
Theoretische Physik III
für das Lehramt L3
(ältere Fassung)

H. van Hees

16. April 2021

Inhaltsverzeichnis

Inhaltsverzeichnis	3
1 Einleitung	5
1.1 Bemerkung zur Interpretation der QT	6
1.2 Lehrbücher	6
2 Grundlagen der Quantentheorie: Polarisationsmessung an Photonen	7
2.1 Polarisationsmessung an klassischen elektromagnetischen Wellen	7
2.2 Polarisationsmessungen an Photonen	10
3 Wellenmechanik	19
3.1 Die Schrödinger-Gleichung für freie Teilchen	19
3.2 Der Hilbert-Raum	24
3.3 Die Observablen-Operatoren	26
3.4 Verträglichkeit von Observablen	30
3.5 Dynamik	35
3.6 Energieeigenvektoren	40
3.7 Formale Herleitung der Wellenmechanik	41
3.8 Teilchen im Potentialtopf	42
3.9 Der harmonische Oszillator	44
3.10 Der symmetrische 3D harmonische Oszillator	49
3.11 Der Bahndrehimpuls	50
3.12 Bahndrehimpulseigenzustände in der Ortsdarstellung	54
3.13 Das Wasserstoffatom	58
3.13.1 Schwerpunkts- und Relativkoordinaten	58
3.13.2 Energieeigenwertproblem für die Relativbewegung	60
3.13.3 Die Radialwellengleichung und Energieeigenwerte	61
3.14 Der Spin	65
3.14.1 Heuristische Einführung des Spins	65
3.14.2 Magnetisches Moment einer klassischen Stromverteilung	67
3.14.3 Teilchen mit $s = 0$ im Magnetfeld	69
3.14.4 Teilchen mit Spin im Magnetfeld	70
3.14.5 Teilchen im homogenen Magnetfeld	71

3.15	Der Stern-Gerlach-Versuch	71
4	Spezielle Relativitätstheorie	75
4.1	Das speziell-relativistische Raum-Zeit-Modell	75
4.2	Kinematische Effekte	81
4.2.1	Relativität der Gleichzeitigkeit	81
4.2.2	Zeitdilatation	82
4.2.3	Längenkontraktion	82
4.2.4	Einsteins Zug-Gedankenexperiment	83
4.2.5	Das Michelson-Morley-Experiment	84
4.2.6	Bemerkungen zur Historie der Lorentz-Transformation	87
4.3	Das Zwillingsparadoxon	88
4.4	Speziell relativistische Mechanik	89
4.5	Zerfall und Stöße von Teilchen	93
4.5.1	Zweiteilchenzerfall	94
4.5.2	Zweiteilchenstöße	95
4.5.3	Kinematik im Schwerpunktssystem	95
4.5.4	Kinematik im Laborsystem	96
4.6	Äquivalenz zwischen Energie und Masse	97
4.6.1	Voll inelastischer Stoß	97
4.6.2	Zerfall eines Pions in zwei Photonen	98
4.7	Geladene Teilchen im elektromagnetischen Feld	98
4.7.1	Teilchen im homogenen elektrischen Feld	100
4.8	Kovariante Formulierung der Elektrodynamik	102
4.8.1	Tensoranalysis und Maxwell-Gleichungen im Minkowski-Raum	102
4.8.2	Die elektromagnetischen Potentiale	105
4.8.3	Transformationsverhalten der Felder	107
4.8.4	Feld einer gleichförmig bewegten Punktladung	107
4.8.5	Dopplereffekt und Aberration bei Licht	108
5	Elementarteilchen	113
A	Spezielle Funktionen	115
A.1	Hermite-Polynome	115
A.2	Kugelflächenfunktionen	115
A.3	Laguerre- und assoziierte Laguerre-Polynome	115
	Literaturverzeichnis	119

Kapitel 1

Einleitung

Die sog. „moderne Physik“ umfasst die Quantentheorie und die Relativitätstheorie. Wir können uns in dieser einsemestrigen Vorlesung allerdings nur mit den wesentlichen Prinzipien dieses sehr umfassenden Gebietes beschäftigen, umfasst doch die Relativitätstheorie die gesamte Physik, indem sie das Raumzeitmodell an die Erfahrungsstatsachen anpasst, denn überraschenderweise erweist sich die Galilei-Newton-Raumzeit der klassischen Mechanik als inkonsistent mit den Erscheinungen des Elektromagnetismus, die wir im vorigen Semester ausführlich besprochen haben. In dieser Vorlesung beschränken wir uns auf die „Spezielle Relativitätstheorie“, die nach einigen Dekaden historischer Verwirrung von Einstein in einem der wohl berühmtesten Artikel der Physikgeschichte [Ein05] zu einer radikalen Korrektur unserer Vorstellungen von Raum und Zeit geführt haben. Zunächst war aber die Relativitätstheorie noch auf die sog. „klassische Physik“ beschränkt, d.h. man musste die Newtonsche Mechanik an die neue Raumzeitstruktur anpassen, und die Maxwell-Theorie des Elektromagnetismus, also die klassische Elektrodynamik, erwies sich als paradigmatisches Beispiel einer klassischen relativistischen Feldtheorie. Abgesehen von der Korrektur des Raumzeitmodells, die mit der Entdeckung der Allgemeinen Relativitätstheorie durch Einstein und Hilbert nochmals verfeinert wurde, konnte allerdings das Grundprinzip eines deterministischen Weltbildes aufrechterhalten werden.

Der wirklich umfassendste Umbruch dieses Weltbildes der klassischen Physik stellt allerdings zweifelsohne die Quantentheorie dar, die ziemlich radikal mit den Vorstellungen über den Aufbau der Materie gebrochen hat. In dieser Vorlesung werden wir die nichtrelativistische Quantenmechanik recht ausführlich besprechen. Leider können wir auf die speziell-relativistische Erweiterung nur recht qualitativ und zum Ende der Vorlesung eingehen. Man kann aber wohl mit einigem Recht behaupten, dass diese relativistische Quantenfeldtheorie und das auf ihr basierende Standardmodell der Elementarteilchen derzeit die umfassendste Theorie der Physik darstellt, die die Phänomene mit zum Teil erstaunlicher Präzision vorhersagt (für manchen Größen, wie das magnetische Moment des Elektrons, auf 12 signifikante Dezimalstellen genau!).

Das Standardmodell erfasst allerdings nur die (bislang) beobachtbaren Teilchen (also auf elementarer Ebene Leptonen, Quarks, Eichbosonen und Higgs-Boson) und nur drei von vier fundamentalen Wechselwirkungen (Elektromagnetismus, starke und schwache Wechselwirkung). Die Gravitation konnte bislang noch nicht quantentheoretisch in befriedigender Weise beschrieben werden. Allerdings macht sich die Gravitation merklich auch nur anhand makroskopischer Objekte bemerkbar, und daher ist eine Beobachtung von „Quantenphänomenen“ im Zusammenhang mit der Gravitationswechselwirkung schwierig. Die Allgemeine Relativitätstheorie (ART) liefert allerdings eine äußerst genaue Beschreibung all dieser makroskopischen Phänomene bis hin zu einem Verständnis des Universums als ganzem im Rahmen der auf der ART beruhenden Kosmologie. Auch dieses faszinierende Thema, das seinerseits eng auch mit der Teilchenphysik zusammenhängt, können wir nur qualitativ streifen.

1.1 Bemerkung zur Interpretation der QT

Die Quantentheorie (QT) ist zwar in ihrer modernen Fassung bereits fast 100 Jahre alt. Trotzdem ist es bezeichnend, dass in Teilen der Physik-Community bis dato noch keine Einigkeit über ihre Interpretation besteht. Wir werden darauf in Kapitel 2 gelegentlich eingehen. Das Problem dabei besteht darin, dass die Interpretationsfragen oft von philosophischen Bedenken unnötig verkompliziert werden, wobei allerdings die philosophisch-erkenntnistheoretische Reflexion über die Physik im Allgemeinen und erst Recht eben auch der QT im Besonderen gerade im Hinblick auf deren didaktische Vermittlung auch wichtig ist.

In dieser Vorlesung folge ich strikt der sog. „minimalen statistischen Interpretation“, die einfach den mathematischen Formalismus der QT als im in der Physik üblichen Wechselspiel zwischen Theorie und Experiment historisch entstanden ist und an „metaphysischen“ Zusatzannahmen nur das unbedingt für die Anwendung des Formalismus auf „reale“ Experimente Benötigte postuliert [Bal70, Bal98]. Demnach sind die sich aus dem Formalismus ergebenden Vorhersagen ausschließlich die Wahrscheinlichkeiten für das Auftreten von Werten von Observablen, wenn eine Messung vorgenommen wird, und deren experimentelle Überprüfung erfordert die Präparation vieler gleichartiger Quantensysteme, also eines Ensembles, an dem die entsprechenden Messungen vorgenommen wird. Das Messproblem, das dabei darin besteht, wie eigentlich durch Wechselwirkung des vermessenen Objekts mit dem Messapparat für jede Einzelmessung ein eindeutiger und damit indiskutabler Messwert resultiert, wird dabei ignoriert, und gerade dieses Messproblem ist die wesentliche Ursache für die Uneinigkeit der Physiker über die Interpretation der QT. Einen gut lesbaren Überblick über diese im Grenzbereich zwischen Physik und Philosophie angesiedelten Probleme findet sich in [FKL⁺18]. Ein eher physikalisches Buch, das auf gut verständliche Weise Einblick in mögliche alternative Interpretationen bzw. Erweiterungen der QT bietet, ist [DL18]. Beide Bücher sind ebenfalls zum kostenlosen Download im Netzwerk der Goethe-Universität verfügbar.

1.2 Lehrbücher

Die Lehrbücher zu den genannten Themen sind Legion. Für die Quantentheorie habe ich mich in Kapitel 2 für die Quantentheorie an [ST93] und [SW10] orientiert. Einen traditionelleren Zugang über die Wellenmechanik, die wir in Kapitel 3 besprechen werden, bietet Bd. 3 der Feynman Lectures [FLS07]. Hervorragend ist auch [Fic79].

Ein sehr schönes neues Lehrbuch der gesamten theoretischen Physik ist das in diesem Jahr in einer Neuauflage in 4 Bänden erschienene Buch [BFK⁺18a, BFK⁺18b, BFK⁺18c, BFK⁺18d]. Der Stoff dieser Vorlesung ist auf die ersten drei Bände verteilt. Dort findet sich insbesondere auch die Relativitätstheorie sehr schön aufbereitet vor. Diese Bücher lassen sich im Netzwerk der Goethe-Universität kostenfrei herunterladen. Die in diesem Skript gegebene Herleitung der Lorentz-Transformation stammt von Einstein [Ein07]. Sehr lesenswert sind auch die Lehrbücher von Einstein [Ein09a, Ein09b], die ebenfalls im Netzwerk der Goethe-Universität kostenfrei heruntergeladen werden können.

Frankfurt, im Sommer 2018.

Kapitel 2

Grundlagen der Quantentheorie: Polarisationsmessung an Photonen

In diesem einleitenden Kapitel besprechen wir anhand des einfachsten denkbaren „Quantensystems“ die formalen Grundlagen der Quantentheorie. Wir verwenden dabei einen naiven Begriff von **Photonen**, also der **Quanten des elektromagnetischen Feldes** wobei wir uns auf die Beobachtung des Polarisationszustandes von ebenen elektromagnetischen Wellen, die sich in x_3 -Richtung ausbreiten, beschränken wollen. Wir gehen von der entsprechenden klassischen elektromagnetischen Welle aus.

2.1 Polarisationsmessung an klassischen elektromagnetischen Wellen

Aus der Vorlesung Theorie 2 (vgl. [Hee18b], Abschnitt 5.1) wissen wir, dass eine ebene elektromagnetische Welle, die sich in x_3 -Richtung ausbreitet, durch das elektrische Feld in komplexer Darstellung in der Form

$$\vec{E} = E_0 \vec{\epsilon} \exp(ikx_3 - i\omega t) \quad (2.1.1)$$

beschrieben werden kann. Dabei ist

$$\vec{\epsilon} = \epsilon_1 \vec{e}_1 + \epsilon_2 \vec{e}_2 \quad (2.1.2)$$

mit i.a. beliebigen komplexen Zahlen $\epsilon_1, \epsilon_2 \in \mathbb{C}$. Wir interessieren uns im Folgenden nur für den Polarisationsvektor $\vec{\epsilon}$.

Wir ändern nun auch gleich die Notation für die Vektoren zu der in der Quantentheorie üblichen sog. **Braket-Schreibweise**, die auf Dirac zurückgeht, ab, da sich mit dieser Schreibweise etwas einfacher rechnen lässt. Statt \vec{e}_j schreiben wir $|e_j\rangle$. Diese Art Vektoren nennt man **Kets**¹. Beliebige Polarisationsvektoren der Art (2.1.2) schreiben wir als

$$|\epsilon\rangle = \epsilon_1 |e_1\rangle + \epsilon_2 |e_2\rangle. \quad (2.1.3)$$

Weiter definieren wir ein **Skalarprodukt**. Da es sich um einen komplexen Vektorraum handelt, ist es etwas anders definiert als in reellen Vektorräumen. Zunächst dürfen wir annehmen, dass die Basisvektoren $|e_j\rangle$ normiert und orthogonal sind (entsprechend unserer Wahl der Polarisationsvektoren \vec{e}_j), d.h. es gilt

$$\langle e_j | e_k \rangle = \delta_{jk} = \begin{cases} 1 & \text{falls } j = k, \\ 0 & \text{falls } j \neq k. \end{cases} \quad (2.1.4)$$

Damit nun beliebige Vektoren (2.1.3) stets eine positive Länge haben, soll das Skalarprodukt durch

$$\langle \delta | \epsilon \rangle = \delta_1^* \epsilon_1 + \delta_2^* \epsilon_2 \quad (2.1.5)$$

¹Die Begründung für diese etwas seltsam anmutende Bezeichnung wird gleich klar werden.

definiert sein, d.h. wir verlangen, dass das Skalarprodukt linear im zweiten Argument ist, d.h.

$$\langle \delta | \lambda_1 \epsilon_1 + \lambda_2 \epsilon_2 \rangle = \lambda_1 \langle \delta | \epsilon_1 \rangle + \lambda_2 \langle \delta | \epsilon_2 \rangle \quad (2.1.6)$$

aber **semilinear** im ersten Argument, d.h. die Zahlen werden konjugiert komplex aus dem Skalarprodukt herausgezogen:

$$\langle \lambda_1 \delta_1 + \lambda_2 \delta_2 | \epsilon \rangle = \lambda_1^* \langle \delta_1 | \epsilon \rangle + \lambda_2^* \langle \delta_2 | \epsilon \rangle. \quad (2.1.7)$$

Entsprechend gelten alle Gesetze des Skalarprodukts, nur ergibt sich beim Vertauschen der Argumente die konjugiert komplexe Zahl:

$$\langle \delta | \epsilon \rangle = \langle \epsilon | \delta \rangle^*. \quad (2.1.8)$$

Folglich ist $\langle \epsilon | \epsilon \rangle \in \mathbb{R}$ und wegen (2.1.5)

$$\langle \epsilon | \epsilon \rangle = \epsilon_1^* \epsilon_1 + \epsilon_2^* \epsilon_2 = |\epsilon_1|^2 + |\epsilon_2|^2 \geq 0, \quad (2.1.9)$$

und es gilt die Bedingung der **positiven Definitheit**, d.h.

$$\langle \epsilon | \epsilon \rangle = 0 \Rightarrow |\epsilon\rangle = |0\rangle, \quad (2.1.10)$$

wobei $|0\rangle$ der Nullvektor ist.

Die Bra-Ket-Schreibweise ist nach dieser Notation des Skalarprodukts als eine Art „Klammer“ (engl. bracket) benannt. Demnach sind die gewöhnlichen Vektoren die Kets $|\epsilon\rangle$. Dann kann man noch die Bras als $\langle \epsilon|$ definieren. Ein Bra ist eine sog. **Linearform**, d.h. eine lineare Abbildung vom Vektorraum in die komplexen Zahlen. Sie ist über das Skalarprodukt durch

$$\langle \epsilon | (|\delta\rangle) = \langle \epsilon | \delta \rangle \quad (2.1.11)$$

definiert. Weiter kann man auch lineare Abbildungen $|\delta\rangle \langle \epsilon|$ definieren, die in naheliegender Weise auf Kets wirken:

$$|\delta\rangle \langle \epsilon | (|\eta\rangle) = |\delta\rangle \langle \epsilon | \eta \rangle. \quad (2.1.12)$$

Weiter ist klar, dass wegen (2.1.4) die Komponenten ϵ_j für einen beliebigen Vektor (2.1.3) durch

$$\epsilon_j = \langle e_j | \epsilon \rangle \quad (2.1.13)$$

gegeben sind. Folglich gilt

$$|\epsilon\rangle = \sum_{j=1}^2 |e_j\rangle \epsilon_j = \sum_{j=1}^2 |e_j\rangle \langle e_j | \epsilon \rangle. \quad (2.1.14)$$

Da dies für jeden Polarisationsvektor $|\epsilon\rangle$ gilt, folgt die sog. **Vollständigkeitsrelation**

$$\sum_{j=1}^2 |e_j\rangle \langle e_j| = \mathbb{1}, \quad (2.1.15)$$

wobei $\mathbb{1}$ den **Einheits-** oder **Identitätsoperator** bezeichnet. Im Folgenden gehen wir stets davon aus, dass alle Polarisationsvektoren auf 1 normiert sind, also

$$\langle \epsilon | \epsilon \rangle = |\epsilon_1|^2 + |\epsilon_2|^2 = 1 \quad (2.1.16)$$

ist.

Jetzt beschäftigen wir uns mit der Frage, wie man den **Polarisationszustand** messen kann. Das führt uns dann auf ziemlich natürliche Weise zum Formalismus der Quantentheorie. Zunächst erinnern wir daran, dass *reelle* Linearkombinationen der Polarisationsvektoren, z.B.

$$|\theta\rangle = \cos\theta |e_1\rangle + \sin\theta |e_2\rangle \quad (2.1.17)$$

beliebige **linear polarisierte Zustände** für in x_3 -Richtung laufende elektromagnetische Wellen sind, wobei die Polarisationsrichtung im Winkel θ zur x_1 -Achse steht.

Das einfachste Hilfsmittel, die Polarisation einer elektromagnetischen Welle zu untersuchen, ist ein **Polarisationsfilter**. Ein Beispiel ist die **Polaroid-Folie**, die aus mit Jod gedoptem Polyvenyl-Plastik besteht, das bei seiner Herstellung gestreckt wurde, sodass die Polymerfasern in einer Richtung ausgerichtet sind. Die Jod-Atome bewirken, dass Licht, das längs der Fasern linear polarisiert ist, komplett absorbiert wird, während dazu senkrecht linear polarisiertes Licht praktisch ungehindert hindurchgelangt. Wir können dies mit einem **linearen Projektionsoperator** in unserem zweidimensionalen Vektorraum beschreiben. Eine ideale Polarisationsfolie, die in θ -Richtung linear polarisiertes Licht komplett durchlässt, während sie dazu senkrecht polarisiertes Licht komplett absorbiert, wird offenbar durch den Operator

$$\mathbf{P}(\theta) = |\theta\rangle \langle \theta| \quad (2.1.18)$$

beschrieben. Diesen können wir auch einfach mit Hilfe der ursprünglichen Basis ausdrücken, indem wir (2.1.18) verwenden

$$\begin{aligned} \mathbf{P}(\theta) &= (\cos\theta |e_1\rangle + \sin\theta |e_2\rangle)(\cos\theta \langle e_1| + \sin\theta \langle e_2|) \\ &= \cos^2\theta |e_1\rangle \langle e_1| + \sin\theta \cos\theta (|e_2\rangle \langle e_1| + |e_1\rangle \langle e_2|) + \sin^2\theta |e_2\rangle \langle e_2|. \end{aligned} \quad (2.1.19)$$

Lassen wir also z.B. in $|\epsilon\rangle = |e_1\rangle$ -Richtung linear polarisiertes Licht durch eine solche im Winkel θ ausgerichtete Polarisationsfolie laufen, wird der Polarisationsvektor des Lichtes nach der Folie durch

$$|\epsilon'\rangle = \mathbf{P}(\theta)|\epsilon\rangle = \mathbf{P}(\theta)|e_1\rangle = \cos^2\theta |e_1\rangle + \sin\theta \cos\theta |e_2\rangle \quad (2.1.20)$$

gegeben sein.

Dieser Vektor ist nun aber nicht mehr auf 1 normiert. Da die Intensität des Lichtes proportional zur Energiedichte $\propto \vec{E}^* \cdot \vec{E}$ ist, bedeutet dies, dass die Intensität um

$$W(\theta) = \langle \epsilon' | \epsilon' \rangle = \cos^4\theta + \sin^2\theta \cos^2\theta = (\cos^2\theta + \sin^2\theta) \cos^2\theta = \cos^2\theta \quad (2.1.21)$$

kleiner geworden ist. Das ist das **Malussche Gesetz**.

Wir können anhand dieses Beispiels noch einige wichtige Rechenregeln für Operatoren erläutern. Sei dazu \mathbf{A} ein beliebiger linearer Operator. Um diesen Operator mit Hilfe unserer Standardbasis $|e_j\rangle$ zu charakterisieren, können wir die Vollständigkeitsrelation (2.1.15) zweimal verwenden. Demnach ist nämlich

$$\mathbf{A} = \sum_{j,k=1}^2 |e_j\rangle \langle e_j | \mathbf{A} | e_k \rangle \langle e_k| = \sum_{j,k=1}^2 A_{jk} |e_j\rangle \langle e_k|. \quad (2.1.22)$$

Dabei heißen die komplexen Zahlen

$$A_{jk} = \langle e_j | \mathbf{A} | e_k \rangle \quad (2.1.23)$$

die **Matrixelemente** des Operators \mathbf{A} bzgl. der vollständigen Orthonormalbasis $|e_j\rangle$. Offenbar können wir den Operator stets eindeutig rekonstruieren, wenn wir alle Matrixelemente des Operators kennen. Wir können die Wirkung des Operators auf einen beliebigen Vektor $|\epsilon\rangle$ ebenfalls aus den Matrixelementen bestimmen, denn offenbar gilt

$$|\epsilon'\rangle = \mathbf{A}|\epsilon\rangle = \sum_{i=1}^2 \epsilon_i \mathbf{A} |e_i\rangle = \sum_{i,j,k=1}^2 \epsilon_i A_{jk} |e_j\rangle \langle e_k | e_i \rangle = \sum_{i,j,k=1}^2 \epsilon_i A_{jk} |e_j\rangle \delta_{ki} = \sum_{i,j=1}^2 A_{ji} \epsilon_i |e_j\rangle. \quad (2.1.24)$$

Damit erhält man die Komponenten von $|\epsilon'\rangle$ durch die übliche Matrix-Vektor-Multiplikationsvorschrift

$$\epsilon'_j = \sum_{i=1}^2 A_{ji} \epsilon_i, \quad (2.1.25)$$

wenn man die Vektorkomponenten als Spaltenvektoren und die Matrixelemente als $\mathbb{C}^{2 \times 2}$ -Matrix anordnet, d.h.

$$\begin{pmatrix} \epsilon'_1 \\ \epsilon'_2 \end{pmatrix} = \hat{A} \begin{pmatrix} \epsilon_1 \\ \epsilon_2 \end{pmatrix} = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix} \begin{pmatrix} \epsilon_1 \\ \epsilon_2 \end{pmatrix} = \begin{pmatrix} A_{11}\epsilon_1 + A_{12}\epsilon_2 \\ A_{21}\epsilon_1 + A_{22}\epsilon_2 \end{pmatrix}. \quad (2.1.26)$$

Schließlich betrachten wir $\langle \epsilon' | = \langle \mathbf{A} \epsilon |$. Verwenden wir (2.1.25), erhalten wir

$$\langle \epsilon' | = \sum_{i,j=1}^2 A_{ji}^* \epsilon_i^* \langle e_j |. \quad (2.1.27)$$

Dabei haben wir berücksichtigt, dass bei der Multiplikation eines Vektors mit einer Zahl

$$\langle \lambda \epsilon | \eta \rangle = \lambda^* \langle \epsilon | \eta \rangle \quad (2.1.28)$$

gilt. Wir bezeichnen die Operation die aus einem Ket $|\epsilon\rangle$ den dazugehörigen Bra $\langle \epsilon|$ macht als **Hermiteische Konjugation** und schreiben

$$(|\epsilon\rangle)^\dagger = \langle \epsilon|. \quad (2.1.29)$$

Entsprechend der Definition der Bra-Vektoren muss demnach gemäß (2.1.28)

$$(\lambda|\epsilon\rangle)^\dagger = \lambda^* \langle \epsilon| \quad (2.1.30)$$

gelten, und das haben wir bei der Rechnung zu (2.1.27) verwendet.

Wir können nun über (2.1.27) den zu **A adjungierten Operator** definieren, indem wir verlangen, dass für beliebige zwei Vektoren $|\epsilon\rangle$ und $|\eta\rangle$

$$\langle \mathbf{A} \epsilon | \eta \rangle = \langle \epsilon | \mathbf{A}^\dagger \eta \rangle \quad (2.1.31)$$

gilt. Mit (2.1.28) folgt

$$\langle \mathbf{A} \epsilon | \eta \rangle = \sum_{j=1}^2 A_{ji}^* \epsilon_i^* \langle e_j | \eta \rangle = \sum_{j=1}^2 A_{ji}^* \epsilon_i^* \eta_j \stackrel{!}{=} \sum_{j=1}^2 (A^\dagger)_{ij} \epsilon_i^* \eta_j. \quad (2.1.32)$$

Da dies für alle Vektoren $|\epsilon\rangle$ und $|\eta\rangle$ gelten soll, folgt

$$(A^\dagger)_{ij} = A_{ji}^*, \quad (2.1.33)$$

d.h. die Matrix von \mathbf{A}^\dagger ist durch die transponierte und konjugiert komplexe Matrix von \mathbf{A} gegeben, d.h.

$$\hat{A}^\dagger = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix}^\dagger = \begin{pmatrix} A_{11}^* & A_{21}^* \\ A_{12}^* & A_{22}^* \end{pmatrix}. \quad (2.1.34)$$

2.2 Polarisationsmessungen an Photonen

Die Notwendigkeit einer von der klassischen Physik abweichenden Quantentheorie wurde von **Max Planck** (1858-1947) in 1900 entdeckt, als er basierend auf Messungen zum Spektrum der Hohlraumstrahlung, das Spektrum von heißen Körpern ausgesandten Lichtes erklären wollte. Dabei hat er sich der statistischen Physik

Boltzmanns bedient und dabei festgestellt, dass er das gemessene Spektrum nur erklären konnte, wenn er annahm, dass elektromagnetische Wellen der Kreisfrequenz ω von der die Strahlung aussendenden Materie in diskreten (bzw. „quantisierten“) Energiepaketen der Größe $\hbar\omega$ emittiert und absorbiert werden kann.

In 1905 hat dann Einstein postuliert, dass elektromagnetische Strahlung neben den durch die klassische Elektrodynamik beschriebenen Welleneigenschaften auch Teilcheneigenschaften zukommt, d.h. eine elektromagnetische Welle besteht aus diskreten teilchenartigen Quanten, die eine Energie $E = \hbar\omega$ und einen Impuls $\vec{p} = \hbar\vec{k}$ besitzen, wobei \vec{k} die Wellenzahl der Strahlung ist. Weiterhin soll auch die Dispersionsrelation der Maxwell-Theorie, also $\omega = c|\vec{k}|$ bzw. $E = c\vec{p}$ gelten. Mit dieser Annahme konnte er auch den **Photoeffekt** erklären, wonach durch auf eine Metalloberfläche fallendes Licht Elektronen herausgelöst werden, jedoch im Widerspruch zur Erwartung von der Maxwell-Theorie die Energie der Elektronen nicht von der Intensität des Lichtes abhängt. Außerdem erfolgte die Emission der Elektronen ohne merkliche Zeitverzögerung, wie es von der klassischen Theorie her zu erwarten gewesen wäre. Die maximale Energie der aus der Metalloberfläche herausgelösten Elektronen ist vielmehr durch

$$E_{\max} = \hbar\omega - W_{\text{aus}} \quad (2.2.1)$$

gegeben, wobei W_{aus} die sog. Austrittsarbeit ist, also die **Bindungsenergie eines Elektrons** im Metall. Es zeigte sich, dass in der Tat \hbar in dieser Formel im Rahmen der Messgenauigkeit mit dem h in Plancks Strahlungsgesetz übereinstimmt. Der moderne Zahlenwert ist

$$\hbar = 1,054\,571\,800(13) \cdot 10^{-34} \text{ J s.}$$

Die etwas ungewöhnliche Bezeichnung mit dem Querstrich (sprich: „h quer“) geht darauf zurück, dass in der Anfangszeit der Quantentheorie statt der Kreisfrequenz der Photonen die gewöhnliche Frequenz verwendet wurde, d.h. die ursprüngliche Planck-Konstante war durch $\hbar\omega = hf$ definiert. Demnach ist

$$h = 2\pi\hbar = 6,626\,070\,040(81) \cdot 10^{-34} \text{ J s.}$$

Freilich wissen wir heute, dass wir uns **Photonen**, also die moderne Version von Einsteins Lichtquanten, nicht naiv als „Teilchen“ vorstellen dürfen. Insbesondere lassen sich Photonen nicht im Sinne von klassischen „Punktteilchen“ lokalisieren. Wir können an dieser Stelle noch nicht en detail auf die schwierige Frage, was ein Photon „eigentlich“ ist, eingehen, denn dazu benötigen wir die relativistische Quantenfeldtheorie, die wir in dieser Vorlesung nicht behandeln können. Ein gut lesbares auf historische und wissenschaftstheoretische Grundlagen des Photonenbegriffs eingehendes Buch ist [Hen17]². Ein sehr gutes populärwissenschaftliches Buch, das auf die aktuelle experimentelle Forschung mit Photonen, insbesondere auch zu den grundlegendsten Fragen der Quantentheorie, eingeht, ist [Zei10].

Für unsere Zwecke ist es allerdings ausreichend, zu bemerken, dass elektromagnetische Strahlung der Frequenz ω mit Wellenvektor \vec{k} stets als **Energie-Impulspakete** in ganzzahligen Vielfachen der Photonenenergie $E = \hbar\omega$ und $\vec{p} = \hbar\vec{k}$ absorbiert und emittiert werden.

Wir können dies zunächst dadurch realisiert denken, dass wir die Lichtquellen, die quasi ebenen Wellen entsprechen (also insbesondere Laserlicht) auf sehr geringe Intensitäten herunterdimmen. Zugleich benötigen wir auch immer empfindlichere Detektoren (z.B. Photomultiplier), um diese immer geringeren Intensitäten noch nachweisen zu können. Bei einer hinreichend geringen Intensität wird nach und nach die diskrete Natur der Absorption und Emission sichtbar. Freilich sind dies noch immer nicht wirklich definitiv einzelne Photonen.

Wirkliche **Einphotonenquellen** konnten erst seit den frühen 1980er Jahren realisiert werden, z.B. durch atomare Kaskaden, wo geeignete Atome angeregt werden und dann zwei Photonen abgeben, sodass man ein Photon verwenden kann, um nachzuweisen, dass auch ein zweites Photon emittiert worden sein muss, und zwar exakt ein Photon, mit dem man dann experimentieren kann.

²Im Netz der Goethe-Universität kostenfrei zum Download verfügbar.

Heute besitzt man wesentlich effizientere Verfahren, um Einzelphotonenzustände zu erzeugen, indem man sich die Kenntnisse der nichtlinearen Optik zunutze macht. Diese technischen Errungenschaften ermöglichen heute sehr interessante Untersuchungen zu den Grundlagen der Quantentheorie, auf die wir nun zu sprechen kommen wollen. Auf die technischen Details wollen wir an dieser Stelle nicht genauer eingehen. Eine sehr schöne populärwissenschaftliche Darstellung dazu ist wieder [Zei10].

Jetzt wollen wir den im vorigen Abschnitt für klassische elektromagnetische Wellen entwickelten Formalismus im Hinblick auf die Lichtquanten umformulieren. Wir denken uns dazu das Polarisationsexperiment mit einem sehr stark gedimmten Laser durchgeführt. Dann bestrahlen wir den Polarisationsfilter immer nur so lange, dass im Mittel im entsprechenden Laserpuls nur die Energie eines Photons abgestrahlt wird. Es zeigt sich nun, dass mit einem Photonendetektor nach dem Polarisationsfilter immer nur ein Photon mit Energie $\hbar\omega$ oder gar nichts registriert wird. Wenn man aber viele solcher Photonen durch die Anordnung laufen lässt, registriert man *zufällig* jeweils immer nur ein Photon oder eben nichts. Welches individuelle Photon durch die Polarisationsfolie hindurchläuft oder nicht, kann man nicht vorhersagen. Die **relative Häufigkeit** registrierter Photonen im Vergleich zu der Gesamtzahl von durch die Apparatur geschickten Photonen entspricht jedoch der **relativen Intensität** der Lichtwelle, die wir oben mit Hilfe unseres aus der klassischen Elektrodynamik motivierten Matrix-Vektor-Formalismus berechnet haben.

Wir haben nun für die Polarisationsmessungen einzelner Photonen folgende Interpretation: Die den **Polarisationszustand** repräsentierenden auf 1 normierten Vektoren $|\epsilon\rangle$ gestatten es, die **Wahrscheinlichkeit** zu berechnen, dass ein einzelnes Photon durch den Polarisationsfilter läuft. Der in θ -Richtung ausgerichtete Polarisationsfilter wird dabei durch den Operator (2.1.19) beschrieben. Dieser Projektionsoperator besitzt dabei offenbar bzgl. unserer Basis $|e_j\rangle$ die Matrixdarstellung

$$\hat{P}(\theta) = \begin{pmatrix} \cos^2 \theta & \sin \theta \cos \theta \\ \sin \theta \cos \theta & \sin^2 \theta \end{pmatrix}. \quad (2.2.2)$$

Die Wahrscheinlichkeit, dass ein in x_1 -Richtung linear polarisiertes Photon durch diesen Filter läuft, wenn also der Polarisationszustand durch $|e_1\rangle$ gegeben ist, ist demnach gemäß (2.1.22) durch

$$W(\theta) = \langle e_1 | \hat{P}(\theta) | e_1 \rangle = P_{11}(\theta) = \cos^2 \theta \quad (2.2.3)$$

gegeben.

Wir können dieses Resultat auch noch ein wenig anders interpretieren. Der um θ gegen die x_1 -Richtung gedrehte Polarisationsfilter lässt ja mit 100% Wahrscheinlichkeit (also mit Sicherheit) ein in diese θ -Richtung linear polarisiertes Photon durch und absorbiert mit 100% Wahrscheinlichkeit ein dazu senkrecht linear polarisiertes Photon. Die beiden Polarisationsvektoren $|\epsilon_1\rangle = |\epsilon(\theta)\rangle$ und $|\epsilon_2\rangle = |\epsilon(\theta + \pi/2)\rangle$ sind demnach zum Einen wieder eine Orthonormalbasis und zum Anderen durch

$$\hat{P}(\theta)|\epsilon_1\rangle = |\epsilon\rangle, \quad \hat{P}(\theta)|\epsilon_2\rangle = |0\rangle = 0|\epsilon_2\rangle \quad (2.2.4)$$

charakterisiert. Man nennt nun einen vom Nullvektor verschiedenen Vektor, der durch einen linearen Operator \mathbf{A} bis auf einen Faktor auf sich selbst abgebildet wird, für den es also eine (i.a. komplexe) Zahl λ gibt, sodass

$$\mathbf{A}|\epsilon\rangle = \lambda|\epsilon\rangle \quad (2.2.5)$$

gilt, einen **Eigenvektor** des Operators \mathbf{A} zum **Eigenwert** λ .

Da in unserem Fall die Eigenvektoren $|\epsilon_j\rangle$ zueinander orthogonal sind und wir sie auf 1 normiert denken können, bildet auch $|\epsilon_j\rangle$ wieder ein Orthonormalsystem, und für den oben betrachteten Operator gilt dann offenbar

$$\hat{P}(\theta) = |\epsilon_1\rangle\langle\epsilon_1|. \quad (2.2.6)$$

Wir können die Wahrscheinlichkeit, dass ein in x_1 -Richtung polarisiertes Photon durch den Filter gelassen wird, also auch durch

$$W(\theta) = \langle \epsilon | \mathbf{P}(\theta) | \epsilon \rangle = \langle \epsilon | \epsilon_1 \rangle \langle \epsilon_1 | \epsilon \rangle = |\langle \epsilon_1 | \epsilon \rangle|^2 \quad (2.2.7)$$

beschreiben. Anders ausgedrückt besagt dies, dass die Wahrscheinlichkeit dafür, dass ein im Polarisationszustand $|\epsilon\rangle$ präpariertes Photon durch den in θ -Richtung orientierten Polarisator gelangt, durch den entsprechenden Eigenvektor des Operators $\mathbf{P}(\theta)$ durch das **Betragsquadrat** in (2.2.7) gegeben ist.

In diesem Sinne repräsentiert $\mathbf{P}(\theta)$ eine **Observable**. Dabei ist das Photon in θ -Richtung linear polarisiert genau dann, wenn es mit der Wahrscheinlichkeit 1=100% durch den entsprechenden Filter läuft und dazu senkrecht linear polarisiert, wenn es mit Wahrscheinlichkeit 0 durch den Filter läuft. Die möglichen Messwerte sind also 0 oder 1, d.h. die Eigenwerte des Operators, und die entsprechenden Zustände werden durch die dazu gehörigen Eigenvektoren repräsentiert.

Wir stellen nun fest, dass offenbar $\mathbf{P}^\dagger(\theta) = \mathbf{P}(\theta)$ gilt. Man sagt, dass der Operator $\mathbf{P}(\theta)$ **selbstadjungiert** ist, und dies ist eine allgemeine Eigenschaft von Operatoren, die Observablen repräsentieren, und zwar aus dem Grund, dass die normierten Eigenvektoren selbstadjungierter Operatoren stets als ein **vollständiges Orthonormalsystem** gewählt werden können. Wir kommen darauf im nächsten Kapitel noch genauer zurück.

Hier wollen wir nun noch einmal unsere recht heuristischen Betrachtungen zu grundlegenden (wenn auch noch unvollständigen) **Postulaten** zusammenfassen:

1. Die möglichen **Zustände** eines Quantensystems werden durch **normierte Vektoren** $|\psi\rangle$ in einem **Hilbert-Raum** repräsentiert. Dabei ist ein Hilbert-Raum ein komplexer Vektorraum mit einem Skalarprodukt, wie wir es oben im zweidimensionalen Raum für Polarisationszustände definiert haben. Wie wir im nächsten Kapitel sehen werden, wo wir uns mit der quantentheoretischen Beschreibung nicht-relativistischer Teilchen beschäftigen werden, ist i.a. der Hilbert-Raum quantenmechanischer Systeme aber „unendlichdimensional“, d.h. es gibt unendlich viele zueinander orthogonaler normierter Vektoren, d.h. die vollständigen Orthonormalbasen bestehen i.a. aus unendlich vielen Vektoren (in der Quantentheorie aber stets nur aus *abzählbar* unendlich vielen). Dabei heißt eine Orthonormalbasis vollständig, wenn man jeden Vektor des Hilbert-Raums durch Linearkombination bzw. entsprechende Verallgemeinerungen zu Linearkombinationen mit unendlich vielen Summanden (und einem geeigneten Konvergenzbegriff) schreiben kann. All dies wird im nächsten Kapitel ausführlich erläutert.
2. Die **Observablen** eines Quantensystems werden durch **selbstadjungierte Operatoren** repräsentiert. Die möglichen Messwerte der Observable sind die **Eigenwerte** dieses Operators, und der Observablen kommt genau dann ein bestimmter möglicher Eigenwert λ zu, wenn es in einem Zustand präpariert ist, der durch einen Eigenvektor $|\mu_\lambda\rangle$ des Operators zum Eigenwert λ repräsentiert wird.
3. Ist ein Quantensystem in einem durch den normierten Vektor $|\psi\rangle$ repräsentierten Zustand präpariert, ist die **Wahrscheinlichkeit**, bei der Messung einer Observablen einen möglichen Messwert λ zu erhalten, durch

$$W(\lambda) = |\langle \mu_\lambda | \psi \rangle|^2 \quad (2.2.8)$$

gegeben. Dabei ist $|\mu_\lambda\rangle$ ein normierter Eigenvektor des die Observable repräsentierenden selbstadjungierten Operators. Nach ihrem Entdecker nennt man (2.2.8) die **Bornsche Regel**.

Diese sehr abstrakten Postulate werden wir im Folgenden noch genauer analysieren. Es ist erfahrungsgemäß recht schwierig, die physikalische Bedeutung des Formalismusses zu verstehen, aber man gewöhnt sich mit der Zeit an diese „quantenphysikalische Denkweise“, die radikal mit den gewohnten Begriffen der klassischen Physik bricht. Insbesondere zeigt unser Beispiel der Polarisation von Photonen, dass Observablen vor ihrer Messung nicht notwendig einen bestimmten Wert besitzen. Vielmehr können wir i.a. für eine bestimmte Observable nur **Wahrscheinlichkeiten** angeben, mit der sich bei einer Messung dieser Observablen ein bestimmter möglicher Messwert ergibt, selbst wenn vor der Messung der Zustandsvektor $|\psi\rangle$ exakt bekannt

ist. Insbesondere nimmt man an, dass die Kenntnis des Zustandsvektors eine **vollständige Kenntnis** des Zustandes des betrachteten Quantensystems ist, d.h. die Unbestimmtheit des Wertes einer Observablen rührt nicht von einer unvollständigen Kenntnis des Systemzustandes her, sondern die Observable besitzt dann *tatsächlich* keinen wohldefinierten Wert, und eine Messung ergibt zufällig einen möglichen Messwert, wobei die Wahrscheinlichkeit durch die Bornsche Regel (2.2.8) gegeben ist.

Um diese Wahrscheinlichkeiten **experimentell** zu bestimmen, müssen wir viele individuelle Quantensysteme im durch den Zustandsket $|\psi\rangle$ beschriebenen Zustand präparieren. Man nennt dies auch ein **Ensemble** von entsprechend präparierten Quantensystemen. Dann nähern sich die relativen Häufigkeiten, mit denen jeder mögliche Messwert bei einer Messung der Observablen resultiert, den berechneten Wahrscheinlichkeiten gemäß der Bornschen Regel an³.

Wir bemerken zum Abschluss dieses einleitenden Kapitels nur noch folgendes:

1. Die selbstadjungierten Operatoren besitzen **nur reelle Eigenwerte**.

Beweis: Es sei \mathbf{A} ein selbstadjungierter Operator, d.h. es gelte $\mathbf{A}^\dagger = \mathbf{A}$, und λ sei ein Eigenwert mit einem auf 1 normierten Eigenvektor $|u_\lambda\rangle$. Dann gilt $\mathbf{A}|u_\lambda\rangle = \lambda|u_\lambda\rangle$ und da $\langle u_\lambda | u_\lambda \rangle = 1$ ist, können wir schreiben

$$\lambda = \langle u_\lambda | \lambda u_\lambda \rangle = \langle u_\lambda | \mathbf{A} u_\lambda \rangle = \langle \mathbf{A}^\dagger u_\lambda | u_\lambda \rangle = \langle \mathbf{A} u_\lambda | u_\lambda \rangle = \langle \lambda u_\lambda | u_\lambda \rangle = \lambda^* \langle u_\lambda | u_\lambda \rangle = \lambda^* \Rightarrow \lambda \in \mathbb{R}. \quad (2.2.9)$$

2. Die Eigenzustände eines selbstadjungierten Operators zu zwei *voneinander verschiedenen* Eigenwerten sind stets orthogonal zueinander.

Beweis: Seien λ und μ zwei voneinander verschiedene Eigenwerte des selbstadjungierten Operators \mathbf{A} . Dann gilt

$$\mu \langle u_\lambda | u_\mu \rangle = \langle u_\lambda | \mathbf{A} u_\mu \rangle = \langle \mathbf{A}^\dagger u_\lambda | u_\mu \rangle = \langle \mathbf{A} u_\lambda | u_\mu \rangle = \lambda \langle u_\lambda | u_\mu \rangle. \quad (2.2.10)$$

Dies ist offenbar gleichwertig mit

$$(\mu - \lambda) \langle u_\lambda | u_\mu \rangle = 0. \quad (2.2.11)$$

Da voraussetzungsgemäß $\mu \neq \lambda$ ist, muss demnach

$$\langle u_\lambda | u_\mu \rangle = 0 \quad (2.2.12)$$

sein, und das war zu zeigen.

Nun müssen wir noch den Fall besprechen, dass es zu einem Eigenwert mehrere linear unabhängige Eigenvektoren gibt. Man sagt dann, dass dieser Eigenwert des selbstadjungierte Operators **entartet** ist. Sei also λ ein Eigenwert des selbstadjungierten Operators \mathbf{A} und $|u_\lambda^{(j)}\rangle$ mit $j \in \{1, \dots, d_\lambda\}$ linear unabhängige Eigenvektoren zu diesem Eigenwert. Wir nehmen weiter an, dass es keinen weiteren zu diesem System linear unabhängigen Eigenvektor zu dem Eigenwert λ gibt. Offenbar kann man dann jeden Eigenvektor zu diesem Eigenwert als Linearkombination dieser Eigenvektoren schreiben

$$|v_\lambda\rangle = \sum_{j=1}^{d_\lambda} v_\lambda^{(j)} |u_\lambda^{(j)}\rangle. \quad (2.2.13)$$

Jede solche Linearkombination ist nämlich ebenfalls ein Eigenvektor von \mathbf{A} zum Eigenwert λ :

$$\mathbf{A}|v_\lambda\rangle = \mathbf{A} \sum_{j=1}^{d_\lambda} v_\lambda^{(j)} |u_\lambda^{(j)}\rangle = \sum_{j=1}^{d_\lambda} v_\lambda^{(j)} \mathbf{A} |u_\lambda^{(j)}\rangle = \sum_{j=1}^{d_\lambda} v_\lambda^{(j)} \lambda |u_\lambda^{(j)}\rangle = \lambda \sum_{j=1}^{d_\lambda} v_\lambda^{(j)} |u_\lambda^{(j)}\rangle = \lambda |v_\lambda\rangle. \quad (2.2.14)$$

³vorausgesetzt die Quantentheorie ist korrekt. Bis dato gibt es keine Widersprüche zwischen Experiment und Quantentheorie bzw. die Quantentheorie ist sogar die am besten im Experiment bestätigte Theorie der Physik.

2.2. Polarisationsmessungen an Photonen

Wir können nun aus den Vektoren $|u_\lambda^{(j)}\rangle$, die voraussetzungsgemäß alle Eigenvektoren von \mathbf{A} zum Eigenwert λ „aufspannen“ (d.h. alle Eigenvektoren lassen sich als Linearkombination gemäß (2.2.13) schreiben), einen neuen Satz Vektoren $|\tilde{u}_\lambda^{(j)}\rangle$ konstruieren, die zueinander orthogonal und normiert sind.

Dazu geht man rekursiv vor. Wir setzen willkürlich

$$|\tilde{u}_\lambda^{(1)}\rangle = \frac{1}{\|u_\lambda^{(1)}\|} |u_\lambda^{(1)}\rangle, \quad (2.2.15)$$

wobei wir die Norm $\|\psi\|$ eines Vektors durch die Definition

$$\|\psi\| = \sqrt{\langle\psi|\psi\rangle} \quad (2.2.16)$$

eingeführt haben. Bzgl. (2.2.15) ist es wichtig zu bemerken, dass wegen der linearen Unabhängigkeit der $|u_\lambda^{(j)}\rangle$ keiner dieser Vektoren der Nullvektor sein kann und daher wegen der positiven Definitheit des Skalarprodukts $\|u_\lambda^{(1)}\| \neq 0$ ist.

Jetzt können wir einen weiteren zu (2.2.15) orthogonalen Eigenvektor von \mathbf{A} zum Eigenwert λ konstruieren. Dazu definieren wir zuerst den i.a. unnormierten zu $|\tilde{u}_\lambda^{(1)}\rangle$ orthogonalen Vektor

$$|\tilde{v}_\lambda^{(2)}\rangle = |u_\lambda^{(2)}\rangle - |\tilde{u}_\lambda^{(1)}\rangle \langle\tilde{u}_\lambda^{(1)}|u_\lambda^{(2)}\rangle. \quad (2.2.17)$$

Man rechnet leicht nach, dass in der Tat $\langle\tilde{u}_\lambda^{(1)}|\tilde{v}_\lambda^{(2)}\rangle = 0$ ist, wie gewünscht. Wegen der linearen Unabhängigkeit der $|u_\lambda^{(j)}\rangle$ ist (2.2.17) nicht der Nullvektor, und wir können daher diesen Vektor auf 1 normieren, d.h. wir setzen

$$|\tilde{u}_\lambda^{(2)}\rangle = \frac{1}{\|\tilde{v}_\lambda^{(2)}\|} |\tilde{v}_\lambda^{(2)}\rangle. \quad (2.2.18)$$

So können wir nun rekursiv fortfahren, d.h. haben wir bereits $k < d_\lambda$ zueinander orthogonale normierte Eigenvektoren unter Verwendung der Vektoren $|u_\lambda^{(j)}\rangle$ mit $j \in \{1, 2, \dots, k\}$ konstruiert, können wir auch einen $(k+1)$ -ten solchen Vektor konstruieren. Dazu definieren wir

$$|\tilde{v}_\lambda^{(k+1)}\rangle = |u_\lambda^{(k+1)}\rangle - \sum_{j=1}^k |\tilde{u}_\lambda^{(j)}\rangle \langle\tilde{u}_\lambda^{(j)}|u_\lambda^{(k+1)}\rangle. \quad (2.2.19)$$

Man rechnet leicht nach, dass dann $|\tilde{v}_\lambda^{(k+1)}\rangle$ orthogonal zu allen Vektoren $|\tilde{u}_\lambda^{(j)}\rangle$ mit $j \in \{1, 2, \dots, k\}$ ist, und da die $u_\lambda^{(j)}$ linear unabhängig sind, ist der Vektor (2.2.20) wieder nicht der Nullvektor. Damit können wir diesen Vektor wieder normieren, d.h. wir setzen

$$|\tilde{u}_\lambda^{(k+1)}\rangle = \frac{1}{\|\tilde{v}_\lambda^{(k+1)}\|} |\tilde{v}_\lambda^{(k+1)}\rangle. \quad (2.2.20)$$

Dieses sog. **Schmidtsche Orthonormierungsverfahren** können wir wiederholen, bis $k+1 = d_\lambda$ ist. Danach bricht das Verfahren ab, denn für $k = d_\lambda$ wird der gemäß (2.2.19) konstruierte Vektor $|\tilde{v}_\lambda^{(d_\lambda+1)}\rangle$ der Nullvektor, weil es ansonsten einen $d_\lambda + 1$ -ten zu den $|u_\lambda^{(j)}\rangle$ linear unabhängigen Eigenvektor von \mathbf{A} zum Eigenwert λ gäbe. Das kann aber nicht sein, da wir vorausgesetzt haben, dass es einen solchen weiteren linear unabhängigen Vektor nicht geben soll.

Wir können also zu jedem entarteten Eigenwert λ des selbstadjungierten Operators \mathbf{A} eine Orthonormalbasis für alle Eigenvektoren zu diesem Eigenwert angeben. Da, wie oben gezeigt, die Eigenvektoren zu voneinander verschiedenen Eigenwerten aufgrund der Selbstadjungiertheit automatisch zueinander orthogonal sind und auf 1 normiert werden können, können wir demzufolge einen Satz von orthonormierten Eigenvektoren des selbstadjungierten Operators \mathbf{A} finden. Für endlichdimensionale Hilbert-Räume ist es leicht zu zeigen, dass dies auch stets **vollständige Orthonormalsysteme** sind (für einen Beweis s. [AHK⁺18]), d.h. jeder Vektor lässt sich als Linearkombination von orthonormierten Eigenvektoren von \mathbf{A} darstellen, d.h. die Orthonormalsysteme aus Eigenvektoren eines selbstadjungierten Operators sind auch **vollständig**, d.h. es gilt die **Vollständigkeitsrelation**

$$\sum_{\lambda} \sum_{j=1}^{d_{\lambda}} |\tilde{u}_{\lambda}^{(j)}\rangle \langle \tilde{u}_{\lambda}^{(j)}| = \mathbb{1}. \quad (2.2.21)$$

Für die Quantentheorie sind diese Eigenschaften sehr wichtig, weil sie die oben aufgestellten Postulate konsistent mit der physikalischen Bedeutung einer Observablen machen.

Wir müssen nun die **Bornsche Regel** für den Fall, dass ein Eigenwert entartet ist, präzisieren. Sei also λ der entartete Eigenwert des eine Observable A repräsentierenden selbstadjungierten Operators \mathbf{A} und $|u_{\lambda}^{(j)}\rangle$ mit $j \in \{1, \dots, d_{\lambda}\}$ ein vollständiges Orthonormalsystem von Eigenvektoren zu diesem Eigenwert. Dann ist die Wahrscheinlichkeit, dass bei einer Messung von A am im Zustand $|\psi\rangle$ präparierten System der Wert λ resultiert, durch

$$W(\lambda) = \sum_{j=1}^{d_{\lambda}} |\langle u_{\lambda}^{(j)} | \psi \rangle|^2 \quad (2.2.22)$$

gegeben.

Zum einen ist mit dieser Interpretation der Zustandsbegriff für den Fall, dass eine Observable einen bestimmten Wert λ besitzt, wobei λ ein (notwendig reeller!) Eigenwert des die Observable repräsentierenden selbstadjungierten Operators \mathbf{A} ist, in sich konsistent, denn in diesem Fall wird der Zustand durch irgendeinen Eigenvektor $|\psi\rangle = |u_{\lambda}\rangle$ von \mathbf{A} zum Eigenwert λ repräsentiert. Die Wahrscheinlichkeit, einen anderen möglichen Wert $\mu \neq \lambda$ zu messen, ist demnach 0, denn diese Wahrscheinlichkeit ist gemäß der Bornschen Regel durch

$$W(\mu) = \sum_{j=1}^{d_{\mu}} |\langle u_{\mu}^{(j)} | \psi \rangle|^2 = 0 \quad (2.2.23)$$

gegeben. Dabei haben wir verwendet, dass alle Eigenvektoren zum Eigenwert λ orthogonal zu den Eigenvektoren zum Eigenwert μ mit $\mu \neq \lambda$ sind.

Außerdem gilt (*Nachrechnen!*)

$$W(\lambda) = \sum_{\lambda'} \sum_{j=1}^{d_{\lambda'}} |\langle u_{\lambda'}^{(j)} | \psi \rangle|^2 = \sum_{j=1}^{d_{\lambda}} \langle \psi | u_{\mu}^{(j)} \rangle \langle u_{\mu}^{(j)} | \psi \rangle = 1, \quad (2.2.24)$$

d.h. es wird in der Tat bei einer Messung der Observablen A mit Sicherheit der Wert λ gemessen. Dabei haben wir die Vollständigkeit des Orthonormalsystems von Eigenvektoren von \mathbf{A} verwendet, also

$$\sum_{\lambda'} \sum_{j=1}^{d_{\lambda'}} |u_{\lambda'}^{(j)}\rangle \langle u_{\lambda'}^{(j)}| = \mathbb{1}. \quad (2.2.25)$$

Falls nun $|\psi\rangle$ kein Eigenvektor von \mathbf{A} ist, gilt nun immer noch

$$\sum_{\lambda'} W(\lambda') = \sum_{\lambda'} \sum_{j=1}^{d_{\lambda'}} \langle \psi | u_{\lambda'}^{(j)} \rangle \langle u_{\lambda'}^{(j)} | \psi \rangle \stackrel{(2.2.21)}{=} \langle \psi | \psi \rangle = 1, \quad (2.2.26)$$

d.h. der Formalismus besagt, dass man bei der Messung der durch \mathbf{A} repräsentierten Observablen A notwendig irgendeinen Eigenwert dieses Operators misst. Damit sind die obigen Postulate konsistent sowohl für den Fall, dass das System in einem Zustand präpariert ist, für den die observable A einen der möglichen Messwerte tatsächlich besitzt (also der Zustand ein Eigenzustand zu \mathbf{A} ist) als auch für den Fall, dass der Zustand vor der Messung kein solcher Eigenzustand ist und demzufolge A keinen definierten Wert besitzt.

Nun können wir auch **Mittelwerte** bzw. **Erwartungswerte** von Observablen berechnen. Definitionsgemäß ist

$$\langle A \rangle = \sum_{\lambda} W(\lambda)\lambda = \sum_{\lambda} \sum_{j=1}^{d_{\lambda}} \langle \psi | u_{\lambda}^{(j)} \rangle \langle \mathbf{A} u_{\lambda}^{(j)} | \psi \rangle = \sum_{\lambda} \sum_{j=1}^{d_{\lambda}} \langle \psi | u_{\lambda}^{(j)} \rangle \langle u_{\lambda}^{(j)} | \mathbf{A} \psi \rangle = \langle \psi | \mathbf{A} \psi \rangle. \quad (2.2.27)$$

Dabei haben wir im letzten Schritt die Vollständigkeitsrelation (2.2.21) verwendet. Wir können also mit Hilfe des die Observable repräsentierenden Operators auch ohne Kenntnis der Eigenwerte und Eigenzustände den Erwartungswert berechnen, wenn wir wissen, wie der Operator auf den auf 1 normierten Zustandsvektor $|\psi\rangle$ wirkt.

Wir haben also einen in sich konsistenten Formalismus zur Beschreibung von Messungen an Quantensystemen. Dabei ist allerdings zu beachten, dass dieses abstrakte Konstrukt, vor allem aus erkenntnistheoretischer Sicht, nicht unproblematisch ist, denn der Zustandsbegriff ist hier offensichtlich „**probabilistisch**“, d.h. i.a. besitzt eine Observable eben keinen bestimmten Wert wie in der klassischen Physik, und die **Präparation in einem Zustand** im obigen Sinne impliziert nur die Kenntnis von Wahrscheinlichkeiten, bei einer Messung dieser Observable, irgendeinen ihrer möglichen Messwerte zu erhalten. Andererseits setzen wir stillschweigend voraus, dass bei einer Messung in der Tat stets ein und nur ein eindeutiger solcher Messwert resultiert. Dies ist das berühmte **Messproblem**, auf das wir in einem späteren Kapitel noch kurz zu sprechen kommen werden. Wie bereits in der Einleitung erwähnt, problematisieren wir diese Frage zunächst einmal nicht weiter und ziehen uns auf den Standpunkt zurück, dass die oben gegebene **minimale statistische Interpretation** des quantenmechanischen Zustandsbegriffs in sehr präziser Übereinstimmung mit allen bislang ausgeführten Experimenten ist.

Mit diesen heuristischen Vorarbeiten sind wir jedenfalls nun gewappnet, uns mit dem eigentlichen Thema dieser Vorlesung, d.h. der Quantentheorie nichtrelativistischer Teilchen zu beschäftigen. Dabei werden wir auch die Postulate der Quantentheorie noch dahingehend ergänzen, dass wir auch die **Dynamik des Quantensystems** beschreiben, d.h. die zeitliche Entwicklung der Wahrscheinlichkeiten für Resultate von Messungen.

Kapitel 3

Wellenmechanik

In diesem Kapitel wenden wir uns nun der eigentlichen **nichtrelativistischen Quantenmechanik** zu. Zunächst betrachten wir ein einzelnes **Teilchen**. Der heuristische Ausgangspunkt ist hier der schon im vorigen Kapitel besprochene naive Photonenbegriff. Im Gegensatz zu den Photonen, deren korrekte Beschreibung die relativistische Quantenfeldtheorie erfordert, können wir aber für **massive Teilchen** in der nichtrelativistischen Näherung auf recht einfache Weise zu einer konsistenten quantentheoretischen Beschreibung gelangen. Der Ausgangspunkt ist die Idee **de Broglies**, dass wie beim Licht auch bei massiven Teilchen eine Art „Welle-Teilchen-Dualismus“ vorliegt, d.h. dass es auch für Teilchen auf mikroskopischer Ebene eine Art Wellenverhalten geben sollte. Freilich wissen wir aus unserer Diskussion über Photonen, dass diese vorläufige Sichtweise sich nicht halten lässt und eine **Wahrscheinlichkeitsbeschreibung** erforderlich sein wird.

Außerdem können wir uns von den **kinematischen Grundpostulaten**, die wir anhand des einfachen Beispiels von Polarisationsmessungen an einzelnen Photonen entwickelt haben, leiten lassen. Dabei muss es für Teilchen freilich um eine Beschreibung ihres Verhaltens in **Raum und Zeit** gehen und der damit zusammenhängenden grundlegenden **kontinuierlichen Observablen Ort und Impuls**. Wir werden also, ausgehend von de Broglies „Materiewellenkonzept“ einen Hilbert-Raum konstruieren, der zunächst **Orts- und Impulsmessungen** im Rahmen der kinematischen quantentheoretischen Postulate formalisiert.

Allerdings werden wir dabei auf recht natürlichem Wege auch zu einer **dynamischen Theorie** geführt, denn die aus diesen Überlegungen resultierende **Schrödinger-Gleichung** beschreibt nichts anderes als die zeitliche Entwicklung der Wellenfunktion, die wir als Elemente des **Hilbert-Raums der quadratintegrablen Funktionen** identifizieren können und somit als eine mögliche konkrete Realisierung des Hilbert-Raums der Zustände für Orts- und Impulsmessungen nichtrelativistischer Teilchen im Sinne der kinematischen Quantenpostulate interpretieren können.

3.1 Die Schrödinger-Gleichung für freie Teilchen

Wir beginnen mit dem einfachsten Fall des freien Teilchens, um uns der Idee der de Broglie-Materiewellen zu nähern. Ausgangspunkt ist die Einstein-Beziehung für Lichtquanten $E = \hbar\omega$ und $\vec{p} = \hbar\vec{k}$, die de Broglie auf materielle Teilchen angewandt hatte.

Ein freies Teilchen zeichnet sich dadurch aus, dass keinerlei äußere Einflüsse auf es stattfinden. Das einzige, was wir über das Teilchen wissen müssen, um eine Wellengleichung aufzustellen, die es beschreiben soll, ist die sog. *Dispersionsrelation*, d.h. den Zusammenhang zwischen Kreisfrequenz ω und Wellenvektor \vec{k} . Diese Beziehung können wir aber vermöge der klassischen Mechanik durch den Zusammenhang zwischen Energie und Impuls unter Zuhilfenahme der Einstein–de Broglie–Beziehung gewinnen. Für ein freies nichtrelativisti-

3. Wellenmechanik

sches Punktteilchen ist die Energie identisch mit der kinetischen Energie, und die ist durch die Beziehung

$$E = \frac{\vec{p}^2}{2m} \quad (3.1.1)$$

gegeben, wobei \vec{p} der Impuls des Teilchens und m seine Masse ist. Hier ist die Masse der einzige das Teilchen näher charakterisierende Parameter, ansonsten ist es als völlig unstrukturiert punktförmig abstrahiert. Setzen wir die Einstein–de Broglie–Beziehung ein, finden wir die gesuchte Dispersionsrelation:

$$\hbar\omega = \frac{(\hbar\vec{k})^2}{2m} \Rightarrow \omega(\vec{k}) = \frac{\hbar\vec{k}^2}{2m}. \quad (3.1.2)$$

Die einfachste Form einer Welle ist nun die aus der allgemeinen Wellenlehre bekannte ebene Welle, die durch eine Sinuswelle beschrieben wird. Aus rechentechnischen Gründen verwenden wir hier die Form der Exponentialfunktion und komplexe Zahlen. Wir werden sogleich sehen, dass die Schrödingersche Wellenfunktion ohnehin am bequemsten mit komplexwertigen Funktionen beschrieben wird. Wir setzen also an

$$\psi(t, \vec{x}) = A \exp(-i\omega t + i\vec{k}\vec{x}). \quad (3.1.3)$$

Die Dispersionsbeziehung (3.1.2) muss nun durch eine Wellengleichung gegeben sein, also eine partielle Differentialgleichung, der ψ identisch genügen muss. Dazu leiten wir den Ansatz (3.1.3) einmal nach t und zweimal nach \vec{x} ab:

$$\partial_t \psi(t, \vec{x}) = -i\omega \psi(t, \vec{x}), \quad \Delta_{\vec{x}} \psi(t, \vec{x}) = -\vec{k}^2 \psi(t, \vec{x}) \quad \text{mit} \quad \Delta = \nabla \cdot \nabla = \partial_x^2 + \partial_y^2 + \partial_z^2. \quad (3.1.4)$$

Vergleichen wir dies mit (3.1.2) sehen wir, dass die Wellenfunktion der Gleichung

$$i\hbar \partial_t \psi = -\frac{\hbar^2 \Delta}{2m} \psi \quad (3.1.5)$$

zu genügen hat, damit die Dispersionsbeziehung identisch erfüllt wird. Dies ist schon die von **Erwin Schrödinger** (1887-1961) in 1926 gefundene **Schrödinger-Gleichung des freien Teilchens**.

Wir können nun die Wellenfunktion im Sinne der Bornschen Wahrscheinlichkeitsinterpretation aus dem vorigen Kapitel interpretieren, indem wir annehmen, dass

$$W(t, \vec{x}) = |\psi(t, \vec{x})|^2 \quad (3.1.6)$$

die Wahrscheinlichkeitsdichte der Teilchen sein, also die Wahrscheinlichkeit pro Volumenelement, zur Zeit t ein Teilchen der Masse m am Ort \vec{x} zu finden. Dies ist insofern gegenüber den Betrachtungen zu Polarisationsmessungen an Photonen erzielte Interpretation eine Verallgemeinerung als wir es hier anstatt mit diskreten Observablen (also z.B. entweder in x_1 - oder x_2 -Richtung linear polarisierten Photonen) beim Ort eines Teilchens mit **kontinuierlichen Observablen** zu tun haben, wobei wir vorläufig davon ausgehen, dass dies auch in der Quantentheorie noch so ist. Wie wir sehen werden, sind in der Tat auch in der Quantentheorie die Komponenten des Ortsvektors eines Teilchens kontinuierliche Observablen. Das Ziel des Rests dieses Abschnitts ist es nun, dies in den allgemeinen Formalismus des vorigen Kapitels zu übersetzen, d.h. wir wollen einen Hilbert-Raum der Zustände sowie selbstadjungierte Operatoren für die Komponenten des Ortsvektors \vec{x} und auch des Impulses \vec{p} definieren, die die entsprechenden „kontinuierlichen Eigenwerte“ im \mathbb{R}^3 besitzen, identifizieren.

Setzen wir nun aber in (3.1.6) unsere ebene Welle (3.1.3) ein, finden wir für W eine Konstante. Damit aber unsere Welle eine Interpretation in dem verallgemeinerten Bornschen Wahrscheinlichkeitssinne (3.1.6) überhaupt besitzen kann, muss die Wahrscheinlichkeit, das Teilchen überhaupt irgendwo im Raum zu finden,

3.1. Die Schrödinger-Gleichung für freie Teilchen

auf 1 normierbar sein. Das ist aber offenbar nicht der Fall, denn eine Konstante lässt sich gewiss nicht mit endlichem Resultat über den ganzen Raum integrieren.

Die ebene Welle entspricht auch ganz und gar nicht unserer Vorstellung von einem Teilchen, ist es doch zu dem einen Zeitpunkt t überall gleich wahrscheinlich, es zu finden, denn es ist für die ebene Welle ja $W = |A|^2 = \text{const.}$ Unserer Teilchenvorstellung käme also ein **Wellenpaket** viel näher, d.h. eine Lösung der Schrödinger-Gleichung, die auf einem relativ kleinen Gebiet im Raum eine große Amplitude besitzt, entsprechend einer hohen Aufenthaltswahrscheinlichkeit in einer mehr oder weniger großen Umgebung¹ eines Punktes. Weiter weg von diesem Punkt soll das Wellenpaket eine vernachlässigbare Amplitude besitzen.

Nun hat die Schrödinger-Gleichung (3.1.5) die sehr angenehme Eigenschaft, dass sie **linear** ist, d.h. für sie gilt das Superpositionsprinzip. Das bedeutet, dass zu zwei Lösungen ψ_1 und ψ_2 der Schrödinger-Gleichung auch jede Linearkombination der Gestalt $c_1\psi_1 + c_2\psi_2$ eine Lösung derselben ergibt, wobei c_1 und c_2 beliebige komplexe Konstanten sind. Das sollte entsprechend unseren Quantenpostulaten auch so sein, denn will man die Wellenfunktion als **Wahrscheinlichkeitsamplitude** für die Ortsmessung interpretieren, sollte es Zustände $|\psi(t)\rangle$ und irgendwie **verallgemeinerte Eigenvektoren** für kontinuierliche Eigenwerte \vec{x} des selbstadjungierten Ortsoperators \vec{x} geben, sodass $\psi(t, \vec{x}) = \langle \vec{x} | \psi(t) \rangle$ ist. Da die $|\psi(t)\rangle$ einen Hilbert-Raum bilden sollen, muss demnach zu zwei Vektoren $|\psi\rangle_1$ und $|\psi\rangle_2$ auch deren Linearkombination, d.h. für die Wellenfunktionen die entsprechende Superposition wieder einen möglichen Zustandsvektor ergeben, d.h. aber, dass Superpositionen beliebiger Lösungen der Bewegungsgleichung für die Wellenfunktion wieder Lösungen sein müssen, d.h. diese Wellengleichung sollte eine **lineare Feldgleichung** sein, und das ist in der Tat für die oben postulierte Schrödinger-Gleichung für eine freies Teilchen auch tatsächlich der Fall.

Nun kann man offenbar auch aus beliebig vielen Lösungen der Schrödinger-Gleichung durch Linearkombination bzw. Superposition und schließlich sogar mit kontinuierlich vielen solcher Lösungen neue Lösungen generieren. Auf diese Weise werden wir sehr natürlich auf die *Fourier-Darstellung* der Wellenfunktion geführt, d.h. wir machen den Ansatz

$$\psi(t, \vec{x}) = \int_{\mathbb{R}^3} \frac{d^3\vec{k}}{(2\pi)^{3/2}} A(\vec{k}) \exp[-i\omega(\vec{k})t + i\vec{k}\vec{x}]. \quad (3.1.7)$$

Dies kennen wir bereits aus der Theorie 2-Vorlesung, wo die Fourier-Darstellung elektromagnetischer Wellen ein wichtiges mathematisches Hilfsmittel zur Lösung der Maxwell-Gleichungen war. Dabei ist in (3.1.7) der Faktor $1/(2\pi)^{(3/2)}$ nur aus Bequemlichkeitsgründen eingeführt. Damit dieser Ansatz die Schrödinger-Gleichung erfüllt, gehen wir davon aus, dass wir Integration und Differentiation vertauschen können. Unter diesen Bedingungen die Schrödinger-Gleichung auf (3.1.7) angewandt ergibt sofort wieder die schon oben benutzte Dispersionsrelation

$$\omega(\vec{k}) = \frac{\hbar\vec{k}^2}{2m}. \quad (3.1.8)$$

Diese Dispersionsrelation in (3.1.7) eingesetzt gewährleistet allein schon die Erfüllung der Schrödinger-Gleichung, und zwar für beliebige **Spektralfunktionen** A ! Wir können also für A irgendeine Funktion einsetzen, für die dieses Fourier-Integral einen Sinn ergibt.

Um unsere Forderung nach einem Wellenpaket zu erfüllen, setzen wir für A die einfachst mögliche Form ein, nämlich eine *Gauß-Verteilung*. Wegen der allgemein großen Bedeutung von Gaußverteilungen in der Physik wollen wir hier die Fouriertransformation (3.1.7) ausführlich vorrechnen.

¹Wir werden weiter unten bemerken, dass die Größe dieser Umgebung durch die prinzipiell stets endliche Auflösung eines Teilchendetektors, also eines Meßgeräts für die Anwesenheit des Teilchens, bestimmt ist. Im Unterschied zur klassischen Physik weist dies darauf hin, dass in der Quantenphysik stets der Messprozess mitberücksichtigt wird, auch wenn von konkreten Messgeräten gar nicht die Rede ist.

3. Wellenmechanik

Es sei also die Spektralverteilung durch

$$A(\vec{k}) = N \exp \left[-\frac{(\vec{k} - \vec{k}_0)^2}{4\alpha} \right] \quad (3.1.9)$$

gegeben. Dabei ist N eine Konstante, die wir unten noch genauer festlegen werden. Wir finden jedenfalls für die Schrödingersche Wellenfunktion gemäß (3.1.7 und (3.1.8)

$$\psi(t, \vec{x}) = N \int_{\mathbb{R}^3} \frac{d^3\vec{k}}{(2\pi)^{3/2}} \exp \left[-\frac{(\vec{k} - \vec{k}_0)^2}{4\alpha} - i\frac{\hbar\vec{k}^2}{2m}t + i\vec{k}\vec{x} \right]. \quad (3.1.10)$$

Wir bemerken als erstes, dass sich in diesem Fall die Wellenfunktion in Form eines Produktes aus Wellenfunktionen für jede der drei Raumrichtungen schreiben lässt, d.h. es ist $\psi(t, \vec{x}) = \psi_1(t, x_1)\psi_2(t, x_2)\psi_3(t, x_3)$ mit

$$\psi_1(t, x_1) = N' \int_{\mathbb{R}} \frac{dk_1}{(2\pi)^{1/2}} \exp \left[-\frac{(k_1 - k_{01})^2}{4\alpha} - i\frac{\hbar k_1^2}{2m}t + ik_1x_1 \right]. \quad (3.1.11)$$

Zur Ausführung dieses Integrals berechnen wir zuerst

$$I = \int_{\mathbb{R}^3} \frac{dk_1}{(2\pi)^{1/2}} \exp(-Ak_1^2 + 2ABk_1 + C). \quad (3.1.12)$$

Dazu führen wir im Argument der Exponentialfunktion eine quadratische Ergänzung und eine Substitution $k'_1 = k_1 - B$ aus. Dann folgt (*Nachrechnen!*)

$$I = \exp(C + AB^2) \int_{\mathbb{R}} \frac{dk'_1}{(2\pi)^{1/2}} \exp(-Ak_1'^2). \quad (3.1.13)$$

Verbleibt das letzte Integral zu berechnen. Dazu bemerken wir, dass dieses Integral für A mit positivem Realteil existiert. Für reelle A ist das Integral positiv, und wir erhalten

$$\begin{aligned} I &= \exp(C + AB^2) \sqrt{\int_{\mathbb{R}} dk_1 \int_{\mathbb{R}} dk_2 \exp[-A(k_1^2 + k_2^2)]} \\ &= \exp(C + AB^2) \sqrt{\int_0^\infty K dK \int_0^{2\pi} d\varphi \exp(-AK^2)} = \sqrt{\frac{\pi}{A}} \exp(C + AB^2). \end{aligned} \quad (3.1.14)$$

Damit haben wir das Integral vollständig berechnet. Es sind nur noch die Werte für die Parameter A , B und C einzusetzen. Die Ausdrücke werden etwas länger, so dass wir hier nur das Ergebnis für die Wahrscheinlichkeitsverteilung angeben wollen:

$$w_1(t, x_1) = |\psi_1(t, x_1)|^2 = \frac{|N'|^2 \sqrt{4\alpha m}}{\sqrt{m^2 + 4\alpha^2 \hbar^2 t^2}} \exp \left\{ -\frac{2\alpha m^2}{m^2 + 4\alpha^2 \hbar^2 t^2} \left[\left(x_1 - \frac{k_{01} \hbar}{m} t \right)^2 \right] \right\}. \quad (3.1.15)$$

Die Normierungskonstante N' ist so zu bestimmen, dass

$$\int_{\mathbb{R}} dx_1 w_1(t, x_1) = 1 \quad (3.1.16)$$

wird, entsprechend der Forderung, dass das Teilchen mit Wahrscheinlichkeit 1 an irgendeinem Ort mit der 1-Komponente $x_1 \in \mathbb{R}$ gefunden wird. Unter Verwendung des Integrals (3.1.12) ergibt sich

$$|N'| = \sqrt[4]{\frac{m}{2\pi}} \quad (3.1.17)$$

3.1. Die Schrödinger-Gleichung für freie Teilchen

Wir bemerken, dass die Normierungskonstante **zeitunabhängig** ist. Weiter ist klar, dass bisher keinerlei Hinweise aus irgendeiner Forderung aufgetreten sind, wie der Phasenfaktor (also eine komplexe Zahl vom Betrag 1) von N bzw. N' bestimmt werden soll. Das haben wir schon im vorigen Kapitel festgestellt: Der absolute Phasenfaktor eines Zustandsvektors spielt aufgrund der Bornschen Regel keine Rolle, denn die demgemäß berechneten Wahrscheinlichkeiten hängen von einem Phasenfaktor offenbar nicht ab.

Bevor wir in der allgemeinen Entwicklung der Quantentheorie fortfahren, die diese Beobachtungen klären wird, wollen wir kurz die Implikationen betrachten, die unser spezielles Resultat ergeben hat. Die Wahrscheinlichkeitsverteilung $w(t, \vec{x})$, die sich durch Multiplikation von drei Faktoren der Form w_1 ausdrückt, ist wieder ein Gaußsches Wellenpaket. Das verwundert weiter nicht, denn Gaußsche Glockenkurven sind invariant unter Fouriertransformationen. Physikalisch interessant sind aber die Parameter dieses Gaußschen Pakets.

Wir erinnern kurz an die Bedeutung dieser Parameter. Die allgemeine Form einer Gaußverteilung ist

$$w_G(x) = \frac{1}{\sqrt{2\pi\sigma}} \exp\left[-\frac{(x-x_0)^2}{2\sigma}\right]. \quad (3.1.18)$$

Der Erwartungswert für x ist

$$\langle x \rangle = \int_{\mathbb{R}} dx x w_G(x) = x_0. \quad (3.1.19)$$

Das rechnet man übrigens am geschicktesten durch Ableiten des Normierungsintegrals nach x_0 aus. Das ergibt natürlich wegen des von x_0 unabhängigen Wertes 0. Die Ableitung ist aber $\langle x - x_0 \rangle / \sigma = 0$. Genauso findet man die Varianz, also die mittlere quadratische Abweichung vom Erwartungswert. Der Trick mit der Ableitung liefert

$$\left\langle \frac{\sigma - (x - x_0)^2}{\sigma^2} \right\rangle = 0 \rightarrow \Delta x^2 = \langle (x - x_0)^2 \rangle = \sigma. \quad (3.1.20)$$

Wenden wir das auf (3.1.15) an, ergibt sich

$$\langle \vec{x} \rangle = \frac{\hbar \vec{k}_0}{m} t, \quad (3.1.21)$$

d.h. der Erwartungswert des Ortes des Teilchens ergibt die freie Bewegung eines Newtonschen Massepunktes mit dem Impuls $\vec{p}_0 = \hbar \vec{k}_0$. Die Ortsunschärfe wächst allerdings mit der Zeit gemäß

$$\Delta x^2 = \frac{m^2 + 4\alpha^2 \hbar^2 t^2}{4\alpha m^2}. \quad (3.1.22)$$

Zur Zeit $t = 0$ beträgt die Ortsunschärfe

$$\Delta x^2|_{t=0} = \frac{1}{4\alpha}. \quad (3.1.23)$$

Für unsere Anfangsverteilung für den Impuls war sie hingegen:

$$\Delta p^2 = \hbar^2 \Delta k^2 = \alpha. \quad (3.1.24)$$

Es gilt also

$$\Delta x \Delta p = \frac{\hbar}{2}. \quad (3.1.25)$$

Wir werden im nächsten Abschnitt zeigen, dass dies die unterste überhaupt mögliche Grenze für den Ausdruck $\Delta x \Delta p$ ist (**Heisenbergsche Unschärferelation**), und dass dies genau für die Gaußverteilung, die wir als Anfangsbedingung für die Schrödinger-Gleichung angegeben hatten, zutrifft. In gewissem Sinne ist unser Gaußsches Wellenpaket das Beispiel, das größtmögliche Annäherung an ein klassisches Teilchen bietet, die

im Rahmen der Quantentheorie möglich ist. Da der Impuls für das freie Teilchen eine Erhaltungsgröße ist, ändert sich auch dessen Wahrscheinlichkeitsverteilung nicht mit der Zeit, und die anfängliche Impulsunschärfe bleibt erhalten. Entsprechend dieser Impulsunschärfe wird jedoch die Ortsunschärfe mit der Zeit immer größer, d.h. die anfängliche „Minimalität“ der Orts-Impulsunschärferelation geht mit der Zeit verloren, und zwar desto schneller, je größer die Impulsunschärfe anfänglich war.

3.2 Der Hilbert-Raum

Wir wollen in diesem Abschnitt unsere eben gewonnenen Erfahrungen mit der einfachen Wellenlösung der Schrödinger-Gleichung nutzen, um die Physik der Schrödinger-Gleichung zu präzisieren und auf Teilchen in **äußeren Potentialfeldern** zu erweitern.

Zunächst bemerken wir, dass die oben heuristisch gefundene Methode zur Gewinnung von Lösungen der freien Schrödinger-Gleichung eine einfache mathematische Erklärung besitzt. Betrachten wir einmal für einen Moment die Lösung dieser Gleichung als rein mathematische Aufgabe, d.h. wir fragen uns, welche Eigenschaften die Lösungen dieser Gleichung bestimmen. Die Schrödinger-Gleichung besitzt die Gestalt:

$$i\hbar\partial_t\psi(t,\vec{x})=-\frac{\hbar^2\Delta_x}{2m}\psi(t,\vec{x}). \quad (3.2.1)$$

Aufgrund der Linearität dieser Gleichung bietet sich ein Ansatz in Form einer Fouriertransformierten an:

$$\psi(t,\vec{x})=\int_{\mathbb{R}^4}\frac{d\omega d^3\vec{k}}{(2\pi)^2}\tilde{\psi}(\omega,\vec{k})\exp(-i\omega t+i\vec{k}\vec{x}). \quad (3.2.2)$$

Diesen Ansatz in die Gleichung eingesetzt ergibt eine rein algebraische Gleichung

$$\left(\omega-\frac{\hbar\vec{k}^2}{2m}\right)\tilde{\psi}(\omega,\vec{k})=0. \quad (3.2.3)$$

Damit muss aber $\tilde{\psi}$ zu einer Diracschen δ -Distribution proportional sein, die das Verschwinden der Klammer auf der rechten Seite sicherstellt:

$$\tilde{\psi}(\omega,\vec{k})=\sqrt{2\pi}A(\vec{k})\delta\left(\omega-\frac{\hbar\vec{k}^2}{2m}\right), \quad (3.2.4)$$

wobei A eine beliebige Funktion des Wellenvektors \vec{k} sein darf². Wir werden durch Resubstitution dieser allgemeinen Lösung in (3.2.2) wieder auf unser schon heuristisch gefundenes Resultat (3.1.7)

$$\psi(t,\vec{x})=\int_{\mathbb{R}^3}\frac{d^3\vec{k}}{(2\pi)^{3/2}}A(\vec{k})\exp\left(-i\frac{\hbar\vec{k}^2}{2m}t+i\vec{k}\vec{x}\right) \quad (3.2.5)$$

geführt, wobei gleich auch die Dispersionsrelation (3.1.8) herauskommt. Wir erkennen, dass die allgemeine Lösung der Schrödinger-Gleichung eine willkürliche Funktion $A(\vec{k})$ enthält. Diese Funktion hatten wir oben als Gauß-Paket gewählt, um ein Beispiel vor Augen zu haben. Hier wollen wir nun die physikalische Bedeutung dieser Funktion näher untersuchen.

Die Schrödinger-Gleichung beschreibt offenbar, ganz unabhängig von der Interpretation der Wellenfunktion, einen dynamischen Prozess in Raum und Zeit. Die Lösungen haben nämlich gemäß (3.2.5) Wellencharakter,

²der Faktor $\sqrt{2\pi}$ dient wieder nur dazu, unsere Konvention einzuhalten, auf deren Wahl wir sogleich noch zurückkommen werden.

beschreiben also einen Bewegungsvorgang. Hier haben wir es mit freien Wellen zu tun, also solchen, die keine Quelle besitzen. Wie wir aus der Elektrodynamik oder der Vorstellung von Wasserwellen her wissen, besitzen solche Gleichungen stets die Freiheit der Wahl der Anfangsbedingungen. Die spezifische Situation der Wellenerscheinung wird durch diese Anfangsbedingungen determiniert, und diese Anfangsbedingungen sind durch die Erregung der Wellen zu einem früheren Zeitpunkt bestimmt. Damit beschreibt die Schrödinger-Gleichung einen **kausalen Vorgang**: Aus der „Wellenerregung“ am Anfang der Ausbreitung derselben lässt sich der gesamte Vorgang nach Beendigung der Erregung vollständig aus der Bewegungsgleichung (hier also der Schrödinger-Gleichung) berechnen.

Kehren wir wieder zur mathematischen Analyse der Schrödinger-Gleichung zurück. Die Freiheit der Wahl der Anfangsbedingung ist bei unserem Zugang zur Beschreibung der Lösung durch eine Fouriertransformation (3.2.5) in der Willkür der Wahl der Wellenzahlverteilungsfunktion $A(\vec{k})$ versteckt. Diesem Mangel können wir aber sofort abhelfen, denn für $t = 0$ gilt gemäß (3.2.5)³:

$$\begin{aligned}\psi(0, \vec{x}) &= \int_{\mathbb{R}^3} \frac{d^3 \vec{k}}{(2\pi)^{3/2}} A(\vec{k}) \exp(i\vec{k}\vec{x}) \Leftrightarrow \\ A(\vec{k}) &= \int_{\mathbb{R}^3} \frac{d^3 \vec{y}}{(2\pi)^{3/2}} \psi(0, \vec{y}) \exp(-i\vec{k}\vec{y}).\end{aligned}\tag{3.2.6}$$

Nun können wir aber auch ohne Umweg über den Wellenzahlbereich der Fouriertransformation die Zeitentwicklung angeben, indem wir dieses Resultat in (3.2.5) einsetzen:

$$\psi(t, \vec{x}) = \int_{\mathbb{R}^3} \frac{d^3 \vec{k}}{(2\pi)^{3/2}} \int_{\mathbb{R}^3} \frac{d^3 \vec{y}}{(2\pi)^{3/2}} \psi(0, \vec{y}) \exp \left[-i \frac{\hbar \vec{k}^2}{2m} t + i\vec{k}(\vec{x} - \vec{y}) \right].\tag{3.2.7}$$

Könnten wir nun die Integration nach \vec{k} mit der nach \vec{y} vertauschen, könnten wir schreiben

$$\psi(t, \vec{x}) = \int_{\mathbb{R}^3} d^3 \vec{y} U(t, \vec{x}, \vec{y}) \psi(0, \vec{y}).\tag{3.2.8}$$

Das bedeutet, dass $\psi(t, \vec{x})$ sich durch eine lineare Abbildung aus der willkürlich vorgegebenden Anfangsbedingung $\psi(0, \vec{y})$ berechnen lässt. Nun zeigt aber der Grenzfall $t \rightarrow 0$ schon, dass U keine gewöhnliche Funktion sein kann. Denn dann muss sich ja auf der rechten Seite $\psi(0, \vec{x})$ ergeben, und zwar für alle Funktionen $\psi(0, \vec{x})$, die bzgl. \vec{x} quadratintegrierbar sind (und die folglich auf 1 gemäß der Bornschen Wahrscheinlichkeitsinterpretation normiert werden können), denn mehr brauchen wir physikalisch von dem Anfangszustand $\psi(0, \vec{x})$ nicht zu fordern. Das bedeutet aber, dass in diesem Grenzfall

$$\lim_{t \rightarrow 0^+} U(t, \vec{x}, \vec{y}) = \delta^{(3)}(\vec{x} - \vec{y})\tag{3.2.9}$$

gilt, wobei δ die Diracsche δ -Distribution ist.

Es ist somit zu erwarten, dass auch für $t > 0$ der **Propagator** $U(t, \vec{x}, \vec{y})$ eine Distribution sein wird. Wir können nun aber durch Regularisierung dieser Distribution ihre konkrete Gestalt ausrechnen. Dazu nutzen wir die Tatsache, dass die hier auftretenden Distributionen als Grenzwerte komplexer Funktionen dargestellt werden können. Betrachten wir nämlich den durch naives Vertauschen der beiden Integrationen in (3.2.7) entstehenden Ausdruck, erkennen wir, dass wir durch die Ersetzung $t \rightarrow t - i\epsilon$ mit $\epsilon > 0$, wieder auf ein wohldefiniertes Gaußintegral zurückgeführt werden:

$$U(t, \vec{x}, \vec{y}) = \lim_{\epsilon \rightarrow +0} \int_{\mathbb{R}^3} \frac{d^3 \vec{k}}{(2\pi \hbar)^3} \exp \left[-\frac{\hbar \vec{k}^2}{2m} (\epsilon + it) + i\vec{k}(\vec{x} - \vec{y}) \right].\tag{3.2.10}$$

³Wir wählen $t = 0$ willkürlich als besonders bequemen „Anfangszeitpunkt“.

Wir wenden wieder unsere Formel (3.1.14) an, und finden für $\epsilon \rightarrow 0$ sofort

$$U(t, \vec{x}, \vec{y}) = \left(\frac{m}{2\pi i \hbar t} \right)^{3/2} \exp \left[\frac{im(\vec{x} - \vec{y})^2}{2\hbar t} \right]. \quad (3.2.11)$$

Jetzt können wir aber die allgemeine mathematische Struktur, die hinter der Schrödinger-Gleichung steckt, klar erkennen: Da die Gleichung linear ist, bilden alle Lösungen zusammen einen **komplexen linearen Raum** (Vektorraum mit \mathbb{C} als Skalarkörper). Damit die **Bornsche Interpretation** der Wellenfunktion als Wahrscheinlichkeitsamplitude sinnvoll ist, muss eine physikalisch sinnvolle Wellenfunktion weiter **quadratintegrierbar** sein, d.h. eine physikalische Wellenfunktion $\psi(t, \vec{x})$ muss nicht nur der Schrödinger-Gleichung genügen, sondern es muss auch das Normierungsintegral $\int_{\mathbb{R}^3} d^3\vec{x} \psi^*(t, \vec{x})\psi(t, \vec{x})$ existieren. Diese Funktionen⁴ bilden einen Funktionenraum, den man als **Hilbertschen Funktionenraum** L^2 bezeichnet. Die Norm einer in diesem Raum gelegenen Funktion ist gerade durch das Normierungsintegral gegeben.

Allerdings besitzt dieser Raum noch eine viel weitergehende Struktur. Seien dazu ψ und ϕ beides L^2 -Funktionen. Dann existiert das Integral

$$\langle \psi | \phi \rangle = \int_{\mathbb{R}^3} d^3\vec{x} \psi^*(\vec{x})\phi(\vec{x}), \quad (3.2.12)$$

und die Klammer $\langle \psi | \phi \rangle$ besitzt alle Eigenschaften einer positiv definiten Sesquilinearform, die den Funktionenraum zu einem **Hilbert-Raum** macht, also einem Vektorraum mit Skalarprodukt.

Wir erwähnen hier ohne Beweis, dass dieser Raum auch vollständig ist, d.h. jede Funktionenfolge, die bzgl. der durch das Skalarprodukt induzierten Norm eine **Cauchy-Folge** ist, besitzt einen in L^2 gelegenen Grenzwert.

3.3 Die Observablen-Operatoren

Wir haben im vorigen Abschnitt nun eine mögliche Realisierung des Hilbert-Raums der Zustände für die Bewegung von nichtrelativistischen Teilchen plausibel gemacht, nämlich als **Hilbert-Raum der quadratintegrierbaren Funktionen** L^2 . Entsprechend unseren anhand des Beispiels von Polarisationsmessungen an elektromagnetischen Wellen bzw. Photonen gewonnenen kinematischen Grundpostulaten müssen wir uns nun noch Gedanken machen, wie wir die der Teilchenbewegung angemessenen Observablen als **selbst-adjungierte Operatoren** in diesem Hilbert-Raum definieren. Aus der klassischen Mechanik ist klar, dass die fundamentalen Observablen für ein Teilchen **Ort und Impuls** sind, denn wir können in der klassischen Mechanik die Bewegung eines Teilchens durch die Vorgabe eines Anfangsortes und -impulses eindeutig bestimmen.

Unsere obige Betrachtung hat nun aber primär nur den Ort des Teilchens im Sinne der Bornschen Wahrscheinlichkeitsinterpretation probabilistisch festgelegt, d.h. durch Wahl der Wellenfunktion zur Zeit $t = 0$ haben wir lediglich die Wahrscheinlichkeitsverteilung für den Ort des Teilchens angegeben. Die Schrödinger-Gleichung ermöglichte dann (hier zunächst nur für freie Teilchen) die Berechnung der Wahrscheinlichkeitsverteilung zu jedem anderen Zeitpunkt. Der Impuls kam nur (zusammen mit der Energie) über die Dispersionsrelation im Zusammenhang mit der Fourier-Transformation ins Spiel. Dies können wir aber nun benutzen, um die selbstadjungierten Operatoren zu finden, die die Komponenten des **Ortsvektors** \vec{x} und des **Impulses** \vec{p} im abstrakten Formalismus der Quantentheorie repräsentieren. Beginnen wir mit dem Ort. Gemäß unserem Grundpostulat, ist die Wahrscheinlichkeitsverteilung für den Ortsvektor durch

$$W(\vec{x}) = |\psi(\vec{x})|^2 \quad (3.3.1)$$

gegeben. Definitionsgemäß ist dann der **Erwartungswert für den Ortsvektor** des Teilchens durch

$$\langle \vec{x} \rangle = \int_{\mathbb{R}^3} d^3x W(\vec{x})\vec{x} = \int_{\mathbb{R}^3} d^3x |\psi(\vec{x})|^2 \vec{x} \quad (3.3.2)$$

⁴Wir gehen nicht auf die Subtilität ein, dass wir eigentlich Klassen von Funktionen, die sich voneinander nur auf Lebesgueschen Nullmengen unterscheiden, betrachten müssten

3.3. Die Observablen-Operatoren

gegeben. Andererseits würden wir dies aber gern in der Form (2.2.27) ausdrücken. Mit der Definition des Skalarprodukts in (3.2.12) bedeutet dies

$$\langle \vec{x} \rangle = \langle \psi | \vec{x} \psi \rangle = \int_{\mathbb{R}^3} d^3 x \psi^*(\vec{x}) [\vec{x} \psi(\vec{x})]. \quad (3.3.3)$$

Schreiben wir in (3.3.2) $|\psi(\vec{x})|^2 = \psi^*(\vec{x})\psi(\vec{x})$, ist klar, dass (3.3.3) mit (3.3.2) übereinstimmt, wenn wir den **Ortsoperator** durch

$$\vec{x} \psi(\vec{x}) = \vec{x} \psi(\vec{x}) \quad (3.3.4)$$

definieren. Es ist leicht zu zeigen, dass \vec{x} ein selbstadjungierter Operator ist (*Übungsaufgabe!*).

Etwas schwieriger wird nun die Argumentation für den Impulsoperator. Dazu erinnern wir uns unserer heuristischen Überlegungen zu den Materiewellen. Wir hatten den Impuls des Teilchens mit den Wellenzahlen in der Fourier-Darstellung der Wellenfunktion über die de Broglie-Beziehung $\vec{p} = \hbar \vec{k}$ assoziiert. Schreiben wir also die Fourier-Transformation (3.2.5) für $t = 0$ zunächst als Integral über die Impulskomponenten um:

$$\psi(\vec{x}) = \frac{1}{(2\pi\hbar)^{3/2}} \int_{\mathbb{R}^3} d^3 p \tilde{\psi}(\vec{p}) \exp\left(\frac{i\vec{p} \cdot \vec{x}}{\hbar}\right). \quad (3.3.5)$$

Dabei wird die spezielle Wahl des Normierungsfaktors vor dem Integral gleich noch klar werden. Da die Fourier-Transformation umkehrbar ist, ist die Wellenfunktion $\tilde{\psi}$ im Impulsraum offenbar eine äquivalente Beschreibung des quantenmechanischen Zustandes. Es liegt nahe, diese Funktion über die Bornsche Wahrscheinlichkeitsinterpretation mit der Wahrscheinlichkeitsverteilung für den Impuls des Teilchens

$$\tilde{W}(\vec{p}) = |\tilde{\psi}(\vec{p})|^2 \quad (3.3.6)$$

in Verbindung zu bringen, wobei freilich die Funktion wieder auf 1 normiert sein muss, d.h.

$$\langle \psi | \psi \rangle = \int_{\mathbb{R}^3} d^3 p |\tilde{\psi}(\vec{p})|^2 = 1. \quad (3.3.7)$$

Dass diese Überlegung in sich konsistent ist, zeigt sich daran, dass $\psi(\vec{x})$ und $\tilde{\psi}(\vec{p})$ über die Fourier-Transformation (3.3.5) verknüpft sind und die Skalarprodukte übereinstimmen, unabhängig davon, ob wir sie in der Orts- oder Impulsdarstellung der Wellenfunktion berechnen.

Um das zu sehen, betrachten wir gleich zwei unterschiedliche Zustände $|\psi\rangle$ und $|\phi\rangle$ und drücken das Skalarprodukt zunächst in der Ortsdarstellung aus und rechnen es vermöge der Fourier-Transformation (3.3.5) (und genau analog für $|\phi\rangle$) auf die Impulswellenfunktionen um:

$$\begin{aligned} \langle \psi | \phi \rangle &= \int_{\mathbb{R}^3} d^3 \vec{x} \psi^*(\vec{x}) \phi(\vec{x}) \\ &= \int_{\mathbb{R}^3} d^3 \vec{x} \frac{1}{(2\pi\hbar)^{3/2}} \int_{\mathbb{R}^3} d^3 p_1 \tilde{\psi}^*(\vec{p}_1) \exp\left(-\frac{i\vec{p}_1 \cdot \vec{x}}{\hbar}\right) \frac{1}{(2\pi\hbar)^{3/2}} \int_{\mathbb{R}^3} d^3 p_2 \tilde{\phi}(\vec{p}_2) \exp\left(\frac{i\vec{p}_2 \cdot \vec{x}}{\hbar}\right). \end{aligned} \quad (3.3.8)$$

Wir fassen nun diesen Ausdruck zusammen und führen das räumliche Integral zuerst aus. Dies liefert

$$\begin{aligned} \langle \psi | \phi \rangle &= \frac{1}{(2\pi\hbar)^3} \int_{\mathbb{R}^3} d^3 p_1 \int_{\mathbb{R}^3} d^3 p_2 \tilde{\psi}^*(\vec{p}_1) \phi(\vec{p}_2) \int_{\mathbb{R}^3} d^3 x \exp\left[\frac{i(\vec{p}_1 - \vec{p}_2) \cdot \vec{x}}{\hbar}\right] \\ &= \frac{1}{(2\pi\hbar)^3} \int_{\mathbb{R}^3} d^3 p_1 \int_{\mathbb{R}^3} d^3 p_2 \tilde{\psi}^*(\vec{p}_1) \phi(\vec{p}_2) (2\pi)^3 \delta^{(3)}\left(\frac{\vec{p}_1 - \vec{p}_2}{\hbar}\right) \\ &= \frac{1}{(2\pi\hbar)^3} \int_{\mathbb{R}^3} d^3 p_1 \int_{\mathbb{R}^3} d^3 p_2 \tilde{\psi}^*(\vec{p}_1) \phi(\vec{p}_2) (2\pi)^3 \hbar^3 \delta^{(3)}(\vec{p}_1 - \vec{p}_2) \\ &= \int_{\mathbb{R}^3} d^3 p_1 \tilde{\psi}^*(\vec{p}_1) \phi(\vec{p}_1). \end{aligned} \quad (3.3.9)$$

3. Wellenmechanik

In der Tat wird also das Skalarprodukt von Hilbert-Raumvektoren durch die Impulswellenfunktionen genauso berechnet wie im Ortsraum. Dies ist auch durch die spezifische Wahl des Normierungsfaktors in der Definition der Fouriertransformation (3.3.5) bedingt.

Wie beim Operator für den Ortsvektor können wir nun auch für den Impulsoperator argumentieren. Für die Impulswellenfunktion gilt demnach

$$\vec{\mathbf{p}}\tilde{\psi}(\vec{p}) = \vec{p}\tilde{\psi}(\vec{p}). \quad (3.3.10)$$

Nun wollen wir aber auch den Impulsoperator für Ortswellenfunktionen bestimmen. Dazu müssen wir nur (3.3.10) durch Fourier-Transformation auf die entsprechende Ortswellenfunktionen umrechnen:

$$\begin{aligned} \vec{\mathbf{p}}\psi(\vec{x}) &= \frac{1}{(2\pi\hbar)^{3/2}} \int_{\mathbb{R}^3} d^3p \vec{p}\tilde{\psi}(\vec{p}) \exp\left(\frac{i\vec{p}\cdot\vec{x}}{\hbar}\right) \\ &= \frac{1}{(2\pi\hbar)^{3/2}} \int_{\mathbb{R}^3} d^3p \tilde{\psi}(\vec{p}) \frac{\hbar}{i} \vec{\nabla} \exp\left(\frac{i\vec{p}\cdot\vec{x}}{\hbar}\right) \\ &= -i\hbar \vec{\nabla} \frac{1}{(2\pi\hbar)^{3/2}} \int_{\mathbb{R}^3} d^3p \tilde{\psi}(\vec{p}) \exp\left(\frac{i\vec{p}\cdot\vec{x}}{\hbar}\right) \\ &= -i\hbar \vec{\nabla} \psi(\vec{x}). \end{aligned} \quad (3.3.11)$$

Damit ist aber der Impulsoperator in seiner Wirkung auf Ortswellenfunktionen bestimmt $\vec{\mathbf{p}} = -i\hbar\vec{\nabla}$.

Nun fügt sich unser Formalismus auch in die Überlegungen im vorigen Abschnitt bzgl. der **Eigenwerte** und **Eigenvektoren** der Operatoren ein. Allerdings ergeben sich einige Schwierigkeiten, da wir es nun mit **kontinuierlichen** Variablen zu tun haben. Die Rechnung zeigt auch, dass $\vec{\mathbf{p}}$ selbstadjungiert ist. In der Darstellung mit Impulswellenfunktionen ist dies selbstverständlich. Im Ortsraum muss man nur partiell integrieren, wobei man annehmen darf, dass die Wellenfunktionen wegen ihrer Quadratintegrabilität hinreichend schnell im Unendlichen verschwinden (*Nachrechnen!*):

$$\begin{aligned} \langle \psi | \vec{\mathbf{p}}\phi \rangle &= \int_{\mathbb{R}^3} d^3x \psi^*(\vec{x}) (-i\hbar\vec{\nabla})\phi(\vec{x}) \\ &= \int_{\mathbb{R}^3} d^3x [-i\hbar\vec{\nabla}\psi(\vec{x})]^* \phi(\vec{x}) \\ &= \langle \vec{\mathbf{p}}\psi | \phi \rangle. \end{aligned} \quad (3.3.12)$$

Das bedeutet aber, dass $\vec{\mathbf{p}}$ in der Tat ein selbstadjungierter Operator ist.

Betrachten nun die **Impulseigenfunktionen**, die durch die Gleichung

$$\vec{\mathbf{p}}u_{\vec{p}}(\vec{x}) = -i\hbar\vec{\nabla}u_{\vec{p}}(\vec{x}) = \vec{p}u_{\vec{p}}(\vec{x}) \quad (3.3.13)$$

definiert sind. Die Eigenwerte \vec{p} sollten dabei reell sein, da ja der Impulsoperator selbstadjungiert ist. Die Gleichung wird offenbar durch

$$u_{\vec{p}}(\vec{x}) = N(\vec{p}) \exp\left(\frac{i\vec{p}\cdot\vec{x}}{\hbar}\right) \quad (3.3.14)$$

gelöst. Dabei ist $N(\vec{p})$ eine zunächst unbestimmte Normierung. Nun ist aber diese Funktion sicher nicht quadratintegrabel, denn $|u_{\vec{p}}(\vec{x})|^2 = |N(\vec{p})|^2 = \text{const}$. Wir haben es also nicht mit einem Element des Hilbert-Raums der quadratintegrablen Funktionen zu tun. Allerdings ergibt die Lösung als **verallgemeinerte Eigenfunktion** dennoch einen Sinn. Dazu berechnen wir formal das Skalarprodukt zwischen zwei verschiedenen

3.3. Die Observablen-Operatoren

Impulseigenfunktionen:

$$\begin{aligned}
 \langle u_{\vec{p}_1} | u_{\vec{p}_2} \rangle &= \int_{\mathbb{R}^3} d^3x u_{\vec{p}_1}^*(\vec{x}) u_{\vec{p}_2}(\vec{x}) \\
 &= \int_{\mathbb{R}^3} d^3x N^*(\vec{p}_1) N(\vec{p}_2) \exp\left(\frac{i(\vec{p}_2 - \vec{p}_1) \cdot \vec{x}}{\hbar}\right) \\
 &= (2\pi)^3 |N(\vec{p}_1)|^2 \delta^{(3)}\left(\frac{\vec{p}_1 - \vec{p}_2}{\hbar}\right) \\
 &= (2\pi \hbar)^3 |N(\vec{p}_1)|^2 \delta^{(3)}(\vec{p}_1 - \vec{p}_2).
 \end{aligned} \tag{3.3.15}$$

Offenbar können wir in gewisser Weise eine Art **verallgemeinertes Orthonormalsystem** von Eigenvektoren des Impulsoperators definieren, wenn wir den Normierungsfaktor so bestimmen, dass

$$(2\pi \hbar)^3 |N(\vec{p})|^2 = 1 \Rightarrow N(\vec{p}) = \frac{1}{(2\pi \hbar)^{3/2}}. \tag{3.3.16}$$

Wobei $N(\vec{p})$ nur bis auf einen Phasenfaktor $\exp(i\varphi)$ mit $\varphi \in \mathbb{R}$ bestimmt ist. Wie wir im vorigen Abschnitt gesehen haben, sind aber wegen der Bornschen Regel, in die nur das Betragsquadrat der Wellenfunktionen eingehen, solche Phasenfaktoren ohnehin unwichtig. Wir definieren also im Folgenden die **Impulseigenfunktionen** durch

$$u_{\vec{p}}(\vec{x}) = \frac{1}{(2\pi \hbar)^{3/2}} \exp\left(\frac{i\vec{p} \cdot \vec{x}}{\hbar}\right). \tag{3.3.17}$$

Zugleich ist nun aber auch die **Vollständigkeit** dieser Eigenvektoren gewährleistet, denn für irgendwelche quadratintegrablen Ortswellenfunktion gilt

$$\begin{aligned}
 \int_{\mathbb{R}^3} d^3p \langle \psi | u_{\vec{p}} \rangle \langle u_{\vec{p}} | \phi \rangle &= \frac{1}{(2\pi \hbar)^3} \int_{\mathbb{R}^3} d^3p \int_{\mathbb{R}^3} d^3x_1 \int_{\mathbb{R}^3} d^3x_2 \psi^*(\vec{x}_1) \phi(\vec{x}_2) \exp\left[\frac{i\vec{p} \cdot (\vec{x}_1 - \vec{x}_2)}{\hbar}\right] \\
 &= \frac{1}{(2\pi \hbar)^3} \int_{\mathbb{R}^3} d^3x_1 \int_{\mathbb{R}^3} d^3x_2 \psi^*(\vec{x}_1) \phi(\vec{x}_2) \int_{\mathbb{R}^3} d^3p \exp\left[\frac{i\vec{p} \cdot (\vec{x}_1 - \vec{x}_2)}{\hbar}\right] \\
 &= \frac{1}{(2\pi \hbar)^3} \int_{\mathbb{R}^3} d^3x_1 \int_{\mathbb{R}^3} d^3x_2 \psi^*(\vec{x}_1) \phi(\vec{x}_2) (2\pi \hbar)^3 \delta^{(3)}(\vec{x}_1 - \vec{x}_2) \\
 &= \int_{\mathbb{R}^3} d^3x_1 \psi^*(\vec{x}_1) \phi(\vec{x}_1) = \langle \psi | \phi \rangle.
 \end{aligned} \tag{3.3.18}$$

Dies bedeutet aber, dass

$$\int_{\mathbb{R}^3} d^3p |u_{\vec{p}}\rangle \langle u_{\vec{p}}| = \mathbb{1} \tag{3.3.19}$$

ist, d.h. dass die verallgemeinerten Impulseigenfunktionen (3.3.17) in dem Sinne vollständig sind, dass (3.3.19) gilt. Dabei wird, entsprechend dem kontinuierlichen Charakter des Impulses, allerdings nicht über alle möglichen Eigenwerte summiert sondern integriert. Für den Impuls sind die verallgemeinerten Eigenwerte demnach der ganze \mathbb{R}^3 .

Berechnen wir entsprechend die verallgemeinerten **Impulskomponenten** des Zustandsvektors, erhalten wir

$$\langle u_{\vec{p}} | \psi \rangle = \int_{\mathbb{R}^3} d^3x u_{\vec{p}}^*(\vec{x}) \psi(\vec{x}) = \frac{1}{(2\pi \hbar)^{3/2}} \int_{\mathbb{R}^3} d^3x \psi(x) \exp\left(-\frac{i\vec{p} \cdot \vec{x}}{\hbar}\right) \psi(\vec{x}). \tag{3.3.20}$$

Dies ist aber genau die Umkehr-Fourier-Transformation zu (3.2.5) (*Nachprüfen!*), d.h. es folgt

$$\langle u_{\vec{p}} | \psi \rangle = \tilde{\psi}(\vec{p}). \tag{3.3.21}$$

Es liegt nun nahe, entsprechende Überlegungen auch für den Ortsoperator anzustellen. Wir können aber aufgrund unserer Erfahrungen mit dem Impulsoperator, den wir über die **Ortsdarstellung** (also unter Verwendung von Ortswellenfunktionen) bestimmt haben, diese Rechnung erheblich abkürzen. Bezeichnen wir die verallgemeinerten Ortseigenvektoren mit $v_{\vec{x}}$ sollten diese im verallgemeinerten Sinne orthonormiert sein, d.h.

$$\langle v_{\vec{x}_1} | v_{\vec{x}_2} \rangle = \delta^{(3)}(\vec{x}_1 - \vec{x}_2) \quad (3.3.22)$$

und die Vollständigkeitsrelation

$$\int_{\mathbb{R}^3} d^3x |v_{\vec{x}}\rangle \langle v_{\vec{x}}| = \mathbb{1} \quad (3.3.23)$$

erfüllen, und die Ortswellenfunktion ist demnach

$$\langle v_{\vec{x}} | \psi \rangle = \psi(\vec{x}). \quad (3.3.24)$$

Für die verallgemeinerten Impulseigenvektoren folgt aus (3.3.17)

$$u_{\vec{p}}(\vec{x}) = \frac{1}{(2\pi\hbar)^{3/2}} \exp\left(\frac{i\vec{p} \cdot \vec{x}}{\hbar}\right) = \langle v_{\vec{x}} | u_{\vec{p}} \rangle. \quad (3.3.25)$$

Entsprechend ist die Impulsdarstellung der verallgemeinerten Ortseigenfunktion durch

$$\tilde{v}_{\vec{x}}(\vec{p}) = \langle u_{\vec{p}} | v_{\vec{x}} \rangle = u_{\vec{p}}^*(\vec{x}) = \frac{1}{(2\pi\hbar)^{3/2}} \exp\left(-\frac{i\vec{p} \cdot \vec{x}}{\hbar}\right) \quad (3.3.26)$$

gegeben.

Entsprechend einfach ist es nun auch, mit diesem Formalismus den Ortsoperator in der Impulsdarstellung zu berechnen. Dazu muss man nur geschickt „Einsoperatoren“ in Form geeigneter Vollständigkeitsrelationen einschieben. Um im Folgenden die Schreibweise übersichtlicher zu gestalten, schreiben wir einfach

$$|\vec{x}\rangle = |v_{\vec{x}}\rangle, \quad |\vec{p}\rangle = |u_{\vec{p}}\rangle. \quad (3.3.27)$$

Für den Ortsoperator in der Impulsdarstellung folgt also mit dieser Schreibweise (*Nachrechnen!*)

$$\begin{aligned} \vec{x}\tilde{\psi}(\vec{p}) &= \langle \vec{p} | \vec{x}\psi \rangle = \int_{\mathbb{R}^3} d^3x \langle \vec{p} | \vec{x} \rangle \langle \vec{x} | \vec{x}\psi \rangle \\ &= \int_{\mathbb{R}^3} d^3x \langle \vec{p} | \vec{x} \rangle \langle \vec{x} | \psi \rangle \vec{x} = \frac{1}{(2\pi\hbar)^{3/2}} \int_{\mathbb{R}^3} d^3x \exp\left(-\frac{i\vec{p} \cdot \vec{x}}{\hbar}\right) \psi(\vec{x}) \vec{x} \\ &= \frac{1}{(2\pi\hbar)^{3/2}} i\hbar \vec{\nabla}_p \int_{\mathbb{R}^3} d^3x \exp\left(-\frac{i\vec{p} \cdot \vec{x}}{\hbar}\right) \psi(\vec{x}) = i\hbar \vec{\nabla}_p \tilde{\psi}(\vec{p}). \end{aligned} \quad (3.3.28)$$

Dabei ist $\vec{\nabla}_{\vec{p}} = (\partial_{p_1}, \partial_{p_2}, \partial_{p_3})$ der Nabla-Operator bzgl. Impulskomponenten.

3.4 Verträglichkeit von Observablen

Wir haben nun gesehen, wie wir die grundlegenden Observablen der Mechanik Ort und Impuls als selbst-adjungierte Vektoren im Hilbert-Raum der Zustände beschreiben können. Entsprechend dem Bornschen Postulat ist die **Wahrscheinlichkeitsverteilung** (oder **Wahrscheinlichkeitsdichte**), ein im Zustand $|\psi\rangle$ präpariertes Teilchen am Ort \vec{x} zu finden, gemäß der Bornschen Regel durch

$$W(\vec{x}) = |\langle \vec{x} | \psi \rangle|^2 = |\psi(\vec{x})|^2 \quad (3.4.1)$$

3.4. Verträglichkeit von Observablen

gegeben ist, d.h. die **Wahrscheinlichkeit**, das Teilchen in einem kleinen Volumenelement d^3x um den Ort \vec{x} zu finden, ist $d^3x W(\vec{x})$. Dabei ergeben sich hier Wahrscheinlichkeitsdichten, weil die Ortsvektorkomponenten kontinuierliche Werte annehmen können, und demnach die verallgemeinerten Ortseigenzustände Distributionen sind, wie anhand der Normierungsbedingung $\langle \vec{x} | \vec{x}' \rangle = \delta^{(3)}(\vec{x} - \vec{x}')$ klar wird.

Genauso ist die Wahrscheinlichkeitsdichte für den Impuls durch

$$\tilde{W}(\vec{p}) = |\langle \vec{p} | \psi \rangle|^2 = |\tilde{\psi}(\vec{p})|^2 \quad (3.4.2)$$

gegeben. Dabei ist die Impulswellenfunktion vermöge (3.3.17) und der Vollständigkeitsrelationen für die verallgemeinerten Orts- und Impulseigenzustände

$$\int_{\mathbb{R}^3} d^3p |\vec{p}\rangle \langle \vec{p}| = \int_{\mathbb{R}^3} d^3x |\vec{x}\rangle \langle \vec{x}| = \mathbb{1} \quad (3.4.3)$$

durch die Fourier-Transformation der Ortswellenfunktion

$$\tilde{\psi}(\vec{p}) = \langle \vec{p} | \psi \rangle = \int_{\mathbb{R}^3} d^3x \langle \vec{p} | \vec{x} \rangle \langle \vec{x} | \psi \rangle = \int_{\mathbb{R}^3} d^3x \frac{1}{(2\pi\hbar)^{3/2}} \exp\left(-i\frac{\vec{x}\cdot\vec{p}}{\hbar}\right) \psi(\vec{x}) \quad (3.4.4)$$

gegeben.

Anhand der Beispiele mit Gaußschen Wellenpaketen in Abschnitt 3.1 ist klar, dass je genauer die Ortskomponente x_1 eines Teilchens bestimmt ist, d.h. je kleiner die Standardabweichung Δx_1 dieser Ortskomponente ist, desto größer ist die Standardabweichung der entsprechenden Impulskomponente Δp_1 und umgekehrt. Es scheint demnach so zu sein, dass man (a) weder Ort noch Impuls eines Teilchens präzise festlegen kann, denn dann müsste man es in einem Orts- bzw. Impulseigenzustand präparieren können. Das ist aber unmöglich, da sowohl die Orts- und Impulseigenzustände keine auf Eins normierbaren Hilbert-Raumvektoren sind. Ist formal das Teilchen im Impulseigenzustand $|\vec{p}\rangle$ präpariert, ist die Ortswellenfunktion $\langle \vec{x} | \vec{p} \rangle \propto \exp(i\vec{x}\cdot\vec{p}/\hbar)$ und dessen Betragsquadrat also eine Konstante. Dies lässt sich bei Integration über den ganzen \mathbb{R}^3 sicher nicht auf Eins normieren. Umgekehrt kann man für ein in einem Ortseigenzustand präpariertes Teilchen hinsichtlich der Impulsmessung argumentieren. Außerdem scheint es so, als wäre (b) bei einer genauen Festlegung des Ortes (d.h. Δx „klein“), also bei genauer Lokalisierung des Teilchens, der Impuls nur sehr ungenau bestimmt (d.h. Δp „groß“). Umgekehrt ist ein Teilchen, dessen Impuls genau festgelegt wird, nur sehr ungenau lokalisiert.

Es erhebt sich nun die Frage, wie man diesen Sachverhalt präzisieren kann, d.h. wann zwei Observablen A und B zugleich präzise Werte annehmen können (man sagt dann, dass A und B zueinander **kompatible Observablen** sind) und wie sich die gegenseitige „Unschärfe“ zueinander **inkompatibler Observabler** quantifizieren lässt.

Dazu erinnern wir uns, dass die Observablen A und B jeweils durch einen selbstadjungierten Operator \mathbf{A} und \mathbf{B} im Hilbertraum beschrieben werden. Die möglichen Messwerte sind dann die Eigenwerte a_j und b_k dieser Observablen, und diese Messwerte werden genau dann mit Sicherheit angenommen, wenn das Teilchen in einem der dazugehörigen Eigenzustände $|a_j\rangle$ bzw. $|b_k\rangle$ präpariert ist. Wir gehen hierbei der Einfachheit halber zunächst davon aus, dass die Eigenvektoren der Operatoren auf Eins normierbar sind und zu jedem Eigenwert nur genau ein linear unabhängiger Eigenvektor existiert. Wir können später leicht auf den allgemeineren Fall von Observablen mit kontinuierlichen Eigenwerten und den Fall, dass es zu Eigenwerten mehr als einen linear unabhängigen Eigenvektor gibt (dann sagt man der betreffende Eigenwert sei **entartet**). Gehen wir also von nichtausgearteten Eigenwerten aus.

Dann kann man offenbar genau dann das Teilchen immer in Zuständen präparieren, wo *beide* Observablen scharf festgelegte Werte besitzen, wenn es ein vollständiges Orthonormalsystem von **simultanen Eigenvek-**

toren $|a_j, b_k\rangle$ gibt, d.h.

$$\mathbf{A}|a_j, b_k\rangle = a_j|a_j, b_k\rangle, \quad \mathbf{B}|a_j, b_k\rangle = b_k|a_j, b_k\rangle, \quad (3.4.5)$$

$$\langle a_j, b_k | a_{j'}, b_{k'} \rangle = \delta_{jj'} \delta_{kk'}, \quad (3.4.6)$$

$$\sum_{j,k} |a_j, b_k\rangle \langle a_j, b_k| = \mathbb{1}. \quad (3.4.7)$$

Dann folgt aber

$$\mathbf{A} = \mathbf{A}\mathbb{1} = \mathbf{A} \sum_{j,k} |a_j, b_k\rangle \langle a_j, b_k| = \sum_{j,k} a_j |a_j, b_k\rangle \langle a_j, b_k| \quad (3.4.8)$$

und genau analog

$$\mathbf{B} = \mathbf{B}\mathbb{1} = \mathbf{B} \sum_{j,k} |a_j, b_k\rangle \langle a_j, b_k| = \sum_{j,k} b_k |a_j, b_k\rangle \langle a_j, b_k| \quad (3.4.9)$$

Es lässt sich dann aber leicht nachrechnen, dass die **Operatoren notwendig kommutieren**, d.h. dass $\mathbf{AB} = \mathbf{BA}$ ist. Um das einzusehen, verwenden wir (3.4.5-3.4.9) (*Nachrechnen!*): Einerseits gilt

$$\begin{aligned} \mathbf{AB} &= \sum_{j,k,j',k'} a_j b_{k'} |a_j, b_k\rangle \langle a_j, b_k | a_{j'}, b_{k'}\rangle \langle a_{j'}, b_{k'}| \\ &= \sum_{j,k,j',k'} a_j b_{k'} |a_j, b_k\rangle \delta_{jj'} \delta_{kk'} \langle a_{j'}, b_{k'}| \\ &= \sum_{j,k} a_j b_k |a_j, b_k\rangle \langle a_j, b_k|. \end{aligned} \quad (3.4.10)$$

Andererseits erhält man dasselbe Resultat, wenn man in dieser Rechnung \mathbf{A} und \mathbf{B} vertauscht, d.h. es ist also in der Tat

$$\mathbf{AB} = \mathbf{BA}. \quad (3.4.11)$$

Dies schreibt man auch gern mit Hilfe des **Kommutators** der Operatoren, d.h.

$$[\mathbf{A}, \mathbf{B}] = \mathbf{AB} - \mathbf{BA} = \mathbf{0}. \quad (3.4.12)$$

Man sagt demnach auch, dass zwei Observablen A und B nur dann kompatibel sein können, wenn die entsprechenden Operatoren **kommutieren**, d.h. wenn ihr Kommutator verschwindet. In der Mathematik zeigt man, dass auch umgekehrt zwei kommutierende selbstadjungierte Operatoren immer ein vollständiges Orthonormalsystem gemeinsamer Eigenvektoren besitzen.

Das ist auch für die Komponenten des Ortsoperators der Fall, denn offenbar ist $|\vec{x}\rangle = |(x_1, x_2, x_3)\rangle$ ein gemeinsamer (verallgemeinerter) Eigenvektor der Komponenten des Ortsoperators. In der Ortsdarstellung bedeutet ja auch der Operator einer Ortskomponente x_j einfach die Multiplikation der Wellenfunktion mit x_j , und demnach kommutieren in der Tat die Komponenten des Ortsvektors. Ähnlich verhält es sich mit den Komponenten des Impulsvektors. In der Ortsdarstellung sind sie nämlich durch die partiellen Ableitungen $\hat{p}_j = -i\hbar\partial_j$ gegeben, die für hinreichend glatte Wellenfunktionen kommutieren.

Orts- und Impulskomponenten vertauschen allerdings i.a. nicht mehr, denn es gilt (in der Ortsdarstellung)

$$[\hat{x}_j, \hat{p}_k] \psi(\vec{x}) = x_k (-i\hbar)\partial_j \psi - (-i\hbar)\partial_j (x_k \psi) = i\hbar (-x_k \partial_j \psi + \delta_{jk} \psi + x_k \partial_j \psi) = i\hbar \delta_{jk} \psi. \quad (3.4.13)$$

Da dies für jede Wellenfunktion, also alle Hilbert-Raumvektoren gilt, folgt die Kommutatorrelation für Ortsvektor- und Impulskomponenten

$$[\mathbf{x}_j, \mathbf{p}_k] = i\hbar \delta_{jk} \mathbb{1}. \quad (3.4.14)$$

3.4. Verträglichkeit von Observablen

Ortsvektor- und Impulskomponenten in gleicher Richtung kommutieren also *nicht*, und demnach können diese beiden Observablen nicht kompatibel sein, d.h. man kann i.a. keinen auf Eins normierbaren gemeinsamen Eigenvektor finden.

Wir können nun aber für zwei Observablen A und B , repräsentiert durch die entsprechenden selbstadjungierten Operatoren \mathbf{A} und \mathbf{B} die sog. **Heisenbergsche Unschärferelation** herleiten. Diese besagt, dass für jeden Zustand $|\psi\rangle$ stets

$$\Delta A \Delta B \geq \frac{1}{2} |\langle [\mathbf{A}, \mathbf{B}] \rangle| \quad (3.4.15)$$

gilt. Dabei sind ΔA und ΔB die Standardabweichungen der Observablen A und B für den vorgegebenen Zustand $|\psi\rangle$, d.h.

$$\Delta A^2 = \langle A^2 \rangle - \langle A \rangle^2, \quad (3.4.16)$$

wobei der Erwartungswert einer beliebigen Observablen C durch

$$\langle C \rangle = \langle \psi | C | \psi \rangle \quad (3.4.17)$$

gegeben ist, wobei wir davon ausgehen, dass $\langle \psi | \psi \rangle = 1$ ist.

Wir wollen nun die Unschärferelation beweisen. Dazu bemerken wir, dass

$$\Delta A^2 = \langle (A - \langle A \rangle)^2 \rangle \quad (3.4.18)$$

ist (*warum?*). Setzen wir also $\mathbf{A}' = \mathbf{A} - \langle A \rangle \mathbb{1}$ und $\mathbf{B}' = \mathbf{B} - \langle B \rangle \mathbb{1}$, ist $\Delta A^2 = \langle A'^2 \rangle$ und $\Delta B^2 = \langle B'^2 \rangle$.

Wir betrachten nun das quadratische Polynom

$$P(\lambda) = \langle (\mathbf{A}' + i\lambda\mathbf{B}')\psi | (\mathbf{A}' + i\lambda\mathbf{B}')\psi \rangle. \quad (3.4.19)$$

Wegen der positiven Semidefinitheit des Skalarprodukts gilt für alle $\lambda \in \mathbb{R}$, dass $P(\lambda) \geq 0$ ist. Da offenbar \mathbf{A}' und \mathbf{B}' selbstadjungiert sind, gilt für $\lambda \in \mathbb{R}$ (*warum?*)

$$P(\lambda) = \langle \psi | (\mathbf{A}' - i\lambda\mathbf{B}')(\mathbf{A}' + i\lambda\mathbf{B}')\psi \rangle \geq 0. \quad (3.4.20)$$

Nun gilt

$$(\mathbf{A}' - i\lambda\mathbf{B}')(\mathbf{A}' + i\lambda\mathbf{B}') = \mathbf{A}'^2 + i\lambda(\mathbf{A}'\mathbf{B}' - \mathbf{B}'\mathbf{A}') + \mathbf{B}'^2\lambda^2. \quad (3.4.21)$$

Damit ist

$$P(\lambda) = \langle A'^2 \rangle + \langle B'^2 \rangle \lambda^2 + i\lambda \langle [\mathbf{A}', \mathbf{B}'] \rangle = \Delta A^2 + \Delta B^2 \lambda^2 + i\lambda \langle [\mathbf{A}', \mathbf{B}'] \rangle \geq 0. \quad (3.4.22)$$

Jetzt nehmen wir an, dass $\langle B'^2 \rangle \neq 0$ ist. Dann bedeutet diese Gleichung, dass das quadratische Polynom $P(\lambda)$ höchstens eine Nullstelle haben kann, d.h. in der Lösungsgleichung für die quadratische Gleichung

$$\lambda_{1,2} = -\frac{i}{2\Delta B^2} \langle [\mathbf{A}', \mathbf{B}'] \rangle \pm \sqrt{\frac{1}{4\Delta B^4} \langle i[\mathbf{A}', \mathbf{B}'] \rangle^2 - \frac{\Delta A^2}{\Delta B^2}} \quad (3.4.23)$$

ist der Ausdruck unter der Wurzel ≤ 0 . Da weiter $[\mathbf{A}', \mathbf{B}'] = [\mathbf{A}, \mathbf{B}]$ ist, folgt somit

$$\frac{1}{4\Delta B^4} \langle [\mathbf{A}, \mathbf{B}] \rangle^2 - \frac{\Delta A^2}{\Delta B^2} \leq 0. \quad (3.4.24)$$

Nach einigen einfachen Umformungen bedeutet dies

$$\Delta A^2 \Delta B^2 \geq \frac{1}{4} \langle [\mathbf{A}, \mathbf{B}] \rangle^2. \quad (3.4.25)$$

3. Wellenmechanik

Wir müssen jetzt nur noch zeigen, dass der Erwartungswert des Kommutators tatsächlich reell ist. Das ist zwar wegen (3.4.22) klar, weil ja für $\lambda \in \mathbb{R}$ gemäß der Herleitung $P(\lambda) \in \mathbb{R}$ ist. Es ist trotzdem hilfreich, es nochmals explizit nachzurechnen. Dazu zeigen wir, dass $i[\mathbf{A}, \mathbf{B}]$ selbstadjungiert ist, wenn \mathbf{A} und \mathbf{B} selbstadjungierte Operatoren sind. Wegen $(\mathbf{AB})^\dagger = \mathbf{B}^\dagger \mathbf{A}^\dagger = \mathbf{BA}$ folgt

$$(i[\mathbf{A}, \mathbf{B}])^\dagger = (-i)(\mathbf{AB} - \mathbf{BA})^\dagger = (-i)(\mathbf{BA} - \mathbf{AB}) = +i(\mathbf{AB} - \mathbf{BA}) = i[\mathbf{A}, \mathbf{B}], \quad (3.4.26)$$

d.h. $i[\mathbf{A}, \mathbf{B}]$ ist tatsächlich selbstadjungiert. Demnach muss der Erwartungswert in der Tat reell sein, denn für einen jeden selbstadjungierten Operator \mathbf{C} gilt

$$\langle \mathbf{C} \rangle^* = \langle \psi | \mathbf{C} \psi \rangle^* = \langle \mathbf{C} \psi | \psi \rangle = \langle \psi | \mathbf{C}^\dagger \psi \rangle = \langle \psi | \mathbf{C} \psi \rangle = \langle \mathbf{C} \rangle \Rightarrow \langle \mathbf{C} \rangle \in \mathbb{R}. \quad (3.4.27)$$

Ziehen wir also aus (3.4.25) die Wurzel, folgt tatsächlich die Heisenbergsche Unschärferelation (3.4.15).

Insbesondere lautet die Heisenbergsche Unschärferelation für Ortsvektor- und Impulskomponenten wegen (3.4.14)

$$\Delta x_j \Delta p_k \geq \frac{\hbar}{2} \delta_{jk}. \quad (3.4.28)$$

Offenbar gilt insbesondere das Gleichheitszeichen in der Unschärferelation genau dann, wenn es ein (und dann nur ein!) $\lambda \in \mathbb{R}$ gibt mit $P(\lambda) = 0$, d.h. für dieses λ ist aufgrund der Definition des Polynoms gemäß (3.4.19)

$$(\mathbf{A}' + i\lambda \mathbf{B}')|\psi\rangle = 0. \quad (3.4.29)$$

Wir wollen für die Orts- und Impulskomponenten in einer Raumrichtung die entsprechenden Zustände „minimaler Unschärfe“ ermitteln. Der Einfachheit betrachten wir dazu ein Teilchen, das sich nur in einer Richtung bewegen kann und nennen die entsprechenden Operatoren für die Ortskoordinate und den Impuls \mathbf{x} und \mathbf{p} . Wir arbeiten am bequemsten in der Ortsdarstellung⁵. Setzen wir $\mathbf{A}' = \mathbf{x} - x_0 \mathbb{1}$ und $\mathbf{B}' = \mathbf{p} - p_0 \mathbb{1}$, wobei x_0 und p_0 die Erwartungswerte von \mathbf{x} bzw. \mathbf{p} bezeichnen sollen. Dann lautet in der Ortsdarstellung die Gleichung (3.4.29)

$$(\hat{x} - x_0)\psi(x) = -i\lambda(\hat{p} - p_0)\psi(x). \quad (3.4.30)$$

Dabei ist $\lambda \in \mathbb{R}$ eine beliebige reelle Zahl. Wegen $\hat{x} = x$ und $\hat{p} = -i\hbar \partial_x$ müssen wir also die Differentialgleichung

$$(x - x_0)\psi(x) = -\lambda\hbar\psi'(x) + i\lambda p_0\psi(x) \quad (3.4.31)$$

lösen. Formen wir etwas um, folgt

$$\lambda\hbar\psi'(x) = [-(x - x_0) + i\lambda p_0]\psi(x). \quad (3.4.32)$$

Dividieren wir durch $\lambda\hbar\psi(x)$ erhalten wir

$$\frac{\psi'(x)}{\psi(x)} = \frac{d}{dx} \ln\left(\frac{\psi(x)}{N}\right) = -\frac{x - x_0}{\lambda\hbar} + \frac{i}{\hbar} p_0. \quad (3.4.33)$$

Dabei ist N eine beliebige Konstante. Die allgemeine Lösung von (3.4.31) ist damit schnell durch Integration gewonnen:

$$\psi(x) = N \exp\left[-\frac{(x - x_0)^2}{2\lambda\hbar} + \frac{i p_0 x}{\hbar}\right]. \quad (3.4.34)$$

Dies ist offenbar eine Gauß-Verteilung. Die Normierungskonstante ergibt sich bis auf einen irrelevanten Phasenfaktor zu

$$N = \frac{1}{(\hbar\lambda\pi)^{1/4}}. \quad (3.4.35)$$

⁵Die Rechnung lässt sich analog auch in der Impulsdarstellung ausführen, was dem Leser zur *Übung* empfohlen sei!

3.5. Dynamik

Dann ist die Wellenfunktion normiert, d.h. es gilt $\langle \psi | \psi \rangle = 1$. Die Ortswahrscheinlichkeitsverteilung $w(x) = |\psi(x)|^2$ zeigt, dass tatsächlich $\langle x \rangle = x_0$ und $\Delta x = \sqrt{\lambda \hbar / 2}$. Für den Erwartungswert und die Standardabweichung von p ergibt sich

$$\langle p \rangle = \int_{-\infty}^{\infty} dx \psi^*(x) (-i\hbar) \psi'(x) = p_0, \quad \Delta p^2 = -p_0^2 + \int_{-\infty}^{\infty} dx \psi^*(x) (-\hbar^2) \psi''(x) = \frac{\hbar}{2\lambda}, \quad (3.4.36)$$

d.h. es gilt in der Tat $\Delta x \Delta p = \hbar/2$, und zwar für jedes $\lambda > 0$ (*vollziehen Sie diese Rechnungen nach!*). Für diese Gaußschen Wellenpakete ist also tatsächlich das Unschärfeprodukt minimal, und man kann dabei offenbar entweder den Ort oder den Impuls beliebig genau festlegen, wobei dann aber die jeweils andere Größe entsprechend ungenauer bestimmt ist.

Wir bemerken noch, dass ein **vollständiger Satz kompatibler Observabler**, repräsentiert durch die selbstadjungierten Operatoren $\mathbf{A}_1, \dots, \mathbf{A}_n$ dadurch charakterisiert ist, dass zum Einen all diese Observablen simultan determinierte Werte besitzen können, d.h. nach den obigen Überlegungen, dass notwendig $[\mathbf{A}_j, \mathbf{A}_k] = 0$ für alle $j, k \in \{1, \dots, n\}$ (Kompatibilität) und zum anderen ein vollständiges Orthonormalsystem simultaner Eigenvektoren $|a_1, \dots, a_n\rangle$ zu den Eigenwerten (a_1, \dots, a_n) existiert, und zwar so, dass zu jedem simultanen Satz von Eigenwerten der entsprechende Eigenvektorraum eindimensional ist, d.h. es existiert nur ein linear unabhängiger simultaner Eigenvektor.

Wie wir am Beispiel von Ortsvektorkomponenten und Impulskomponenten gesehen haben, kann es sich bei den Eigenvektoren auch um verallgemeinerte Eigenvektoren handeln, d.h. verallgemeinerte Eigenvektoren zu verallgemeinerten Eigenwerten im kontinuierlichen Spektrum der selbstadjungierten Operatoren sind nicht auf 1 normierbar und repräsentieren daher auch keine Zustände. Die Eigenvektoren lassen sich dann aber i.a. „auf δ -Distributionen“ normieren.

3.5 Dynamik

Nun haben wir die quantentheoretische **Kinematik** für den Fall eines „quantisierten klassischen Teilchens“ vollständig durch die Festlegung der selbstadjungierten Operatoren für die Ortsvektorkomponenten und die Impulskomponenten sowie den dazugehörigen Hilbert-Raum, der in der Orts- oder Impulsdarstellung durch den Hilbert-Raum der quadratintegrierbaren komplexen Funktionen $L^2(\mathbb{R}^3)$ eindeutig charakterisiert ist.

Kommen wir nun auf die Schrödinger-Gleichung (3.2.1) für ein freies Teilchen zurück. Die rechte Seite können wir dabei mit Hilfe der Impulsoperatoren ausdrücken, denn in der in (3.2.1) offenbar zur Anwendung kommenden Ortsdarstellung ist $\hat{\vec{p}} = -i\hbar \vec{\nabla}$ und damit $-\hbar^2 \Delta = \hat{\vec{p}}^2$. Damit können wir die Zeitabhängigkeit des Zustandes auch darstellungsunabhängig, also für die abstrakten Ket-Vektoren charakterisieren:

$$i\hbar \partial_t |\psi(t)\rangle = \frac{\hat{\vec{p}}^2}{2m} |\psi(t)\rangle. \quad (3.5.1)$$

Der Operator auf der rechten Seite ist nun offenbar der **Operator der kinetischen Energie**. Das ist eigentlich angesichts unserer heuristischen Herleitung der Schrödinger-Gleichung anhand der de Broglie-Welle-Teilchen-Dualismus-Argumentation wenig überraschend, denn wir hatten ja die Dispersionsrelation so gewählt, dass für die ebenen Wellen eines freien Teilchens $E = \hbar\omega$ und $\vec{p} = \hbar\vec{k}$ gilt. In diesem Zusammenhang bezeichnet man die **Gesamtenergie** auch als **Hamilton-Funktion** bzw. entsprechend im abstrakten Formalismus der Quantentheorie den entsprechenden Operator der Gesamtenergie als **Hamilton-Operator**. Für ein freies Teilchen ist demnach der Hamilton-Operator

$$\mathbf{H} = \frac{\hat{\vec{p}}^2}{2m}, \quad (3.5.2)$$

und die Schrödinger-Gleichung lautet entsprechend (3.5.2)

$$i\hbar\partial_t|\psi(t)\rangle = \mathbf{H}|\psi(t)\rangle. \quad (3.5.3)$$

Dies ist in zweierlei Hinsicht von großer Bedeutung für die weitere Entwicklung der Theorie. Zum einen können wir jetzt auch die quantenmechanische Dynamik für kompliziertere Fälle als für ein freies Teilchen erraten. Betrachtet man z.B. ein Teilchen, das sich in einem äußeren Kraftfeld befindet, das in der klassischen Mechanik durch ein Potential $V(\vec{x})$ beschrieben wird, so liegt es nahe, dass dann auch der quantenmechanische Hamilton-Operator entsprechend der Gesamtenergie des Systems durch

$$\mathbf{H} = \frac{\vec{\mathbf{p}}^2}{2m} + V(\vec{x}) \quad (3.5.4)$$

gegeben sein wird.

Dabei müssen wir uns nur noch darüber klar werden, wie man **Funktionen von Operatoren** zu interpretieren hat. Dabei sind vor allem zwei Ideen wichtig. Falls die Funktionen Polynome sind, wie bei (3.5.2) ist unmittelbar klar, was mit der entsprechenden Operatorfunktion gemeint ist, denn hier ist $\vec{\mathbf{p}}^2$ einfach die zweimalige Anwendung der entsprechenden Impulsoperatoren auf den Zustandsket, also die Hintereinanderausführung der entsprechenden linearen Abbildungen auf die Hilbert-Raumvektoren. Entsprechend können wir Funktionen, die sich in **Potenzreihen** entwickeln lassen, zumindest formal auf die entsprechenden Potenzreihen mit Operatoren in den (i.a. unendlich vielen) Summanden auffassen.

Hierbei ergibt sich allerdings zuweilen ein Problem, wenn man Funktionen für das klassische Teilchen vorliegen hat, die von Observablen abhängen, für die die entsprechenden selbstadjungierten Operatoren nicht kommutieren. Dann muss man noch irgendwie eine Reihenfolge für die Operatoren festlegen, was i.a. nicht eindeutig ist. Für den Fall des Hamilton-Operators eines Teilchens in einem äußeren Kraftpotential (3.5.4) tritt dieses Problem nicht auf, denn sowohl die Impulsoperatoren als auch die Ortsoperatoren kommutieren untereinander.

Eine weitere Möglichkeit, operatorwertige Funktionen zu definieren, ist es, in einer spezifischen Darstellung zu arbeiten, also mit geeigneten (verallgemeinerten) vollständigen Orthonormalbasen von Eigenvektoren der im Problem vorkommenden Operatoren. Für (3.5.4) bietet sich als besonders einfach die Ortsdarstellung an, d.h. die Verwendung der simultanen Eigenvektoren der drei Ortsvektorkomponenten \mathbf{x}_j ($j \in \{1, 2, 3\}$), denn dann ist offenbar

$$\hat{H}\psi(t, \vec{x}) = \langle \vec{x} | \mathbf{H} | \psi(t) \rangle = \left\langle \vec{x} \left| \frac{\vec{\mathbf{p}}^2}{2m} + V(\vec{x}) \right| \psi(t) \right\rangle. \quad (3.5.5)$$

Nun kennen wir bereits die Ortsdarstellung des Impulsoperators:

$$\hat{p}\psi(t, \vec{x}) = \langle \vec{x} | \vec{\mathbf{p}} | \psi(t) \rangle = -i\hbar\vec{\nabla}\psi(t, \vec{x}) \Rightarrow \hat{p}^2\psi(t, \vec{x}) = -\hbar^2\Delta\psi(t, \vec{x}). \quad (3.5.6)$$

Für den Operator der potentiellen Energie ist offenbar

$$V(\hat{x})\psi(t, \vec{x}) = \langle \vec{x} | V(\vec{x}) | \psi(t) \rangle = \langle V(\vec{x})\vec{x} | \psi(t) \rangle. \quad (3.5.7)$$

Dabei haben wir verwendet, dass $V(\vec{x})$ ein **selbstadjungierter Operator** ist, denn mit $\mathbf{x}_k = \mathbf{x}_k^\dagger$ ist auch jedes Produkt von Ortsvektorkomponentenoperatoren selbstadjungiert, weil die Operatoren allesamt miteinander kommutieren. Z.B. gilt $(\mathbf{x}_k\mathbf{x}_j)^\dagger = \mathbf{x}_j^\dagger\mathbf{x}_k^\dagger = \mathbf{x}_j\mathbf{x}_k = \mathbf{x}_k\mathbf{x}_j$, d.h. $\mathbf{x}_k\mathbf{x}_j$ ist tatsächlich selbstadjungiert. Da $V(\vec{x})$ als Funktion für gewöhnliche Vektoren der klassischen Mechanik stets reell ist, kommen in der entsprechenden Potenzreihe nur reelle Koeffizienten vor, und damit ist auch die formale Potenzreihe mit Operatoren statt gewöhnlicher Vektoren ein selbstadjungierter Operator. Nun ist aber schließlich offenbar

$$V(\vec{x})|\vec{x}\rangle = V(\vec{x})|\vec{x}\rangle, \quad (3.5.8)$$

3.5. Dynamik

d.h. die (verallgemeinerten) Ortseigenvektoren sind auch Eigenvektoren der entsprechenden Funktion dieser Operatoren. Es genügt wieder, sich dies anhand eines Operatorprodukts aus zwei Ortskomponentenoperatoren klar zu machen:

$$\mathbf{x}_j \mathbf{x}_k |\vec{x}\rangle = \mathbf{x}_j x_k |\vec{x}\rangle = x_k \mathbf{x}_j |\vec{x}\rangle = x_k x_j |\vec{x}\rangle. \quad (3.5.9)$$

Damit ist aber einfach aufgrund von (3.5.7)

$$V(\hat{\vec{x}})\psi(t, \vec{x}) = V(\vec{x})\psi(t, \vec{x}), \quad (3.5.10)$$

denn da $V(\vec{x}) \in \mathbb{R}$, ist

$$\langle V(\vec{x})\vec{x} | \psi(t) \rangle = \langle V(\vec{x})\vec{x} | \psi(t) \rangle = V^*(\vec{x}) \langle \vec{x} | \psi(t) \rangle = V(\vec{x})\psi(t, \vec{x}). \quad (3.5.11)$$

Schließlich können wir damit auch Funktionen $V(\vec{x})$ die sich nicht (oder zumindest nicht bequem) in Potenzreihen entwickeln lassen, als Funktionen $V(\vec{x})$ mit den Ortskomponentenoperatoren interpretieren. Freilich sind all diese Betrachtungen nicht mathematisch streng begründet. Letztlich erweist sich aber dieser pragmatisch heuristische Zugang zur **Quantisierung klassischer mechanischer Systeme** als äußerst erfolgreich, wie wir im Folgenden noch sehen werden. Jedenfalls ist schließlich durch diese Überlegungen der Hamilton-Operator (3.5.4) in der Ortsdarstellung mathematisch wohldefiniert:

$$\hat{H}\psi(t, \vec{x}) = -\frac{\hbar^2}{2m}\Delta\psi(t, \vec{x}) + V(\vec{x})\psi(t, \vec{x}). \quad (3.5.12)$$

Schließlich können wir nun auch die Schrödingergleichung im abstrakten Hilbert-Raum

$$i\hbar\partial_t |\psi(t)\rangle = \mathbf{H} |\psi(t)\rangle \quad (3.5.13)$$

formal leicht lösen. Dazu behandeln wir wieder den Hamilton-Operator als wäre er eine ganz gewöhnliche Zahl. Dann ist offenbar eine Lösung für die Schrödinger-Gleichung bei vorgegebener Anfangsbedingung $|\psi(t_0)\rangle = |\psi_0\rangle$, wobei $|\psi_0\rangle$ der Zustand des Teilchens ist, in dem es zur Anfangszeit t_0 präpariert wurde

$$|\psi(t)\rangle = \exp\left[-\frac{i\mathbf{H}}{\hbar}(t-t_0)\right] |\psi_0\rangle = \mathbf{U}(t-t_0) |\psi_0\rangle. \quad (3.5.14)$$

Der Operator $\mathbf{U}(t-t_0)$ heißt **Zeitentwicklungsoperator**.

Diese formale Lösung der Schrödinger-Gleichung zeigt weiter die Wichtigkeit der Tatsache, dass \mathbf{H} als Repräsentant einer Observablen (nämlich der Gesamtenergie des Teilchens) ein selbstadjungierter Operator ist, denn es gilt

$$\mathbf{U}^\dagger(t-t_0) = \exp\left[+\frac{i\mathbf{H}^\dagger}{\hbar}(t-t_0)\right] = \exp\left[+\frac{i\mathbf{H}}{\hbar}(t-t_0)\right]. \quad (3.5.15)$$

Da trivialerweise \mathbf{H} mit sich selbst vertauscht, kann man mit dieser **Operatorexponentialfunktion** wie mit einer gewöhnlichen Funktion rechnen, und daher gilt

$$\mathbf{U}^\dagger(t-t_0)\mathbf{U}(t-t_0) = \mathbf{U}(t-t_0)\mathbf{U}^\dagger(t-t_0) = \mathbb{1}. \quad (3.5.16)$$

Folglich ist

$$\mathbf{U}^\dagger(t-t_0) = \mathbf{U}^{-1}(t-t_0). \quad (3.5.17)$$

Man nennt einen Operator, dessen adjungierter Operator der inverse Operator ist, **unitär**. Die wichtigste Eigenschaft solcher unitärer Operatoren ist, dass sie Skalarprodukte beliebiger Hilbert-Raumvektoren ungedändert lassen, denn ist \mathbf{U} ein beliebiger unitärer Operator, so gilt

$$\langle \mathbf{U}\phi | \mathbf{U}\psi \rangle = \langle \phi | \mathbf{U}^\dagger \mathbf{U}\psi \rangle = \langle \phi | \mathbf{U}^{-1} \mathbf{U}\psi \rangle = \langle \phi | \mathbb{1}\psi \rangle = \langle \phi | \psi \rangle. \quad (3.5.18)$$

3. Wellenmechanik

Damit ergibt sich aus (3.5.15) insbesondere, dass

$$\langle \psi(t) | \psi(t) \rangle = \langle \psi_0 | \psi_0 \rangle, \quad (3.5.19)$$

d.h. normiert man $|\psi_0\rangle$ anfangs (zur Zeit t_0) auf Eins, entsprechend der Bornschen Wahrscheinlichkeitsinterpretation, so bleibt der Vektor auch zu jeder späteren Zeit auf Eins normiert.

Schließlich können wir auch die Frage beantworten, wie zu einer beliebigen Observablen A , repräsentiert durch den selbstadjungierten Operator \mathbf{A} , der Operator definiert werden kann, der der Zeitableitung \dot{A} entspricht. Diesen Operator wollen wir mit $\dot{\mathbf{A}}$ bezeichnen. Dazu betrachten wir den Erwartungswert der Observablen A als Funktion der Zeit, der durch

$$\langle A \rangle(t) = \langle \psi(t) | \mathbf{A} | \psi(t) \rangle \quad (3.5.20)$$

gegeben ist. Dabei gehen wir der Einfachheit davon aus, dass $\mathbf{A} = \mathbf{A}(\vec{x}, \vec{p})$ nicht irgendwie explizit von der Zeit abhängt. Wir definieren nun $\dot{\mathbf{A}}$ durch die Forderung, dass

$$\langle \dot{A} \rangle(t) = \frac{d}{dt} \langle A \rangle(t) = \langle \psi(t) | \dot{\mathbf{A}} | \psi(t) \rangle \quad (3.5.21)$$

sein soll. Dazu müssen wir nur (3.5.20) mit Hilfe der Produktregel nach der Zeit ableiten

$$\frac{d}{dt} \langle A \rangle(t) = \langle \partial_t \psi(t) | \mathbf{A} | \psi(t) \rangle + \langle \psi(t) | \mathbf{A} | \partial_t \psi(t) \rangle \quad (3.5.22)$$

und die Schrödinger-Gleichung (3.5.13) anwenden. Es gilt demnach

$$\partial_t |\psi(t)\rangle = -\frac{i}{\hbar} \mathbf{H} |\psi(t)\rangle \quad (3.5.23)$$

und durch Adjungieren

$$\partial_t \langle \psi(t) | = +\frac{i}{\hbar} \langle \psi(t) | \mathbf{H}^\dagger = +\frac{i}{\hbar} \langle \psi(t) | \mathbf{H}, \quad (3.5.24)$$

wobei wir im letzten Schritt wieder die Selbstadjungiertheit von \mathbf{H} verwendet haben. Setzen wir also (3.5.23) und (3.5.24) in (3.5.22) ein, erhalten wir nach einigen einfachen Umformungen

$$\frac{d}{dt} \langle A \rangle(t) = \left\langle \psi(t) \left| \frac{1}{i\hbar} [\mathbf{A}, \mathbf{H}] \right| \psi(t) \right\rangle. \quad (3.5.25)$$

Da dies für alle Zustandsvektoren $|\psi(t)\rangle$ gelten soll, ergibt der Vergleich mit der rechten Seite von (3.5.21)

$$\dot{\mathbf{A}} = \frac{1}{i\hbar} [\mathbf{A}, \mathbf{H}]. \quad (3.5.26)$$

Dabei haben wir vorausgesetzt, dass $\mathbf{A} = \mathbf{A}(\vec{x}, \vec{p})$ ist. Falls die Observable auch explizit von der Zeit abhängt, erhält man beim Ableiten noch einen Zusatzterm, der der Ableitung des Operators nach der expliziten Zeitabhängigkeit entspricht (*Nachrechnen!*)

$$\dot{\mathbf{A}} = \frac{1}{i\hbar} [\mathbf{A}, \mathbf{H}] + \left(\frac{\partial \mathbf{A}}{\partial t} \right)_{\text{expl}}. \quad (3.5.27)$$

Betrachten wir als einfache Beispiele die Zeitableitungen des Ortsvektors und des Impulses, woraus sich die Operatoren für **Geschwindigkeit** und **Kraft** ergeben. Dabei gehen wir vom Hamilton-Operator (3.5.4) aus.

3.5. Dynamik

Um (3.5.26) anwenden zu können, müssen wir dazu zunächst einige **Rechenregeln für Kommutatoren** von Operatoren bzw. Funktionen von Operatoren herleiten. Zunächst ist klar, dass der Kommutator in beiden Argumenten linear ist, d.h. für irgendwelche Operatoren \mathbf{A} , \mathbf{B} und \mathbf{C} und komplexe Zahlen λ , μ gilt

$$[\lambda\mathbf{A} + \mu\mathbf{B}, \mathbf{C}] = \lambda[\mathbf{A}, \mathbf{C}] + \mu[\mathbf{B}, \mathbf{C}]. \quad (3.5.28)$$

Außerdem ist

$$[\mathbf{AB}, \mathbf{C}] = \mathbf{ABC} - \mathbf{CAB} = \mathbf{A}[\mathbf{B}, \mathbf{C}] + (\mathbf{AC} - \mathbf{CA})\mathbf{B} \quad (3.5.29)$$

oder in eine griffige Formel gebracht

$$[\mathbf{AB}, \mathbf{C}] = \mathbf{A}[\mathbf{B}, \mathbf{C}] + [\mathbf{A}, \mathbf{C}]\mathbf{B}. \quad (3.5.30)$$

Da nun alle Ortskomponentenoperatoren untereinander vertauschen, gilt offenbar

$$[\vec{\mathbf{x}}, V(\vec{\mathbf{x}})] = 0. \quad (3.5.31)$$

Damit bleibt

$$\vec{\mathbf{v}} = \overset{\circ}{\vec{\mathbf{x}}} = \frac{1}{i\hbar} \left[\vec{\mathbf{x}}, \frac{1}{2m} \vec{\mathbf{p}}^2 \right] = \frac{1}{2mi\hbar} [\vec{\mathbf{x}}, \vec{\mathbf{p}}^2] \quad (3.5.32)$$

Um die verbliebene Kommutatorklammer besser auswerten zu können schreiben wir sie mit Hilfe des Ricci-Kalküls aus (s. Skript zur Vorlesung „Theorie 2 für das Lehramt L3“ [Hee18b])⁶

$$[\mathbf{x}_j, \vec{\mathbf{p}}^2] = [\mathbf{x}_j, \mathbf{p}_k \mathbf{p}_k] = \mathbf{p}_k [\mathbf{x}_j, \mathbf{p}_k] + [\mathbf{x}_j, \mathbf{p}_k] \mathbf{p}_k = \mathbf{p}_k 2i\hbar \delta_{jk} \mathbb{1} = 2i\hbar \mathbf{p}_j. \quad (3.5.33)$$

Setzt man dies wieder in (3.5.32) ein, ergibt dies

$$\vec{\mathbf{v}} = \frac{1}{m} \vec{\mathbf{p}}. \quad (3.5.34)$$

Dies ist kein besonders überraschendes Ergebnis, entspricht dies doch genau der Definition des Impulses in der klassischen Mechanik. Es ist allerdings auf der anderen Seite schon bemerkenswert, dass dies im Sinne von Beziehungen zwischen den die Observablen darstellenden Operatoren auch in der Quantentheorie genauso verhält.

Für den Operator der Kraft stehen wir allerdings vor einem neuen Problem: Da die Impulskomponenten untereinander kommutieren, gilt

$$\vec{\mathbf{F}} = \overset{\circ}{\vec{\mathbf{p}}} = \frac{1}{i\hbar} [\vec{\mathbf{p}}, \mathbf{H}] = \frac{1}{i\hbar} [\vec{\mathbf{p}}, V(\vec{\mathbf{x}})], \quad (3.5.35)$$

und wir benötigen den Kommutator der Impulskomponenten mit einer beliebigen Funktion der Ortskomponenten.

Um diesen Kommutator auszurechnen, verwenden wir die **Ortsdarstellung** und die entsprechenden Operatoren für Ort und Impuls. Es ist zum Einen für eine beliebige Wellenfunktion $\psi(\vec{\mathbf{x}}) = \langle \vec{\mathbf{x}} | \psi \rangle$

$$\begin{aligned} \hat{p}_j V(\vec{\mathbf{x}}) \psi(\vec{\mathbf{x}}) &= \hat{p}_j V(\vec{\mathbf{x}}) \psi(\vec{\mathbf{x}}) \\ &= -i\hbar \partial_j [V(\vec{\mathbf{x}}) \psi(\vec{\mathbf{x}})] \\ &= -i\hbar [\psi(\vec{\mathbf{x}}) \partial_j V(\vec{\mathbf{x}}) + V(\vec{\mathbf{x}}) \partial_j \psi(\vec{\mathbf{x}})] \\ &= -i\hbar \psi(\vec{\mathbf{x}}) \partial_j V(\vec{\mathbf{x}}) + V(\vec{\mathbf{x}}) \hat{p}_j \psi(\vec{\mathbf{x}}). \end{aligned} \quad (3.5.36)$$

⁶Man beachte insbesondere, dass über Paare gleichnamiger Indizes, wie hier k , gemäß der Einsteinschen Summenkonvention über $k \in \{1, 2, 3\}$ summiert wird!

und damit

$$\left[\hat{p}_j, V(\hat{\vec{x}}) \right] \psi(\vec{x}) = \hat{p}_j [V(\vec{x})\psi(x)] - V(\vec{x})\hat{p}_j \psi(\vec{x}) = -i\hbar \psi(\vec{x}) \partial_j V(\vec{x}). \quad (3.5.37)$$

Da dies für alle Wellenfunktion ψ gilt, folgt daraus

$$\left[\mathbf{p}_j, V(\vec{x}) \right] = -i\hbar \partial_j V(\vec{x}). \quad (3.5.38)$$

Setzen wir dies in (3.5.35) ein, und kehren dabei wieder zur Vektorschreibeweise zurück, folgt

$$\vec{\mathbf{F}} = -\vec{\nabla} V(\vec{x}), \quad (3.5.39)$$

d.h. es gilt wieder die Analogie zwischen den klassischen Beziehungen zwischen den verschiedenen Observablen und den entsprechenden Operatoren in der Quantentheorie.

3.6 Energieeigenvektoren

Die Energieeigenvektoren, d.h. Eigenvektoren des Hamiltonoperators spielen eine besondere Rolle. Zum einen stellen sie die **stationären Zustände** des Teilchens dar. Dies sieht man wie folgt ein: Angenommen der Zustand des Teilchens zur Zeit $t = t_0$ ist ein Eigenvektor des Hamiltonoperators $|\psi_0\rangle = |E\rangle$ mit $\mathbf{H}|E\rangle = E|E\rangle$. Mit (3.5.14) folgt dann für die Zeitentwicklung des Zustandes

$$|\psi(t)\rangle = \exp\left(-\frac{i\mathbf{H}}{\hbar}(t-t_0)\right)|\psi_0\rangle = \exp\left(-\frac{iE}{\hbar}(t-t_0)\right)|\psi_0\rangle. \quad (3.6.1)$$

Die zeitliche Änderung besteht also lediglich aus der Multiplikation mit einem Phasenfaktor. Dies ändert aber den Zustand prinzipiell nicht, denn beobachtbar sind nur die Wahrscheinlichkeiten für die Messung von Observablen. Betrachtet man beispielsweise die simultane Messung eines vollständigen Satzes von Observablen A_1, \dots, A_n , so besteht die einzige physikalische Information des Zustandes darin, dass die Wahrscheinlichkeit (oder Wahrscheinlichkeitsdichte, wenn kontinuierliche Observablen beteiligt sind), bei der Messung zur Zeit t bestimmte Werte (a_1, \dots, a_n) zu erhalten, durch

$$\begin{aligned} W(t, a_1, \dots, a_n) &= |\langle a_1, \dots, a_n | \psi(t) \rangle|^2 \\ &= \left| \langle a_1, \dots, a_n | \psi_0 \rangle \exp\left(-\frac{iE}{\hbar}(t-t_0)\right) \right|^2 \\ &= |\langle a_1, \dots, a_n | \psi_0 \rangle|^2 \end{aligned} \quad (3.6.2)$$

gegeben ist. Ist also der Systemzustand anfangs ein Energieeigenzustand zum Eigenwert E , verbleibt das System für alle Zeiten in diesem Zustand, und die Wahrscheinlichkeiten für beliebige Messungen hängen nicht von der Zeit ab. Daher sind die Energieeigenzustände die **stationären Zustände** des Systems.

Da der Hamiltonoperator \mathbf{H} selbstadjungiert sein muss, weil er die Observable Energie repräsentiert, gibt es auch ein vollständiges Orthonormalsystem von Energieeigenvektoren. I.a. bildet \mathbf{H} allein keinen vollständigen Satz von Operatoren, d.h. i.a. sind die Energieeigenwerte „entartet“. Dem tragen wir Rechnung, dass wir den vollständigen Satz von Energieeigenvektoren mit $|E, \alpha\rangle$ bezeichnen. Dabei bedeutet das zusätzliche Argument α einen oder mehrere reelle Parameter, die die Energiezustände zum gleichen Eigenwert identifizieren. Dabei kann α i.a. auch kontinuierlich sein. Im folgenden schreiben wir einfach Summen. Wenn α kontinuierliche Werte annehmen kann ist die entsprechende Summe durch ein Integral zu ersetzen.

Angenommen, wir haben ein VONS von Energieeigenvektoren für einen gegebenen Hamiltonoperator berechnet, lässt sich mit deren Hilfe sehr leicht die Zeitentwicklung eines beliebigen Anfangszustandes berechnen. Dazu müssen wir nur die Vollständigkeitsrelation

$$\sum_{E, \alpha} |E, \alpha\rangle \langle E, \alpha| = \mathbb{1} \quad (3.6.3)$$

3.7. Formale Herleitung der Wellenmechanik

und die formale Lösung der zeitabhängigen Schrödingergleichung (3.5.14) verwenden. Demnach ist nämlich

$$\begin{aligned}
 |\psi(t)\rangle &= \exp\left(-\frac{i\mathbf{H}}{\hbar}(t-t_0)\right)|\psi_0\rangle \\
 &= \sum_{E,\alpha} \exp\left(-\frac{i\mathbf{H}}{\hbar}(t-t_0)\right)|E,\alpha\rangle \langle E,\alpha|\psi_0\rangle \\
 &= \sum_{E,\alpha} \exp\left(-\frac{iE}{\hbar}(t-t_0)\right)|E,\alpha\rangle \langle E,\alpha|\psi_0\rangle.
 \end{aligned} \tag{3.6.4}$$

Die Lösung der Schrödinger-Gleichung für die Zeitentwicklung des Zustandes ist damit in Form einer Entwicklung nach Energieeigenfunktionen gegeben.

Betrachten wir als Beispiel ein freies Teilchen mit dem Hamiltonoperator

$$\mathbf{H} = \frac{1}{2m}\vec{p}^2. \tag{3.6.5}$$

Es ist sofort klar, dass ein vollständiger Satz von Energieeigenvektoren offenbar durch die verallgemeinerten Impulseigenzustände $|\vec{p}\rangle$ gegeben sind. Der Energieeigenwert ist dann natürlich als Funktion der Impulswerte einfach durch $E(\vec{p}) = \vec{p}^2/(2m)$ gegeben. Die Gleichung (3.6.4) kann man demnach direkt auf die Impulsdarstellung, d.h. die Impulswellenfunktion hinschreiben:

$$\tilde{\psi}(t, \vec{p}) = \langle \vec{p} | \psi(t) \rangle = \int_{\mathbb{R}^3} d^3 p \exp\left(-\frac{iE(\vec{p})}{\hbar}(t-t_0)\right) \tilde{\psi}_0(\vec{p}), \tag{3.6.6}$$

wobei $\tilde{\psi}_0$ die Impulswellenfunktion zur Anfangszeit t_0 ist. Dies stimmt vollständig mit unseren Überlegungen in Abschnitt 3.1 überein, wo wir von der heuristisch postulierten Schrödingergleichung für die Ortswellenfunktion ausgegangen sind.

3.7 Formale Herleitung der Wellenmechanik

Wir können diese „Wellenmechanik“ nun auch nochmals aus unserer jetzigen allgemeineren Perspektive herleiten. Der Einfachheit halber betrachten wir dazu die Bewegung eines Teilchens in einer Raumdimension. Wir gehen von den Kommutatorrelationen für die Operatoren der Ortskoordinate x und der Impulskoordinate p , also

$$[x, p] = i\hbar \mathbb{1}, \tag{3.7.1}$$

aus.

Der Schlüssel für die Formulierung in der Ortsdarstellung ist, wie schon in Abschnitt 3.3, die Herleitung der Impulseigenfunktionen, also die Komponenten der Impulseigenzustände bzgl. der verallgemeinerten Eigenbasis der Ortskoordinate $u_p(x) = \langle x | p \rangle$. Wir wissen bereits, dass (3.3.17) dabei herauskommen muss. Um dies zu beweisen, müssen wir herausfinden, wie der Impulsoperator auf die Ortseigenzustände wirkt. Wir können den dazu benötigten Trick aus diesem Ergebnis gewinnen, denn offenbar gilt

$$\begin{aligned}
 u_p(x) = \langle x | p \rangle &= u_p(0) \exp\left(\frac{ipx}{\hbar}\right) \\
 &= \langle x=0 | p \rangle \exp\left(\frac{ipx}{\hbar}\right) \\
 &= \left\langle x=0 \left| \exp\left(\frac{ipx}{\hbar}\right) \right| p \right\rangle \\
 &= \left\langle x=0 \left| \exp\left(\frac{ipx}{\hbar}\right) \right| p \right\rangle.
 \end{aligned} \tag{3.7.2}$$

Wir müssen also nur zeigen, dass

$$\langle x=0 | \exp\left(\frac{i\mathbf{p}x}{\hbar}\right) = \langle x | \Rightarrow |x\rangle = \exp\left(-\frac{i\mathbf{p}x}{\hbar}\right) |x=0\rangle \quad (3.7.3)$$

ist. Dazu benötigen wir den Kommutator $[\mathbf{x}, \exp(-i\mathbf{p}x/\hbar)]$. Dieser lässt sich in der Tat allein aus der Kommutatorrelation (3.7.1) bestimmen. Dazu denken wir uns die Operatorexponentialfunktion durch ihre Potenzreihe definiert. Dann benötigen wir zur Berechnung des Kommutators nur noch die Kommutatoren $[\mathbf{x}, \mathbf{p}^k]$ für $k \in \mathbb{N}_0$. Dies gelingt aber leicht dadurch, dass wir

$$[\mathbf{x}, \mathbf{p}^k] = i\hbar k \mathbf{p}^{k-1} \quad (3.7.4)$$

durch vollständige Induktion beweisen. Offenbar gilt die Behauptung für $k=0$, denn es ist $[\mathbf{x}, \mathbf{p}^0] = [\mathbf{x}, \mathbb{1}] = 0$. Angenommen, sie gälte auch für $k=n$ dann folgt mit (3.5.29)

$$[\mathbf{x}, \mathbf{p}^{n+1}] = [\mathbf{x}, \mathbf{p}\mathbf{p}^n] = \mathbf{p}[\mathbf{x}, \mathbf{p}^n] + [\mathbf{x}, \mathbf{p}]\mathbf{p}^n = i\hbar n \mathbf{p}^n + i\hbar \mathbf{p}^n = i\hbar(n+1)\mathbf{p}^n, \quad (3.7.5)$$

und das ist die Behauptung für $k=n+1$, d.h. die Formeln (3.7.4) gilt für alle $k \in \mathbb{N}_0$. Damit ist aber gezeigt, dass für eine durch eine Potenzreihe definierte Funktion des Impulsoperators

$$[\mathbf{x}, f(\mathbf{p})] = i\hbar f'(\mathbf{p}) \quad (3.7.6)$$

gilt. Damit erhalten wir aber sofort

$$\left[\mathbf{x}, \exp\left(-\frac{i\mathbf{p}x}{\hbar}\right) \right] = i\hbar \partial_{\mathbf{p}} \exp\left(-\frac{i\mathbf{p}x}{\hbar}\right) = x \exp\left(-\frac{i\mathbf{p}x}{\hbar}\right). \quad (3.7.7)$$

Damit folgt aber endlich

$$\mathbf{x} \exp\left(-\frac{i\mathbf{p}x}{\hbar}\right) |x=0\rangle = \left\{ \left[\mathbf{x}, \exp\left(-\frac{i\mathbf{p}x}{\hbar}\right) \right] + \exp\left(-\frac{i\mathbf{p}x}{\hbar}\right) \mathbf{x} \right\} |x=0\rangle = x \exp\left(-\frac{i\mathbf{p}x}{\hbar}\right) |x=0\rangle, \quad (3.7.8)$$

d.h. aber, dass in der Tat $\exp(-i\mathbf{p}x/\hbar)|x=0\rangle$ ein Eigenvektor des Ortsoperators zum Eigenwert x ist, wie in (3.7.3) behauptet! Folglich ist (3.7.2) erfüllt, und man kann schließlich den Normierungs $u_p(0)$ wie in Abschnitt 3.3 aus der Normierungsbedingung $\langle p | p' \rangle = \delta(p-p')$ bestimmen, was wieder (3.3.7) ergibt.

3.8 Teilchen im Potentialtopf

Als einfachstes Beispiel für Energieeigenprobleme betrachten wir ein Teilchen in einem eindimensionalen Potentialtopf, d.h. der Hamilton-Operator ist durch

$$\mathbf{H} = \frac{1}{2m} \mathbf{p}^2 + V(\mathbf{x}) \quad (3.8.1)$$

gegeben, wobei das Potential in der Ortsdarstellung durch

$$V(x) = \begin{cases} 0 & \text{für } x \in [-a/2, a/2], \\ \infty & \text{für } |x| > a/2. \end{cases} \quad (3.8.2)$$

Das bedeutet, dass das Teilchen sich nie außerhalb des Potentialtopfes befinden kann, und die Ortswellenfunktionen müssen die **Randbedingungen**

$$\psi(a/2) = \psi(-a/2) = 0 \quad (3.8.3)$$

3.8. Teilchen im Potentialtopf

erfüllen. Im Folgenden beschränken wir uns auf das Innere des Potentialtopfes, d.h. nehmen *immer* an, dass $x \in [-a/2, a/2]$ ist.

Wir können nun das Eigenwertproblem für den Hamilton-Operator in der Ortsdarstellung leicht lösen. Dabei bietet sich hier die Ortsdarstellung natürlicherweise an, weil das (etwas künstliche) Potential sich in der Impuls-Darstellung nicht vernünftig definieren lässt. Jedenfalls ist das Eigenwertproblem für den Hamilton-Operator in der Ortsdarstellung durch

$$-\frac{\hbar^2}{2m} u_E''(x) = E u_E(x) \quad (3.8.4)$$

gegeben. Für jedes $E > 0$ gibt es nun offenbar zwei linear unabhängige Lösungen im Hilbert-Raum $L^2([-a/2, a/2])$ d.h. dem Hilbert-Raum der quadratintegriblen Funktionen $\psi(x)$, wobei $x \in [-a/2, a/2]$ ist:

$$u_E^{(1)}(x) = N \cos(k_1 x), \quad u_E^{(2)} = N \sin(k_2 x). \quad (3.8.5)$$

Setzt man dies in (3.8.4) ein, erhält man

$$k_{1,2} = \sqrt{\frac{2mE}{\hbar^2}}. \quad (3.8.6)$$

Nun müssen aber die „Wellenzahlen“ $k_{1,2}$ in (3.8.5) so gewählt werden, dass diese Funktionen auch die Randbedingungen (3.8.3) erfüllen. Für $u_E^{(1)}$ verlangt dies

$$\frac{k_1 a}{2} = \frac{2j+1}{2} \pi \Rightarrow k = \frac{(2j+1)\pi}{a} \quad \text{mit } j \in \mathbb{N}_0 = \{0, 1, 2, \dots\}, \quad (3.8.7)$$

denn genau dort besitzt der \cos seine Nullstellen. Negative Werte von j brauchen wir nicht zu berücksichtigen, da dies wegen $\cos(-\alpha) = \cos(\alpha)$ nur die gleichen Funktionen liefert wie die positiven Werte.

Für $u_E^{(2)}$ liefern die Randbedingungen

$$\frac{k_2 a}{2} = j \pi \Rightarrow k_2 = \frac{2j \pi}{a} \quad \text{mit } j \in \mathbb{N} = \{1, 2, \dots\}. \quad (3.8.8)$$

Negative Werte und $j = 0$ kommen hier nicht in Frage, weil wegen $\sin(-\alpha) = -\sin \alpha$ negative Werte bis auf das Vorzeichen nur wieder die gleichen Eigenfunktionen liefern wie die mit positiven Werten, und für $j = 0$ wird $u_E^{(2)}(x) = 0 = \text{const}$, und definitionsgemäß ist der Nullvektor des Hilbert-Raums keine Eigenfunktion des Hamilton-Operators.

Es ist nun für das Folgende bequemer, die Energie-Eigenfunktionen mit einem Index $n \in \mathbb{N}_0$ durchzunummerieren, denn offenbar durchlaufen dann möglichen Werte für k_1 und k_2 die Werte $k_n = n\pi/a$, und es ist

$$u_n(x) = \begin{cases} N_n \cos(k_n x) & \text{falls } n \in \mathbb{N} \text{ gerade ist,} \\ N_n \sin(k_n x) & \text{falls } n \in \mathbb{N} \text{ ungerade ist.} \end{cases} \quad (3.8.9)$$

Die entsprechenden Energie-Eigenwerte sind dann gemäß (3.8.6)

$$E_n = \frac{\hbar^2}{2m} k_n^2. \quad (3.8.10)$$

Die Normierungskonstanten berechnen sich aus der Forderung, dass die Wellenfunktionen normiert sein sollen, d.h.

$$\langle u_n | u_n \rangle = \|u_n\|^2 = \int_{-a/2}^{a/2} dx |u_n(x)|^2 = 1. \quad (3.8.11)$$

Zur Berechnung dieser Integrale verwenden wir das Doppelwinkeltheorem für den cos:

$$\cos(2\alpha) = \cos^2 \alpha - \sin^2 \alpha = 1 - 2 \sin^2 \alpha = 2 \cos^2 \alpha - 1. \quad (3.8.12)$$

Dies können wir leicht umformen zu

$$\sin^2 \alpha = \frac{1}{2}[1 - \cos(2\alpha)], \quad \cos^2 \alpha = \frac{1}{2}[1 + \cos(2\alpha)]. \quad (3.8.13)$$

Demnach wird das Normierungsintegral für gerade n

$$\int_{-a/2}^{a/2} dx |u_n(x)|^2 = |N_n|^2 \int_{-a/2}^{a/2} dx \cos^2(k_n x) = \frac{1}{2} |N_n|^2 \int_{-a/2}^{a/2} dx [1 + \cos(2k_n x)] = \frac{a}{2} |N_n|^2. \quad (3.8.14)$$

Dabei fällt der Beitrag zum Integral von $\cos(2k_n x)$ wegen $\pm 2k_n a/2 = 2n\pi$ weg. Genauso folgt, dass (3.8.14) auch für ungerade n gilt, d.h. wir haben

$$N_n = \sqrt{\frac{2}{a}}. \quad (3.8.15)$$

Man rechnet leicht nach, dass die u_n für $n \in \mathbb{N}$ ein Orthonormalsystem bilden. Aus der Theorie der Fourier-Reihen [CH10] weiß man, dass dieses Orthonormalsystem auch vollständig ist, d.h. man kann jede über das Intervall $x \in [-a/2, a/2]$ quadratintegrale Funktion $\psi(x)$, die die Randbedingungen (3.8.8) erfüllt, nach den Energie-Eigenfunktionen (3.8.9) entwickeln, d.h. es gilt

$$\psi(x) = \sum_{n=1}^{\infty} \psi_n u_n(x), \quad (3.8.16)$$

wobei die Fourier-Koeffizienten durch

$$\langle u_n | \psi \rangle = \int_{-a/2}^{a/2} dx u_n^*(x) \psi(x) = \int_{-a/2}^{a/2} dx u_n(x) \psi(x) \quad (3.8.17)$$

gegeben sind. Dabei gilt die letzte Gleichung, weil in diesem Fall die Eigenfunktionen reell sind.

3.9 Der harmonische Oszillator

Wir betrachten ein Teilchen, das sich in einem äußeren harmonischen Potential entlang der x -Achse bewegen möge. Der Hamiltonoperator besitzt die Gestalt

$$\mathbf{H} = \frac{\mathbf{p}^2}{2m} + \frac{m\omega^2}{2} \mathbf{x}^2. \quad (3.9.1)$$

Wir stellen uns die Aufgabe, ein vollständiges Orthonormalsystem von Energieeigenvektoren zu finden, und zwar zunächst ohne von der Orts- oder Impulsdarstellung Gebrauch zu machen, also rein algebraisch.

Dazu führt man die folgenden *nicht selbstadjungierten (!)* Operatoren

$$\mathbf{a} = \sqrt{\frac{m\omega}{2\hbar}} \mathbf{x} + \frac{i}{\sqrt{2m\hbar\omega}} \mathbf{p} \quad (3.9.2)$$

ein. Die Struktur des Hamiltonoperators legt die Berechnung des selbstadjungierten Operators

$$\mathbf{a}^\dagger \mathbf{a} = \frac{m\omega}{2\hbar} \mathbf{x}^2 + \frac{1}{2m\omega\hbar} \mathbf{p}^2 - \frac{1}{2} \mathbb{1} \quad (3.9.3)$$

3.9. Der harmonische Oszillator

nahe, wobei wir das Operatorprodukt einfach ausmultipliziert und die Vertauschungsrelationen für Ort und Impuls

$$[\mathbf{x}, \mathbf{p}] = i\hbar \mathbb{1} \quad (3.9.4)$$

verwendet haben. Der Vergleich mit (3.9.1) ergibt

$$\mathbf{H} = \frac{\hbar\omega}{2} (2\mathbf{a}^\dagger \mathbf{a} + \mathbb{1}). \quad (3.9.5)$$

Wir können also das Eigenwertproblem für \mathbf{H} lösen, indem wir das Eigenwertproblem für den *selbstadjungierten* Operator

$$\mathbf{N} = \mathbf{a}^\dagger \mathbf{a} \quad (3.9.6)$$

betrachten.

Die Kommutatorrelationen der Operatoren \mathbf{a} und \mathbf{a}^\dagger berechnet man sofort mit Hilfe der Kommutatorrelationen für Ort und Impuls 3.9.4 (*Nachrechnen!*)

$$[\mathbf{a}, \mathbf{a}] = [\mathbf{a}^\dagger, \mathbf{a}^\dagger] = 0, \quad [\mathbf{a}, \mathbf{a}^\dagger] = \mathbb{1}. \quad (3.9.7)$$

Schließlich sind im Folgenden noch die Kommutatoren

$$[\mathbf{a}, \mathbf{N}] = \mathbf{a}, \quad [\mathbf{a}^\dagger, \mathbf{N}] = -\mathbf{a}^\dagger \quad (3.9.8)$$

nützlich.

Sei nun $|n\rangle$ irgendein Eigenvektor von \mathbf{N} zum Eigenwert n . Dann gilt wegen (3.9.8)

$$\mathbf{N}\mathbf{a}|n\rangle = \{[\mathbf{N}, \mathbf{a}] + \mathbf{a}\mathbf{N}\}|n\rangle = (n-1)\mathbf{a}|n\rangle, \quad (3.9.9)$$

d.h. $\mathbf{a}|n\rangle$ ist also entweder ein Eigenvektor von \mathbf{N} zum Eigenwert $n-1$ oder der Nullvektor.

Weiter ist \mathbf{N} ein „positiv semidefiniter Operator“, d.h. für jeden Vektor $|\psi\rangle$ gilt

$$\langle \psi | \mathbf{N} | \psi \rangle = \langle \mathbf{a}\psi | \mathbf{a}\psi \rangle \geq 0. \quad (3.9.10)$$

Setzen wir hierin insbesondere $|\psi\rangle = |n\rangle$, folgt, daß $n \geq 0$ sein muß. Wenden wir also \mathbf{a}^k auf $|n\rangle$ an, muß für ein $k \in \mathbb{N}$ gelten $n-k=0$, weil es andernfalls Eigenvektoren mit beliebig kleinen Eigenwerten von \mathbf{N} gäbe, was aber eben wegen (3.9.10) nicht möglich ist. Existiert also überhaupt ein simultaner Eigenvektor von \mathbf{N} , dann muß es (wenigstens) einen Eigenvektor $|\Omega\rangle$ mit $n=0$ geben, und für diesen muß

$$\mathbf{a}|\Omega\rangle = |0\rangle \quad (3.9.11)$$

gelten. Weiter gilt aber auch

$$\mathbf{N}\mathbf{a}^\dagger |n\rangle = \{[\mathbf{N}, \mathbf{a}^\dagger] + \mathbf{a}^\dagger \mathbf{N}\}|n\rangle = (n+1)\mathbf{a}^\dagger |n\rangle, \quad (3.9.12)$$

und das beweist, dass $\mathbf{a}^\dagger |n\rangle$ ein Eigenvektor von \mathbf{N} zum Eigenwert $n+1$ sein muß, und also die Eigenwerte \mathbf{N} gerade $n \in \mathbb{N}_0$ sind.

Nehmen wir an, $|\Omega\rangle$ sei auf 1 normiert, ergeben sich gemäß (3.9.12) weiter die auf 1 normierten simultanen Eigenvektoren von \mathbf{N} durch fortgesetzte Anwendung von \mathbf{a}^\dagger auf $|\Omega\rangle$ und nachfolgende Normierung. Zur Bestimmung des Normierungsfaktors nehmen wir nun also an, die $|n\rangle$ seien auf 1 normiert. Dann folgt

$$\langle \mathbf{a}^\dagger n | \mathbf{a}^\dagger n \rangle = \langle n | \mathbf{a}\mathbf{a}^\dagger | n \rangle = \langle n | \mathbf{N} + [\mathbf{a}, \mathbf{a}^\dagger] | n \rangle = (n+1) |n+1\rangle. \quad (3.9.13)$$

Durch Rekursion erhalten wir also schließlich

$$|n\rangle = \sqrt{\frac{1}{n!}} (\mathbf{a}^\dagger)^n |\Omega\rangle, \quad (3.9.14)$$

3. Wellenmechanik

wobei wir die Normierungskonstanten willkürlich reell gewählt haben, da Phasenfaktoren ohnehin irrelevant sind. Daraus folgt

$$|n\rangle = \frac{1}{\sqrt{n}} \mathbf{a}^\dagger |n-1\rangle, \quad (3.9.15)$$

und mit der Kommutatorrelation (3.9.7) finden wir

$$\mathbf{a}|n\rangle = \frac{1}{\sqrt{n}} \mathbf{a} \mathbf{a}^\dagger |n-1\rangle = \frac{1}{\sqrt{n}} (\mathbf{N} + [\mathbf{a}, \mathbf{a}^\dagger]) |n-1\rangle = \frac{1}{\sqrt{n}} (\mathbf{N} + \mathbb{1}) |n-1\rangle = \sqrt{n} |n-1\rangle. \quad (3.9.16)$$

Wir stellen weiter fest, daß damit auch das Eigenwertproblem für \mathbf{H} gelöst ist, denn gemäß (3.9.5) ist $|n\rangle$ ein Eigenvektor von \mathbf{H}

$$\mathbf{H}|n\rangle = E_n |n\rangle \quad (3.9.17)$$

zum Eigenwert

$$E_n = \frac{\hbar\omega}{2}(2n+1), \quad n \in \mathbb{N}. \quad (3.9.18)$$

Es ist klar, daß der Zustand $|\Omega\rangle$ der Grundzustand, d.h. der Zustand der niedrigstmöglichen Energie des Teilchens ist.

Nun müssen wir noch die Existenz und Eindeutigkeit des Grundzustands nachweisen sowie zeigen, dass durch (3.9.14) wieder normierbare Hilbertraumvektoren entstehen. Dies lässt sich am einfachsten in der Ortsdarstellung bewerkstelligen. Betrachten wir also die Wellenfunktionen

$$u_n(x) = \langle x | n \rangle. \quad (3.9.19)$$

Erinnern wir uns, dass die Orts- und Impulsoperatoren in der Ortsdarstellung durch $\mathbf{x}\psi(x) = x\psi(x)$ und $\mathbf{p}\psi(x) = -i\hbar\partial_x\psi(x)$ gegeben sind, folgt für die Grundzustandswellenfunktion

$$\mathbf{a}u_0(x) = \left(\sqrt{\frac{\hbar}{2m\omega}} \frac{\partial}{\partial x} + \sqrt{\frac{m\omega}{2\hbar}} x \right) u_0(x) = 0. \quad (3.9.20)$$

Setzen wir zur Abkürzung

$$a^2 = \frac{\hbar}{m\omega}, \quad (3.9.21)$$

schreibt sich diese Gleichung als

$$\frac{\partial}{\partial x} u_0 = -\frac{x}{a^2} u_0. \quad (3.9.22)$$

Die Lösung ergibt sich sofort *eindeutig* zu

$$u_0(x) = \mathcal{N} \exp\left(-\frac{x^2}{2a^2}\right). \quad (3.9.23)$$

Die Normierung ist (bis auf einen unerheblichen Phasenfaktor) durch

$$\langle 0 | 0 \rangle = \int dx |u_0(x)|^2 = |\mathcal{N}|^2 (\pi a^2)^{1/2} \stackrel{!}{=} 1 \Rightarrow \mathcal{N} = \left(\frac{1}{\pi a^2}\right)^{1/4} \quad (3.9.24)$$

gegeben.

Wir erhalten also schließlich

$$u_0(x) = \left(\frac{1}{\pi a^2}\right)^{1/4} \exp\left(-\frac{x^2}{2a^2}\right). \quad (3.9.25)$$

3.9. Der harmonische Oszillator

Es existiert also genau eine quadratintegrale Wellenfunktion zum Eigenwertproblem von \mathbf{N} mit dem Eigenwert $n = 0$. Mittels (3.9.14) sehen wir, dass dann auch alle übrigen Eigenfunktionen quadratintegrabel sind, denn sie ergeben sich als Produkt aus der Gauß-Funktion (3.9.25) mit einem Polynom, das sich durch fortgesetzte Anwendung des Differentialoperators \mathbf{a}^\dagger bestimmen läßt:

$$u_n(x) = \sqrt{\frac{1}{n!}} (\mathbf{a}^\dagger)^n u_0(x) = \sqrt{\frac{1}{2^n n!}} \left(\frac{x}{a} - a \frac{d}{dx} \right)^n u_0(x). \quad (3.9.26)$$

Führen wir die dimensionslose Variable

$$\xi = \frac{x}{a} \quad (3.9.27)$$

ein, können wir gem. (3.9.25) schreiben

$$u_n(x) = \left(\frac{1}{\pi a^2} \right)^{1/4} \sqrt{\frac{1}{2^n n!}} \left[\left(\xi - \frac{d}{d\xi} \right)^n \exp\left(-\frac{\xi^2}{2}\right) \right]_{\xi=\frac{x}{a}}. \quad (3.9.28)$$

Offensichtlich ist

$$u_n(x) = \left(\frac{1}{\pi a^2} \right)^{1/4} \sqrt{\frac{1}{2^n n!}} \left[H_n(\xi) \exp\left(-\frac{\xi^2}{2}\right) \right]_{\xi=\frac{x}{a}} \quad (3.9.29)$$

Dabei sind

$$H_n(\xi) := \exp\left(\frac{\xi^2}{2}\right) \left(\xi - \frac{d}{d\xi} \right)^n \exp\left(-\frac{\xi^2}{2}\right) \quad (3.9.30)$$

die sog. *Hermite-Polynome* (Charles Hermite, 1822-1901). In Abb. 3.1 sind einige dieser Energieeigenfunktionen dargestellt.

Es läßt sich durch Anwendung von (3.9.30) leicht zeigen, dass sie der Rekursionsformel

$$H_{n+1}(\xi) = 2\xi H_n(\xi) - H_n'(\xi) \quad (3.9.31)$$

genügen, aus der durch vollständige Induktion sofort die alternative Darstellung

$$H_n(\xi) = (-1)^n \exp(\xi^2) \frac{d^n}{d\xi^n} \exp(-\xi^2), \quad (3.9.32)$$

die *Rodrigues-Formel* (Benjamin Olinde Rodrigues, 1795-1851), folgt (*Übungsaufgabe!*). Eine nützliche Darstellung als Integral entsteht, indem man die Gaußfunktion in der folgenden Fourierdarstellung schreibt

$$\exp(-\xi^2) = \frac{1}{\sqrt{\pi}} \int_{\mathbb{R}} du \exp(-u^2 + 2iu\xi) \quad (3.9.33)$$

und in die Rodrigues-Formel einsetzt:

$$H_n(\xi) = \frac{\exp(\xi^2)}{\sqrt{\pi}} \int_{\mathbb{R}} du (-2iu)^n \exp(-u^2 + 2iu\xi). \quad (3.9.34)$$

Aus dem Vorstehenden ergibt sich dann weiter unmittelbar, dass die u_n ein orthonormiertes Funktionensystem bilden, dass also

$$\int_{\mathbb{R}} dx u_{n'}^*(x) u_n(x) = \delta_{n'n} \quad (3.9.35)$$

gilt. Substituieren wir hierin $x = a\xi$, folgt für die Hermitepolynome die Orthogonalitätsbedingung

$$\int_{\mathbb{R}} d\xi \exp(-\xi^2) H_{n'}(\xi) H_n(\xi) = 2^n n! \sqrt{\pi} \delta_{n'n}. \quad (3.9.36)$$

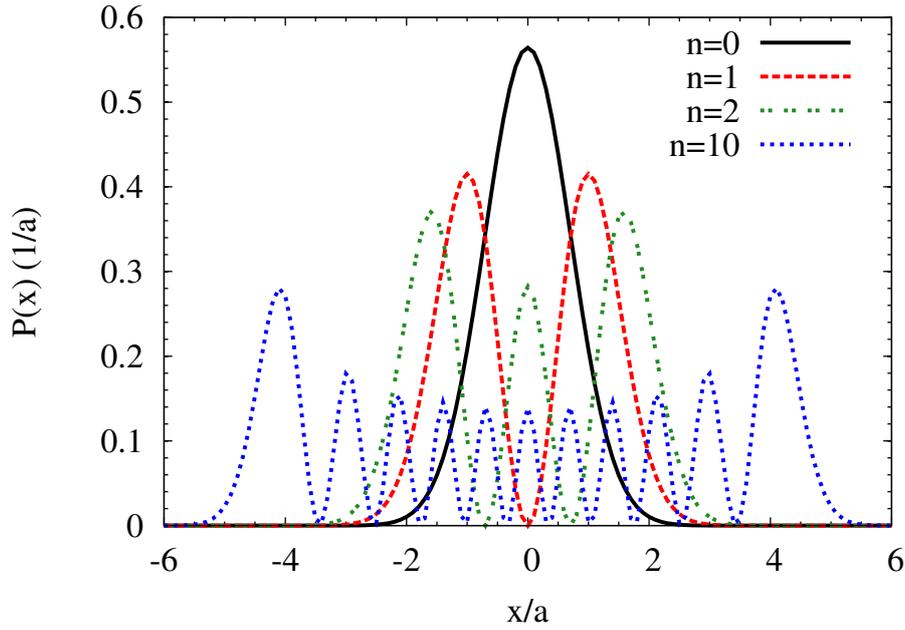


Abbildung 3.1: Die Energieeigenfunktionen des eindimensionalen harmonischen Oszillators zu den Eigenwerten $E_n = \hbar\omega(n + 1/2)$ für $n = 0, 1, 2, 10$. Besonders bemerkenswert ist die Tatsache, dass in dieser Situation, die einem Teilchen mit scharf bestimmter Energie entsprechen, die Aufenthaltswahrscheinlichkeit für beliebig große Ortskomponenten x nicht verschwindet. In der klassischen Situation verbietet der Energiesatz, dass sich das Teilchen außerhalb der durch die Energie vorgegebenen Umkehrpunkte $\pm x_0$ bewegt, die durch $E = m\omega^2 x_0^2/2$ bestimmt sind. Diese Eigenschaft, dass Teilchen mit scharf bestimmter Energie sich in Regionen aufhalten können, in die sie sich gemäß der klassischen Mechanik aufgrund des Energiesatzes nicht gelangen können, bezeichnet man als **Tunneleffekt**.

Andere nützliche Beziehungen ergeben sich aus der „erzeugenden Funktion“ der Hermitepolynome. Sie ist definiert als

$$F(\xi, z) = \sum_{n=0}^{\infty} \frac{1}{n!} H_n(\xi) z^n. \quad (3.9.37)$$

Verwenden wir (3.9.32) und setzen einstweilen $z = -(y - \xi)$, ergibt sich aufgrund der Taylorentwicklung

$$F(\xi, z) = \exp(\xi^2) \sum_{n=0}^{\infty} \frac{1}{n!} \left[\frac{d^n}{d\xi^n} \exp(-\xi^2) \right] (y - \xi)^n = \exp(\xi^2 - y^2) = \exp(2\xi z - z^2). \quad (3.9.38)$$

Es ist also

$$\exp(2\xi z - z^2) = \sum_{n=0}^{\infty} \frac{1}{n!} H_n(\xi) z^n. \quad (3.9.39)$$

Differenzieren wir diese Gleichung nach z , finden wir

$$2(\xi - z) \exp(2\xi z - z^2) = \sum_{n=1}^{\infty} \frac{1}{(n-1)!} H_n(\xi) z^{n-1} = \sum_{n=0}^{\infty} \frac{1}{n!} H_{n+1}(\xi) z^n. \quad (3.9.40)$$

Multiplizieren wir andererseits (3.9.39) mit $2(\xi - z)$ und vergleichen die Koeffizienten von z^n , finden wir die Rekursionsformel

$$H_1(\xi) = 2\xi H_0(\xi), \quad H_{n+1}(\xi) = 2\xi H_n(\xi) - 2n H_{n-1}(\xi), \quad (3.9.41)$$

die sich (zusammen mit der Startbedingung $H_0(\xi) \equiv 1$) gut zur numerischen Berechnung der Hermitepolynome eignet.

Auf dieselbe Weise erhalten wir durch Ableiten von (3.9.39) nach ξ die Ableitung der Hermitepolynome

$$H'_n(\xi) = \begin{cases} 0 & \text{für } n = 0, \\ 2nH_{n-1}(\xi) & \text{für } n \geq 1. \end{cases} \quad (3.9.42)$$

3.10 Der symmetrische 3D harmonische Oszillator

Ohne weitere Rechnungen kann man nun das Verfahren des vorigen Abschnitts für den eindimensionalen harmonischen Oszillator auf den symmetrischen dreidimensionalen harmonischen Oszillator verallgemeinern. Der Hamilton-Operator lautet

$$\mathbf{H} = \frac{1}{2m} \vec{\mathbf{p}}^2 + \frac{m\omega^2}{2} \vec{\mathbf{x}}^2. \quad (3.10.1)$$

Wie im eindimensionalen Fall können wir Leiteroperatoren einführen, und zwar für jede der drei Raumrichtungen

$$\mathbf{a}_j = \sqrt{\frac{m\omega}{2\hbar}} \mathbf{x}_j + \frac{i}{\sqrt{2m\hbar\omega}} \mathbf{p}_j. \quad (3.10.2)$$

Aufgrund der Vertauschungsrelationen für die Operatoren der Orts- und Impulskomponenten

$$[\mathbf{x}_j, \mathbf{x}_k] = [\mathbf{p}_j, \mathbf{p}_k] = 0, \quad [\mathbf{x}_j, \mathbf{p}_k] = i\hbar \delta_{jk} \mathbb{1} \quad (3.10.3)$$

erhält man die Kommutatoren der Leiteroperatoren zu

$$[\mathbf{a}_j, \mathbf{a}_k] = [\mathbf{a}_j^\dagger, \mathbf{a}_k^\dagger] = 0, \quad [\mathbf{a}_j, \mathbf{a}_k^\dagger] = \delta_{jk} \mathbb{1}. \quad (3.10.4)$$

Man hat es aufgrund dieser Vertauschungsrelationen also formal mit drei völlig voneinander unabhängigen Oszillatoren zu tun, d.h. wir können ein vollständiges orthonormales System von simultanen Eigenvektoren $|n_1, n_2, n_3\rangle$ der drei untereinander kommutierenden „Phononenzahloperatoren“

$$\mathbf{N}_j = \mathbf{a}_j^\dagger \mathbf{a}_j \quad (3.10.5)$$

konstruieren, wie für den eindimensionalen Fall im vorigen Abschnitt gezeigt, d.h. es gibt einen eindeutig bestimmten Zustand mit $n_1 = n_2 = n_3 = 0$, also $|0, 0, 0\rangle$, und für jedes Tripel $(n_1, n_2, n_3) \in \mathbb{N}_0^3$ ist

$$|n_1, n_2, n_3\rangle = \frac{1}{\sqrt{n_1! n_2! n_3!}} (\mathbf{a}_1^\dagger)^{n_1} (\mathbf{a}_2^\dagger)^{n_2} (\mathbf{a}_3^\dagger)^{n_3} |0, 0, 0\rangle. \quad (3.10.6)$$

Die Phononvernichtungs- und Erzeugungsoperatoren wirken ebenfalls analog wie im eindimensionalen Fall durch

$$\mathbf{a}_1 |n_1, n_2, n_3\rangle = \sqrt{n_1 - 1} |n_1 - 1, n_2, n_3\rangle, \quad \mathbf{a}_1^\dagger |n_1, n_2, n_3\rangle = \sqrt{n_1} |n_1 + 1, n_2, n_3\rangle \quad (3.10.7)$$

und entsprechend für \mathbf{a}_2 und \mathbf{a}_3 bzw. \mathbf{a}_2^\dagger und \mathbf{a}_3^\dagger .

Der Hamiltonoperator lässt sich als

$$\mathbf{H} = \hbar\omega \left(\mathbf{N}_1 + \mathbf{N}_2 + \mathbf{N}_3 + \frac{3}{2} \right) \quad (3.10.8)$$

schreiben (*Nachrechnen!*). Damit lösen die simultanen Phononenzahleigenvektoren auch das Energieeigenwertproblem, d.h. sie bilden auch ein VONS von Energieeigenzuständen.

3.11 Der Bahndrehimpuls

Wir können nun die Ergebnisse des vorigen Abschnittes zum dreidimensionalen symmetrischen harmonischen Oszillator auch verwenden, um auf elegante Weise eine andere nützliche vollständige Orthonormalbasis einzuführen, und zwar eines, das mit dem **Bahndrehimpuls** zu tun hat. Wie in der klassischen Mechanik ist der Bahndrehimpuls eines Teilchens durch

$$\vec{\mathbf{L}} = \vec{\mathbf{x}} \times \vec{\mathbf{p}} \quad (3.11.1)$$

definiert. Dabei sind nur Produkte von Orts- und Impulsoperatoren in verschiedenen Richtungen involviert, und es gibt demnach kein Anordnungsproblem. Die einzelnen Komponenten können mit Hilfe des Levi-Civita-Symbols (s. [Hee18b]) als

$$\mathbf{L}_j = \epsilon_{jkl} \mathbf{x}_k \mathbf{p}_l \quad (3.11.2)$$

ausgedrückt werden, wobei wir die Einsteinsche Summenkonvention anwenden, d.h. über doppelt auftretende Vektorindizes wird summiert.

Wir können nun mit Hilfe von (3.10.3) die Kommutatorregeln für die Drehimpulskomponenten berechnen:

$$\begin{aligned} [\mathbf{L}_a, \mathbf{L}_b] &= \epsilon_{acd} \epsilon_{bef} [\mathbf{x}_c \mathbf{p}_d, \mathbf{x}_e \mathbf{p}_f] \\ &= \epsilon_{acd} \epsilon_{bef} (\mathbf{x}_c [\mathbf{p}_d, \mathbf{x}_e \mathbf{p}_f] + [\mathbf{x}_c, \mathbf{x}_e \mathbf{p}_f] \mathbf{p}_d) \\ &= \epsilon_{acd} \epsilon_{bef} (\mathbf{x}_c [\mathbf{p}_d, \mathbf{x}_e] \mathbf{p}_f + \mathbf{x}_e [\mathbf{x}_c, \mathbf{p}_f] \mathbf{p}_d) \\ &= i\hbar \epsilon_{acd} \epsilon_{bef} (-\delta_{de} \mathbf{x}_c \mathbf{p}_f + \delta_{cf} \mathbf{x}_e \mathbf{p}_d) \\ &= i\hbar [(\delta_{ab} \delta_{cf} - \delta_{af} \delta_{cb}) \mathbf{x}_c \mathbf{p}_f - (\delta_{ab} \delta_{de} - \delta_{ae} \delta_{db}) \mathbf{x}_e \mathbf{p}_d] \\ &= i\hbar (\mathbf{x}_a \mathbf{p}_b - \mathbf{x}_b \mathbf{p}_a) = \\ &= i\hbar \epsilon_{abc} \mathbf{L}_c. \end{aligned} \quad (3.11.3)$$

Die Drehimpulskomponenten sind also *nicht kommutativ*, und wir können nur Eigenvektoren für eine Komponente definieren, nicht jedoch für zwei oder gar alle drei Komponenten.

Es ist allerdings leicht zu sehen, dass alle Drehimpulskomponenten mit dem Dreimpulsbetragsquadrat kommutieren, denn es ist

$$[\mathbf{L}_j, \vec{\mathbf{L}}^2] = [\mathbf{L}_j, \mathbf{L}_k \mathbf{L}_k] = \mathbf{L}_k [\mathbf{L}_j, \mathbf{L}_k] + [\mathbf{L}_j, \mathbf{L}_k] \mathbf{L}_k = i\hbar \epsilon_{jkl} (\mathbf{L}_k \mathbf{L}_l + \mathbf{L}_l \mathbf{L}_k) = 0. \quad (3.11.4)$$

Wir können also simultane Eigenvektoren einer Drehimpulskomponente, wozu man traditionellerweise \mathbf{L}_3 wählt und von \mathbf{L}^2 definieren.

Dazu ist es nun bequem, statt \mathbf{L}_1 und \mathbf{L}_2 die *nicht selbstadjungierten* Operatoren

$$\mathbf{L}_+ = \mathbf{L}_1 + i\mathbf{L}_2, \quad \mathbf{L}_- = \mathbf{L}_1 - i\mathbf{L}_2 = \mathbf{L}_+^\dagger \quad (3.11.5)$$

einzuführen. Dann folgt

$$[\mathbf{L}_3, \mathbf{L}_+] = [\mathbf{L}_3, \mathbf{L}_1 + i\mathbf{L}_2] = i\hbar \mathbf{L}_2 + \hbar \mathbf{L}_1 = \hbar \mathbf{L}_+. \quad (3.11.6)$$

Bildet man davon das Adjungierte, folgt

$$[\mathbf{L}_+^\dagger, \mathbf{L}_3] = [\mathbf{L}_-, \mathbf{L}_3] = \hbar \mathbf{L}_+^\dagger = \hbar \mathbf{L}_- \Rightarrow [\mathbf{L}_3, \mathbf{L}_-] = -\hbar \mathbf{L}_-. \quad (3.11.7)$$

Weiter folgt

$$\mathbf{L}_- \mathbf{L}_+ = (\mathbf{L}_1 - i\mathbf{L}_2)(\mathbf{L}_1 + i\mathbf{L}_2) = \mathbf{L}_1^2 + \mathbf{L}_2^2 + i[\mathbf{L}_1, \mathbf{L}_2] = \mathbf{L}_1^2 + \mathbf{L}_2^2 - \hbar \mathbf{L}_3 = \vec{\mathbf{L}}^2 - \mathbf{L}_3^2 - \hbar \mathbf{L}_3. \quad (3.11.8)$$

3.11. Der Bahndrehimpuls

Daraus folgt

$$\vec{\mathbf{L}}^2 = \mathbf{L}_- \mathbf{L}_+ + \mathbf{L}_3^2 + \hbar \mathbf{L}_3. \quad (3.11.9)$$

Wir schreiben $\hbar^2 \alpha$ für die Eigenwerte von $\vec{\mathbf{L}}^2$ und $\hbar m$ für die von \mathbf{L}_3 . Damit sind α und m offenbar dimensionslose Zahlen. Im folgenden seien $|\alpha, m\rangle$ die gemeinsamen Eigenvektoren von $\vec{\mathbf{L}}^2$ und \mathbf{L}_3 , d.h.

$$\vec{\mathbf{L}}^2 |\alpha, m\rangle = \hbar^2 \alpha |\alpha, m\rangle, \quad \mathbf{L}_3 |\alpha, m\rangle = \hbar m |\alpha, m\rangle. \quad (3.11.10)$$

Jetzt können wir die Kommutatorregel (3.11.6) anwenden:

$$\mathbf{L}_3 \mathbf{L}_+ |\alpha, m\rangle = ([\mathbf{L}_3, \mathbf{L}_+] + \mathbf{L}_+ \mathbf{L}_3) |\alpha, m\rangle = (\hbar \mathbf{L}_+ + \hbar m \mathbf{L}_+) |\alpha, m\rangle = \hbar(m+1) \mathbf{L}_+ |\alpha, m\rangle. \quad (3.11.11)$$

Damit ist $\mathbf{L}_+ |\alpha, m\rangle$ entweder Eigenvektor von \mathbf{L}_3 zum Eigenwert $\hbar(m+1)$ oder der Nullvektor. Ebenso folgt aus (3.11.7)

$$\mathbf{L}_3 \mathbf{L}_- |\alpha, m\rangle = \hbar(m-1) \mathbf{L}_- |\alpha, m\rangle, \quad (3.11.12)$$

d.h. $\mathbf{L}_- |\alpha, m\rangle$ ist entweder Eigenvektor von \mathbf{L}_3 zum Eigenwert $\hbar(m-1)$ oder der Nullvektor. Die Operatoren \mathbf{L}_\pm wirken also bzgl. der \mathbf{L}_3 -Eigenzustände ähnlich wie ein Leiteroperator beim harmonischen Oszillator, d.h. er führt zu einem Eigenvektor zu einem um 1 höheren bzw. niedrigeren Wert für m .

Wegen (3.11.4) kommutiert nun $\vec{\mathbf{L}}^2$ mit \mathbf{L}_\pm und folglich gilt

$$\vec{\mathbf{L}}^2 \mathbf{L}_\pm |\alpha, m\rangle = \mathbf{L}_\pm \vec{\mathbf{L}}^2 |\alpha, m\rangle = \hbar^2 \alpha \mathbf{L}_\pm |\alpha, m\rangle. \quad (3.11.13)$$

Das bedeutet, dass die Vektoren $\mathbf{L}_\pm |\alpha, m\rangle$ Eigenvektoren von $\vec{\mathbf{L}}^2$ zum Eigenwert $\hbar^2 \alpha$ sind, falls sie nicht Nullvektoren sind.

Nun sind sowohl $\vec{\mathbf{L}}^2$ als auch $\vec{\mathbf{L}}^2 - \mathbf{L}_3^2 = \mathbf{L}_1^2 + \mathbf{L}_2^2$ positiv semidefinite Operatoren, d.h. für jeden Vektor $|\psi\rangle$ gilt $\langle \psi | \vec{\mathbf{L}}^2 | \psi \rangle \geq 0$. Dies folgt sofort aus der Selbstadjungiertheit der Drehimpulsoperatoren. Z.B. ist

$$\langle \psi | \mathbf{L}_j^2 | \psi \rangle = \langle \mathbf{L}_j \psi | \mathbf{L}_j \psi \rangle \geq 0, \quad (3.11.14)$$

weil das Skalarprodukt des Hilbert-Raums definitionsgemäß positiv definit ist. Damit ist

$$\langle \alpha, m | \vec{\mathbf{L}}^2 | \alpha, m \rangle = \alpha \hbar^2 \langle \alpha, m | \alpha, m \rangle \geq 0 \Rightarrow \alpha \hbar^2 \geq 0 \quad (3.11.15)$$

und

$$\langle \alpha, m | \vec{\mathbf{L}}^2 - \mathbf{L}_3^2 | \alpha, m \rangle = \hbar^2(\alpha - m^2) \langle \alpha, m | \alpha, m \rangle \geq 0 \Rightarrow m^2 \leq \alpha. \quad (3.11.16)$$

Wie wir oben gesehen haben, liefert $\mathbf{L}_+ |\alpha, m\rangle$ einen Eigenvektor von \mathbf{L}_3 zum Eigenwert $\hbar(m+1)$ und zu $\vec{\mathbf{L}}^2$ zum Eigenwert $\alpha \hbar^2$ oder den Nullvektor. Offenbar muss es demnach für ein vorgegebenes α einen maximalen Eigenwert l von \mathbf{L}_3 geben, weil ansonsten irgendwann (3.11.16) verletzt würde. Damit muss

$$\mathbf{L}_+ |\alpha, l\rangle = |0\rangle \quad (3.11.17)$$

sein, denn sonst hätte man ja einen Eigenvektor von \mathbf{L}_3 zum Eigenwert $\hbar(l+1) > \sqrt{\alpha}$, was (3.11.16) widerspräche.

Wir können nun ausgehend von dem Eigenvektor $|\alpha, l\rangle$ durch wiederholtes Anwenden von \mathbf{L}_- Eigenvektoren von \mathbf{L}_3 zu den Eigenwerten $\hbar m$ mit $m = l - n$, wobei $n \in \mathbb{N}_0$ ist, finden. Allerdings kann dies auch nicht

3. Wellenmechanik

beliebig oft funktionieren, weil man sonst irgendwann $\hbar^2(l-n)^2 > \alpha \hbar^2$ im Widerspruch zu (3.11.16) erhalte. Es muss also zu vorgegebenem α einen minimalen Eigenwert $m_{\min} = l - n$ geben, so dass

$$\mathbf{L}_- |\alpha, m_{\min}\rangle = |0\rangle. \quad (3.11.18)$$

Jetzt wenden wir (3.11.9) an. Demnach ist wegen (3.11.17)

$$\begin{aligned} \vec{\mathbf{L}}^2 |\alpha, l\rangle &= \alpha \hbar^2 |\alpha, l\rangle \\ &= (\mathbf{L}_- \mathbf{L}_+ + \mathbf{L}_3^2 + \hbar \mathbf{L}_3) |\alpha, l\rangle = (\mathbf{L}_3^2 + \hbar \mathbf{L}_3) |\alpha, l\rangle \\ &= \hbar^2 l(l+1) |\alpha, l\rangle. \end{aligned} \quad (3.11.19)$$

Damit folgt

$$\alpha = l(l+1). \quad (3.11.20)$$

Nach unserer obigen Überlegung gibt es nun aber eine Zahl $n \in \mathbb{N}_0$ mit $m_{\min} = l - n$.

Nun weist man aber wie bei der Rechnung zu (3.11.9) nach, dass

$$\mathbf{L}_+ \mathbf{L}_- = \mathbf{L}^2 - \mathbf{L}_3^2 + \hbar \mathbf{L}_3 \quad (3.11.21)$$

ist. Dann folgt aus $\mathbf{L}_- |\alpha, m_{\min}\rangle = |0\rangle$ genau wie in der Rechnung (3.11.19)

$$\alpha = l(l+1) = m_{\min}(m_{\min} - 1) = (l-n)(l-n-1) \Rightarrow l = \frac{n}{2}. \quad (3.11.22)$$

Damit durchläuft m für vorgegebenes l die Werte $m \in \{-l, -l+1, \dots, l-1, l\}$, und es ist $l \in \{0, 1/2, 1, \dots\}$. Aus den Drehimpulsvertauschungsregeln allein folgt also, dass es sowohl ganz- als auch halbzahlige Drehimpulszahlen l und m gibt. Allerdings haben wir hier noch nicht dem speziellen Aufbau des **Bahndrehimpulsoperators** $\vec{\mathbf{L}} = \vec{\mathbf{x}} \times \vec{\mathbf{p}}$ aus den Orts- und Impulskomponenten Rechnung getragen. Es zeigt sich nämlich, dass für den Bahndrehimpuls nur **ganzzahlige Drehimpulsquantenzahlen** l und m auftreten können.

Um das zu zeigen, betrachten wir zunächst die Leiteroperatoren für einen zweidimensionalen harmonischen Oszillator

$$\mathbf{a}_j = \sqrt{\frac{C}{2\hbar}} \mathbf{x}_j + \frac{i}{\sqrt{2C\hbar}} \mathbf{p}_j, \quad C = m_0 \omega. \quad (3.11.23)$$

Dabei haben wir jetzt m_0 für die Masse des Teilchens im harmonischen Oszillator geschrieben, um Verwechslungen mit der Drehimpulsquantenzahl m zu vermeiden. Jedenfalls wissen wir aus dem vorigen Kapitel, dass $\mathbf{n}_j = \mathbf{a}_j^\dagger \mathbf{a}_j$ Operatoren mit Eigenwerten $n_j \in \mathbb{N}_0$ sind und die simultanen Eigenvektoren $|n_1, n_2\rangle$ dieser Phononenzahloperatoren ein vollständiges Orthonormalsystem des Hilbertraums für ein Teilchen, das sich in zwei Raumdimensionen bewegen kann, ergeben. Nun können wir aber noch einen weiteren Leiteroperator vermöge

$$\mathbf{b} = \frac{1}{\sqrt{2}} (\mathbf{a}_1 + i\mathbf{a}_2) \quad (3.11.24)$$

definieren. Dass dieser Operator die Kommutatoreigenschaft eines Leiteroperators besitzt, also $[\mathbf{b}, \mathbf{b}^\dagger] = \mathbb{1}$ erfüllt, rechnet man leicht nach (*Übung!*), indem man die entsprechende Eigenschaft $[\mathbf{a}_j, \mathbf{a}_k^\dagger] = \delta_{jk} \mathbb{1}$ für die ursprünglichen Leiteroperatoren verwendet. Damit sind aber auch die Eigenwerte von $\mathbf{n}_b = \mathbf{b}^\dagger \mathbf{b}$ die ganzen nichtnegativen Zahlen, und wir können statt der Eigenvektoren $|n_1, n_2\rangle$ auch die simultanen Eigenvektoren von $n = n_1 + n_2$ und n_b verwenden, denn man rechnet wieder leicht nach, dass in der Tat $[\mathbf{n}, \mathbf{n}_b] = [\mathbf{n}_1 + \mathbf{n}_2, \mathbf{n}_b] = 0$ ist, und folglich \mathbf{n} und \mathbf{n}_b gemeinsame Eigenvektoren besitzen. Eine einfache, wenn auch etwas längere Rechnung (*Übung!*) zeigt nun, dass

$$\mathbf{L}_3 = \mathbf{x}_1 \mathbf{p}_2 - \mathbf{x}_2 \mathbf{p}_1 = \hbar (\mathbf{n} - 2\mathbf{n}_b) \quad (3.11.25)$$

3.11. Der Bahndrehimpuls

ist. Offenbar durchlaufen nun die Eigenwerte von $\mathbf{n} = \mathbf{n}_1 + \mathbf{n}_2$ die Werte $n = n_1 + n_2 \in \mathbb{N}_0$. Da nun auch $n_b \in \mathbb{N}_0$ ist, folgt aus (3.11.25), dass die Drehimpulsquantenzahl $m = n - 2n_b \in \mathbb{Z}$ ist, d.h. es treten nur ganzzahlige m -Drehimpulsquantenzahlen auf. Folglich muss aber auch $l \in \mathbb{N}_0$ sein, denn für halbzahlige l sind sämtliche möglichen m -Quantenzahlen ebenfalls halbzahlige.

Zum Schluss betrachten wir noch die Wirkung der Operatoren \mathbf{L}_\pm auf die Drehimpulseigenzustände, die wir nach den obigen Überlegungen wie folgt konstruieren können. Wir beginnen für ein beliebiges $l \in \mathbb{N}_0$ mit einem normierten Eigenvektor $|l, l\rangle$, $\langle l, l | l, l\rangle$ und definieren alle übrigen Eigenvektoren zu gegebenem l über den Absteigeoperator \mathbf{L}_- . Dazu müssen wir uns nur die Normierungsfaktoren überlegen. Angenommen, wir hätten für ein $m \in \{-l, -l+1, \dots, l-1, l\}$ bereits den normierten Eigenvektor $|l, m\rangle$ konstruiert. Dann definieren wir

$$\mathbf{L}_- |l, m\rangle = N |l, m-1\rangle, \quad (3.11.26)$$

wobei wir den Normierungsfaktor $N \geq 0$ annehmen dürfen⁷. Da definitionsgemäß alle $|l, m\rangle$ normiert sein sollen, gilt aber wegen (3.11.9)

$$\begin{aligned} N^2 &= \langle \mathbf{L}_- l, m | \mathbf{L}_- l, m \rangle \\ &= \langle l, m | \mathbf{L}_+ \mathbf{L}_- | l, m \rangle \\ &= \langle l, m | \tilde{\mathbf{L}}^2 - \mathbf{L}_3^2 + \hbar \mathbf{L}_3 | l, m \rangle \\ &= \hbar^2 [l(l+1) - m(m-1)] \\ &= \hbar^2 (l+m)(l-m+1) \Rightarrow N = \hbar \sqrt{(l+m)(l-m+1)}. \end{aligned} \quad (3.11.27)$$

Es ist also

$$\mathbf{L}_- |l, m\rangle = \hbar \sqrt{(l+m)(l-m+1)} |l, m-1\rangle. \quad (3.11.28)$$

Wir sehen, dass auf diese Weise für $m = -l$ tatsächlich der Nullvektor herauskommt, wie es sein muss.

Genauso weist man nach, dass

$$\mathbf{L}_+ |l, m\rangle = \hbar \exp(i\varphi_{lm}) \sqrt{(l-m)(l+m+1)} |l, m+1\rangle \quad (3.11.29)$$

gilt. Dabei ist zunächst die Phase $\exp(i\varphi_{lm})$ noch unbestimmt, denn wir haben uns mit der Wahl der relativen Phasen zwischen den Eigenvektoren mit (3.11.28) bereits festgelegt. Mit dieser Gleichung folgt nun aber eindeutig auch die Phase. Dazu wenden (3.11.28) und (3.11.21) an, indem wir (3.11.29) von links mit \mathbf{L}_- multiplizieren:

$$\begin{aligned} \mathbf{L}_- \mathbf{L}_+ |l, m\rangle &= \hbar^2 [l(l+1) - m(m+1)] |l, m\rangle \\ &= \hbar^2 \exp(i\varphi_{lm}) \sqrt{(l-m)(l+m+1)(l+m+1)(l-m)} |l, m\rangle \\ &= \hbar^2 \exp(i\varphi_{lm}) [l(l+1) - m(m+1)] |l, m\rangle \\ \Rightarrow \exp(i\varphi_{lm}) &= 1. \end{aligned} \quad (3.11.30)$$

Folglich ist also einfach

$$\mathbf{L}_+ |l, m\rangle = \hbar \sqrt{(l-m)(l+m+1)} |l, m+1\rangle. \quad (3.11.31)$$

Auch hier ergibt sich korrekterweise, dass $\mathbf{L}_+ |l, l\rangle = |0\rangle$ ist.

⁷Offenbar kommt für $m = -l$ nach der obigen Überlegung in (3.11.27) der Nullvektor heraus!

3.12 Bahndrehimpulseigenzustände in der Ortsdarstellung

Im Folgenden benötigt man oft die simultanen Eigenzustände von \vec{L}^2 und \vec{L}_3 in der **Ortsdarstellung**. Wir bemerken dazu, dass als dritter unabhängiger zu diesen beiden Observablen kompatibler Operator der Operator $\mathbf{r} = \sqrt{\vec{x}^2}$ gewählt werden kann, denn \vec{L} vertauscht mit \vec{x}^2 . Das beweist man leicht durch direkte Rechnung. Berechnen wir dazu zunächst

$$[\mathbf{L}_j, \mathbf{x}_k] = \epsilon_{jab} [\mathbf{x}_a \mathbf{p}_b, \mathbf{x}_k] = \epsilon_{jab} (\mathbf{x}_a [\mathbf{p}_b, \mathbf{x}_k] + [\mathbf{x}_a, \mathbf{x}_k] \mathbf{p}_b) = -i\hbar \epsilon_{jak} \mathbf{x}_a. \quad (3.12.1)$$

Dann folgt

$$[\mathbf{L}_j, \mathbf{x}_k \mathbf{x}_k] = \mathbf{x}_k [\mathbf{L}_j, \mathbf{x}_k] + [\mathbf{L}_j, \mathbf{x}_k] \mathbf{x}_k = -i\hbar \epsilon_{jak} (\mathbf{x}_k \mathbf{x}_a + \mathbf{x}_a \mathbf{x}_k) = -2i\hbar \epsilon_{jak} \mathbf{x}_a \mathbf{x}_k = 0. \quad (3.12.2)$$

denn ϵ_{jak} ist antisymmetrisch bzgl. Vertauschen des Indexpaares (a, k) . Da die Ortskomponentenoperatoren vertauschen ist hingegen $\mathbf{x}_a \mathbf{x}_k$ symmetrisch und folglich der ganze Ausdruck 0.

Wir können also simultane Eigenvektoren von \mathbf{r} , \vec{L}^2 und \vec{L}_3 finden. Offenbar durchlaufen die Eigenwerte des Operators \mathbf{r} alle Werte $r \in \mathbb{R}_{\geq 0}$, denn \vec{x}^2 durchläuft diese Werte, weil die Operatoren \vec{x} simultane Eigenzustände mit Eigenwerten $\vec{x} \in \mathbb{R}^3$ besitzen.

Offenbar bietet es sich nun an, die Wellenfunktionen in der Ortsdarstellung in Kugelkoordinaten (r, ϑ, φ) darzustellen. Wir schreiben entsprechend $|r, \vartheta, \varphi\rangle$ für den Ortseigenzustand $|\vec{x}\rangle = |x_1, x_2, x_3\rangle$ und

$$u_{r'l m}(\vec{x}) = \langle r, \vartheta, \varphi | r', l, m \rangle, \quad (3.12.3)$$

wobei

$$\mathbf{r} |r', l, m\rangle = r' |r', l, m\rangle, \quad \vec{L}^2 |r', l, m\rangle = \hbar^2 l(l+1) |r', l, m\rangle, \quad \mathbf{L}_3 |r', l, m\rangle = \hbar m |r', l, m\rangle. \quad (3.12.4)$$

Dabei haben wir das Resultat aus dem vorigen Abschnitt verwendet, dass \vec{L}^2 die Eigenwerte $\hbar^2 l(l+1)$ mit $l \in \mathbb{N}_0 = \{0, 1, 2, \dots\}$ besitzt und für jedes l der Operator \mathbf{L}_3 die Eigenwerte $\hbar m$ mit $m \in \{-l, -l+1, \dots, l-1, l\}$ besitzt.

Jetzt müssen wir die Drehimpulsoperatoren in der Ortsdarstellung in Kugelkoordinaten ausdrücken. Dazu erinnern wir uns, dass (vgl. [Hee18b]) der Gradient in Kugelkoordinaten die Darstellung

$$\vec{\nabla} = \vec{e}_r \partial_r + \frac{1}{r} \vec{e}_\vartheta \partial_\vartheta + \frac{1}{r \sin \vartheta} \vec{e}_\varphi \partial_\varphi \quad (3.12.5)$$

besitzt. Wegen $\vec{x} = r \vec{e}_r$ folgt

$$\vec{L} = \vec{x} \times \vec{p} = -i\hbar r \vec{e}_r \times \vec{\nabla} = -i\hbar \left(\vec{e}_r \times \vec{e}_\vartheta \partial_\vartheta + \frac{1}{\sin \vartheta} \vec{e}_r \times \vec{e}_\varphi \partial_\varphi \right) = -i\hbar \left(\vec{e}_\varphi \partial_\vartheta - \frac{1}{\sin \vartheta} \vec{e}_\vartheta \partial_\varphi \right). \quad (3.12.6)$$

Dabei haben wir verwendet, dass $(\vec{e}_r, \vec{e}_\vartheta, \vec{e}_\varphi)$ ein rechtshändiges kartesisches Koordinatensystem bilden. In kartesischen Koordinaten sind diese Vektoren

$$\vec{e}_r = \begin{pmatrix} \sin \vartheta \cos \varphi \\ \sin \vartheta \sin \varphi \\ \cos \vartheta \end{pmatrix}, \quad \vec{e}_\vartheta = \begin{pmatrix} \cos \vartheta \cos \varphi \\ \cos \vartheta \sin \varphi \\ -\sin \vartheta \end{pmatrix}, \quad \vec{e}_\varphi = \begin{pmatrix} -\sin \varphi \\ \cos \varphi \\ 0 \end{pmatrix}. \quad (3.12.7)$$

Damit erhalten wir für die kartesischen Komponenten des Drehimpulsoperators (*Nachrechnen!*)

$$\begin{aligned} \mathbf{L}_1 &= \vec{e}_1 \cdot \vec{L} = +i\hbar (\sin \varphi \partial_\vartheta + \cot \vartheta \cos \varphi \partial_\varphi), \\ \mathbf{L}_2 &= \vec{e}_2 \cdot \vec{L} = +i\hbar (-\cos \varphi \partial_\vartheta + \cot \vartheta \sin \varphi \partial_\varphi), \\ \mathbf{L}_3 &= \vec{e}_3 \cdot \vec{L} = -i\hbar \partial_\varphi, \end{aligned} \quad (3.12.8)$$

$$\vec{L}^2 \psi = -\hbar^2 \left[\frac{1}{\sin \vartheta} \frac{\partial}{\partial \vartheta} \left(\sin \vartheta \frac{\partial \psi}{\partial \vartheta} \right) + \frac{1}{\sin^2 \vartheta} \frac{\partial^2 \psi}{\partial \varphi^2} \right].$$

3.12. Bahndrehimpulseigenzustände in der Ortsdarstellung

Damit können wir bereits die Eigenwertgleichung für L_3 hinschreiben und lösen. Es gilt

$$L_3 u_{r',l,m}(\vec{x}) = -i\hbar \partial_\varphi u_{r',l,m}(\vec{x}) = m\hbar u_{r',l,m}(\vec{x}) \Rightarrow u_{r',l,m}(\vec{x}) = v_{r',l}(r, \vartheta) \exp(im\varphi). \quad (3.12.9)$$

Für die Auf- und Absteigeoperatoren für die Magnetquantenzahl m folgt aus (3.12.9)

$$L_\pm = (L_1 \pm iL_2) = \hbar \exp(\pm i\varphi) (i \cot \vartheta \partial_\varphi \pm \partial_\vartheta). \quad (3.12.10)$$

Damit können wir unter Verwendung von (3.11.17) bereits $u_{r',l,l}$ aus

$$L_+ u_{r',l,l} = 0 \quad (3.12.11)$$

bestimmen. Setzen wir hierin (3.12.9) ein, folgt nach einigen einfachen Umformungen (*Nachrechnen!*)

$$\partial_\vartheta v_{r',l,l}(r, \vartheta) = l \cot \vartheta v_{r',l,l}(r, \vartheta). \quad (3.12.12)$$

Die Lösung dieser Differentialgleichung lautet

$$v_{r',l,l}(r, \vartheta) = w_{r',l,l}(r) \sin^l \vartheta \Rightarrow u_{r',l,l}(r, \vartheta, \varphi) = w_{r',l,l}(r) \sin^l \vartheta \exp(il\varphi). \quad (3.12.13)$$

Da nun für gegebenes l alle Eigenfunktionen von L_3 durch den Absteigeoperator L_- gemäß (3.11.28) berechnet werden können, ist die allgemeine Struktur der $u_{r',l,m}$ durch

$$u_{r',l,m}(r, \vartheta) = w_{r',l,m}(r) Y_{lm}(\vartheta, \varphi) \quad (3.12.14)$$

gegeben, denn L_- enthält nur Ableitungen nach den Winkel ϑ und φ und hängt nicht von r oder ∂_r ab. Die $Y_{lm}(\vartheta, \varphi)$ sind offenbar Funktionen von Punkten auf der Einheitskugelfläche $r = 1$. Sie heißen deshalb **Kugelflächenfunktionen**.

Der verbliebene Faktor $w_{r',l,m}$ ist trivial, denn offenbar muss wegen

$$r u_{r',l,m}(r, \vartheta, \varphi) = r u_{r',l,m}(r, \vartheta, \varphi) = r' u_{r',l,m}(r, \vartheta, \varphi) \quad (3.12.15)$$

und der Orthonormalitätsbedingung

$$\langle r'_1, l_1, m_1 | r'_2, l_2, m_2 \rangle = \delta(r'_1 - r'_2) \delta_{l_1 l_2} \delta_{m_1 m_2} \quad (3.12.16)$$

für die Wellenfunktion

$$u_{r',l,m}(r, \vartheta, \varphi) = N'_{r',l,m} \delta(r - r') Y_{lm}(\vartheta, \varphi) \quad (3.12.17)$$

folgen. In der Ortsdarstellung schreibt sich (3.12.16) nämlich

$$\begin{aligned} \langle r'_1, l_1, m_1 | r'_2, l_2, m_2 \rangle &= \int_{\mathbb{R}^3} d^3x u_1^*(\vec{x}) u_2(\vec{x}) \\ &= \int_0^\infty dr \int_0^\pi d\vartheta \int_0^{2\pi} d\varphi r^2 \sin \vartheta |N'_1|^2 \delta(r - r_1) \delta(r - r_2) \\ &\quad Y_{l_1, m_1}^*(\vartheta, \varphi) Y_{l_2, m_2}(\vartheta, \varphi) \\ &= r_1^2 |N'_1|^2 \delta(r_1 - r_2) \int_0^\vartheta d\vartheta \int_0^{2\pi} d\varphi \sin \vartheta Y_{l_1, m_1}^*(\vartheta, \varphi) Y_{l_2, m_2}(\vartheta, \varphi). \end{aligned} \quad (3.12.18)$$

Dabei haben wir zur Abkürzung $u_1 = u_{r'_1, l_1, m_1}$ usw. geschrieben. Offenbar ist dann

$$N'_{r',l,m} = \frac{1}{r'}, \quad (3.12.19)$$

3. Wellenmechanik

und die Normierungsbedingung für die Kugelflächenfunktionen lautet

$$\int_0^\pi d\vartheta \int_0^{2\pi} d\varphi \sin \vartheta Y_{l_1, m_1}^*(\vartheta, \varphi) Y_{l_2, m_2}(\vartheta, \varphi) = \delta_{l_1 l_2} \delta_{m_1 m_2}. \quad (3.12.20)$$

Wir kennen aus (3.12.13) bereits

$$Y_{ll}(\vartheta, \varphi) = N_l \sin^l \vartheta \exp(il\varphi). \quad (3.12.21)$$

Wir können die übrigen Kugelflächenfunktionen mit (3.11.28) bestimmen, indem wir den Absteigeoperator für m, \mathbf{L}_- , in der Ortsdarstellung wiederholt auf (3.12.21) anwenden.

Zunächst bestimmen wir aber die Normierungskonstante N_l . Wegen (3.12.20) mit $l_1 = l_2 = l$ und $m_1 = m_2 = m$ gilt nämlich

$$|N_l|^2 \int_0^\pi d\vartheta \int_0^{2\pi} d\varphi \sin \vartheta \sin^{2l}(\vartheta) = 2\pi N_l^2 \int_{-1}^1 du (1-u^2)^l = 2\pi |N_l|^2 J_l \stackrel{!}{=} 1. \quad (3.12.22)$$

Um das Integral

$$J_l = \int_{-1}^1 du (1-u^2)^l \quad (3.12.23)$$

zu berechnen verwenden wir die Formel für die partielle Integration

$$\int_{-1}^1 du f'(u)g(u) = f(u)g(u)|_{u=-1}^{u=1} - \int_{-1}^1 du f(u)g'(u) \quad (3.12.24)$$

mit $f(u) = (1-u^2)^l$ und $g'(u) = 1$. Dann ist $f'(u) = -2lu(1-u^2)^{l-1}$ und $g(u) = u$. Damit wird

$$J_l = \int_{-1}^1 du 2lu^2(1-u^2)^{l-1} = \int_{-1}^1 du 2l[1-(1-u^2)](1-u^2)^{l-1} = 2lJ_{l-1} - 2lJ_l \Rightarrow J_l = \frac{2l}{2l+1}J_{l-1}. \quad (3.12.25)$$

Mit $J_0 = 2$ folgt

$$J_1 = 2 \cdot \frac{2}{3}, \quad J_2 = 2 \cdot \frac{2}{3} \cdot \frac{4}{5}, \quad \dots, \quad J_l = 2 \frac{2 \cdot 4 \cdots (2l)}{3 \cdot 5 \cdots (2l+1)}. \quad (3.12.26)$$

Nun ist aber

$$2 \cdot 4 \cdots (2l) = 2^l l!, \quad 3 \cdot 5 \cdots (2l+1) = \frac{(2l+1)!}{2 \cdot 4 \cdots (2l)} = \frac{(2l+1)!}{2^l l!} \quad (3.12.27)$$

und damit schließlich

$$J_l = 2 \frac{(2^l l!)^2}{(2l+1)!}. \quad (3.12.28)$$

Dies in (3.12.22) liefert schließlich

$$4\pi |N_l|^2 \frac{(2^l l!)^2}{(2l+1)!} = 1. \quad (3.12.29)$$

Die Normierungskonstante ist bis auf eine Phase eindeutig bestimmt. Die in der Literatur am meisten verbreitete Konvention ist

$$N_l = \sqrt{\frac{(2l+1)!}{4\pi} \frac{(-1)^l}{2^l l!}}. \quad (3.12.30)$$

Damit ist schließlich gemäß (3.12.21)

$$Y_{ll}(\vartheta, \varphi) = \sqrt{\frac{(2l+1)!}{4\pi} \frac{(-1)^l}{2^l l!}} \sin^l \vartheta \exp(il\varphi). \quad (3.12.31)$$

3.12. Bahndrehimpulseigenzustände in der Ortsdarstellung

Daraus erhalten wir nun mit Hilfe der Rekursionsformel (3.11.28), ausgeschrieben in der Ortsdarstellung, alle Y_{lm} . Zur Berechnung der Ableitungen nach ϑ im Leiteroperator (3.12.10) ist es nun bequemer $u = \cos \vartheta$ statt ϑ als unabhängige Variable einzuführen. Es gilt

$$\partial_{\vartheta} = \frac{\partial u}{\partial \vartheta} \partial_u = -\sin \vartheta \partial_u = -\sqrt{1-u^2} \partial_u. \quad (3.12.32)$$

Dabei haben wir verwendet, dass für Kugelkoordinaten $\vartheta \in (0, \pi)$ ist und daher $\sin \vartheta > 0$ ist und folglich $\sin \vartheta = +\sqrt{1-\cos^2 \vartheta} = \sqrt{1-u^2}$ gilt. Damit wird gemäß (3.12.10)

$$\mathbf{L}_{\pm} = \hbar \exp(\pm i\varphi) \left(i \frac{u}{\sqrt{1-u^2}} \partial_{\varphi} \mp \sqrt{1-u^2} \partial_u \right). \quad (3.12.33)$$

Wir schreiben nun

$$Y_{lm}(\vartheta, \varphi) = U_{lm}(u) \exp(im\varphi), \quad u = \cos \vartheta. \quad (3.12.34)$$

Daraus ergibt sich in der Ortsdarstellung (*Nachrechnen*)

$$\begin{aligned} \mathbf{L}_- Y_{lm} &= \hbar \exp[i(m-1)\varphi] \left[-m \frac{u}{\sqrt{1-u^2}} U_{lm}(u) + \sqrt{1-u^2} U'_{lm}(u) \right] \\ &= \hbar \exp[i(m-1)\varphi] \frac{1}{\sqrt{1-u^2}^{m-1}} \frac{d}{du} \left[\sqrt{1-u^2}^m U_{lm}(u) \right]. \end{aligned} \quad (3.12.35)$$

Andererseits ist wegen (3.12.10)

$$\mathbf{L}_- Y_{lm} = \hbar \sqrt{(l+m)(l-m+1)} Y_{l,m-1}. \quad (3.12.36)$$

Zum Einen zeigt (3.12.35), dass tatsächlich der richtige Faktor $\exp[i(m-1)\varphi]$ resultiert und zum Anderen folgt

$$\sqrt{(l+m)(l-m+1)} U_{l,(m-1)}(u) = \frac{1}{\sqrt{1-u^2}^{m-1}} \frac{d}{du} \left[\sqrt{1-u^2}^m U_{l,m}(u) \right]. \quad (3.12.37)$$

Beginnend mit

$$U_{ll}(u) = \sqrt{\frac{(2l+1)!}{4\pi} \frac{(-1)^l}{2^l l!}} \sqrt{1-u^2}^l, \quad (3.12.38)$$

was aus (3.12.31) und der Definition (3.12.34) der Funktionen U_{lm} folgt. Verwendet man nun (3.12.37) wiederholt, beginnend mit $m = l$, erhält man durch Iteration (*Nachrechnen!*)

$$U_{lm}(u) = \frac{(-1)^l}{2^l l!} \sqrt{\frac{2l+1}{4\pi} \frac{(l+m)!}{(l-m)!}} \frac{1}{\sqrt{1-u^2}^m} \frac{d^{l-m}}{du^{l-m}} (1-u^2)^l, \quad m \in \{-l, -l+1, \dots, l-1, l\}. \quad (3.12.39)$$

Damit sind die Kugelflächenfunktionen vollständig bestimmt, denn es ist gemäß (3.12.34)

$$Y_{lm}(\vartheta, \varphi) = U_{lm}(\cos \vartheta) \exp(im\varphi). \quad (3.12.40)$$

Wir halten noch einmal die Orthonormalität der Kugelflächenfunktionen, die gemäß (3.12.20)

$$\langle l_1 m_1 | l_2 m_2 \rangle = \int_0^{\pi} d\vartheta \int_0^{2\pi} d\varphi \sin \vartheta Y_{l_1 m_1}^*(\vartheta, \varphi) Y_{l_2 m_2}(\vartheta, \varphi) = \delta_{l_1 l_2} \delta_{m_1 m_2} \quad (3.12.41)$$

lautet, fest. Setzen wir hierin $m_1 = m_2 = m$ und verwenden (3.12.34), folgt durch Substitution von $u = \cos \vartheta$ im Integral über ϑ

$$\int_{-1}^1 du U_{l_1 m}(u) U_{l_2 m} = \frac{1}{2\pi} \delta_{l_1 l_2}. \quad (3.12.42)$$

3. Wellenmechanik

Man kann nun auch umgekehrt durch fortgesetzte Anwendung der Formel (3.11.31) die U_{lm} aus $U_{l,-l}$ gewinnen. Die Berechnung von $U_{l,-l}$ aus (3.12.39) ist einfach, denn für $m = -l$ benötigt man nur die Ableitung

$$\frac{d^{2l}}{du^{2l}}(1-u^2)^l = \frac{d^{2l}}{du^{2l}}[(-u^2)^l + \dots] = (-1)^l (2l)! \quad (3.12.43)$$

Setzt man dies in (3.12.40) für $m = -l$ ein, folgt sofort die einfache Beziehung

$$U_{l,-l}(u) = (-1)^l U_{ll}(u). \quad (3.12.44)$$

Wendet man darauf fortgesetzt den Operator L_+ an, findet man allgemein (*Nachrechnen!*)

$$U_{l,-m}(u) = (-1)^m U_{lm}(u) \quad (3.12.45)$$

oder für die Kugelflächenfunktionen

$$Y_{l,-m}(\vartheta, \varphi) = (-1)^m Y_{lm}^*(\vartheta, \varphi). \quad (3.12.46)$$

Der Vollständigkeit halber erwähnen wir noch, dass für $m = 0$ die U_{l0} noch die **Legendre-Polynome** $P_l(u)$ definieren:

$$U_{l0}(u) = \sqrt{\frac{2l+1}{4\pi}} P_l(u). \quad (3.12.47)$$

Dabei ist der Vorfaktor Konvention. Aus (3.12.39) folgt

$$P_l(u) = \frac{1}{2^l l!} \frac{d^l}{du^l} (u^2 - 1)^l. \quad (3.12.48)$$

Aus (3.12.42) und (3.12.47) folgt für $m = 0$, dass die $P_l(u)$ ein System von orthogonalen Polynomen im Hilbert-Raum der auf dem Intervall $[-1, 1]$ definierten quadratintegrablen Funktionen bilden, und es ergibt sich daraus die Normierung zu

$$\int_{-1}^1 du P_{l_1}(u) P_{l_2}(u) = \frac{2}{2l_1+1} \delta_{l_1 l_2}. \quad (3.12.49)$$

3.13 Das Wasserstoffatom

In diesem Abschnitt beschäftigen wir uns mit der einfachsten (nichtrelativistischen) Behandlung des **Wasserstoffatoms** als einem der wenigen physikalisch interessanten analytisch lösbaren Energieeigenwertproblem.

3.13.1 Schwerpunkts- und Relativkoordinaten

Wir betrachten die gemeinsame Bewegung eines Elektrons und eines Atomkerns mit der Kernladungszahl Z , die wir hier beide zu spinlosen geladenen Punktteilchen vereinfachen. Für die Wechselwirkung nehmen wir ein einfaches Coulombpotential an, vernachlässigen also auch die Retardierung der Wechselwirkung, welche allerdings auch ein relativistischer Effekt ist. Wir kommen auf die Rechtfertigung dieser Annahme weiter unten noch kurz zurück. Bezeichnen wir die Observablen des Elektrons mit dem Index 1, die des Kerns mit Index 2, so ergibt sich der Hamiltonoperator

$$\mathbf{H}_{12} = \frac{\vec{\mathbf{p}}_1^2}{2m_1} + \frac{\vec{\mathbf{p}}_2^2}{2m_2} + V(|\vec{\mathbf{x}}_1 - \vec{\mathbf{x}}_2|). \quad (3.13.1)$$

3.13. Das Wasserstoffatom

Es ist dabei klar, daß die Orts- und Impulsoperatoren für die verschiedenen Teilchen miteinander vertauschen. Insgesamt haben wir also die folgende Operatorenalgebra vorliegen

$$[\mathbf{x}_{ij}, \mathbf{x}_{i'j'}] = [\mathbf{p}_{ij}, \mathbf{p}_{i'j'}] = 0, \quad [\mathbf{x}_{ij}, \mathbf{p}_{i'j'}] = i\hbar \delta_{ii'} \delta_{jj'}. \quad (3.13.2)$$

Dabei bezieht sich der erste Index $i, i' \in \{1, 2\}$ auf das Teilchen, während der zweite $j, j' \in \{1, 2, 3\}$ die Orts- bzw. Impulskomponenten durchnummeriert.

In Analogie zum klassischen Fall erwarten wir, daß der **Gesamt- bzw. Schwerpunktsimpuls**

$$\vec{\mathbf{P}} = \vec{\mathbf{p}}_1 + \vec{\mathbf{p}}_2 \quad (3.13.3)$$

mit \mathbf{H} kommutiert, also eine **Erhaltungsgröße** ist. Um dies zu zeigen, führen wir als weitere unabhängige Variablen noch Schwerpunkts- und Relativkoordinaten sowie den Relativimpuls wie in der klassischen Mechanik ein (vgl. Skript zur Theorie 1 [Hee18a]):

$$\vec{\mathbf{X}} = \frac{m_1 \vec{\mathbf{x}}_1 + m_2 \vec{\mathbf{x}}_2}{M}, \quad \vec{\mathbf{r}} = \vec{\mathbf{x}}_1 - \vec{\mathbf{x}}_2, \quad \vec{\mathbf{p}} = \frac{m_2 \vec{\mathbf{p}}_1 - m_1 \vec{\mathbf{p}}_2}{M}. \quad (3.13.4)$$

Dabei ist $M = m_1 + m_2$. Die Einteilchenimpulsoperatoren ergeben sich umgekehrt nach einigen einfachen algebraischen Umformungen zu

$$\vec{\mathbf{p}}_1 = \frac{m_1}{M} \vec{\mathbf{P}} + \vec{\mathbf{p}}, \quad \vec{\mathbf{p}}_2 = \frac{m_2}{M} \vec{\mathbf{P}} - \vec{\mathbf{p}}. \quad (3.13.5)$$

Man zeigt unter Verwendung von (3.13.2) durch einfaches Nachrechnen, dass die folgenden Kommutatorregeln gelten (*Übung!*)

$$\begin{aligned} [\mathbf{X}_i, \mathbf{X}_j] &= [\mathbf{P}_i, \mathbf{P}_j] = 0, & [\mathbf{X}_i, \mathbf{P}_j] &= i\hbar \delta_{ij} \mathbb{1}, \\ [\mathbf{X}_i, \mathbf{r}_j] &= [\mathbf{P}_i, \mathbf{r}_j] = [\mathbf{P}_i, \mathbf{p}_j] = 0, \\ [\mathbf{r}_i, \mathbf{r}_j] &= [\mathbf{p}_i, \mathbf{p}_j] = 0, & [\mathbf{r}_i, \mathbf{p}_j] &= i\hbar \delta_{ij} \mathbb{1}. \end{aligned} \quad (3.13.6)$$

Setzen wir (3.13.5) in den Hamiltonoperator (3.13.1) ein, erhalten wir (wieder analog zur klassischen Mechanik)

$$\mathbf{H}_{12} = \frac{1}{2M} \vec{\mathbf{P}}^2 + \frac{1}{2\mu} \vec{\mathbf{p}}^2 + V(\mathbf{r}) \quad \text{mit} \quad \mu = \frac{m_1 m_2}{M}, \quad \mathbf{r} = |\vec{\mathbf{x}}|. \quad (3.13.7)$$

Als vollständigen Satz kommutierender Observabler können wir also die drei Schwerpunktsimpulskomponenten $\vec{\mathbf{P}}$, den Hamiltonoperator der Relativbewegung

$$\mathbf{H} = \frac{1}{2\mu} \vec{\mathbf{p}}^2 + V(\mathbf{r}), \quad (3.13.8)$$

das Relativbahndrehimpulsbetragsquadrat $\vec{\mathbf{L}}^2 = (\vec{\mathbf{r}} \times \vec{\mathbf{p}})^2$ und dessen 3-Komponente \mathbf{L}_3 verwenden. Wir suchen also simultane Eigenvektoren $|\vec{\mathbf{P}}, E, l, m\rangle$ mit

$$\begin{aligned} \vec{\mathbf{P}} |\vec{\mathbf{P}}, E, l, m\rangle &= \vec{\mathbf{P}} |\vec{\mathbf{P}}, E, l, m\rangle, & \mathbf{H} |\vec{\mathbf{P}}, E, l, m\rangle &= E |\vec{\mathbf{P}}, E, l, m\rangle, \\ \vec{\mathbf{L}}^2 |\vec{\mathbf{P}}, E, l, m\rangle &= \hbar^2 l(l+1) |\vec{\mathbf{P}}, E, l, m\rangle, & \mathbf{L}_3 |\vec{\mathbf{P}}, E, l, m\rangle &= \hbar m |\vec{\mathbf{P}}, E, l, m\rangle. \end{aligned} \quad (3.13.9)$$

Offenbar verhält sich der Schwerpunkt wie ein freies Teilchen, und aus der Behandlung des Eigenwertproblems für Impulse ist klar, dass $\vec{\mathbf{P}} \in \mathbb{R}^3$ ist. Aus der Behandlung des simultanen Eigenwertproblems für $\vec{\mathbf{L}}^2$ und \mathbf{L}_3 im vorigen Abschnitt wissen wir, dass $l \in \mathbb{N}_0$ und für jedes l dann $m \in \{-l, -l+1, \dots, l-1, l\}$ ist.

Die Relativbewegung entspricht der Bewegung eines Teilchens mit der **reduzierten Masse** μ im Zentralpotential $V(\mathbf{r})$. Wir betrachten nun den Fall „wasserstoffartiger“ Systeme, wo dieses Potential durch ein **attraktives Coulombpotential**

$$V(\mathbf{r}) = -\frac{Ze^2}{4\pi\epsilon_0 r} \quad (3.13.10)$$

gegeben ist, wobei wir uns einen Atomkern der Ladungszahl Z vorstellen, um den sich ein Elektron bewegt. Wir bemerken dazu, dass ein Atomkern seinerseits ein zusammengesetztes System aus Protonen und Neutronen ist. Da die Protonen positiv geladen und die Neutronen elektrisch neutral sind, ist die Gesamtladung des Atomkerns $+Ze$, wobei die Ladungszahl Z der Anzahl der Protonen im Kern entspricht. Das Elektron trägt eine negative Elementarladung $q_{\text{Elektron}} = -e$.

Bequemer als die hier verwendeten SI-Einheiten ist die Einführung der dimensionslosen **Sommerfeld-Feinstrukturkonstanten** (Arnold Sommerfeld, 1868-1951), die durch

$$\alpha = \frac{e^2}{4\pi\epsilon_0 \hbar c} \simeq \frac{1}{137,036} \quad (3.13.11)$$

gegeben ist⁸. Auf die physikalische Bedeutung dieser dimensionslosen Konstanten kommen wir später noch zu sprechen.

3.13.2 Energieeigenwertproblem für die Relativbewegung

Es bleibt uns also, das simultane Eigenwertproblem für drei die Observablen der Relativbewegung \mathbf{H} , $\vec{\mathbf{L}}^2$ und L_3 zu lösen. Der Hamiltonoperator lautet gemäß (3.13.8) mit (3.13.10) und (3.13.11)

$$\mathbf{H} = \frac{1}{2\mu} \vec{\mathbf{p}}^2 - \frac{Z\alpha \hbar c}{r}. \quad (3.13.12)$$

Es ist nun am bequemsten, in der Ortsdarstellung weiterzuarbeiten. Die natürliche Wahl für die Ortskoordinaten sind aufgrund der Rotationssymmetrie Kugelkoordinaten (r, ϑ, φ) .

Zur Lösung des besagten simultanen Eigenwertproblems schreiben wir die zeitunabhängige Schrödinger-Gleichung in Kugelkoordinaten hin

$$\mathbf{H}\psi(\vec{x}) = -\frac{\hbar^2}{2\mu} \Delta\psi(\vec{x}) - \frac{Z\alpha \hbar c}{r} \psi(\vec{x}) = E\psi(\vec{x}). \quad (3.13.13)$$

Drücken wir den Laplace-Operator nun in Kugelkoordinaten aus und vergleichen mit (3.12.8), können wir diese Gleichung in die Form

$$\begin{aligned} \mathbf{H}\psi(r, \vartheta, \varphi) &= -\frac{\hbar^2}{2\mu r} \frac{\partial^2}{\partial r^2} [r\psi(r, \vartheta, \varphi)] + \frac{\mathbf{L}^2}{2\mu r^2} \psi(r, \vartheta, \varphi) - \frac{Z\alpha \hbar c}{r} \psi(r, \vartheta, \varphi) \\ &\stackrel{!}{=} E\psi(r, \vartheta, \varphi) \end{aligned} \quad (3.13.14)$$

bringen. Das simultane Eigenwertproblem für $\vec{\mathbf{L}}^2 := \mathbf{L}^2$ und L_z haben wir in Abschnitt 3.12 gelöst. Demnach muß sich die simultane Eigenfunktion für \mathbf{H} , \mathbf{L}^2 und L_z in der Form

$$\psi_{Elm}(r, \vartheta, \varphi) = \frac{1}{r} R_{El}(r) Y_{lm}(\vartheta, \varphi) \quad (3.13.15)$$

schreiben lassen, und zwar für beliebige Zentralpotentiale (nicht nur für das hier betrachtete Coulomb-Potential). Dabei sind die Y_{lm} die in Abschnitt 3.12 ausführlich besprochenen Kugelflächenfunktionen (s. auch [CH10]).

⁸In SI-Einheiten gilt: $e \simeq 1,602 \cdot 10^{-19}$ C, $\hbar c \simeq 3,162 \cdot 10^{-26}$ J m, $\epsilon_0 = 1/(\mu_0 c^2) \simeq 8,854 \cdot 10^{-12}$ C²/(J m). Alle Größen sind nach [T⁺18] auf die angegebene Stelle gerundet angegeben.

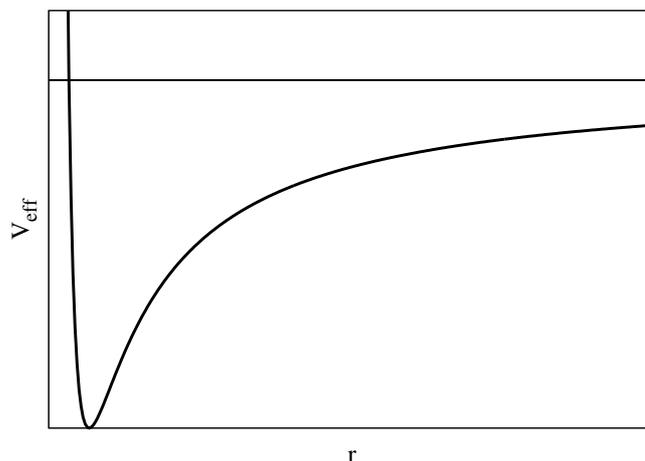


Abbildung 3.2: Das effektive Radialpotential (3.13.18) für das Coulombproblem.

3.13.3 Die Radialwellengleichung und Energieeigenwerte

Wegen

$$\mathbf{L}^2 \psi_{Elm}(r, \vartheta, \varphi) = \hbar^2 l(l+1) \psi_{Elm}(r, \vartheta, \varphi) \quad (3.13.16)$$

ergibt sich für die Radialwellenfunktion die Gleichung

$$-\frac{\hbar^2}{2\mu} R''_{El}(r) + \left[\frac{l(l+1)\hbar^2}{2\mu r^2} - \frac{Z\alpha\hbar c}{r} \right] R_{El}(r) = ER_{El}(r). \quad (3.13.17)$$

Diese Gleichung ist offenbar die Schrödingergleichung für die eindimensionale Bewegung im **effektiven Potential**

$$V_{\text{eff}}(r) = \frac{l(l+1)\hbar^2}{2\mu r^2} - \frac{Z\alpha\hbar c}{r}. \quad (3.13.18)$$

Allerdings ist hierbei zu beachten, dass r auf $r \geq 0$ eingeschränkt ist, und wir müssen daher zunächst Randbedingungen für die Radialwellenfunktion bei $r = 0$ finden. Dabei gilt es sicherzustellen, daß der Differentialoperator auf der linken Seite von (3.13.17) selbstadjungiert ist bzgl. des entsprechenden Skalarproduktes. Letzteres ergibt sich daraus, dass die Gesamtwellenfunktion (3.13.15) im $L^2(\mathbb{R}^3)$ liegt. Das Skalarprodukt für die Radialwellenfunktionen ist also

$$\langle R_1 | R_2 \rangle = \int_0^\infty dr R_1^*(r) R_2(r). \quad (3.13.19)$$

Damit der Differentialoperator $(d/dr)^2$ selbstadjungiert ist, muss offenbar für alle Radialwellenfunktionen die Randbedingung

$$R(0) = 0 \quad (3.13.20)$$

gefordert werden.

Wir wollen nun die verbliebene Radialgleichung (3.13.17) lösen und beschränken uns in diesem Kapitel auf die **gebundenen Zustände**. Da das effektive Potential (3.13.18) für $r \rightarrow 0^+$ gegen $+\infty$ divergiert und für $r \rightarrow \infty$ gegen 0 strebt, erwarten wir gebundene Zustände für $E < 0$ (vgl. Abb. 3.2). Wir suchen also Lösungen für $E < 0$. Dazu führen wir zunächst die Wellenzahl

$$k = \sqrt{-\frac{2\mu E}{\hbar^2}} \quad (3.13.21)$$

3. Wellenmechanik

und die dimensionslose Radialkoordinate

$$\xi = 2kr \quad (3.13.22)$$

ein. Setzen wir dann $R_{El}(r) = \chi_{nl}(\xi)$, ergibt sich nach einigen einfachen Umformungen die Radialgleichung (*Nachrechnen!*)

$$\left(\frac{d^2}{d\xi^2} - \frac{1}{4} + \frac{n}{\xi} - \frac{l(l+1)}{\xi^2} \right) \chi_{nl}(\xi) = 0 \quad \text{mit} \quad n = Z\alpha \sqrt{-\frac{\mu c^2}{2E}}. \quad (3.13.23)$$

Wir betrachten zunächst das Verhalten dieser Funktion für $\xi \rightarrow 0$. Dann kann man in der Klammer die Terme $-1/4 + n/\xi$ gegenüber $l(l+1)/\xi^2$ vernachlässigen, und man erhält

$$\chi_{nl}''(\xi) - \frac{l(l+1)}{\xi^2} \chi_{nl}(\xi) = 0. \quad (3.13.24)$$

Diese Gleichung lässt sich offenbar mit dem Ansatz $\chi_{nl}(\xi) = A\xi^{\lambda+1}$ lösen. Setzt man diesen Ansatz ein, folgt

$$\lambda(\lambda+1) = l(l+1) \Rightarrow \lambda_1 = l, \quad \lambda_2 = -(l+1). \quad (3.13.25)$$

Wegen der Randbedingung (3.13.20) kommt nur die Lösung mit $\lambda = \lambda_1 = l$ in Frage.

Weiter bestimmen wir das Verhalten von χ_{nl} für $\xi \rightarrow \infty$. Dann kann man in der Klammer in (3.13.23) die Terme $n/\xi - l(l+1)/\xi^2$ gegenüber $1/4$ vernachlässigen, und man erhält

$$\chi_{nl}''(\xi) = \frac{1}{4} \chi_{nl}(\xi) \quad (3.13.26)$$

Die Lösungen sind offenbar $\chi_{nl}(\xi) = A \exp(\pm \xi/2)$. Damit χ_{nl} quadratintegrabel ist, kommt nur das untere Vorzeichen in Frage.

Wir arbeiten nun die asymptotischen Lösungen für $\xi \rightarrow 0$ bzw. $\xi \rightarrow \infty$ in den folgenden Ansatz für χ_{nl} ein:

$$\chi_{nl}(\xi) = \xi^{l+1} \exp\left(-\frac{\xi}{2}\right) w_{nl}(\xi). \quad (3.13.27)$$

Setzen wir dies in (3.13.22) ein, ergibt sich für w_{nl} die DGL

$$\xi w_{nl}'' + (2l+2-\xi)w_{nl}' + (n-l-1)w_{nl} = 0. \quad (3.13.28)$$

Diese Gleichung lösen wir nun mit einem Potenzreihenansatz. Damit das asymptotische Verhalten für $\xi \rightarrow 0$ erhalten bleibt, muss offenbar

$$w_{nl}(\xi) = \sum_{j=0}^{\infty} A_j \xi^j \quad \text{mit} \quad A_0 \neq 0 \quad (3.13.29)$$

sein. Um die A_j zu bestimmen, setzen wir diesen Ansatz in (3.13.29) ein und bringen die dabei entstehenden Reihen für die verschiedenen Terme in eine einheitliche Form. Es gilt

$$\xi w_{nl}''(\xi) = \sum_{j=2}^{\infty} j(j-1)A_j \xi^{j-1}. \quad (3.13.30)$$

Substituieren wir hierin $j' = j - 1$, folgt

$$\xi w_{nl}''(\xi) = \sum_{j'=1}^{\infty} (j'+1)j'A_{j'+1} \xi^{j'}. \quad (3.13.31)$$

3.13. Das Wasserstoffatom

Ebenso folgt

$$(2l+2)w'_{nl}(\xi) = \sum_{j'=0}^{\infty} (2l+2)(j'+1)A_{j'+1}\xi^{j'}, \quad (3.13.32)$$

und

$$-\xi w'_{nl}(\xi) = -\sum_{j'=1}^{\infty} jA_j\xi^j \quad (3.13.33)$$

und

$$(n-l-1)w_{nl}(\xi) = \sum_{j=0}^{\infty} (n-l-1)A_j\xi^j. \quad (3.13.34)$$

Fassen wir alles zusammen, wobei wir alle Summationsindizes wieder j nennen, folgt aus (3.13.28)

$$\sum_{j=0}^{\infty} [(2l+j+2)(j+1)A_{j+1} + (n-l-1-j)A_j] \xi^j = 0. \quad (3.13.35)$$

Folglich muss für alle $j \in \mathbb{N}_0 = \{0, 1, 2, 3, \dots\}$ gelten

$$(2l+j+2)(j+1)A_{j+1} + (n-l-1-j)A_j = 0 \Rightarrow A_{j+1} = -\frac{n-l-1-j}{(2l+j+2)(j+1)}A_j. \quad (3.13.36)$$

Damit kann man alle Koeffizienten rekursiv berechnen, wobei $A_0 \neq 0$ beliebig ist.

Jetzt müssen wir aber beachten, dass die Funktion $w_{nl}(\xi)$ für $\xi \rightarrow \infty$ nicht zu schnell anwächst. Für $\xi \rightarrow \infty$ sind nun die Koeffizienten A_j für $j \rightarrow \infty$ wichtig. In diesem Limes ist aber gemäß (3.13.36)

$$A_{j+1} \underset{j \rightarrow \infty}{\cong} \frac{1}{j+1}A_j \Rightarrow A_j \underset{j \rightarrow \infty}{\cong} \frac{C}{j!}. \quad (3.13.37)$$

Bricht also die Reihe nicht für $j > n_r$ mit $n_r \in \mathbb{N}_0$ ab, verhält sich $w_{nl} \underset{\xi \rightarrow \infty}{\cong} C \exp \xi$. Dann würde aber wegen (3.13.27) $\chi_{nl}(\xi)$ für $\xi \rightarrow \infty$ wie $\exp(\xi/2)$ anwachsen und wäre nicht mehr quadratintegrabel. Folglich muss die Reihe bei einem endlichen n_r abbrechen, d.h. wir müssen

$$n_r = n - l - 1 \quad \text{mit} \quad n_r \in \mathbb{N}_0 \quad (3.13.38)$$

setzen. Dies ergibt die möglichen Eigenwerte für E , denn gemäß (3.13.23) ist

$$E_n = -\frac{Z^2 \alpha^2 \mu c^2}{2n^2}. \quad (3.13.39)$$

Dabei durchläuft die **Hauptquantenzahl** $n = n_r + l + 1$ offenbar die Werte $n \in \mathbb{N} = \{1, 2, \dots\}$, denn es ist $n_r \in \mathbb{N}_0$ und $l \in \mathbb{N}_0$.

Die Funktion w_{nl} ist nun ein Polynom n_r -ten Grades. Wenn man die entsprechende Rekursionsformel (3.13.36, die wir jetzt einfacher in der Form

$$A_{j+1} = -\frac{n_r - j}{(2l+j+2)(j+1)}A_j \quad (3.13.40)$$

schreiben können, ausgewertet, erhält man nach einiger Rechnung (*Nachprüfen!*)

$$A_j = (-1)^j \frac{n_r!}{(n_r - j)!} \frac{(2l+1)!}{(2l+1+j)!} A_0. \quad (3.13.41)$$

3. Wellenmechanik

Mit Hilfe von Binomialkoeffizienten $\binom{n}{k} = n!/(k!(n-k)!)$ können wir dies in die Form

$$A_j = \frac{(-1)^j}{j!} \binom{n_r + 2l + 1}{n_r - j} \frac{n_r!}{(n_r + 2l + 1)!} A_0 \quad (3.13.42)$$

bringen. Den Bruch vor dem Koeffizienten A_0 können wir in die Normierungskonstante der Wellenfunktion schieben, denn er hängt nicht von j ab. Damit wird

$$A_j = \frac{(-1)^j}{j!} \binom{n_r + 2l + 1}{n_r - j} N_{nl} \quad (3.13.43)$$

mit $N = \text{const.}$

Die Polynome

$$L_{n_r}^{2l+1}(\xi) = \sum_{j=0}^{n_r} \frac{(-1)^j}{j!} \binom{n_r + 2l + 1}{n_r - j} \xi^j \quad (3.13.44)$$

heißen **Laguerre-Polynome**, wobei n_r und $2l + 1$ Indizes sind (d.h. der obere Index ist nicht als Potenz misszuverstehen!). Damit wird schließlich gemäß (3.13.15), (3.13.22) und (3.13.27)

$$\psi_{nlm}(r, \vartheta, \varphi) = \frac{N_{nl}}{r} (2kr)^{l+1} L_{n-l-1}^{2l+1}(2kr) \exp(-kr) Y_{lm}(\vartheta, \varphi). \quad (3.13.45)$$

Dabei haben wir statt E die Wellenfunktionen mit den üblichen Quantenzahlen $n = n_r + l + 1$, l und m (Hauptquantenzahl, Drehimpulsquantenzahl und magnetische Quantenzahl genannt) durchnumeriert. Die Normierungskonstante bestimmt sich mit einigem Aufwand (vgl. Anh. A.3) zu

$$N_{nl} = \sqrt{\frac{k(n-l-1)!}{n(n+l)!}}. \quad (3.13.46)$$

Es ist im Folgenden bequem, die verschiedenen Parameter in (3.13.45) mit noch etwas anderen Größen auszudrücken. Zunächst ist es sinnvoll, den **Bohrschen Radius** a_B einzuführen, und zwar als

$$a_B = \frac{\hbar}{\alpha \mu c}. \quad (3.13.47)$$

Dann wird gemäß (3.13.23) und (3.13.39)

$$k = \frac{Z}{na_B}, \quad E_n = -\frac{Z^2 \hbar^2}{2\mu a_B^2 n^2} = -\frac{\mu c^2 \alpha^2 Z^2}{2n^2} \simeq -\frac{13,6 \text{ eV} Z^2}{n^2}. \quad (3.13.48)$$

Setzt man all dies in (3.13.45) ein, erhält man schließlich endgültig

$$\psi_{nlm}(r, \vartheta, \varphi) = \frac{2}{n^2} \sqrt{\frac{Z^2(n-l-1)!}{(n+l)! a_B^3}} \left(\frac{2Zr}{na_B}\right)^l \exp\left(-\frac{Zr}{na_B}\right) L_{n-l-1}^{2l+1}\left(\frac{2Zr}{na_B}\right) Y_{lm}(\vartheta, \varphi). \quad (3.13.49)$$

Diese Eigenfunktionen sind dann ein Orthonormalsystem. Es ist allerdings *kein vollständiges Orthonormalsystem*, da es noch Lösungen des Energieeigenwertproblems im kontinuierlichen Spektrum $E \geq 0$ gibt. Diese Zustände entsprechen **Streuzuständen**, also nicht gebundenen Zuständen.

Jedenfalls erfüllen die Eigenfunktionen zu den gebundenen Zuständen (3.13.49) die Orthonormierungsbedingungen

$$\langle nlm | n'l'm' \rangle = \int_0^\infty dr \int_0^\pi d\vartheta \int_0^{2\pi} d\varphi r^2 \sin \vartheta \psi_{nlm}^*(r, \vartheta, \varphi) \psi_{n'l'm'}(r, \vartheta, \varphi) = \delta_{nn'} \delta_{ll'} \delta_{mm'}. \quad (3.13.50)$$

Wir bemerken noch, dass es eine spezielle Eigenschaft des Coulomb-Potentials ist, dass die Energieeigenwerte nur von der Hauptquantenzahl n und nicht noch zusätzlich von l abhängen. Entsprechend gibt es zu jedem Energieeigenwert (außer dem Grundzustand entsprechend $n = 1$) n^2 verschiedene Eigenzustände. Man sagt auch, das Energieniveau E_n ist n^2 -fach entartet. Dies sieht man wie folgt ein: Zu vorgegebenem $n \in \mathbb{N}$ hat man $n = n_r - l - 1$. Wegen $n_r \in \mathbb{N}_0$ kann zu gegebenem n folglich l die Werte $0, 1, \dots, n-1$ annehmen. Zu jedem l gehören die $(2l+1)$ möglichen Werte für $m \in \{-l, -l+1, \dots, l-1, l\}$. Damit ist der Entartungsgrad des zu n gehörigen Energieniveaus

$$\sum_{l=0}^{n-1} (2l+1) = n^2. \quad (3.13.51)$$

Die Summenformel kann man leicht durch vollständige Induktion beweisen (*Übung!*).

3.14 Der Spin

3.14.1 Heuristische Einführung des Spins

Wir haben in Abschnitt 3.11 gesehen, dass aus den Drehimpulsvertauschungsrelationen allein für die simultanen Eigenwerte von \vec{J}^2 und J_3 die Werte $\hbar^2 j(j+1)$ mit $j \in \{0, 1/2, 1, \dots\}$ bzw. (für gegebenes j) $m\hbar$ mit $m \in \{-j, -j+1, \dots, j-1, j\}$ resultieren. Nur die spezielle Form $\vec{L} = \vec{x} \times \vec{p}$ des Bahndrehimpulsoperators hat dann die Eigenwerte für \vec{L}^2 auf $\hbar^2 l(l+1)$ mit $l \in \mathbb{N}_0$ eingeschränkt. Zunächst sieht es so aus, als gäbe es für die Möglichkeit **halbzahliger Drehimpulsquantenzahlen** keine Verwendung. Dies ist jedoch ein Irrtum, der auf einer allzu sehr an der klassischen Mechanik orientierten Analogiebildung beruht! Vielmehr ermöglicht die Quantentheorie das Vorhandensein einer Art „intrinsischen Drehimpulses“ von (Elementar-)Teilchen, den man **Spin** nennt.

In der klassischen Mechanik kennt man eine solche Möglichkeit nur von ausgedehnten Objekten, z.B. dem starren Körper. Außer drei Freiheitsgraden, die die Lage des Körpers im Raum beschreiben (z.B. die Koordinaten des Schwerpunkts dieses Körpers), kann er sich noch um einen beliebigen körperfesten Punkt drehen, und dieser Drehung entspricht ein zusätzlicher zu einem ggf. vorhandenen Bahndrehimpuls zu addierender intrinsischer Drehimpuls (Drall oder eben engl. Spin).

In der Tat gab es schon vor der Entdeckung der eigentlichen modernen Quantentheorie bereits Hinweise auf solche intrinsischen Spindrehimpulse von Teilchen wie insbesondere dem Elektron und tatsächlich wurde der **Spin des Elektrons** im Wesentlichen bereits 1922 durch Stern und Gerlach in Frankfurt (im Gebäude, das heute den physikalischen Verein in Bockenheim beherbergt) entdeckt. Das war gut 3 Jahre vor Entdeckung der modernen Quantentheorie durch Heisenberg, Born, Schrödinger und Dirac, und die Deutung des Versuchs war entsprechend auch noch fehlerhaft. Wir kommen darauf weiter unten zurück.

Betrachten wir nun aber gleich die moderne Theorie des Spins und beginnen wieder mit einem freien Teilchen. Wie wir wissen, ist für ein freies Teilchen ohne Spin, wie wir es bislang betrachtet haben, ein vollständiger Satz von kompatiblen Observablen durch die drei **Impulskomponenten** \vec{p} gegeben.

Stellen wir uns nun einen klassischen starren Körper vor, so manifestiert sich dessen intrinsischer Drehimpuls, also die Drehung des Körpers um z.B. seinen Schwerpunkt, der der natürliche Bezugspunkt für einen freien starren Körpers ist, als ein Drehimpulsvektor \vec{s} , den ein im **inertialen Ruhesystem des Schwerpunkts** ruhender Beobachter misst.

Für ein quantenmechanisches Punktteilchen können wir nun einfach entsprechende **Spin-Drehimpulsoperatoren** \vec{s} einführen. Wir nehmen an, dass diese Spin-Drehimpulsoperatoren ähnlich wie im klassischen Fall

⁹Wir schreiben jetzt für eine allgemeinere Art Drehimpuls als den in Abschnitt 3.11 behandelten Bahndrehimpuls \vec{J} . Dies sollen drei Operatoren sein, die die Kommutatorrelationen $[\mathbf{J}_j, \mathbf{J}_k] = i\hbar \epsilon_{jkl} \mathbf{J}_l$ erfüllen.

3. Wellenmechanik

des starren Körpers von den Translationsfreiheitsgraden (also Ort \vec{x} und Impuls \vec{p}) unabhängig sind und daher die entsprechenden Operatoren mit den Spinkomponenten kommutieren. Wir haben dann also folgende Vertauschungsrelationen für die fundamentalen Observablen eines Teilchens mit Spin:

$$[\mathbf{x}_j, \mathbf{x}_k] = 0, \quad [\mathbf{p}_j, \mathbf{p}_k] = 0, \quad [\mathbf{x}_j, \mathbf{p}_k] = i\hbar \delta_{jk} \mathbb{1}, \quad [\mathbf{s}_j, \mathbf{x}_k] = 0, \quad [\mathbf{s}_j, \mathbf{p}_k] = 0, \quad [\mathbf{s}_j, \mathbf{s}_k] = i\hbar \epsilon_{jkl} \mathbf{s}_l. \quad (3.14.1)$$

Entsprechend kann man als (verallgemeinerte) vollständige Systeme von Eigenvektoren \vec{s}^2 und s_3 zusammen mit entweder \vec{x} oder \vec{p} .

Ein freies Elementarteilchen ist dann durch seine Masse und seine Spin-Quantenzahl, die den Eigenwert von \vec{s}^2 zu $\hbar^2 s(s+1)$ festlegt, definiert. Dabei gibt es prinzipiell die Möglichkeiten $s \in \{0, 1/2, 1, \dots\}$. Dann sind die möglichen Eigenwerte von s_3 durch $\sigma \hbar$ mit $\sigma \in \{-s, -s+1, \dots, s-1, s\}$ gegeben. Z.B. ist ein Elektron nach dem gegenwärtigen Stand des Wissens ein Elementarteilchen, d.h. man hat noch keinerlei Konstituenten, aus denen es sich zusammensetzen könnte, entdeckt. Es besitzt den Spin $s = 1/2$, und entsprechend sind die Eigenwerte von s_3 durch $\pm \hbar/2$ gegeben („spin up“ und „spin down“).

Die Wellenfunktionen in der Orts- bzw. Impulsdarstellung sind demnach für ein Elektron **zweikomponentige Spinoren**, auch **Weyl-Spinoren**¹⁰ genannt, dessen Komponenten durch

$$\psi_\sigma(\vec{x}) = \langle \vec{x}, \sigma | \psi \rangle \quad \text{bzw.} \quad \tilde{\psi}_\sigma(\vec{p}) = \langle \vec{p}, \sigma | \psi \rangle, \quad \sigma \in \{-1/2, +1/2\} \quad (3.14.2)$$

gegeben sind. Gewöhnlich ordnet man die Komponenten in Spaltenvektoren, die man in dem Fall **Spinoren** nennt, $(\psi_{+1/2}, \psi_{-1/2})^T$ an. Die (verallgemeinerten Eigenvektoren) sind gemäß

$$\langle \vec{x}, \sigma | \vec{x}', \sigma' \rangle = \delta^{(3)}(\vec{x} - \vec{x}') \delta_{\sigma\sigma'}, \quad \langle \vec{p}, \sigma | \vec{p}', \sigma' \rangle = \delta^{(3)}(\vec{p} - \vec{p}') \delta_{\sigma\sigma'}, \quad (3.14.3)$$

normiert, und das Skalarprodukt lautet in der Orts- bzw. Impulsdarstellung entsprechend

$$\langle \psi_1 | \psi_2 \rangle = \sum_{\sigma=\pm 1/2} \int_{\mathbb{R}^3} d^3x \psi_{1\sigma}^*(\vec{x}) \psi_{2\sigma}(\vec{x}) \equiv \int_{\mathbb{R}^3} d^3x \psi_1^\dagger(\vec{x}) \psi_2(\vec{x}), \quad (3.14.4)$$

wobei

$$\psi(\vec{x}) = \begin{pmatrix} \psi_{+1/2}(\vec{x}) \\ \psi_{-1/2}(\vec{x}) \end{pmatrix}, \quad \psi^\dagger(\vec{x}) = (\psi_{+1/2}^*(\vec{x}), \psi_{-1/2}^*(\vec{x})). \quad (3.14.5)$$

die abkürzende Schreibweise mit **spinorwertigen Wellenfunktionen** ist. Analoge Formeln wie (3.14.4) und (3.14.5) gelten freilich auch für die Impulsdarstellungen $\tilde{\psi}(\vec{p})$ (Spinorschreibweise) bzw. die Spinorkomponenten $\tilde{\psi}_\sigma(\vec{p})$.

Die Darstellungen der Spinoperatoren \vec{s} in dieser Orts-Spin- bzw. Impuls-Spin-Darstellung der Wellenfunktionen sind demnach **Matrizen im Spinorraum**, die wir mit der allgemeinen Drehimpulsalgebra wie in Abschnitt 3.11 bestimmen können. Dies ist für $s = 1/2$ sogar besonders einfach. Für s_3 gilt

$$\mathbf{s}_3 |\vec{x}, \sigma\rangle = \hbar \sigma |\vec{x}, \sigma\rangle, \quad \sigma = \pm \frac{1}{2}, \quad (3.14.6)$$

d.h. in der Ortsdarstellung ist

$$[\hat{s}_3 \psi(\vec{x})]_\sigma = \langle \vec{x}, \sigma | \mathbf{s}_3 \psi \rangle = \langle \mathbf{s}_3 \vec{x}, \sigma | \psi \rangle = \hbar \sigma \langle \vec{x}, \sigma | \psi \rangle = \hbar \sigma \psi_\sigma(\vec{x}). \quad (3.14.7)$$

In der Spinor-Schreibweise können wir dies als

$$\hat{s}_3 \psi = \begin{pmatrix} \hbar/2 & 0 \\ 0 & -\hbar/2 \end{pmatrix} \begin{pmatrix} \psi_{+1/2}(\vec{x}) \\ \psi_{-1/2}(\vec{x}) \end{pmatrix} \quad (3.14.8)$$

¹⁰Benannt nach dem Mathematiker Hermann Weyl, der sich u.a. auch mit der Mathematik der Quantentheorie und insbesondere des Drehimpulses ausführlich beschäftigt hat.

3.14. Der Spin

schreiben. Die Spinoperatoren in der Orts-Spin-Darstellungen sind also einfach $\mathbb{C}^{2 \times 2}$ -Matrizen, die auf die Spinoren, geschrieben als Spaltenvektoren, wirken.

Um nun auch \hat{s}_1 und \hat{s}_2 zu finden, verwenden wir die Leiteroperatoren und die zu (3.11.29) und (3.11.31) analogen Beziehungen, wobei wir hier $l = s = 1/2$ und $m = \sigma \in \{1/2, -1/2\}$ zu setzen haben. Es gilt also

$$\mathbf{s}_+ |\vec{x}, \sigma\rangle = \hbar \sqrt{(1/2 - \sigma)(3/2 + \sigma)} |\vec{x}, \sigma + 1\rangle, \quad \mathbf{s}_- |\vec{x}, \sigma\rangle = \hbar \sqrt{(1/2 + \sigma)(3/2 - \sigma)} |\vec{x}, \sigma - 1\rangle, \quad (3.14.9)$$

und damit ist

$$\begin{aligned} \mathbf{s}_+ |\vec{x}, +1/2\rangle &= 0, & \mathbf{s}_+ |\vec{x}, -1/2\rangle &= \hbar |\vec{x}, +1/2\rangle \\ \mathbf{s}_- |\vec{x}, +1/2\rangle &= \hbar |\vec{x}, -1/2\rangle, & \mathbf{s}_- |\vec{x}, -1/2\rangle &= 0. \end{aligned} \quad (3.14.10)$$

Damit erhält man analog wie bei \hat{s}_3

$$\hat{s}_+ = \begin{pmatrix} 0 & \hbar \\ 0 & 0 \end{pmatrix}, \quad \hat{s}_- = \begin{pmatrix} 0 & 0 \\ \hbar & 0 \end{pmatrix} \quad (3.14.11)$$

und wegen $\hat{s}_\pm = \hat{s}_1 \pm i\hat{s}_2$

$$\begin{aligned} \hat{s}_1 &= \frac{1}{2}(\hat{s}_+ + \hat{s}_-) = \frac{\hbar}{2} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \\ \hat{s}_2 &= \frac{1}{2i}(\hat{s}_+ - \hat{s}_-) = \frac{\hbar}{2} \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \end{aligned} \quad (3.14.12)$$

Zur Vereinfachung schreibt man allgemein $\hat{s}_j = \hbar/2\hat{\sigma}_j$ mit den **Pauli-Matrizen**

$$\hat{\sigma}_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \hat{\sigma}_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \hat{\sigma}_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \quad (3.14.13)$$

3.14.2 Magnetisches Moment einer klassischen Stromverteilung

Die Möglichkeit, dass in der Quantentheorie ein elementares Teilchen einen Spin haben kann, eröffnet die Möglichkeit, dass es auch ein **magnetisches Moment** besitzt, das proportional zum Spin ist.

Um einen Hinweis zu erhalten, wie wir diese Idee beschreiben können, betrachten wir zunächst als Modell eine auf ein (kleines) Gebiet beschränkte stationäre klassische Stromverteilung $\vec{j}(\vec{x})$. Wegen der Ladungserhaltung muss

$$\vec{\nabla} \cdot \vec{j} = 0 \quad (3.14.14)$$

gelten. Die Ladungsverteilung sei auf einen Raumbereich innerhalb einer Kugel vom Radius R um den Ursprung des Koordinatensystems K_R beschränkt, und wir wollen nun die Kraft, die ein äußeres Magnetfeld $\vec{B}(\vec{x})$ auf diese Stromverteilung ausübt, berechnen. Nach der Formel für die **Lorentz-Kraft** auf Teilchen im Magnetfeld gilt

$$\vec{F} = \int_{K_R} d^3x \vec{j}(\vec{x}) \times \vec{B}(\vec{x}). \quad (3.14.15)$$

Wir nehmen nun an, dass sich \vec{B} über die Kugel K_R um den Koordinatenursprung wenig ändert und es daher ausreicht, die Taylor-Entwicklung bis zur 1. Ordnung

$$\vec{B}(\vec{x}) = \vec{B}(\vec{0}) + (\vec{x} \cdot \vec{\nabla}) \vec{B}(\vec{0}) \quad (3.14.16)$$

3. Wellenmechanik

im Integral (3.14.16) zu verwenden. Dann erhalten wir

$$\vec{F} = \int_{K_R} d^3x \vec{j}(\vec{x}) \times [\vec{B}(\vec{0}) + (\vec{x} \cdot \vec{\nabla})\vec{B}(\vec{0})]. \quad (3.14.17)$$

Wir wollen nun zeigen, dass der erste Term verschwindet.

Sei dazu $f(\vec{x})$ eine beliebige Funktion. Dann gilt

$$\vec{\nabla} \cdot (f \vec{j}) = \partial_k (f j_k) = j_k \partial_k f + f \partial_k j_k = j_k \partial_k f, \quad (3.14.18)$$

wobei wir im letzten Schritt (3.14.14) verwendet haben. Integrieren wir diese Gleichung über die Kugel K_R , folgt mit dem Gaußschen Integralsatz und wegen $\vec{j}(\vec{x}) = 0$ für $r = |\vec{x}| \geq R$

$$\int_{K_R} d^3x \vec{j} \cdot \vec{\nabla} f = \int_{\partial K_R} d^2\vec{f} \cdot (f \vec{j}) = 0. \quad (3.14.19)$$

Setzen wir hier $f(\vec{x}) = x_k$ ($k \in \{1, 2, 3\}$), folgt wegen $\vec{j} \cdot \vec{\nabla} x_k = j_l \partial_l x_k = j_l \delta_{lk} = j_k$ aus (3.14.19)

$$\int_{K_R} d^3\vec{x} j_k = 0. \quad (3.14.20)$$

Dies zeigt, dass nur der zweite Term in (3.14.17) einen Beitrag zur Kraft liefert. Wir schreiben diesen Term nun noch ein wenig um:

$$F_k = \int_{K_R} d^3x \epsilon_{klm} j_l x_n \partial_n B_m(\vec{0}). \quad (3.14.21)$$

Setzen wir nun in (3.14.19) $f = x_l x_n$, liefert (3.14.19)

$$\int_{K_R} d^3x j_k \partial_k (x_l x_n) = \int_{K_R} d^3x j_k (\delta_{kl} x_n + \delta_{kn} x_l) = \int_{K_R} d^3x (j_l x_n + j_n x_l) = 0. \quad (3.14.22)$$

Wir verwenden dies nun, um (3.14.20) weiter umzuformen:

$$F_k = \frac{1}{2} \int_{K_R} d^3x \epsilon_{klm} (x_n j_l - x_l j_n) \partial_n B_m(\vec{0}) = \frac{1}{2} \int_{K_R} d^3x \epsilon_{klm} \epsilon_{nla} (\vec{x} \times \vec{j})_a \partial_n B_m(\vec{0}). \quad (3.14.23)$$

Wir definieren nun das **magnetische Moment** der Stromverteilung durch

$$\vec{\mu} = \frac{1}{2} \int_{K_R} d^3x \vec{x} \times \vec{j}. \quad (3.14.24)$$

Dann wird (3.14.23) zu

$$F_k = \epsilon_{lkm} \epsilon_{lna} \mu_a \partial_n B_m(\vec{0}) = (\delta_{kn} \delta_{ma} - \delta_{ka} \delta_{mn}) \partial_k (\mu_a B_a)_{\vec{x}=\vec{0}} \Rightarrow \vec{F} = \vec{\nabla} [\vec{\mu} \cdot \vec{B}(\vec{x})]_{\vec{x}=\vec{0}}. \quad (3.14.25)$$

Dabei haben wir verwendet, dass $\vec{\mu} = \text{const}$ ist. Damit ist also die Kraft auf die am Koordinatenursprung lokalisierte Stromverteilung eine konservative Kraft mit dem Potential

$$V(\vec{x}) = -\vec{\mu} \cdot \vec{B}(\vec{x}). \quad (3.14.26)$$

Demnach ist

$$\vec{F}(\vec{x}) = -\vec{\nabla} V(\vec{x}). \quad (3.14.27)$$

die Kraft für eine Stromverteilung, die um den Ort \vec{x} herum lokalisiert ist.

Im Folgenden betrachten wir noch die Stromverteilung als Verteilung von Punktladungen der Größe q mit der Teilchenzahldichte $n(\vec{x})$ und dem Geschwindigkeitsfeld $\vec{v}(\vec{x})$. Dann ist $\vec{j}(\vec{x}) = qn(\vec{x})\vec{v}(\vec{x})$ und

$$\vec{\mu} = \frac{q}{2} \int_{K_R} d^3x n(\vec{x}) \vec{x} \times \vec{v}(\vec{x}) = \frac{q}{2m} \vec{L}, \quad (3.14.28)$$

wobei

$$\vec{L} = \int_{K_R} d^3x m n(\vec{x}) \vec{x} \times \vec{v}(\vec{x}) \quad (3.14.29)$$

der **Gesamtbahndrehimpuls** der die Stromverteilung erzeugenden Teilchen ist.

Schließlich betrachten wir noch das **Drehmoment** auf die Stromverteilung, die wir uns wieder im Koordinatenursprung lokalisiert denken. Wir machen wieder die Näherung, dass sich \vec{B} über die Kugel K_R , die diese Stromverteilung enthält, nur wenig ändert. Für das Drehmoment kommen wir mit der führenden Ordnung $\vec{B}(\vec{x}) = \vec{B}(\vec{0}) = \vec{B}_0$ aus. Damit wird das Drehmoment

$$\vec{M} = \int_{K_R} d^3x \vec{x} \times (\vec{j} \times \vec{B}_0) = \int_{K_R} d^3x [j(\vec{x} \cdot \vec{B}_0) - \vec{B}_0(\vec{x} \cdot \vec{j})]. \quad (3.14.30)$$

Verwenden wir nun (3.14.19) mit $f(\vec{x}) = r^2 = \vec{x}^2$, sehen wir, dass wegen $\vec{\nabla}f = 2\vec{x}$ der zweite Term verschwindet, d.h. es gilt

$$M_k = \int_{K_R} d^3x j_k x_l B_{0l} = \frac{1}{2} \int_{K_R} d^3x (x_l j_k - x_k j_l) B_{0l} = \frac{1}{2} \int_{K_R} d^3x \epsilon_{lkm} (\vec{x} \times \vec{j})_m B_{0l} = (\vec{\mu} \times \vec{B}_0)_k. \quad (3.14.31)$$

Dabei haben wir im letzten Schritt die Definition des magnetischen Moments (3.14.24) verwendet. Es ist also

$$\vec{M} = \vec{\mu} \times \vec{B}(\vec{x}), \quad (3.14.32)$$

wenn wir uns die Stromverteilung wieder allgemein um den Ort \vec{x} verteilt denken.

Verwenden wir nun (3.14.28), folgt für die Bewegungsgleichung für den Drehimpuls der Stromverteilung

$$\dot{\vec{L}} = \frac{2m}{q} \dot{\vec{\mu}} = \vec{\mu} \times \vec{B} \Rightarrow \dot{\vec{\mu}} = \frac{q}{2m} \vec{\mu} \times \vec{B}. \quad (3.14.33)$$

Das bedeutet, dass in einem Magnetfeld das magnetische Moment der Stromverteilung mit der **klassischen Larmor-Frequenz** (Joseph Larmor 1857-1942)

$$\omega_{\text{Larmor}} = \frac{q|\vec{B}|}{2m} \quad (3.14.34)$$

um die Richtung dieses Magnetfeldes präzediert (für $q > 0$ im Uhrzeigersinn). Auf ein Teilchen mit magnetischem Moment in einem *inhomogenen* Magnetfeld wirkt also eine Kraft gemäß (3.14.27) und ein Drehmoment gemäß (3.14.32).

3.14.3 Teilchen mit $s = 0$ im Magnetfeld

Wir betrachten nun zunächst ein Teilchen mit Spin 0 im Magnetfeld. Wir geben dazu den Hamilton-Operator an:

$$\mathbf{H} = \frac{1}{2M} [\vec{\mathbf{p}} - q\vec{A}(\vec{x})]^2. \quad (3.14.35)$$

Dabei ist \vec{A} das Vektorpotential des Magnetfeldes, d.h.

$$\vec{B} = \vec{\nabla} \times \vec{A}. \quad (3.14.36)$$

Dass dies wirklich der Hamilton-Operator für ein Teilchen mit Masse M und Ladung q ist, zeigen wir in den Übungen. Dort berechnen wir mittels der Formel (3.5.26) für den Operator, der die Zeitableitung einer Observablen repräsentiert, die den Operator für die Geschwindigkeit des Teilchens repräsentiert:

$$\vec{v} = \dot{\vec{x}} = \frac{1}{i\hbar} [\vec{x}, \mathbf{H}] = \frac{1}{m} [\vec{p} - q\vec{A}(\vec{x})]. \quad (3.14.37)$$

Wir sehen also, dass hier der Impulsoperator nicht die übliche mechanische Bedeutung $\vec{p}_{\text{mech}} = m\vec{v}$ besitzt. Vielmehr ist bei Anwesenheit eines Magnetfeldes $\vec{p} = m\vec{v} + q\vec{A}(\vec{x})$ der Impuls ein sogenannter generalisierter oder **kanonischer Impuls**. In der klassischen Mechanik ergibt sich dies im Rahmen der Formulierung der Dynamik des Systems mittels des sog. **Hamiltonschen Prinzips** der kleinsten Wirkung (vgl. z.B. [Hee18a, Hee08, BFK⁺18a, LL97]). Jedenfalls ergibt sich mit für den Operator der auf das Teilchen wirkenden Kraft

$$\vec{F} = \frac{1}{i\hbar} [m\vec{v}, \mathbf{H}] = \frac{q}{2} [\vec{v} \times \vec{B}(\vec{x}) - \vec{B}(\vec{x}) \times \vec{v}]. \quad (3.14.38)$$

Hätten wir es mit gewöhnlichen Vektoren zu tun, ergäbe sich $\vec{F} = q\vec{v} \times \vec{B}$, also die **Lorentz-Kraft** auf ein geladenes Teilchen im Magnetfeld. Da es sich bei \vec{v} und \vec{x} um Operatoren handelt, dürfen wir in (3.14.38) die Faktoren in den Vektorprodukten nicht einfach vertauschen. Allerdings ist (3.14.38) insofern eine plausible Verallgemeinerung der klassischen Formel für die Lorentz-Kraft zu einem Operatorausdruck für die Observable der Kraft, als aufgrund der „Symmetrisierung“ der nichtkommutierenden Operatorprodukte einen **selbstadjungierte Operator** für den Operator der Kraft ergibt, wie es für Repräsentanten von Observablen im quantentheoretischen Formalismus sein muss.

Wir bemerken noch, dass in der Ortsdarstellung der *kanonische* Impuls durch $\hat{p} = -i\hbar\vec{\nabla}$ repräsentiert wird. Der Hamilton-Operator (3.14.35) lautet also in der Ortsdarstellung

$$\hat{H} = \frac{1}{2M} [-i\hbar\vec{\nabla} - q\vec{A}(\vec{x})]^2. \quad (3.14.39)$$

3.14.4 Teilchen mit Spin im Magnetfeld

Besitzt ein Teilchen einen Spin $s \neq 0$, kann es zusätzlich zur Ladung auch noch ein **magnetisches Moment** besitzen. Dann besitzt es nicht nur eine elektrische Ladung wie ein klassisches geladenes (Punkt-)Teilchen sondern auch die Eigenschaften eines klassischen **Permanentmagneten**. Es gibt auch ungeladene Teilchen, die ein magnetisches Moment besitzen, z.B. das Neutron, das ein ungeladenes Teilchen mit Spin 1/2 ist. Heute wissen wir allerdings, dass es aus geladenen Quarks zusammengesetzt ist. In Abschnitt 3.14.2 haben wir gesehen, dass eine klassische Stromverteilung ein magnetisches Moment besitzt, das proportional zum Drehimpuls dieser Stromverteilung ist. Entsprechend gilt für den Operator des magnetische Moments eines Teilchens mit Spin

$$\vec{\mu} = g \frac{q}{2m} \vec{s}. \quad (3.14.40)$$

Der gegenüber dem klassischen Resultat, für das gemäß (3.14.28) $g = 1$ gilt, zusätzliche Faktor g heißt **Landé-Faktor** (Alfred Landé 1888-1976) oder **Gyrofaktor**.

Wir bemerken, dass sich in der relativistischen Theorie für ein elementares Teilchen mit Spin 1/2 (für ein Dirac-Teilchen) $g = 2$ ergibt. Dies ist auch für das **Elektron** und das **Myon**, die nach heutigem Kenntnisstand tatsächlich elementare Teilchen sind (sog. **Leptonen**), in sehr guter Näherung erfüllt. Die Abweichung

von diesem Wert $g = 2$ nennt man das „anomale magnetische Moment“ des Elektrons bzw. Myons. Theoretisch ergeben sich diese Abweichungen daraus, dass nicht nur die Teilchen im Rahmen der Quantentheorie behandelt werden müssen sondern auch das elektromagnetische Feld. Dabei ist die natürliche Beschreibung für relativistische Teilchen und Felder die **Quantenfeldtheorie**, die u.a. eine Beschreibung von Streuprozessen, bei denen Teilchenerzeugungs- und -vernichtungsprozesse vorkommen, ermöglicht. In der Tat sind Teilchenerzeugungs- und -vernichtungsprozesse bei relativistischen Stoßenergien zwischen Teilchen (z.B. am Large Hardon Collider am CERN) der Normalfall. Aufgrund der Quantisierung des elektromagnetischen Feldes, ergeben sich **Quantenfluktuationen**, die zu sog. **Strahlungskorrekturen** führen, die den Gyrofaktor der Leptonen korrigieren. Die Vorhersage dieser Korrekturen durch die störungstheoretische Behandlung der **Quantenelektrodynamik** (also der quantisierten Theorie von wechselwirkenden geladenen Teilchen und dem elektromagnetischen Feld) gehört zu den am genauesten experimentell bestätigten Vorhersagen der theoretischen Physik. Für das Elektron gilt $g_e = 2,002\,319\,304\,361\,82(52)$. Dabei gibt die Klammer die Messungsgenauigkeit der letzten Nachkommastellen an.

Für zusammengesetzte Teilchen wie Proton und Neutron, beides wie das Elektron Spin-1/2-Teilchen, ergeben sich davon erheblich abweichende Werte: $g_p = 5,585\,694\,702(17)$, $g_n = -3,826\,085\,45(90)$. Beim Proton und ungeladenen Neutron bezieht sich der Wert des Gyrofaktors auf das sog. **Kernmagneton**, also $\mu_N = e\hbar/(2M_p)$, d.h. $\mu_p = g_p\mu_N$ bzw. $\mu_n = g_n\mu_N$. Beim Elektron definiert man entsprechend das **Bohr-Magneton** $\mu_B = e\hbar/(2M_e)$ und das magnetische Moment ist entsprechend dem Gyrofaktor 2 ein (negatives) Bohr-Magneton.

Der Hamilton-Operator für ein geladenes Teilchen mit Spin im Magnetfeld $\vec{B} = \vec{\nabla} \times \vec{A}$ ergibt sich dann zu

$$\mathbf{H} = \frac{1}{2m} [\vec{p} - q\vec{A}(\vec{x})]^2 - \vec{\mu} \cdot \vec{B}(x), \quad (3.14.41)$$

d.h. zum Hamilton-Operator eines Teilchens ohne Spin im Magnetfeld (3.14.35) ist noch die potentielle Energie der Kraft auf das magnetische Moment des Teilchens (3.14.26) zu addieren, freilich in quantisierter Form. Da $\vec{\mu} \propto \vec{s}$ ist und \vec{s} mit \vec{x} kommutiert, treten dabei keine Probleme mit der Operatorordnung auf.

3.14.5 Teilchen im homogenen Magnetfeld

TBD

3.15 Der Stern-Gerlach-Versuch

Der Stern-Gerlach-Versuch zur „Richtungsquantelung“ wurde 1922, also schon vor der Entwicklung der modernen Quantentheorie ausgeführt¹¹. Das Bohr-Sommerfeld-Atommodell sagte allerdings bereits damals voraus, dass der Drehimpuls quantisiert sein sollte. Außerdem war bekannt, dass ein Atom mit einer Elektronenkonfiguration in einem Zustand zum Bahndrehimpuls $l \neq 0$ (in moderner Sprache ausgedrückt) ein entsprechendes magnetisches Moment aufweisen muß. Der Spin war allerdings noch unbekannt. Gleichwohl sollte ein Atomstrahl mit Atomen mit magnetischem Moment in einem inhomogenen Magnetfeld abgelenkt werden. Während die klassische Physik einfach eine kontinuierliche Aufweitung des Strahls vorhersagte, mußte nach dem Bohrschen Atommodell der Strahl in einem Magnetfeld mit Feldgradient in z-Richtung entsprechend der Quantelung der z-Komponente des Bahndrehimpulses in diskreter Weise abgelenkt werden. Stern und Gerlach gelang der Nachweis dieser Richtungsquantelung eines Atomstrahls von Silberatomen. Sie fanden eine Aufspaltung des Strahls entsprechend zweier Einstellmöglichkeiten der z-Komponente des Drehimpulses, was mit der damaligen Vorhersage des Bohr-Sommerfeld-Modells verträglich war. Die moderne Quantentheorie würde allerdings bei einem Teilchen ohne Spin nur Aufspaltungen in eine ungerade Anzahl

¹¹Eine äußerst vergnügliche Darstellung der Historie des Stern-Gerlach-Versuchs findet sich in [FH03].

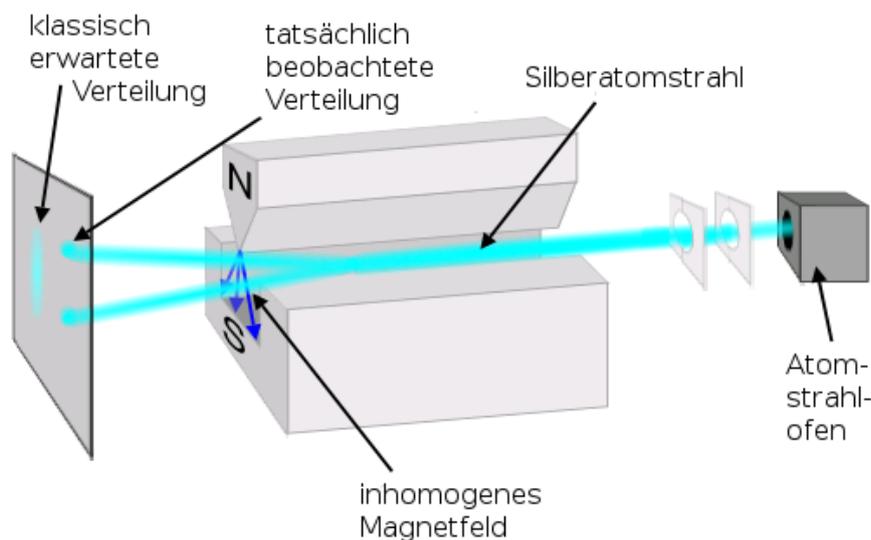


Abbildung 3.3: Schematischer Aufbau des Stern-Gerlach-Versuchs (Wikipedia: Theresa Knott - aus dem Englischen ins Deutsche übersetzt, CC BY-SA 3.0, <https://commons.wikimedia.org/w/index.php?curid=10344788>).

von Strahlen vorhersagen, denn der Bahndrehimpuls gehört immer zu Darstellungen der Drehgruppe mit Drehimpulsquantenzahl $l \in \mathbb{N}_0$, und die z -Komponente des Bahndrehimpulses kann immer nur die $(2l + 1)$ Werte $m_z = \{0, \pm\hbar, \dots, \pm l\hbar\}$ annehmen.

Im Lichte der modernen Quantentheorie betrachtet ist allerdings die Aufspaltung in zwei Strahlen für Silberatome leicht erklärbar: Das Silberatom besitzt ein Valenzelektron, das sich im Grundzustand in einem Bahndrehimpulszustand $l = 0$ (s -Orbital) befindet. Die übrigen Elektronen füllen ihre entsprechenden Orbitale vollständig auf, so dass deren Gesamtdrehimpuls 0 ist. Entsprechend ist der Gesamtdrehimpuls des Atoms $J = 1/2$. Es verhält sich also wie ein neutrales Teilchen mit einer großen Masse und einem magnetischen Moment aufgrund des Spins $1/2$ dieses Elektrons, solange das Atom nicht auf irgendeine Weise in angeregte Zustände übergeht. Dies ist aber unter den Versuchsbedingungen von Stern und Gerlach nur höchst unwahrscheinlich. Daher können wir bei der Analyse des Experiments das Silberatom einfach als neutrales Teilchen mit einem magnetischen Moment entsprechend dem vom Spin des Valenzelektrons behandeln. Entsprechend der beiden Einstellungsmöglichkeiten für die z -Komponente des Spins ($\sigma = \pm 1/2$) spaltet sich also der Strahl in der Tat in zwei Teilstrahlen auf, von denen einer aus Teilchen mit $\sigma = \hbar/2$ und einer aus solchen mit $\sigma = -\hbar/2$ besteht.

Schon diese qualitative Beschreibung zeigt, warum der Stern-Gerlach-Versuch auch heute noch als Musterbeispiel für den quantenmechanischen Messprozeß dient: Er zeigt alle Charakteristika einer quantenmechanischen Messung, und wir wollen daher dieses Experiment aus quantentheoretischer Sicht genauer betrachten. Der Versuchsaufbau ist schematisch in Abb. 3.3 dargestellt. In einem Ofen wird Silber geschmolzen. Durch eine Öffnung tritt ein Atomstrahl aus und wird durch Blenden auf eine bestimmte Richtung fokussiert. Dies können wir im Sinne der Quantentheorie als **Präparation der Silberatome** auffassen. Die Atome durchlaufen nun ein **inhomogenes zeitlich konstantes Magnetfeld**, welches wir in der Nähe des Strahls durch die Entwicklung bis zur ersten Ordnung in den Raumkoordinaten approximieren können:

$$\vec{B} = (B_0 + \beta z)\vec{e}_z - \beta y\vec{e}_y. \quad (3.15.1)$$

Dieses Feld erfüllt offenbar die Maxwell-Gleichungen

$$\vec{\nabla} \cdot \vec{B} = 0, \quad \vec{\nabla} \times \vec{B} = 0 \quad (3.15.2)$$

3.15. Der Stern-Gerlach-Versuch

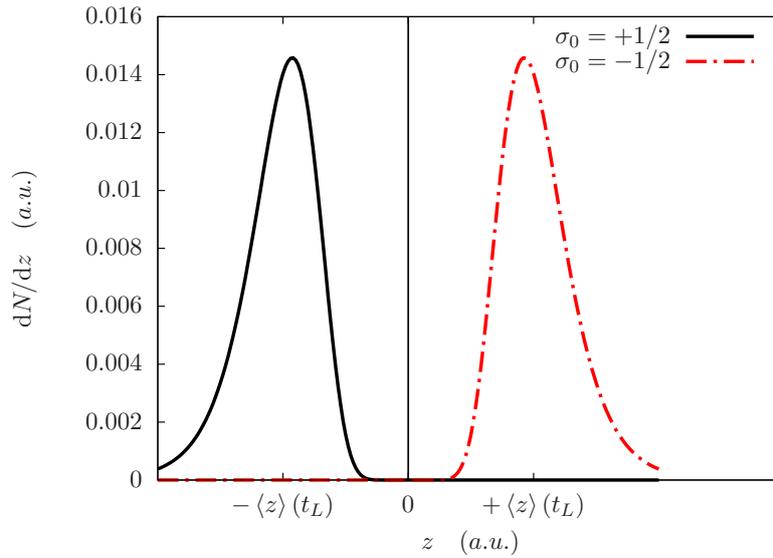


Abbildung 3.4: Wahrscheinlichkeitsverteilung für die z -Position eines Teilchens nach Durchlaufen eines inhomogenen Magnetfeldes bei $x = L$. Anfangs lag eine bei $\bar{x} = 0$ gepeakte Gaußverteilung vor, wobei die Teilchen mit gleicher Wahrscheinlichkeit eine Spinkomponente $\sigma = +1/2$ oder $\sigma = -1/2$ besitzen.

für ein statisches Magnetfeld in einem quellenfreien Raumbereich. Der Hamiltonoperator (3.14.41) vereinfacht sich im gegebenen Falle wegen $q = 0$ zu¹²

$$\mathbf{H} = \frac{\vec{\mathbf{p}}^2}{2M} + \mu_B g_s \vec{\mathbf{S}} \cdot \vec{\mathbf{B}}. \quad (3.15.3)$$

Setzen wir (3.15.1) ein, erhalten wir ein vom Spin abhängiges Potential, dessen klassisches Analogon einer konstanten Kraft entspricht. Um die Analyse weiter zu vereinfachen, nehmen wir an, daß $|B_0| \gg |\beta| |y|$ ist, d.h. der Strahl sei hinreichend in xy -Richtung um $x = 0$, $y = 0$ fokussiert. Dann können wir zunächst den einfacheren Hamiltonoperator

$$\mathbf{H}' = \frac{\vec{\mathbf{p}}^2}{2M} + \mu_B g_s (B_0 + \beta z) \mathbf{S}_z \quad (3.15.4)$$

verwenden. Betrachten wir Teilchen in einem \mathbf{S}_z -Eigenzustand, haben wir es also mit der Bewegung in einem konstanten Kraftfeld zu tun, und die Atome werden für $\sigma = \hbar/2$ nach unten, für $\sigma = -\hbar/2$ nach oben abgelenkt. Ein Atomstrahl, dessen Zustand durch eine beliebige Superposition aus solchen Eigenzuständen beschrieben wird, spaltet sich also entsprechend in zwei Teilstrahlen auf.

Falls die Strahlen in der yz -Ebene hinreichend fokussiert bleiben, bilden sich zwei wohlseparierte Teilstrahlen, die in der hier betrachteten Näherung aus reinen \mathbf{S}_z -Eigenzuständen bestehen, d.h. in diesen Teilstrahlen haben wir es mit Teilchen zu tun, die eine wohldefinierte z -Komponente des Spins besitzen. Vorher war diese Spinkomponente unbestimmt. Durch das inhomogene Magnetfeld können wir nun einen der Teilstrahlen ausfiltern und so Teilchen mit determinierter Spinkomponente präparieren. Der Zustand des Gesamtensembles wird aber immer noch durch eine Superposition beschrieben. Allerdings sind nach Durchlaufen des Magnetfeldes **Ort und Spin-3-Komponente** verschränkt, d.h. eine hinreichend genaue Ortsmessung liefert zugleich auch einen wohlbestimmten Spinzustand. Man kann also durch Nachweis eines Silberatoms am Schirm mit einer nahezu 100%-Wahrscheinlichkeit sagen, welchen Wert $\sigma = \pm \hbar/2$ die Spin-3-Komponente dieses Silberatoms besitzt. Voraussetzung dafür ist allerdings, dass der Ort der Teilchen zumindest in z -Richtung so scharf bestimmt ist, dass die beiden Teilstrahlen als wohlsepariert angesehen werden können. In diesem Zusammenhang nennt man die Ortskomponente z auch eine **Zeigervariable**, denn sie wird bei dem betrachteten

¹²Das Vorzeichen des Spinterms kehrt sich wegen der negativen Ladung des Elektrons um.

3. Wellenmechanik

Versuchsaufbau zur Messung der eigentlich interessierenden Observable, nämlich der Spin-3-Komponente, im Sinne des Zeigers eines Messgerätes verwendet.

Man kann die zeitabhängige Schrödinger-Gleichung mit dem Hamilton-Operator (3.15.4) relativ einfach lösen, allerdings mit Methoden, die wir in dieser Vorlesung aus Zeitgründen nicht genauer behandeln können (s. [Hee15]). Das Resultat bestätigt die oben gegebene qualitative Überlegung. In Abb. 3.4 ist das Ergebnis für $|\psi|^2$ für einen Anfangszustand

$$\psi_{\sigma}(t=0, \vec{x}) = \frac{c_{\sigma}}{(2\pi\Delta^2)^{3/4}} \exp\left(-\frac{\vec{x}^2}{4\Delta^2} + ip_0x\right), \quad |c_{1/2}|^2 + |c_{-1/2}|^2 = 1. \quad (3.15.5)$$

mit $c_{1/2} = c_{-1/2} = 1/\sqrt{2}$, also gleicher Wahrscheinlichkeit von $1/2 = 50\%$, dass das Silberatom anfangs eine Spin-z-Komponente mit $\sigma = \hbar/2$ oder $\sigma = -\hbar/2$ besitzt.

Kapitel 4

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 (1888), fortleiten sollte, ähnlich wie die Schwingungen der Luft Schallwellen entsprechen. Allerdings hatte diese Substanz sehr ungewöhnliche 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 ; in den in diesem Skript verwendeten SI-Einheiten ist $c = 1/\sqrt{\mu_0\epsilon_0}$. 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 (Albert Einstein, 1879-1955) 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.

4.1 Das speziell-relativistische Raum-Zeit-Modell

In seiner berühmten Arbeit „Zur Elektrodynamik bewegter Körper“ [Ein05] ging Einstein von zwei Postulaten aus:

1. **Spezielles Relativitätsprinzip:** 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. **Invarianz der Lichtgeschwindigkeit:** Die Ausbreitungsgeschwindigkeit **elektromagnetischer Wellen** (insbesondere des **Lichtes**) im Vakuum ist unabhängig von der Geschwindigkeit der Lichtquelle und des Empfängers (Beobachters) dieser Wellen relativ zum Bezugssystem.

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

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 bewegtes 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.

Einsteins wesentliche Idee war, dass die beiden Grundpostulate nur dadurch miteinander vereinbar sein können, wenn man statt des absoluten Raumes und der absoluten Zeit und dem entsprechenden **Galilei-Newton-schen Raum-Zeit-Modell** eine neue Beschreibung von Zeit und Raum einführt. Insbesondere kann offenbar nicht zugleich das spezielle Relativitätsprinzip und die Invarianz der Lichtgeschwindigkeit gelten, solange man von der Existenz einer **absoluten Zeit** wie in der Galilei-Newton'schen Mechanik ausgeht, d.h. bei der Transformation der Beschreibung eines physikalischen Vorgangs (z.B. die Trajektorie eines Massenpunktes mittels Zeit t und dem Ortsvektor $\vec{x}(t)$ dieses Teilchens), muss neben den Ortskoordinaten auch die Zeit transformiert werden, wenn man von einem Inertialsystem Σ zu einem anderen Inertialsystem Σ' übergeht, das sich gegenüber Σ mit der Geschwindigkeit \vec{v} bewegt. Dann kann man wegen der Absolutheit von Raum und Zeit im Rahmen des Galilei-Newton'schen Raum-Zeit-Modells die gleichen kartesischen Basisvektoren für beide Bezugssysteme verwenden, und die absolute Zeit ist ohnehin in beiden Bezugssystemen gleich (wobei wir hier davon ausgehen wollen, dass der Ursprung der Zeitachse gleich gewählt wird $t = t' = 0$ und zu dieser Zeit auch die Ursprünge der Bezugssysteme zusammenfallen sollen). Damit ergibt sich die Transformationsformel eines solchen **Galilei-Boosts** zu

$$t' = t, \quad \vec{x}' = \vec{x} - \vec{v}t \quad (\text{Galilei-Newton}). \quad (4.1.1)$$

Wir können nun mit den Postulaten die speziell-relativistische Raum-Zeit-Struktur (bzw. die entsprechend modifizierten Transformationsformeln für Raum- und Zeitkoordinaten eines Teilchens) leicht mittels eines Gedankenexperiments finden.

Dazu denken wir uns eine punktförmige Lichtquelle, die in einem Inertialsystem Σ 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

$$c^2 t^2 - \vec{x}^2 = 0. \quad (4.1.2)$$

Dabei gehen wir stillschweigend davon aus, dass hinsichtlich der Geometrie des Raumes für einen in einem Inertialsystem ruhenden Beobachter die Euklidischen Axiome gelten. Wir betrachten nun denselben Vorgang von einem weiteren Inertialsystem Σ' aus, das sich gegenüber Σ mit der konstanten Geschwindigkeit $\vec{v} = c\vec{\beta} = c\beta\vec{e}_1$ bewegt. Die Uhren und Koordinatenursprünge der beiden Systeme mögen zur Zeit $t = t' = 0$ zusammenfallen, und die räumlichen kartesischen Basisvektoren beider Bezugssysteme seien nicht gegeneinander verdreht. Gemäß dem zweiten Postulat muss nun auch bzgl. Σ' für die Raum-Zeit-Koordinaten des Lichtpulses (4.1.2) gelten, d.h. es gilt auch

$$c^2 t'^2 - \vec{x}'^2 = 0. \quad (4.1.3)$$

Wegen des speziellen Relativitätsprinzips muss nun die Transformation zwischen den Zeiten und der räumlichen Koordinaten bzgl. zweier Inertialsysteme 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 und zunächst **vierdimensionale Vektoren** (kurz **Vierervektoren** genannt) mit Komponenten $\underline{x} = (x^\mu) = (x^0, x^1, x^2, x^3)^T$ einzuführen, wobei $x^0 = ct$ ist. Dabei ist μ ein Index, den wir aus Gründen, die weiter unten klar werden, als oberen Index schreiben. Es handelt sich bei diesen Indizes also *nicht* um Potenzen! Solche Vierervektorkomponenten mit oberen Indizes schreiben wir auch als **Spaltenvektoren**.

Nun führen wir eine Bilinearform, das **Minkowski-Produkt**

$$\underline{x} \cdot \underline{y} = x^0 y^0 - \vec{x} \cdot \vec{y} \quad (4.1.4)$$

4.1. Das speziell-relativistische Raum-Zeit-Modell

ein, so dass Gleichung (4.1.2) in der Form

$$\underline{x} \cdot \underline{x} =: x^2 = 0 \quad (4.1.5)$$

geschrieben werden kann.

Dies ist einem Skalarprodukt recht ähnlich, nur dass wegen der Vorzeichen in der Definition (4.1.4) des Minkowski-Produkts $\underline{x} \cdot \underline{x}$ auch negativ werden kann, d.h. die Bilinearform ist **nicht positiv definit**.

Führen wir nun eine Diagonalmatrix $\hat{\eta} = \text{diag}(1, -1, -1, -1)$ mit den Komponenten

$$\eta_{\mu\nu} := \begin{cases} 1 & \text{for } \mu = \nu = 0, \\ -1 & \text{for } \mu = \nu, \quad \mu, \nu \in \{1, 2, 3\}, \\ 0 & \text{for } \mu \neq \nu, \quad \mu, \nu \in \{0, 1, 2, 3\} \end{cases} \quad (4.1.6)$$

ein, können wir das das **Minkowski-Produkt** in der Form

$$\underline{x} \cdot \underline{y} = \eta_{\mu\nu} x^\mu y^\nu \quad (4.1.7)$$

schreiben. Dabei ist es für das Folgende sehr wichtig zu bemerken, dass die Indizes bei $\hat{\eta}$ als *untere Indizes* notiert sind. Weiter verwenden wir hier und im Folgenden wieder die **Einsteinsche Summenkonvention**, wonach über doppelt auftretende Indizes stillschweigend summiert wird, wobei die griechischen Indizes stets die Werte $\{0, 1, 2, 3\}$ annehmen. Dabei muss bei der hier eingeführten Konvention mit oberen und unteren Indizes bei solchen doppelt vorkommenden Indizes *immer* einer oben und einer unten stehen.

Wir betrachten nun die Transformation zwischen den beiden oben spezifizierten Inertialsystemen Σ und Σ' , einen sog. **Lorentz-Boost** in x^1 -Richtung (Hendrik Antoon Lorentz, 1853-1928). Es ist klar, dass in diesem Fall $x'^2 = x^2$ und $x'^3 = x^3$ ist und wir daher nur die Transformation in der (01)-Ebene zu betrachten brauchen. Die Forderung, dass für die Wellenfront der elektromagnetische Kugelwelle sowohl (4.1.2) als auch (4.1.3) gelten soll, ist sicher garantiert, wenn für *alle* Vierervektoren \underline{x} und \underline{y} gilt

$$\underline{x}' \cdot \underline{y}' = \eta_{\mu\nu} x'^\mu y'^\nu = \eta_{\mu\nu} x^\mu y^\nu = \underline{x} \cdot \underline{y} \quad (4.1.8)$$

Sei nun $\hat{\Lambda}$ die entsprechende $\mathbb{R}^{4 \times 4}$ -Transformationsmatrix mit

$$\underline{x}' = \hat{\Lambda} \underline{x} \Leftrightarrow x'^\mu = \Lambda^\mu_\nu x^\nu. \quad (4.1.9)$$

In der Matrix-Vektorschreibweise ist nun

$$\underline{x}' \cdot \underline{y}' = \underline{x}'^T \hat{\eta} \underline{y}' = (\hat{\Lambda} \underline{x})^T \hat{\eta} (\hat{\Lambda} \underline{y}) = \underline{x}^T (\hat{\Lambda}^T \hat{\eta} \hat{\Lambda}) \underline{y}. \quad (4.1.10)$$

Nun soll die Matrix $\hat{\Lambda}$ so beschaffen sein, dass diese Gleichung für *alle* Vierervektoren \underline{x} und \underline{y} gilt. Dies ist aber nur möglich, wenn die Matrix in der Klammer die Matrix $\hat{\eta}$ ist. Es muss also

$$\hat{\Lambda}^T \hat{\eta} \hat{\Lambda} = \hat{\eta} \quad (4.1.11)$$

sein.

Betrachten wir dies nun für unseren Fall, dass die Transformation nur Komponenten in der (01)-Ebene ändert. Dann können wir zwischenzeitlich mit zweidimensionalen Vektoren bzw. 2×2 -Matrizen rechnen. Schreiben wir für die unbekannte Matrix Λ

$$\hat{\Lambda} = \begin{pmatrix} A & B \\ C & D \end{pmatrix} \quad (4.1.12)$$

wird (4.1.11) zu (*Nachrechnen*)

$$\begin{pmatrix} A & C \\ B & D \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} A & B \\ C & D \end{pmatrix} = \begin{pmatrix} A & -C \\ B & -D \end{pmatrix} \begin{pmatrix} A & B \\ C & D \end{pmatrix} = \begin{pmatrix} A^2 - C^2 & AB - CD \\ AB - CD & B^2 - D^2 \end{pmatrix} \stackrel{!}{=} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \quad (4.1.13)$$

4. Spezielle Relativitätstheorie

Die Matrixelemente A, B, C und D müssen also die Gleichungen

$$A^2 - C^2 = 1, \quad AB - CD = 0, \quad D^2 - B^2 = 1 \quad (4.1.14)$$

erfüllen. Es ist nun klar, dass $\hat{\Lambda}$ bzw. die vier Matrixelemente nur von einem reellen Parameter abhängen, der eindeutig mit der Relativgeschwindigkeit $v = c\beta \in \mathbb{R}$ zusammenhängt. Wir erinnern uns nun daran, dass für die Hyperbelfunktionen

$$\cosh \alpha = \frac{1}{2}[\exp(\alpha) + \exp(-\alpha)], \quad \sinh \alpha = \frac{1}{2}[\exp(\alpha) - \exp(-\alpha)] \quad (4.1.15)$$

die Beziehung

$$\cosh^2(\alpha) - \sinh^2(\alpha) = 1 \quad (4.1.16)$$

gilt. Wir setzen nun also

$$A = D = \cosh \alpha, \quad B = C = -\sinh \alpha, \quad (4.1.17)$$

wobei sich das zusätzliche Minuszeichen im letzten Ausdruck gleich als nützlich herausstellen wird. In der Tat sind dann alle Bedingungsgleichungen (4.1.14) erfüllt (*Nachrechnen!*). Die Transformationsmatrix lautet also

$$\hat{\Lambda} = \begin{pmatrix} \cosh \alpha & -\sinh \alpha \\ -\sinh \alpha & \cosh \alpha \end{pmatrix}. \quad (4.1.18)$$

Um dies physikalisch interpretieren zu können, schreiben wir auch (4.1.9) explizit hin:

$$\underline{x}' = \begin{pmatrix} ct' \\ x'^1 \end{pmatrix} = \hat{\Lambda} \underline{x} = \begin{pmatrix} ct \cosh \alpha - x^1 \sinh \alpha \\ -ct \sinh \alpha + x^1 \cosh \alpha \end{pmatrix}. \quad (4.1.19)$$

Um nun die Beziehung des Parameters α zur Relativgeschwindigkeit $v = \beta c$ zwischen den Inertialsystemen Σ und Σ' zu bestimmen, betrachten wir den Ursprung von Σ' , $x'^1 = 0$. Gemäß (4.1.19) ergibt sich für diesen Punkt, von Σ aus betrachtet:

$$x^1 \cosh \alpha - ct \sinh \alpha = 0 \Rightarrow x^1 = \frac{\sinh \alpha}{\cosh \alpha} ct = ct \tanh \alpha \stackrel{!}{=} vt = \beta ct. \quad (4.1.20)$$

Damit ist aber die gesuchte Beziehung zwischen α und $v = \beta c$ gefunden:

$$\beta = \tanh \alpha. \quad (4.1.21)$$

Zunächst ist daran bemerkenswert, dass stets $|\beta| < 1$ bzw. $v < c$ sein muss, d.h. zwei Inertialsysteme sind nur dann konsistent mit den beiden Einsteinschen Postulaten, wenn die **Relativgeschwindigkeit** zwischen den Bezugssystemen kleiner als die Vakuumlichtgeschwindigkeit ist.

Nun wollen wir die Transformationsgleichungen (4.1.19) noch mit Hilfe der Relativgeschwindigkeit $v = c\beta$ ausdrücken. Nun gilt

$$\begin{aligned} \tanh^2 \alpha &= \frac{\sinh^2 \alpha}{\cosh^2 \alpha} = \frac{\cosh^2 \alpha - 1}{\cosh^2 \alpha} \Rightarrow \cosh^2 \alpha = \frac{1}{1 - \tanh^2 \alpha} \\ \Rightarrow \cosh \alpha &= \frac{1}{\sqrt{1 - \tanh^2 \alpha}} = \frac{1}{\sqrt{1 - \beta^2}} = \gamma. \end{aligned} \quad (4.1.22)$$

Dabei haben wir im letzten Schritt verwendet, dass $\cosh \alpha \geq 1 > 0$ für alle $\alpha \in \mathbb{R}$ ist und (4.1.21) gilt. Daraus folgt sofort

$$\sinh \alpha = \tanh \alpha \cosh \alpha = \beta \gamma = \frac{\beta}{\sqrt{1 - \beta^2}}. \quad (4.1.23)$$

4.1. Das speziell-relativistische Raum-Zeit-Modell

Die Transformation zwischen zwei Inertialsystemen, die sich mit einer Relativgeschwindigkeit $v = \beta c$ bewegen, lautet also gemäß (4.1.19) mit (4.1.22) und (4.1.23)

$$x'^0 = ct' = \gamma(x^0 - \beta x^1) = \frac{ct - \beta x}{\sqrt{1 - \beta^2}}, \quad x'^1 = \gamma(x^1 - \beta x^0) = \frac{x^1 - vt}{\sqrt{1 - \beta^2}}. \quad (4.1.24)$$

Die entsprechende Transformationsmatrix (4.1.18) ergibt sich gemäß (4.1.22) und (4.1.23), jetzt wieder ausführlich für alle vier Raum-Zeit-Komponenten geschrieben, zu

$$\hat{\Lambda} = (\Lambda^\mu{}_\nu) = \begin{pmatrix} \gamma & -\beta\gamma & 0 & 0 \\ -\beta\gamma & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}. \quad (4.1.25)$$

Die entsprechende Umkehrtransformation erhalten wir sehr einfach durch Anwendung von (4.1.11). Dazu bemerken wir, dass $\hat{\eta}^2 = \mathbb{1}$ ist. Multiplizieren wir also (4.1.11) von links mit $\hat{\eta}$, folgt

$$\hat{\eta} \hat{\Lambda}^T \hat{\eta} \hat{\Lambda} = \mathbb{1} \Rightarrow \hat{\Lambda}^{-1} = \hat{\eta} \hat{\Lambda}^T \hat{\eta} = \begin{pmatrix} \gamma & \beta\gamma & 0 & 0 \\ \beta\gamma & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}. \quad (4.1.26)$$

Die Umkehrtransformation zu (4.1.24) ist also durch

$$\underline{x} = \hat{\Lambda}^{-1} \underline{x}' \Rightarrow ct = \gamma(ct' + \beta x'^1), \quad x^1 = \gamma(x'^1 + \beta ct) \quad (4.1.27)$$

gegeben. Das entspricht einem Lorentz-Boost mit $\vec{v} = -v\vec{e}_1$. Dies leuchtet auch physikalisch sofort ein: Wenn sich Σ' gegenüber Σ mit der Geschwindigkeit $+\vec{v}$ bewegt, bewegt sich umgekehrt Σ gegenüber Σ' mit der Geschwindigkeit $-\vec{v}$.

Man kann nun, zumindest für den Fall, dass man es nur mit Bewegungen entlang einer Achse zu tun hat, die Minkowski-Geometrie in **Minkowski-Diagrammen** veranschaulichen. Man muss dabei nur bedenken, dass die Zeichenebene nun nicht mehr die gleiche Bedeutung wie in der gewöhnlichen euklidischen Geometrie besitzt. Vielmehr wird die Geometrie durch das nicht positiv definite Minkowskiprodukt bestimmt. Wie in Abb. 4.1 gezeigt, wird die Zeitachse des Systems Σ , also die $x^0 = ct$ -Achse, nach oben und die räumliche Achse x^1 nach rechts aufgetragen. Um die Koordinatenachsen des Systems Σ' zu konstruieren, können wir den Lorentz-Boost (4.1.24) bzw. die entsprechende Umkehrtransformation (4.1.27) verwenden: Der Koordinatenursprung von Σ' ist durch $x'^1 = 0$ gegeben. Im Minkowski-Diagramm entspricht der Koordinatenursprung einer Weltlinie, die durch die Zeit t' im System Σ' parametrisiert wird. Gemäß (4.1.24) ist diese Weltlinie durch

$$ct = \gamma ct', \quad x^1 = \gamma \beta ct' = \beta ct \quad (4.1.28)$$

gegeben. Entsprechend ist die x'^1 -Achse durch $ct' = 0 = \text{const}$ gegeben. Die entsprechenden Koordinaten in Σ sind wieder durch (4.1.27) gegeben:

$$x^1 = \gamma x'^1, \quad ct = \gamma \beta x'^1 = \beta x'^1 \quad (4.1.29)$$

gegeben. Dies ist offenbar genau die an der Winkelhalbierenden $ct = x^1$, die die Wellenfront eines zur Zeit $t = 0$ ausgesandten Lichtimpulses in x^1 -Richtung beschreibt und daher als **Lichtkegel** bezeichnet wird, gespiegelte ct' -Achse (4.1.28). Das muss auch so sein, denn dadurch wird sichergestellt, dass die Weltlinie eines Lichtsignals auch in Σ' durch $x'^1 = ct'$ beschrieben wird. Dies entspricht aber genau dem Einsteinschen Postulat von der Konstanz der Lichtgeschwindigkeit, aus dem wir ja die Lorentz-Boost-Transformation (4.1.24) bzw. (4.1.27) hergeleitet haben.

4. Spezielle Relativitätstheorie

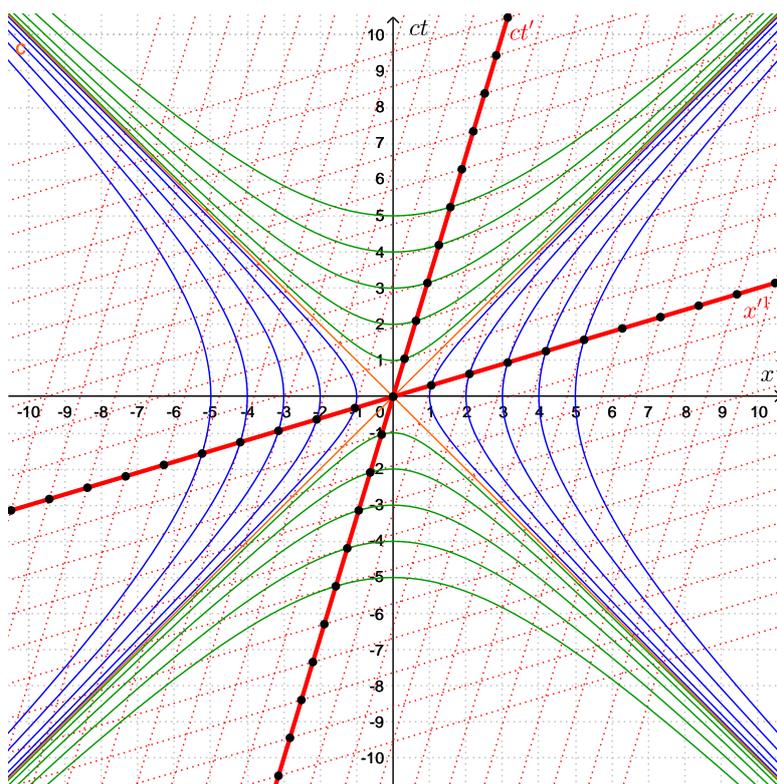


Abbildung 4.1: Minkowski-Diagramm, das die Konstruktion der Raum-Zeit-Achsen des Systems Σ' illustriert: Die dicken roten Linien sind entsprechend die ct' und x'^1 -Achsen. Die Einheitslängen werden auf den Achsen durch die Schnittpunkte mit den grünen bzw. blauen Hyperbeln $(ct)^2 - (x^1)^2 = \pm(n \text{ Ls})^2$ mit $n \in \{1, 2, \dots\}$ bestimmt. Die orangenen Linien $ct = \pm x^1$ definieren die Ausbreitung von Lichtsignalen in $\pm x^1$ - und auch $\pm x'^1$ -Richtung. Sie sind die Winkelhalbierende zwischen den Zeit- und Raum-Achsen sowohl in Σ als auch in Σ' d.h. für beide Beobachter bewegt sich die Front eines Lichtsignals mit derselben Geschwindigkeit c , wie es vom 2. Postulat Einsteins verlangt wird.

Es bleibt nun noch zu klären, wie die Einheiten der Zeit- und der räumlichen Achse in Σ' zu wählen sind. Die geeignete Längeneinheit für unsere Zwecke ist die **Lichtsekunde** $1 \text{ Ls} = c \cdot 1 \text{ s} = 299\,792\,458 \text{ m}$. Wir haben nun die Lorentz-Transformation so bestimmt, dass Minkowski-Produkte beliebiger Vierervektoren sich nicht ändern. Entsprechend ergibt sich die Längeneinheit auf der ct' -Achse daraus, dass der Punkt $ct' = 1 \text{ Ls}$, $x'^0 = 0$ auf der entsprechenden **Einheitshyperbel**

$$(ct)^2 - (x^1)^2 = 1 \text{ Ls}^2 \quad (4.1.30)$$

liegt, d.h. im Minkowski-Diagramm durch den entsprechenden Schnittpunkt der ct' -Achse mit dieser Hyperbel gegeben ist.

Genauso konstruiert man die Längeneinheit auf der räumlichen x'^1 -Achse. Der Punkt $ct' = 0$, $x'^1 = 1 \text{ Ls}$ muss nämlich offenbar auf der Hyperbel

$$(ct)^2 - (x^1)^2 = -1 \text{ Ls}^2 \quad (4.1.31)$$

liegen, und der Schnittpunkt mit der oben konstruierten x'^1 -Achse bestimmt folglich die Längeneinheit. In der Abb. 4.1 sind zur Verdeutlichung auch noch die entsprechenden zeit- und raumartigen blauen Hyperbeln für die Abstände 2 Ls, 3 Ls usw. und das entsprechende Koordinatennetz für Σ' eingezeichnet.

4.2. Kinematische Effekte

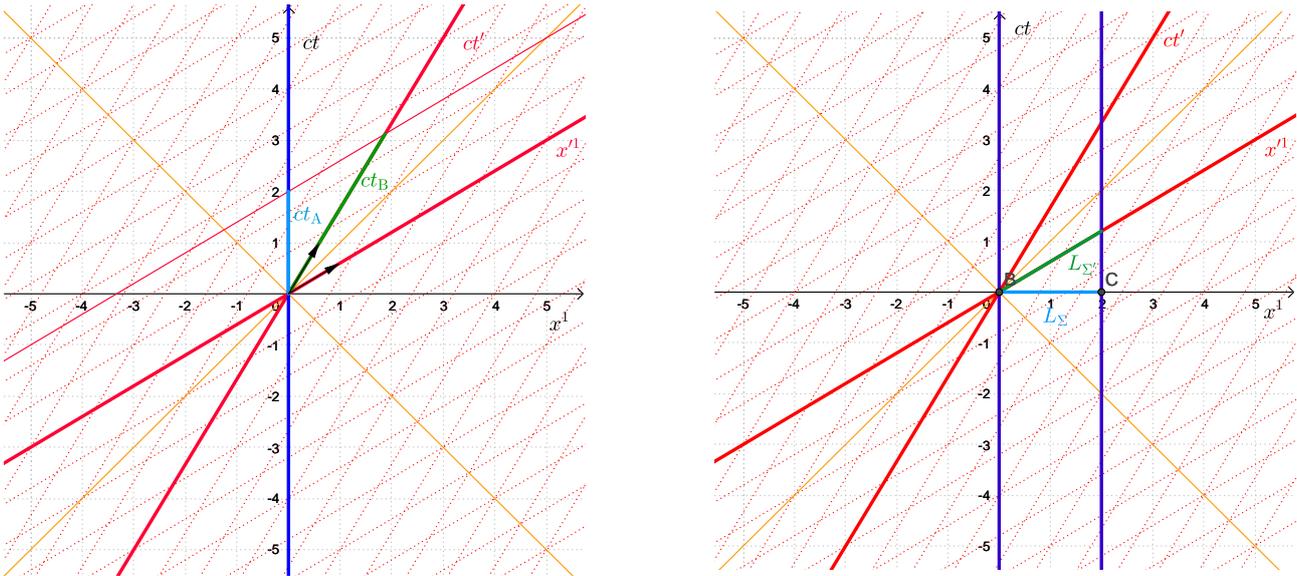


Abbildung 4.2: **Links:** Minkowski-Diagramm zur Zeitdilatation. Die Zeit, die zwischen den Endpunkten der hellblau eingezeichneten Weltlinie bzgl. Σ am gleichen Ort $x^1 = 0$ vergeht (entsprechend zwei Längeneinheiten ct), erscheint bzgl. Σ' länger als bzgl. Σ . Die Rechnung (s. Text) ergibt, dass $ct' = ct\gamma > ct$ ist. **Rechts:** Minkowski-Diagramm zur Längenkontraktion: Die blauen Linien sind die Weltlinien der Endpunkte der Stange, die relativ zu Σ ruht. Die in Σ ruhende Beobachterin Alice misst die Länge, indem sie zu ihrer Zeit $t = 0$ die Koordinaten der Stangenenden auf ihrer x^1 -Achse abliest. In Σ' Bezugssystem bewegt sich die Stange, und der bzgl. Σ' ruhende Beobachter Bob misst dessen Länge, indem er seine x'^1 -Koordinaten der Endpunkte zur für ihn gleichen Zeit $t' = 0$ markiert. Entsprechend misst er eine kürzere Länge als Alice. Dabei muss man beachten, dass die Längen für Alice und Bob 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. 4.1 konstruiert, also in den hier gezeigten Diagrammen durch die gepunkteten Koordinatenlinien. Offensichtlich besitzt für Alice die Stange eine Länge L_Σ von 2 Längeneinheiten und für Bob eine kürzere Länge $L_{\Sigma'}$. Wie die Rechnung im Text zeigt, ist die Länge gerade um den inversen Lorentz-Faktor $1/\gamma = \sqrt{1 - \beta^2}$ verkürzt.

4.2 Kinematische Effekte

Entsprechend des gegenüber der Newtonschen Physik geänderten Raum-Zeit-Modells ergeben sich einige der Intuition zuwiderlaufende Folgerungen, die wir in diesem Abschnitt besprechen wollen.

4.2.1 Relativität der Gleichzeitigkeit

Zuerst stellen wir fest, dass die **Gleichzeitigkeit von Ereignissen** vom Bezugssystem abhängt (**Relativität der Gleichzeitigkeit**). Zwei Ereignisse \underline{x} und \underline{y} sind definitionsgemäß im Inertialsystem Σ gleich, wenn $x^0 = y^0$ bzw. $x^0 - y^0 = 0$ gilt. Mittels des Lorentz-Boosts (4.1.24) folgt für den zeitlichen Abstand dieser Ereignisse bzgl. Σ'

$$x'^0 - y'^0 = \gamma[(x^0 - y^0) - \beta(x^1 - y^1)] = -\gamma\beta(x^1 - y^1). \quad (4.2.1)$$

D.h. bzgl. Σ' finden die beiden Ereignisse i.a. *nicht* gleichzeitig statt.

4.2.2 Zeitdilatation

Als nächstes betrachten wir eine in Σ ruhende Uhr. Deren Weltlinie ist durch $(x^0, x^1) = (ct, 0)$ gegeben. Ein bzgl. Σ' ruhender Beobachter beschreibt dieselbe Weltlinie gemäß der Lorentztransformation (4.1.27) durch

$$ct' = \gamma ct, \quad x'^1 = \gamma \beta ct = \beta ct' = vt'. \quad (4.2.2)$$

Abgesehen davon, dass sich selbstverständlich die Uhr bzgl. Σ' mit der Geschwindigkeit $v = \beta c$ in x'^1 -Richtung bewegt, liest der in Σ' ruhende Beobachter auch eine andere Zeit ab. Ist im Ruhssystem der Uhr Σ die Zeit $t_A = 1$ Ls vergangen, ist bzgl. Σ' gemäß (4.2.2) die längere Zeit $t' = \gamma$ Ls vergangen. Die Zeit in Σ' erscheint also gegenüber der Zeit im Ruhssystem der Uhr Σ also um den **Lorentz-Faktor** $\gamma = 1/\sqrt{1-\beta^2}$ gedehnt. Dieses Phänomen nennt man **Zeitdilatation**.

4.2.3 Längenkontraktion

Als nächstes betrachten wir eine Stange entlang der x^1 -Richtung, die in Σ ruht. Bzgl. der Koordinaten in Σ sind die Weltlinien der Endpunkte der Stange durch

$$\underline{x} = \begin{pmatrix} \lambda_1 \\ x^1 \end{pmatrix}, \quad \underline{y} = \begin{pmatrix} \lambda_2 \\ y^1 \end{pmatrix}, \quad \lambda_1, \lambda_2 \in \mathbb{R} \quad (4.2.3)$$

gegeben. Dabei sind λ_1 und λ_2 beliebige Parameter dieser Weltlinien. Ein in Σ ruhender Beobachter misst die Länge der Stange, indem er die räumlichen Koordinaten zur gleichen Zeit t abliest, d.h. für $\lambda_1 = \lambda_2$. Die Relativkoordinaten der Enden für diese „Messereignisse“ der Stange und die entsprechende Länge sind dann

$$\underline{y} - \underline{x} = \begin{pmatrix} 0 \\ y^1 - x^1 \end{pmatrix} \Rightarrow L_\Sigma = |y^1 - x^1|. \quad (4.2.4)$$

Für einen bzgl. Σ' ruhenden Beobachter bewegt sich die Stange in negative 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. Mit der Lorentztransformation (4.1.24) folgt aus (4.2.3)

$$\underline{y}' - \underline{x}' = \gamma \begin{pmatrix} (\lambda_2 - \lambda_1) - \beta(y^1 - x^1) \\ y^1 - x^1 - \beta(\lambda_2 - \lambda_1) \end{pmatrix}. \quad (4.2.5)$$

Nun liest der in Σ' ruhende Beobachter zur Längenmessung der Stange deren Anfangs- und Endpunkt zur gleichen Zeit $t' = 0$ bzgl. Σ' ab, d.h.

$$\lambda_2 - \lambda_1 - \beta(y^1 - x^1) = 0 \Rightarrow \lambda_2 - \lambda_1 = \beta(y^1 - x^1) = \beta L_\Sigma. \quad (4.2.6)$$

Der räumliche Abstand ergibt sich nun, indem wir dies in die räumliche Komponente von (4.2.5) einsetzen

$$L_{\Sigma'} = |y'^1 - x'^1| = \gamma(1 - \beta^2)|y^1 - x^1| = \sqrt{1 - \beta^2} L_\Sigma = \frac{1}{\gamma} L_\Sigma. \quad (4.2.7)$$

Bzgl. Σ' erscheint also der sich dort mit der Geschwindigkeit $v = \beta c$ bewegende Stab um den Faktor $1/\gamma = \sqrt{1-\beta^2}$ verkürzt. Dieses Phänomen bezeichnet man als **Lorentz-FitzGerald-Kontraktion** (George Francis FitzGerald, 1851-1901). Lorentz und FitzGerald haben nämlich bereits vor Einstein das Nullresultat des Michelson-Morley-Experiment mit der Annahme erklärt, dass sich Abmessungen von Körpern in Richtung der Bewegung durch den Äther um den Faktor $1/\gamma$ verkürzen. Mit Einsteins Interpretation einer geänderten Raumzeitbeschreibung hat sich die Annahme eines ausgezeichneten Bezugssystems jedoch erübrigt, und die entsprechende Verkürzung ergibt sich als rein kinematischer Effekt, der aus der Relativität der Gleichzeitigkeit resultiert: Die beiden Messereignisse des Beobachters zur Längenbestimmung des Stabes in dessen

4.2. Kinematische Effekte

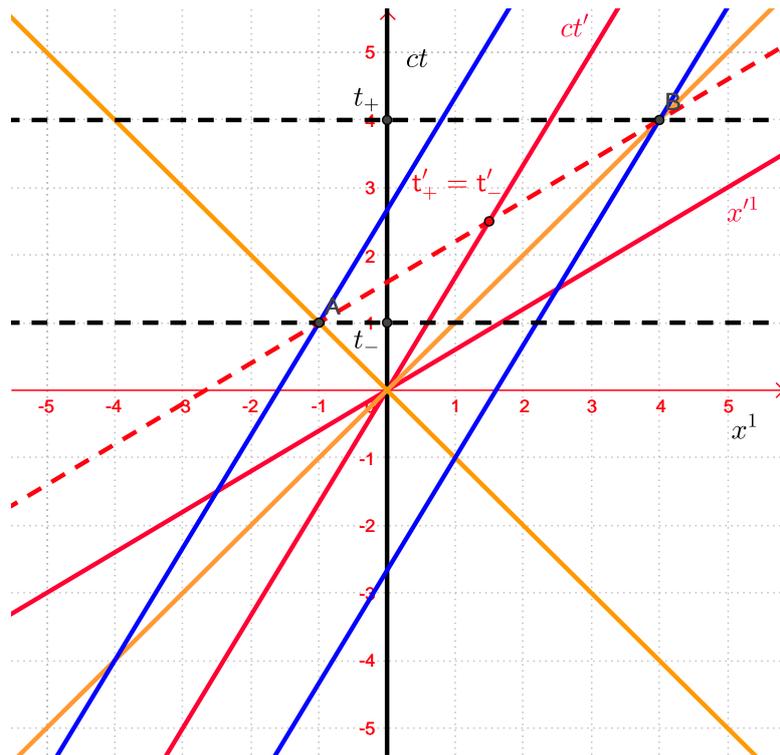


Abbildung 4.3: Zum Einsteinschen Zug-Gedankenexperiment: Die ct' -Achse ist die Weltlinie von Alice, die sich in der Mitte des Zuges befindet. Die Zugenden beschreiben die beiden blauen Weltlinien. Die Weltlinien der Lichtsignale, die Alice zur Zeit t aussendet treffen bei den Ereignissen A und B bei den Zugenden ein. Bzgl. Alices Bezugssystem geschieht dies gleichzeitig bei $t'_+ = t'_-$, wie die zur x'^1 -Achse parallele rot-gestrichelte Linie durch die Ereignisse A und B zeigt. Bei Bob treffen die Lichtsignale hingegen nicht gleichzeitig ein, wie die durch die gestrichelten schwarzen Linien gegebenen Zeiten t_{\pm} zeigen.

Ruhsystem Σ erfolgen bzgl. Σ gleichzeitig, nicht aber bzgl. Σ' . Daher wird ein Beobachter in Σ' zu einem anderen Ergebnis bei der Längenmessung kommen. Denn dieser Beobachter liest die räumlichen Koordinaten von Anfangs- und Endpunkt zur gleichen Zeit bzgl. Σ' ab. Diese Messereignisse wird umgekehrt ein in Σ ruhender Beobachter nicht als gleichzeitig beobachten, und die in Σ' abgelesene Länge entspricht daher nicht der bzgl. Σ gemessenen Länge. Die Relativität der Gleichzeitigkeit verhindert also, dass hier kein Widerspruch entsteht.

4.2.4 Einsteins Zug-Gedankenexperiment

Wir betrachten noch ein weiteres berühmtes Gedankenexperiment von Einstein im Zusammenhang mit der Relativität der Gleichzeitigkeit. Wir betrachten dazu einen Zug, der sich bzgl. des Inertialsystems Σ , das Ruhssystem des Bahnhofs, mit der konstanten Geschwindigkeit $\vec{v} = \beta c \vec{e}_1$ bewegt. Das Ruhssystem des Zuges ist also Σ' . 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. Zur Zeit $t' = t = 0$ befindet sich Alice auf der Höhe des auf dem Bahnsteig stehenden Beobachter Bob. Die Weltlinien der beiden Lichtsignale bzgl. Alices Raumzeit-Koordinaten (parametrisiert mit Alices Zeit t') lauten

$$\underline{x}'_{\pm} = \begin{pmatrix} ct' \\ \pm ct' \end{pmatrix}. \quad (4.2.8)$$

Besitzt der Zug in seinem Ruhssystem die Länge L' , kommt demnach das Lichtsignal bzgl. Alice *gleichzeitig* zur Zeit $t' = L'/(2c)$ bei den beiden Enden des Zuges an.

4. Spezielle Relativitätstheorie

Betrachten wir nun die Weltlinien bzgl. des am Bahnhof stehenden Beobachters Bob, der also bzgl. Σ ruht. Dazu müssen wir den entsprechenden Lorentz-Boost (4.1.27) anwenden:

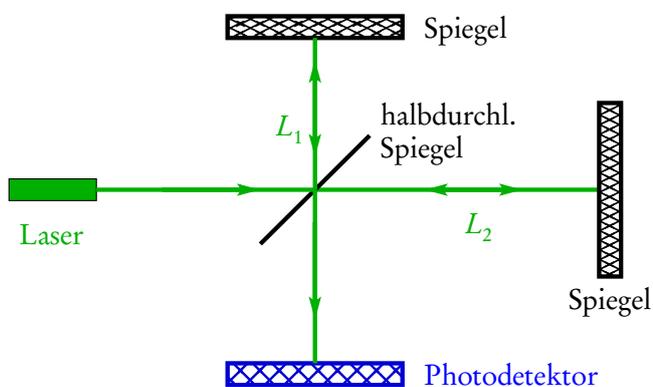
$$\underline{x}_{\pm} = \begin{pmatrix} \gamma & \beta\gamma \\ \beta\gamma & \gamma \end{pmatrix} \begin{pmatrix} ct' \\ \pm ct' \end{pmatrix} = \gamma ct' \begin{pmatrix} 1 \pm \beta \\ 1 \pm \beta \end{pmatrix} = \gamma \frac{L'}{2} \begin{pmatrix} 1 \pm \beta \\ 1 \pm \beta \end{pmatrix}. \quad (4.2.9)$$

Die beiden Lichtsignale erreichen also von Bob aus betrachtet die Enden des Zuges zu den Zeiten

$$t_{\pm} = \frac{\gamma(1 \pm \beta)L'}{2c}. \quad (4.2.10)$$

Dies erklärt sich dadurch, dass sich von Bob aus betrachtet das Lichtsignal 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 bzw. entgegen der Ausbreitungsrichtung des jeweiligen Lichtsignals bewegt, wobei aber bzgl. Bob diese Signale sich (wie für Alice) mit der Lichtgeschwindigkeit c ausbreiten.

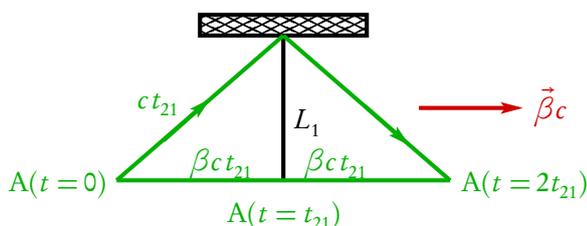
4.2.5 Das Michelson-Morley-Experiment



Das **Michelson-Morley-Experiment** ist eines der wesentlichen Schlüsselexperimente zur speziellen Relativitätstheorie. Die ursprüngliche Motivation zu seiner Durchführung war der Nachweis eines absoluten Bezugssystems, das als **Ruhsystem des Äthers** gedacht wurde. Dies beruhte auf der durch die Relativitätstheorie widerlegte Annahme, dass die Maxwell-Gleichungen, also die physikalischen Grundgesetze der elektromagnetischen und optischen Phänomene, in der üblichen Form nur in einem bestimmten Inertialsystem gelten sollten, in dem der Äther ruht. Dabei war der Äther eine hypothetische Substanz, deren Schwingungen elektromagnetischen Wellen entsprechen in analogem Sinne wie z.B. die Schwingungen der Luft Schallwellen entsprechen.

Allerdings waren die Eigenschaften dieses Äthers schon für die damaligen Zeitgenossen sehr rätselhaft. Wir wollen uns nicht näher mit dieser nunmehr widerlegten Theorie beschäftigen. Für das Verständnis des Michelson-Morley-Experiments genügt es zu wissen, dass die Äthertheorie in ihrer ursprünglichen Form davon ausgeht, dass (a) ein bevorzugtes Inertialsystem existiert, in dem der Äther ruht und das mit **Newtons absolutem Raum** identifiziert wurde und (b) das Galilei-Newtonsche Raum-Zeitmodell exakt gilt. Hinsichtlich der Theorie der elektromagnetischen Wellen genügt es zu wissen, dass deren **Vakuumphasengeschwindigkeit** nur im Ätherruhsystem c ist und in allen anderen Inertialsystemen entsprechend der Galileische Geschwindigkeitsaddition gegeben sein soll.

Analyse des Experiments im Rahmen der Äthertheorie



Wir betrachten nun den Aufbau des Michelson-Morley-Experiments. Es ist in der obigen Abbildung in seinem Ruhsystem dargestellt. Im Rahmen der Äthertheorie, für die ja wie oben gesagt die Galilei-Newton-Raumzeit als gültig angenommen wird, ist dies freilich auch die Anordnung zu einem jeden festen Zeitpunkt, wo sich die Apparatur bewegt. In der Tat nahmen Michelson und Morley

an, dass das Ätherruhsystem durch das Ruhsystem des Erde-Sonnensystems gegeben sei. Da das Experiment auf der Erde aufgebaut ist, bewegt es sich relativ zum Ätherruhsystem mit der Geschwindigkeit des Umlaufs der Erde um die Sonne (die viel kleinere Geschwindigkeit aufgrund der Erdrotation können wir hier vernachlässigen). Diese Geschwindigkeit beträgt etwa $v = 3 \cdot 10^4 \text{ m/s}$, also $\beta = v/c = 10^{-4}$.

4.2. Kinematische Effekte

Betrachten wir aber zunächst das Experiment im Ruhssystem des Äthers. Dann können wir die obige Abbildung verwenden. Es läuft also von rechts ein Laserstrahl auf den halbdurchlässigen Spiegel, der als Strahlteiler dient. Der eine Teilstrahl wird reflektiert und trifft auf den oberen Spiegel, wo er wieder reflektiert wird und wieder zum Strahlteiler zurückläuft. Ein Teil läuft dann durch den Strahlteiler zum Photodetektor. Die Laufzeit dieses Teilstrahls vom Strahlteiler zum Spiegel und wieder zurück beträgt nun offenbar $t_1 = 2L_1/c$, denn im Ätherruhesystem bewegt sich ja ein Lichtsignal (oder hier, wo ebene Wellen zum Einsatz kommen, ein beliebiger Punkt fester Phase z.B. ein Wellenberg der ebenen Welle) mit c . Der zweite Teilstrahl bewegt sich durch den Strahlteiler hindurch, gelangt zum rechten Spiegel, wird dort wieder zum Strahlteiler zurückreflektiert und gelangt dann im Falle einer Reflexion am Strahlteiler zum Detektor. Die Gesamtlaufzeit vom Strahlteiler zum Spiegel und zurück beträgt nun $t_2 = 2L_2/c$. Die Laufzeitdifferenz zwischen den beim Photodetektor auftreffenden Teilstrahlen ist also $t_1 - t_2 = 2(L_1 - L_2)/c$. Diese Laufzeitdifferenz hat ein Interferenzmuster am Photodetektor zur Folge. Dreht man die Apparatur um 90° , vertauschen die Abstände L_2 und L_1 ihre Rollen, und es wird die Laufzeitdifferenz zwischen vertikalem und horizontalem Teilstrahl $t'_2 - t'_1 = 2(L_1 - L_2)/c$, d.h. das Interferenzmuster wird sich gegenüber der ursprünglichen Anordnung nicht ändern.

Betrachten wir nun das Experiment, wenn sich die Apparatur im erdfesten System befindet. Wir analysieren es jedoch auch diesmal im Ruhesystem des Äthers, d.h. in einem Bezugssystem, in dem sich die ganze Apparatur mit der Geschwindigkeit $v = \beta c$ in x^1 -Richtung bewegt. Die Geometrie ist im Rahmen des hier angewandten Galilei-Newtonschen Raumzeitmodells zu jedem Zeitpunkt gleich. Man muss nur berücksichtigen, dass für die Lichtwege die Bewegung der Spiegel zu berücksichtigen ist. Da wir uns im Ruhesystem des Äthers befinden ist die Lichtgeschwindigkeit immer noch c .

Betrachten wir nun erst den Teilstrahl, der sich vom Strahlteiler in x^1 -Richtung zum Spiegel bewegt. Dazu benötige es die Zeit t_{11} . Da sich in dieser Zeit der Spiegel um $\beta c t_{11}$ weiter bewegt hat, legt das Licht insgesamt den Weg $c t_{11} = L_2 + \beta c t_{11}$ zurück. Von dort gelangt es zurück zum Strahlteiler. Dafür benötige es die Zeit t_{12} . Da sich nun aber das Licht auf den Strahlteiler zubewegt, gilt für die Gesamtweglänge $c t_{12} = L_2 - \beta c t_{12}$. Insgesamt ist die Laufzeit vom Strahlteiler zum Spiegel und zurück für den sich in x^1 -Richtung hin- und herbewegenden Teilstrahl also

$$t_1 = t_{11} + t_{12} = \frac{L_2}{c(1+\beta)} + \frac{L_2}{c(1-\beta)} = \frac{2L_2}{c(1-\beta^2)} = \frac{2L_2\gamma^2}{c}. \quad (4.2.11)$$

Bei dem zweiten Teilstrahl, der sich im Ruhesystem der Apparatur senkrecht zur x^1 -Richtung zwischen Strahlteiler und Spiegel hin- und herbewegt, gilt (s. S. 11) für jeweils einen Weg $c^2 t_{21}^2 = L_1^2 + \beta^2 c^2 t_{21}^2$, d.h.

$$t_2 = 2t_{21} = \frac{2L_1}{c} \frac{1}{\sqrt{1-\beta^2}} = \frac{2L_1\gamma}{c}. \quad (4.2.12)$$

Die Laufzeitdifferenz ist also

$$\Delta t_A = \frac{2}{c}(L_1\gamma - L_2\gamma^2). \quad (4.2.13)$$

Dreht man nun die Apparatur um 90° , erhält man hingegen

$$\Delta t_B = \frac{2}{c}(L_1\gamma^2 - L_2\gamma), \quad (4.2.14)$$

d.h. das Interferenzmuster verschiebt sich im Vergleich zur vorherigen Ausrichtung entsprechend einer Änderung der Laufzeitdifferenz um

$$\Delta t_A - \Delta t_B = \frac{2(L_1 + L_2)}{c}(\gamma^2 - \gamma). \quad (4.2.15)$$

Da im Galilei-Newtonschen Raumzeitmodell die Zeit absolut ist, ist dies auch die Laufzeitdifferenz für einen auf der Erde ruhenden Beobachter, d.h. man sollte beim Drehen der Apparatur um 90° eine entsprechende Verschiebung des Interferenzmusters sehen. Da in diesem Fall $|\beta| \simeq 10^{-4} \ll 1$ ist, können wir den zweiten Faktor via Taylor-Entwicklung noch nähern zu

$$\gamma^2 - \gamma = (1 - \beta^2)^{-1} - (1 - \beta^2)^{-1/2} = 1 + \beta^2 - (1 + \beta^2/2) + \mathcal{O}(\beta^4) = \frac{\beta^2}{2}. \quad (4.2.16)$$

Die Laufzeitverschiebung und damit die Änderung des Interferenzmusters ist also ein Effekt der Ordnung β^2 :

$$\Delta t_A - \Delta t_B \simeq \frac{L_1 + L_2}{c} \beta^2. \quad (4.2.17)$$

Es zeigt sich übrigens, dass im Allgemeinen die Abweichungen zwischen der Äthertheorie und der speziellen Relativitätstheorie sich eben nur in der Ordnung β^2 bemerkbar machen. Für Effekte zur Ordnung β ergeben beide Theorien hingegen die gleichen Ergebnisse. Die Herausforderung für die Experimentatoren war also, dass man Effekte der Ordnung β^2 nachweisen musste, was wegen der i.a. verfügbaren Geschwindigkeiten sehr schwierig ist, denn ist schon $|\beta| \ll 1$ ist erst Recht $\beta^2 \ll 1$. Beim Michelson-Morley-Experiment ist $|\beta| \simeq 10^{-4}$ und $\beta^2 \simeq 10^{-8}$.

Die relative Verschiebung d der Interferenzstreifen ergibt sich zu

$$\frac{d}{\lambda} = \frac{c}{\lambda} (\Delta t_A - \Delta t_B) = \frac{L_1 + L_2}{\lambda} \beta^2, \quad (4.2.18)$$

wo λ die Wellenlänge des verwendeten Lichtes ist. Michelson und Morleys Apparatur von 1887 hatte die Abmessungen $L_1 \simeq L_2 = 11$ m und für das Licht können wir $\lambda \simeq 500$ nm = $5 \cdot 10^{-7}$ m annehmen. Daraus ergibt sich nach der Äthertheorie gemäß (4.2.18) eine relative Verschiebung des Interferenzmusters zu $d/\lambda \simeq 0,44$. Die Apparatur war genau genug, um zu zeigen, dass tatsächlich $d/\lambda < 0,02$ ist. Das war ein recht klarer Widerspruch zur Äthertheorie.

Analyse im Rahmen der Speziellen Relativitätstheorie (SRT)

Hier ergibt es natürlich keinen Sinn, von einem Ätherruhesystem zu reden, denn der Äther ist ja aus Sicht der SRT in Übereinstimmung mit der Tatsache, dass bis dato keinerlei Hinweise für eine derartige Substanz vorliegen, überflüssig, denn gemäß Einstein-Minkowski-Raumzeitmodell gehorcht die Elektrodynamik dem speziellen Relativitätsprinzip, d.h. es kann keine absolute Bewegung mittels elektromagnetischer Phänomene (einschließlich der Optik) nachweisbar sein, denn es gibt ja eben kein irgendwie privilegiertes Inertialsystem.

Wir können nur fragen, ob aus Sicht eines Beobachters, in dem die Versuchsapparatur sich geradlinig gleichförmig bewegt, eine Verschiebung des Interferenzmusters bei Drehung der Apparatur zu erwarten ist. Erstaunlicherweise wird dies in kaum einem Lehrbuch erörtert und leider zum Teil auch nicht ganz korrekt erklärt. Die einzige Referenz für eine korrekte Erklärung, die mir bekannt ist, ist [Sch94]. In diesem Paper werden die einzelnen Betrachtungen zur Geometrie der Apparatur und der Lichtwege bzgl. eines bewegten Beobachters nochmals ausführlich erklärt. Hier beschränken wir uns auf qualitative Bemerkungen zu den diversen Phänomenen.

Für das quantitative Verständnis genügt es bereits, mit der aus der Lorentz-Transformation folgenden Längenkontraktion zu argumentieren. Demnach ist der Abstand zwischen dem Strahlteiler und dem Spiegel in Richtung der Bewegung der Apparatur um den inversen Lorentz-Faktor $1/\gamma$ verkürzt gegenüber der entsprechenden Distanz für einen ruhenden Beobachter, während der Abstand des Spiegels in dazu senkrechter Richtung für den bewegten Beobachter gleich dem für den ruhenden Beobachter ist.

Wir müssen also in der obigen Analyse der Phasenunterschiede für die beiden Teilstrahlen im Rahmen der SRT gegenüber der im Rahmen der Äthertheorie lediglich diese Lorentz-Kontraktion berücksichtigen. In der ursprünglichen Orientierung des Interferometers ist also in (4.2.13) lediglich L_2 durch die Lorentz-kontrahierete Länge $L'_2 = L_2/\gamma$ zu ersetzen, während $L'_1 = L_1$ bleibt. Bei der um 90° gedrehten Apparatur ist hingegen

4.2. Kinematische Effekte

$L'_1 = L_1/\gamma$ statt L_1 in (4.2.14) und $L'_2 = L_2$ zu setzen. Dies liefert entsprechend

$$\Delta t_A = \Delta t_B = \frac{L_1 - L_2}{c} \gamma, \quad (4.2.19)$$

d.h. $\Delta t_A - \Delta t_B = 0$ in Übereinstimmung mit dem Nullresultat des Michelson-Morley-Experiments.

Zum Abschluss unserer Diskussion des Michelson-Morley-Experiments machen wir noch einige qualitative Anmerkungen: Es ist zu beachten, dass aus Sicht eines Beobachters, bzgl. dessen sich das Michelson-Interferometer bewegt, die Geometrie eine andere ist als im Ruhesystem der Apparatur. Auch dies kann man leicht allein mit der Längenkontraktion verstehen. Es ist aufgrund der Lorentz-Transformation klar, dass die beiden Spiegel in beiden Bezugssystemen jeweils in x^1 -Richtung bzw. senkrecht dazu ausgerichtet bleiben. Allerdings ist der Winkel des strahlteilenden halbdurchlässigen Spiegels nicht mehr im Winkel von 45° relativ zur x^1 -Achse ausgerichtet. Dies werden wir in einer Übungsaufgabe mittels ausführlicher Verwendung der Lorentz-Transformation auch nachrechnen.

Hier genüge die Argumentation unter Verwendung der Längenkontraktion: Sei also ein Stab in seinem Ruhesystem in der x^1 - x^2 -Ebene so orientiert, dass er einen Winkel θ mit der x^1 -Achse einnimmt. Der entsprechende Einheitsrichtungsvektor ist also $\vec{n} = (\cos \theta, \sin \theta)$. In einem Bezugssystem, das sich gegen dieses Ruhesystem mit der Geschwindigkeit βc in x^1 -Richtung bewegt, ist die x^1 -Koordinate dieses Richtungsvektors längenkontrahiert; die x^2 -Koordinate bleibt hingegen gleich. Der Stab nimmt also für einen Beobachter, in dem sich der Stab mit der Geschwindigkeit βc in x^1 -Richtung bewegt einen anderen Winkel θ' zur x^1 -Richtung ein. Der entsprechende (nicht auf Länge 1 normierte) Richtungsvektor ist demnach $\vec{n}' = (\cos \theta/\gamma, \sin \theta)$, und der Winkel θ' ist demnach durch

$$\tan \theta' = \frac{\sin \theta}{\cos \theta/\gamma} = \gamma \tan \theta \quad (4.2.20)$$

gegeben. Da $\gamma > 1$ ist also i.a. $\theta' > \theta$. Für $\theta = 0$ ist allerdings $\tan \theta = 0$ und folglich auch $\tan \theta' = 0$ und also $\theta' = 0$. Für $\theta = \pi/2$ wird $\tan \theta \rightarrow \infty$, und damit ist auch $\tan \theta' \rightarrow \infty$, d.h. es ist auch $\theta' = \pi/2$. Für einen Winkel $\theta \in (0, \pi)$ wird hingegen $\theta' > \theta$.

Allerdings ändert sich auch die Geometrie der Lichtstrahlen (bzw. genauer gesagt der Wellenvektoren) im Rahmen der SRT auf nichttriviale Weise. Wir gehen darauf weiter unten noch genauer ein, wenn wir die relativistische Formulierung der Elektrodynamik besprechen.

Schließlich könnte man meinen, dass nun ein Widerspruch zu den Reflexions- und Brechungsgesetzen vorliegen könnte. Dies ist freilich nicht der Fall, denn die Elektrodynamik ist ja invariant unter Lorentz-Transformationen, und es kann ein solcher Widerspruch nicht entstehen. In der Tat ermöglicht es uns, die Gesetze für die Reflexion von Licht an sich bewegenden Spiegeln und die Lichtbrechung in sich bewegenden Medien einfach dadurch herzuleiten, dass wir die bekannten Brechungs- und Reflexionsgesetze in einem Bezugssystem, wo das Medium und die spiegelnde Oberfläche ruhen, verwenden und dann eine Lorentz-Transformation durchführen.

Genauere Ausführungen zu diesen geometrischen Aspekten des Michelson-Morley-Experiments in seiner Behandlung im Rahmen der SRT findet sich auf recht elementarem Level in dem bereits oben zitierten Paper [Sch94].

4.2.6 Bemerkungen zur Historie der Lorentz-Transformation

Die einfache Tatsache, dass die Lorentz-Kontraktion eben diesen Widerspruch zwischen dem Nullresultat des Michelson-Morley-Experiments und der Erwartung aufgrund der ursprünglichen Äthertheorie, behebt, hat ja vor Aufstellung der SRT durch Einstein zu eben der modifizierten Äthertheorie, insbesondere durch FitzGerald und Lorentz geführt: Sie haben angenommen, dass die Bewegung materieller Körper durch den Äther Kraftwirkungen hervorruft, die zur Verkürzung um genau den Lorentz-Faktor führen. Deshalb heißt

zwar der Faktor der Längenkontraktion ironischerweise Lorentz-Kontraktion, obwohl der ursprüngliche Erklärungsansatz mit der modifizierten Äthertheorie aus Sicht der SRT falsch ist. Lorentz selbst hat dies freilich auch sehr schnell bemerkt, denn auch in der modifizierten Äthertheorie wurde die Zeit noch als absolut angenommen. Die Längenkontraktionshypothese unter Beibehaltung der Galilei-Newtonschen absoluten Zeit liefert beim Wechsel von einem Inertialsystem zu einem anderen aber keine Transformationen zwischen den Koordinaten und der Zeit (wobei letztere ja eben nicht transformiert wird!) bzgl. der die Maxwell-Gleichungen forminvariant bleiben. Dies ist erst mit der oben besprochenen Lorentz-Transformation der Fall, und Lorentz hat diese Transformation auch bereits vor Einstein gefunden, und obwohl Einstein in seinem berühmten Paper von 1905 [Ein05] kein einziges Zitat bringt, nennt er ja dort auch entsprechend diese Transformationen Lorentz-Transformation. Während allerdings Lorentz die Transformation durch rein mathematische Überlegungen über mögliche Raumzeit-Transformationen, die die Maxwell-Gleichungen forminvariant lassen, gefunden hat und zunächst eher als mathematische Kuriosität als als tieferliegende physikalische Eigenschaften der Raumzeit-Beschreibung, hat eben Einstein genau diesen für die gesamte moderne Physik entscheidenden Schluss gezogen: Es muss die Raumzeit-Beschreibung gegenüber der Galilei-Newtonschen Raumzeit-Beschreibung so geändert werden wie oben beschrieben, und dies führt eben bereits ohne komplizierte Analyse der Maxwell-Gleichungen aus rein kinematischen Überlegungen (im Originalpaper übrigens sehr ähnlich wie oben im Skript) auf die Lorentz-Transformation und über Minkowskis Analyse schließlich zur geometrischen Beschreibung der Raumzeit als Minkowski-Raum. Was zu dieser Zeit übrigens in Vergessenheit geraten war, ist die Tatsache, dass bereits Woldemar Voigt in 1887 Transformationen sehr ähnlich den Lorentz-Transformationen hergeleitet hat, die die Maxwell-Gleichungen für den Fall, dass keine Ladungs- und Stromverteilungen vorhanden sind, invariant lassen. Diese „Voigt-Transformationen“ unterschieden sich dabei von den Lorentz-Transformationen lediglich dadurch, dass der γ -Faktor nicht auftritt¹. Diese Transformation läßt nun nicht alle Minkowski-Produkte invariant sondern nur die Beziehung $(ct)^2 - \vec{x}^2 = 0$. Dies genügt aber dafür, dass die quellenfreie Wellengleichung $\square\Phi = 0$ (bzw. für die freien elektromagnetischen Wellen $\square\vec{E} = 0, \square\vec{B} = 0$) bzgl. dieser Transformation invariant ist.

4.3 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

$$[x^\mu(t)] = \begin{pmatrix} ct \\ R[\cos(\omega t) - 1] \\ R \sin \omega t \\ 0 \end{pmatrix} \quad (4.3.1)$$

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 $(dx'^\mu) = (cdt', 0, 0, 0) = (cd\tau, 0, 0, 0)$ gilt. Damit ist aber $c^2 d\tau^2 = \eta_{\mu\nu} dx'^\mu dx'^\nu = \eta_{\mu\nu} dx^\mu dx^\nu$, 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

¹

S. den sehr lesenswerten Wikipedia-Artikel zur Historie der Lorentz-Transformation https://de.wikipedia.org/wiki/Geschichte_der_Lorentz-Transformation.

Ruhsystem von Bob vergeht, zu

$$\tau = \frac{1}{c} \int_0^t d\tilde{t} \sqrt{\frac{dx^\mu(\tilde{t})}{d\tilde{t}} \frac{dx^\mu(\tilde{t})}{d\tilde{t}}} = \int_0^t d\tilde{t} \sqrt{1 - \frac{\omega^2 R^2}{c^2}} = t \sqrt{1 - \frac{\omega^2 R^2}{c^2}}. \quad (4.3.2)$$

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 anhand der Messung der Lebensdauer von instabilen Teilchen in Speicherringen überprüft².

4.4 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 relativistische Raum-Zeit-Struktur, dass sich keine Signale mit einer schnelleren als der Lichtgeschwindigkeit ausbreiten können, denn wäre eine solche Wirkung gleichzeitig in einem Bezugssystem Σ , wäre sie dies nicht mehr in einem relativ dazu bewegten Bezugssystem Σ' . Seien nämlich die Raum-Zeit-Koordinaten der Ereignisse bzgl. Σ , die bzgl. dieses Bezugssystems gleichzeitig zur Zeit $t = 0$ stattfinden $\underline{x} = (0, \vec{x})^T$ und $\underline{y} = (0, \vec{y})^T$. Dann gilt für die entsprechende Differenz

$$(\underline{x} - \underline{y}) \cdot (\underline{x} - \underline{y}) = -(\vec{x} - \vec{y})^2 < 0. \quad (4.4.1)$$

Man nennt entsprechend Ereignisse, für die die Differenz der entsprechenden Vierervektoren ein negatives Lorentz-Produkt besitzen, **raumartig**. Bewegt sich wieder Σ' gegenüber Σ mit der Relativgeschwindigkeit $\vec{v} = \beta c \vec{e}_1$, gilt gemäß (4.1.24)

$$\begin{pmatrix} c(t'_x - t'_y) \\ x'^1 - y'^1 \\ x'^2 - y'^1 \\ x'^3 - y'^3 \end{pmatrix} = \begin{pmatrix} -\gamma\beta(x^1 - y^1) \\ \gamma(x^1 - y^1) \\ x^2 - y^2 \\ x^3 - y^3 \end{pmatrix}. \quad (4.4.2)$$

Die Ereignisse sind also bzgl. Σ' *nicht* mehr gleichzeitig. Außerdem ist das Vorzeichen der Zeitkomponente abhängig vom Vorzeichen der Relativgeschwindigkeiten der Bezugssysteme. Je nach dem Vorzeichen von β kann bzgl. Σ' Ereignis \underline{x} vor ($\beta > 0$) dem Ereignis \underline{y} stattgefunden haben oder danach ($\beta < 0$), entsprechend dem Vorzeichen der entsprechenden Zeitdifferenz $t'_x - t'_y$.

Daraus folgt auch, dass zwei Ereignisse, die raumartig zueinander liegen, nicht **kausal** verknüpft sein können, d.h. es kann das eine Ereignisse nicht Ursache des anderen sein, denn dann müsste die zeitliche Reihenfolge der Ereignisse unabhängig vom Bezugssystem sein.

Charakterisieren nun die beiden bzgl. Σ gleichzeitigen Ereignisse \underline{x} und \underline{y} eine Fernwirkung zwischen zwei Punktteilchen, finden diese Ereignisse nach der obigen Betrachtung je nach Boostgeschwindigkeit bzgl. Σ' in unterschiedlicher zeitlicher Reihenfolge statt, d.h. es kann sich nicht um einen *kausalen* Vorgang handeln, wie wir es für eine Wechselwirkung verlangen.

²Vgl. dazu den entsprechenden Wikipedia-Artikel

https://en.wikipedia.org/wiki/Time_dilation_of_moving_particles

4. Spezielle Relativitätstheorie

Dies ist für zueinander **zeitartig** gelegene Ereignisse, für die $(\underline{x} - \underline{y}) \cdot (\underline{x} - \underline{y}) > 0$ sicher der Fall, denn dann gilt für die Koordinaten der Ereignisse bzgl. Σ'

$$\begin{pmatrix} c(t'_x - t'_y) \\ x'^1 - y'^1 \\ x'^2 - y'^2 \\ x'^3 - y'^3 \end{pmatrix} = \begin{pmatrix} \gamma[c(t_x - t_y) - \beta(x^1 - y^1)] \\ \gamma[(x^1 - y^1) - \beta c(t_x - t_y)] \\ x^2 - y^2 \\ x^3 - y^3 \end{pmatrix}. \quad (4.4.3)$$

Sei nun $t_x > t_y$. Wir wollen zeigen, dass für zeitartig zueinander gelegene Ereignisse auch $t'_x > t'_y$ sein muss. Da $(\underline{x} - \underline{y})$ zeitartig sein soll, gilt definitionsgemäß $c^2(t_x - t_y)^2 > (\vec{x} - \vec{y})^2$, d.h. mit $t_x - t_y > 0$ ist $c(t_x - t_y) > |\vec{x} - \vec{y}|$. Da notwendig $|\beta| < 1$ ist, folgt damit

$$\begin{aligned} c(t'_x - t'_y) &= \gamma[c(t_x - t_y) - \beta(x^1 - y^1)] \\ &\geq \gamma[c(t_x - t_y) - |\beta||x^1 - y^1|] \\ &\geq \gamma[c(t_x - t_y) - |\beta||\vec{x} - \vec{y}|] > \gamma[c(t_x - t_y) - |\vec{x} - \vec{y}|] > 0. \end{aligned} \quad (4.4.4)$$

Der Ausweg aus diesem Dilemma ist es, die Kraftwirkungen zwischen Punktteilchen 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 bemerken dazu, dass auch zwei **lichtartig** zueinander gelegene Ereignisse kausal verknüpft sein können. Die obige Betrachtung in (4.4.4) gilt nämlich offenbar auch noch, falls $(\underline{x} - \underline{y}) \cdot (\underline{x} - \underline{y}) = (\underline{x}' - \underline{y}') \cdot (\underline{x}' - \underline{y}') = 0$ ist. Demnach folgt auch in diesem Fall aus $t_x - t_y > 0$, dass $t'_x - t'_y > 0$ ist, d.h. die zeitliche Reihenfolge für zueinander zeit- oder licht-artige Ereignisse ist unabhängig vom Inertialsystem.

Wir wollen die relativistischen Bewegungsgesetze plausibel machen, indem wir zunächst mit allein Lorentz-kovarianten Größen, also Skalaren und Vierervektoren arbeiten. Offenbar kann man die Bewegung eines Teilchens dadurch kovariant beschreiben, dass man seine Trajektorie im Minkowski-Raum, also die Weltlinie $\underline{x}(\lambda)$ betrachtet. Dabei ist λ ein beliebiger Parameter.

Andererseits ist klar, dass wie in der Newtonschen Mechanik die Bewegung durch die Trajektorie bzgl. jedes Bezugssystems im gewöhnlichen dreidimensionalen physikalischen Raum als Funktion der Zeit beschreibbar sein muss. Dabei sollten die bewährten Newtonschen Gesetze der Bewegung im Grenzfall von Geschwindigkeiten $\vec{v} = d\vec{x}/dt$ mit $|\vec{v}| \ll c$ zumindest näherungsweise korrekt sein. Betrachten wir nun die momentane Bewegung des Teilchens zur Zeit t . Da in jedem Bezugssystem die Weltlinie des Teilchens zeitartig sein muss, können wir durch eine geeignete Lorentz-Transformation stets in ein Bezugssystem Σ' transformieren, wo das Teilchen momentan ruht. Ein infinitesimales Zeitinkrement in diesem **momentanen Ruhesystem** können wir nun **kovariant** schreiben. Offenbar ist nämlich in diesem Bezugssystem

$$\frac{d}{d\lambda} \begin{pmatrix} ct'(\lambda) \\ \vec{x}' \end{pmatrix} = \begin{pmatrix} cd t'/d\lambda \\ \vec{0} \end{pmatrix}, \quad (4.4.5)$$

denn voraussetzungsgemäß ist die Geschwindigkeit in diesem Bezugssystem $d\vec{x}'/dt' = (d\vec{x}'/d\lambda)(d\lambda/dt') = \vec{0}$. Damit ist in diesem Bezugssystem

$$dt' = d\lambda \frac{1}{c} \sqrt{\frac{d\underline{x}'}{d\lambda} \cdot \frac{d\underline{x}'}{d\lambda}}. \quad (4.4.6)$$

Da sich aber Lorentz-Produkte unter Lorentz-Transformationen nicht ändern, können wir dies auch bzgl. der Koordinaten im festen Inertialsystem Σ ausdrücken, d.h. wir können das Zeitinkrement im momentanen Ruhesystem des Teilchens $dt' = d\tau$ kovariant schreiben. In diesem Zusammenhang nennt man τ auch die **Eigenzeit** des Teilchens:

$$d\tau = d\lambda \frac{1}{c} \sqrt{\frac{dx}{d\lambda} \cdot \frac{dx}{d\lambda}}. \quad (4.4.7)$$

Nun ist der Weltlinienparameter λ beliebig wählbar, und wir können demnach für λ auch die Zeit t im Bezugssystem Σ verwenden:

$$d\tau = dt \frac{1}{c} \sqrt{\frac{dx}{dt} \cdot \frac{dx}{dt}} = dt \sqrt{1 - \vec{\beta}^2} \quad (4.4.8)$$

gegeben. Dabei haben wir $\vec{\beta} = \vec{v}/c = \dot{\vec{x}}/c$ geschrieben.

In der Newtonschen Mechanik wurde nun der **Impuls** eines Teilchens durch $\vec{p} = m\dot{\vec{x}}$ definiert. Dies sollte nun zu jedem Zeitpunkt auch noch im jeweiligen momentanen Ruhesystem des Teilchens in der Relativitätstheorie die korrekte Definition sein. Dies legt es nahe, in der Relativitätstheorie den **Vierer-Impulsvektor** durch

$$\underline{p} = m \frac{dx}{d\tau} \quad (4.4.9)$$

zu definieren. Da nämlich gemäß (4.4.6) $d\tau$, also das **Eigenzeitinkrement** ein Lorentz-Skalar ist, ist demnach (4.4.9) ein Lorentz-Vektor, wobei wir annehmen, dass die **Masse** m ein Lorentz-Skalar ist³.

Betrachten wir nun die Komponenten des Viererimpulses in einem beliebigen Inertialsystem, so gilt bei einer Parametrisierung der Weltlinie mit der Zeit t in diesem Inertialsystem

$$\underline{p} = m \frac{dx}{d\tau} = m \frac{dx}{dt} \frac{dt}{d\tau}. \quad (4.4.10)$$

Nun gilt wegen der Definition (4.4.8) des Eigenzeitinkrements

$$\frac{dt}{d\tau} = \left(\frac{d\tau}{dt}\right)^{-1} = \frac{1}{\sqrt{1 - \dot{\vec{x}}^2/c^2}} = \frac{1}{\sqrt{1 - \vec{\beta}^2}} = \gamma \quad (4.4.11)$$

und folglich

$$(p^\mu) = m\gamma \begin{pmatrix} c \\ \dot{\vec{x}} \end{pmatrix}. \quad (4.4.12)$$

Betrachten wir nun den nichtrelativistischen Grenzfall, d.h. $|\vec{\beta}| = |\dot{\vec{x}}|/c \ll 1$. Dann ist $\gamma = 1 + \beta^2/2 + \mathcal{O}(\beta^4)$, und die räumlichen Komponenten von (4.4.9) sind in der Tat

$$\vec{p} = m\gamma\dot{\vec{x}} = m\dot{\vec{x}}[1 + \mathcal{O}(\beta^2)] \simeq m\dot{\vec{x}}, \quad (4.4.13)$$

d.h. der gewöhnliche nichtrelativistische Impuls. Die zeitliche Komponente ist

$$p^0 = mc\gamma = mc \left[1 + \frac{\beta^2}{2} + \mathcal{O}(\beta^4) \right] \simeq \frac{1}{c} \left(mc^2 + \frac{m}{2} \dot{\vec{x}}^2 \right). \quad (4.4.14)$$

³In einigen älteren Lehrbüchern und leider auch in den meisten neueren Schulbüchern 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. Vgl. dazu auch [Fli18].

4. Spezielle Relativitätstheorie

In der Klammer entsteht also bis auf eine additive Konstante $E_0 = mc^2$ die **kinetische Energie** $mv^2/2$ des Teilchens wie sie in der Newtonschen Mechanik definiert ist. Es liegt daher nahe,

$$\mathcal{E} = p^0 c = mc^2 \gamma = E_0 + E_{\text{kin}}, \quad \text{mit} \quad E_0 = mc^2, \quad E_{\text{kin}} = mc^2(\gamma - 1) \quad (4.4.15)$$

als die **relativistische Energie** des Teilchens zu definieren⁴. Die räumliche Komponente des Viererimpulses ist dann $p^0 = \mathcal{E}/c$. Wir bemerken, dass es aufgrund dieser Betrachtung in der relativistischen Mechanik sinnvoll ist, die **Ruheenergie** E_0 in die Definition der Energie eines Teilchens einzubeziehen, weil dann die Energie (bis auf einen Faktor) die Zeitkomponente des Impulsvierervektors $\underline{p} = (\mathcal{E}/c, \vec{p})$ ist und damit leicht die Transformationseigenschaften von Energie und Impuls unter Lorentz-Transformation folgen, nämlich als Komponenten eines Vierervektors:

$$p'^{\mu} = \Lambda^{\mu}_{\nu} p^{\nu}. \quad (4.4.16)$$

Mit dem Viererimpuls können wir nun den Lorentz-Skalar $\underline{p} \cdot \underline{p}$ bilden. Mit (4.4.12) folgt unter Verwendung von (4.4.15)

$$\underline{p} \cdot \underline{p} = \left(\frac{\mathcal{E}}{c}\right)^2 - \vec{p}^2 = m^2 \gamma^2 c^2 (1 - \beta^2) = m^2 c^2. \quad (4.4.17)$$

Daraus ergibt sich sofort die relativistische Beziehung zwischen Energie und Impuls zu

$$\mathcal{E} = c \sqrt{m^2 c^2 + \vec{p}^2}. \quad (4.4.18)$$

Im Folgenden ist es auch bequem, die dimensionslose **Vierergeschwindigkeit**

$$\underline{u} = \frac{1}{c} \frac{d\underline{x}}{d\tau} = \begin{pmatrix} \gamma \\ \gamma \vec{\beta} \end{pmatrix} \quad (4.4.19)$$

einzuführen. Dann ist (*nachrechnen!*)

$$\underline{p} = mc \underline{u} \quad \text{und} \quad \underline{u} \cdot \underline{u} = 1. \quad (4.4.20)$$

Wir bemerken noch, dass wir aus (4.4.19) und (4.4.20) die Dreiergeschwindigkeit in einem gegebenen Inertialsystem durch die Vierergeschwindigkeit bzw. den Viererimpuls gemäß

$$\vec{\beta} = \frac{1}{c} \frac{d\vec{x}}{dt} = \frac{\vec{u}}{u^0} = \frac{mc \vec{u}}{mc u^0} \Rightarrow \vec{\beta} = \frac{\vec{p}}{\mathcal{E}/c} \quad (4.4.21)$$

ausdrücken können.

Jetzt können wir auch Newtons zweites Postulat relativistisch erweitern, indem wir schreiben

$$\frac{d}{d\tau} p^{\mu} = mc \frac{d}{d\tau} u^{\mu} = m \frac{d^2}{d\tau^2} x^{\mu} = K^{\mu}, \quad (4.4.22)$$

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, sodass wieder eine Vierervektorgleichung entsteht. Man nennt den Vierervektor \underline{K} die **Minkowski-Kraft**. Die vier Komponenten sind aber nicht unabhängig voneinander. Aufgrund der Definition der Vierergeschwindigkeit (4.4.10) gilt nämlich $\underline{u} \cdot \underline{u} = 1$. Leitet man diese Bedingung nach τ ab, folgt

$$\frac{d\underline{u}}{d\tau} \cdot \underline{u} = 0, \quad (4.4.23)$$

⁴Wir 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.

4.5. Zerfall und Stöße von Teilchen

d.h. multipliziert man (4.4.22) mit u^μ , erhält man die Bedingung

$$\underline{K} \cdot \underline{u} = \eta_{\mu\nu} K^\mu u^\nu = 0 \quad (4.4.24)$$

für die Minkowski-Kraft. Damit lässt sich auch die Zeitkomponente der Bewegungsgleichung (4.4.22) interpretieren, denn aus (4.4.24) folgt

$$\eta_{\mu\nu} K^\mu u^\nu = u^0 K^0 - \vec{u} \cdot \vec{k} = 0 \Rightarrow u^0 K^0 = \vec{u} \cdot \vec{k}. \quad (4.4.25)$$

Damit folgt für die Zeitkomponente von (4.4.22)

$$m c \frac{d u^0}{d \tau} = K^0 = \frac{\vec{u}}{u^0} \cdot \vec{K} = \vec{\beta} \cdot \vec{K}. \quad (4.4.26)$$

Wegen $d\tau = \sqrt{1 - \vec{\beta}^2} dt = dt / \gamma$ erhalten wir schließlich

$$m c \frac{d u^0}{d t} = \vec{\beta} \cdot \frac{\vec{K}}{\gamma}. \quad (4.4.27)$$

Multiplizieren wir dies mit c und schreiben $\vec{K} = u^0 \vec{F} = \gamma \vec{F}$ erhalten wir wegen (4.4.15)

$$\frac{d \mathcal{E}}{d t} = \frac{d \mathcal{E}_{\text{kin}}}{d t} = \vec{v} \cdot \vec{F}. \quad (4.4.28)$$

Für den räumlichen Teil von (4.4.22) erhalten wir die nicht manifest kovariante Form der Bewegungsgleichung

$$\frac{d}{d t} \vec{p} = \frac{1}{\gamma} \frac{d}{d \tau} \vec{p} = \frac{1}{\gamma} \vec{K} = \vec{F}, \quad (4.4.29)$$

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 \underline{K} zu betrachten. Jedenfalls erweist sich die Zeitkomponente der Bewegungsgleichung (4.4.22), geschrieben in der Form (4.4.29) einfach als die übliche Form des **Energie-Arbeits-Theorems**, die notwendig aus den drei voneinander unabhängigen Bewegungsgleichungen für die räumlichen Komponenten in (4.4.22) folgen muss (in genauer Analogie zur Newtonschen Mechanik).

4.5 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 Quantenfeldtheorie**. Allerdings gelten auch dort stets **Energie- und Impulserhaltung** und die Energie-Impulsbeziehung (4.4.17). 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 → 2)-Stöße**.

4.5.1 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 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 \underline{p} für das zerfallende Teilchen mit \underline{p}'_1 und \underline{p}'_2 , d.h.

$$\underline{p} = \underline{p}'_1 + \underline{p}'_2. \quad (4.5.1)$$

Weiter gelten die Energie-Impulsbeziehungen für jeden dieser Viererimpulse. Diese können wir am einfachsten in der kovarianten Form (4.4.17) berücksichtigen:

$$\underline{p} \cdot \underline{p} = M^2 c^2, \quad \underline{p}'_1 \cdot \underline{p}'_1 = m_1^2 c^2, \quad \underline{p}'_2 \cdot \underline{p}'_2 = m_2^2 c^2. \quad (4.5.2)$$

Gehen wir nun in das Ruhssystem des zerfallenden Teilchens, ist

$$(p^\mu) = \begin{pmatrix} Mc \\ 0 \end{pmatrix}, \quad (p_1'^\mu) = \begin{pmatrix} \mathcal{E}_1/c \\ \vec{p}' \end{pmatrix}, \quad (p_2'^\mu) = \begin{pmatrix} \mathcal{E}_2/c \\ -\vec{p}' \end{pmatrix}. \quad (4.5.3)$$

Offenbar ist durch diese Schreibweise der räumliche Teil von (4.5.1), also die Impulserhaltung, identisch erfüllt. Die Zeitkomponente ergibt

$$\mathcal{E}_1 + \mathcal{E}_2 = M c^2. \quad (4.5.4)$$

Wir wollen nun $|\vec{p}'|$ bestimmen. Wegen der Energie-Impuls-Beziehungen (4.5.2) und der Energieerhaltungsgleichung (4.5.4) ist dies nämlich der einzige freie Parameter. Dazu müssen wir aber nur (4.5.1) im Sinne des Minkowski-Produktes quadrieren. Es folgt dann

$$\underline{p} \cdot \underline{p} = M^2 c^2 = \underline{p}'_1 \cdot \underline{p}'_1 + \underline{p}'_2 \cdot \underline{p}'_2 + 2 \underline{p}'_1 \cdot \underline{p}'_2 = (m_1^2 + m_2^2) c^2 + 2 \underline{p}'_1 \cdot \underline{p}'_2. \quad (4.5.5)$$

Im obigen Bezugssystem, wo das zerfallende Teilchen vor dem Zerfall ruht, gilt

$$\underline{p}'_1 \cdot \underline{p}'_2 = \frac{\mathcal{E}_1 \mathcal{E}_2}{c^2} + \vec{p}'^2. \quad (4.5.6)$$

Mit (4.5.3) folgt

$$\underline{p}'_1 \cdot \underline{p}'_2 = \frac{\mathcal{E}_1 (M c^2 - \mathcal{E}_1)}{c^2} + \vec{p}'^2 = M \mathcal{E}_1 - (m_1^2 c^2 + \vec{p}'^2) + \vec{p}'^2 = M \mathcal{E}_1 - m_1^2 c^2. \quad (4.5.7)$$

Setzen wir dies in (4.5.5) ein, erhalten wir

$$M^2 c^2 = (m_1^2 + m_2^2) c^2 + 2(M \mathcal{E}_1 - m_1^2 c^2) = (m_2^2 - m_1^2) c^2 + 2M \mathcal{E}_1 \quad (4.5.8)$$

bzw.

$$\mathcal{E}_1 = \frac{(M^2 + m_1^2 - m_2^2) c^2}{2M} \quad (4.5.9)$$

Mit der Energie-Impulsbeziehung erhalten wir daraus schließlich (*nachrechnen!*)

$$\vec{p}'^2 = \frac{\mathcal{E}_1^2}{c^2} - m_1^2 c^2 = \frac{(M^2 + m_1^2 - m_2^2)^2 c^2}{4M^2} - m_1^2 c^2 = \frac{[M^2 - (m_1 + m_2)^2][M^2 - (m_1 - m_2)^2] c^2}{4M^2}. \quad (4.5.10)$$

Da $\vec{p}'^2 \geq 0$ sein muss, kann der Zerfall nur unter der Bedingung stattfinden, dass

$$M \geq m_1 + m_2 \quad (4.5.11)$$

ist. Mit (4.5.11) ist die Kinematik des Zerfalls in dem betrachteten Inertialsystem, in dem das Teilchen anfangs ruht, vollständig bestimmt.

4.5.2 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 \underline{p}_1 und \underline{p}_2 und nach dem Stoß \underline{p}_3 und \underline{p}_4 . Es gilt beim Stoß wieder die Energie- und Impulserhaltung, d.h.

$$\underline{p}_1 + \underline{p}_2 = \underline{p}_3 + \underline{p}_4. \quad (4.5.12)$$

Wir können nun aus diesen vier Impulsen drei Invarianten, die **Mandelstam-Variablen**

$$\begin{aligned} s &= (\underline{p}_1 + \underline{p}_2)^2 = (\underline{p}_3 + \underline{p}_4)^2, \\ t &= (\underline{p}_1 - \underline{p}_3)^2 = (\underline{p}_2 - \underline{p}_4)^2, \\ u &= (\underline{p}_1 - \underline{p}_4)^2 = (\underline{p}_2 - \underline{p}_3)^2 \end{aligned} \quad (4.5.13)$$

bilden. Diese sind aber nicht voneinander unabhängig. Multipliziert man die Minkowskiprodukte aus und benutzt die Energie-Impulsbeziehungen $\underline{p}_j^2 = m_j^2 c^2$ sowie (4.5.12), erhält man (*nachrechnen!*)

$$s + t + u = (m_1^2 + m_2^2 + m_3^2 + m_4^2)c^2. \quad (4.5.14)$$

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 \underline{p}_2 wegen $\underline{p}_2^2 = m_2^2 c^2 > 0$ 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 $\underline{p}_1 + \underline{p}_2$ 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

$$p_1^0 + p_2^0 = \sqrt{m_1^2 c^2 + \vec{p}_1^2} + \sqrt{m_2^2 c^2 + \vec{p}_2^2} > |\vec{p}_1| + |\vec{p}_2| \geq |\vec{p}_1 + \vec{p}_2| \Rightarrow s = (p_1^0 + p_2^0)^2 - (\vec{p}_1 + \vec{p}_2)^2 > 0. \quad (4.5.15)$$

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 (*Nachrechnen!*).

4.5.3 Kinematik im Schwerpunktsystem

Wir beginnen mit der näheren Analyse der Kinematik im **Schwerpunktsystem**. Dann gilt

$$(p_1^\mu) = \begin{pmatrix} \mathcal{E}_1/c \\ \vec{p} \end{pmatrix}, \quad (p_2^\mu) = \begin{pmatrix} \mathcal{E}_2/c \\ -\vec{p} \end{pmatrix}, \quad (p_3^\mu) = \begin{pmatrix} \mathcal{E}_3/c \\ \vec{p}' \end{pmatrix}, \quad (p_4^\mu) = \begin{pmatrix} \mathcal{E}_4/c \\ -\vec{p}' \end{pmatrix}, \quad (4.5.16)$$

und folglich ergibt sich aus (4.5.13)

$$\sqrt{s} = \frac{\mathcal{E}_1 + \mathcal{E}_2}{c} = \frac{\mathcal{E}_3 + \mathcal{E}_4}{c}, \quad (4.5.17)$$

d.h. $c\sqrt{s}$ ist die Gesamtenergie der Teilchen vor und nach dem Stoß im Schwerpunktsystem.

Wir können nun die Energien und Impulsbeträge im Schwerpunktsystem mit Hilfe der Mandelstam-Variablen ausdrücken. Zunächst ist

$$\begin{aligned} s &= (\underline{p}_1 + \underline{p}_2)^2 = (m_1^2 + m_2^2)c^2 + 2\left(\frac{\mathcal{E}_1\mathcal{E}_2}{c^2} + \vec{p}^2\right) \\ &\stackrel{(4.5.17)}{=} (m_1^2 + m_2^2)c^2 + 2\left(\frac{\mathcal{E}_1(c\sqrt{s} - \mathcal{E}_1)}{c^2} + \vec{p}^2\right) \\ &= (m_2^2 - m_1^2)c^2 + \frac{2\mathcal{E}_1\sqrt{s}}{c}. \end{aligned} \quad (4.5.18)$$

Dabei haben wir im letzten Schritt die Energie-Impulsbeziehung $\mathcal{E}_1^2/c^2 - \vec{p}^2 = m_1^2c^2$ verwendet. Aus (4.5.18) erhält man schließlich

$$\frac{\mathcal{E}_1}{c} = \frac{s + (m_1^2 - m_2^2)c^2}{2\sqrt{s}}. \quad (4.5.19)$$

Eliminieren wir in (4.5.18) \mathcal{E}_1 zugunsten von \mathcal{E}_2 , erhalten wir

$$\frac{\mathcal{E}_2}{c} = \frac{s + (m_2^2 - m_1^2)c^2}{2\sqrt{s}}. \quad (4.5.20)$$

Die Energien im Endzustand ergeben sich durch Ersetzen der Indizes 1 und 2 durch 3 bzw. 4:

$$\begin{aligned} \frac{\mathcal{E}_3}{c} &= \frac{s + (m_3^2 - m_4^2)c^2}{2\sqrt{s}}, \\ \frac{\mathcal{E}_4}{c} &= \frac{s + (m_4^2 - m_3^2)c^2}{2\sqrt{s}}. \end{aligned} \quad (4.5.21)$$

Die Impulsbeträge im Anfangs- bzw. Endzustand ergeben sich daraus mit Hilfe der Energie-Impuls-Beziehung nach einigen Umformungen (*nachrechnen*)

$$\begin{aligned} \vec{p}^2 &= \frac{\mathcal{E}_1^2}{c^2} - m_1^2c^2 = \frac{1}{4s}[s - (m_1 + m_2)^2c^2][s - (m_1 - m_2)^2c^2], \\ \vec{p}'^2 &= \frac{\mathcal{E}_3^2}{c^2} - m_3^2c^2 = \frac{1}{4s}[s - (m_3 + m_4)^2c^2][s - (m_3 - m_4)^2c^2] \end{aligned} \quad (4.5.22)$$

4.5.4 Kinematik im Laborsystem

Im Laborsystem haben wir

$$p_1 = \begin{pmatrix} \mathcal{E}_1/c \\ \vec{p}_1 \end{pmatrix}, \quad p_2 = \begin{pmatrix} m_2c \\ 0 \end{pmatrix}. \quad (4.5.23)$$

Damit ist

$$s = \left(\frac{\mathcal{E}_1 + m_2c^2}{c}\right)^2 - \vec{p}_1^2 = \left(\frac{\mathcal{E}_1 + m_2c^2}{c}\right)^2 + m_1^2c^2 - \frac{\mathcal{E}_1^2}{c^2} \quad (4.5.24)$$

Nach einigen Umformungen erhält man

$$\mathcal{E}_1 = \frac{s - (m_1^2 + m_2^2)c^2}{2m_2}. \quad (4.5.25)$$

Weiter folgt (*nachrechnen!*)

$$\vec{p}_1^2 = \frac{\mathcal{E}_1^2}{c^2} - m_1^2c^2 = \frac{[s - (m_1 + m_2)^2c^2][s - (m_1 - m_2)^2c^2]}{4m_2^2c^2}, \quad (4.5.26)$$

und der Vergleich mit (4.5.22) liefert

$$\vec{p}_{\text{SPS}}^2 = \frac{m_2^2 c^2}{s} \vec{p}_{\text{LS}}^2. \quad (4.5.27)$$

Für die Mandelstam-Variable t gilt (*nachrechnen!*)

$$t = (p_2 - p_4)^2 = \left(\frac{\mathcal{E}_4}{c} - m_2 c \right)^2 - \vec{p}_4^2 = (m_4^2 + m_2^2) c^2 - 2 m_2 \mathcal{E}_4 \quad (4.5.28)$$

bzw.

$$\mathcal{E}_4 = \frac{(m_2^2 + m_4^2) c^2 - t}{2 m_2}. \quad (4.5.29)$$

Ebenso erhält man aus der Definition der Mandelstam-Variablen u

$$\mathcal{E}_3 = \frac{(m_2^2 + m_3^2) c^2 - u}{2 m_2} \quad (4.5.30)$$

4.6 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.

4.6.1 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 Geschoss 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 Stoß 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

$$(p_1^\mu + p_2^\mu) = \begin{pmatrix} (\mathcal{E}_1 + \mathcal{E}_2)/c \\ 0 \end{pmatrix} = \begin{pmatrix} Mc \\ 0 \end{pmatrix}. \quad (4.6.1)$$

Es ist also

$$M = \frac{\mathcal{E}_1 + \mathcal{E}_2}{c^2} = m_1 + m_2 + \frac{\mathcal{E}_{1\text{kin}} + \mathcal{E}_{2\text{kin}}}{c^2}. \quad (4.6.2)$$

Da die kinetische Energie des ruhenden Teilchens im Endzustand 0 ist, bedeutet dies, dass die gesamte kinetische Energie im Anfangszustand in Wärmeenergie $\Delta Q = \mathcal{E}_{1\text{kin}} + \mathcal{E}_{2\text{kin}}$ umgewandelt worden ist. Gemäß

(4.6.2) 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 Wärmeenergie) zur Gesamtmasse des Körpers beiträgt.

4.6.2 Zerfall eines Pions in zwei Photonen

Ein weiteres Beispiel für die Äquivalenz von Energie und Masse 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 Teilchen- und Kernphysik 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 handele sich im „Teilchen“ eigentlich etwas irreführend ist. 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 (4.4.17) mit $m = m_\gamma = 0$ zusammenhängen

$$\mathcal{E}_\gamma = c |\vec{p}_\gamma|. \quad (4.6.3)$$

Daraus ergibt sich für die Geschwindigkeit in jedem Inertialsystem

$$|\beta_\gamma| = \frac{c |\vec{p}_\gamma|}{\mathcal{E}_\gamma} = 1. \quad (4.6.4)$$

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 (4.6.3) aber sehr wohl, und nur diese Beziehung benötigt man, um die Energie-Impulsbilanz auszuwerten.

Wir können direkt die Beziehung (4.5.10) im Schwerpunktssystem der Photonen (also dem Ruhssystem des zerfallenden Pions) mit $m_1 = m_2 = 0$ verwenden, um den Photonenimpuls zu

$$|\vec{p}_\gamma| = \frac{Mc}{2} \quad (4.6.5)$$

zu bestimmen. Wir haben hier den Extremfall vorliegen, dass die invarianten Massen der Zerfallsprodukte 0 sind aber die Masse im Ausgangszustand durch

$$Mc^2 = \mathcal{E}'_1 + \mathcal{E}'_2 = 2c |\vec{p}_\gamma| \quad (4.6.6)$$

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.

4.7 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

4.7. Geladene Teilchen im elektromagnetischen Feld

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

$$F'^{\mu\nu}(x'^\rho) = \Lambda^\mu_\alpha \Lambda^\nu_\beta F^{\alpha\beta}(x^\gamma) \quad \text{mit} \quad x^\gamma = (\Lambda^{-1})^\gamma_\rho x'^\rho \quad (4.7.1)$$

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 (4.4.25) erfüllt. Da $F^{\mu\nu}$ antisymmetrisch ist, brauchen wir nur den Ansatz

$$m \frac{d^2 x^\mu}{d\tau^2} = K^\mu = q F^{\mu\nu}(x^\rho) \frac{dx_\nu}{d\tau} = q F^{\mu\nu} c u_\nu \quad \text{mit} \quad x_\nu = \eta_{\mu\nu} x^\mu, \quad \underline{u} = \frac{1}{c} \frac{dx}{d\tau} \quad (4.7.2)$$

zu machen. Überschiebt man diese Gleichung mit u_μ ergibt sich sofort (4.4.24), 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 (4.4.25) für die Minkowski-Kraft (4.7.1). Wegen (4.7.2) transformiert sich diese Minkowski-Kraft unter Lorentz-Transformationen auch wie ein Vierervektor (*nachrechnen*), wie es sein muss.

Um die Bewegungsgleichung (4.7.2) besser interpretieren zu können, schreiben wir den Feldstärketensor mit Hilfe der üblichen elektrischen und magnetischen Feldkomponenten \vec{E} und \vec{B} :

$$(F^{\mu\nu}) = \begin{pmatrix} 0 & -E^1/c & -E^2/c & -E^3/c \\ E^1/c & 0 & -B^3 & B^2 \\ E^2/c & B^3 & 0 & -B^1 \\ E^3/c & -B^2 & B^1 & 0 \end{pmatrix}. \quad (4.7.3)$$

Um zu sehen, dass diese Wahl der Anordnung der Feldkomponenten einen Sinn ergibt, schreiben wir die Minkowski-Kraft (4.7.2) explizit hin

$$(K^\mu) = q F^{\mu\nu} c u_\nu = q \begin{pmatrix} \vec{E} \cdot d\vec{x}/d\tau \\ u_0 \vec{E} + d\vec{x}/d\tau \times \vec{B} \end{pmatrix}. \quad (4.7.4)$$

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:

$$\frac{d}{d\tau} \underline{p} = m \frac{d^2}{d\tau^2} \underline{x} = \underline{K} = q \begin{pmatrix} \vec{E} \cdot \vec{u} \\ u_0 \vec{E} + d\vec{x}/d\tau \times \vec{B} \end{pmatrix}, \quad u^\mu = \frac{p^\mu}{mc} = \frac{1}{c} \frac{d}{d\tau} x^\mu. \quad (4.7.5)$$

Da nun $u^0 = u_0 = \gamma = (1 - \beta^2)^{-1/2}$ mit $\vec{\beta} = \vec{v}/c$, $\vec{v} = d\vec{x}/dt = c\vec{u}/u^0$, ergibt sich für den räumlichen Anteil der ersten Gleichung

$$\frac{1}{u_0} \frac{d}{d\tau} \vec{p} = \frac{d}{dt} \vec{p} = q (\vec{E} + \vec{v} \times \vec{B}), \quad (4.7.6)$$

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, \vec{E} die elektrischen und \vec{B} die magnetischen Feldkomponenten. Für die räumlichen Koordinaten der zweiten Gleichung in (4.7.5) bemerken wir, dass

$$\vec{u} = \frac{1}{mc} \vec{p} = \frac{1}{c} \frac{d}{d\tau} \vec{x}. \quad (4.7.7)$$

Dividiert man dies durch $u^0 = p^0/(mc)$ erhalten wir

$$\vec{v} = \frac{d}{dt} \vec{x} = c \frac{\vec{p}}{p^0}. \quad (4.7.8)$$

Wie wir oben bei der allgemeinen Betrachtung gesehen haben, ist die Zeitkomponente in (4.7.2) automatisch für die Lösungen der Bewegungsgleichungen (4.7.6) erfüllt, denn multiplizieren wir (4.7.6) mit \vec{v} , folgt

$$\vec{v} \cdot \frac{d}{dt} \vec{p} = q \vec{v} \cdot \vec{E}, \quad (4.7.9)$$

und dies ist genau die durch u^0 dividierte zeitliche Komponente von (4.7.2). Damit ist klar, dass die Bewegungsgleichung mit der Energie-Impuls-Beziehung (4.4.17) verträglich ist, wie es aufgrund unserer Konstruktion der Minkowski-Kraft zu erwarten ist.

4.7.1 Teilchen im homogenen elektrischen Feld

Eines des 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 ist⁵ $\vec{E} = (E, 0, 0)^T = \text{const}$ 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 (4.7.6) lässt sich dann sehr leicht lösen

$$\frac{d}{dt} p^1 = \frac{q}{m} E \Rightarrow p^1(t) = qEt. \quad (4.7.10)$$

Die beiden anderen Komponenten des Impulses sind Erhaltungsgrößen und daher aufgrund der Anfangsbedingungen $p^2 = p^3 = 0$. Aus der Energie-Impuls-Beziehung (4.4.17) folgt für die zeitliche Komponente des Viererimpulses

$$p^0(t) = \frac{\mathcal{E}}{c} = \sqrt{m^2 c^2 + (qEt)^2} \quad (4.7.11)$$

und daher

$$v^1(t) = \frac{dx^1}{dt} = \frac{c p^1}{p^0} = \frac{qcEt}{\sqrt{m^2 c^2 + q^2 E^2 t^2}}. \quad (4.7.12)$$

Dies lässt sich unter Berücksichtigung der Anfangsbedingung leicht nochmals nach der Zeit integrieren, wodurch wir die Lösung

$$x^1(t) = \frac{mc^2}{qE} \left(\sqrt{1 + \frac{q^2 E^2 t^2}{m^2 c^2}} - 1 \right) \quad (4.7.13)$$

erhalten.

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 (4.7.13) entwickeln (*nachrechnen*) und erhalten

$$x^1(t) = \frac{mc^2}{qE} \left(\frac{q^2 E^2 t^2}{2m^2 c^2} + \mathcal{O} \left[\left(\frac{qEt}{mc} \right)^4 \right] \right) = \frac{qEt^2}{2m} + \dots \quad (4.7.14)$$

In führender Ordnung erhalten wir die erwartete Zeitabhängigkeit für die nichtrelativistische Bewegung in einem konstanten Kraftfeld.

⁵Man 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.

4.7. Geladene Teilchen im elektromagnetischen Feld

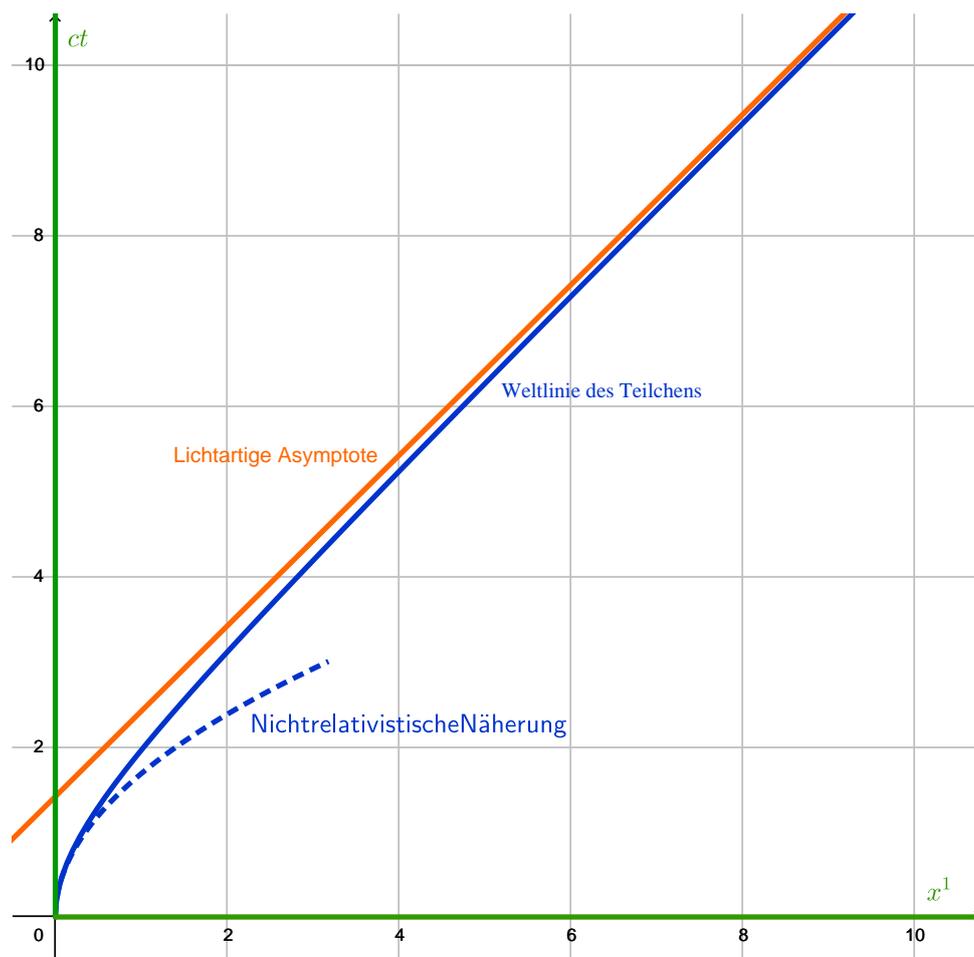


Abbildung 4.4: Minkowski-Diagramm für die „hyperbolische Bewegung“ in einem homogenen elektrischen Feld. Die entsprechende Weltlinie ist eine Hyperbel und besitzt eine lichtartige Asymptote. Dies zeigt, dass ein Lichtsignal, das vom Ursprung zu einer Zeit $t > t_H = \frac{mc}{qE}$ ausgesandt wird, das 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).

Wir können auch den relativistischen Limes betrachten, wo $t \gg mc/(qE)$ ist. Dann folgt durch die entsprechend angepasste Entwicklung der Wurzel

$$x^1(t) = \frac{mc^2}{qE} \left(\frac{qEt}{mc} \sqrt{1 + \frac{m^2 c^2}{q^2 E^2 t^2}} - 1 \right) \underset{t \rightarrow \infty}{\cong} c \left(t - \frac{mc}{qE} \right). \quad (4.7.15)$$

Das impliziert auch, dass ein in Sect. 4.1 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. 4.4).

4.8 Kovariante Formulierung der Elektrodynamik

Einsteins Ausgangspunkt zur Entwicklung der speziellen Relativitätstheorie war die mangelnde Kovarianz der **Maxwell-Gleichungen** unter Galilei-Transformationen. Wie wir oben gesehen haben, führt die Annahme, dass die Elektrodynamik in Form der Maxwell-Gleichungen dem speziellen Relativitätsprinzip genügen, zu dem Schluss, dass die Beschreibung von Raum und Zeit geändert werden und am bequemsten mittels des oben eingeführten vierdimensionalen **Minkowski-Raums** formuliert wird.

In diesem Abschnitt wollen wir nun zeigen, dass die Maxwell-Gleichungen tatsächlich forminvariant unter Lorentz-Transformationen sind. Dies geschieht dadurch, dass wir die Maxwell-Gleichungen mit Hilfe von **Vierertensorfeldern** formulieren. Wie wir bereits im vorigen Abschnitt gesehen haben, können nämlich die elektromagnetischen Kraftwirkungen auf Probeladungen im Viererformalismus der relativistischen geschrieben werden, indem die Felder \vec{E} und \vec{B} der herkömmlichen mit euklidischen Dreivektoren formulierten Elektrodynamik zu dem antisymmetrischen Faraday-Tensor (4.7.3) zusammengefasst werden.

Die Maxwell-Gleichungen im Vakuum lauten im üblichen 3D-Formalismus bekanntlich

$$\vec{\nabla} \times \vec{E} + \partial_t \vec{B} = \vec{0}, \quad (4.8.1)$$

$$\vec{\nabla} \cdot \vec{B} = 0, \quad (4.8.2)$$

$$\vec{\nabla} \times \vec{B} - \epsilon_0 \mu_0 \partial_t \vec{E} = \mu_0 \vec{j}, \quad (4.8.3)$$

$$\vec{\nabla} \cdot \vec{E} = \frac{1}{\epsilon_0} \rho, \quad (4.8.4)$$

wobei ρ und \vec{j} die elektrischen **Ladungs- und Stromdichten** bezeichnen.

Um diese Gleichungen in eine Form bringen zu können, die nur mit Vierertensorfeldern operiert, müssen wir uns zunächst mit der **vierdimensionalen Vektor- bzw. Tensoranalysis** im Minkowski-Raum beschäftigen.

4.8.1 Tensoranalysis und Maxwell-Gleichungen im Minkowski-Raum

Die Tensoranalysis im vierdimensionalen Minkowskiraum lässt sich in weiten Teilen analog zur Tensoranalysis im dreidimensionalen euklidischen Raum entwickeln. Dabei bietet es sich an, die Gleichungen im **Ricci-Kalkül** zu formulieren. Im Gegensatz zum euklidischen Raum ist dabei jedoch auf die Index-Stellung zu achten.

Wir gehen wieder vom Raum-Zeit-Vierervektor für Ereignisse mit Komponenten $(x^\mu) = (x^0, x^1, x^2, x^3)^T$ bzgl. eines Inertialsystems Σ aus. Dabei ist $x^0 = ct$. Wie wir oben ausführlich erörtert haben, transformieren sich diese Komponenten eines Vierervektors mit oberen Indizes, die sog. **kontravarianten Komponente** unter Lorentz-Transformationen als

$$x'^\mu = \Lambda^\mu_\nu x^\nu. \quad (4.8.5)$$

Dabei bilden die Λ^μ_ν eine Matrix, die das Minkowskiprodukt beliebiger Vierervektoren \underline{x} und \underline{y} ungeändert lässt, d.h.

$$\eta_{\mu\nu} x^\mu y^\nu = \eta_{\mu\nu} x'^\mu y'^\nu, \quad (\eta_{\mu\nu}) = \text{diag}(1, -1, -1, -1). \quad (4.8.6)$$

Daraus ergibt sich die Bedingung an die Matrix $\hat{\Lambda}$, eine Lorentz-Transformation zu sein, zu

$$\eta_{\mu\nu} \Lambda^\mu_\rho \Lambda^\nu_\sigma = \eta_{\rho\sigma}. \quad (4.8.7)$$

Zur Bequemlichkeit führt man nun auch Vektorkomponenten mit unteren Indizes, die **kovarianten Komponenten** ein, indem man definiert

$$x_\mu = \eta_{\mu\nu} x^\nu. \quad (4.8.8)$$

4.8. Kovariante Formulierung der Elektrodynamik

Da $\hat{\eta}^{-1} = \hat{\eta}$ ist, kann man mit $\eta^{\mu\nu} = \eta_{\mu\nu}$ umgekehrt die Indizes auch wieder heraufziehen:

$$x^\mu = \eta^{\mu\nu} x_\nu. \quad (4.8.9)$$

Entsprechend kann man das Minkowskiprodukt auf verschiedene Arten mit den ko- bzw. kontravarianten Indizes schreiben:

$$\underline{x} \cdot \underline{y} = \eta_{\mu\nu} x^\mu y^\nu = x_\nu y^\nu = x^\mu y_\mu = \eta^{\mu\nu} x_\mu y_\nu. \quad (4.8.10)$$

Jetzt wollen wir das Transformationsverhalten der kovarianten Vektorkomponenten bestimmen:

$$x'_\mu = \eta_{\mu\rho} x'^\rho = \eta_{\mu\rho} \Lambda^\rho_\sigma x^\sigma = \eta_{\mu\rho} \Lambda^\rho_\sigma \eta^{\sigma\nu} x_\nu. \quad (4.8.11)$$

Wendet man die Regeln zum Indexziehen, auch auf die Komponenten der Lorentz-Transformationsmatrix an, folgt

$$x'_\mu = \Lambda_\mu^\nu x_\nu. \quad (4.8.12)$$

Andererseits folgt in dieser Schreibweise aus (4.8.7)

$$\Lambda_{\nu\rho} \Lambda^\nu_\sigma = \eta_{\rho\sigma}. \quad (4.8.13)$$

Zieht man nun den Index ρ mittels $\eta^{\rho\alpha}$ nach oben, ergibt sich daraus wiederum

$$\Lambda_\nu^\alpha \Lambda^\nu_\sigma = \eta^{\rho\alpha} \eta_{\rho\sigma} = \delta^\alpha_\sigma \Rightarrow (\hat{\Lambda}^{-1})^\alpha_\nu = \Lambda_\nu^\alpha. \quad (4.8.14)$$

Demnach gilt

$$x^\alpha = \delta^\alpha_\sigma x^\sigma = \Lambda_\nu^\alpha \Lambda^\nu_\sigma x^\sigma \stackrel{(4.8.5)}{=} \Lambda_\nu^\alpha x'^\nu. \quad (4.8.15)$$

Wollen wir nun **partielle Ableitungen** irgendwelcher Funktionen $\Phi(x^\mu)$ durch partielle Ableitungen nach den x'^μ ausdrücken, folgt aus der Kettenregel

$$\frac{\partial \Phi}{\partial x'^\mu} = \frac{\partial \Phi}{\partial x^\nu} \frac{\partial x^\nu}{\partial x'^\mu} \stackrel{(4.8.15)}{=} \frac{\partial \Phi}{\partial x^\nu} \Lambda_\mu^\nu. \quad (4.8.16)$$

Vergleichen wir dies mit der Transformationsformel (4.8.12), ergibt sich, dass die partiellen Ableitungen eines **skalaren Feldes** nach den kontravarianten Komponenten offenbar kovariante Komponenten von Vektorfeldern ergeben, d.h. wir müssen den vierdimensionalen Gradienten im Ricci-Kalkül entsprechend mit einem *unteren Indizes* notieren:

$$\frac{\partial \Phi}{\partial x^\nu} = \partial_\nu \Phi. \quad (4.8.17)$$

Ein viererskalares Feld beschreibt dabei eine Größe, die nicht vom Bezugssystem abhängt, z.B. die Temperatur als Funktion des Ortes und der Zeit. Demnach gilt für eine solche Größe die Transformationsformel

$$\Phi'(\underline{x}') = \Phi(\underline{x}) = \Phi(\hat{\Lambda}^{-1} \underline{x}'). \quad (4.8.18)$$

Entsprechend gilt für ein **Vierervektorfeld**

$$A'^\mu(\underline{x}') = \Lambda^\mu_\rho A^\rho(\underline{x}) = \Lambda^\mu_\rho A^\rho(\hat{\Lambda}^{-1} \underline{x}'). \quad (4.8.19)$$

Demnach ist

$$T_\mu^\nu = \partial_\mu A^\nu \quad (4.8.20)$$

ein **Tensorfeld 2. Stufe**. Die Transformationvorschrift ergibt sich aus der Indexstellung. Das rechnet man auch leicht direkt nach:

$$T'^{\rho\sigma} = \partial'_\rho A'^\sigma = \frac{\partial x^\mu}{\partial x'^\rho} \Lambda^\sigma_\nu \partial_\mu A^\nu = \Lambda_\rho^\mu \Lambda^\sigma_\nu T_\mu^\nu. \quad (4.8.21)$$

Wendet man also partielle Ableitungen auf irgendwelche Komponenten von Tensorfeldern n -ter Stufe an, ergeben sich wieder Komponenten eines Tensors $(n + 1)$ -ter Stufe. Leitet man nach x^μ ab, ergibt sich ein unterer zusätzlicher Index μ . Entsprechend liefert Ableiten nach x_μ einen zusätzlichen oberen Index μ .

Weiter darf man natürlich auch über Indexpaare kontrahieren. So ergibt sich die **vierdimensionale Divergenz** eines Vektorfeldes zu

$$\Phi = \partial_\mu A^\mu. \quad (4.8.22)$$

Dies sollte nun natürlich ein Skalarfeld sein. In der Tat ergibt sich aus (4.8.21) (*Nachrechnen!*)

$$\Phi' = \partial'_\rho A'^\rho = \Lambda_\rho^\mu \Lambda^\rho_\nu \partial_\mu A^\nu \stackrel{(4.8.14)}{=} \delta^\mu_\nu \partial_\mu A^\nu = \partial_\mu A^\mu = \Phi. \quad (4.8.23)$$

Jetzt können wir bereits die erste kovariante Beziehung für die Maxwell-Gleichungen erraten. Im Dreierformalismus sind dies Differentialgleichungen erster Ordnung der Feldkomponenten (\vec{E}, \vec{B}) . Diese sind im Viererformalismus zu dem Faraday-Tensor (4.7.3) zusammengefasst. Nun liefert zum Einen $\partial_\mu F^{\mu\nu}$ einen Differentialausdruck mit 1. Ableitungen und zum Anderen Vektorkomponenten. Schreiben wir dies nach Raum- und Zeit-Komponenten aus und beachten, dass $\partial_0 = \partial / \partial x^0 = (1/c) \partial_t$, folgt⁶

$$(\partial_\mu F^{\mu\nu}) = \begin{pmatrix} \partial_m F^{m0} \\ \partial_0 F^{01} - \partial_m F^{m1} \\ \partial_0 F^{02} - \partial_m F^{m2} \\ \partial_0 F^{03} - \partial_m F^{m3} \end{pmatrix} = \begin{pmatrix} \vec{\nabla} \cdot \vec{E} / c \\ -\partial_t E^1 / c^2 + \partial_2 B^3 - \partial_3 B^2 \\ -\partial_t E^2 / c^2 + \partial_3 B^1 - \partial_1 B^3 \\ -\partial_t E^3 / c^2 + \partial_1 B^2 - \partial_2 B^3 \end{pmatrix} = \begin{pmatrix} \vec{\nabla} \cdot \vec{E} / c \\ -\partial_t E^1 / c^2 + (\vec{\nabla} \times \vec{B})^1 \\ -\partial_t E^2 / c^2 + (\vec{\nabla} \times \vec{B})^2 \\ -\partial_t E^3 / c^2 + (\vec{\nabla} \times \vec{B})^3 \end{pmatrix}. \quad (4.8.24)$$

Vergleicht man dies nun mit den beiden inhomogenen Maxwell-Gleichungen (4.8.3) und (4.8.4) und beachtet, dass $c^2 = 1/(\epsilon_0 \mu_0)$ ist, folgt

$$(\partial_\mu F^{\mu\nu}) = \begin{pmatrix} c\rho / (\epsilon_0 c^2) \\ \mu_0 j^1 \\ \mu_0 j^2 \\ \mu_0 j^3 \end{pmatrix} = \mu_0 \begin{pmatrix} c\rho \\ j^1 \\ j^2 \\ j^3 \end{pmatrix}. \quad (4.8.25)$$

Damit wird klar, dass offenbar

$$\partial_\mu F^{\mu\nu} = \mu_0 j^\nu \quad (4.8.26)$$

eine kovariante Gleichung ist. Dazu muss man nur annehmen, dass sich $F^{\mu\nu}$ entsprechend Tensorfeldkomponenten transformiert. Dies darf man selbstverständlich annehmen, denn es definiert lediglich, wie sich die Komponenten des elektromagnetischen Feldes unter Lorentz-Transformationen transformieren. Aus (4.8.26) folgt dann zwingend, dass die Komponenten des **elektrischen Viererstroms** $\underline{j} = (j^\nu) = (c\rho, \vec{j})^T$ eben Komponenten eines Vierervektorfeldes bilden müssen und sich entsprechend transformieren. Wir werden diese Transformationen im nächsten Kapitel noch ausführlich betrachten.

Um die Viererformulierung der Maxwell-Gleichungen zu komplettieren, benötigen wir noch Gleichungen, die den homogenen Maxwell-Gleichungen (4.8.1) und (4.8.2) entsprechen. Diese sind ebenfalls vier lineare Differentialgleichungen 1. Ordnung bzgl. der Ableitungen ∂_μ und involviert nur die Feldkomponenten \vec{E} und \vec{B} .

Nun erinnern wir uns, dass im Dreierformalismus das Levi-Civita-Symbol ϵ_{jkl} invariante Tensorkomponenten unter orientierungserhaltenden Drehungen sind. Entsprechend definieren wir das vierdimensionale **Levi-Civita-Symbol** dadurch, dass $\epsilon^{0123} = 1$ ist und $\epsilon^{\mu\nu\rho\sigma}$ unter beliebigen Vertauschungen zweier Indizes das Vorzeichen wechselt. Nehmen zwei Indizes den gleichen Wert an, ergibt sich insbesondere $\epsilon^{\mu\nu\rho\sigma} = 0$. Das Levi-Civita-Symbol mit unteren Indizes ergibt sich definitionsgemäß nach der Regel zum Indexziehen zu

$$\epsilon_{\alpha\beta\gamma\delta} = \eta_{\alpha\mu} \eta_{\beta\nu} \eta_{\gamma\rho} \eta_{\delta\sigma} \epsilon^{\mu\nu\rho\sigma} = \det(\hat{\eta}) \epsilon^{\alpha\beta\gamma\delta} = -\epsilon^{\alpha\beta\gamma\delta}. \quad (4.8.27)$$

⁶Im Folgenden laufen definitionsgemäß mit lateinischen Buchstaben bezeichnete Indizes von 1 bis 3.

Das Transformationsverhalten unter Lorentz-Transformationen ergibt sich entsprechend zu (*Nachrechnen!*)

$$\epsilon'^{\alpha\beta\gamma\delta} = \det(\hat{\Lambda})\epsilon^{\alpha\beta\gamma\delta}, \quad (4.8.28)$$

d.h. das Levi-Civita-Symbol verhält sich unter Lorentztransformationen mit $\det(\hat{\Lambda}) = +1$ wie die invarianten Komponenten eines Tensors 4. Stufe, wechselt jedoch das Vorzeichen für Lorentz-Transformationen mit $\det(\hat{\Lambda}) = -1$. Insbesondere sind aber reine Lorentz-Boosts der Art (4.1.25) Lorentz-Transformationen mit $\det(\hat{\Lambda}) = +1$, und für diese transformiert sich das Levi-Civita-Symbol demnach wie Tensorkomponenten.

Mit dem Levi-Civita-Symbol können wir demnach aus dem Faraday-Tensor den dazu **dualen Faraday-Tensor** bilden:

$$[(\dagger F)_{\mu\nu}] = \frac{1}{2}(\epsilon_{\mu\nu\rho\sigma}F^{\rho\sigma}) = \begin{pmatrix} 0 & B^1 & B^2 & B^3 \\ -B^1 & 0 & E^3/c & E^2/c \\ -B^2 & -E^3/c & 0 & E^1/c \\ -B^3 & E^2/c & -E^1/c & 0 \end{pmatrix}. \quad (4.8.29)$$

Daraus erhalten wir

$$[\partial^\mu(\dagger F)_{\mu\nu}] = \begin{pmatrix} \vec{\nabla} \cdot \vec{B} \\ \partial_t B^1/c - \partial_2 E^3/c + \partial_3 E^2/c \\ \partial_t B^2/c - \partial_1 E^3/c + \partial_3 E^1/c \\ \partial_t B^3/c + \partial_1 E^2/c - \partial_2 E^1/c \end{pmatrix} = \begin{pmatrix} \vec{\nabla} \cdot \vec{B} \\ \partial_t \vec{B}/c + \vec{\nabla} \times \vec{E}/c \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}. \quad (4.8.30)$$

Dabei haben wir im letzten Schritt die homogenen Maxwell-Gleichungen (4.8.1) und (4.8.2) verwendet.

Die vier im Dreierformalismus geschriebenen Maxwell-Gleichungen (4.8.1-4.8.4) sind also zu den beiden Gleichungen (4.8.28) und (4.8.30)

$$\partial_\mu F^{\mu\nu} = \mu_0 j^\nu, \quad (4.8.31)$$

$$\partial^\mu(\dagger F)_{\mu\nu} = 0 \quad (4.8.32)$$

äquivalent.

Da partielle Ableitungen kommutieren, also $\partial_\mu \partial_\nu = \partial_\nu \partial_\mu$ gilt und $F^{\mu\nu} = -F^{\nu\mu}$ ist, folgt aus (4.8.31)

$$\partial_\nu \partial_\mu F^{\mu\nu} = \mu_0 \partial_\nu j^\nu = 0. \quad (4.8.33)$$

Damit also die Maxwell-Gleichungen überhaupt eine Lösung haben können, muss für den elektrischen Viererstrom die Bedingung

$$\partial_\nu j^\nu = \partial_0(c\rho) + \vec{\nabla} \cdot \vec{j} = \partial_t + \vec{\nabla} \cdot \vec{j} = 0 \quad (4.8.34)$$

gelten, also die **Kontinuitätsgleichung**. Aus dem Dreierformalismus wissen wir, dass dies der Erhaltung der Gesamtladung entspricht. Wir kommen darauf unten bei der Behandlung der diversen Integrale im Viererformalismus noch zurück.

4.8.2 Die elektromagnetischen Potentiale

Vom Dreierformalismus her wissen wir, dass die homogenen Maxwellgleichungen die Existenz eines skalaren und eines Vektorpotentials (ϕ bzw. \vec{A}) garantieren, d.h. dass die Felder durch

$$\vec{E} = -\partial_t \vec{A} - \vec{\nabla} \phi, \quad \vec{B} = \vec{\nabla} \times \vec{A} \quad (4.8.35)$$

dargestellt werden. Betrachten wir die Gleichung für das \vec{B} -Feld, liegt es nahe, dass man dies in den Viererformalismus übersetzen kann, indem man ein Vierervektorfeld A^μ einführt und den Ansatz

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu \quad (4.8.36)$$

macht.

Trennen wir dies in Zeit- und Raum-Komponenten, ergibt sich mit (4.7.3)⁷

$$F_{0n} = -F^{0n} = E^n/c = \partial_t A_n/c - \partial_n A_0 = -\partial_t A^n/c - \partial_n A_0, \quad (4.8.37)$$

$$F_{mn} = F^{mn} = -\epsilon^{lmn} B^l = -\epsilon^{lmn} (\vec{\nabla} \times \vec{A})^l = \partial_m A_n - \partial_n A_m = -(\partial_m A^n - \partial_n A^m). \quad (4.8.38)$$

In der Tat ergeben sich demnach die vier Komponenten des **Viererpotentials** \underline{A} aus dem Skalar- und Dreiervektorpotential des Dreierformalismus:

$$\underline{A} = (A^\mu) = \begin{pmatrix} \phi/c \\ \vec{A} \end{pmatrix}. \quad (4.8.39)$$

Dabei ist \underline{A} nur bis auf den Vierergradienten eines Viererskalarfeldes bestimmt, denn setzen wir

$$A'_\mu = A_\mu + \partial_\mu \chi, \quad (4.8.40)$$

folgt wegen der Vertauschbarkeit der partiellen Ableitungen

$$F'_{\mu\nu} = \partial_\mu A'_\nu - \partial_\nu A'_\mu = (\partial_\mu A_\nu - \partial_\nu A_\mu) + (\partial_\mu \partial_\nu \chi - \partial_\nu \partial_\mu \chi) = F_{\mu\nu}, \quad (4.8.41)$$

d.h. für jedes beliebige Skalarfeld χ beschreiben die Viererpotentiale A'_μ und A_μ dieselben physikalisch beobachtbaren Felder $F_{\mu\nu}$ (bzw. \vec{E} und \vec{B}). Auch dies kennen wir bereits als **Eichinvarianz** der Elektrodynamik.

Setzen wir nun (4.8.36) in die homogenen Maxwell-Gleichungen (4.8.32) ein, folgt

$$\partial^\mu F_{\mu\nu} = \partial^\mu (\partial_\mu A_\nu - \partial_\nu A_\mu) = \square A_\nu - \partial_\nu \partial_\mu A^\mu = \mu_0 j_\nu. \quad (4.8.42)$$

Dabei ist

$$\square = \partial_\mu \partial^\mu = \partial_0^2 - \vec{\nabla}^2 = \frac{1}{c^2} \partial_t^2 - \Delta \quad (4.8.43)$$

der bereits vom Dreierformalismus her bekannte **d' Alembert-Operator**.

Im Dreierformalismus hatten wir die Integration der Gleichungen (4.8.42) für die Potentiale wesentlich vereinfacht, indem wir die Eichinvarianz der Potentiale ausgenutzt haben, um an die Potentiale geeignete Nebenbedingungen zu stellen bzw. eine **Eichbedingung** zu stellen. In der Tat separieren die Komponenten des Viererpotentials in (4.8.41), wenn wir die **Lorenz-Eichbedingung**

$$\partial_\mu A^\mu = 0 \Leftrightarrow \frac{1}{c^2} \partial_t \phi + \vec{\nabla} \cdot \vec{A} = 0. \quad (4.8.44)$$

erfüllen. Dann vereinfacht sich nämlich zum Einen (4.8.42) zu vier getrennten **Wellengleichungen** für die Komponenten des Viererpotentials,

$$(\square A^\nu) = \begin{pmatrix} \square \phi/c \\ \square \vec{A} \end{pmatrix} = \mu_0 \begin{pmatrix} c\rho \\ \vec{j} \end{pmatrix} \quad (4.8.45)$$

bzw.

$$\square \phi = \mu_0 c^2 \rho = \frac{1}{\epsilon_0} \rho, \quad \square \vec{A} = \mu_0 \vec{j}. \quad (4.8.46)$$

Dies stimmt mit dem entsprechenden Resultat im Dreierformalismus überein (vgl. [Hee18b], Abschnitt 5.6). Zum Anderen ist der Vorteil der Lorenz-Eichbedingung (4.8.44), dass es sich um eine vom Bezugssystem unabhängige Größe handelt, denn offensichtlich ist $\partial_\mu A^\mu$ ein Skalarfeld, d.h. gilt die Lorenz-Eichbedingung in einem Inertialsystem, so ist sie automatisch auch in jedem anderen Inertialsystem erfüllt.

⁷Wir bemerken, dass für den Levi-Civita-Tensor im dreidimensionalen Euklidischen Raum nach wie vor $\epsilon_{lmn} = \epsilon^{lmn}$ mit $\epsilon_{123} = +1$ gilt.

4.8.3 Transformationsverhalten der Felder

Jetzt betrachten wir das Transformationsverhalten der Felder unter Lorentz-Boosts in x^1 -Richtung. Dazu gehen wir vom Faraday-Tensor (4.7.3) aus:

$$(F^{\mu\nu}) = \begin{pmatrix} 0 & -E^1/c & -E^2/c & -E^3/c \\ E^1/c & 0 & -B^3 & B^2 \\ E^2/c & B^3 & 0 & -B^1 \\ E^3/c & -B^2 & B^1 & 0 \end{pmatrix}. \quad (4.8.47)$$

Die Tensorkomponenten transformieren sich gemäß der entsprechenden Vorschrift analog zu (4.8.21), nur dass wir jetzt zwei kontravariante (obere) Indizes vorliegen haben:

$$F'^{\mu\nu}(\underline{x}') = \Lambda^\mu{}_\rho \Lambda^\nu{}_\sigma F^{\rho\sigma}(\underbrace{\hat{\Lambda}^{-1}\underline{x}'}_{=\underline{x}}). \quad (4.8.48)$$

Dabei ist $\hat{\Lambda} = (\Lambda^\mu{}_\nu)$ die Lorentz-Boost-Matrix (4.1.25). In Matrix-Vektor-Schreibweise entspricht dies der Formel

$$\hat{F}' = \hat{\Lambda} \hat{F} \hat{\Lambda}^T. \quad (4.8.49)$$

Führt man die Matrizenmultiplikationen aus (*Übung!*), kann man die Komponenten im Dreierformalismus ablesen:

$$\begin{aligned} E'^1 &= cF'^{10} = E^1, & E'^2 &= cF'^{20} = \gamma(E^2 - \beta c B^3), & E'^3 &= cF'^{30} = \gamma(E^3 + \beta c B^2), \\ B'^1(\underline{x}') &= F'^{32} = B^1, & B'^2 &= F'^{13} = \gamma(B^2 + \beta E^3/c), & B'^3 &= F'^{21} = \gamma(B^3 - \beta E^2/c). \end{aligned} \quad (4.8.50)$$

Manchmal ist es einfacher, zuerst das Viererpotential (4.8.39) zu transformieren:

$$\underline{A}'(\underline{x}') = \begin{pmatrix} \phi'/c \\ A'^1 \\ A'^2 \\ A'^3 \end{pmatrix} = \hat{\Lambda} \underline{A}(\underline{x}) = \begin{pmatrix} \gamma(A^0 - \beta A^1) \\ \gamma(A^1 - \beta A^0) \\ A^2 \\ A^3 \end{pmatrix} \quad (4.8.51)$$

und dann die entsprechenden Ableitungen des Dreierformalismus zu verwenden

$$\vec{E}'(\underline{x}') = -\partial'_\nu \vec{A}'(\underline{x}') - \vec{\nabla}' \phi'(\underline{x}'), \quad \vec{B}'(\underline{x}') = \vec{\nabla}' \times \vec{A}'(\underline{x}'). \quad (4.8.52)$$

4.8.4 Feld einer gleichförmig bewegten Punktladung

Mit der Lorentz-Transformation lässt sich das Feld einer Punktladung, die sich mit der konstanten Geschwindigkeit $v = \beta c$ in x^1 -Richtung bewegt, leicht berechnen. Wir kennen nämlich das Feld im Bezugssystem Σ' , das sich gegenüber Σ mit eben der Geschwindigkeit v in x^1 -Richtung bewegt, denn in Σ' ruht offenbar die Ladung. Wir legen den Ursprung des räumlichen Koordinatensystems von Σ' an den Ort der Ladung. Dort ist das elektromagnetische Feld durch das **Coulomb-Feld** gegeben:

$$\vec{E}'(\underline{x}') = \frac{q}{4\pi\epsilon_0} \frac{\vec{x}'}{r'^3}, \quad \vec{B} = \vec{0}, \quad r' = |\vec{x}'|. \quad (4.8.53)$$

Nun benötigen wir die Umkehrtransformation von (4.8.50). Diese erhalten wir aus (4.8.50), indem wir dort überall β durch $-\beta$ ersetzen. Damit wird

$$\vec{E} = \begin{pmatrix} E'^1 \\ \gamma E'^2 \\ \gamma E'^3 \end{pmatrix}, \quad \vec{B} = \begin{pmatrix} 0 \\ -\gamma\beta E'^3/c \\ \gamma\beta E'^2/c \end{pmatrix} = \vec{\beta} \times \vec{E}/c. \quad (4.8.54)$$

Dabei müssen wir noch die Argumente der Felder durch die Raumzeit-Koordinaten \underline{x} ausdrücken:

$$\underline{x}' = \begin{pmatrix} ct' \\ x'^1 \\ x'^2 \\ x'^3 \end{pmatrix} = \begin{pmatrix} \gamma(ct - \beta x^1) \\ \gamma(x^1 - \beta ct) \\ x^2 \\ x^3 \end{pmatrix}. \quad (4.8.55)$$

Demnach wird

$$r' = \sqrt{\gamma^2(x^1 - \beta ct)^2 + (x^2)^2 + (x^3)^2} \quad (4.8.56)$$

und somit

$$\begin{aligned} \vec{E}' &= \frac{q\gamma}{4\pi\epsilon_0\sqrt{\gamma^2(x^1 - \beta ct)^2 + (x^2)^2 + (x^3)^2}} \begin{pmatrix} x^1 - \beta ct \\ x^2 \\ x^3 \end{pmatrix}, \\ \vec{B}' &= \frac{q\gamma\beta}{4\pi\epsilon_0 c\sqrt{\gamma^2(x^1 - \beta ct)^2 + (x^2)^2 + (x^3)^2}} \begin{pmatrix} 0 \\ -x^3 \\ x^2 \end{pmatrix}. \end{aligned} \quad (4.8.57)$$

4.8.5 Dopplereffekt und Aberration bei Licht

Wir können nun leicht die Effekte der Bewegung von Lichtquellen auf das emittierte Licht studieren. Dazu betrachten wir eine sehr weit entfernte Lichtquelle, die sich bzgl. des Systems Σ mit der Geschwindigkeit $v = \beta c$ bewegt und dementsprechend im System Σ' ruht. Wir denken uns Alice (Ruhsystem Σ) und Bob (Ruhsystem Σ') als sehr weit von der Lichtquelle entfernte Beobachter, sodass beide das Licht als ebene elektromagnetische Welle beschreiben können.

Im vorigen Semester haben wir gesehen, dass eine ebene elektromagnetische Welle durch einen Wellenvektor $\vec{k} = k\vec{n}$, der die Ausbreitungsrichtung der Welle charakterisiert, bestimmt ist. Es ist klar, dass in unserem Szenario \vec{n} der Einheitsvektor ist, der von der Quelle zum Beobachter weist. Für die Kreisfrequenz der Welle ergab sich dann aus der Wellengleichung für die Feldkomponenten $\omega = ck$. Das elektrische und magnetische Feld sind beide transversal. Für eine linear polarisierte Welle ist (vgl. [Hee18b], Abschnitt 5.1)

$$\vec{E}(t, \vec{x}) \equiv \vec{E}(\underline{x}) = \vec{E}_0 \cos[k(ct - \vec{n} \cdot \vec{x})], \quad \vec{B}(\underline{x}) = \frac{1}{c} \vec{n} \times \vec{E}(\underline{x}). \quad (4.8.58)$$

Dabei ist $\vec{E}_0 \mathbb{R}^3$ ein beliebiger zu \vec{n} orthogonaler Vektor, d.h. es gilt $\vec{n} \cdot \vec{E}_0 = 0$. Wir bemerken, dass wir das Argument im cos, offenbar kovariant schreiben können, indem wir den Viererwellenvektor

$$\underline{k} = (k^\mu) = \begin{pmatrix} k \\ k\vec{n} \end{pmatrix} = \begin{pmatrix} \omega/c \\ k\vec{n} \end{pmatrix} \quad (4.8.59)$$

definieren. Wir bemerken, dass $\underline{k} \cdot \underline{k} = k^2 - k^2\vec{n}^2 = 0$ ist, d.h. \underline{k} ist ein **lichtartiger Vektor**. Das muss auch so sein, denn dies bedeutet, dass die Phasengeschwindigkeit der elektromagnetischen Welle c ist, wie es im Vakuum der Fall sein muss. Mit dem Wellenvektor \underline{k} wird (4.8.59)

$$\vec{E}(\underline{x}) = \vec{E}_0 \cos(\underline{k} \cdot \underline{x}), \quad \vec{B}(\underline{x}) = \frac{1}{c} \vec{n} \times \vec{E}(\underline{x}). \quad (4.8.60)$$

Betrachten wir nun die Situation von Bobs Standpunkt aus, also im Ruhsystem der Lichtquelle. Er verwendet zur Beschreibung des elektromagnetischen Feldes die Komponenten

$$\vec{E}'(\underline{x}') = \vec{E}'_0 \cos(\underline{k}' \cdot \underline{x}'). \quad (4.8.61)$$

4.8. Kovariante Formulierung der Elektrodynamik

Ohne Beschränkung der Allgemeinheit können wir annehmen, dass er sein Koordinatensystem so wählt, dass $\vec{n}' = (\cos \alpha', \sin \alpha', 0)^T$ und $\vec{E}'_0 = E'_{01} \vec{\epsilon}'_1 + E'_{02} \vec{\epsilon}'_2$ ist. Dabei sind

$$\vec{\epsilon}'_1 = \begin{pmatrix} -\sin \alpha' \\ \cos \alpha' \\ 0 \end{pmatrix}, \quad \vec{\epsilon}'_2 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \quad (4.8.62)$$

zwei zu \vec{n}' senkrechte Einheitsvektoren mit $\vec{\epsilon}'_1 \times \vec{\epsilon}'_2 = \vec{n}'$ (Nachrechnen!). Dann ist $\vec{B}'_0 = \vec{n}' \times \vec{E}' = (E'_{01} \vec{\epsilon}'_2 - E'_{02} \vec{\epsilon}'_1)/c$ und

$$\vec{E}'(\underline{x}') = \vec{E}'_0 \cos(\underline{k}' \cdot \underline{x}'), \quad \vec{B}'(\underline{x}') = \vec{B}'_0 \cos(\underline{k}' \cdot \underline{x}'). \quad (4.8.63)$$

Da weiter $\underline{k}' \cdot \underline{x}' = \underline{k} \cdot \underline{x}$ ist, gilt gemäß (4.8.50)

$$\vec{E}(\underline{x}) = \vec{E}_0 \cos(\underline{k} \cdot \underline{x}), \quad \vec{B}(\underline{x}) = \vec{B}_0 \cos(\underline{k} \cdot \underline{x}) \quad (4.8.64)$$

mit

$$\vec{E}_0 = \begin{pmatrix} -E'_{01} \sin \alpha' \\ \gamma E'_{01} (\beta + \cos \alpha') \\ E'_{02} \gamma (1 + \beta \cos \alpha') \end{pmatrix}, \quad \vec{B}_0 = \frac{1}{c} \begin{pmatrix} E'_{02} \sin \alpha' \\ -E'_{02} \gamma (\beta + \cos \alpha') \\ E'_{01} \gamma (1 + \beta \cos \alpha') \end{pmatrix}. \quad (4.8.65)$$

Nun ist

$$\underline{k} = \hat{\Lambda}^{-1} \underline{k}' = \hat{\Lambda}^{-1} \begin{pmatrix} k' \\ k' \cos \alpha' \\ k' \sin \alpha' \\ 0 \end{pmatrix} = \begin{pmatrix} \gamma k' (1 + \beta \cos \alpha') \\ \gamma k' (\beta + \cos \alpha') \\ k' \sin \alpha' \\ 0 \end{pmatrix} = \begin{pmatrix} k \\ k \vec{n} \end{pmatrix}. \quad (4.8.66)$$

Es ist also

$$k = \gamma k' (1 + \beta \cos \alpha'), \quad \vec{n} = \begin{pmatrix} (\beta + \cos \alpha') / (1 + \beta \cos \alpha') \\ \sqrt{1 - \beta^2} \sin \alpha' / (1 + \beta \cos \alpha') \\ 0 \end{pmatrix} = \begin{pmatrix} \cos \alpha \\ \sin \alpha \\ 0 \end{pmatrix}. \quad (4.8.67)$$

Die Felder (4.8.65) besitzen dann exakt die gleiche Form wie im System Σ' , wobei

$$\vec{\epsilon}_1 = \begin{pmatrix} -\sin \alpha \\ \cos \alpha \\ 0 \end{pmatrix}, \quad \vec{\epsilon}_2 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}, \quad \vec{E} = (E_{01} \vec{\epsilon}_1 + E_{02} \vec{\epsilon}_2) \cos(\underline{k} \cdot \underline{x}), \quad \vec{B} = \frac{1}{c} \vec{n} \times \vec{E}. \quad (4.8.68)$$

Es sind also wieder ebene Wellen wie im Ruhssystem der Quelle, und ausgedrückt durch die entsprechenden Parameter E_{01} , E_{02} und α , die sich auf Σ beziehen, gleich wie im System Σ' mit den dortigen Parametern E'_{01} , E'_{02} und α' . Dies ist zu erwarten, weil die Forminvarianz der Maxwell-Gleichungen unter Lorentz-Transformationen bedeutet, dass die Lorentz-transformierten Felder wieder Lösungen der Maxwell-Gleichungen sind, und die obigen Parameter bestimmen die Form der ebenen Welle eindeutig. Freilich ändern sich diese Parameter im System Σ , wo sich die Lichtquelle bewegt, gegenüber denen im Ruhssystem der Lichtquelle Σ' .

So ändert sich die Frequenz gemäß (4.8.66) zu

$$\omega = ck = \gamma \omega' (1 + \beta \cos \alpha'). \quad (4.8.69)$$

Dies ist der **Doppler-Effekt** für Licht, denn die Frequenz ändert sich in Abhängigkeit von der Beobachtungsrichtung relativ zur Richtung der Geschwindigkeit der Lichtquelle (parametrisiert durch die Winkel α bzw. α'). Der Effekt wird am größten, wenn sich die Lichtquelle direkt auf den Beobachter in Σ zu- oder weg bewegt, entsprechend $\alpha = 0$ bzw. $\alpha = \pi$. Dann erhält man aus (4.8.67) auch $\alpha' = 0$ bzw. $\alpha' = \pi$ (warum?)

$$\omega_{\pm} = \gamma \omega' (1 \pm \beta) = \omega' \sqrt{\frac{1 \pm \beta}{1 \mp \beta}}. \quad (4.8.70)$$

4. Spezielle Relativitätstheorie

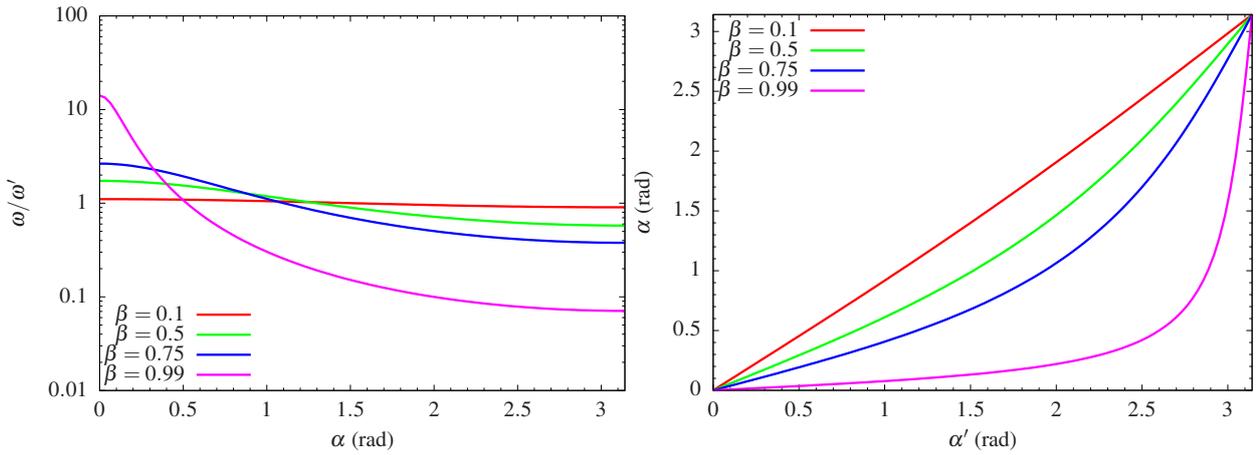


Abbildung 4.5: Links: Dopplereffekt beim Licht gemäß (4.8.74); rechts: Aberration des Lichts gemäß (4.8.72).

Kommt also die Lichtquelle auf den Beobachter zu ($\alpha = 0$, oberes Vorzeichen) ist $\omega = \omega_+ > \omega'$, bewegt sie sich vom Beobachter weg ($\alpha = \pi$, unteres Vorzeichen), ist $\omega = \omega_- < \omega'$. Da uns im sichtbaren Spektrum Licht mit kleinen Frequenzen rot und mit großen Frequenzen blau erscheint, spricht man entsprechend von Blau- bzw. Rotverschiebung.

Interessant ist noch der Spezialfall, dass sich die Lichtquelle senkrecht zur Ausbreitungsrichtung des Lichtes bewegt. Das entspricht $\cos \alpha' = -\beta$ und $\sin \alpha' = \sqrt{1 - \beta^2}$. Dann wird gemäß (4.8.69)

$$\omega = \gamma \omega' (1 - \beta^2) = \omega' \sqrt{1 - \beta^2}, \quad (4.8.71)$$

d.h. im Gegensatz zum analogen Sachverhalt bei Schallwellen, gibt es also auch einen **transversalen Dopplereffekt**. Da hier $\omega < \omega'$ ist, handelt es sich dabei um eine Rotverschiebung.

Weiter zeigt sich, dass sich der Winkel α im Bezugssystem, in dem sich die Lichtquelle bewegt, gegenüber dem Winkel α' im Ruhssystem der Quelle ändert. Dies nennt man die **Aberrations des Lichts**. Definiert man $\alpha, \alpha' \in]-\pi, \pi]$, errechnet sich gemäß (4.8.67) der Winkel α zu

$$\alpha = \text{sign}(\sin \alpha') \arccos \left(\frac{\beta + \cos \alpha'}{1 + \beta \cos \alpha'} \right). \quad (4.8.72)$$

Es zeigt sich an Abb. 4.5, dass i.a. $\alpha < \alpha'$ (außer für die oben diskutierten Fälle $\alpha = \alpha' = 0$ und $\alpha = \alpha' = \pi$). Entsprechend ändert sich auch die Richtung der Polarisationsvektoren gemäß (4.8.68).

Natürlich können wir auch umgekehrt den Dopplereffekt direkt durch den Winkel α bzgl. Σ ausdrücken, denn offenbar gilt

$$\underline{k}' = \hat{\Lambda} \underline{k} = \hat{\Lambda} \begin{pmatrix} k \\ k \cos \alpha \\ k \sin \alpha \\ 0 \end{pmatrix} = \begin{pmatrix} \gamma k (1 - \beta \cos \alpha) \\ \gamma k (-\beta + \cos \alpha) \\ k \sin \alpha \\ 0 \end{pmatrix}. \quad (4.8.73)$$

Wie zu erwarten, unterscheidet sich dies von (4.8.66) nur durch das Ersetzen der Größen $(k, \alpha) \leftrightarrow (k', \alpha')$ und $\beta \rightarrow -\beta$. Aus der Zeitkomponente der Gl. (4.8.73) liest man ab (vgl. Abb. 4.5, rechts)

$$\omega' = \omega \frac{1 - \beta \cos \alpha}{\sqrt{1 - \beta^2}} \Rightarrow \omega = \omega' \frac{\sqrt{1 - \beta^2}}{1 - \beta \cos \alpha}. \quad (4.8.74)$$

Schließlich berechnen wir noch die Intensität (also die Energiedichte) der elektromagnetischen Welle in Alices Bezugssystem und vergleichen sie mit der Intensität im Ruhssystem der Quelle. Es ergibt sich nach einfacher

4.8. Kovariante Formulierung der Elektrodynamik

Rechnung

$$\varepsilon = \frac{\epsilon_0}{2} \vec{E}_0^2 = \frac{\epsilon_0}{2} \vec{E}'^2 \gamma^2 (1 + \beta \cos \alpha')^2 = \varepsilon' \left(\frac{\omega}{\omega'} \right)^2. \quad (4.8.75)$$

Die Intensität ändert sich also im Vergleich zum Ruhssystem der Quelle mit dem Quadrat des „Doppler-Faktors“ ω/ω' .

4. Spezielle Relativitätstheorie

Kapitel 5

Elementarteilchen

5. Elementarteilchen

Anhang A

Spezielle Funktionen

In diesem Anhang stellen wir einige spezielle Funktionen zusammen, die uns im Haupttext aus Lösungen von Eigenwertgleichungen begegnen.

A.1 Hermite-Polynome

A.2 Kugelflächenfunktionen

A.3 Laguerre- und assoziierte Laguerre-Polynome

Wir haben die **assozierten Laguerre-Polynome** bei der Lösung des Energieeigenwertproblems des Wasserstoffatoms in (3.13.44) eingeführt. Mit $n, k \in \mathbb{N}_0$ lautet diese Definition (wir setzen in (3.13.44) $n_r = n$ und $2l + 1 = k$)

$$L_n^k(x) = \sum_{j=0}^n \frac{(-1)^j}{j!} \binom{n+k}{n-j} x^j. \quad (\text{A.3.1})$$

Es ist wichtig zu bemerken, dass in der Literatur unterschiedliche Normierungskonventionen für die assoziierten Laguerre-Polynome gebräuchlich sind.

Wir benötigen oft Integrale mit Laguerre-Polynomen, z.B. bei der Berechnung der Normierungskonstanten für die Energieeigenfunktionen des Wasserstoffatoms. Hierzu ist es nützlich, eine sog. **Rodrigues-Formel** für die L_n^k herzuleiten. Dazu benötigen wir zuerst die **Leibnizsche Produktformel**:

$$\frac{d^a}{dx^a} [f(x)g(x)] = \sum_{j=0}^a \binom{a}{j} f^{(j)}(x)g^{(a-j)}(x) \quad (\text{A.3.2})$$

mit den Binomialkoeffizienten

$$\binom{a}{j} = \frac{a!}{j!(a-j)!}. \quad (\text{A.3.3})$$

Außerdem bedeutet $f^{(j)}$ die j -te Ableitung der Funktion f . Für $a = 1$ erhält man die übliche Produktregel

$$\frac{d}{dx} [f(x)g(x)] = f'(x)g(x) + f(x)g'(x). \quad (\text{A.3.4})$$

Für beliebige a folgt dann der Beweis von (A.3.2) durch vollständige Induktion. Sei also die Formel für a

gültig. Dann ist

$$\begin{aligned} \frac{d^{a+1}}{dx^{a+1}}[f(x)g(x)] &= \frac{d}{dx} \sum_{j=0}^a \binom{a}{j} f^{(j)}(x) g^{(a-j)}(x) \\ &= \sum_{j=0}^a \binom{a}{j} [f^{(j+1)}(x) g^{(a-j)}(x) + f^{(j)}(x) g^{(a-j+1)}(x)]. \end{aligned} \quad (\text{A.3.5})$$

Dabei haben wir im letzten Schritt die übliche Produktformel für Ableitungen verwendet. Wir nummerieren nun in der ersten Summe den Summationsindex zu $j' = j + 1$ um. Dann erhält man

$$\sum_{j=0}^a \binom{a}{j} f^{(j+1)}(x) g^{(a-j)}(x) = \sum_{j'=1}^{a+1} \binom{a}{j'-1} f^{(j')}(x) g^{(a+1-j')}(x). \quad (\text{A.3.6})$$

Benennen wir dann j' wieder in j um und setzen das Resultat in (A.3.5) ein, erhalten wir

$$\frac{d^{a+1}}{dx^{a+1}}[f(x)g(x)] = f(x)g^{(a+1)}(x) + \sum_{j=1}^a \left[\binom{a}{j-1} + \binom{a}{j} \right] f^{(j)}(x) g^{(a+1-j)}(x) + f^{(a+1)}(x)g(x). \quad (\text{A.3.7})$$

Nun ist

$$\binom{a}{j-1} + \binom{a}{j} = \frac{a!}{(j-1)!(a+1-j)!} + \frac{a!}{j!(a-j)!} = a! \left[\frac{j+(a+1-j)}{j!(a+1-j)!} \right] = \frac{(a+1)!}{j!(a+1-j)!} = \binom{a+1}{j}. \quad (\text{A.3.8})$$

Da weiter $\binom{a+1}{0} = \binom{a+1}{a+1} = 1$ ist also schließlich

$$\frac{d^{a+1}}{dx^{a+1}}[f(x)g(x)] = \sum_{j=0}^{a+1} \binom{a+1}{j} f^{(j)}(x) g^{(a+1-j)}(x), \quad (\text{A.3.9})$$

und das ist die Produktformel für die $(a+1)$ -te Ableitung.

Jetzt wenden wir die Produktformel auf $f(x) = \exp(-x)$ und $g(x) = x^b$ an. Dann ist offenbar (*Nachrechnen!*)

$$f^{(j)}(x) = (-1)^j \exp(-x), \quad g^{(a-j)}(x) = \frac{b!}{(b-a+j)!} x^{b-a+j} \quad (\text{A.3.10})$$

und damit vermöge (A.3.2)

$$\frac{d^a}{dx^a} [x^b \exp(-x)] = \exp(-x) \sum_{j=0}^a \binom{a}{j} \frac{b!}{(b-a+j)!} (-1)^j x^{b-a+j}. \quad (\text{A.3.11})$$

Setzt man hierin $a = n$ und $b = n + k$, erhalten wir

$$\begin{aligned} \frac{d^n}{dx^n} [x^{n+k} \exp(-x)] &= \exp(-x) \sum_{j=0}^n (-1)^j \binom{n}{j} \frac{(n+k)!}{(k+j)!} x^{k+j} \\ &= x^k \exp(-x) \sum_{j=0}^n (-1)^j \frac{n!(n+k)!}{j!(n-j)!(k+j)!}. \end{aligned} \quad (\text{A.3.12})$$

Nun ist der in (A.3.1) vorkommende Faktor

$$\frac{1}{j!} \binom{n+k}{n-j} = \frac{(n+k)!}{j!(n-j)!(k+j)!} \quad (\text{A.3.13})$$

Vergleicht man demnach (A.3.1) mit (A.3.12), ergibt sich die **Rodrigues-Formel** für die assoziierten Laguerre-Polynome

$$L_n^k(x) = \frac{1}{n!} x^{-k} \exp x \frac{d^n}{dx^n} \left[x^{n+k} \exp(-x) \right]. \quad (\text{A.3.14})$$

Definieren wir nun ein Skalarprodukt auf dem Hilbertraum von Funktionen durch

$$\langle \psi | \phi \rangle_{L,k} = \int_0^\infty dx x^k \exp(-x) \psi^*(x) \phi(x), \quad (\text{A.3.15})$$

So bilden die assoziierten Laguerre-Polynome L_n^k zu festem k mit $n \in \mathbb{N}_0$ ein orthogonales Funktionensystem. Man kann auch zeigen, dass es sogar vollständig ist. Um das zu sehen, berechnen wir die entsprechenden Skalarprodukte. Sei zunächst $0 \leq n < n'$. Dann gilt wegen (A.3.14)

$$\langle L_n^k | L_{n'}^k \rangle_{L,k} = \frac{1}{n!} \int_0^\infty dx L_n^k(x) \frac{d^{n'}}{dx^{n'}} \left[x^{n'+k} \exp(-x) \right]. \quad (\text{A.3.16})$$

Nun können wir n' -mal partiell integrieren, wobei alle „integralfreien“ Terme verschwinden, weil an der unteren Grenze stets mindestens 1 Faktor x übrig bleibt. Es entsteht schließlich

$$\langle L_n^k | L_{n'}^k \rangle_{L,k} = \frac{1}{n!} (-1)^{n'} \int_0^\infty dx \frac{d^{n'}}{dx^{n'}} \left[L_n^k(x) \right] x^{n'+k} \exp(-x) = 0. \quad (\text{A.3.17})$$

Letzteres folgt daraus, dass L_n^k ein Polynom n -ten Grades ist. Da $n' > n$ verschwindet folglich dessen n' -te Ableitung. Dieselbe Rechnung ist aber auch für $n' = n$ gültig. Bei der Ableitung müssen wir nur den Term in $L_n^k(x)$ mit der höchsten Potenz x^n berücksichtigen. Aus (A.3.1) folgt, dass $L_n^k(x) = (-1)^n x^n / n! + \dots$ und folglich $d_x^n L_n^k(x) = (-1)^n$. Damit erhalten wir

$$\langle L_n^k | L_n^k \rangle_{L,k} = \frac{1}{n!} (-1)^n \int_0^\infty \frac{d^n}{dx^n} \left[L_n^k(x) \right] x^{n+k} \exp(-x) = \frac{1}{n!} \int_0^\infty dx x^{n+k} \exp(-x) = \frac{(n+k)!}{n!}. \quad (\text{A.3.18})$$

Dabei ergibt sich das Integral im Letzten Schritt mit Hilfe einer erzeugenden Funktion. Definieren wir

$$F(y) = \int_0^\infty dx \exp(-yx) = -\frac{1}{y} \exp(-yx) \Big|_{x=0}^\infty = \frac{1}{y}, \quad (\text{A.3.19})$$

folgt

$$\int_0^\infty dx x^k \exp(-x) = (-1)^k F^{(k)}(1). \quad (\text{A.3.20})$$

Nun ist

$$F'(y) = -\frac{1}{y^2}, \quad F''(y) = +\frac{2}{y^3}, \dots, F^{(k)}(y) = (-1)^k \frac{k!}{y^{k+1}} \Rightarrow F^{(k)}(1) = (-1)^k k!. \quad (\text{A.3.21})$$

Damit wird aber

$$\int_0^\infty dx x^k \exp(-x) = k!. \quad (\text{A.3.22})$$

A. Spezielle Funktionen

Literaturverzeichnis

- [AHK⁺18] T. Arens, et al., *Mathematik*, 4. Aufl., Springer-Verlag, Berlin, Heidelberg (2018).
URL <https://doi.org/10.1007/978-3-662-56741-8>
- [Bal70] L. E. Ballentine, The Statistical Interpretation of Quantum Mechanics, *Rev. Mod. Phys.* **42**, 358 (1970).
URL <https://doi.org/10.1103/RevModPhys.42.358>
- [Bal98] L. E. Ballentine, *Quantum Mechanics*, World Scientific, Singapore, New Jersey, London, Hong Kong (1998).
- [BFK⁺18a] M. Bartelmann, et al., *Theoretische Physik 1 – Mechanik*, Springer-Verlag, Berlin, Heidelberg (2018).
URL <https://doi.org/10.1007/978-3-662-56115-7>
- [BFK⁺18b] M. Bartelmann, et al., *Theoretische Physik 2 – Elektrodynamik*, Springer-Verlag, Berlin, Heidelberg (2018).
URL <https://doi.org/10.1007/978-3-662-56117-1>
- [BFK⁺18c] M. Bartelmann, et al., *Theoretische Physik 3 – Quantenmechanik*, Springer-Verlag, Berlin, Heidelberg (2018).
URL <https://doi.org/10.1007/978-3-662-56072-3>
- [BFK⁺18d] M. Bartelmann, et al., *Theoretische Physik 4 – Thermodynamik und Statistische Physik*, Springer-Verlag, Berlin, Heidelberg (2018).
URL <https://doi.org/10.1007/978-3-662-56113-3>
- [CH10] W. Cassing, H. v. Hees, *Mathematische Methoden für Physiker*, Universität Gießen (2010).
URL <https://itp.uni-frankfurt.de/~hees/publ/maphy.pdf>
- [DL18] D. Dürr, D. Lazarovici, *Verständliche Quantenmechanik*, Springer Spektrum, Berlin (2018).
URL <https://doi.org/10.1007/978-3-662-55888-1>
- [Ein05] A. Einstein, Zur Elektrodynamik bewegter Körper, *Ann. d. Phys.* **322**, 891 (1905).
URL <https://doi.org/10.1002/andp.19053221004>
- [Ein07] A. Einstein, Über das Relativitätsprinzip und die aus demselben gezogenen Folgerungen, *Jahrbuch der Radioaktivität und Elektronik* **4**, 411 (1907).
- [Ein09a] A. Einstein, *Grundzüge der Relativitätstheorie*, 7. Aufl., Springer, Berlin, Heidelberg (2009).
URL <https://doi.org/10.1007/978-3-540-87847-6>
- [Ein09b] A. Einstein, *Über die spezielle und die allgemeine Relativitätstheorie*, 24. Aufl., Springer, Berlin, Heidelberg (2009).
URL <https://doi.org/10.1007/978-3-540-87777-6>

- [FH03] B. Friedrich, D. Herschbach, Stern and Gerlach: How a Bad Cigar Helped to Reorient Atomic Physics, *Physics Today* **12/2003**, 53 (2003).
URL https://scitation.aip.org/journals/doc/PHTOAD-ft/vol_56/iss_12/53_1.shtml
- [Fic79] E. Fick, *Einführung in die Grundlagen der Quantentheorie*, 4. Aufl., Aula-Verlag, Wiesbaden (1979).
- [FKL⁺18] C. Friebe, et al., *Philosophie der Quantenphysik: Zentrale Begriffe, Probleme, Positionen*, 2. Aufl., Springer Spektrum, Berlin (2018).
URL <https://doi.org/10.1007/978-3-662-54276-7>
- [Fli18] T. Fließbach, *Die relativistische Masse*, Springer Spektrum, Berlin (2018).
URL <https://doi.org/10.1007/978-3-662-58084-4>
- [FLS07] R. P. Feynman, R. B. Leighton, M. Sands, *Feynman-Vorlesungen über Physik*, Bd. 3, 5. Aufl., Oldenbourg Verlag, München, Wien (2007).
- [Hee08] H. van Hees, *Klassische Mechanik* (2008).
URL <https://itp.uni-frankfurt.de/~hees/faq-pdf/mech.pdf>
- [Hee15] H. v. Hees, Grundlagen der Quantentheorie, I. Teil: Nichtrelativistische Quantentheorie, FAQ der Newsgroup de.sci.physik (2015).
URL <https://itp.uni-frankfurt.de/~hees/faq-pdf/quant.pdf>
- [Hee18a] H. van Hees, *Theoretische Physik 1 für das Lehramt L3*, Goethe-Universität Frankfurt (2018).
URL <https://itp.uni-frankfurt.de/~hees/publ/theo1-13.pdf>
- [Hee18b] H. van Hees, *Theoretische Physik 2 für das Lehramt L3*, Goethe-Universität Frankfurt (2018).
URL <https://itp.uni-frankfurt.de/~hees/publ/theo2-13.pdf>
- [Hen17] K. Hentschel, *Lichtquanten*, Springer Spektrum, Berlin, Heidelberg (2017).
URL <https://doi.org/10.1007/978-3-662-55273-5>
- [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).
- [Sch94] R. A. Schumacher, Special relativity and the Michelson–Morley interferometer, *Am. Jour. Phys* **62**, 609 (1994).
URL <https://doi.org/10.1119/1.17535>
- [ST93] J. J. Sakurai, S. Tuan, *Modern Quantum Mechanics*, Addison Wesley (1993).
- [SW10] B. Schumacher, M. Westmoreland, *Quantum Processes, Systems, and Information*, Cambridge University Press, Cambridge, New York, Melbourne, Madrid, Cape Town, Singapore, Sao Paulo, Delhi, Dubai, Tokyo (2010).
URL <https://doi.org/10.1017/CB09780511814006>
- [T⁺18] M. Tanabashi, et al., The Review of Particle Physics (2018), *Phys. Rev. D* **98**, 030001 (2018).
URL <https://pdg.lbl.gov/>
- [Zei10] A. Zeilinger, *Dance of the photons: from Einstein to quantum teleportation*, Farrar, Straus and Giroux, New York (2010).