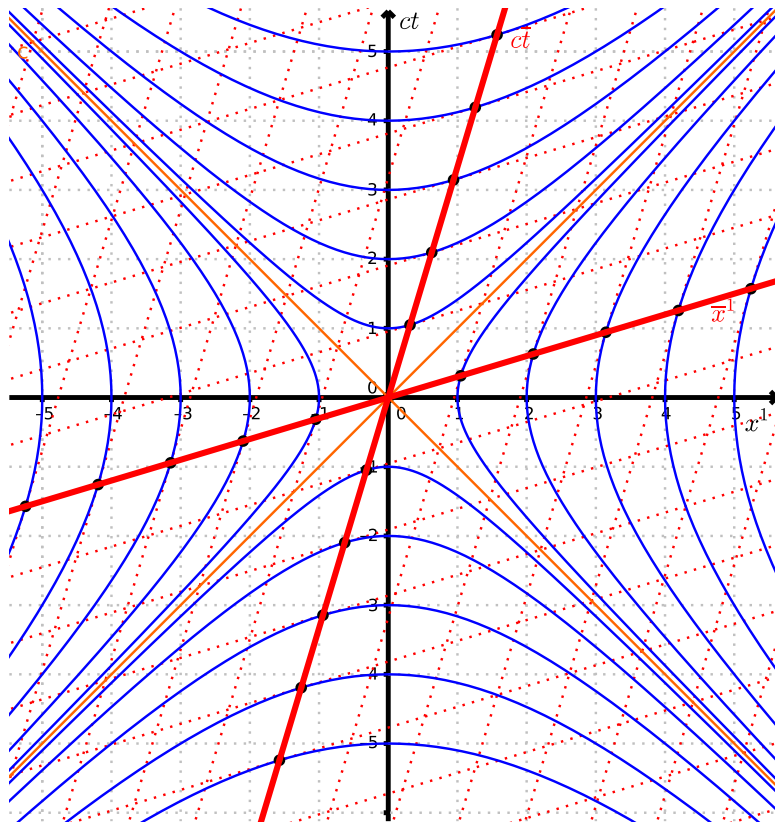


Abbildung 2.1: Minkowski-Diagramm, das die Konstruktion von B's Raum-Zeit-Achsen illustriert: Die dicken grünen Linien sind B's Zeit- und Raum-Achse. Die Einheitslängen werden auf den Achsen durch die Schnittpunkte mit den blauen Hyperbeln $(ct)^2 - (x^1)^2 = \pm n^2$ mit $n \in \{1, 2, \dots\}$ bestimmt. Die orangefarbenen Linien $ct = \pm x^1$ definieren die Ausbreitung von Lichtsignalen in x^1 - bzw. \bar{x}^1 -Richtung. Sie sind die Winkelhalbierende zwischen den Zeit- und Raum-Achsen sowohl für A als auch für B, d.h. für beide Beobachter bewegt sich die Front eines Lichtsignals mit derselben Geschwindigkeit c , wie es vom 2. Postulat Einsteins verlangt wird.

Für diesen Vektor gilt $x_L \cdot x_L = 0$, und das bedeutet, daß dieser Vektor die Ausbreitung eines Lichtsignals beschreibt. Dies ist aber wegen (2.1.12) auch die Winkelhalbierende zwischen B's zeit- und raumartigem Basisvektor (s. auch das Minkowski-Diagramm in Abb. 2.1). Das bedeutet, daß in der Tat durch unsere Rechnung das Postulat 2 erfüllt ist, wonach die Lichtgeschwindigkeit des Lichtsignals unabhängig von der Bewegung des Beobachters und der Quelle ist.

Um die Transformationsmatrix Λ^μ_ν zu bestimmen, benötigen wir gemäß (2.1.7) die umgekehrte Form, d.h. wir müssen (2.1.11) und (2.1.12) nach den Basisvektoren von A umrechnen (*Übung*):

$$e_0 = \gamma(\bar{e}_0 - \beta\bar{e}_1), \quad e_1 = \gamma(-\beta\bar{e}_0 + \bar{e}_1), \quad e_2 = \bar{e}_2, \quad e_3 = \bar{e}_3. \quad (2.1.13)$$

Wegen (2.1.7) lesen wir aus (2.1.13) für die Transformationsmatrix für einen **Lorentz-Boost in x^1 -Richtung**

$$\hat{\Lambda} = (\Lambda^\mu_\nu) = \begin{pmatrix} \gamma & -\beta\gamma & 0 & 0 \\ -\beta\gamma & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \quad (2.1.14)$$

ab.

Nun ist für jeden Vierervektor x

$$x = x^\nu e_\nu = x^\nu \Lambda^\mu_\nu \bar{e}_\mu = \bar{x}^\mu \bar{e}_\mu \Rightarrow \bar{x}^\mu = \Lambda^\mu_\nu x^\nu. \quad (2.1.15)$$

Daraus ergibt sich nun für die Umrechnung zwischen den Komponenten bzgl. A 's und B 's Bezugssystem

$$\bar{x}^0 = c\bar{t} = \gamma(x^0 - \beta x^1), \quad \bar{x}^1 = \gamma(x^1 - \beta x^0), \quad \bar{x}^2 = x^2, \quad \bar{x}^3 = x^3. \quad (2.1.16)$$

Dieser Lorentz-Boost und ersetzt die entsprechende Formel für einen **Galilei-Boost** in der Newtonschen Physik. Entwickelt man diese Formeln für kleine $|\beta| = |v|/c$, erhält man in der Tat

$$\bar{t} = \gamma(t - vx^1/c^2) \simeq t + \mathcal{O}(x^1 v/c^2), \quad \bar{x}^1 \simeq x^1 - vt + \mathcal{O}(\beta^2), \quad (2.1.17)$$

d.h. für kleine Relativgeschwindigkeiten zwischen den Beobachtern A und B erhält man näherungsweise wieder den **Galilei-Boost** der Newtonschen Mechanik.

Der gegenüber der Newtonschen Raum-Zeit augenfälligste Unterschied ist, daß sich hier nicht nur die räumlichen Koordinaten transformieren sondern auch die Zeit. Dies wird durch Postulat 2 erzwungen: Damit die Lichtgeschwindigkeit für beide Beobachter dieselbe sein kann, muß sich auch die Zeit mittransformieren. Dies hat einige ungewohnte Konsequenzen bzgl. des zeitlichen Ablaufs und der Messung von Längen für Beobachter in verschiedenen gegeneinander gleichförmig geradlinig bewegten Inertialsystemen.

Wir benötigen dazu nur die oben hergeleiteten Beziehungen (2.1.11-2.1.18) zwischen den Einheitsvektoren. Betrachten wir als erstes den Fall, daß zwei durch Vierervektoren gekennzeichnete x und y Ereignisse für A am gleichen Ort stattfinden (z.B. Lichtblitze). Es gilt also $x = ct_x e_0 + x^j e_j$ ² und $y = ct_y e_0 + x^j e_j$. Offenbar vergehen für A und B dann verschiedene Zeiten zwischen diesen Ereignissen, denn es gilt wegen (2.1.11)

$$\Delta t_A = \frac{1}{c} e_0 \cdot (y - x) = t_y - t_x, \quad \Delta t_B = \frac{1}{c} \bar{e}_0 \cdot (y - x) = \gamma \Delta t_{\text{Alice}}, \quad (2.1.18)$$

d.h. bzgl. Bob vergeht eine um den **Lorentz-Faktor** $\gamma = 1/\sqrt{1-\beta^2}$ längere Zeit. Dies nennt man die **Zeitdilatation**. Natürlich finden die Ereignisse für B nicht am gleichen Ort ab, denn er bewegt sich ja gegenüber A , so daß er sich in dem Zeitintervall Δt_B weiterbewegt. Die entsprechende Distanz muß natürlich $\beta c \Delta t_B$ sein, denn B bewegt sich relativ zu A mit der Geschwindigkeit βc . Formal finden wir dies unter Verwendung von (2.1.13) bestätigt:

$$(\bar{y} - \bar{x})^1 = -\bar{e}_1 \cdot (y - x) = -\beta c \gamma \Delta t_A = -\beta c \Delta t_B. \quad (2.1.19)$$

Als nächstes betrachten wir eine Stange entlang der x^1 -Richtung, die relativ zu A ruht. Bzgl. A 's Koordinaten sind die Weltlinien der Endpunkte der Stange durch

$$(x^\mu) = \begin{pmatrix} \lambda_1 \\ x \end{pmatrix}, \quad (y^\mu) = \begin{pmatrix} \lambda_2 \\ y \end{pmatrix}, \quad \lambda_1, \lambda_2 \in \mathbb{R} \quad (2.1.20)$$

gegeben. Dabei sind λ_1 und λ_2 beliebige Parameter dieser Weltlinie. Beobachter A mißt die Länge der Stange, indem er die Koordinaten zur gleichen Zeit t_A abliest, d.h. für $\lambda_1 = \lambda_2$. Die Relativkoordinaten der Enden für diese „Meßereignisse“ der Stange und die entsprechende Länge sind dann

$$((y - x)^\mu) = \begin{pmatrix} 0 \\ y - x \end{pmatrix} \Rightarrow L_A = |y^1 - x^1|. \quad (2.1.21)$$

Für B bewegt sich die Stange in negative \bar{x}^1 -Richtung. Er wird die Länge der Stange aber ebenfalls bestimmen indem er zu einem Zeitpunkt bzgl. seiner (!) Zeit die räumlichen Koordinaten der Stangenenden abliest. Für die Zeitkomponente des Differenzvektors der Stangenenden bzgl. B gilt aber

$$\begin{aligned} \Delta t_B &= \bar{e}_B \cdot (y - x) = \gamma[(\lambda_2 - \lambda_1) - \beta L_A] \stackrel{!}{=} 0 \\ &\Rightarrow \lambda_2 - \lambda_1 = \beta L_A. \end{aligned} \quad (2.1.22)$$

²Hier und im folgenden laufen lateinische Indizes stets von 1 bis 3, griechische von 0 bis 3.

2. Spezielle Relativitätstheorie

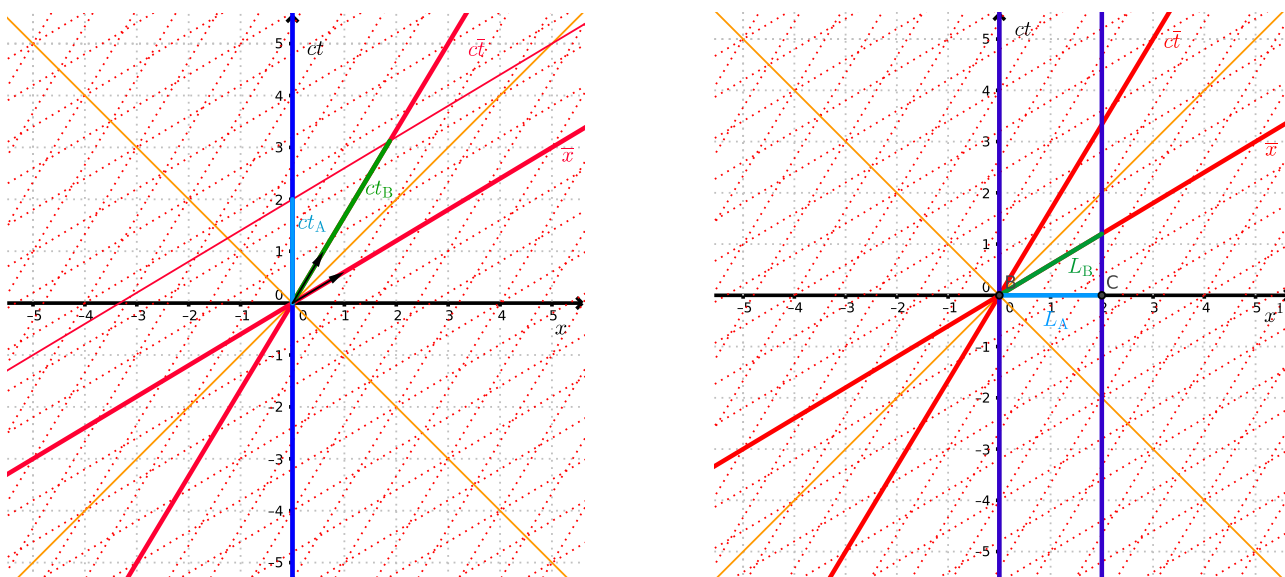


Abbildung 2.2: **Links:** Minkowski-Diagramm zur Zeitdilatation. Die Zeit, die zwischen den beiden bzgl. A am gleichen Ort $x^1 = 0$ vergeht, erscheint für B länger als für A. **Rechts:** Minkowski-Diagramm zur Längenkontraktion: Die blauen Linien sind die Weltlinien der Endpunkte der Stange, die relativ zu A ruht. In B's Bezugssystem bewegt sich die Stange, und B mißt dessen Länge, indem er seine Koordinaten der Endpunkte zur für ihn gleichen Zeit $\bar{t} = 0$ markiert. Entsprechend mißt er eine kürzere Länge als A. Dabei muß man beachten, daß die Längen für A und B im Minkowski-Diagramm nicht durch die üblichen euklidischen Abstände in der Zeichenebene gegeben sind sondern durch die pseudo-euklidische Geometrie im Minkowski-Raum, wie in Abb. 2.1 konstruiert, also in den hier gezeigten Diagrammen durch die gepunkteten Koordinatenlinien. Offensichtlich besitzt für A die Stange eine Länge von 2 Längeneinheiten und für B eine kürzere Länge. Wie die Rechnung zeigt, ist die Länge gerade um den inversen Lorentz-Faktor $1/\gamma = \sqrt{1 - \beta^2}$ verkürzt.

D.h. wir bestimmen die Weltlinienparameter der Enden durch die Forderung, daß B die räumlichen Koordinaten der Stangenenden bzgl. seiner Zeit gleichzeitig abliest. Dies ergibt schließlich die durch B gemessene Länge

$$L_B = -e_1 \cdot (\mathbf{y} - \mathbf{x}) = \gamma L_A (1 - \beta^2) = \frac{L_A}{\gamma} = \sqrt{1 - \beta^2} L_A. \quad (2.1.23)$$

Dieses kinematische Phänomen bezeichnet man als **Längenkontraktion**, denn B mißt gemäß (2.1.23) eine um den inversen Lorentz-Faktor $1/\gamma = \sqrt{1 - \beta^2}$ verkürzte Länge vgl. mit der Länge der Stange in ihrem Ruhssystem.

Wir sehen an diesem Beispiel auch, daß zwei bzgl. A gleichzeitige Ereignisse (wie hier die Ablesung der Stangenenden durch A), die an verschiedenen Orten stattfinden, für den gegen A bewegten Beobachter B nicht gleichzeitig sind. Umgekehrt sind die Ereignisse des gleichzeitigen Ablesens der räumlichen Koordinaten der Stangenenden durch B von A aus gesehen ebenfalls nicht gleichzeitig. Man bezeichnet dieses Phänomen auch als **Relativität der Gleichzeitigkeit**.

Wir können nun auch die Messung der Länge einer Stange für beliebige Orientierung derselben relativ zur Boost-Geschwindigkeit $\vec{v} = \beta c \mathbf{e}_1$ von B bzgl. A betrachten. Die Rechnung ist vollständig analog wie zuvor. Die Stange ruhe also in A's Bezugssystem, d.h. die Weltlinien der beiden Enden seien durch

$$(x^\mu) = \begin{pmatrix} \lambda_1 \\ \vec{x} \end{pmatrix}, \quad (y^\mu) = \begin{pmatrix} \lambda_2 \\ \vec{y} \end{pmatrix} \quad (2.1.24)$$

gegeben. A mißt wieder die Länge, indem sie die räumlichen Koordinaten zur gleichen Zeit, also für $\lambda_1 = \lambda_2$

abliest, d.h.

$$L_A^2 = (\vec{y} - \vec{x}) \cdot (\vec{y} - \vec{x}). \quad (2.1.25)$$

Für B 's Messung müssen wir λ_1 und λ_2 wieder so bestimmen, daß er die räumlichen Koordinaten der Enden der Stange zeitgleich bzgl. seines Referenzsystems abliest, d.h.

$$\bar{e}_0 \cdot (\underline{y} - \underline{x}) = \gamma[(\lambda_2 - \lambda_1) - \beta(y^1 - x^1)] \stackrel{!}{=} 0 \Rightarrow (\lambda_2 - \lambda_1) = \beta(y^1 - x^1), \quad (2.1.26)$$

und daraus ergeben sich die entsprechend gleichzeitig abgelesenen räumlichen Koordinaten zu

$$\begin{aligned} \bar{y}^1 - \bar{x}^1 &= -\bar{e}_1 \cdot (\underline{y} - \underline{x}) = \gamma[(y^1 - x^1) - \beta(\lambda_2 - \lambda_1)] = \frac{y^1 - x^1}{\gamma}, \\ \bar{y}^2 - \bar{x}^2 &= -\bar{e}_2 \cdot (\underline{y} - \underline{x}) = -e_2 \cdot (\underline{y} - \underline{x}) = y^2 - x^2, \\ \bar{y}^3 - \bar{x}^3 &= -\bar{e}_3 \cdot (\underline{y} - \underline{x}) = -e_3 \cdot (\underline{y} - \underline{x}) = y^3 - x^3. \end{aligned} \quad (2.1.27)$$

Daraus folgt für die durch B gemessene Länge

$$L_B^2 = \frac{(y^1 - x^1)^2}{\gamma^2} + (y^2 - x^2)^2 + (y^3 - x^3)^2. \quad (2.1.28)$$

Dies und (2.1.27) bedeutet, daß nur die Komponente der Länge der Stange in Richtung der Boost-Geschwindigkeit längenkontrahiert ist. Für die Komponenten in senkrechter Richtung ändern sich die scheinbaren Werte nicht.

Wir betrachten nun schon die oben angedeutete Bemerkung, daß die Gleichzeitigkeit von Ereignissen vom Bezugssystem abhängt. Nehmen wir an, zwei Ereignisse finden bzgl. A gleichzeitig statt, d.h. es gilt

$$(x^\mu) = \begin{pmatrix} ct \\ \vec{x} \end{pmatrix}, \quad (y^\mu) = \begin{pmatrix} ct \\ \vec{y} \end{pmatrix}. \quad (2.1.29)$$

Bzgl. B finden wir dann, daß diese Ereignisse nicht gleichzeitig stattfinden sondern um die Zeit

$$\Delta \bar{t} = \frac{1}{c} \bar{e}_0 \cdot (\underline{y} - \underline{x}) = \frac{\gamma \beta}{c} (y^1 - x^1) \quad (2.1.30)$$

versetzt.

Wir betrachten noch ein berühmtes Gedankenexperiment von Einstein im Zusammenhang mit der Relativität der Gleichzeitigkeit. Wir betrachten dazu einen Zug, der sich mit der konstanten Geschwindigkeit $\vec{v} = \beta c \vec{e}_1$ bewegt. In diesem Zug befinde sich Alice in der Mitte und sende von dort aus zur Zeit $t = 0$ ein Lichtsignal zu den beiden Enden des Zuges (s. Abb. 2.3). Die Weltlinien der beiden Lichtsignale bzgl. Alices Raumzeit-Koordinaten (parametrisiert mit Alices Zeit t) lauten

$$(x_\pm^\mu) = \begin{pmatrix} ct \\ \pm ct \end{pmatrix}, \quad (2.1.31)$$

wobei wir wieder nur die relevanten Raumzeit-Koordinaten (x^0, x^1) aufschreiben. Besitzt der Zug die Länge L , erreicht demnach das Lichtsignal bzgl. Alice *gleichzeitig* zur Zeit $t = L/(2c)$ an den beiden Enden.

Betrachten wir nun die Weltlinie bzgl. Bobs Bezugssystem. Dazu müssen wir den entsprechenden Lorentz-Boost anwenden. Da jetzt Alice relativ zu Bob mit $\vec{v} = \beta c \vec{e}_1$ bewegt ist, lautet die Transformation jetzt

$$(\bar{x}_\pm^\mu) = \begin{pmatrix} \gamma & \beta\gamma \\ \beta\gamma & \gamma \end{pmatrix} \begin{pmatrix} ct \\ \pm ct \end{pmatrix} = \begin{pmatrix} c(1 \pm \beta)t \\ c(1 \pm \beta)t \end{pmatrix}. \quad (2.1.32)$$

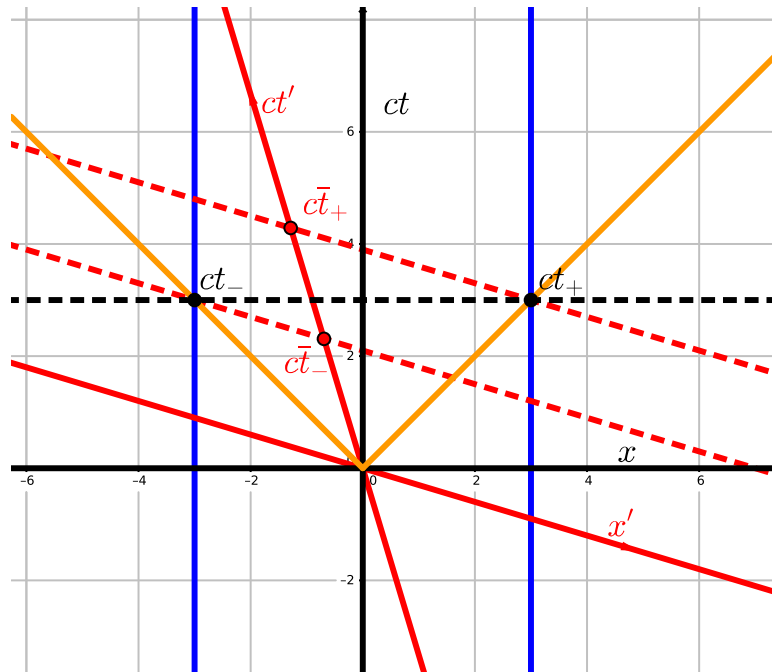


Abbildung 2.3: Zum Einsteinschen Zug-Gedankenexperiment. Erklärung s. Text.

Die beiden Lichtsignale erreichen also von Bob aus betrachtet die Enden des Zuges zu den Zeiten

$$\bar{t}_{\pm} = \frac{(1 \pm \beta)L}{2c}. \quad (2.1.33)$$

Dies erklärt sich dadurch, daß sich von Bob aus betrachtet das Licht zum rechten (linken) Ende einen längeren (kürzeren) Weg zurücklegen muß als bzgl. Alice, weil sich ja der Zug mit der entsprechenden Geschwindigkeit $v = \beta c$ in Ausbreitungsrichtung des Lichtsignals bewegt, dabei aber bzgl. Bob die gleiche Lichtgeschwindigkeit c besitzt.

2.2 Das Zwillingsparadoxon

Das Zwillingsparadoxon ist ein Beispiel für die ungewohnte Relativität der Zeitdauer zwischen Ereignissen in verschiedenen Bezugssystemen, also der Zeitdilatation. Wir nehmen an, daß der Zwilling Bob mit großer Geschwindigkeit auf einer Kreisbahn bzgl. Alices Bezugssystem reist. Bobs Weltlinie wird also in Alices Bezugssystem durch

$$[x^{\mu}(t)] = \begin{pmatrix} ct \\ R[\cos(\omega t) - 1] \\ R \sin \omega t \\ 0 \end{pmatrix} \quad (2.2.1)$$

beschrieben. Wir können nun zu jedem Zeitpunkt durch einen geeigneten momentanen Lorentz-Boost in ein momentanes Ruh-System von Bob transformieren. Vergeht in Alices Bezugssystem eine infinitesimale Zeit dt' , so vergeht in Bobs Bezugssystem nur dessen Eigenzeit $d\tau = dt' \sqrt{\eta_{\mu\nu} \dot{x}^{\mu} \dot{x}^{\nu}}/c$. Dies wird dadurch klar, daß die Eigenzeit in Bobs momentanen Ruhsystem definiert ist, wo $(d\bar{x}^{\mu}) = (cd\bar{t}, 0, 0, 0) = (cd\tau, 0, 0, 0)$ gilt. Damit ist aber $c^2 d\tau^2 = \eta_{\mu\nu} d\bar{x}^{\mu} d\bar{x}^{\nu} = \eta_{\mu\nu} dx^{\mu} dx^{\nu}$, denn ein Lorentz-Boost zwischen Alices und Bobs Bezugssystem läßt ja das Minkowski-Linienelement invariant. Damit ergibt sich Bobs **Eigenzeit**, also die Zeit, die für

Bob, gemessen als die Summe über die infinitesimalen Zeitinkremente im jeweils momentanen Ruhssystem von Bob vergeht, zu

$$\tau = \frac{1}{c} \int_0^t dt' \sqrt{\frac{dx^\mu(t')}{dt'} \frac{dx^\mu(t')}{dt'}} = \int_0^t dt' \sqrt{1 - \frac{\omega^2 R^2}{c^2}} = t \sqrt{1 - \frac{\omega^2 R^2}{c^2}}. \quad (2.2.2)$$

Das bedeutet aber, daß Bob im Sinne von Alices Zeit weniger schnell altert, d.h. wenn Bob und Alice sich nach einer in Alices System gemessenen Umlaufzeit $T = 2\pi/\omega$ wieder begegnen, ist Bob nur um $\tau = 2\pi/\omega \sqrt{1 - \omega^2 R^2/c^2}$, also um den inversen Lorentz-Faktor weniger, gealtert als Alice.

Dieser Effekt wurde mehrfach für die Lebensdauer von instabilen Teilchen in Speicherringen überprüft³.

2.3 Allgemeine Lorentz-Transformationen

Wir betrachten nun die Lorentz-Transformationen etwas genauer. Bisher haben wir den **Lorentz-Boost** in x^1 -Richtung betrachtet, d.h. die Transformation von den Koordinaten x^μ für ein Ereignis im Inertialsystem des Beobachters A zu denen eines beliebigen anderen Beobachters B , der sich gegenüber A mit der konstanten Geschwindigkeit βc in x^1 -Richtung bewegt. Die Zeitzählung erfolge in beiden Bezugssystemen so, daß die Ursprünge der räumlichen Koordinaten für beide Beobachter zur Zeit $t = \bar{t} = 0$ zusammenfallen. Die Koordinaten von B 's pseudo-kartesischer Basis bzgl. A 's Basis lauten dann gemäß (2.1.11) und (2.1.12)

$$(\bar{e}_0^\mu) = \gamma \begin{pmatrix} 1 \\ \beta \\ 0 \\ 0 \end{pmatrix}, \quad (\bar{e}_1^\mu) = \gamma \begin{pmatrix} \beta \\ 1 \\ 0 \\ 0 \end{pmatrix}, \quad (\bar{e}_2^\mu) = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \quad (\bar{e}_3^\mu) = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}. \quad (2.3.1)$$

Dabei ist $\gamma = 1/\sqrt{1 - \beta^2}$. Da weiter voraussetzungsgemäß beide Basissysteme pseudo-kartesisch sein sollen, d.h. daß $e_\mu \cdot e_\nu = \bar{e}_\mu \cdot \bar{e}_\nu = \eta_{\mu\nu}$ mit den Komponenten des Minkowski-Produkts bzgl. solcher Basen ($\eta_{\mu\nu} = \text{diag}(1, -1, -1, -1)$) gilt, folgt daraus sofort.

$$(\bar{x}^\mu) = \begin{pmatrix} \bar{e}_0 \cdot \mathbf{x} \\ -\bar{e}_1 \cdot \mathbf{x} \\ -\bar{e}_2 \cdot \mathbf{x} \\ -\bar{e}_3 \cdot \mathbf{x} \end{pmatrix} = \begin{pmatrix} \gamma(ct - \beta x^1) \\ \gamma(x^1 - \beta ct) \\ x^2 \\ x^3 \end{pmatrix}. \quad (2.3.2)$$

Damit gelangen wir wieder zur Lorentz-Transformationsmatrix (2.1.14)

$$\bar{x}^\mu = \Lambda^\mu{}_\nu x^\nu \quad \text{mit} \quad (\Lambda^\mu{}_\nu) = \hat{\Lambda} = \begin{pmatrix} \gamma & -\beta\gamma & 0 & 0 \\ -\beta\gamma & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}. \quad (2.3.3)$$

Es ist nun leicht, dies für eine Bewegung von B mit beliebiger konstanter Geschwindigkeit $\vec{v} = c\vec{\beta}$ gegenüber A zu verallgemeinern. In offensichtlicher „(1+3)-dimensionaler Schreibweise“ ergibt sich

$$\hat{B}(\vec{\beta}) = \begin{pmatrix} \gamma & -\gamma\vec{\beta}^T \\ -\gamma\vec{\beta} & (\gamma - 1)\hat{\beta}\hat{\beta}^T + \mathbb{1}_3 \end{pmatrix} \quad (2.3.4)$$

³Vgl. dazu den entsprechenden Wikipedia-Artikel

https://en.wikipedia.org/wiki/Time_dilation_of_moving_particles

2. Spezielle Relativitätstheorie

Dabei ist $\hat{\beta} = \vec{\beta}/|\vec{\beta}|$, und es gilt weiter

$$(\hat{\beta}\hat{\beta}^T)^j_k = \hat{\beta}^j\hat{\beta}^k. \quad (2.3.5)$$

Diese **rotationsfreien Lorentz-Boosts** sind nun offensichtlich nicht die einzigen Transformationen, die das Minkowski-Produkt im Sinne von

$$\eta_{\mu\nu}x^\mu x^\nu = \eta_{\mu\nu}\bar{x}^\mu\bar{x}^\nu \quad \text{mit} \quad \bar{x}^\rho = \Lambda^\rho_\mu x^\mu \quad (2.3.6)$$

invariant lassen. Da dies für alle $(x^\mu) \in \mathbb{R}^4$ gelten muß, ergibt sich als notwendige und hinreichende Bedingung an die Lorentz-Transformationsmatrix $\hat{\Lambda} = (\Lambda^\mu_\nu$

$$\eta_{\rho\sigma}\Lambda^\rho_\mu\Lambda^\sigma_\nu = \eta_{\mu\nu} \quad (2.3.7)$$

bzw. in Matrixnotation

$$\hat{\Lambda}^T \hat{\eta} \hat{\Lambda} = \hat{\eta}. \quad (2.3.8)$$

Dies impliziert, daß alle diese Matrizen zusammengenommen eine **Gruppe** bilden, denn offenbar erfüllt das Produkt irgendwelcher zwei Lorentz-Transformationsmatrizen wieder (2.3.8). In der Tat folgt wegen $\hat{\eta}^1 = \mathbb{1}$

$$(\hat{\Lambda}_1 \hat{\Lambda}_2)^T \hat{\eta} (\hat{\Lambda}_1 \hat{\Lambda}_2) \hat{\eta} = \hat{\Lambda}_2^T \hat{\Lambda}_1^T \hat{\eta} \hat{\Lambda}_1 \hat{\Lambda}_2 = \hat{\Lambda}_2^T \hat{\eta} \hat{\Lambda}_2 = \hat{\eta}. \quad (2.3.9)$$

Offenbar sind nun neben den allgemeinen drehungsfreien Boosts (2.3.4) auch alle Transformationen, die einer Drehung oder Raumspiegelung entsprechen, also Matrizen der Form

$$\hat{R} = \begin{pmatrix} 1 & \vec{0}^T \\ \vec{0} & \hat{O} \end{pmatrix}, \quad (2.3.10)$$

wobei $\vec{0}^T = (0, 0, 0)$ und $\hat{O} \in O(3)$ eine orthogonale $\mathbb{R}^{3 \times 3}$ -Matrix ist. Ebenso kann man Lorentztransformationen durch Multiplikation beliebiger Lorentz-Boost- und Rotationsmatrizen erzeugen.

orthogonal $\mathbb{R}^{3 \times 3}$ matrix and also all combinations of such matrices. Man kann zeigen, daß die allgemeinste Lorentz-Transformationsmatrix in der Form

$$\hat{\Lambda} = \hat{R}_1 \hat{B}(\beta \vec{e}_1) \hat{R}_2. \quad (2.3.11)$$

geschrieben werden kann, wobei \hat{R}_1 und \hat{R}_2 Rotationsmatrizen der Form (2.3.10) und $\hat{B}(\beta \vec{e}_1)$ durch (2.3.4) definiert ist und in dem Fall mit (2.3.3) übereinstimmt. Auf den ersten Blick sieht es nun so aus, als benötigte man 7 Parameter, um eine allgemeine Lorentz-Transformation zu charakterisieren, nämlich jeweils 3 für die beiden Drehungen und die Relativgeschwindigkeit β zwischen den Inertialsystemen. Dies ist aber nicht korrekt, wie wir nun zeigen wollen.

Dazu stellen wir eine geometrische Betrachtung an und denken wieder an zwei inertielle Beobachter A und B , die jeweils beliebige pseudokartesische Basisvektoren e_μ bzw. \bar{e}_ν verwenden. Der einzige Unterschied zu dem oben verwendeten Spezialfall ist, daß nun beide Beobachter beliebig zueinander orientierte räumliche kartesische Basisvektoren verwenden. Zur Einfachheit nehmen wir an, daß beide räumlichen kartesischen Basissysteme *rechtshändig* orientiert sind. Wir können nun die Lorentz-Matrix durch sukzessive Transformationen der Art (2.3.11) gewinnen. Die Bewegung von B 's räumlichem Koordinatenursprung relativ zu A 's räumlichem Koordinatenursprung sei durch den Vektor $\vec{v} = c\vec{\beta}$ gegeben. Als erstes verwenden wir eine Rotationsmatrix $(R_2^{-1})^\rho_\mu$, die Alices räumliches Basissystem so verdreht, daß ihre neue 1-Achse in die Richtung von \vec{v} fällt. Dies liefert das neue Basissystem

$$e'_\mu = (R_2^{-1})^\rho_\mu e_\rho, \quad (2.3.12)$$

das immer noch bzgl. A ruht. Dafür benötigen wir aber nur *zwei* Winkel, um diese Rotation zu, nämlich den Polar- und Azimutal-Winkel von A 's alter 1-Achse relativ zur Richtung $\hat{\beta}$ von B 's Dreiergeschwindigkeit.

Dann führt A den Boost $\hat{B}^{-1}(\beta\vec{e}'_1)$ aus, um in B 's Ruhssystem zu transformieren, was zu einer weiteren pseudo-kartesischen Basis

$$\mathbf{e}''_{\mu} = (B^{-1}(\beta\vec{e}'_1))_{\mu}^{\rho} \mathbf{e}'_{\rho} \quad (2.3.13)$$

führt. In dem dazugehörigen Bezugssystem ist nun B in Ruhe, aber die räumlichen Vektoren dieser Basis brauchen nicht notwendig mit den von B gewählten räumlichen Basisvektoren übereinzustimmen. Dies kann aber durch eine weitere beliebige Rotation, für die wir drei Drehwinkel (oder drei Euler-Winkel, vgl. [Hee14], Abschnitt 1.7.3) benötigen:

$$\bar{\mathbf{e}}_{\mu} = (R_1^{-1})_{\mu}^{\rho} \mathbf{e}''_{\rho}. \quad (2.3.14)$$

Bringen wir (2.3.12-2.3.13) zusammen, erhalten wir

$$\bar{\mathbf{e}}_{\mu} = (\hat{R}_2^{-1} \hat{B}(\beta\vec{e}'_1) \hat{R}_1^{-1})_{\mu}^{\nu} \mathbf{e}_{\nu} \Rightarrow \mathbf{e}_{\mu} = (\hat{R}_1 \hat{B}(\beta\vec{e}'_1) \hat{R}_2)_{\mu}^{\nu} \bar{\mathbf{e}}_{\nu} = \Lambda^{\nu}_{\mu} \bar{\mathbf{e}}_{\nu}, \quad (2.3.15)$$

wie in (2.3.11) behauptet. Bei der Drehung \hat{R}_2 benötigen wir aber nur *zwei* Drehwinkel. Insgesamt ist also eine allgemeine Lorentztransformation durch insgesamt **sechs Parameter**, nämlich zwei Drehwinkel für \hat{R}_2 , drei Drehwinkel für \hat{R}_1 und die Boostgeschwindigkeit $v = c\beta$ mit $0 \leq \beta < 1$.

Eine andere Art, einen rotationsfreien Boost zu charakterisieren ist die sog. **Rapidity**. Die Idee dabei ist, daß man einen Boost in A 's 1-Richtung als eine Art „hyperbolischer Drehung“ der Basisvektoren der 01-Ebene, die wir oben in den Minkowski-Diagrammen dargestellt haben, auffassen kann. Für den Lorentz-Boost in 1-Richtung (2.3.3) führen wir die Rapidity $\eta \in \mathbb{R}$ ein, indem wir

$$\Lambda^0_0 = \gamma = \cosh \eta, \quad \Lambda^0_1 = \Lambda^1_0 = -\beta\gamma = -\sinh \eta, \quad \Lambda^1_1 = \cosh \eta. \quad (2.3.16)$$

setzen. Dann gilt offenbar $\eta_{\mu\nu} \Lambda^{\mu}_{\rho} \Lambda^{\nu}_{\sigma} = \eta_{\rho\sigma}$, denn es ist $\cosh^2 \eta - \sinh^2 \eta = 1$. Der Zusammenhang zu B 's Daraus folgt für die Geschwindigkeit $v = \beta c$ von B relativ zu A

$$\beta = -\frac{\Lambda^1_0}{\Lambda^0_0} = \frac{\sinh \eta}{\cosh \eta} = \tanh \eta, \quad (2.3.17)$$

und die Lorentz-Transformationsmatrix ergibt sich zu

$$(B^{\mu}_{\nu}) = \begin{pmatrix} \cosh \eta & -\sinh \eta & 0 & 0 \\ -\sinh \eta & \cosh \eta & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}. \quad (2.3.18)$$

Mit dieser Parametrisierung ist es leicht, nachzurechnen, daß die Hintereinanderausführung zweier Lorentz-Boosts in *dieselbe* Richtung wieder einen Lorentz-Boost in diese Richtung ergibt, und zwar ist (*nachrechnen!*)

$$\hat{B}_1(\eta_1) \hat{B}_1(\eta_2) = \hat{B}_1(\eta_1 + \eta_2), \quad (2.3.19)$$

wobei wir die Additionstheoreme der Hyperbelfunktionen

$$\begin{aligned} \sinh(\eta_1 + \eta_2) &= \cosh \eta_1 \sinh \eta_2 + \sinh \eta_1 \cosh \eta_2, \\ \cosh(\eta_1 + \eta_2) &= \cosh \eta_1 \cosh \eta_2 + \sinh \eta_1 \sinh \eta_2 \end{aligned} \quad (2.3.20)$$

verwendet haben. Wir bestimmen noch die Dreiergeschwindigkeit β des zusammengesetzten Boosts. Wegen (2.3.17) und (2.3.19) folgt

$$\beta = \tanh(\eta_1 + \eta_2) = \frac{\sinh(\eta_1 + \eta_2)}{\cosh(\eta_1 + \eta_2)} \stackrel{(2.3.20)}{=} \frac{\sinh \eta_1 \cosh \eta_2 + \sinh \eta_2 \cosh \eta_1}{\cosh \eta_1 \cosh \eta_2 + \sinh \eta_1 \sinh \eta_2}. \quad (2.3.21)$$

Nun kürzen wir den Bruch mit $\cosh \eta_1 \cosh \eta_2$ und erhalten damit schließlich

$$\beta = \frac{\tanh \eta_1 + \tanh \eta_2}{1 + \tanh \eta_1 \tanh \eta_2} = \frac{\beta_1 + \beta_2}{1 + \beta_1 \beta_2}. \quad (2.3.22)$$

Wie wir sehen, geht dies für $|\beta_1 \beta_2| \ll 1$ in die aus der Newtonschen Physik zu erwartende einfache Addition der Boost-Geschwindigkeiten über

$$\beta = (\beta_1 + \beta_2)[1 + \mathcal{O}(\beta_1 \beta_2)]. \quad (2.3.23)$$

Die Boosts *in einer Richtung* ergeben also insgesamt eine Abelsche Untergruppe der Lorentz-Gruppe, denn wegen (2.3.19) gilt

$$\hat{B}_1(\eta_1)\hat{B}_1(\eta_2) = \hat{B}_1(\eta_2)\hat{B}_1(\eta_1) = \hat{B}_1(\eta_1 + \eta_2). \quad (2.3.24)$$

Wir bemerken ohne im Detail darauf weiter einzugehen, daß die Boosts insgesamt *keine Untergruppe* der Lorentz-Gruppe bilden. Die Hintereinanderausführung zweier Boosts in unterschiedlicher Richtung ist weder kommutativ noch ergibt sie wieder einen reinen Boost sondern einen Boost mit einer zusätzlichen Rotation, die sog. **Wigner-Rotation**.

2.4 Addition von Geschwindigkeiten

Jetzt betrachten wir die Geschwindigkeit eines Punktteilchens in verschiedenen Inertialsystemen. Sei dazu \vec{w} die Geschwindigkeit des Punktes relativ zu B 's Inertialsystem, das sich relativ zu A 's System mit der Geschwindigkeit $\vec{v} = \beta c \vec{e}_1$ bewegt. Dazu führen wir zunächst Lorentz-Vektoren für die Geschwindigkeit ein, die **Vierergeschwindigkeit**. Dazu führen wir zunächst die **Eigenzeit** des Punktteilchens ein. Ist $x(\lambda)$ die Weltlinie des Teilchens im Minkowski-Raum, wobei λ ein beliebiger Parameter ist, so läßt sich zunächst statt der Koordinatenzeit bzgl. eines willkürlichen Inertialsystems das Inkrement der Eigenzeit des Teilchens durch

$$d\tau^2 = \frac{1}{c^2} d\lambda^2 \frac{dx}{d\lambda} \cdot \frac{dx}{d\lambda} \quad (2.4.1)$$

definieren. Offenbar ist diese Definition unabhängig von der Wahl des gewählten Parameters λ . Verwenden wir die Zeit eines beliebigen Inertialsystems, folgt

$$d\tau = dt \sqrt{1 - \dot{\vec{x}}^2/c^2} = \frac{dt}{\gamma}. \quad (2.4.2)$$

Dabei muß freilich stets gewährleistet sein, daß $|\dot{\vec{x}}| < c$ ist, d.h. ein Teilchen kann sich relativ zu einem beliebigen Inertialsystem nur mit einer Geschwindigkeit unterhalb der Lichtgeschwindigkeit bewegen. Statt der Dreiergeschwindigkeit $\dot{\vec{x}}$ verwenden wir nun die **Vierergeschwindigkeit**

$$\mathbf{u} = \frac{1}{c} \frac{dx}{d\tau}, \quad (u^\mu) = \gamma \begin{pmatrix} 1 \\ \dot{\vec{x}}/c \end{pmatrix}, \quad (2.4.3)$$

die offensichtlich einen **Vierervektor** bildet, denn $d\tau$ ist gemäß (2.4.1) invariant unter Lorentz-Transformationen und definiert das Zeitinkrement, die im momentanen inertialen Ruhssystem des Teilchens vergeht.

Betrachten wir nun wieder die obige Fragestellung. Die Vierergeschwindigkeit des Teilchens besitzt bzgl. B 's Inertialsystem gemäß (2.4.3) die Komponenten

$$(\overline{W}^\mu) = \gamma_w \begin{pmatrix} 1 \\ \vec{\beta}_w \end{pmatrix} \quad \text{mit} \quad \vec{\beta}_w = \frac{1}{c} \vec{w}, \quad \gamma_w = \frac{1}{\sqrt{1 - \beta_w^2}}. \quad (2.4.4)$$

Wir können mit (2.3.6) sofort die Komponenten in A 's Inertialsystem angeben, wobei $\hat{\Lambda} = \hat{B}_1(v)$

$$(W^\mu) = \hat{\Lambda}^{-1}(\overline{W}^\mu) = \hat{B}_1(-v)\overline{W}^\mu = \gamma_w \begin{pmatrix} \gamma_v(1 + \beta_v \overline{\beta}_w^1) \\ \gamma_v(\beta_v + \overline{\beta}_w^1) \\ \overline{\beta}_w^2 \\ \overline{\beta}_w^3 \end{pmatrix}. \quad (2.4.5)$$

Für die gewöhnliche Dreiergeschwindigkeit in A 's Bezugssystem ergibt sich daraus

$$\vec{w} = c \frac{\vec{W}}{W^0} = \frac{c}{1 + \beta_v \overline{\beta}_w^1} \begin{pmatrix} \beta_v + \overline{\beta}_w^1 \\ \overline{\beta}_w^2 / \gamma_v \\ \overline{\beta}_w^3 / \gamma_v \end{pmatrix}. \quad (2.4.6)$$

Wir wir sehen, ergibt sich ein recht kompliziertes **Additionsgesetz für Geschwindigkeiten**, was daran liegt, daß hier bzgl. der Kinematik des Teilchens alle drei kinematischen Effekte der Raum-Zeit-Struktur auftreten: Für die Komponente der Geschwindigkeit in Boost-Richtung haben wir die Längenkontraktion, die Relativität der Gleichzeitigkeit und die Zeitdilatation bei der Umrechnung zu beachten, während für die Komponenten senkrecht zur Boost-Richtung die Längenkontraktion wegfällt.

Drücken wir (2.4.6) mit Hilfe der Dreiervektoren aus, folgt

$$\vec{\beta}_w = \frac{1}{c} \vec{w} = \frac{1}{1 + \vec{\beta}_v \cdot \vec{\beta}_w} \left[\vec{\beta}_v + \frac{1}{\gamma_v} \vec{\beta}_w + \left(1 - \frac{1}{\gamma_v}\right) \vec{\beta}_v \frac{\vec{\beta}_v \cdot \vec{\beta}_w}{\beta_v^2} \right], \quad (2.4.7)$$

und das läßt sich mit einiger Vektoralgebra in (*nachprüfen!*)

$$\vec{\beta}_w = \frac{1}{1 + \vec{\beta}_v \cdot \vec{\beta}_w} \left[\vec{\beta}_v + \vec{\beta}_w + \frac{\gamma_v}{1 + \gamma_v} \vec{\beta}_v \times (\vec{\beta}_v \times \vec{\beta}_w) \right] \quad (2.4.8)$$

umformen. Für den Betrag der Geschwindigkeit bzgl. A 's Inertialsystem ergibt sich nach einiger Rechnung (*Übung!*)

$$|\vec{\beta}_w| = \frac{1}{1 + \vec{\beta}_v \cdot \vec{\beta}_w} \sqrt{\left(\vec{\beta}_v + \vec{\beta}_w\right)^2 - \left(\vec{\beta}_v \times \vec{\beta}_w\right)^2}. \quad (2.4.9)$$

2.5 Relativgeschwindigkeit zweier Teilchen

In engem Zusammenhang mit dem im vorigen Abschnitt hergeleiteten Additionstheorem der Geschwindigkeiten steht die Definition der **Relativgeschwindigkeit zweier Teilchen**. Dies spielt u.a. eine wichtige Rolle, um eine Lorentz-invariante Definition von Streuquerschnitten für Zweiteilchenstreuungen definieren zu können. Dabei ist die Relativgeschwindigkeit eines Teilchens 2 gegenüber einem Teilchen 1 definiert als die Geschwindigkeit des Teilchens 2 im (momentanen) Ruhssystem des Teilchens 1. Um diese Relativgeschwindigkeit aus den Komponenten der Geschwindigkeit bzgl. eines beliebigen Inertialsystems zu berechnen, ist es wieder bequemer zunächst mit den Vierergeschwindigkeiten

$$u_1^\mu = \frac{1}{\sqrt{1 - \beta_1^2}} \begin{pmatrix} 1 \\ \vec{\beta}_1 \end{pmatrix}, \quad u_2^\mu = \frac{1}{\sqrt{1 - \beta_2^2}} \begin{pmatrix} 1 \\ \vec{\beta}_2 \end{pmatrix} \quad (2.5.1)$$

zu arbeiten. Dann ist der Lorentz-Boost ins Ruhssystem von Teilchen 1 durch

$$(\Lambda^\mu{}_\nu) = \hat{B}(\vec{\beta}_1) = \begin{pmatrix} \gamma_1 & -\gamma_1 \vec{\beta}_1^T \\ -\gamma_1 \vec{\beta}_1 & \mathbb{1}_3 + (\gamma_1 - 1) \hat{\beta}_1 \hat{\beta}_1^T \end{pmatrix} \quad (2.5.2)$$

gegeben., denn dann gilt in der Tat

$$(\tilde{u}_1^\mu) = (\Lambda^\mu{}_\nu u_1^\nu) = (1, 0, 0, 0)^T. \quad (2.5.3)$$

Wir können also das Resultat des vorigen Abschnitts zur Addition von Geschwindigkeiten benutzen, indem wir $\vec{W}^\mu = u_2^\mu$ bzw. $\vec{\beta}_w = \vec{\beta}_2$ und $\vec{\beta}_v = -\vec{\beta}_1$ setzen. Dann folgt mit

$$\vec{\beta}_{\text{rel}} = \frac{1}{1 - \vec{\beta}_1 \cdot \vec{\beta}_2} \left[\vec{\beta}_2 - \vec{\beta}_1 + \frac{\gamma_1}{1 + \gamma_1} \vec{\beta}_1 \times (\vec{\beta}_1 \times \vec{\beta}_2) \right] \quad (2.5.4)$$

and from (2.4.9)

$$|\vec{\beta}_{\text{rel}}| = \frac{1}{1 - \vec{\beta}_1 \cdot \vec{\beta}_2} \sqrt{(\vec{\beta}_2 - \vec{\beta}_1)^2 - (\vec{\beta}_1 \times \vec{\beta}_2)^2}. \quad (2.5.5)$$

2.6 Die Lorentz-Gruppe als Lie-Gruppe

Die Lorentz-Gruppe hängt ähnlich wie die Drehgruppe, die ja eine Untergruppe der Lorentz-Gruppe ist, kontinuierlich von Parametern wie Drehwinkeln und Boost-Geschwindigkeiten ab. Daher ergibt die Frage Sinn, inwiefern man Lorentz-Transformationen als Hintereinanderausführung sehr vieler „infinitesimaler“ Lorentz-Transformationen aufgebaut denken kann. Dazu betrachten wir eine allgemeine Lorentz-Transformation, die sich von der Einheits-Matrix nur sehr wenig unterscheidet:

$$\Lambda^\mu{}_\nu = \delta^\mu{}_\nu + \delta \omega^\mu{}_\nu. \quad (2.6.1)$$

Die Lorentz-Transformation läßt sich gemäß (2.3.8) als **pseudoorthogonale Transformation** charakterisieren, d.h. für die infinitesimal von der Einheitsmatrix abweichende Lorentz-Transformation gilt bis zur Ordnung $\mathcal{O}(\delta \hat{\omega})$

$$\eta_{\mu\nu} \Lambda^\mu{}_\rho \Lambda^\nu{}_\sigma = \eta_{\rho\sigma} + (\delta \omega_{\rho\sigma} + \delta \omega_{\sigma\rho}) + \mathcal{O}(\delta^2) \stackrel{!}{=} \eta_{\rho\sigma}, \quad (2.6.2)$$

wobei wir die Übliche Regel zum „Indezziehen“

$$\delta \omega_{\rho\sigma} = \eta_{\rho\mu} \delta \omega^\mu{}_\sigma \quad (2.6.3)$$

verwendet haben; d.h. aus (2.6.2) folgt bis zur ersten Ordnung in $\delta \hat{\omega}$

$$\delta \omega_{\rho\sigma} = -\delta \omega_{\sigma\rho}. \quad (2.6.4)$$

Eine infinitesimale Lorentz-Transformation ist demnach durch eine **antisymmetrische Matrix** $\delta \omega_{\rho\sigma}$, also $(4 \cdot 4 - 4)/2 = 6$ unabhängige reelle Parameter charakterisiert.

Für die infinitesimalen Matrizen mit einem ersten oberen und einem zweiten unteren Index gilt

$$\delta \omega^\mu{}_\nu = \eta^{\mu\rho} \delta \omega_{\rho\nu} = -\eta^{\mu\rho} \delta \omega_{\nu\rho}. \quad (2.6.5)$$

Wir zerlegen nun diese Matrizen nach zeitlichen und räumlichen Komponenten. Offenbar ist

$$\delta \omega^0{}_j = \eta^{0\rho} \delta \omega_{\rho j} = \delta \omega_{0j}, \quad \delta \omega^j{}_0 = \eta^{j\rho} \delta \omega_{\rho 0} = \eta^{jk} \delta \omega_{k0} = -\omega_{j0} = +\omega_{0j}. \quad (2.6.6)$$

Das bedeutet, daß $\delta \omega^0_j = \delta \omega^j_0$. Die gemischt zeitlich-räumlichen Komponenten der Matrix lassen sich also als Linearkombination der drei **symmetrischen Matrizen**

$$[(\hat{K}^1)^\mu_\nu] = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \quad [(\hat{K}^2)^\mu_\nu] = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \quad [(\hat{K}^3)^\mu_\nu] = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix} \quad (2.6.7)$$

schreiben.

Wir können nun diese matrixwertigen Komponenten in einen Dreivektor $\hat{\vec{K}}$ anordnen, d.h. die zeit-räumlichen Komponenten der infinitesimalen Lorentz-Transformationsmatrix können in der Form

$$\hat{B}(\delta \vec{\beta}) = \mathbb{1}_4 - \delta \vec{\beta} \cdot \hat{\vec{K}} \quad (2.6.8)$$

geschrieben werden. Intuitiv ist klar, daß dies **infinitesimalen Boosts** mit einem infinitesimalen Boost-Geschwindigkeitsvektor $\delta \vec{\beta}$ entspricht. Um dies zu verifizieren, entwickeln wir (2.3.4) bis zur Ordnung $\mathcal{O}(\delta \vec{\beta})$ um $\vec{\beta} = 0$:

$$\hat{B}(\delta \vec{\beta}) = \mathbb{1}_4 + \delta \vec{\beta} \cdot \left(\frac{\partial}{\partial \vec{\beta}} \hat{B}(\vec{\beta}) \right)_{\vec{\beta}=0} + \mathcal{O}(\delta^2). \quad (2.6.9)$$

Es gilt

$$\gamma = \frac{1}{\sqrt{1 - \delta \vec{\beta}^2}} = 1 + \mathcal{O}(\delta^2), \quad (2.6.10)$$

und daher folgt in der Tat

$$B(\delta \vec{\beta}) = \begin{pmatrix} 1 & -\delta \vec{\beta}^T \\ -\delta \vec{\beta} & \mathbb{1}_3 \end{pmatrix} + \mathcal{O}(\delta^2) \equiv \mathbb{1}_4 - \delta \vec{\beta} \cdot \hat{\vec{K}} + \mathcal{O}(\delta^2). \quad (2.6.11)$$

Für die rein räumlichen Komponenten der infinitesimalen Matrix $\delta \hat{\omega}$ folgt

$$\delta \omega^j_k = \eta^{j\rho} \delta \omega_{\rho k} = -\delta \omega_{jk}, \quad \Rightarrow \delta \omega^j_k = -\delta \omega_{kj} = \delta \omega_{jk} = -\delta \omega^j_k, \quad (2.6.12)$$

d.h. der rein räumliche Anteil der infinitesimalen Lorentz-Transformation ist durch **antisymmetrische Matrizen** $\delta \hat{\omega} = -\delta \hat{\omega}^T$ charakterisiert und entspricht **Rotationen**. Diese antisymmetrische Matrizen lassen sich als Linearkombinationen der drei antisymmetrischen Matrizen

$$[(J^1)^\mu_\nu] = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 \end{pmatrix}, \quad [(J^2)^\mu_\nu] = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{pmatrix}, \quad [(J^3)^\mu_\nu] = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \quad (2.6.13)$$

schreiben.

Wir untersuchen nun, wie wir aus den infinitesimalen Transformationen endliche Transformationen gewinnen können. Dazu betrachten wir die Rotation $\hat{R}(\varphi \vec{n})$ um eine durch den Einheitsvektor \vec{n} gegebene Achse um den Drehwinkel φ (im Sinne der Rechten-Handregel)

$$x'(\varphi) = \hat{R}(\varphi \vec{n})x. \quad (2.6.14)$$

2. Spezielle Relativitätstheorie

Nehmen wir nun an, daß eine Rotation um den infinitesimalen Drehwinkel $\delta\varphi$ durch $-\delta\varphi\vec{n}\cdot\hat{J}$ gegeben ist, sollte für (2.6.14) die Differentialgleichung

$$\frac{d}{d\varphi}x' = -\vec{n}\cdot\hat{J}x \quad (2.6.15)$$

gelten. Nehmen wir nun weiter an, daß wir diese Gleichung formal so lösen können als wäre die konstante Matrix $\vec{n}\cdot\hat{J}$ eine Zahl, ergibt sich eine formale Lösung in der Form

$$x'(\varphi) = \exp(-\varphi\vec{n}\cdot\hat{J})x. \quad (2.6.16)$$

Dabei haben wir auch benutzt, daß die Anfangsbedingung $x'(\varphi=0) = x$ gelten soll. Rechts steht nun die Exponentialfunktion einer Matrix. Diese definieren wir durch die entsprechende Taylor-Reihe mit matrixwertigen Summanden:

$$\exp\hat{M} = \sum_{k=0}^{\infty} \frac{1}{k!}\hat{M}^k, \quad \hat{M}^0 = \mathbb{1}. \quad (2.6.17)$$

Wir können nun (2.6.16) auswerten, indem wir $(\vec{n}\cdot\hat{J})^k$ für $k \in \mathbb{N}$ berechnen. Wir brauchen natürlich nur die rein räumlichen Komponenten der Matrix zu betrachten, also 3×3 -dimensionale Untermatrizen der 4×4 -Minkowski-Matrizen (2.6.13) Für diese Untermatrizen können wir offenbar

$$(\vec{n}\cdot\hat{J})^j_k = n^i \epsilon^{ijk} \quad (2.6.18)$$

schreiben, wobei wir für die rein räumlichen Komponenten nur obere Indizes schreiben. Hierbei ist ϵ^{ijk} das gewöhnliche **Levi-Civita-Symbol** im dreidimensionalen euklidischen Raum. Es ist dadurch definiert, daß $\epsilon^{123} = 1$ ist. Ansonsten ist es total antisymmetrisch unter Vertauschung seiner Indizes. Dann folgt

$$[(\vec{n}\cdot\hat{J})^2]^j_k = n^i \epsilon^{ijk'} n^{i'} \epsilon^{i'k'k} = -\delta^{jk} + n^i n^j. \quad (2.6.19)$$

Da Potenzen dieser Matrix bis auf das Vorzeichen diese Matrix reproduzieren (*nachrechnen*), folgt

$$(\vec{n}\cdot\hat{J})^{2k} = (-1)^k \begin{pmatrix} 0 & 0 \\ 0 & \mathbb{1}_3 - \vec{n}\vec{n}^T \end{pmatrix} = (-1)^k \hat{P}_{\perp}(\vec{n}), \quad k \in \{1, 2, 3, \dots\} \quad (2.6.20)$$

und

$$(-\vec{n}\cdot\hat{J})^{2k+1} = (-1)^{k+1} \vec{n}\cdot\hat{J}, \quad k \in \{0, 1, 2, \dots\}. \quad (2.6.21)$$

Setzen wir dies in (2.6.16) ein und schreiben die Reihe (2.6.17) für die Matrixexponentialfunktion aus, ergibt sich

$$\hat{D}(\vec{\varphi}) = \exp(-\varphi\vec{n}\cdot\hat{J}) = \hat{P}_{\parallel}(\vec{n}) + \cos\varphi\hat{P}_{\perp}(\vec{n}) - \sin\varphi(\vec{n}\cdot\hat{J}). \quad (2.6.22)$$

Angewandt auf einen Vierervektor ergibt sich schließlich (*nachrechnen!*)

$$\hat{D}(\vec{\varphi})x = \begin{pmatrix} x^0 \\ \vec{n}(\vec{n}\cdot\vec{x}) + \cos\varphi[\vec{x} - \vec{n}(\vec{n}\cdot\vec{x})] - \sin\varphi(\vec{n}\times\vec{x}) \end{pmatrix}, \quad (2.6.23)$$

und dies beschreibt in der Tat die Änderung der Komponenten des räumlichen Anteils unter einer Drehung um eine Achse in Richtung \vec{n} um den Drehwinkel φ .

Auf analoge Weise lassen sich auch die endlichen Lorentz-Boosts mit Hilfe der Generatoren (2.6.7) gewinnen. Als erstes bemerken wir, daß

$$\vec{n}\cdot\hat{K} = \begin{pmatrix} 0 & \vec{n}^T \\ \vec{n} & 0_3 \end{pmatrix}, \quad (\vec{n}\cdot\hat{K})^2 = \hat{P}_0 + \hat{P}_{\parallel}(\vec{n}) \quad (2.6.24)$$

mit

$$\hat{P}_0 = \begin{pmatrix} 1 & 0 \\ 0 & 0_3 \end{pmatrix}. \quad (2.6.25)$$

Daraus folgt

$$(\vec{n} \cdot \hat{K})^k = \begin{cases} \vec{n} \cdot \hat{K} & \text{for } k \in \{1, 3, 5, \dots\}, \\ \hat{P}_0 + \hat{P}_{||}(\vec{n}) & \text{for } k \in \{2, 4, 6, \dots\}. \end{cases} \quad (2.6.26)$$

Daraus ergibt sich für die Exponentialreihe (*nachrechnen*)

$$\exp(-\eta \vec{n} \cdot \hat{K}) = \cosh \eta \hat{P}_0 - \sinh \eta \vec{n} \cdot \hat{K} + (\cosh \eta - 1) \hat{P}_{||}(\vec{n}). \quad (2.6.27)$$

Setzen wir nun $\gamma = \cosh \eta$ und $\beta = \tanh \eta$, finden wir

$$\hat{B}(\vec{\beta}) = \exp[-\eta(|\vec{\beta}|) \vec{n} \cdot \hat{K}] \quad \text{mit} \quad \eta(\beta) = \operatorname{artanh} \beta. \quad (2.6.28)$$

Damit ist klar, daß die **Rapidity** $\eta = \operatorname{artanh} \beta$ der natürliche Parameter für den Boost ist, denn bei der Hintereinanderausführung zweier Boosts *in die gleiche Richtung* addieren sich gemäß (2.3.19) die Rapiditäten. Andererseits folgt für einen infinitesimalen Boost

$$\eta(\delta \beta) \vec{n} = \operatorname{artanh} \delta \beta \vec{n} = [\delta \beta + \mathcal{O}(\delta \beta^3)] \vec{n} = \delta \vec{\beta} + \mathcal{O}(\delta \beta^3). \quad (2.6.29)$$

Damit ist also bis auf Terme der Ordnung $\mathcal{O}(\delta \beta^3)$ $\eta \vec{n} \simeq \delta \vec{\beta}$ in Übereinstimmung mit (2.6.11).

Wir bemerken noch, daß diese aus der Identität stetig hervorgehenden Lorentz-Transformationen, d.h. diejenigen Matrizen $\hat{A}(\lambda)$, deren Komponenten als Funktion eines Parameters λ stetig sind und für die $\hat{A}(0) = \mathbb{1}$ gilt, nicht alle Lorentz-Transformationen umfaßt. Dazu betrachten wir zunächst nochmals die Bedingung (2.3.8) dafür, daß eine Matrix \hat{A} eine Lorentz-Transformation ist. Bilden wir die Determinante auf beiden Seiten der Gleichung, ergibt sich

$$\det(\hat{A}^T \hat{\eta} \hat{A}) = \det \hat{\eta} = -1 \Rightarrow [\det(\hat{A})]^2 = 1, \quad (2.6.30)$$

wobei wir benutzt haben, daß für beliebige Matrizen $\det(\hat{A}\hat{B}) = \det(\hat{A})\det(\hat{B})$ und $\det(\hat{A}^T) = \det \hat{A}$ gilt. Die Bedingung (2.3.8) besagt, daß sich die darstellende Matrix $\hat{\eta}$ des Minkowskiprodukts unter Lorentz-Transformationen nicht ändert. Die Lorentz-Transformationen bilden mit der Hintereinanderausführung (entsprechend der Multiplikation der Lorentz-Matrizen) als Produkt eine Gruppe, die **Lorentz-Gruppe**. Da *hat* η einen positiven und drei negative Eigenwerte besitzt, nennt man die Lorentz-Gruppe auch die (pseudo-)orthogonale Gruppe des Minkowskiprodukts, das durch die **Signatur** (1, 3), die die Anzahl der positiven und negativen Eigenwerte ihrer darstellenden Matrix bzgl. einer beliebigen Basis des \mathbb{R}^4 angibt⁴, charakterisiert ist und bezeichnet sie mit $O(1, 3)$.

Jedenfalls folgt aus (2.6.30), daß eine Lorentz-Transformation entweder die Determinante +1 oder -1 haben muß. Da die Determinante eine stetige Funktion der Komponenten ist, können nur diejenigen Lorentz-Matrizen stetig aus der Einheitsmatrix hervorgehen, für die $\det \hat{A} = +1$ ist. Die Lorentz-Matrizen mit Determinante 1 bilden wieder eine Gruppe (*warum?*), und sie sind eine echte **Untergruppe** der Lorentz-Gruppe. Man nennt sie $SO(1, 3)$, die **spezielle (pseudo-)orthogonale Gruppe** zur Bilinearform mit Signatur (1, 3).

Selbst die $SO(1, 3)$ charakterisiert aber noch nicht die Zusammenhangskomponente mit der Gruppenidentität vollständig. Die Lorentz-Bedingung lautet in Komponentenschreibweise geschrieben

$$\eta_{\rho\sigma} \Lambda^\rho_\mu \Lambda^\sigma_\nu = \eta_{\mu\nu}. \quad (2.6.31)$$

⁴Wir bemerken, daß die Anzahl der positiven und negativen Eigenwerte der darstellenden Matrix einer Bilinearform für *alle* Basen gleich bleibt, auch für solche Basen, die nicht (pseudo-)kartesisch sind. Dies ist als **Sylvesterscher Trägheitssatz** bekannt. Für einen Beweis s. Lehrbücher zur linearen Algebra, z.B. [Fis10, Art91].

Für $\mu = \nu = 0$ folgt

$$(\Lambda^0_0)^2 - [(\Lambda^1_0)^2 + (\Lambda^2_0)^2 + (\Lambda^3_0)^2] = 1 \Rightarrow (\Lambda^0_0)^2 \geq 1. \quad (2.6.32)$$

Das bedeutet aber, daß für eine Lorentz-Matrix entweder $\Lambda^0_0 \geq 1$ oder $\Lambda^0_0 \leq -1$ ist. Für eine Lorentz-Matrix die stetig aus der Identität hervorgeht, wobei die Matrizen alle Lorentz-Matrizen sind, muß demnach $\Lambda^0_0 \geq 1$ sein, d.h. z.B. die Matrix $\text{diag}(-1, -1, -1, -1) \in \text{SO}(1, 3)$ kann nicht stetig mit der Identität zusammenhängen, weil sonst Λ^0_0 von 1 auf einen negativen Wert springen müßte und damit die Stetigkeit der Abbildung verletzt wäre.

Wir bemerken, daß auch die Lorentz-Transformationen mit $\Lambda^0_0 \geq 1$ eine Untergruppe der Lorentz-Transformationen bilden, die man als **orthochrone Lorentz-Gruppe** $\text{O}(1, 3)^\uparrow$ bezeichnet. Der Name rührt daher, daß solche Lorentz-Transformationen die Zeitordnung von zeit- oder lichtartig zueinander gelegenen Ereignissen ungeändert lassen, d.h. für einen Vierer-Vektor x mit $x_\mu x^\mu \geq 0$ und $x^0 > 0$ ist auch $x' = \hat{\Lambda}x$, wenn $\hat{\Lambda} \in \text{O}(1, 3)^\uparrow$. Schließlich sei noch erwähnt, daß die **eigentlich orthochronen Lorentz-Transformationen**, also die orthochronen Lorentz-Transformationen mit Determinante +1, die $\text{SO}(1, 3)^\uparrow$ eine Untergruppe der Lorentz-Gruppe bildet und die stetig mit der Identität zusammenhängende Untergruppe der $\text{O}(1, 3)$ ist.

Formuliert man die Forderung der **Lorentz-Invarianz der Naturgesetze** strikt, muß man nur Invarianz oder Kovarianz bzgl. der eigentlich orthochronen **Lorentz-Transformationen** beachten, denn nur die Zusammenhangskomponente mit der Identität entspricht tatsächlich realisierbaren Transformationen zwischen den Raumzeitkoordinaten inertialer Beobachter. In der Tat zeigt sich in der Hochenergiephysik, daß die Naturgesetze nur unter dieser Untergruppe kovariant sind. Die **schwache Wechselwirkung** verletzt sowohl die Symmetrie unter **räumlichen Spiegelungen**, die durch die Matrix $\hat{P} = \text{diag}(1, -1, -1, -1) \in \text{O}(1, 3)^\uparrow$ (aber $\hat{P} \notin \text{SO}(1, 3)^\uparrow$, weil zwar $P^0_0 = 1$ aber $\det(\hat{P}) = -1$ ist) repräsentiert wird als auch die Symmetrie unter **Zeitumkehrtransformationen**, die durch die Matrix $\hat{T} = \text{diag}(-1, 1, 1, 1) \in \text{O}(1, 3)$ repräsentiert wird, und diese Matrix ist in der Tat wegen $T^0_0 = -1$ nicht orthochron und auch keine spezielle Lorentz-Transformation, da $\det(\hat{T}) = -1$ ist. Die schwache Wechselwirkung verletzt auch die Symmetrie unter der „großen Raumzeit-Spiegelung“ $\hat{T}\hat{P} = \hat{P}\hat{T} = \text{diag}(-1, -1, -1, -1) \in \text{SO}(1, 3)$. Diese Transformation besitzt zwar die Determinante +1 ist aber wegen $(\hat{T}\hat{P})^0_0 = -1$ nicht orthochron.

2.7 Speziell relativistische Mechanik

In diesem Abschnitt wenden wir uns der **Dynamik von Punktteilchen** im Rahmen der speziellen Relativitätstheorie zu. Wie in der Newtonschen Mechanik ist auch hier die Bewegung eines Teilchens in einem gegebenen Feld (z.B. eines geladenen Teilchens in einem elektromagnetischen Feld) das einfachste Beispiel. Allerdings wird dieses Beispiel für den relativistischen Fall weitaus bedeutender als in der Newtonschen Mechanik, denn es ist offensichtlich, daß schon die Wechselwirkung zweier Punktteilchen nicht mehr wie in der Newtonschen Mechanik als eine instantane Fernwirkung erfolgen kann, denn wie wir oben gesehen haben, impliziert die reativistische Raum-Zeit-Struktur, daß sich keine Signale mit einer schnelleren als der Lichtgeschwindigkeit ausbreiten können, denn wäre eine solche Wirkung gleichzeitig in einem Bezugssystem A wäre sie dies nicht mehr in einem relativ dazu bewegten Bezugssystem B . Da die beiden Ereignisse, die eine solche instantane Fernwirkung charakterisieren raumartig zueinander liegen, finden sie je nach Boost bzgl. B in unterschiedlicher zeitlicher Reihenfolge statt, d.h. es kann sich nicht um einen *kausalen* Vorgang handeln, wie wir es für eine Wechselwirkung verlangen.

Der Ausweg aus diesem Dilemma ist es, die Kraftwirkungen zwischen Punktteilchens über **Felder** zu vermitteln. Das Paradebeispiel ist die elektromagnetische Wechselwirkung zwischen geladenen Teilchen. Im Feldbild erfolgt die Wechselwirkung dadurch, daß aufgrund des einen Teilchens ein elektromagnetisches Feld, das sich (maximal mit Lichtgeschwindigkeit!) ausbreitet, erzeugt wird und auf das andere Teilchen aufgrund dieses Feldes an seinem Ort, also **lokal**, die entsprechende Kraft wirkt. In dieser Vorlesung werden wir dieses

schwierige Problem allerdings nicht behandeln sondern uns auf die Bewegung von Teilchen in vorgegebenen Feldern beschränken.

Wir wollen die relativistischen Bewegungsgesetze plausibel machen, in dem wir zunächst mit allein Lorentz-kovarianten Größen, also Skalaren und Vierervektoren arbeiten. Wie schon in den vorigen beiden Abschnitten charakterisieren wir daher zunächst die Bewegung eines Teilchens durch seine **Weltlinie im Minkowski-Raum** $x(\lambda)$, wobei λ ein beliebiger Parameter ist.

Wir für alle auf Symmetrien basierende Betrachtungen ist weiter hier das **Hamiltonsche Prinzip** sehr hilfreich. Betrachten wir also zunächst ein freies Teilchen. Die Wirkung sollte einerseits nur von der Dreiergeschwindigkeit abhängen, damit die Wirkung translationsinvariant ist aber andererseits auch ein Lorentz-Skalar sein⁵. Der einzige Lorentz-Skalar den wir für eine vorgegebene Weltlinie ist durch die **Eigenzeit**

$$ds = c d\tau = d\lambda \sqrt{\frac{dx}{d\lambda} \cdot \frac{dx}{d\lambda}} = dt \sqrt{\eta_{\mu\nu} \frac{dx^\mu}{dt} \frac{dx^\nu}{dt}} \quad (2.7.1)$$

gegeben.

In der Newtonschen Mechanik wurde nun der **Impuls** eines Teilchens durch $\vec{p} = m \dot{\vec{x}}$ definiert. Dies legt es nahe, den **Vierer-Impulsvektor** durch

$$p = m \frac{dx}{d\tau} \quad (2.7.2)$$

zu definieren. Da nämlich gemäß (2.7.1) $d\tau$, also das **Eigenzeitinkrement** ein Lorentz-Skalar ist, ist demnach (2.7.2) ein Lorentz-Vektor, wobei wir annehmen, daß die **Masse** m ein Lorentz-Skalar ist⁶.

Betrachten wir nun die Komponenten des Viererimpulses in einem beliebigen Inertialsystem:

$$(p^\mu) = \left(m \frac{dx^\mu}{d\tau} \right) = \left(m \frac{dx^\mu}{dt} \frac{dt}{d\tau} \right). \quad (2.7.3)$$

Nun gilt wegen der Definition (2.7.1) des Eigenzeitinkrements

$$\frac{dt}{d\tau} = \left(\frac{d\tau}{dt} \right)^{-1} = \frac{1}{\sqrt{1 - \dot{\vec{x}}^2/c^2}} = \frac{1}{\gamma} \quad (2.7.4)$$

und folglich

$$(p^\mu) = m\gamma \begin{pmatrix} c \\ \dot{\vec{x}} \end{pmatrix}. \quad (2.7.5)$$

Betrachten wir nun den nichtrelativistischen Grenzfall, d.h. $|\dot{\vec{x}}|/c \ll 1$. Dann ist $\gamma = 1 + \beta^2/2 + \mathcal{O}(\beta^4)$, und die räumlichen Komponenten von (2.7.5) sind in der Tat

$$\vec{p} = m\gamma \dot{\vec{x}} = m\dot{\vec{x}} [1 + \mathcal{O}(\beta^2)] \simeq m\dot{\vec{x}}, \quad (2.7.6)$$

d.h. der gewöhnliche nichtrelativistische Impuls. Die zeitliche Komponente ist

$$p^0 = mc\gamma = mc \left[1 + \frac{\beta^2}{2} + \mathcal{O}(\beta^4) \right] \simeq \frac{1}{c} \left(mc^2 + \frac{m}{2} \dot{\vec{x}}^2 \right). \quad (2.7.7)$$

⁵Streng genommen genügt es, daß die Variations der Wirkung ein Lorentz-Skalar ist, aber dieser allgemeinere Fall wird wieder dadurch abgedeckt, daß wir zu der hier hergeleiteten Lagrangefunktion eine beliebige totale Zeitableitung $d/dt \Omega(t, \vec{x})$ addieren.

⁶In einigen älteren Lehrbüchern findet man oft auch die Definition einer sog. relativistischen Masse $m_{\text{rel}} = m\gamma$. Dies vermeiden wir in diesem Manuskript strikt, da diese Definition immer wieder zu Mißverständnissen Anlaß gibt. Wir verwenden ausschließlich die als Skalar definierte Masse m , die oft auch als **invariante Masse** bezeichnet wird, um sie klar von der relativistischen Masse m_{rel} zu unterscheiden.

2. Spezielle Relativitätstheorie

In der Klammer entsteht also bis auf eine additive Konstante $E_0 = mc^2$ die **kinetische Energie** des Teilchens wie sie in der Newtonschen Mechanik definiert ist. Es liegt daher nahe,

$$\mathcal{E} = p^0 c = mc^2 \gamma = E_0 + E_{\text{kin}}, \quad \text{mit} \quad E_0 = mc^2, \quad E_{\text{kin}} = mc^2(\gamma - 1) \quad (2.7.8)$$

als die **relativistische Energie** des Teilchens zu definieren⁷. Die räumliche Komponente des Viererimpulses ist dann $p^0 = \mathcal{E}/c$. Wir bemerken, daß es in der relativistischen Mechanik sinnvoll ist, die Ruheenergie in die Definition der Energie eines Teilchens einzubeziehen, weil dann die Energie (bis auf einen Faktor) die Zeitkomponente des Impulsvierervektors $(p^\mu) = (\mathcal{E}/c, \vec{p})$ ist und damit leicht die Transformationseigenschaften von Energie und Impuls unter Lorentz-Transformation folgen, nämlich als Komponenten eines Vierervektors:

$$\bar{p}^\mu = \Lambda^\mu_\nu p^\nu. \quad (2.7.9)$$

Mit dem Viererimpuls können wir nun den Lorentz-Skalar $p \cdot p$ bilden. Mit (2.7.5) folgt

$$p \cdot p = p_\mu p^\mu = \left(\frac{\mathcal{E}}{c}\right)^2 - \vec{p}^2 = m^2 \gamma^2 c^2 (1 - \beta^2) = m^2 c^2. \quad (2.7.10)$$

Daraus ergibt sich sofort die Beziehung zwischen Energie und Impuls zu

$$\mathcal{E} = c \sqrt{m^2 c^2 + \vec{p}^2}. \quad (2.7.11)$$

Im folgenden ist es auch bequem, die dimensionslose **Vierergeschwindigkeit**

$$\mathbf{u} = \frac{1}{c} \frac{d\mathbf{x}}{d\tau} \quad (2.7.12)$$

einzuführen. Dann ist (*nachrechnen!*)

$$\mathbf{p} = mc \mathbf{u} \quad \text{und} \quad \mathbf{u} \cdot \mathbf{u} = 1. \quad (2.7.13)$$

Wir bemerken noch, daß wir aus (2.7.12) die Dreiergeschwindigkeit in einem gegebenen Inertialsystem durch die Vierergeschwindigkeit bzw. den Viererimpuls gemäß

$$(u^\mu) = \left(\frac{dt/d\tau}{1/c \cdot d\vec{x}/d\tau} \right) \Rightarrow \vec{\beta} = \frac{1}{c} \frac{d\vec{x}}{dt} = \frac{\vec{u}}{u^0} \Rightarrow \vec{\beta} = \frac{\vec{p}}{\mathcal{E}} \quad (2.7.14)$$

ausdrücken können.

Jetzt können wir auch Newtons zweites Postulat relativistisch erweitern, indem wir schreiben

$$\frac{d}{d\tau} p^\mu = mc \frac{d}{d\tau} u^\mu = m \frac{d^2}{d\tau^2} x^\mu = K^\mu, \quad (2.7.15)$$

wobei die räumlichen Komponenten \vec{K} der Kraft entsprechen. Dabei ist allerdings zu beachten, daß wir auf der linken Seite nach der **Eigenzeit des Teilchens** τ abgeleitet haben und nicht nach der Zeit t eines Inertialsystems, so daß wieder eine Vierervektorgleichung entsteht. Man nennt den Vierervektor \mathbf{K} die **Minkowski-Kraft**. Die vier Komponenten sind aber nicht unabhängig voneinander. Aufgrund der Definition der Vierergeschwindigkeit (2.7.12) gilt nämlich $\mathbf{u} \cdot \mathbf{u} = 1$. Leitet man diese Bedingung nach τ ab, folgt

$$\frac{d\mathbf{u}}{d\tau} \cdot \mathbf{u} = 0, \quad (2.7.16)$$

⁷Wir bezeichnen die Energie mit dem kalligraphischen \mathcal{E} , um Verwechslungen mit dem elektrischen Feld, das bei der Bewegung geladener Teilchen eingeführt wird (s.u.), zu vermeiden.

d.h. multipliziert man (2.7.15) mit u^μ , erhält man die Bedingung

$$\mathbf{K} \cdot \mathbf{u} = \eta_{\mu\nu} K^\mu u^\nu = 0 \quad (2.7.17)$$

für die Minkowskikraft. Damit läßt sich auch die Zeitkomponente der Bewegungsgleichung (2.7.15) interpretieren, denn aus (2.7.17) folgt

$$\eta_{\mu\nu} K^\mu u^\nu = u^0 K^0 - \vec{u} \cdot \vec{k} = 0 \Rightarrow u^0 K^0 = \vec{u} \cdot \vec{k}. \quad (2.7.18)$$

Damit folgt für die Zeitkomponente von (2.7.15)

$$m c \frac{d u^0}{d \tau} = K^0 = \frac{\vec{u}}{u^0} \cdot \vec{K} = \vec{\beta} \cdot \vec{K}. \quad (2.7.19)$$

Wegen $d\tau = \sqrt{1 - \vec{\beta}^2} dt = dt / \gamma = dt / u^0$ erhalten wir schließlich

$$m c \frac{d u^0}{d t} = \vec{\beta} \cdot \frac{\vec{K}}{\gamma} \quad (2.7.20)$$

Multiplizieren wir dies mit c und schreiben $\vec{K} = u^0 \vec{F} = \gamma \vec{F}$ erhalten wir wegen (2.7.8)

$$\frac{d \mathcal{E}}{d t} = \frac{d \mathcal{E}_{\text{kin}}}{d t} = \vec{v} \cdot \vec{F}. \quad (2.7.21)$$

Für den räumlichen Teil von (2.7.15) erhalten wir die nicht manifest kovariante Form der Bewegungsgleichung

$$\frac{d \vec{p}}{d t} = \frac{1}{\gamma} \frac{d \vec{p}}{d \tau} = \frac{1}{\gamma} \vec{K} = \vec{F}, \quad (2.7.22)$$

wobei \vec{F} die Kraft im Sinne der Newtonschen Mechanik in dem willkürlich gewählten Inertialsystem ist, wobei dies aber nicht die räumlichen Komponenten eines Vierervektors darstellen. Für Betrachtungen zum Transformationsverhalten von Kräften ist es daher stets einfacher, die Minkowski-Viererkraft K^μ zu betrachten. Jedenfalls erweist sich die Zeitkomponente der Bewegungsgleichung (2.7.15), geschrieben in der Form (2.7.21) einfach als die übliche Form des Energie-Arbeits-Theorems in genauer Analogie zur Newtonschen Mechanik.

2.8 Zerfall und Stöße von Teilchen

Besonders wichtig werden Energie- und Impuls und deren Zusammenfassung zum Viererimpulsvektor in der relativistischen **Kern- und Elementarteilchenphysik**, wo man Zerfälle von und Stoßprozesse zwischen Teilchen betrachtet. Dabei kommt es leicht auch zur Erzeugung bzw. Vernichtung neuer Teilchen. Um solche Prozesse genauer zu verstehen, benötigt man eigentlich die **relativistische Quantentheorie**. Allerdings gelten auch dort stets **Energie- und Impulserhaltung** und die Energie-Impulsbeziehung (2.7.11). Daher können wir Energie und Impuls der Endprodukte eines Zerfalls oder Stoßes aufgrund der Erhaltungssätze bestimmen. Wir betrachten nur die einfachsten Fälle des **Zweiteilchenzerfalls** und ($2 \rightarrow 2$)-**Stöße**.

2.8.1 Zweiteilchenzerfall

Wir betrachten den Zerfall eines Teilchens mit der Masse M in zwei Teilchen mit Massen m_1 und m_2 . Um die Zerfallskinetik zu charakterisieren, ist es geschickt, möglichst mit **invarianten** Größen zu arbeiten. Diese

2. Spezielle Relativitätstheorie

Größen lassen sich, wie wir gleich sehen werden, besonders einfach in dem Bezugssystem interpretieren, in dem das Teilchen vor dem Zerfall ruht.

Zunächst gilt die Erhaltung von Energie und Impuls, d.h. mit den Viererimpulsvektoren \mathbf{p} für das zerfallende Teilchen mit \mathbf{p}'_1 und \mathbf{p}'_2 , d.h.

$$\mathbf{p} = \mathbf{p}'_1 + \mathbf{p}'_2. \quad (2.8.1)$$

Weiter gelten die Energie-Impulsbeziehungen für jeden dieser Viererimpulse. Diese können wir am einfachsten in der kovarianten Form (2.7.10) berücksichtigen:

$$\mathbf{p} \cdot \mathbf{p} = M^2 c^2, \quad \mathbf{p}'_1 \cdot \mathbf{p}'_1 = m_1^2 c^2, \quad \mathbf{p}'_2 \cdot \mathbf{p}'_2 = m_2^2 c^2. \quad (2.8.2)$$

Gehen wir nun in das Ruhssystem des zerfallenden Teilchens, ist

$$(\mathbf{p}^\mu) = \begin{pmatrix} Mc \\ 0 \end{pmatrix}, \quad (\mathbf{p}'_1{}^\mu) = \begin{pmatrix} \mathcal{E}_1/c \\ \vec{p}'_1 \end{pmatrix}, \quad (\mathbf{p}'_2{}^\mu) = \begin{pmatrix} \mathcal{E}_2/c \\ -\vec{p}'_1 \end{pmatrix}. \quad (2.8.3)$$

Offenbar ist durch diese Schreibweise der räumliche Teil von (2.8.1), also die Impulserhaltung, identisch erfüllt. Die Zeitkomponente ergibt

$$\mathcal{E}_1 + \mathcal{E}_2 = Mc^2. \quad (2.8.4)$$

Wir wollen nun $|\vec{p}'_1|$ bestimmen. Wegen der Energie-Impuls-Beziehungen (2.8.2) und der Energieerhaltungsgleichung (2.8.4) ist dies nämlich der einzige freie Parameter. Dzu müssen wir aber nur (2.8.1) im Sinne des Minkowski-Produktes quadrieren. Es folgt dann

$$\mathbf{p} \cdot \mathbf{p} = M^2 c^2 = \mathbf{p}_1 \cdot \mathbf{p}'_1 + \mathbf{p}'_2 \cdot \mathbf{p}'_2 + 2\mathbf{p}'_1 \cdot \mathbf{p}'_2 = (m_1^2 + m_2^2)c^2 + 2\mathbf{p}'_1 \cdot \mathbf{p}'_2. \quad (2.8.5)$$

Im obigen Bezugssystem, wo das zerfallende Teilchen vor dem Zerfall ruht, gilt

$$\mathbf{p}'_1 \cdot \mathbf{p}'_2 = \frac{\mathcal{E}_1 \mathcal{E}_2}{c^2} + \vec{p}'_1{}^2. \quad (2.8.6)$$

Mit (2.8.3) folgt

$$\mathbf{p}'_1 \cdot \mathbf{p}'_2 = \frac{\mathcal{E}_1(Mc^2 - \mathcal{E}_1)}{c^2} + \vec{p}'_1{}^2 = M\mathcal{E}_1 - (m_1^2 c^2 + \vec{p}'_1{}^2) + \vec{p}'_1{}^2 = M\mathcal{E}_1 - m_1^2 c^2. \quad (2.8.7)$$

Setzen wir dies in (2.8.5) ein, erhalten wir

$$M^2 c^2 = (m_1^2 + m_2^2)c^2 + 2(M\mathcal{E}_1 - m_1^2 c^2) = (m_2^2 - m_1^2)c^2 + 2M\mathcal{E}_1 \quad (2.8.8)$$

bzw.

$$\mathcal{E}_1 = \frac{(M^2 + m_1^2 - m_2^2)c^2}{2M} \quad (2.8.9)$$

Mit der Energie-Impulsbeziehung erhalten wir daraus schließlich (*nachrechnen!*)

$$\vec{p}'_1{}^2 = \frac{\mathcal{E}_1^2}{c^2} - m_1^2 c^2 = \frac{(M^2 + m_1^2 - m_2^2)^2 c^2}{4M^2} - m_1^2 c^2 = \frac{[M^2 - (m_1 + m_2)^2][M^2 - (m_1 - m_2)^2] c^2}{4M^2}. \quad (2.8.10)$$

Da $\vec{p}'_1{}^2 \geq 0$ sein muß, kann der Zerfall nur unter der Bedingung stattfinden, daß

$$M \geq m_1 + m_2 \quad (2.8.11)$$

ist. Mit (2.8.11) ist die Kinematik des Zerfalls in dem betrachteten Inertialsystem, in dem das Teilchen anfangs ruht, vollständig bestimmt.

2.8.2 Zweiteilchenstöße

Als nächstes betrachten wir den Stoß zweier Teilchen mit Massen m_1 und m_2 zu zwei Teilchen mit Massen m_3 und m_4 . Die Viererimpulse der Teilchen vor dem Stoß seien entsprechend \boldsymbol{p}_1 und \boldsymbol{p}_2 und nach dem Stoß \boldsymbol{p}_3 und \boldsymbol{p}_4 . Es gilt beim Stoß wieder die Energie- und Impulserhaltung, d.h.

$$\boldsymbol{p}_1 + \boldsymbol{p}_2 = \boldsymbol{p}_3 + \boldsymbol{p}_4. \quad (2.8.12)$$

Wir können nun aus diesen vier Impulsen drei Invarianten, die **Mandelstam-Variablen**

$$\begin{aligned} s &= (\boldsymbol{p}_1 + \boldsymbol{p}_2)^2 = (\boldsymbol{p}_3 + \boldsymbol{p}_4)^2, \\ t &= (\boldsymbol{p}_1 - \boldsymbol{p}_3)^2 = (\boldsymbol{p}_2 - \boldsymbol{p}_4)^2, \\ u &= (\boldsymbol{p}_1 - \boldsymbol{p}_4)^2 = (\boldsymbol{p}_2 - \boldsymbol{p}_3)^2 \end{aligned} \quad (2.8.13)$$

bilden. Diese sind aber nicht voneinander unabhängig. Multipliziert man die Minkowskiprodukte aus und benutzt die Energie-Impulsbeziehungen $p_j^2 = m_j^2 c^2$ sowie (2.8.12), erhält man (*nachrechnen!*)

$$s + t + u = m_1^2 + m_2^2 + m_3^2 + m_4^2. \quad (2.8.14)$$

Es sind nun zwei Bezugssysteme besonders ausgezeichnet, und zwar das **Schwerpunktssystem**, wo $\vec{p}_1^{(\text{SPS})} = -\vec{p}_2^{(\text{SPS})}$ gilt sowie das sog. **Laborsystem**, in dem im Ausgangszustand ein Teilchen ruht, also z.B. $\vec{p}_2^{(\text{LS})} = 0$ gilt. Dabei geht die Bezeichnung Laborsystem auf die Anfangszeit der Teilchenbeschleuniger zurück, in der man gewöhnlich einen Strahl von Teilchen (wie Elektronen oder Protonen) auf ein ruhendes Ziel (z.B. Metallfolien) gerichtet hat.

Wir bemerken zuerst, daß es für massive Teilchen stets sowohl ein Schwerpunkts- als auch ein Laborsystem gibt. Für das Laborsystem ist dies besonders leicht zu erkennen, denn es ist \boldsymbol{p}_2 wegen $p_2^2 = m_2^2 > 0$ ein zeitartiger Vektor. Ein Lorentz-Boost mit $\vec{\beta} = \vec{p}_2 / p_2^0 = \vec{p}_2 / \sqrt{m_2^2 c^2 + \vec{p}_2^2}$, woraus sich offensichtlich $|\vec{\beta}| < 1$ ergibt, liefert dann (*nachrechnen!*) in der Tat $(\vec{p}_2^\mu) = (m_2 c, 0, 0, 0)$.

Ebenso können wir für $\boldsymbol{p}_1 + \boldsymbol{p}_2$ argumentieren, wenn wir zeigen können, daß dieser Vektor ebenfalls zeitartig ist, d.h. $s > 0$ gilt. Dies ist aber sicher der Fall, denn es ist wegen der Dreiecksungleichung für die gewöhnlichen Euklidischen Beträge der Dreierimpulse

$$p_1^0 + p_2^0 = \sqrt{m_1^2 c^2 + \vec{p}_1^2} + \sqrt{m_2^2 c^2 + \vec{p}_2^2} > |\vec{p}_1| + |\vec{p}_2| \geq |\vec{p}_1 + \vec{p}_2| \Rightarrow s = (p_1^0 + p_2^0)^2 - (\vec{p}_1 + \vec{p}_2)^2 > 0. \quad (2.8.15)$$

Wir müssen also nur mit einem Boost mit der Geschwindigkeit $\vec{\beta} = (\vec{p}_1 + \vec{p}_2) / (p_1^0 + p_2^0)$ boosten, um $\vec{p}_1 + \vec{p}_2 = 0$ zu erreichen.

2.8.3 Kinematik im Schwerpunktssystem

Wir beginnen mit der näheren Analyse der Kinematik im **Schwerpunktssystem**. Dann gilt

$$(p_1^\mu) = \begin{pmatrix} \mathcal{E}_1/c \\ \vec{p} \end{pmatrix}, \quad (p_2^\mu) = \begin{pmatrix} \mathcal{E}_2/c \\ -\vec{p} \end{pmatrix}, \quad (p_3^\mu) = \begin{pmatrix} \mathcal{E}_3/c \\ \vec{p}' \end{pmatrix}, \quad (p_4^\mu) = \begin{pmatrix} \mathcal{E}_4/c \\ \vec{p}' \end{pmatrix}, \quad (2.8.16)$$

und folglich ergibt sich aus (2.8.13)

$$\sqrt{s} = \frac{\mathcal{E}_1 + \mathcal{E}_2}{c} = \frac{\mathcal{E}_3 + \mathcal{E}_4}{c}, \quad (2.8.17)$$

d.h. $c\sqrt{s}$ ist die Gesamtenergie der Teilchen vor und nach dem Stoß.

2. Spezielle Relativitätstheorie

Wir können nun die Energien und Impulsbeträge im Schwerpunktsystem mit Hilfe der Mandelstamvariablen ausdrücken. Zunächst ist

$$\begin{aligned}
 s &= (\mathbf{p}_1 + \mathbf{p}_2)^2 = (m_1^2 + m_2^2)c^2 + 2\left(\frac{\mathcal{E}_1\mathcal{E}_2}{c^2} + \vec{p}^2\right) \\
 &\stackrel{(2.8.17)}{=} (m_1^2 + m_2^2)c^2 + 2\left(\frac{\mathcal{E}_1(c\sqrt{s} - \mathcal{E}_1)}{c^2} + \vec{p}^2\right) \\
 &= (m_2^2 - m_1^2)c^2 + \frac{2\mathcal{E}_1\sqrt{s}}{c}.
 \end{aligned} \tag{2.8.18}$$

Dabei haben wir im letzten Schritt die Energie-Impulsbeziehung $\mathcal{E}_1^2/c^2 - \vec{p}^2 = m_1^2c^2$ verwendet. Aus (2.8.18) erhält man schließlich

$$\frac{\mathcal{E}_1}{c} = \frac{s + (m_1^2 - m_2^2)c^2}{2\sqrt{s}}. \tag{2.8.19}$$

Eliminieren wir in (2.8.18) \mathcal{E}_1 zugunsten von \mathcal{E}_2 , erhalten wir

$$\frac{\mathcal{E}_2}{c} = \frac{s + (m_2^2 - m_1^2)c^2}{2\sqrt{s}}. \tag{2.8.20}$$

Die Energien im Endzustand ergeben sich durch Ersetzen der Indizes 1 und 2 durch 3 bzw. 4:

$$\begin{aligned}
 \frac{\mathcal{E}_3}{c} &= \frac{s + (m_3^2 - m_4^2)c^2}{2\sqrt{s}}, \\
 \frac{\mathcal{E}_4}{c} &= \frac{s + (m_4^2 - m_3^2)c^2}{2\sqrt{s}}.
 \end{aligned} \tag{2.8.21}$$

Die Impulsbeträge im Anfangs- bzw. Endzustand ergeben sich daraus mit Hilfe der Energie-Impuls-Beziehung nach einigen Umformungen (*nachrechnen*)

$$\begin{aligned}
 \vec{p}^2 &= \frac{\mathcal{E}_1^2}{c^2} - m_1^2c^2 = \frac{1}{4s}[s - (m_1 + m_2)^2c^2][s - (m_1 - m_2)^2c^2], \\
 \vec{p}'^2 &= \frac{\mathcal{E}_3^2}{c^2} - m_3^2c^2 = \frac{1}{4s}[s - (m_3 + m_4)^2c^2][s - (m_3 - m_4)^2c^2]
 \end{aligned} \tag{2.8.22}$$

2.8.4 Kinematik im Laborsystem

Im Laborsystem haben wir

$$p_1 = \begin{pmatrix} \mathcal{E}_1/c \\ \vec{p}_1 \end{pmatrix}, \quad p_2 = \begin{pmatrix} m_2c \\ 0 \end{pmatrix}. \tag{2.8.23}$$

Damit ist

$$\begin{aligned}
 s &= \left(\frac{\mathcal{E}_1 + m_2c^2}{c}\right)^2 - \vec{p}_1^2 \\
 &= \left(\frac{\mathcal{E}_1 + m_2c^2}{c}\right)^2 + m_1^2c^2 - \frac{\mathcal{E}_1^2}{c^2}
 \end{aligned} \tag{2.8.24}$$

Nach einigen Umformungen erhält man

$$\mathcal{E}_1 = \frac{s - (m_1^2 + m_2^2)c^2}{2m_2}. \tag{2.8.25}$$

Weiter folgt (*nachrechnen!*)

$$\vec{p}_1^2 = \frac{\mathcal{E}_1^2}{c^2} - m_1^2 c^2 = \frac{[s - (m_1 + m_2)^2 c^2][s - (m_1 - m_2)^2 c^2]}{4m_2^2 c^2}, \quad (2.8.26)$$

und der Vergleich mit (2.8.22) liefert

$$\vec{p}_{\text{SPS}}^2 = \frac{m_2^2 c^2}{s} \vec{p}_{\text{LS}}^2 \quad (2.8.27)$$

Für die Mandelstam-Variable t gilt (*nachrechnen!*)

$$t = (p_2 - p_4)^2 = \left(\frac{\mathcal{E}_4}{c} - m_2 c \right)^2 - \vec{p}_4^2 = (m_4^2 + m_2^2) c^2 - 2m_2 \mathcal{E}_4 \quad (2.8.28)$$

bzw.

$$\mathcal{E}_4 = \frac{(m_2^2 + m_4^2) c^2 - t}{2m_2}. \quad (2.8.29)$$

Ebenso erhält man aus der Definition der Mandelstam-Variablen u

$$\mathcal{E}_3 = \frac{(m_2^2 + m_3)^2 c^2 - u}{2m_2} \quad (2.8.30)$$

2.9 Äquivalenz zwischen Energie und Masse

Eine weitere wichtige Anwendung der relativistischen Kinematik ist die **Äquivalenz zwischen Energie und Masse**, die in der populärwissenschaftlichen Literatur oft in der „berühmten Formel“ $E = mc^2$ zusammengefaßt wird. Wir wissen aus den obigen Betrachtungen, daß dies etwas ungenau ist. Für ein massives Teilchen ist dies lediglich die **Ruheenergie** $\mathcal{E}_0 = mc^2$, wenn wir die unter der Masse, wie üblich, die **invariante Masse** verstehen und den Energienullpunkt so legen, daß die Energie $\mathcal{E} = \mathcal{E}_0 + \mathcal{E}_{\text{kin}} = mc^2 \gamma = c \sqrt{m^2 c^2 + \vec{p}^2} = p^0 c$ die Zeitkomponente eines Vierervektors im gerade betrachteten Bezugssystem wird.

Die eigentliche Bedeutung des Prinzips der Äquivalenz zwischen Energie und Masse wird für **zusammengesetzte** Systeme klar. Wir betrachten dazu drei Beispiele.

2.9.1 Voll inelastischer Stoß

Wir betrachten den Fall eines total inelastischen Stoßes. Dazu nehmen wir an, zwei massive Körper mit Massen m_1 und m_2 stoßen zusammen und bewegen sich dann als ein Körper mit der Masse M weiter. Man kann sich z.B. ein Geschloß vorstellen, das auf einen Körper abgefeuert wird und in diesem stecken bleibt. In der nichtrelativistischen Physik ist klar, daß $M = m_1 + m_2$ ist. Wir zeigen nun, daß dies in der Relativitätstheorie nicht der Fall sein kann. Es ist klar, daß dieser inelastische Stoßvorgang kinematisch nur möglich ist, wenn Energie dissipiert wird, d.h. wenn der Gesamtkörper nach dem Stoß sich erwärmt hat, d.h. es wird kinetische Energie der Körper vor dem Stoß teilweise in die intrinsische Wärmebewegung des Gesamtkörpers nach dem Körper umgewandelt. Dabei gilt allerdings immer noch die Erhaltung des Gesamtimpulses.

Es ist wieder am einfachsten, die Kinematik im Schwerpunktssystem zu betrachten, d.h. die beiden Körper besitzen in diesem Bezugssystem vor dem Stoß entgegengesetzt gleiche Impulse $\vec{p}_2 = -\vec{p}_1 = -\vec{p}$ und nach dem Stoß liegt ein einziger ruhender Körper vor. Die Energie-Impulsbilanz dieses Stoßes ergibt dann

$$(p_1^\mu + p_2^\mu) = \begin{pmatrix} (\mathcal{E}_1 + \mathcal{E}_2)/c \\ 0 \end{pmatrix} = \begin{pmatrix} Mc \\ 0 \end{pmatrix}. \quad (2.9.1)$$

Es ist also

$$M = \frac{\mathcal{E}_1 + \mathcal{E}_2}{c^2} = m_1 + m_2 + \frac{\mathcal{E}_{1\text{kin}} + \mathcal{E}_{2\text{kin}}}{c^2}. \quad (2.9.2)$$

Da die kinetische Energie des ruhenden Teilchens im Endzustand 0 ist, bedeutet dies, daß die gesamte kinetische Energie im Anfangszustand in Wärmenergie $\Delta Q = \mathcal{E}_{1\text{kin}} + \mathcal{E}_{2\text{kin}}$ umgewandelt worden ist. Gemäß (2.9.2) ist die invariante Masse des Körpers nach dem Stoß um $\Delta Q/c^2$ größer als $m_1 + m_2$, und dies ist ein Beispiel dafür, daß die intrinsische Energie ΔQ (in diesem Fall Wärme) zur Gesamtmasse des Körpers beiträgt.

2.9.2 Zerfall eines Pions in zwei Photonen

Ein weiteres Beispiel für die Äquivalenz ist der Zerfall eines neutralen Pions (Masse $m_{\pi_0} c^2 = 134,98 \text{ MeV}$). Dabei ist MeV eine Energieeinheit, die in der Teilchenphysik besonders bequem ist. Wir können sie im nächsten Abschnitt genauer definieren. Hier genügt es zu bemerken, daß 1 eV (ein Elektronenvolt) der Änderung der kinetischen Energie eines Elektrons entspricht, das eine elektrostatische Potentialdifferenz von 1 V durchläuft. Es ist also $1 \text{ eV} = |e| \cdot 1 \text{ V}$. Mit der Elektronenladung (in SI-Einheiten) $e = -1,602 \cdot 10^{-19} \text{ C}$ folgt daraus, daß $1 \text{ eV} = 1,602 \cdot 10^{-19} \text{ J}$ entspricht. Dies zeigt, warum man in der gerne die Einheit $\text{MeV} = 10^6 \text{ eV}$ oder $\text{GeV} = 10^9 \text{ eV}$ verwendet und dabei die Massen als ihre Ruheenergien mc^2 (m : invariante Masse!) angibt.

Photonen sind strikt genommen die Quanten des elektromagnetischen Feldes, und die Vorstellung, es handle sich im „Teilchen“ eigentlich etwas irreführend, und man kann sie erst im Rahmen der relativistischen Quantenfeldtheorie wirklich vollständig korrekt definieren. Für die Energie-Impulsbilanz bei Teilchenreaktionen oder Zerfällen, in denen Photonen auftreten, können wir sie aber wie Teilchen mit der **invarianten Masse** $m_\gamma = 0$ behandeln. Ein Photon besitzt also Energie und Impuls, die in jedem Inertialsystem über die Energie-Impulsbeziehung (2.7.11) mit $m = m_\gamma = 0$ zusammenhängen

$$\mathcal{E}_\gamma = c |\vec{p}_\gamma|. \quad (2.9.3)$$

Daraus ergibt sich für die Geschwindigkeit in jedem Inertialsystem

$$|\beta_\gamma| = \frac{c |\vec{p}_\gamma|}{\mathcal{E}_\gamma} = 1. \quad (2.9.4)$$

Daraus ist ersichtlich, daß für solche „masselosen Teilchen“ wegen $\gamma \rightarrow \infty$ für $|\beta| \rightarrow 1$ eine Vierergeschwindigkeit keinen Sinn ergibt, der Zusammenhang zwischen Energie und Impuls (2.9.3) aber sehr wohl, und nur diese Beziehung benötigt man, um die Energie-Impulsbilanz auszunutzen.

Wir können direkt die Beziehung (2.8.10) im Schwerpunktssystem der Photonen (also dem Ruhssystem des zerfallenden Pions) mit $m_1 = m_2 = 0$ verwenden, um den Photonenimpuls zu

$$|\vec{p}_\gamma| = \frac{Mc}{2} \quad (2.9.5)$$

zu bestimmen. Wir haben hier den Extremfall vorliegen, daß die invarianten Massen der Zerfallsprodukte 0 sind aber die Masse im Ausgangszustand durch

$$Mc^2 = \mathcal{E}'_1 + \mathcal{E}'_2 = 2c |\vec{p}_\gamma| \quad (2.9.6)$$

bestimmt ist. In gewissem Sinne können wir sagen, daß im Zerfall des neutralen Pions in zwei Photonen die Masse vollständig in „kinetische Energie“ zweier Photonen „umgewandelt“ wird.

2.10 Geladene Teilchen im elektromagnetischen Feld

Jetzt können wir recht einfach Bewegungsgleichungen für Teilchen in äußeren Feldern postulieren. Ein wichtiges Beispiel ist die Bewegung geladener Teilchen in einem elektromagnetischen Feld. Wie im 3. Semester in der Vorlesung zur Elektrodynamik ausführlich erläutert werden wird, sind die **Maxwell-Gleichungen** des elektromagnetischen Feldes ein Paradebeispiel für eine **relativistische Feldtheorie**, und sie läßt sich entsprechend auch in vierdimensionaler Form mit Hilfe von **Minkowski-Tensoren** schreiben. Das elektromagnetische Feld wird in einem beliebigen Inertialsystem beispielsweise durch die 6 Komponenten eines antisymmetrischen Tensorfeldes 2. Stufe $F^{\mu\nu}(x^\rho) = -F^{\nu\mu}(x^\rho)$ beschrieben. Das einzige, was wir über Tensoren wissen müssen ist, daß sich diese Komponenten unter Lorentz-Transformationen gemäß

$$\bar{F}^{\mu\nu}(\bar{x}^\rho) = \Lambda^\mu_\alpha \Lambda^\nu_\beta F^{\alpha\beta}(x^\gamma) \quad \text{mit} \quad x^\gamma = (\Lambda^{-1})^\gamma_\rho \bar{x}^\rho \quad (2.10.1)$$

transformieren. Man nennt $F^{\mu\nu}$ auch den **elektromagnetischen Feldstärketensor** oder **Faraday-Tensor**.

Man kann mit solch einem antisymmetrischen Tensorfeld sehr leicht eine Minkowski-Kraft postulieren, die die Einschränkung (2.7.17) erfüllt. Da $F^{\mu\nu}$ antisymmetrisch ist, brauchen wir nur den Ansatz

$$m \frac{d^2 x^\mu}{d\tau^2} = K^\mu = q F^{\mu\nu}(x^\rho) u_\nu \quad \text{with} \quad u^\mu = \frac{1}{c} \frac{dx^\mu}{d\tau}, \quad u_\nu = \eta_{\mu\nu} u^\mu \quad (2.10.2)$$

zu machen. Überschiebt man diese Gleichung mit u_μ ergibt sich sofort (2.7.17), weil $F^{\mu\nu} u_\mu u_\nu = -F^{\nu\mu} u_\nu u_\mu$ ist. Andererseits sind die beiden vollständig kontrahierten Tensoren gleich, und damit muß $F^{\mu\nu} u_\mu u_\nu = 0$ sein, und dies entspricht gerade (2.7.17) für die Minkowski-Kraft (2.10.2). Wegen (2.10.1) transformiert sich diese Minkowski-Kraft unter Lorentztransformationen auch wie ein Vierervektor (*nachrechnen*), wie es sein muß. Um die Bewegungsgleichung (2.10.2) besser interpretieren zu können schreiben wir den Feldstärketensor mit Hilfe der üblichen elektrischen und magnetischen Feldkomponenten \vec{E} und \vec{B} :

$$(F^{\mu\nu}) = \begin{pmatrix} 0 & -E^1 & -E^2 & -E^3 \\ E^1 & 0 & -B^3 & B^2 \\ E^2 & B^3 & 0 & -B^1 \\ E^3 & -B^2 & B^1 & 0 \end{pmatrix}. \quad (2.10.3)$$

Um zu sehen, daß diese Wahl der Anordnung der Feldkomponenten einen Sinn ergibt, schreiben wir die Minkowski-Kraft (2.10.2) explizit hin

$$(K^\mu) = q F^{\mu\nu} u_\nu = q \begin{pmatrix} \vec{E} \cdot \vec{u} \\ u_0 \vec{E} + \vec{u} \times \vec{B} \end{pmatrix}. \quad (2.10.4)$$

Dies schreiben wir, wie oben im allgemeinen Fall bereits besprochen noch um in die nichtkovariante Form mit Ableitungen nach der gewöhnlichen Zeit im hier betrachteten Inertialsystem. Dazu schreiben wir die Bewegungsgleichung als System erster Ordnung, indem wir den Viererimpuls einführen:

$$\left(\frac{d}{d\tau} p^\mu \right) = m \frac{d^2}{d\tau^2} x^\mu = (K^\mu) = q \begin{pmatrix} \vec{E} \cdot \vec{u} \\ u_0 \vec{E} + \vec{u} \times \vec{B} \end{pmatrix}, \quad u^\mu = \frac{p^\mu}{mc} = \frac{1}{c} \frac{d}{d\tau} x^\mu. \quad (2.10.5)$$

Da nun $u^0 = u_0 = \gamma = (1 - \beta^2)^{-1/2}$ with $\vec{\beta} = \vec{v}/c$, $\vec{v} = \vec{u}/u^0$, ergibt sich für den räumlichen Anteil der ersten Gleichung

$$\frac{1}{u_0} \frac{d}{d\tau} \vec{p} = \frac{d}{dt} \vec{p} = q \left(\vec{E} + \frac{\vec{v}}{c} \times \vec{B} \right), \quad (2.10.6)$$

und das entspricht genau der üblichen Bewegungsgleichung für ein Teilchen in einem elektromagnetischen Feld mit den Komponenten (\vec{E}, \vec{B}) in der nichtrelativistischen Mechanik. Dabei ist q die Ladung des Teilchens

2. Spezielle Relativitätstheorie

\vec{E} die elektrischen und \vec{B} die magnetischen Feldkomponenten⁸. Für die räumlichen Koordinaten von (2.10.5) bemerken wir, daß

$$\vec{u} = \frac{1}{mc} \vec{p} = \frac{1}{c} \frac{d}{d\tau} \vec{x}. \quad (2.10.7)$$

Dividiert man dies durch $u^0 = p^0/(mc)$ erhalten wir

$$\frac{d}{dt} \vec{x} = c \frac{\vec{p}}{p^0}. \quad (2.10.8)$$

Wie wir oben bei der allgemeinen Betrachtung gesehen haben, ist die Zeitkomponente in (2.10.5) automatisch für die Lösungen der Bewegungsgleichungen (2.10.6) und (2.10.8) erfüllt, denn multiplizieren wir (2.10.6) mit \vec{v} , folgt

$$\vec{v} \cdot \frac{d}{dt} \vec{p} = q \vec{v} \cdot \vec{E}, \quad (2.10.9)$$

und dies ist genau die durch u^0 dividierte zeitliche Komponente von (2.10.5). Damit ist klar, daß die Bewegungsgleichung mit der Energie-Impuls-Beziehung (2.7.11) verträglich ist, wie es aufgrund unserer Konstruktion der Minkowski-Kraft zu erwarten ist.

2.10.1 Teilchen im homogenen elektrischen Feld

Eines der einfachsten Beispiele für die relativistische Bewegung eines geladenen Teilchens ist die Bewegung in einem homogenen elektrischen Feld, wie es (näherungsweise für einen begrenzten räumlichen Bereich) in einem **Plattenkondensator**, der auf konstanter Gleichspannung gehalten wird, realisiert werden kann. In dem Fall ist⁹ $\vec{E} = (E, 0, 0)^T = \text{const}$ und $\vec{B} = 0$. Wir betrachten als einfachsten Fall das Anfangswertproblem $\vec{p}(0) = 0$, $\vec{x}(0) = 0$, d.h. das Teilchen startet aus der Ruhe im Ursprung des räumlichen Koordinatensystems. Die Gleichung (2.10.6) läßt sich dann sehr leicht lösen

$$\frac{d}{dt} p^1 = \frac{q}{m} E \Rightarrow p^1(t) = qEt. \quad (2.10.10)$$

Die beiden anderen Komponenten des Impulses sind Erhaltungsgrößen und daher aufgrund der Anfangsbedingungen $p^2 = p^3 = 0$. Aus der Energie-Impuls-Beziehung (2.7.11) folgt für die zeitliche Komponente des Viererimpulses

$$p^0(t) = \frac{\mathcal{E}}{c} = \sqrt{m^2 c^2 + (qEt)^2} \quad (2.10.11)$$

und daher

$$v^1(t) = \frac{dx^1}{dt} = \frac{cp^1}{p^0} = \frac{qcEt}{\sqrt{m^2 c^2 + q^2 E^2 t^2}}. \quad (2.10.12)$$

Dies läßt sich unter Berücksichtigung der Anfangsbedingung leicht nochmals nach der Zeit integrieren, wodurch wir die Lösung

$$x^1(t) = \frac{mc^2}{qE} \left(\sqrt{1 + \frac{q^2 E^2 t^2}{m^2 c^2}} - 1 \right) \quad (2.10.13)$$

erhalten.

⁸Man beachte, daß hier statt der üblicherweise verwendeten Einheiten des Internationalen Einheitensystems (SI) das sog. Gaußsche Einheitensystem verwendet wird, welches für die relativistische Darstellung der Theorie etwas bequemer als die SI-Einheiten ist.

⁹Man beachte, daß hier E die 1-Komponente des elektrischen Feldes und nicht die Energie bezeichnet. Deshalb haben wir oben \mathcal{E} als Bezeichnung für die Energie gewählt.

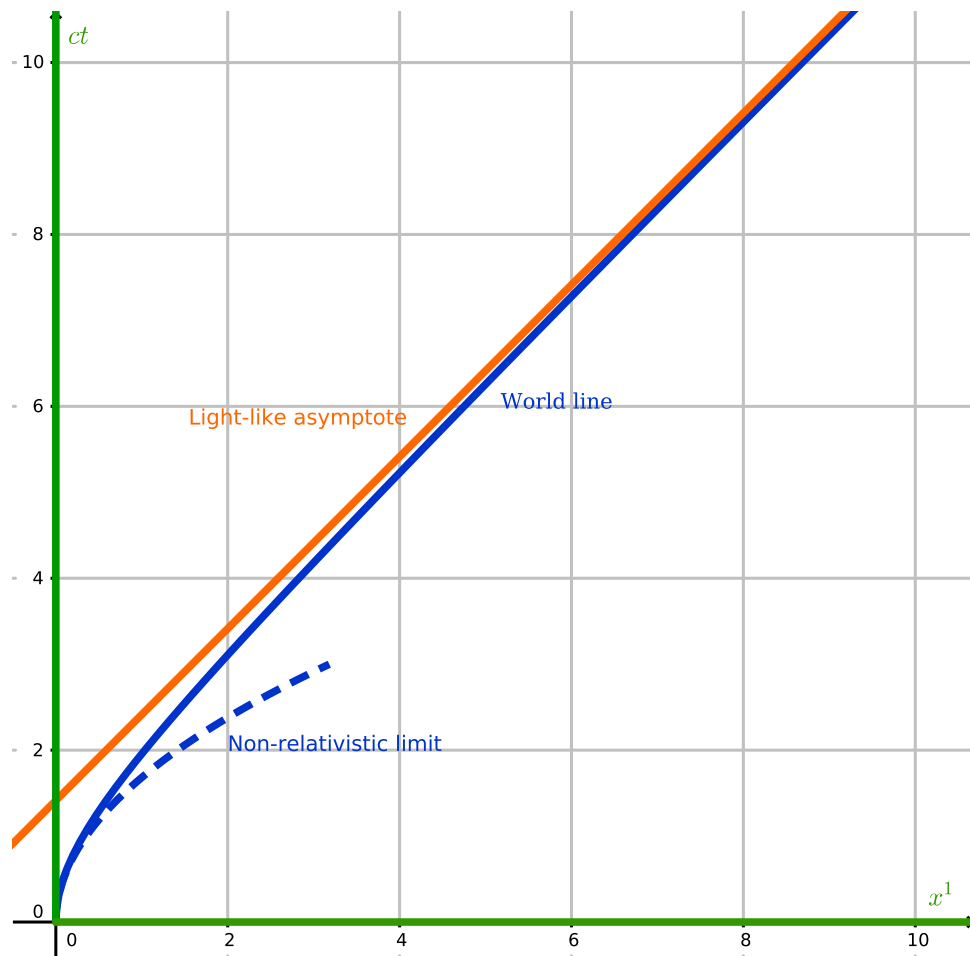


Abbildung 2.4: Minkowski-Diagramm für die „hyperbolische Bewegung“ in einem homogenen elektrischen Feld. Die entsprechende Weltlinie ist eine Hyperbel und besitzt eine lichtartige Asymptote. Dies zeigt, daß ein Lichtsignal, das vom Ursprung zu einer Zeit $t > t_H = \frac{mc}{qE}$ ausgesandt wird, das Teilchen nicht erreichen kann. In diesem Sinne ist die lichtartige Asymptote ein **Ereignishorizont** für einen mit dem Teilchen mitbeschleunigten Beobachter, denn ihn können vom Ursprung zu Zeiten $t > t_H$ ausgesandte Lichtsignale (und damit auch keine anderen Signale jedweder Art) erreichen. Man nennt diesen Horizont auch den **Rindler-Horizont** (benannt nach dem theoretischen Physiker Wolfgang Rindler).

In dieser Form läßt sich auch leicht der nichtrelativistische Limes der Lösung diskutieren. Diese gilt für $|v^1| \ll c$, also für $t \ll mc/(qE)$. Dann können wir die Wurzel in (2.10.13) entwickeln (*nachrechnen*) und erhalten

$$x^1(t) = \frac{mc^2}{qE} \left(\frac{q^2 E^2 t^2}{2m^2 c^2} + \mathcal{O} \left[\left(\frac{qEt}{mc} \right)^4 \right] \right) = \frac{qEt^2}{2m} + \dots \quad (2.10.14)$$

In führender Ordnung erhalten wir die erwartete Zeitabhängigkeit für die Bewegung in einem konstanten Kraftfeld.

Wir können auch den relativistischen Limes betrachten, wo $t \gg mc/(qE)$ ist dann folgt durch die entsprechend angepaßte Entwicklung der Wurzel

$$x^1(t) = \frac{mc^2}{qE} \left(\frac{qEt}{mc} \sqrt{1 + \frac{m^2 c^2}{q^2 E^2 t^2}} - 1 \right) \underset{t \rightarrow \infty}{\cong} c \left(t - \frac{mc}{qE} \right). \quad (2.10.15)$$

Das impliziert auch, daß ein in Sect. 2.1 diskutiertes vom Ursprung ausgesendetes Lichtsignal das Teilchen nur erreichen kann, wenn es zu einer Zeit $t < t_H$ mit $t_H = \frac{mc}{qE}$ ausgesandt wird (vgl. das Minkowskidiagramm in Abb. 2.4).

2.10.2 Teilchen in einem homogenen Magnetfeld

Jetzt betrachten wir die Bewegung in einem homogenen Magnetfeld, wie es (näherungsweise) im Inneren einer langen stromdurchflossenen Spule erzeugt werden kann. Hier ist es im Gegensatz zu unserem Vorgehen bei der Bewegung im homogenen elektrischen Feld bequemer, wenn wir von der manifest kovarianten Formel (2.10.5) ausgehen, wobei wir nur den räumlichen Teil der Bewegungsgleichung verwenden müssen, denn die Zeitkomponente der Gleichung wird ja dann automatisch erfüllt. Diese räumliche Bewegungsgleichung lautet für $\vec{E} = 0$ und $\vec{B} = B\vec{e}_3$

$$\frac{d}{d\tau} \vec{u} = \frac{q}{mc} \vec{u} \times \vec{B} = \frac{q}{mc} \begin{pmatrix} Bu^2 \\ -Bu^1 \\ 0 \end{pmatrix}. \quad (2.10.16)$$

Die 3-Komponente der Gleichung läßt sich sofort integrieren zu

$$u^3(\tau) = u_0^3 = \text{const.} \quad (2.10.17)$$

Die Bewegung in der (12)-Ebene läßt sich elegant finden, indem man die komplexe Hilfsvariable

$$\xi = u^1 + iu^2 \Rightarrow \frac{d\xi}{d\tau} = \frac{qB}{mc}(u^2 - iu^1) = -i\omega\xi \quad (2.10.18)$$

eingführt, wobei wir die **Zyklotronfrequenz** $\omega = qB/(mc)$ eingeführt haben. Diese Gleichung wird offenbar durch

$$\xi(\tau) = \xi_0 \exp(-i\omega\tau) \quad (2.10.19)$$

gelöst. Wir erhalten die Komponenten der Vierergeschwindigkeit, indem wir diese Lösung wieder in Real- und Imaginärteil aufspalten, zu

$$\begin{aligned} u^1(\tau) &= \text{Re } \xi(\tau) = u_0^1 \cos(\omega\tau) + u_0^2 \sin(\omega\tau), \\ u^2(\tau) &= \text{Im } \xi(\tau) = -u_0^1 \sin(\omega\tau) + u_0^2 \cos(\omega\tau). \end{aligned} \quad (2.10.20)$$

Die räumliche Komponente der Vierergeschwindigkeit ist wieder durch die Identität $u_\mu u^\mu = 1$, d.h.

$$u^0(\tau) = \sqrt{1 + \vec{u}(\tau)^2} = \sqrt{1 + |\xi(\tau)|^2 + (u_0^3)^2} = \sqrt{1 + \vec{u}_0^2} = \text{const} \quad (2.10.21)$$

gegeben. Dies bedeutet, daß in der Tat die Zeitkomponente der Bewegungsgleichung (2.10.5) automatisch erfüllt wird, wie es sein muß.

Durch eine weitere Integration von (2.10.20) und (2.10.18) nach der Eigenzeit τ ergibt sich schließlich die Weltlinie zu

$$[x^\mu(\tau)] = c \begin{pmatrix} \sqrt{1 + \vec{u}_0^2} \tau \\ -u_0^1/\omega \sin(\omega\tau) + u_0^2/\omega [\cos(\omega\tau) - 1] + x_0^1 \\ u_0^1/\omega [\cos(\omega\tau) - 1] + u_0^2/\omega \sin(\omega\tau) + x_0^2 \\ u_0^3 \tau + x_0^3 \end{pmatrix}. \quad (2.10.22)$$

Dies läßt sich noch einfacher interpretieren, indem wir die komplexe Form der Bewegung für die 1- und 2-Komponenten (2.10.19) betrachten. Integriert man diese Gleichung, erhält man

$$\zeta(\tau) = x^1(\tau) + ix^2(\tau) = \frac{i\xi_0 c}{\omega} [\exp(-i\omega\tau) - 1] + \zeta_0, \quad (2.10.23)$$

Und dies beschreibt einen Kreis vom Radius

$$R = \frac{c|\xi_0|}{|\omega|} = \frac{c\sqrt{(u_0^1)^2 + (u_0^2)^2}}{|\omega|}. \quad (2.10.24)$$

Man kann die Lösung (2.10.20) leicht mittels der Zeit t im hier verwendeten Inertialsystem ausdrücken, denn es ist

$$\frac{dt}{d\tau} = u^0 \Rightarrow t = \tau \sqrt{1 + \vec{u}_0^2} = \text{const.} \quad (2.10.25)$$

bzw. mit der Dreiergeschwindigkeit $\vec{\beta} = \vec{v}/c = \vec{u}/u^0$

$$u^0 = \frac{1}{\sqrt{1 - \vec{\beta}^2}} = \gamma = \frac{1}{\sqrt{1 - \vec{\beta}_0^2}} \Rightarrow \tau = t \sqrt{1 - \vec{\beta}_0^2}. \quad (2.10.26)$$

Mit $\vec{v} = c\vec{u}/u^0$ führt dies gemäß (2.10.17) und (2.10.20) zu

$$\vec{v}(t) = \begin{pmatrix} v_0^1 \cos(\omega_{\text{lab}} t) + v_0^2 \sin(\omega_{\text{lab}} t) \\ -v_0^1 \sin(\omega_{\text{lab}} t) + v_0^2 \cos(\omega_{\text{lab}} t) \\ v_0^3 \end{pmatrix} \quad \text{with} \quad \omega_{\text{lab}} = \omega \sqrt{1 - \vec{\beta}_0^2}, \quad (2.10.27)$$

und

$$[x^\mu(t)] = \begin{pmatrix} ct \\ -v_0^1/\omega_{\text{lab}} \sin(\omega_{\text{lab}} t) + v_0^2/\omega_{\text{lab}} [\cos(\omega_{\text{lab}} t) - 1] + x_0^1 \\ v_0^1/\omega_{\text{lab}} [\cos(\omega_{\text{lab}} t) - 1] + v_0^2/\omega_{\text{lab}} \sin(\omega_{\text{lab}} t) + x_0^2 \\ v_0^3 t + x_0^3 \end{pmatrix}. \quad (2.10.28)$$

Der Radius des Kreises der Projektion der Trajektorie auf die (12)-Ebene ergibt sich daraus zu

$$R = \frac{\sqrt{(v_0^1)^2 + (v_0^2)^2}}{|\omega_{\text{lab}}|} = \frac{mc \sqrt{(v_0^1)^2 + (v_0^2)^2}}{qB \sqrt{1 - (\beta_0^1)^2 + (\beta_0^2)^2}}. \quad (2.10.29)$$

2.10.3 Teilchen in parallelen homogenen elektrischen und magnetischen Feldern

Wir betrachten weiter das Beispiel der Bewegung eines massiven Teilchens in beliebigen homogenen parallelen elektrischen und magnetischen Feldern. Es sei

$$\vec{E} = E\vec{e}_3, \quad \vec{B} = B\vec{e}_3. \quad (2.10.30)$$

Wir schreiben die Bewegungsgleichungen wieder in ihrer kovarianten Form (2.10.5) hin. Für den Viererimpuls gilt

$$\frac{d}{d\tau}(p^\mu) = \frac{q}{mc} \begin{pmatrix} E p^3 \\ B p^2 \\ -B p^1 \\ E p^0 \end{pmatrix}. \quad (2.10.31)$$

Wir sehen, daß die Bewegungsgleichungen für (p^0, p^3) einerseits und für (p^1, p^2) entkoppeln, so daß wir sie getrennt voneinander lösen können. Zunächst ist

$$\frac{d}{d\tau} \begin{pmatrix} p^0 \\ p^3 \end{pmatrix} = \frac{qE}{mc} \begin{pmatrix} p^3 \\ p^0 \end{pmatrix} = \omega_{\parallel} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} p^0 \\ p^3 \end{pmatrix} = \hat{M} \begin{pmatrix} p^0 \\ p^3 \end{pmatrix}. \quad (2.10.32)$$

2. Spezielle Relativitätstheorie

Dies ist die Bewegung eines Teilchens auf der x^3 -Achse unter dem Einfluß eines elektrischen Feldes in x^3 -Richtung. Dieses Problem hatten wir zwar oben in Abschnitt 2.10.1 bereits in der nicht-kovarianten Form gelöst. Hier können wir aber die elegante Methode der Matrixexponentialfunktion aus Abschnitt 2.6 anwenden. Offenbar ist nämlich die Lösung von (2.10.32)

$$\begin{pmatrix} p^0(\tau) \\ p^3(\tau) \end{pmatrix} = \exp(\tau \hat{M}) \begin{pmatrix} p_0^0 \\ p_0^3 \end{pmatrix}, \quad (2.10.33)$$

wobei p_0^0 und p_0^3 beliebige Anfangswerte für die entsprechenden Viererimpulskomponenten sind. Man rechnet leicht direkt nach, daß

$$(\tau \hat{M})^2 = (\omega_{\parallel} \tau)^2 \mathbb{1} \Rightarrow (\tau \hat{M})^{2k} = (\omega_{\parallel} \tau)^{2k} \mathbb{1}, \quad (\tau \hat{M})^{2k+1} = (\omega_{\parallel} \tau)^k \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}. \quad (2.10.34)$$

Setzt man dies in die Exponentialreihe ein, erhält man (*nachrechnen!*)

$$\exp(\tau \hat{M}) = \cosh(\omega_{\parallel} \tau) \mathbb{1} + \sinh(\omega_{\parallel} \tau) \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} \cosh(\omega_{\parallel} \tau) & \sinh(\omega_{\parallel} \tau) \\ \sinh(\omega_{\parallel} \tau) & \cosh(\omega_{\parallel} \tau) \end{pmatrix}. \quad (2.10.35)$$

Ebenso können wir für die beiden übrigen Komponenten vorgehen. Die Bewegungsgleichungen (2.10.31) lassen sich in die Form (*nachprüfen!*)

$$\frac{d}{d\tau} \begin{pmatrix} p^1 \\ p^2 \end{pmatrix} = \omega_{\perp} \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \begin{pmatrix} p^1 \\ p^2 \end{pmatrix} \quad \text{mit} \quad \omega_{\perp} = \frac{qB}{mc} \quad (2.10.36)$$

bringen. Auch in diesem Fall läßt sich die Matrixexponentialfunktion explizit angeben (*Übung!*)

$$\begin{pmatrix} p^1(\tau) \\ p^2(\tau) \end{pmatrix} = \begin{pmatrix} \cos(\omega_{\perp} \tau) & \sin(\omega_{\perp} \tau) \\ -\sin(\omega_{\perp} \tau) & \cos(\omega_{\perp} \tau) \end{pmatrix} \begin{pmatrix} p_0^1 \\ p_0^2 \end{pmatrix}. \quad (2.10.37)$$

Dabei sind p_0^1 und p_0^2 die Anfangsimpulskomponenten bei $\tau = 0$. Wir bemerken, daß $\vec{p}_{\perp}^2 = (p^1)^2 + (p^2)^2 = (p_0^1)^2 + (p_0^2)^2 = \text{const}$ und ebenso $(p^0)^2 - (p^3)^2 = (p_0^0)^2 - (p_0^3)^2 = \text{const}$. Es ist also

$$p_{\mu} p^{\mu} \stackrel{!}{=} m^2 c^2 \Rightarrow p_0^0 = \sqrt{m^2 c^2 + \vec{p}_0^2} = \mathcal{E}_0 / c. \quad (2.10.38)$$

Wegen $dx/d\tau = p/m$ folgt die Weltlinie durch einfaches Integrieren von (2.10.33) und (2.10.37) nach τ . Wir nehmen der Einfachheit halber an, daß $\vec{p}(\tau = 0) = p_0^1 \vec{e}_1$ (was wir durch einen geeigneten Lorentz-Boost in x^3 Richtung und eine Drehung um die x^3 -Richtung stets erreichen können). Unter Berücksichtigung der Anfangsbedingungen für die Raumzeit-Koordinaten $t(\tau = 0) = 0$ und $\vec{x}(\tau = 0) = \vec{x}_0$ folgt dann

$$\begin{aligned} t &= \frac{p_0^0}{mc\omega_{\parallel}} \sinh(\omega_{\parallel} \tau), \\ \begin{pmatrix} x^1(\tau) \\ x^2(\tau) \end{pmatrix} &= \frac{p_0^1}{m\omega_{\perp}} \begin{pmatrix} \sin(\omega_{\perp} \tau) \\ \cos(\omega_{\perp} \tau) - 1 \end{pmatrix} + \begin{pmatrix} x_0^1 \\ x_0^2 \end{pmatrix}, \\ x^3(\tau) &= \frac{p_0^0}{m\omega_{\parallel} \tau} [\cosh(\omega_{\parallel} \tau) - 1] + x_0^3. \end{aligned} \quad (2.10.39)$$

Die Trajektorie im Ortsraum ist also eine Helix mit einem Kreis vom Radius $p_0^1 / (m\omega_{\text{perp}})$ als Projektion auf die $(x^1 x^2)$ -Ebene.

2.10.4 Teilchen im gekreuzten homogenen elektromagnetischen Feld

Wir betrachten nun noch einen speziellen Fall für ein homogenes elektromagnetisches Feld mit $|\vec{E}| = |\vec{B}| = E$ und $\vec{E} = E\vec{e}_2$ und $\vec{B} = E\vec{e}_3$. Im folgenden Abschnitt werden wir dann zeigen, daß wir aus den vier Beispielen in diesem Kapitel alle möglichen Bewegungen in homogenen elektromagnetischen Feldern behandelt haben, weil wir durch eine geeignete Lorentz-Transformation alle anderen Fälle auf die hier behandelten Spezialfälle abbilden können. Behandeln wir aber zunächst das oben beschriebene spezielle Problem.

Die Bewegungsgleichungen (2.10.5) lauten in diesem Fall

$$\frac{d}{d\tau}(p^\mu) = \omega \begin{pmatrix} p^2 \\ p^2 \\ p^0 - p^1 \\ 0 \end{pmatrix} \quad \text{mit} \quad \omega = \frac{qE}{mc}. \quad (2.10.40)$$

Wir könnten leicht wieder das Verfahren mit Hilfe der Matrixexponentialabbildung anwenden, um dieses Gleichungssystem zu lösen (*Übung*). Es ist aber wesentlich einfacher zu bemerken, daß aus den ersten beiden Komponenten der Gleichung

$$\frac{d}{d\tau}(p^0 - p^1) = 0 \Rightarrow p^0 - p^1 = \alpha = p_0^0 - p_0^1 = \text{const} \quad (2.10.41)$$

folgt. Für die dritte und vierte Komponente ergibt sich dann

$$p^2(\tau) = \omega\alpha\tau + p_0^2, \quad p^3(\tau) = p_0^3 = \text{const} \quad (2.10.42)$$

und daraus durch eine weitere Integration

$$p^0 = \frac{\alpha\omega}{2}\tau^2 + \omega p_0^2\tau + p_0^0, \quad p^1 = \frac{\alpha\omega}{2}\tau^2 + \omega p_0^2\tau + p_0^1. \quad (2.10.43)$$

Insgesamt ist also

$$(p^\mu) = \begin{pmatrix} \frac{\alpha\omega}{2}\tau^2 + \omega p_0^2\tau + p_0^0 \\ \frac{\alpha\omega}{2}\tau^2 + \omega p_0^2\tau + p_0^1 \\ \omega\alpha\tau + p_0^2 \\ p_0^3 \end{pmatrix}. \quad (2.10.44)$$

Man rechnet leicht nach (*Übung*), daß die Lösung, wie zu erwarten, mit der relativistischen Energie-Impuls-Beziehung kompatibel ist, denn es gilt

$$\mathbf{p} \cdot \mathbf{p} = \mathbf{p}_0 \cdot \mathbf{p}_0 = \text{const} \stackrel{!}{=} m^2 c^2, \quad (2.10.45)$$

und es ergibt sich daraus, daß nur die Anfangsbedingungen für den Dreierimpuls unabhängig vorgegeben werden können. Für die räumliche Komponente gilt

$$p_0^0 = \sqrt{m^2 c^2 + \vec{p}_0^2}. \quad (2.10.46)$$

Schließlich müssen wir nur (2.10.44) noch einmal nach τ integrieren, um die Weltlinie des Teilchens

$$(x^\mu) = \begin{pmatrix} \frac{\alpha\omega}{6m}\tau^3 + \frac{\omega p_0^2}{2m}\tau^2 + p_0^0\tau \\ \frac{\alpha\omega}{6m}\tau^3 + \frac{\omega p_0^2}{2m}\tau^2 + p_0^1\tau + x_0^1 \\ \frac{\omega\alpha}{2m}\tau^2 + p_0^2\tau + x_0^2 \\ p_0^3\tau + x_0^3 \end{pmatrix} \quad (2.10.47)$$

2.11 Relativistische geladene Teilchen und das Wirkungsprinzip

Im folgenden untersuchen wir die Frage, wie man die Bewegung geladener Teilchen in vorgegebenen elektromagnetischen Feldern mit Hilfe des Hamilton-Prinzips der kleinsten Wirkung behandeln kann. Dazu bemerken wir, daß die nichtkovariant geschriebene Bewegungsgleichung (2.10.6) proportional zu dx^μ/dt ist. Wir betrachten zunächst die nichtkovariante Formulierung in einem beliebig gewählten Inertialsystem, im sog. (3 + 1)-Formalismus und untersuchen dann die Möglichkeit eines manifest kovarianten Wirkungsprinzips, d.h. daß man die Wirkung und Bewegungsgleichungen mit Hilfe von Vierertensoren schreibt.

2.11.1 (3+1)-Formalismus

Zusammen mit der Lagrangefunktion für freie Teilchen legt dies den Ansatz

$$L = -mc^2 \sqrt{1 - \dot{\vec{x}}^2/c^2} - \frac{q}{c} \dot{x}_\mu A^\mu(x^\rho) = -mc \sqrt{\eta_{\mu\nu} \dot{x}^\mu \dot{x}^\nu} - \frac{q}{c} \dot{x}_\mu A^\mu(x^\rho), \quad \dot{x}^\mu = \frac{dx^\mu}{dt}, \quad (2.11.1)$$

nahe, wobei $A^\mu(x^\rho)$ die Komponenten eines **Vierervektorfeldes** sind. Als Vierervektorfeld bezeichnen wir dabei eine Größe deren Komponenten sich unter Lorentz-Transformationen $\bar{x}^\mu = \Lambda^\mu{}_\nu x^\nu$ gemäß

$$\bar{A}^\mu(\bar{x}^\rho) = \Lambda^\mu{}_\nu A^\nu(x^\sigma) = \Lambda^\mu{}_\nu A^\nu[(\Lambda^{-1})^\sigma{}_\rho \bar{x}^\rho] \quad (2.11.2)$$

transformieren. Daß (2.11.1) auf letztlich relativistische Bewegungsgleichungen führt, folgt daraus, daß die Wirkung

$$S[\vec{x}] = \int_{t_1}^{t_2} dt L(t, \vec{x}, \dot{\vec{x}}) \quad (2.11.3)$$

ein Lorentz-Skalar ist.

Die Bewegungsgleichungen ergeben sich mit den Euler-Lagrange-Gleichungen. Zunächst ist es wichtig zu bemerken, daß hier ein Fall vorliegt, wo die kanonischen Impulse von dem oben definierten mechanischen Impuls abweichen, denn es ist

$$\vec{p}_{\text{can}} = \frac{\partial L}{\partial \dot{\vec{x}}} = m\gamma \dot{\vec{x}} + \frac{q}{c} \vec{A} = \vec{p} + \frac{q}{c} \vec{A}. \quad (2.11.4)$$

Dabei haben wir beachtet, daß

$$\dot{x}_\mu A^\mu = A^0 - \dot{\vec{x}} \cdot \vec{A} \quad (2.11.5)$$

ist. Die Euler-Lagrange-Gleichungen lauten nun

$$\dot{\vec{p}}_{\text{can}} = \frac{\partial L}{\partial \vec{x}}. \quad (2.11.6)$$

Berechnen wir zunächst die linke Seite, d.h. wir leiten (2.11.4) nach der Zeit ab.

$$\dot{p}_{\text{can}}^k = \dot{p}^k + \frac{q}{c} (\partial_t A^k + \dot{x}^j \partial_j A^k). \quad (2.11.7)$$

Dabei definieren wir für irgendeine Funktion

$$\partial_j f(t, \vec{x}) = \frac{\partial}{\partial x^j} f(t, \vec{x}). \quad (2.11.8)$$

Wir kommen auf die Tatsache, daß hier ein unten stehender Index (entsprechend einer kovarianten Tensor-komponente) stehen muß, gleich noch zurück. Die rechte Seite von (2.11.6) ergibt

$$\frac{\partial L}{\partial x^k} = \partial_k L = -\frac{q}{c} \partial_k (A^0 - \dot{x}^i A^i \delta_{ij}) = -\frac{q}{c} (c \partial_k A^0 - \dot{x}^i \partial_k A^i \delta_{ij}). \quad (2.11.9)$$

Schreiben wir nun (2.11.7) aus und formen ein wenig um, ergibt sich als Bewegungsgleichung

$$\dot{p}^k = -q \left(\frac{1}{c} \partial_t A^k + \partial_k A^0 \right) - \frac{q}{c} \dot{x}^j (\partial_j A^k - \partial_k A^j). \quad (2.11.10)$$

Jetzt erinnern wir uns an die Vektoranalysis zurück. In kartesischen Koordinaten können wir nämlich den letzten Ausdruck mit Hilfe der **Rotation** des Dreivektorfeldes \vec{A} ausdrücken

$$(\vec{\nabla} \times \vec{A})^i = \epsilon_{ijk} \partial_j A^k. \quad (2.11.11)$$

Dabei ist ϵ_{ijk} das **Levi-Civita-Symbol**, das total antisymmetrisch unter Vertauschen seiner Argumente ist und für das $\epsilon_{123} = 1$ gilt. Umgekehrt folgt (*nachrechnen!*)

$$\partial_j A^k - \partial_k A^j = \epsilon_{ijk} (\vec{\nabla} \times \vec{A})^i. \quad (2.11.12)$$

Setzen wir dann

$$\vec{E} = -\frac{1}{c} \partial_t \vec{A} - \vec{\nabla} A^0, \quad \vec{B} = \vec{\nabla} \times \vec{A}, \quad (2.11.13)$$

können wir die Bewegungsgleichung (2.11.10) in die Form

$$\dot{\vec{p}} = q \left(\vec{E} + \frac{1}{c} \dot{\vec{x}} \times \vec{B} \right), \quad (2.11.14)$$

und das stimmt mit (2.10.6) überein.

Daß sich jedes elektromagnetische Feld mit Hilfe der Komponenten des Vierervektors gemäß (2.11.13) dargestellt werden kann, wird ausführlich in der Vorlesung zur Elektrodynamik (Theoretische Physik 3) behandelt.

2.11.2 Manifest kovariantes Hamilton-Prinzip

Die obige Herleitung ist an einigen Stellen wenig elegant, denn sie macht von der Tatsache, daß man die Bewegungsgleichungen für ein Teilchen in einem elektromagnetischen Feld auch manifest kovariant, also gemäß (2.10.2) mit Hilfe von Vierervektoren und Tensoren formulieren kann. Dazu bemerken wir, daß die Wirkung ein Lorentz-Skalar ist, und wir können einen beliebigen skalaren Parameter λ verwenden, um die Wirkung zu berechnen. Andererseits ist die Wahl dieses Parameters willkürlich und führt nicht zu einfach interpretierbaren Bewegungsgleichungen. Allerdings ist mit der **Eigenzeit** τ ein physikalischer Parameter für die Weltlinie ausgezeichnet. Allerdings ist bei der Variation zu beachten, daß in Wirklichkeit nur die drei räumlichen Koordinaten unabhängige Größen sind und die aus der Definition der Eigenzeit folgende Nebenbedingung

$$\eta_{\mu\nu} \frac{dx^\mu}{d\tau} \frac{dx^\nu}{d\tau} = c^2 \quad (2.11.15)$$

erfüllt werden muß. Wir können dies als holonome Zwangsbedingung auffassen und mit einem Lagrange-Parameter im Variationsprinzip berücksichtigen. Die Wirkung des (3 + 1)-Formalismus (2.11.3) können wir auch in der Form

$$\tilde{S}[x^\mu] = \int_{\tau_1}^{\tau_2} d\tau \left[mc \sqrt{\dot{x}^\mu \dot{x}_\mu} + \frac{q}{c} A_\mu \dot{x}^\mu \right] = \int_{\tau_1}^{\tau_2} d\tau \tilde{L}(x^\mu, \dot{x}^\mu). \quad (2.11.16)$$

schreiben, wobei wir jetzt definieren $\dot{x}^\mu = dx^\mu/d\tau$, d.h. der Punkt über einer Größe bedeutet nun die Ableitung nach der *Eigenzeit* und nicht nach der Koordinatenzeit t wie im vorigen Unterabschnitt. Außerdem haben wir bequemlichkeitshalber das Vorzeichen der Wirkung umgekehrt. Dies ist unerheblich für die Bewegungsgleichungen, da es sich nur um einen multiplikativen Faktor -1 handelt, und die Stationaritätsbedingung der neuen Wirkung daher zu denselben Bewegungsgleichungen führt wie mit der ursprünglichen Wirkung.

2. Spezielle Relativitätstheorie

Nun ergibt sich aber die Schwierigkeit, daß bei der Wahl der Eigenzeit τ als Integrationsparameter dieser selbst ein Funktional der Trajektorie darstellt und wir daher bei der Variation von x^μ auch τ mitvariieren müßten. Wir können aber auch die Bedingung, daß τ die Eigenzeit ist, als Zwangsbedingung $\dot{x}^\mu \dot{x}_\mu = c^2$ betrachten und mit Hilfe eines Lagrange-Multiplikators ins Variationsprinzip einbauen, d.h. wir schreiben

$$\tilde{L}'(x^\mu, \dot{x}^\mu, \mu) = mc \sqrt{\dot{x}^\mu \dot{x}_\mu} + \frac{q}{c} A_\mu \dot{x}^\mu + \frac{\mu}{2} (\dot{x}_\mu \dot{x}^\mu - c^2) = mc^2 + \frac{q}{c} A_\mu \dot{x}^\mu + \frac{\mu}{2} (\dot{x}_\mu \dot{x}^\mu - c^2). \quad (2.11.17)$$

Damit können wir die $x_\mu(\tau)$ und $\mu(\tau)$ unabhängig voneinander variieren, und τ als freien Parameter für diese erweiterte Trajektorie betrachten und im ersten Term ausnutzen, daß wir τ als Eigenzeit für beliebige Trajektorien ansehen, also nicht nur für die Lösungen der Bewegungsgleichungen. Freilich müssen wir den Term mit dem Lagrange-Parameter beibehalten, denn dessen Variation soll ja die entsprechende Bedingung als dynamische Gleichung generieren.

Dann erhalten wir die Euler-Lagrange-Gleichungen (*nachrechnen!*)

$$\begin{aligned} \delta x^\mu : \quad \frac{d}{d\tau}(\mu \dot{x}^\mu) &= \frac{q}{c} F_{\mu\nu} \dot{x}^\nu, \\ \delta \mu : \quad \dot{x}_\mu \dot{x}^\mu &= c^2. \end{aligned} \quad (2.11.18)$$

Nun ist in unserer Lagrange-Funktion (2.11.17) weiterhin der Wechselwirkungsterm

$$L_{\text{int}} = \frac{q}{c} A_\mu \dot{x}^\mu \quad (2.11.19)$$

eine bzgl. \dot{x}^μ homogene Funktion ersten Grades, d.h. es gilt

$$\dot{x}^\mu \frac{\partial L_{\text{int}}}{\partial \dot{x}^\mu} = L_{\text{int}}. \quad (2.11.20)$$

Weiter hängt die Lagrange-Funktion (2.11.17) nicht explizit von τ ab, d.h. es gilt die Erhaltung der entsprechenden Hamilton-Funktion

$$\tilde{H} = \dot{x}^\mu p_\mu - \tilde{L}', \quad p_\mu = \frac{\partial \tilde{L}'}{\partial \dot{x}^\mu}. \quad (2.11.21)$$

Wegen (2.11.20) ist

$$\tilde{H} = \frac{\mu}{2} \dot{x}_\mu \dot{x}^\mu + \frac{\mu}{2} c^2 - mc^2 = (\mu - m)c^2 = \text{const} \quad (2.11.22)$$

Für die Lösungen der Bewegungsgleichungen (2.11.18). Damit ist aber $\mu = \text{const}$, und wir erhalten als Bewegungsgleichung endgültig

$$\mu \ddot{x}^\mu = \frac{\partial L_{\text{int}}}{\partial x_\mu} - \frac{d}{d\tau} \frac{\partial L_{\text{int}}}{\partial \dot{x}_\mu} = \frac{q}{c} F_{\mu\nu} \dot{x}^\nu. \quad (2.11.23)$$

Vergleichen wir dies mit den ursprünglichen Bewegungsgleichungen, ergibt sich für den konstanten Lagrangeparameter schließlich noch

$$\mu = m. \quad (2.11.24)$$

Da wir nun Konstanten in der Lagrange-Funktion weglassen können, weil dies die Euler-Lagrange-Gleichungen nicht ändert und wir wissen, daß mit einer Wechselwirkungs-Lagrange-Funktion, die die Homogenitätsbedingung (2.11.20) erfüllt, $\mu = m = \text{const}$ automatisch kompatibel mit der Euler-Lagrange-Gleichung für die Variation der x^μ ist, können wir als Lagrange-Funktion folglich auch einfach

$$L' = \frac{m}{2} \dot{x}^\mu \dot{x}_\mu + L_{\text{int}} \quad (2.11.25)$$

zusammen mit der Nebenbedingung

$$\dot{x}_\mu \dot{x}^\mu = c^2, \quad (2.11.26)$$

welche den Weltlinienparameter τ als die Eigenzeit des Teilchens festlegt, verwenden, wobei die x^μ unabhängig variiert werden dürfen.

2.12 Tensoranalysis im Minkowski-Raum

Wir wollen der Vollständigkeit halber noch die einfachsten Grundlagen der Tensoranalysis im Minkowski-Raum betrachten, wobei wir uns auf pseudokartesische Koordinaten beschränken wollen. Seien also e_μ und \bar{e}_ν zwei pseudokartesische Basissysteme, die

$$e_\mu \cdot e_\nu = \bar{e}_\mu \cdot \bar{e}_\nu = \eta_{\mu\nu} \quad (2.12.1)$$

erfüllen. Es sei weiter Λ^μ_ν die entsprechende Lorentz-Transformationsmatrix

$$e_\nu = \bar{e}_\mu \Lambda^\mu_\nu. \quad (2.12.2)$$

Die Komponenten eines Vektors transformieren sich dann stets gemäß

$$\bar{x}^\mu = \Lambda^\mu_\nu x^\nu \Rightarrow \Lambda^\mu_\nu = \frac{\partial \bar{x}^\mu}{\partial x^\nu} = \partial_\nu \bar{x}^\mu. \quad (2.12.3)$$

Wegen (2.12.1) folgt, daß eine Lorentz-Transformation dann und nur dann vorliegt, wenn

$$\eta_{\mu\nu} \Lambda^\mu_\rho \Lambda^\nu_\sigma = \eta_{\rho\sigma} \quad (2.12.4)$$

gilt (*nachrechnen!*). Kontrahieren wir dies mit $(\Lambda^{-1})^\sigma_\alpha \eta^{\rho\beta}$, folgt daraus (*nachrechnen!*)

$$\eta_{\mu\alpha} \eta^{\rho\beta} \Lambda^\mu_\rho = \Lambda^\beta_\alpha = (\Lambda^{-1})^\beta_\alpha. \quad (2.12.5)$$

Wegen (2.12.3) folgt, daß

$$(\Lambda^{-1})^\mu_\nu = \Lambda_\nu^\mu = \frac{\partial x^\mu}{\partial \bar{x}^\nu} = \bar{\partial}_\nu x^\mu. \quad (2.12.6)$$

Wir wollen nun herausfinden wie sich die Komponenten

$$x_\mu = \eta_{\mu\nu} x^\nu \quad (2.12.7)$$

mit unteren Indizes transformieren. Es gilt

$$\bar{x}_\mu = \eta_{\mu\nu} \bar{x}^\nu = \eta_{\mu\nu} \Lambda^\nu_\rho x^\rho = \eta_{\mu\nu} \eta^{\rho\sigma} \Lambda^\nu_\rho x_\sigma = \Lambda_\mu^\sigma x_\sigma. \quad (2.12.8)$$

Tensorobjekte mit beliebigen oberen und unteren Indizes transformieren sich dann für jede Komponente wie die Komponenten des Raum-Zeit-Vierervektors für obere Indizes gemäß (2.12.3) **kontravariant** bzw. für untere Indizes gemäß (2.12.8) **kovariant**. Indizes werden stets mit der **Minkowski-Pseudometrik** $\eta_{\mu\nu}$ von oben nach unten bzw. mit der inversen Pseudometrix $\eta^{\mu\nu}$ von unten nach oben gezogen. Vermöge dieser Vorschrift kann man also jeden Tensor sowohl in ko- als auch kontravarianten Komponenten ausdrücken.

Durch Kontraktion erhält man aus Tensoren neue Tensoren. Z.B. sind

$$A_{\alpha\beta} = B_{\alpha\gamma\delta} C^\gamma_\beta = B_\alpha^\gamma C_{\gamma\beta} \quad (2.12.9)$$

die kovarianten Komponenten eines Tensors 2. Stufe, wenn $B_{\alpha\beta\gamma}$ die kovarianten Komponenten eines Tensors 3. und C^γ_β die gemischt kontra- und kovarianten Komponenten eines Tensors 2. Stufe sind. Als Beispiel für das Transformationsverhalten betrachten wir den letzteren Fall. Offenbar gilt

$$\bar{C}^\gamma_\beta = \Lambda^\gamma_\nu \Lambda_\beta^\nu C^\mu_\nu. \quad (2.12.10)$$

2. Spezielle Relativitätstheorie

Betrachten wir nun ein Skalarfeld Φ . Es transformiert sich definitionsgemäß wie folgt:

$$\overline{\Phi}(\overline{x}^\mu) = \Phi(x^\nu) = \Phi[(\Lambda^{-1})^\nu{}_\mu \overline{x}^\mu]. \quad (2.12.11)$$

Wir zeigen dann, daß $\partial_\mu \Phi$ die kovarianten Komponenten eines Vektorfeldes sind, wie es die Indexstellung an dem Ableitungssymbol bereits andeutet, denn es ist

$$\overline{\partial}_\mu \overline{\Phi} = \overline{\partial}_\mu \Phi = (\overline{\partial}_\mu x^\nu) \partial_\nu \Phi = \Lambda_\mu{}^\nu \partial_\nu \Phi, \quad (2.12.12)$$

und das ist genau das Transformationsverhalten eines kovarianten Tensors. Ebenso zeigt man, daß für Komponenten A_μ eines Vektorfeldes die Ableitung $\partial_\nu A_\mu$ Komponenten eines Tensors 2. Stufe sind (*nachrechnen!*). Insbesondere sind also auch $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$ tatsächlich kovariante Komponenten eines Tensorfeldes 2. Stufe, denn die A_μ sind kovariante Komponenten eines Vektorfeldes.

2.13 Lorentz-Transformationen des elektromagnetischen Feldes

Wir betrachten nun explizit die Lorentz-Transformationen des elektromagnetischen Feldes. Dazu müssen wir nur beachten, daß sich die Komponenten des Faraday-Tensors $F^{\mu\nu}$ (2.10.3) gemäß

$$F'^{\mu\nu}(x'^\alpha) = \Lambda^\mu{}_\rho \Lambda^\nu{}_\sigma F^{\rho\sigma}(x^\alpha). \quad (2.13.1)$$

transformieren.

Wir können nun zunächst aus dem antisymmetrischen Tensor 2. Stufe zwei Skalar (Invarianten) unter Lorentz-Transformationen bilden. Die erste ist

$$F_{\mu\nu} F^{\mu\nu} = -2(\vec{E}^2 - \vec{B}^2). \quad (2.13.2)$$

Um die zweite zu finden, definieren wir uns die vierdimensionale Version des **Levi-Civita-Symbols** $\epsilon^{\mu\nu\rho\sigma}$ als total antisymmetrisch unter Vertauschung von Indizes mit $\epsilon^{0123} = 1$. Es ist wichtig zu bemerken, daß nicht alle Lehrbücher und Paper dieser Konvention folgen, und man muß sich immer vergewissern, welche Konvention hinsichtlich der Vorzeichen benutzt wird. Es ist offenbar

$$\epsilon_{\mu\nu\rho\sigma} = \eta_{\alpha\mu} \eta_{\beta\nu} \eta_{\gamma\rho} \eta_{\delta\sigma} \epsilon^{\alpha\beta\gamma\delta} = \det \eta \epsilon^{\mu\nu\rho\sigma} = -\epsilon^{\mu\nu\rho\sigma}. \quad (2.13.3)$$

Das Levi-Civita-Symbol bildet invariante Tensorkomponenten bzgl. *spezieller* Lorentz-Transformationen, also für $\hat{\Lambda} = (\Lambda^\mu{}_\nu) \in \text{SO}(1,3)$, denn es gilt

$$\epsilon'^{\mu\nu\rho\sigma} = \Lambda^\mu{}_\alpha \Lambda^\nu{}_\beta \Lambda^\rho{}_\gamma \Lambda^\sigma{}_\delta \epsilon^{\alpha\beta\gamma\delta} = \det \hat{\Lambda} \epsilon^{\mu\nu\rho\sigma}. \quad (2.13.4)$$

Wir können nun das sog. **Hodge-Dual** des Faraday-Tensors

$$\dagger F^{\mu\nu} = \frac{1}{2} \epsilon^{\mu\nu\rho\sigma} F_{\rho\sigma} = \begin{pmatrix} 0 & -B^1 & -B^2 & -B^3 \\ B^1 & 0 & E^3 & -E^2 \\ B^2 & -E^3 & 0 & E^1 \\ B^3 & E^2 & -E^1 & 0 \end{pmatrix} \quad (2.13.5)$$

bilden. Wir können nun die Formeln aus Abschnitt 3.5 in [Hee14] für Kontraktionen von zwei Levi-Civita-Symbolen verwenden, müssen aber die Vorzeichenänderung in (2.13.3) beachten, die daraus resultiert, daß die Minkowski-Pseudometrik eine negative Determinante besitzt. Im folgenden benötigen wir die Formel

$$\epsilon^{\mu\nu\alpha\beta} \epsilon_{\rho\sigma\alpha\beta} = -2(\delta_\rho^\mu \delta_\sigma^\nu - \delta_\sigma^\mu \delta_\rho^\nu). \quad (2.13.6)$$

Wir bilden nun das Hodge-Dual von (2.13.5)

$$\begin{aligned}
 \dagger\dagger F_{\rho\sigma} &= \frac{1}{2} \epsilon_{\rho\sigma\mu\nu} \dagger F^{\mu\nu} \\
 &= \frac{1}{2} \epsilon_{\rho\sigma\mu\nu} \epsilon^{\mu\nu\alpha\beta} F_{\alpha\beta} \\
 &= \frac{1}{4} \epsilon_{\rho\sigma\mu\nu} \epsilon^{\alpha\beta\mu\nu} \\
 &= -\frac{1}{2} (\delta_{\rho}^{\alpha} \delta_{\sigma}^{\beta} - \delta_{\sigma}^{\alpha} \delta_{\rho}^{\beta}) F_{\alpha\beta} \\
 &= -\frac{1}{2} (F_{\rho\sigma} - F_{\sigma\rho}) = -F_{\rho\sigma}.
 \end{aligned} \tag{2.13.7}$$

Bis auf das Vorzeichen ist also das zweifache Hodge-Dual eines vollständig antisymmetrischen Tensors wieder derselbe antisymmetrische Tensor.

Wir können aus dem Faraday-Tensor also nur noch einen weiteren Skalar unter $SO(1,3)$ -Transformationen bilden:

$$\dagger F_{\mu\nu} F^{\mu\nu} = -4 \vec{E} \cdot \vec{B}. \tag{2.13.8}$$

Die beiden Invarianten (2.13.2) und (2.13.8) legen es nun nahe, den *komplexen Riemann-Silberstein-Vektor*

$$\vec{F} = \vec{E} + i\vec{B} \in \mathbb{C}^3 \tag{2.13.9}$$

einzuführen [Sil07]. Wir behalten für diese Vektoren allerdings das gewöhnliche euklidische Skalarprodukt bei, denn dann ist

$$\vec{F} \cdot \vec{F} = \vec{E} \cdot \vec{E} - \vec{B} \cdot \vec{B} + 2i\vec{E} \cdot \vec{B}, \tag{2.13.10}$$

und dieses Produkt ist invariant unter $SO(1,3)$ -Transformationen (2.13.1) des elektromagnetischen Feldes. Dies wiederum legt es nahe, zu vermuten, daß wir diese Transformationen als komplexwertige orthogonale Transformationen für die Riemann-Silberstein-Vektoren auffassen können. Wir werden gleich sehen, daß wir dadurch in der Tat eine umkehrbar eindeutige Abbildung zwischen den Gruppen $SO(1,3)^{\uparrow}$ und $SO(3, \mathbb{C})$ erhalten. Dabei ist die $SO(3, \mathbb{C})$ diejenige Gruppe, die aus Matrizen $\hat{R} \in \mathbb{C}^{3 \times 3}$ gebildet wird, die das oben definierte Skalarprodukt für die komplexen Vektoren invariant lassen, d.h. für die $\hat{R}^T \hat{R} = \hat{R}^{\dagger} \hat{R} = \mathbb{1}_3$ und $\det \hat{R} = 1$ gilt. Wenn nämlich für alle $\vec{F} \in \mathbb{C}^{3 \times 3}$ (2.13.10) gilt, folgt wegen

$$(\vec{F}_1 + \vec{F}_2) \cdot (\vec{F}_1 + \vec{F}_2) = \vec{F}_1 \cdot \vec{F}_1 + \vec{F}_2 \cdot \vec{F}_2 + 2\vec{F}_1 \cdot \vec{F}_2, \tag{2.13.11}$$

daß auch $\vec{F}_1 \cdot \vec{F}_2$ für beliebige $\vec{F}_1, \vec{F}_2 \in \mathbb{C}^3$ unter diesen Transformationen invariant bleibt.

Um diesen Sachverhalt zu beweisen wählen wir für \hat{R} zunächst einen Boost in x^1 -Richtung (2.3.18). Führt man (2.13.1) für diesen Fall explizit aus, ergibt sich (*nachrechnen!*)

$$\begin{aligned}
 \vec{E}' &= \begin{pmatrix} E^1 \\ \cosh \eta E^2 - \sinh \eta B^3 \\ \cosh \eta E^3 + \sinh \eta B^2 \end{pmatrix}, \\
 \vec{B}' &= \begin{pmatrix} B^1 \\ \cosh \eta B^2 + \sinh \eta E^3 \\ \cosh \eta B^3 - \sinh \eta E^2 \end{pmatrix}.
 \end{aligned} \tag{2.13.12}$$

Nun ist aber (*warum?*)

$$\cosh \eta = \frac{\exp \eta + \exp(-\eta)}{2} = \cos(i\eta), \quad \sinh \eta = \frac{\exp \eta - \exp(-\eta)}{2} = -i \sin(i\eta). \tag{2.13.13}$$

Damit wird (2.13.12) zu

$$\begin{aligned}\vec{E}' &= \begin{pmatrix} E^1 \\ \cos(i\eta)E^2 + i\sin(i\eta)B^3 \\ \cos(i\eta)E^3 - i\sin(i\eta)B^2 \end{pmatrix}, \\ \vec{B}' &= \begin{pmatrix} B^1 \\ \cos(i\eta)B^2 - i\sin(i\eta)E^3 \\ \cos(i\eta)B^3 + i\sin(i\eta)E^2 \end{pmatrix}.\end{aligned}\tag{2.13.14}$$

Fassen wir dies nun zu dem entsprechenden Riemann-Silberstein-Vektor zusammen, erhalten wir

$$\begin{aligned}\vec{F}' = \vec{E}' + i\vec{B}' &= \begin{pmatrix} F^1 \\ \cos(i\eta)F^2 + \sin(i\eta)F^3 \\ \cos(i\eta)F^3 - \sin(i\eta)F^2 \end{pmatrix} \\ &= \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos(i\eta) & \sin(i\eta) \\ 0 & -\sin(i\eta) & \cos(i\eta) \end{pmatrix} \vec{F} \\ &= \hat{R}_1(i\eta) \vec{F} \\ &= \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cosh \eta & i \sinh \eta \\ 0 & -i \sinh \eta & \cosh \eta \end{pmatrix} \vec{F}\end{aligned}\tag{2.13.15}$$

Die im letzten Ausdruck auftretende Matrix ist also tatsächlich $\hat{R}_1(i\eta) \in \text{SO}(3, \mathbb{C})$. Genauso zeigt man, daß eine Drehung in beliebiger Richtung

$$\hat{\Lambda}_{\text{rot}}(\vec{\varphi}) = \begin{pmatrix} 1 & 0 \\ 0 & \hat{R}(\vec{\varphi}) \end{pmatrix} \quad \text{mit} \quad \hat{R}(\vec{\varphi}) \in \text{SO}(3),\tag{2.13.16}$$

angewandt auf $F^{\mu\nu}$ gemäß (2.13.1) für den entsprechenden Riemann-Silberstein-Vektor auf

$$\vec{F}' = \hat{R}(\vec{\varphi}) \vec{F}\tag{2.13.17}$$

führt, d.h. wie zu erwarten, transformieren sich \vec{E} und \vec{B} unter gewöhnlichen $\text{SO}(3)$ -Transformationen wie Vektoren. Da die reinen Drehungen rein reell sind und selbstverständlich $\text{SO}(3) \subset \text{SO}(3, \mathbb{C})$ ist, ist damit die Behauptung bewiesen. Die $\text{SO}(1, 3)^\uparrow$ -Matrizen ergeben also für das elektromagnetische Feld in der Form von Riemann-Silberstein-Vektoren ausgedrückt umkehrbar eindeutig Transformationen mit $\text{SO}(3, \mathbb{C})$ -Matrizen, wobei Boosts rein imaginären Drehwinkeln $i\eta$ und Drehungen rein reellen Drehwinkeln φ entsprechen. D.h. jedem $\hat{\Lambda} \in \text{SO}(1, 3)^\uparrow$ entspricht umkehrbar eindeutig ein $\mathcal{L} \in \text{SO}(3, \mathbb{C})$, und für $\hat{\Lambda}_1, \hat{\Lambda}_2 \in \text{SO}(1, 3)^\uparrow$ wird $\hat{\Lambda}_1 \hat{\Lambda}_2$ auf $\hat{\mathcal{L}}_1 \hat{\mathcal{L}}_2$ abgebildet. Man nennt solche Abbildungen zwischen Gruppen, die die Produkte in der einen auf die entsprechenden Produkte in der anderen Gruppe abbilden, **Gruppenhomomorphismen**. In unserem Fall liegt sogar ein **Gruppenisomorphismus** vor, d.h. die Abbildung ist umkehrbar eindeutig.

Als Beispiel für eine Anwendung solcher gruppentheoretischer Formalismen zeigen wir, daß wir in Abschnitt (2.10) alle möglichen Fälle für die Bewegung in homogenen elektromagnetischen Feldern in dem Sinne erfaßt haben, daß wir alle anderen Fälle mittels Lorentz-Transformationen aus den angegebenen Lösungen erhalten können. Dazu betrachten wir die beiden Invarianten (2.13.2) und (2.13.8) hinsichtlich ihres Vorzeichens bzw. Verschwindens oder Nichtverschwindens. Betrachten wir zunächst den Fall $\vec{E} \cdot \vec{B} = 0$. Dann können wir offenbar die räumliche Basis so wählen, daß $\vec{E} = E\vec{e}_3$ und $\vec{B} = B\vec{e}_2$ ist. Dann ist $\vec{F} = (0, iB, E)^\text{T}$. Wenden wir dann den Boost (2.13.15) an, ergibt sich

$$\vec{F}' = \begin{pmatrix} 0 \\ i(B \cosh \eta + E \sinh \eta) \\ B \sinh \eta + E \cosh \eta \end{pmatrix}.\tag{2.13.18}$$

Ist dann $|B| < |E|$ bzw. $E^2 - B^2 > 0$, können wir durch eine geeignete Wahl von $\eta \in \mathbb{R}$ die zweite Komponente zum Verschwinden bringen:

$$B \cosh \eta + E \sinh \eta \Rightarrow \frac{B}{E} = -\tanh \eta \in (-1, 1) \Rightarrow \eta = -\operatorname{artanh}\left(\frac{B}{E}\right), \quad (2.13.19)$$

und dann gilt

$$\vec{F}' = \vec{E}' = \begin{pmatrix} 0 \\ 0 \\ E / \cosh \eta \end{pmatrix}, \quad (2.13.20)$$

d.h. im Bezugssystem Σ' handelt es sich um die Bewegung eines Teilchens im rein elektrostatischen Feld (vgl. Abschnitt 2.10.1).

Ist hingegen $\vec{E} \cdot \vec{B} = 0$ und $E^2 - B^2 < 0$, können wir die dritte Komponente von (2.13.18) durch die Wahl $\eta = -\operatorname{artanh}(E/B)$ zum Verschwinden bringen, und es wird

$$\vec{F}' = i\vec{B}' = \begin{pmatrix} 0 \\ iB / \cosh \eta \\ 0 \end{pmatrix}, \quad (2.13.21)$$

und es liegt im Bezugssystem Σ' der Fall eines rein magnetischen Feldes vor, den wir in Abschnitt 2.10.2 behandelt haben.

Ist $\vec{E} \cdot \vec{B} = 0$ und $E^2 - B^2 = 0$, handelt es sich um ein elektrisches Feld, das zu einem magnetischen Feld senkrecht steht. Diesen Fall haben wir in Abschnitt 2.10.4 behandelt.

Es bleibt uns schließlich noch, die Fälle mit $\vec{E} \cdot \vec{B} \neq 0$ zu behandeln. Dann können wir offenbar ohne Beschränkung der Allgemeinheit $\vec{E} = E\vec{e}_2$ und $\vec{B} = B^2\vec{e}_2 + B^3\vec{e}_3$ setzen. Dann gilt für den Riemann-Silberstein-Vektor in einem entlang der x^1 -Achse geboosteten Bezugssystem Σ'

$$\vec{F}' = \begin{pmatrix} 0 \\ \cosh \eta(E + iB^2) - \sinh \eta B^3 \\ -i \sinh \eta(E + iB^2) + \cosh \eta B^3 \end{pmatrix}. \quad (2.13.22)$$

Da $\vec{E}\vec{B} = EB^2 \neq 0$ ist, ist $E + iB^2 \neq 0$, und folglich können wir durch Wahl von $\eta = \operatorname{artanh}[-iB^3/(E + iB^2)] \in \mathbb{C}$ diesen Fall auf den Fall, daß $\vec{E}' = E'\vec{e}_2$ und $\vec{B}' = B'\vec{e}_2$, also $\vec{E}' \parallel \vec{B}'$ zurückführen, und die Bewegung eines geladenen Teilchens in einem solchen Feld haben wir in Abschnitt 2.10.3 behandelt. Damit haben wir aber alle möglichen Fälle für alle möglichen Werte für die Invarianten $\vec{E}^2 - \vec{B}^2$ und $\vec{E} \cdot \vec{B}$ erfaßt und damit gezeigt, daß sich die Bewegung von geladenen Teilchen in beliebigen homogenen elektromagnetischen Feldern durch Lorentz-Transformationen auf einen der in Abschnitt 2.10 behandelten Fälle zurückführen läßt.

2. Spezielle Relativitätstheorie

Anhang A

Matrix-Lie-Gruppen und -Lie-Algebren

Die sog. „klassischen Lie-Gruppen“ spielen, wie dieses Skript bereits anhand der klassischen Mechanik zeigt, eine herausragende Rolle in der Physik. Sie sind alle durch lineare Abbildungen in endlichdimensionalen reellen oder komplexen Vektorräumen bzw. nach Einführung einer Basis von Matrizen. Daher stellen wir in diesem Anhang kurz die wichtigsten Grundlagen der entsprechenden Mathematik zusammen. Wir betrachten auch die Anfangsgründe der **Darstellungstheorie** von Gruppen.

A.1 Unitäre Vektorräume

Hier führen wir den wichtigen Begriff des sog. **unitären Vektorraums** ein, also der endlichdimensionalen Vektorräume über dem Körper der komplexen Zahlen mit einem Skalarprodukt. Wir bedienen uns dabei der bequemen in der Quantentheorie üblichen Notation des Diracschen **Bra-Ket-Formalismus**.

Ein **unitärer Vektorraum** V ist definiert durch eine Menge mit Elementen $|\psi\rangle$, auf der zunächst zwei algebraische Verknüpfungen definiert sind. Die **Addition zweier Vektoren** $|\psi_1\rangle, |\psi_2\rangle \mapsto |\psi_1\rangle + |\psi_2\rangle$, die die Rechenregeln

$$(|\psi_1\rangle + |\psi_2\rangle) + |\psi_3\rangle = |\psi_1\rangle + (|\psi_2\rangle + |\psi_3\rangle), \quad (\text{A.1.1})$$

$$|\psi_1\rangle + |\psi_2\rangle = |\psi_2\rangle + |\psi_1\rangle, \quad (\text{A.1.2})$$

$$\exists |0\rangle \forall |\psi\rangle \in V: |\psi\rangle + |0\rangle = |\psi\rangle, \quad (\text{A.1.3})$$

$$\forall |\psi\rangle \in V \exists -|\psi\rangle \in V: |\psi\rangle + (-|\psi\rangle) := |\psi\rangle - |\psi\rangle = |0\rangle. \quad (\text{A.1.4})$$

Kurz gesagt, ist der Vektorraum bzgl. der Addition als Verknüpfung eine abelsche Gruppe.

Weiter ist auch eine Verknüpfung $\lambda, |\psi\rangle \mapsto \lambda|\psi\rangle$ für $\lambda \in \mathbb{C}$ und $|\psi\rangle \in V$ definiert, das folgende Axiome erfüllt

$$\lambda(|\psi_1\rangle + |\psi_2\rangle) = \lambda|\psi_1\rangle + \lambda|\psi_2\rangle, \quad (\text{A.1.5})$$

$$(\lambda_1 + \lambda_2)|\psi\rangle = \lambda_1|\psi\rangle + \lambda_2|\psi\rangle, \quad (\text{A.1.6})$$

$$(\lambda_1\lambda_2)|\psi\rangle = \lambda_1(\lambda_2|\psi\rangle), \quad (\text{A.1.7})$$

$$1|\psi\rangle = |\psi\rangle. \quad (\text{A.1.8})$$

Wie bei reellen Vektorräumen definiert man eine Teilmenge $U \subset V$ als **linear unabhängig**, wenn für beliebige endliche Mengen $|\psi_j\rangle \in U, j \in \{1, \dots, n\}$ aus

$$\sum_{j=1}^n \lambda_j |\psi_j\rangle = |0\rangle \quad (\text{A.1.9})$$

notwendig $\lambda_j = 0$ für $j \in \{1, \dots, n\}$ gilt.

Eine Teilmenge $B \subset V$ linear unabhängiger Vektoren heißt **Basis des Vektorraums**, wenn es zu jedem $|\psi\rangle \in V$ Zahlen λ_j und Vektoren $|b_j\rangle \in B$ ($j \in \{1, \dots, n\}$) gibt, so daß

$$|\psi\rangle = \sum_{j=1}^n \lambda_j |b_j\rangle. \quad (\text{A.1.10})$$

Man sagt auch, daß man jeden Vektor durch **Linearkombination** aus Vektoren in B darstellen kann. Die Darstellung jedes Vektors $|\psi\rangle$ als Linearkombination von Vektoren einer Basis ist **eindeutig**. Nehmen wir dazu an, es sei auch

$$|\psi\rangle = \sum_{j=1}^{n'} \lambda'_j |b'_j\rangle, \quad |b'_j\rangle \in B, \quad \lambda'_j \in \mathbb{C}. \quad (\text{A.1.11})$$

Dann ist aber auch

$$|\psi\rangle - |\psi\rangle = |0\rangle = \sum_{j=1}^n \lambda_j |b_j\rangle - \sum_{j=1}^{n'} \lambda'_j |b'_j\rangle \quad (\text{A.1.12})$$

eine Linearkombination von Vektoren aus B . Da die darin enthaltenen Vektoren definitionsgemäß linear unabhängig sind, müssen die Koeffizienten allesamt verschwinden. Ist nun für ein $j, k \in \mathbb{N}$ $|b_j\rangle = |b'_k\rangle$ ist also $\lambda_j = \lambda'_k$, und für evtl. vorhandene voneinander verschiedene Vektoren in den Mengen $\{|b_j\rangle\}_{j \in \{1, \dots, n\}}$ und $\{|b'_j\rangle\}_{j \in \{1, \dots, n'\}}$ müssen die entsprechenden λ_j bzw. λ'_j einzeln verschwinden. Dies impliziert aber die behauptete Eindeutigkeit der Linearkombination eines jeden beliebigen Vektors aus Vektoren der Basis B .

Im folgenden betrachten wir nur **endlichdimensionale Vektorräume**, d.h. wir nehmen an, es gibt eine Basis B mit genau $d \in \mathbb{N}$ Vektoren. Dann besitzen *alle* Basen von V genau d Vektoren, und man nennt V einen **d -dimensionalen komplexen Vektorraum**. Um zu zeigen, daß jede Basis eines d -dimensionalen Vektorraums genau d Vektoren beinhaltet, beweisen wir den **Austauschsatz** von Steinitz:

Sei also $B = \{|b_j\rangle\}_{j \in \{1, \dots, d\}}$ eine Basis und $|\psi\rangle \neq |0\rangle$ ein beliebiger Vektor. Dann gibt es ein $i \in \{1, \dots, d\}$, so daß

$$\{|b_1\rangle, \dots, |b_{i-1}\rangle, |\psi\rangle, |b_{i+1}\rangle, \dots, |b_d\rangle\} \quad (\text{A.1.13})$$

wieder eine Basis ist.

Beweis: Da B eine Basis ist, gibt es Zahlen $\psi_j \in \mathbb{C}$ ($j \in \{1, \dots, d\}$), so daß

$$|\psi\rangle = \sum_{j=1}^d \psi_j |b_j\rangle. \quad (\text{A.1.14})$$

Da $|\psi\rangle \neq 0$ ist wenigstens eine der Komponenten $\psi_i \neq 0$. Durch Umordnung können wir erreichen, daß $\psi_1 \neq 0$. Folglich ist

$$|b_1\rangle = \frac{1}{\psi_1} \left(|\psi\rangle - \sum_{j=2}^d \psi_j |b_j\rangle \right). \quad (\text{A.1.15})$$

Damit ist klar, daß wir jeden Vektor auch als Linearkombination der Vektoren $B' = \{|\psi\rangle, |b_2\rangle, \dots, |b_d\rangle\}$ darstellen können, denn dies ist für die Vektoren aus B der Fall, und wir können $|b_1\rangle$ mit Hilfe von (A.1.15) aus der entsprechenden Linearkombination eliminieren. Bleibt zu zeigen, daß B' linear unabhängig ist. Sei also

$$\lambda_1 |\psi\rangle + \sum_{j=2}^d \lambda_j |b_j\rangle = |0\rangle. \quad (\text{A.1.16})$$

Wegen (A.1.14) folgt daraus

$$\lambda_1 \psi_1 |b_1\rangle + \sum_{j=2}^d (\lambda_j + \lambda_1 \psi_j) |b_j\rangle. \quad (\text{A.1.17})$$

Da B als Basis linear unabhängig ist, folgt $\lambda_1 \psi_1 = 0$. Da voraussetzungsgemäß $\psi_1 \neq 0$ folgt $\lambda_1 = 0$. Daraus ergibt sich wiederum, daß $0 = \lambda_j + \lambda_1 \psi_j = \lambda_j$ auch für $j \in \{2, \dots, d\}$ gilt. Damit ist B' in der Tat eine Basis, denn wie wir eben gezeigt haben, kann man mit den in B' enthaltenen Vektoren jeden anderen Vektor als Linearkombination darstellen, und sie sind linear unabhängig.

Betrachten wir nun eine beliebige Menge $M = \left\{ |b'_j\rangle \right\}_{j \in \{1, \dots, n\}}$ linear unabhängiger Vektoren. Dann ist $n \leq d$.

In der Tat können wir nach dem eben bewiesenen Austauschatz von Steinitz einen der Vektoren $|b_j\rangle$ aus der Basis B durch $|b'_1\rangle$ austauschen. Wir können (evtl. nach Ummumerieren der Vektoren $|b'_j\rangle$) annehmen, dies sei $|b_1\rangle$. Dadurch haben wir eine neue Basis B_1 . Genauso können wir jetzt mit $|b'_2\rangle$ und der Basis B_2 verfahren. Da M eine Menge linear unabhängiger Vektoren ist, kann man einen der Vektoren $\{|b_2\rangle, \dots, |b_d\rangle\}$ austauschen, denn in der Linearkombination $|b'_2\rangle = \lambda_1 |b'_1\rangle + \sum_j \lambda_j |b_j\rangle$ ist mindestens ein $\lambda_j \neq 0$ für $j \in \{2, \dots, d\}$, weil sonst andernfalls $|b'_2\rangle = \lambda_1 |b'_1\rangle$ wäre, aber dann wäre M keine Menge linear unabhängiger Vektoren. So können wir offenbar verfahren, bis alle Vektoren der Menge M in einer durch Austausch neu konstruierten Basis B_n vorhanden sind. Folglich muß aber $n \neq d$ sein. Ist nun M sogar Basis, kann man die Rollen von M und B in der obigen Argumentation vertauschen, woraus sich $d \leq n$ ergibt. Ist also M ebenfalls Basis, muß $n = d$ sein. Damit ist gezeigt, daß alle Basen eines endlichdimensionalen Vektorraums die gleiche Anzahl von Basisvektoren beinhalten. Diese Zahl nennt man die **Dimension des Vektorraums**.

Zu einem unitären Vektorraum fehlt uns nun nur noch der Begriff des **Skalarprodukts**. Bei reellen Vektorräumen kann dieses durch irgendeine positiv definite Bilinearform definiert werden. Für einen komplexen Vektorraum benötigen wir stattdessen eine **positiv definite Sesquilinearform**, die wir jetzt definieren wollen. Es handelt sich um eine Abbildung zweier Vektoren $|\phi\rangle, |\psi\rangle \mapsto \langle \phi | \psi \rangle$, das die folgenden Eigenschaften besitzt: viii

$$\langle \phi | \lambda_1 \psi_1 + \lambda_2 \psi_2 \rangle = \lambda_1 \langle \phi | \psi_1 \rangle + \lambda_2 \langle \phi | \psi_2 \rangle, \quad (\text{A.1.18})$$

$$\langle \phi | \psi \rangle = \langle \psi | \phi \rangle^*, \quad (\text{A.1.19})$$

$$\langle \psi | \psi \rangle \geq 0, \quad (\text{A.1.20})$$

$$\langle \psi | \psi \rangle = 0 \Leftrightarrow |\psi\rangle = |0\rangle. \quad (\text{A.1.21})$$

Wir bemerken sofort, daß sich aus (A.1.18) und (A.1.19) ergibt, daß das Skalarprodukt bzgl. des ersten Arguments **semilinear** ist, d.h.

$$\begin{aligned} \langle \lambda_1 \phi_1 + \lambda_2 \phi_2 | \psi \rangle &= \langle \psi | \lambda_1 \phi_1 + \lambda_2 \phi_2 \rangle^* \\ &= (\lambda_1 \langle \psi | \phi_1 \rangle + \lambda_2 \langle \psi | \phi_2 \rangle)^* \\ &= \lambda_1^* \langle \psi | \phi_1 \rangle^* + \lambda_2^* \langle \psi | \phi_2 \rangle^* \\ &= \lambda_1^* \langle \phi_1 | \psi \rangle + \lambda_2^* \langle \phi_2 | \psi \rangle. \end{aligned} \quad (\text{A.1.22})$$

Wie bei reellen Vektorräumen können wir nun die **Norm** („Länge“) eines Vektors durch

$$\|\psi\| = \sqrt{\langle \psi | \psi \rangle} \quad (\text{A.1.23})$$

definieren.

Im Zusammenhang mit dem Skalarprodukt ist wie im Reellen die **Schwarzsche Ungleichung**:

$$|\langle \psi_1 | \psi_2 \rangle| \leq \|\psi_1\| \|\psi_2\|. \quad (\text{A.1.24})$$

Falls $|\psi_2\rangle = |0\rangle$ ist $\langle\psi_1|\psi_2\rangle = \langle\psi_1|0\rangle = 0$, und es gilt damit die Ungleichung mit dem Gleichheitszeichen. Sei nun im folgenden $|\psi\rangle \neq |0\rangle$. Dann betrachten verwenden wir die positive Definitheit (A.1.20) für $\psi = |\psi_1\rangle + \lambda|\psi_2\rangle$ für $\lambda \in \mathbb{C}$:

$$\langle\psi|\psi\rangle = \|\psi_1\|^2 + |\lambda|^2\|\psi_2\|^2 + \lambda\langle\psi_1|\psi_2\rangle + \lambda^*\langle\psi_2|\psi_1\rangle \geq 0. \quad (\text{A.1.25})$$

Insbesondere können wir hierin

$$\lambda = -\frac{\langle\psi_2|\psi_1\rangle}{\|\psi_2\|^2} \quad (\text{A.1.26})$$

setzen, denn wegen $|\psi_2\rangle \neq |0\rangle$ ist $\|\psi_2\|^2 > 0$. Dann ergibt sich aus (A.1.25)

$$\|\psi_1\|^2 - \frac{|\langle\psi_1|\psi_2\rangle|^2}{\|\psi_2\|^2} \geq 0, \quad (\text{A.1.27})$$

und daraus ergibt sich sofort (A.1.24).

Besonders bequeme Basen sind nun, wie bei den reellen Euklidischen Vektorräumen, die **Orthonormalbasen**. Eine Basis $|e_j\rangle$ ($j \in \{1, \dots, d\}$) heißt Orthonormalbasis, wenn

$$\langle e_j | e_k \rangle = \delta_{jk} \quad (\text{A.1.28})$$

ist. Wir hatten oben vorausgesetzt, daß der Vektorraum eine Basis $B = \{|b_j\rangle\}_{j \in \{1, \dots, d\}}$ besitzt. Daraus folgt aber sofort, daß es auch eine Orthonormalbasis gibt. Diese läßt sich mit dem **Schmidtschen Orthonormalisierungsverfahren** rekursiv aus einer beliebigen Basis konstruieren. Dazu wählen wir

$$|e_1\rangle = \frac{1}{\|b_1\|} |b_1\rangle, \quad (\text{A.1.29})$$

was problemlos möglich ist, denn es ist $|b_1\rangle \neq |0\rangle$, weil ansonsten B keine Basis sein könnte. Dann gilt $\|e_1\| = 1$. Wir können dann den Vektor

$$|b'_2\rangle = |b_2\rangle - |e_1\rangle \langle e_1 | b_2 \rangle \quad (\text{A.1.30})$$

definieren. Es ist offensichtlich $|b'_2\rangle \neq |0\rangle$, weil ansonsten $|b_1\rangle$ und $|b_2\rangle$ linear abhängig wären. Wir können also

$$|e_2\rangle = \frac{1}{\|b'_2\|} |b'_2\rangle \quad (\text{A.1.31})$$

definieren. Man rechnet leicht nach, daß dann $\langle e_1 | e_2 \rangle = 0$ und $\|e_2\| = 1$ ist. Auf diese Weise fahren wir nun fort. Angenommen, wir hätten die paarweise zueinander orthogonalen Einheitsvektoren $|e_1\rangle, \dots, |e_k\rangle$ konstruiert. Dann ergibt sich

$$|b'_{k+1}\rangle = |b_{k+1}\rangle - \sum_{j=1}^k |e_j\rangle \langle e_j | b_{k+1} \rangle, \quad |e_{k+1}\rangle = \frac{1}{\|b'_k\|} |b'_k\rangle. \quad (\text{A.1.32})$$

Man rechnet auch hier leicht nach, daß $\langle e_{k+1} | e_j \rangle = 0$ für $j \leq k$ und $\|e_{k+1}\| = 1$ (*Übung*). Offensichtlich bricht dieses Verfahren mit $k = d$ ab, d.h. man hat dann eine **Orthonormalbasis**, auch **vollständiges Orthonormalsystem (VONS)** genannt, konstruiert.

Sei nun $|\psi\rangle \in V$ beliebig. Dann können wir die Komponenten bzgl. des VONS leicht mit Hilfe des Skalarprodukts berechnen. Da nach der obigen Konstruktion das VONS eine Basis ist, läßt sich $|\psi\rangle$ als Linearkombination

$$|\psi\rangle = \sum_{j=1}^d |e_j\rangle \langle e_j | \psi \rangle \quad (\text{A.1.33})$$

darstellen. Aus der Orthonormalität der $|e_j\rangle$ folgt dann aber daraus sofort

$$\langle e_k | \psi \rangle = \sum_{j=1}^d \langle e_k | e_j \rangle \psi_j = \sum_{j=1}^d \delta_{kj} \psi_j = \psi_k. \quad (\text{A.1.34})$$

Wir können also (A.1.33) auch in der Form

$$|\psi\rangle = \sum_{j=1}^d |e_j\rangle \langle e_j | \psi \rangle \quad (\text{A.1.35})$$

schreiben. Definieren wir nun $|\alpha\rangle\langle\beta|$ als die lineare Abbildung (oft auch **linearer Operator** genannt)

$$|\psi\rangle \mapsto |\alpha\rangle\langle\beta | \psi \rangle \quad (\text{A.1.36})$$

und definieren die Summe von Operatoren punktweise, können wir (A.1.35) auch in der Form

$$\sum_{j=1}^d |e_j\rangle\langle e_j| = \mathbb{1} \quad (\text{A.1.37})$$

schreiben. Man bezeichnet (A.1.38) als die **Vollständigkeitsrelation**.

Es zeigt sich, daß alle unitären Vektorräume mit \mathbb{C}^d äquivalent sind. Mit dem oben definierten VONS brauchen wir nur $\langle\psi | \mapsto \rangle(\psi_j)$ mit $\psi_j = \langle e_j | \psi \rangle$ abzubilden. Für Vektoren $|\psi\rangle$ und $|\phi\rangle$ gilt dann offenbar $|\lambda_1\psi + \lambda_2\phi\rangle \mapsto (\lambda_1\psi_j + \lambda_2\phi_j)$ und mit Hilfe der Vollständigkeitsrelation (A.1.38)

$$\langle\psi | \phi \rangle = \sum_{j=1}^d \langle\psi | e_j\rangle \langle e_j | \phi \rangle = \sum_{j=1}^d \psi_j^* \phi_j. \quad (\text{A.1.38})$$

Dies definiert das **kanonische Skalarprodukt** auf \mathbb{C}^d .

A.2 Unitäre und selbstadjungierte Abbildungen

Wir betrachten nun **lineare Abbildung** auf einem d -dimensionalen unitären Vektorraum. Man spricht, insbesondere auch in der Quantentheorie, wo allerdings die Theorie der unitären Vektorräume zur Theorie des unendlich-dimensionalen **Hilbert-Raums** erweitert wird, von **linearen Operatoren**. Auch dieser Begriff ist vollständig analog zu dem in reellen Vektorräumen. Eine Abbildung $\hat{A} : V \rightarrow V$ heißt lineare Abbildung, wenn für alle $\lambda_1, \lambda_2 \in \mathbb{C}$ und alle $|\psi_1\rangle, |\psi_2\rangle \in V$

$$\hat{A}(\lambda_1 |\psi_1\rangle + \lambda_2 |\psi_2\rangle) = \lambda_1 \hat{A} |\psi_1\rangle + \lambda_2 \hat{A} |\psi_2\rangle \quad (\text{A.2.1})$$

gilt.

Als erstes zeigen wir, daß zu jedem linearen Operator \hat{A} auch ein dazugehöriger **hermitesch adjungierter Operator** \hat{A}^\dagger definiert werden kann, für den für alle $|\psi_1\rangle, |\psi_2\rangle \in V$

$$\langle\psi_1 | \hat{A} \psi_2 \rangle = \langle \hat{A}^\dagger \psi_1 | \psi_2 \rangle. \quad (\text{A.2.2})$$

Um dies zu beweisen, verwenden wir ein beliebiges VONS $\{|e_j\rangle\}_{j \in \{1, \dots, d\}}$. Verwenden wir zweimal die Vollständigkeitsrelation (A.1.37) folgt

$$\hat{A} = \sum_{j,k=1}^d |e_j\rangle\langle e_j | \hat{A} | e_k \rangle \langle e_k| = \sum_{j,k=1}^d A_{jk} |e_j\rangle\langle e_k|. \quad (\text{A.2.3})$$

Man nennt

$$A_{jk} = \langle e_j | \hat{A} | e_k \rangle \quad (\text{A.2.4})$$

die **Matrixelemente** des linearen Operators \hat{A} bzgl. des VONS $\{|e_j\rangle\}_{j \in \{1, \dots, d\}}$.

Nun gilt wegen (A.2.3)

$$\langle \psi_1 | \hat{A} \psi_2 \rangle = \sum_{j,k=1}^d \langle \psi_1 | e_j \rangle A_{jk} \langle e_j | \psi_2 \rangle = \sum_{j,k=1}^d \psi_{1,j}^* A_{jk} \psi_k = \sum_{j,k=1}^d (A_{jk}^* \psi_{1,j})^* \psi_k. \quad (\text{A.2.5})$$

Definieren wir also

$$\hat{A}^\dagger = \sum_{j,k=1}^d A_{jk}^* |e_k\rangle \langle e_j|, \quad (\text{A.2.6})$$

folgt

$$\hat{A}^\dagger |\psi_1\rangle = \sum_{j,k=1}^d A_{jk}^* \psi_{1,j} |e_k\rangle. \quad (\text{A.2.7})$$

und damit

$$\begin{aligned} \langle \psi_2 | \hat{A}^\dagger \psi_1 \rangle &= \sum_{j,k=1}^d A_{jk}^* \psi_{1,j} \psi_{2,k}^* \\ \Rightarrow \langle \hat{A}^\dagger \psi_1 | \psi_2 \rangle &= (\langle \psi_2 | \hat{A}^\dagger \psi_1 \rangle)^* = \sum_{j,k=1}^d A_{jk} \psi_{1,j}^* \psi_{2,k} \stackrel{(\text{A.2.5})}{=} \langle \psi_1 | \hat{A} \psi_2 \rangle. \end{aligned} \quad (\text{A.2.8})$$

Damit erfüllt aber der durch (A.2.6) definierte lineare Operator in der Tat die Definition (A.2.2) des zu \hat{A} hermitesch adjungierten Operators, und damit ist dessen Existenz sichergestellt.

Betrachten wir die Komponenten ψ_j als Spaltenvektor $\tilde{\psi}$ und definieren $\tilde{\psi}^\dagger = (\psi_1^*, \dots, \psi_d^*)$ als den entsprechenden hermitesch adjungierten Vektor, schreibt sich das Skalarprodukt in der Form

$$\langle \psi_1 | \psi_2 \rangle = \tilde{\psi}_1^\dagger \psi_2 \quad (\text{A.2.9})$$

im Sinne eines Matrixprodukts.

Genauso können wir die lineare Abbildung \hat{A} als Komponentenmatrix $\tilde{A} = (A_{jk})$ darstellen. Dann ist

$$\hat{A} |\psi\rangle = |\phi\rangle \Rightarrow \phi_j = \langle e_j | \hat{A} \psi \rangle = \sum_{k=1}^d \langle e_j | \hat{A} | e_k \rangle \langle e_k | \psi \rangle = \sum_{k=1}^d A_{jk} \psi_k. \quad (\text{A.2.10})$$

Also gilt in Matrix-Vektorschreibweise für die \mathbb{C}^n -Spaltenvektoren

$$\tilde{\phi} = \tilde{A} \tilde{\psi}. \quad (\text{A.2.11})$$

Aus (A.2.6) folgt, daß

$$(\hat{A}^\dagger)_{jk} = A_{kj}^*, \quad (\text{A.2.12})$$

d.h. \hat{A}^\dagger entspricht im \mathbb{C}^n die Matrix $\tilde{A}^\dagger = (\tilde{A}^*)^T$, d.h. die Matrix, die aus der Matrix $\tilde{A}^* = (A_{jk}^*)$ und anschließende Transposition (also „Spiegelung an der Hauptdiagonalen“) hervorgeht. Daraus folgt auch unmittelbar, daß

$$(\tilde{A}^\dagger)^\dagger = \tilde{A}. \quad (\text{A.2.13})$$

Dies überträgt sich auch auf die Operatoren, d.h. es ist auch stets

$$(\hat{A}^\dagger)^\dagger = \hat{A}. \quad (\text{A.2.14})$$

Das folgt auch aus der allgemeinen Definition der hermiteschen Adjungation eines Operators. Es gilt nämlich

$$\langle (\hat{A}^\dagger)^\dagger \psi_1 | \psi_2 \rangle = \langle \psi_1 | \hat{A}^\dagger \psi_2 \rangle = (\langle \hat{A}^\dagger \psi_2 | \psi_1 \rangle)^* = (\langle \psi_2 | \hat{A} \psi_1 \rangle)^* = \langle \hat{A} \psi_1 | \psi_2 \rangle \quad (\text{A.2.15})$$

für alle $|\psi_1\rangle, |\psi_2\rangle \in V$. Damit muß aber (A.2.14) gelten.

Jetzt können wir weitere spezielle lineare Abbildungen definieren: Eine lineare Abbildung \hat{A} bzw. ein linearer Operator heißt **selbstadjungiert** falls $\hat{A}^\dagger = \hat{A}$ ist. Entsprechend nennt man auch die dazugehörige Matrix \tilde{A} selbstadjungiert, wenn $\tilde{A}^\dagger = \tilde{A}^{*\text{T}} = \tilde{A}$ gilt. Selbstadjungierte Matrizen sind das komplexe Analogon zu den symmetrischen Matrizen für reelle Vektorräume.

Eine lineare Abbildung \hat{U} heißt **unitär**, wenn sie invertierbar ist und wenn $\hat{U}^{-1} = \hat{U}^\dagger$ gilt. Es ist klar, daß solche linearen Abbildungen Skalarprodukte beliebiger Vektoren invariant lassen, denn es gilt

$$\langle \hat{U} \psi_1 | \hat{U} \psi_2 \rangle = \langle \hat{U}^\dagger \hat{U} \psi_1 | \psi_2 \rangle = \langle \psi_1 | \psi_2 \rangle. \quad (\text{A.2.16})$$

Die entsprechenden Matrizen bzgl. eines beliebigen VONS nennt man entsprechend auch unitäre Matrizen, $\tilde{U}^{-1} = \tilde{U}^\dagger = \tilde{U}^{*\text{T}}$.

Wir bemerken, daß auch der Wechsel von einem VONS $\{|e_j\rangle\}$ zu einem anderen VONS $\{|e'_j\rangle\}$ durch eine unitäre Transformation gegeben ist, denn es gilt

$$|e'_j\rangle = \sum_{k=1}^d |e_k\rangle \langle e_k | e'_j \rangle = \sum_{k=1}^d |e_k\rangle U_{kj} \quad \text{mit} \quad U_{jk} = \langle e_k | e'_j \rangle. \quad (\text{A.2.17})$$

Damit folgt aus der Orthonormalität der Vektoren $|e_j\rangle$ und $|e'_j\rangle$

$$\delta_{jk} = \langle e'_j | e'_k \rangle = \sum_{i=1}^d \langle e'_j | e_i \rangle \langle e_i | e'_k \rangle = \sum_{i=1}^d U_{ij}^* U_{ik} \Rightarrow \tilde{U}^\dagger \tilde{U} = \mathbb{1}. \quad (\text{A.2.18})$$

Es ist auch klar, daß die unitären Abbildungen eine **Gruppe** bilden, denn sind \hat{U}_1 und \hat{U}_2 unitäre Operatoren, so gilt

$$(\hat{U}_1 \hat{U}_2)^\dagger (\hat{U}_1 \hat{U}_2) = \hat{U}_2^\dagger \hat{U}_1^\dagger \hat{U}_1 \hat{U}_2 = \hat{U}_2^\dagger \hat{U}_2 = \mathbb{1}, \quad (\text{A.2.19})$$

d.h. $(\hat{U}_1 \hat{U}_2)^\dagger = (\hat{U}_1 \hat{U}_2)^{-1}$, d.h. auch $\hat{U}_1 \hat{U}_2$ ist ein unitärer Operator.

A.3 Spektralsatz für selbstadjungierte Operatoren

Als nächstes beschäftigen wir uns mit dem **Eigenwertproblem** für einen selbstadjungierten Operator \hat{A} . Sei also \hat{A} selbstadjungiert. Seien nun $|a_j\rangle$ die **Eigenvektoren** von \hat{A} zu **Eigenwerten** a_j (vorausgesetzt es existieren überhaupt Eigenvektoren, was wir im folgenden beweisen wollen). Wir setzen im folgenden voraus, daß $\|a_j\| = 1$. Da voraussetzungsgemäß ein Eigenvektor nicht der Nullvektor sein darf, können wir dies durch Normieren stets erreichen.

Als erstes zeigen wir, daß die Eigenwerte reell sind ($a_j \in \mathbb{R}$) und daß Eigenvektoren zu verschiedenen Eigenwerten orthogonal zueinander sind. Dazu betrachten wir

$$a_j = \langle a_j | \hat{A} a_j \rangle = \langle \hat{A}^\dagger a_j | a_j \rangle = \langle \hat{A} a_j | a_j \rangle = a_j^* \langle a_j | a_j \rangle = a_j^* \Rightarrow a_j \in \mathbb{R}. \quad (\text{A.3.1})$$

Weiter gilt einerseits

$$\langle a_j | \hat{A} a_k \rangle = a_k \langle a_j | a_k \rangle \quad (\text{A.3.2})$$

und andererseits

$$\langle a_j | \hat{A} a_k \rangle = \langle \hat{A}^\dagger a_j | a_k \rangle = \langle \hat{A} a_j | a_k \rangle = a_j^* \langle a_j | a_k \rangle = a_j \langle a_j | a_k \rangle. \quad (\text{A.3.3})$$

Ziehen wir die Gleichungen (A.3.3) und (A.3.2) voneinander ab, folgt

$$(a_j - a_k) \langle a_j | a_k \rangle = 0. \quad (\text{A.3.4})$$

Falls also $a_j \neq a_k$ ist, folgt daraus $\langle a_j | a_k \rangle = 0$, d.h. die Eigenvektoren zu verschiedenen Eigenwerten sind orthogonal zueinander.

Falls es zu einem Eigenwert a_j mehr als einen linear unabhängigen Eigenvektor gibt, können wir mittels des Schmidtschen Orthonormalisierungsverfahrens diese linear unabhängigen Eigenvektoren durch eine Menge orthonormierter Eigenvektoren zu diesem Eigenwert ersetzen. Wir dürfen im folgenden also davon ausgehen, daß wir eine Menge orthonormierter Eigenvektoren vorliegen haben, d.h. wir dürfen annehmen, daß

$$\langle a_j | a_k \rangle = \delta_{jk} \quad (\text{A.3.5})$$

gilt.

Wir wollen nun zeigen, daß sich alle selbstadjungierten Operatoren diagonalisieren lassen, d.h. daß es ein VONS aus Eigenvektoren dieses Operators gibt. Dazu definieren wir die **Determinante** des Operators als die Determinante der entsprechenden Matrix bzgl. eines beliebigen VONS $|e_j\rangle$. Die Determinante ist unabhängig von der Wahl des VONS. Sei dazu $|e'_j\rangle$ ein beliebiges anderes VONS. Wie oben gezeigt, gibt es dann eine unitäre Abbildung \hat{U} , so daß $|e'_j\rangle = \hat{U} |e_j\rangle$. Dann folgt

$$A'_{jk} = \langle e'_j | \hat{A} e'_k \rangle = \langle \hat{U} e_j | \hat{A} \hat{U} e_k \rangle = \langle e_j | \hat{U}^\dagger \hat{A} \hat{U} e_k \rangle. \quad (\text{A.3.6})$$

Für die entsprechenden Matrizen gilt also

$$\tilde{A}' = \tilde{U}^\dagger \tilde{A} \tilde{U} = \tilde{U}^{-1} \tilde{A} \tilde{U} \Rightarrow \det \tilde{A}' = \det \tilde{U}^{-1} \det \tilde{A} \det \tilde{U} = \det \tilde{A}, \quad (\text{A.3.7})$$

d.h. die Determinante ist unabhängig vom gewählten VONS.

Damit ein Eigenvektor des Operators \hat{A} existiert, ist es offenbar notwendig und hinreichend, daß dies für die entsprechende Matrix \tilde{A} bzgl. eines beliebigen VONS der Fall ist, d.h. es existiert dann und nur dann ein Vektor $|a_j\rangle \neq |0\rangle$ mit $\hat{A} |a_j\rangle = a_j |a_j\rangle$, wenn a_j Nullstelle des **charakteristischen Polynoms** d -ten

$$P_A(\lambda) = \det(\tilde{A} - \lambda \mathbb{1}_d) = \det(\hat{A} - \lambda \mathbb{1}) \quad (\text{A.3.8})$$

ist (*Beweis als Übung*). Nach dem Fundamentalsatz der Algebra besitzt nun jedes Polynom d -ten Grades d Nullstellen (wobei es natürlich sein kann, daß es mehrfache Nullstellen gibt). Es existiert also wenigstens ein Eigenwert und ein dazugehöriger Eigenvektor. Daraus können wir nun einen Induktionsbeweis konstruieren.

Zunächst gilt der Satz für $d = 1$. Dann ist die Matrix \tilde{A} einfach eine Zahl, und die ist Eigenwert zu jedem von $|0\rangle$ verschiedenen Vektor in dem eindimensionalen unitären Raum. Nehmen wir nun an, die Behauptung gilt für $d \leq n$ und betrachten wir den Fall $d = n + 1$. Dann existiert nach der eben geführten Argumentation wenigstens ein Eigenwert a_1 und wenigstens ein dazugehöriger Eigenvektor $|a_1\rangle \neq |0\rangle$, den wir der Einfachheit halber als normiert annehmen, d.h. $\langle a_1 | a_1 \rangle = 1$.

Nun betrachten wir das **orthogonale Komplement** zu $|a_1\rangle$, d.h. den $d - 1$ -dimensionalen Untervektorraum U aller zu $|a_1\rangle$ orthogonalen Vektoren. Daß dies ein Untervektorraum ist, macht man sich leicht klar. Wir können nämlich mit dem Schmidtschen Orthonormalisierungsverfahren aus jeder Basis des Vektorraums ein VONS $\{|e_j\rangle\}$ mit $|e_1\rangle = |a_1\rangle$ konstruieren. Dann ist jeder Vektor, der orthogonal zu $|a_1\rangle$ durch

$$|\psi\rangle \in U \Leftrightarrow \langle \psi | a_1 \rangle = 0 \Rightarrow |\psi\rangle = \sum_{j=2}^d \psi_j |e_j\rangle \quad (\text{A.3.9})$$

darstellbar. Es ist also U tatsächlich ein $(d - 1)$ -dimensionaler Untervektorraum von V . Dieser Untervektorraum ist aber auch invariant unter der linearen Abbildung \hat{A} , d.h. für $|\psi\rangle \in U$ ist auch $\hat{A}|\psi\rangle \in U$. Das folgt aus

$$\langle a_1 | \hat{A}\psi \rangle = \langle \hat{A}^\dagger a_1 | \psi \rangle = \langle \hat{A} a_1 | \psi \rangle = a_1^* \langle a_1 | \psi \rangle = 0. \quad (\text{A.3.10})$$

Die Restriktion von \hat{A} auf den Untervektorraum U ist also eine selbstadjungierte lineare Abbildung im Untervektorraum U , und folglich besitzt dieser Operator gemäß der Induktionsvoraussetzung ein VONS aus Eigenvektoren $\{|a_j\rangle\}_{j \in \{2, \dots, d\}}$ in U . Zusammen mit $|a_1\rangle$ haben wir dann aber ein VONS für ganz V aus Eigenvektoren von \hat{A} . MaW. \hat{A} ist diagonalisierbar, und es existiert folglich zur entsprechenden Matrix \tilde{A} bzgl. irgendeines VONS von V eine unitäre Matrix \tilde{U} , so daß $\tilde{U}\tilde{A}\tilde{U}^\dagger = \tilde{A}' = \text{diag}(a_1, \dots, a_d)$, wobei die $a_j \in \mathbb{R}$ die Einwerte von \hat{A} bzw. der Matrix \tilde{A} sind.

Man bezeichnet die Menge aller Eigenwerte eines selbstadjungierten Operators \hat{A} auch als das **Spektrum** des Operators und den eben hergeleiteten Satz, wonach stets ein VONS aus Eigenvektoren dieses Operators existiert, d.h. für das die dazugehörige Matrix \tilde{A} diagonal wird als **Spektralsatz**.

Übrigens folgt aus diesem Satz für komplexe unitäre Vektorräume auch der entsprechende Satz für reelle euklidische Vektorräume: Für einen symmetrischen Operator \hat{A} existiert stets eine **kartesische Basis** aus Eigenvektoren dieses Operators bzgl. der die darstellende Matrix diagonal mit den Eigenwerten des Operators als Diagonalelementen wird. Dies ist unmittelbar klar, da wir die darstellende symmetrische Matrix $\tilde{A} \in \mathbb{R}^{d \times d}$ auch als selbstadjungierten Operator im unitären Vektorraum \mathbb{C}^d mit dem kanonischen Skalarprodukt ansehen können. Nach dem obigen Satz besitzt diese Matrix lauter reelle Eigenwerte und entsprechend auch reelle Eigenvektoren, so daß das VONS in \mathbb{C}^d aus Eigenvektoren tatsächlich auch eine kartesische Basis im \mathbb{R}^d ist.

Es sei dem Leser zur *Übung* überlassen, daß auch jede unitäre Abbildung diagonalisierbar ist. Die Eigenwerte sind dabei „reine Phasen“, d.h. von der Form $\exp(i\varphi)$ mit $\varphi \in \mathbb{R}$.

A.4 Frobenius-Skalarprodukt für lineare Abbildungen

Im folgenden wollen wir **Analysis** mit linearen Abbildungen bzw. den entsprechenden Matrizen betreiben. Dazu benötigen wir eine **Norm**, um einen für uns geeigneten Konvergenzbegriff für operatorwertige Folgen, Reihen und Ableitungen nach reellen oder komplexen Parametern definieren zu können. Für endlichdimensionale Vektorräume stellt dies kein größeres Problem dar, wie wir gleich sehen werden.

Zunächst stellen wir fest, daß beliebige lineare Operatoren $\hat{A}: V \rightarrow V$, wobei V ein d -dimensionaler unitärer Vektorraum ist, selbst einen Vektorraum bilden, wobei wir die Multiplikation mit einer komplexen Zahl und die Summe zweier Operatoren „punktweise“ definieren, d.h. sind \hat{A}_1 und \hat{A}_2 beliebige Operatoren auf V und $\lambda_1, \lambda_2 \in \mathbb{C}$, so definieren wir den Operator $\lambda_1\hat{A}_1 + \lambda_2\hat{A}_2$ dadurch, daß für jeden Vektor $|\psi\rangle \in V$

$$(\lambda_1\hat{A}_1 + \lambda_2\hat{A}_2)|\psi\rangle = \lambda_1\hat{A}_1|\psi\rangle + \lambda_2\hat{A}_2|\psi\rangle \quad (\text{A.4.1})$$

gelten soll. Man weist schnell nach (*Übung*), daß die linearen Operatoren einen komplexen Vektorraum bilden. Diesen Vektorraum nennt man $\text{End}(V)$ ¹.

Da nun jedem Operator umkehrbar eindeutig eine $\mathbb{C}^{d \times d}$ -Matrix entspricht, indem man für einen gegebenen Operator \hat{A} die Matrixelemente $A_{jk} = \langle e_j | \hat{A} | e_k \rangle$ bzgl. eines beliebigen VONS $\{|e_j\rangle\}_{j \in \{1, \dots, d\}}$ berechnet. Hat man umgekehrt eine Matrix $\tilde{A} = (A_{jk}) \in \mathbb{C}^{d \times d}$ gegeben, definiert

$$\hat{A} = \sum_{j,k=1}^d A_{jk} |j\rangle \langle k| \quad (\text{A.4.2})$$

einen Operator \hat{A} mit den vorgegebenen Matrixelementen A_{jk} (*nachprüfen!*).

Wir können nun über diese darstellenden Matrizen von Operatoren einige nützliche Funktionale definieren, also lineare Abbildungen $\text{End}(V) \rightarrow \mathbb{C}$. Natürlich sollten diese Abbildungen unabhängig vom gewählten Orthonormalsystem (noch besser sogar unabhängig von irgendeiner nicht notwendig orthonormalen Basis) von V sein. Es ist klar, daß dies auf die **Determinante** zutrifft, d.h. es definiert

$$\det \hat{A} = \det \tilde{A} \quad (\text{A.4.3})$$

eine von der gewählten Basis unabhängige Abbildung $\det : \text{End}(V) \rightarrow \mathbb{C}$.

Eine weitere im folgenden wichtige Abbildung ist die **Spur** eines Operators, die durch die entsprechende Spur der Matrix (also die Summe der Diagonalelemente) definiert ist²:

$$\text{tr} \hat{A} = \text{tr} \tilde{A} = \sum_{j=1}^d A_{jj} = \sum_{j=1}^d \langle e_j | \hat{A} | e_j \rangle. \quad (\text{A.4.4})$$

Wir zeigen, daß die Spur unabhängig von der Wahl des VONS ist. Ist nämlich $\{|e'_k\rangle\}_{k \in \{1, \dots, d\}}$ ebenfalls ein VONS, so gilt wegen der Vollständigkeitsrelation

$$\begin{aligned} \text{tr} \tilde{A}' &= \sum_{j=1}^d \langle e'_j | \hat{A} | e'_j \rangle \\ &= \sum_{j,k,l=1}^d \langle e'_j | e_k \rangle \langle e_k | \hat{A} | e_l \rangle \langle e_l | e'_j \rangle \\ &= \sum_{j,k,l=1}^d \langle e_l | e'_j \rangle \langle e'_j | e_k \rangle \langle e_k | \hat{A} | e_l \rangle \\ &= \sum_{k,l=1}^d \delta_{lk} \langle e_k | \hat{A} | e_l \rangle \\ &= \sum_{k=1}^d \langle e_k | \hat{A} | e_k \rangle = \text{tr} \tilde{A}. \end{aligned} \quad (\text{A.4.5})$$

Die Spur ist also tatsächlich unabhängig von dem zu ihrer Berechnung verwendeten VONS.

Als nächstes wollen wir das Ziel verwirklichen, eine Norm für Operatoren zu definieren. Da die linearen Operatoren über V den komplexen $d \times d$ -dimensionalen Vektorraum $\text{End}(V)$ bilden, liegt es nahe, gleich

¹Dabei steht End für **Endomorphismen**, was einfach lineare Abbildung bzw. linearer Operator bedeutet.

²Das Formelzeichen tr steht für engl. Trace=Spur.

ein **Skalarprodukt** zu definieren. Für zwei Operatoren \hat{A}_1 und \hat{A}_2 definieren wir dazu

$$\langle \hat{A}_1 | \hat{A}_2 \rangle = \text{tr}(\hat{A}_1^\dagger \hat{A}_2) = \sum_{j,k=1}^d A_{1,jk}^* A_{2,jk}. \quad (\text{A.4.6})$$

Daß tatsächlich ein Skalarprodukt (also eine positiv definite Sesquilinearform) vorliegt, ist anhand des zuletzt angegebenen expliziten Ausdrucks über Matrixelemente klar, denn es handelt sich ja um nichts anderes als das kanonische Skalarprodukt des unitären Vektorraums $\mathbb{C}^{d \times d}$. Entsprechend definieren wir die **Frobenius-Norm** auf $\text{End}(V)$ durch

$$\|\hat{A}\| = \sqrt{\langle \hat{A} | \hat{A} \rangle} = \sum_{j,k=1}^d |A_{jk}|^2. \quad (\text{A.4.7})$$

Nun läßt sich der Vektorraum $\text{End}(V)$ noch zu einer **Algebra**³ erweitern, indem wir als Produkt die Hintereinanderausführung von linearen Operatoren definieren. Wir bemerken, daß man unter der Spur die Reihenfolge der Operatoren vertauschen darf, auch wenn i.a. $\hat{A}\hat{B} \neq \hat{B}\hat{A}$ ist:

$$\begin{aligned} \text{tr}(\hat{A}\hat{B}) &= \sum_{j=1}^d \langle e_j | \hat{A}\hat{B} | e_j \rangle = \sum_{j,k=1}^d \langle e_j | \hat{A} | e_k \rangle \langle e_k | \hat{B} | e_j \rangle \\ &= \sum_{j,k=1}^d \langle e_k | \hat{B} | e_j \rangle \langle e_j | \hat{A} | e_k \rangle = \sum_{k=1}^d \langle e_k | \hat{B}\hat{A} | e_k \rangle = \text{tr}(\hat{B}\hat{A}). \end{aligned} \quad (\text{A.4.8})$$

Für das folgende ist es wichtig, daß wir nicht nur eine Norm im Sinne eines Vektorraumes vorliegen haben, sondern auch die **Submultiplikativität**

$$\|\hat{A}\hat{B}\| \leq \|\hat{A}\| \|\hat{B}\| \quad (\text{A.4.9})$$

gilt. Zum Beweis schreiben wir die Norm mit Hilfe der darstellenden Matrizen aus

$$\|\hat{A}\hat{B}\|^2 = \sum_{j,l=1}^d \left| \sum_{k=1}^d A_{jk} B_{kl} \right|^2 \quad (\text{A.4.10})$$

Wir definieren nun die folgenden Vektoren

$$|\bar{A}_j\rangle = \sum_{k=1}^d |e_k\rangle A_{jk}^*, \quad |B_l\rangle = \sum_{k=1}^d |e_k\rangle B_{kl}. \quad (\text{A.4.11})$$

Dann ist

$$\langle \bar{A}_j | B_l \rangle = \sum_{k=1}^d A_{jk} B_{kl}, \quad (\text{A.4.12})$$

und mit der Cauchy-Schwarzschen Ungleichung für das Skalarprodukt folgt damit

$$\left| \sum_{k=1}^d A_{jk} B_{kl} \right|^2 = \left| \langle \bar{A}_j | B_l \rangle \right|^2 \leq \|\bar{A}_j\|^2 \|B_l\|^2. \quad (\text{A.4.13})$$

Dies in (A.4.10) verwendet liefert

$$\|\hat{A}\hat{B}\|^2 \leq \sum_{j=1}^d \|\bar{A}_j\|^2 \sum_{l=1}^d \|B_l\|^2 = \|\hat{A}\|^2 \|\hat{B}\|^2, \quad (\text{A.4.14})$$

³genauer gesagt zu einem Ring, aber dies ist für das folgende unwichtig

und das war zu zeigen.

Genauso beweist man (*Übung*), daß für jeden Vektor $|\psi\rangle$ und jeden Operator \hat{A} auch

$$\|\hat{A}\psi\| \leq \|A\| \|\psi\| \quad (\text{A.4.15})$$

gilt.

Nun lassen sich alle Begriffe bzgl. **Konvergenz** von Folgen, Reihen und operatorwertigen Funktionen wie für die analogen reell- bzw. komplexwertigen Objekte definieren. Sei z.B. (\hat{A}_n) eine Folge von Operatoren. Dann besitzt diese Folge den Grenzwert $\hat{A} \in \text{End}(V)$ definitionsgemäß genau dann, wenn zu jedem $\epsilon > 0$ ein $N \in \mathbb{N}$ existiert, so daß für alle $n > N$

$$\|\hat{A}_n - \hat{A}\| < \epsilon \quad (\text{A.4.16})$$

ist. Unendliche Reihen werden durch ihre **Teilsammenfolgen**

$$\hat{S}_n = \sum_{j=1}^n \hat{A}_j \quad (\text{A.4.17})$$

definiert und ihr evtl. existierender Grenzwert als der Grenzwert dieser Teilsammenfolge.

Im folgenden benötigen wir nur Potenzreihen von Matrizen, also Reihen der Form⁴

$$\hat{f}(\hat{A}) = \sum_{j=0}^{\infty} a_j \hat{A}^j. \quad (\text{A.4.18})$$

Wegen der Submultiplikativität der Frobenius-Norm für Matrizen folgt, daß diese Potenzreihe sicher für $\|\hat{A}\| < R_f$, wobei R_f der **Konvergenzradius** der entsprechenden Potenzreihe für reelle oder komplexe Funktionen

$$f(z) = \sum_{j=0}^{\infty} a_j z^j. \quad (\text{A.4.19})$$

Der Konvergenzradius läßt sich durch

$$R_f = \lim_{j \rightarrow \infty} \frac{a_j}{a_{j+1}} = \lim_{j \rightarrow \infty} a_j^{-1/j} \quad (\text{A.4.20})$$

berechnen, falls der jeweilige Grenzwert existiert. Vgl. dazu [Hee14].

A.5 Die Matrixexponentialfunktion

Als wichtigste Anwendung können wir die **Matrixexponentialfunktion** durch die entsprechende Potenzreihe

$$\exp \hat{A} = \sum_{j=0}^{\infty} \frac{1}{j!} \hat{A}^j \quad (\text{A.5.1})$$

definieren. Der Konvergenzradius ist offenbar ∞ , d.h. die Exponentialfunktion existiert für alle Matrizen \hat{A} :

$$R_{\text{exp}} = \lim_{j \rightarrow \infty} \frac{1/j!}{1/(j+1)!} = \lim_{j \rightarrow \infty} \frac{(j+1)!}{j!} = \lim_{j \rightarrow \infty} (j+1) = \infty. \quad (\text{A.5.2})$$

Wir betrachten einige wichtige Eigenschaften der Matrixexponentialfunktion.

⁴Wir definieren dabei, daß für alle Matrizen $\hat{A}^0 = \mathbb{1}$ ist.

Wichtig sind im folgenden Matrixexponentialfunktionen, die von einem (reellen oder komplexen) Parameter t gemäß

$$\hat{B}(t) = \exp(t\hat{A}) \quad (\text{A.5.3})$$

abhängen. Wie bei der reellen oder komplexen Exponentialfunktion kann man zeigen, daß man Potenzreihen innerhalb ihres Konvergenzradiuses differenzieren kann, und daß die Ableitung der Potenzreihe durch die Ableitung ihrer Reihenglieder gegeben ist, d.h. die Summation und die Differentiation darf auch für die Matrixexponentialfunktion vertauscht werden. Daraus folgt

$$\begin{aligned} \frac{d}{dt}\hat{B}(t) &= \frac{d}{dt} \sum_{j=0}^{\infty} \frac{1}{j!} t^j \hat{A}^j \\ &= \sum_{j=1}^{\infty} \frac{1}{(j-1)!} t^{j-1} \hat{A}^j \\ &= \hat{A} \sum_{j=1}^{\infty} \frac{1}{(j-1)!} t^{j-1} \hat{A}^{j-1} \\ &= \hat{A} \sum_{k=0}^{\infty} \frac{1}{k!} t^k \hat{A}^k \\ &= \hat{A} \exp(t\hat{A}) = \hat{A}\hat{B}(t). \end{aligned} \quad (\text{A.5.4})$$

Weiter gilt für jede invertierbare Matrix \hat{C}

$$\exp(\hat{C}\hat{A}\hat{C}^{-1}) = \hat{C} \exp(\hat{A}) \hat{C}^{-1}, \quad (\text{A.5.5})$$

denn es ist für $j \in \mathbb{N}_0$

$$(\hat{C}\hat{A}\hat{C}^{-1})^j = \hat{C}\hat{A}^j\hat{C}^{-1} \quad (\text{A.5.6})$$

Dies läßt sich durch vollständige Induktion leicht beweisen, denn die Behauptung ist sicher wahr für $j = 0$, denn dann steht auf beiden Seiten der Gleichung die Einheitsmatrix. Angenommen, die Behauptung gilt für $j = k$, so folgt

$$(\hat{C}\hat{A}\hat{C}^{-1})^{k+1} = (\hat{C}\hat{A}\hat{C}^{-1})^k \hat{C}\hat{A}\hat{C}^{-1} = \hat{C}\hat{A}^k\hat{C}^{-1}\hat{C}\hat{A}\hat{C}^{-1} = \hat{C}\hat{A}^k\hat{A}\hat{C}^{-1} = \hat{C}\hat{A}^{k+1}\hat{C}^{-1}, \quad (\text{A.5.7})$$

und das ist die Behauptung für $j = k + 1$ und damit (A.5.6) bewiesen. Damit folgt aber auch (A.5.5), denn wir können (A.5.6) auf die Glieder der Exponentialreihe anwenden. Daraus ergibt sich eine Möglichkeit, für eine gegebene Matrix \hat{A} die Matrixexponentialfunktion $\exp \hat{A}$ auszurechnen. Angenommen, wir können \hat{A} diagonalisieren. Dann gibt es eine Matrix \hat{C} , so daß

$$\hat{A}' = \hat{C}\hat{A}\hat{C}^{-1} = \text{diag}(\lambda_1, \dots, \lambda_d). \quad (\text{A.5.8})$$

Nun ist offenbar

$$\hat{A}'^j = \text{diag}(\lambda_1^j, \dots, \lambda_d^j) \Rightarrow \exp(\hat{A}') = \text{diag}[\exp(\lambda_1), \dots, \exp(\lambda_d)]. \quad (\text{A.5.9})$$

Andererseits folgt aus (A.5.8)

$$\hat{A} = \hat{C}^{-1}\hat{A}'\hat{C} \Rightarrow \exp \hat{A} = \hat{C}^{-1} \exp(\hat{A}') \hat{C} = \hat{C}^{-1} \text{diag}[\exp(\lambda_1), \dots, \exp(\lambda_d)] \hat{C}. \quad (\text{A.5.10})$$

Oft ist auch noch die folgende Formel nützlich:

$$\exp(\hat{A})\hat{B}\exp(-\hat{A}) = \sum_{j=0}^{\infty} \frac{1}{j!} [\hat{A}, \hat{B}]_j. \quad (\text{A.5.11})$$

Dabei ist der sog. **Multikommutator** rekursiv durch

$$[\hat{A}, \hat{B}]_0 = \hat{B}, \quad [\hat{A}, \hat{B}]_{j+1} = [\hat{A}, [\hat{A}, \hat{B}]_j] \quad (\text{A.5.12})$$

definiert. Demnach sind die ersten Ausdrücke

$$[\hat{A}, \hat{B}]_1 = [\hat{A}, \hat{B}], \quad [\hat{A}, \hat{B}]_2 = [\hat{A}, [\hat{A}, \hat{B}]] \quad \text{usw.} \quad (\text{A.5.13})$$

Zum Beweis von (A.5.11) betrachten wir die Funktion

$$\hat{F}(t) = \exp(t\hat{A})\hat{B}\exp(-t\hat{A}). \quad (\text{A.5.14})$$

Bilden wir die Ableitung nach t , ergibt sich

$$\frac{d}{dt}\hat{F}(t) = \dot{\hat{F}}(t) = \exp(t\hat{A})[\hat{A}, \hat{B}]\exp(-t\hat{A}). \quad (\text{A.5.15})$$

Nun ist

$$\exp(t\hat{A})\hat{A}\exp(-t\hat{A}) = \exp(t\hat{A})\exp(-t\hat{A})\hat{A} = \hat{A}, \quad (\text{A.5.16})$$

weil offensichtlich \hat{A} mit jeder Funktion von \hat{A} kommutiert. Daraus folgt nun

$$\dot{\hat{F}}(t) = [\exp(t\hat{A})\hat{A}\exp(-t\hat{A}), \exp(t\hat{A})\hat{B}\exp(-t\hat{A})] = [\hat{A}, \hat{F}(t)]. \quad (\text{A.5.17})$$

Damit ergibt sich dann weiter durch Iteration

$$\frac{d^j}{dt^j}\hat{F}(t) = \hat{F}^{(j)}(t) = [\hat{A}, \hat{F}(t)]_j \Rightarrow \hat{F}^{(j)}(0) = [\hat{A}, \hat{B}]_j. \quad (\text{A.5.18})$$

Nun können wir die Taylorsche Formel verwenden

$$\hat{F}(t) = \sum_{j=0}^{\infty} \frac{1}{j!} \hat{F}^{(j)}(0) t^j = \sum_{j=0}^{\infty} \frac{1}{j!} [\hat{A}, \hat{B}]_j t^j. \quad (\text{A.5.19})$$

Setzen wir hierin $t = 1$, erhalten wir (A.5.11).

Als eine Anwendung dieser Formel betrachten wir nun zwei Matrizen \hat{A} und \hat{B} , für die

$$[[\hat{A}, \hat{B}], \hat{A}] = [[\hat{A}, \hat{B}], \hat{B}] = 0 \quad (\text{A.5.20})$$

gilt. Wir wollen beweisen, daß dann

$$\exp(\hat{A} + \hat{B}) = \exp(\hat{A})\exp(\hat{B})\exp\left(-\frac{1}{2}[\hat{A}, \hat{B}]\right) \quad (\text{A.5.21})$$

gilt.

Dazu betrachten wir die matrixwertige Funktion

$$\hat{F}(t) = \exp[t(\hat{A} + \hat{B})]. \quad (\text{A.5.22})$$

Wenden wir nun (A.5.19) an, folgt wegen (A.5.20)

$$\hat{F}(t)\hat{A}\hat{F}^{-1}(t) = \hat{A} + t[\hat{A} + \hat{B}, \hat{A}] = \hat{A} - t[\hat{A}, \hat{B}]. \quad (\text{A.5.23})$$

Multiplizieren wir diese Gleichung von rechts mit $\hat{F}(t)$, folgt

$$\hat{F}(t)\hat{A} = \hat{A}\hat{F}(t) - t[\hat{A}, \hat{B}]\hat{F}(t). \quad (\text{A.5.24})$$

Leiten wir andererseits (A.5.22) nach t ab, folgt

$$\dot{\hat{F}}(t) = \exp[t(\hat{A} + \hat{B})](\hat{A} + \hat{B}) = \hat{F}(t)(\hat{A} + \hat{B}). \quad (\text{A.5.25})$$

Mit (A.5.25) folgt weiter

$$\dot{\hat{F}} = \hat{A}\hat{F}(t) + \hat{F}(t)\hat{B} - t[\hat{A}, \hat{B}]\hat{F}(t). \quad (\text{A.5.26})$$

Wegen (A.5.20) wird diese Differentialgleichung für $\hat{F}(t)$ offenbar durch

$$\hat{F}(t) = \exp(t\hat{A})\exp(t\hat{B})\exp\left(-\frac{t^2}{2}[\hat{A}, \hat{B}]\right) \quad (\text{A.5.27})$$

gelöst (*nachprüfen durch Bilden der Ableitung dieser Gleichung nach t !*). Für $t = 1$ folgt (A.5.21). Wir bemerken, daß aus dieser Gleichung

$$\exp(\hat{A} + \hat{B}) = \exp(\hat{A})\exp(\hat{B}) \quad \text{falls} \quad [\hat{A}, \hat{B}] = 0 \quad (\text{A.5.28})$$

gilt. Es ist wichtig, sich daran zu erinnern, daß diese für reelle und komplexe Zahlen geltende Summenformel der Exponentialfunktion für die Matrixexponentialfunktion dann und nur dann erfüllt ist, wenn die involvierten Matrizen kommutieren!

Daraus folgt dann auch für $\hat{B} = -\hat{A}$, daß

$$\exp(\hat{A} - \hat{A}) = \mathbb{1} = \exp(\hat{A})\exp(-\hat{A}) \Rightarrow [\exp(\hat{A})]^{-1} = \exp(-\hat{A}). \quad (\text{A.5.29})$$

Eine Matrix, die das Matrixexponential einer beliebigen Matrix \hat{A} ist, ist also stets invertierbar, und es gilt (A.5.29).

Schließlich betrachten wir noch die Determinante

$$D = \det \exp(\hat{A}). \quad (\text{A.5.30})$$

Wir definieren dazu wieder die matrixwertige Funktion

$$\hat{F}(t) = \exp(t\hat{A}) \quad (\text{A.5.31})$$

und

$$D(t) = \det \hat{F}(t). \quad (\text{A.5.32})$$

Nun gilt

$$\begin{aligned} D(t + \delta t) &= \det \hat{F}(t + \delta t) \\ &= \det[\hat{F}(t) + \delta t \dot{\hat{F}}(t)] + \mathcal{O}(\delta t^2) \\ &= \det\{\hat{F}(t)[\mathbb{1} + \delta t \hat{A}]\} + \mathcal{O}(\delta t^2) \\ &= \det \hat{F}(t) \det(\mathbb{1} + \delta t \hat{A}) + \mathcal{O}(\delta t^2) \\ &= D(t) \det(\mathbb{1} + \delta t \hat{A}) + \mathcal{O}(\delta t^2). \end{aligned} \quad (\text{A.5.33})$$

Nun wenden wir die Definition der Determinante auf den zweiten Faktor an, wobei wir das Resultat nur bis zur Ordnung $\mathcal{O}(\delta t)$ benötigen. Es gilt

$$\det(\mathbb{1} + \delta t \hat{A}) = \sum_{P \in S_d} \prod_{k=1}^d (-1)^{\text{sign} P} \left(\delta_{kP(k)} + \delta t A_{kP(k)} \right). \quad (\text{A.5.34})$$

Dabei durchläuft P die Menge der Permutationen von $(1, 2, \dots, d)$. Der Beitrag zur Ordnung δt rührt offenbar nur vom Produkt der d Diagonalelemente her, d.h. für die identische Permutation $P(k) = k$. Also gilt

$$\begin{aligned} \det(\mathbb{1} + \delta t \hat{A}) &= \prod_{k=1}^d \left(\delta_{kP(k)} + \delta t A_{kP(k)} \right) + \mathcal{O}(\delta t^2) \\ &= 1 + \delta t \sum_{k=1}^d A_{kk} + \mathcal{O}(\delta t^2) \\ &= 1 + \delta t \text{tr} \hat{A} + \mathcal{O}(\delta t^2). \end{aligned} \quad (\text{A.5.35})$$

Setzen wir dies in (A.5.33) ein, und formen ein wenig um, folgt im Limes $\delta t \rightarrow 0$ die Differentialgleichung

$$\dot{D}(t) = \text{tr} \hat{A} D(t). \quad (\text{A.5.36})$$

Wegen $D(0) = 1$ ergibt sich als Lösung

$$D(t) = \exp(t \text{tr} \hat{A}). \quad (\text{A.5.37})$$

Für $t = 1$ ergibt sich damit die wichtige Formel

$$\det \exp(\hat{A}) = \exp(\text{tr} \hat{A}). \quad (\text{A.5.38})$$

A.6 Grundlagen der Gruppentheorie

Wir fassen hier einige grundlegende Eigenschaften von **Gruppen** zusammen. Zunächst einmal ist eine Gruppe eine Menge G mit einer Abbildung $g_1, g_2 \in G \mapsto g_1 g_2 \in G$ mit den folgenden Eigenschaften

$$(g_1 g_2) g_3 = g_1 (g_2 g_3), \quad (\text{A.6.1})$$

$$\exists e \in G \forall g \in G: \quad e g = g, \quad (\text{A.6.2})$$

$$\forall g \in G \exists g^{-1} \in G: \quad g^{-1} g = e. \quad (\text{A.6.3})$$

Man nennt e ein **linksneutrales Element** der Gruppe, und g^{-1} ein **linksinverses Element** von $g \in G$.

Wir zeigen nun einige elementare Eigenschaften von Gruppen. Zunächst gilt

$$g_1 g_2 = e \Rightarrow g_2 g_1 = e. \quad (\text{A.6.4})$$

Sei $g_1 g_2 = e$ Es gibt voraussetzungsgemäß ein linksinverses Element g_1^{-1} zu g_1 . Damit folgt

$$g_2 g_1 = (e g_2) g_1 = ((g_1^{-1} g_1) g_2) g_1 = (g_1^{-1} (g_1 g_2)) g_1 = (g_1^{-1} e) g_1 = g_1^{-1} (e g_1) = g_1^{-1} g_1 = e, \quad (\text{A.6.5})$$

und das war zu zeigen.

Weiter zeigen wir, daß das linksinverse Element g^{-1} zu jedem $g \in G$ eindeutig bestimmt ist. Sei dazu x irgendein weiteres Linksinverses zu g , d.h. sei $x g = e$. Dann gilt wegen (A.6.4) $g g^{-1} = e$ und damit

$$(x g) g^{-1} = e g^{-1} = g^{-1} = x (g g^{-1}) = x, \quad (\text{A.6.6})$$

d.h. $x = g^{-1}$.

Ebenso folgt, daß e auch ein rechtsneutrales Element ist, denn für $g \in G$ folgt

$$ge = g(g^{-1}g) = (gg^{-1})g = eg = g. \quad (\text{A.6.7})$$

Außerdem gibt es in einer Gruppe auch genau ein neutrales Element. Seien nämlich e und e' neutrale Elemente, so folgt sofort

$$e' = ee' = e \quad (\text{A.6.8})$$

und $e^{-1} = e$, denn für alle $g \in G$ ist

$$e^{-1}g = e^{-1}(eg) = (e^{-1}e)g = eg = g. \quad (\text{A.6.9})$$

Weiter ist

$$(g_1g_2)^{-1} = g_2^{-1}g_1^{-1}, \quad (\text{A.6.10})$$

denn es ist $(g_2^{-1}g_1^{-1})(g_1g_2) = g_2^{-1}[g_1^{-1}(g_1g_2)] = g_2^{-1}[(g_1^{-1}g_1)g_2] = g_2^{-1}(eg_2) = g_2^{-1}g_2 = e$, und damit folgt aus der Eindeutigkeit des Elements zu $(g_1g_2) \in G$ (A.6.10).

A.7 Die Gruppen $SU(n)$ und ihre Lie-Algebren

Literaturverzeichnis

- [ABH⁺13] T. Arens, et al., *Grundwissen Mathematikstudium*, Springer-Verlag, Berlin, Heidelberg (2013).
<http://dx.doi.org/10.1007/978-3-8274-2309-2>
- [AHK⁺15] T. Arens, et al., *Mathematik*, Springer-Verlag, Berlin, Heidelberg (2015).
<http://dx.doi.org/10.1007/978-3-642-44919-2>
- [Art91] B. Artmann, *Lineare Algebra*, Birkhäuser Verlag, Basel (1991).
- [BFK⁺15] M. Bartelmann, et al., *Theoretische Physik*, Springer-Verlag, Berlin, Heidelberg (2015).
<http://dx.doi.org/10.1007/978-3-642-54618-1>
- [CH24] R. Courant, D. Hilbert, *Methoden der mathematischen Physik 1*, Springer Verlag, Berlin (1924).
<http://gdz.sub.uni-goettingen.de/dms/load/toc/?PID=PPN380672502>
- [Dir58] P. A. Dirac, *Generalized Hamiltonian dynamics*, Proc. Roy. Soc. Lond. **A246**, 326 (1958).
<http://dx.doi.org/10.1098/rspa.1958.0141>
- [Ein05] A. Einstein, *Zur Elektrodynamik bewegter Körper*, Ann. Phys. (Leipzig) **17**, 891 (1905).
<http://dx.doi.org/10.1002/andp.19053221004>
- [Fis10] G. Fischer, *Lineare Algebra*, 17. Aufl., Vieweg+Teubner Verlag, Stuttgart (2010).
<http://dx.doi.org/10.1007/978-3-8348-9365-9>
- [Gre03] W. Greiner, *Klassische Mechanik 1*, 7. Aufl., Verlag Harri Deutsch, Frankfurt/Main (2003).
- [Hee08] H. van Hees, *Klassische Mechanik* (2008).
<http://theory.gsi.de/~vanhees/faq-pdf/mech.pdf>
- [Hee14] H. van Hees, *Mathematische Ergänzungen zu Theoretische Physik 1* (2014).
<http://fias.uni-frankfurt.de/~hees/math-erg1-ws14/matherg1.pdf>
- [Joo89] G. Joos, *Lehrbuch der theoretischen Physik*, 15. Aufl., Aulaverlag, Wiesbaden (1989).
- [LL97] L. D. Landau, E. M. Lifschitz, *Lehrbuch der Theoretischen Physik, Bd. 1, Mechanik*, 14. Aufl., Verlag Harri Deutsch, Frankfurt/Main (1997).
- [Min09] H. Minkowski, *Raum und zeit*, Jahresbericht der Deutschen Mathematiker-Vereinigung, vol. 18, p. 75-88 **18**, 75 (1909).
- [Nol13] W. Nolting, *Grundkurs Theoretische Physik 1: Klassische Mechanik*, 10. Aufl., Springer Spektrum, Berlin, Heidelberg (2013).
<http://dx.doi.org/10.1007/978-3-642-29937-7>
- [Sch07] F. Scheck, *Theoretische Physik 1: Mechanik*, 8. Aufl., Springer-Verlag, Berlin, Heidelberg, New York (2007).

Literaturverzeichnis

- [Sil07] L. Silberstein, *Elektromagnetische Grundgleichungen in bivectorieller Behandlung*, Annalen der Physik **327**, 579 (1907).
<http://dx.doi.org/10.1002/andp.19073270313>