Theoretische Physik III
für das Lehramt L3

H. van Hees

4. Juli 2025

Inhaltsverzeichnis

Inhaltsverzeichnis **3**

1 Einleitung **7**
1.1 Bemerkung zur Interpretation der QT 8
1.2 Lehrbücher 8

2 Das Versagen der klassischen Physik **9**
2.1 Vorstellungen zum Aufbau der Materie um 1900 9
2.2 1900: Planck und die Hohlraumstrahlung 13
2.2.1 Hohlraumstrahlung im Rahmen der klassischen Physik 14
2.2.2 Folgerungen aus der phänomenologischen Thermodynamik 16
2.2.3 Plancks „Verzweiflungstat“ 18
2.3 1905: Einstein und der Photoeffekt 20
2.4 1923: Compton und die Streuung von Röntgenstrahlung 23
2.5 1909-1913: Rutherford-Streuung und das Bohrsche Atommodell 25
2.6 1923-1926: de Broglie und der Welle-Teilchen-Dualismus 28
2.7 1926: Schrödinger und die Wellenmechanik 28
2.7.1 Die Schrödinger-Gleichung für ein freies Teilchen 29
2.7.2 Die Schrödinger-Gleichung für ein Teilchen in einem Potential 36
2.7.3 Messwerte von Observablen 36
2.7.4 Die Hilbert-Raum-Struktur der Wellenfunktionen 39
2.7.5 Selbstadjungierte Operatoren 40
2.7.6 Energieeigenfunktionen 44
2.7.7 Kompatible und inkompatible Observable 45

3 Quantenmechanik **49**
3.1 Die Grundpostulate der Quantentheorie 49
3.1.1 Anmerkungen zum Begriff des quantenmechanischen Zustandes 50
3.2 Der Hilbert-Raum 52
3.3 Lineare Operatoren im Hilbertraum und Observablen 57
3.4 Verträglichkeit von Observablen 59
3.5 Die Heisenbergsche Unschärferelation 61
3.6 Zeitentwicklung des Zustands 63
3.7 Diracs Zugang zur Quantenmechanik 65

3.7.1 Erinnerung an die Hamiltonsche Mechanik . . . 65
3.7.2 Quantentheorie für ein Teilchen mit Spin 0 . . . 66
3.8 Das freie Teilchen . . . 72
3.9 Der eindimensionale harmonische Oszillator . . . 74
3.10 Eigenwertproblem für beliebige Drehimpulsoperatoren . . . 77
3.11 Der Bahndrehimpuls . . . 80
3.12 Freies Teilchen (Energie-Drehimpuls-Eigenzustände) . . . 89
3.13 Das Wasserstoffatom . . . 94
3.14 Der Spin . . . 101
3.15 Magnetisches Moment einer klassischen Stromverteilung . . . 105
3.16 Teilchen mit $s = 0$ im Magnetfeld . . . 108
3.17 Teilchen mit Spin im Magnetfeld . . . 108
3.18 Teilchen im homogenen Magnetfeld . . . 109
3.19 Der Stern-Gerlach-Versuch . . . 114
3.20 EPR-Paradoxon und Bell-Ungleichung . . . 117

4 Spezielle Relativitätstheorie 125
4.1 Das speziell-relativistische Raum-Zeit-Modell . . . 125
4.1.1 Herleitung der Lorentz-Transformation (physikalische Formulierung) . . . 126
4.1.2 Herleitung der Lorentz-Transformationen (geometrische Formlierung) . . . 130
4.2 Kinematische Effekte . . . 135
4.2.1 Relativität der Gleichzeitigkeit . . . 135
4.2.2 Zeitdilatation . . . 135
4.2.3 Längenkontraktion . . . 135
4.2.4 Einsteins Zug-Gedankenexperiment . . . 137
4.2.5 Das Michelson-Morley-Experiment . . . 138
4.2.6 Bemerkungen zur Historie der Lorentz-Transformation . . . 142
4.3 Das Zwillingsparadoxon . . . 143
4.4 Speziell relativistische Mechanik . . . 143
4.5 Zerfall und Stöße von Teilchen . . . 151
4.5.1 Zweiteilchenzerfall . . . 151
4.5.2 Zweiteilchenstöße . . . 152
4.5.3 Kinematik im Schwerpunktssystem . . . 153
4.5.4 Kinematik im Laborsystem . . . 154
4.6 Äquivalenz zwischen Energie und Masse . . . 154
4.6.1 Voll inelastischer Stoß . . . 155
4.6.2 Zerfall eines Pions in zwei Photonen . . . 155
4.7 Geladene Teilchen im elektromagnetischen Feld . . . 156
4.7.1 Teilchen im homogenen elektrischen Feld . . . 159
4.8 Kovariante Formulierung der Elektrodynamik . . . 161
4.8.1 Tensoranalysis und Maxwell-Gleichungen im Minkowski-Raum . . . 161
4.8.2 Die elektromagnetischen Potentiale . . . 165
4.8.3 Transformationsverhalten der Felder . . . 166

4.8.4 Feld einer gleichförmig bewegten Punktladung 167
4.8.5 Dopplereffekt und Aberration bei Licht 168
4.8.6 Der Energie-Impuls-Tensor 171

5 Elementarteilchen **175**

A Die Baker-Campbell-Hausdorff-Formel **177**
A.1 Herleitung der BCH-Formel 177
A.2 Anwendung auf Exponentialfunktion einer Summe 177

B Spezielle Funktionen **179**
B.1 Hermite-Poloynome 179
B.1.1 Erzeugende Funktion 179
B.1.2 Die Green-Funktion des harmonischen Oszillators 181
B.2 Sphärische Bessel-, Neumann- und Hankel-Funktionen 183
B.3 Kugelflächenfunktionen 186
B.3.1 Green-Funktion des Laplace-Operators 187
B.3.2 Green-Funktion des Helmholtz-Operators 191
B.4 Laguerre- und assoziierte Laguerre-Polynome 194

C Gauß-Integrale **197**
C.1 Grundlegende Formeln 197
C.2 Fourier-Transformation der Gauß-Verteilung 199

D Zur Ganzzahligkeit der Bahndrehimpulseigenwerte **201**

E Teilchen im elektromagnetischen Feld und Eichinvarianz **203**
E.1 Klassische Theorie 203
E.2 Quantenmechanik und Eichtransformationen 206
E.3 Eichtheoretischer Zugang zur Quantentheorie geladener Teilchen 208

Literaturverzeichnis **211**

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 Erfahrungstatsachen 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 [Ein05b] 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 Kaptitel 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 „metaphysikalischen" Zusatzannahmen nur das unbedingt für die Anwendung des Formalismusses 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 werden. 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 verstänliche Weise Einblick in mögliche alternative Interpretationen bzw. Erweiterungen der QT bietet, ist [DL18]. Beide Bücher sind ebenfalls zum kostenlosen Download im Netzwerkt der Goethe-Universität verfügbar.

1.2 Lehrbücher

Mit dem Wintersemester 2020/2021 habe ich mich entschlossen, mich zur Einführung der Quantentheorie des traditionellen Zugangs über die **Wellenmechanik** zu bedienen, da nach zwei Semestern Lehrerfahrung in der Kürze der Zeit eine Vermittlung des modernen „darstellungsfreien" Dirac-Formalismusses eher zur Verwirrung als zum Verständnis der Theorie führt. Dieser traditionelle Zugang über die Wellenmechanik, die wir in Kapitel 4 besprechen werden, bietet Bd. 3 der Feynman Lectures [FLS07]. Hervorragend ist auch [Fic79].

Ein sehr schönes neues Lehrbuch der gesamten theoretischen Physik ist das sowohl in 1 Band (1. Auflage) als auch in 4 Teilbänden (2. Auflage) 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 selbst [Ein09a, Ein09b], die ebenfalls im Netzwerk der Goethe-Universität kostenfrei heruntergeladen werden können.

Frankfurt, im Sommer 2020.

Kapitel 2

Das Versagen der klassischen Physik

In diesem Kapitel besprechen wir kurz, wie es historisch zur Entwicklung der Quantentheorie gekommen ist. Sie ist nämlich entstanden, weil einige Phänomene, die aufgrund der fortschreitenden Technik in der zweiten Hälfte des 19. und zu Beginn des 20. Jahrhunderts der Beobachtung zugänglich geworden sind, nicht mit der bekannten „klassischen Physik“ erklärbar sind.

Wir können dabei auf einige wichtige Details nicht allzu genau eingehen, weil dazu eine detaillierte Behandlung der Thermodynamik und klassischen statistischen Physik notwendig wäre.

Aus didaktischer Sicht kann man auch gegen diesen „historischen Zugang“ einwenden, dass gerade im Hinblick auf die Quantentheorie die Verfolgung der historischen Entwicklung der Theoriebildung oft zunächst falsche Vorstellungen beim Lernenden hervorruft. Dies gilt insbesondere für die sich zwischen 1900-1920 entwickelte sog. „**alte Quantentheorie**“, die im Wesentlichen die Idee des **Welle-Teilchen-Dualismus'** und das **Bohr-Sommerfeldsche Atommodell** mit seinen auf Bahnen um den Atomkern kreisenden Elektronen zur theoretischen Erklärung der **Spektrallinien** der von Atomen bzw. Molekülen abgestrahlten elektromagnetischen Wellen. Es muss auch vor dem **naiven Photonenbild** gewarnt werden, das auf Einsteins berühmtes Paper [Ein05a] zurückgeht: Hier wird die (falsche) Vorstellung impliziert, es handele sich bei Photonen um lokalisierte teilchenartige Objekte, was der modernen Auffassung dieses Begriffs im Rahmen der **relativistischen Quantenfeldtheorie** widerspricht. Wir können leider darauf in dieser Vorlesung nicht genauer eingehen. Eine gut lesbare auf die Historie konzentrierte Darstellung des Photonenbegriffs findet sich in [Hen17].

Wir gehen daher in jedem der folgenden Teilabschnitte am Ende auch auf die moderne Sichtweise genauer ein, um solche Fehler in der Begriffsbildung im Zusammenhang mit der QT vorzubeugen. Die eigentliche Schwierigkeit der QT besteht nämlich eher in einem Verständnis ihrer grundlegenden physikalischen Bedeutung als in der Mathematik, die im Vergleich zur Elektrodynamik sogar etwas einfacher ist, kommen wir doch in den weitesten Teilen dieses Manuskripts mit einer einzelnen komplexwertigen **skalaren Wellenfunktion**, wie sie von **Schrödinger** in 1926, basierend auf Vorarbeiten von **de Broglie** in seiner Dissertation von 1923, als eine der möglichen Grundlegungen der Theorie in die theoretische Physik eingeführt und von **Born** im modernen Sinne als **Wahrscheinlichkeitsamplitude** interpretiert wurde, aus.

2.1 Vorstellungen zum Aufbau der Materie um 1900

In der zweiten Hälfte des 19. Jh. erschien den Physikern die Natur als weitgehend vollständig verstanden und die Physik in diesem Sinne abgeschlossen. Insgesamt betrachtet wurde demnach die Welt durch ein Theoriegebäude beschrieben, das wir heute als **klassische Physik** beschreiben. Sie bestand in groben Zügen systematisiert aus den Teilgebieten der **Mechanik, Elektrodynamik und Thermodynamik**, von denen wir in diesem Vorlesungszyklus die ersten beiden ausführlich besprochen haben.

Die **Mechanik** beschreibt zum einen, wie in der Vorlesung Theorie 1 für Lehramt L3 besprochen, die Bewe-

gung von „Punktteilchen“ aufgrund der Wirkung von Wechselwirkungskräften unter Verwendung der **drei Newtonschen Postulate**, die zuerst das **Raumzeitmodell** mit der Auszeichnung der **Inertialsysteme** bereitstellen, innerhalb dessen die Beobachtungen über die Welt erst formulierbar und quantifizierbar werden, kann man doch Längen und Zeiten erst dann sinnvoll definieren, wenn man ein solches Raumzeitmodell zur Verfügung hat. Als dritte für die Mechanik wesentliche fundamentale Größe kommt dann noch die Masse der Materie hinzu, die zum einen die **Trägheit** der Materie quantifiziert, wonach $m\vec{a} = \vec{F}$ gilt (2. Newtonsches Postulat), d.h. die Beschleunigung eines Massenpunktes ist proportional zur totalen auf diesen Körper einwirkenden Kraft und umgekehrt proportional zur Masse. Dabei charakterisiert die Masse auch die Menge einer jeden spezifischen Art von Materie, und das 2. Newtonsche Postulat gibt die qualitative Erfahrung wieder, dass zur Erzielung einer bestimmten Beschleunigung für eine größere Menge einer Substanz eine größere Kraft benötigt wird, und Newton quantifiziert dies, indem er die **träge Masse** als Proportionalitätskonstante zwischen Kraft und Beschleunigung einführt.

Die Betrachtungen zur Bewegung der Himmelskörper umeinander, insbesondere der Planeten um die Sonne unter der Kenntnis der empirisch gefundenen **Keplerschen Gesetzen** führte Newton dann auch zu seinem **universellen Gravitationsgesetz**, wonach zwischen zwei Punktmassen stets eine entlang der Verbindungslinie der beiden Massen anziehend wirkende Wechselwirkungskraft vom Betrag $F_{\text{grav}} = Gm_1m_2/r^2$ wirkt, wobei $G = 6{,}67430(15) \cdot 10^{-11}\text{m}^3/(\text{kgs}^2)$ die **Newtonsche Gravitationskonstante** ist und m_1 und m_2 die Massen der beiden Körper. Bemerkenswert dabei ist, dass diese **schweren Massen** bis auf eine im Prinzip willkürlich zu wählende Proportionalitätskonstante genau den trägen Massen aus dem 2. Newtonschen Postulat entspricht. Der Einfachheit halber misst man die träge und schwere Masse in der gleichen Einheit kg und passt die Gravitationskonstante entsprechend an die experimentellen Befunde (Coulombscher Drehwaagenversuch) an.

Basierend auf diesen Grundlagen entwickelte sich auch die **Kontinuumsmechanik**, die die Bewegung von ausgedehnter Materie, also Gasen, Flüssigkeiten und Festkörpern beschreiben. Dabei wurden auch einige empirisch zu bestimmende Kenngrößen für jeden Stoff der Materie eingeführt, z.B. bei Gasen und Flüssigkeiten die **Viskosität**, die die Reibung zwischen Flüssigkeitselementen beschreibt oder der Elastizitätsmodul für elastische Festkörper, aus dem sich z.B. die Federkonstante einer Feder ergibt, wonach die Rückstellkraft proportional zur Auslenkung relativ zur Länge im „entspannten Zustand“ ergibt, wobei dies nur in einem begrenzten Gültigkeitsbereich für kleine Auslenkungen gilt. Außerdem benötigte man auch noch Beziehungen zwischen Größen wie Druck und Dichte von Gasen oder Flüssigkeiten, die Erhaltung der Masse usw. sowie schließlich eine Beschreibung der Bewegung, was dann über Feldgleichungen erfolgt (z.B. die Eulersche Bewegungsgleichung für ein ideales Fluid oder die Navier-Stokes-Gleichung für viskose Flüssigkeiten und Gase), die aber im wesentlichen auf den grundlegenden Bewegungsgesetzen für Punktteilchen sowie der Annahme, dass die Gesamtmasse der Materie bei allen Naturvorgängen erhalten bleibt, beruhen. Letzteres war auch aus der Chemie im Zusammenhang mit Stoffumwandlungen empirisch verifiziert worden (Lavoisier).

Als zunächst von der Mechanik unabhängiges Gebiet entwickelte sich auch die **Thermodynamik**, die Phänomene der Wärme mit einbezog. Dabei spielt naturgemäß der gegenüber der Mechanik neu einzuführende Begriff der Temperatur als einer **Zustandsgröße** eine herausragende Rolle. Dabei ergab sich, dass Wärme eine Energieform ist, die sich von der aus den Mechanik her bekannten Begriffen der kinetischen und Potentiellen Energie unterscheidet. Ein wesentliches Prinzip, das auch unter Einbeziehung der Wärme als neuartiger Energieform weiter gilt, wurde der **Energieerhaltungssatz** erkannt. Dieser wird im Zusammenhang mit der Thermodynamik als **1. Hauptsatz der Thermodynamik** bezeichnet. Die Wärmemengen, die bei Temperaturunterschieden zwischen zwei Körpern ausgetauscht werden, sind dabei keine Zustandsgrößen sondern hängen vom spezifischen Prozess, mit dem dieser Wärmeaustausch einhergeht. Die Annahme, dass die Temperatur eine Zustandsgröße ist, also nicht vom spezifischen Weg abhängt, wie der Körper durch Wärmeaustausch nach einiger Zeit in den Gleichgewichtszustand übergeht, wird dabei auch als **0. Hauptsatz der Thermodynamik** bezeichnet. Der Gleichgewichtszustand hinsichtlich Wärmeaustausch ist demnach dann erreicht, wenn die Temperaturen der beteiligten Körper überall gleich sind. Dabei zeigt sich, dass stets Wärmeenergie vom

Körper der höheren Temperatur zum niedrigeren übergeht, d.h. der wärmere Körper (also der mit der höheren Temperatur) kühlt sich stets ab und der Körper mit der niedrigen Temperatur wird dabei wärmer. Die Annahme, dass ohne Verrichtung äußerer Arbeit nie Wärme von einem kälteren auf einen wärmeren Körper übergehen kann, nennt man den **2. Hauptsatz der Thermodynamik**. Mit diesem hängt auch der wichtige Begriff der **Entropie** zusammen. Dabei wird die Entropie als weitere Erhaltungsgröße eingeführt, wonach die Änderung δQ der Wärmeenergie stets $\delta Q \geq T\mathrm{d}S$ erfüllt, wobei T die Temperatur des betrachteten Körpers und S dessen Entropie bezeichnet. Dabei gilt das Gleichheitszeichen für sog. **reversible Zustandsänderungen**. Ebenso ergibt sich der 2. Hauptsatz auch aus der Formulierung, dass für ein abgeschlossenes System die Entropie nie abnimmt, woraus dann sofort folgt, dass der Gleichgewichtszustand durch den Zustand der maximalen Entropie des Systems unter den durch die Erhaltungssätze (insbesondere den Energieerhaltungssatz) vorgegebenen Einschränkungen charakterisiert ist.

So erfolgreich diese **phänomenologische Thermodynamik** auch war, so waren die Physiker doch bemüht, sie aus den Grundprinzipien der Mechanik herzuleiten, denn die Mechanik galt als die fundamentale Grundlage der gesamten Physik, und man wollte alle Naturvorgänge **mechanistisch** verstehen. Nun war unter den Chemikern der damaligen Zeit das Postulat von der Existenz von **Atomen** weitgehend unbestritten. Dabei galt ein Atom (griech. für „das Unteilbare") als eine kleinstmögliche **diskrete** Menge eines bestimmten elementaren Stoffes. Die Elemente konnten dabei aufgrund ihrer chemischen Eigenschaften ins **Periodensystem der Elemente** eingruppiert werden, und alle übrigen Stoffe wurden als aus Atomen zusammengesetzte Moleküle verstanden, die sich in chemischen Reaktionen untereinander umwandeln. Dabei war die Atomvorstellung vor allem dadurch begründet, dass an chemischen Reaktionen jeweils ganzzahlige Vielfache von elementaren Bestandteilen beteiligt sind. Z.B. gibt es ja die Reaktion $2\mathrm{H}+\mathrm{O} \rightarrow \mathrm{H_2O}$, bzw. $2\mathrm{H_2}+\mathrm{O_2} \rightarrow 2\mathrm{H_2O}$, d.h. Wasser ist ein Molekül, das stets aus zwei Wasserstoffatomen und einem Sauerstoffatom besteht, d.h. es gibt ein festes ganzzahliges Verhältnis der atomaren Bestandteile eines Moleküls eines Stoffes. Bei den Physikern wurde die Vorstellung diskreter atomarer oder molekularer Einheiten eines Stoffes allerdings bis zum Beginn des 20. Jh. eher abgelehnt, weil es keinen direkten Nachweis einer solchen „Körnung" der Materie gab. Allerdings gab es auch seit der 2. Hälfte des 19. Jh. Physiker (wie insbesondere Maxwell und Boltzmann), die diese Atomvorstellung verwendeten, um die Gesetze der phänomenologischen Thermodynamik aus den Grundlagen der Mechanik herzuleiten, wobei man sich die Materie als System einer sehr großen Zahl von Atomen oder Molekülen vorstellt. Entsprechend der atomistischen Vorstellung der Materie wurde auch das Mol als ein neues Maß für die Menge eines bestimmten Stoffes eingeführt, und zwar als die Menge von fundamentalen Atomen oder Molekülen, die der Anzahl von solchen Atomen und Molekülen, die in einer bestimmten Masse einer Referenzsubstanz enthalten sind. Anfangs wählte man den Wasserstoff, später dann den Kohlenstoff. Demnach ist 1 mol die Menge eines Stoffes, die der Anzahl der Kohlenstoffatome, die in 12 g Kohlenstoff (wobei man genauer sagen muss, dass es sich dabei um Kohlenstoff mit den Atomkernen $^{12}\mathrm{C}$ handeln soll) enthalten sind. Dabei nutzte man aus, dass die Atome vieler Elemente eine Masse besitzen, die einem ganzzahligen Vielfachen der Masse eines Wasserstoffatoms entspricht, was eine atomare Masseneinheit u begründet. Natürlich vorkommender Kohlenstoff besitzt dabei eine Masse von 12,0107u. Heute wissen wir, dass die nicht genaue Ganzzahligkeit darauf zurückzuführen ist, dass Atome aus einem Atomkern mit positiver Ladung, die der Anzahl der in diesem Kern enthaltenen **Protonen** entspricht, also ein ganzzahliges Vielfaches der Elementarladung $e = 1{,}602176634 \cdot 10^{-19}$ C, besteht. Die Atomkerne enthalten aber auch noch eine bestimmte Anzahl **Neutronen**, die ziemlich genau die gleiche Masse wie ein Proton besitzen. Dabei gibt es oft Kerne, die die gleiche Protonenzahl aber eine unterschiedliche Neutronenzahl besitzen. Diese bezeichnet man als **Isotope**, denn die Chemie der Elemente wird im Wesentlichen durch die Anzahl der um den Atomkern angeordneten Elektronen bestimmt, und ein neutrales Atom besitzt genauso viele negativ geladene Elektronen, die die Ladung $q_{\mathrm{e}^-} = -e$ tragen. Isotop steht für griech. „am gleichen Ort". Das bezieht sich darauf, dass die Chemie von Isotopen des gleichen Elements gleich viele Protonen aber eine unterschiedliche Zahl von Neutronen enthalten und sich demzufolge chemisch gleich verhalten und folglich an derselben Stelle im Periodensystem der Elemente einzuordnen sind. Natürlicher Kohlenstoff ist nun ein Gemisch verschiedener Isotope. Ein Kohlenstoffkern besitzt immer 6 Protonen. Das am häufigsten in natürlichem Kohlenstoff enthaltene Isotop

(98,9%) besitzt zusätzlich 6 Neutronen, besitzt also die Massenzahl $A = 12$ (6 Protonen + 6 Neutronen). Fast der gesamte Rest wird durch das Kohlenstoffisotop mit einem Neutron mehr im Kern, also ^{13}C, gebildet.

Einem mol, also 12 g aus isotopenreinem ^{12}C bestehendem Kohlenstoff, entsprechen dabei der unvorstellbar großen Zahl $N_A = 6{,}02214076 \cdot 10^{23}/\text{mol}$ Atomen bzw. Molekülen. Es versteht sich von selbst, dass es unmöglich ist, ein System von $\sim 10^{24}$ Teilchen im Rahmen der Punktmechanik zu beschreiben. Dazu müsste man schließlich die Orte und Impulse all dieser Teilchen am Anfang der Bewegung exakt kennen (Anfangsbedingungen für die Newtonschen Bewegungsgleichungen), die Wechselwirkungskräfte zwischen all diesen Teilchen und schließlich ein gekoppeltes System von $\sim 3 \cdot 10^{24}$ Newtonschen Bewegungsgleichungen lösen. Das ist selbst mit den leistungsfähigsten Großcomputern nicht möglich. Andererseits ist eine solch detaillierte Beschreibung aber auch gar nicht notwendig, denn aus Sicht der Mechanik erhält man ja mit der Punktmechanik (wenn die Ausdehnung der Körper in einem Problem weitgehend vernachlässigbar ist, wie bei der Bewegung von Planeten um die Sonne), der Mechanik starrer Körper (z.B. Kreiseltheorie) oder auch der Kontinuumsmechanik (Gase, Flüssigkeiten und Festkörper) usw. schon eine hinreichend befriedigende Beschreibung des Verhaltens dieser **Vielteilchensysteme**. Dabei betrachtet man eine begrenzte Zahl **kollektiver makroskopischer Freiheitsgrade**, z.B. in der Punktmechanik für jeden Massenpunkt den Schwerpunkt des entsprechenden tatsächlichen ausgedehnten Körpers oder in der Kontinuumsmechanik das Geschwindigkeitsfeld $\vec{v}(t,\vec{x})$ von Fluidelementen (die selbst aus einer sehr großen Zahl mikroskopischer Teilchen bestehen, die aber auf den makroskopisch relevanten Längenskalen als punktförmig betrachtet werden können).

Der Zusammenhang zwischen der praktisch unmöglichen und auch gar nicht notwendigen mikroskopischen Beschreibung eines Vielteilchensystems mit $\sim 10^{24}$ Freiheitsgraden wird nun durch die **statistische Mechanik** hergestellt. Dabei betrachtet man nicht die $\sim 6 \cdot 10^{24}$ Orts- und Impulsfreiheitsgrade, die zur detaillierten Beschreibung im Sinne der Punktmechanik nötig wären, sondern zunächst einmal **Einteilchenphasenraumverteilungen** $f(t,\vec{x},\vec{p})$. Dabei ist $\mathrm{d}N = \mathrm{d}^3x\mathrm{d}^3p f(t,\vec{x},\vec{p})$ die Anzahl von Teilchen, die sich zur Zeit t um den Ort $\vec{x}$ innerhalb eines kleinen Volumens d^3x befinden und einen Impuls innerhalb eines Impulsraumvolumens d^3p um den Impuls $\vec{p}$ herum besitzen. Dann kann man die zeitliche Änderung dieser Verteilung beschreiben, indem man zum Einen berücksichtigt, dass aufgrund der momentanen Orte und Impulse die Teilchen im durch $(\vec{x},\vec{p})$ gegebenen Einteilchenphasenraum strömen und zum Anderen auch aneinander stoßen können. Für ein nicht zu dichtes System von Teilchen (einem nicht zu stark komprimierten Gas) darf man dabei annehmen, dass im wesentlichen **elastische Zweiteilchenstöße** relevant sein werden. Dies wird dann durch Übergangswahrscheinlichkeiten beschrieben, dass bei einem Stoß die Impulse $(\vec{p}_1,\vec{p}_2)$ zu Impulsen $(\vec{p}_1',\vec{p}_2')$ übergehen, wobei freilich die Impulserhaltung $\vec{p}_1+\vec{p}_2 \to \vec{p}_1'+\vec{p}_2'$ gilt. Dazu braucht man eigentlich die sog. Zweiteilchenphasenraumverteilung $f^{(2)}(t,\vec{x}_1,\vec{p}_1,\vec{x}_2,\vec{p}_2)$, die die Phasenraumdichte für das Auftreten von Teilchenpaaren angibt. Hier kommt nun die geniale Idee **Boltzmanns** ins Spiel, wonach wir selbst die genauen Korrelationen zwischen solchen Teilchenpaaren i.a. nicht erfassen können. Andererseits erfolgen auf einer makroskopischen Zeitskala ungeheuer viele Stöße, so dass man diese statistisch beschreiben kann, wobei man annehmen darf, dass das Herauspicken zweier Teilchen im Phasenraum durch statistisch unabhängige Verteilungen beschrieben werden darf. Diese **Hypothese vom molekularen Chaos** führt dann dazu, dass man für die Zweiteilchenphasenraumverteilung einfach das Produkt zweier Einteilchenphasenraumverteilungen verwenden darf: $f^{(2)}(t,\vec{x}_1,\vec{p}_1,\vec{x}_2,\vec{p}_2) = f(t,\vec{x}_1,\vec{p}_1)f(t,\vec{x}_2,\vec{p}_2)$. Zusammen mit den entsprechenden Übergangsraten bzw. **Streuquerschnitten**, die man bei bekannter Paarwechselwirkung zwischen den Teilchen aus der Punktmechanik berechnen kann, erhält man dann eine sog. **Transportgleichung**, auf die wir aber nicht näher eingehen wollen.

Wichtig ist nur zu bemerken, dass man mit dieser Gleichung in der Tat große Teile der phänomenologischen Thermodynamik als statistische Mittelwertbildung über viele mikroskopische Freiheitsgrade verstehen kann. Außerdem ergibt sich interessanterweise aus Boltzmanns Betrachtungen auch das **Anwachsen der Entropie**, was eine bestimmte **„thermodynamische Zeitrichtung“** auszeichnet. Dieses Anwachsen der Entropie kommt dabei durch die Annahme vom **molekularen Chaos** zustande, und es wird klar, dass die **thermodynamische Zeitrichtung** mit der **kausalen Zeitrichtung**, die wir ja bereits zu Beginn in den Annahmen über die Grundlagen der Mechanik zugrundegelegt haben, übereinstimmt, und das obwohl die meisten Wechselwir-

kungen in der Natur **zeitumkehrinvariant** sind, d.h. ein Vorgang, wie er sich darstellt, wenn man etwa einen Film dieses Vorgangs rückwärts laufen lässt, ist aufgrund der mechanischen Naturgesetze genauso möglich wie der tatsächliche Vorgang. Dass wir also bemerken, dass bei Alltagsvorgängen ein Film rückwärts läuft, weil z.B. statt dass eine Tasse vom Tisch fällt und am Boden in viele Bruchstücke zersplittert, der Scherbenhaufen einer Tasse spontan vom Boden auf dem Tisch als ganze Tasse wieder ankommt, kann also nicht durch die fundamentalen Naturgesetze erklärt werden. Vielmehr ergibt die statistische Betrachtung, dass es äußerst unwahrscheinlich ist, dass durch eine Vielzahl von mikroskopischen Stößen zwischen Molekülen „zufällig" genau die Anfangsbedingungen für all diese Moleküle zustandekommen, die diesen Vorgang bewirken würden. Dabei stellt sich heraus, dass die Entropie auch als **Maß für die Unordnung** angesehen werden kann, und das Anwachsen der Entropie bedeutet einfach, dass ein Übergang von einem eher ungeordneten Zustand (Scherbenhaufen auf dem Boden) zu einem sehr geordneten Zustand (ganze Tasse auf dem Tisch) extrem unwahrscheinlich ist. Typische Vorgänge auf makroskopischer Betrachtungsebene erfolgen also umgekehrt gesagt mit einer immensen Wahrscheinlichkeit in die Richtung, die von anfangs geordneten Zuständen durch viele (quasi) zufällige Stöße zwischen den mikroskopischen Konstituenten zu einem ungeordneten Zustand übergehen. Das thermodynamische Gleichgewicht ist demnach ein **Zustand größtmöglicher Unordnung** unter den durch die Erhaltungssätze (Energie, Impuls, Drehimpuls, Schwerpunktsgeschwindigkeit) bedingten Einschränkungen.

Als weiteren fundamentalen Baustein des **klassischen Weltbildes** der Physik haben wir in der Vorlesung Theorie 2 für Lehramt L3 auch die **Maxwell-Theorie** der elektromagnetischen Phänomene ausführlich kennengelernt. Dies umfasst die elektromagnetischen Kräfte zwischen geladenen Teilchen oder kontinuumsmechanischer Systeme (Lorentz-Kraft), die Erzeugung von elektromagnetischen Feldern aus den Ladungs- und Stromverteilungen der Materie, die Existenz und das Verhalten elektromagnetischer Wellen/Strahlung und schließlich die Phänomene der Optik, denn Licht wurde ja durch die Maxwellsche Theorie als elektromagnetische Wellen (in einem bestimmten Frequenz- bzw. Wellenlängenbereich, auf den unsere Augen empfindlich ist) erkannt. Wir haben auch gesehen, dass die Wechselwirkung zwischen Materie und elektromagnetischen Feldern durch einige wenige **Materialgrößen** wie Dielektrizitätszahlen $\epsilon = \epsilon_r \epsilon_0$, magnetische Permeabilitäten $\mu = \mu_r \mu_0$ und elektrische Leitfähigkeiten σ charakterisiert werden können. Diese müssen im Rahmen dieser makroskopischen phänomenologischen Elektrodynamik aus Messungen für jedes Material bestimmt werden. Im Rahmen der klassischen statistischen Physik kann man in manchen Fällen auch zumindest qualitativ solche Größen berechnen. Z.B. kann man die Leitfähigkeit eine Metalls verstehen, indem man annimmt, dass im Metall quasi frei bewegliche Elektronen vorhanden sind, die sich im positiv geladenen starren Kristallgitter des Metalls befinden, wobei es immer wieder zu Stößen der Elektronen an Fehlstellen des Kristallgitters bzw. durch kleine Ladungsdichteschwankungen auf Grund der thermischen Schwingungen des Gitters kommt. Dadurch kommt es zu einem Energie-Impuls-Austausch zwischen Elektronen und Gitter, also zu einer Bewegung der Elektronen mit Reibung. Dies ermöglicht im Prinzip die Berechnung der elektrischen Leitfähigkeit, indem man die Bewegungsgleichung für ein Elektron bei vorgegebenen äußeren elektrischen Feldern analysiert (vgl. [Hee18c]).

Es ergaben sich allerdings gerade im Hinblick auf die Wechselwirkung zwischen geladener Materie und elektromagnetischer Strahlung auch einige Diskrepanzen zwischen der aufgrund dieser klassischen Theorien erwarteten und den tatsächlich beobachteten Phänomenen, von denen wir im Folgenden einige für die Entdeckung der Quantentheorie wichtige Beispiele anführen.

2.2 1900: Planck und die Hohlraumstrahlung

In diesem Abschnitt benötigen wir einige grundlegenden Vorkenntnisse über statistische Physik, d.h. er ist nicht vollständig abgeschlossen im Hinblick auf die Theorievorlesungsreihe für L3. Ich füge ihn wegen der (nicht nur historischen) Wichtigkeit dieses Themas trotzdem hier ein. Für diejenigen, die den Abschnitt überspringen wollen, sei bemerkt, dass für die weitere Entwicklung der Quantentheorie die wesentliche Erkennt-

nis aus Plancks Herleitung der Strahlungsformel war, dass bei der Wechselwirkung der Materie mit elektromagnetischer Strahlung Energie nur in diskreten Portionen der Größe $E_\omega = \hbar\omega$ ausgetauscht werden kann, was im klaren Widerspruch zur klassischen Maxwell-Theorie steht. Diese Erkenntnis wurde dann von Einstein benutzt, um die Idee des **Welle-Teilchen-Dualismus'** in die Physik einzuführen und der elektromagnetischen Strahlung neben Wellen- auch Teilcheneigenschaften zuzuschreiben, was allerdings gerade im Hinblick auf die später **Photonen** genannten Quanten des quantisierten elektromagnetischen Feldes im Sinne der derzeit gültigen Behandlung im Rahmen der relativistischen Quantenfeldtheorie nicht besonders zutreffend ist. Insbesondere stellt sich heraus, dass man Photonen nicht lokalisieren kann, d.h. man kann im strengen Sinne noch nicht einmal die Observable „Ort eines Photons" definieren. Trotzdem waren die Ideen im Zusammenhang mit der durch Einstein und Bohr begründeten „alten Quantentheorie" wichtig für die Entwicklung der bis heute gültigen „modernen Quantentheorie", und die Physiker der damaligen Zeit (ca. 1900-1925) waren sich der Vorläufigkeit dieser Ideen auch bewusst.

2.2.1 Hohlraumstrahlung im Rahmen der klassischen Physik

Die Frage nach der Spektralverteilung der **Hohlraumstrahlung** war für die Physiker in der zweiten Hälfte des 19. Jhs. von besonderem Interesse, weil aufgrund rein thermodynamischer Überlegungen dieses Spektrum **universell** ist, d.h. nicht abhängig ist von den Eigenschaften des den Hohlraum bildenden Materials sondern einzig von der Temperatur, die durch die Temperatur der Wände des Hohlraums festgelegt sein muss, und diese rein thermodynamischen Betrachtungen behalten auch im Rahmen der Quantentheorie ihre Gültigkeit. Qualitativ handelt es sich um die Frage, wie man die unterschiedlichen „Farben" des Lichtes, das von heißen Körpern abgestrahlt wird, quantitativ verstehen kann. In der Tat beobachtet man ja bei einem Stück Metall bei Zimmertemperatur gar keine sichtbare Strahlung. Erhitzt man es dann immer mehr, beginnt es zunächst rot, dann gelb und schließlich weiß zu glühen. I.a., hängt bei dieser Betrachtung das Spektrum auch noch von der Materialbeschaffenheit insofern ab als verschiedene Materialien ein verschiedenes Absorptionsvermögen besitzen. Um ein wirklich universelles Strahlungsgesetz zu erhalten, betrachtet man daher den Idealfall eines **schwarzen Körpers**, der alle auf ihn treffende Strahlung komplett absorbiert. Im atomistischen Bild vom Aufbau der Materie aus geladenen Teilchen (positiv geladenen Atomkernen und negativ geladenen Elektronen) ist es allerdings klar, dass bei einer endlichen Temperatur der Körper auch Strahlung abgeben muss, denn die chaotische Wärmebewegung entspricht ja beschleunigten Ladungen, die demzufolge elektromagnetische Wellen abstrahlen.

Die beste Näherung für einen idealen schwarzen Strahler ist demnach ein Hohlraum mit einer kleinen Öffnung mit (fast) ideal reflektierenden Wänden. Gelangt nämlich em. Strahlung in diesen Hohlraum, wird sie an den Wänden dauernd hin- und herreflektiert, und es gelangt kaum noch Strahlung nach außen. Daher nennt man die Strahlung eines schwarzen Körpers auch **Hohlraumstrahlung**. Wir suchen also die **spektrale Verteilung** für die em. Strahlung im thermodynamischen Gleichgewicht mit den Wänden eines Hohlraums, dessen Wände sehr lange auf einer konstanten Temperatur gehalten werden. Thermodynamisches Gleichgewicht bedeutet dabei, dass pro Zeiteinheit die Wände genausoviel Energie des Strahlungsfeldes absorbieren wie emittieren, d.h. die Strahlungsenergie im ganzen Hohlraum zeitlich konstant und auch räumlich homogen verteilt ist.

In diesem Abschnitt betrachten wir nun die Theorie dieser Strahlung im Rahmen der klassischen Elektrodynamik unter Zuhilfenahme der Thermodynamik und statistischen Physik im thermodynamischen Gleichgewicht.

Zunächst betrachten wir das Strahlungsfeld. Wie jedes elektromagnetische Feld können wir es durch Fourier-Entwicklung nach ebenen Wellen darstellen. Dazu benötigen wir zunächst einmal die durch die **Randbedingungen** bestimmten entsprechenden Ebenen-Wellen-Lösungen der Maxwell-Gleichungen, die **Feldmoden**. Wir betrachten dazu der Einfachheit halber einen würfelförmigen Hohlraum (Kantenlänge L) mit quasi ideal leitenden Wänden. Es genügt, sich zu erinnern, dass die Tangentialkomponenten des elektrischen Feldes an den Wänden verschwinden müssen, da wegen $\vec{j} = \sigma\vec{E}$ sonst sehr große Ströme in der Wand fließen würden.

Dieses Problem wird in [Hee18c] ausführlich gelöst. Demnach gibt es Feldmoden mit diskreten Wellenvektoren

$$\vec{k} = \frac{\pi}{L}\begin{pmatrix} n_1 \\ n_2 \\ n_3 \end{pmatrix}, \tag{2.2.1}$$

wobei im Wesentlichen $n_i \in \mathbb{N} = \{1, 2, 3, \ldots\}$, und zu jedem dieser „erlaubten Wellenvektoren" gibt es noch zwei unabhängige Polarisationszustände. Dabei sind es nur zwei und nicht drei Polarisationszustände, weil elektromagnetische Wellen strikt **transversal** sind. Es gibt zwar auch noch Moden, wo eine der ganzzahligen Komponenten $n_i = 0$ wird, und für diese Moden gibt es sogar nur einen Polarisationszustand. Allerdings sind dies im Grenzfall eines sehr großen Hohlraums (groß im Vergleich zur typischen Wellenlänge der Strahlung, s.u.) vernachlässigbar wenige Zustände.

Wie ebenfalls in [Hee18c] gezeigt, entspricht jeder Wellenmode zum Wellenvektor $\vec{k}$ zwei **harmonische Oszillatoren** mit Kreisfrequenz $\omega_{\vec{k}} = c|\vec{k}|$. Um die Spektralverteilung der Hohlraumstrahlung im Gleichgewicht zu berechnen, benötigen wir also nur noch (a) die mittlere Energie eines harmonischen Oszillators im thermodynamischen Gleichgewicht bei gegebener Temperatur T und (b) die Anzahl der Moden im Frequenzintervall $(\omega, \omega + \mathrm{d}\omega)$.

Für die erste Frage müssen wir nun auf die statistische Physik zurückgreifen. In diesem Abschnitt betrachten wir die harmonischen Oszillatoren klassisch. Nach den Gesetzen der statistischen Physik ist die mittlere Energie eines harmonischen Oszillators, der sich im thermischen Gleichgewicht mit einem Wärmebad der Temperatur T befindet durch

$$\langle E \rangle = \frac{1}{Z}\int_{\mathbb{R}^3} \mathrm{d}^3x \int_{\mathbb{R}^3} \mathrm{d}^3p\, H(x,p)\exp[-\beta H(x,p)], \quad Z = \int_{\mathbb{R}^3} \mathrm{d}^3x \int_{\mathbb{R}^3} \mathrm{d}^3p \exp[-\beta H(x,p)] \tag{2.2.2}$$

gegeben. Dabei ist

$$H(x,p) = \frac{1}{2m}p^2 + \frac{m\omega^2}{2}x^2 \tag{2.2.3}$$

die **Hamilton-Funktion** für einen harmonischen Oszillator (s. Theoretische Physik 1) und $\beta = 1/(k_\mathrm{B}T)$ mit der Boltzmann-Konstanten $k_\mathrm{B} = 1{,}38064852 \cdot 10^{-23}\,\mathrm{J/K}$, wobei K die SI-Einheit für die absolute Temperatur ist. Seit der Neudefinition des SI in 2019 ist dieser Wert für die Boltzmann-Konstante exakt. Die Auswertung der Gauß-Integrale in (2.2.2) liefert schließlich

$$\langle E \rangle = k_\mathrm{B}T, \tag{2.2.4}$$

d.h. jeder Wellenmode (2.2.1) des elektromagnetischen Feldes entspricht im Hohlraum bei einer Temperatur T eine mittlere Energie $\langle E \rangle$.

Wir wollen nun die mittlere Energiedichte im Frequenzintervall $[\omega, \omega + \mathrm{d}\omega]$ bestimmen. Dazu müssen wir nur die Anzahl der in diesem Frequenzintervall enthaltenen Moden berechnen. Dabei entspricht gemäß (2.2.1) diese Anzahl der Anzahl von Punkten in einem diskreten Gitter im $\vec{k}$-Raum mit Gitterabstand π/L, und es gilt $\omega = c|\vec{k}|$, d.h. dem Frequenzintervall entspricht im $\vec{k}$-Raum eine Kugelschale der Dicke $\mathrm{d}\omega/c$ mit Radius ω/c. Dabei ist aber nur der Oktant der Kugelschale mit positiven Komponenten k_j zu berücksichtigen. Die Anzahl der Moden in dieser Kugelschale im $\vec{k}$-Raum ist somit

$$\mathrm{d}N = \mathrm{d}\omega 2\frac{1}{8}4\pi(\omega/c)^2\frac{L^3}{\pi^3c^3} = \mathrm{d}\omega\omega^2\frac{V}{c^3\pi^2}. \tag{2.2.5}$$

Dabei rührt der Faktor 2 daher, dass zu jeder Mode mit gegebenem $\vec{k}$ zwei unabhängige harmonische Oszillatoren gleicher Frequenz entsprechen, weil zu jedem $\vec{k}$ zwei Polarisationsfreiheitsgrade mit $\omega = c|\vec{k}|$ existieren.

Damit ist die spektrale Energiedichte durch

$$u(\omega, T) = \frac{\mathrm{d}N}{\mathrm{d}\omega}\langle E\rangle = \frac{\omega^2}{c^3\pi^2}k_\mathrm{B}T. \tag{2.2.6}$$

gegeben. Dieses Resultat für die spektrale Energiedichte der Hohlraumstrahlung, die auf der klassischen statistischen Physik beruht, heißt **Rayleigh-Jeans-Gesetz**.

Allerdings endet dieses Gesetz in einer „Katastrophe“, denn berechnet man die gesamte Energie, müssen wir ja über alle $\omega > 0$ integrieren, und dieses Integral divergiert.

2.2.2 Folgerungen aus der phänomenologischen Thermodynamik

Bevor wir zu Plancks Lösung durch „einen Verzweiflungsakt“, der die „Quantenrevolution“ im Jahre 1900 eingeleitet hat, kommen, stellen wir zunächst eine Betrachtung an, die auf der Anwendung der phänomenologischen Thermodynamik beruht. Dazu gehen wir vom 1. Hauptsatz der Thermodynamik aus. Offenbar wird die Energie der Wärmestrahlung im Hohlraum nur durch das Volumen und die Temperatur bestimmt. Der erste Hauptsatz lautet also bei Zustandsänderungen

$$\mathrm{d}U = T\mathrm{d}S - p\mathrm{d}V = T\left[\left(\frac{\partial S}{\partial T}\right)_V \mathrm{d}T + \left(\frac{\partial S}{\partial V}\right)_T \mathrm{d}V\right] - P\mathrm{d}V, \tag{2.2.7}$$

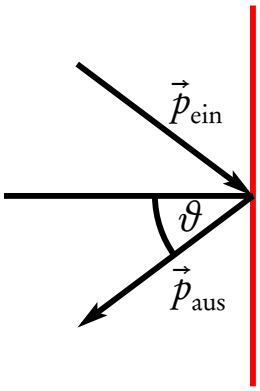
wobei U die gesamte innere Energie der Hohlraumstrahlung ist und P der durch diese Strahlung auf die Wände ausgeübte Druck ist.

Wir benötigen für das Folgende zum Glück nur die Beziehung zwischen der Energiedichte u der Strahlung und dem Strahlungsdruck. Dazu erinnern wir uns aus der Theorie 2-Vorlesung, dass für ebene Wellen die (zeitgemittelte) Energiedichte des elektromagnetischen Feldes durch

$$u = \left\langle \frac{\epsilon_0}{2}\vec{E}^2 + \frac{1}{2\mu_0}\vec{B}^2 \right\rangle = \frac{\epsilon_0}{2}\hat{\vec{E}}^2, \tag{2.2.8}$$

gegeben ist, wobei $\hat{\vec{E}}$ die Amplitude des elektrischen Feldes ist, und für die Impulsdichte

$$\vec{g}^{(\mathrm{em})} = \epsilon_0\left\langle \vec{E}\times\vec{B}\right\rangle = \frac{1}{c}u\frac{\vec{k}}{k}. \tag{2.2.9}$$



Jeder Schwingungsmode im Hohlraum entspricht nun eine stehende ebene Welle. Den Druck auf die Wände erhalten wir daraus, dass wir uns diese stehende Welle dadurch entstanden denken können, dass eine ebene Welle mit Wellenvektor $\vec{k}$ in Richtung auf eine Wand zuläuft und dort reflektiert wird. Ist der Winkel zwischen $\vec{k}$ und der Normale der Wand θ, wird die Welle nach dem Reflexionsgesetz gestreut. Der Impulsübertrag auf die Wand ist wegen der Impulserhaltung in Richtung senkrecht zur Wand vom Betrag $|\delta\vec{g}| = 2/c\,u\cos\theta$ und daher der Impulsübertrag pro Zeit und Fläche $|\delta\vec{g}|c\cos\theta = 2u\cos^2\theta$. Dies mitteln wir nun über alle Richtungen, d.h. der Druck wird

$$P = \frac{1}{4\pi}\int_{\pi/2}^{\pi}\mathrm{d}\vartheta\int_0^{2\pi}\mathrm{d}\varphi\sin\vartheta\, 2u\cos^2\vartheta = \frac{1}{3}u. \tag{2.2.10}$$

Wir wollen nun $U(T,V) = Vu(T)$ als funktion von V und T bestimmen. Dazu gehen wir vom 1. Hauptsatz (2.2.7) aus. Zunächst folgt

$$\left(\frac{\partial U}{\partial V}\right)_T = T\left(\frac{\partial S}{\partial V}\right)_T - P. \tag{2.2.11}$$

Wir haben die Zustandsgleichung (2.2.10) für Strahlung aber als Funktion von P, d.h. wir müssen noch S in (2.2.11) eliminieren. Die entsprechende Maxwell-Beziehung finden wir durch die Betrachtung der freien Energie $F = U - TS$. Aus (2.2.7) folgt nämlich

$$\mathrm{d}F = \mathrm{d}U - T\mathrm{d}S - S\mathrm{d}T = -P\mathrm{d}V - S\mathrm{d}T. \tag{2.2.12}$$

Da $F(T,V)$ eine Zustandsgröße und folglich $\mathrm{d}F$ ein totales Differential ist, folgt aus (2.2.12) $S = -(\partial F/\partial T)_V$

$$\left(\frac{\partial S}{\partial V}\right)_T = -\partial_V \partial_T F = -\partial_T \partial_V F = \left(\frac{\partial P}{\partial T}\right)_V. \tag{2.2.13}$$

Damit wird zusammen mit (2.2.10)

$$\left(\frac{\partial U}{\partial V}\right)_T = u = T\left(\frac{\partial P}{\partial T}\right)_V - P = \frac{T}{3}u'(T) - \frac{u}{3} \Rightarrow u'(T) = \frac{4u}{T} \Rightarrow u(T) = aT^4. \tag{2.2.14}$$

Dabei ist a eine Integrationskonstante, die wir ohne ein verbessertes Modell für die statistische Behandlung des Problems nicht gewinnen können. Mit dem 1. Hauptsatz können wir aber auch $S = S(V,T)$ finden. Demnach ist

$$\mathrm{d}S = \frac{1}{T}\mathrm{d}U + \frac{P}{T}\mathrm{d}V = \frac{u}{T}\mathrm{d}V + \frac{V}{T}u'(T)\mathrm{d}T + \frac{u}{3T}\mathrm{d}V = \frac{4a}{3}T^3\mathrm{d}V + 4aVT^2\mathrm{d}T. \tag{2.2.15}$$

Daraus erhalten wir sofort, dass

$$S = \frac{4}{3}aVT^3 \tag{2.2.16}$$

sein muss, wobei wir noch den 3. Hauptsatz der Thermodynamik benutzt haben, wonach $S = 0$ für $T = 0$ gelten muss.

Betrachten wir nun eine adiabatische Zustandsänderung, d.h. eine Änderung des Volumens ohne Wärmeaustausch und betrachten wieder die spektrale Energiedichte $u(\omega, T)$. Diese Funktion beinhaltet als einzige Länge die Wellenlänge der Strahlung mit der Kreisfrequenz ω. Ändern wir also die Kantenlänge des oben betrachteten Hohlraums von L auf $L' = \alpha L$ muss $\lambda' = \alpha\lambda$ bzw. $\omega' = \omega/\alpha$ sein. Bei der adiabatischen Zustandsänderung ist weiter $S = \text{const}$ und wegen (2.2.16

$$\alpha^3 VT'^3 = VT^3 \Rightarrow \frac{T'}{T} = \frac{1}{\alpha} \Rightarrow \frac{\omega'}{T'} = \frac{\omega}{T}. \tag{2.2.17}$$

Andererseits gilt für die spektrale Energiedichte

$$\frac{u(\omega', T')}{u(\omega, T)} = \frac{\mathrm{d}U'/\mathrm{d}\omega'}{\mathrm{d}U/\mathrm{d}\omega} = \alpha\frac{U'}{U} = \alpha\frac{T'^4}{T^4} = \frac{1}{\alpha^3}. \tag{2.2.18}$$

Multiplizieren wir dies mit $\omega^3/\omega'^3 = \alpha^3$, folgt schließlich

$$\frac{u(\omega', T')}{\omega'^3} = \frac{u(\omega, T)}{\omega^3}, \tag{2.2.19}$$

d.h. $u(\omega, T)/\omega^3$ ist eine adiabatische Invariante und daher wegen (2.2.17)

$$u(\omega, T) = \omega^3 F\left(\frac{\omega}{T}\right). \tag{2.2.20}$$

Dies ist das **Wiensche Gesetz**.

Suchen wir nun für fest vorgegebenes T diejenige Kreisfrequenz $\omega_{\max}$, für die $u(\omega, T)$ maximal wird. Das verlangt

$$\partial_\omega u(\omega_{\max}, T) = 3\omega_{\max}^2 F(\omega_{\max}/T) + \frac{\omega_{\max}^3}{T} F'(\omega_{\max}/T) = 0. \tag{2.2.21}$$

Multipliziert man dies mit $T/\omega_{\max}^3$ und $\omega_{\max}/T = x_{\max}$, folgt als Bedingung

$$3F(x_{\max})/x_{\max} = F'(x_{\max}). \tag{2.2.22}$$

Es ist also

$$\frac{\omega_{\max}}{T} = x_{\max} = \text{const.} \tag{2.2.23}$$

Dies ist als das **Wiensche Verschiebungsgesetz** bekannt, d.h. $\omega_{\max}$ verschiebt sich proportional zur Temperaturänderung, wenn man die Temperatur der Hohlraumwände ändert.

Wir können die unbekannte Funktion F nun wieder nur durch eine Betrachtung über die statistische Physik erhalten. Es ist klar, dass auch für das Rayleigh-Jeans-Gesetz (2.2.20) gelten muss. In der Tat gilt gemäß (2.2.6)

$$\frac{u_{\mathrm{RJ}}(\omega, T)}{\omega^3} = F_{\mathrm{RJ}}\left(\frac{\omega}{T}\right) = \frac{k_{\mathrm{B}}}{c^3\pi^2}\left(\frac{\omega}{T}\right)^{-1}. \tag{2.2.24}$$

Wie aber oben schon erläutert, kann dieses auf der klassischen Physik beruhende Gesetz nicht gelten, weil die Gesamtenergie unendlich wird.

2.2.3 Plancks „Verzweiflungstat“

Planck hatte nun im Jahre 1900 sehr genaue experimentelle Daten von Kurlbaum und Rubens zur Spektralverteilung der Hohlraumstrahlung für einen weiten Frequenzbereich erhalten. Die Motivation für diese Messungen an der Technischen Reichsanstalt war, durch das universelle damals noch unbekannte Strahlungsgesetz ein sehr genaues universelles Maß für Lichtstärken zu gewinnen, um genaue Normen für die Beleuchtungsindustrie bereitzustellen. Planck hatte nun nach genauer Auswertung der Daten festgestellt, dass das Rayleigh-Jeans-Gesetz für kleine Frequenzen ω eine gute Näherung ist.

In Anbetracht der Ultraviolettkatastrophe der klassischen Rayleigh-Jeans-Näherung hatte Wien durch Vergleich mit früheren Messdaten zum Hohraumstrahlungsspektrum den Ansatz

$$F_{\mathrm{W}}\left(\frac{\omega}{T}\right) = c_1 \exp\left(-\frac{c_2\omega}{T}\right) \tag{2.2.25}$$

gefunden. Diese beschrieb die Messdaten für große Frequenzen ω recht gut. Nun kann man aus jeder Annahme für F einen Ausdruck für die mittlere Entropie eines harmonischen Oszillators als Funktion der mittleren Energie des Oszillators bestimmen, und aus den entsprechenden Ausdrücken für die Rayleigh-Jeans-Lösung und den Ansatz von Wien für $u(T, \omega)$ interpolierte Planck zunächst einfach und erhielt damit auf rein empirischem Wege seine exakt die Messdaten beschreibende **Planck-Verteilung**.

Als gewissenhafter theoretischer Physiker konnte sich Planck aber damit nicht zufrieden geben, denn es musste eine theoretische Herleitung basierend auf den Prinzipien der statistischen Physik gewonnen werden, um zu verstehen, wie die korrekte Strahlungsformel zu verstehen ist, d.h. wie eine endliche Gesamtstrahlungsenergie im Hohlraum theoretisch zustandekommt.

Planck fand eine Lösung, die er bis zum Ende seines Lebens zwar nicht mochte, er konnte aber auch keine Lösung im Rahmen der von ihm favorisierten klassischen Physik finden, und das ist auch bis heute nicht gelungen. Stattdessen war die Lösung der Anstoß für die Entwicklung der Quantentheorie, die die bislang umfassendste Theorie der Physik ist und alle Phänomene korrekt beschreibt. Allerdings ist bis heute noch keine befriedigende Quantentheorie der Gravitation gefunden, so dass auch heute die theoretische Physik als noch immer nicht abgeschlossen betrachtet wird.

Plancks „revolutionärer Verzweiflungsakt“ bestand darin, anzunehmen, dass das Strahlungsfeld bzw. die dessen Moden zur Frequenz ω im Hohlraum entsprechenden Oszillatoren mit den Hohlraumwänden nicht beliebige Energieportionen austauschen können sondern lediglich ganzzahlige Vielfache einer kleinsten Menge $\epsilon = \epsilon(\omega)$, einem **Energie-Quantum**. Die mittlere Energie für jeden Oszillator im Gleichgewicht mit den Oszillatorwänden der Temperatur T sind dann nach den Prinzipien der statistischen Physik nämlich

$$\langle E\rangle = \frac{1}{Z}\sum_{n=0}^{\infty} n\epsilon \exp\left(-n\frac{\epsilon}{k_\mathrm{B}T}\right), \quad Z = \sum_{n=0}^{\infty}\exp\left(-n\frac{\epsilon}{k_\mathrm{B}T}\right). \tag{2.2.26}$$

Dabei entspricht $n = 0$ dem Fall, dass gar keine Schwingungsmode angeregt ist, was freilich der Energie $0 = 0\epsilon$ im Hohlraum entspricht.

Hier lässt sich nun $\langle E\rangle$ recht einfach ausrechnen. Die Zustandssumme ist nämlich einfach eine geometrische Reihe, deren Summe wir kennen:

$$Z = \frac{1}{1-\exp[-\epsilon/(k_\mathrm{B}T)]}. \tag{2.2.27}$$

Die andere Summe in (2.2.27) erhalten wir, indem wir Z nach ϵ ableiten, denn daraus ergibt sich

$$\partial_\epsilon Z = -\frac{1}{k_\mathrm{B}T}\sum_{n=0}^{\infty} n\exp\left(-n\frac{\epsilon}{k_\mathrm{B}T}\right) = -\frac{1}{k_\mathrm{B}T}\frac{\exp[-\epsilon/(k_\mathrm{B}t)]}{\{1-\exp[-\epsilon/(k_\mathrm{B}T)]\}^2}. \tag{2.2.28}$$

Damit wird

$$\langle E\rangle = -\frac{k_\mathrm{B}T\epsilon}{Z}\partial_\epsilon Z = \frac{\epsilon\exp[-\epsilon/(k_\mathrm{B}T)]}{1-\exp[-\epsilon/(k_\mathrm{B}T)]} = \frac{\epsilon}{\exp[\epsilon/(k_\mathrm{B}T)]-1}. \tag{2.2.29}$$

Setzen wir dies nun in die immer noch gültige Formel (2.2.5) für die Anzahldichte der Moden bei der Frequenz ω ein, erhalten wir

$$u_\mathrm{Pl}(\omega,T) = \frac{\omega^2}{c^3\pi^2}\langle E\rangle = \frac{\omega^2}{c^3\pi^2}\frac{\epsilon}{\exp[\epsilon/(k_\mathrm{B}T)]-1}. \tag{2.2.30}$$

Nun muss aber immer noch das allgemeingültige Wiensche Gesetz (2.2.20) gelten, d.h. es muss $u_\mathrm{Pl}/\omega^3 = F_\mathrm{Pl}(\omega/T)$ sein. Dies verlangt nun zwingend, dass $\epsilon/\omega = \hbar = \mathrm{const}$ ist. Dabei muss $\hbar$ eine im Rahmen der *klassischen Physik unbekannte fundamentale Naturkonstante sein*, denn das Spektrum eines Hohlraums im thermodynamischen Gleichgewicht muss nach Kirchhoffs Überlegungen vom Beginn dieses Kapitels eine vom Material der Wände unabhängige Funktion von ω und T sein, d.h. außer diesen Größen *können* in der Formel nur Naturkonstanten vorkommen. Es ist also

$$\epsilon = \hbar\omega, \tag{2.2.31}$$

und die Dimension von $\hbar$ ist demzufolge „Energie × Zeit“, und das entspricht der Dimension der **Wirkung** im **Hamiltonschen Prinzip der kleinsten Wirkung** (vgl. die Theorie 1-Vorlesung). Daher heißt $\hbar$ **modifiziertes Plancksches Wirkungsquantum**. Historisch verwendete Planck bei seinen Überlegungen über den harmonischen Oszillator bzw. die Feldmoden in einem Hohlraum statt der Kreisfrequenz ω die gewöhnliche Frequenz $\nu = 1/T_\mathrm{Osz}$, wobei T_Osz die Schwingungsdauer des Oszillators bzw. der Feldmoden ist. Nun ist aber einfach $\omega = 2\pi f$ und damit $\epsilon = hf = 2\pi\hbar f$, d.h. es ist $\hbar = h/(2\pi)$, wobei h das eigentliche **Plancksche Wirkungsquantum** ist.

Da es eine sehr genau bestimmbare universelle Naturkonstante ist, wird der Zahlenwert zur Definition der SI-Einheiten festgelegt, und zwar zu

$$h = 6{,}62607015\cdot 10^{-34}\,\mathrm{J\,s} \Rightarrow \hbar = \frac{h}{2\pi} = 1{,}05457182\cdot 10^{-34}\,\mathrm{J\,s}. \tag{2.2.32}$$

Damit wird die **Plancksche Strahlungsformel** (2.2.30)

$$u_{\mathrm{Pl}}(\omega,T)=\frac{\hbar\omega^3}{c^3\pi^2}\frac{1}{\exp[\hbar\omega/(k_{\mathrm{B}}T)]-1}. \tag{2.2.33}$$

In der Tat nimmt dieses Gesetz für $\hbar\omega \ll k_{\mathrm{B}}T$ die Form des Rayleigh-Jeans-Gesetzes (2.2.6) an, denn dann können wir im Nenner $\exp[\hbar\omega/(k_{\mathrm{B}}T)]-1 \simeq \hbar\omega/(k_{\mathrm{B}}T)$ nähern. Für $\hbar\omega \gg k_{\mathrm{B}}T$ wird hingegen die 1 im Nenner vernachlässigbar, und man erhält das Wiensche Strahlungsgesetz in der Form

$$u_{\mathrm{W}}(\omega,T)=\frac{\hbar\omega^3}{c^3\pi^2}\exp\left(-\frac{\hbar\omega}{k_{\mathrm{B}}T}\right). \tag{2.2.34}$$

Schließlich können wir noch die Konstante in (2.2.14) berechnen, denn es ist

$$u(T)=\int_0^\infty \mathrm{d}\omega\, u_{\mathrm{Pl}}(\omega,T). \tag{2.2.35}$$

Substituieren wir $x=\hbar\omega/(k_{\mathrm{B}}T)$ bzw. $\omega=k_{\mathrm{B}}T/\hbar$ folgt

$$u(T)=\frac{(k_{\mathrm{B}}T)^4}{c^3\pi^2\hbar^3}\int_0^\infty \mathrm{d}x\frac{x^3}{\exp x-1}=\frac{\pi^2k_{\mathrm{B}}^4}{15c^3\hbar^3}T^4. \tag{2.2.36}$$

Dabei haben wir ohne Beweis das Integral

$$\int_0^\infty \mathrm{d}x\frac{x^3}{\exp x-1}=\frac{\pi^4}{15} \tag{2.2.37}$$

verwendet. Oft wird auch der Strahlungsfluss pro Raumwinkel und Frequenz angegeben, für den gilt

$$B(\omega,T)=\frac{c}{4\pi}u_{\mathrm{Pl}}(\omega,T)=\frac{\hbar\omega^3}{4\pi^3c^2}\frac{1}{\exp[\hbar\omega/(k_{\mathrm{B}}T)]-1}. \tag{2.2.38}$$

Schließlich betrachtet man auch die gesamte auf eine ebene Fläche pro Flächeneinheit abgestrahlte Strahlungsleistung

$$P_{\mathrm{rad}}=2\pi\int_0^\infty \mathrm{d}\omega\int_0^{\pi/2}\mathrm{d}\theta\sin\theta\cos\theta B(\omega,T)=\frac{c}{4}u(T)=\frac{\pi^2k_{\mathrm{B}}^4}{60c^2\hbar^3}T^4. \tag{2.2.39}$$

Den Faktor

$$\sigma=\frac{\pi^2k_{\mathrm{B}}^4}{60c^2\hbar^3}\simeq 5{,}6704\cdot 10^{-8}\,\frac{\mathrm{J}}{\mathrm{m^2\,s\,K^4}} \tag{2.2.40}$$

heißt **Stefan-Boltzmann-Konstante**.

2.3 1905: Einstein und der Photoeffekt

Einstein zog aus Plancks Erkenntnis bei der Herleitung der korrekten Formel für das Spektrum der Hohlraumstrahlung den „revolutionären Schluss“, dass Licht neben den damals wohletablierten Welleneigenschaften auch **Teilcheneigenschaften** besitzt. Dabei ließ er sich von der Planckschen Erkenntnis leiten, dass etwa bei der Absorption von elektromagnetischen Wellen der Kreisfrequenz ω durch Materie immer eine ganzzahlige Anzahl **diskreter Energiequanten** der Größe $E_\gamma=\hbar\omega$ auf die Materie übertragen wird, wobei ω die

Kreisfrequenz einer elektromagnetischen Feldmode ist. Umgekehrt erfolgt auch die Emission von elektromagnetischer Strahlung in diesen diskreten Energiequanten. Wie schon oben bei der Herleitung des Strahlungsdrucks im Hohlraum gesehen, können wir analog auch die entsprechende Formel (2.2.9) für den mittleren Impuls der Strahlung so interpretieren, denn demnach ist

$$\vec{p}_{\text{em}} = V\vec{g}_{\text{em}} = \frac{U_{\text{em}}}{ck}\vec{k} = \frac{U_{\text{em}}}{\omega}\vec{k}. \tag{2.3.1}$$

Jetzt ist aber gemäß der obigen Überlegung $U_{\text{em}} = NE_\omega = N\hbar\omega$, wobei N einer mittlere Anzahl von Energiequanten entspricht. Demnach gilt

$$\vec{p}_{\text{em}} = \hbar N\vec{k}, \tag{2.3.2}$$

und das legt es nahe, anzunehmen, dass auch der Impulsübertrag zwischen elektromagnetischer Strahlen in ganzzahligen Portionen

$$\vec{p}_\gamma = \hbar\vec{k} \tag{2.3.3}$$

erfolgt. Dies hat nun in der Tat einiges gemeinsam mit Teilchen. Ein Teilchen in der Newtonschen Mechanik mit Impuls $\vec{p}$ besitzt ja die kinetische Energie $E = \vec{p}^2/(2m)$, und in einem total inelastischen Stoß mit Materie (z.B. ein Geschoss, das in einem Sandsack steckenbleibt) überträgt es diese gesamte Energie und diesen gesamten Impuls auf die Materie.

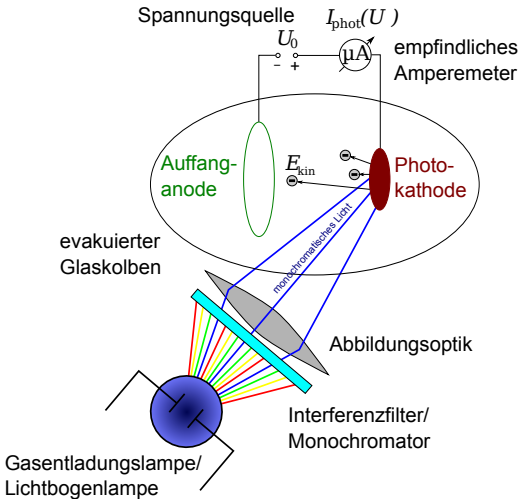
Andererseits sind aber die „Lichtteilchen“ Einsteins, oder **Photonen** wie sie bald genannt wurden, auch recht seltsam insofern als ihre Energie und Impuls durch die Wellengrößen Kreisfrequenz ω und Wellenvektor $\vec{k}$ definiert werden. Eine Masse können wir zunächst einmal nicht so einfach definieren. Charakteristisch ist aber eben auch das Auftreten der Planck-Konstante $\hbar$, was darauf hinweist, dass wir es hier mit „Teilchen“ zu tun haben, die nicht wirklich klassisch gedeutet werden können sondern quantentheoretisch verstanden werden müssen. Es ist auch nicht klar, inwiefern ein klassisches Teilchenbild zum Erfolg des Wellenbildes und den vielen experimentellen Tatsachen, die zeigen, dass Licht ein elektromagnetisches sich in Wellen ausbreitendes elektromagnetisches Feld ist, passt.

Andererseits gibt es aber eben auch Experimente, die die klassische Wellentheorie nicht korrekt beschreiben kann. Eines der einfachsten ist der sogenannte (äußere) **lichtelektrische Effekt**, der von Hertz und Hallwachs entdeckt wurde: Bestrahlten sie nämlich negativ geladenes Zinkblech mit Licht, entlud es sich. Etwas später hat Lennard mit seinen Kathodenröhren gezeigt, dass beim Bestrahlen der Metallkathode aus derselben Elektronen herausgelöst werden.

Es war ja schon bekannt, dass sich im Metall Elektronen befinden, die an das Metall gebunden sind, d.h. es wird eine bestimmte Austrittsarbeit Φ_{aus} (also zu überwindende Bindungsenergie) benötigt, um die Elektronen aus dem Metall zu befreien. Im klassischen Wellenbild würde sich der Photoeffekt dann so erklären, dass die Elektronen im Metall einfach hinreichend viel elektromagnetische Strahlungsenergie aufnehmen bis sich genügend Energie angesammelt hat, dass sie die Bindungsenergie überwinden können. Manche Elektronen sollten auch noch etwas mehr Energie aufnehmen, sodass sie mit einer bestimmten Geschwindigkeit von der Kathode wegfliegen. Bei höherer Intensität der Strahlung sollten sie im Mittel mehr kinetische Energie besitzen. Auch sollte es einige Zeit dauern, bis die Austrittsarbeit durch Übertragung von em. Energie des Lichtes aufgebracht ist. Beides widerspricht aber den Beobachtungstatsachen.

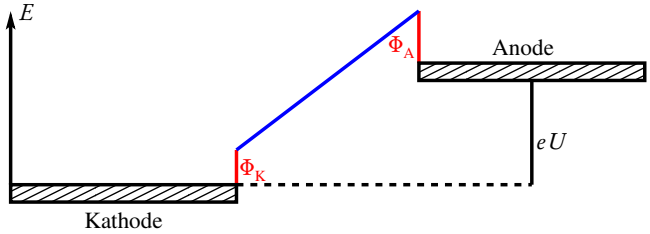
Zum Einen hängt es nicht von der Intensität des Lichtes ab, ob mehr oder weniger schnell, Elektronen aus der Metalloberfläche herausgelöst werden, sondern von der Frequenz des Lichts, d.h. wenn die Kreisfrequenz ω von monochromatischem Licht kleiner ist als eine bestimmte Grenzfrequenz, werden unabhängig davon wie groß man die Intensität dieses Lichts auch macht, überhaupt keine Elektronen freigesetzt. Ist die Frequenz hoch genug, werden allerdings auch bei sehr kleinen Lichtintensitäten sofort, also ohne Zeitverzögerung, Elektronen frei. Einzig die Anzahl der pro Zeiteinheit herausgelösten Photoelektronen ist proportional zur Lichtintensität.

Einstein konnte dies nun mit seiner Lichtquantenhypothese sehr einfach erklären. Demnach kommt hier die Photonenvorstellung zum Zuge: Um ein Elektron aus dem Metall herauszulösen, muss ja wenigstens ein Strahlungsquant mit einer Energie $\hbar\omega = E_{\text{aus}}$ vorliegen, aber dann genügt auch schon ein einziges solches Photon, d.h. das Elektron wird wie bei einem Stoß mit einem anderen Teilchen aus dem Metall herausgeschlagen. Dass mehrere Lichtquanten, die jedes für sich nicht genügend Energie hierzu besitzen, so zusammenwirken, dass mehrere das gleiche Elektron so treffen, dass es sich aus dem Metall löst, ist hingegen eher unwahrscheinlich, weshalb auch bei hohen Intensitäten solcher Strahlung praktisch gar keine Elektronen aus dem Metall herausgeschlagen werden.



Die Lichtquantenhypothese konnte mit dem Photoeffekt auch quantitativ untersucht werden. Hierzu hat **Millikan** die berühmte **Gegenfeldmethode** erfunden [Mil16a, Mil16b]. Er verwendete dazu zwei Metallplatten in einem evakuierten Glasgefäß, die er mit einer Spannung verbinden konnte. Dabei polte er die Spannung so, dass die Platte, die mit Licht bestrahlt wird, ein höheres Potential besitzt, d.h. die vom Licht herausgeschlagenen Elektronen müssen neben der Austrittsarbeit auch noch die entsprechende Gegenspannung überwinden. Nun maß Millikan den Strom der in dem unten gezeichneten Stromkreis fließt, offenbar kann nur ein Strom fließen, wenn die Elektronen, die aus der **Photokathode** durch das Licht herausgelöst werden, eine so hohe kinetische Energie besitzen, dass sie bis zur **Auffanganode** gelangen und dort in den Stromkreis eintreten. Wenn Millikan die Gegenspannung so einstellte, dass gerade kein Strom mehr floss, mussten gerade die Elektronen mit der maximalen kinetischen Energie durch die Gegenspannung vollständig abgebremst sein[1].

Wir wollen nun Einsteins Lichtquantenhypothese verwenden, um quantitativ zu verstehen, was Millikan dabei gemessen haben muss. Es sei gleich darauf hingewiesen, dass leider sehr viele Lehrbücher (sowohl Schulphysikbücher als auch einführende Universitätslehrbücher) die jetzt folgende Überlegung *falsch darstellen*. Der wichtige Punkt ist, dass die angelegte Spannung der Differenz der elektromagnetischen Potentiale *im Inneren* der Emitterkathode und der Auffanganode entspricht. In der folgenden Energiebilanz müssen wir aber die Bindungsenergie der Elektronen *sowohl in der Photokathode als auch in der Auffanganode* berücksichtigen.



Dazu betrachten wir das nebenstehende Diagramm, das die Energie des Elektrons an den verschiedenen Stellen im Stromkreis darstellt. Die außen anliegende Spannungsdifferenz U liegt dabei zwischen Kathode und Anode an. Ein Photostrom beginnt gerade dann zu fließen, wenn die Photonen eine Energie $E_\gamma = \hbar\omega$ besitzen, so dass das Elektron den höchstgelegenen Punkt in dem Energiediagramm erreichen kann. Dieser sitzt aber bei $E = eU + \Phi_A = \hbar\omega$.

Ausführlicher betrachtet muss zunächst das Elektron durch ein eintreffendes Photon mit der Kreisfrequenz ω aus der Kathode ausgelöst werden. Seine kinetische Energie direkt außerhalb der Kathode ist also

$$E_{\text{kin}} = \hbar\omega - \Phi_K. \tag{2.3.4}$$

Damit es mit dieser kinetischen Energie die Spannungsdifferenz zwischen der Oberfläche der Kathode bis zur Oberfläche der Anode, für die

$$U_{AK} = U - \Phi_K/e + \Phi_A/e \tag{2.3.5}$$

[1] Abbildung aus Wikipedia: Jkrieger - Eigenes Werk, CC BY-SA 3.0, `https://commons.wikimedia.org/w/index.php?curid=18130611`

gilt, überwinden kann, muss

$$E_{\text{kin}} \geq eU + (\Phi_{\text{A}} - \Phi_{\text{K}}) \tag{2.3.6}$$

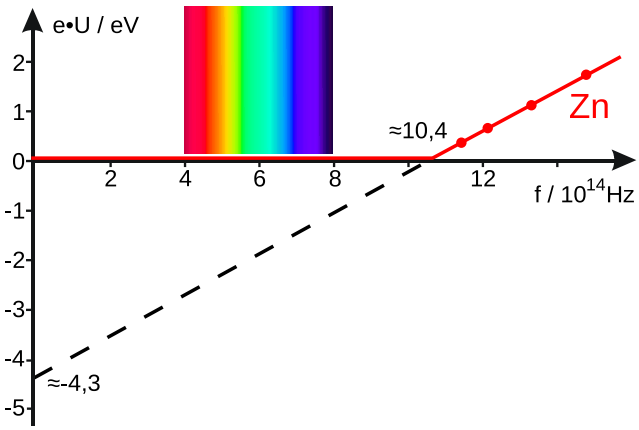
gelten. Gemäß (2.3.4) muss für die außen angelegte Spannung also

$$eU \leq \hbar\omega - \Phi_{\text{A}} \tag{2.3.7}$$

gelten. Damit ist also die Grenzspannung, bei der gerade keine Elektronen mehr zur Anode gelangen, wo also der Photostrom gerade verschwindet, durch

$$eU_{\text{grenz}} = \hbar\omega - \Phi_{\text{A}} \tag{2.3.8}$$

gegeben, d.h. entgegen der Behauptung in vielen Lehrbüchern steht auf der rechten Seite dieser **Einsteinschen Formel für den Photoeffekt** beim Gegenfeldexperiment die **Austrittsarbeit der Auffanganode** und *nicht, wie dort behauptet, die der Photokathode* [RT76]. Das ist umso verwunderlicher als Millikan sich dieser Tatsache wohl bewusst war und dies auch in einem späteren Experiment verifiziert hat [Mil21].



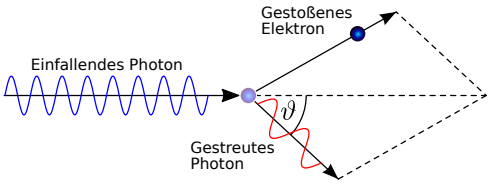
Ansonsten ist freilich allemal die Schlussfolgerung korrekt, dass sich für die Gegenspannung als Funktion der Kreisfrequenz des verwendeten Lichtes eine Gerade mit der Steigung $\hbar/e$, bzw. als Funktion der Frequenz f eine Gerade mit der Steigung h/e, ergibt und aus dem bekannten Wert der Elementarladung e (ebenfalls von Millikan in seinem berühmten Öltröpfchenversuch bereits recht genau gemessen) schließlich $\hbar$ ermittelt werden kann. Es ergab sich dabei ein Wert, der mit dem aus dem Planckschen Strahlungsgesetz ermittelten Wert im Rahmen der Messgenauigkeit übereinstimmte.[2]

Bzgl. der modernen Theorie des Photoeffekts ist zu sagen, dass die obige Argumentation für eine „Teilchennatur“ von elektromagnetischer Strahlung nicht mehr zutreffend ist, denn auf dem hier erörterten experimentellen Niveau lässt sich die Phänomenologie des Photoeffekts auch durch die **semiklassische Näherung** erklären. Darunter versteht man die Näherung, dass das elektromagnetische Feld im Rahmen der klassischen Maxwell-Theorie beschrieben wird, die im Metall gebundenen Elektronen jedoch im Rahmen der modernen Quantentheorie, dabei kann man die Wechselwirkung zwischen dem klassischen elektromagnetischen Wellenfeld und den Elektronen in der 1. Ordnung der zeitabhängigen Störungstheorie behandeln (s. z.B. [LL92]).

2.4 1923: Compton und die Streuung von Röntgenstrahlung

Die Lichtquantenhypothese stieß, wie zu erwarten, zunächst auf große Vorbehalte. Trotz der sehr erfolgreichen experimentellen Bestätigung der Einsteinschen Erklärung des Photoeffekts durch Millikan, war doch klar, dass es sich dabei in der Tat um einen sehr „revolutionären Akt“ handelt, und Einstein bezeichnete die Idee ja selbst lediglich als „heuristischen Gesichtspunkt“. Das änderte sich erst ab 1923 als **Compton** die Streuung von **Röntgenstrahlung** an (quasi) freien Elektronen gemessen und im Rahmen der Einsteinschen Quantentheorie der Strahlung analysiert hat [Com23].

[2] Abbildung: Wikipedia, Klaus-Dieter Keller, CC BY 3.0, https://commons.wikimedia.org/w/index.php?curid=24751457



Compton streute dabei Röntgen-Strahlung an Graphit. Dabei kann man die Auslösearbeit für die Elektronen gegenüber der einstrahlenden Energie der Lichtquanten vernachlässigen. Im Gegensatz zum Photoeffekt, wo das Strahlungsquant vollständig absorbiert wird, handelt es sich bei der Compton-Streuung um eine **elastische Streuung**, d.h. im naiven Photonenbild Einsteins handelt es sich um die elastische Streuung eines Lichtquants und eines Elektrons, also eine Reaktion der Art $\gamma + e^- \to \gamma + e^-$. Dabei gilt Energie- und Impulserhaltung. Dabei müssen wir allerdings wegen der Energieverhältnisse die kinetische Energie und den Impuls des Elektrons *relativistisch* behandeln. Dies besprechen wir genau im zweiten Teil der Vorlesung. Wie in Abschnitt 4.4 ausführlich besprochen, gilt zwischen Energie und Impuls eines massiven Teilchens (hier dem Elektron) die Beziehung

$$E = c\sqrt{m^2c^2 + \vec{p}^2}. \tag{2.4.1}$$

Diese Beziehung gilt auch für Photonen, wobei deren Masse $m_\gamma = 0$ ist. Wir können nun das Elektron im Graphit als ruhend ansehen, d.h. im Stoß ist $E_1 = mc^2$ und $\vec{p}_1 = 0$. Das Lichtquant besitzt die Energie $E_2 = \hbar\omega_2$ und einen Impuls $\vec{p}_2 = \hbar\vec{k}_2$. Wir betrachten nun einen Streuprozess, bei dem das Photon um einen Streuwinkel ϑ abgelenkt wird. Im Folgenden wollen wir nun die Energie (bzw. Frequenz) E_2' dieses gestreuten Photons in Abhängigkeit vom Streuwinkel bestimmen. Der Impuls des gestreuten Photons sei $\vec{p}_2'$. Es gilt wieder $E_2' = c|\vec{p}_2'|$. Es ist klar, dass das Elektron nach der Streuung ebenfalls einen Impuls besitzen muss, den wir mit $\vec{p}_1'$ bezeichnen für die Energie des gestreuten Elektrons gilt wieder (2.4.1). Aus der Impulserhaltung folgt also

$$\vec{p}_2 = \vec{p}_1' + \vec{p}_2' \Rightarrow \vec{p}_2 - \vec{p}_2' = \vec{p}_1'. \tag{2.4.2}$$

Quadrieren liefert

$$(\vec{p}_2 - \vec{p}_2')^2 = p_2^2 + p_2'^2 - 2p_2p_2'\cos\vartheta = p_1'^2. \tag{2.4.3}$$

Die Energieerhaltung liefert

$$E_1 + E_2 = E_1' + E_2' \Rightarrow mc^2 + cp_2 = c\sqrt{m^2c^2 + \vec{p}_1'^2} + cp_2. \tag{2.4.4}$$

Mit Hilfe dieser Gleichung lässt sich mit einer einfachen aber etwas umfangreichen Rechnung p_1' in (2.4.3) eliminieren. Es ergibt sich daraus die Beziehung zwischen den Beträgen des Photonenimpulses vor und nach dem Stoß in Abhängigkeit vom Streuwinkel ϑ

$$p_2p_2'(1 - \cos\vartheta) = mc(p_2 - p_2'). \tag{2.4.5}$$

Diese Beziehung ist einfacher zu formulieren, wenn wir die Photonenbeträge durch die entsprechende Wellenlänge ausdrücken:

$$p_2 = \hbar k_2 = \frac{2\pi\hbar}{\lambda_2}, \quad p_2' = \hbar k_2' = \frac{2\pi\hbar}{\lambda_2'}. \tag{2.4.6}$$

Setzt man dies in (2.4.5) ein, erhält man wieder nach einfachen Umformungen

$$\Delta\lambda = \lambda_2' - \lambda_2 = \lambda_C(1 - \cos\vartheta). \tag{2.4.7}$$

Dabei ist $\lambda_C = 2\pi\hbar/(mc)$ die sog. **Compton-Wellenlänge** des Elektrons. Es ist $\lambda_C \simeq 2{,}426 \cdot 10^{-12}$m. Dies erklärt auch, warum man den Compton-Effekt nur im Bereich kurzwelliger Röntgen- bzw. γ-Strahlung beobachten kann. Bei sichtbarem Licht ist nämlich z.B. $\lambda_{\text{sichtbar}} \simeq 600$ nm $= 6 \cdot 10^{-7}$ m $\gg \lambda_C$ und damit die Änderung der Wellenlänge gemäß (2.4.7) kaum bemerkbar.

Es ist historisch noch interessant zu bemerken, dass bei der weiteren Suche nach einer in sich schlüssigen Quantentheorie zwischenzeitlich Zweifel an der Energie- und Impulserhaltung im Einzelprozess aufkamen

und **Bohr, Kramers und Slater** eine Theorie entwickelt haben, die die Gültigkeit des Energie- und Impulserhaltungssatzes nur im statistischen Mittel annahm.. Diese Frage wurde durch **Bothe und Geiger** in 1924 geklärt. Durch die Erfindung der **Koinzidenzzählmethode**, wo man in jedem Compton-Streuprozess mit einem einzelnen Röntgenphoton sowohl das gestreute Röntgenphoton als auch das gestreute Elektron messen konnte, also „in Koinzidenz".

Es sei auch hier noch abschließend bemerkt, dass der Compton-Effekt in niedrigster Ordnung der Störungstheorie der quantenmechanischen Streutheorie, wie auch der Photoeffekt, mit der semiklassischen Theorie erklärt werden kann, d.h. unter Beibehaltung der klassischen Maxwell-Elektrodynamik für das elektromagnetische Feld aber unter Verwendung der modernen Quantentheorie für die Beschreibung des Elektrons.

Streng genommen benötigt man für den experimentellen Nachweis der Notwendigkeit, dass auch das elektromagnetische Feld als Quantensystem behandelt werden muss, andere Sachverhalte. Der einfachste ist die **spontane Emission**. Betrachtet man nämlich in der semiklassischen modernen Quantentheorie des Elektrons z.B. Atome, ergeben sich diskrete Energieniveaus für die an den Atomkern eines Atoms gebundenen Elektronen. Daraus erklärt sich das charakteristische Spektrum jedes Atoms: Die Frequenzen der von Atomen ausgesandten bzw. absorbierten Strahlung entspricht dem Übergang der Elektronen im Atom von einem diskreten Energieniveau zu einem anderen. In der semiklassischen Theorie entspricht aber einem angeregten Energiezustand des Atoms ein exakt zeitunabhängiger Zustand, d.h. ohne Anwesenheit äußerer Strahlung kann das Atom nicht unter Aussendung eines entsprechenden Photons spontan in einen niedrigeren Energiezustand übergehen.

Allerdings wird die spontane Emission beobachtet, und tatsächlich stabil ist nur der **Grundzustand** des Atoms, also der Zustand niedrigster Energie. Die moderne Erklärung dafür ist, dass auch das elektromagnetische Feld „quantisiert werden muss". Die entsprechende **Quantenfeldtheorie** liefert dann die spontane Emission. Qualitativ lässt sich das daraus erklären, dass das elektromagnetische Feld auch im Vakuum-Zustand fluktuiert, und durch diese Fluktuationen des elektromagnetischen Feldes erfolgt mit einer bestimmten Wahrscheinlichkeit die spontane Emission eines Photons durch ein angeregtes Atom, wobei dieses in einen niedrigeren Energiezustand übergeht.

2.5 1909-1913: Rutherford-Streuung und das Bohrsche Atommodell

Ein für die Geschichte der Atomphysik entscheidendes Experiment war der **Rutherfordsche Goldfolienversuch**. Dabei verwendete er α-Teilchen (also ^{4}He-Kerne) aus dem Zerfall von Radium in einem Bleibehälter mit einer Öffnung. Die α-Teilchen ließ er an einer Goldfolie streuen. Entgegen der Erwartung ergab sich kaum eine Ablenkung, aber einige wenige α-Teilchen wurden auch sehr stark abgelenkt. Rutherford gelangte dadurch zu dem Schluss, dass die Goldfolie nicht aus einer kontinuierlichen Verteilung positiver Ladung, in die die negativ geladenen Elektronen eingebettet sind, wie man damals meinte (**Thomsonsches „Rosinenkuchenmodell"**) sondern im wesentlichen aus sehr kleinen schweren positiv geladenen **Atomkernen** besteht, um die die Elektronen kreisen. Dies wurde quantitativ dadurch untermauert, dass der **Streuquerschnitt** für die α-Teilchen genau der Streuung im Coulomb-Feld der schweren Kerne entsprach.

Die Vorstellung, dass sich die Elektronen im Coulomb-Feld des Kerns bewegen in mathematisch exakter Analogie zur Planetenbewegung um die Sonne, hat allerdings den Fehler, dass um den Kern kreisende Elektronen zufolge der klassischen Elektrodynamik elektromagnetische Wellen abstrahlen und dadurch sehr schnell ihre Energie verlieren und schließlich in den Kern stürzen müssten. Die klassische Theorie der im Coulomb-Feld kreisenden Elektronen widerspricht also eklatant der beobachteten Stabilität der Materie.

Bohr hatte dann 1913 die Idee, dass man die Atome quantenmechanisch behandeln müsste. Dazu griff er zunächst die Plancksche Idee auf, wonach ein harmonischer Oszillator (im Fall der Hohlraumstrahlung waren die den Oszillatoren entsprechenden Freiheitsgrade die Moden des elektromagnetischen Feldes, wie in Abschnitt 2.2 besprochen) nur die diskreten Energieniveaus $E_n = \hbar\omega n$ mit $n \in \mathbb{N}_0$ annehmen kann. Die Frage war nur, wie man diese Idee auf allgemeinere Bewegungen übertragen könnte.

Bohrs Idee war, den harmonischen Oszillator im Phasenraum der analytischen Mechanik zu betrachten, wie es das Hamiltonsche Prinzip der kleinsten Wirkung nahelegt (vgl. die Vorlesung Theorie 1 [Hee18c]). Die Lagrange-Funktion für den harmonischen Oszillator lautet

$$L = \frac{m}{2}\dot{x}^2 - \frac{m\omega^2}{2}x^2. \tag{2.5.1}$$

Der kanonisch konjugierte Impuls ist

$$p = \frac{\partial L}{\partial \dot{x}} = m\dot{x}. \tag{2.5.2}$$

Eine beliebige Bahn ergibt sich aus der Lösung der Euler-Lagrange-Gleichung. Wir können sie aber auch anders analysieren. Zunächst ist die Energie erhalten, weil L nicht explizit von der Zeit abhängt, so dass die Hamilton-Funktion

$$H = p\dot{x} - L = \frac{m}{2}\dot{x}^2 + \frac{m\omega^2}{2}x^2 = \frac{p^2}{2m} + \frac{m\omega^2}{2}x^2 = E = \text{const} \tag{2.5.3}$$

ist. Teilen wir diese Gleichung durch E, sehen wir, dass die Bahn im Phasenraum eine Ellipse ist:

$$\frac{1}{2mE}p^2 + \frac{m\omega^2}{2E}x^2 = 1. \tag{2.5.4}$$

Die Halbachsen sind demnach $\sqrt{2mE}$ und $\sqrt{2E/(m\omega^2)}$. Betrachtet man dann die Fläche der Ellipse im Phasenraum, also

$$A = \oint \mathrm{d}x\, p = \pi\sqrt{2mE}\sqrt{\frac{2E}{m\omega^2}} = \frac{2\pi E}{\omega}, \tag{2.5.5}$$

wobei $\oint$ dem Integral eines Umlaufs um die Ellipse im Phasenraum entspricht. Entsprechend der Planckschen Quantisierungshypothese, dass der harmonische Oszillator nur Energien $E_n = n\hbar\omega$ annehmen kann, ergibt sich daraus für **periodische Bewegungen** die Quantisierungsbedingung

$$\oint \mathrm{d}x\, p = 2\pi\hbar n, \quad n \in \mathbb{N}. \tag{2.5.6}$$

Nun ist auch die Bewegung eines Elektrons im **Coulomb-Potential** eines Protons eine solche periodische Bewegung, wenn man von dem oben erwähnten Problem der Abstrahlung elektromagnetischer Wellen einfach absieht, und das war Bohrs Hypothese: Das Elektronen darf sich auf bestimmten Bahnen, die durch die Quantisierungsbedingung der periodischen Bewegung (2.5.6) ausgewählt werden, bewegen, ohne elektromagnetische Wellen abzustrahlen.

Im Prinzip könnten wir hier unsere Kenntnisse aus der Theorie-Vorlesung 1 verwenden, um das vollständige Bohr-Sommerfeld-Modell zu behandeln, denn wir kennen ja die Lösung des Keplerproblems. Allerdings ist das für den allgemeinen Fall recht aufwändig, und wir begnügen uns daher hier auf den von Bohr betrachteten Spezialfall von **Kreisbahnen**. Dazu genügt es, die Gleichung für die Zentripetalkraft mit der Coulomb-Kraft gleich zu setzen. Dies ergibt

$$-mr\omega^2 = -\frac{e^2}{4\pi\epsilon_0 r^2}, \tag{2.5.7}$$

wobei m die Masse des Elektrons ist. Auf der rechten Seite haben wir die Ladungen $-e$ für das Elektron und die Ladung $+e$ für das Proton in die Formel für die Coulomb-Kraft eingesetzt. Es folgt für den Bahnradius

$$\omega^2 = \frac{e^2}{4\pi\epsilon_0 m r^3}. \tag{2.5.8}$$

Betrachten wir nun die Lagrangefunktion für dieses Problem, ergibt sich für die Bewegung in der Ebene senkrecht zum erhaltenen Drehimpuls in Polarkoordinaten geschrieben

$$L = \frac{m}{2}(\dot{r}^2 + r^2\dot{\varphi}^2) + \frac{e^2}{4\pi\epsilon_0 r}. \tag{2.5.9}$$

Für unsere Kreisbahnen benötigen wir hier nur eine Quantisierungsbedingung, und zwar die für die Bewegung von φ. Der kanonisch konjugierte Impuls von φ ist nun die Drehimpulskomponente

$$l_z = \frac{\partial L}{\partial \dot{\varphi}} = m r^2 \dot{\varphi} = m r^2 \omega = \text{const.} \tag{2.5.10}$$

Dass l_z konstant ist für die Lösung der Bewegungsgleichung folgt daraus, dass gemäß der entsprechenden Euler-Lagrange-Gleichung für $\dot{l}_z = \partial_\varphi L = 0$ ist. Die Quantisierungsbedingung für das Paar φ und l_z lautet nun

$$\oint \mathrm{d}\varphi l_z = 2\pi l_z = 2\pi n\hbar, \quad n \in \mathbb{N} \Rightarrow l_z = m r^2 \omega = n\hbar. \tag{2.5.11}$$

Es ist also $\omega = n\hbar/(m r^2)$. Setzt man dies in (2.5.8) ein, folgt

$$r = \frac{4\pi\epsilon_0 \hbar^2 n^2}{e^2 m}. \tag{2.5.12}$$

Andererseits ist

$$E_{\text{kin}} = \frac{m}{2} r^2 \dot{\varphi}^2 = \frac{l_z^2}{2m r^2} = \frac{n^2\hbar^2}{2m r^2} = \frac{e^4 m}{2(4\pi\epsilon_0\hbar)^2 n^2} \tag{2.5.13}$$

und

$$E_{\text{pot}} = -\frac{e^2}{4\pi\epsilon_0 r} = -\frac{e^4 m}{(4\pi\epsilon_0\hbar)^2 n^2}. \tag{2.5.14}$$

Damit ergeben sich die **quantisierten Energien** für die Kreisbahnen des Wasserstoffatoms

$$E_n = E_{\text{kin}} + E_{\text{pot}} = -\frac{e^4 m}{2(4\pi\epsilon_0\hbar)^2 n^2}. \tag{2.5.15}$$

Eine genaue Analyse Sommerfelds des allgemeinen Problems beliebiger Ellipsenbahnen ergibt, dass auch die weiteren möglichen Bahnen stets Energien der Form (2.5.15) liefern.

Nach dieser Bohr-Sommerfeld-Quantisierung waren nun diese diskreten Bahnen die Bahnen, auf denen das Elektron nicht strahlt und also der Grundzustand (kleinstmögliche Energie entsprechend $n = 1$) und die angeregten Zustände $n > 1$ eines stabilen Wasserstoffatoms. Übergänge von einem Zustand zum anderen konnte dadurch erfolgen, dass Photonen der entsprechenden Frequenz absorbiert oder emittiert werden, d.h. die für Wasserstoff charakteristischen **Spektrallinien** entspricht den **Energiedifferenzen** der entsprechenden stabilen Elektronenbahnen

$$\hbar\omega_{n_1 n_2} = E_{n_1} - E_{n_2}. \tag{2.5.16}$$

Für $n_1 < n_2$ entspricht dies einer Emission, also einem Übergang von einem Zustand des Atoms größerer zu kleinerer Energie. Die sehr genaue Übereinstimmung dieser Vorhersage mit den beobachteten Spektrallinien war ein großer Erfolg für die so entstandene „alte Quantentheorie“.

Andererseits waren sich die Physiker auch bewusst, dass die Annahme stabiler Bahnen, die durch Annahme einer recht willkürlichen Quantisierungsbedingung zustandekommt, doch recht unbefriedigend ist. Außerdem zeigte sich bald, dass das Bohr-Sommerfeldsche Quantisierungsverfahren nur für das Wasserstoffatom

korrekte Vorhersagen für die von den Atomen abgestrahlten Spektrallinien ergab. Die übrigen Atome (angefangen bereits mit dem nächstkomplizierteren Helium mit zwei Elektronen) konnten nur durch weitere ad-hoc-Annahmen mühsam verstanden werden. Es war also klar, dass es einer konsistenten „Quantentheorie" bedurfte, um auch die komplizierteren Atome quantitativ korrekt beschreiben zu können.

Vom didaktischen Standpunkt muss man sagen, dass eine ausführliche Behandlung des Bohr-Sommerfeld-Atommodells eher kontraproduktiv ist, denn es zementiert qualitativ falsche Vorstellungen über die Quantenmechanik. Insbesondere die Idee, dass Elektronen sich auf Bahnen um den Atomkern bewegen. Die in sich konsistente Lösung des damit zusammenhängenden Problems mit der Stabilität der Atome ist vielmehr im Rahmen der modernen Quantentheorie, dass es sich um **statische Lösungen** der Schrödinger-Gleichung handelt, wie wir im nächsten Kapitel ausführlich besprechen werden. In diesem Bild handelt es sich bei einem Atom in einem solchen stationären Zustand um ein elektrostatisches Gebilde, das demzufolge auch nicht im Sinne klassischer Synchrotronstrahlung elektromagnetische Wellen abstrahlt und demzufolge stabil sein kann. Dabei besitzt das Elektron keinen definierten Ort und auch keinen definierten Impuls. Vielmehr gibt Schrödingers **Wellenfunktion** (genauer gesagt deren Betragsquadrat) lediglich die Wahrscheinlichkeitsdichte für das Auffinden des Elektrons an jedem Ort an bzw. die Wahrscheinlichkeitsdichte seines Impulses, wenn man *entweder* den Ort *oder* den Impuls des Elektrons genau misst.

2.6 1923-1926: de Broglie und der Welle-Teilchen-Dualismus

Im Jahre 1923 verfasste **de Broglie** seine Dissertation über die Idee, dass Teilchen auch Welleneigenschaften aufweisen sollten. Die Idee dabei war, dass ja im Sinne der Einsteinschen Vorstellungen zu Lichtquanten das Licht im oben erörterten Sinne Teilcheneigenschaften aufweist, d.h. dass der **Welle-Teilchen-Dualismus** eine allgemeine Eigenschaft sowohl von „Teilchen" wie Elektronen als auch der elektromagnetischen „Wellen" ist.

Dabei ging er davon aus, dass die Beziehung zwischen Wellen- und Teilcheneigenschaften in Analogie zu den Einsteinschen Beziehungen beim Licht (bzw. allgemein elektromagnetischen Wellen) durch

$$E = \hbar\omega, \quad \vec{p} = \hbar\vec{k} \tag{2.6.1}$$

gegeben ist. Diese Beziehungen werden seither als **Einstein-de Broglie-Beziehungen** bezeichnet. Interessanterweise waren die Prüfer an der Sorbonne, Perrin und Langevin, ein wenig unsicher, wie sie diese Dissertation bewerten sollten und schickten sie daher an Einstein als einem der führenden Experten der Quantentheorie. Einstein zeigte sich von der Arbeit begeistert, und de Broglie wurde daraufhin promoviert.

Solche Welleneigenschaften von Teilchen sollten nun wie beim Licht durch **Interferenzerscheinungen** nachweisbar sein. Typische Wellenlängen ergaben sich für Elektronen und Protonen typischerweise in dem Bereich von Röntgenstrahlen beim Licht. Entsprechend sollten dieselben Methoden, die zum Nachweis der Wellennatur der Röntgenstrahlung, also die **Streuung an Kristallen** auch die Wellennatur der Teilchen nachweisen lassen. Dies gelang bereits 1921-1923 **Davisson und Germer** durch Streuung an dünnen Nickelfolien. In neuerer Zeit gibt es auch schöne Experimente in kompletter Analogie zum **Doppelspaltexperiment** mit Licht [TEM+89]. Wir werden auf Beugungserscheinungen für Teilchen weiter unten noch genauer zurückkommen, wenn wir die moderne Quantentheorie in ihrer Schrödingerschen Form als **Wellenmechanik** eingeführt haben.

2.7 1926: Schrödinger und die Wellenmechanik

In diesem den historischen Überblick abschließenden Abschnitt wollen wir uns mit einer der drei fast zeitgleich entstandenen Formulierungen der Quantentheorie, der durch **Schrödinger** 1926 begründeten **Wellenmechanik**. Die Idee geht auf einen Besuch Schrödingers Ende 1925 im Physikseminar der ETH Zürich zurück, wo er über de Broglies Ideen einen Vortrag hielt und Debye dazu nur bemerkte, dass er nur dann

vernünftig über Wellen nachdenken könne, wenn er eine **Wellengleichung** zur Verfügung hätte. Das nahm dann Schrödinger zum Anlass, seine nach ihm benannte Gleichung aufzustellen. Es ist klar, dass man eine fundamental neue Gleichung nicht irgendwie mathematisch herleiten sondern nur geschickt raten kann. Wir beginnen zunächst mit **freien Teilchen**. Es ist interessant, dass Schrödinger zunächst eine relativistische Wellengleichung aufgestellt hat, deren Lösung für das Wasserstoffatom allerdings die falsche Vorhersage für die Feinstruktur der Spektrallinien lieferte. Diese heute als **Klein-Gordon-Gleichung** bekannte Gleichung berücksichtigt allerdings nicht den Spin 1/2 des Elektrons, und die richtige Feinstruktur ergibt sich erst mit der entsprechenden **Dirac-Gleichung**, also der relativistischen Wellengleichung für Teilchen mit Spin 1/2. Die Klein-Gordon-Gleichung beschreibt allerdings Teilchen mit Spin 0 (z.B. Pionen) richtig. Allerdings ist die physikalische Interpretation relativistischen Wellengleichungen um noch einiges verzwickter als für nichtrelativistische. Daher beschränken wir uns in dieser Vorlesung auf die **nichtrelativistische Schrödinger-Gleichung** und die entsprechend erweiterte nichtrelativistische Gleichung für Teilchen mit Spin 1/2, die **Pauli-Gleichung**, worauf wir allerdings erst später genauer eingehen können.

2.7.1 Die Schrödinger-Gleichung für ein freies Teilchen

Die Idee Schrödingers bei der Aufstellung der nichtrelativistischen Wellengleichung für Teilchen verwendet zum Einen die Einstein-de Broglie-Beziehung (2.6.1) zwischen Teilchen- und Welleneigenschaften von „Quanten“ und zum Anderen die daraus folgende **Dispersionsrelation** der Wellen. Dazu verwendete Schrödinger die Beziehung zwischen der kinetischen Energie und dem Impuls nichtrelativistischer Teilchen der Masse m, also

$$E = \frac{1}{2m}\vec{p}^2 \Rightarrow \hbar\omega = \frac{1}{2m}\hbar^2\vec{k}^2. \tag{2.7.1}$$

Wie bei den Elektromagnetischen Wellen muss diese Dispersionsrelation zwischen Frequenz und Wellenzahl aus der entsprechenden Lösung der Wellengleichung für eine **ebene Welle**

$$\psi(t,\vec{x}) = A\exp(-\mathrm{i}\omega t + \mathrm{i}\vec{k}\cdot\vec{x}) \tag{2.7.2}$$

ergeben. Offenbar gilt für diese ebene Welle

$$\partial_t\psi(t,\vec{x}) = -\mathrm{i}\omega\psi(t,\vec{x}), \quad \vec{\nabla}\psi(t,\vec{x}) = \mathrm{i}\vec{k}\psi(t,\vec{x}). \tag{2.7.3}$$

Wendet man auf die letztere Gleichung nochmals den Nabla-Operator an, finden wir

$$\Delta\psi(t,\vec{x}) = -\vec{k}^2\psi(t,\vec{x}). \tag{2.7.4}$$

Multiplizieren wir diese Gleichung mit $-\hbar^2/(2m)$ und verwenden die Dispersionsrelation (2.7.1), folgt daraus

$$-\frac{\hbar^2}{2m}\Delta\psi(t,\vec{x}) = \frac{\hbar^2}{2m}\vec{k}^2\psi(t,\vec{x}) \stackrel{!}{=} \hbar\omega\psi(t,\vec{x}) = \mathrm{i}\hbar\partial_t\psi(t,\vec{x}). \tag{2.7.5}$$

Dabei haben wir im letzten Schritt die 1. Gleichung in (2.7.3) verwendet, und dies ist bereits die **Schrödinger-Gleichung** für ein freies Teilchen

$$\mathrm{i}\hbar\partial_t\psi(t,\vec{x}) = -\frac{\hbar^2}{2m}\Delta\psi(t,\vec{x}). \tag{2.7.6}$$

Das Hauptproblem bestand nun für Schrödinger darin, wie er diese Gleichung *physikalisch interpretieren* sollte. Zunächst ging er davon aus, dass (wieder in Analogie zu den elektromagnetischen Feldern) $|\psi(t,\vec{x})|^2$ die Intensität der Wellen ergibt. Dabei ist klar, dass man das *Betrags*quadrat schreiben muss, weil offenbar ψ i.a. eine komplexwertige Funktion sein muss, denn in (2.7.6) tritt ja die imaginäre Einheit i explizit auf. Für die ebene Welle (2.7.2) ergibt sich aber einfach $|\psi(t,\vec{x})|^2 = |A|^2 = \text{const}$. Wollte man nun einfach Teilchen als

Felderscheinung umdeuten, wie Schrödinger anfangs gedacht hat, würde diese einfachste Lösung bedeuten, dass dieses Teilchen über alle Orte gleichmäßig verschmiert sein müsste.

Schrödingers Idee war also, nach räumlich lokalisierten Lösungen seiner Gleichung (2.7.6) zu suchen. Das ist wegen der Linearität der Schrödinger-Gleichung auch nicht schwierig, denn wir können auf die aus der Elektrodynamik bewährte Methode der **Fourier-Darstellung** zurückgreifen, d.h. wir machen den Ansatz

$$\psi(t,\vec{x}) = \int_{\mathbb{R}^3} \frac{\mathrm{d}^3 p}{(2\pi\hbar)^{3/2}} \tilde{\psi}(t,\vec{p}) \exp\left(\frac{\mathrm{i}\vec{p}\cdot\vec{x}}{\hbar}\right). \tag{2.7.7}$$

Dabei haben wir statt des Wellenvektors $\vec{k}$ den Impuls $\vec{p} = \hbar\vec{k}$ der ebenen Schrödingerwelle verwendet. Der Normierungsfaktor $1/(2\pi\hbar)^{3/2}$ erweist sich gleich noch als bequem.

Setzen wir nun diesen Ansatz in (2.7.6) ein, folgt für $\tilde{\psi}(t,\vec{p})$ die Gleichung (*Nachrechnen!*)

$$\mathrm{i}\hbar\partial_t \tilde{\psi}(t,\vec{p}) = \frac{1}{2m}\vec{p}^{\,2}\tilde{\psi}(t,\vec{p}) = E(\vec{p})\tilde{\psi}(t,\vec{p}). \tag{2.7.8}$$

Diese lineare Differentialgleichung 1. Ordnung (!) lässt sich sehr leicht durch „Trennung der Variablen“ lösen:

$$\tilde{\psi}(t,\vec{p}) = A(\vec{p}) \exp\left(-\frac{\mathrm{i}E(\vec{p})t}{\hbar}\right). \tag{2.7.9}$$

Dabei ist $A(\vec{p})$ eine beliebige Funktion. Damit nimmt unsere allgemeine Lösung der Schrödinger-Gleichung die folgende Form an

$$\psi(t,\vec{x}) = \int_{\mathbb{R}^3} \frac{\mathrm{d}^3 p}{(2\pi\hbar)^{3/2}} A(\vec{p}) \exp\left(-\frac{\mathrm{i}E(\vec{p})t}{\hbar} + \frac{\mathrm{i}\vec{p}\cdot\vec{x}}{\hbar}\right). \tag{2.7.10}$$

Betrachten wir nun zunächst ein Wellenpaket für den Fall, dass $A(\vec{p})$ nur in einem kleinen Bereich um $\vec{p} = \vec{p}_0$ von 0 verschieden ist. Verwenden wir der Einfachheit halber eine „Kastenfunktion“

$$A(\vec{p}) = \begin{cases} A_0 & \text{für} \quad Q_p \\ 0 & \text{sonst,} \end{cases} \tag{2.7.11}$$

wobei $Q_p = [p_{01} - \Delta p/2, p_{01} + \Delta p/2, p_{03} - \Delta p/2, p_{03} + \Delta p/2, p_{03} - \Delta p/2, p_{03} + \Delta p/2]$ ein Würfel der Kantenlänge Δp im Impulsraum mit Mittelpunkt $\vec{p}_0$ ist. Für $\Delta p \ll |\vec{p}_0|$ können wir das Argument in der Exponentialfunktion in (2.7.10) um $\vec{p} = \vec{p}_0$ entwickeln, wobei wir nur Terme bis zur linearen Ordnung mitnehmen müssen. Daraus ergibt sich

$$\begin{aligned} \psi(t,\vec{x}) &\simeq A_0 \exp\left(-\frac{\mathrm{i}tE(\vec{p}_0)}{\hbar} + \frac{\mathrm{i}\vec{p}_0\cdot\vec{x}}{\hbar}\right) \int_{Q_p} \frac{\mathrm{d}^3 p}{(2\pi\hbar)^{3/2}} \exp\left(\frac{\mathrm{i}(\vec{x} - \vec{p}_0 t/m)\cdot(\vec{p} - \vec{p}_0)}{\hbar}\right) \\ &= A_0 \exp\left(-\frac{\mathrm{i}tE(\vec{p}_0)}{\hbar} + \frac{\mathrm{i}\vec{p}_0\cdot\vec{x}}{\hbar}\right)\left(\frac{2\hbar}{\pi}\right)^{3/2} \prod_{j=1}^{3} \frac{\sin[\Delta p(x_j - p_{0j}t/m)/\hbar]}{(x_j - p_{0j}t/m)}. \end{aligned} \tag{2.7.12}$$

Wie wir sehen, handelt es sich um ein Wellenpaket mit einem Maximum bei $\vec{x} = \vec{p}_0 t/m$. Das erste Minimum liegt jeweils bei $|x_j - p_{0j}t/m| = \hbar\pi/\Delta p$, d.h. die **Breite** des Wellenpakets im Ortsraum um das Maximum wird klein, wenn Δp relativ groß ist. Dann ist das Teilchen offenbar im Ortsraum recht gut lokalisiert, und der **Schwerpunkt** des Wellenpakets bewegt sich mit der Geschwindigkeit $\vec{v}_0 = \vec{\nabla}_p E(\vec{p})|_{\vec{p}=\vec{p}_0} = \vec{p}_0/m$. Dieses Wellenpaket entspricht also einem auf einem kleinen Raumgebiet ausgedehntes Teilchen.

Wie wir anhand des Beispiels eines sog. **Gaußschen Wellenpakets**, für das man das Fourier-Integral (2.7.10) exakt ausrechnen kann in einer Übung zeigen werden, wird aber ein anfangs gut lokalisiertes Wellenpaket

mit der Zeit recht schnell breiter, d.h. wartet man nur lange genug, müsste man bei der Messung des Ortes eines solchen ausgedehnten Teilchens ein auf immer größere Bereiche „verschmiertes“ Objekt vorfinden, auch wenn es anfangs in einem recht engen Bereich lokalisiert war. Dies entspricht aber keineswegs den Beobachtungen. Vielmehr wird man ein Elektron, wenn man es z.B. mittels einer CCD-Kamera detektiert immer als ein punktförmiges Teilchen an einem bestimmten Ort finden, wobei die „Ausdehnung des Teilchens“ durch die Auflösung des Detektors (bei der CCD-Kamera in der Größenordnung der Pixel-Größe) bestimmt wird. Man kann also die Wellenfunktion *nicht* als eine Art kontinuumsmechanischer Beschreibung quasipunktförmiger Teilchen deuten. Hier manifestiert sich die Problematik des **Welle-Teilchen-Dualismus'** der Quantentheorie. Die naive Interpretation der Wellenfunktion als eine Art kontinuumsmechanische Beschreibung für das Teilchen, z.B. die Annahme Schrödingers, dass $|\psi(t,\vec{x})|^2 \propto \rho(t,\vec{x})$, wobei ρ als eine klassische Größe wie die Ladungs- oder Massenverteilung eines endlich ausgedehnten Körpers (oder eines Fluidelements, wenn man sich das „Teilchen“ sich eher als eine Art Flüssigkeit vorstellt, was eher zu dem Ergebnis einer sich verbreiternden Materieverteilung passt) interpretiert, führt zu Widersprüchen mit den experimentellen Fakten.

Denkt man sich also das Teilchen doch wieder als diskretes Punktteilchen und betrachtet wieder ein typisches Wellenexperiment wie ein Doppelspaltexperiment, wird der Widerspruch noch deutlicher: Einerseits ergibt das Wellenbild hier erwartungsgemäß ein Interferenzmuster des Doppelspalts. Deckt man jeweils einen Spalt zu erhält man allenfalls ein Beugungsmuster des Einzelspalts. Für klassische Teilchenverteilungen würde man dann erwarten, dass sich bei beiden geöffneten Spalten einfach die beiden Einzelspaltergebnisse addieren. Das ist aber nicht der Fall. Vielmehr sieht man klar die Interferenzmuster, wie man sie vom Doppelspalt erwartet, wenn man viele Teilchen durch die Anordnung schickt. Andererseits liefert jedes Teilchen am Beobachtungsschirm (z.B. kann man wieder eine CCD-Kamera verwenden, die die Auftrefforte der einzelnen Teilchen speichert) einen einzelnen Punkt. Allerdings lässt sich für ein bestimmtes Teilchen nicht vorhersagen, wo es auftreffen wird, bevor man es durch die Spalten schickt. Nach einiger Zeit bildet sich aber bei Wiederholung des Experiments mit hinreichend vielen Teilchen das von der Wellenmechanik erwartete Interferenzmuster aus. Insbesondere kommen in den Beugungsminima so gut wie gar keine Teilchen am Beobachtungsschirm an.

Eine ähnliche Beobachtung macht man auch bei Streuexperimenten mit Teilchen: Auch hier ist nicht vorherzusehen, wohin ein einzelnes Teilchen gestreut werden wird. Führt man jedoch das Streuexperiment hinreichend oft unter stets gleichen Anfangsbedingungen durch, ergibt sich eine dem differentiellen Streuquerschnitt entsprechende Verteilung. Auch für solche Streuexperimente liefert die Wellentheorie einen im Prinzip einfachen Ansatz: Man berechnet die Wellenfunktion für ein anfangs mit einem ziemlich gut bestimmten Impuls auf ein anderes ruhendes Teilchen zufliegendes Teilchen und löst die Schrödinger-Gleichung unter Berücksichtigung der Wechselwirkung zwischen den beiden Teilchen, z.B. kann man beim Rutherfordschen Gold-Folienversuch einfach annehmen, dass jedes α-Teilchen an einem festgehaltenen Goldatomkern vorbeifliegt, wobei dann der Goldatomkern einfach durch ein Coulomb-Potential für die elektromagnetische Abstoßung zwischen Goldatomkern und α-Teilchen modelliert werden kann.

Bei der Lösung dieser Art **Streuprobleme** hat **Born** in 1926 die bis heute von der größten Mehrheit der Physiker akzeptierte Interpretation gefunden, wonach $|\psi(t,\vec{x})|^2$ proportional zur **Aufenthaltswahrscheinlichkeitsdichte** des Teilchens ist. Normiert man also die Gesamtwahrscheinlichkeit, wie in der Wahrscheinlichkeitsrechnung üblich auf 1, d.h. wählt man die Normierung der Wellenfunktion so, dass

$$\int_{\mathbb{R}^3} \mathrm{d}^3x |\psi(t,\vec{x})|^2 = 1 \tag{2.7.13}$$

ist, so ist $\mathrm{d}^3x|\psi(t,\vec{x})|^2$ die Wahrscheinlichkeit dafür, das durch die Wellenfunktion $\psi(t,\vec{x})$ beschriebene Teilchen zur Zeit t in einem kleinen Volumen d^3x in der Umgebung des Ortes $\vec{x}$ zu finden. Dies ist die **Bornsche Regel**, und wir folgen in dieser Vorlesung dieser **Standardinterpretation** der Bedeutung der Wellenfunktion.

Experimentell lassen sich die Wahrscheinlichkeitsverteilungen bestimmen, wenn man das Experiment mit jeweils einem Teilchen unter immer denselben Bedingungen sehr oft wiederholt, d.h. man **präpariert** das

Teilchen zu einem Anfangszeitpunkt $t = 0$ so, wie es der Wellenfunktion $\psi(t = 0, \vec{x})$ entspricht. Dann wird die Wellenfunktion zu jedem späteren Zeitpunkt durch die Lösung der Schrödinger-Gleichung eindeutig bestimmt (wobei wir im Folgenden noch klären müssen, wie die Schrödinger-Gleichung für den Fall aussieht, dass das Teilchen nicht frei ist sondern sich z.B. in einem elektrischen Feld bewegt). Um die Vorhersage der entsprechenden Aufenthaltswahrscheinlichkeitsdichte zu überprüfen müssen wir mit einem Detektor die Orte jedes der gleichartig präparierten Teilchens messen. Für sehr viele Wiederholungen dieses Experiments ergeben sich dann die Wahrscheinlichkeiten zur Zeit t als die relativen Häufigkeiten für jedes kleine Volumenelement d^3x um jeden Punkt $\vec{x}$. In dieser sog. **minimalen statistischen Interpretation** der Bedeutung der Wellenfunktion bezieht sich also die Beschreibung des Teilchens durch die Wellenfunktion auf die nach der Bornschen Regel bestimmten Wahrscheinlichkeitsverteilungen bzw. im Experiment relativen Häufigkeiten von Ensembles stets gleich präparierter voneinander unabhängiger Teilchen. Daher nennt man diese Interpretation der Quantenmechanik auch die **Ensemble-Interpretation**. Sie entspricht auch der einfachsten Art der sog. **Kopenhagener Deutungen**, die so benannt wurde, weil sie in den 20er-30er Jahren von den Physikern am Niels-Bohr-Institut in Kopenhagen entwickelt worden ist, wobei auffallenderweise verschiedene Physiker leicht unterschiedliche Vorstellungen über die Bedeutung der Wellenfunktion hatten und auch heute noch haben.

Es besteht nämlich von einem prinzipiell philosophischen Standpunkt aus betrachtet ein bestimmtes Problem darin, dass eben die Physik im Rahmen dieser statistischen Minimalinterpretation keine genauen Vorhersagen über Observablen an einem einzelnen Teilchen machen kann sondern nur Wahrscheinlichkeiten dafür angeben kann, dass bei der Messung einer Observablen ein bestimmter Messwert resultiert. Andererseits ergeben freilich nur Messungen einen Sinn, die für jede Beobachtung einen möglichst eindeutigen Messwert ergeben, wobei die Genauigkeit der Messung lediglich durch die technischen Eigenschaften des Messinstruments bedingt sein kann. Es werden also bis heute noch alternative Deutungen des im Folgenden entwickelten mathematischen Formalismus' diskutiert.

Aus physikalischer Sicht ist die minimale statistische Interpretation allerdings die einfachst mögliche Deutung. Außerdem gibt es keine inneren Widersprüche mit dem mathematischen Formalismus, und die statistischen Vorhersagen der Quantentheorie haben sich bei der experimentellen Überprüfung stets als erfolgreich erwiesen.

Im Rest dieses Abschnittes entwickeln wir nun die Theorie so weit, dass wir sie dann im nächsten Kapitel in Form klarer Postulate endgültig formulieren können. Wir klären dazu unter Verwendung der eben entwickelten Wahrscheinlichkeitsinterpretation der Wellenmechanik folgende Fragen:

- Wie werden die **Observablen** beschrieben?
- Wie ergeben sich die Wahrscheinlichkeiten für Messergebnisse einer beliebigen Observablen (wie bei einem Teilchen Ort, Impuls, Drehimpuls) bei vorgegebener Wellenfunktion, die den **Zustand** des Teilchens festlegt?
- Wie berechnet man die **Zeitentwicklung** der Wellenfunktion, wenn wir es nicht mit einem freien Teilchen zu tun haben sondern mit der Bewegung in elektromagnetischen Feldern (im einfachsten Fall einem elektrostatischen Feld)?

Analysieren wir dazu zunächst noch ein wenig den Fall freier Teilchen, wobei die Zeitentwicklung durch die diesem Fall entsprechende Schrödinger-Gleichung (2.7.6) definiert ist. Wir setzen im Folgenden voraus, dass die Wellenfunktion die Normierungsbedingung (2.7.11) erfüllt.

Als erstes weisen wir nach, dass die Schrödinger-Gleichung **normerhaltend** ist, d.h. dass wenn $\psi(t = 0, \vec{x})$ gemäß (2.7.11) auf 1 normiert ist, bleibt sie automatisch unter der durch (2.7.6) definierten Zeitentwicklung auf 1 normiert. Betrachten wir also die Zeitableitung des Normierungsintegrals:

$$\frac{\mathrm{d}}{\mathrm{d}t}\int_{\mathbb{R}^3}\mathrm{d}^3x|\psi(t,\vec{x})|^2 = \frac{\mathrm{d}}{\mathrm{d}t}\int_{\mathbb{R}^3}\mathrm{d}^3x\,\psi^*(t,\vec{x})\psi(t,\vec{x}) = \int_{\mathbb{R}^3}\mathrm{d}^3x\,[\psi^*(t,\vec{x})\partial_t\psi(t,\vec{x}) + \psi(t,\vec{x})\partial_t\psi^*(t,\vec{x})]. \quad (2.7.14)$$

Dabei haben wir im letzten Schritt Ableitung in Integration vertauscht und die Produktregel verwendet. Die Zeitableitungen können wir nun durch die Schrödinger-Gleichung (2.7.7) und die konjugiert komplexe Gleichung

$$-\mathrm{i}\hbar\partial_t\psi^*(t,\vec{x}) = -\frac{\hbar^2}{2m}\Delta\psi^*(x) \tag{2.7.15}$$

ausdrücken. Dies liefert nach einfachen Umformungen

$$\frac{\mathrm{d}}{\mathrm{d}t}\int_{\mathbb{R}^3}\mathrm{d}^3x|\psi(t,\vec{x})|^2 = \frac{\mathrm{i}\hbar}{2m}\int_{\mathbb{R}^3}\mathrm{d}^3x\,[\psi^*(t,\vec{x})\Delta\psi(t,\vec{x}) - \psi(t,\vec{x})\Delta\psi^*(t,\vec{x})]. \tag{2.7.16}$$

Mit der Produktregel kann man nun sofort nachrechnen, dass der Integrand eine totale Divergenz ist, denn es gilt (*Nachrechnen!*)

$$\psi^*(t,\vec{x})\Delta\psi(t,\vec{x}) - \psi(t,\vec{x})\Delta\psi^*(t,\vec{x}) = \vec{\nabla}\cdot\Big[\psi^*(t,\vec{x})\vec{\nabla}\psi(t,\vec{x}) - \psi(t,\vec{x})\vec{\nabla}\psi^*(t,\vec{x})\Big]. \tag{2.7.17}$$

Demnach ist aber nach dem Gaußschen Integralsatz das Integral auf der rechten Seite (2.7.16) ein Integral über den im Unendlichen Gelegenen Rand des Integrationsbereichs $\mathbb{R}^3$. Nehmen wir nun an, dass die auftretenden Integranden hinreichend schnell im Unendlichen gegen Null streben, verschwindet also die rechte Seite. Damit ist

$$\frac{\mathrm{d}}{\mathrm{d}t}\int_{\mathbb{R}^3}\mathrm{d}^3x|\psi(t,\vec{x})|^2 = 0. \tag{2.7.18}$$

Wir können dieses Ergebnis auch noch anders verstehen. Definieren wir nämlich die Aufenthaltswahrscheinlichkeitsdichte gemäß der Bornschen Regel zu

$$P(t,\vec{x}) = |\psi(t,\vec{x})|^2 = \psi^*(t,\vec{x})\psi(t,\vec{x}), \tag{2.7.19}$$

folgt mit genau den obigen Rechenschritten

$$\partial_t P(t,\vec{x}) = -\vec{\nabla}\vec{j}(t,\vec{x}), \tag{2.7.20}$$

mit der **Wahrscheinlichkeitsstromdichte**

$$\vec{j}(t,\vec{x}) = \frac{1}{2m\mathrm{i}}\Big[\psi^*(t,\vec{x})\vec{\nabla}\psi(t,\vec{x}) - \psi(t,\vec{x})\vec{\nabla}\psi^*(t,\vec{x})\Big]. \tag{2.7.21}$$

Nun ist aber (2.7.20) eine Kontinuitätsgleichung, also die lokale Form eines Erhaltungssatzes, d.h. dass die „Gesamtladung“ $Q = \int_{\mathbb{R}^3}\mathrm{d}^3xP(t,\vec{x}) = \text{const}$ ist. Ist also die Wellenfunktion zur Zeit $t = 0$ gemäß 2.7.13 auf 1 normiert, bleibt sie dies auch für alle späteren Zeiten vermöge der durch die Schrödinger-Gleichung gegebenen Zeitentwicklung.

Nun nähern wir uns einer Idee, wie mathematisch die Observable $\vec{x}$ beschrieben werden kann dadurch, das wir bedenken, dass der **Erwartungswert** (oder Mittelwert) einer kontinuierlichen Zufallsvariablen über ihre Wahrscheinlichkeitsverteilung durch

$$\langle\vec{x}(t)\rangle = \int_{\mathbb{R}^3}\mathrm{d}^3xP(t,\vec{x})\vec{x} \tag{2.7.22}$$

gegeben ist. Daraus lässt sich nun zunächst wenig ablesen.

Betrachten wir nun aber zunächst die Observable $\vec{p}$, also den Impuls des Teilchens. Es liegt nun nahe, dass die Wahrscheinlichkeitsverteilung für den Impuls durch die Fouriertransformierte der Wellenfunktion $\tilde{\psi}(t,\vec{p})$ in (2.7.7) gegeben ist, d.h. wir nehmen an, dass die Impulswahrscheinlichkeitsverteilung gemäß der Bornschen Regel durch

$$\tilde{P}(t,\vec{p}) = |\tilde{\psi}(t,\vec{p})|^2 \tag{2.7.23}$$

gegeben ist. Gemäß der Umkehrung der Fourier-Transformation gilt

$$\tilde{\psi}(t,\vec{p}) = \int_{\mathbb{R}^3} \frac{\mathrm{d}^3 x'}{(2\pi\hbar)^{3/2}} \psi(t,\vec{x}') \exp\left(-\frac{\mathrm{i}\vec{x}'\cdot\vec{p}}{\hbar}\right). \tag{2.7.24}$$

Zum Beweis setzen wir diese Behauptung in (2.7.7) ein:

$$\int_{\mathbb{R}^3} \frac{\mathrm{d}^3\vec{p}}{(2\pi\hbar)^{3/2}} \tilde{\psi}(t,\vec{p}) \exp\left(\frac{\mathrm{i}\vec{x}\cdot\vec{p}}{\hbar}\right) = \int_{\mathbb{R}^3} \mathrm{d}^3 p \int_{\mathbb{R}^3} \mathrm{d}^3 x' \frac{1}{(2\pi\hbar)^3} \psi(t,\vec{x}') \exp\left(\frac{\mathrm{i}(\vec{x}-\vec{x}')\cdot\vec{p}}{\hbar}\right). \tag{2.7.25}$$

Vertauschen wir hier nun im Sinne der Distributionentheorie die beiden Integrale, ergibt sich

$$\begin{aligned}
\int_{\mathbb{R}^3} \frac{\mathrm{d}^3\vec{p}}{(2\pi\hbar)^{3/2}} \tilde{\psi}(t,\vec{p}) \exp\left(\frac{\mathrm{i}\vec{x}\cdot\vec{p}}{\hbar}\right) &= \int_{\mathbb{R}^3} \mathrm{d}^3 x' \psi(t,\vec{x}') \int_{\mathbb{R}^3} \mathrm{d}^3 p \frac{1}{(2\pi\hbar)^3} \exp\left(\frac{\mathrm{i}(\vec{x}-\vec{x}')\cdot\vec{p}}{\hbar}\right) \\
&= \int_{\mathbb{R}^3} \mathrm{d}^3 x' \psi(t,\vec{x}') \frac{1}{(2\pi\hbar)^3} (2\pi)^3 \delta\left(\frac{\vec{x}-\vec{x}'}{\hbar}\right) \\
&= \int_{\mathbb{R}^3} \mathrm{d}^3 x' \psi(t,\vec{x}') \delta(\vec{x}-\vec{x}') = \psi(t,\vec{x}),
\end{aligned} \tag{2.7.26}$$

und das war zu zeigen.

Als nächstes zeigen wir, dass mit $\psi(t,\vec{x})$ auch $\tilde{\psi}(t,\vec{p})$ automatisch auf 1 normiert ist. Es ist nämlich wegen (2.7.24)

$$\begin{aligned}
\int_{\mathbb{R}^3} \mathrm{d}^3 p |\tilde{\psi}(t,\vec{p})|^2 &= \int_{\mathbb{R}^3} \mathrm{d}^3 p \int_{\mathbb{R}^3} \mathrm{d}^3 x \int_{\mathbb{R}^3} \mathrm{d}^3 y \frac{1}{(2\pi\hbar)^3} \psi^*(t,\vec{x}) \psi(t,\vec{y}) \exp\left(\frac{\mathrm{i}(\vec{y}-\vec{x})\cdot\vec{p}}{\hbar}\right) \\
&= \int_{\mathbb{R}^3} \mathrm{d}^3 x \int_{\mathbb{R}^3} \mathrm{d}^3 y \frac{1}{(2\pi\hbar)^3} \psi^*(t,\vec{x}) \psi(t,\vec{y}) \int_{\mathbb{R}^3} \mathrm{d}^3 p \exp\left(\frac{\mathrm{i}(\vec{y}-\vec{x})\cdot\vec{p}}{\hbar}\right) \\
&= \int_{\mathbb{R}^3} \mathrm{d}^3 x \int_{\mathbb{R}^3} \mathrm{d}^3 y \psi^*(t,\vec{x}) \psi(t,\vec{y}) \delta^{(3)}(\vec{y}-\vec{x}) \\
&= \int_{\mathbb{R}^3} \mathrm{d}^3 x \psi^*(t,\vec{x}) \psi(t,\vec{x}) = \int_{\mathbb{R}^3} \mathrm{d}^3 x |\psi(t,\vec{x})|^2 = 1.
\end{aligned} \tag{2.7.27}$$

Ist also die **Ortswellenfunktion** $\psi(t,\vec{x})$ korrekt auf 1 normiert, so trifft dies mit unserer Definition der Fourier-Transformationen (2.7.7) und (2.7.24) (insbesondere durch die Wahl der Faktoren $1/(2\pi\hbar)^{3/2}$ in den Integralen) auch auf die **Impulswellenfunktion** $\tilde{\psi}(t,\vec{p})$ zu. Unsere obigen Betrachtungen zeigen, dass die Interpretation der Impulswellenfunktion im Sinne der Bornschen Regel als **Wahrscheinlichkeitsamplitude** für Impulswerte, d.h. dass (2.7.23) die Impulswahrscheinlichkeitsverteilung ist, in sich schlüssig ist.

Es ist klar, das für eine vorgegebene Ortswellenfunktion $\psi(t,\vec{x})$ über (2.7.24) die Impulswellenfunktion eindeutig bestimmt ist. Dann ist der Erwartungswert für den Impuls zur Zeit t durch

$$\left\langle \vec{p}(t) \right\rangle = \int_{\mathbb{R}^3} \mathrm{d}^3 p \tilde{P}(t,\vec{p}) \vec{p} \tag{2.7.28}$$

gegeben. Nun ist es i.a. nicht so einfach aus der Ortswellenfunktion durch Fouriertransformation die Impulswellenfunktion zu gewinnen, und wir wollen daher eine Formel für den Impulserwartungswert herleiten, die

sich direkt mit der Ortswellenfunktion berechnen lässt. Dazu verwenden wir (2.7.23) wie folgt:

$$\begin{aligned}
\left\langle \vec{p}(t) \right\rangle &= \int_{\mathbb{R}^3} \mathrm{d}^3 p \tilde{\psi}^*(t,\vec{p}) \vec{p} \tilde{\psi}(t,\vec{p}) \\
&= \int_{\mathbb{R}^3} \mathrm{d}^3 p \tilde{\psi}^*(t,\vec{p}) \vec{p} \int_{\mathbb{R}^3} \frac{\mathrm{d}^3 x}{(2\pi\hbar)^{3/2}} \psi(t,\vec{x}) \exp\left(-\frac{\mathrm{i}\vec{x}\cdot\vec{p}}{\hbar}\right) \\
&= \int_{\mathbb{R}^3} \mathrm{d}^3 p \tilde{\psi}^*(t,\vec{p}) \int_{\mathbb{R}^3} \frac{\mathrm{d}^3 x}{(2\pi\hbar)^{3/2}} \psi(t,\vec{x}) \mathrm{i}\hbar\vec{\nabla}_x \exp\left(-\frac{\mathrm{i}\vec{x}\cdot\vec{p}}{\hbar}\right) \\
&= \int_{\mathbb{R}^3} \mathrm{d}^3 p \tilde{\psi}^*(t,\vec{p}) \int_{\mathbb{R}^3} \frac{\mathrm{d}^3 x}{(2\pi\hbar)^{3/2}} \exp\left(-\frac{\mathrm{i}\vec{x}\cdot\vec{p}}{\hbar}\right) (-\mathrm{i}\hbar)\vec{\nabla}_x \psi(t,\vec{x}) \\
&= \int_{\mathbb{R}^3} \mathrm{d}^3 x (-\mathrm{i}\hbar)\vec{\nabla}_x \psi(t,\vec{x}) \int_{\mathbb{R}^3} \frac{\mathrm{d}^3 p}{(2\pi\hbar)^{3/2}} \tilde{\psi}^*(t,\vec{p}) \exp\left(-\frac{\mathrm{i}\vec{x}\cdot\vec{p}}{\hbar}\right) \\
&= \int_{\mathbb{R}^3} \mathrm{d}^3 x (-\mathrm{i}\hbar)\vec{\nabla}_x \psi(t,\vec{x}) \left[\int_{\mathbb{R}^3} \frac{\mathrm{d}^3 p}{(2\pi\hbar)^{3/2}} \tilde{\psi}(t,\vec{p}) \exp\left(+\frac{\mathrm{i}\vec{x}\cdot\vec{p}}{\hbar}\right)\right]^* \\
&= \int_{\mathbb{R}^3} \mathrm{d}^3 x \psi^*(t,\vec{x}) (-\mathrm{i}\hbar)\vec{\nabla}_x \psi(t,\vec{x}).
\end{aligned} \tag{2.7.29}$$

Diese Formel deuten wir nun so, dass der auf ψ wirkende Operator den Impuls „repräsentiert". Wir definieren also als **Impulsoperator**, der auf die Ortswellenfunktion wirkt,

$$\vec{\mathbf{p}}\psi(t,\vec{x}) = -\mathrm{i}\hbar\vec{\nabla}\psi(t,\vec{x}). \tag{2.7.30}$$

Damit können wir gemäß (2.7.29) den Erwartungswert des Impulses in der Form

$$\left\langle \vec{p}(t) \right\rangle = \int_{\mathbb{R}^3} \mathrm{d}^3 x \psi^*(t,\vec{p}) \vec{\mathbf{p}} \psi(t,\vec{x}) \tag{2.7.31}$$

schreiben.
Genauso können wir nun auch den **Ortsoperator** aus (2.7.22) ablesen, denn diese Gleichung können wir als

$$\langle \vec{x}(t) \rangle = \int_{\mathbb{R}^3} \mathrm{d}^3 x \psi^*(t,\vec{x}) \vec{x} \psi(t,\vec{x}), \tag{2.7.32}$$

d.h. wir definieren als **Ortsoperator** einfach

$$\vec{\mathbf{x}}\psi(t,\vec{x}) = \vec{x}\psi(t,\vec{x}), \tag{2.7.33}$$

d.h. der Ortsoperator bedeutet in Anwendung auf die Ortswellenfunktion einfach Multiplikation mit $\vec{x}$.
Umgekehrt folgt genauso für den Impulsoperator in Anwendung auf die Impulswellenfunktion

$$\vec{\mathbf{p}}\tilde{\psi}(t,\vec{p}) = \vec{p}\tilde{\psi}(t,\vec{p}). \tag{2.7.34}$$

In einer ähnlichen Rechnung wie oben in (2.7.29) ergibt sich für den Ortsoperator in Anwendung auf die Impulswellenfunktion (*Nachrechnen!*)

$$\vec{\mathbf{x}}\tilde{\psi}(t,\vec{p}) = +\mathrm{i}\hbar\vec{\nabla}_p \tilde{\psi}(t,\vec{p}). \tag{2.7.35}$$

2.7.2 Die Schrödinger-Gleichung für ein Teilchen in einem Potential

Wir können nun die Schrödinger-Gleichung für freie Teilchen in der Form

$$\mathrm{i}\hbar\partial_t\psi(t,\vec{x}) = \frac{1}{2m}\vec{\mathbf{p}}^2\psi(t,\vec{x}) = \mathbf{H}\psi(t,\vec{x}) \tag{2.7.36}$$

schreiben, wobei **H** der **Hamilton-Operator** des freien Teilchens

$$\mathbf{H} = \frac{1}{2m}\vec{\mathbf{p}}^2 \tag{2.7.37}$$

ist. Hätten wir es nicht mit Operatoren, die auf Funktionen wirken zu tun, sähe das genau aus wie in der klassischen Mechanik. Es ist nun einfach, die Schrödinger-Gleichung auf allgemeinere Situationen zu übertragen. Für ein Teilchen, das sich in einem **Potential äußerer Kräfte** bewegt, z.B. ein Elektron mit Ladung $(-e)$ im Coulomb-Potential eines als ruhend betrachteten Protons mit Ladung $(+e)$, ist die Gesamtenergie bzw. quantenmechanisch der Hamilton-Operator

$$\mathbf{H} = \frac{1}{2m}\vec{\mathbf{p}}^2 + V(\vec{\mathbf{x}}). \tag{2.7.38}$$

Dabei verwendet man für das Potential einfach die aus der klassischen Mechanik bekannte Funktion, also für das Elektron im Coulomb-Potential eines ruhenden Protons

$$V(\vec{\mathbf{x}}) = -\frac{e^2}{4\pi\epsilon_0|\vec{\mathbf{x}}|}. \tag{2.7.39}$$

Dabei ist die Definition der Funktion des Ortsoperators auf die Ortswellenfunktion sehr einfach, denn die Wirkung des Ortsoperators ist ja einfach durch die Multiplikation mit $\vec{x}$ definiert. Folglich ist hier einfach

$$V(\vec{\mathbf{x}})\psi(t,\vec{x}) = V(\vec{x})\psi(t,\vec{x}) = -\frac{e^2}{4\pi\epsilon_0|\vec{x}|}\psi(t,\vec{x}). \tag{2.7.40}$$

Freilich ist damit kein Beweis gegeben, dass dann die Schrödinger-Gleichung mit dem Hamilton-Operator (2.7.38)

$$\mathrm{i}\hbar\partial_t\psi(t,\vec{x}) = \mathbf{H}\psi(t,\vec{x}) = \left[\frac{1}{2m}\vec{\mathbf{p}}^2 + V(\vec{x})\right]\psi(t,\vec{x}) = -\frac{\hbar^2}{2m}\Delta\psi(t,\vec{x}) + V(\vec{x})\psi(t,\vec{x}) \tag{2.7.41}$$

korrekte Ergebnisse liefert. Wie wir im Folgenden allerdings sehen werden, liefert diese Annahme hervorragend mit der Erfahrung übereinstimmende Ergebnisse. Wie wir sehen werden, liefert die Quantentheorie mit dem Hamilton-Operator (2.7.38) und dem Coulomb-Potential (2.7.39) hervorragende Vorhersagen für das Spektrum des Wasserstoffatoms.

2.7.3 Messwerte von Observablen

Wir haben nun bereits einige charakteristische Eigenschaften des quantenmechanischen Formalismus' kennengelernt: Der Zustand eines quantenmechanischen Systems wird demnach durch eine **Wellenfunktion** $\psi(t,\vec{x})$ beschrieben, deren Betragsquadrat integrierbar über $\vec{x} \in \mathbb{R}^3$ ist und daher gemäß (2.7.13) auf 1 normiert werden kann. Dann ist gemäß der **Bornschen Regel** $P(t,\vec{x}) = |\psi(t,\vec{x})|^2$ die **Wahrscheinlichkeitsdichte** für das Vorfinden des Teilchens am Ort $\vec{x}$.

Observablen werden durch **lineare Operatoren** repräsentiert. Für den Ort und Impuls gilt

$$\vec{\mathbf{x}}\psi(t,\vec{x}) = \vec{x}\psi(t,\vec{x}), \quad \vec{\mathbf{p}}\psi(t,\vec{x}) = -\mathrm{i}\hbar\vec{\nabla}_x\psi(t,\vec{x}). \tag{2.7.42}$$

Wir können Erwartungswerte von Observablen durch Integrale der Art

$$\langle O(t)\rangle = \int_{\mathbb{R}^3} \mathrm{d}^3x\,\psi^*(t,\vec{x})\mathbf{O}\psi(t,\vec{x}) \tag{2.7.43}$$

berechnen. Dabei ist **O** irgendeine Funktion der Operatoren $\vec{\mathbf{x}}$ und $\vec{\mathbf{p}}$, wobei man allerdings aufpassen muss, wenn man irgendwelche Funktionen aus der klassischen Mechanik in die Quantenmechanik übernimmt, weil i.a. die Operatoren **nicht vertauschen**. Wir kommen darauf im Folgenden noch ausführlich zu sprechen.

Die **Zeitentwicklung** der Wellenfunktion, also des Zustandes des Teilchens, ist durch die **Schrödinger-Gleichung**

$$\mathrm{i}\hbar\partial_t\psi(t,\vec{x}) = \mathbf{H}\psi(t,\vec{x}) \tag{2.7.44}$$

gegeben, wobei der **Hamilton-Operator** in Analogie zur klassischen Mechanik durch

$$\mathbf{H} = \frac{1}{2m}\vec{\mathbf{p}}^2 + V(\vec{\mathbf{x}}) \tag{2.7.45}$$

gegeben ist.

Wir müssen uns nun fragen, inwiefern die linearen Operatoren, die Observablen beschreiben, mit den möglichen **Messwerten**, die diese Observablen annehmen können, zusammenhängen. Dazu fragen wir uns, welche Wellenfunktionen die Situation beschreiben, dass eine Observable einen vorherbestimmten oder **determinierten Wert** besitzt.

Sei also **O** irgendein Operator, der eine Observable beschreibt. Dann definieren wir eine Wellenfunktion $u_o(\vec{x})$ als einen Zustand, für den diese Observable den exakt bestimmten Wert o besitzt, wenn für beliebige Funktionen $f(O)$ bzgl. des Erwartungswertes

$$\langle f(O)\rangle_{u_o} = \int_{\mathbb{R}^3} u_o^*(\vec{x})f(\mathbf{O})u_o(\vec{x}) = f(o) \tag{2.7.46}$$

gilt.

Aus der Experimentalphysik kennen wir z.B. die **Standardabweichung** als **Maß für die Unsicherheit eines Messergebnisses**. Da i.a. Observablen in der Quantentheorie keine bestimmten Werte annehmen, sondern lediglich Wahrscheinlichkeiten für das Auftreten eines bestimmten Wertes bei einer Messung der Observable bekannt sind, ist die Standardabweichung hier ein Maß für die Unbestimmtheit des **Mittelwerts**, d.h.

$$\Delta O^2 = \left\langle (O - \langle O\rangle)^2\right\rangle = \left\langle O^2\right\rangle - \langle O\rangle^2 \geq 0 \tag{2.7.47}$$

ist ein Maß, um wieviel die Resultate einer Messung der Observablen O um den Mittelwert **fluktuieren**, wenn man die Messung sehr oft wiederholt. Nach der obigen Definition für den Zustand u_o, der beschreibt, dass O mit Sicherheit den bestimmten Wert o annimmt folgt in der Tat $\Delta O^2 = 0$, d.h. bei jeder Messung erhält man den Wert o.

Es ist nun klar, das dies für *alle* Funktionen $f(O)$ nur erfüllt sein kann, wenn

$$f(\mathbf{O})u_o(\vec{x}) = f(o)u_0(\vec{x}) \tag{2.7.48}$$

ist. Das ist aber der Fall, wenn

$$\mathbf{O}u_o(\vec{x}) = o\,u_o(\vec{x}) \tag{2.7.49}$$

ist. Man nennt eine solche Funktion eine **Eigenfunktion** des Operators **O** zum Eigenwert o. Es ist klar, dass physikalische Observablen **reelle Messwerte** besitzen sollten. Wir müssen also gleich noch herausfinden, welche Eigenschaften die Operatoren haben müssen, damit es nur Eigenfunktionen zu **reellen Eigenwerten** gibt. Jedenfalls können wir dann postulieren, dass die möglichen Messwerte einer Observablen die (reellen) Eigenwerte des Operators sind, der diese Observable im quantenmechanischen Formalismus repräsentiert.

Wir betrachten zunächst als Beispiel den **Impulsoperator**. Die Eigenwertgleichung lautet in diesem Fall

$$\vec{\mathbf{p}} u_{\vec{p}}(\vec{x}) = \vec{p} u_{\vec{p}}(\vec{x}). \tag{2.7.50}$$

Es muss für jede Komponente des Impulses also

$$\mathbf{p}_j u_{\vec{p}}(\vec{x}) = -\mathrm{i}\hbar \partial_j u_{\vec{p}}(\vec{x}) = p_j u_{\vec{p}}(\vec{x}) \tag{2.7.51}$$

gelten. Dieses System partieller Differentialgleichungen lässt sich nun für beliebige komplexe (!) Zahlen p_j leicht lösen:

$$u_{\vec{p}}(\vec{x}) = A(\vec{p}) \exp\left(\frac{\mathrm{i}\vec{x}\cdot\vec{p}}{\hbar}\right), \quad A(\vec{p}) = \text{const.} \tag{2.7.52}$$

Das sind die ebenen Wellen, falls $\vec{p} \in \mathbb{R}^3$ ist. Nun ergibt sich allerdings ein Problem. Sollte $u_{\vec{p}}(\vec{x})$ einen „richtigen Zustand" repräsentieren, müsste $|u_{\vec{p}}(\vec{x})|^2$ über $\vec{x} \in \mathbb{R}^3$ integrierbar sein, damit wir die Konstante $A(\vec{p})$ so bestimmen können, dass die Normierungsbedingung (2.7.13) erfüllt ist. Das ist aber offenbar für keine Werte für $\vec{p}$ der Fall. Für allgemeine komplexe Zahlen wird $u_{\vec{p}}(\vec{x})$ sogar exponentiell divergieren für $\vec{x} \to \infty$.

Allerdings zeichnen sich die sich die *reellen* Eigenwerte, also $\vec{p} \in \mathbb{R}^3$ dadurch aus, dass $|u_{\vec{p}}(\vec{x})|^2 = |A|^2 = \text{const}$ ist, d.h. das Betragsquadrat bleibt für alle $\vec{x} \in \mathbb{R}^3$ wenigstens beschränkt. Freilich können wir noch immer nicht die Normierungsbedingung erfüllen. Es gibt also keinen Zustand, in dem ein Teilchen einen determinierten Impuls $\vec{p}$ annimmt!

Allerdings beobachten wir noch folgendes: Betrachten wir statt des Normierungsintegrals das Integral

$$\begin{aligned}\left\langle u_{\vec{p}} \,\middle|\, u_{\vec{p}'} \right\rangle &:= \int_{\mathbb{R}^3} \mathrm{d}^3x\, u^*_{\vec{p}}(\vec{x}) u_{\vec{p}'}(\vec{x}) \\ &= \int_{\mathbb{R}^3} \mathrm{d}^3x\, A^*(\vec{p}) A(\vec{p}') \exp\left[\frac{\mathrm{i}\vec{x}(\vec{p}'-\vec{p})}{\hbar}\right] \\ &= (2\pi\hbar)^3 |A(\vec{p})|^2 \delta^{(3)}(\vec{p}-\vec{p}').\end{aligned} \tag{2.7.53}$$

Wir definieren nun die **verallgemeinerte Impulseigenfunktion** durch die Lösungen der entsprechenden Eigenwertgleichung (2.7.52) indem wir $A(\vec{p}) = 1/(2\pi\hbar)^{3/2}$ setzen, also

$$u_{\vec{p}}(\vec{x}) = \frac{1}{(2\pi\hbar)^{3/2}} \exp\left(\frac{\mathrm{i}\vec{p}\cdot\vec{x}}{\hbar}\right). \tag{2.7.54}$$

Dann ist bequemerweise das obige Integral (2.7.53) „auf eine δ-Distribution normiert":

$$\left\langle u_{\vec{p}} \,\middle|\, u_{\vec{p}'} \right\rangle = \delta^{(3)}(\vec{p}-\vec{p}') \tag{2.7.55}$$

Damit wird nun aber auch unsere obige heuristische Betrachtung zu Orts- und Impulswellenfunktionen klar. Wir können zwar keine wirkliche Impulseigenfunktion definieren, weil die Ebenewellenlösung (2.7.54) der Eigenwertgleichung nicht **quadratintegrabel** ist, wir können aber durch **Superposition** unserer verallgemeinerten Impulseigenfunktionen beliebige echte Zustände definieren, d.h. alle **quadratintegrablen Ortswellenfunktionen** lassen sich durch

$$\psi(t,\vec{x}) = \int_{\mathbb{R}^3} \mathrm{d}^3p\, \tilde{\psi}(t,\vec{p}) u_{\vec{p}}(\vec{x}) \tag{2.7.56}$$

mit beliebigen **quadratintegrablen Impulswellenfunktionen** $\tilde{\psi}(t,\vec{p})$ bestimmen. Dies ist freilich nichts anderes als die oben ausführlich besprochene **Fouriertransformation** zwischen Orts- und Impulswellenfunktionen. Es ist auch klar, dass der durch die Ortswellenfunktion $\psi(t,\vec{x})$ beschriebene Zustand genauso gut

auch durch die Impulswellenfunktion $\tilde{\psi}(t,\vec{p})$ beschrieben werden kann. Schließlich ist die Fourier-Transformation umkehrbar, d.h. Orts- und Impulswellenfunktionen bestimmen sich mit Hilfe der verallgemeinerten Impulseigenfunktionen gegenseitig. Die Umkehrung von (2.7.56) lässt sich nämlich sehr einfach durch

$$\tilde{\psi}(t,\vec{p}) = \left\langle u_{\vec{p}} \,\middle|\, \psi(t) \right\rangle = \int_{\mathbb{R}^3} \mathrm{d}^3x\, u^*_{\vec{p}}(\vec{x})\psi(t,\vec{x}) \tag{2.7.57}$$

finden. Setzt man hier nämlich (2.7.54) ein, erhält man tatsächlich die Umkehrtransformation von (2.7.56), also (2.7.24).

2.7.4 Die Hilbert-Raum-Struktur der Wellenfunktionen

Jetzt können wir noch einige mathematische Eigenschaften des bislang gefundenen Formalismus' besprechen, die wir benötigen, um die Quantentheorie im nächsten Kapitel durch einige griffige Postulate formulieren zu können.

Wir stellen zunächst fest, dass alle wichtigen Operationen mit Wellenfunktionen **linear** sind. Insbesondere ist die **Schrödinger-Gleichung** für die Zeitentwicklung der Wellenfunktionen eine **lineare partielle Differentialgleichung**. Auch die Operatoren, die Observablen beschreiben sind **lineare Operatoren**.

Wegen der Linearität der Schrödingergleichung (2.7.44) gilt für beliebige Lösungen ψ_1 und ψ_2 der Schrödinger-Gleichung das **Superpositionsprinzip**, d.h. für beliebige Zahlen λ_1, λ_2 ist auch die Linearkombination

$$\psi(t,\vec{x}) = \lambda_1\psi_1(t,\vec{x}) + \lambda_2\psi_2(t,\vec{x}), \tag{2.7.58}$$

d.h. die Lösungen der Schrödinger-Gleichung bilden einen **komplexen Vektorraum von Funktionen**.

Im Folgenden ist es nun bequem für die typischen Integrale, wie sie im letzten Abschnitt öfter aufgetreten sind, die Schreibweise

$$\langle \psi_1 | \psi_2 \rangle = \int_{\mathbb{R}^3} \mathrm{d}^3x\, \psi_1^*(\vec{x})\psi_2(\vec{x}). \tag{2.7.59}$$

Dies definiert ein sog. **Skalarprodukt** auf dem Funktionenvektorraum. Dabei ist zu beachten, dass das linke Argument in das definierende Integral durch die *komplex konjugierte Wellenfunktion* ψ_1^* eingeht, d.h. das Skalarprodukt ist nur im rechten Argument linear, d.h. es gilt stets

$$\langle \psi_1 | \lambda_1\psi_2 + \lambda_2\psi_3 \rangle = \lambda_2 \langle \psi_1 | \psi_2 \rangle + \lambda_2 \langle \psi_2 | \psi_3 \rangle . \tag{2.7.60}$$

Im ersten Argument ist das Skalarprodukt **semilinear**, d.h. es gilt

$$\langle \lambda_1\psi_1 + \lambda_2\psi_2 | \psi_3 \rangle = \lambda_1^* \langle \psi_1 | \psi_3 \rangle + \lambda_2^* \langle \psi_2 | \psi_3 \rangle , \tag{2.7.61}$$

d.h. die komplexen Zahlen in der Linearkombination von Wellenfunktionen im linken Argument sind beim „Herausziehen" aus dem Skalarprodukt komplex zu konjugieren.

Schließlich gilt noch beim Vertauschen der Argumente

$$\langle \psi_2 | \psi_1 \rangle = \langle \psi_1 | \psi_2 \rangle^* . \tag{2.7.62}$$

Weiter ist das Skalarprodukt auch **positiv definit**, denn es gilt

$$\langle \psi | \psi \rangle = \int_{\mathbb{R}^3} \mathrm{d}^3x\, \psi^*(\vec{x})\psi(\vec{x}) = \int_{\mathbb{R}^3} \mathrm{d}^3x\, |\psi(\vec{x})|^2 \geq 0. \tag{2.7.63}$$

Ist nun $\langle \psi | \psi \rangle = 0$ ist die Wellenfunktion notwendig ebenfalls fast überall 0, und wir identifizieren solche Funktionen einfach mit der Nullfunktion.

Es müssen dabei freilich die betrachteten Wellenfunktionen quadratintegrabel sein, damit die Integrale oben überhaupt einen Sinn ergeben, d.h. einen endlichen Wert besitzen. Erfreulicherweise kann man zeigen, dass für zwei quadratintegrable Funktionen ψ_1 und ψ_2 stets auch das Integral (2.7.59) wohldefiniert ist.

Die quadratintegrablen Funktionen bilden also einen komplexen Vektorraum mit Skalarprodukt, einen sog. **Hilbert-Raum**. Der hier für die Quantenmechanik konstruierte Hilbertraum der quadratintegrablen Funktionen wird als $\mathrm{L}^2(\mathbb{R}^3)$ bezeichnet.

2.7.5 Selbstadjungierte Operatoren

Als nächstes beantworten wir die Frage, wie wir sicherstellen, dass die Operatoren, die Observablen repräsentieren nur reelle Werte besitzen. Wir haben oben gesehen, dass die möglichen Messwerte der Observablen O durch die Eigenwerte des entsprechenden Operators $\mathbf{O}$ gegeben ist, d.h. die Eigenwerte des Operators sollten reell sein. Außerdem müssen wir noch einen Formalismus finden, der es erlaubt, aus den Eigenfunktionen zu den Eigenwerten und der gegebenen Wellenfunktion $\psi(t,\vec{x})$, die den Zustand des Teilchens beschreibt, die Wahrscheinlichkeiten (bzw. Wahrscheinlichkeitsverteilungen), dass der entsprechende Eigenwert bei einer Messung der Observablen als Messergebnis resultiert.

Wir geben hier einfach die mathematische Lösung vor und zeigen dann, dass dies zu einer konsistenten Theorie in diesem Sinne führt. Zunächst definieren wir, dass zu einem Operator $\mathbf{O}$ der **adjungierte Operator** $\mathbf{O}^\dagger$ dadurch gegeben ist, dass für beliebige Wellenfunktionen ψ_1 und ψ_2

$$\langle \psi_1 | \mathbf{O}\psi_2 \rangle = \left\langle \mathbf{O}^\dagger \psi_1 \middle| \psi_2 \right\rangle \tag{2.7.64}$$

gilt.

Betrachten wir als Beispiele den Orts- und Impulsoperator. Für den Ortsoperator $\vec{\mathbf{x}}$ folgt

$$\begin{aligned}
\langle \psi_1 | \vec{\mathbf{x}}\psi_2 \rangle &= \int_{\mathbb{R}^3} \mathrm{d}^3x\, \psi_1^*(\vec{x})\vec{\mathbf{x}}\psi_2(\vec{x}) \\
&= \int_{\mathbb{R}^3} \mathrm{d}^3x\, \psi_1^*(x)\vec{x}\psi_2(\vec{x}) \\
&= \int_{\mathbb{R}^3} \mathrm{d}^3x\, [\vec{x}\psi_1(\vec{x})]^*\psi_2(\vec{x}) \\
&= \int_{\mathbb{R}^3} \mathrm{d}^3x\, [\vec{\mathbf{x}}\psi_1(\vec{x})]^*\psi_2(\vec{x}) \\
&= \langle \vec{\mathbf{x}}\psi_1 | \psi_2 \rangle .
\end{aligned} \tag{2.7.65}$$

Das bedeutet, dass für den Ortsoperator

$$\vec{\mathbf{x}}^\dagger = \vec{\mathbf{x}} \tag{2.7.66}$$

gilt. Man nennt Operatoren, für die $\mathbf{O}^\dagger = \mathbf{O}$ gilt **selbstadjungiert**.

Betrachten wir als nächstes den Impulsoperator:

$$\begin{aligned}
\langle \psi_1 | \vec{\mathbf{p}}\psi_2 \rangle &= \int_{\mathbb{R}^3} \mathrm{d}^3x\, \psi_1^*(\vec{x})(-\mathrm{i}\hbar)\vec{\nabla}\psi_2(\vec{x}) \\
&= \int_{\mathbb{R}^3} \mathrm{d}^3x \left[-\mathrm{i}\hbar\vec{\nabla}[\psi_1^*(\vec{x})\psi_2(\vec{x})] + \mathrm{i}\hbar[\vec{\nabla}\psi_1^*(\vec{x})]\psi_2(\vec{x}) \right]
\end{aligned} \tag{2.7.67}$$

Auf den ersten Term können wir den Gaußschen Integralsatz anwenden. Da die Wellenfunktionen quadratintegrabel sind, dürfen wir davon aus gehen, dass das Oberflächenintegral im Unendlichen verschwindet. Es

ist also

$$\begin{aligned}\langle\psi_1|\vec{\mathbf{p}}\psi_2\rangle &= \int_{\mathbb{R}^3} \mathrm{d}^3x\,\mathrm{i}\hbar[\vec{\nabla}\psi_1^*(\vec{x})]\psi_2(\vec{x}) \\ &= \int_{\mathbb{R}^3} \mathrm{d}^3x[-\mathrm{i}\hbar\vec{\nabla}\psi_1(\vec{x})]^*\psi_2(\vec{x}) \\ &= \langle\vec{\mathbf{p}}\psi_1|\psi_2\rangle\,.\end{aligned} \tag{2.7.68}$$

Es ist also auch

$$\vec{\mathbf{p}}^\dagger = \vec{\mathbf{p}}. \tag{2.7.69}$$

Wir gehen nun davon aus, dass allgemein die Operatoren, die Observablen repräsentieren **selbstadjungierte Operatoren** sein sollen. Es sei also $\mathbf{O} = \mathbf{O}^\dagger$. Wir nehmen nun an, $\mathbf{O}$ besitze normierbare Eigenvektoren, um einige wichtige Eigenschaften selbstadjungierter Operatoren herzuleiten. Bei Orts- und Impulsoperatoren hatten wir allerdings kontinuierliche Eigenwerte und entsprechend keine normierbaren Eigenfunktionen. Vielmehr ließen sich z.B. die Impulseigenfunktionen nur auf δ-Distributionen normieren, und wir haben diese als verallgemeinerte Eigenfunktionen bezeichnet. Ähnlich verhält es sich auch für andere Operatoren, die Observablen repräsentieren, und es ist in jedem Einzelfall zu prüfen, wie die Lösungen des Eigenwertproblems jeweils aussehen.

Nehmen wir aber der Einfachheit halber zunächst an, dass $\mathbf{O}$ einen diskreten Satz von Eigenwerten o besitzt und dazugehörige normierbare Eigenfunktionen $u_{o,k}(\vec{x})$. Dabei nummeriert k möglicherweise existierende voneinander linear unabhängige Eigenvektoren von $\mathbf{O}$ zum gleichen Eigenwert o durch. Wir nehmen an, dass auch k jeweils eine kontinuierliche Menge durchläuft. Wir nehmen im Folgenden an, dass die Eigenfunktionen alle normiert sind, d.h. $\left\langle u_{o,k} \middle| u_{o,k} \right\rangle = 1$ gilt. Sie erfüllen definitionsgemäß die Eigenwertgleichungen

$$\mathbf{O}u_{o,k}(\vec{x}) = o\,u_{o,k}(\vec{x}). \tag{2.7.70}$$

Wir zeigen nun zuerst, dass **die Eigenwerte selbstadjungierter Operatoren** reell sind. Dazu rechnen wir

$$\left\langle u_{o,k} \middle| \mathbf{O}u_{o,k} \right\rangle = \left\langle u_{o,k} \middle| o\,u_{o,k} \right\rangle = o\left\langle u_{o,k} \middle| u_{o,k} \right\rangle = o. \tag{2.7.71}$$

Andererseits ist

$$o = \left\langle u_{o,k} \middle| \mathbf{O}u_{o,k} \right\rangle = \left\langle \mathbf{O}^\dagger u_{o,k} \middle| u_{o,k} \right\rangle = \left\langle \mathbf{O}u_{o,k} \middle| u_{o,k} \right\rangle = \left\langle o\,u_{o,k} \middle| u_{o,k} \right\rangle = o^*\left\langle u_{o,k} \middle| u_{o,k} \right\rangle = o^*. \tag{2.7.72}$$

Es ist also $o = o^*$ und damit $o \in \mathbb{R}$.

Es seien nun o und o' zwei *verschiedene* Eigenwerte von $\mathbf{O}$ dann gilt

$$\begin{aligned}\left\langle u_{o,k} \middle| \mathbf{O}u_{o',k'} \right\rangle &= o'\left\langle u_{o,k} \middle| u_{o',k'} \right\rangle \\ &= \left\langle \mathbf{O}^\dagger u_{o,k} \middle| u_{o',k'} \right\rangle \\ &= \left\langle \mathbf{O}u_{o,k} \middle| u_{o',k'} \right\rangle \\ &= \left\langle o\,u_{o,k} \middle| u_{o',k'} \right\rangle \\ &= o^*\left\langle u_{o,k} \middle| u_{o',k'} \right\rangle \\ &= o\left\langle u_{o,k} \middle| u_{o',k'} \right\rangle.\end{aligned} \tag{2.7.73}$$

Fassen wir die erste mit der zweiten Zeile zusammen, ergibt sich

$$(o' - o)\left\langle u_{o,k} \middle| u_{o',k'} \right\rangle = 0. \tag{2.7.74}$$

Da voraussetzungsgemäß $o \neq o'$ gilt, ist also notwendig

$$\left\langle u_{o,k} \middle| u_{o',k'} \right\rangle = 0. \tag{2.7.75}$$

In Analogie zu gewöhnlichen reellen Vektorräumen mit Skalarprodukt sagt man, die beiden **Eigenvektoren zu verschiedenen Eigenwerten sind orthogonal** zueinander.

Betrachten wir nun den Fall, dass es zu einem Eigenwert o mehrere linear unabhängige Eigenfunktionen gibt. Diese haben wir ja bereits mit dem Index k durchnummeriert. Es ist klar, dass diese Vektoren einen **Untervektorraum**, den wir $\mathrm{Eig}(\mathbf{O}, o)$ nennen wollen, aufspannen, denn mit zwei Eigenfunktionen zu Eigenwert o ist auch jede Linearkombination eine Eigenfunktion zu diesem Eigenwert, denn es gilt

$$\begin{aligned} \mathbf{O}[\lambda_1 u_{o,k_1}(\vec{x}) + \lambda_2 u_{o,k_2}(\vec{x})] &= \lambda_1 \mathbf{O} u_{o,k_1}(\vec{x}) + \lambda_2 \mathbf{O} u_{o,k_2}(\vec{x}) \\ &= \lambda_1 o u_{o,k_1}(\vec{x}) + \lambda_2 o u_{o,k_2}(\vec{x}) \\ &= o[\lambda_1 u_{o,k_1}(\vec{x}) + \lambda_2 u_{o,k_2}(\vec{x})]. \end{aligned} \tag{2.7.76}$$

Wir können nun die verschiedenen Eigenvektoren zum gleichen Eigenwert **orthonormalisieren**, d.h. wir können eine Basis von $\mathrm{Eig}(\mathbf{O}, o)$ finden, so dass $\left\langle u_{o,k_1} \middle| u_{o,k_2} \right\rangle = \delta_{k_1 k_2}$ ist.

Wir führen das am Beispiel, dass drei linear unabhängige Eigenvektoren vorliegen, vor. Es sei $u_{o,1}$ auf 1 normiert. Dann berechnen wir

$$u'_{o,2}(\vec{x}) = u_{o,2}(\vec{x}) - \left\langle u_{o,1} \middle| u_{o,2} \right\rangle u_{o,1}(\vec{x}). \tag{2.7.77}$$

Diese Eigenfunktion ist mit Sicherheit von 0 verschieden, denn sonst würden $u_{o,1}$ und $u_{o,2}$ keine linear unabhängigen Funktionen, was wir aber vorausgesetzt haben. Wir können also auch diesen Vektor normieren

$$\tilde{u}_{o,2}(\vec{x}) = \frac{1}{\sqrt{\left\langle u'_{o,2} \middle| u'_{o,2} \right\rangle}} u'_{o,2}(\vec{x}). \tag{2.7.78}$$

Offensichtlich ist nun $\left\langle u_{o,1} \middle| \tilde{u}_{o,2} \right\rangle = 0$ und $\left\langle \tilde{u}_{o,2} \middle| \tilde{u}_{o,2} \right\rangle = 1$. Nun müssen wir schließlich noch dritten zu $u_{o,1}$ und $u_{o,2}$ orthogonalen normierten Eigenfunktion finden. Das lässt sich aber genau analog wie eben ausführen:

$$u'_{o,3}(\vec{x}) = u_{o,3}(\vec{x}) - \left\langle u_{o,1} \middle| u_{o,3} \right\rangle u_{o,1}(\vec{x}) - \left\langle \tilde{u}_{o,2} \middle| u_{o,3} \right\rangle \tilde{u}_{o,2}. \tag{2.7.79}$$

Es gilt wieder $u_{o,3} \neq 0$, da die drei ursprünglichen Vektoren linear unabhängig sein sollten. Daher können wir auch diese Eigenfunktion normieren, d.h.

$$\tilde{u}_{o,3}(\vec{x}) = \frac{1}{\sqrt{\left\langle u'_{o,3} \middle| u'_{o,3} \right\rangle}} u'_{o,3}(\vec{x}). \tag{2.7.80}$$

Setzen wir schließlich noch $\tilde{u}_{o,1}(\vec{x}) = u_{o,1}(\vec{x})$, so gilt für diese drei Vektoren

$$\left\langle \tilde{u}_{o,k_1} \middle| \tilde{u}_{o,k_2} \right\rangle = \delta_{k_1,k_2}. \tag{2.7.81}$$

Es ist klar, dass wir dieses sog. **Schmidtsche Orthonormierungsverfahren** auch für den Fall, dass mehr als drei linear unabhängige Eigenvektoren vorliegen, ausführen können.

Wir können also im Folgenden davon ausgehen, dass wir zu jedem selbstadjungierten Operator $\mathbf{O}$ ein **Orthonormalsystem** von Eigenfunktionen finden können, d.h. Eigenfunktionen $u_{o,k}$, für die

$$\left\langle u_{o_1,k_1} \middle| u_{o_2,k_2} \right\rangle = \delta_{o_1,o_2} \delta_{k_1,k_2} \tag{2.7.82}$$

gilt.
Wir können nun überlegen, wie wir für eine gegebene Zustandswellenfunktion $\psi(t,\vec{x})$ die Wahrscheinlichkeiten, dass bei einer Messung der Observablen O der Eigenwert o des repräsentierenden Operators $\mathbf{O}$ als Messergebnis herauskommt.
Für die Impulsmessungen haben wir gesehen, dass die Wahrscheinlichkeiten durch

$$P(t,\vec{p}) = \left|\tilde{\psi}(t,\vec{p})\right|^2 = \left|\left\langle u_{\vec{p}} \,\middle|\, \psi(t)\right\rangle\right|^2 \tag{2.7.83}$$

gegeben sind. Wir *postulieren*, dass diese **Bornsche Regel** analog auch für beliebige Messungen gilt, wobei wir die **Entartung** berücksichtigen müssen, also die Möglichkeit, dass es zu einem Eigenwert mehrere linear unabhängige Eigenfunktionen und das entsprechende Orthonormalsystem von $u_{o,k}$ gibt:

$$P(t,o) = \sum_k \left|\left\langle u_{o,k} \,\middle|\, \psi(t)\right\rangle\right|^2. \tag{2.7.84}$$

Damit dies mit der Wahrscheinlichkeitsinterpretation konsistent ist, muss

$$\sum_o P(t,o) = 1 \tag{2.7.85}$$

gelten, denn die Wahrscheinlichkeit, irgendeinen der möglichen Messwerte bei einer Messung der Observablen zu erhalten, muss natürlich 1 sein, und das muss für beliebige auf 1 normierte Zustandswellenfunktionen $\psi(t,\vec{x})$ gelten, also für jeden möglichen Zustand, in dem sich das Teilchen bei der Messung befindet.
Dies impliziert, dass die Orthonormalsysteme aus Eigenfunktionen von **O vollständig** sein müssen, d.h. für jede Wellenfunktion $\psi(t,\vec{x})$ muss es möglich sein, diese nach den Eigenfunktionen von $\mathbf{O}$ zu entwickeln, d.h. es muss Zahlen $\psi_{o,k}(t)$ geben, so dass

$$\psi(t,\vec{x}) = \sum_{o,k} \psi_{o,k}(t) u_{o,k}(\vec{x}) \tag{2.7.86}$$

ist. Wegen der Orthonormalität der $u_{o,k}(\vec{x})$ folgt daraus, dass

$$\left\langle u_{o',k'} \,\middle|\, \psi(t)\right\rangle = \sum_{o,k} \psi_{o,k}(t) \left\langle u_{o',k'} \,\middle|\, u_{o,k}\right\rangle = \sum_{o,k} \psi_{o,k}(t) \delta_{o,o'} \delta_{k,k'} = \psi_{o',k'}(t) \tag{2.7.87}$$

ist. Die $\psi_{o,k}(t)$ nennt man auch die **Zustandswellenfunktionen** in der O-Darstellung. Die Vollständigkeitsrelation ist formal durch

$$\begin{aligned}
\psi(t,\vec{x}) &= \sum_{o,k} \left\langle u_{o,k} \,\middle|\, \psi(t)\right\rangle u_{o,k}(\vec{x}) = \sum_{o,k} u_{o,k}(\vec{x}) \int_{\mathbb{R}^3} \mathrm{d}^3x' \, u^*_{o,k}(\vec{x}') \psi(t,\vec{x}') \\
&= \int_{\mathbb{R}^3} \mathrm{d}^3x' \, \psi(t,\vec{x}') \sum_{o,k} u^*_{o,k}(\vec{x}') u_{o,k}(\vec{x}) \\
\Rightarrow \sum_{o,k} u^*_{o,k}(\vec{x}') u_{o,k}(\vec{x}) &= \delta^{(3)}(\vec{x}-\vec{x}').
\end{aligned} \tag{2.7.88}$$

Aus dieser **Vollständigkeitsrelation** der Eigenfunktionen von $\mathbf{O}$ folgt dann auch, dass

$$\begin{aligned}
1 = \langle \psi(t) | \psi(t) \rangle &= \sum_{o_1,k_1} \sum_{o_2,x_3} \left\langle u_{o_1,k_1} \,\middle|\, \psi(t)\right\rangle^* \left\langle u_{o_2,k_2} \,\middle|\, \psi(t)\right\rangle \left\langle u_{o_1,k_1} \,\middle|\, u_{o_2,k_2}\right\rangle \\
&= \sum_{o_1,k_1} \sum_{o_2,x_3} \left\langle u_{o_1,k_1} \,\middle|\, \psi(t)\right\rangle^* \left\langle u_{o_2,k_2} \,\middle|\, \psi(t)\right\rangle \delta_{o_1,o_2} \delta_{k_1,k_2} \\
&= \sum_{o_1,k_1} \left|\left\langle u_{o_1,k_1} \,\middle|\, \psi(t)\right\rangle\right|^2.
\end{aligned} \tag{2.7.89}$$

Damit ist gezeigt, dass die Vollständigkeit des Orthonormalsystems von Eigenfunktionen die Bedingung (2.7.85) impliziert.

Man kann zeigen, dass die in der Physik auftretenden selbstadjungierten Operatoren i.a. tatsächlich vollständige Orthonormalsysteme von Eigenfunktionen besitzen. Dies ist aber ein recht schwieriges mathematisches Problem, und wir werden im Folgenden die Vollständigkeit der orthonormierten Eigenfunktionen einfach voraussetzen.

Es ist auch klar, dass in dem Fall, dass die selbstadjungierten Operatoren kontinuierliche Eigenwerte besitzen (z.B. sind die Eigenwerte der Impulsoperatoren $\mathbf{p}_j$ alle reellen Zahlen $p_j \in \mathbb{R}$), die Summe über diese Eigenwerte durch ein Integral zu ersetzen ist, also z.B.

$$\psi(t,\vec{x}) = \int_{\mathbb{R}^3} \mathrm{d}^3 p \left\langle u_{\vec{p}} \middle| \psi(t) \right\rangle u_{\vec{p}}(\vec{x}) = \int_{\mathbb{R}^3} \mathrm{d}^3 p \tilde{\psi}(t,\vec{p}) \frac{1}{\sqrt{(2\pi\hbar)^3}} \exp\left(\frac{\mathrm{i}\vec{p}\cdot\vec{x}}{\hbar}\right). \tag{2.7.90}$$

2.7.6 Energieeigenfunktionen

Eine wichtige Observable ist die Energie des Teilchens, die durch den Hamilton-Operator

$$\mathbf{H} = \frac{1}{2m}\vec{\mathbf{p}}^2 + V(\vec{\mathbf{x}}) \tag{2.7.91}$$

gegeben ist. Dieser Operator nimmt insofern eine Sonderstellung ein, weil er nach dem oben motivierten Postulat auch die **Zeitentwicklung** der Wellenfunktion bestimmt, die durch die Lösungen der Schrödinger-Gleichung

$$\mathrm{i}\hbar\partial_t \psi(t,\vec{x}) = \mathbf{H}\psi(t,\vec{x}) = -\frac{\hbar^2}{2m}\Delta\psi(t,\vec{x}) + V(\vec{x})\psi(t,\vec{x}) \tag{2.7.92}$$

bestimmt wird. Dabei ist die Lösung eindeutig durch die Angabe der Wellenfunktion zur Anfangszeit $t = 0$ bestimmt. Dies entspricht der **Präparation** des Systems zur Zeit $t = 0$ in dem durch die Wellenfunktion beschriebenen Zustand.

Wir betrachten nun die **Energie-Eigenfunktion** $u_n(\vec{x})$, d.h. die Eigenfunktionen des Hamilton-Operators: $\mathbf{H}u_n(\vec{x}) = E_n u_n(\vec{x})$. Wir nehmen der Einfachheit halber an, dass die dazugehörigen Energieeigenwerte $E_n \in \mathbb{R}$ eine diskrete Menge durchlaufen die wir mit $n \in \mathbb{N}$ durchnummerieren. Weiter nehmen wir an, dass diese Eigenfunktionen eine vollständige Orthonormalbasis bilden, d.h. wir können die Wellenfunktion nach diesen Eigenfunktionen gemäß

$$\psi(t,\vec{x}) = \sum_{n=1}^{\infty} \psi_n(t) u_n(\vec{x}). \tag{2.7.93}$$

Setzen wir diesen Ansatz in die Schrödinger-Gleichung (2.7.92) ein, folgt

$$\begin{aligned} \mathrm{i}\hbar\dot{\psi}(t,\vec{x}) &= \sum_{n=1}^{\infty} \mathrm{i}\hbar\dot{\psi}_n(t) u_n(\vec{x}) = \mathbf{H}\psi(t,\vec{x}) \\ &= \sum_{n=1}^{\infty} \psi_n(t)\mathbf{H}u_n(\vec{x}) = \sum_{n=1}^{\infty} E_n \psi_n(t) u_n(\vec{x}). \end{aligned} \tag{2.7.94}$$

Im letzten Schritt haben wir verwendet, dass voraussetztungsgemäß u_n eine Eigenfunktion des Hamilton-Operators zum Eigenwert E_n ist. Da die Entwicklung nach Energie-Eigenfunktionen wegen deren Orthonormalität eindeutig ist, müssen die $\psi_n(t)$

$$\mathrm{i}\hbar\dot{\psi}_n(t) = E_n\psi_n(t) \Rightarrow \psi_n(t) = \psi_{n0}\exp\left(-\frac{\mathrm{i}E_n t}{\hbar}\right) \tag{2.7.95}$$

erfüllen. Es gilt also

$$\psi(t,\vec{x}) = \sum_{n=1}^{\infty} \psi_{n0} \exp\left(-\frac{\mathrm{i}E_n t}{\hbar}\right) u_n(\vec{x}). \tag{2.7.96}$$

Die Konstanten ψ_{n0} ergeben sich aus der Anfangsbedingung $\psi(0,\vec{x}) = \psi_0(\vec{x})$, denn wegen (2.7.96) gilt ja

$$\psi_0(\vec{x}) = \psi(0,\vec{x}) = \sum_{n=1}^{\infty} \psi_{n0} u_n(\vec{x}). \tag{2.7.97}$$

Wegen der Orthonormalität der Energieeigenfunktionen $\left\langle u_j \middle| u_n \right\rangle = \delta_{jn}$ folgt

$$\left\langle u_j \middle| \psi_0 \right\rangle = \sum_{n=1}^{\infty} \left\langle u_j \middle| u_n \right\rangle \psi_{n0} = \sum_{n=1}^{\infty} \delta_{jn} \psi_{n0} = \psi_{j0}. \tag{2.7.98}$$

Ein spezieller Fall ist nun, dass zur Zeit $t = 0$ die Wellenfunktion ein Energieeigenzustand ist, d.h. wenn $\psi_0(\vec{x}) = u_k(\vec{x})$ ist. Dann folgt

$$\psi_{n0} = \langle u_n | \psi_0 \rangle = \langle u_n | u_k \rangle = \delta_{nk}. \tag{2.7.99}$$

Setzen wir dies in (2.7.96) ein, folgt

$$\psi(t,\vec{x}) = \exp\left(-\frac{\mathrm{i}E_k t}{\hbar}\right) u_k(\vec{x}). \tag{2.7.100}$$

Die Wellenfunktion ist also bis auf den Phasenfaktor $\exp(-\mathrm{i}E_k t/\hbar)$ durch die Energieeigenfunktion u_k gegeben, und die Aufenthaltswahrscheinlichkeitsverteilung ist durch

$$P(t,\vec{x}) = |\psi(t,\vec{x})|^2 = |u_k(\vec{x})|^2 = P(t=0,\vec{x}) \tag{2.7.101}$$

gegeben.

Das bedeutet, dass sich die Wahrscheinlichkeitsverteilung nicht mit der Zeit ändert, d.h. die **Energieeigenfunktionen beschreiben die stationären Zustände** des Teilchens. Allgemein beschreiben Wellenfunktionen, die sich nur um einen Phasenfaktor voneinander unterscheiden, nämlich offenbar denselben physikalischen Zustand des Systems, da alle physikalisch bedeutsamen Größen wie die Aufenthaltswahrscheinlichkeitsdichte und Erwartungswerte beliebiger Observabler von diesem Phasenfaktor unabhängig sind (*warum?*).

2.7.7 Kompatible und inkompatible Observable

Aus den Überlegungen im vorigen Abschnitt geht hervor, dass man i.a. ein Teilchen nicht so präparieren kann, dass irgendwelche willkürlich ausgewählten Observablen bestimmte Werte annehmen. Dies ist wohl eine der erstaunlichsten Folgerungen der Quantentheorie. Wir untersuchen in diesem Abschnitt, welche Observablen eines Teilchens simultan exakt bestimmte Werte annehmen können.

Betrachten wir also irgendwelche zwei Observable A und B und die selbstadjungierten Operatoren $\mathbf{A}$ und $\mathbf{B}$, die diese in der Quantenmechanik repräsentieren. Aufgrund der Betrachtungen im vorigen Abschnitt muss ein Zustand, für den sowohl A als auch B scharfe Werte a und b annehmen, ein **simultaner Eigenzustand** der beiden Operatoren $\mathbf{A}$ und $\mathbf{B}$ zu den Eigenwerten a und b sein. Damit also für beliebige mögliche Messwerte a und b, die notwendig Eigenwerte der Operatoren $\mathbf{A}$ und $\mathbf{B}$ sein müssen, stets Zustände existieren, für die diese Observablen zugleich scharfe Werte für A und B besitzt, muss es ein vollständiges Orthonormalsystem $u_{a,b,k}(\vec{x})$ **simultaner Eigenfunktionen** der selbstadjungierten Operatoren $\mathbf{A}$ und $\mathbf{B}$ geben. Dabei nummeriert wieder k, wie im vorigen Abschnitt, mögliche linear unabhängige simultane Eigenvektoren zu den gleichen Eigenwerten a und b durch. Man nennt solche Observablen auch **miteinander verträglich**.

Was bedeutet dies nun aber für die Operatoren **A** und **B**? Betrachten wir dazu die Anwendung beider Operatoren auf die Eigenzustände:

$$\mathbf{AB}u_{a,b,k}(\vec{x}) = \mathbf{A}b\,u_{a,b,k} = b\mathbf{A}u_{a,b,k} = a\,b\,u_{a,b,k} = \mathbf{BA}u_{a,b,k}, \tag{2.7.102}$$

d.h. in ihrer Wirkung auf das vollständige Orthonormalsystem der simultanen Eigenvektoren ist die Reihenfolge der Hintereinanderanwendung der Operatoren irrelevant, d.h. es gilt $\mathbf{AB} = \mathbf{BA}$. Man sagt auch, dass die **Operatoren kommutieren**. Da die Kommutativität (oder eben auch die Nichtkommutativität!) von Operatorprodukten eine wichtige Rolle spielt, führt man den sog. **Kommutator zweier Operatoren** durch

$$[\mathbf{A},\mathbf{B}] = \mathbf{AB} - \mathbf{BA} \tag{2.7.103}$$

ein.

Da nun das Orthonormalsystem simultaner Eigenfunktionen vorausssetzungsgemäß auch vollständig ist, können wir jede Wellenfunktion $\psi(t,\vec{x})$ nach diesen simultanen Eigenfunktionen entwickeln, d.h.

$$\psi(t,\vec{x}) = \sum_{a,b,k} \tilde{\psi}(t,a,b,k)u_{a,b,k}(\vec{x}), \quad \tilde{\psi}_{a,b,k} = \left\langle u_{a,b,k} \middle| \psi(t) \right\rangle. \tag{2.7.104}$$

Damit folgt aber, dass die Operatoren **A** und **B** in Anwendung auf *jede* Wellenfunktion kommutieren, denn es ist ja

$$\begin{aligned}
\mathbf{AB}\psi(t,\vec{x}) &= \sum_{a,b,k} \tilde{\psi}(t,a,b,k)\mathbf{AB}u_{a,b,k}(\vec{x}) \\
&= \sum_{a,b,k} \tilde{\psi}(t,a,b,k)\mathbf{BA}u_{a,b,k}(\vec{x}) \\
&= \mathbf{BA}\sum_{a,b,k} \tilde{\psi}(t,a,b,k)u_{a,b,k}(\vec{x}) \\
&= \mathbf{BA}\psi(t,\vec{x}).
\end{aligned} \tag{2.7.105}$$

Das bedeutet, dass die beiden Observablen nur dann miteinander verträglich sein können, wenn die entsprechenden selbstadjungierten Operatoren kommutieren, d.h. wenn

$$[\mathbf{A},\mathbf{B}] = 0 \tag{2.7.106}$$

ist. Man kann mathematisch zeigen, dass auch umgekehrt zwei kommutierende selbstadjungierte Operatoren auch stets ein simultanes vollständiges System von simultanen Eigenfunktionen besitzen, d.h. zwei Observablen sind genau dann miteinander kompatibel, wenn die sie repräsentierenden selbstadjungierten Operatoren kommutieren.

Wir stellen z.B. fest, dass die drei Ortskomponenten allesamt kommutieren, d.h. die drei Ortskomponenten können im Prinzip zugleich scharfe Werte annehmen. In der Tat ist ja $\mathbf{x}_j = x_j$ und damit sicher $\left[\mathbf{x}_j,\mathbf{x}_k\right] = 0$ für $j,k \in \{1,2,3\}$.

Dabei muss man hier aber beachten, dass die Ortskomponenten kontinuierliche Observable sind und daher keine echten Eigenvektoren besitzen sondern nur auf die δ-Distribution normierte, d.h. für diese verallgemeinerten Eigenvektoren des Ortsoperators gilt

$$u_{\vec{x}_1}(\vec{x}) = \delta^{(3)}(\vec{x} - \vec{x}_1), \tag{2.7.107}$$

denn es muss ja wegen der Vollständigkeit dieses verallgemeinerten Orthonormalsystems für jede Wellenfunktion gelten

$$\psi(t,\vec{x}) = \int_{\mathbb{R}^3} \mathrm{d}^3\vec{x}'\, u_{\vec{x}'}(\vec{x})\tilde{\psi}(t,\vec{x}'), \tag{2.7.108}$$

wobei

$$\tilde{\psi}(t,\vec{x}') = \langle u_{\vec{x}'} \,|\, \psi(t)\rangle = \int_{\mathbb{R}^3} \mathrm{d}^3\vec{x}\, u^*_{\vec{x}'}(\vec{x})\psi(t,\vec{x}) \tag{2.7.109}$$

ist. Andererseits ist aber eben $\psi(t,\vec{x})$ bereits der Zustand in der Ortsdarstellung, d.h. es muss $\tilde{\psi}(t,\vec{x}') = \psi(t,\vec{x}')$ für beliebige Wellenfunktionen $\psi(t,\vec{x})$ gelten, und das kann nur erfüllt sein, wenn (2.7.107) gilt.

Die Kompatibilität der drei Komponenten des Ortsvektors eines Teilchens bedeutet also, dass man alle drei Komponenten simultan beliebig genau bestimmen kann, aber doch immer noch endliche Unsicherheiten Δx_j für jeden tatsächlich realisierbaren Zustand, also für jede quadratintegrable Wellenfunktion $\psi(t,\vec{x})$ bestehen müssen. Dies ist eine Besonderheit von Observablen mit kontinuierlichen Werten.

Natürlich kann man dieselben Überlegungen mit dem Impulsoperator anstellen. In der Tat kommutieren ja auch die Operatoren für die Impulskomponenten, denn es ist ja $\mathbf{p}_j = -\mathrm{i}\hbar\partial_j$, und die partiellen Ableitungen kommutieren ja in Anwendung zumindest auf Wellenfunktionen, die stetige partielle Ableitungen besitzen.

Allerdings kommutieren Orts- und Impulskomponenten i.a. nicht, d.h. **Ort und Impuls eines Teilchens sind nichtkompatible Observable**, d.h. sie können nicht zugleich beliebig scharf determiniert sein. Das rechnet man wie folgt nach. Für eine beliebige Wellenfunktion gilt

$$\mathbf{x}_j\mathbf{p}_k\psi(t,\vec{x}) = x_j(-\mathrm{i}\hbar\partial_k)\psi(t,\vec{x}), \quad \mathbf{p}_k\mathbf{x}_j = (-\mathrm{i}\hbar\partial_k)[x_j\psi(t,\vec{x}) = -\mathrm{i}\hbar[\delta_{jk}\psi(t,\vec{x}) + x_j\partial_k\psi(t,\vec{x})]. \tag{2.7.110}$$

Zieht man beide Gleichungen voneinander ab, erhält man

$$\left[\mathbf{x}_j,\mathbf{p}_k\right]\psi(t,\vec{x}) = \mathrm{i}\hbar\delta_{jk}\psi(t,\vec{x}) \Rightarrow \left[\mathbf{x}_j,\mathbf{p}_k\right] = \mathrm{i}\hbar\delta_{jk}\mathbb{1}, \tag{2.7.111}$$

wobei $\mathbb{1}$ der **Identitätsoperator** ist, der definitionsgemäß auf beliebige Wellenfunktionen angewandt einfach nichts bewirkt, d.h.

$$\mathbb{1}\psi(t,\vec{x}) = \psi(t,\vec{x}). \tag{2.7.112}$$

Schließlich bemerken wir noch, dass definitionsgemäß ein **vollständiger Satz voneinander unabhängiger verträglicher Observabler** A_k ($k \in \{1,2,\ldots,n\}$) solche verträglichen Observablen sind, für die die simultanen Eigenräume eindimensional sind, d.h. dass für die möglichen simultanen Eigenwerte $(a_1,\ldots,a_n)$ nur eine linear unabhängige simultane Eigenfunktion $u_{a_1,\ldots,a_n}(\vec{x})$ existiert. Die verträglichen Observablen heißen dabei voneinander unabhängig, wenn keine dieser Observablen eine Funktion der übrigen ist.

Z.B. sind definitionsgemäß die drei Ortskomponenten oder auch die drei Impulskomponenten jeweils vollständige Sätze voneinander unabhängiger verträglicher Observabler.

Das bedeutet, dass der Zustand eines Teilchens dadurch **vollständig bestimmt werden** kann, dass er durch eine simultane Eigenfunktion eines vollständigen Satzes voneinander unabhängiger kompatibler Observabler gegeben ist, d.h. praktisch, dass man das Teilchen so präpariert, dass ein solcher Satz von Observablen bestimmte Werte $(a_1,\ldots,a_n)$ annimmt.

Auf die Frage, was man aus der Inkompatibilität von Observablen schließen kann, kommen wir später noch zurück. Dabei werden wir auf die sog. **Heisenbergsche Unbestimmtheitsrelation** geführt. Demnach können z.B. Orts- und Impulskomponenten bzgl. der gleichen Koordinatenachse nicht simultan beliebig scharf bestimmt werden, es gilt vielmehr die besagte Unbestimmtheitsrelation

$$\Delta x_j \Delta p_j \geq \frac{\hbar}{2}. \tag{2.7.113}$$

Wir werden diese wichtige Schlussfolgerung Heisenbergs von 1927 im nächsten Kapitel beweisen.

Jedenfalls haben wir nun die nötigen mathematischen Grundlagen für die griffige Formulierung der **Postulate für die Quantentheorie**, so dass wir die Theorie im nächsten Kapitel systematisch formulieren und anwenden können.

Kapitel 3

Quantenmechanik

In diesem Kapitel formulieren wir nun die am Ende des vorigen Kapitels heuristischen Überlegungen als übersichtliche Postulate. Dabei bedienen wir uns der Formulierung der Theorie mit einem abstrakten Hilbert-Raum anstelle von (Ort-)Wellenfunktionen, die auf **Dirac** zurückgeht. Dies hat den Vorteil, dass die konzeptionellen Grundlagen der Theorie sich besonders prägnant beschreiben lassen. Es zeigt sich auch, dass die meisten Rechnungen, die in vielen Lehrbüchern durch die Lösung der Schrödinger-Gleichung behandelt werden, sich auch im abstrakten Formalismus viel transparenter darstellen lassen und auch die wellenmechanische Herleitung erheblich vereinfachen.

Basierend auf den heuristischen Betrachtungen anhand der (freilich reichlich verkürzten) Historie zur Entwicklung der Quantenmechanik formulieren wir zunächst die grundlegenden **Postulate**. Im Hinblick auf die **physikalische Interpretation** des Formalismus' folgen wir dabei der **minimalen statistische Interpretation**, die auch als **Ensemble-Interpretation** bekannt ist [Bal70, Bal98]. Zunächst formulieren wir die Quantenmechanik für ein Teilchen mit Spin 0, das sich in einem äußeren Potential bewegt.

Wir schreiben nun zunächst die **Grundpostulate der Quantentheorie** in übersichtlicher Form hin und erläutern die mathematischen Grundlagen, auf denen diese basieren, in den folgenden Abschnitten

3.1 Die Grundpostulate der Quantentheorie

Wir stellen zunächst die Struktur der Quantentheorie in einigen **Grundpostulaten** zusammen und erläutern sie in den folgenden Abschnitten genauer:

1. Der Zustand eines quantenmechanischen Systems wird durch einen normierten Vektor $|\psi\rangle$ eines Hilbertraums $\mathscr{H}$ repräsentiert.

2. Jede physikalische Observable O wird durch einen (auf einem dichten Teilraum von $\mathscr{H}$ definierten) selbstadjungierten Operator $\mathbf{O}$ repräsentiert.

   Die möglichen Messwerte der Observablen sind durch die (verallgemeinerten) Eigenwerte des ihr zugeordneten Operators gegeben.

3. Die (verallgemeinerten) Eigenvektoren $|o,\alpha\rangle$ des Operators $\mathbf{O}$ zum (verallgemeinerten) Eigenwert o können normiert und zueinander orthogonal bzw. auf die δ-Distribution normiert gewählt werden, d.h. so, dass
$$\left\langle o',\alpha' \,\middle|\, o,\alpha \right\rangle = \delta(o'-o)\delta(\alpha'-\alpha) \tag{3.1.1}$$
   gilt Dabei bezeichnet α einen oder mehrere Parameter, die im Falle einer Entartung des Eigenraums die Eigenvektoren durchnumerieren. Diese Parameter können sowohl kontinuierliche als auch diskrete Werte durchlaufen, und die δ-Symbole in Gl. (3.1.1) bezeichnen entsprechend δ-Distributionen oder

Kronecker-δ-Symbole. Ist das System bei einer Messung der Observablen O im normierten Zustand $|\psi\rangle$ präpariert, so ist die Wahrscheinlichkeit, bei der Messung der Observablen O den Eigenwert o des ihr zugeordneten Operators $\mathbf{O}$ zu finden, durch die **Bornsche Regel**

$$w_\psi(o) = \sum\!\!\!\!\!\!\int \mathrm{d}\alpha\, |\langle o, \alpha \,|\, \psi \rangle|^2 \tag{3.1.2}$$

gegeben, wobei das kombinierte Summations-Integrationssymbol das Integral über den kontinuierlichen und die Summe über den diskreten Teil des Parameters α bedeutet.

4. Die Dynamik des Systems wird eindeutig durch die Zuordnung eines selbstadjungierten nach unten beschränkten Operators $\mathbf{H}$, des **Hamiltonoperators** des Systems, bestimmt.

   Ist $\mathbf{O}$ der die Observable O repräsentierende selbstadjungierte Operator, so repräsentiert die kovariante Zeitableitung

$$\mathring{\mathbf{O}} = \frac{1}{\mathrm{i}\hbar}[\mathbf{O}, \mathbf{H}] \tag{3.1.3}$$

   die zeitliche Ableitung $\dot{O}$ der Observablen O.

3.1.1 Anmerkungen zum Begriff des quantenmechanischen Zustandes

Aus der physikalischen Bedeutung des Zustandsvektors $|\psi\rangle$, der durch die Bornsche Regel (3.1.2) gegeben ist, geht hervor, dass für jeden normierten Zustandsvektor $|\psi\rangle$ die Zustandsvektoren

$$\left|\psi_\varphi\right\rangle = \exp(\mathrm{i}\varphi)\,|\psi\rangle \quad \text{mit} \quad \varphi \in \mathbb{R} \tag{3.1.4}$$

die gleichen Wahrscheinlichkeitsverteilungen (3.1.2) ergeben, d.h. all diese Zustandsvektoren, die sog. **Einheitsstrahlen**, sind die eigentlichen eindeutig den Zustand bestimmenden mathematischen Objekte, d.h. wir müssen streng genommen das Postulat 1 wie folgt umformulieren:

1. 1.' Der Zustand eines quantenmechanischen Systems wird durch einen Einheitsstrahl im Hilbertraum repräsentiert.

Dazu äquivalent ist offenbar die Definition

1. 1." Der Zustand eines quantenmechanischen Systems wird durch den Projektionsoperator

$$\mathbf{R}_\psi = |\psi\rangle\langle\psi| \tag{3.1.5}$$

   mit einem normierten Hilbertraumvektor $|\psi\rangle$ eindeutig bestimmt.

Es ist nämlich

$$\mathbf{R}_{\psi_\varphi} = |\exp(\mathrm{i}\varphi)\psi\rangle\langle\exp(\mathrm{i}\varphi)\psi| = \exp(\mathrm{i}\varphi)\exp(-\mathrm{i}\varphi)\langle\psi\,|\,\psi\rangle = \mathbf{R}_\psi. \tag{3.1.6}$$

Außerdem lassen sich mit diesem Projektionsoperator auch die Wahrscheinlichkeitsverteilungen (3.1.2) bestimmen:

$$w_\psi(o) = \sum\!\!\!\!\!\!\int \mathrm{d}\alpha |\langle o, \alpha \,|\, \psi\rangle|^2 = \sum\!\!\!\!\!\!\int \mathrm{d}\alpha\, \langle o, \alpha \,|\, \psi\rangle\langle\psi \,|\, o, \alpha\rangle = \sum\!\!\!\!\!\!\int \mathrm{d}\alpha \left\langle o, \alpha \middle| \mathbf{R}_\psi \middle| o, \alpha \right\rangle. \tag{3.1.7}$$

Schließlich ist noch zu beachten, dass diese Zustände nur denjenigen Zuständen entsprechen, die einer **vollständigen Kenntnis des Zustandes des Systems entspricht**, also einer möglichen vollständigen **Präparation**

des quantenmechanischen Systems. Die Wahrscheinlichkeiten, die sich für die möglichen Messresultate o der Observablen O gemäß (3.1.7) ergeben, rühren also nicht von einer Unkenntnis über den Systemzustand her, sondern sind **objektiv**, d.h. i.a. besitzen Observablen **keinen determinierten Wert**, auch wenn der Zustand vollständig durch (3.1.5) spezifiziert worden ist. Man bezeichnet solche Zustände, die eine **vollständige Systempräparation** beschreiben, auch als **reine Zustände**.

Es ist wichtig zu bemerken, dass man den Zustandbegriff noch verallgemeinern kann, indem man Situationen betrachtet, in denen der Zustand des Systems in diesem quantenmechanischen Sinne nicht vollständig bekannt ist. Um auch diesen Fall zu erfassen, nehmen wir an, dass Alice ein Ensemble von Quantensystemen erzeugt, für das mit einer einer Wahrscheinlichkeit (bzw. Wahrscheinlichkeitsdichte) P_k das System in den reinen Zuständen $\mathbf{R}_{\psi_k}$ präpariert wird. Misst dann Bob an diesem Ensemble die Observable O ergeben sich offenbar die Wahrscheinlichkeiten bzw. Wahrscheinlichkeitsdichten für den Messwert o zu

$$w_{\mathbf{R}}(o) = \sum\!\!\!\!\!\!\int \mathrm{d}k P_k \sum\!\!\!\!\!\!\int \mathrm{d}\alpha \left\langle o,\alpha \middle| \mathbf{R}_{\psi_k} \middle| o,\alpha \right\rangle = \sum\!\!\!\!\!\!\int \mathrm{d}\alpha \left\langle o,\alpha \middle| \sum\!\!\!\!\!\!\int \mathrm{d}k P_k \mathbf{R}_{\psi_k} \middle| o,\alpha \right\rangle. \tag{3.1.8}$$

Dieser **gemischte Zustand** wird also durch den **statistischen Operator**

$$\mathbf{R} = \sum\!\!\!\!\!\!\int \mathrm{d}k P_k \mathbf{R}_{\psi_k} = \sum\!\!\!\!\!\!\int \mathrm{d}k P_k \left|\psi_k\right\rangle \left\langle\psi_k\right| \tag{3.1.9}$$

repräsentiert. Der Spezialfall eines reinen Zustands ist natürlich in dieser Beschreibung allgemeiner gemischter Zustände enthalten, wenn $P_{k=1} = 1$ und $P_k = 0$ für alle $k \neq 1$ ist, denn dann ist offenbar $\mathbf{R} = \mathbf{R}_{\psi_1}$. Damit erhalten wir die allgemeinst mögliche Definition für einen quantenmechanischen Zustand

1.''' Ein (reiner oder gemischter) Zustand eines quantenmechanischen Systems ist durch einen selbstadjungierten positiv semidefiniten Operator $\mathbf{R}$ mit

$$\sum\!\!\!\!\!\!\int \mathrm{d}o\, w_{\mathbf{R}}(o) = \sum\!\!\!\!\!\!\int \mathrm{d}o \mathrm{d}\alpha \left\langle o,\alpha \middle| \mathbf{R} \middle| o,\alpha \right\rangle = \mathrm{Tr}\,\mathbf{R} = 1 \tag{3.1.10}$$

definiert.

Man macht sich leicht klar, dass die **Spur** eines Operators, die mit Hilfe eines beliebigen vollständigen Orthonormalsystems $|u_n\rangle$ durch

$$\mathrm{Tr}\,\mathbf{R} = \sum\!\!\!\!\!\!\int \mathrm{d}n \left\langle u_n \middle| \mathbf{R} \middle| u_n \right\rangle \tag{3.1.11}$$

definiert ist, unabhängig vom verwendeten Orthonormalsystem ist. Ist nämlich auch $|v_k\rangle$ ein vollständiges

Orthonormalsystem, folgt

$$\begin{aligned}
\sum\!\!\!\!\!\!\int \mathrm{d}k\, \langle v_k | \mathbf{R} | v_k \rangle &= \sum\!\!\!\!\!\!\int \mathrm{d}n \sum\!\!\!\!\!\!\int \mathrm{d}n' \sum\!\!\!\!\!\!\int \mathrm{d}k\, \langle v_k | u_n \rangle \langle u_n | \mathbf{R} | u_{n'} \rangle \\
&= \sum\!\!\!\!\!\!\int \mathrm{d}n \sum\!\!\!\!\!\!\int \mathrm{d}n' \langle u_n | \mathbf{R} | u_{n'} \rangle \sum\!\!\!\!\!\!\int \mathrm{d}k\, \langle u_{n'} | v_k \rangle \langle v_k | u_n \rangle \\
&= \sum\!\!\!\!\!\!\int \mathrm{d}n \sum\!\!\!\!\!\!\int \mathrm{d}n' \langle u_n | \mathbf{R} | u_{n'} \rangle \langle u_{n'} | u_n \rangle \\
&= \sum\!\!\!\!\!\!\int \mathrm{d}n \sum\!\!\!\!\!\!\int \mathrm{d}n' \langle u_n | \mathbf{R} | u_{n'} \rangle \delta_{nn'} \\
&= \sum\!\!\!\!\!\!\int \mathrm{d}n \langle u_n | \mathbf{R} | u_n \rangle = \mathrm{Tr}\, \mathbf{R},
\end{aligned} \tag{3.1.12}$$

d.h. man kann die Spur eines Operators mit einem beliebigen vollständigen Orthonormalsystem berechnen. Die Spur charakterisiert also einen Operator unabhängig von der Wahl des Orthonormalsystems.

3.2 Der Hilbert-Raum

Das Verhalten eines quantenmechanischen System wird mathematisch durch einen komplexen **Hilbert-Raum** $\mathscr{H}$ (David Hilbert 1862-1943) beschrieben. Dabei handelt es sich um einen i.a. unendlichdimensionalen Vektorraum mit den Skalaren in $\mathbb{C}$ und einem positiv definiten Skalarprodukt. Nach der von Dirac eingeführten „**Bra-Ket-Schreibweise**“ werden die Vektoren mit $|\psi\rangle$ bezeichnet. Für sie gelten die üblichen Axiome eines Vektorraums für die Addition von Vektoren und Multiplikation von Vektoren mit komplexen Zahlen:

$$|\psi_1\rangle + |\psi_2\rangle = |\psi_2\rangle + |\psi_1\rangle \tag{3.2.1}$$
$$(|\psi_1\rangle + |\psi_2\rangle) + |\psi_3\rangle = |\psi_1\rangle + (|\psi_2\rangle + |\psi_3\rangle) \tag{3.2.2}$$
$$\exists\, |0\rangle \in \mathscr{H}: \quad |\psi_1\rangle + 0 = |\psi_1\rangle, \tag{3.2.3}$$
$$\lambda_1(|\psi_1\rangle + |\psi_2\rangle) = \lambda_1 |\psi_1\rangle + \lambda_1 |\psi_2\rangle, \tag{3.2.4}$$
$$(\lambda_1 + \lambda_2)|\psi_1\rangle = \lambda_1 |\psi_1\rangle + \lambda_2 |\psi_1\rangle, \tag{3.2.5}$$
$$0|\psi_1\rangle = |0\rangle, \quad 1|\psi\rangle = |\psi\rangle. \tag{3.2.6}$$

Daraus ergibt sich unter anderem auch, dass es zu jedem Vektor $|\psi\rangle$ einen Vektor $|-\psi\rangle = (-1)|\psi\rangle = -|\psi\rangle$ gibt, so dass $|\psi\rangle + |-\psi\rangle = 0$ gilt. Wegen (3.2.6) und (3.2.5) ist nämlich

$$|0\rangle = (1-1)|\psi\rangle = 1|\psi\rangle + (-1)|\psi\rangle =: |\psi\rangle + |-\psi\rangle. \tag{3.2.7}$$

Weiter ist auf dem Hilbertraum noch ein **Skalarprodukt**, das zwei Vektoren $|\psi_1\rangle, |\psi_2\rangle \in \mathscr{H}$ auf eine komplexe Zahl $\langle \psi_1 | \psi_2 \rangle$ abbildet, definiert. Es besitzt die folgenden Eigenschaften:

$$\langle \psi_1 | \lambda_1 \psi_2 + \lambda_2 \psi_3 \rangle = \lambda_1 \langle \psi_1 | \psi_2 \rangle + \lambda_2 \langle \psi_1 | \psi_3 \rangle, \tag{3.2.8}$$
$$\langle \psi_1 | \psi_2 \rangle = \langle \psi_2 | \psi_1 \rangle^* \tag{3.2.9}$$
$$\langle \psi | \psi \rangle \geq 0, \tag{3.2.10}$$
$$\langle \psi | \psi \rangle = 0 \Leftrightarrow |\psi\rangle = 0. \tag{3.2.11}$$

Eine wichtige Folgerung aus (3.2.8) und (3.2.9) ist

$$\langle \lambda_1 \psi_1 + \lambda_2 \psi_2 | \psi_3 \rangle = \lambda_1^* \langle \psi_1 | \psi_3 \rangle + \lambda_2^* \langle \psi_2 | \psi_3 \rangle, \tag{3.2.12}$$

d.h. das Skalarprodukt ist bzgl. des zweiten Arguments **linear** aber bzgl. des ersten Arguments **semilinear**, d.h. die Zahlenfaktoren in der Linearkombination sind beim Herausziehen aus dem ersten Argument komplex zu konjugieren. Man sagt auch, das Skalarprodukt sei eine **Sesquilinearform**. Der Beweis folgt einfach aus (3.2.9) und (3.2.8) sowie einfachen Eigenschaften der komplexen Konjugation für komplexe Zahlen:

$$\begin{aligned}\langle \lambda_1\psi_1+\lambda_2\psi_2 \,|\, \psi_3\rangle &= (\langle \psi_3 \,|\, \lambda_1\psi_1+\lambda_2\psi_2\rangle)^* = (\lambda_1 \langle \psi_3 \,|\, \psi_1\rangle + \lambda_2 \langle \psi_3 \,|\, \psi_2\rangle)^* \\ &= \lambda_1^* \langle \psi_3 \,|\, \psi_1\rangle^* + \lambda_2^* \langle \psi_3 \,|\, \psi_2\rangle^* = \lambda_1^* \langle \psi_1 \,|\, \psi_3\rangle + \lambda_2^* \langle \psi_2 \,|\, \psi_3\rangle\,.\end{aligned} \tag{3.2.13}$$

Auf dem Hilbertraum wird mit dem Skalarprodukt zugleich auch eine **Norm** definiert:

$$||\psi|| = \sqrt{\langle \psi \,|\, \psi\rangle} \geq 0. \tag{3.2.14}$$

Diese Definition erfüllt in der Tat die Eigenschaften einer Vektorraumnorm, d.h. es gilt

$$||\lambda\psi|| = |\lambda|\,||\psi||. \tag{3.2.15}$$

Der Beweis ist eine sehr einfache *Übungsaufgabe*.
Etwas schwieriger ist der Beweis der **Dreiecksungleichung**

$$||\psi_1+\psi_2|| \leq ||\psi_1|| + ||\psi_2||. \tag{3.2.16}$$

Dazu betrachten wir

$$\begin{aligned}||\psi_1+\psi_2||^2 &= \langle \psi_1+\psi_2 \,|\, \psi_1+\psi_2\rangle = \langle \psi_1 \,|\, \psi_1\rangle + \langle \psi_1 \,|\, \psi_2\rangle + \langle \psi_2 \,|\, \psi_1\rangle + \langle \psi_2 \,|\, \psi_2\rangle \\ &= ||\psi_1||^2 + ||\psi_2||^2 + \langle \psi_1 \,|\, \psi_2\rangle + \langle \psi_2 \,|\, \psi_1\rangle\,.\end{aligned} \tag{3.2.17}$$

Nun ist (3.2.16) offenbar gleichbedeutend mit

$$||\psi_1+\psi_2||^2 \overset{?}{\leq} ||\psi_1||^2 + ||\psi_2||^2 + 2||\psi_1||\,||\psi_2||. \tag{3.2.18}$$

Wir müssen also nachweisen, dass

$$\langle \psi_1 \,|\, \psi_2\rangle + \langle \psi_2 \,|\, \psi_1\rangle = 2\,\mathrm{Re}(\langle \psi_1 \,|\, \psi_2\rangle) \overset{?}{\leq} 2||\psi_1||\,||\psi_2||. \tag{3.2.19}$$

Dazu beweisen wir die auch für sich genommen wichtige **Cauchy-Schwarzsche Ungleichung**

$$|\langle \psi_1 \,|\, \psi_2\rangle| \leq ||\psi_1||\,||\psi_2||. \tag{3.2.20}$$

Wir können dabei annehmen, dass $|\psi_1\rangle \neq 0$ und $|\psi_2\rangle \neq 0$, denn andernfalls wären beide Seiten der Ungleichung $= 0$, und somit die Behauptung erfüllt.
Zum Beweis von (3.2.20) setzen wir

$$|\psi\rangle = |\psi_1\rangle - \frac{\langle \psi_2 \,|\, \psi_1\rangle}{||\psi_2||^2}\,|\psi_2\rangle\,. \tag{3.2.21}$$

Dann folgt aus der positiven Definitheit des Skalarprodukts (3.2.10)

$$0 \leq \langle \psi \,|\, \psi\rangle = ||\psi_1||^2 + \frac{|\langle \psi_2 \,|\, \psi_1\rangle|^2}{||\psi_2||^2} - 2\frac{|\langle \psi_2 \,|\, \psi_1\rangle|^2}{||\psi_2||^2} = ||\psi_1||^2 - \frac{|\langle \psi_2 \,|\, \psi_1\rangle|^2}{||\psi_2||^2}, \tag{3.2.22}$$

und daraus folgt durch einige einfache Umformungen (3.2.20). Damit ergibt sich aber sofort auch (3.2.19), denn es gilt

$$\langle \psi_1 \,|\, \psi_2\rangle + \langle \psi_2 \,|\, \psi_1\rangle \leq |\langle \psi_1 \,|\, \psi_2\rangle + \langle \psi_2 \,|\, \psi_1\rangle| \leq 2|\langle \psi_1 \,|\, \psi_2\rangle| \leq 2||\psi_1||\,||\psi_2||. \tag{3.2.23}$$

Damit ist die Dreiecksungleichung (3.2.18) bewiesen.

Wichtige **Beispiele** für konkrete Hilberträume, die in der Quantentheorie eine Rollen spielen, sind der **Hilbertsche Funktionenraum der quadratintegrablen Funktionen** $\mathrm{L}^2(\mathbb{R}^3)$ und der **Hilbertsche Folgenraum der quadratsummierbaren Folgen** ℓ^2.

Der erste Fall $\mathrm{L}^2(\mathbb{R}^3)$ entspricht der Formulierung der Quantentheorie als **Schrödingersche Wellenmechanik** (s. Kap. 2.7). Wie wir dort gesehen haben, werden dabei die quantenmechanischen **Zustände** durch Funktionen $\psi : \mathbb{R}^3 \to \mathbb{C}$ repräsentiert, für die das Integral

$$\langle \psi \,|\, \psi \rangle = ||\psi||^2 = \int_{\mathbb{R}^3} \mathrm{d}^3x \, |\psi(\vec{x})|^2 \tag{3.2.24}$$

existiert. Für zwei solcher Funktionen existiert dann auch das Skalarprodukt

$$\langle \psi_1 \,|\, \psi_2 \rangle := \int_{\mathbb{R}^3} \mathrm{d}^3x \, \psi_1^*(\vec{x})\psi_2(\vec{x}). \tag{3.2.25}$$

Es ist eine einfache *Übungsaufgabe* nachzuweisen, dass die Axiome (3.2.1-3.2.6) und (3.2.8-3.2.11) gelten. Hinsichtlich (3.2.11) müssen wir allerdings vereinbaren, dass wir Funktionen, für die (3.2.24) verschwindet, mit der Funktion $\psi(\vec{x}) \equiv 0$ identifizieren. Das bedeutet anders ausgedrückt, dass zwei Funktionen bereits als gleich angesehen werden, wenn sie sich nur in abzählbar vielen Stellen des $\mathbb{R}^3$ voneinander unterscheiden.

Wie wir unten noch systematische formulieren werden, impliziert diese Realisierbarkeit des Hilbert-Raums als quadratintegrable Funktionen und die physikalische Interpretation im Rahmen der Wellenmechanik, dass im Dirac-Formalismus der **reine Zustand** eines Teilchens durch einen abstrakten Vektor $|\psi\rangle$ im Hilbert-Raum repräsentiert wird. Dabei wird von der Bedeutung der Wellenfunktion als Wahrscheinlichkeitsdichteamplitude für den Ort des Teilchens, d.h. das $P(\vec{x}) = |\psi(\vec{x})|^2$ die Wahrscheinlichkeitsdichte, bei einer Ortsmessung das Teilchen am Ort $\vec{x}$ zu finden, zunächst abstrahiert, d.h. der Zustand $|\psi\rangle$ ist zunächst unabhängig von der Entscheidung, welche Observable gemessen werden soll, als Repräsentant für einen **Präparationsprozess** zu einem Anfangszeitpunkt definiert. Den Zusammenhang mit den Wahrscheinlichkeiten für das Auftreten von Messergebnissen werden wir weiter unten noch formulieren.

Ein anderes Beispiel für eine konkrete Realisierung des Hilbertraums ist der Folgenraum ℓ^2 aus allen Folgen $\psi = (\psi_n)_{n\in\mathbb{N}}$, für die

$$\langle \psi \,|\, \psi \rangle = ||\psi||^2 = \sum_{n=1}^{\infty} |\psi_n|^2 \tag{3.2.26}$$

existiert, und das Skalarprodukt wird durch

$$\langle \psi_1 \,|\, \psi_2 \rangle = \sum_{n=1}^{\infty} \psi_{1n}^* \psi_{2n} \tag{3.2.27}$$

definiert. Die Darstellung der quantenmechanischen Zustandsvektoren als solche Folgen führt zur Formulierung der Quantentheorie als **Heisenberg-Born-Jordansche Matrizenmechanik** (Werner Heisenberg 1901-1976, Ernst Pascual Jordan 1902-1980, Max Born 1882-1970).

Auch der Zusammenhang zwischen diesen verschiedenen Darstellungen der Quantentheorie ist durch den **Diracschen darstellungsfreien Formalismus** im abstrakten Hilbertraum $\mathscr{H}$, wie wir ihn hier zusammenfassen, gegeben. Wir kommen darauf weiter unten noch zurück.

Betrachten wir also wieder den abstrakten Hilbertraum $\mathscr{H}$. Die wichtigste Begriffsbildung, die wir aus dem gegebenen Axiomensystem aufbauen können, ist der der **Konvergenz** und der damit zusammenhängenden **vollständigen Orthonormalsysteme**. Eine Folge von Vektoren $(|\psi_n\rangle)_{n\in\mathbb{N}}$ heißt konvergent gegen einen Vektor $|\psi\rangle$ im Sinne der Hilbertraum-Norm (3.2.14), wenn

$$\lim_{n\to\infty} ||\psi_n - \psi|| = 0 \tag{3.2.28}$$

gilt.
Im Folgenden wird weiterhin postuliert, dass der **Hilbertraum vollständig** ist, d.h. jede **Cauchy-Folge** zu einem Vektor im Hilbertraum konvergiert. Dabei heißt $(|\psi_n\rangle)_{n\in\mathbb{N}}$ definitionsgemäß Cauchy-Folge genau dann, wenn zu jedem $\epsilon > 0$ eine natürliche Zahl $N \in \mathbb{N}$ existiert, so dass für alle $m, n \in \mathbb{N}$ mit $m, n > N$

$$||\psi_m - \psi_n|| < \epsilon \tag{3.2.29}$$

gilt. Wir bemerken ohne Beweis, dass sowohl der Hilbertsche Funktionenraum L^2 als auch der Hilbertsche Folgenraum ℓ^2 vollständig ist (zu solchen eher mathematischen Fragestellungen sei auf [FK07, FK08, FK06] verwiesen).
Eine Folge von Vektoren $(|u_n\rangle)_{n\in\mathbb{N}}$ heißt **Orthonormalsystem**, wenn für alle $m, n \in \mathbb{N}$

$$\langle u_m \,|\, u_n \rangle = \delta_{mn} := \begin{cases} 1 & \text{falls} \quad m = n, \\ 0 & \text{falls} \quad m \neq n \end{cases} \tag{3.2.30}$$

ist. Falls die Reihe

$$|\psi\rangle = \sum_{n=1}^{\infty} \psi_n \,|u_n\rangle \tag{3.2.31}$$

konvergiert, gilt offenbar

$$\langle u_m \,|\, \psi \rangle = \sum_{n=1}^{\infty} \psi_n \,\langle u_m \,|\, u_n \rangle = \sum_{n=1}^{\infty} \psi_n \delta_{mn} = \psi_m. \tag{3.2.32}$$

Ist umgekehrt ein beliebiger Vektor $|\psi\rangle$ gegeben und definieren wir

$$\psi_n = \langle u_n \,|\, \psi \rangle, \tag{3.2.33}$$

so ist die Reihe (3.2.31) konvergent, denn für jede Partialsumme gilt

$$0 \leq \left\| \psi - \sum_{n=1}^{N} \psi_n \,|u_n\rangle \right\|^2 = ||\psi||^2 + \left\| \sum_n \psi_n \,|u_n\rangle \right\|^2 - \left\langle \psi \,\middle|\, \sum_{n=1}^{N} \psi_n u_n \right\rangle - \left\langle \sum_{n=1}^{N} \psi_n u_n \,\middle|\, \psi \right\rangle \tag{3.2.34}$$

Nun ist aber

$$\left\| \sum_n \psi_n \,|u_n\rangle \right\|^2 = \sum_{n,m=1}^{N} \psi_m^* \psi_n \underbrace{\langle u_m \,|\, u_n \rangle}_{=\delta_{mn}} = \sum_{n=1}^{n} |\psi_n|^2 \tag{3.2.35}$$

und

$$\left\langle \psi \,\middle|\, \sum_{n=1}^{N} \psi_n u_n \right\rangle = \sum_{n=1}^{N} \psi_n \,\langle \psi \,|\, u_n \rangle = \sum_{n=1}^{N} |\psi_n|^2 = \left\langle \psi \,\middle|\, \sum_{n=1}^{N} \psi_n u_n \right\rangle^* = \left\langle \sum_{n=1}^{N} \psi_n u_n \,\middle|\, \psi \right\rangle. \tag{3.2.36}$$

Dies in (3.2.35) eingesetzt liefert die **Besselsche Ungleichung**

$$\left\| \sum_{n=1}^{N} \psi_n \,|u_n\rangle \right\|^2 = \sum_{n=1}^{N} |\psi_n|^2 \leq ||\psi||^2. \tag{3.2.37}$$

Die Teilsummenfolge der aus den positiven Gliedern $|\psi_n|^2$ gebildeten Reihe ist also beschränkt und diese folglich konvergent. Nennen wir den entsprechenden Grenzwert

$$\sum_{n=1}^{\infty} |\psi_n|^2 = L^2. \tag{3.2.38}$$

Daraus folgt, dass die Teilsummenfolge

$$|S_k\rangle = \sum_{n=1}^{k} \psi_n |u_n\rangle \tag{3.2.39}$$

eine Cauchyfolge ist, denn zu $\epsilon > 0$ können wir ein $N > 0$ angeben, so dass für alle $n > N$

$$\left| L^2 - \sum_{k=1}^{n} |\psi_k|^2 \right| < \frac{\epsilon^2}{2} \tag{3.2.40}$$

ist. Dann gilt aber für die Teilsummenfolge (3.2.39) für alle $m > n > N$:

$$||S_m - S_n||^2 = \left\| \sum_{k=n+1}^{m} \psi_k |u_k\rangle \right\|^2 = \sum_{k=n+1}^{m} |\psi_k|^2 \leq \left| L^2 - \sum_{k=1}^{n} |\psi_k|^2 \right| + \left| L^2 - \sum_{k=1}^{m} |\psi_k|^2 \right| \overset{(3.2.40)}{\leq} \epsilon^2, \tag{3.2.41}$$

d.h. für alle $m > n > N$ gilt

$$||S_m - S_n|| < \epsilon. \tag{3.2.42}$$

Da wir oben $\epsilon > 0$ beliebig wählen konnten, ist also (3.2.39) eine Cauchyfolge und folglich (wegen der Vollständigkeit des Hilbertraums) die Reihe (3.2.31) mit den Koeffizienten (3.2.33) gegen einen Vektor $|\psi'\rangle$ konvergent:

$$|\psi'\rangle = \sum_{n=1}^{\infty} \psi_n |u_n\rangle = \sum_{n=1}^{\infty} |u_n\rangle \langle u_n | \psi \rangle. \tag{3.2.43}$$

Ein Orthonormalsystem heißt **vollständig**, wenn für jeden Vektor $|\psi\rangle$ die Reihe gegen diesen Vektor konvergiert, wenn also in (3.2.43) $|\psi'\rangle = |\psi\rangle$ ist. Dies können wir symbolisch auch dadurch ausdrücken, dass wir

$$\sum_{n=1}^{\infty} |u_n\rangle \langle u_n| = \mathbb{1} \tag{3.2.44}$$

schreiben. Dabei ist $\mathbb{1}$ der **Einheitsoperator im Hilbertraum** $\mathscr{H}$, d.h. für jeden Vektor $|\psi\rangle \in \mathscr{H}$ gilt

$$\mathbb{1} |\psi\rangle = |\psi\rangle. \tag{3.2.45}$$

Weiter definieren wir das **dyadische Produkt** zweier Vektoren $|\psi_1\rangle$ und $|\psi_2\rangle$ als den Operator $|\psi_1\rangle \langle \psi_2|$, der einen beliebigen Vektor $|\psi_3\rangle$ in den Vektor

$$(|\psi_1\rangle \langle \psi_2|)|\psi_3\rangle = |\psi_1\rangle \langle \psi_2 | \psi_3 \rangle \tag{3.2.46}$$

abbildet. Eine Summe oder unendliche Reihe dyadischer Produkte wirkt entsprechend auf Vektoren gemäß

$$\left(\sum_n |\psi_1\rangle \langle \phi_n| \right) |\psi_2\rangle = \sum_n |\psi_1\rangle \langle \phi_n | \psi_2 \rangle. \tag{3.2.47}$$

Damit ist die Gültigkeit von (3.2.44) in der Tat gleichbedeutend mit

$$|\psi\rangle = \sum_{n=1}^{\infty} |u_n\rangle \langle u_n | \psi \rangle \tag{3.2.48}$$

also mit der **Vollständigkeit** des Orthonormalsystems $(|u\rangle_n)_{n\in\mathbb{N}}$.
Es ist sehr leicht, ein **Beispiel** für ein vollständiges Orthonormalsystem (VONS) im Folgenraum ℓ^2 anzugeben. Offenbar ist ein solches nämlich durch

$$u_n = (0, 0, \ldots, \underbrace{1}_{\text{n-ter Eintrag}}, 0, \ldots) \tag{3.2.49}$$

gegeben. Es gilt jedenfalls

$$\langle u_n \,|\, u_m \rangle = \delta_{nm}, \tag{3.2.50}$$

und ist dann

$$\psi = (\psi_1, \psi_2, \ldots,) \in \ell^2, \tag{3.2.51}$$

so ist offenbar in der Tat

$$\psi_n = \langle u_n \,|\, \psi \rangle \tag{3.2.52}$$

und weiter

$$\sum_{n=1}^{\infty} \psi_n u_n = (\psi_1, \psi_2, \ldots) = \psi. \tag{3.2.53}$$

Also ist (3.2.49) tatsächlich ein VONS.

Wir bemerken noch, dass ein Hilbertraum, in dem es wenigstens ein **VONS aus abzählbar vielen Vektoren** gibt, genauer als **separabler Hilbertraum** bezeichnet wird. Da die Hilberträume ℓ^2 und L^2 separabel sind und die Quantentheorie eines Teilchens in diesen Hilberträumen formulierbar ist, ist also der quantenmechanische Hilbertraum in diesem Fall separabel. Für die praktische Anwendung der Quantentheorie spielt dies allerdings eher eine untergeordnete Rolle.

3.3 Lineare Operatoren im Hilbertraum und Observablen

Mit diesen Betrachtungen haben wir das 1. Postulat für unsere Zwecke hinreichend erläutert. Wenden wir uns also dem 2. Postulat zu. Dazu rekapitulieren wir zunächst einmal den Begriff des **linearen Operators**. Es sei $\mathscr{D} \subseteq \mathscr{H}$ ein Untervektorraum von $\mathscr{H}$. Ein Operator $\mathbf{O} : \mathscr{D} \to \mathscr{H}$ heißt linear, wenn für alle Vektoren $|\psi_1\rangle$, $|\psi_2\rangle \in \mathscr{D}$ und alle Zahlen $\lambda_1, \lambda_2 \in \mathbb{C}$ (für die voraussetzungsgemäß auch der Vektor $\lambda_1 |\psi_1\rangle + \lambda_2 |\psi_2\rangle \in \mathscr{D}$ ist)

$$\mathbf{O}(\lambda_1 |\psi_1\rangle + \lambda_2 |\psi_2\rangle) = \lambda_1 \mathbf{O} |\psi_1\rangle + \lambda_2 \mathbf{O} |\psi_2\rangle \tag{3.3.1}$$

gilt.

Existiert zu dem linearen Operator $\mathbf{O}$ ein linearer Operator $\mathbf{O}^\dagger : \mathscr{D} \to \mathscr{H}$, so dass für alle $|\psi_1\rangle, |\psi_2\rangle \in \mathscr{D}$

$$\langle \psi_1 \,|\, \mathbf{O}\psi_2 \rangle = \left\langle \mathbf{O}^\dagger \psi_1 \,\middle|\, \psi_2 \right\rangle \tag{3.3.2}$$

gilt, so heißt $\mathbf{O}^\dagger$ der zu $\mathbf{O}$ **adjungierte Operator**. Gilt für einen Operator $\mathbf{O}^\dagger = \mathbf{O}$, so heißt er **hermitesch**. Ist für jeden Vektor $|\psi\rangle \in \mathscr{D}$ auch $\mathbf{O}|\psi\rangle \in \mathscr{D}$ und ist $\mathscr{D}$ **dicht im Hilbertraum** $\mathscr{H}$, so heißt $\mathbf{O}$ **selbstadjungiert**. Ein Untervektorraum $\mathscr{D}$ heißt dabei dicht im Hilbertraum, wenn es zu jedem $|\psi\rangle \in \mathscr{H}$ eine Folge $(|\psi_n\rangle)_{n\in\mathbb{N}}$ mit $|\psi_n\rangle \in \mathscr{D}$ gibt, so dass

$$\lim_{n\to\infty} |\psi_n\rangle = |\psi\rangle \tag{3.3.3}$$

gilt. Postulat 2 besagt nun, dass die Observablen im quantenmechanischen Formalismus durch selbstadjungierte Operatoren im Hilbertraum repräsentiert werden.

Ist nun $\mathbf{O}$ ein linearer Operator, so heißt $|u_o\rangle$ **Eigenvektor** von $\mathbf{O}$ zum Eigenwert o, wenn

$$\mathbf{O}|u_o\rangle = o\,|u_o\rangle \tag{3.3.4}$$

ist.

Wie wir oben bei der Betrachtung zur Schrödingerschen Wellenmechanik gesehen haben, müssen wir den Begriff des Eigenvektors hierbei im verallgemeinerten Sinne verstehen. Z.B.

Wie wir anhand dieser Beispiele gesehen haben, besitzen selbstadjungierte Operatoren stets ein **reelles Spektrum**[1], und die dazugehörigen (verallgemeinerten) Eigenvektoren zu verschiedenen Spektralwerten sind zueinander (im verallgemeinerten Sinne) orthogonal. Ohne Beweis nehmen wir an, dass diese (verallgemeinerten) Eigenzustände nach geeigneter Normierung insgesamt ein **vollständiges (verallgemeinertes) Orthogonalsystem** bilden. Im Folgenden schreiben wir bei allgemeinen Betrachtungen die Gleichungen stets für diskrete Eigenwerte. Für verallgemeinerte Eigenwerte gelten die Gleichungen entsprechend im Sinne von Distributionen. Sind z.B. o_1 und o_2 echte voneinander verschiedene reelle Eigenwerte eines selbstadjungierten Operators, so gilt

$$\langle o_1 \,|\, o_2 \rangle = 0 \quad \text{falls} \quad o_1 \neq o_2. \tag{3.3.5}$$

Falls die Werte hingegen zum kontinuierlichen Teil des Spektrums des Operators gehören, gilt (nach entsprechender Normierung)

$$\langle o_1 \,|\, o_2 \rangle = \delta(o_1 - o_2). \tag{3.3.6}$$

Wir wollen die Orthogonalität (3.3.5) der Eigenvektoren und die Realität der Spektralwerte selbstadjungierter Operatoren beweisen. Seien also $|o_k\rangle$ Eigenvektoren des selbstadjungierten Operators $\mathbf{O}$ zu den Eigenwerten o_k. Wir nehmen an, dass all diese Vektoren auf 1 normiert sind, d.h.

$$\langle o_k \,|\, o_k \rangle = 1. \tag{3.3.7}$$

Zum Beweis, dass die Eigenwerte reell sind, verwenden wir die Selbstadjungiertheit des Operators $\mathbf{O}$:

$$o_1 = \langle o_1 \,|\, \mathbf{O} o_1 \rangle = \left\langle \mathbf{O}^\dagger o_1 \,\middle|\, o_1 \right\rangle = \langle \mathbf{O} o_1 \,|\, o_1 \rangle = o_1^* \langle o_1 \,|\, o_1 \rangle = o_1^*. \tag{3.3.8}$$

Das bedeutet aber in der Tat, dass $o_1 \in \mathbb{R}$ ist. Weiter gilt einerseits

$$\langle o_1 \,|\, \mathbf{O} o_2 \rangle = o_2 \langle o_1 \,|\, o_2 \rangle, \tag{3.3.9}$$

denn $|o_2\rangle$ ist Eigenvektor von $\mathbf{O}$ zum Eigenwert o_2. Andererseits gilt aber wegen der Selbstadjungiertheit von $\mathbf{O}$ und der Realität von o_1

$$\langle o_1 \,|\, \mathbf{O} o_2 \rangle = \left\langle \mathbf{O}^\dagger o_1 \,\middle|\, o_2 \right\rangle = \langle \mathbf{O} o_1 \,|\, o_2 \rangle = o_1 \langle o_1 \,|\, o_2 \rangle. \tag{3.3.10}$$

Ziehen wir dieses Resultat von (3.3.9) ab, erhalten wir schließlich

$$(o_2 - o_1) \langle o_1 \,|\, o_2 \rangle = 0. \tag{3.3.11}$$

Falls nun $o_1 \neq o_2$, d.h. $o_2 - o_1 \neq 0$ ist, folgt daraus in der Tat (3.3.6).

Falls es zu einem Eigenwert o_k von $\mathbf{O}$ mehr als einen linear unabhängigen Eigenvektor gibt, bezeichnen wir die Eigenzustände mit $|o_k, \alpha\rangle$, wobei α eine diskrete oder kontinuierliche Variable ist, welche die verschiedenen Eigenvektoren zu diesem gleichen Eigenwert durchnumeriert. Man nennt diesen Eigenwert dann **entartet**. Diese Vektoren spannen den **Eigenraum** des Operators zum Eigenwert o_k auf. Wir können dann mit Hilfe des **Schmittschen Orthonormierungsverfahrens** [CH10] dafür sorgen, dass die Eigenvektoren in diesem Unterraum wieder ein Orthonormalsystem bilden, d.h. im Falle diskreter Werte α

$$\left\langle o_k, \alpha \,\middle|\, o_k, \alpha' \right\rangle = \delta_{\alpha,\alpha'} = \begin{cases} 1 & \text{falls} \quad \alpha = \alpha', \\ 0 & \text{falls} \quad \alpha \neq \alpha' \end{cases} \tag{3.3.12}$$

und für kontinuierliche Werte

$$\left\langle o_k, \alpha \,\middle|\, o_k, \alpha' \right\rangle = \delta(\alpha - \alpha'). \tag{3.3.13}$$

[1] Man nennt die Menge aller echten und verallgemeinerten Eigenwerte eines Operators sein **Spektrum**.

3.4 Verträglichkeit von Observablen

Damit haben wir alle Vorbereitungen getroffen, um die physikalische Bedeutung der Zustandsvektoren, die durch das **Bornsche Postulat (3.1.2)** gegeben ist, auszuarbeiten. Zunächst wollen wir klären, wie eine **vollständige Präparation eines Zustandes** erfolgen kann. Wir werden sehen, dass wir dazu einen **vollständigen Satz kompatibler Observabler** für das betrachtete System festlegen müssen. Wir wollen also zunächst überlegen, wann zwei oder mehrere Observable ein VONS von *gemeinsamen* (verallgemeinerten) Eigenvekoren besitzen.

Nach dem Postulat (3.1.2) besitzt eine Observable A einen wohlbestimmten Wert genau dann, wenn der Systemzustand $|\psi\rangle = \left|a_k, \alpha_j\right\rangle$ ist, wobei a_k ein Eigenwert des der Observablen zugeordneten selbstadjungierten Operators $\mathbf{A}$ und $\left|a_k, \alpha_j\right\rangle$ ein entsprechender Eigenvektor ist. Wir nehmen der Einfachheit halber wieder an, dass die Eigenwerte dieses Operators (und die Werte α_j im Fall der Entartung) nur diskrete Werte annehmen. Wir gehen weiter unten noch auf den Fall des kontinuierlichen Spektrums näher ein. Ist das System nämlich im Zustand $|\psi\rangle$ mit $||\psi|| = 1$ präpariert, so ist die Wahrscheinlichkeit, bei einer Messung von A den Wert a_k zu finden

$$w_\psi(a_k) = \sum_j \left|\left\langle a_k, \alpha_j \,\middle|\, \psi \right\rangle\right|^2 . \tag{3.4.1}$$

Da nun voraussetzungsgemäß die Vollständigkeitsrelation

$$\sum_{jk} \left|a_k, \alpha_j\right\rangle\left\langle a_k, \alpha_j\right| = \mathbb{1} \tag{3.4.2}$$

gilt und folglich

$$\sum_{jk} \left\langle \psi \,\middle|\, a_k, \alpha_j \right\rangle\left\langle a_k, \alpha_j \,\middle|\, \psi \right\rangle = \sum_k w_\psi(a_k) = \langle \psi \,|\, \psi \rangle = 1 \tag{3.4.3}$$

ist und $w_\psi(a_k) \geq 0$ ist, ist also $w_\psi(a_k) = 1$ genau dann, wenn

$$|\psi\rangle = \sum_j c_j \left|a_k, \alpha_j\right\rangle \quad \text{mit} \quad \sum_j |c_j|^2 = 1 \tag{3.4.4}$$

für genau einen Eigenwert a_k ist. Für alle anderen Eigenwerte $a_{k'}$ muss dann außerdem $w_\psi(a_{k'}) = 0$ sein, und nur genau in diesem Falle besitzt die Observable A aufgrund der Präparation des Systems im Zustand $|\psi\rangle$ den wohlbestimmten Wert a_k. Es muss also $|\psi\rangle$ tatsächlich ein Eigenvektor zu diesem Messwert a_k sein.

In dem Fall, dass a_k ein entarteter Eigenwert ist, d.h. wenn es mehrere linear unabhängige Eigenvektoren zu diesem Eigenwert gibt, genügt eine Festlegung der Observablen A auf diesen Wert nicht, um den Zustand eindeutig festzulegen, und wir müssen eine weitere Observable B messen, um den Zustand genauer zu bestimmen. Dabei müssen wir allerdings darauf achten, dass diese Messung mit der Festlegung des Messwertes der Observablen A **kompatibel** ist. Es muss also für jeden möglichen Messwert a_k der Observablen A und jeden möglichen Messwert b_l der Observablen B wenigstens ein **gemeinsamer Eigenvektor** der dazugehörigen Operatoren $\mathbf{A}$ und $\mathbf{B}$ existieren. Nehmen wir also an, dass dies der Fall ist und bezeichnen diese gemeinsamen Eigenvektoren mit $|a_k, b_l, \beta_m\rangle$, wobei β_m wieder die, bei einer möglicherweise immer noch bestehenden Entartung dieser gemeinsamen Eigenwerte, zueinander orthonormiert gewählten Eigenvektoren durchnumeriert. Wir wollen nun herausfinden, was dies für die Operatoren $\mathbf{A}$ und $\mathbf{B}$ bedeutet.

Dazu bemerken wir, dass wir wegen der Vollständigkeit der gemeinsamen Eigenvektoren

$$\begin{aligned} \mathbf{A} &= \sum_{klm} \mathbf{A}\,|a_k, b_l, \beta_m\rangle\,\langle a_k, b_l, \beta_m| = \sum_{klm} a_k\,|a_k, b_l, \beta_m\rangle\,\langle a_k, b_l, \beta_m|\,, \\ \mathbf{B} &= \sum_{klm} \mathbf{B}\,|a_k, b_l, \beta_m\rangle\,\langle a_k, b_l, \beta_m| = \sum_{klm} b_l\,|a_k, b_l, \beta_m\rangle\,\langle a_k, b_l, \beta_m|\,. \end{aligned} \tag{3.4.5}$$

schreiben können. Das bedeutet aber

$$\begin{aligned}\mathbf{AB} &= \sum_{klm}\sum_{k'l'm'} a_k b_{l'} \left|a_k, b_l, \beta_m\right\rangle \underbrace{\left\langle a_k, b_l, \beta_m \middle| a_{k'}, b_{l'}, \beta_{m'}\right\rangle}_{\delta_{kk'}\delta_{ll'}\delta_{mm'}} \left\langle a_{k'}, b_{l'}, \beta_{m'}\right| \\ &= \sum_{klm} a_k b_l \left|a_k, b_l, \beta_m\right\rangle \left\langle a_k b_l, \beta_m\right| = \mathbf{BA}.\end{aligned} \tag{3.4.6}$$

Die Reihenfolge der Operatormultiplikation ist in diesem Fall also unerheblich, d.h. die Operatoren **kommutieren**. Definieren wir also den Kommutator zweier beliebiger Operatoren **A** und **B** vermöge

$$[\mathbf{A}, \mathbf{B}] := \mathbf{AB} - \mathbf{BA}, \tag{3.4.7}$$

bedeutet unsere obige Rechnung, dass es für das Vorliegen eines vollständigen Orthonormalsystems von gemeinsamen Eigenvektoren zweier selbstadjungierter Operatoren *notwendig* ist, dass der Kommutator dieser Operatoren verschwindet:

$$[\mathbf{A}, \mathbf{B}] = 0. \tag{3.4.8}$$

Man kann zeigen, dass diese Bedingung auch *hinreichend* ist.

Um nun also den Zustand des Systems $|\psi\rangle$ vollständig festzulegen, müssen wir die Werte eines **vollständigen Satzes voneinander unabhängiger miteinander kompatibler Observabler** $A, B, C, \ldots$ bestimmen. Dabei heißt ein Satz von Observablen **kompatibel**, wenn die dazugehörigen selbstadjungierten Operatoren untereinander kommutieren, so dass ein vollständiges Orthonormalsystem von simultanen Eigenzuständen dieser Operatoren existiert. Ein Satz solcher kompatibler Observablen heißt **vollständig**, wenn es zu allen möglichen Tupeln von Eigenwerten $(a, b, c, \ldots)$ **genau einen linear unabhängigen simultanen Eigenvektor** gibt. Die **Unabhängigkeit** der Observablen bedeutet, dass nicht ein Operator **Z** in dem Satz als Funktion der übrigen Operatoren geschrieben werden kann, d.h. $\mathbf{Z} \neq f(\mathbf{A}, \mathbf{B}, \ldots, \mathbf{Y})$.

Wir müssen nun noch kurz auf die Besonderheiten eingehen, die sich für eventuell auftretende kontinuierliche Spektralwerte von Operatoren ergeben. Formal sind dann in dem oben zusammengefassten Formalismus zunächst lediglich die Summen durch die entsprechenden Integrale zu ersetzen. Die wesentlichste Änderung liegt eher in der Interpretation der Resultate.

Nehmen wir der Einfachheit halber an, wir hätten ein System von nur einem Freiheitsgrad vorliegen, d.h. schon ein Operator **A** bildet einen vollständigen Satz. Um ein konkretes Beispiel vor Augen zu haben, betrachten wir wieder ein Teilchen, das sich nur entlang der x-Achse bewegt, und wir können dann die Ortskoordinate x als diese Observable wählen. Wir gelangen dann, wie oben ausgeführt, zur Formulierung der Quantenmechanik als Wellenmechanik in der Ortsdarstellung. Wie wir ebenfalls oben gesehen haben, besitzt der dazugehörige Ortsoperator **x** ganz $\mathbb{R}$ als Spektrum. Es liegen hier also keine diskreten Eigenwerte vor. Die dazugehörigen verallgemeinerten Eigenvektoren $|x\rangle$ sind **keine Hilbertraumvektoren** sondern Distributionen über dem entsprechenden dicht definierten Unterraum $\mathscr{D}$, der den Definitionsbereich (und auch Wertebereich) des Operators **x** bildet. Entsprechend können wir das Teilchen **nie präzise lokalisieren**. Ein echter Hilbertraumzustand, der ein Teilchen beschreibt, das sich „in der Nähe" des Ortes x_0 aufhält, wird durch

$$\left|\psi_{x_0}\right\rangle = \int_{\mathbb{R}} \mathrm{d}x\, A_{x_0}(x) |x\rangle \tag{3.4.9}$$

gegeben sein. Die Wellenfunktion ist

$$\psi_{x_0}(x) = \left\langle x \middle| \psi_{x_0}\right\rangle = \int_{\mathbb{R}} \mathrm{d}x' A_{x_0}(x) \underbrace{\left\langle x \middle| x'\right\rangle}_{\delta(x-x')} = A_{x_0}(x). \tag{3.4.10}$$

Dabei muß $A_{x_0} : \mathbb{R} \to \mathbb{C}$ eine quadratintegrable Funktion sein. Damit $\left|\psi_{x_0}\right\rangle$ auf 1 normiert ist, verlangen wir

$$\left\langle \psi_{x_0} \middle| \psi_{x_0}\right\rangle = \int_{\mathbb{R}} \mathrm{d}x\, \left|\psi_{x_0}(x)\right|^2 = \int_{\mathbb{R}} \mathrm{d}x\, \left|A_{x_0}(x)\right|^2 = 1. \tag{3.4.11}$$

Da gemäß dem Bornschen Postulat (3.1.2) die Wahrscheinlichkeitsverteilung, das Teilchen am Ort x zu finden, durch

$$w(x) = \left|\psi_{x_0}(x)\right|^2 \tag{3.4.12}$$

gegeben ist, wird die Lokalisierung „in der Nähe von x_0" lediglich bedeuten, dass diese Wahrscheinlichkeitsverteilung um x_0 stark gepeakt ist. Sie wird aber eine gewisse **Breite** aufweisen. Entsprechend wird der **Erwartungswert** für den Ort

$$\langle x\rangle = \int_{\mathbb{R}} \mathrm{d}x\; x \left|\psi_{x_0}(x)\right|^2 = \int_{\mathbb{R}} \left\langle \psi_{x_0} \middle| \mathbf{x}x \right\rangle \left\langle x \middle| \psi_{x_0} \right\rangle = \left\langle \psi_{x_0} \middle| \mathbf{x}\psi_{x_0} \right\rangle \tag{3.4.13}$$

mit einer gewissen **statistischen Unsicherheit** Δx „in der Nähe von x_0" liegen. Diese Unsicherheit kann, wie in der **Statistik** üblich, durch die **Standardabweichung** definiert werden:

$$\Delta x = \sqrt{\langle x^2\rangle - \langle x\rangle^2} = \sqrt{\langle (x-\langle x\rangle)^2\rangle} = \sqrt{\left\langle \psi_{x_0} \middle| \mathbf{x}^2\psi_{x_0} \right\rangle - \langle x\rangle^2}. \tag{3.4.14}$$

Wir bemerken, dass allgemein Erwartungswerte von Observablen bzw. von Funktionen von Observablen durch

$$\langle f(A_1)\rangle = \langle \psi | f(\mathbf{A}_1) | \psi\rangle \tag{3.4.15}$$

gegeben ist, wenn $|\psi\rangle$ der auf 1 normierte Zustandsvektor des Teilchens ist.
Zum Beweis sei $A_1, A_2, \ldots, A_n$ ein vollständiger Satz kompatibler Observabler und $|a_1, \ldots, a_n\rangle$ ein VONS aus entsprechenden simultanen Eigenvektoren. Dann gilt wegen der Vollständigkeit dieser Orthonormalbasis

$$\begin{aligned}
\langle \psi | f(\mathbf{A}_1) | \psi\rangle &= \sum_{a_1,\ldots,a_n} \langle \psi | f(\mathbf{A}_1) | a_1,\ldots,a_n\rangle \langle a_1,\ldots,a_n | \psi\rangle \\
&= \sum_{a_1,\ldots,a_n} \langle \psi | f(a_1) a_1,\ldots,a_n\rangle \langle a_1,\ldots,a_n | \psi\rangle \\
&= \sum_{a_1,\ldots,a_n} \langle \psi | a_1,\ldots,a_n\rangle \langle a_1,\ldots,a_n | \psi\rangle f(a_1) \\
&= \sum_{a_1,\ldots,a_n} |\psi(a_1,\ldots,a_n)|^2 f(a_1) \\
&= \sum_{a_1,\ldots,a_n} P(a_1,\ldots,a_n) f(a_1) = \langle f(A_1)\rangle .
\end{aligned} \tag{3.4.16}$$

3.5 Die Heisenbergsche Unschärferelation

Eine wichtige Folgerung aus der statistischen Interpretation des quantentheoretischen Zustandsbegriffs über die Bornsche Regel (3.1.2) ist die **Heisenbergsche Unschärferelation**. Seien dazu A und B zwei Observablen, die zueinander kompatibel oder inkompatibel sein können, und $|\psi\rangle$ irgendein Zustand[2] des Systems. Dann gibt es eine **untere Schranke** für das Unschärfeprodukt $\Delta A \Delta B$.
Heisenberg ist auf diese Folgerung anhand des Beispiels von Ort und Impuls gekommen. Haben wir nämlich, wie in dem gerade besprochenen Beispiel der Lokalisierung eines Teilchens in der Nähe des Ortes x_0, eine Wellenfunktion $\psi_{x_0}(x)$, die scharf um diesen Ort gepeakt ist, so wird die entsprechende Impulsverteilung durch die Fouriertransformierte der Wellenfunktion gegeben sein (vgl. (2.7.90)). Die daraus resultierende Impulsverteilung wird aber desto breiter und entsprechend Δp desto größer sein je schärfer die Ortsverteilung (also je kleiner Δx) ist.

[2] Es ist hier wichtig, dass es sich um einen „echten" Hilbertraumvektor handelt und nicht um einen verallgemeinerten Eigenvektor zu einem Wert im kontinuierlichen Spektrum eines Operators!

Der bis jetzt entwickelte quantentheoretische Formalismus lässt bereits eine Quantifizierung der Schranke für ΔA und ΔB zu. Um diese zu finden, definieren wir hilfsweise die neuen Operatoren

$$\mathbf{A}' = \mathbf{A} - \langle A \rangle \, \mathbb{1}, \quad \mathbf{B}' = \mathbf{B} - \langle B \rangle \, \mathbb{1}. \tag{3.5.1}$$

Die Erwartungswerte sind dabei bzgl. des Zustandes $|\psi\rangle$, in dem das Teilchen präpariert ist, zu bilden. Dann gilt nämlich

$$\langle A' \rangle = \langle B' \rangle = 0, \quad \Delta A^2 = \langle A'^2 \rangle, \quad \Delta B^2 = \langle B'^2 \rangle, \quad [\mathbf{A}', \mathbf{B}'] = [\mathbf{A}, \mathbf{B}]. \tag{3.5.2}$$

Es sei weiter $\lambda \in \mathbb{R}$. Dann definieren wir das quadratische Polynom

$$P(\lambda) = \langle (\mathbf{A}' + \mathrm{i}\lambda \mathbf{B}')\psi \,|\, (\mathbf{A}' + \mathrm{i}\lambda \mathbf{B}')\psi \rangle = \langle \psi \,|\, (\mathbf{A}' - \mathrm{i}\lambda \mathbf{B}')(\mathbf{A}' + \mathrm{i}\lambda \mathbf{B}')\psi \rangle. \tag{3.5.3}$$

Ausmultiplizieren des Operatorprodukts liefert dann unter Verwendung von (3.5.2)

$$P(\lambda) = \Delta A^2 + \lambda^2 \Delta B^2 + \lambda \langle \psi \,|\, \mathrm{i}[\mathbf{A}, \mathbf{B}] \, \psi \rangle. \tag{3.5.4}$$

Da $\mathbf{A}$ und $\mathbf{B}$ selbstadjungiert sind, gilt

$$\{\mathrm{i}[\mathbf{A}, \mathbf{B}]\}^\dagger = -\mathrm{i}(\mathbf{AB} - \mathbf{BA})^\dagger = -\mathrm{i}(\mathbf{B}^\dagger \mathbf{A}^\dagger - \mathbf{A}^\dagger \mathbf{B}^\dagger) = -\mathrm{i}(\mathbf{BA} - \mathbf{AB}) = +\mathrm{i}[\mathbf{A}, \mathbf{B}]. \tag{3.5.5}$$

Es ist also auch $\mathrm{i}[\mathbf{A}, \mathbf{B}]$ selbstadjungiert und folglich der Koeffizient von λ in (3.5.4) reell:

$$\langle \psi \,|\, \mathrm{i}[\mathbf{A}, \mathbf{B}] \, \psi \rangle \in \mathbb{R}. \tag{3.5.6}$$

Das quadratische Polynom (3.5.4) ist also reell und wegen der positiven Definitheit des Skalarprodukts gilt für alle $\lambda \in \mathbb{R}$

$$P(\lambda) \geq 0. \tag{3.5.7}$$

Demnach besitzt dieses Polynom entweder eine einzige doppelte reelle Nullstelle oder zwei verschiedene zueinander komplex konjugierte Nullstellen. Nach der Lösungsformel für quadratische Gleichungen muss also für die Diskriminante des Polynoms

$$\frac{1}{4} \langle \psi \,|\, \mathrm{i}[\mathbf{A}, \mathbf{B}] \, \psi \rangle^2 - \Delta A^2 \Delta B^2 \leq 0 \tag{3.5.8}$$

oder

$$\Delta A \, \Delta B \geq \frac{1}{2} |\langle \psi \,|\, \mathrm{i}[\mathbf{A}, \mathbf{B}] \, \psi \rangle| \tag{3.5.9}$$

gelten. Dies ist die **Heisenbergsche Unschärferelation** für irgendwelche Observablen A und B. Sind insbesondere A und B kompatibel, können also deren Werte simultan scharf festgelegt werden, so kommutieren die entsprechenden Operatoren $\mathbf{A}$ und $\mathbf{B}$, und die rechte Seite der Ungleichung verschwindet. Die Ungleichung ergibt dann keine echte Einschränkung für das Produkt der Standardabweichungen.

Betrachten wir die Unschärferelation insbesondere für **Ort und Impuls**. Aus der konkreten Darstellung der entsprechenden Operatoren für die Orts- und Impulskomponenten im Ortsraum (2.7.42) ergeben sich die Kommutatorrelationen (**Heisenberg-Algebra**)

$$\left[\mathbf{x}_j, \mathbf{x}_k\right] = \left[\mathbf{p}_j, \mathbf{p}_k\right] = 0, \quad \left[\mathbf{x}_j, \mathbf{p}_k\right] = \mathrm{i}\hbar \delta_{jk}. \tag{3.5.10}$$

Dies in (3.5.9) eingesetzt ergibt die bekannte Heisenbergsche Unschärferelation für Ort und Impuls

$$\Delta x_j \Delta p_k \geq \frac{\hbar}{2} \delta_{jk}. \tag{3.5.11}$$

Es können also nur Komponenten von Ort und Impuls in zueinander senkrechten Richtungen gleichzeitig scharf bestimmt sein. Ein vollständiger Satz kompatibler Observabler kann in diesem Falle als die drei Orts- oder die drei Impulskomponenten oder z.B. x_1 und p_2, p_3 etc. gewählt werden.

3.6 Zeitentwicklung des Zustands

Die Zeitentwicklung der Wellenunktion $\psi(t,\vec{x})$, ist gemäß unserer heuristischen Betrachtungen im vorigen Kapitel durch die **Schrödinger-Gleichung**

$$\mathrm{i}\hbar\partial_t\psi(t,\vec{x}) = \mathbf{H}\psi(t,\vec{x}), \tag{3.6.1}$$

gegeben, wobei $\mathbf{H} = \mathbf{H}(\vec{\mathbf{x}},\vec{\mathbf{p}})$ der selbstadjungierte **Hamilton-Operator** ist, der die Energie des Systems beschreibt. Für die Bewegung eines Teilchens in einem Kräftepotential $V(\vec{x})$ gilt in Analogie zur klassischen Mechanik

$$\mathbf{H} = \frac{1}{2m}\vec{\mathbf{p}}^2 + V(\vec{\mathbf{x}}). \tag{3.6.2}$$

Dies legt es nun nahe, im abstrakten Formalismus die Zeitentwicklung des Systems durch die entsprechende Schrödinger-Gleichung für den abstrakten Zustandsvektor $|\psi(t)\rangle$ zu postulieren:

$$\mathrm{i}\hbar\partial_t|\psi(t)\rangle = \mathbf{H}|\psi(t)\rangle. \tag{3.6.3}$$

Daraus ergibt sich sofort eine formale Lösung. Behandeln wir nämlich $\mathbf{H}$ wie eine gewöhnliche Zahl, erhalten wir

$$|\psi(t)\rangle = \exp\left(-\frac{\mathrm{i}}{\hbar}\mathbf{H}t\right)|\psi(0)\rangle. \tag{3.6.4}$$

D.h. die Schrödinger-Gleichung bestimmt den Zustand des Systems zur Zeit t, also $|\psi(t)\rangle$, wenn der Zustand zur Anfangszeit $t = 0$, zu der das Teilchen im Zustand $|\psi(0)\rangle$ präpariert wurde, gegeben ist.

Die **Operatorexponentialfunktion** spielt nun eine herausragende Rolle im Formalismus der Quantentheorie. Dazu bemerken wir zuerst, dass sie einfach durch die operatorwertige Taylor-Reihe der Exponentialfunktion definiert ist. Für einen beliebigen linearen Operator $\mathbf{A}$ ist also

$$\exp\mathbf{A} = \sum_{j=0}^{\infty}\frac{1}{j!}\mathbf{A}^j. \tag{3.6.5}$$

Es ist damit klar, dass für die Adjunktion

$$[\exp\mathbf{A}]^\dagger = \exp\mathbf{A}^\dagger \tag{3.6.6}$$

gilt. Außerdem ergibt sich wie für die Exponentialfunktion von Zahlen, dass für **kommutierende Observable**

$$\exp\mathbf{A}\exp\mathbf{B} = \exp(\mathbf{A}+\mathbf{B}) \quad \text{falls} \quad [\mathbf{A},\mathbf{B}] = 0 \tag{3.6.7}$$

gilt.

Betrachten wir nun den oben eingeführten **Zeitentwicklungsoperator**

$$\mathbf{U}(t) = \exp\left(-\frac{\mathrm{i}}{\hbar}\mathbf{H}t\right) \tag{3.6.8}$$

für den Zustand genauer, so folgt mit (3.6.6)

$$\mathbf{U}^\dagger(t) = \exp\left(+\frac{\mathrm{i}}{\hbar}\mathbf{H}t\right), \tag{3.6.9}$$

wobei wir die Selbstadjungiertheit des Hamilton-Operators verwendet haben. Mit (3.6.7) folgt daraus aber sofort, dass

$$\mathbf{U}^\dagger(t)\mathbf{U}(t) = \mathbf{U}(t)\mathbf{U}^\dagger(t) = \mathbb{1} \Rightarrow \mathbf{U}^\dagger(t) = \mathbf{U}^{-1}(t) \tag{3.6.10}$$

ist. Man nennt einen Operator $\mathbf{U}$, für den $\mathbf{U}^\dagger = \mathbf{U}^{-1}$ ist, **unitär**. Die Zeitentwicklung des Zustandsvektors ist also gemäß (3.6.4) eine sog. **unitäre Transformation**

$$|\psi(t)\rangle = \mathbf{U}(t)|\psi(t)\rangle\,. \tag{3.6.11}$$

Daraus ergibt sich nun, dass die Schrödingergleichung die Normierung des Zustandes ungeändert lässt, denn es folgt

$$\langle\psi(t)|\,\psi(t)\rangle = \langle\mathbf{U}(t)\psi_0|\,\mathbf{U}(t)\psi_0\rangle = \left\langle\psi_0\middle|\,\mathbf{U}^\dagger(t)\mathbf{U}(t)\psi_0\right\rangle = \langle\psi_0|\,\psi_0\rangle = 1, \tag{3.6.12}$$

wobei wir vorausgesetzt haben, dass wir den Anfangszustandsvektor auf 1 normiert haben.

Wir zeigen nun noch, inwiefern das 4. Postulat (3.1.3) mit der Schrödinger-Gleichung für den abstrakten Zustandsvektor verträglich ist. Dazu betrachten wir eine beliebige Observable O. Für den Erwartungswert gilt gemäß (3.4.16) und (3.6.11)

$$\langle O\rangle(t) = \langle\psi(t)|\,\mathbf{O}\psi(t)\rangle = \langle\mathbf{U}(t)\psi_0|\,\mathbf{O}\mathbf{U}(t)\psi_0\rangle = \left\langle\psi_0\middle|\,\mathbf{U}^\dagger(t)\mathbf{O}\mathbf{U}(t)\psi_0\right\rangle. \tag{3.6.13}$$

Bilden wir nun die Zeitableitung, folgt mit (3.6.8) und (3.6.9)

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}t}\langle O\rangle(t) &= \left\langle\psi_0\middle|\left[\dot{\mathbf{U}}^\dagger(t)\mathbf{O}\mathbf{U}(t) + \mathbf{U}^\dagger(t)\mathbf{O}\dot{\mathbf{U}}(t)\right]\psi_0\right\rangle \\
&= \left\langle\psi_0\middle|\left[\frac{i}{\hbar}\mathbf{U}^\dagger(t)\mathbf{H}\mathbf{O}\mathbf{U}(t) - \frac{\mathrm{i}}{\hbar}\mathbf{U}^\dagger(t)\mathbf{O}\mathbf{H}\mathbf{U}(t)\right]\psi_0\right\rangle \\
&= \left\langle\psi_0\middle|\left[\frac{-\mathrm{i}}{\hbar}\mathbf{U}^\dagger(t)[\mathbf{O},\mathbf{H}]\right]\mathbf{U}(t)\psi_0\right\rangle \\
&= \left\langle\mathbf{U}(t)\psi_0\middle|\left(\frac{1}{\mathrm{i}\hbar}[\mathbf{O},\mathbf{H}]\right)\mathbf{U}(t)\psi_0\right\rangle \\
&= \left\langle\psi(t)\middle|\left(\frac{1}{\mathrm{i}\hbar}[\mathbf{O},\mathbf{H}]\right)\psi(t)\right\rangle \\
&= \left\langle\mathring{\mathbf{O}}\right\rangle.
\end{aligned} \tag{3.6.14}$$

Damit ist die Intepretation des Operators

$$\mathring{\mathbf{O}} = \frac{1}{\mathrm{i}\hbar}[\mathbf{O},\mathbf{H}] \tag{3.6.15}$$

als der selbstadjungierte Operator, der die Zeitableitung $\dot{O}$ der Observablen O beschreibt, gerechtfertigt.

Mittels der formalen Lösung (3.6.4) ist es nun weiter sehr einfach, die besondere Bedeutung von **Energie-Eigenvektoren** zu diskutieren. Nehmen wir dazu an, es sei $|E,\alpha\rangle$ ein VONS von Energie-Eigenvektoren, wobei α die entarteten Zustände durchnumeriert. Wir nehmen wieder an, dass sowohl E als auch α diskret sind. Für kontinuierliche verallgemeinerte Eigenwerte muss im Folgenden wieder nur die Summe durch entsprechende Integrale ersetzt werden. Sei nun also das Teilchen anfangs in einem Energie-Eigenzustand $|\psi_0\rangle = |E,\alpha\rangle$ präpariert. Dann folgt aus (3.6.4) sofort

$$|\psi(t)\rangle\rangle = \exp\left(-\frac{\mathrm{i}}{\hbar}\mathbf{H}t\right)|\psi_0\rangle = \exp\left(-\frac{\mathrm{i}}{\hbar}Et\right)|\psi_0\rangle\,, \tag{3.6.16}$$

d.h. der Zustandsvektor ändert sich bis auf einen Phasenfaktor nicht. Physikalische Bedeutung besitzen nun allerdings nur die Wahrscheinlichkeiten für das Auftreten von Messwerten eines vollständigen Satzes von unabhängigen kompatiblen Observablen $(A_1,\ldots,A_n)$, d.h.

$$P(t,a_1,\ldots,a_n) = |\langle a_1,\ldots,a_n\,|\,\psi(t)\rangle|^2 = |\langle a_1,\ldots,a_n\,|\,\psi_0\rangle|^2\,. \tag{3.6.17}$$

Dieses Beispiel zeigt allgemein, dass Phasenfaktoren für die physikalische Bedeutung des Zustandsvektors keine Bedeutung haben und im Besonderen, dass Energieeigenzustände die **stationären Zustände** des Systems beschreiben, d.h. ein in einem Energieeigenzustand präpariertes Teilchen verbleibt für alle späteren Zeiten in diesem Energieeigenzustand.

Wir bemerken in diesem Zusammenhang, dass wir streng genommen den Zustandsbegriff präzisieren müssen, wie bereits in Abschnitt 3.1.1 erörtert. Die obige Betrachtung zeigt nämlich, dass normierte Hilbert-Raumvektoren, die sich nur um einen Phasenfaktor $\exp(\mathrm{i}\varphi)$ mit $\varphi \in \mathbb{R}$ unterscheiden, denselben Zustand beschreiben. Wir können dem Rechnung tragen, indem wir als eindeutige Charakterisierung eines (reinen) Zustands nicht einen normierten Vektor im Hilbert-Raum sondern den Projektionsoperator

$$\mathbf{P}_\psi(t) = |\psi(t)\rangle\langle\psi(t)| \tag{3.6.18}$$

betrachten. Die Bornsche Regel ist dann so zu formulieren, dass die Wahrscheinlichkeit, dass bei einer Messung des vollständigen unabhängigen Satzes kompatibler Observablen $(A_1,\ldots,A_n)$ die Wahrscheinlichkeit dafür, die simultanen Messwerte $(a_1,\ldots,a_n)$ zu erhalten,

$$P(t,a_1,\ldots,a_n) = \left\langle a_1,\ldots,a_n \middle| \mathbf{P}_\psi(t) \middle| a_1,\ldots,a_n \right\rangle \tag{3.6.19}$$

ist.

3.7 Diracs Zugang zur Quantenmechanik

Bevor wir uns den wichtigsten, exakt lösbaren Problemen der Quantenmechanik zuwenden, wollen wir in diesem Abschnitt noch Diracs heuristische Konstruktion der Quantentheorie betrachten. Dieser ist zwar wesentlich abstrakter als die Argumente, die Schrödinger zur Formulierung der Quantenmechanik als Wellenmechanik geführt haben, zeigt aber auch, wie man die Quantenmechanik aus den Symmetrieprinzipien des **Noether-Theorems** der klassischen Mechanik konstruieren kann. Diese auf Symmetrien begründete Heuristik hat sich für die gesamte moderne Physik als äußerst fruchtbar erwiesen. So kann u.a. auch die **relativistische Quantenfeldtheorie**, die die Grundlage der gesamten **Elementarteilchenphysik** bildet, aus reinen Symmetriebetrachtungen begründet werden, wobei allerdings die relativistische Raum-Zeitbeschreibung, die wir im zweiten Teil dieses Manuskripts ausführlich besprechen werden, verwendet werden muss. Hier begnügen wir uns mit der Newtonschen Raumzeitbeschreibung und der entsprechenden Symmetrie unter Raum-Zeit-Translationen, Drehungen und Galilei-Transformationen.

3.7.1 Erinnerung an die Hamiltonsche Mechanik

Diracs grundlegende Idee ist es, den oben eingeführten abstrakten Hilbert-Raumformalismus mit den Postulaten (1-4) in Abschnitt 3.1 zugrunde zu legen. Als einzige weitere heuristische Regel verwendet Dirac dann die **Hamiltonsche Formulierung** der klassischen Mechanik mit Poisson-Klammern und insbesondere das Noether-Theorems. Um das Prinzip zu verstehen, genügt wieder die Betrachtung eines Teilchens in einem äußeren Potential.

Im Hamilton-Formalismus (s. [Hee18b]) beginnt man mit der Aufstellung der **Hamilton-Funktion**, also der Gesamtenergie, als Funktion der (generalisierten) Koordinaten und der entsprechenden kanonisch konjugierten Impulse. Wir verwenden hier der Einfachheit halber kartesische Koordinaten:

$$H(\vec{x},\vec{p}) = \frac{1}{2m}\vec{p}^2 + V(\vec{x}). \tag{3.7.1}$$

Die Bewegungsgleichungen ergeben sich dann unter Verwendung des Hamiltonschen Prinzips der kleinsten Wirkung als die **Hamiltonschen Kanonischen Gleichungen**

$$\dot{x}_k = \frac{\partial H}{\partial p_k}, \quad \dot{p}_k = -\frac{\partial H}{\partial q_k}. \tag{3.7.2}$$

In der Tat liefert das die richtigen Bewegungsgleichungen, denn mit (3.7.1) ergibt sich

$$\dot{x}_k = \frac{1}{m} p_k, \quad \dot{p}_k = m\ddot{x}_k = -\frac{\partial V}{\partial x_k}, \tag{3.7.3}$$

also die korrekte Newtonsche Bewegungsgleichung mit der Kraft $\vec{F} = -\vec{\nabla} V$.
Betrachten wir nun eine allgemeine Funktion $f(\vec{x}, \vec{p})$ der Phasenraumkoordinaten $(\vec{x}, \vec{p})$, folgt für die Zeitableitung für eine beliebige Lösung der Bewegungsgleichungen (3.7.2)

$$\frac{\mathrm{d}}{\mathrm{d}t} f(\vec{x}, \vec{p}) = \dot{x}_k \frac{\partial f}{\partial x_k} + \dot{p}_k \frac{\partial f}{\partial p_k} = \frac{\partial f}{\partial x_k}\frac{\partial H}{\partial p_k} - \frac{\partial H}{\partial x_k}\frac{\partial f}{\partial p_k} = \{f, H\}_{\mathrm{pb}}, \tag{3.7.4}$$

wobei die **Poisson-Klammer** für zwei Phasenraumfunktion $f(\vec{x}, \vec{p})$ und $g(\vec{x}, \vec{p})$ durch

$$\{f, g\}_{\mathrm{pb}} = \frac{\partial f}{\partial x_k}\frac{\partial g}{\partial p_k} - \frac{\partial g}{\partial x_k}\frac{\partial f}{\partial p_k} \tag{3.7.5}$$

definiert ist.
Weiter haben wir gesehen, dass eine **kanonische Transformation** mit einer erzeugenden Funktion $G(\vec{x}, \vec{p}, t)$ durch die Differentialgleichung

$$\begin{aligned} &\frac{\mathrm{d}}{\mathrm{d}\lambda} x_k(\lambda) = \{x_k, G\}_{\mathrm{pb}} = \frac{\partial G}{\partial p_k}, \quad \frac{\mathrm{d}}{\mathrm{d}\lambda} p_k(\lambda) = \{x_k, G\}_{\mathrm{pb}} = -\frac{\partial G}{\partial x_k}, \\ &H'(X_k, P_k, t) = H(x_k, p_k, t) + \partial_t G, \end{aligned} \tag{3.7.6}$$

mit den „Anfangsbedingungen“ $x_k(0) = x_k$, $p_k(0) = p_k$, wobei die $X_k = x_k(\lambda)$ und $P_k = p_k(\lambda)$ für ein festes λ als neue Phasenraumvariablen interpretiert werden, erzeugt wird.
Eine Symmetrietransformation liegt dann vor, wenn

$$H'(X_k, P_k, t) - H(X_k, P_k, t) = H(x_k, p_k, t) - H(X_k, P_k, t) + \partial_t G = 0 \tag{3.7.7}$$

ist. Leitet man diese Bedingung nach λ ab, folgt

$$-\{H, G\}_{\mathrm{pb}} + \partial_t G = \{G, H\}_{\mathrm{pb}} + \partial_t G = 0. \tag{3.7.8}$$

Für die Lösungen der Bewegungsgleichungen ist also $G(x_k, p_k, t) = \text{const}$. Folglich ist die Erzeugende einer kanonischen Symmetrietransformation stets eine Erhaltungsgröße des Systems und umgekehrt.

3.7.2 Quantentheorie für ein Teilchen mit Spin 0

Dirac postuliert nun einfach, dass die Quantentheorie durch die in Abschnitt 3.1 angegebenen Postulate definiert ist und dass die Observablen-Operatoren die Kommutatorregeln in Analogie zu den entsprechenden Poisson-Klammern der klassischen Theorie erfüllen. Dabei legt der Vergleich der klassischen Regel für die Zeitableitung einer beliebigen Phasenraumfunktion mit Posulat 4 der Quantentheorie nahe, dass der klassischen Poisson-Klammer $\{A, B\}_{\mathrm{pb}}$ für beliebige Phasenraumfunktionen A und B der Kommutator $[\mathbf{A}, \mathbf{B}]/(\mathrm{i}\hbar)$ entspricht.
Insbesondere ergeben sich für die kartesischen Orts- und Impuls-Operatoren die Kommutatorregeln

$$\left[\mathbf{x}_j, \mathbf{x}_k\right] = \left[\mathbf{p}_j, \mathbf{p}_k\right] = 0, \quad \left[\mathbf{x}_j, \mathbf{p}_k\right] = \mathrm{i}\hbar\delta_{jk}. \tag{3.7.9}$$

Daraus ergibt sich auch, dass die Impulse die Generatoren für räumliche Translationen sind. Für eine Translation um $a\vec{n}$ ergibt sich dies aus

$$\frac{\mathrm{d}}{\mathrm{d}a}\vec{\mathbf{X}}(a) = -\frac{\mathrm{i}}{\hbar}[\mathbf{X}(a), \vec{n}\cdot\vec{\mathbf{p}}] = \vec{n}, \quad \vec{\mathbf{X}}(a=0) = \vec{\mathbf{x}}. \tag{3.7.10}$$

Integriert ergibt sich nämlich einerseits

$$\vec{\mathbf{X}}(a) = \vec{\mathbf{x}} + a\vec{n} \tag{3.7.11}$$

Andererseits ist aber auch

$$\vec{\mathbf{X}}(a) = \exp(\mathrm{i}a\vec{n}\cdot\vec{\mathbf{p}}/\hbar)\vec{\mathbf{x}}\exp(-\mathrm{i}a\vec{n}\cdot\vec{\mathbf{p}}/\hbar), \tag{3.7.12}$$

denn die Ableitung dieses Ausdrucks ergibt

$$\frac{\mathrm{d}}{\mathrm{d}a}\vec{\mathbf{X}}(a) = \exp(\mathrm{i}a\vec{n}\cdot\vec{\mathbf{p}}/\hbar)\frac{-\mathrm{i}}{\hbar}[\vec{\mathbf{x}}, \vec{n}\cdot\vec{\mathbf{p}}]\exp(-\mathrm{i}a\vec{n}\cdot\vec{\mathbf{p}}/\hbar) = \vec{n}. \tag{3.7.13}$$

Die Translation um $\vec{a} = a\vec{n}$ ist also durch die **unitäre Transformation**

$$\mathbf{T}(\vec{a}) = \exp\left(-\frac{\mathrm{i}}{\hbar}\vec{a}\cdot\vec{\mathbf{p}}\right) \tag{3.7.14}$$

gegeben. D.h. ist $|\psi\rangle$ der Zustand eines Teilchens, dann ist $\mathbf{T}(\vec{a})|\psi\rangle$ der Zustand eines Teilchens, das um $\vec{a}$ verschoben ist.
Wir können auch die verallgemeinerten Eigenzustände von $\vec{\mathbf{x}}$ dadurch definieren, dass wir

$$|\vec{x}\rangle = \mathbf{T}(\vec{x})|\vec{x}=0\rangle \tag{3.7.15}$$

setzen. Denn offenbar gilt

$$\vec{\mathbf{x}}\mathbf{T}(\vec{x})|\vec{x}=0\rangle = \mathbf{T}(\vec{x})\mathbf{T}^{-1}(\vec{x})\vec{\mathbf{x}}\mathbf{T}(\vec{x})|\vec{x}=0\rangle = \mathbf{T}(\vec{x})(\vec{\mathbf{x}}+\vec{x})|\vec{x}=0\rangle = \mathbf{T}(\vec{x})\vec{x}|\vec{x}=0\rangle = \vec{x}\mathbf{T}|\vec{x}=0\rangle. \tag{3.7.16}$$

Es ist also in der Tat (3.7.15) ein Ortseigenvektor zum Eigenwert $\vec{x}$.
Ist $|\vec{p}\rangle$ nun ein verallgemeinerter Impulseigenvektor, so folgt

$$\begin{aligned}\langle\vec{p}\,|\,\vec{x}\rangle &= \langle\vec{p}\,|\mathbf{T}(\vec{x})|\,\vec{x}=0\rangle \\ &= \left\langle\mathbf{T}^{\dagger}(\vec{x})\vec{p}\,\middle|\,\vec{x}=0\right\rangle \\ &= \langle\exp(\mathrm{i}\vec{p}\cdot\vec{x}/\hbar)\vec{p}\,|\,\vec{x}=0\rangle \\ &= \exp(-\mathrm{i}\vec{p}\cdot\vec{x}/\hbar)\langle\vec{p}\,|\,\vec{x}=0\rangle \\ &= N(\vec{p})\exp(-\mathrm{i}\vec{p}\cdot\vec{x}/\hbar).\end{aligned} \tag{3.7.17}$$

Der Normierungsfaktor folgt nun daraus, dass wir die verallgemeinerten Eigenzustände für Ort und Impuls jeweils auf eine δ-Distribution normieren wollen und dass diese Eigenzustände vollständig sein sollen. Es ist also

$$\begin{aligned}\langle\vec{p}\,'\,|\,\vec{p}\rangle = \delta^{(3)}(\vec{p}\,'-\vec{p}) &= \int_{\mathbb{R}^3}\mathrm{d}^3x\,\langle\vec{p}\,'\,|\,\vec{x}\rangle\langle\vec{x}\,|\,\vec{p}\rangle \\ &= \int_{\mathbb{R}^3}\mathrm{d}^3x N(\vec{p}\,')N^*(\vec{p})\exp\left[\frac{\mathrm{i}}{\hbar}(\vec{p}-\vec{p}\,')\cdot\vec{x}\right] \\ &= |N(\vec{p})|^2(2\pi\hbar)^3\delta^{(3)}(\vec{p}-\vec{p}\,').\end{aligned} \tag{3.7.18}$$

Bis auf eine irrelevante Phase können wir also $N(\vec{p}) = 1/(2\pi\hbar)^{3/2}$ setzen. Damit ergibt sich für die Impulseigenfunktion in der Ortsdarstellung

$$u_{\vec{p}}(\vec{x}) = \langle \vec{x} \,|\, \vec{p} \,\rangle = \langle \vec{p} \,|\, \vec{x} \,\rangle^* = \frac{1}{(2\pi\hbar)^{3/2}} \exp\left(\frac{\mathrm{i}}{\hbar}\vec{x}\cdot\vec{p}\right). \tag{3.7.19}$$

Wir können mit (3.7.15) auch leicht die Wirkung des Impulsoperators auf den Ortseigenvektor berechnen:

$$\vec{\mathbf{p}}\,|\vec{x}\rangle = \vec{\mathbf{p}} \exp\left(-\frac{\mathrm{i}}{\hbar}\vec{x}\cdot\vec{\mathbf{p}}\right)|\vec{x}=0\rangle = \mathrm{i}\hbar\vec{\nabla}_x \exp\left(-\frac{\mathrm{i}}{\hbar}\vec{x}\cdot\vec{\mathbf{p}}\right)|\vec{x}=0\rangle = \mathrm{i}\hbar\vec{\nabla}\,|\vec{x}\rangle\,. \tag{3.7.20}$$

Durch Hermitesche Adjunktion folgt daraus

$$\langle\vec{x}|\,\vec{\mathbf{p}}^\dagger = \langle\vec{x}|\,\vec{\mathbf{p}} = -\mathrm{i}\hbar\vec{\nabla}\,\langle\vec{x}| \tag{3.7.21}$$

und damit für den Impulsoperator in der Ortsdarstellung der Wellenfunktion

$$\vec{\mathbf{p}}\psi(\vec{x}) = \langle x \,|\, \mathbf{p}\psi \,\rangle = -\mathrm{i}\hbar\vec{\nabla}\,\langle \vec{x} \,|\, \psi \,\rangle = -\mathrm{i}\hbar\psi(\vec{x}). \tag{3.7.22}$$

Es lässt sich also die Schrödingersche Wellenmechanik aus der abstrakten Dirac-Formulierung der Quantenmechanik allein aufgrund der Kommutatorregeln zwischen den Observablen-Operatoren konstruieren. Bezeichnen wir nun mit $\left|\vec{p}\right\rangle$ die verallgemeinerten **Impulseigenvektoren**, folgt

$$\begin{aligned}
\vec{\mathbf{x}}\,\left|\vec{p}\right\rangle &= \int_{\mathbb{R}^3} \mathrm{d}^3x\,|\vec{x}\rangle\,\vec{x}\,\left\langle \vec{x} \,\middle|\, \vec{p} \,\right\rangle \\
&= \int_{\mathbb{R}^3} \mathrm{d}^3x\,|\vec{x}\rangle\,\vec{x}\,\frac{1}{(2\pi\hbar)^{3/2}}\exp(\mathrm{i}\vec{x}\cdot\vec{p}/\hbar) \\
&= \int_{\mathbb{R}^3} \mathrm{d}^3x\,|\vec{x}\rangle \int_{\mathbb{R}^3} \mathrm{d}^3x\,|\vec{x}\rangle\,(-\mathrm{i}\hbar)\vec{\nabla}_p\,\frac{1}{(2\pi\hbar)^{3/2}}\exp(\mathrm{i}\vec{x}\cdot\vec{p}/\hbar) \\
&= \frac{1}{(2\pi\hbar)^{3/2}}\exp(\mathrm{i}\vec{x}\cdot\vec{p}/\hbar) \\
&= -\mathrm{i}\hbar\vec{\nabla}_p \int_{\mathbb{R}^3} \mathrm{d}^3x\,\left\langle \vec{x} \,\middle|\, \vec{p} \,\right\rangle \\
&= -\mathrm{i}\hbar\vec{\nabla}_p\,\left|\vec{p}\right\rangle
\end{aligned} \tag{3.7.23}$$

und durch Hermitesche Adjunktion

$$\left\langle\vec{p}\right|\vec{\mathbf{x}}^\dagger = \left\langle\vec{p}\right|\vec{\mathbf{x}} = +\mathrm{i}\hbar\vec{\nabla}_p\,\left\langle\vec{p}\right| \tag{3.7.24}$$

und damit für die Wellenfunktion in der Impulsdarstellung

$$\vec{\mathbf{x}}\tilde{\psi}(\vec{p}) = \left\langle \vec{p} \,\middle|\, \vec{\mathbf{x}}\psi \,\right\rangle = \mathrm{i}\hbar\vec{\nabla}_p\,\left\langle \vec{p} \,\middle|\, \psi \,\right\rangle = \mathrm{i}\hbar\vec{\nabla}_p\tilde{\psi}(\vec{p}). \tag{3.7.25}$$

Weiter können wir vermuten, dass in Analogie zum Impuls als Erzeuger der Translationen der Drehimpuls Drehungen erzeugt, denn auch in der klassischen Mechanik war ja der Drehimpuls die zur Isotropie des Raumes, also die Symmetrie des Euklidischen Raums unter Drehungen um einen beliebigen Punkt, gehörige Erhaltungsgröße im Sinne des Noether-Theorems. Für unsere hier betrachtete Quantenmechanik, die einfach die Analogie zur klassischen Mechanik verwendet, sollte der Operator für den Drehimpuls dadurch gegeben sein, dass man im entsprechenden klassischen Ausdruck einfach die entsprechenden Operatoren einsetzt:

$$\vec{\mathbf{L}} = \vec{\mathbf{x}} \times \vec{\mathbf{p}} \tag{3.7.26}$$

bzw. in Komponenten

$$\mathbf{L}_j = \epsilon_{jkl}\mathbf{x}_k\mathbf{p}_l. \tag{3.7.27}$$

Um zu zeigen, dass $\vec{n}\cdot\vec{\mathbf{L}}$ tatsächlich Drehungen um die Drehachse $\vec{n}$ erzeugt, berechnen wir zunächst die Kommutatoren der Drehimpulskomponenten mit den Orts- und Impulskomponenten:

$$\left[\mathbf{L}_j,\mathbf{x}_m\right] = \epsilon_{jkl}\left[\mathbf{x}_k\mathbf{p}_l,\mathbf{x}_m\right] = \epsilon_{jkl}\mathbf{x}_k\left[\mathbf{p}_l,\mathbf{x}_m\right] = -\mathrm{i}\hbar\epsilon_{jkl}\mathbf{x}_k\delta_{lm} = -\mathrm{i}\hbar\epsilon_{jkm}\mathbf{x}_k = +\mathrm{i}\hbar\epsilon_{jmk}\mathbf{x}_k, \tag{3.7.28}$$

$$\left[\mathbf{L}_j,\mathbf{p}_m\right] = \epsilon_{jkl}\left[\mathbf{x}_k\mathbf{p}_l,\mathbf{p}_m\right] = \epsilon_{jkl}\left[\mathbf{x}_k,\mathbf{p}_m\right]\mathbf{p}_l = \mathrm{i}\hbar\epsilon_{jkl}\delta_{km}p_l = \mathrm{i}\hbar\epsilon_{jml}\mathbf{p}_l. \tag{3.7.29}$$

Allgemein bezeichnet man irgendwelche drei Operatoren $\mathbf{V}_k$ als **Vektoroperatoren**, wenn die Kommutatorrelationen

$$\left[\mathbf{L}_j,\mathbf{V}_m\right] = \mathrm{i}\hbar\epsilon_{jmk}\mathbf{V}_k \tag{3.7.30}$$

gelten. Mit Hilfe der Kommutatorrelationen (3.7.29) und (3.7.30) weist man z.B. leicht nach, dass wie zu erwarten auch die $\mathbf{L}_j$ selbst in diesem Sinne Vektoroperatoren sind, d.h. es gilt

$$\left[\mathbf{L}_j,\mathbf{L}_k\right] = \mathrm{i}\hbar\epsilon_{jkl}\mathbf{L}_l. \tag{3.7.31}$$

Um zu zeigen, dass in der Tat für beliebige Vektoroperatoren $\vec{n}\cdot\vec{\mathbf{L}}$ der Erzeuger von Drehungen um die Drehachse $\vec{n}$ ist, betrachten wir den Operator

$$\vec{\mathbf{V}}(\alpha) = \exp(\mathrm{i}\alpha\vec{n}\cdot\vec{\mathbf{L}}/\hbar)\vec{\mathbf{v}}\exp(-\mathrm{i}\alpha\vec{n}\cdot\vec{\mathbf{L}}/\hbar), \tag{3.7.32}$$

wobei $\vec{\mathbf{v}}$ ein beliebiger Vektoroperator ist. Um den expliziten Ausdruck zu finden, entwickeln wir diesen Ausdruck in eine Taylor-Reihe

$$\vec{\mathbf{V}}(\alpha) = \sum_{j=0}^{\infty}\frac{\alpha^j}{j!}\left[\frac{\mathrm{d}^j}{\mathrm{d}\alpha^j}\vec{\mathbf{V}}(\alpha)\right]_{\alpha=0}. \tag{3.7.33}$$

Die Ableitung von (3.7.32) ergibt zusammen mit den Kommutatorregeln (3.7.30)

$$\frac{\mathrm{d}}{\mathrm{d}\alpha}\vec{\mathbf{V}} = \left[\mathrm{i}\vec{n}\cdot\vec{\mathbf{L}}/\hbar,\vec{\mathbf{V}}\right] = \vec{n}\times\vec{\mathbf{V}}. \tag{3.7.34}$$

Entsprechend folgt für die zweite Ableitung

$$\frac{\mathrm{d}^2}{\mathrm{d}\alpha^2}\vec{\mathbf{V}} = \vec{n}\times[\vec{n}\times\vec{\mathbf{V}}] = \vec{n}(\vec{n}\cdot\vec{V}) - \vec{\mathbf{V}} = -\vec{\mathbf{V}}_\perp \quad \text{mit} \quad \vec{\mathbf{V}}_\perp = (\vec{n}\times\vec{\mathbf{V}})\times\vec{n} = \vec{\mathbf{V}} - \vec{n}(\vec{n}\cdot\vec{\mathbf{V}}) \tag{3.7.35}$$

und daraus weiter

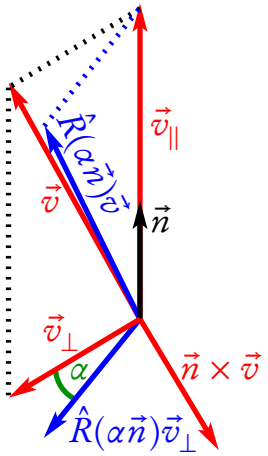
$$\frac{\mathrm{d}^3}{\mathrm{d}\alpha^3}\vec{\mathbf{V}} = -\vec{n}\times\vec{\mathbf{V}}. \tag{3.7.36}$$

Durch Iteration finden wir also

$$\frac{\mathrm{d}^{2j-1}}{\mathrm{d}\alpha^{2j-1}}\vec{\mathbf{V}} = (-1)^{j+1}\vec{n}\times\vec{\mathbf{V}}, \quad \frac{\mathrm{d}^{2j}}{\mathrm{d}\alpha^{2j}} = (-1)^j(\vec{n}\times\vec{\mathbf{V}})\times\vec{n} = (-1)^j\vec{\mathbf{V}}_\perp. \tag{3.7.37}$$

Setzen wir dies in die Reihe (3.7.33) ein, erhalten wir schließlich

$$\vec{\mathbf{V}}(\alpha) = \vec{n}(\vec{n}\cdot\vec{\mathbf{v}}) + \vec{\mathbf{v}}_\perp\cos\alpha + \vec{n}\times\vec{\mathbf{v}}\sin\alpha. \tag{3.7.38}$$



Wie aus der nebenstehenden Skizze ersichtlich, beschreibt dies in der Tat eine Drehung um die Drehachse $\vec{n}$ mit Drehwinkel α.
Wir bezeichnen den unitären Operator für Drehungen mit

$$\mathbf{D}(\vec{\alpha}) = \exp\left(-\frac{\mathrm{i}}{\hbar}\vec{\alpha}\cdot\vec{\mathbf{L}}\right). \tag{3.7.39}$$

Dabei ist $\vec{\alpha} = \alpha\vec{n}$ mit $\alpha \in [0,\pi]$ und $\vec{n}$ einem beliebigen Einheitsvektor, also $\vec{n} \in \Omega$, wobei $\Omega = \{\vec{n}|\vec{n}^2 = 1\}$ die Einheitskugeloberfläche bezeichnet. Analog wie oben für die Translationen gezeigt, folgt für die Wirkung auf Orts- bzw. Impulseigenzustände

$$\mathbf{D}(\vec{\alpha})|\vec{x}\rangle = \left|\hat{R}(\vec{\alpha})\vec{x}\right\rangle, \quad \mathbf{D}(\vec{\alpha})|\vec{p}\rangle = \left|\hat{R}(\vec{\alpha})\vec{p}\right\rangle. \tag{3.7.40}$$

Dabei ist $\hat{R}(\vec{\alpha}) \in \mathrm{SO}(3)$ die Drehmatrix, die einer Drehung um den Winkel $\alpha = |\vec{\alpha}|$ um die Drehachse $\vec{n} = \vec{\alpha}/\alpha$ entspricht. Für einen Vektor $\vec{a} \in \mathbb{R}^3$ ist dabei analog we für einen Vektoroperator (3.7.38)

$$\hat{R}(\vec{\alpha})\vec{a} = \vec{n}(\vec{n}\cdot\vec{a}) + (\vec{n}\times\vec{a})\times\vec{n}\cos\alpha + \vec{n}\times\vec{a}\sin\alpha. \tag{3.7.41}$$

Daraus folgt auch das Verhalten einer beliebigen Ortswellenfunktion unter Drehungen:

$$\mathbf{D}(\vec{\alpha})\psi(\vec{x}) = \langle\vec{x}\,|\,\mathbf{D}(\vec{\alpha})\psi\rangle = \left\langle\mathbf{D}^\dagger(\vec{\alpha})\vec{x}\,\middle|\,\psi\right\rangle = \left\langle\hat{R}^{-1}(\vec{\alpha})\vec{x}\,\middle|\,\psi\right\rangle = \psi[\hat{R}^{-1}(\vec{\alpha})\vec{x}]. \tag{3.7.42}$$

Wie zu erwarten, transformiert sich also eine Wellenfunktion unter Drehungen wie ein **Skalarfeld**.
Der Vollständigkeit halber betrachten wir auch noch **Galilei-Boosts**. Wir bezeichnen die Generatoren für einen Galilei-Boost mit $\vec{\mathbf{K}}$ und den unitären Operator für einen Boost mit Geschwindigkeit $\vec{v}$

$$\mathbf{B}(\vec{v}) = \exp(\mathrm{i}m\vec{v}\cdot\vec{\mathbf{K}}/\hbar). \tag{3.7.43}$$

Dabei ist m die Masse des Teilchens. Wir werden gleich sehen, warum wir die angegebene Konvention hinsichtlich der Faktoren im Exponenten wie angegeben gewählt haben. Damit nun die Wirkung dieses Operators auf die Orts- und Impuls-Eigenvektoren einem Boost entspricht, setzen wir $\vec{v} = v\vec{n}$ und verlangen, dass

$$\vec{\mathbf{X}}(v) = \mathbf{B}^\dagger(v\vec{n})\vec{\mathbf{x}}\mathbf{B}(v\vec{n}) = \vec{\mathbf{x}} + v\vec{n}t\,\mathbb{1}, \tag{3.7.44}$$

$$\vec{\mathbf{P}}(v) = \mathbf{B}^\dagger(v\vec{n})\vec{\mathbf{p}}\mathbf{B}(v\vec{n}) = \vec{\mathbf{p}} + m\vec{v}\,\mathbb{1} \tag{3.7.45}$$

gilt. Um die Operatoren $\vec{\mathbf{K}}$ zu finden, leiten wir (3.7.44) nach v ab (*Übung!*):

$$\frac{\mathrm{d}}{\mathrm{d}v}\vec{\mathbf{X}}(v) = \mathbf{B}^\dagger(v\vec{n})\left[\vec{\mathbf{x}},\vec{n}\cdot\vec{\mathbf{K}}\right]\mathbf{B}^\dagger = -\frac{\mathrm{i}\hbar}{m}t\vec{n}\mathbb{1} \overset{!}{=} \vec{n}t\,\mathbb{1} \Rightarrow [\mathbf{x}_a,\mathbf{K}_b] = -\frac{\mathrm{i}\hbar}{m}t\delta_{ab}. \tag{3.7.46}$$

Dies wird wegen (3.7.9) offenbar durch

$$\vec{\mathbf{K}} = \vec{\mathbf{K}}'(\vec{\mathbf{x}}) - \frac{t}{m}\vec{\mathbf{p}} \tag{3.7.47}$$

erfüllt. Leiten wir nun (3.7.45) nach v ab:

$$\frac{\mathrm{d}}{\mathrm{d}v}\vec{\mathbf{P}}(v) = \frac{\mathrm{i}m}{\hbar}\mathbf{B}^\dagger(v\vec{n})\left[\vec{\mathbf{p}},\vec{n}\cdot\vec{\mathbf{K}}\right]\mathbf{B}(v\vec{n}) \overset{!}{=} m\vec{n}\mathbb{1} \Rightarrow [\mathbf{p}_a,\mathbf{K}_b] = -\mathrm{i}\hbar\delta_{ab}\mathbb{1}. \tag{3.7.48}$$

Setzen wir hierin (3.7.47) ein, erhalten wir schließlich

$$\vec{\mathbf{K}} = \vec{\mathbf{x}} - \frac{t}{m}\vec{\mathbf{p}}. \tag{3.7.49}$$

Wir erhalten damit für den Boost-Operator

$$\mathbf{B}(\vec{v}) = \exp\left[-\frac{\mathrm{i}}{\hbar}\vec{v}\cdot(m\vec{\mathbf{x}} - t\vec{\mathbf{p}})\right]. \tag{3.7.50}$$

Im Folgenden ist es nüzlich, diesen Operator noch umzuformen. Dazu bemerken wir, dass die Bedingungen für die Anwendung von (A.2.2) mit

$$\mathbf{A} = \mathrm{i}\vec{v}\cdot\vec{\mathbf{p}}/\hbar, \quad \mathbf{B} = -\mathrm{i}mt\vec{v}\cdot\vec{\mathbf{x}}/\hbar \tag{3.7.51}$$

erfüllt sind, denn man rechnet leicht nach, dass

$$[\mathbf{A},\mathbf{B}] = -\frac{\mathrm{i}mt}{\hbar}\vec{v}^2\mathbb{1} \tag{3.7.52}$$

ist, so dass in der Tat die Bedingungen (A.2.1) erfüllt sind. Damit liefert (A.2.2)

$$\mathbf{B}(\vec{v}) = \exp\left(-\frac{\mathrm{i}t}{\hbar}\vec{v}\cdot\vec{\mathbf{p}}\right)\exp\left(\frac{\mathrm{i}m}{\hbar}\vec{v}\cdot\vec{\mathbf{x}}\right)\exp\left(\frac{\mathrm{i}mt}{2\hbar}\vec{v}^2\right) \tag{3.7.53}$$

Vertauscht man in (A.2.2) die Operatoren **A** und **B** erhält man die alternative Formel

$$\mathbf{B}(\vec{v}) = \exp\left(\frac{\mathrm{i}m}{\hbar}\vec{v}\cdot\vec{\mathbf{x}}\right)\exp\left(-\frac{\mathrm{i}t}{\hbar}\vec{v}\cdot\vec{\mathbf{p}}\right)\exp\left(-\frac{\mathrm{i}mt}{2\hbar}\vec{v}^2\right). \tag{3.7.54}$$

Nun können wir die Wirkung des Boost-Operators auf die verallgemeinerten Ortseigenvektoren berechnen. Verwenden wir den Boost-Operator in der Form (3.7.53), erhalten wir unter Verwendung von (3.7.15)

$$\mathbf{B}(\vec{v})|\vec{x}\rangle = \exp\left[\frac{\mathrm{i}m}{\hbar}\left(\vec{v}\cdot\vec{x} + \frac{1}{2}\vec{v}^2 t\right)\right]|\vec{x} + \vec{v}t\rangle. \tag{3.7.55}$$

Beim Impulseigenvektor bietet sich die Form (3.7.54) des Boost-Operators an:

$$\mathbf{B}(\vec{v})\left|\vec{p}\right\rangle = \exp\left[-\frac{\mathrm{i}t}{\hbar}\left(\vec{v}\cdot\vec{p} + \frac{m}{2}\vec{v}^2\right)\right]\exp\left(\frac{\mathrm{i}m}{\hbar}\vec{v}\cdot\vec{\mathbf{x}}\right)\left|\vec{p}\right\rangle. \tag{3.7.56}$$

Berechnen wir also unter Verwendung von (3.7.19)

$$\begin{aligned}
\exp\left(\frac{\mathrm{i}m}{\hbar}\vec{v}\cdot\vec{\mathbf{x}}\right)\left|\vec{p}\right\rangle &= \int_{\mathbb{R}^3} \mathrm{d}^3x \exp\left(\frac{\mathrm{i}m}{\hbar}\vec{v}\cdot\vec{\mathbf{x}}\right)|\vec{x}\rangle\left\langle\vec{x}\,\middle|\,\vec{p}\right\rangle \\
&= \int_{\mathbb{R}^3} \mathrm{d}^3x \exp(\mathrm{i}m\vec{v}\cdot\vec{x}/\hbar)|\vec{x}\rangle\left\langle\vec{x}\,\middle|\,\vec{p}\right\rangle \\
&= \int_{\mathbb{R}^3} \mathrm{d}^3x \left(\frac{1}{2\pi\hbar}\right)^{3/2} \exp[\mathrm{i}(\vec{p} + m\vec{v})\cdot\vec{x}/\hbar]|\vec{x}\rangle \\
&= \int_{\mathbb{R}^3} \mathrm{d}^3x\,|\vec{x}\rangle\left\langle\vec{x}\,\middle|\,\vec{p} + m\vec{v}\right\rangle \\
&= \left|\vec{p} + m\vec{v}\right\rangle.
\end{aligned} \tag{3.7.57}$$

Wie zu erwarten ist also (3.7.56) ein Impulseigenvektor zum Eigenwert $(\vec{p} + m\vec{v})$:

$$\mathbf{B}(\vec{v})\left|\vec{p}\right\rangle = \exp\left[-\frac{\mathrm{i}t}{\hbar}\left(\vec{v}\cdot\vec{p} + \frac{m}{2}\vec{v}^2\right)\right]\left|\vec{p} + m\vec{v}\right\rangle. \tag{3.7.58}$$

Aus (3.7.56) ergibt sich auch das Verhalten der Ortswellenfunktion unter Galilei-Boosts

$$\begin{aligned}\mathbf{B}(\vec{v})\psi(\vec{x}) &= \langle \vec{x} \,|\, \mathbf{B}(\vec{v})\psi \rangle = \left\langle \mathbf{B}^\dagger(\vec{v})\vec{x} \,\middle|\, \psi \right\rangle \\ &= \langle \mathbf{B}(-\vec{v})\vec{x} \,|\, \psi \rangle \\ &= \left\langle \exp\left[\frac{\mathrm{i}m}{\hbar}\left(-\vec{v}\cdot\vec{x} + \frac{1}{2}\vec{v}^2 t\right)\right](\vec{x}-\vec{v}t) \,\middle|\, \psi \right\rangle \\ &= \exp\left[\frac{\mathrm{i}m}{\hbar}\left(\vec{v}\cdot\vec{x} - \frac{1}{2}\vec{v}^2 t\right)\right]\psi(\vec{x}-\vec{v}t).\end{aligned} \tag{3.7.59}$$

Gegenüber der naiven Erwartung eines für ein Feld üblichen Transformationsverhaltens, für das man wegen $\vec{x}' \to \vec{x}+\vec{v}t$ erwartet hätte, dass $\vec{\psi}'(\vec{x}') = \psi(\vec{x}) = \psi(\vec{x}'-\vec{v}t)$ folgt, erhält man für die geboostete Wellenfunktion einen zusätzlichen Phasenfaktor. Dieser ist freilich physikalisch irrelevant. Das Transformationsverhalten für die Aufenthaltswahrscheinlichkeitsdichte, also $|\psi(\vec{x})|^2$ ist offensichtlich das für ein Feld übliche.

3.8 Das freie Teilchen

In diesem Abschnitt betrachten wir nochmals das freie Teilchen aus dem Blickwinkel des nunmehr vollständig entwickelten Formalismus' der Quantenmechanik. Wir konzentrieren uns dabei insbesondere auf die Lösung der zeitabhängigen Schrödinger-Gleichung.
Da für ein freies Teilchen

$$\mathbf{H} = \frac{1}{2m}\vec{\mathbf{p}}^2 \tag{3.8.1}$$

ist, ist es klar, dass die generalisierten **Impuls-Eigenzustände**

$$u_{\vec{p}}(\vec{x}) = \frac{1}{(2\pi\hbar)^{3/2}} \exp\left(\frac{\mathrm{i}\vec{p}\cdot\vec{x}}{\hbar}\right) \tag{3.8.2}$$

zugleich einen vollständigen Satz generalisierter **Energie-Eigenzustände** bilden. In der Tat ist

$$\mathbf{H}u_{\vec{p}}(\vec{x}) = -\frac{\hbar^2}{2m}\Delta u_{\vec{p}}(\vec{x}) = \frac{1}{2m}\vec{p}^2 u_{\vec{p}}(\vec{x}). \tag{3.8.3}$$

Die Energieeigenwerte sind

$$E_{\vec{p}} = \frac{1}{2m}\vec{p}^2. \tag{3.8.4}$$

Wir können nun leicht die Zeitentwicklung der Wellenfunktion aus der formalen Lösung (3.6.6) gewinnen. Demnach ist nämlich

$$\psi(t,\vec{x}) = \mathbf{U}(t)\psi_0(\vec{x}) \quad \text{mit} \quad \mathbf{U}(t) = \exp\left(-\frac{\mathrm{i}\mathbf{H}t}{\hbar}\right), \tag{3.8.5}$$

wobei $\psi_0(\vec{x})$ die Wellenfunktion ist, die den Zustand beschreibt, in dem das Teilchen zur Zeit $t = 0$ präpariert worden ist. Dies kann eine beliebige quadratintegrable komplexwertige Funktion sein. Die explizite Lösung können wir leicht angeben, wenn wir die Wellenfunktion nach den Impulseigenzuständen (3.8.2) entwickeln

$$\psi(t,\vec{x}) = \int_{\mathbb{R}^3} \mathrm{d}^3p\, u_{\vec{p}}(\vec{x})\left\langle u_{\vec{p}} \,\middle|\, \psi(t) \right\rangle = \int_{\mathbb{R}^3} \mathrm{d}^3p\, u_{\vec{p}}(\vec{x})\left\langle u_{\vec{p}} \,\middle|\, \mathbf{U}(t)\psi_0 \right\rangle = \int_{\mathbb{R}^3} \mathrm{d}^3p\, u_{\vec{p}}(\vec{x})\left\langle \mathbf{U}^\dagger(t)u_{\vec{p}} \,\middle|\, \psi_0 \right\rangle. \tag{3.8.6}$$

Nun gilt aber

$$\mathbf{U}(t) = \exp\left(-\frac{\mathrm{i}\mathbf{H}t}{\hbar}\right) \Rightarrow \mathbf{U}^\dagger(t) = \exp\left(+\frac{\mathrm{i}\mathbf{H}t}{\hbar}\right) \tag{3.8.7}$$

und damit

$$\mathbf{U}^\dagger(t)u_{\vec{p}}(\vec{x}) = \exp\left(\frac{\mathrm{i}E_{\vec{p}}t}{\hbar}\right)u_{\vec{p}}(\vec{x}). \tag{3.8.8}$$

Setzen wir dies in (3.8.6) ein, folgt

$$\psi(t,\vec{x}) = \int_{\mathbb{R}^3} \mathrm{d}^3p\, u_{\vec{p}}(\vec{x})\left\langle \exp\left(+\frac{\mathrm{i}E_{\vec{p}}t}{\hbar}\right)u_{\vec{p}}\,\middle|\,\psi_0\right\rangle = \int_{\mathbb{R}^3} \mathrm{d}^3p\, u_{\vec{p}}(\vec{x})\exp\left(-\frac{\mathrm{i}E_{\vec{p}}t}{\hbar}\right)\left\langle u_{\vec{p}}\,\middle|\,\psi_0\right\rangle. \tag{3.8.9}$$

Wir müssen also bei gegebener Anfangswellenfunktion nur deren Impulsdarstellung

$$\tilde{\psi}_0(\vec{p}) = \left\langle u_{\vec{p}}\,\middle|\,\psi_0\right\rangle = \int_{\mathbb{R}^3} \mathrm{d}^3x\, u^*_{\vec{p}}(\vec{x})\psi_0(\vec{x}) = \int_{\mathbb{R}^3} \mathrm{d}^3x \frac{1}{(2\pi\hbar)^{3/2}}\exp\left(-\frac{\mathrm{i}\vec{p}\cdot\vec{x}}{\hbar}\right)\psi_0(\vec{x}) \tag{3.8.10}$$

berechnen, und in (3.8.9) einsetzen:

$$\psi(t,\vec{x}) = \int_{\mathbb{R}^3} \mathrm{d}^3p \frac{1}{(2\pi\hbar)^{3/2}}\exp\left(\frac{\mathrm{i}\vec{p}\cdot\vec{x} - \mathrm{i}E(\vec{p})t}{\hbar}\right)\tilde{\psi}_0(\vec{p}). \tag{3.8.11}$$

Dies ist natürlich das gleiche Resultat, das wir im vorigen Kapitel bereits mit Hilfe der Fourier-Transformation in (2.7.10) gefunden hatten. In den Übungen werden wir eine konkrete Lösung, das Gaußsche Wellenpaket, ausführlich besprechen.

Wir können (3.8.11) allerdings noch in eine Form bringen, die direkt mit den Ortswellenfunktionen auskommt. Dazu müssen wir nur (3.8.10) in (3.8.11) einsetzen:

$$\begin{aligned}\psi(t,\vec{x}) &= \int_{\mathbb{R}^3} \mathrm{d}^3p \frac{1}{(2\pi\hbar)^{3/2}}\exp\left(\frac{\mathrm{i}\vec{p}\cdot\vec{x} - \mathrm{i}E(\vec{p})t}{\hbar}\right)\int_{\mathbb{R}^3} \mathrm{d}^3\vec{x}' \frac{1}{(2\pi\hbar)^{3/2}}\exp\left(-\frac{\mathrm{i}\vec{p}\cdot\vec{x}'}{\hbar}\right)\psi_0(\vec{x}') \\ &= \int_{\mathbb{R}^3} \mathrm{d}^3\vec{x}' \frac{1}{(2\pi\hbar)^{3/2}}\psi_0(\vec{x}')\int_{\mathbb{R}^3} \mathrm{d}^3p \frac{1}{(2\pi\hbar)^{3/2}}\exp\left(\frac{\mathrm{i}\vec{p}\cdot(\vec{x}-\vec{x}') - \mathrm{i}E(\vec{p})t}{\hbar}\right)\end{aligned} \tag{3.8.12}$$

Definieren wir den **Propagator**

$$U(t,\vec{x},\vec{x}') = \int_{\mathbb{R}^3} \mathrm{d}^3p \frac{1}{(2\pi\hbar)^3}\exp\left(\frac{\mathrm{i}\vec{p}\cdot(\vec{x}-\vec{x}') - \mathrm{i}E(\vec{p})t}{\hbar}\right), \tag{3.8.13}$$

können wir (3.8.12) in der Form

$$\psi(t,\vec{x}) = \int_{\mathbb{R}^3} \mathrm{d}^3x' U(t,\vec{x},\vec{x}')\psi_0(\vec{x}') \tag{3.8.14}$$

schreiben. Dabei ist zu beachten, dass der Propagator eine verallgemeinerte Funktion (Distribution) ist. Das wird u.a. dadurch deutlich, dass

$$U(t=0,\vec{x},\vec{x}') = \delta^{(3)}(\vec{x}-\vec{x}') \tag{3.8.15}$$

ist (*nachrechnen!*). Das muss freilich auch so sein, damit sich aus (3.8.14) $\psi(t=0,\vec{x}) = \psi_0(\vec{x})$ ergibt.

Wir können (3.8.13) noch geschlossen ausrechnen. Dabei verwendent wir die Ergebnisse von Anhang C, insbesondere (C.1.22). Damit (3.8.13) konvergiert, denken wir uns t durch $t_\epsilon = t - \mathrm{i}\epsilon$ mit $\epsilon > 0$ ersetzt. Am Schluss der Rechnung können wir dann den Limes $\epsilon \to 0$ bilden. Jedenfalls ergibt (C.1.22), angewandt auf die Integrale bzgl. der drei Komponenten von $\vec{p}$ (*Nachrechnen!*)

$$U(t,\vec{x},\vec{x}') = \left(\frac{m}{2\pi\mathrm{i}\hbar t}\right)^{3/2}\exp\left[\frac{\mathrm{i}m(\vec{x}-\vec{x}')^2}{2\hbar t}\right]. \tag{3.8.16}$$

3.9 Der eindimensionale harmonische Oszillator

In diesem Abschnitt betrachten wir als eines der wichtigsten Beispiele für die Quantenmechanik den **harmonischen Oszillator**. Wie wir sehen werden, können wir das **Energieeigenwertproblem** fast vollständig algebraisch lösen, indem wir nur auf die **Kommutatorrelationen** zwischen Orts- und Impulsoperatoren zurückgreifen. Wir betrachten nacheinander die Fälle eines harmonischen Oszillators in ein, zwei und drei Dimensionen.

Wir beginnen mit dem einfachsten Fall eines Teilchens, das sich entlang der x-Achse in einem **Oszillatorpotential** bewegt. Der Hamilton-Operator lautet

$$\mathbf{H} = \frac{1}{2m}\mathbf{p}^2 + \frac{m\omega^2}{2}\mathbf{x}^2. \tag{3.9.1}$$

Dabei ist ω die Schwingungsfrequenz der klassischen Lösungen. Wir wollen im Folgenden die Energieeigenwerte und -eigenfunktionen für diesen Hamilton-Operator finden. Erstaunlicherweise können wir dieses Problem mit rein algebraischen Methoden lösen, ohne die zeitunabhängige Schrödinger-Gleichung direkt lösen zu müssen. Freilich werden wir diese Lösungen aus den algebraischen Betrachtungen auch explizit angeben können.

Dazu führen wir zunächst den folgenden *nicht selbstadjungierten (!)* Operator

$$\mathbf{a} = \sqrt{\frac{m\omega}{2\hbar}}\mathbf{x} + \frac{\mathrm{i}}{\sqrt{2m\hbar\omega}}\mathbf{p} \tag{3.9.2}$$

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

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

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

$$[\mathbf{x},\mathbf{p}] = \mathrm{i}\hbar\mathbb{1} \tag{3.9.4}$$

verwendet haben. Der Vergleich mit (3.9.1) ergibt

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

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

$$\mathbf{N} = \mathbf{a}^\dagger\mathbf{a} \tag{3.9.6}$$

betrachten.

Die Kommutatorrelationen der Operatoren $\mathbf{a}$ und $\mathbf{a}^\dagger$ berechnet man sofort mit Hilfe der Kommutatorrelationen für Ort und Impuls 3.7.9 (*Nachrechnen!*)

$$[\mathbf{a},\mathbf{a}] = \left[\mathbf{a}^\dagger,\mathbf{a}^\dagger\right] = 0, \quad \left[\mathbf{a},\mathbf{a}^\dagger\right] = \mathbb{1}. \tag{3.9.7}$$

Schließlich sind im Folgenden noch die Kommutatoren

$$[\mathbf{a},\mathbf{N}] = \mathbf{a}, \quad \left[\mathbf{a}^\dagger,\mathbf{N}\right] = -\mathbf{a}^\dagger \tag{3.9.8}$$

nützlich (*Nachrechnen!*).

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

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

d.h. $\mathbf{a}|u_n\rangle$ ist also entweder ein Eigenvektor von $\mathbf{N}$ zum Eigenwert $(n-1)$ oder $\mathbf{a}|u_n\rangle = 0$.
Weiter ist $\mathbf{N}$ ein „positiv semidefiniter Operator", d.h. für jede Wellenfunktion $\psi(x)$ gilt

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

Setzen wir hierin insbesondere $|\psi\rangle = |u_n\rangle$, folgt, dass $n \geq 0$ sein muss. Wenden wir also $\mathbf{a}^k$ auf $|u_n\rangle$ an, muss für ein $k \in \mathbb{N}$ gelten $n-k=0$, weil es andernfalls Eigenfunktionen mit beliebig kleinen (negativen) Eigenwerten von $\mathbf{N}$ gäbe, was aber eben wegen (3.9.10) nicht möglich ist. Existiert also überhaupt ein Eigenvektor von $\mathbf{N}$, dann muss es (wenigstens) einen Eigenvektor $|u_0\rangle$ zum Eigenwert $n=0$ geben, und für diesen muss folglich

$$\mathbf{a}|u_0\rangle = 0 \tag{3.9.11}$$

gelten. Weiter gilt aber auch

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

und das beweist, dass $\mathbf{a}^\dagger|u_n\rangle$ ein Eigenvektor von $\mathbf{N}$ zum Eigenwert $(n+1)$ sein muss. Damit sind die Eigenwerte von $\mathbf{N}$ die natürlichen Zahlen (einschließlich 0), also $n \in \mathbb{N}_0$.
Nehmen wir an, $|u_0\rangle$ 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 $|u_0\rangle$ und nachfolgende Normierung. Zur Bestimmung des Normierungsfaktors nehmen wir nun also an, die $|u_n\rangle$ seien auf 1 normiert. Dann folgt

$$\left\langle \mathbf{a}^\dagger u_n \middle| \mathbf{a}^\dagger u_n \right\rangle = \left\langle u_n \middle| \mathbf{a}\mathbf{a}^\dagger \middle| u_n \right\rangle = \left\langle u_n \middle| \mathbf{N} + \left[\mathbf{a},\mathbf{a}^\dagger\right] \middle| u_n \right\rangle = (n+1). \tag{3.9.13}$$

Durch Rekursion erhalten wir also schließlich

$$|u_n\rangle = \sqrt{\frac{1}{n!}}(\mathbf{a}^\dagger)^n|u_0\rangle, \tag{3.9.14}$$

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

$$|u_n\rangle = \frac{1}{\sqrt{n}}\mathbf{a}^\dagger\left|u_{n-1}\right\rangle, \tag{3.9.15}$$

und mit der Kommutatorrelation (3.9.7) finden wir

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

Wir stellen weiter fest, dass damit auch das Eigenwertproblem für $\mathbf{H}$ gelöst ist, denn gemäß (3.9.6) ist $|u_n\rangle$ ein Eigenvektor von $\mathbf{H}$, d.h. es gilt

$$\mathbf{H}|u_n\rangle = E_n|u_n\rangle \tag{3.9.17}$$

mit den Eigenwerten

$$E_n = \frac{\hbar\omega}{2}(2n+1), \quad n \in \mathbb{N}_0. \tag{3.9.18}$$

Es ist klar, dass der Zustand $|u\rangle_0$ der Grundzustand, d.h. der Zustand der niedrigstmöglichen Energie des Teilchens ist.
Wir können mit den obigen Betrachtungen ziemlich einfach alle Energieeigenfunktionen $u_n(x) = \langle x|u_n\rangle$ berechnen. Dazu müssen wir in die Definition (3.9.2) für $\mathbf{a}$ nur $\mathbf{x} = x$ und $\mathbf{p} = -\mathrm{i}\hbar\partial_x$ einsetzen und die Bestimmungsgleichung (3.9.11) der Grundzustandswellenfunktion hinschreiben:

$$\mathbf{a}u_0(x) = \left(\sqrt{\frac{\hbar}{2m\omega}}\frac{\partial}{\partial x} + \sqrt{\frac{m\omega}{2\hbar}}x\right)u_0(x) = 0. \tag{3.9.19}$$

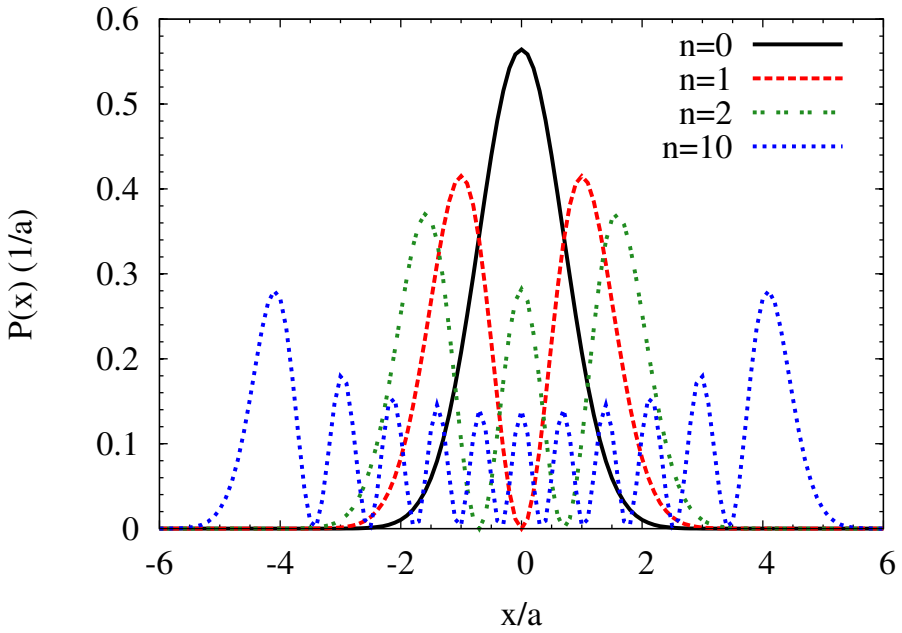


Abbildung 3.1: *Die Energieeigenfunktionen des eindimensionalen harmonischen Oszillators zu den Eigenwerten $E_n = \hbar\omega(n+1/2)$ für $n = 0, 1, 2, 10$. Besonders bemerkenswert ist die Tatsache, dass in dieser Situation, die einem Teilchen mit scharf bestimmter Energie entsprechen, die Aufenthaltswahrscheinlichkeit 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****.*

Setzen wir zur Abkürzung

$$a^2 = \frac{\hbar}{m\omega}, \tag{3.9.20}$$

schreibt sich diese Gleichung als

$$\frac{\partial}{\partial x} u_0 = -\frac{x}{a^2} u_0. \tag{3.9.21}$$

Die Lösung ergibt sich sofort *eindeutig* zu

$$u_0(x) = \mathcal{N} \exp\left(-\frac{x^2}{2a^2}\right). \tag{3.9.22}$$

Die Normierung ist (bis auf einen unerheblichen Phasenfaktor) durch

$$\langle u_0 \,|\, u_0 \rangle = \int \mathrm{d}x \, |u_0(x)|^2 = |\mathcal{N}|^2 (\pi a^2)^{1/2} \stackrel{!}{=} 1 \Rightarrow \mathcal{N} = \left(\frac{1}{\pi a^2}\right)^{1/4} \tag{3.9.23}$$

gegeben.
Wir erhalten also schließlich

$$u_0(x) = \left(\frac{1}{\pi a^2}\right)^{1/4} \exp\left(-\frac{x^2}{2a^2}\right). \tag{3.9.24}$$

Es existiert also genau eine quadratintegrable Wellenfunktion zum Eigenwertproblem von **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 (3.9.24) mit einem Polynom, das sich durch fortgesetzte Anwendung des Differentialoperators $\mathbf{a}^\dagger$ bestimmen lässt:

$$u_n(x) = \sqrt{\frac{1}{n!}}(\mathbf{a}^\dagger)^n u_0(x) = \sqrt{\frac{1}{2^n n!}}\left(\frac{x}{a} - a\frac{\mathrm{d}}{\mathrm{d}x}\right)^n u_0(x). \tag{3.9.25}$$

Führen wir die dimensionslose Variable

$$\xi = \frac{x}{a} \tag{3.9.26}$$

ein, können wir gem. (3.9.24) schreiben

$$u_n(x) = \left(\frac{1}{\pi a^2}\right)^{1/4} \sqrt{\frac{1}{2^n n!}}\left[\left(\xi - \frac{\mathrm{d}}{\mathrm{d}\xi}\right)^n \exp\left(-\frac{\xi^2}{2}\right)\right]_{\xi=\frac{x}{a}}. \tag{3.9.27}$$

Offensichtlich ist

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

Dabei sind

$$\mathrm{H}_n(\xi) := \exp\left(\frac{\xi^2}{2}\right)\left(\xi - \frac{\mathrm{d}}{\mathrm{d}\xi}\right)^n \exp\left(-\frac{\xi^2}{2}\right) \tag{3.9.29}$$

die sog. *Hermite-Polynome* (Charles Hermite, 1822-1901). In Abb. 3.1 sind einige dieser Energieeigenfunktionen dargestellt. Wir leiten im Anhang B.1 noch einige weitere nützliche Eigenschaften der Hermite-Polynome sowie den **Propagator** für den harmonischen Oszillator her.

3.10 Eigenwertproblem für beliebige Drehimpulsoperatoren

Als nächstes Beispiel für die schlagkräftigen Methoden der abstrakten Hilbert-Raumdarstellung beschäftigen wir uns mit dem Eigenwertproblem für Drehimpulsoperatoren. Dabei machen wir zunächst nur von den Kommutatorrelationen (3.7.31) für die Bahndrehimpulsoperatoren aus, sehen aber von der Realisierung $\vec{\mathbf{L}} = \vec{\mathbf{x}} \times \vec{\mathbf{p}}$ ab. Wie wir später noch sehen werden, gibt es nämlich in der Quantentheorie eine spezifische der klassischen Physik fremde Art Drehimpuls für ein einzelnes „Punktteilchen“, den Spin.

Wir betrachten also allgemeine selbstadjungierte Drehimpulsoperatoren, die einzig dadurch charakterisiert sind, dass sie die zu (3.7.31) analogen Kommutatorregeln

$$[\mathbf{J}_a, \mathbf{J}_b] = \mathrm{i}\hbar\epsilon_{abc}\mathbf{J}_c \tag{3.10.1}$$

erfüllen. Im Folgenden ist es bequemer mit dimensionslosen Drehimpulsoperatoren $\mathbf{j}_a = \mathbf{J}_a/\hbar$ zu arbeiten. Für sie gelten offenbar die Kommutatorrelationen

$$[\mathbf{j}_a, \mathbf{j}_b] = \mathrm{i}\epsilon_{abc}\mathbf{j}_c. \tag{3.10.2}$$

Diese mögen in irgendeinem abstrakten Hilbert-Raum operieren, und wir suchen einen vollständigen Satz von Drehimpulseigenvektoren. Offenbar kommutieren nun aber die Drehimpulsoperatoren nicht, und wir können daher nur Eigenvektoren für eine Drehimpulskomponente suchen. Üblicherweise wählt man die Komponente $\vec{j}_3$.

Nun folgen aber die Kommutatorregeln (3.10.1) bzw. (3.10.2) daraus, dass die $\vec{\mathbf{j}}$ die Erzeugenden von Drehungen sind. Es liegt daher nahe, dass $\vec{\mathbf{j}}^2$ mit allen Drehimpulskomponenten kommutiert, da dies einem Skalar entspricht, der sich unter Drehungen nicht ändert. In der Tat gilt

$$\left[\mathbf{j}_a, \vec{\mathbf{j}}^2\right] = [\mathbf{j}_a, \mathbf{j}_b\mathbf{j}_b] = [\mathbf{j}_a, \mathbf{j}_b]\mathbf{j}_b + \mathbf{j}_b[\mathbf{j}_a, \mathbf{j}_b] = \mathrm{i}\epsilon_{abc}(\mathbf{j}_c\mathbf{j}_b + \mathbf{j}_b\mathbf{j}_c) = 0. \tag{3.10.3}$$

Letzteres folgt daraus, dass der Ausdruck in der Klammer symmetrisch unter Vertauschung der Indizes a und b symmetrisch, das Levi-Civita-Symbol aber antisymmetrisch ist. Wir können also ein vollständiges System simultaner Eigenvektoren für $\vec{\mathbf{j}}^2$ und $\mathbf{j}_3$ finden. Offenbar können wir keine weiteren skalaren Operatoren mit den $\vec{\mathbf{j}}$-Operatoren bilden, so dass hinsichtlich des die Drehimpulse realisierenden Hilbertraums die gemeinsamen Eigenvektoren von $\vec{\mathbf{j}}^2$ und $\mathbf{j}_3$ bis auf die Normierung eindeutig bestimmt sein sollten.

Wir suchen also gemeinsame Eigenvektoren $u_{\alpha m}$ zu $\vec{\mathbf{j}}^2$ und $\mathbf{j}_3$:

$$\vec{\mathbf{j}}^2 |u_{\alpha m}\rangle = \alpha |u_{\alpha m}\rangle, \quad \mathbf{j}_3 |u_{\alpha m}\rangle = m |u_{\alpha m}\rangle. \tag{3.10.4}$$

Wie wir im Folgenden sehen werden, erweist sich in Analogie zum Vorgehen beim harmonischen Oszillator die Einführung der Operatoren

$$\mathbf{j}_+ = \mathbf{j}_1 + \mathrm{i}\mathbf{j}_2, \quad \mathbf{j}_- = \mathbf{j}_+^\dagger = \mathbf{j}_1 - \mathrm{i}\mathbf{j}_2 \tag{3.10.5}$$

als nützlich. Dazu zeigen wir, dass diese Operatoren (ähnlich wie $\mathbf{a}^\dagger$ und $\mathbf{a}$ beim harmonischen Oszillator) den Eigenwert m von $\mathbf{j}_3$ um 1 erhöhen bzw. erniedrigen. Um das zu sehen, benötigen wir die Kommutatoren mit $\mathbf{j}_3$. Das ist sehr leicht mit Hilfe von (3.10.2) erledigt (*Nachrechnen!*):

$$\left[\mathbf{j}_3, \mathbf{j}_+\right] = \mathbf{j}_+, \quad \left[\mathbf{j}_3, \mathbf{j}_-\right] = -\mathbf{j}_-. \tag{3.10.6}$$

Daraus folgt nun

$$\mathbf{j}_3\mathbf{j}_+ |u_{\alpha m}\rangle = \left\{\left[\mathbf{j}_3, \mathbf{j}_+\right] + \mathbf{j}_+\mathbf{j}_3\right\} |u_{\alpha m}\rangle = (\mathbf{j}_+ + \mathbf{j}_+\mathbf{j}_3) |u_{\alpha m}\rangle = (m+1)\mathbf{j}_+ |u_{\alpha m}\rangle. \tag{3.10.7}$$

Da außerdem $\vec{\mathbf{j}}^2$ sowohl mit $\mathbf{j}_+$ als auch mit $\mathbf{j}_-$ vertauscht, folgt

$$\vec{\mathbf{j}}^2 \mathbf{j}_+ |u_{\alpha m}\rangle = \mathbf{j}_+ \vec{\mathbf{j}}^2 u_{\alpha m} = \alpha \mathbf{j}_+ |u_{\alpha m}\rangle. \tag{3.10.8}$$

Falls nun $\mathbf{j}_+ |u_{\alpha m}\rangle \neq 0$ ist, bedeutet (3.10.7), dass dies ein Eigenvektor von $\mathbf{j}_3$ zum Eigenwert $(m+1)$ und wegen (3.10.8), auch auch ein Eigenvektor von $\vec{\mathbf{j}}^2$ zum Eigenwert α ist. Anders gesagt, liefert $\mathbf{j}_+ |u_{\alpha m}\rangle$ entweder einen simultanen Eigenvektor $u_{\alpha,(m+1)}$ von $\vec{\mathbf{j}}^2$ und $\mathbf{j}_3$ oder den Nullvektor. Auf genau analoge Weise zeigt man (*Übung!*), dass

$$\mathbf{j}_3\mathbf{j}_- |u_{\alpha m}\rangle = (m-1)\mathbf{j}_- |u_{\alpha m}\rangle, \quad \vec{\mathbf{j}}^2 \mathbf{j}_- |u_{\alpha m}\rangle = \alpha \mathbf{j}_- |u_{\alpha m}\rangle \tag{3.10.9}$$

ist. Das bedeutet, dass $\mathbf{j}_- |u_{\alpha m}\rangle$ entweder einen simultanen Eigenvektor von $\vec{\mathbf{j}}^2$ und $\mathbf{j}_3$ zu den Eigenwerten α bzw. $(m-1)$ oder den Nullvektor liefert.

Als nächstes bemerken wir, dass für jede Drehimpulskomponente $\mathbf{j}_a$ und jeden Vektor $|\psi\rangle$

$$\left\langle \psi \middle| \mathbf{j}_a^2 \psi \right\rangle = \langle \mathbf{j}_a \psi | \mathbf{j}_a \psi \rangle \geq 0 \tag{3.10.10}$$

gilt. Dabei haben wir im ersten Schritt die Selbstadjungiertheit von $\mathbf{j}_a$ und im zweiten Schritt die positive Definitheit des Skalarprodukts verwendet. Daraus folgt aber

$$0 \leq \left\langle |u_{\alpha m}\rangle \middle| \mathbf{j}_3^2 |u_{\alpha m}\rangle \right\rangle = m^2 \leq \left\langle |u_{\alpha m}\rangle \middle| \vec{\mathbf{j}}^2 |u_{\alpha m}\rangle \right\rangle = \alpha. \tag{3.10.11}$$

Damit gilt

$$-\sqrt{\alpha} \leq m \leq \sqrt{\alpha}. \tag{3.10.12}$$

Demnach muss es für jedes α einen Eigenvektor von $\mathbf{j}_3$ mit maximalem Eigenwert $m = j$ geben, und für diesen muss

$$\mathbf{j}_+ u_{\alpha j} = 0 \tag{3.10.13}$$

gelten, weil andernfalls dieser Vektor ein Eigenvektor von $\mathbf{j}_3$ zum Eigenwert $j+1$ wäre, aber das kann nicht sein, weil j bereits der maximale Eigenwert von $\mathbf{j}_3$ bei vorgegebenem Eigenwert α von $\vec{\mathbf{j}}^2$ ist.
Wir können nun wiederholt $\mathbf{j}_-$ auf $u_{\alpha j}$ anwenden. Dabei erhält man nacheinander Eigenvektoren mit dem gleichen α aber $m=j-1$, $m=j-2$ usw. Allerdings kann dies nicht beliebig so weitergehen, denn wegen (3.10.12) muss es auch einen minimalen Eigenwert $m=j'$ von $\mathbf{j}_3$ geben, und für den dazugehörigen Eigenvektor gilt

$$\mathbf{j}_- \left| u_{\alpha j'} \right\rangle = 0. \tag{3.10.14}$$

Nun berechnen wir

$$\mathbf{j}_-\mathbf{j}_+ = (\mathbf{j}_1 - \mathrm{i}\mathbf{j}_2)(\mathbf{j}_1 + \mathrm{i}\mathbf{j}_2) = \mathbf{j}_1^2 + \mathbf{j}_2^2 + \mathrm{i}(\mathbf{j}_1\mathbf{j}_2 - \mathbf{j}_2\mathbf{j}_1) = \vec{\mathbf{j}}^2 - \mathbf{j}_3^2 - \mathbf{j}_3 \tag{3.10.15}$$

und analog (*nachrechnen*!)

$$\mathbf{j}_+\mathbf{j}_- = \vec{\mathbf{j}}^2 - \mathbf{j}_3^2 + \mathbf{j}_3. \tag{3.10.16}$$

Mit (3.10.13) folgt dann aus (3.10.15)

$$0 = \mathbf{j}_-\mathbf{j}_+ \left| u_{\alpha j} \right\rangle = [\alpha - j(j+1)] \left| u_{\alpha j} \right\rangle \Rightarrow \alpha = j(j+1). \tag{3.10.17}$$

Mit (3.10.14) und (3.10.16) erhalten wir

$$0 = \mathbf{j}_+\mathbf{j}_- \left| u_{\alpha j'} \right\rangle = [\alpha - j'(j'-1)] \left| u_{\alpha j'} \right\rangle \Rightarrow \alpha = j'(j'-1). \tag{3.10.18}$$

Daraus folgt

$$j(j+1) = j'(j'-1) \Rightarrow j_1' = j+1, \quad j_2' = -j. \tag{3.10.19}$$

Da aber $j' \leq j$ sein muss, folgt $j \geq 0$ und $j' = -j$.
Nun können wir aber durch wiederholte Anwendung von $\mathbf{j}_-$ auf $u_{\alpha j}$ alle Eigenfunktionen $|u_{\alpha m}\rangle$ mit $m \in \{j, j-1, \ldots, -j\}$ finden. Dabei müssen wir nach n-maliger Anwendung auf $m = j' = -j$ kommen, denn es muss bei nochmaliger Anwendung von $\mathbf{j}_-$ Null herauskommen, weil man ja sonst beliebig kleine Eigenwerte m von $\mathbf{j}_3$ erhielte, was aber nicht sein kann. Demnach gibt es also ein $n \in \mathbb{N}_0$ mit

$$j - n = -j \Rightarrow n = 2j \Rightarrow j = \frac{n}{2}, \tag{3.10.20}$$

d.h. $j \in \{0, 1/2, 1, 3/2, \ldots\}$, und für jedes j durchläuft $m \in \{j, j-1, \ldots, -j+1, -j\}$. Zu jedem j gibt es also $2j+1$ Eigenwerte m zu $\mathbf{j}_3$.

Zusammenfassung: Es ist üblich, statt dem Eigenwert α von $\vec{\mathbf{j}}^2$ die **Drehimpulsbetragsquantenzahl** $j = m_{\max}$ anzugeben. Demnach ist

$$\vec{\mathbf{j}}^2 \left| u_{jm} \right\rangle = j(j+1) \left| u_{jm} \right\rangle, \quad \mathbf{j}_3 u_{jm} = m \left| u_{jm} \right\rangle, \quad j \in \{0, 1/2, 1, 3/2, \ldots\}, \quad m \in \{-j, -j+1, \ldots, j-1, j\}. \tag{3.10.21}$$

Wir erhalten alle $\left| u_{jm} \right\rangle$ für gegebenes j aus der Rekursionsvorschrift

$$\mathbf{j}_+ \left| u_{jj} \right\rangle = 0, \quad \mathbf{j}_- \left| u_{jm} \right\rangle = N_{jm} \left| u_{j,m-1} \right\rangle. \tag{3.10.22}$$

Dabei gehen wir davon aus, dass alle $\left| u_{j,m} \right\rangle$ auf 1 normiert sind. Dann gilt einerseits

$$\left\langle \mathbf{j}_- u_{jm} \middle| \mathbf{j}_- u_{jm} \right\rangle = \left\langle N_{jm} u_{j,m-1} \middle| N_{jm} u_{j,m-1} \right\rangle = |N_{jm}|^2 \left\langle u_{j,m-1} \middle| u_{j,m-1} \right\rangle = |N_{jm}|^2 \tag{3.10.23}$$

und andererseits

$$\left\langle \mathbf{j}_- u_{jm} \middle| \mathbf{j}_- u_{jm} \right\rangle = \left\langle u_{jm} \middle| \mathbf{j}_+ \mathbf{j}_- u_{jm} \right\rangle \overset{(3.10.16)}{=} j(j+1) - m(m-1). \tag{3.10.24}$$

Bis auf einen unbestimmten Phasenfaktor ist also

$$\begin{aligned} N_{jm} &= \sqrt{j(j+1) - m(m-1)} = \sqrt{(j+m)(j-m+1)} \Rightarrow \\ \mathbf{j}_- \left| u_{jm} \right\rangle &= \sqrt{(j+m)(j-m+1)} \left| u_{j,m-1} \right\rangle. \end{aligned} \tag{3.10.25}$$

Für $m = -j$ wird $N_{jj} = 0$, wie es sein muss.

Mit diesen Definitionen können wir auch die Wirkung von $\mathbf{j}_+$ auf die $\left| u_{jm} \right\rangle$ erhalten. Aufgrund der obigen Überlegungen folgt nämlich

$$\mathbf{j}_+ \left| u_{jm} \right\rangle = N'_{jm} \left| u_{j,(m+1)} \right\rangle. \tag{3.10.26}$$

Um N'_{jm} zu erhalten, verwenden wir (3.10.15):

$$\begin{aligned} \mathbf{j}_- \mathbf{j}_+ \left| u_{jm} \right\rangle &= (\vec{\mathbf{j}}^2 - \mathbf{j}_3^2 - \mathbf{j}_3) \left| u_{jm} \right\rangle = [j(j+1) - m(m+1)] u_{jm} \\ &= N'_{jm} \mathbf{j}_- \left| u_{j,(m+1)} \right\rangle \overset{(3.10.22,3.10.25)}{=} N'_{jm} \sqrt{j(j+1) - m(m+1)} u_{jm}. \end{aligned} \tag{3.10.27}$$

Wegen $j(j+1) - m(m+1) = (j+m+1)(j-m)$ ergibt sich

$$\begin{aligned} N'_{jm} &= \sqrt{j(j+1) - m(m+1)} = \sqrt{(j+m+1)(j-m)} \Rightarrow \\ \mathbf{j}_+ \left| u_{jm} \right\rangle &= \sqrt{(j+m+1)(j-m)} \left| u_{j,m+1} \right\rangle. \end{aligned} \tag{3.10.28}$$

3.11 Der Bahndrehimpuls

Wir wenden uns nun wieder dem Teilchen ohne Spin zu, wo der einzige Drehimpuls der **Bahndrehimpuls** $\vec{\mathbf{L}} = \vec{\mathbf{x}} \times \vec{\mathbf{p}}$ ist. Auch hier ist es bequem, die dimensionslosen Drehimpulsoperatoren $\vec{\mathbf{l}} = \vec{\mathbf{L}}/\hbar$ einzuführen, und auch hier können wir freilich simultane Eigenvektoren $|u\rangle_{\ell m}$ von $\vec{\mathbf{L}}^2$ und $\mathbf{L}_3$ betrachten. Dabei soll

$$\vec{\mathbf{l}}^2 |u_{\ell m}\rangle = \ell(\ell+1) |u_{\ell m}\rangle, \quad \mathbf{l}_3 |u_{\ell m}\rangle = m |u_{\ell m}\rangle \tag{3.11.1}$$

gelten. Als erstes wollen wir nun zeigen, dass es aufgrund der Definition von $\vec{\mathbf{l}}$ aus Orts- und Impulsoperatoren für den Bahndrehimpuls keine halbzahligen Drehimpulsquantenzahlen gibt, d.h. es ist $\ell \in \mathbb{N}_0 = \{0, 1, 2, \ldots\}$. Dazu konstruieren wir ein VONS im Hilbertraum des Teilchens. Um es eindeutig zu charakterisieren, benötigen wir neben $\vec{\mathbf{l}}^2$ und $\mathbf{l}_3$ noch eine dritte dazu kompatible Observable. Wir verwenden hier $\mathbf{r} = \sqrt{\vec{\mathbf{x}}^2}$. Wegen (3.10.25) kommutieren nämlich, wie für einen Skalar zu erwarten, alle $\mathbf{l}_a$ mit $\vec{\mathbf{x}}^2$:

$$[\mathbf{l}_a, \mathbf{x}_b \mathbf{x}_b] = \mathrm{i}\epsilon_{abc}(\mathbf{x}_b \mathbf{x}_c + \mathbf{x}_c \mathbf{x}_b) = 0, \tag{3.11.2}$$

denn der Ausdruck in der Klammer ist symmetrisch unter Vertauschen der Indizes b und c, das Levi-Civita-Symbol aber antisymmetrisch. Es kommutiert also in der Tat $\mathbf{r}$ sowohl mit $\vec{\mathbf{l}}^2$ als auch mit $\vec{o}_3$. Wir können also davon ausgehen, dass die entsprechenden simultanen Eigenvektoren $|u_{r\ell m}\rangle$ ein VONS bilden. Demnach lassen sich alle verallgemeinerten Ortseigenvektoren in der Form

$$|\vec{x}\rangle = \int_0^\infty \mathrm{d}r \sum_\ell \sum_{m=-\ell}^{\ell} |u_{r\ell m}\rangle \langle u_{r\ell m} | \vec{x}\rangle \tag{3.11.3}$$

darstellen. Dabei haben wir verwendet, dass wegen $\vec{x} \in \mathbb{R}^3$ offenbar die Eigenwerte von $\mathbf{r}$ kontinuierlich sind und $r \geq 0$ ist. Wir haben uns hier auch *noch nicht* festgelegt, ob ℓ auch über die halbzahligen Werte laufen muss oder nicht.

Nun betrachten wir den speziellen Ortseigenvektor $|r\vec{e}_3\rangle$. Gemäß (3.7.40) gilt für eine Drehung um die 3-Achse

$$\mathbf{D}(\alpha\vec{e}_3)|r\vec{e}_3\rangle = \left|\hat{R}(\alpha\vec{e}_3)r\vec{e}_3\right\rangle = |r\vec{e}_3\rangle\,. \tag{3.11.4}$$

Nun ist aber gemäß (3.7.39)

$$\mathbf{D}(\alpha\vec{e}_3) = \exp(-\mathrm{i}\alpha\mathbf{j}_3) \tag{3.11.5}$$

und damit

$$0 = \frac{\mathrm{d}}{\mathrm{d}\alpha}\mathbf{D}(\alpha\vec{e}_3)|r\vec{e}_3\rangle = \mathbf{D}(\alpha\vec{e}_3)(-\mathrm{i})\mathbf{j}_3|r\vec{e}_3\rangle \;\Rightarrow\; \mathbf{j}_3|r\vec{e}_3\rangle = 0. \tag{3.11.6}$$

Damit ist aber gemäß (3.11.3)

$$\mathbf{j}_3|r\vec{e}_3\rangle = \int_0^\infty \mathrm{d}r \sum_\ell \sum_{m=-\ell}^{\ell} |u_{r\ell m}\rangle\, m\, \langle u_{r\ell m}\,|\,r\vec{e}_3\rangle = 0. \tag{3.11.7}$$

Da die $|u_{r\ell m}\rangle$ ein VONS bilden, muss demnach

$$m\,\langle u_{r\ell m}\,|\,r\vec{e}_3\rangle = 0 \tag{3.11.8}$$

sein. Demnach ist $\langle u_{r\ell m}\,|\,r\vec{e}_3\rangle \neq 0$ nur für $m = 0$, d.h. es gilt

$$|r\vec{e}_3\rangle = \int_0^\infty \mathrm{d}r \sum_{\ell=0}^{\infty} \left|u_{r\ell,m=0}\right\rangle\left\langle u_{r\ell,m=0}\,\middle|\,r\vec{e}_3\right\rangle = \int_0^\infty \mathrm{d}r \sum_{\ell=0}^{\infty} C_\ell(r)\left|u_{r\ell,m=0}\right\rangle. \tag{3.11.9}$$

Dabei haben wir verwendet, dass der Eigenwert $m = 0$ wegen $m \in \{\ell, \ell-1, \ldots, -\ell+1, -\ell\}$ nur für ganzzahlige, nicht aber für halbzahlige ℓ auftreten kann.

Nun kann man aber jeden Vektor $\vec{x}$ durch eine geeignete Drehung des Vektors $r\vec{e}_3$ erreichen, d.h. es gibt einen „Drehvektor“ $\vec{\phi}$, so dass $\vec{x} = \hat{R}(\vec{\phi})r\vec{e}_3$ mit $r = |\vec{x}|$ ist. Damit gilt

$$|\vec{x}\rangle = \mathbf{D}(\vec{\phi})|r\vec{e}_3\rangle = \int_0^\infty \mathrm{d}r \sum_{\ell=0}^{\infty} C_\ell(r)\mathbf{D}(\vec{\phi})\left|u_{r,\ell,m=0}\right\rangle. \tag{3.11.10}$$

Da sowohl $\mathbf{r}$ als auch $\vec{\mathbf{l}}^2$ mit $\vec{\mathbf{l}}$ und damit auch mit $\mathbf{D}(\vec{\phi}) = \exp(-\mathrm{i}\vec{\phi}\cdot\vec{\mathbf{l}})$ vertauscht, ist $\mathbf{D}(\vec{\phi})\left|u_{r,\ell,m=0}\right\rangle$ ein Eigenvektor von $\mathbf{r}$ zum Eigenwert r und von $\vec{\mathbf{l}}^2$ zum Eigenwert $\ell(\ell+1)$ und damit

$$\mathbf{D}(\vec{\phi})\left|u_{r,\ell,m=0}\right\rangle = \sum_{m=-\ell}^{\ell} \mathrm{Y}_{\ell m}\left|u_{r,\ell,m}\right\rangle. \tag{3.11.11}$$

Dies in (3.11.10) eingesetzt ergibt nun

$$|\vec{x}\rangle = \int_0^\infty \mathrm{d}r \sum_{\ell=0}^{\infty} \sum_{m=-\ell}^{\ell} C_\ell(r)\mathrm{Y}_{\ell m}\left|u_{r,\ell,m}\right\rangle. \tag{3.11.12}$$

Da die verallgemeinerten Ortseigenvektoren $|\vec{x}\rangle$ ein verallgemeinertes VONS bilden, sind auch die $\left|u_{r,\ell,m}\right\rangle$ vollständig. Das bedeutet, dass die Bahndrehimpulseigenzustände nur ganzzahlige Drehimpulsbetragsquantenzahlen ℓ besitzen.

Im Folgenden sind wir besonders an Energieeigenzuständen interessiert, um z.B. die Energieniveaus des Wasserstoffatoms zu berechnen und damit die in der Atomphysik genau vermessenen **Spektrallinien** zu erklären. Betrachten wir also das Eigenwertproblem für den Hamilton-Operator für ein Teilchen, das sich in einem beliebigen Radialpotential bewegt. Der Hamilton-Operator lautet[3]

$$\mathbf{H} = \frac{1}{2\mu}\vec{\mathbf{p}}^2 + V(\mathbf{r}), \quad \mathbf{r} = |\vec{\mathbf{x}}|. \tag{3.11.13}$$

Für das Wasserstoffatom ist z.B. V das Coulomb-Potential des elektrostatischen Feldes des Protons. Da die Protonenmasse $m_\mathrm{p} \simeq 938\,\mathrm{MeV}/c^2$ viel größer ist als die Elektronenmasse $m_\mathrm{e} \simeq 0{,}511\,\mathrm{MeV}/c^2$ können wir nämlich annehmen, dass das Proton fest im Ursprung des Koordinatensystems sitzt und sich das Elektron in dessen elektrostatischen Feld bewegt. Dabei kann man in sehr guter Näherung ein einfaches Coulombpotential

$$V(\mathbf{r}) = -\frac{e^2}{4\pi\epsilon_0 \mathbf{r}} \tag{3.11.14}$$

annehmen.

Offensichtlich ist der Hamilton-Operator (3.11.13) symmetrisch unter Drehungen, da in ihm nur skalare Operatoren vorkommen, d.h. der Hamilton-Operator vertauscht mit allen Drehimpulskomponenten:

$$[\mathbf{L}_a, \mathbf{H}] = 0. \tag{3.11.15}$$

Das bedeutet, dass wir in diesem Fall stets einen kompletten orthonormierten Satz von gemeinsamen Eigenvektoren $|u_{E\ell m}\rangle$ für $\mathbf{H}, \vec{\mathbf{l}}^2$ und $\mathbf{l}_3$ finden können:

$$\mathbf{H}\,|u_{E\ell m}\rangle = E\,|u_{E\ell m}\rangle, \quad \vec{\mathbf{l}}^2\,|u_{E\ell m}\rangle = \ell(\ell+1)\,|u_{E\ell m}\rangle, \quad \mathbf{l}_3\,|u_{E\ell m}\rangle = m\,|u_{E\ell m}\rangle. \tag{3.11.16}$$

Wir kennen auch bereits die Lösungen für das Eigenwertproblem für $\vec{\mathbf{l}}^2$ und $\mathbf{l}_3$. Wir wollen im Folgenden allerdings die entsprechenden Eigenfunktionen in der Ortsdarstellung finden. Wegen der Rotationssymmetrie des Hamilton-Operators bietet es sich dabei an, mit Kugelkoordinaten zu arbeiten (s. [Hee18a]). Es gilt

$$\underline{x} = r\begin{pmatrix} \sin\vartheta\cos\varphi \\ \sin\vartheta\sin\varphi \\ \cos\vartheta \end{pmatrix}. \tag{3.11.17}$$

Die Kugelkoordinaten sind orthogonale Koordinaten, d.h. die Koordinatenlinien schneiden sich alle in rechten Winkel. Die entsprechenden rechtshändigen kartesischen Tangentenvektoren an die Koordinatenlinien sind

$$\underline{e}_r = \begin{pmatrix} \sin\vartheta\cos\varphi \\ \sin\vartheta\sin\varphi \\ \cos\vartheta \end{pmatrix}, \quad \underline{e}_\vartheta = \begin{pmatrix} \cos\vartheta\cos\varphi \\ \cos\vartheta\sin\varphi \\ -\sin\vartheta \end{pmatrix}, \quad \underline{e}_\varphi = \begin{pmatrix} -\sin\varphi \\ \cos\varphi \\ 0 \end{pmatrix}. \tag{3.11.18}$$

Da in der Ortsdarstellung $\vec{\mathbf{p}} = -\mathrm{i}\hbar\vec{\nabla}$ und folglich $\vec{\mathbf{p}}^2 = -\hbar^2\Delta$ ist, benötigen wir nun lediglich die Differentialoperatoren in Kugelkoordinaten in ihrer Wirkung auf eine beliebige Ortswellenfunktion $\psi(\vec{x}) = \langle\vec{x}\,|\,\psi\rangle$, (vgl. [Hee18a]):

$$\vec{\nabla}\psi = \vec{e}_r\partial_r\psi + \vec{e}_\vartheta\frac{1}{r}\partial_\vartheta\psi + \vec{e}_\varphi\frac{1}{r\sin\vartheta}\partial_\varphi\psi, \tag{3.11.19}$$

$$\Delta\psi = \frac{1}{r}\partial_r^2(r\psi) + \frac{1}{r^2\sin\vartheta}\partial_\vartheta(\psi\sin\vartheta) + \frac{1}{r^2\sin^2\vartheta}\partial_\varphi^2\psi. \tag{3.11.20}$$

[3] Wir bezeichnen die Masse des Teilchens in Zentralkraftproblemen mit μ, um Verwechslungen mit der „magnetischen Quantenzahl“ m, also den Eigenwerten von $\mathbf{l}_3$ zu vermeiden.

Der dimensionslose Drehimpulsoperator ergibt sich damit zu

$$\vec{\mathrm{l}}\psi = -\mathrm{i}\vec{x}\times\vec{\nabla}\psi = -\mathrm{i}r\vec{e}_r\times\vec{\nabla}\psi = -\mathrm{i}\vec{e}_r\times\left(\vec{e}_\vartheta\partial_\vartheta\psi + \vec{e}_\varphi\frac{1}{\sin\vartheta}\partial_\varphi\psi\right) = -\mathrm{i}\left(\vec{e}_\varphi\partial_\vartheta - \frac{1}{\sin\vartheta}\vec{e}_\vartheta\partial_\varphi\right)\psi. \quad (3.11.21)$$

Wir bemerken, dass die drei Drehimpulskomponentenoperatoren nur Ableitungen nach den Winkeln ϑ und φ enthalten. Das ist auch zu erwarten, da der Drehimpuls Drehungen erzeugt und daher keinen Einfluss auf die Länge des Ortsvektors haben kann.

Um $\vec{\mathrm{l}}^2$ zu berechnen, ist es bequemer, zunächst die kartesischen Komponenten zu berechnen. Dazu benötigen wir nur, dass gemäß (3.11.18

$$\vec{e}_\vartheta = (\vec{e}_1\cos\varphi + \vec{e}_2\sin\varphi)\cos\vartheta - \vec{e}_3\sin\vartheta, \quad \vec{e}_\varphi = -\vec{e}_1\sin\varphi + \vec{e}_2\cos\varphi \quad (3.11.22)$$

ist. Damit erhalten wir aus (3.11.21)

$$\mathrm{l}_1\psi = \vec{e}_1\cdot\vec{\mathrm{l}}\psi = +\mathrm{i}(\sin\varphi\partial_\vartheta + \cot\vartheta\cos\varphi\partial_\varphi)\psi, \quad (3.11.23)$$

$$\mathrm{l}_2\psi = \vec{e}_2\cdot\vec{\mathrm{l}}\psi = +\mathrm{i}(-\cos\varphi\partial_\vartheta + \cot\vartheta\sin\varphi\partial_\varphi)\psi, \quad (3.11.24)$$

$$\mathrm{l}_3\psi = \vec{e}_3\cdot\vec{\mathrm{l}}\psi = -\mathrm{i}\partial_\varphi. \quad (3.11.25)$$

Daraus ergibt sich nach einiger Rechnung (*Nachprüfen!*)

$$\vec{\mathrm{l}}^2\psi = -\left[\frac{1}{\sin\vartheta}\partial_\vartheta(\sin\vartheta\partial_\vartheta\psi) + \frac{1}{\sin^2\vartheta}\partial_\varphi^2\psi\right]. \quad (3.11.26)$$

Mit (3.11.21) erhalten wir

$$\Delta\psi = \frac{1}{r}\partial_r^2(r\psi) - \frac{1}{r^2}\vec{\mathrm{l}}^2\psi. \quad (3.11.27)$$

Betrachten wir nun die Gleichungen für das simultane Eigenwertproblem (3.11.16) in der Ortsdarstellung. Es ist im Folgenden $u_{E\ell m} = \langle\vec{x}\,|\,u_{E\ell m}\rangle$. Für l_3 ergibt sich

$$\mathrm{l}_3 u_{E\alpha m} = -\mathrm{i}\partial_\varphi u_{E,\alpha,m} = m u_{E,\alpha,m} \Rightarrow u_{E,\alpha,m} = v_{E\ell m}(r,\vartheta)\exp(\mathrm{i}m\varphi). \quad (3.11.28)$$

Weiter wird

$$\vec{\mathrm{l}}^2 u_{E\ell m} = \ell(\ell+1)u_{E\ell m} \Rightarrow -\left[\frac{1}{\sin\vartheta}\partial_\vartheta(\sin\vartheta\partial_\vartheta v_{E\ell m}) - \frac{m^2}{\sin^2\vartheta}v_{E\ell m}\right] = \ell(\ell+1)v_{E\ell m}. \quad (3.11.29)$$

Daraus ergibt sich, dass auch $v_{E\ell m}$ in eine Funktion von ϑ und r faktorisiert:

$$v_{E\ell m} = R_{E\ell}(r)\Theta_{\ell m}(\vartheta) \overset{(3.11.28)}{\Rightarrow} u_{E\ell m} = R_{E\ell}(r)\Theta_{\ell m}(\vartheta)\exp(\mathrm{i}m\varphi). \quad (3.11.30)$$

Damit folgt nämlich aus (3.11.29)

$$-\left[\frac{1}{\sin\vartheta}\partial_\vartheta(\sin\vartheta\partial_\vartheta\Theta_{\ell m}) - \frac{m^2}{\sin^2\vartheta}\Theta_{\ell m}\right] = \ell(\ell+1)\Theta_{\ell m}(\vartheta) \quad (3.11.31)$$

und mit (3.11.13) und (3.11.27)

$$\mathbf{H}u_{E\ell m} = E u_{E\ell m} \Rightarrow -\frac{\hbar^2}{2\mu r}\partial_r^2(rR_{E\ell}) + \left[\frac{\alpha}{2\mu r^2} + V(r)\right]R_{E\ell} = ER_{E\ell}. \quad (3.11.32)$$

Wir können also zunächst das simultane Eigenwertproblem von $\vec{\mathbf{l}}^2$ und $\mathbf{l}_3$ separat lösen, d.h. wir betrachten die entsprechende simultane Eigenfunktion

$$\mathrm{Y}_{\ell m}(\vartheta,\varphi)=\Theta_{\ell m}(\vartheta)\exp(\mathrm{i}m\varphi). \tag{3.11.33}$$

Dazu bemerken wir, dass das Skalarprodukt für Funktionen im Hilbertraum in Kugelkoordinaten durch

$$\langle\psi_1|\psi_2\rangle=\int_{\mathbb{R}^3}\mathrm{d}^3x\,\psi_1^*(\vec{x})\psi_2(\vec{x})=\int_0^\infty\mathrm{d}r\int_0^\pi\mathrm{d}\vartheta\int_0^{2\pi}\mathrm{d}\varphi\,r^2\sin\vartheta\,\psi_1^*(r,\vartheta,\varphi)\psi_2(r,\vartheta,\varphi) \tag{3.11.34}$$

gegeben ist.
Wir definieren daher das Skalarprodukt für Funktionen $f(\vartheta,\varphi)$, also für Funktionen, die auf der **Einheitskugelschale** Ω um den Ursprung des Koordinatensystems definiert sind, durch

$$\langle f_1|f_2\rangle_\Omega=\int_0^\pi\mathrm{d}\vartheta\int_0^{2\pi}\mathrm{d}\varphi\sin\vartheta\,f_1^*(\vartheta,\varphi)f_2(\vartheta,\varphi). \tag{3.11.35}$$

Da $\mathbf{l}_3$ und $\vec{\mathbf{l}}^2$ selbstadjungiert sind, bilden die $\mathrm{Y}_{\ell m}$ ein vollständiges System von simultanen Eigenfunktionen dieser beiden Operatoren, d.h. es gilt

$$\langle\mathrm{Y}_{\ell'm'}|\mathrm{Y}_{\ell m}\rangle_\Omega=\delta_{\ell'\ell}\delta_{m'm}. \tag{3.11.36}$$

Mit Hilfe des allgemeinen Formalismus' aus Abschnitt 3.10 können wir nun diese simultanen Eigenfunktionen von $\vec{\mathbf{l}}^2$ und $\mathbf{l}_z$ bestimmen. Dazu verwenden wir (3.10.22) und die Rekursionsvorschrift (3.10.25) in der konkreten Realisierung der Bahndrehimpulseigenfunktionen an, d.h. wir bestimmen zuerst $\mathrm{Y}_{\ell\ell}(\vartheta,\varphi)$ aus

$$\mathbf{l}_+\mathrm{Y}_{\ell\ell}=0 \tag{3.11.37}$$

und berechnen dann die übrigen $\mathrm{Y}_{\ell m}$ mittels (3.10.25)

$$\mathrm{Y}_{\ell,m-1}=\frac{1}{\sqrt{(\ell+m)(\ell-m+1)}}\mathbf{l}_-\mathrm{Y}_{\ell m}. \tag{3.11.38}$$

Dazu verwenden wir wieder Kugelkoordinaten, wobe es sich als nüzlich erweist, statt $\vartheta\in[0,\pi]$ vorübergehend die Variable $\xi=\cos\vartheta\in[-1,1]$ einzuführen. Um bequem rechnen zu können, rechnen wir zunächst die kartesischen Komponenten des Bahndrehimpulsoperators (3.11.23) in die neuen Koordinaten (ξ,φ) um. Dazu benötigen wir wegen $\partial_\vartheta\xi=\partial_\vartheta\cos\vartheta=-\sin\vartheta=-\sqrt{1-\xi^2}$ nur

$$\partial_\vartheta\psi=\frac{\partial\xi}{\partial\vartheta}\partial_\xi\psi=-\sin\vartheta\,\partial_\xi\psi=-\sqrt{1-\xi^2}\partial_\xi\psi. \tag{3.11.39}$$

Dabei haben wir verwendet, dass für $\vartheta\in[0,\pi]$ stets $\sin\vartheta\geq 0$ ist und folglich eindeutig $\sin\vartheta=\sqrt{1-\cos^2\vartheta}=\sqrt{1-\xi^2}$ ist. Damit erhält man aus (3.11.23) nach einiger Rechnung (*Übung*)

$$\mathbf{l}_+\psi=\exp(\mathrm{i}\varphi)\left(\mathrm{i}\frac{\xi}{\sqrt{1-\xi^2}}\partial_\varphi-\sqrt{1-\xi^2}\partial_\xi\right)\psi, \tag{3.11.40}$$

$$\mathbf{l}_-\psi=\exp(-\mathrm{i}\varphi)\left(\mathrm{i}\frac{\xi}{\sqrt{1-\xi^2}}\partial_\varphi+\sqrt{1-\xi^2}\partial_\xi\right)\psi, \tag{3.11.41}$$

$$\mathbf{l}_3\psi=-\mathrm{i}\partial_\varphi\psi. \tag{3.11.42}$$

Das Eigenwertproblem für $\mathbf{l}_3$ ist sehr leicht zu lösen. Daraus hat sich oben (3.11.33) ergeben. Diese Gleichung schreiben wir nun nach Einführung der neuen Variablen $\xi = \cos\vartheta$ in der Form

$$\mathrm{Y}_{\ell m}(\vartheta,\varphi) = \Xi_{\ell m}(\xi)\exp(\mathrm{i}m\varphi). \tag{3.11.43}$$

Nun verwenden wir als erstes (3.11.37). Nach einer einfachen Rechnung (*Übung!*) ergibt sich daraus eine Differentialgleichung für $\Xi_{\ell\ell}$:

$$\mathbf{l}_+\mathrm{Y}_{\ell\ell} = 0 \Rightarrow \left(\frac{\xi}{\sqrt{1-\xi^2}}\ell + \sqrt{1-\xi^2}\partial_\xi\right)\Xi_{\ell\ell} = 0. \tag{3.11.44}$$

Diese Differentialgleichung lässt sich leicht durch Trennung der Variablen lösen:

$$\ln\left(\frac{\Xi_{\ell\ell}}{C}\right) = \frac{\ell}{2}\ln(1-\xi^2) \Rightarrow \Xi_{\ell\ell} = C(1-\xi^2)^{\ell/2}. \tag{3.11.45}$$

Die Integrationskonstante C ergibt sich (abgesehen von einem beliebigen Phasenfaktor) aus der Normierungsbedingung für die Kugelflächenfunktionen:

$$\int_0^\pi \mathrm{d}\vartheta \int_0^{2\pi} \mathrm{d}\varphi \sin\vartheta |\mathrm{Y}_{\ell\ell}(\vartheta,\varphi)|^2 = 2\pi\int_{-1}^1 \mathrm{d}\xi |\Xi_{\ell\ell}(\xi)|^2 = 2\pi|C|^2\int_{-1}^1 \mathrm{d}\xi(1-\xi^2)^\ell \overset{!}{=} 1\,. \tag{3.11.46}$$

Für das Integral über ξ lässt sich mit Hilfe der partiellen Integration eine Rekursionsformel herleiten

$$\begin{aligned} I_\ell &= \int_{-1}^1 \mathrm{d}\xi\, 1\cdot(1-\xi^2)^\ell = \xi(1-\xi^2)^\ell\Big|_{u=-1}^1 + 2\ell\int_{-1}^1 \mathrm{d}\xi\,\xi^2(1-\xi^2)^{\ell-1} \\ &= 2\ell\int_{-1}^1 \mathrm{d}\xi[1-(1-\xi^2)](1-\xi^2)^{\ell-1} = 2\ell(I_{\ell-1}-I_\ell) \end{aligned} \tag{3.11.47}$$

und damit

$$I_\ell = \frac{2\ell}{2\ell+1}I_{\ell-1}. \tag{3.11.48}$$

Für $\ell = 0$ erhält man $I_0 = 2$ und damit durch Iteration

$$I_\ell = \frac{2(2^\ell\ell!)^2}{(2\ell+1)!}. \tag{3.11.49}$$

Setzt man dies in die Normierungsbedingung (3.11.46) ein, erhält man schließlich

$$C = \frac{(-1)^\ell}{2^\ell\ell!}\sqrt{\frac{(2\ell+1)!}{4\pi}}. \tag{3.11.50}$$

Dabei ist der Faktor $(-1)^\ell = \pm 1$ die in der Literatur übliche Phasenwahl.
Es ist also

$$\Xi_{\ell\ell}(\xi) = \frac{(-1)^\ell}{2^\ell\ell!}\sqrt{\frac{(2\ell+1)!}{4\pi}}(1-\xi^2)^{\ell/2}. \tag{3.11.51}$$

Um daraus die $\Xi_{\ell m}$ für die übrigen $m \in \{-\ell, -\ell+1, \ldots, \ell-1, \ell\}$ zu erhalten, müssen wir die Rekursionsrelation (3.11.38) anwenden. Beginnen wir mit (3.11.51), finden wir damit (*Nachrechnen!*)

$$\Xi_{\ell m}(\xi) = \frac{(-1)^\ell}{2^\ell\ell!}\sqrt{\frac{2\ell+1}{4\pi}\frac{(\ell+m)!}{(\ell-m)!}}(1-\xi^2)^{-m/2}\partial_\xi^{\ell-m}(1-\xi^2)^\ell. \tag{3.11.52}$$

Mit dieser Definition sind die Kugelflächenfunktionen automatisch normiert. Da sie die simultanen Eigenfunktionen der selbstadjungierten Operatoren $\vec{\mathrm{l}}^2$ und l_3 sind, sind sie auch orthogonal zueinander, d.h. es gilt

$$\begin{aligned}\left\langle \mathrm{Y}_{\ell_1 m_1} \middle| \mathrm{Y}_{\ell_2 m_2} \right\rangle_\Omega &= \int_0^\pi \mathrm{d}\vartheta \int_0^{2\pi} \mathrm{d}\varphi \sin\vartheta \Xi_{\ell_1 m_1}(\cos\vartheta)\Xi_{\ell_2 m_2}(\cos\vartheta)\exp[\mathrm{i}(m_2-m_1)] \\ &= 2\pi\delta_{m_1 m_2}\int_{-1}^1 \mathrm{d}\xi\, \Xi_{\ell_1 m_1}(\xi)\Xi_{\ell_2 m_2}(\xi) \\ &= \delta_{\ell_1\ell_2}\delta_{m_1 m_2}.\end{aligned} \tag{3.11.53}$$

Dabei haben wir verwendet, dass gemäß (3.11.52) $\Xi_{\ell m}(\xi) \in \mathbb{R}$ ist. Aus (3.11.53) folgt für die Normierung der $\Xi_{\ell m}$

$$\int_{-1}^1 \mathrm{d}\xi\, \Xi_{\ell_1 m}(\xi)\Xi_{\ell_2 m} = \frac{1}{2\pi}\delta_{\ell_1\ell_2}. \tag{3.11.54}$$

Man kann nun freilich auch von $\mathrm{Y}_{\ell,-\ell}$ ausgehen und durch wiederholte Anwendung von l_+ die übrigen $\mathrm{Y}_{\ell m}$ gewinnen. Um $\Xi_{\ell,-\ell}$ zu berechnen, verwenden wir (3.11.52) mit $m=-\ell$:

$$\Xi_{\ell,-\ell}(\xi) = \frac{(-1)^\ell}{2^\ell \ell!}\sqrt{\frac{2\ell+1}{4\pi(2\ell)!}}(1-\xi^2)^{\ell/2}\partial_\xi^{2\ell}(1-\xi^2)^\ell. \tag{3.11.55}$$

Die Ableitung ist leicht berechnet, denn $(1-\xi^2)^\ell$ ist ein Polynom 2ℓ-ten Grades, und leitet man dieses 2ℓ mal nach u ab, bleibt nur der Beitrag vom Term $\propto \xi^{2\ell}$ übrig, also $\partial_\xi^{2\ell}(-1)^\ell\xi^{2\ell} = (-1)^\ell(2\ell)!$. Damit wird

$$\Xi_{\ell,-\ell} = \frac{1}{2^\ell \ell!}\sqrt{\frac{(2\ell+1)!}{4\pi}}(1-\xi^2)^{\ell/2} = (-1)^\ell \Xi_{\ell\ell}. \tag{3.11.56}$$

Wendet man nun l_+ fortgesetzt auf $\mathrm{Y}_{\ell,-\ell} = \exp(-\mathrm{i}\ell\varphi)\Xi_{\ell,-\ell}(\cos\vartheta)$ an, ergibt sich durch Iteration die Beziehungen (*Nachrechnen*)

$$\begin{aligned}\Xi_{\ell,m} &= \frac{(-1)^{\ell+m}}{2^\ell \ell!}\sqrt{\frac{2\ell+1}{4\pi}\frac{(\ell-m)!}{(\ell+m)!}}(1-\xi^2)^{m/2}\partial_\xi^{\ell+m}(1-\xi^2)^\ell, \\ \mathrm{Y}_{\ell,-m} &= (-1)^m \mathrm{Y}^*_{\ell m}.\end{aligned} \tag{3.11.57}$$

Wichtig ist auch noch der Spezialfall $m=0$. Man definiert damit die sog. **Legendre-Polynome** P_ℓ vermöge

$$\Xi_{\ell 0}(\xi) = \sqrt{\frac{2\ell+1}{4\pi}}P_\ell(\xi). \tag{3.11.58}$$

Mit (3.11.52) für $m=0$ erhält man

$$P_\ell(\xi) = \frac{1}{2^\ell \ell!}\partial_\xi^\ell(\xi^2-1)^\ell. \tag{3.11.59}$$

Aus der allgemeinen Beziehung (3.11.53) folgt

$$\int_{-1}^1 \mathrm{d}\xi\, P_{\ell_1}(\xi)P_{\ell_2}(\xi) = \frac{2}{2\ell_1+1}\delta_{\ell_1\ell_2}. \tag{3.11.60}$$

Nun können wir (3.11.31) in der Form

$$-\partial_\xi[(1-\xi^2)\partial_\xi\Xi_{\ell m}] + \frac{m^2}{1-\xi^2}\Xi_{\ell m} = \ell(\ell+1)\Xi_{\ell m} \tag{3.11.61}$$

schreiben. Setzt man hierin $m = 0$, folgt die Differentialgleichung für die Legendre-Polynome

$$\partial_\xi[(1-\xi^2)\partial_\xi P_\ell] = -\ell(\ell+1)P_\ell. \tag{3.11.62}$$

Das ist die Eigenwertgleichung für den Operator $\partial_\xi[(1-\xi^2)\partial_\xi]$. Wir können leicht zeigen, dass dies ein selbstadjungierter Operator auf dem Hilbertraum der über das Intervall $[-1,1]$ quadratintegrablen Funktionen ist, denn für zwei beliebige Funktionen $\Xi_1(\xi)$ und $\Xi_2(\xi)$ findet man durch partielle Integration

$$\begin{aligned}\left\langle \Xi_1 \middle| \partial_\xi[(1-\xi^2)\partial_\xi \Xi_2] \right\rangle &= \int_{-1}^{1} \mathrm{d}\xi\, \Xi_1^*(\xi)\partial_\xi[(1-\xi^2)\partial_\xi \Xi_2(\xi)] \\ &= -\int_{-1}^{1} \mathrm{d}\xi\, [\partial_\xi \Xi_1^*(\xi)](1-\xi^2)\partial_\xi \Xi_2(\xi) \\ &= +\int_{-1}^{1} \mathrm{d}\xi\, \left\{\partial_\xi[(1-\xi^2)\partial_\xi \Xi_1(\xi)\right\}^* \Xi_2(\xi) \\ &= \left\langle \partial_\xi[(1-\xi^2)\partial_\xi \Xi_1] \middle| \Xi_2 \right\rangle.\end{aligned} \tag{3.11.63}$$

Damit bilden die P_ℓ ein vollständiges orthogonales System auf dem Hilbertraum $\mathrm{L}^2([-1,1])$ der über das Intervall $[-1,1]$ quadratintegrablen Funktionen.

Zum Schluss betrachten wir noch das oft wichtige Verhalten der Kugelflächenfunktionen unter **räumlichen Spiegelungen**. Dazu definieren wir die Kugelflächenfunktionen als Funktion eines beliebigen Einheitsvektors $\mathrm{Y}_{\ell m}(\vec{n}) \equiv \mathrm{Y}_{\ell m}(\vartheta,\varphi)$ mit $\vec{n} = (\sin\vartheta\cos\varphi, \sin\vartheta\sin\varphi, \cos\vartheta)$ und fragen, wie sich diese Funktionen unter einer Raumspiegelung $\vec{n} \to -\vec{n}$ verhalten. Dazu betrachten wir (3.11.43). Zunächst bedeutet die Raumspiegelung, dass $\xi = \cos\vartheta \to -\xi$ (oder $\vartheta \to \pi - \vartheta$) sein muss, was $\sin\vartheta = \sqrt{1-\xi^2} \to \sin\vartheta$ impliziert und $\varphi \to \varphi + \pi$, d.h. $\cos\varphi \to -\cos\varphi$ und $\sin\varphi \to -\sin\varphi$, so dass in der Tat $\vec{n} \to -\vec{n}$ wird. Zunächst folgt aus (3.11.57), dass

$$\Xi_{\ell m}(-\xi) = (-1)^{\ell+m}\Xi_{\ell m}(\xi) \tag{3.11.64}$$

ist, denn $(1-\xi^2)^\ell$ ist eine gerade Funktion. Leitet man diese Funktion nach $\xi = \cos\vartheta$ ab, entsteht eine ungerade Funktion. Leitet man diese Funktion nochmals ab, erhält man wieder eine gerade Funktion usw. Damit folgt aber in der Tat (3.11.64). Damit ist vermöge (3.11.43)

$$\begin{aligned}\mathrm{Y}_{\ell m}(-\vec{n}) &= \mathrm{Y}_{\ell m}(\pi-\vartheta, \varphi+\pi) = \Xi_{\ell m}(-\xi)\exp[\mathrm{i}m(\varphi+\pi)] \\ &= (-1)^{\ell+m}\Xi_{\ell m}(\xi)\exp(\mathrm{i}m\varphi)(-1)^m = (-1)^\ell \mathrm{Y}_{\ell m}(\vartheta,\varphi) \Rightarrow \\ \mathrm{Y}_{\ell m}(-\vec{n}) &= (-1)^\ell \mathrm{Y}_{\ell m}(\vec{n}).\end{aligned} \tag{3.11.65}$$

Mehr interessante Eigenschaften der Kugelflächenfunktionen und Legendre-Polynome, sowie wichtige Anwendungen in der Physik, werden in Anhang B.3 zusammengestellt.

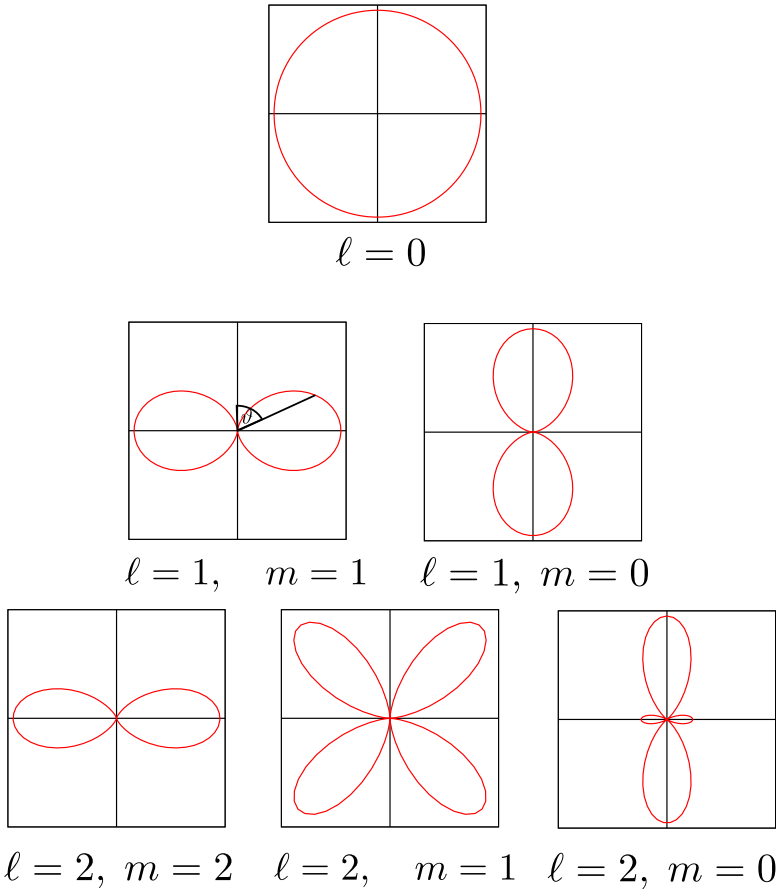


Abbildung 3.2: *Polardiagramme für* $|Y_{\ell m}|^2$. *Die Diagramme sind in 3 Dimensionen um die vertikale* z*-Achse rotiert zu denken. Die Länge zu einem Punkt der jeweiligen Kurve entspricht* $|Y_{\ell m}(\vartheta,\varphi)|^2$ *in Abhängigkeit zum Winkel zwischen der* z*-Achse und diesem Punkt.*

Wir geben noch die niedrigsten Ordnungen für $\ell \in \{0,1,2\}$ der Kugelflächenfunktionen explizit an:

$$\mathrm{Y}_{00}(\vartheta,\varphi) = \frac{1}{\sqrt{4\pi}}, \tag{3.11.66}$$

$$\mathrm{Y}_{11}(\vartheta,\varphi) = -\mathrm{Y}^*_{1,-1}(\vartheta,\varphi) = -\sqrt{\frac{3}{8\pi}}\sin\vartheta\exp(\mathrm{i}\varphi), \tag{3.11.67}$$

$$\mathrm{Y}_{10}(\vartheta,\varphi) = \sqrt{\frac{3}{4\pi}}\cos\vartheta, \tag{3.11.68}$$

$$\mathrm{Y}_{22}(\vartheta,\varphi) = \mathrm{Y}^*_{2,-2}(\vartheta,\varphi) = \sqrt{\frac{15}{32\pi}}\sin^2\vartheta\exp(2\mathrm{i}\varphi), \tag{3.11.69}$$

$$\mathrm{Y}_{21}(\vartheta,\varphi) = -\mathrm{Y}^*_{2,-1}(\vartheta,\varphi) = -\sqrt{\frac{15}{8\pi}}\cos\vartheta\sin\vartheta\exp(\mathrm{i}\varphi), \tag{3.11.70}$$

$$\mathrm{Y}_{20}(\vartheta,\varphi) = \sqrt{\frac{5}{16\pi}}(3\cos^2\vartheta - 1). \tag{3.11.71}$$

Schließlich geben wir noch die Entwicklung der Ortseigenvektoren $|\vec{x}\rangle$ nach den simultanenen Eigenvektoren von $r = |\vec{x}|$, $\vec{\mathrm{l}}^2$ und l_3, also $|r,\ell,m\rangle$ an. Dazu bemerken wir, dass offenbar

$$\langle \vec{x} \,|\, u_{r\ell m}\rangle = C_\ell(r)\delta(r - |\vec{x}|)\mathrm{Y}_{\ell m}(\vartheta,\varphi) \tag{3.11.72}$$

sein muss, wobei ϑ und φ die entsprechenden Winkel der Kugelkoordinaten von $\vec{x}$ sind. Außerdem haben wir verwendet, dass $\mathbf{r} = \sqrt{\vec{\mathbf{x}}^2}$ ist. Um $C_\ell(r)$ zu bestimmen, verwenden wir, dass

$$\left\langle \vec{x} \,\middle|\, \vec{x}' \right\rangle = \delta^{(3)}(\vec{x} - \vec{x}') = \frac{1}{r^2 \sin\vartheta} \delta(r - r')\delta(\vartheta - \vartheta')\delta(\varphi - \varphi') \tag{3.11.73}$$

ist, denn für eine beliebige Funktion $\psi(\vec{x}) \equiv \tilde{\psi}(r, \vartheta, \varphi)$ gilt dann

$$\begin{aligned} \int_{\mathbb{R}^3} \mathrm{d}^3 x' \psi(\vec{x}')\delta(\vec{x} - \vec{x}') &= \psi(\vec{x}) \\ &= \int_0^\infty \mathrm{d}r' \int_0^\pi \mathrm{d}\vartheta' \int_0^{2\pi} \mathrm{d}\varphi r'^2 \sin\vartheta' \frac{1}{r^2 \sin\vartheta} \delta(r - r')\delta(\vartheta - \vartheta')\delta(\varphi - \varphi')\tilde{\psi}(r', \vartheta', \varphi') \\ &= \tilde{\psi}(r, \vartheta, \varphi). \end{aligned} \tag{3.11.74}$$

Weiter verwenden wir die Vollständigkeit der Kugelflächenfunktionen für quadratintegrable Funktionen $f(\vartheta, \varphi)$, also

$$\sum_{\ell=0}^{\infty} \sum_{m=-\ell}^{\ell} \mathrm{Y}^*_{\ell m}(\vartheta, \varphi) \mathrm{Y}_{\ell m}(\vartheta', \varphi') = \frac{1}{\sin\vartheta} \delta(\vartheta - \vartheta')\delta(\varphi - \varphi'). \tag{3.11.75}$$

Demnach ist mit (3.11.72)

$$\begin{aligned} \left\langle \vec{x}' \,\middle|\, \vec{x} \right\rangle = \delta^{(3)}(\vec{x} - \vec{x}') &= \int_0^\infty \mathrm{d}r \sum_{\ell=0}^{\infty} \sum_{m=-\ell}^{\ell} \left\langle \vec{x} \,\middle|\, u_{r\ell m} \right\rangle \left\langle u_{r\ell m} \,\middle|\, \vec{x} \right\rangle \\ &= \int_0^\infty \mathrm{d}r \sum_{\ell=0}^{\infty} \sum_{m=-\ell}^{\ell} C_\ell(r)\delta(r - |\vec{x}|)\mathrm{Y}_\ell(\vartheta', \varphi') C^*_\ell(r)\delta(r - |\vec{x}'|)\mathrm{Y}^*_{\ell m}(\vartheta, \varphi) |u_{r\ell m}\rangle \\ &= \delta(|\vec{x}| - |\vec{x}'|) \sum_{\ell=0}^{\infty} \sum_{m=-\ell}^{\ell} |C_\ell(r)|^2 \mathrm{Y}^*_{\ell m}(\vartheta, \varphi)\mathrm{Y}_{\ell m}(\vartheta', \varphi') \\ &\overset{!}{=} \frac{1}{r^2 \sin\vartheta} \delta(r - r')\delta(\vartheta - \vartheta')\delta(\varphi - \varphi'). \end{aligned} \tag{3.11.76}$$

Mit (3.11.76) folgt daraus, dass

$$|C_\ell(r)|^2 = \frac{1}{r^2} \tag{3.11.77}$$

gelten muss. Da ein Phasenfaktor keine Rolle spielt, können wir die Phase von $|u_{r\ell m}\rangle$ so definieren, dass

$$u_{r\ell m}(\vec{x}) = \left\langle \vec{x} \,\middle|\, u_{r\ell m} \right\rangle = \frac{1}{r} \delta(r - |\vec{x}|)\mathrm{Y}_{\ell m}(\vartheta, \varphi) \tag{3.11.78}$$

ist.

3.12 Freies Teilchen (Energie-Drehimpuls-Eigenzustände)

Wir haben das Energieeigenwertproblem für das freie Teilchen bereits in Abschnitt 3.8 gelöst, indem wir die Impulskomponenten $\vec{\mathbf{p}}$ als vollständigen Satz von mit der Energie kompatiblen Observablen zur Definition verwendet haben. Dies waren die ebenen Wellen

$$u_{\vec{p}}(\vec{x}) = N_{\vec{p}} \exp(\mathrm{i}\vec{p} \cdot \vec{x}/\hbar) \tag{3.12.1}$$

und die Energieeigenwerte sind $E(\vec{p}) = \vec{p}^2/(2\mu)$. Da die simultanen Eigenwerte der Impulsoperatoren $\vec{\mathbf{p}}$ durch $\vec{p} \in \mathbb{R}^3$ gegeben sind, sind die Energieeigenwerte durch $E(\vec{p}) \in \mathbb{R}_{\geq 0}$ gegeben. Dabei ist $N_{\vec{p}} = 1/(2\pi\hbar)^{3/2}$ ein Normierungsfaktor, der so gewählt wurde, dass $\left\langle u_{\vec{p}} \,\middle|\, u_{\vec{p}'} \right\rangle = \delta^{(3)}(\vec{p} - \vec{p}')$ ist.

Manchmal ist es aber auch nützlich, als Energieeigenbasis die simultanen Eigenvektoren von $\mathbf{H}$, $\vec{\mathbf{L}}^2$ und $\mathbf{L}_3$ zu verwenden (z.B. wenn man die Streuung eines Teilchens durch ein Radialpotential betrachtet). Es ist klar, dass diese Observablen in der Tat ein vollständiges System kompatibler Observabler darstellen, denn

$$\mathbf{H} = \vec{\mathbf{p}}^2/(2\mu) \tag{3.12.2}$$

vertauscht mit allen drei Bahndrehimpulskomponenten $\vec{\mathbf{L}}$ und folglich mit $\vec{\mathbf{L}}^2$ und $\mathbf{L}_3$. Diese beiden Operatoren kommutieren auch untereinander, und wir haben das gemeinsame Eigenwertproblem im vorigen Abschnitt ausführlich erläutert. Suchen wir also die Eigenfunktionen $u_{E\ell m}(\vec{x})$, die wir natürlich wieder am bequemsten in Abhängigkeit von den Kugelkoordinaten (r, ϑ, φ) von $\vec{x}$ schreiben. Die Lösung des simultanen Eigenwertproblems für $\vec{\mathbf{L}}^2$ und $\mathbf{L}_3$ hat auf die Kugelflächenfunktionen $\mathrm{Y}_{\ell m}(\vartheta, \varphi)$ geführt, und es ist klar, dass

$$u_{E\ell m}(\vec{x}) = R_{E\ell}(r)\mathrm{Y}_{\ell m}(\vartheta, \varphi) \tag{3.12.3}$$

sein muss. In der Tat enthalten alle Drehimpulskomponentenoperatoren $\vec{\mathbf{L}}$ nur die Winkel und Ableitungen nach den Winkeln, so dass

$$\begin{aligned} \vec{\mathbf{L}}^2 u_{E\ell m} &= R_{E\ell}(r)\vec{\mathbf{L}}^2 \mathrm{Y}_{\ell m}(\vartheta, \varphi) \\ &= \hbar^2 \ell(\ell+1) R_{E\ell}(r)\mathrm{Y}_{\ell m}(\vartheta, \varphi) \\ &= \hbar^2 \ell(\ell+1) u_{E\ell m}, \\ \vec{\mathbf{L}}_z u_{E\ell m} &= m\hbar u_{E\ell m} \end{aligned} \tag{3.12.4}$$

ist. Weiter ist gemäß (3.11.27)

$$\mathbf{H}\psi = \frac{1}{2m}\vec{\mathbf{p}}^2\psi = -\frac{\hbar^2}{2m}\Delta\psi = -\frac{\hbar^2}{2\mu}\frac{1}{r}\partial_r^2(r\psi) + \frac{\vec{\mathbf{L}}^2}{2\mu r^2}\psi. \tag{3.12.5}$$

Die Energie-Eigenwertgleichung lautet

$$\mathbf{H}u_{E\ell m} = E u_{E\ell m}, \tag{3.12.6}$$

und wenden wir (3.12.5) auf den Ansatz (3.12.3) an, folgt nach einfachen Umformungen die Gleichung für die Radialwellenfunktion

$$-\frac{\hbar^2}{2\mu r}(rR_{E\ell})'' + \frac{\ell(\ell+1)\hbar^2}{2\mu r^2}R_{E\ell} = ER_{E\ell}. \tag{3.12.7}$$

Dabei bedeutet der Strich Ableitungen nach r. Da $E \geq 0$ ist, können wir zur Vereinfachung den Parameter

$$k = \sqrt{2\mu E/\hbar^2} \tag{3.12.8}$$

einführen und künftig mit k statt E zu rechnen, wodurch sich mit $R_{k\ell}(r) \equiv R_{E\ell}(r)$ (3.12.7) zu

$$\frac{1}{r}(rR_{k\ell})'' - \frac{\ell(\ell+1)}{r^2}R_{k\ell} = -k^2 R_{k\ell} \tag{3.12.9}$$

vereinfacht. Führen wir schließlich noch als dimensionslose Radialkoordinate $\rho = kr$ ein und schreiben $R_{E\ell}(r) = u_\ell(\rho)$, wird diese Gleichung schließlich zu

$$\frac{1}{\rho}(\rho u_\ell)'' - \frac{\ell(\ell+1)}{\rho^2}u_\ell + u_\ell = 0. \tag{3.12.10}$$

Dabei bedeutet jetzt der Strich natürlich die Ableitung nach ρ. Dies ist die Differentialgleichung (B.2.1) für die sphärischen Bessel-, Neumann- und Hankel-Funktionen.
Wir müssen nun Lösungen dieser Gleichung finden, für die der Operator in (3.12.9) auf der linken Seite bzgl. des Skalarprodukts für Radialwellenfunktionen

$$\langle R_1 | R_2 \rangle_2 = \int_0^\infty \mathrm{d}r\, r^2 R_1^*(r) R_2(r) \tag{3.12.11}$$

selbstadjungiert ist. Das ist nur für den ersten Term in (3.12.9) nicht trivial. Betrachten wir also diesen Operator

$$\left\langle R_1 \middle| (rR_2)''/r \right\rangle_2 = \int_0^\infty \mathrm{d}r\, r R_1^*(r) [r R_2(r)]''. \tag{3.12.12}$$

Um den Operator auf den Faktor mit R_1^* „überzuwälzen“, müssen wir zweimal partiell integrieren, wobei die Randterme jeweils wegfallen müssen. Dabei gehen wir für $r \to \infty$ davon aus, dass die Funktionen so schnell im Unendlichen abfallen, dass die Integrale existieren. Im ersten Schritt erhalten wir

$$\left\langle R_1 \middle| (rR_2)''/r \right\rangle_2 = -\int_0^\infty \mathrm{d}r [r R_1^*(r)]' [r R_2(r)]', \quad r R_1^*(r) [r R_2(r)]' = 0 \quad \text{für} \quad r = 0 \tag{3.12.13}$$

und im zweiten Schritt

$$\left\langle R_1 \middle| (rR_2)''/r \right\rangle_2 = \int_0^\infty \mathrm{d}r [r R_1(r)]''^* [r R_2(r)], \quad [r R_1^*(r)]' r R_2(r) = 0 \quad \text{für} \quad r = 0. \tag{3.12.14}$$

Wir müssen also für die Radialwellenfunktionen offenbar verlangen, dass

$$\lim_{r \to 0^+} r R(r) = 0 \tag{3.12.15}$$

ist. Bzgl. $r \to \infty$ müssen die Funktionen $rR(r)$ entweder hinreichend schnell verschwinden, so dass $\langle R|R \rangle$ existiert oder für verallgemeinerte Eigenfunktionen von selbstadjungierten Operatoren mit kontinuierlichen Eigenwerten wie in unserem Fall „auf eine δ-Distribution“ normierbar ist. Wir legen die Normierung zu $\langle R_{k\ell} | R_{k'\ell} \rangle = N_{k\ell} \delta(k - k')/k^2$ fest.
Wir müssen also (3.12.10) mit der Randbedingung

$$\lim_{\rho \to 0^+} \rho u_\ell(\rho) = 0 \tag{3.12.16}$$

lösen. Aus Anhang B.2 wissen wir, dass damit nur $u_\ell(\rho) = C_\ell \mathrm{j}_\ell(\rho)$ mit einer Normierungskonstanten C_ℓ in Frage kommt.
Damit ist die Lösung des Energieeigenwertproblems für das freie Teilchen in Form von simultanen Drehimpulseigenlösungen gefunden:

$$R_{k\ell}(r) = N_{k\ell} \mathrm{j}_\ell(kr), \quad E = \frac{\hbar^2 k^2}{2\mu}. \tag{3.12.17}$$

Das betreffende Normierungsintegral erhalten wir mit Hilfe der Formel (B.3.68)

$$\begin{aligned} \langle R_{k\ell} | R_{k'\ell} \rangle &= \int_0^\infty \mathrm{d}r\, r^2 R_{k\ell}^*(r) R_{k'\ell}(r) \\ &= |N_{k\ell}|^2 \int_0^\infty \mathrm{d}r\, r^2 \mathrm{j}_\ell(kr) \mathrm{j}_\ell(k'r) = |N_{k\ell}|^2 \frac{\pi}{2k^2} \delta(k - k'). \end{aligned} \tag{3.12.18}$$

Setzen wir also $N_{k\ell} = \sqrt{2/\pi}$, ist die Gesamtwellenfunktion

$$u_{klm}(\vec{x}) = u_{klm}(r,\vartheta,\varphi) = \sqrt{\frac{2}{\pi}} \mathrm{j}_\ell(kr) \mathrm{Y}_{\ell m}(\vartheta,\varphi) \tag{3.12.19}$$

so normiert, dass

$$\int_{\mathbb{R}^3} \mathrm{d}^3x\, u^*_{k\ell m}(\vec{x}) u_{k'\ell' m'}(\vec{x}) = \frac{1}{k^2} \delta(k-k') \delta_{\ell\ell'} \delta_{mm'} \tag{3.12.20}$$

gilt. Der zur Energieeigenfunktion (3.12.19) gehörige Energieeigenwert ist dabei $E = \hbar^2 k^2/(2\mu)$.
Wir merken der Vollständigkeit halber an, dass wir das Normierungsintegral (3.12.18) auch anders berechnen können. Dazu betrachten wir die Wronski-Determinante der Funktionen $R_{k\ell}(r)$ für zwei $k_1, k_2 > 0$:

$$W_{k_1k_2}(r) = \det \begin{pmatrix} R_{k_1\ell}(r) & R_{k_2\ell}(r) \\ R'_{k_1\ell}(r) & R'_{k_2\ell}(r) \end{pmatrix}. \tag{3.12.21}$$

Leiten wir diese Funktion nach r ab folgt

$$W'_{k_1k_2}(r) = \begin{pmatrix} R_{k_1\ell}(r) & R_{k_2\ell}(r) \\ R''_{k_1\ell}(r) & R''_{k_2\ell}(r) \end{pmatrix}. \tag{3.12.22}$$

In der letzten Reihe können wir die $R''_{k\ell}$ mit Hilfe der DGL (3.12.9) eliminieren:

$$R''_{k\ell}(r) = -\frac{2}{r} R'_{k\ell}(r) + \left[\frac{\ell(\ell+1)}{r^2} - k^2 \right] R_{k\ell}(r). \tag{3.12.23}$$

Daraus folgt (*Nachrechnen*)

$$W'_{k_1k_2}(r) = -\frac{2}{r} W_{k_1k_2}(r) + (k_1^2 - k_2^2) R_{k_1\ell}(r) R_{k_2\ell}(r). \tag{3.12.24}$$

Bringt man den ersten Term auf die linke Seite, folgt

$$R_{k_1\ell}(r) R_{k_2\ell}(r) = \frac{1}{r^2(k_1^2 - k_2^2)} \frac{\mathrm{d}}{\mathrm{d}r} \left[r^2 W_{k_1k_2}(r) \right]. \tag{3.12.25}$$

Daraus folgt

$$\int_0^R \mathrm{d}r\, r^2 R_{k_1\ell}(r) R_{k_2\ell}(r) = \frac{R^2}{k_1^2 - k_2^2} W_{k_1k_2}(R) = \frac{R^2}{k_1^2 - k_2^2} [R_{k_1\ell}(R) R'_{k_2\ell}(R) - R_{k_2\ell}(R) R'_{k_1\ell}(R)]. \tag{3.12.26}$$

Dabei haben wir verwendet, dass $rR_{k\ell}(r) \to 0$ für $r \to 0$ ist. Da wir das Normierungsintegral für die $R_{k\ell}$ bestimmen wollen, benötigen wir den (schwachen) Limes des Integrals für $R \to \infty$. Wir können also annehmen, dass $k_1 R \gg 1$ und $k_2 R \gg 1$ sind und folglich in (3.12.17) die asymptotische Formel (B.2.27) für die sphärischen Besselfunktionen verwenden können. Dann folgt

$$\begin{aligned} \int_0^R \mathrm{d}r\, r^2 R_{k_1\ell}(r) R_{k_2\ell}(r) \underset{R\to\infty}{\cong} \frac{N_{k_1\ell} N_{k_2\ell}}{k_1^2 - k_2^2} \Bigg[& \frac{\sin(k_1 R - \ell\pi/2)\cos(k_2 R - \ell\pi/2)}{k_1} \\ & - \frac{\sin(k_2 R - \ell\pi/2)\cos(k_1 R - \ell\pi/2)}{k_2} \Bigg]. \end{aligned} \tag{3.12.27}$$

Nun wenden wir im Nenner die trigonometrische Identität

$$\sin\alpha\cos\beta = \frac{1}{2}[\sin(\alpha+\beta)+\sin(\alpha-\beta)] \tag{3.12.28}$$

an. Das ergibt

$$\int_0^R \mathrm{d}r\, r^2 R_{k_1\ell}(r) R_{k_2\ell}(r) \underset{R\to\infty}{\cong} \frac{N_{k_1\ell}N_{k_2\ell}}{2k_1k_2}\frac{\sin[(k_1-k_2)R]}{k_1-k_2} \xrightarrow[R\to\infty]{} \frac{N_{k_1\ell}^2\pi}{2k_1^2}\delta(k_1-k_2). \tag{3.12.29}$$

Den Beweis des Limes für $R\to\infty$ verschieben wir auf das Ende dieses Abschnitts. Setzt man

$$N_{k\ell} = \sqrt{\frac{2}{\pi}} \tag{3.12.30}$$

sind also die entsprechenden Radialwellenfunktionen

$$R_{k\ell}(r) = \sqrt{\frac{2}{\pi}} j_\ell(kr) \tag{3.12.31}$$

bzgl. des Skalarproduktes mit dem Gewicht r^2 im Orts- und k^2 im k-Raum auf δ-Distributionen normiert, d.h. es gilt

$$\int_0^\infty \mathrm{d}r\, r^2 R_{k_1\ell}(r) R_{k_2\ell}(r) = \frac{\delta(k_1-k_2)}{k_1^2} \tag{3.12.32}$$

und

$$\int_0^\infty \mathrm{d}r k^2 R_{k\ell}(r_1) R_{k\ell}(r_2) = \frac{\delta(r_1-r_2)}{r_1^2}. \tag{3.12.33}$$

Entsprechend hat man im Hilbert-Raum $\mathrm{L}^2(\mathbb{R}_{\geq 0}, r^2)$

$$\langle k\ell \,|\, r\ell\rangle = \langle r\ell \,|\, k\ell\rangle_2 = \sqrt{\frac{2}{p}} j_\ell(kr), \tag{3.12.34}$$

d.h. die Umrechnung zwischen Orts- und k-Darstellungen von Radialwellenfunktionen bzgl. der verallgemeinerten Orthonormalbasis (3.12.31) ergibt sich zu

$$\begin{aligned} \psi(r) &= \langle r\ell \,|\, \psi\rangle_2 = \sqrt{\frac{2}{\pi}}\int_0^\infty \mathrm{d}k k^2 \tilde{\psi}(k) j_\ell(kr) \\ \tilde{\psi}(k) &= \langle k\ell \,|\, \psi\rangle_2 = \sqrt{\frac{2}{\pi}}\int_0^\infty \mathrm{d}r\, r^2 \psi(r) j_\ell(kr). \end{aligned} \tag{3.12.35}$$

Jetzt holen wir noch den Beweis der Limesbildung in (3.12.29) nach, d.h. wir wollen zeigen, dass

$$\lim_{R\to\infty} \frac{\sin[(k_1-k_2)R)}{k_1-k_2} = \pi\delta(k_1-k_2) \tag{3.12.36}$$

für auf $\mathbb{R}_{>0}$ definierte Funktionen ist. Wir müssen also zeigen, dass für beliebige „Testfunktionen“ $f:\mathbb{R}_{\geq 0}\to\mathbb{C}$

$$\lim_{R\to\infty}\int_0^\infty \mathrm{d}k f(k_2)\frac{\sin[(k_1-k_2)R)}{k_1-k_2} = \pi f(k_1) \tag{3.12.37}$$

ist. Substituieren wir im Integral $k' = (k_2 - k_1)R$, folgt zunächst

$$\lim_{R\to\infty} \int_{-k_1 R}^{\infty} \mathrm{d}k' \frac{\sin k'}{k'} f(k_1 + k'/R) = f(k_1) \int_{-\infty}^{\infty} \mathrm{d}k' \frac{\sin k'}{k'}. \tag{3.12.38}$$

Wir müssen also noch das verbliebene Integral berechnen. Dazu betrachten wir die Funktion

$$F(s) = 2\int_0^{\infty} \mathrm{d}k' \exp(-sk') \frac{\sin k'}{k'}, \quad s > 0. \tag{3.12.39}$$

Dann ist das gesuchte Integral $F(0^+)$. Um (3.12.39) zu berechnen, leiten wir die Funktion nach s ab:

$$\begin{aligned} F'(s) &= -2\int_0^{\infty} \exp(-sk')\sin k' = -\frac{1}{2\mathrm{i}}\int_0^{\infty} \mathrm{d}k' \exp(-sk')[\exp(\mathrm{i}k') - \exp(-\mathrm{i}k')] \\ &= -\frac{1}{\mathrm{i}}\int_0^{\infty} \mathrm{d}k' \left\{\exp[k'(\mathrm{i}-s)] - \exp[k'(-\mathrm{i}-s)]\right\} \\ &= \frac{1}{\mathrm{i}}\left[\frac{1}{\mathrm{i}-s} + \frac{1}{\mathrm{i}+s}\right] = -\frac{2}{s^2+1}. \end{aligned} \tag{3.12.40}$$

Es ist also

$$F(s) = -\int \mathrm{d}s \frac{2}{s^2+1} = -2\arctan s + C \tag{3.12.41}$$

Da offenbar $F(s) \to 0$ für $s \to \infty$ ist, folgt

$$F(s) = \pi - 2\arctan s \;\Rightarrow\; F(0^+) = \pi, \tag{3.12.42}$$

d.h. es gilt

$$\int_{-\infty}^{\infty} \mathrm{d}k' \frac{\sin k'}{k'} = 2\int_0^{\infty} \mathrm{d}k' \frac{\sin k'}{k'} = 2F(0^+) = \pi. \tag{3.12.43}$$

3.13 Das Wasserstoffatom

In diesem Abschnitt beschäftigen wir uns mit einem der wenigen realistischen Energieeigenwertprobleme der Wellenmechanik, die vollständig analytisch gelöst werden können. Dabei handelt es sich um die **gebundenen Zustände** für die Bewegung eines Elektrons um ein Proton. Da die Masse des Protons $m_\mathrm{p} \simeq 1836 m_\mathrm{e}$ ist, können wir in erster Näherung annehmen, dass das Proton als im Ursprung des Koordinatensystems ruhend angenommen werden darf. Außerdem reduziert sich dann das Problem auf die Beschreibung der Bewegung des Elektrons im entsprechenden Coulomb-Feld des Protons. Der Hamilton-Operator für das Elektron lautet demnach

$$\mathbf{H} = \frac{\vec{\mathbf{p}}^2}{2m_\mathrm{e}} - \frac{e^2}{4\pi\epsilon_0 r}. \tag{3.13.1}$$

Dabei ist e die Ladung des Protons und $-e$ die Ladung des Elektrons. Diese **Elementarladung** ist eine fundamentale Naturkonstante. In den hier verwendeten SI-Einheiten ist sie *exakt* festgelegt zu $e = 1{,}602176634 \cdot 10^{-19}$ C, was implizit die Einheit Coulomb (C) für die elektrische Ladung definiert.

Wir interessieren uns nun für die **Eigenwerte** und die **Eigenfunktionen** des Hamilton-Operators. Da es sich hier um ein **Zentralpotentialproblem** handelt, kommutiert der Hamilton-Operator mit dem Bahndrehimpulsoperator. Daher können wir als vollständiges System von kompatiblen Observablen $\mathbf{H}$, $\vec{\mathbf{L}}^2$ und $\mathbf{L}_z$ wählen. Wie wir aus Abschnitt (3.11) wissen, sind die simultanen orthonormierten Eigenfunktionen von $\vec{\mathbf{L}}^2$ und

$\mathbf{L}_z$ die Kugelflächenfunktionen $\mathrm{Y}_{\ell m}$ zu den Eigenwerten $\ell(\ell+1)\hbar^2$ für $\vec{\mathbf{L}}^2$, wobei $\ell \in \mathbb{N}_0 = \{0,1,2,\ldots\}$ und den Eigenwerten $m\hbar$, wobei für jedes ℓ gilt: $m \in \{-\ell,-\ell+1,\ldots,\ell-1,\ell\}$.

Es ist daher von vornherein klar, dass die simultanen Eigenfunktionen für $\mathbf{H}$, $\vec{\mathbf{L}}^2$ und $\mathbf{L}_z$ die Form

$$u_{E,\ell,m}(\vec{x}) = R_{E\ell}(r)\mathrm{Y}_{\ell m}(\vartheta,\varphi) \tag{3.13.2}$$

besitzen wobei (r,ϑ,φ) die üblichen Kugelkoordinaten des Ortsvektors $\vec{r}$ bezeichnen. Dass dieser Ansatz zum Ziel führt, ergibt sich wieder daraus, dass wir für den Laplace-Operator in Kugelkoordinaten

$$-\Delta\psi = -\frac{1}{r}\partial_r^2(r\psi) + \frac{1}{r^2}\vec{\mathbf{l}}^2\psi \quad \text{mit} \quad \vec{\mathbf{l}} = \frac{1}{\hbar}\vec{\mathbf{L}} \tag{3.13.3}$$

schreiben können (s. Abschnitt 3.11).

Mit (3.13.1) lautet also die Eigenwertgleichung für $\mathbf{H}$

$$\mathbf{H}u_{E,\ell,m} = -\frac{\hbar^2}{2m_\mathrm{e}}\Delta u_{E,\ell,m} - \frac{e^2}{4\pi\epsilon_0 r}u_{E,\ell,m} = E u_{E,\ell,m}. \tag{3.13.4}$$

Setzen wir hier (3.13.3) ein und verwenden den Ansatz (3.13.4), erhalten wir wegen $\vec{\mathbf{l}}^2\mathrm{Y}_{\ell m} = \ell(\ell+1)\mathrm{Y}_{\ell m}$ und weil $\vec{\mathbf{L}}^2$ nur auf die Winkel ϑ und φ, nicht aber auf die Radialkoordinate r wirkt nach Kürzen durch $\mathrm{Y}_{\ell m}$ (*nachrechnen!*)

$$-\frac{\hbar^2}{2m_\mathrm{e}r}\mathrm{d}_r^2[rR_{E,\ell}(r)] + \left(\frac{\ell(\ell+1)\hbar^2}{2m_\mathrm{e}r^2} - \frac{e^2}{4\pi\epsilon_0 r}\right)R_{E,\ell}(r) = ER_{E,\ell}(r). \tag{3.13.5}$$

Zum Einen ersehen wir daraus, dass in der Tat $R_{E,\ell}$ nicht von der Drehimpulsquantenzahl m abhängt, sondern nur vom Energieeigenwert E und der Drehimpulsbetragsquantenzahl ℓ und dass der Ansatz (3.13.2) wirklich mit dem simultanen Eigenwertproblem kompatibel ist.

Es bleibt nun noch, die **Radialgleichung** (3.13.5) zu lösen. Dabei müssen wir für jedes $\ell \in \mathbb{N}_0$ die Energieeigenwerte E und die dazugehörigen Eigenfunktionen $R_{E,\ell}$ finden.

Dazu müssen wir beachten, dass $R_{E,\ell}$, wie beim analogen Problem für das freie Teilchen die Randbedingung (3.12.15), also

$$\lim_{r\to 0} rR_{E,\ell}(r) = 0 \tag{3.13.6}$$

erfüllen muss, denn auch hier ist der Operator (3.13.5) offenbar selbstadjungiert, wenn der Operator im ersten Term selbstadjungiert ist, und diese Bedingung führt auf die besagte Randbedingung (3.13.6).

Weiter sind wir hier nur an den sog. **gebundenen Zuständen** interessiert. Beim analogen Bewegungsproblem in der klassischen Mechanik, also der Kepler-Bewegung der Planeten um die Sonne, gab es ja sowohl gebundene Lösungen (Ellipsenbahnen) als auch ungebundene Lösungen (Parabel- und Hyperbelbahnen). Die letzteren Lösungen entsprechen für das Elektron der **Streuung an einem Proton aufgrund der Coulomb-Wechselwirkung**. Wir wollen hier aber nur die Lösungen betrachten, die gebundenen Bahnen entsprechen, also den Ellipsen des klassischen Problems. Natürlich gibt es in der Quantentheorie keine Bahnen sondern lediglich die hier interessanten Energieeigenfunktionen, die **stationären Zuständen** entsprechen. Diesen Wellenfunktionen entsprechen demnach Zuständen, in denen das Elektron ruht, und das Betragsquadrat gibt die Aufenthaltswahrscheinlichkeitsverteilung für dieses ruhende an das Proton gebundene Elektron an. Damit ein gebundener Zustand vorliegt, muss die Wellenfunktion wirklich quadratintegrabel sein, d.h. es müssen die gebundenen Energieeigenfunktionen auf 1 normierbar sein:

$$\langle u_{E,\ell,m}|u_{E,\ell,m}\rangle = \int_{\mathbb{R}^3}\mathrm{d}^3x\, u^*_{E,\ell,m}(x)u_{E,\ell,m}(x) = \int_0^\infty \mathrm{d}r\, r^2|R_{E,\ell}(r)|^2 = 1 \tag{3.13.7}$$

sein. Dabei haben wir im letzten Schritt den Ansatz (3.13.2) eingesetzt und das Integral in Kugelkoordinaten umgeschrieben. Die Winkelintegration liefert 1 wegen der entsprechenden Normierung der Kugelflächenfunktionen (*nachrechnen!*). Jedenfalls folgt aus (3.13.7), dass

$$\lim_{r\to\infty} rR_{E,\ell}(r) = 0 \tag{3.13.8}$$

sein muss. Wir werden nun zeigen, dass unter den Randbedingungen (3.13.6) und (3.13.8) die Radialgleichung für jedes $\ell \in \mathbb{N}_0$ (3.13.11) nur für bestimmte **diskrete Energieeigenwerte** E_n (mit einem Index $n \in \mathbb{N}$) Lösungen besitzt. Wie wir sehen werden, stimmen diese diskreten Energieniveaus mit dem Resultat der alten Bohr-Sommerfeld-Quantentheorie (wie in Abschnitt 2.5 kurz behandelt) überein. In der Tat lassen sich diese Vorhersagen des Energiespektrums experimentell überprüfen, denn sie führen zu den **charakteristischen Spektrallinien** des Wasserstoffatoms, wenn die Elektronen durch irgendwelche Anregungen von einem Energieniveau zu einem anderen „springen" und dabei die entsprechenden elektromagnetischen Wellen mit den Frequenzen $\omega_{nn'}$ abstrahlen bzw. absorbieren, wobei für die entsprechenden Frequenzen für die emittierten bzw. absorbierten Photonen $\hbar\omega_{nn'} = E_n - E_{n'}$ gelten muss (Energieerhaltung!).

Wenden wir uns also nun ausführlich der Lösung des Eigenwertproblems (3.13.5) zu. Dazu führen wir als erstes die **Sommerfeldsche Feinstrukturkonstante**

$$\alpha = \frac{e^2}{4\pi\epsilon_0\hbar c} \simeq \frac{1}{137{,}036} \tag{3.13.9}$$

ein und beachten, dass wir an Energie-Eigenwerten $E < 0$, also gebundenen Zuständen, interessiert sind. Schreiben wir dann noch

$$k = \sqrt{-\frac{2m_\mathrm{e}E}{\hbar^2}}, \quad \xi = 2kr \quad und \quad R_{E\ell} = \frac{1}{r}\chi_{n\ell}(\xi) \tag{3.13.10}$$

und setzen all diese Substitutionen in (3.13.5) ein, schreibt sich diese Gleichung in der Form (*nachrechnen!*)

$$\left(\frac{\mathrm{d}^2}{\mathrm{d}\xi^2} - \frac{1}{4} + \frac{n}{\xi} - \frac{\ell(\ell+1)}{\xi^2}\right)\chi_{n\ell}(\xi) = 0 \quad \text{mit} \quad n = \alpha\sqrt{-\frac{m_\mathrm{e}c^2}{2E}}. \tag{3.13.11}$$

Wir betrachten zunächst das Verhalten dieser Funktion für $\xi \to 0$. Dann kann man in der Klammer die Terme $-1/4 + n/\xi$ gegenüber $\ell(\ell+1)/\xi^2$ vernachlässigen, und man erhält

$$\chi''_{n\ell}(\xi) - \frac{\ell(\ell+1)}{\xi^2}\chi_{n\ell}(\xi) = 0. \tag{3.13.12}$$

Diese Gleichung lässt sich offenbar mit dem Ansatz $\chi_{n\ell}(\xi) = A\xi^{\lambda+1}$ lösen. Setzt man diesen Ansatz ein, folgt

$$\lambda(\lambda+1) = \ell(\ell+1) \Rightarrow \lambda_1 = \ell, \quad \lambda_2 = -(\ell+1). \tag{3.13.13}$$

Wegen der Randbedingung (3.13.6) kommt nur die Lösung mit $\lambda = \lambda_1 = \ell$ in Frage.

Weiter bestimmen wir das Verhalten von $\chi_{n\ell}$ für $\xi \to \infty$. Dann kann man in der Klammer in (3.13.11) die Terme $n/\xi - \ell(\ell+1)/\xi^2$ gegenüber $1/4$ vernachlässigen, und man erhält

$$\chi''_{n\ell}(\xi) = \frac{1}{4}\chi_{n\ell}(\xi) \tag{3.13.14}$$

Die Lösungen sind offenbar $\chi_{n\ell}(\xi) = A\exp(\pm\xi/2)$. Damit $\chi_{n\ell}$ quadratintegrabel ist, kommt nur das untere Vorzeichen in Frage.

Wir arbeiten nun die asymptotischen Lösungen für $\xi \to 0$ bzw. $\xi \to \infty$ in den Ansatz für $\chi_{n\ell}$ ein, indem wir eine neue Funktion $w_{n\ell}$ durch

$$\chi_{n\ell}(\xi) = \xi^{l+1} \exp\left(-\frac{\xi}{2}\right) w_{n\ell}(\xi) \tag{3.13.15}$$

definieren. Setzen wir dies in (3.13.12) ein, ergibt sich für $w_{n\ell}$ die DGL

$$\xi w''_{n\ell} + (2l+2-\xi) w'_{n\ell} + (n-\ell-1) w_{n\ell} = 0. \tag{3.13.16}$$

Diese Gleichung lösen wir nun mit einem Potenzreihenansatz. Damit das asymptotische Verhalten für $\xi \to 0$ erhalten bleibt, muss offenbar

$$w_{n\ell}(\xi) = \sum_{j=0}^{\infty} A_j \xi^j \quad \text{mit} \quad A_0 \neq 0 \tag{3.13.17}$$

sein. Um die A_j zu bestimmen, setzen wir diesen Ansatz in (3.13.16) ein und bringen die dabei entstehenden Reihen für die verschiedenen Terme in eine einheitliche Form. Es gilt

$$\xi w''_{n\ell}(\xi) = \sum_{j=2}^{\infty} j(j-1) A_j \xi^{j-1} = \sum_{j=0}^{\infty} j(j-1) A_j \xi^{j-1}. \tag{3.13.18}$$

Substituieren wir hierin $j' = j-1$, folgt

$$\xi w''_{n\ell}(\xi) = \sum_{j'=1}^{\infty} (j'+1) j' A_{j'+1} \xi^{j'} = \sum_{j'=0}^{\infty} (j'+1) j' A_{j'+1} \xi^{j'}. \tag{3.13.19}$$

Ebenso folgt

$$(2\ell+2) w'_{n\ell}(\xi) = \sum_{j'=0}^{\infty} (2\ell+2)(j'+1) A_{j'+1} \xi^{j'} \tag{3.13.20}$$

und

$$-\xi w'_{n\ell}(\xi) = -\sum_{j=1}^{\infty} j A_j \xi^j = -\sum_{j=0}^{\infty} j A_j \xi^j \tag{3.13.21}$$

und

$$(n-\ell-1) w_{n\ell}(\xi) = \sum_{j=0}^{\infty} (n-\ell-1) A_j \xi^j. \tag{3.13.22}$$

Fassen wir alles zusammen, wobei wir alle Summationsindizes wieder j nennen, folgt aus (3.13.16)

$$\sum_{j=0}^{\infty} \Big[(2\ell+j+2)(j+1) A_{j+1} + (n-\ell-1-j) A_j \Big] \xi^j = 0. \tag{3.13.23}$$

Folglich muss für alle $j \in \mathbb{N}_0 = \{0,1,2,3\ldots\}$ gelten

$$(2\ell+j+2)(j+1) A_{j+1} + (n-\ell-1-j) A_j = 0 \Rightarrow A_{j+1} = -\frac{n-\ell-1-j}{(2\ell+j+2)(j+1)} A_j. \tag{3.13.24}$$

Damit kann man alle Koeffizienten rekursiv berechnen, wobei $A_0 \neq 0$ beliebig ist.
Jetzt müssen wir aber beachten, dass die Funktion $w_{n\ell}(\xi)$ für $\xi \to \infty$ nicht zu schnell anwächst. Für $\xi \to \infty$ sind nun die Koeffizienten A_j für $j \to \infty$ wichtig. In diesem Limes ist aber gemäß (3.13.24)

$$A_{j+1} \underset{j\to\infty}{\cong} \frac{1}{j+1} A_j \Rightarrow A_j \underset{j\to\infty}{\cong} \frac{C}{j!}. \tag{3.13.25}$$

Bricht also die Reihe nicht für $j > n_r$ mit $n_r \in \mathbb{N}_0$ ab, verhält sich $w_{n\ell} \cong_{\xi\to\infty} C \exp \xi$. Dann würde aber wegen (3.13.15) $\chi_{n\ell}(\xi)$ für $\xi \to \infty$ wie $\exp(\xi/2)$ anwachsen und wäre nicht mehr quadratintegrabel. Folglich muss die Reihe bei einem endlichen n_r abbrechen, d.h. wir müssen

$$n_r = n - \ell - 1 \quad \text{mit} \quad n_r \in \mathbb{N}_0 \tag{3.13.26}$$

setzen. Dies ergibt die möglichen Eigenwerte für E, denn gemäß (3.13.11) ist

$$E_n = -\frac{\alpha^2 m_e c^2}{2n^2}. \tag{3.13.27}$$

Dabei durchläuft die **Hauptquantenzahl** $n = n_r + \ell + 1$ offenbar die Werte $n \in \mathbb{N} = \{1, 2, \ldots\}$, denn es ist $n_r \in \mathbb{N}_0$ und $l \in \mathbb{N}_0$.
Die Funktion $w_{n\ell}$ ist somit ein Polynom n_r-ten Grades. Wenn man die entsprechende Rekursionsformel (3.13.24, die wir jetzt einfacher in der Form

$$A_{j+1} = -\frac{n_r - j}{(2\ell + j + 2)(j+1)} A_j \tag{3.13.28}$$

schreiben können, auswertet, erhält man nach einiger Rechnung (*Nachprüfen!*)

$$A_j = (-1)^j \frac{n_r!}{(n_r - j)!} \frac{(2\ell+1)!}{(2\ell + 1 + j)! j!} A_0. \tag{3.13.29}$$

Mit Hilfe von Binomialkoeffizienten $\binom{n}{k} = n!/(k!(n-k)!)$ können wir dies in die Form

$$A_j = \frac{(-1)^j}{j!} \binom{n_r + 2\ell + 1}{n_r - j} \frac{n_r!}{(n_r + 2\ell + 1)!} A_0 \tag{3.13.30}$$

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

$$A_j = \frac{(-1)^j}{j!} \binom{n_r + 2\ell + 1}{n_r - j} N_{n\ell} \tag{3.13.31}$$

mit $N_{n\ell} = \text{const.}$
Die Polynome

$$L_{n_r}^{2\ell+1}(\xi) = \sum_{j=0}^{n_r} \frac{(-1)^j}{j!} \binom{n_r + 2\ell + 1}{n_r - j} \xi^j \tag{3.13.32}$$

heißen **assoziierte Laguerre-Polynome**, wobei n_r und $2\ell + 1$ Indizes sind (d.h. der obere Index ist nicht als Potenz misszuverstehen!). Damit wird schließlich gemäß (3.13.2) und (3.13.10)

$$u_{n\ell m}(r, \vartheta, \varphi) = \frac{N_{n\ell}}{r} (2kr)^{\ell+1} \mathrm{L}_{n-\ell-1}^{2\ell+1}(2kr) \exp(-kr) \mathrm{Y}_{\ell m}(\vartheta, \varphi). \tag{3.13.33}$$

Dabei haben wir die Wellenfunktionen mit den üblichen Quantenzahlen $n = n_r + \ell + 1$, ℓ und m (Hauptquantenzahl, Drehimpulsquantenzahl und magnetische Quantenzahl genannt) durchnummeriert. Die Normierungskonstante bestimmt sich mit einigem Aufwand (vgl. Anh. B.4) zu

$$N_{n\ell} = \sqrt{\frac{k(n-\ell-1)!}{n(n+\ell)!}}. \tag{3.13.34}$$

Es ist im Folgenden bequem, die verschiedenen Parameter in (3.13.33) mit noch etwas anderen Größen auszudrücken. Zunächst ist es sinnvoll, den **Bohrschen Radius** a_B einzuführen, und zwar als

$$a_\mathrm{B} = \frac{\hbar}{\alpha m_\mathrm{e} c} = 5{,}29177210903(80) \cdot 10^{-11}\mathrm{m} \simeq 0{,}53\mathring{\mathrm{A}}. \tag{3.13.35}$$

Dabei ist 1 $\mathring{\mathrm{A}} = 10^{-10}$ m eine in der Atomphysik und Kistallographie gebräuchliche Einheit, benannt nach dem Physiker Jonas Anders Ångström (1813-1874). Dann wird gemäß (3.13.10) und (3.13.11)

$$k = \frac{1}{n a_\mathrm{B}}, \quad E_n = -\frac{\hbar^2}{2 m_\mathrm{e} a_\mathrm{B}^2 n^2} = -\frac{m_\mathrm{e} c^2 \alpha^2}{2n^2} \simeq -\frac{13{,}6\,\mathrm{eV}}{n^2}. \tag{3.13.36}$$

Setzt man all dies in (3.13.33) ein, erhält man schließlich endgültig

$$\psi_{n\ell m}(r,\vartheta,\varphi) = \frac{2}{n^2}\sqrt{\frac{(n-\ell-1)!}{(n+\ell)! a_\mathrm{B}^3}}\left(\frac{2r}{n a_\mathrm{B}}\right)^\ell \exp\left(-\frac{r}{n a_\mathrm{B}}\right) \mathrm{L}_{n-\ell-1}^{2\ell+1}\left(\frac{2r}{n a_\mathrm{B}}\right) \mathrm{Y}_{\ell m}(\vartheta,\varphi). \tag{3.13.37}$$

Diese Eigenfunktionen bilden 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 (3.13.37) die Orthonormierungsbedingungen

$$\langle \psi_{n\ell m} | \psi_{n'\ell' m'} \rangle = \int_0^\infty \mathrm{d}r \int_0^\pi \mathrm{d}\vartheta \int_0^{2\pi} \mathrm{d}\varphi\, r^2 \sin\vartheta\, \psi^*_{n\ell m}(r,\vartheta,\varphi) \psi_{n'\ell' m'}(r,\vartheta,\varphi) = \delta_{nn'}\delta_{\ell\ell'}\delta_{mm'}. \tag{3.13.38}$$

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 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 $n = n_r + \ell + 1$. Wegen $n_r \in \mathbb{N}_0$ kann zu gegebenem n folglich ℓ die Werte $0, 1, \ldots, n-1$ annehmen. Zu jedem ℓ gehören die $(2\ell+1)$ möglichen Werte für $m \in \{-\ell, -\ell+1, \ldots, \ell-1, \ell\}$. Damit ist der Entartungsgrad des zu n gehörigen Energieniveaus

$$\sum_{\ell=0}^{n-1} (2\ell+1) = n^2. \tag{3.13.39}$$

Die Summenformel kann man leicht durch vollständige Induktion beweisen (*Übung!*).

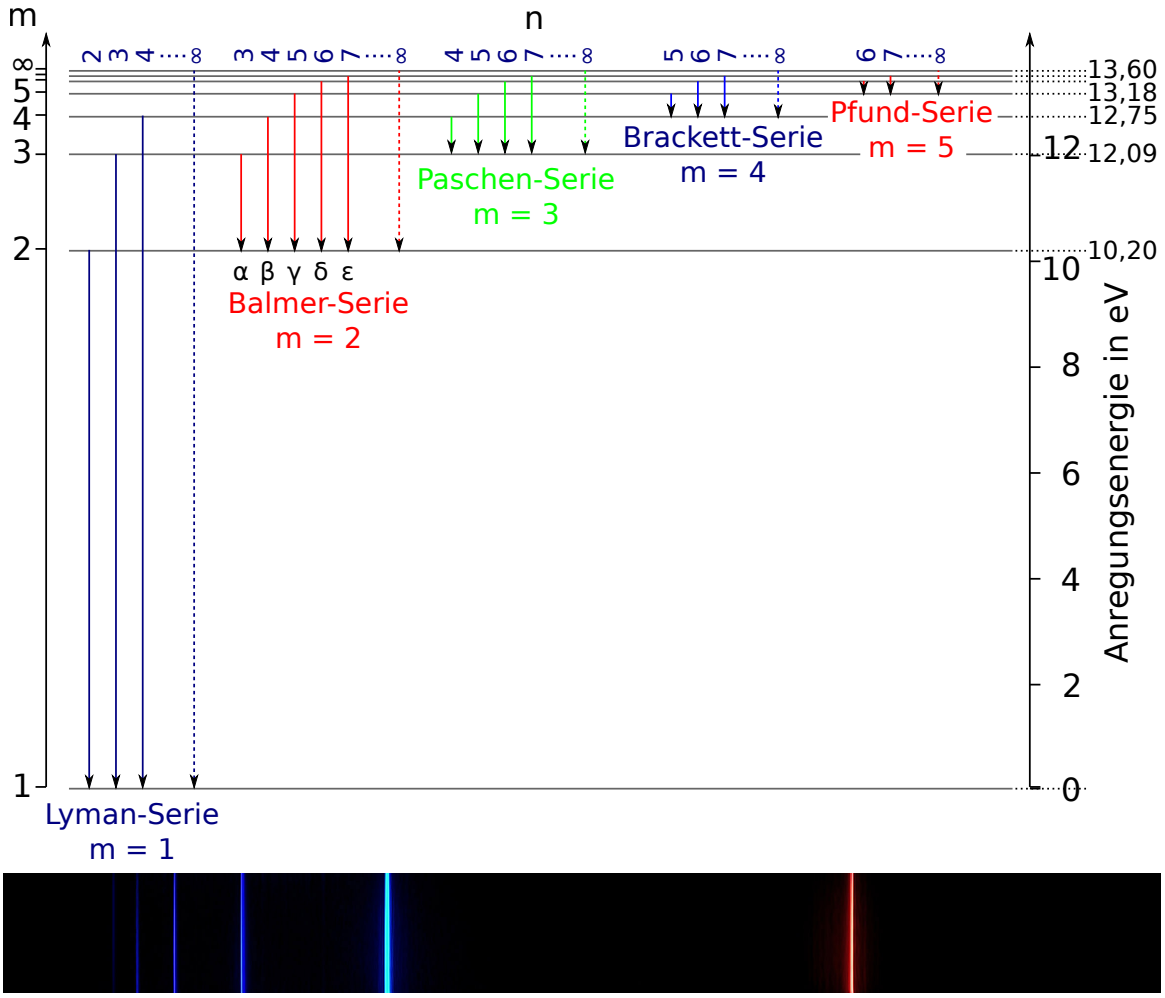


Abbildung 3.3: *Das Termschema des Wasserstoffatoms. Die Spektrallinien, also die Frequenzen des abgestrahlten Lichtes wenn man Wasserstoff verbrennt, ergeben sich aus den Energiedifferenzen beim Übergang von einem angeregten zu einem niedrigeren Energieeigenzustand,* $\hbar\omega = E_n - E_{n'}$. *Die im unteren Bild gezeigten sichtbaren Spektrallinien sind durch die Übergänge von* $n \in \{3, 4, \ldots\}$ *zu* $n = 2$, *die Balmer-Serie (benannt nach dem Lehrer* ***Johann Jakob Balmer (1825-1898)****), gegeben. Die Wellenlängen der gezeigten 6 linien (im Bild von links nach rechts angeordnet) sind 389 nm (kaum erkennbar), 397 nm (schwach violett), 410 nm (violett), 434 nm (indigo), 486 nm (türkis) und 656 nm (rot)*
Bildquellen (Wikipedia): Von PNG: w:de:user:Kiko2000; SVG: Cepheiden - SVG-conversion: Eigenes Werk (own work), CC BY-SA 3.0, https://commons.wikimedia.org/w/index.php?curid=6501211
Von Jan Homann - Eigenes Werk, CC BY-SA 3.0, https://commons.wikimedia.org/w/index.php?curid=6504291

3.14 Der Spin

Wir haben in Abschnitt 3.10 gesehen, dass aus den Drehimpulsvertauschungsrelationen allein für die simultanen Eigenwerte von[4] $\vec{\mathbf{J}}^2$ und $\mathbf{J}_3$ die Werte $\hbar^2 j(j+1)$ mit $j \in \{0, 1/2, 1, \ldots\}$ bzw. (für gegebenes j) $m\hbar$ mit $m \in \{-j, -j+1, \ldots, j-1, j\}$resultieren. Nur die spezielle Form $\vec{\mathbf{L}} = \vec{\mathbf{x}} \times \vec{\mathbf{p}}$ des Bahndrehimpulsoperators hat dann die Eigenwerte für $\vec{\mathbf{L}}^2$ auf $\hbar^2 \ell(\ell+1)$ 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 Elementarteilchen, den man **Spin** nennt.

In der klassischen Mechanik kennt man eine solche Möglichkeit nur von ausgedehnten Objekten, z.B. dem starren Körper (s. [Hee18b]). 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 bereits vor der Entdeckung der eigentlichen modernen Quantentheorie 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 der Entdeckung der modernen Quantentheorie durch Heisenberg, Born, Schrödinger und Dirac, und die Deutung des Versuchs war entsprechend auch noch fehlerhaft [FH03]. 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** $\vec{\mathbf{p}}$ 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 Ruhsystem des Schwerpunkts** ruhender Beobachter misst.

Für ein quantenmechanisches Punktteilchen können wir nun einfach entsprechende **Spin-Drehimpulsoperatoren** $\vec{\mathbf{S}}$ einführen. Wir nehmen an, dass diese Spin-Drehimpulsoperatoren ähnlich wie im klassischen Fall des freien starren Körpers von den Translationsfreiheitsgraden (also Ort $\vec{\mathbf{x}}$ und Impuls $\vec{\mathbf{p}}$) 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:

$$\begin{aligned} \left[\mathbf{x}_j, \mathbf{x}_k\right] &= 0, \quad \left[\mathbf{p}_j, \mathbf{p}_k\right] = 0, \quad \left[\mathbf{x}_j, \mathbf{p}_k\right] = \mathrm{i}\hbar\delta_{jk}\mathbb{1}, \\ \left[\mathbf{S}_j, \mathbf{x}_k\right] &= 0, \quad \left[\mathbf{S}_j, \mathbf{p}_k\right] = 0, \quad \left[\mathbf{S}_j, \mathbf{S}_k\right] = \mathrm{i}\hbar\epsilon_{jkl}\mathbf{S}_l. \end{aligned} \tag{3.14.1}$$

Entsprechend kann man als (verallgemeinerte) vollständige Orthonormalsysteme die simultanen Eigenvektoren von $\vec{\mathbf{S}}^2$ und $\mathbf{S}_3$ und entweder $\vec{\mathbf{x}}$ („Orts-Spin-Darstellung“) oder $\vec{\mathbf{p}}$ („Impuls-Spin-Darstellung“) verwenden.

Ein freies Elementarteilchen ist dann durch seine Masse und seine Spin-Quantenzahl, die den Eigenwert von $\vec{\mathbf{S}}^2$ zu $\hbar^2 s(s+1)$ festlegt, definiert. Dabei gibt es prinzipiell die Möglichkeiten $s \in \{0, 1/2, 1, \ldots\}$. Dann sind die möglichen Eigenwerte von $\mathbf{S}_3$ durch $m_s\hbar$ mit $m_s \in \{-s, -s+1, \ldots, s-1, s\}$ 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 $\mathbf{S}_3$ durch $\pm\hbar/2$ gegeben („ spin up“ und „spin down“).

[4] Wir schreiben jetzt für eine allgemeinere Art Drehimpuls als den in Abschnitt 3.10 behandelten Bahndrehimpuls $\vec{\mathbf{J}}$. Dies sollen drei Operatoren sein, die die Kommutatorrelationen $\left[\mathbf{J}_j, \mathbf{J}_k\right] = \mathrm{i}\hbar\epsilon_{jkl}\mathbf{J}_l$ erfüllen.

Betrachten wir also die simultanen Eigenvektoren der Ortskomponenten $\vec{\mathbf{x}}$ und $\mathbf{S}_3$. Da $\mathbf{S}_3$ mit allen drei Ortskomponenten $\vec{\mathbf{x}}$ kommutiert, können wir wie bei den Teilchen mit Spin 0 die Orts-Spin-Eigenvektoren durch

$$|\vec{x}, m_s\rangle = \exp\left(-\frac{\mathrm{i}}{\hbar}\vec{x}\cdot\vec{\mathbf{p}}\right)|\vec{x}=0, m_s\rangle \tag{3.14.2}$$

definieren. Dabei gilt

$$\left\langle \vec{x}, m_s \,\middle|\, \vec{x}', m_s' \right\rangle = \delta^{(3)}(\vec{x}-\vec{x}')\delta_{m_s m_s'} \tag{3.14.3}$$

und

$$\left\langle \vec{x}, m_s' \,\middle|\, \vec{p}, m_s \right\rangle = u_{\vec{p},m_s}(\vec{x}, m_s') = \frac{1}{(2\pi\hbar)^{3/2}}\exp\left(\frac{\mathrm{i}}{\hbar}\vec{x}\cdot\vec{p}\right)\delta_{m_s m_s'}. \tag{3.14.4}$$

Die Wellenfunktionen in der Orts- bzw. Impulsdarstellung sind demnach für ein Elektron **zweikomponentige Spinoren**, auch **Weyl-Spinoren**[5] genannt, dessen Komponenten durch

$$\psi_{m_s}(\vec{x}) = \langle \vec{x}, m_s \,|\, \psi\rangle \quad \text{bzw.} \quad \tilde{\psi}_{m_s}(\vec{p}) = \left\langle \vec{p}, m_s \,\middle|\, \psi \right\rangle, \quad m_s \in \{-1/2, +1/2\} \tag{3.14.5}$$

gegeben sind. Gewöhnlich ordnet man die Komponenten in Spaltenvektoren, die man in dem Fall **Spinoren** nennt, $(\psi_{+1/2}, \psi_{-1/2})^{\mathrm{T}}$ an. Die (verallgemeinerten Eigenvektoren) sind gemäß

$$\left\langle \vec{x}, m_s \,\middle|\, \vec{x}', m_s' \right\rangle = \delta^{(3)}(\vec{x}-\vec{x}')\delta_{m_s m_s'}, \quad \left\langle \vec{p}, m_s \,\middle|\, \vec{p}', m_s' \right\rangle = \delta^{(3)}(\vec{p}-\vec{p}')\delta_{m_s m_s'}, \tag{3.14.6}$$

normiert, und das Skalarprodukt lautet in der Ortsdarstellung entsprechend

$$\langle \psi_1 \,|\, \psi_2\rangle = \sum_{m_s=\pm 1/2}\int_{\mathbb{R}^3} \mathrm{d}^3x\, \psi^*_{1m_s}(\vec{x})\psi_{2m_s}(\vec{x}) \equiv \int_{\mathbb{R}^3} \mathrm{d}^3x\, \psi_1^\dagger(\vec{x})\psi_2(\vec{x}), \tag{3.14.7}$$

wobei

$$\psi(\vec{x}) = \begin{pmatrix} \psi_{+1/2}(\vec{x}) \\ \psi_{-1/2}(\vec{x}) \end{pmatrix}, \quad \psi^\dagger(\vec{x}) = \left(\psi^*_{+1/2}(\vec{x}), \psi^*_{-1/2}(\vec{x})\right). \tag{3.14.8}$$

die abkürzende Schreibweise mit **spinorwertigen Wellenfunktionen** ist. Analoge Formeln wie (3.14.7) und (3.14.8) gelten freilich auch für die Impulsdarstellungen $\tilde{\psi}(\vec{p})$ (Spinorschreibweise) bzw. die Spinorkomponenten $\tilde{\psi}_{m_s}(\vec{p})$.

Die Darstellungen der Spinoperatoren $\vec{\mathbf{S}}$ in dieser Orts-Spin- bzw. Impuls-Spin-Darstellung der Wellenfunktionen sind demnach **Matrizen im Spinorraum**, die wir mit der allgemeinen Drehimpulsalgebra wie in Abschnitt 3.10 bestimmen können. Dies ist für $s = 1/2$ sogar besonders einfach. Seien also $\boldsymbol{u}_{m_s} \in \mathbb{C}^2$ die $\mathbf{S}_3$-Eigenspinoren $\boldsymbol{u}_{1/2} = (1,0)^{\mathrm{T}}$ und $\boldsymbol{u}_{-1/2} = (0,1)^{\mathrm{T}}$, d.h. gilt

$$\hat{s}_3 u_{m_s} = \hbar m_s u_{m_s}, \quad m_s = \pm\frac{1}{2}, \tag{3.14.9}$$

so ist

$$\hat{s}_3 = \frac{\hbar}{2}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{3.14.10}$$

Um nun auch $\hat{s}_1$ und $\hat{s}_2$ zu finden, verwenden wir die Leiteroperatoren $\mathbf{S}_\pm = \mathbf{S}_1 + \pm\mathrm{i}\mathbf{S}_2$ und die Beziehungen (3.10.22) und (3.10.28), wobei wir hier

$$j = s = 1/2 \quad \text{und} \quad m = m_s \in \{1/2, -1/2\} \tag{3.14.11}$$

[5] Benannt nach dem Mathematiker Hermann Weyl, 1885-1955, der sich u.a. auch mit der Mathematik der Quantentheorie und insbesondere des Drehimpulses ausführlich beschäftigt hat.

setzen haben. Es gilt also

$$\mathbf{S}_+ \boldsymbol{u}_{m_s} = \hbar\sqrt{(1/2-m_s)(3/2+m_s)}\boldsymbol{u}_{m_s}, \quad \mathbf{S}_- \boldsymbol{u}_{m_s} = \hbar\sqrt{(1/2+m_s)(3/2-m_s)}\boldsymbol{u}_{m_s}, \tag{3.14.12}$$

und damit ist

$$\begin{aligned} &\mathbf{S}_+ \boldsymbol{u}_{1/2} = 0, \quad \mathbf{S}_+ \boldsymbol{u}_{-1/2} = \hbar \boldsymbol{u}_{1/2} \\ &\mathbf{S}_- \boldsymbol{u}_{1/2} = \hbar \boldsymbol{u}_{-1/2}, \quad \mathbf{S}_- \boldsymbol{u}_{-1/2} = 0. \end{aligned} \tag{3.14.13}$$

Damit erhält man analog wie bei $\hat{s}_3$

$$\hat{s}_+ = \begin{pmatrix} 0 & \hbar \\ 0 & 0 \end{pmatrix}, \quad \hat{s}_- = \begin{pmatrix} 0 & 0 \\ \hbar & 0 \end{pmatrix} \tag{3.14.14}$$

und wegen $\hat{s}_\pm = \hat{s}_1 \pm \mathrm{i}\hat{s}_2$

$$\begin{aligned} \hat{s}_1 &= \frac{1}{2}(\hat{s}_+ + \hat{s}_-) = \frac{\hbar}{2}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \\ \hat{s}_2 &= \frac{1}{2\mathrm{i}}(\hat{s}_+ - \hat{s}_-) = \frac{\hbar}{2}\begin{pmatrix} 0 & -\mathrm{i} \\ \mathrm{i} & 0 \end{pmatrix} \end{aligned} \tag{3.14.15}$$

Zur Vereinfachung schreibt man allgemein $\hat{s}_j = (\hbar/2)\hat{\sigma}_j$ ($j \in \{1,2,3\}$ mit den **Pauli-Matrizen** (Wolfgang Pauli, 1900-1958)

$$\hat{\sigma}_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \hat{\sigma}_2 = \begin{pmatrix} 0 & -\mathrm{i} \\ \mathrm{i} & 0 \end{pmatrix}, \quad \hat{\sigma}_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{3.14.16}$$

Betrachten wir nun die Drehungen. Die Orts-Spin-Eigenvektoren transformieren sich unter Drehungen offenbar gemäß

$$\mathbf{D}(\vec{\phi}) = \exp\left[-\frac{\mathrm{i}}{\hbar}\vec{\phi}\cdot(\vec{\mathbf{L}}+\vec{\mathbf{s}})\right] = \exp\left(-\frac{\mathrm{i}}{\hbar}\vec{\phi}\cdot\vec{\mathbf{L}}\right)\exp\left(-\frac{\mathrm{i}}{\hbar}\cdot\vec{\mathbf{s}}\right) = \mathbf{D}^{(l)}(\vec{\phi})\mathbf{D}^{(s)}(\vec{\phi}). \tag{3.14.17}$$

Letzteres folgt daraus, dass die Spinkomponenten mit $\vec{\mathbf{x}}$ und $\vec{\mathbf{p}}$ und damit auch mit $\vec{\mathbf{L}}$ vertauschen. Für die Orts-Spin-Eigenvektoren gilt gemäß (3.7.40) für die Wirkung des Bahndrehimpulsanteils der Drehungen auf den Ortseigenvektor

$$\mathbf{D}(\vec{\phi})|\vec{x}, m_s\rangle = \sum_{m_s' \in \{1/2,-1/2\}} \left|\hat{R}(\varphi)\vec{x}, m_s'\right\rangle D^{(s)}_{m_s' m_s}(\vec{\phi}), \tag{3.14.18}$$

In der Orts-Spin-Darstellung transformiert sich demnach die Wellenfunktion unter Drehungen gemäß

$$\begin{aligned} \mathbf{D}(\varphi)\psi_{m_s}(\vec{x}) &= \left\langle \vec{x}, m_s \middle| \mathbf{D}(\vec{\phi}) \middle| \psi \right\rangle \\ &= \left\langle \mathbf{D}^\dagger(\vec{\phi})\vec{x}, m_s \middle| \psi \right\rangle \\ &= \sum_{m_s'} \left\langle D^\dagger_{m_s' m_s} \hat{R}^{-1}\vec{x}, m_s' \middle| \psi \right\rangle \\ &= \sum_{m_s'} \left\langle D^*_{m_s m_s'} \hat{R}^{-1}\vec{x}, m_s' \middle| \psi \right\rangle \\ &= \sum_{m_s'} D_{m_s m_s'}(\vec{\phi})\psi_{m_s'}[\hat{R}^{-1}(\vec{\phi})\vec{x}]. \end{aligned} \tag{3.14.19}$$

In der Spinorschreibweise ist also

$$\mathbf{D}(\vec{\phi})\psi(\vec{x}) = \hat{D}(\vec{\phi})\psi[\hat{R}^{-1}(\vec{\phi})]. \tag{3.14.20}$$

Dabei ist $\hat{D}$ aufgrund von (3.14.17) durch

$$\hat{D}(\vec{\phi}) = \exp\left(-\frac{\mathrm{i}}{\hbar}\vec{\phi}\cdot\hat{\vec{S}}\right) = \exp\left(-\frac{\mathrm{i}}{2}\vec{\phi}\cdot\hat{\vec{\sigma}}\right) \tag{3.14.21}$$

definiert. Wir können nun die Matrixexponentialfunktion mit Hilfe der Pauli=Matrizen (3.14.16) durch die entsprechende Reihenentwicklung explizit berechnen. Dazu setzen wir $\vec{\phi} = \phi\vec{n}$ mit $\vec{n}^2 = 1$. Dann enötigen wir die Potenzen $(\vec{n}\cdot\hat{\vec{\sigma}})^k$ für $k \in \mathbb{N}_0$. Offenbar ist, unter Verwendung der Einsteinschen Summenkonvention,

$$(\vec{n}\cdot\hat{\vec{\sigma}})^2 = n_a n_b \hat{\sigma}_a \hat{\sigma}_b = \frac{1}{2} n_a n_b \{\hat{\sigma}_a, \hat{\sigma}_b\}. \tag{3.14.22}$$

Dabei haben wir den **Antikommutator** für Matrizen (bzw. Operatoren im Hilbertraum), der durch

$$\{\mathbf{A}, \mathbf{B}\} = \mathbf{AB} + \mathbf{BA} \tag{3.14.23}$$

definiert ist, verwendet. Durch direktes Nachrechnen ergibt sich für die Pauli-Matrizen

$$\{\hat{\sigma}_a, \hat{\sigma}_b\} = 2\delta_{ab}\hat{1}. \tag{3.14.24}$$

Demnach ist

$$(\vec{n}\cdot\hat{\vec{\sigma}})^2 = \vec{n}^2\hat{1} = \hat{1}. \tag{3.14.25}$$

Damit folgt durch Iteration

$$(\vec{n}\cdot\hat{\vec{\sigma}})^{2j} = 1, \quad (\vec{n}\cdot\hat{\vec{\sigma}})^{2j+1} = \vec{n}\cdot\hat{\vec{\sigma}} \quad \text{für} \quad j \in \mathbb{N}_0. \tag{3.14.26}$$

Schließlich ergibt sich daraus für (3.14.27)

$$\begin{aligned}
\hat{D}(\vec{\phi}) &= \sum_{j=0}^{\infty} \frac{1}{(2j)!}\left(\frac{-\mathrm{i}\phi}{2}\right)^{2j}\hat{1} + \sum_{j=0}^{\infty}\frac{1}{(2j+1)!}\left(\frac{-\mathrm{i}\phi}{2}\right)^{2j+1}\vec{n}\cdot\hat{\vec{\sigma}} \\
&= \sum_{j=0}^{\infty}\frac{(-1)^j}{(2j)!}\left(\frac{\phi}{2}\right)^{2j}\hat{1} - \mathrm{i}\vec{n}\cdot\hat{\vec{\sigma}}\sum_{j=0}^{\infty}\frac{(-1)^j}{(2j+1)!}\left(\frac{\phi}{2}\right)^{2j+1} \\
&= \cos\left(\frac{\phi}{2}\right)\hat{1} - \mathrm{i}\vec{n}\cdot\hat{\vec{\sigma}}\sin\left(\frac{\phi}{2}\right).
\end{aligned} \tag{3.14.27}$$

Zur späteren Verwendung berechnen wir noch die Eigenvektoren für Spinkomponenten in einer beliebigen Richtung $\vec{n}$. Dazu seien die entsprechenden Polar- und Asimuthwinkel von $\vec{n}$ durch α und β gegeben:

$$\vec{n} = \begin{pmatrix} n_1 \\ n_2 \\ n_3 \end{pmatrix} = \begin{pmatrix} \sin\beta\cos\alpha \\ \sin\beta\sin\alpha \\ \cos\beta \end{pmatrix}. \tag{3.14.28}$$

Diesen erhalten wir offenbar, indem wir zuerst $\vec{e}_3$ um $\vec{e}_2$ um den Winkel β und dann um $\vec{e}_3$ um den Winkel α drehen. Mit (3.7.41) folgt in der Tat

$$\begin{aligned}
\hat{R}(\alpha\vec{e}_3)\hat{R}(\beta\vec{e}_2)\vec{e}_3 &= \hat{R}(\alpha\vec{e}_3)[(\vec{e}_2\times\vec{e}_3)\times\vec{e}_2\cos\beta + \vec{e}_2\times\vec{e}_3\sin\beta \\
&= \hat{R}(\alpha\vec{e}_3)(\vec{e}_3\cos\beta + \vec{e}_1\sin\beta) \\
&= \vec{e}_3\cos\beta + \vec{e}_1\sin\beta\cos\phi + \vec{e}_2\sin\beta\sin\phi \\
&= \begin{pmatrix} \sin\beta\cos\alpha \\ \sin\beta\sin\alpha \\ \cos\beta \end{pmatrix} = \vec{n}.
\end{aligned} \tag{3.14.29}$$

Für die Eigenspinoren $u_{1/2} = (1,0)^{\mathrm{T}}$ und $u_{-1/2} = (0,1)^{\mathrm{T}}$ von $\hat{s}_3 = \hbar\hat{\sigma}_3/2$ zu den Eigenwerten $\pm\hbar/2$ sollten demnach

$$u_{m_s}(\vec{n}) = \hat{D}^{(s)}(\alpha\vec{e}_3)\hat{D}^{(s)}(\beta\vec{e}_2)u_{m_s} \tag{3.14.30}$$

die entsprechenden Eigenvektoren von $\vec{n}\cdot\hat{s}_3$ sein. Mit (3.14.27) erhalten wir für die entsprechende Matrix

$$\begin{aligned}\hat{D}^{(s)}(\alpha\vec{e}_3)\hat{D}^{(s)}(\beta\vec{e}_2) &= [\cos(\alpha/2) - \mathrm{i}\hat{\sigma}_3\sin(\alpha/2)][\cos(\beta/2) - \mathrm{i}\hat{\sigma}_2\sin(\beta 2)\\ &= \begin{pmatrix}\cos(\beta/2)\exp(-\mathrm{i}\alpha/2) & -\sin(\beta/2)\exp(-\mathrm{i}\alpha/2)\\ \sin(\beta/2)\exp(\mathrm{i}\alpha/2) & \cos(\beta/2)\exp(\mathrm{i}\alpha/2)\end{pmatrix}.\end{aligned} \tag{3.14.31}$$

Es ist also

$$u_{1/2}(\vec{n}) = \begin{pmatrix}\cos(\beta/2)\exp(-\mathrm{i}\alpha/2)\\ \sin(\beta/2)\exp(\mathrm{i}\alpha/2)\end{pmatrix}, \quad u_{-1/2}(\vec{n}) = \begin{pmatrix}-\sin(\beta/2)\exp(-\mathrm{i}\alpha/2)\\ \cos(\beta/2)\exp(\mathrm{i}\alpha/2)\end{pmatrix}. \tag{3.14.32}$$

Die bemerkenswerteste Eigenschaft der Wellenfunktionen für Teilchen mit halbzahligem Spin ist, dass bei einer Rotation um einen vollen Winkel 2π um eine beliebige Achse die Wellenfunktion das Vorzeichen wechselt. Für Spin 1/2 erkennt man das direkt aus dem Transformationsverhalten unter Drehungen (3.14.20) und der expliziten Form der entsprechenden Spinor-Drehmatrix (3.14.27). Für beliebiges $\vec{n}$ ist offenbar $\hat{D}(2\pi\vec{n}) = -\mathbb{1}$, d.h. die Spinorwellenfunktion ändert unter beliebigen Drehungen um den Drehwinkel 2π ihr Vorzeichen, und erst bei zwei Umläufen, also einer Drehung um 4π nimmt sie wieder ihren ursprünglichen Wert an. Physikalisch ist der Faktor (-1) nun aber unerheblich, da ja wieder nur $P(\vec{x}, m_s) = |\psi_{m_s}(\vec{x})|^2$ die physikalische Bedeutung der Ortswahrscheinlichkeitsdichte und Wahrscheinlichkeit, s_3-Messung den Wert m_s zu erhalten, besitzt.

Wichtig ist nur, dass man keine Superpositionen von Zuständen von Teilchen mit halb- und ganzzahligem Spin bilden darf, da sonst eine Drehung um 2π nicht mehr einfach auf dieselbe Wellenfunktion mit einem zusätzlichen Phasenfaktor ± 1 führt. Dann ergibt sich ein Widerspruch zur Eindeutigkeit der Ortswahrscheinlichkeitsdichte unter Drehungen um 2π. Dies ist ein Beispiel für eine sog. **Superauswahlregel**, die bestimmte Superpositionen aufgrund von Symmetrien verbietet, hier also eine Superposition von Zuständen von Teilchen mit ganz- und halbzahligem Spin.

3.15 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

$$\vec{\nabla}\cdot\vec{j} = 0 \tag{3.15.1}$$

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

$$\vec{F} = \int_{K_R} \mathrm{d}^3x\, \vec{j}(\vec{x}) \times \vec{B}(\vec{x}). \tag{3.15.2}$$

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

$$\vec{B}(\vec{x}) = \vec{B}(\vec{0}) + (\vec{x}\cdot\vec{\nabla})\vec{B}(\vec{0}) \tag{3.15.3}$$

im Integral (3.15.3) zu verwenden. Dann erhalten wir

$$\vec{F} = \int_{K_R} \mathrm{d}^3x\, \vec{j}(\vec{x}) \times [\vec{B}(\vec{0}) + (\vec{x}\cdot\vec{\nabla})\vec{B}(0)]. \tag{3.15.4}$$

Wir wollen nun zeigen, dass der erste Term verschwindet.
Sei dazu $f(\vec{x})$ eine beliebige Funktion. Dann gilt

$$\vec{\nabla}\cdot(f\vec{j}) = \partial_k(f\, j_k) = j_k\partial_k f + f\,\partial_k j_k = j_k\partial_k f\,, \tag{3.15.5}$$

wobei wir im letzten Schritt (3.15.1) verwendet haben. Integrieren wir diese Gleichung über die Kugel K_R, folgt mit dem Gaußschen Integralsatz und wegen $\vec{j}(\vec{x}) = 0$ für $r = |\vec{x}| \geq R$

$$\int_{K_R} \mathrm{d}^3x\, \vec{j}\cdot\vec{\nabla} f = \int_{\partial K_R} \mathrm{d}^2\vec{f}\cdot(f\vec{j}) = 0. \tag{3.15.6}$$

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 (3.15.6)

$$\int_{K_R} \mathrm{d}^3\vec{x}\, j_k = 0. \tag{3.15.7}$$

Dies zeigt, dass nur der zweite Term in (3.15.4) einen Beitrag zur Kraft liefert. Wir schreiben diesen Term nun noch ein wenig um:

$$F_k = \int_{K_R} \mathrm{d}^3x\, \epsilon_{klm} j_l x_n \partial_n B_m(\vec{0}). \tag{3.15.8}$$

Setzen wir nun in (3.15.6) $f = x_l x_n$, liefert (3.15.6)

$$\int_{K_R} \mathrm{d}^3x\, j_k\partial_k(x_l x_n) = \int_{K_R} \mathrm{d}^3x\, j_k(\delta_{kl}x_n + \delta_{kn}x_l) = \int_{K_R} \mathrm{d}^3x\,(j_l x_n + j_n x_l) = 0. \tag{3.15.9}$$

Wir verwenden dies nun, um (3.15.7) weiter umzuformen:

$$F_k = \frac{1}{2}\int_{K_R} \mathrm{d}^3x\epsilon_{klm}(x_n j_l - x_l j_n)\partial_n B_m(\vec{0}) = \frac{1}{2}\int_{K_R} \mathrm{d}^3x\epsilon_{klm}\epsilon_{nla}(\vec{x}\times\vec{j})_a\partial_n B_m(\vec{0}). \tag{3.15.10}$$

Wir definieren nun das **magnetische Moment** der Stromverteilung durch

$$\vec{\mu} = \frac{1}{2}\int_{K_R} \mathrm{d}^3x\, \vec{x}\times\vec{j}. \tag{3.15.11}$$

Dann wird (3.15.10) zu

$$\begin{aligned} F_k &= \epsilon_{lkm}\epsilon_{lna}\mu_a\partial_n B_m(\vec{0}) \\ &= (\delta_{kn}\delta_{ma} - \delta_{ka}\delta_{mn})\mu_a\partial_n B_m(\vec{0}) \\ &= \partial_k(\mu_a B_a)_{\vec{x}=\vec{0}} \\ \Rightarrow\ \vec{F} &= \vec{\nabla}[\vec{\mu}\cdot\vec{B}(\vec{x})]_{\vec{x}=\vec{0}}. \end{aligned} \tag{3.15.12}$$

Dabei haben wir verwendet, dass $\vec{\mu} = \text{const}$ und $\vec{\nabla}\cdot\vec{B} = 0$ ist. Damit ist also die Kraft auf die am Koordinatenursprung lokalisierte Stromverteilung eine konservative Kraft mit dem Potential

$$V(\vec{x}) = -\vec{\mu}\cdot\vec{B}(\vec{x}). \tag{3.15.13}$$

Demnach ist

$$\vec{F}(\vec{x}) = -\vec{\nabla} V(\vec{x}). \tag{3.15.14}$$

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

$$\vec{\mu} = \frac{q}{2}\int_{K_R} \mathrm{d}^3x\, n(\vec{x})\vec{x} \times \vec{v}(\vec{x}) = \frac{q}{2m}\vec{L}, \tag{3.15.15}$$

wobei

$$\vec{L} = \int_{K_r} \mathrm{d}^3x\, mn(\vec{x})\vec{x} \times \vec{v}(\vec{x}) \tag{3.15.16}$$

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

$$\vec{M} = \int_{K_R} \mathrm{d}^3x\, \vec{x} \times (\vec{j} \times \vec{B}_0) = \int_{K_R} \mathrm{d}^2x\, [\vec{j}(\vec{x} \cdot \vec{B}_0) - \vec{B}_0(\vec{x} \cdot \vec{j})]. \tag{3.15.17}$$

Verwenden wir nun (3.15.6) 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

$$M_k = \int_{K_R} \mathrm{d}^3x\, j_k x_l B_{0l} = \frac{1}{2}\int_{K_R} \mathrm{d}^3x (x_l j_k - x_k j_l) B_{0l} = \frac{1}{2}\int_{K_R} \mathrm{d}^3x\, \epsilon_{lkm} (\vec{x} \times \vec{j})_m B_{0l} = (\vec{\mu} \times \vec{B}_0)_k. \tag{3.15.18}$$

Dabei haben wir im letzten Schritt die Definition des magnetischen Moments (3.15.11) verwendet. Es ist also

$$\vec{M} = \vec{\mu} \times \vec{B}(\vec{x}), \tag{3.15.19}$$

wenn wir uns die Stromverteilung wieder allgmein um den Ort $\vec{x}$ verteilt denken.
Verwenden wir nun (3.15.15), folgt für die Bewegungsgleichung für den Drehimpuls der Stromverteilung

$$\dot{\vec{L}} = \frac{2m}{q}\dot{\vec{\mu}} = \vec{\mu} \times \vec{B} \Rightarrow \dot{\vec{\mu}} = \frac{q}{2m}\vec{\mu} \times \vec{B}. \tag{3.15.20}$$

Das bedeutet, dass in einem Magnetfeld das magnetische Moment der Stromverteilung mit der **klassischen Larmor-Frequenz** (Joseph Larmor 1857-1942)

$$\omega_{\text{Larmor}} = \frac{q|\vec{B}|}{2m} \tag{3.15.21}$$

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 (3.15.14) und ein Drehmoment gemäß (3.15.19).

3.16 Teilchen mit $s = 0$ im Magnetfeld

Wir betrachten nun zunächst ein Teilchen mit Spin 0 im Magnetfeld. Wir geben dazu den Hamilton-Operator an:

$$\mathbf{H} = \frac{1}{2M}[\vec{\mathbf{p}} - q\vec{A}(\vec{\mathbf{x}})]^2. \tag{3.16.1}$$

Dabei ist $\vec{A}$ das Vektorpotential des Magnetfeldes, d.h.

$$\vec{B} = \vec{\nabla} \times \vec{A}. \tag{3.16.2}$$

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 (3.1.3) für den Operator, der die Zeitableitung einer Observablen repräsentiert, den Operator für die Geschwindigkeit des Teilchens:

$$\vec{\mathbf{v}} = \overset{\circ}{\vec{\mathbf{x}}} = \frac{1}{\mathrm{i}\hbar}[\vec{\mathbf{x}}, \mathbf{H}] = \frac{1}{M}[\vec{\mathbf{p}} - q\vec{A}(\vec{\mathbf{x}})]. \tag{3.16.3}$$

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 $\vec{\mathbf{p}} = M\vec{\mathbf{v}} + q\vec{A}(\vec{\mathbf{x}})$ 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. Anhang E und [Hee18b, Hee08, BFK$^+$18a, LL97]). Jedenfalls ergibt sich mit (3.1.3) für den Operator der auf das Teilchen wirkenden Kraft

$$\vec{\mathbf{F}} = M\overset{\circ}{\vec{\mathbf{v}}} = \frac{1}{\mathrm{i}\hbar}[M\vec{\mathbf{v}}, \mathbf{H}] = \frac{q}{2}[\vec{\mathbf{v}} \times \vec{B}(\vec{\mathbf{x}}) - \vec{B}(\vec{\mathbf{x}}) \times \vec{\mathbf{v}}]. \tag{3.16.4}$$

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 $\vec{\mathbf{v}}$ und $\vec{\mathbf{x}}$ um Operatoren handelt, dürfen wir in (3.16.4) die Faktoren in den Vektorprodukten nicht einfach vertauschen. Allerdings ist (3.16.4) 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 $\hat{\vec{p}} = -\mathrm{i}\hbar\vec{\nabla}$ repräsentiert wird. Der Hamilton-Operator (3.16.1) lautet also in der Ortsdarstellung

$$\hat{H} = \frac{1}{2M}[-\mathrm{i}\hbar\vec{\nabla} - q\vec{A}(\vec{x})]^2. \tag{3.16.5}$$

3.17 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 3.15 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

$$\vec{\mu} = g\frac{q}{2M}\vec{\mathbf{S}}. \tag{3.17.1}$$

Der gegenüber dem klassischen Resultat, für das gemäß (3.15.15) $g = 1$ gilt, zusätzliche Faktor g heißt **Landé-Faktor** (Alfred Landé 1888-1976) oder **Gyrofaktor**. Die typische Größenordnung des Betrags des magnetischen Moments für das Elektron ist das **Bohrsche Magneton** $\mu_B = e\hbar/(2M_e) = 5{,}7883818060(17)\cdot 10^{-5}\mathrm{eV/T}$. Setzt man $g = 2$, erhält man, dass eine Komponente des magnetischen Moments eines Elektrons in einer Richtung die beiden Werte $\pm\mu_B$ annehmen kann („Richtungsquantelung").

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 Messungenauigkeit 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

$$\mathbf{H} = \frac{1}{2M}[\vec{\mathbf{p}} - q\vec{A}(\vec{\mathbf{x}})]^2 - \vec{\mu}\cdot\vec{B}(x), \tag{3.17.2}$$

d.h. zum Hamilton-Operator eines Teilchens ohne Spin im Magnetfeld (3.16.1) ist noch das Potential der Kraft aufgrund des magnetische Moments des Teilchens (3.15.13) zu addieren, freilich in quantisierter Form. Da $\vec{\mu} \propto \vec{\mathbf{S}}$ ist und $\vec{\mathbf{S}}$ mit $\vec{\mathbf{x}}$ kommutiert, treten dabei keine Probleme mit der Operatorordnung auf.

3.18 Teilchen im homogenen Magnetfeld

Wir betrachten nun ein Teilchen mit Spin 0 und $\vec{\mu} = 0$ im homogenen Magnetfeld $\vec{B} = B\vec{e}_3$ ($B = \mathrm{const}$) und suchen die Energieeigenfunktionen. Wir benötigen zunächst ein Vektorpotential, das wegen der Eichinvarianz der Elektrodynamik (vgl. das Skript zur Elektrodynamik [Hee18c]) nur bis auf ein Gradientenfeld bestimmt ist.

Wir wollen im Folgenden zeigen, dass die Energieeigenwerte unabhängig von der Eichung sind und die Energieeigenfunktionen für verschiedene Eichungen jeweils äquivalente orthonormierte Basen bilden.

Für diesen Fall wird (3.17.2) zu

$$\mathbf{H} = \frac{1}{2M}[\vec{\mathbf{p}} - q\vec{A}(\vec{\mathbf{x}})]^2 \tag{3.18.1}$$

Durch Vergleich mit (3.16.3) sehen wir, dass dies der klassischen Formel $\mathbf{H} = M\vec{\mathbf{v}}^2/2$ entspricht. Der Operator

für den gewöhnlichen mechanischen Impuls ist

$$\vec{\pi} = M\vec{\mathbf{v}} = \vec{\mathbf{p}} - q\vec{A}(\vec{\mathbf{x}}) = -\mathrm{i}\hbar\vec{\nabla} - q\vec{A}(\vec{\mathbf{x}}). \tag{3.18.2}$$

Allerdings kommutieren diese Operatoren nicht, denn es gilt offenbar für eine beliebige Wellenfunktion $\psi = \psi(\vec{x})$

$$\begin{aligned}\pi_j\pi_k\psi &= (-\mathrm{i}\hbar\partial_j - qA_j)(-\mathrm{i}\hbar\partial_k - qA_k)\psi(\vec{x}) \\ &= -\hbar^2\partial_j\partial_k\psi + \mathrm{i}q\hbar\partial_j(A_k\psi) + \mathrm{i}qA_j\partial_k\psi + q^2A_jA_k\psi \\ &= -\hbar^2\partial_j\partial_k\psi + \mathrm{i}q\hbar(\partial_jA_k)\psi + \mathrm{i}qA_k\partial_j\psi + \mathrm{i}qA_j\partial_k\psi + q^2A_jA_k\psi.\end{aligned} \tag{3.18.3}$$

Vertauschen wir in dieser Formel j und k, folgt

$$\pi_k\pi_j = -\hbar^2\partial_k\partial_j\psi + \mathrm{i}q\hbar(\partial_kA_j)\psi + \mathrm{i}qA_j\partial_k\psi + \mathrm{i}qA_k\partial_j\psi + q^2A_kA_j\psi. \tag{3.18.4}$$

Ziehen wir dies nun von (3.18.3) ab, ergibt sich

$$\left[\pi_j, \pi_k\right] = \mathrm{i}q\hbar(\partial_jA_k - \partial_kA_j) = \mathrm{i}q\hbar\epsilon_{jkl}B_l. \tag{3.18.5}$$

Für unseren Fall mit $B_1 = B_2 = 0$ und $B_3 = B$ ist also

$$\begin{aligned}[\pi_1, \pi_2] &= \mathrm{i}q\hbar\epsilon_{12l}B_l = \mathrm{i}q\hbar B, \\ [\pi_2, \pi_3] &= \mathrm{i}q\hbar\epsilon_{23l}B_l = \mathrm{i}q\hbar B_1 = 0, \\ [\pi_3, \pi_1] &= \mathrm{i}q\hbar\epsilon_{31l}B_l = \mathrm{i}q\hbar B_2 = 0.\end{aligned} \tag{3.18.6}$$

Daraus folgt weiter (unter Verwendung der Einsteinschen Summenkonvention)

$$\left[\vec{\pi}^2, \pi_k\right] = \left[\pi_j\pi_j, \pi_k\right] = \pi_j\left[\pi_j, \pi_k\right] + \left[\pi_j, \pi_k\right]\pi_k = 2\mathrm{i}q\hbar\epsilon_{jkl}B_l\pi_j. \tag{3.18.7}$$

Dabei haben wir im letzten Schritt (3.18.5) verwendet und beachtet, dass in unserem Fall eines *konstanten* Magnetfeldes B_l eine von $\vec{x}$ unabhängige Zahl ist. Insbesondere folgt für $k = 3$ und wegen $\mathbf{H} = \vec{\pi}^2/(2m)$

$$\left[\vec{\pi}^2, \pi_3\right] = 0 \Rightarrow [\mathbf{H}, \pi_3] = 0. \tag{3.18.8}$$

Weiter kommutiert demnach noch

$$\mathbf{H}_\perp = \frac{1}{2m}(\pi_1^2 + \pi_2^2) = \mathbf{H} - \frac{1}{2m}\pi_3^2 \tag{3.18.9}$$

sowohl mit $\mathbf{H}$ als auch mit π_3.

Wir können also simultane Eigenfunktionen der beiden Operatoren $\mathbf{H}_\perp$ und π_3 finden, und zwar *unabhängig* von der konkreten Wahl des Vektorpotentials $\vec{A}$, das ja nur bis auf ein Gradientenfeld bestimmt ist. Dann ist offenbar auch $\mathbf{H}$ diagonalisiert. Wir benötigen noch eine dritte unabhängige kompatible Observable, um die Eigenfunktionen eindeutig festlegen zu können. Wir werden diese dritte Observable unten noch bestimmen.

Wir können nun die Eigenwerte und Eigenfunktionen finden, indem wir für $\mathbf{H}_\perp$ Erzeugungs- und Vernichtungsoperatoren analog zum Vorgehen beim harmonischen Oszillator einführen. Dies ist in unserem Spezialfall eines homogenen Magnetfeldes aufgrund der Vertauschungsrelationen (3.18.6) möglich. Bilden wir nämlich

$$\mathbf{a} = \frac{1}{\sqrt{2q\hbar B}}(\pi_1 + \mathrm{i}\pi_2), \tag{3.18.10}$$

folgt

$$\left[\mathbf{a},\mathbf{a}^\dagger\right]=\frac{1}{2q\hbar B}\left[\pi_1+\mathrm{i}\pi_2,\pi_1-\mathrm{i}\pi_2\right]=\frac{-2\mathrm{i}}{2q\hbar B}\left[\pi_1,\pi_2\right]=\mathbb{1}. \tag{3.18.11}$$

Außerdem ist

$$\mathbf{N}_a=\mathbf{a}^\dagger\mathbf{a}=\frac{1}{2q\hbar B}(\pi_1-\mathrm{i}\pi_2)(\pi_1+\mathrm{i}\pi_2)=\frac{1}{2q\hbar B}(\pi_1^2+\pi_2^2+\mathrm{i}[\pi_1,\pi_2])=\frac{M}{q\hbar B}\mathbf{H}_\perp-\frac{1}{2}\mathbb{1}. \tag{3.18.12}$$

Es ist also

$$\mathbf{H}_\perp=\hbar\omega\left(\mathbf{N}_a+\frac{1}{2}\right)\quad\text{mit}\quad\omega=\frac{qB}{M}. \tag{3.18.13}$$

Die Kreisfrequenz ω ist dabei die **Zyklotronfrequenz** mit der das Teilchen im entsprechenden klassischen Problem um das konstante Magnetfeld kreist. $\mathbf{H}_\perp$ ist der entsprechende Hamilton-Operator für einen harmonischen Oszillator, der mit dieser Kreisfrequenz schwingt. Die möglichen Eigenwerte von $\mathbf{H}_\perp$ sind folglich

$$E_{\perp,n_a}=\hbar\omega\left(n_a+\frac{1}{2}\right)\quad\text{mit}\quad n_a\in\mathbb{N}_0=\{0,1,2,\ldots\}. \tag{3.18.14}$$

Diese Energieeigenwerte nennt man **Landau-Level** (benannt nach L. D. Landau[6], der als erster dieses Eigenwertproblem gelöst hat).

Damit ist aber der Anteil der Bewegung in der Ebene senkrecht zu $\vec{B}$ noch nicht vollständig charakterisiert. Es muss noch eine zweite zu $\mathbf{H}_\perp$ kompatible Observable für die Bewegung in der xy-Ebene geben. Um diese zu finden, betrachten wir die Kommutatoren von $\mathbf{x}_1$ und $\mathbf{x}_2$ mit $\mathbf{a}$. Es ist offenbar (wieder unabhängig von der konkreten Wahl von $\vec{A}(\vec{\mathbf{x}})$!)

$$\left[\mathbf{x}_j,\pi_k\right]=\left[\mathbf{x}_j,\mathbf{p}_k-qA_k(\vec{\mathbf{x}})\right]=\left[\mathbf{x}_j,\mathbf{p}_k\right]=\mathrm{i}\hbar\delta_{jk} \tag{3.18.15}$$

und damit

$$\begin{aligned}[\mathbf{x}_1,\mathbf{a}]&=\frac{1}{\sqrt{2q\hbar B}}[\mathbf{x}_1,\pi_1+\mathrm{i}\pi_2]=\mathrm{i}\sqrt{\frac{\hbar}{2qB}}\mathbb{1},\\ [\mathbf{x}_2,\mathbf{a}]&=\frac{1}{\sqrt{2q\hbar B}}[\mathbf{x}_2,\pi_1+\mathrm{i}\pi_2]=-\sqrt{\frac{\hbar}{2qB}}\mathbb{1}.\end{aligned} \tag{3.18.16}$$

Adjungieren dieser Gleichungen liefert

$$\begin{aligned}\left[\mathbf{a}^\dagger,\mathbf{x}_1\right]&=-\mathrm{i}\sqrt{\frac{\hbar}{2qB}}\mathbb{1}\\ \left[\mathbf{a}^\dagger,\mathbf{x}_2\right]&=-\sqrt{\frac{\hbar}{2qB}}\mathbb{1}.\end{aligned} \tag{3.18.17}$$

Definieren wir nun den Operator

$$\mathbf{b}=\mathrm{i}\mathbf{a}^\dagger+\sqrt{\frac{qB}{2\hbar}}(\mathbf{x}_1-\mathrm{i}\mathbf{x}_2), \tag{3.18.18}$$

ergibt sich aus (3.18.16) und (3.18.17)

$$\begin{aligned}[\mathbf{b},\mathbf{a}]&=\mathrm{i}\left[\mathbf{a}^\dagger,\mathbf{a}\right]+\sqrt{\frac{qB}{2\hbar}}[\mathbf{x}_1-\mathrm{i}\mathbf{x}_2,\mathbf{a}]=0,\\ \left[\mathbf{b},\mathbf{a}^\dagger\right]&=\sqrt{\frac{qB}{2\hbar}}\left[\mathbf{x}_1-\mathrm{i}\mathbf{x}_2,\mathbf{a}^\dagger\right]=0.\end{aligned} \tag{3.18.19}$$

[6] Lew Dawidowitsch Landau (1908-1968)

Weiter gilt

$$\begin{aligned}\left[\mathbf{b},\mathbf{b}^\dagger\right] &= \left[\mathrm{i}\mathbf{a}^\dagger + \sqrt{\frac{qB}{2\hbar}}(\mathbf{x}_1 - \mathrm{i}\mathbf{x}_2), -\mathrm{i}\mathbf{a} + \sqrt{\frac{qB}{2\hbar}}(\mathbf{x}_1 + \mathrm{i}\mathbf{x}_2)\right] \\ &= \left[\mathbf{a}^\dagger,\mathbf{a}\right] + \mathrm{i}\sqrt{\frac{qB}{2\hbar}}\left[\mathbf{a}^\dagger,\mathbf{x}_1 + \mathrm{i}\mathbf{x}_2\right] - \mathrm{i}\sqrt{\frac{qB}{2\hbar}}\left[\mathbf{x}_1 - \mathrm{i}\mathbf{x}_2,\mathbf{a}\right] = \mathbb{1}.\end{aligned} \tag{3.18.20}$$

Damit ist

$$\mathbf{N}_b = \mathbf{b}^\dagger\mathbf{b} \tag{3.18.21}$$

ein Zahloperator für die Erzeugungs- und Vernichtungsoperatoren $\mathbf{b}^\dagger$ und $\mathbf{b}$, der offenbar mit $\mathbf{H}_\perp$ und π_3 kommutiert. Als vollständigen Satz kompatibler Observabler, die zudem auch unabhängig von der Eichung des Vektorpotentials sind, können wir also in der Tat $\mathbf{N}_a$ (oder $\mathbf{H}_\perp$), $\mathbf{N}_b$ und π_3 wählen. Die Eigenwerte von $\mathbf{N}_b$ sind $n_b \in \mathbb{N}_0 = \{0,1,2,\ldots\}$.
Um die physikalische Bedeutung dieser weiteren Erhaltungsgröße zu verstehen, bemerken wir, dass (*Nachrechnen!*)

$$\mathbf{N}_b = \mathbf{N}_a + \frac{1}{\hbar}\mathbf{\Lambda}_3 + \frac{m\omega}{2\hbar}(\mathbf{x}_1^2 + \mathbf{x}_2^2) \quad \text{mit} \quad \vec{\mathbf{\Lambda}} = \vec{\mathbf{x}} \times \vec{\pi} \tag{3.18.22}$$

ist. Unsere Rechnung zeigt, dass $\mathbf{N}_b - \mathbf{N}_a$ eine für die Bewegung eines geladenen Teilchens im Magnetfeld spezifische drehimpulsähnliche Erhaltungsgröße ist, die eine eichinvariante, physikalische Observable repräsentiert. In der Tat erzeugt $\mathbf{\Lambda}_3$ für den Ortsoperator Drehungen um die 3-Achse, da

$$\left[\mathbf{x}_j,\mathbf{\Lambda}_k\right] = \left[\mathbf{x}_j,\mathbf{L}_k\right] = \mathrm{i}\hbar\epsilon_{jkl}\mathbf{x}_l \tag{3.18.23}$$

ist. Dabei ist $\vec{\mathbf{L}} = \vec{\mathbf{x}} \times \vec{\mathbf{p}}$ der *kanonische* Bahndrehimpuls, der Drehungen beliebiger Vektoroperatoren erzeugt. Allerdings repräsentiert $\vec{\mathbf{L}}$ offenbar keine eichunabhängige, physikalische Observable.
Wir müssen nun noch das Eigenwertproblem für π_3 betrachten. Dazu suchen wir zunächst ein Vektorpotential für $\vec{B}$. Da $\vec{B} = B\vec{e}_3$ mit $B = \text{const}$ ist, können wir offenbar den Ansatz

$$\mathbf{A}_\mathrm{L}(\vec{x}) = A_2(x_1)\vec{e}_2 \tag{3.18.24}$$

wählen. Daraus folgt

$$\vec{B} = B\vec{e}_3 = \vec{\nabla} \times \vec{A}_\mathrm{L} = \partial_1 A_2 \vec{e}_3 \Rightarrow A_2(x_1) = Bx_1 \Rightarrow \vec{A}_\mathrm{L} = Bx_1\vec{e}_2. \tag{3.18.25}$$

Man nennt dieses Vektorpotential das Potential in der Landau-Eichung. Da $A_{\mathrm{L}3} = 0$ ist, ist in dieser Eichung $\pi_{\mathrm{L}3} = \mathbf{p}_3 = -\mathrm{i}\hbar\partial_3$. Jedes andere mögliche Vektorpotential unterscheidet sich von diesem Potential durch ein reines Gradientenfeld, d.h. mit einem beliebigen Skalarfeld $\chi(\vec{x})$ gilt

$$\vec{A}(\vec{x}) = \vec{A}_\mathrm{L}(\vec{x}) - \vec{\nabla}\chi(\vec{x}). \tag{3.18.26}$$

Damit ist also für ein allgemeines Vektorpotential für unser $\vec{B}$

$$\pi_3 = \mathbf{p}_3 - \partial_3\chi = -\mathrm{i}\hbar\partial_3 + q\partial_3\chi. \tag{3.18.27}$$

Die Eigenfunktionen zu diesem Operator sind nun offenbar durch

$$\pi_3 u_{n_a,n_b,p_3} = p_3 u_{n_a,n_b,p_3} \tag{3.18.28}$$

definiert, also durch die Differentialgleichung

$$-\mathrm{i}\hbar\partial_3 u_{n_a,n_b,p_3} = (p_3 - q\partial_3\chi)u_{n_a,n_b,p_3}. \tag{3.18.29}$$

Die Lösung ist offenbar

$$u_{n_a,n_b,p_3}(\vec{x}) = u^{(\perp)}_{n_a,n_b}(x_1,x_2)\exp\left(\frac{\mathrm{i}p_3x_3 - \mathrm{i}q\chi(\vec{x})}{\hbar}\right) = u^{(\mathrm{L})}_{n_a,n_b,p_3}(\vec{x})\exp\left[-\mathrm{i}\frac{q\chi(\vec{x})}{\hbar}\right], \tag{3.18.30}$$

Der verbliebene Faktor $u^{(\perp)}_{n_a,n_b}(x_1,x_2)$ ergibt sich bis auf einen konstanten Normierungsfaktor eindeutig aus dem simultanen Eigenwertproblem von $\mathbf{N}_a$ und $\mathbf{N}_b$. Die Eigenwerte p_3 dürfen offenbar wie beim gewöhnlichen Impulsoperator beliebig $p_3 \in \mathbb{R}$ sein. Jedes Landau-Level ist also unendlichfach entartet. Zu jedem $E_{n_a,p_3} = \hbar\omega(n_a + 1/2) + p_3^2/(2M)$ gibt es beliebig viele Eigenfunktionen, die sich durch $n_b \in \mathbb{N}_0$ unterscheiden. Dass die Bewegung in 3-Richtung zu kontinuierlichen Eigenwerten führt, ist klar, weil sich das Teilchen entlang der 3-Richtung kräftefrei bewegt. Es ist auch klar, dass p_3 der Eigenwert der 3-Komponente des mechanischen Impulses $\pi_3 = m\mathbf{v}_3$ ist, und dieser ist unabhängig von der Eichung.

Das simultane Eigenwertproblem für $\mathbf{N}_a$ und $\mathbf{N}_b$ bestimmt nun $u^{(\perp)}_{n_a,n_b}(x_1,x_2)$, und die Eigenfunktion für $n_a = n_b = 0$ ist bis auf einen Faktor eindeutig durch

$$\mathbf{a}u^{(\perp)}_{0,0} = 0, \quad \mathbf{b}u^{(\perp)}_{0,0} = 0 \tag{3.18.31}$$

bestimmt, und daraus ergeben sich alle übrigen Eigenfunktionen von $\mathbf{N}_a$ und $\mathbf{N}_b$ gemäß

$$u^{\perp}_{n_a,n_b} = \sqrt{\frac{1}{n_a!n_b!}}\left(\mathbf{a}^\dagger\right)^{n_a}\left(\mathbf{b}^\dagger\right)^{n_b} u^{(\perp)}_{0,0}. \tag{3.18.32}$$

Wir wollen noch zeigen, dass alle Eichungen zu äquivalenten Resultaten führen, d.h. die Erwartungswerte und Wahrscheinlichkeiten für die Messung von **eichunabhängigen** (also **physikalischen**) Observablen unabhängig von der Eichung sind. Offenbar können wir alle physikalischen Observablen des Teilchens, durch die von der Eichung unabhängigen Observablen $\vec{\mathbf{x}}$ und $\vec{\pi} = m\vec{\mathbf{v}}$ ausdrücken. Die $\vec{\pi}$ hängen allerdings vom Vektorpotential ab. Wir vergleichen nun die Situation in einer allgemeinen Eichung für das Vektorpotential mit der in der Landau-Eichung. Es gilt wegen (3.18.26)

$$\vec{\pi} = \vec{\pi}_{\mathrm{L}} + q\vec{\nabla}\chi. \tag{3.18.33}$$

Sei nun ψ_{L} eine Wellenfunktion, die sich auf die Beschreibung in der Landau-Eichung bezieht, dann legt der Zusammenhang zwischen den oben berechneten Eigenfunktionen (3.18.30) in einer allgemeinen Eichung und denen der Eigenfunktionen in der Landau-Eichung (entsprechend $\chi = 0$) nahe, dass dann derselbe Zustand in einer beliebigen Eichung durch

$$\psi(\vec{x}) = \psi_{\mathrm{L}}(\vec{x})\exp[-\mathrm{i}q\chi(\vec{x})/\hbar] \tag{3.18.34}$$

beschrieben wird. Da sich die Wellenfunktionen nur um einen von dieser Funktion unabhängigen Phasenfaktor unterscheiden, beschreiben diese in allen Eichungen denselben Zustand. Entscheidend ist nun, dass wegen

$$\vec{\pi} = -\mathrm{i}\hbar\vec{\nabla} - q\vec{A}_{\mathrm{L}} + q\vec{\nabla}\chi = \vec{\pi}_{\mathrm{L}} + q\vec{\nabla}\chi. \tag{3.18.35}$$

Nun ist für beliebige Wellenfunktionen (*Nachrechnen*)

$$\vec{\pi}\psi = \exp[-\mathrm{i}q\chi(\vec{x})/\hbar]\vec{\pi}_{\mathrm{L}}\exp[+\mathrm{i}q\chi(\vec{x})/\hbar]\psi, \tag{3.18.36}$$

und damit folgt für die Anwendung dieses Operators auf (3.18.34)

$$\vec{\pi}\psi(\vec{x}) = \exp[-\mathrm{i}q\chi(\vec{x})/\hbar]\vec{\pi}_{\mathrm{L}}\psi_{\mathrm{L}}(\vec{x}) \tag{3.18.37}$$

folgt. Da offenbar auch $\vec{\mathbf{x}}\psi(\vec{x}) = \exp(-\mathrm{i}q\chi/\hbar)\vec{x}\psi_{\mathrm{L}}(\vec{x})$ gilt, ergibt sich für beliebige *von der Eichung unabhängige* Observablen

$$\begin{aligned}\mathbf{O} = \mathbf{O}(\vec{\mathbf{x}}, \vec{\pi}) &= \exp[-\mathrm{i}q\chi(\vec{x})/\hbar]\mathbf{O}(\vec{\mathbf{x}}, \vec{\pi}_{\mathrm{L}})\exp[+\mathrm{i}q\chi(\vec{x})/\hbar] \\ &= \exp[-\mathrm{i}q\chi(\vec{x})/\hbar]\mathbf{O}_{\mathrm{L}}\exp[+\mathrm{i}q\chi(\vec{x})/\hbar]\end{aligned} \tag{3.18.38}$$

und damit

$$\langle O\rangle = \langle\psi\,|\,\mathbf{O}\psi\rangle = \langle\exp(-\mathrm{i}q\chi/\hbar)\psi_{\mathrm{L}}\,|\exp(-\mathrm{i}q\chi/\hbar)\mathbf{O}_{\mathrm{L}}\psi_{\mathrm{L}}\rangle = \langle\psi_{\mathrm{L}}\,|\,\mathbf{O}_{\mathrm{L}}\psi_{\mathrm{L}}\rangle\,, \tag{3.18.39}$$

d.h. die Erwartungswerte beliebiger beobachtbarer (und also eichunabhängiger) Größen stimmen mit den Erwartungswerten, die sich in der Landau-Eichung ergeben, überein, d.h. sie sind eichunabhängig.

Ebenso sind die Eigenwerte von $\mathbf{O}$ und $\mathbf{O}_{\mathrm{L}}$ dieselben. Bilden nämlich $u_o^{(\mathrm{L})}(\vec{x})$ einen vollständigen Satz orthonormaler Eigenzustände zu $\mathbf{O}_{\mathrm{L}}$ mit Eigenwerten o, so bilden wegen (3.18.38)

$$u_o(\vec{x}) = \exp[-\mathrm{i}q\chi(\vec{x})/\hbar]u_o^{(\mathrm{L})}(\vec{x}) \tag{3.18.40}$$

einen vollständigen Satz von Eigenvektoren von $\mathbf{O}$ mit denselben Eigenwerten o. Die quantenmechanische Beschreibung von Observablen liefert also mit den Operatoren einer beliebigen Eichung des Vektorpotentials dieselben Ergebnisse wie in der Landau-Eichung. Dabei ist zu beachten, dass nur solche selbstadjungierten Operatoren Observablen repräsentieren können, die sich unter Eichtransformationen **eichkovariant**, also gemäß (3.18.38) transformieren.

Da wir die Energieeigenzustände durch die drei *eichkovarianten* Observablen $\mathbf{N}_a$, $\mathbf{N}_b$ und π_3 bestimmt haben, stimmen auch diese Eigenfunktionen bis auf den Phasenfaktor mit denen der Landau-Eichung überein und besitzen auch die gleichen Eigenwerte, d.h. mit beliebigen Eigenlösungen $u^{(\mathrm{L})}_{n_a,n_b,p_3}$ in der Landau-Eichung sind die

$$u_{n_a,n_b,p_3}(\vec{x}) = \exp[-\mathrm{i}q\chi(\vec{x})/\hbar]u^{(\mathrm{L})}_{n_a,n_b,p_3}(\vec{x}) \tag{3.18.41}$$

die Eigenlösungen in einer allgemeinen Eichung zu den Eigenwerten n_a, n_b und p_3. Da sie sich von den Eigenlösungen der Landau-Eichung nur um einen Phasefaktor unterscheiden, beschreiben sie auch den gleichen Zustand im Hinblick auf alle eichinvarianten Größen und sind folglich physikalisch äquivalent.

3.19 Der Stern-Gerlach-Versuch

Der Stern-Gerlach-Versuch zur „Richtungsquantelung“ wurde 1922, also schon vor der Entwicklung der modernen Quantentheorie ausgeführt[7]. 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 $\ell \neq 0$ (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, müsste 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 von Strahlen vorhersagen, denn der Bahndrehimpuls gehört immer zu Darstellungen der Drehgruppe mit

[7] Eine äußerst vergnügliche Darstellung der Historie des Stern-Gerlach-Versuchs findet sich in [FH03].

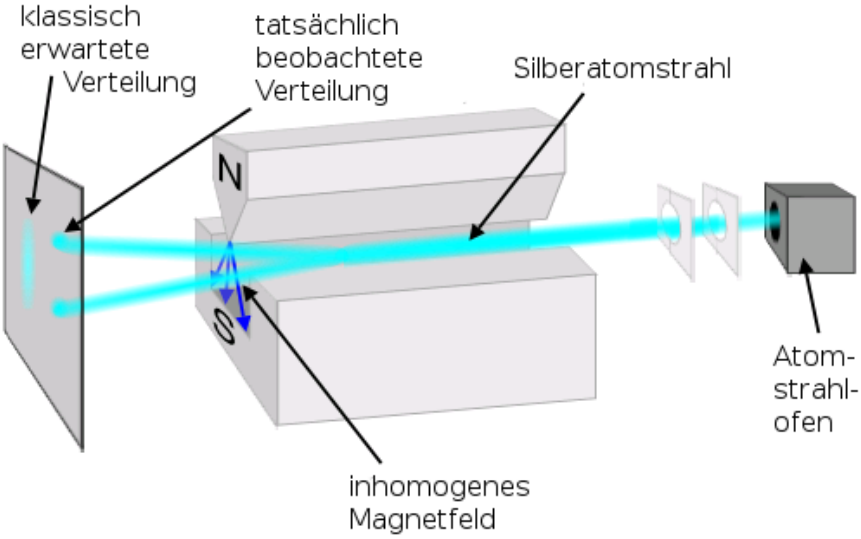


Abbildung 3.4: *Schematischer Aufbau des Stern-Gerlach-Versuchs (Wikipedia: Theresa Knott - aus dem Englischen ins Deutsche übersetzt, CC BY-SA 3.0,* `https://commons.wikimedia.org/w/index.php?curid=10344788`*).*

Drehimpulsquantenzahl $l \in \mathbb{N}_0$, und die z-Komponente des Bahndrehimpulses kann immer nur die $(2l+1)$ Werte $m_z = \{0, \pm\hbar, \ldots, \pm l\,\hbar\}$ 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 ($m_s = \pm 1/2$) spaltet sich also der Strahl in der Tat in zwei Teilstrahlen auf, von denen einer aus Teilchen mit $m_s = \hbar/2$ und einer aus solchen mit $m_s = -\hbar/2$ 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. 3.4 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:

$$\vec{B} = (B_0 + \beta z)\vec{e}_z - \beta y \vec{e}_y. \tag{3.19.1}$$

Dieses Feld erfüllt offenbar die magnetostatischen Maxwell-Gleichungen

$$\vec{\nabla} \cdot \vec{B} = 0, \quad \vec{\nabla} \times \vec{B} = 0 \tag{3.19.2}$$

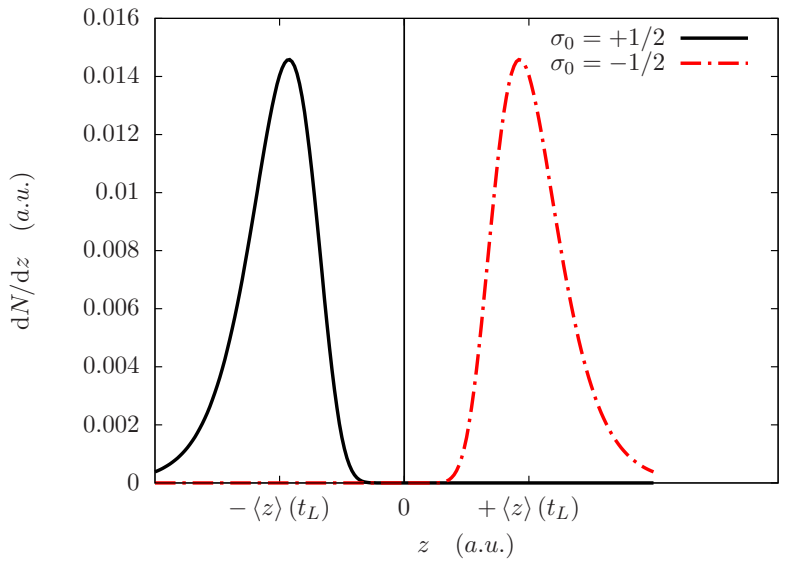


Abbildung 3.5: *Wahrscheinlichkeitsverteilung für die z-Position eines Teilchens nach Durchlaufen eines inhomogenen Magnetfeldes bei $x = L$. Anfangs lag eine bei $\vec{x} = 0$ gepeakte Gaußverteilung vor, wobei die Teilchen mit gleicher Wahrscheinlichkeit eine Spinkomponente $m_s = +1/2$ oder $m_s = -1/2$ besitzen.*

für ein statischen Magnetfeld in einem quellenfreien Raumbereich. Der Hamiltonoperator (3.17.2) vereinfacht sich im gegebenen Falle wegen $q = 0$ zu[8]

$$\mathbf{H} = \frac{\vec{\mathbf{p}}^2}{2M} + \mu_B g_s \vec{\mathbf{S}} \cdot \vec{B}(\vec{\mathbf{x}}). \tag{3.19.3}$$

Setzen wir (3.19.1) 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 $|B_0| \gg \beta |\langle y\rangle|$ ist, d.h. der Strahl sei hinreichend in xy-Richtung um $x = 0$, $y = 0$ fokussiert. Dann können wir zunächst den einfacheren Hamiltonoperator

$$\mathbf{H}' = \frac{\vec{\mathbf{p}}^2}{2M} + \mu_B g_s (B_0 + \beta \mathbf{z}) \mathbf{S}_z \tag{3.19.4}$$

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 $m_s = \hbar/2$ nach unten, für $m_s = -\hbar/2$ 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 $\hbar m_s = \pm\hbar/2$ 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.

[8] Das Vorzeichen des Spinterms kehrt sich wegen der negativen Ladung des Elektrons um.

In diesem Zusammenhang nennt man die Ortskomponente z auch eine **Zeigervariable**, denn sie wird bei dem betrachteten Versuchsaufbau zur Messung der eigentlich interessierenden Observable, nämlich der Spin-3-Komponente, im Sinne des Zeigers eines Messgeräts verwendet.

Man kann die zeitabhängige Schrödinger-Gleichung mit dem Hamilton-Operator (3.19.4) 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. 3.5 ist das Ergebnis für $|\psi|^2$ für einen Anfangszustand

$$\psi_\sigma(t=0,\vec{x}) = \frac{c_\sigma}{(2\pi\Delta^2)^{3/4}} \exp\left(-\frac{\vec{x}^2}{4\Delta^2} + \mathrm{i}p_0 x\right), \quad |c_{1/2}|^2 + |c_{-1/2}|^2 = 1. \tag{3.19.5}$$

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.

3.20 EPR-Paradoxon und Bell-Ungleichung

In diesem Abschnitt wollen wir noch sehr kurz auf die berühmte Debatte über die **Vollständigkeit** der Quantentheorie als Beschreibung der Natur eingehen. Berühmt geworden ist dabei vor allem die Debatte zwischen **Niels Bohr** und **Albert Einstein**. Niels Bohr ist dabei nicht nur als Mitbegründer der sog. „alten Quantentheorie“ aufgrund seines Atommodells von 1913 (vgl. Abschnitt 2.5) wie auch Einstein mit seiner Lichtquantenhypothese von 1905 (vgl. Abschnitt 2.3) berühmt sondern auch als einer der wichtigsten Begründer für die **Kopenhagener Deutung** der Quantentheorie mit der **Komplementarität**. Dabei bedeutet Komplementarität, dass z.B. der **Welle-Teilchen-Dualismus** durch die **Wahrscheinlichkeitsdeutung** des quantenmechanischen Zustands durch Born zu einer konsistenten Naturbeschreibung aufgelöst werden kann: Vor der Messung liegen demnach die Werte der möglichen Messwerte i.a. nicht fest, und erst bei der Messung einer gegebenen Observablen resultieren festgelegte Werte für diese Observablen, wobei die möglichen Messwerte die Eigenwerte der die gemessene Observable repräsentierenden selbstadjungierten Operatoren sind. Selbst die vollständige Kenntnis des Zustandes des zu vermessenden Systems aufgrund der Präparation des Systems in einem reinen Zustand (z.B. durch Festlegung der Werte eines vollständigen Satzes kompatibler Observabler, so dass dann die Wellenfunktion ψ dem bis auf einen Phasenfaktor eindeutigen simultanen Eigenzustand dieser Observablen entspricht) impliziert i.a. nur Wahrscheinlichkeiten für das Resultat bei einer Messung.

Mit genau dieser Deutung waren nun einige Physiker nicht einverstanden; darunter waren neben Einstein auch Schrödinger und de Broglie. Ihrer Meinung nach sollte eine **vollständige Naturbeschreibung** implizieren, dass alle Observablen eines Systems stets wohldeterminierte Werte haben sollten, und Wahrscheinlichkeiten ergeben sich dann nur aus einer unvollständigen Kenntnis dieser Werte wie in der **klassischen Statistischen Gastheorie**, wo man die genauen Orte und Impulse der Gasmoleküle aufgrund der großen Zahl (Größenordnung der Avogadro-Zahl von ca. $6 \cdot 10^{23}$ Teilchen für ein Mol eines Stoffes, also z.B. 12 g Kohlenstoff) nicht kennt und stattdessen das Gas statistisch durch eine Phasenraumverteilung beschreibt (u.a. vor allem durch **Ludwig Boltzmann** und **James Clerk Maxwell** begründet).

Die berühmteste Arbeit zu diesem Themenkomplex ist die von A. Einstein, Boris Podolsky und Nathan Rosen (EPR) [EPR35] mit dem Titel „Can quantum-mechanical description of physical reality be considered complete?“ und der gleichfalls berühmten Antwort von N. Bohr [Boh35] mit demselben Titel. Beide Arbeiten sind allerdings aufgrund ihrer eher philosophischen Ausdrucksweise nicht einfach zu lesen. Einsteins eigentliche Bedenken gegen die Vollständigkeit der Quantentheorie sind besser verständlich in [Ein48] zusammengefasst. Das EPR-Paper ist deshalb von so grundlegender Bedeutung, weil die aufgebrachten erkenntnistheoretischen Probleme zu einem immens besseren Verständnis der diesbezüglichen Implikationen der Quantentheorie geführt haben. Letztlich hat die eingehende Beschäftigung mit diesen Grundlagenfragen zu modernen Anwendungen wie Quantenkryptographie, Quanteninformatik und der Idee für die Entwicklung von Quantencomputern geführt. Der eigentliche Durchbruch waren allerdings die Arbeiten von John Bell (1928-1990),

der die philosophischen Argumente zu naturwissenschaftlich überprüfbare Aussagen über sogenannte **lokale realistische Theorien**, die den Voraussagen der Quantentheorie widersprechen [Bel64, Bel66, Bel87], konkretisiert hat. In diesem Abschnitt beschäftigen wir uns mit diesem erstaunlichsten Aspekt der Quantentheorie anhand eines sehr einfachen Beispiels [ST93].

Zunächst müssen wir uns dazu mit der Kernthese des EPR-Papers beschäftigen, die definiert, was diese Autoren unter einer „vollständigen physikalischen Theorie" verstehen. Die berühmten Schlüsselsätze dazu lauten:

„If, without in any way disturbing a system, we can predict with certainty (i.e., with probability unity) the value of a physical quantity, then there exists an element of physical reality corresponding to this quantity."[9]

und weiter bzgl. der „Realität" *inkmpatibler* Observabler (im EPR-Paper anhand von Ort und Impuls diskutiert)

„From this it follows that either (1) the quantum mechanical description of reality by the wave function is not complete or (2) when the operators corresponding to two physical quantities do not commute the two quantities cannot have simultaneous reality."[10]

Einstein, Podolsky und Rosen (EPR) kamen bekanntermaßen zu dem Schluss, dass (1) korrekt sein sollte, während wir in dieser Vorlesung der Auffassung (2) gefolgt sind. Das Hauptproblem für einige Jahrzehnte nach Erscheinen des Papers war, dass man keine Idee hatte, wie man zwischen den beiden Möglichkeiten durch ein **physikalisches Experiment** entscheiden kann. Es war sogar für einige Jahre für Physiker karriereschädigend, sich überhaupt mit diesen Grundlagenfragen zu beschäftigen. Aufgrund des großen Erfolgs der Quantentheorie in ihrer Interpretation durch die sog. **Kopenhagener Deutung**, der wir auch in dieser Vorlesung weitgehend gefolgt sind, galt das „Unvollständigkeitsargument" von EPR als weitgehend irrelevant und die aufgeworfenen Fragen für nicht durch physikalische Argumente bzw. Experimente entscheidbar. Der Durchbruch gelang erst durch Bells oben zitierte bahnbrechende Arbeiten, die die eher philosophisch-erkenntnistheoretischen Betrachtungen des EPR-Papers in eine physikalisch überprüfbare quantitative Hypothese überführten, die der Quantentheorie widerspricht [Kai11].

Dazu definierte Bell zunächst, was unter „Realismus" im EPR-Sinne zu verstehen sei. Auf Ideen Einsteins zurückgreifend postulierte Bell, dass die probabilistische Beschreibung der Phänomene, die in der Quantentheorie unvermeidlich ist, „in Wirklichkeit" darauf zurückzuführen sei, dass es sog. „**versteckte Variablen**" gibt, deren Werte wir nicht kennen oder die gar der Messung gänzlich unzugänglich sind, und wir daher „nur" statistische Eigenschaften dieser Variablen kennen. Eine **realistische Theorie** ist nun eine solche Beschreibung der Physik, in der *alle* Observablen und versteckten Variablen „in Wirklichkeit" stets bestimmte Werte besitzen. Die aufgrund unserer Unkenntnis der versteckten Variablen nötigen Wahrscheinlichkeiten folgen dann den üblichen Eigenschaften von Wahrscheinlichkeiten, die dieser Annahme der **Determiniertheit aller Observablen** entsprechen.

In der Quantentheorie verhält es sich freilich anders: Wie wir stets betont haben, ist eine *vollständige* Kenntnis durch die **Präparation** des Systems in einem Zustand gegeben, in dem lediglich die Observablen in einem **vollständiger Satz voneinander unabhängiger kompatibler Observabler** simultan bestimmte Werte annehmen. Dann wird der Zustand durch eine bis auf einen physikalisch irrelevanten Phasenfaktor **eindeutige** normierte simultane Eigenfunktion der den kompabiblen Observablen entsprechenden selbstadjungierten Operatoren beschrieben. Nur diesen Observablen kommt dann im Sinne von EPR „physikalische Realität" zu. Alle Observablen, die zumindest zu einer dieser zur Präparation gewählten kompatiblen Observablen inkompatibel ist, hat i.a. keinen bestimmten Wert, und der Zustand sagt „nur" die Wahrscheinlichkeiten für das Auftreten der möglichen Messwerte vorher, wenn man diese Observable misst. Davon abgesehen ändert aufgrund der notwendigen Wechselwirkung des vermessenen Systems mit dem Messapparat eine solche Mes-

[9] Wenn wir mit Sicherheit (d.h. mit Wahrscheinlichkeit 1) den Wert einer physikalischen Größe [Observablen] vorhersagen können, ohne in irgendeiner Weise das System zu stören, dann existiert ein dieser Größe entsprechendes Element physikalischer Realität.

[10] Daraus [der Existenz inkompatibler Observabler] folgt, dass entweder (1) die quantenmechanische Beschreibung der Realität durch die Wellenfunktion unvollständig ist oder (2) wenn die beiden physikalischen Größen [Observablen] repräsentierenden [selbstadjungierten] Operatoren nicht kommutieren, diesen beiden Observablen nicht zugleich Realität zukommen kann.

sung i.a. auch die Festlegung der ursprünglich determinierten Observablen, d.h. i.a. ist eine **störungsfreie Messung** der inkompatiblen Observable zufolge der Quantentheorie nicht möglich.

Wir wollen nun die Idee Bells an einem einfachen auf David Bohm (1917-1992) und Yakir Aharonov (geb. 1932) [BA58, Boh89] zurückgehendes Beispiel mit Spinkomponenten für zwei Spin-1/2-Teilchen diskutieren. Dazu stellen wir uns ein ruhendes instabiles Teilchen mit Spin 0 („skalares Teilchen") vor, das spontan in zwei Teilchen mit spin 1/2 zerfällt. Ein konkretes Beispiel ist der Zerfallsprozess eines η-Mesons in ein Myon und ein Antimyon ($\eta \to \mu^- + \mu^+$). Aufgrund der Impulserhaltung sind die Impulse der Myonen stets entgegengesetzt gleich und folglich wegen der Drehimpulserhaltung der Gesamtspin von μ^- und μ^+ ebenfalls 0. Im Folgenden interessieren uns nun nur die Messungen von Spinkomponenten der beiden Myonen, d.h. wir betrachten als Observablen die Messung von Spinkomponenten $\mathbf{S}_1(\vec{a}) = \vec{a} \cdot \vec{\mathbf{S}}_1$ und $\mathbf{S}_2(\vec{b}) = \vec{b} \cdot \vec{\mathbf{S}}_2$. Dabei sind $\vec{a}$ und $\vec{b}$ beliebige Einheitsvektoren und $\vec{\mathbf{S}}_k$ mit $k \in \{1,2\}$ die Spins des Myons ($k = 1$) bzw. Anti-Myons ($k = 2$). Die Spinzustände des Paars werden dabei in einem Hilbertraum beschrieben, die durch die vier simultanen Eigenzustände von $\mathbf{S}_j(\vec{e}_3) = \mathbf{S}_{j3}$, also $|\sigma_1, \sigma_2\rangle$, beschrieben werden. Da für Myon und Anti-Myon $s_1 = s_2 = 1/2$ ist, gilt bekanntermaßen für die möglichen $\mathbf{S}_{j3}$-Eigenwerte $\sigma_{j3} \in \{\hbar/2, -\hbar/2\}$. Die Spinkomponenten erfüllen dabei die Vertauschungsrelationen

$$\left[\mathbf{S}_{k_1}(\vec{e}_a), \mathbf{S}_{k_2}(\vec{e}_b)\right] = \mathrm{i}\hbar\epsilon_{abc}\mathbf{S}_{k_1}(\vec{e}_c)\delta_{k_1k_2}. \tag{3.20.1}$$

Bzgl. der Forderung von EPR, dass die Messungen an den Spins der Teilchen das System nicht stören sollen, bemerken wir, dass wir die Messungen an Myon und Anti-Myon an weit entfernten Orten im Abstand R vornehmen können. Dabei stellen wir sicher, dass die Spinmessungen zu Zeiten t_1 bzw. t_2 erfolgen, so dass $c|t_1 - t_2| < R$, wobei c die Lichtgeschwindigkeit ist, erfüllt ist. Gemäß der Einsteinschen Relativitätstheorie können wir dann davon ausgehen, dass die eine Messung die andere nicht kausal beeinflussen kann, da sich kein Signal mit einer größeren als der Lichtgeschwindigkeit ausbreiten kann.

Nun analysieren wir eine Messung der Spin-Komponenten in drei Richtungen, die durch die Einheitsvetkoren $\vec{a}$, $\vec{b}$ und $\vec{c}$ vorgegeben werden bzw. der umskalierten Observablen $\boldsymbol{\sigma}_k(\vec{n}) = 2\mathbf{S}_k(\vec{n})/\hbar$ mit $k \in \{1,2\}$ und $\vec{n} \in \{\vec{a}, \vec{b}, \vec{c}\}$, die stets die Werte $\alpha, \beta, \gamma \in \{1, -1\}$ annehmen.

Wir nehmen dabei zunächst an, es gälte nicht die Quantentheorie sondern irgendeine **lokal realistische Theorie** (LRT), d.h. wir nehmen an, alle Spinkomponenten nehmen stets wohldefinierte Werte an, und die zufälligen Messresultate kommen nur aufgrund unserer Unkenntnis irgendwelcher versteckter Variabler zustande. Aufgrund der Drehimpulserhaltung beim Zerfall des η-Mesons wissen wir, dass dabei stets $\vec{s}_1(\vec{n}) = -\vec{s}_2(\vec{n})$ gelten muss. Es gibt also im Rahmen einer LRT die 8 Möglichkeiten für die Werte der Spinkomponenten:

Wahrscheinlichkeit	Teilchen 1	Teilchen 2
	$\alpha_1, \beta_1, \gamma_1$	$\alpha_2, \beta_2, \gamma_2$
p(1,1,1)	$+++$	$---$
p(1,1,-1)	$++-$	$--+$
p(1,-1,1)	$+-+$	$-+-$
p(-1,1,1)	$-++$	$+--$
p(1,-1,-1)	$+--$	$-++$
p(-1,1,-1)	$-+-$	$+-+$
p(-1,-1,1)	$--+$	$++-$
p(-1,-1,-1)	$---$	$+++$

Es ist klar, dass wir bei jeder Spinmessung an einem Teilchen stets nur die Spinkomponente in einer Richtung messen können. Verwenden wir nämlich einen Stern-Gerlach-Apparat, wie im vorigen Kapitel beschrieben, bestimmt die Richtung des $\vec{B}$-Feldes, welche Spinkomponente gemessen wird. Wir können also stets die Spinkomponente von Teilchen 1 in einer der drei Richtungen $\vec{a}$, $\vec{b}$ und $\vec{c}$ messen und ebenso für Teilchen 2. In

einer lokal realistischen Theorie ergibt sich dann z.B. die Wahrscheinlichkeit, dass bei einer Messung von $\boldsymbol{\sigma}_1(\vec{a})$ und $\boldsymbol{\sigma}_2(\vec{b})$ die Werte $\alpha_1 = +1$ und $\beta_2 = -1$ gefunden werden, zu

$$p(\vec{a},1;\vec{b},-1) = p(1,1,1) + p(1,1,-1) \tag{3.20.2}$$

usw.
Bell betrachtet nun die Korrelationsfunktion

$$C(\vec{n}_1,\vec{n}_2) = \langle \boldsymbol{\sigma}_1(\vec{n}_1)\boldsymbol{\sigma}_2(\vec{n}_2)\rangle\,. \tag{3.20.3}$$

Die für die Entscheidung zwischen einer beliebigen LRT und der Quantenmechanik bedeutsame Eigenschaft dieser Korrelationsfunktion ist die **Bellsche Ungleichung**:

$$|C(\vec{a},\vec{b}) - C(\vec{a},\vec{c})| \overset{\text{LRT}}{\leq} 1 + C(\vec{b},\vec{c}). \tag{3.20.4}$$

Zum **Beweis** bemerken wir, dass die Korrelationsfunktionen im Rahmen der LRT aufgrund von $\vec{s}_1 = -\vec{s}_2$ durch

$$\begin{aligned} C(\vec{a},\vec{b}) &= \left\langle \boldsymbol{\sigma}_1(\vec{a})\boldsymbol{\sigma}_2(\vec{b})\right\rangle = -\left\langle \boldsymbol{\sigma}_1(\vec{a})\boldsymbol{\sigma}_1(\vec{b})\right\rangle = -\sum_{\alpha,\beta,\gamma} p(\alpha,\beta,\gamma)\alpha\beta \\ C(\vec{a},\vec{c}) &= \langle \boldsymbol{\sigma}_1(\vec{a})\boldsymbol{\sigma}_2(\vec{c})\rangle = -\langle \boldsymbol{\sigma}_1(\vec{a})\boldsymbol{\sigma}_1(\vec{c})\rangle = -\sum_{\alpha,\beta,\gamma} p(\alpha,\beta,\gamma)\alpha\gamma \end{aligned} \tag{3.20.5}$$

gegeben ist. Daraus folgt unter Verwendung von $\alpha^2 = \beta^2 = \gamma^2 = (\pm 1)^2 = +1$

$$\begin{aligned} C(\vec{a},\vec{b}) - C(\vec{a},\vec{c}) &= -\sum_{\alpha,\beta,\gamma} p(\alpha,\beta,\gamma)(\alpha\beta - \alpha\gamma) \\ &= -\sum_{\alpha,\beta,\gamma} p(\alpha,\beta,\gamma)(\alpha\beta - \alpha\beta^2\gamma) \\ &= -\sum_{\alpha,\beta,\gamma} p(\alpha,\beta,\gamma)\alpha\beta(1-\beta\gamma). \end{aligned} \tag{3.20.6}$$

Mit der Dreiecksungleichung, sowie wegen $|\alpha\beta| = 1$ und $1-\beta\gamma \in \{0,2\}$, d.h. $1-\beta\gamma \geq 0$ folgt daraus

$$\left|C(\vec{a},\vec{b}) - C(\vec{a},\vec{c})\right| \leq \sum_{\alpha,\beta,\gamma} p(\alpha,\beta,\gamma)|\alpha\beta||1-\beta\gamma| = \sum_{\alpha,\beta,\gamma} p(\alpha,\beta,\gamma)(1-\beta\gamma) = 1 + C(\vec{b},\vec{c}), \tag{3.20.7}$$

und das ist die zu beweisende Bell-Ungleichung (3.20.4).
Nun wollen wir die Vorhersage der **Quantentheorie** für die Korrelationsfunktion (3.20.3) ausrechnen. Dazu müssen wir zunächst den Eigenzustand mit $S = 0$ für den Gesamtspin des Zweiteilchensystems finden. Es handelt sich um ein einfaches Beispiel zur **Addition von Drehimpulsen**. Es ist klar, dass die Spinzustände des Zweiteilchensystems durch die gemeinsamen Eigenvektoren $\left|u_{\sigma_1,\sigma_2}\right\rangle$ der kompatiblen Observablen $\vec{\mathbf{S}}_1^2$, $\vec{\mathbf{S}}_2$, $\mathbf{S}_{13}$ und $\mathbf{S}_{23}$ aufgespannt werden. Da $s_1 = s_2 = 1/2$ ist, ist der entsprechende Vektorraum vierdimensional. Es ist weiter klar, dass es auch ein vollständiges gemeinsames Orthonormalsystem von Eigenvektoren von $\vec{\mathbf{S}}^2$ und $\mathbf{S}_3$ für den Gesamtspin $\vec{\mathbf{S}} = \vec{\mathbf{S}}_1 + \vec{\mathbf{S}}_2$ geben muss. Als die notwendigen weiteren kompatiblen Observablen können wieder $\vec{\mathbf{S}}_1^2$ und $\vec{\mathbf{S}}_2^2$ dienen. Wir bezeichnen die entsprechenden Eigenvektoren als $\left|v_{S,\Sigma_3}\right\rangle$. Nun gilt

$$\mathbf{S}_3\left|u_{\sigma_1,\sigma_2}\right\rangle = (\mathbf{S}_{13} + \mathbf{S}_{23})\left|u_{\sigma_1,\sigma_2}\right\rangle = (\sigma_1 + \sigma_2)\hbar\left|u_{\sigma_1+\sigma_2}\right\rangle. \tag{3.20.8}$$

Die Vektoren $\left|u_{\sigma_1,\sigma_2}\right\rangle$ sind also auch Eigenvektoren von $\mathbf{S}_3$. Daher können die Eigenwerte wegen $\sigma_1,\sigma_2 \in \{1/2,-1/2\}$ nur $\Sigma \in \{1,0,-1\}$ sein. Der größte Eigenwert von $\mathbf{S}_3$ ist also $1\hbar$, d.h. es müssen Eigenvektoren von $\vec{\mathbf{S}}^2$ zum Eigenwert $S(S+1)\hbar^2$ mit $S \geq 1$ vorkommen. Es kann nun aber nicht $S > 1$ vorkommen, weil man dann mit Hilfe des Aufsteigeoperators $\mathbf{S}_+$ entsprechend auch Eigenvektoren zu Eigenwerten $\Sigma_3 > \hbar$ erhalten würde. Wir müssen also Eigenvektoren mit $S = 1$ haben, und der eindeutige Eigenvektor zu $\Sigma = 1$ muss einer dieser Eigenvektoren sein. Wir können also

$$\left|v_{1,1}\right\rangle = \left|u_{1/2,1/2}\right\rangle \tag{3.20.9}$$

setzen. Die übrigen beiden zu $S = 1$ gehörigen Basisvektoren ergeben sich dann mit dem Absteigeoperator $\mathbf{S}_-$ mit Hilfe der allgemeinen Formel (3.10.25)

$$\begin{aligned}\left|v_{1,0}\right\rangle &= (\mathbf{S}_{1-} + \mathbf{S}_{2-})\left|v_{1,1}\right\rangle \\ &= (\mathbf{S}_{1-} + \mathbf{S}_{2-})\left|u_{1/2,1/2}\right\rangle \\ &= \sqrt{\frac{1}{2}}\left(\left|u_{1/2,-1/2}\right\rangle + \left|u_{-1/2,1/2}\right\rangle\right).\end{aligned} \tag{3.20.10}$$

Nochmalige Anwendung des Absteigeoperators auf diesen Vektor liefert schließlich

$$\left|v_{1,-1}\right\rangle = \left|u_{-1/2,-1/2}\right\rangle, \tag{3.20.11}$$

was zu erwarten war, da $-\hbar$ der kleinste vorkommende Eigenwert von $\mathbf{S}_3$ ist und zu diesem Eigenwert der Eigenvektor bis auf einen unerheblichen Phasenfaktor folglich nur $\left|u_{-1/2,-1/2}\right\rangle$ sein kann. Es fehlt nun nur noch ein Eigenvektor von $\mathbf{S}_3$ zum Eigenwert 0, denn zu diesem Eigenwert gibt es die beiden orthogonalen Eigenvektoren $|\sigma_1 = 1/2, \sigma_2 = -1/2\rangle$ und $|\sigma_1 = -1/2, \sigma_2 = 1/2\rangle$. Oben haben wir die entsprechende Linearkombination zum simultanen Eigenvektor mit $S = 1$ und $\Sigma = 0$ gefunden. Der verbliebene weitere Eigenvektor muss folglich zu $S = 0$ gehören und daher orthogonal zu $\left|v_{1,0}\right\rangle$ sein. Damit ist (bis auf einen unerheblichen Phasenfaktor)

$$\left|v_{0,0}\right\rangle = \sqrt{\frac{1}{2}}\left(\left|u_{1/2,-1/2}\right\rangle - \left|u_{-1/2,1/2}\right\rangle\right), \tag{3.20.12}$$

und das muss demnach der entsprechende Spinzustand für das Myon-Antimyon-Paar des Zerfalls $\eta \to \mu^+ + \mu^-$ sein.

Berechnen wir also die Korrelationsfunktion (3.20.3). Dazu benötigen wir die Matrixelemente für $\boldsymbol{\sigma}_k(\vec{a})$ bzgl. der Basis $|\sigma_1, \sigma_2\rangle$. Diese sind aber durch die Pauli-Matrizen (3.14.16) gegeben. In dieser Basis werden also die besagten Operatoren durch

$$\hat{\boldsymbol{\sigma}}_k(\vec{a}) = \vec{a} \cdot \hat{\vec{\sigma}} = \begin{pmatrix} a_3 & a_1 - \mathrm{i}a_2 \\ a_1 + \mathrm{i}a_2 & -a_3 \end{pmatrix} \tag{3.20.13}$$

dargestellt. Damit erhalten wir für die Korrelationsfunktion

$$\begin{aligned}C(\vec{a},\vec{b}) &= \frac{1}{2}\left\langle u_{1/2,-1/2} - u_{-1/2,1/2}\left|\boldsymbol{\sigma}_1(\vec{a})\boldsymbol{\sigma}_2(\vec{b})\right|u_{1/2,-1/2} - u_{-1/2,1/2}\right\rangle \\ &= \frac{1}{2}[-a_3b_3 - (a_1 - \mathrm{i}a_2)(b_1 + \mathrm{i}b_2) - (a_1 + \mathrm{i}a_2)(b_1 - \mathrm{i}b_2) - a_3b_3] \\ &= -\vec{a} \cdot \vec{b}.\end{aligned} \tag{3.20.14}$$

Wir erhalten nun einen Widerspruch zur Bell-Ungleichung (3.20.4), also einen Widerspruch zu jeder lokalen realistischen Theorie, wenn wir die drei Einheitsvektoren

$$\vec{a} = (1,0,0), \quad \vec{b} = (0,1,0) \quad \text{und} \quad \vec{c} = \frac{1}{\sqrt{2}}(1,1,0) \tag{3.20.15}$$

wählen, denn dann ist

$$C(\vec{a},\vec{b})=0, \quad C(\vec{a},\vec{c})=C(\vec{b},\vec{c})=-\frac{1}{\sqrt{2}}\simeq -0{,}707 \tag{3.20.16}$$

und

$$|C(\vec{a},\vec{b})-C(\vec{a},\vec{c})| \overset{\text{QM}}{=} \frac{1}{\sqrt{2}}\simeq 0{,}707 \quad \text{und} \quad 1+C(\vec{b},\vec{c})=1-\frac{1}{\sqrt{2}}\simeq 0{,}293, \tag{3.20.17}$$

so dass die Bell-Ungleichung (3.20.4) in der Tat deutlich verletzt ist.

Damit ergibt sich, zumindest im Prinzip, die Möglichkeit, experimentell zwischen der Gültigkeit beliebiger lokal-realistischer Theorien, die das EPR-Kriterium an eine „vollständige physikalische Theorie" erfüllen, und den Vorhersagen der Quantentheorie, die den Vorhersagen der lokal-realistischen Theorien (wie z.B. der eben betrachtete Bell-Ungleichung) widersprechen, zu entscheiden. Die ersten Experimente dazu wurden erfolgreich von John Clauser (geb. 1942) und Stewart Friedmann (1944-2012) ausgeführt. Dabei wurden allerdings Polarisationszustände von Photonen anstelle von Spinzuständen von Teilchen verwendet, was aber keinen prinzipiellen Unterschied ergibt. Der Grund ist, dass es für Photonen wesentlich einfacher ist, die „**verschränkten Zustände**" wie den $S=0$-Zustand zweier $s=1/2$-Spins (3.20.12) stabil zu halten, um die Polarisation der Einzelphotonen, die ein solch polarisationsverschränktes Photonenpaar bilden, zu messen, und zwar auch in großen Entfernungen, so dass die Messprozesse sich aufgrund der Endlichkeit kausaler Signalausbreitung aufgrund der Relativitätstheorie (s. 4) nicht gegenseitig beeinflussen können. Bereits diese ersten Experimente konnten klar die quantentheoretische Vorhersage und entsprechend die Verletzung der Bell-Ungleichung bestätigen und damit die Möglichkeit lokaler realistischer Theorien á la EPR widerlegen.

Dank der Entdeckung der sog. „spontaneous parametric down conversion"[11] lassen sich heute polarisationsverschränkte Photonenpaare in großer Intensität leicht herstellen und experimentell nutzen. Dabei wird mit intensivem Laser-Licht eine bestimmte Art doppelbrechender Kristalle (z.B. sog. β-Bariumborat) angeregt, jeweils ein Photon aus dem Laser-Licht zu absorbieren und zwei Photonen niedrigerer Frequenz zu emittieren. Diese Photonen erfüllen Energie- und Impuls-Erhaltung. In dem Fall, dass beide Photonen die gleiche Energie besitzen (entsprechend der halben Energie des Photons aus dem Laser-Licht) sind diese Photonen im Prinzip ununterscheidbar. Aufgrund der doppelbrechenden Eigenschaften des Kristalls laufen zueinander senkrecht polarisierte Photonen in verschiedene Richtungen, die entsprechend zwei Kegel bilden. An zwei Stellen schneiden sich diese beiden Kegel, und man erhält zwei tatsächlich ununterscheidbare Photonen mit zueinander senkrechter Polarisation, was in verschränkten Polarisationszuständen resultiert (Type-II-PDC), die genau analog zu dem oben besprochenen $S=0$-Zustand zweier $s=1/2$-Spins sind. Mit solchen Photonenpaaren wurde das Experiment wiederholt und ebenfalls die Verletzung der Bellschen Ungleichung mit größter Präzision bestätigt. Ein gut lesbarer Übersichtsartikel zu diesen **EPR-Bell-Tests** ist [Gen05].

Wir gehen noch kurz auf die recht faszinierenden Eigenschaften verschränkter Zustände ein. Dazu betrachten wir zunächst den *nicht verschränkten* Zustand $\left|u_{1/2,1/2}\right\rangle$, also den Zustand zu $S=1$ und $\sigma_1=\sigma_2=1/2$ (und folglich $\Sigma=1$). Hier besitzen die Einzelspinkomponenten $\mathbf{S}_{13}$ und $\mathbf{S}_{23}$ die determinierten Werte $\hbar/2$, denn es handelt sich ja um einen simultanen Eigenzustand von $\mathbf{S}_{13}$ und $\mathbf{S}_{23}$.

Im Gegensatz dazu ist der $S=0$-Zustand

$$\left|v_{0,0}\right\rangle=\frac{1}{\sqrt{2}}\left(\left|u_{1/2,-1/2}\right\rangle-\left|u_{-1/2,1/2}\right\rangle\right) \tag{3.20.18}$$

zwar ein Eigenzustand von $\mathbf{S}_3$ zum Eigenwert $\Sigma=0$, aber kein Eigenzustand der Spinkomponenten $\mathbf{S}_{13}$ oder $\mathbf{S}_{23}$, d.h. diese nehmen unbestimmte Werte an. Z.B. ist die Wahrscheinlichkeit, bei einer Messung von $\mathbf{S}_{13}$ (also der Spinkomponente in 3-Richtung des 1. Teilchens), jeweils einen der möglichen Werte $\pm\hbar/2$ zu messen, durch

$$P_1(\hbar/2)=P_1(-\hbar/2)=\sum_{\sigma_2}\left|\left\langle u_{1/2,\sigma_2}\,\middle|\,v_{0,0}\right\rangle\right|^2=\frac{1}{2} \tag{3.20.19}$$

[11] auf Deutsch manchmal als „parametrische Fluoreszenz" bezeichnet

gegeben. Das gleiche Resultat erhält man auch für die Messung von $\mathbf{S}_{23}$ (also der Spinkomponente in 3-Richtung des 2. Teilchens). Beide Teilchen sind also **unpolarisiert**, d.h. man findet bei der Messung der Einteilchenspinkomponenten völlig zufällig mit jeweils 50% Chance entweder das Resultat $+\hbar/2$ oder $-\hbar/2$. Gemäß der Quantentheorie und aufgrund der Widerlegung der Möglichkeit lokal realistischer Theorien handelt es sich hierbei nicht um einen Zufall, der allein durch unsere Unkenntnis der gemessenen Spinkomponenten herrührt, sondern die Spinkomponenten nehmen vor der Messung tatsächlich keinen bestimmten Wert an. Dabei ist das Gesamtsystem von zwei Teilchen in einem vollständigst möglichen „reinen Zustand", der durch den Zustandsvektor (3.20.18) repräsentiert wird, präpariert. In der Tat besitzt das Gesamtsystem determinierte Werte für einen vollständig bestimmten Satz von kompatiblen Spinobservablen, nämlich $\vec{\mathbf{S}}_1^2$, $\vec{\mathbf{S}}_2^2$, $\vec{\mathbf{S}}^2$ und $\mathbf{S}_3$, und dem entsprechend ist (3.20.18) auch der (bis auf einen Phasenfaktor) eindeutig bestimmte simultane Eigenvektor dieser drei Operatoren.

Obwohl nun die Einzelspinkomponenten $\mathbf{S}_{13}$ und $\mathbf{S}_{23}$ völlig zufällig sind, besteht doch eine starke Korrelation. Betrachten wir nämlich die vier möglichen Ausgänge einer simultanen Messung der Spinkomponenten beider Teilchen, erhalten wir (*nachrechnen*)

$$\begin{aligned} P_{12}(\sigma_1 = \hbar/2, \sigma_2 = \hbar/2) = P_{12}(\sigma_1 = -\hbar/2, \sigma_2 = -\hbar/2) = 0, \\ P_{12}(\sigma_1 = \hbar/2, \sigma_2 = -\hbar/2) = P_{12}(\sigma_1 = -\hbar/2, \sigma_2 = \hbar/2) = \frac{1}{2}. \end{aligned} \tag{3.20.20}$$

Die Spinkomponenten nehmen also nie den gleichen Wert an, und wenn man am ersten Teilchen $\sigma_1 = \hbar/2$ misst, misst man folglich am zweiten Teilchen notwendigerweise $\sigma_2 = -\hbar/2$ und umgekehrt. Misst man also Spinkomponenten der beiden Teilchen in gleicher Richtung hat man eine 100%-Korrelation für die Ausgänge der Messung. Dabei kann die Messung wieder so erfolgen, dass eine mögliche gegenseitige kausale Beeinflussung der beiden Messungen ausgeschlossen ist.

Einsteins Haupteinwand gegen die Vollständigkeit der Quantentheorie war diese Eigenschaft verschränkter Zustände, dass beliebig weit voneinander entfernte Teile eines Quantensystems (in diesem Beispiel zweier Teilchen, an denen Messungen, z.B. von Spinkomponenten, an weit voneinander entfernten Orten ausgeführt werden) stark korreliert sein können (**Inseparabilität**), obwohl diese gemessenen Spinkomponenten vor der Messung prinzipiell keine bestimmten Werte besitzen [Ein48].

Entsprechend der in diesem Skript dargestellten minimalen statistischen Interpretation erfordert die Erklärung der Korrelation auch keine „spukhaften Fernwirkungen" (Einstein), die einen der Relativitätstheorie widersprechenden überlichtschnellen Signalaustausch zwischen den beiden Messapparaten implizieren würde. Vielmehr ist diese Korrelation, genau wie auch die völlige Indeterminiertheit der Einzelspinkomponenten, durch die anfängliche Präparation in dem Zustand (3.20.18) bedingt.

Man muss allerdings betonen, dass diese Grundlagenfragen auch heute noch zum Teil kontrovers diskutiert werden. Z.B. kommt Steven Weinberg (1933-2021), einer der bedeutendsten theoretischen Hochenergiephysiker (Nobelpreis 1979 gemeinsam mit Sheldon Glashow und Abdus Salam für die Entwicklung des Standardmodells der elektroschwachen Wechselwirkung) in seinem sehr empfehlenswerten Lehrbuch der Quantentheorie [Wei15] zu dem Schluss, dass die interpretatorischen Grundlagenprobleme der Quantentheorie noch (immer) nicht gelöst seien.

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 zum Einen 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 Ruhsystem 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“ [Ein05b] 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-Newtonschen 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-Newtonschen 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-Newtonschen 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

$$t' = t, \quad \vec{x}' = \vec{x} - \vec{v}t \quad \text{(Galilei-Newton).} \tag{4.1.1}$$

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 mehrerer Gedankenexperimente finden. Diese Herleitung findet sich in dem hervorragenden Schulbuch [GK20].

4.1.1 Herleitung der Lorentz-Transformation (physikalische Formulierung)

Im Folgenden betrachten wir zwei Inertialsysteme $\Sigma(t, \vec{x})$ und $\Sigma'(t', \vec{x}')$. Wir verwenden für beide Bezugssysteme die gleichen Basisvektoren $\vec{e}_j' = \vec{e}_j$ für die jeweiligen Ortsvektoren und nehmen an, dass sich das System Σ' mit der konstanten Geschwindigkeit $\vec{v} = v\vec{e}_1$ gegenüber dem System Σ bewegt. Wir postulieren dabei, dass für jeden Beobachter in einem Inertialsystem der Raum wie in der klassischen Physik durch einen **euklidischen Raum** beschrieben wird. Aus Symmetriegründen sollte die mit den Einsteinschen Postulaten verträgliche **Lorentz-Transformation** die zu $\vec{v}$ senkrechten Komponenten des Ortsvektors $\vec{x}$ ungeändert lassen, d.h.

$$x_2' = x_2, \quad x_3' = x_3. \tag{4.1.2}$$

Da gemäß dem ersten Postulat das spezielle Relativitätsprinzip gilt, müssen die Transformationen der Komponente in Richtung von $\vec{v}$ sowie der Zeit **linear** sein, d.h.

$$t' = At + Bx_1, \quad x_1' = Ct + Dx_1, \tag{4.1.3}$$

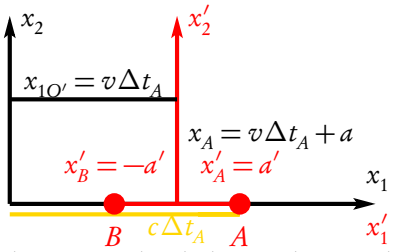
und wir wollen die Konstanten A, B, C und D finden, so dass diese Transformation das 2. Einstein-Postulat von der Unabhängigkeit der Lichtgeschwindigkeit c von der Geschwindigkeit der Lichtquelle gewährleistet. Dies erreichen wir durch die Analyse mehrerer physikalischer Gedankenexperimente.

Zunächst müssen wir die Zeitmessung in einem Inertialsystem so definieren, dass dieses 2. Postulat erfüllt ist. Wir führen dazu zuerst eine **Uhrensynchronisation** von Uhren im Inertialsystem Σ' durch. Dazu verwenden wir **Lichtsignale**, denn sie sind durch das 2. Einstein-Postulat durch ihre Ausbreitungsgeschwindigkeit c ausgezeichnet. Wir denken uns also eine Uhr im Ursprung $\vec{x}'$ ruhend, und wir wollen die Zeit in allen weiteren Punkten im Raum durch am jeweiligen Ort bzgl. Σ' ruhende weitere Uhren angezeigt denken. Wir müssen

nun all diese Uhren mit der Uhr im Ursprung so synchronisieren, dass ein zur Zeit $t' = 0$ vom Ursprung ausgesandtes Lichtsignal bei einer Uhr, die im Abstand r zum Ursprung befestigt ist, zur Zeit $t_r = r/c$ dort ankommt. Im Folgenden genügt es, Uhren entlang der x_1-Achse zu betrachten. Die Uhr an diesem Ort muss also zum Zeitpunkt des Eintreffens des Lichtsignals die Zeit $t_r = r/c$ anzeigen. Dann sind die an jedem Ort befindlichen bzgl. Σ' befindlichen Uhren zueinander synchronisiert.

Wir betrachten also nun eine bei x_1' in Σ' ruhende Uhr. Der Zeitpunkt, bei dem das zur Zeit $t' = 0$ am Ursprung ausgesandte Signal bei dieser Uhr eintrifft, ist gemäß dieser Synchronisationskonvention $\Delta t' = x_1'/c$. Genauso synchronisieren wir freilich auch die bzgl. Σ ruhenden Uhren an jedem Punkt im Raum mit der im Ursprung von Σ ruhenden Uhr. Wir nehmen im Folgenden der Einfachheit halber an, dass zur Zeit $t = 0$ der Ursprung von Σ' mit dem Ursprung von Σ zusammenfällt und dann auch $t' = 0$ ist.

Als nächstes machen wir uns klar, dass die in Σ' synchronisierten Uhren nicht mit den Uhren in Σ synchronisiert sein können. Dazu betrachten wir die in Σ' bei $x_1' = a'$ (Punkt A in der nebenstehenden Abbildung) und bei $x_1' = -a$ (Punkt B) ruhenden Uhren und die zur Zeit $t' = 0$ bei $x' = 0$ zur Synchronisation dieser Uhren ausgesandten Lichtsignale. Es ist klar, dass aufgrund unserer Synchronisationsvorschrift beide Uhren die gleiche Zeit $\Delta t' = a'/c$ anzeigen wenn die Signale bei den jeweiligen Uhren eintreffen, d.h. das Lichtsignal kommt im System Σ' gleichzeitig bei den beiden Uhren an.



Bzgl. Σ besitzen zur Zeit $t = 0$ die Punkte A und B die Koordinaten $x_{1A} = a$ und $x_{1B} = -a$. Wie wir weiter unten sehen werden, ist $a \neq a'$, da A und B sich bzgl. Σ mit der Geschwindigkeit v in x_1-Richtung bewegen. Das vom Ursprung von Σ' zur Zeit $t' = t = 0$ zum Punkt A ausgesandte Lichtsignal bewegt sich wegen des 2. Postulats nun auch bzgl. Σ mit der Lichtgeschwindigkeit c. Das Signal benötigt nun aber verschiedene Zeiten um nach der Zeit Δt_A bei A und nach der Zeit Δt_B bei B einzutreffen, denn A und B haben sich ja in diesen Zeiten mit der Geschwindigkeit v weiterbewegt. Im nebenstehenden Bild haben wir den Zeitpunkt festgehalten, zu dem das zu A gesandte Lichtsignal bei A eintrifft. Es gilt also

$$c\Delta t_A = a + v\Delta t_A, \quad c\Delta t_B = a - v\Delta t_B. \tag{4.1.4}$$

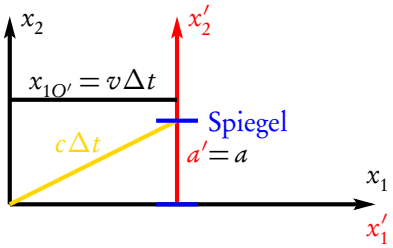
Die Situation beim Eintreffen des Lichtsignals nach der Zeit Δt_B lässt sich in einer analogen Skizze darstellen, aus der das genannte Resultat für Δt_B folgt (*Übung!*). Daraus ergibt sich

$$\Delta t_A = \frac{a}{c-v}, \quad \Delta t_B = \frac{a}{c+v}, \tag{4.1.5}$$

d.h. die in Σ' ruhenden Uhren sind bzgl. Σ nicht synchronisiert.

Man nennt dieses in der Newtonschen Physik unbekannte Phänomen auch die **Relativität der Gleichzeitigkeit**, denn wir haben gesehen, dass das Lichtsignal von Σ' aus gesehen bei A und B *gleichzeitig* ankommt. Von Σ aus betrachtet erreicht das Signal A später als B, d.h. bzgl. Σ erreicht das Signal die Punkte A und B *nicht gleichzeitig*. Man kann also die Einsteinschen Postulate nur erfüllen, indem man eine Synchronisation von gegeneinander bewegten Uhren aufgibt. Wegen des 2. Postulats zeigt dies formal, dass sich die Zeit im Gegensatz zur Newtonschen Raumzeitbeschreibung von einem Inertialsystem zu einem anderen ändern muss, damit die im jeweiligen Bezugssystem ruhenden Uhren untereinander synchronisiert sind. Außerdem sind dann notwendig die in Σ ruhenden Uhren *nicht* mit den Σ' ruheden Uhren synchron.

Als nächstes betrachten wir eine in Σ' bei $x' = 0$ ruhende Uhr. Wir konstruieren uns dabei in einem Gedankenexperiment eine solche Uhr mittels zwei parallel zur x_1'-Achse angeordnete Spiegeln bei $x_2' = 0$ und $x_2' = a'$. Wir senden nun vom unteren Spiegel zur Zeit $t' = 0$ ein Lichtsignal aus. Die Lichtlaufzeit in Σ' ist $\Delta t' = a'/c$. Dann wird das Signal am oberen Spiegel reflektiert und gelangt nach einer weiteren Zeit $\Delta t'$ am unteren Spiegel wieder an usw. Dadurch haben wir mit den regelmäßigen „Ticks“ der Zeitdauer $\Delta t'$ eine Uhr realisiert, die die Zeit im System Σ' anzeigt.



Wir können diese Ticks nun auch als Uhr im System Σ verwenden. Allerdings stimmen die Ticks dieser in Σ' ruhenden Uhr nicht mit den Ticks einer baugleichen in Σ ruhenden Lichtuhr überein. Im nebenstehenden Bild stellen wir die Situation zu der Zeit Δt dar, wenn das bei $t = t' = 0$ bei $x_1 = x_1' = 0$ ausgesandte Lichtsignal am oberen Spiegel eintrifft, aus Sicht des Inertialsystems Σ dar. Wegen des obigen Symmetriearguments, dass sich bei einer Relativbewegung der Inertialsysteme in x_1-Richtung die Komponenten von Vektoren in x_2-Richtung nicht ändern, besitzen die beiden in Σ' ruhenden und also in Σ mit der Geschwindigkeit v in x_1-Richtung bewegten Spiegel auch in Σ den Abstand $a = a'$. Da sich gemäß dem 2. Einstein-Postulat auch in Σ das Lichtsignal mit der Geschwindigkeit c bewegt, muss gemäß der Abbildung und dem Satz des Phythagoras

$$(c\Delta t)^2 = a^2 + v^2\Delta t^2 \Rightarrow \Delta t = \frac{a}{\sqrt{c^2 - v^2}} = \frac{c\Delta t'}{\sqrt{c^2 - v^2}} = \frac{\Delta t'}{\sqrt{1-\beta^2}} \quad \text{mit} \quad \beta = \frac{v}{c} \tag{4.1.6}$$

sein. Die Ticks der in Σ' ruhenden Uhr sind also bzgl. in Σ ruhenden Uhren um $\Delta t/\Delta t' = \gamma$ um den **Lorentz-Faktor**

$$\gamma = \frac{1}{\sqrt{1-\beta^2}} = \frac{1}{\sqrt{1 - v^2/c^2}} \tag{4.1.7}$$

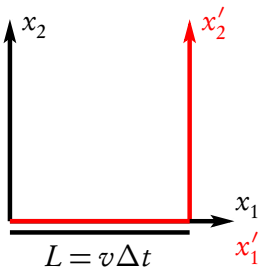
gedehnt. Dieses in der Newtonschen Physik ebenfalls unbekannte Phänomen nennt man **Zeitdilation**.

Es ist wichtig zu bemerken, dass man die **Eigenzeit** $\Delta\tau$ einer Uhr, die definiert ist als diejenige Zeit, die im Ruhsystem dieser Uhr vergeht und entsprechend von dieser Uhr angezeigt wird, mittels der Zeit in jedem beliebigen Inertialsystem angeben kann. In unserem Fall gilt

$$\Delta\tau = \Delta t' = \frac{\Delta t}{\gamma} = \Delta t\sqrt{1-\beta^2}. \tag{4.1.8}$$

Schließlich stellen wir fest, dass all diese Betrachtungen nur Sinn ergeben, wenn $|\beta| < 1$ also $|v| < c$ ist, d.h. in der speziellen Relativitätstheorie kann es keine Relativgeschwindigkeiten $\geq c$ zwischen zwei Inertialsystemen geben! In diesem Sinne ist die Lichtgeschwindigkeit eine universelle **Grenzgeschwindigkeit**.

Als nächstes zeigen wir, dass auch die Längen von Objekten von deren Bewegungszustand abhängen. Betrachten wir dazu einen parallel zur x_1'-Achse im System Σ' ruhenden Stab der Länge L'. Die Länge des Stabes in Σ, wo er sich mit der Geschwindigkeit v entlang der x_1-Achse bewegt, muss nun dadurch bestimmt werden, dass man gleichzeit, z.B. zur Zeit t die Koordinaten der Punkte A und B abliest. Es ist klar, dass die Ablesung von Σ' aus betrachtet nicht gleichzeitig erfolgt, woraus sich bzgl. Σ eine andere Länge $L \neq L'$ ergibt.

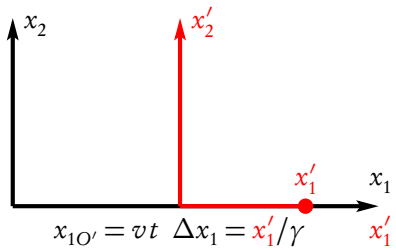


Diese Länge können wir nun mit Hilfe einer bzgl. Σ im Ursprung ruhenden Uhr bestimmen. Zur Zeit $t = 0$ sei das Stabende A bei der Uhr ($x_1 = x_1' = 0$) und das Stabende B sei zur Zeit Δt bei der Uhr eingetroffen (s. nebenstehende Skizze). Da sich der Stab bzgl. Σ mit der Geschwindigkeit v entlang der x_1-Achse bewegt, ist $\Delta t = L/v$, wobei L die in Σ gemessene Länge des Stabes ist. Betrachten wir denselben Vorgang in Σ', benötigt die in Σ ruhende gegenüber Σ' mit $-v$ entlang der x_1'-Achse bewegte Uhr die Zeit $\Delta t' = L'/v$, um an der Uhr vorbeizufliegen. Die Eigenzeit $\Delta\tau = \Delta t$ dieser Uhr ergibt sich gemäß (4.1.8) nun zu

$$\Delta\tau = \Delta t = \frac{L}{v} = \Delta t'\sqrt{1-\beta^2} = \frac{L'}{v}\sqrt{1-\beta^2} \Rightarrow L = L'\sqrt{1-\beta^2} = \frac{L'}{\gamma}. \tag{4.1.9}$$

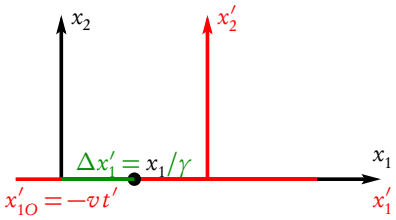
Die Länge erscheint also in Σ, wo sich der Stab mit der Geschwindigkeit $v = \beta c$ bewegt, um $1/\gamma = \sqrt{1-\beta^2}$ verkürzt.

Dieses in der Newtonschen Physik unbekannte Phänomene nennt man **Längenkontraktion**.



Mittels der Längenkontraktion können wir nun auch die **Lorentz-Transformation** herleiten, also die Konstanten A, B, C und D in (4.1.3) bestimmen. Dazu betrachten wir den bzgl. Σ' bei x_1' ruhenden Punkt. Zur Zeit t stellt sich die Situation in Σ wie in der nebenstehenden Skizze gezeigt dar. Wir lesen in Σ die Koordinate x_1 des Punktes zur Zeit t ab und erhalten für die in Σ' ruhende Länge in Σ die Länge $\Delta x_1 = x_1'/\gamma$. Der Ursprung von Σ' befindet sich dann bzgl. Σ bei $x_{1O'} = vt$. Es ist also

$$x_1 = vt + \Delta x_1 = vt + \frac{x_1'}{\gamma} \Rightarrow x_1' = \gamma(x_1 - vt). \tag{4.1.10}$$



Betrachten wir umgekehrt einen bzgl. Σ bei x_1 ruhenden Punkt in Σ' folgt analog, weil sich dieser Punkt bzgl. Σ' mit der Geschwindigkeit $(-v)$ entlang der x_1'-Achse bewegt und folglich wegen der Längenkontraktion $\Delta x_1' = x_1/\gamma$ ist,

$$x_1' = -(vt' - \Delta x_1') = \frac{x_1}{\gamma} - vt' \Rightarrow x_1 = \gamma(x_1' + vt'). \tag{4.1.11}$$

Setzen wir hier für x_1' das in (4.1.10) hergeleitete Resultat ein, können wir x_1' eliminieren und die resultierende Gleichung nach t' auflösen:

$$x_1 = \gamma^2(x_1 - vt) + \gamma v t' \Rightarrow t' = \gamma t - \frac{\gamma^2 - 1}{\gamma v} x_1. \tag{4.1.12}$$

Nun ist

$$\gamma^2 - 1 = \frac{1}{1-\beta^2} - 1 = \frac{1-(1-\beta^2)}{1-\beta^2} = \frac{\beta^2}{1-\beta^2} = \gamma^2\beta^2 \Rightarrow \frac{\gamma^2-1}{\gamma v} = \frac{\gamma^2-1}{\gamma\beta c} = \frac{\gamma\beta}{c} \tag{4.1.13}$$

und damit

$$t' = \gamma\left(t - \frac{\beta}{c}x_1\right). \tag{4.1.14}$$

Eliminieren wir umgekehrt mittels (4.1.11) x aus der Gleichung (4.1.10) und lösen nach t auf, erhalten wir

$$t = \gamma\left(t' + \frac{\beta}{c}x_1'\right). \tag{4.1.15}$$

Fassen wir (4.1.10) und (4.1.15) zusammen, erhalten wir für die Lorentz-Transformation also

$$\begin{aligned} t' &= \gamma\left(t - \frac{\beta}{c}x_1\right) \\ x_1' &= \gamma(x_1 - vt), \\ x_2' &= x_2, \\ x_3' &= x_3 \end{aligned} \tag{4.1.16}$$

und mit (4.1.11) und (4.1.14) die Umkehrtransformation

$$\begin{aligned} t &= \gamma\left(t' + \frac{\beta}{c}x_1'\right) \\ x_1 &= \gamma(x_1' + vt'), \\ x_2 &= x_2', \\ x_3 &= x_3'. \end{aligned} \tag{4.1.17}$$

4.1.2 Herleitung der Lorentz-Transformationen (geometrische Formlierung)

Wir wollen die eben aus physikalischen Betrachtungen mittels der Einsteinschen Postulate hergeleitete Lorentz-Transformation noch einmal unter einen anderen wichtigen Aspekt betrachten. Daraus ergibt sich eine sehr elegante geometrische Beschreibung des speziell relativistischen **Raum-Zeit-Modells** als ein vierdimensionaler „pseudo-Euklidischen" Raum, der sog. **Minkowski-Raum**. Auch hier gehen wir natürlich von den beiden Einsteinschen Postulaten und der Euklidizität des Raumes für jeden inertialen Beoabachter aus.

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

$$c^2t^2 - \vec{x}^2 = 0. \tag{4.1.18}$$

Dabei gehen wir stillschweigend davon aus, dass hinsichtlich der Geometrie des Raumes für einen in einem Inertialsystem ruhenden Beobachter die Euklidische Geometrie gilt. 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 (4.1.18) gelten, d.h. es gilt auch

$$c^2t'^2 - \vec{x}'^2 = 0. \tag{4.1.19}$$

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 $\underline{x} = (x^\mu) = (x^0, x^1, x^2, x^3)^{\mathrm{T}}$ 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**

$$\underline{x} \cdot \underline{y} = x^0y^0 - \vec{x} \cdot \vec{y} \tag{4.1.20}$$

ein, so dass Gleichung (4.1.18) in der Form

$$\underline{x} \cdot \underline{x} =: \boldsymbol{x}^2 = 0 \tag{4.1.21}$$

geschrieben werden kann.

Dies ist einem Skalarprodukt recht ähnlich, nur dass wegen der Vorzeichen in der Definition (4.1.20) 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 $\hat{\eta} = \mathrm{diag}(1, -1, -1, -1)$ mit den Komponenten

$$\eta_{\mu\nu} := \begin{cases} 1 & \text{for} \quad \mu = \nu = 0, \\ -1 & \text{for} \quad \mu = \nu, \quad \mu, \nu \in \{1, 2, 3\}, \\ 0 & \text{for} \quad \mu \neq \nu, \quad \mu, \nu \in \{0, 1, 2, 3\} \end{cases} \tag{4.1.22}$$

ein, können wir das das **Minkowski-Produkt** in der Form

$$\underline{x} \cdot \underline{y} = \eta_{\mu\nu}x^\mu y^\nu \tag{4.1.23}$$

schreiben. Dabei ist es für das Folgende sehr wichtig zu bemerken, das 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 (4.1.18) als auch (4.1.19) gelten soll, ist sicher garantiert, wenn für *alle* Vierervektoren $\underline{x}$ und $\underline{y}$ gilt

$$\underline{x}' \cdot \underline{y}' = \eta_{\mu\nu} x'^{\mu} y'^{\nu} = \eta_{\mu\nu} x^{\mu} y^{\nu} = \underline{x} \cdot \underline{y} \tag{4.1.24}$$

Sei nun $\hat{\Lambda}$ die entsprechende $\mathbb{R}^{4\times 4}$-Transformationsmatrix mit

$$\underline{x}' = \hat{\Lambda}\underline{x} \Leftrightarrow x'^{\mu} = \Lambda^{\mu}{}_{\nu} x^{\nu}. \tag{4.1.25}$$

In der Matrix-Vektorschreibweise ist nun

$$\underline{x}' \cdot \underline{y}' = \underline{x}'^{\mathrm{T}} \hat{\eta} \underline{y}' = (\hat{\Lambda}\underline{x})^{\mathrm{T}} \hat{\eta} (\hat{\Lambda}\underline{y}) = \underline{x}^{\mathrm{T}} (\hat{\Lambda}^{\mathrm{T}} \hat{\eta} \hat{\Lambda}) \underline{y}. \tag{4.1.26}$$

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

$$\hat{\Lambda}^{\mathrm{T}} \hat{\eta} \hat{\Lambda} = \hat{\eta} \tag{4.1.27}$$

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

$$\hat{\Lambda} = \begin{pmatrix} A & B \\ C & D \end{pmatrix} \tag{4.1.28}$$

wird (4.1.27) zu (*Nachrechnen*)

$$\begin{pmatrix} A & C \\ B & D \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} A & B \\ C & D \end{pmatrix} = \begin{pmatrix} A & -C \\ B & -D \end{pmatrix} \begin{pmatrix} A & B \\ C & D \end{pmatrix} = \begin{pmatrix} A^2 - C^2 & AB - CD \\ AB - CD & B^2 - D^2 \end{pmatrix} \overset{!}{=} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{4.1.29}$$

Die Matrixelemente A, B, C und D müssen also die Gleichungen

$$A^2 - C^2 = 1, \quad AB - CD = 0, \quad D^2 - B^2 = 1 \tag{4.1.30}$$

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

$$\cosh\alpha = \frac{1}{2}[\exp(\alpha) + \exp(-\alpha)], \quad \sinh\alpha = \frac{1}{2}[\exp(\alpha) - \exp(-\alpha)] \tag{4.1.31}$$

die Beziehung

$$\cosh^2(\alpha) - \sinh^2(\alpha) = 1 \tag{4.1.32}$$

gilt. Wir setzen nun also

$$A = D = \cosh\alpha, \quad B = C = -\sinh\alpha, \tag{4.1.33}$$

wobei sich das zusätzliche Minuszeichen im letzten Ausdruck gleich als nützlich herausstellen wird. In der Tat sind dann alle Bedingungsgleichungen (4.1.30) erfüllt (*Nachrechnen!*). Die Transformationsmatrix lautet also

$$\hat{\Lambda} = \begin{pmatrix} \cosh\alpha & -\sinh\alpha \\ -\sinh\alpha & \cosh\alpha \end{pmatrix}. \tag{4.1.34}$$

Um dies physikalisch interpretieren zu können, schreiben wir auch (4.1.25) explizit hin:

$$\underline{x}' = \begin{pmatrix} ct' \\ x'^1 \end{pmatrix} = \hat{\Lambda}\underline{x} = \begin{pmatrix} ct\cosh\alpha - x^1\sinh\alpha \\ -ct\sinh\alpha + x^1\cosh\alpha \end{pmatrix}. \tag{4.1.35}$$

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 (4.1.35) ergibt sich für diesen Punkt, von Σ aus betrachtet:

$$x^1\cosh\alpha - ct\sinh\alpha = 0 \Rightarrow x^1 = \frac{\sinh\alpha}{\cosh\alpha}ct = ct\tanh\alpha \overset{!}{=} vt = \beta ct. \tag{4.1.36}$$

Damit ist aber die gesuchte Beziehung zwischen α und $v = \beta c$ gefunden:

$$\beta = \tanh\alpha. \tag{4.1.37}$$

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 (4.1.35) noch mit Hilfe der Relativgeschwindigkeit $v = c\beta$ ausdrücken. Nun gilt

$$\begin{aligned} \tanh^2\alpha &= \frac{\sinh^2\alpha}{\cosh^2\alpha} = \frac{\cosh^2\alpha - 1}{\cosh^2\alpha} \Rightarrow \cosh^2\alpha = \frac{1}{1-\tanh^2\alpha} \\ \Rightarrow \cosh\alpha &= \frac{1}{\sqrt{1-\tanh^2\alpha}} = \frac{1}{\sqrt{1-\beta^2}} = \gamma. \end{aligned} \tag{4.1.38}$$

Dabei haben wir im letzten Schritt verwendet, dass $\cosh\alpha \geq 1 > 0$ für alle $\alpha \in \mathbb{R}$ ist und (4.1.37) gilt. Daraus folgt sofort

$$\sinh\alpha = \tanh\alpha\cosh\alpha = \beta\gamma = \frac{\beta}{\sqrt{1-\beta^2}}. \tag{4.1.39}$$

Die Transformation zwischen zwei Inertialsystemen, die sich mit einer Relativgeschwindigkeit $v = \beta c$ bewegen, lautet also gemäß (4.1.35) mit (4.1.38) und (4.1.39)

$$\begin{aligned} x'^0 &= ct' = \gamma(x^0 - \beta x^1) = \frac{ct - \beta x^1}{\sqrt{1-\beta^2}}, \\ x'^1 &= \gamma(x^1 - \beta x^0) = \frac{x^1 - vt}{\sqrt{1-\beta^2}}. \end{aligned} \tag{4.1.40}$$

Die entsprechende Transformationsmatrix (4.1.34) ergibt sich gemäß (4.1.38) und (4.1.39), jetzt wieder ausführlich für alle vier Raum-Zeit-Komponenten geschrieben, zu

$$\hat{\Lambda} = (\Lambda^\mu{}_\nu) = \begin{pmatrix} \gamma & -\beta\gamma & 0 & 0 \\ -\beta\gamma & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}. \tag{4.1.41}$$

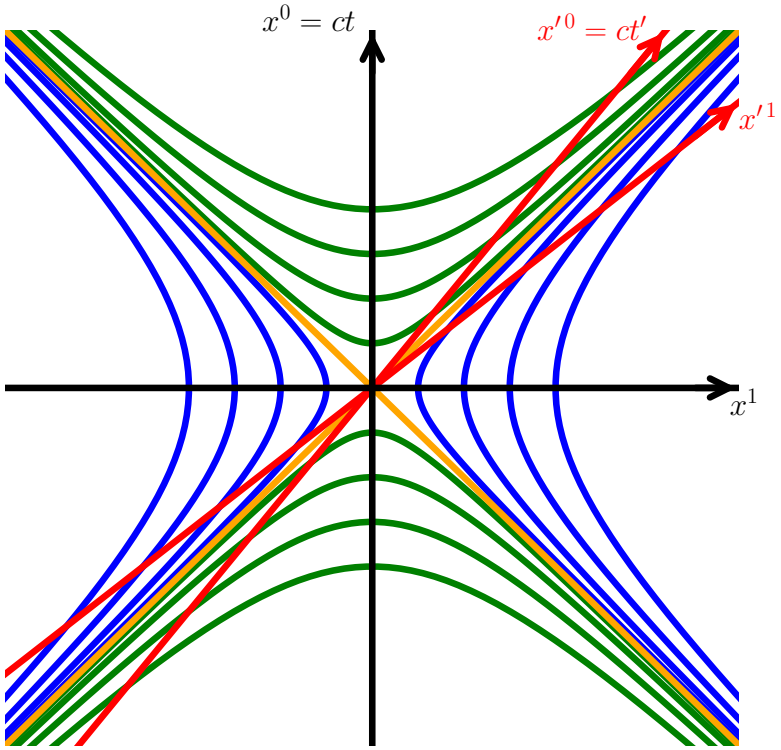


Abbildung 4.1: *Minkowski-Diagramm, das die Konstruktion der Raum-Zeit-Achsen des Systems Σ' illustriert: Die roten Linien sind entsprechend die ct' und x'^1-Achsen. Die Einheitslängen auf den Achsen werden durch die Schnittpunkte mit den grünen bzw. blauen Hyperbeln $(ct)^2 - (x^1)^2 = \pm(n\,Ls)^2$ mit $n \in \{1, 2, \ldots\}$ 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 Winkelhalbierenden 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.*

Die entsprechende Umkehrtransformation erhalten wir sehr einfach durch Anwendung von (4.1.27). Dazu bemerken wir, dass $\hat{\eta}^2 = \mathbb{1}$ ist. Multiplizieren wir also (4.1.27) von links mit $\hat{\eta}$, folgt

$$\hat{\eta}\hat{\Lambda}^{\mathrm{T}}\hat{\eta}\hat{\Lambda} = \mathbb{1} \Rightarrow \hat{\Lambda}^{-1} = \hat{\eta}\hat{\Lambda}^{\mathrm{T}}\hat{\eta} = \begin{pmatrix} \gamma & \beta\gamma & 0 & 0 \\ \beta\gamma & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}. \tag{4.1.42}$$

Die Umkehrtransformation zu (4.1.40) ist also durch

$$\underline{x} = \hat{\Lambda}^{-1}\underline{x}' \Rightarrow ct = \gamma(ct' + \beta x'^1), \quad x^1 = \gamma(x'^1 + \beta ct') \tag{4.1.43}$$

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 (4.1.40) bzw. die entsprechende Umkehrtransformation (4.1.43) 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 (4.1.40) ist diese Weltlinie durch

$$ct = \gamma ct', \quad x^1 = \gamma\beta ct' = \beta ct \tag{4.1.44}$$

gegeben. Entsprechend ist die x'^1-Achse durch $ct' = 0 = \text{const}$ gegeben. Die entsprechenden Koordinaten in Σ sind wieder durch (4.1.43) gegeben:

$$x^1 = \gamma x'^1, \quad ct = \gamma\beta x'^1 = \beta x'^1 \tag{4.1.45}$$

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 (4.1.44). 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 (4.1.40) bzw. (4.1.43) hergeleitet haben.

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 L s $= c\,1$ s $= 299\,792\,458$ 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 $ct' = 1$ Ls, $x'^0 = 0$ auf der entsprechenden **Einheitshyperbel**

$$(ct)^2 - (x^1)^2 = 1\,\text{Ls}^2 \tag{4.1.46}$$

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 $ct' = 0$, $x'^1 = 1$ Ls muss nämlich offenbar auf der Hyperbel

$$(ct)^2 - (x^1)^2 = -1\,\text{Ls}^2 \tag{4.1.47}$$

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.

Wir bemerken schließlich noch, dass das Transformationsverhalten der Komponenten Vierervektoren unter beliebigen Lorentz-Transformationen durch eine Matrix $\hat{\Lambda} = (\Lambda^\mu{}_\nu)$ durch

$$x'^\mu = \Lambda^\mu{}_\nu x^\nu. \tag{4.1.48}$$

Damit das Minkowski-Produkt ungeändert bleibt, muss für beliebige Vierervektoren $\underline{x}$ und $\underline{y}$

$$\underline{x}' \cdot \underline{y}' = \eta_{\mu\nu} x'^\mu y'^\nu = \eta_{\mu\nu} \Lambda^\mu{}_\rho x^\rho \Lambda^\nu{}_\sigma y^\sigma \stackrel{!}{=} \underline{x} \cdot \underline{y} = \eta_{\rho\sigma} x^\rho y^\sigma \tag{4.1.49}$$

sein. Damit folgt, dass $\hat{\Lambda}$ eine Lorentz-Transformationsmatrix ist, wenn

$$\eta_{\mu\nu} \Lambda^\mu{}_\rho \Lambda^\mu{}_\sigma = \eta_{\rho\sigma} \tag{4.1.50}$$

ist.

Weiter führt man vermöge

$$x_\mu = \eta_{\mu\nu} x^\nu \Rightarrow (x_\mu) = \begin{pmatrix} x^0 \\ -x^1 \\ -x^2 \\ -x^3 \end{pmatrix} \tag{4.1.51}$$

zur Abkürzung noch Vektorkomponenten mit unteren Indizes ein. Umgekehrt kann man auch einen unteren Index wieder „nach oben ziehen", indem man die zu $(\eta_{\mu\nu})$ inversen Matrix $(\eta^{\mu\nu})$, die aber wegen $\hat{\eta}^2 = \mathbb{1}$ ebenfalls die Form $(\eta^{\mu\nu}) = \mathrm{diag}(1,-1,-1-1)$ besitzt, anwendet:

$$x^\mu = \eta^{\mu\nu} x_\nu. \tag{4.1.52}$$

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 gleichzeitig, wenn $x^0 = y^0$ bzw. $x^0 - y^0 = 0$ gilt. Mittels des Lorentz-Boosts (4.1.40) folgt für den zeitlichen Abstand dieser Ereignisse bzgl. Σ'

$$x'^0 - y'^0 = \gamma[(x^0 - y^0) - \beta(x^1 - y^1)] = -\gamma\beta(x^1 - y^1). \tag{4.2.1}$$

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 (4.1.43) durch

$$ct' = \gamma ct, \quad x'^1 = -\gamma\beta ct = -\beta ct' = -vt'. \tag{4.2.2}$$

Abgesehen davon, dass sich selbstverständlich die Uhr bzgl. Σ' mit der Geschwindigkeit $-v = -\beta c$ in (negative) x'^1=Richtung bewegt, liest der in Σ' ruhende Beobachter auch eine andere Zeit ab. Ist im Ruhsystem der Uhr Σ die Zeit $t_A = 1$ Ls vergangen, ist bzgl. Σ' gemäß (4.2.2) die längere Zeit $t' = \gamma 1$ Ls vergangen. Die Zeit in Σ' erscheint also gegenüber der Zeit im Ruhsystem 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

$$\underline{x} = \begin{pmatrix} \lambda_1 \\ x^1 \end{pmatrix}, \quad \underline{y} = \begin{pmatrix} \lambda_2 \\ y^1 \end{pmatrix}, \quad \lambda_1, \lambda_2 \in \mathbb{R} \tag{4.2.3}$$

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

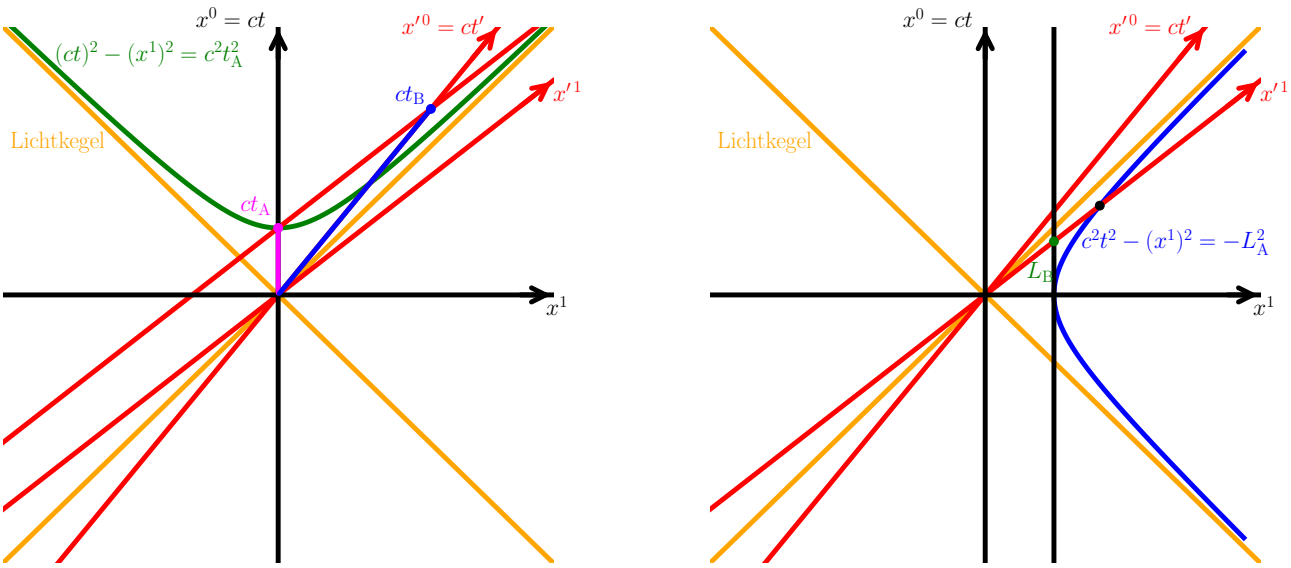


Abbildung 4.2: ***Links:*** *Minkowski-Diagramm zur Zeitdilatation. Die Zeit, die zwischen den Endpunkten der in magenta eingezeichneten Weltlinie bzgl. Σ am gleichen Ort $x^1 = 0$ vergeht, erscheint bzgl. Σ' länger als bzgl. Σ. Die Rechnung (s. Text) ergibt, dass $ct' = ct\gamma > ct$ ist. Im Minkowskidiagramm liest man das daran ab, dass der Schnittpunkt der grünen Hyperbel $c^2t^2 - (x^1)^2 = c^2t_A^2$ die x'^1-Achse an einem Zeitpunkt $ct' < ct_B$ schneidet.* ***Rechts:*** *Minkowski-Diagramm zur Längenkontraktion: Die schwarzen 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. Im Bezugssystem Σ' 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 im Minkowski-Diagramm nicht durch die üblichen euklidischen Abstände in der Zeichenebene gegeben sind sondern durch die pseudo-euklidische Geometrie im Minkowski-Raum, wie in Abb. 4.1 konstruiert. Die Weltlinie des rechten Stabendes schneidet Bobs räumliche Achse bei L_B, während die Länge L_A auf dieser Achse durch den Schnittpunkt der entsprechenden blauen raumartigen Hyperbel $c^2t^2 - x^2 = -L_A^2$ markiert wird. Offensichtlich ist also $L_B < L_A$, d.h. für Alice, bzgl. der die Stange ruht, besitzt sie eine Länge L_A und für Bob die kürzere Länge L_B. Wie die Rechnung im Text zeigt, ist die Länge gerade um den inversen Lorentz-Faktor $1/\gamma = \sqrt{1-\beta^2}$ verkürzt.*

Relativkoordinaten der Enden für diese „Messereignisse“ der Stange und die entsprechende Länge sind dann

$$\underline{y} - \underline{x} = \begin{pmatrix} 0 \\ y^1 - x^1 \end{pmatrix} \Rightarrow L_\Sigma = |y^1 - x^1|. \tag{4.2.4}$$

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 (4.1.40) folgt aus (4.2.3)

$$\underline{y}' - \underline{x}' = \gamma \begin{pmatrix} (\lambda_2 - \lambda_1) - \beta(y^1 - x^1) \\ y^1 - x^1 - \beta(\lambda_2 - \lambda_1) \end{pmatrix}. \tag{4.2.5}$$

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.

$$\lambda_2 - \lambda_1 - \beta(y^1 - x^1) = 0 \Rightarrow \lambda_2 - \lambda_1 = \beta(y^1 - x^1) = \beta L_\Sigma. \tag{4.2.6}$$

Der räumliche Abstand ergibt sich nun, indem wir dies in die räumliche Komponente von (4.2.5) einsetzen

$$L_{\Sigma'} = |y'^1 - x'^1| = \gamma(1-\beta^2)|y^1 - x^1| = \sqrt{1-\beta^2}L_\Sigma = \frac{1}{\gamma}L_\Sigma. \tag{4.2.7}$$

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 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 ein 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 Ruhsystem des Bahnhofs, mit der konstanten Geschwindigkeit $\vec{v} = \beta c \vec{e}_1$ bewegt. Das Ruhsystem 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

$$\underline{x}'_\pm = \begin{pmatrix} ct'_\pm \\ \pm ct'_\pm \end{pmatrix}. \tag{4.2.8}$$

Besitzt der Zug in seinem Ruhsystem die Länge L', kommt demnach das Lichtsignal bzgl. Alice *gleichzeitig* zur Zeit $t' = t'_+ = t'_- = L'/(2c)$ bei den beiden Enden des Zuges an.

Betrachten wir nun die Weltlinien bzgl. des am Bahnhof stehenden Beobachters Bob, der also bzgl. Σ ruht. Dazu müssen wir den entsprechenden Lorentz-Boost (4.1.43) anwenden:

$$\underline{x}_\pm = \begin{pmatrix} \gamma & \beta\gamma \\ \beta\gamma & \gamma \end{pmatrix} \begin{pmatrix} ct' \\ \pm ct' \end{pmatrix} = \gamma c t' \begin{pmatrix} 1\pm\beta \\ 1\pm\beta \end{pmatrix} = \gamma \frac{L'}{2} \begin{pmatrix} 1\pm\beta \\ 1\pm\beta \end{pmatrix}. \tag{4.2.9}$$

Die beiden Lichtsignale erreichen also von Bob aus betrachtet die Enden des Zuges zu den Zeiten

$$t_\pm = \frac{\gamma(1\pm\beta)L'}{2c} = \frac{L'}{2c}\frac{1\pm\beta}{\sqrt{1-\beta^2}} = \frac{L'}{2c}\sqrt{\frac{(1\pm\beta)^2}{1-\beta^2}} = \frac{L'}{2c}\sqrt{\frac{1\pm\beta}{1\mp\beta}}. \tag{4.2.10}$$

Dies erklärt sich dadurch, dass von Bob aus betrachtet das Lichtsignal zum rechten (linken) Ende einen längeren (kürzeren) Weg zurücklegen muss als bzgl. Alice, weil sich in Σ 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.

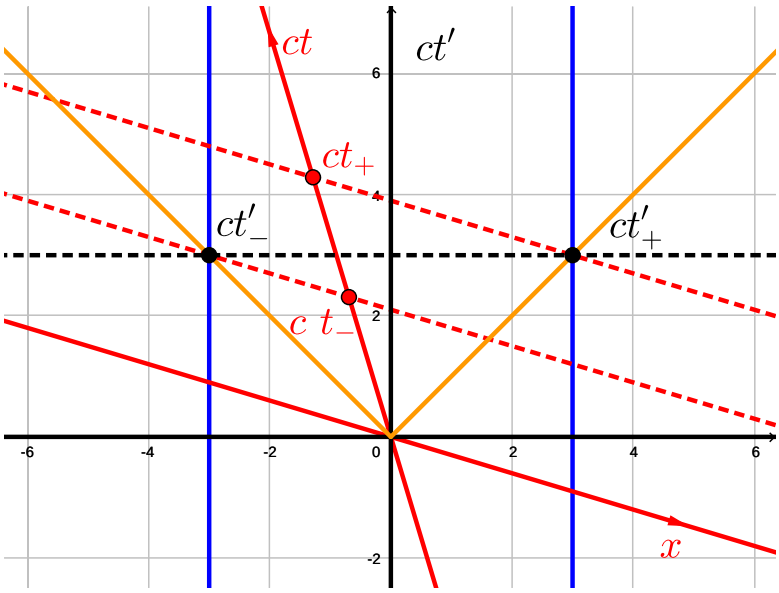


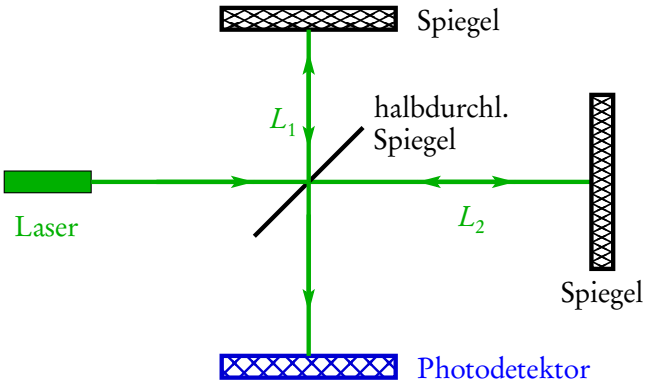
Abbildung 4.3: *Zum Einsteinschen Zug-Gedankenexperiment: Die ct'-Achse ist die Weltlinie von Alice, die sich in der Mitte des Zuges befindet. Die beiden blauen Weltlinien beschreiben die Zugenden. Die Weltlinien der Lichtsignale, die Alice zur Zeit $t = 0$ aussendet, treffen bei den Ereignissen A und 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.*

Mit dieser Überlegung kann man auch direkt in Bobs Bezugssystem argumentieren. Die Länge des Zuges ist in diesem Bezugssystem aufgrund der Längenkontraktion $L = L'/\gamma$, d.h. das Lichtsignal, das beim rechten (linken) Zugende eintrifft, hat den Weg

$$c\,t_\pm = \frac{L}{2} \pm \beta c\,t_\pm \Rightarrow t_\pm = \frac{L}{2c(1\mp\beta)} = \frac{L'}{2c\gamma(1\mp\beta)} = \frac{L'}{2c}\frac{\sqrt{1-\beta^2}}{1\mp\beta} = \frac{L'}{2c}\sqrt{\frac{1\pm\beta}{1\mp\beta}}, \tag{4.2.11}$$

und das stimmt mit (4.2.10) überein.

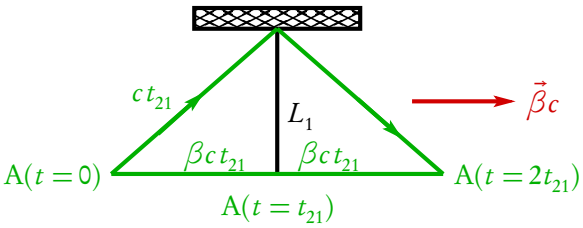
4.2.5 Das Michelson-Morley-Experiment



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 Ätherruhesystem 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 $v = 3 \cdot 10^4 \mathrm{m/s}$, also $\beta = v/c = 10^{-4}$.

Betrachten wir aber zunächst das Experiment im Ruhsystem 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 zwischwen 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 $ct_{11} = L_2 + \beta c t_{11}$ 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 $ct_{12} = L_2 - \beta c t_{12}$. Insgesamt ist die Laufzeit vom Strahlteiler zum Spiegel und zurück für den sich in x^1-Richtung hin- und herbewegenden Teilstrahl also

$$t_1 = t_{11} + t_{12} = \frac{L_2}{c(1+\beta)} + \frac{L_2}{c(1-\beta)} = \frac{2L_2}{c(1-\beta^2)} = \frac{2L_2\gamma^2}{c}. \tag{4.2.12}$$

Bei dem zweiten Teilstrahl, der sich im Ruhsystem der Apparatur senkrecht zur x^1-Richtung zwischen Strahlteiler und Spiegel hin- und herbewegt, gilt (s. Sizze) für jeweils einen Weg $c^2t_{21}^2 = L_1^2 + \beta^2c^2t_{21}^2$, d.h.

$$t_2 = 2t_{21} = \frac{2L_1}{c}\frac{1}{\sqrt{1-\beta^2}} = \frac{2L_1\gamma}{c}. \tag{4.2.13}$$

Die Laufzeitdifferenz ist also

$$\Delta t_{\mathrm{A}} = \frac{2}{c}(L_1\gamma - L_2\gamma^2). \tag{4.2.14}$$

Dreht man nun die Apparatur um 90°, erhält man hingegen

$$\Delta t_{\mathrm{B}} = \frac{2}{c}(L_1\gamma^2 - L_2\gamma), \tag{4.2.15}$$

d.h. das Interferenzmuster verschiebt sich im Vergleich zur vorherigen Ausrichtung entsprechend einer Änderung der Laufzeitdifferenz um

$$\Delta t_{\mathrm{A}} - \Delta t_{\mathrm{B}} = \frac{2(L_1+L_2)}{c}(\gamma^2-\gamma). \tag{4.2.16}$$

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

$$\gamma^2 - \gamma = (1-\beta^2)^{-1} - (1-\beta^2)^{-1/2} = 1 + \beta^2 - (1+\beta^2/2) + \mathcal{O}(\beta^4) = \frac{\beta^2}{2}. \tag{4.2.17}$$

Die Laufzeitverschiebung und damit die Änderung des Interferenzmusters ist also ein Effekt der Ordnung β^2:

$$\Delta t_{\mathrm{A}} - \Delta t_{\mathrm{B}} \simeq \frac{L_1+L_2}{c}\beta^2. \tag{4.2.18}$$

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

$$\frac{d}{\lambda} = \frac{c}{\lambda}(\Delta t_{\mathrm{A}} - \Delta t_{\mathrm{B}}) = \frac{L_1+L_2}{\lambda}\beta^2, \tag{4.2.19}$$

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 (4.2.19) eine relative Verschiebung des Interferenzmusters zu $d/\lambda \simeq 0{,}44$. Die Apparatur war genau genug, um zu zeigen, dass tatsaächlich $d/\lambda < 0{,}02$ 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 auführlich 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 (4.2.14) lediglich L_2 durch die Lorentz-kontrahierte Länge $L_2' = L_2/\gamma$ zu ersetzen, während $L_1' = L_1$ bleibt. Bei der um 90° gedrehten Apparatur ist hingegen $L_1' = L_1/\gamma$ statt L_1 in (4.2.15) und $L_2' = L_2$ zu setzen. Dies liefert entsprechend

$$\Delta t_{\mathrm{A}} = \Delta t_{\mathrm{B}} = \frac{L_1 - L_2}{c}\gamma, \tag{4.2.20}$$

d.h. $\Delta t_{\mathrm{A}} - \Delta t_{\mathrm{B}} = 0$ in Übereinstimmung mit dem Nullresultat des Michelson-Morley-Experiments.

Zum Abschluss unserer Diskussion des Michelson-Moreley-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-Transformationauch 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 entsprechnde (nicht auf Länge 1 normierte) Richtungsvektor ist demnach $\vec{n}' = (\cos\theta/\gamma, \sin\theta)$, und der Winkel θ' ist demnach durch

$$\tan\theta' = \frac{\sin\theta}{\cos\theta/\gamma} = \gamma\tan\theta \tag{4.2.21}$$

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 $\tan\theta \to \infty$, und damit ist auch $\tan\theta' \to \infty$, 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 Längenkontraktionshypthese 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 [Ein05b] 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 auftritt. S. dazu den sehr lesenswerten Wikipedia-Artikel zur Historie der Lorentz-Transformation [1]. Diese Transformation läßt nun nicht alle Minkowksi-Produkte invariant sondern nur die Beziehung $(ct)^2 - \vec{x}^2 = 0$. Dies genügt aber dafür, dass die quellenfreie Wellengleichung $\Box\Phi = 0$ (bzw. für die freien elektromagnetischen Wellen $\Box\vec{E} = 0$, $\Box\vec{B} = 0$) bzgl. dieser Transformation invariant ist.

[1] https://de.wikipedia.org/wiki/Geschichte_der_Lorentz-Transformation.

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

$$[x^\mu(t)] = \begin{pmatrix} ct \\ R[\cos(\omega t)-1] \\ R\sin\omega t \\ 0 \end{pmatrix}, \quad \frac{\mathrm{d}}{\mathrm{d}t}x^\mu(t) = \begin{pmatrix} c \\ -R\omega\sin(\omega t) \\ R\omega\cos(\omega t) \\ 0 \end{pmatrix} \tag{4.3.1}$$

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 $\mathrm{d}t$, so vergeht in Bobs Bezugssystem nur dessen Eigenzeit $\mathrm{d}\tau = \mathrm{d}t\sqrt{\eta_{\mu\nu}\dot{x}^\mu\dot{x}^\nu}/c$. Dies wird dadurch klar, dass die Eigenzeit in Bobs momentanen Ruhsystem definiert ist, wo $(\mathrm{d}x'^\mu) = (c\mathrm{d}t', 0, 0, 0) = (c\mathrm{d}\tau, 0, 0, 0)$ gilt. Damit ist aber $c^2\mathrm{d}\tau^2 = \eta_{\mu\nu}\mathrm{d}'^\mu\mathrm{d}x'^\nu = \eta_{\mu\nu}\mathrm{d}x^\mu\mathrm{d}x^\nu$, 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 Ruhsystem von Bob vergeht, zu

$$\tau = \frac{1}{c}\int_0^t \mathrm{d}\tilde{t}\sqrt{\frac{\mathrm{d}x^\mu(\tilde{t})}{\mathrm{d}\tilde{t}}\frac{\mathrm{d}x_\mu(\tilde{t})}{\mathrm{d}\tilde{t}}} = \int_0^t \mathrm{d}\tilde{t}\sqrt{1-\frac{\omega^2R^2}{c^2}} = t\sqrt{1-\frac{\omega^2R^2}{c^2}}. \tag{4.3.2}$$

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 $\tau = 2\pi/\omega\sqrt{1-\omega^2R^2/c^2}$, also um den inversen Lorentz-Faktor weniger, gealtert als Alice.
Dieser Effekt wurde mehrfach anhand der Messung der Lebensdauer von instabilen Teilchen in Speicherringen überprüft. Vgl. dazu den entsprechenden Wikipedia-Artikel[2].

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 $\underline{x} = (0, \vec{x})^\mathrm{T}$ und $\underline{y} = (0, \vec{y})^\mathrm{T}$. Dann gilt für die entsprechende Differenz

$$(\underline{x}-\underline{y})\cdot(\underline{x}-\underline{y}) = -(\vec{x}-\vec{y})^2 < 0. \tag{4.4.1}$$

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

[2] `https://en.wikipedia.org/wiki/Time_dilation_of_moving_particles`

$\vec{v} = \beta c \vec{e}_1$, gilt gemäß (4.1.40)

$$\begin{pmatrix} c(t'_x - t'_y) \\ x'^1 - y'^1 \\ x'^2 - y'^1 \\ x'^3 - y'^3 \end{pmatrix} = \begin{pmatrix} -\gamma\beta(x^1 - y^1) \\ \gamma(x^1 - y^1) \\ x^2 - y^2 \\ x^3 - y^3 \end{pmatrix}. \tag{4.4.2}$$

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.

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'

$$\begin{pmatrix} c(t'_x - t'_y) \\ x'^1 - y'^1 \\ x'^2 - y'^1 \\ x'^3 - y'^3 \end{pmatrix} = \begin{pmatrix} \gamma[c(t_x - t_y) - \beta(x^1 - y^1)] \\ \gamma[(x^1 - y^1) - \beta c(t_x - t_y)] \\ x^2 - y^2 \\ x^3 - y^3 \end{pmatrix}. \tag{4.4.3}$$

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

$$\begin{aligned} c(t'_x - t'_y) &= \gamma[c(t_x - t_y) - \beta(x^1 - y^1)] \\ &\geq \gamma[c(t_x - t_y) - |\beta||x^1 - y^1|] \\ &\geq \gamma[c(t_x - t_y) - |\beta||\vec{x} - \vec{y}|] > \gamma[c(t_x - t_y) - |\vec{x} - \vec{y}|] > 0. \end{aligned} \tag{4.4.4}$$

Der Ausweg aus diesem Dilemma ist es, die Kraftwirkungen zwischen Punktteilchens über **Felder** zu vermitteln. Das Paradebeispiel ist die elektromagnetische Wechselwirkung zwischen geladenen Teilchen. Im Feldbild erfolgt die Wechselwirkung dadurch, dass aufgrund des einen Teilchens ein elektromagnetisches Feld, das sich (maximal mit Lichtgeschwindigkeit!) ausbreitet, erzeugt wird und auf das andere Teilchen aufgrund dieses Feldes an seinem Ort, also **lokal**, die entsprechende Kraft wirkt. In dieser Vorlesung werden wir dieses schwierige Problem allerdings nicht behandeln sondern uns auf die Bewegung von Teilchen in vorgegebenen Feldern beschränken.

Wir 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 $\vec{v} = \mathrm{d}\vec{x}/\mathrm{d}t$ 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 Ruhsystem** können wir nun **kovariant** schreiben. Offenbar ist nämlich in diesem Bezugssystem

$$\frac{\mathrm{d}}{\mathrm{d}\lambda}\begin{pmatrix} ct'(\lambda) \\ \vec{x}' \end{pmatrix} = \begin{pmatrix} c\mathrm{d}t'/\mathrm{d}\lambda \\ \vec{0} \end{pmatrix}, \tag{4.4.5}$$

denn voraussetzungsgemäß ist die Geschwindigkeit in diesem Bezugssystem $\mathrm{d}\vec{x}'/\mathrm{d}t' = (\mathrm{d}\vec{x}/\mathrm{d}\lambda)(\mathrm{d}\lambda/\mathrm{d}t') = 0$. Damit ist in diesem Bezugssystem

$$\mathrm{d}t' = \mathrm{d}\lambda \frac{1}{c}\sqrt{\frac{\mathrm{d}\underline{x}'}{\mathrm{d}\lambda}\cdot\frac{\mathrm{d}\underline{x}'}{\mathrm{d}\lambda}}. \tag{4.4.6}$$

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 Ruhsystem des Teilchens $\mathrm{d}t' = \mathrm{d}\tau$ kovariant schreiben. In diesem Zusammenhang nennt man τ auch die **Eigenzeit** des Teilchens:

$$\mathrm{d}\tau = \mathrm{d}\lambda \frac{1}{c}\sqrt{\frac{\mathrm{d}\underline{x}}{\mathrm{d}\lambda}\cdot\frac{\mathrm{d}\underline{x}}{\mathrm{d}\lambda}}. \tag{4.4.7}$$

Nun ist der Weltlinienparameter λ beliebig wählbar, und wir können demnach für λ auch die Zeit t im Bezugssystem Σ verwenden:

$$\mathrm{d}\tau = \mathrm{d}t \frac{1}{c}\sqrt{\frac{\mathrm{d}\underline{x}}{\mathrm{d}t}\cdot\frac{\mathrm{d}\underline{x}}{\mathrm{d}t}} = \mathrm{d}t\sqrt{1-\vec{\beta}^2} \tag{4.4.8}$$

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 Ruhsystem des Teilchens in der Relativitätstheorie die korrekte Definition sein. Dies legt es nahe, in der Relativitätstheorie den **Vierer-Impulsvektor** durch

$$\underline{p} = m\frac{\mathrm{d}\underline{x}}{\mathrm{d}\tau} \tag{4.4.9}$$

zu definieren. Da nämlich gemäß (4.4.6) $\mathrm{d}\tau$, 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 ist[3].

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

$$\underline{p} = m\frac{\mathrm{d}\underline{x}}{\mathrm{d}\tau} = m\frac{\mathrm{d}\underline{x}}{\mathrm{d}t}\frac{\mathrm{d}t}{\mathrm{d}\tau}. \tag{4.4.10}$$

Nun gilt wegen der Definition (4.4.8) des Eigenzeitinkrements

$$\frac{\mathrm{d}t}{\mathrm{d}\tau} = \left(\frac{\mathrm{d}\tau}{\mathrm{d}t}\right)^{-1} = \frac{1}{\sqrt{1-\dot{\vec{x}}^2/c^2}} = \frac{1}{\sqrt{1-\vec{\beta}^2}} = \gamma \tag{4.4.11}$$

[3] In einigen älteren Lehrbüchern und leider auch in den meisten neueren Schulbüchern findet man oft auch die Definition einer sog. relativistischen Masse $m_{\mathrm{rel}} = m\gamma$. 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].

und folglich

$$(p^\mu) = m\gamma \begin{pmatrix} c \\ \dot{\vec{x}} \end{pmatrix}. \tag{4.4.12}$$

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

$$\vec{p} = m\gamma\dot{\vec{x}} = m\dot{\vec{x}}[1 + \mathcal{O}(\beta^2)] \simeq m\dot{\vec{x}}, \tag{4.4.13}$$

d.h. der gewöhnliche nichtrelativistische Impuls. Die zeitliche Komponente ist

$$p^0 = mc\gamma = mc\left[1 + \frac{\beta^2}{2} + \mathcal{O}(\beta^4)\right] \simeq \frac{1}{c}\left(mc^2 + \frac{m}{2}\dot{\vec{x}}^2\right). \tag{4.4.14}$$

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,

$$\mathcal{E} = p^0 c = mc^2\gamma = E_0 + E_{\text{kin}}, \quad \text{mit} \quad E_0 = mc^2, \quad E_{\text{kin}} = mc^2(\gamma - 1) \tag{4.4.15}$$

als die **relativistische Energie** des Teilchens zu definieren[4]. 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:

$$p'^\mu = \Lambda^\mu{}_\nu p^\nu. \tag{4.4.16}$$

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)

$$\underline{p} \cdot \underline{p} = \left(\frac{\mathcal{E}}{c}\right)^2 - \vec{p}^2 = m^2\gamma^2c^2(1 - \beta^2) = m^2c^2. \tag{4.4.17}$$

Daraus ergibt sich sofort die relativistische Beziehung zwischen Energie und Impuls zu

$$\mathcal{E} = c\sqrt{m^2c^2 + \vec{p}^2}. \tag{4.4.18}$$

Im Folgenden ist es auch bequem, die dimensionslose **Vierergeschwindigkeit**

$$\underline{u} = \frac{1}{c}\frac{\mathrm{d}\underline{x}}{\mathrm{d}\tau} = \begin{pmatrix} \gamma \\ \gamma\vec{\beta} \end{pmatrix} \tag{4.4.19}$$

einzuführen. Dann ist (*nachrechnen!*)

$$\underline{p} = mc\underline{u} \quad \text{und} \quad \underline{u} \cdot \underline{u} = 1. \tag{4.4.20}$$

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

$$\vec{\beta} = \frac{1}{c}\frac{\mathrm{d}\vec{x}}{\mathrm{d}\vec{t}} = \frac{\vec{u}}{u^0} = \frac{mc\vec{u}}{mcu^0} \Rightarrow \vec{\beta} = \frac{\vec{p}}{\mathcal{E}/c} \tag{4.4.21}$$

[4] Wir 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.

ausdrücken können.
Jetzt können wir auch Newtons zweites Postulat relativistisch erweitern, indem wir schreiben

$$\frac{\mathrm{d}}{\mathrm{d}\tau}p^\mu = mc\frac{\mathrm{d}}{\mathrm{d}\tau}u^\mu = m\frac{\mathrm{d}^2}{\mathrm{d}\tau^2}x^\mu = K^\mu, \tag{4.4.22}$$

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 Viergeschwindigkeit (4.4.10) gilt nämlich $\underline{u}\cdot\underline{u} = 1$. Leitet man diese Bedingung nach τ ab, folgt

$$\frac{\mathrm{d}\underline{u}}{\mathrm{d}\tau}\cdot\underline{u} = 0, \tag{4.4.23}$$

d.h. multipliziert man (4.4.22) mit u^μ, erhält man die Bedingung

$$\underline{K}\cdot\underline{u} = \eta_{\mu\nu}K^\mu u^\nu = 0 \tag{4.4.24}$$

für die Minkowski-Kraft. Damit lässt sich auch die Zeitkomponente der Bewegungsgleichung (4.4.22) interpretieren, denn aus (4.4.24) folgt

$$\eta_{\mu\nu}K^\mu u^\nu = u^0K^0 - \vec{u}\cdot\vec{k} = 0 \Rightarrow u^0K^0 = \vec{u}\cdot\vec{k}. \tag{4.4.25}$$

Damit folgt für die Zeitkomponente von (4.4.22)

$$mc\frac{\mathrm{d}u^0}{\mathrm{d}\tau} = K^0 = \frac{\vec{u}}{u^0}\cdot\vec{K} = \vec{\beta}\cdot\vec{K}. \tag{4.4.26}$$

Wegen $\mathrm{d}\tau = \sqrt{1-\vec{\beta}^2}\mathrm{d}t = \mathrm{d}t/u^0 = \mathrm{d}t/\gamma$ erhalten wir schließlich

$$mc\frac{\mathrm{d}u^0}{\mathrm{d}t} = \vec{\beta}\cdot\frac{\vec{K}}{\gamma}. \tag{4.4.27}$$

Multiplizieren wir dies mit c und schreiben $\vec{K} = u^0\vec{F} = \gamma\vec{F}$ erhalten wir wegen (4.4.15)

$$\frac{\mathrm{d}\mathcal{E}}{\mathrm{d}t} = \frac{\mathrm{d}\mathcal{E}_{\mathrm{kin}}}{\mathrm{d}t} = \vec{v}\cdot\vec{F}. \tag{4.4.28}$$

Für den räumlichen Teil von (4.4.22) erhalten wir die nicht manifest kovariante Form der Bewegungsgleichung

$$\frac{\mathrm{d}}{\mathrm{d}t}\vec{p} = \frac{1}{\gamma}\frac{\mathrm{d}}{\mathrm{d}\tau}\vec{p} = \frac{1}{\gamma}\vec{K} = \vec{F}, \tag{4.4.29}$$

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).

Wir bemerken noch, dass das in der Vorlesung Theoretische Physik 1 eingeführte **Hamiltonsche Wirkungsprinzip** und das damit verbundene **Noether-Theorem** einen weiteren eleganten Zugang zur Definition des relativistischen Impulses und der Energie eines freien Teilchens und zur Formulierung relativistischer Kraftgesetze darstellt.

Dazu knüpfen wir an die Konstruktion der Lagrange-Funktion für ein einzelnes freies Teilchen in der Newtonschen Mechanik an [Hee18b]. Von den fundamentalen Symmetrien der Galilei-Newton-Raumzeit gelten in der speziellen Relativitätstheorie die Homogenität und Isotropie des Raumes unverändert sowie die Homogenität der Zeit ungeändert weiter, d.h. wir können aus der Analyse vermöge des Noether-Theorems in Abschnitt 3.5 übernehmen, dass die Lagrange-Funktion die Form

$$L = L(\dot{\vec{x}}^2) \tag{4.4.30}$$

besitzen muss, wobei wir uns hier des aus der Newtonschen Mechanik bekannten Lagrange-Formalismusses des Wirkungsprinzips bedienen, wo wir die Trajektorie des Teilchens als Funktion der Koordinatenzeit t beschreiben und entsprechend $\dot{\vec{x}} = \mathrm{d}\vec{x}/\mathrm{d}t$ die übliche nichtkovariante Dreiergeschwindigkeit bezeichnet.

Während wir für die endgültige Bestimmung der Form der Lagrange-Funktion in der Newtonschen Physik nun noch die Invarianz unter Galilei-Boosts verwendet haben, was dann auf $L = T = m\dot{x}^2/2$ für die Lagrangefunktion des freien Teilchens geführt hat, müssen wir nun freilich einen Lorentz-Boost in beliebiger Richtung verwenden. Für die Analyse mit Hilfe des Noether-Theorems benötigen wir wie in [Hee18b] gezeigt zunächst nur eine infinitesimale Transformation, d.h.

$$x'^{\mu} = (\delta^{\mu}_{\nu} + \epsilon\omega^{\mu}{}_{\nu})x^{\nu}. \tag{4.4.31}$$

Die Bedingung an die Matrix $\hat{\omega}$ erhalten wir aus der Entwicklung von (4.1.50) bis zu linearen Ordnung in ϵ:

$$\eta_{\mu\nu}(\delta^{\mu}_{\rho} + \epsilon\omega^{\mu}{}_{\rho})(\delta^{\nu}_{\sigma} + \epsilon\omega^{\nu}{}_{\sigma}) = \eta_{\rho\sigma} + \epsilon\omega_{\sigma\rho} + \epsilon\omega_{\rho\sigma} + \mathcal{O}(\epsilon^2) \stackrel{!}{=} \eta_{\rho\sigma}. \tag{4.4.32}$$

Daraus folgt, dass

$$\omega_{\rho\sigma} = \eta_{\mu\rho}\omega^{\mu}{}_{\sigma} = -\omega_{\sigma\rho}. \tag{4.4.33}$$

Wir sind nun nur an den Boosts interessiert und nicht an Drehungen der drei räumlichen kartesischen Koordinatenachsen, d.h. wir setzen $\omega^{j}{}_{k} = 0$ für $j,k \in \{1,2,3\}$. Hier und im Folgenden laufen lateinische Indizes stets von 1 bis 3. Dann bleiben nur noch die drei unabhängigen Parameter $\omega_{0j} = -\omega_{j0}$ bzw. $\omega^{j}{}_{0}$. Weiter ist $\omega^{0}{}_{j} = \eta^{0\rho}\omega_{\rho j} = \omega_{0j} = -\omega_{j0} = +\omega^{j}{}_{0}$. Im Folgenden schreiben wir $(\omega^{j}{}_{0}) = \vec{\omega}$. Außerdem ist $\omega^{0}{}_{0} = \eta^{0\rho}\omega_{\rho 0} = \omega_{00} = 0$.

In [Hee18b] haben wir die infinitesimale Transformation in der Form

$$t' = t + \epsilon T(t,\vec{x},\dot{\vec{x}}), \quad \vec{x}' = \vec{x} + \epsilon\vec{Q}(t,\vec{x},\dot{\vec{x}}) \tag{4.4.34}$$

geschrieben. Mit (4.4.31) und $x^0 = ct$ bzw. $x'^0 = ct'$ ergibt sich

$$cT = \omega^{0}{}_{\rho}x^{\rho} = \omega^{0}{}_{j}x^{j} = \vec{\omega}\cdot\vec{x}, \quad (Q^j) = (\omega^{j}{}_{\rho}x^{\rho}) = (\omega^{j}{}_{0}x^{0}) = \vec{\omega}ct. \tag{4.4.35}$$

Dann ergab sich, dass eine Symmetrie der Wirkung vorliegt, wenn die Lagrangefunktion die Bedingung

$$\vec{Q}\cdot\frac{\partial L}{\partial\vec{Q}} + \left(\frac{\mathrm{d}\vec{Q}}{\mathrm{d}t} - \frac{\mathrm{d}T}{\mathrm{d}t}\dot{\vec{x}}\right)\cdot\frac{\partial L}{\partial\dot{\vec{x}}} + L\frac{\mathrm{d}T}{\mathrm{d}t} + T\frac{\partial L}{\partial t} + \frac{\mathrm{d}}{\mathrm{d}t}\tilde{\Omega}(t,\vec{x}) = 0 \tag{4.4.36}$$

erfüllt. Dabei darf $\tilde{\Omega}$ eine beliebige Funktion von t und $\vec{x}$ sein. Gemäs (4.4.35) ist

$$\frac{\mathrm{d}T}{\mathrm{d}t} = \frac{1}{c}\vec{\omega}\cdot\dot{\vec{x}}, \quad \frac{\mathrm{d}\vec{Q}}{\mathrm{d}t} = c\vec{\omega}. \tag{4.4.37}$$

Da wir weiter aus den übrigen Symmtrien bereits die Form (4.4.30) für die Lagrange-Funktion hergeleitet haben, woraus $\partial L/\partial \vec{Q} = \partial L/\partial t = 0$ folgt, und wir versuchen, (4.4.36) mit dem Ansatz $\tilde{\Omega} = 0$ zu erfüllen, ergibt sich nach einigen einfachen Umformungen aus (4.4.36) die Bedingung dafür, dass die Wirkung symmetrisch unter Lorentz-Boost-Transformationen ist

$$\frac{1}{c}\vec{\omega}\cdot\dot{\vec{x}}[2(c^2-\dot{\vec{x}}^2)L'(\dot{\vec{x}}^2)+L]=0. \tag{4.4.38}$$

Da $\vec{\omega}$ beliebig gewählt werden kann, muss der Ausdruck in der eckigen Klammer verschwinden. Das bedeutet

$$\frac{L'}{L}=\frac{1}{L}\frac{\mathrm{d}L}{\mathrm{d}\dot{\vec{x}}^2}=\frac{1}{2(\dot{\vec{x}}^2-c^2)}. \tag{4.4.39}$$

Wir beachten nun noch, dass für die Bewegung eines relativistischen Teilchens aus den oben erörterten Gründen stets $|\dot{\vec{x}}| < c$ sein muss. Damit können wir (4.4.39) eindeutig über das Intervall $\dot{\vec{x}}^2 \in (0, \beta^2 c^2)$ integrieren:

$$\ln\left(\frac{L}{L_0}\right)=\frac{1}{2}\ln\left(\frac{c^2-\beta^2c^2}{c^2}\right)=\ln\sqrt{1-\beta^2}. \tag{4.4.40}$$

Damit ist also

$$L=L_0\sqrt{1-\beta^2}. \tag{4.4.41}$$

Die Integrationskonstante L_0 bestimmen wir aus dem Newtonschen Grenzfall, der gilt, wenn $|\beta| \ll 1$, also $|\dot{\vec{x}}| \ll c$ ist. Dann ergibt die Taylor-Entwicklung von (4.4.41)

$$L=L_0\left(1-\frac{1}{2}\beta^2\right)=L_0-\frac{L_0}{2c^2}\dot{\vec{x}}^2. \tag{4.4.42}$$

Für die Newtonsche Mechanik ist $L = T + L_0 = m\dot{\vec{x}}^2/2 + L_0$, wobei wir ausnutzen, dass wir zu L eine beliebige Konstante addieren können, ohne dass dies etwas and den durch die Euler-Lagrange-Gleichungen gegebenen Bewegungsgleichungen ändert. Durch Vergleich mit (4.4.42) ergibt sich $L_0 = -mc^2$. Damit erhalten wir die Lagrangefunktion für ein freies relativistisches Teilchen:

$$L=-mc^2\sqrt{1-\beta^2}=-mc^2\sqrt{1-\frac{\dot{\vec{x}}^2}{c^2}} \tag{4.4.43}$$

und der kanonische Impuls

$$\vec{p}=\frac{\partial L}{\partial\dot{\vec{x}}}=\frac{m\dot{\vec{x}}}{\sqrt{1-\beta^2}}=m\gamma\dot{\vec{x}}=m\frac{\mathrm{d}\vec{x}}{\mathrm{d}\tau}, \tag{4.4.44}$$

wobei τ wieder die Eigenzeit des Teilchens bezeichnet. Damit erhalten wir die räulichen Komponenten des relativistischen Impulses auch aus dem Hamiltonschen Prinzip. Der kanonische Impuls ist zugleich die Erhaltungsgröße, die sich über das Noether-Theorem aus der Symmetrie unter räumlichen Translationen ergibt. Auch dies ist eine weitere formale Rechtfertigung dafür, dass (4.4.44) der korrekte relativistische Impuls eines Teilchens ist.

Schließlich können wir auch den Ausdruck für die Energie herleiten, die wir als diejenige Erhaltungsgröße definieren, die sich aus dem Noether-Theorem für die Symmetrie unter zeitlichen Translationen ergibt. Wie wir aus [Hee18b] wissen, ist diese Erhaltungsgröße die Hamilton-Funktion. Es ergibt sich nach einigen einfachen Umformungen (*Übung!*)

$$H=\vec{p}\cdot\dot{\vec{x}}-L=\frac{mc^2}{\sqrt{1-\dot{\vec{v}}^2/c^2}}=mc^2\gamma=c\sqrt{m^2c^2+\vec{p}^2}. \tag{4.4.45}$$

Auch dies stimmt mit den obigen Überlegungen überein.
Wir bemerken nun noch, dass die Lagrange-Funktion (4.4.43) nicht invariant unter Lorentz-Boosts ist. Das liegt daran, dass sich bei Lorentz-Boosts die Zeit transformiert. Allerdings ist das Wirkungsfunktional

$$S[\vec{x}] = \int_{t_1}^{t_2} \mathrm{d}t L = -mc^2 \int_{t_1}^{t_2} \mathrm{d}t \sqrt{1 - \frac{\dot{\vec{x}}^2}{c^2}} = -mc^2 \int_{t_1}^{t_2} \mathrm{d}\tau \tag{4.4.46}$$

invariant unter Lorentz-Transformationen. Dabei sind $\vec{x}(t)$ beliebige Kurven, für die die Anfangs- und Endpunkte fest vorgegeben sind: $\vec{x}(t_1) = \vec{x}_1$ und $\vec{x}(t_2) = \vec{x}_2$.
Das Integral ist nun aber auch invariant unter beliebigen Umparametrisierungen der Weltlinie, denn ist die Weltlinie durch $\underline{x}(\lambda)$ mit einem beliebigen Parameter λ gegeben, so dass $x^0(\lambda_1) = c t_1$, $\vec{x}(\lambda_1) = \vec{x}_1$ und $x^0(\lambda_2) = c t_2$, $\vec{x}(\lambda_2) = \vec{x}_2$ ist, können wir (4.4.46) in die Form

$$\begin{aligned} S &= -mc^2 \int_{\lambda_1}^{\lambda_2} \mathrm{d}\lambda \frac{\mathrm{d}t}{\mathrm{d}\lambda} \sqrt{1 - (\mathrm{d}_t \vec{x})^2/c^2} \\ &= -mc^2 \int_{\lambda_1}^{\lambda_2} \mathrm{d}\lambda \frac{1}{c} \sqrt{(\mathrm{d}_\lambda x^0)^2 - (\mathrm{d}_\lambda \vec{x})^2} \\ &= -mc \int_{\lambda_1}^{\lambda_2} \mathrm{d}\lambda \sqrt{\eta_{\mu\nu} \frac{\mathrm{d}x^\mu}{\mathrm{d}\lambda} \frac{\mathrm{d}x^\nu}{\mathrm{d}\lambda}} \end{aligned} \tag{4.4.47}$$

bringen. Dies ist nun offensichtlich manifest kovariant, denn wir können λ als einen Skalar unter Lorentz-Transformationen interpretieren.
Offensichtlich können wir nun die Bewegungsgleichungen auch in dieser manifest kovarianten Form aus dem sog. erweiterten Hamiltonschen Prinzip herleiten, d.h. wir verlangen, dass die Weltlinie des Teilchens dadurch charakterisiert ist, dass S extremal bzgl. beliebiger Variationen δx^μ mit festgehaltenen Endpunkten $\delta x^\mu(\lambda_1) = \delta x^\mu(\lambda_2) = 0$ ist. Wir erhalten dann wieder die entsprechenden Euler-Lagrange-Gleichungen

$$\frac{\mathrm{d}}{\mathrm{d}\lambda} \frac{\partial L}{\partial \dot{x}^\mu} = \frac{\partial L}{\partial x^\mu}, \tag{4.4.48}$$

wobei hier der Punkt die Ableitung nach dem Weltlinienparameter λ bedeuten soll. Nun ist

$$\frac{\partial L}{\partial \dot{x}^\mu} = -mc \frac{\eta_{\mu\nu} \dot{x}^\nu}{\sqrt{\eta_{\rho\sigma} \dot{x}^\rho \dot{x}^\sigma}} = -m \frac{\mathrm{d}x_\mu}{\mathrm{d}\tau} = -p_\mu. \tag{4.4.49}$$

Dabei haben wir im letzten Schritt den Ausdruck wieder mit Hilfe der Eigenzeit des Teilchens, die durch

$$\mathrm{d}\tau = \frac{1}{c} \mathrm{d}\lambda \sqrt{\eta_{\rho\sigma} \dot{x}^\rho \dot{x}^\sigma} \tag{4.4.50}$$

definiert werden kann.
Die Euler-Lagrange-Gleichungen liefern in diesem Fall

$$\dot{p}_\mu = \frac{\mathrm{d}\tau}{\mathrm{d}\lambda} \frac{\mathrm{d}p_\mu}{\mathrm{d}\tau} = 0 \Rightarrow \frac{\mathrm{d}p^\mu}{\mathrm{d}\tau} = 0. \tag{4.4.51}$$

Wir erhalten also, wie für ein freies Teilchen in einem Inertialsystem durch das Trägheitsprinzip zu erwarten ist, $p^\mu = \text{const}$ und damit auch $\vec{v} = \mathrm{d}\vec{x}/\mathrm{d}t = c\vec{p}/p^0 = \text{const}$.

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 $(2 \to 2)$**-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 Zerfallskinematik 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.

$$\underline{p} = \underline{p}_1' + \underline{p}_2'. \tag{4.5.1}$$

Weiter gelten die Energie-Impulsbeziehungen für jeden dieser Viererimpulse. Diese können wir am einfachsten in der kovarianten Form (4.4.17) berücksichtigen:

$$\underline{p} \cdot \underline{p} = M^2c^2, \quad \underline{p}_1' \cdot \underline{p}_1' = m_1^2c^2, \quad \underline{p}_2' \cdot \underline{p}_2' = m_2^2c^2. \tag{4.5.2}$$

Gehen wir nun in das Ruhsystem des zerfallenden Teilchens, ist

$$(p^\mu) = \begin{pmatrix} Mc \\ 0 \end{pmatrix}, \quad (p_1'^\mu) = \begin{pmatrix} \mathcal{E}_1/c \\ \vec{p}\,' \end{pmatrix}, \quad p_2'^\mu = \begin{pmatrix} \mathcal{E}_2/c \\ -\vec{p}\,' \end{pmatrix}. \tag{4.5.3}$$

Offenbar ist durch diese Schreibweise der räumliche Teil von (4.5.1), also die Impulserhaltung, identisch erfüllt. Die Zeitkomponente ergibt

$$\mathcal{E}_1 + \mathcal{E}_2 = Mc^2. \tag{4.5.4}$$

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

$$\underline{p} \cdot \underline{p} = M^2c^2 = \underline{p}_1 \cdot \underline{p}_1' + \underline{p}_2' \cdot \underline{p}_2' + 2\underline{p}_1' \cdot \underline{p}_2' = (m_1^2 + m_2^2)c^2 + 2\underline{p}_1' \cdot \underline{p}_2'. \tag{4.5.5}$$

Im obigen Bezugssystem, wo das zerfallende Teilchen vor dem Zerfall ruht, gilt

$$\underline{p}_1' \cdot \underline{p}_2' = \frac{\mathcal{E}_1\mathcal{E}_2}{c^2} + \vec{p}\,'^2. \tag{4.5.6}$$

Mit (4.5.3) folgt

$$\underline{p}_1' \cdot \underline{p}_2' = \frac{\mathcal{E}_1(Mc^2 - \mathcal{E}_1)}{c^2} + \vec{p}\,'^2 = M\mathcal{E}_1 - (m_1^2c^2 + \vec{p}\,'^2) + \vec{p}\,'^2 = M\mathcal{E}_1 - m_1^2c^2. \tag{4.5.7}$$

Setzen wir dies in (4.5.5) ein, erhalten wir

$$M^2c^2 = (m_1^2 + m_2^2)c^2 + 2(M\mathcal{E}_1 - m_1^2c^2) = (m_2^2 - m_1^2)c^2 + 2M\mathcal{E}_1 \tag{4.5.8}$$

bzw.

$$\mathcal{E}_1 = \frac{(M^2 + m_1^2 - m_2^2)c^2}{2M} \tag{4.5.9}$$

Mit der Energie-Impulsbeziehung erhalten wir daraus schließlich (*nachrechnen!*)

$$\begin{aligned}
\vec{p}'^2 &= \frac{\mathcal{E}_1^2}{c^2} - m_1^2 c^2 = \frac{(M^2 + m_1^2 - m_2^2)^2 c^2}{4M^2} - m_1^2 c^2 \\
&= \frac{c^2}{4M^2}\left[M^4 - 2M^2(m_1^2 + m_2^2) + (m_1^2 - m_2^2)^2\right] \\
&= \frac{c^2}{4M^2}\left[M^4 - 2M^2(m_1^2 + m_2^2) + (M_1^2 + m_2^2)^2 - 4m_1^2 m_2^2\right] \\
&= \frac{c^2}{4M^2}\left[(M^2 - m_1^2 - m_2^2)^2 - 4m_1^2 m_2^2\right] \\
&= \frac{c^2}{4M^2}(M^2 - m_1^2 - M_2^2 - 2m_1 m_2)(M^2 - m_1^2 - m_2^2 + 2m_1 m_2) \\
&= \frac{[M^2 - (m_1 + m_2)^2][M^2 - (m_1 - m_2)^2]c^2}{4M^2}.
\end{aligned} \tag{4.5.10}$$

Da $\vec{p}'^2 \geq 0$ sein muss, kann der Zerfall nur unter der Bedingung stattfinden, dass

$$M \geq m_1 + m_2 \tag{4.5.11}$$

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.

$$\underline{p}_1 + \underline{p}_2 = \underline{p}_3 + \underline{p}_4. \tag{4.5.12}$$

Wir können nun aus diesen vier Impulsen drei Invarianten, die **Mandelstam-Variablen**

$$\begin{aligned}
s &= (\underline{p}_1 + \underline{p}_2)^2 = (\underline{p}_3 + \underline{p}_4)^2, \\
t &= (\underline{p}_1 - \underline{p}_3)^2 = (\underline{p}_2 - \underline{p}_4)^2, \\
u &= (\underline{p}_1 - \underline{p}_4)^2 = (\underline{p}_2 - \underline{p}_3)^2
\end{aligned} \tag{4.5.13}$$

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!*)

$$s + t + u = (m_1^2 + m_2^2 + m_3^2 + m_4^2)c^2. \tag{4.5.14}$$

Es sind nun zwei Bezugssysteme besonders ausgezeichnet, und zwar das **Schwerpunktssystem**, wo $\vec{p}_1^{(\mathrm{SPS})} = -\vec{p}_2^{(\mathrm{SPS})}$ gilt, sowie das sog. **Laborsystem**, in dem im Ausgangszustand ein Teilchen ruht, also z.B. $\vec{p}_2^{(\mathrm{LS})} = 0$ 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 $\underline{p}_2^2 = m_2^2c^2 > 0$ ein zeitartiger Vektor. Ein Lorentz-Boost mit $\vec{\beta} = \vec{p}_2/p_2^0 = \vec{p}_2/\sqrt{m_2^2c^2+\vec{p}_2^2}$, woraus sich offensichtlich $|\vec{\beta}| < 1$ ergibt, liefert dann (*nachrechnen!*) in der Tat $(\overline{p}_2^\mu) = (m_2c, 0, 0, 0)$.

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

$$p_1^0 + p_2^0 = \sqrt{m_1^2c^2+\vec{p}_1^2} + \sqrt{m_2^2c^2+\vec{p}_2^2} > |\vec{p}_1| + |\vec{p}_2| \geq |\vec{p}_1+\vec{p}_2| \Rightarrow s = (p_1^0+p_2^0)^2 - (\vec{p}_1+\vec{p}_2)^2 > 0. \tag{4.5.15}$$

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 Schwerpunktssystem

Wir beginnen mit der näheren Analyse der Kinematik im **Schwerpunktssystem**. Dann gilt

$$(p_1^\mu) = \begin{pmatrix} \mathscr{E}_1/c \\ \vec{p} \end{pmatrix}, \quad (p_2^\mu) = \begin{pmatrix} \mathscr{E}_2/c \\ -\vec{p} \end{pmatrix}, \quad (p_3^\mu) = \begin{pmatrix} \mathscr{E}_3/c \\ \vec{p}' \end{pmatrix}, \quad (p_4^\mu) = \begin{pmatrix} \mathscr{E}_4/c \\ -\vec{p}' \end{pmatrix}, \tag{4.5.16}$$

und folglich ergibt sich aus (4.5.13)

$$\sqrt{s} = \frac{\mathscr{E}_1+\mathscr{E}_2}{c} = \frac{\mathscr{E}_3+\mathscr{E}_4}{c}, \tag{4.5.17}$$

d.h. $c\sqrt{s}$ ist die Gesamtenergie der Teilchen vor und nach dem Stoß im Schwerpunktssystem.

Wir können nun die Energien und Impulsbeträge im Schwerpunktssystem mit Hilfe der Mandelstam-Variablen ausdrücken. Zunächst ist

$$\begin{aligned} s &= (\underline{p}_1+\underline{p}_2)^2 = (m_1^2+m_2^2)c^2 + 2\left(\frac{\mathscr{E}_1\mathscr{E}_2}{c^2} + \vec{p}^2\right) \\ &\overset{(4.5.17)}{=} (m_1^2+m_2^2)c^2 + 2\left(\frac{\mathscr{E}_1(c\sqrt{s}-\mathscr{E}_1)}{c^2} + \vec{p}^2\right) \\ &= (m_2^2-m_1^2)c^2 + \frac{2\mathscr{E}_1\sqrt{s}}{c}. \end{aligned} \tag{4.5.18}$$

Dabei haben wir im letzten Schritt die Energie-Impulsbeziehung $\mathscr{E}_1^2/c^2 - \vec{p}^2 = m_1^2c^2$ verwendet. Aus (4.5.18) erhält man schließlich

$$\frac{\mathscr{E}_1}{c} = \frac{s+(m_1^2-m_2^2)c^2}{2\sqrt{s}}. \tag{4.5.19}$$

Eliminieren wir in (4.5.18) $\mathscr{E}_1$ zugunsten von $\mathscr{E}_2$, erhalten wir

$$\frac{\mathscr{E}_2}{c} = \frac{s+(m_2^2-m_1^2)c^2}{2\sqrt{s}}. \tag{4.5.20}$$

Die Energien im Endzustand ergeben sich durch Ersetzen der Indizes 1 und 2 durch 3 bzw. 4:

$$\begin{aligned} \frac{\mathscr{E}_3}{c} &= \frac{s+(m_3^2-m_4^2)c^2}{2\sqrt{s}}, \\ \frac{\mathscr{E}_4}{c} &= \frac{s+(m_4^2-m_3^2)c^2}{2\sqrt{s}}. \end{aligned} \tag{4.5.21}$$

Die Impulsbeträge im Anfangs- bzw. Endzustand ergeben sich daraus mit Hilfe der Energie-Impuls-Beziehung nach einigen Umformungen (*nachrechnen*)

$$\begin{aligned}\vec{p}^{\,2} &= \frac{\mathscr{E}_1^2}{c^2} - m_1^2c^2 = \frac{1}{4s}[s-(m_1+m_2)^2c^2][s-(m_1-m_2)^2c^2],\\ \vec{p}^{\,\prime 2} &= \frac{\mathscr{E}_3^2}{c^2} - m_3^2c^2 = \frac{1}{4s}[s-(m_3+m_4)^2c^2][s-(m_3-m_4)^2c^2]\end{aligned} \tag{4.5.22}$$

4.5.4 Kinematik im Laborsystem

Im Laborsystem haben wir

$$p_1 = \begin{pmatrix}\mathscr{E}_1/c\\ \vec{p}_1\end{pmatrix}, \quad p_2 = \begin{pmatrix}m_2c\\ 0\end{pmatrix}. \tag{4.5.23}$$

Damit ist

$$s = \left(\frac{\mathscr{E}_1 + m_2c^2}{c}\right)^2 - \vec{p}_1^{\,2} = \left(\frac{\mathscr{E}_1 + m_2c^2}{c}\right)^2 + m_1^2c^2 - \frac{\mathscr{E}_1^2}{c^2} \tag{4.5.24}$$

Nach einigen Umformungen erhält man

$$\mathscr{E}_1 = \frac{s-(m_1^2+m_2^2)c^2}{2m_2}. \tag{4.5.25}$$

Weiter folgt (*nachrechnen!*)

$$\vec{p}_1^{\,2} = \frac{\mathscr{E}_1^2}{c^2} - m_1^2c^2 = \frac{[s-(m_1+m_2)^2c^2][s-(m_1-m_2)^2c^2]}{4m_2^2c^2}, \tag{4.5.26}$$

und der Vergleich mit (4.5.22) liefert

$$\vec{p}_{\mathrm{SPS}}^{\,2} = \frac{m_2^2c^2}{s}\vec{p}_{\mathrm{LS}}^{\,2}. \tag{4.5.27}$$

Für die Mandelstam-Variable t gilt (*nachrechnen!*)

$$t = (p_2-p_4)^2 = \left(\frac{\mathscr{E}_4}{c} - m_2c\right)^2 - \vec{p}_4^{\,2} = (m_4^2+m_2^2)c^2 - 2m_2\mathscr{E}_4 \tag{4.5.28}$$

bzw.

$$\mathscr{E}_4 = \frac{(m_2^2+m_4^2)c^2 - t}{2m_2}. \tag{4.5.29}$$

Ebenso erhält man aus der Definition der Mandelstam-Variablen u

$$\mathscr{E}_3 = \frac{(m_2^2+m_3)^2c^2 - u}{2m_2} \tag{4.5.30}$$

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** $\mathscr{E}_0 = mc^2$, wenn wir die unter der Masse, wie üblich, die **invariante Masse** verstehen und den Energienullpunkt so legen, dass die Energie $\mathscr{E} = \mathscr{E}_0 + \mathscr{E}_{\mathrm{kin}} = mc^2\gamma = c\sqrt{m^2c^2 + \vec{p}^{\,2}} = p^0c$ 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 Körper umgewandelt. Dabei gilt allerdings immer noch die Erhaltung des Gesamtimpulses.

Es ist wieder am einfachsten, die Kinematik im Schwerpunktssystem zu betrachten, d.h. die beiden Körper besitzen in diesem Bezugssystem vor dem Stoß entgegengesetzt gleiche Impulse $\vec{p}_2 = -\vec{p}_1 = -\vec{p}$ und nach dem Stoß liegt ein einziger ruhender Körper vor. Die Energie-Impulsbilanz dieses Stoßes ergibt dann

$$(p_1^\mu + p_2^\mu) = \begin{pmatrix} (\mathscr{E}_1 + \mathscr{E}_2)/c \\ 0 \end{pmatrix} = \begin{pmatrix} Mc \\ 0 \end{pmatrix}. \tag{4.6.1}$$

Es ist also

$$M = \frac{\mathscr{E}_1 + \mathscr{E}_2}{c^2} = m_1 + m_2 + \frac{\mathscr{E}_{1\mathrm{kin}} + \mathscr{E}_{2\mathrm{kin}}}{c^2}. \tag{4.6.2}$$

Da die kinetische Energie des ruhenden Teilchens im Endzustand 0 ist, bedeutet dies, dass die gesamte kinetische Energie im Anfangszustand in Wärmenergie $\Delta Q = \mathscr{E}_{1\mathrm{kin}} + \mathscr{E}_{2\mathrm{kin}}$ 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 $m_{\pi_0} c^2 = 134{,}98\,\mathrm{MeV}$). 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 $1\,\mathrm{eV} = |e| \cdot 1\,\mathrm{V}$. Mit der Elektronenladung (in SI-Einheiten) $e = -1{,}602 \cdot 10^{-19}\,\mathrm{C}$ folgt daraus, dass $1\,\mathrm{eV} = 1{,}602 \cdot 10^{-19}\,\mathrm{J}$ entspricht. Dies zeigt, warum man in der Teilchen- und Kernphysik gerne die Einheit $\mathrm{MeV} = 10^6\,\mathrm{eV}$ oder $\mathrm{GeV} = 10^9\,\mathrm{eV}$ 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

$$\mathscr{E}_\gamma = c|\vec{p}_\gamma|. \tag{4.6.3}$$

Daraus ergibt sich für die Geschwindigkeit in jedem Inertialsystem

$$|\beta_\gamma| = \frac{c|\vec{p}_\gamma|}{\mathscr{E}_\gamma} = 1. \tag{4.6.4}$$

Daraus ist ersichtlich, dass für solche „masselosen Teilchen“ wegen $\gamma \to \infty$ für $|\beta| \to 1$ 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 Ruhsystem des zerfallenden Pions) mit $m_1 = m_2 = 0$ verwenden, um den Photonenimpuls zu

$$|\vec{p}_\gamma| = \frac{Mc}{2} \tag{4.6.5}$$

zu bestimmen. Wir haben hier den Extremfall vorliegen, dass die invarianten Massen der Zerfallsprodukte 0 sind aber die Masse im Ausgangszustand durch

$$Mc^2 = \mathscr{E}_1' + \mathscr{E}_2' = 2c|\vec{p}_\gamma| \tag{4.6.6}$$

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 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

$$F'^{\mu\nu}(x'^\rho) = \Lambda^\mu{}_\alpha \Lambda^\nu{}_\beta F^{\alpha\beta}(x^\gamma) \quad \text{mit} \quad x^\gamma = (\Lambda^{-1})^\gamma{}_\rho x'^\rho \tag{4.7.1}$$

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

$$m\frac{\mathrm{d}^2 x^\mu}{\mathrm{d}\tau^2} = K^\mu = qF^{\mu\nu}(x^\rho)\frac{\mathrm{d}x_\nu}{\mathrm{d}\tau} = qF^{\mu\nu}cu_\nu \quad \text{mit} \quad x_\nu = \eta_{\mu\nu}x^\mu, \quad \underline{u} = \frac{1}{c}\frac{\mathrm{d}\underline{x}}{\mathrm{d}\tau} \tag{4.7.2}$$

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 $\vec{B}$:

$$(F^{\mu\nu}) = \begin{pmatrix} 0 & -E^1/c & -E^2/c & -E^3/c \\ E^1/c & 0 & -B^3 & B^2 \\ E^2/c & B^3 & 0 & -B^1 \\ E^3/c & -B^2 & B^1 & 0 \end{pmatrix}. \tag{4.7.3}$$

Um zu sehen, dass diese Wahl der Anordnung der Feldkomponenten einen Sinn ergibt, schreiben wir die Minkowski-Kraft (4.7.2) explizit hin

$$(K^\mu) = qF^{\mu\nu}cu_\nu = q\begin{pmatrix} \vec{E}\cdot \mathrm{d}\vec{u} \\ u_0\vec{E} + \mathrm{d}\vec{x}/\mathrm{d}\tau \times \vec{B} \end{pmatrix}. \tag{4.7.4}$$

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:

$$\frac{\mathrm{d}}{\mathrm{d}\tau}\underline{p} = m\frac{\mathrm{d}^2}{\mathrm{d}\tau^2}\underline{x} = \underline{K} = q\begin{pmatrix} \vec{E}\cdot\vec{u} \\ u_0\vec{E} + \mathrm{d}\vec{x}/\mathrm{d}\tau \times \vec{B} \end{pmatrix}, \quad u^\mu = \frac{p^\mu}{mc} = \frac{1}{c}\frac{\mathrm{d}}{\mathrm{d}\tau}x^\mu. \tag{4.7.5}$$

Da nun $u^0 = u_0 = \gamma = (1-\beta^2)^{-1/2}$ mit $\vec{\beta} = \vec{v}/c$, $\vec{v} = \mathrm{d}\vec{x}/\mathrm{d}t = c\vec{u}/u^0$, ergibt sich für den räumlichen Anteil der ersten Gleichung

$$\frac{1}{u_0}\frac{\mathrm{d}}{\mathrm{d}\tau}\vec{p} = \frac{\mathrm{d}}{\mathrm{d}t}\vec{p} = q\left(\vec{E} + \vec{v}\times\vec{B}\right), \tag{4.7.6}$$

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

$$\vec{u} = \frac{1}{mc}\vec{p} = \frac{1}{c}\frac{\mathrm{d}}{\mathrm{d}\tau}\vec{x}. \tag{4.7.7}$$

Dividiert man dies durch $u^0 = p^0/(mc)$ erhalten wir

$$\vec{v} = \frac{\mathrm{d}}{\mathrm{d}t}\vec{x} = c\frac{\vec{p}}{p^0}. \tag{4.7.8}$$

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

$$\vec{v}\cdot\frac{\mathrm{d}}{\mathrm{d}t}\vec{p} = q\vec{v}\cdot\vec{E}, \tag{4.7.9}$$

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.

Wir können auf noch elegantere Weise zum gleichen Resultat für die relativistische Bewegungsgleichung eines geladenen Teilchens in einem elektromagnetischen Feld gelangen, wenn wir wieder das Hamiltonsche Prinzip verwenden. Wir wählen zuerst wieder die Koordinatenzeit t als unabhängige Variable. Damit der Beitrag der Wechselwirkungsterme in der Lagrangefunktion zum Wirkungsintegral wieder unabhängig von beliebigen Umparametrisierungen der Weltlinie sowie invariant unter Lorentz-Transformationen ist, machen wir den Ansatz

$$\begin{aligned} L &= -mc^2\sqrt{1-\frac{(\mathrm{d}_t\vec{x})^2}{c^2}} - q\frac{\mathrm{d}x^\mu}{\mathrm{d}t}A_\mu(\underline{x}) \\ &= -mc^2\sqrt{1-\frac{(\mathrm{d}_t\vec{x})^2}{c^2}} - qcA^0 + q\mathrm{d}_t\vec{x}\cdot\vec{A}(\underline{x}), \quad \text{mit} \quad (A^\mu) = \begin{pmatrix} A^0 \\ \vec{A} \end{pmatrix}. \end{aligned} \tag{4.7.10}$$

Die Bewegungsgleichungen zu dieser Lagrange-Funktion erhalten wir wieder mittels der Euler-Lagrange-Gleichungen. Der kanonische Impuls ist

$$\vec{P} = \frac{\partial L}{\partial(\mathrm{d}_t \vec{x})} = \frac{m\mathrm{d}_t \vec{x}}{\sqrt{1-\frac{(\mathrm{d}_t \vec{x})^2}{c^2}}} + q\vec{A} = \vec{p} + q\vec{A}. \tag{4.7.11}$$

Dabei haben wir für den kanonischen Impuls $\vec{P}$ geschrieben, um ihn vom gewöhnlichen „mechanischen Impuls“ $\vec{p} = m\gamma \mathrm{d}_t \vec{x} = m\mathrm{d}_\tau \vec{x}$ zu unterscheiden. Die Bewegungsgleichung ist dann gemäß den Euler-Lagrange-Gleichungen

$$\mathrm{d}_t \vec{P} = \mathrm{d}_t \vec{p} + q(\partial_t \vec{A} + (\dot{\vec{x}} \cdot \vec{\nabla})\vec{A}) = \frac{\partial L}{\partial \vec{x}} = -qc\vec{\nabla}A^0 + q\vec{\nabla}(\mathrm{d}_t \vec{x} \cdot \vec{A}). \tag{4.7.12}$$

Daraus folgt

$$\mathrm{d}_t \vec{p} = -q(\partial_t \vec{A} + c\vec{\nabla}A^0) + q[\vec{\nabla}(\partial_t \vec{x} \cdot \vec{A}) - (\dot{\vec{x}} \cdot \vec{\nabla})\vec{A}]. \tag{4.7.13}$$

Da im letzten Term die Geschwindigkeit $\mathrm{d}_t \vec{x}$ bzgl. des Nabla-Operators, also bzgl. partiellen Ableitungen nach den Komponenten von $\vec{x}$, als konstant zu behandeln ist, wie beim Hamiltonschen Prinzip üblich, können wir den letzten Term noch etwas anders schreiben, denn offenbar gilt nach der „bac-cab-Formel“

$$\mathrm{d}_t \vec{x} \times (\vec{\nabla} \times \vec{A}) = \vec{\nabla}(\mathrm{d}_t \vec{x} \cdot \vec{A}) - [(\mathrm{d}_t \vec{x}) \cdot \vec{\nabla}]\vec{A} \tag{4.7.14}$$

und damit ergibt sich für (4.7.13)

$$\mathrm{d}_t \vec{p} = -q(\partial_t \vec{A} + c\vec{\nabla}A^0) + q(\mathrm{d}_t \vec{x}) \times (\vec{\nabla} \times \vec{A}). \tag{4.7.15}$$

Nun erinnern wir uns, dass wir jedes elektromagnetische Feld mit Hilfe eines skalaren und eines Vektorpotential in der Form

$$\vec{E} = -\vec{\nabla}\Phi - \partial_t \vec{A}, \quad \vec{B} = \vec{\nabla} \times \vec{A} \tag{4.7.16}$$

ausdrücken können. Damit ist (4.7.15) tatsächlich die relativistische Bewegungsgleichung Bewegungsgleichung für ein geladenes Teilchen im Feld, wenn wir $\Phi = cA^0$ setzen, denn dann ist

$$\mathrm{d}_t \vec{p} = q\vec{E} + q(\mathrm{d}_t \vec{x}) \times \vec{B}, \tag{4.7.17}$$

und das stimmt mit der nicht manifest kovariant geschriebenen relativistischen Bewegungsgleichung (4.7.6) überein. Dabei ist zu beachten, dass für den relativistischen (mechanischen) Impuls $\vec{p} = m\gamma \mathrm{d}_t \vec{x}$ gilt.

Es ist nun weiter klar, dass wir auch den manifest kovarianten Formalismus mit einem beliebigen Weltlinienparameter λ und das erweiterte Hamiltonsche Prinzip verwenden können. Die entsprechende Lagrange-Funktion lautet

$$L = -mc\sqrt{\eta_{\rho\sigma}\dot{x}^\rho \dot{x}^\sigma} - q\dot{x}^\nu A_\nu, \tag{4.7.18}$$

wobei der Punkt über einem Symbol wieder die Ableitung nach dem Weltlinienparameter λ bedeuten soll. Die Euler-Lagrange-Gleichungen des erweiterten Hamiltonschen Prinzips ergeben dann

$$\frac{\mathrm{d}}{\mathrm{d}\lambda}\frac{\partial L}{\partial \dot{x}^\mu} = -mc\frac{\mathrm{d}}{\mathrm{d}\lambda}\frac{\dot{x}^\mu}{\sqrt{\eta_{\rho\sigma}\dot{x}^\rho \dot{x}^\sigma}} - q\frac{\mathrm{d}}{\mathrm{d}\lambda}A_\mu = -\dot{p}_\mu - q\dot{x}^\nu \partial_\nu A_\mu. \tag{4.7.19}$$

Dabei bedeutet

$$\partial_\nu A_\mu = \frac{\partial A_\mu}{\partial x^\mu}. \tag{4.7.20}$$

Wir gehen weiter unten in Abschnitt 4.8 noch genauer auf die vierdimensional kovariante Vektoranalysis ein. Mit (4.7.19) erhalten wir für die Euler-Lagrangegleichungen

$$-\dot{p}_\mu - q\dot{x}^\nu \partial_\nu A_\mu = \frac{\partial L}{\partial x^\mu} = -q\dot{x}^\nu \partial_\mu A_\nu \tag{4.7.21}$$

bzw.

$$\dot{p}_\mu = q(\partial_\mu A_\nu - \partial_\nu A_\mu)\dot{x}^\nu. \tag{4.7.22}$$

Dividieren wir diese Gleichung schließlich noch durch $\sqrt{\eta_{\rho\sigma}\dot{x}^\rho \dot{x}^\sigma}$ leads to

$$\frac{\mathrm{d}p_\mu}{\mathrm{d}\tau} = m\frac{\mathrm{d}^2 x_\mu}{\mathrm{d}\tau^2} = qF_{\mu\nu}\frac{\mathrm{d}x^\nu}{\mathrm{d}\tau} \quad \text{with} \quad F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu. \tag{4.7.23}$$

Das stimmt mit (4.7.2) überein, denn es ist

$$F_{0j} = -F_{j0} = \partial_0 A_j - \partial_j A_0 = \frac{1}{c}\partial_t A_j - \partial_j A_0 = -\frac{1}{c}(\partial_t A^j - \partial_j \Phi) = \frac{1}{c}E^j. \tag{4.7.24}$$

Dabei haben wir die oben hergeleitete Beziehcung $A^0 = c\Phi$ verwendet. Für die räumlichen Komponenten von $F_{\mu\nu}$ folgt

$$F_{jk} = \partial_j A_k \partial_k A_j = -\partial_j A^k + \partial_k A^j = -\epsilon_{jkl}(\vec{\nabla} \times \vec{A})^l = -\epsilon_{jkl}B^l. \tag{4.7.25}$$

Zum Vergleich mit (4.7.3) ist zu beachten, dass

$$\begin{aligned} F^{\mu\nu} = \eta^{\mu\rho}\eta^{\nu\sigma}F_{\rho\sigma} \Rightarrow F^{0j} = -F^{j0} = \eta^{0\rho}\eta^{j\sigma}F_{\rho\sigma} = \eta^{j\sigma}F_{0\sigma} = \eta^{jk}F_{0k} = \eta^{jk}E^k = -\frac{1}{c}E^j \\ F^{jk} = -F^{kj} = \eta^{j\rho}\eta^{k\sigma}F_{\rho\sigma} = \eta^{jl}\eta^{km}F_{lm} = +F_{jk} = -\epsilon^{jkl}B^l \end{aligned} \tag{4.7.26}$$

ist, was in der Tat mit (4.7.3) übereinstimmt. Es ist klar, dass wir den freien Index in (4.7.23) nach oben ziehen können. Dann folgt in der Tat

$$\frac{\mathrm{d}p^\mu}{\mathrm{d}\tau} = m\frac{\mathrm{d}^2 x^\mu}{\mathrm{d}\tau} = qF^{\mu\nu}\frac{\mathrm{d}x_\nu}{\mathrm{d}\tau}, \tag{4.7.27}$$

und das ist die oben bereits hergeleitete Bewegungsgleichung (4.7.2).

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 ist[5] $\vec{E} = (E, 0, 0)^{\mathrm{T}} = \text{const}$ 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

$$\frac{\mathrm{d}}{\mathrm{d}t}p^1 = qE \Rightarrow p^1(t) = qEt. \tag{4.7.28}$$

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

$$p^0(t) = \frac{\mathscr{E}}{c} = \sqrt{m^2c^2 + (qEt)^2} \tag{4.7.29}$$

[5] Man beachte, dass hier E die 1-Komponente des elektrischen Feldes und nicht die Energie bezeichnet. Deshalb haben wir oben $\mathscr{E}$ als Bezeichnung für die Energie gewählt.

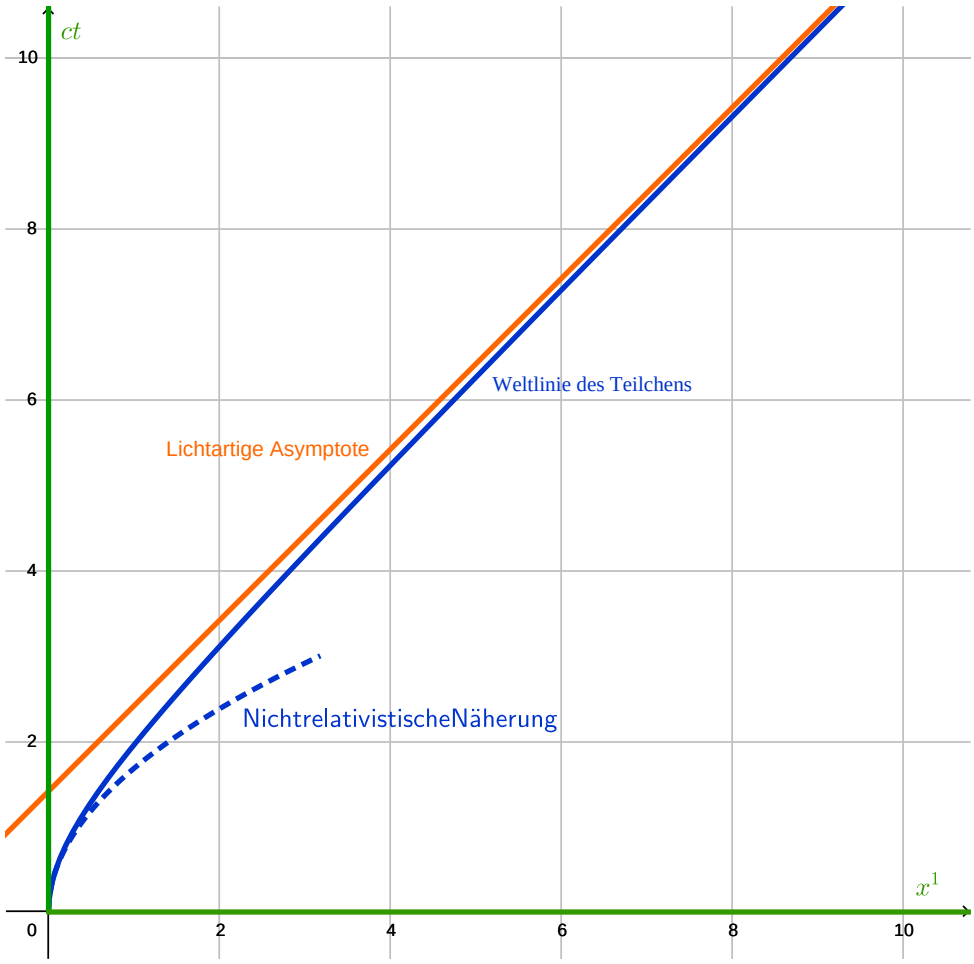


Abbildung 4.4: *Minkowski-Diagramm für die „hyperbolische Bewegung" in einem homogenen elektrischen Feld. Die entsprechende Weltlinie ist eine Hyperbel und besitzt eine lichtartige Asymptote. Dies zeigt, dass ein Lichtsignal, das vom Ursprung zu einer Zeit* $t > t_H = \frac{mc}{qE}$ *ausgesandt wird, 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).*

und daher

$$v^1(t) = \frac{\mathrm{d}x^1}{\mathrm{d}t} = \frac{c\,p^1}{p^0} = \frac{qcEt}{\sqrt{m^2c^2 + q^2E^2t^2}}. \tag{4.7.30}$$

Dies lässt sich unter Berücksichtigung der Anfangsbedingung leicht nochmals nach der Zeit integrieren, wodurch wir die Lösung

$$x^1(t) = \frac{mc^2}{qE}\left(\sqrt{1 + \frac{q^2E^2t^2}{m^2c^2}} - 1\right) \tag{4.7.31}$$

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 (4.7.31) entwickeln (*nachrechnen*) und erhalten

$$x^1(t) = \frac{mc^2}{qE}\left(\frac{q^2E^2t^2}{2m^2c^2} + \mathcal{O}\left[\left(\frac{qEt}{mc}\right)^4\right]\right) = \frac{qEt^2}{2m} + \cdots \tag{4.7.32}$$

In führender Ordnung erhalten wir die erwartete Zeitabhängigkeit für die nichtrelativistische Bewegung in einem konstanten Kraftfeld.

Wir können auch den relativistischen Limes betrachten, wo $t \gg mc/(qE)$ ist. Dann folgt durch die entsprechend angepasste Entwicklung der Wurzel

$$x^1(t) = \frac{mc^2}{qE}\left(\frac{qEt}{mc}\sqrt{1+\frac{m^2c^2}{q^2E^2t^2}} - 1\right) \underset{t\to\infty}{\cong} c\left(t - \frac{mc}{qE}\right). \tag{4.7.33}$$

Das impliziert auch, dass ein in Sect. 4.1 diskutiertes vom Ursrung ausgesendetes Lichtsignal das Teilchen nur erreichen kann, wenn es zu einer Zeit $t < t_{\mathrm{H}}$ mit $t_{\mathrm{H}} = \frac{mc}{qE}$ 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 **Minkowksi-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 Dreiervektoren formulierten Elektrodynamik zu dem antisymmetrischen Faraday-Tensor (4.7.3) zusammengefasst werden.

Die Maxwell-Gleichungen im Vakuum lauten im üblichen 3D-Formalismus bekanntlich

$$\vec{\nabla} \times \vec{E} + \partial_t \vec{B} = \vec{0}, \tag{4.8.1}$$

$$\vec{\nabla} \cdot \vec{B} = 0, \tag{4.8.2}$$

$$\vec{\nabla} \times \vec{B} - \epsilon_0 \mu_0 \partial_t \vec{E} = \mu_0 \vec{j}, \tag{4.8.3}$$

$$\vec{\nabla} \cdot \vec{E} = \frac{1}{\epsilon_0}\rho, \tag{4.8.4}$$

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 $(x^\mu) = (x^0, x^1, x^2, x^3)^\mathrm{T}$ 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

$$x'^\mu = \Lambda^\mu{}_\nu x^\nu. \tag{4.8.5}$$

Dabei bilden die $\Lambda^\mu{}_\nu$ eine Matrix, die das Minkowskiprodukt beliebiger Vierervektoren $\underline{x}$ und $\underline{y}$ ungeändert lässt, d.h.

$$\eta_{\mu\nu} x^\mu y^\nu = \eta_{\mu\nu} x'^\mu y'^\nu, \quad (\eta_{\mu\nu}) = \mathrm{diag}(1,-1,-1,-1). \tag{4.8.6}$$

Daraus ergibt sich die Bedingung an die Matrix $\hat{\Lambda}$, eine Lorentz-Transformation zu sein, zu

$$\eta_{\mu\nu} \Lambda^\mu{}_\rho \Lambda^\nu{}_\sigma = \eta_{\rho\sigma}. \tag{4.8.7}$$

Zur Bequemlichkeit führt man nun auch Vektorkomponenten mit unteren Indizes, die **kovarianten Komponenten** ein, indem man definiert

$$x_\mu = \eta_{\mu\nu} x^\nu. \tag{4.8.8}$$

Da $\hat{\eta}^{-1} = \hat{\eta}$ ist, kann man mit $\eta^{\mu\nu} = \eta_{\mu\nu}$ umgekehrt die Indizes auch wieder heraufziehen:

$$x^\mu = \eta^{\mu\nu} x_\nu. \tag{4.8.9}$$

Entsprechend kann man das Minkowskiprodukt auf verschiedene Arten mit den ko- bzw. kontravarianten Indizes schreiben:

$$\underline{x} \cdot \underline{y} = \eta_{\mu\nu} x^\mu y^\nu = x_\nu y^\nu = x^\mu y_\mu = \eta^{\mu\nu} x_\mu y_\nu. \tag{4.8.10}$$

Jetzt wollen wir das Transformationsverhalten der kovarianten Vektorkomponenten bestimmen:

$$x'_\mu = \eta_{\mu\rho} x'^\rho = \eta_{\mu\rho} \Lambda^\rho{}_\sigma x^\sigma = \eta_{\mu\rho} \Lambda^\rho{}_\sigma \eta^{\sigma\nu} x_\nu. \tag{4.8.11}$$

Wendet man die Regeln zum Indexziehen, auch auf die Komponenten der Lorentz-Transformationsmatrix an, folgt

$$x'_\mu = \Lambda_\mu{}^\nu x_\nu. \tag{4.8.12}$$

Andererseits folgt in dieser Schreibweise aus (4.8.7)

$$\Lambda_{\nu\rho} \Lambda^\nu{}_\sigma = \eta_{\rho\sigma}. \tag{4.8.13}$$

Zieht man nun den Index ρ mittels $\eta^{\rho\alpha}$ nach oben, ergibt sich daraus wiederum

$$\Lambda_\nu{}^\alpha \Lambda^\nu{}_\sigma = \eta^{\rho\alpha} \eta_{\rho\sigma} = \delta^\alpha{}_\sigma \Rightarrow (\hat{\Lambda}^{-1})^\alpha{}_\nu = \Lambda_\nu{}^\alpha. \tag{4.8.14}$$

Demnach gilt

$$x^\alpha = \delta^\alpha_\sigma x^\sigma = \Lambda_\nu{}^\alpha \Lambda^\nu{}_\sigma x^\sigma \overset{(4.8.5)}{=} \Lambda_\nu{}^\alpha x'^\nu. \tag{4.8.15}$$

Wollen wir nun **partielle Ableitungen** irgendwelcher Funktionen $\Phi(x^\mu)$ durch partielle Ableitungen nach den x'^μ ausdrücken, folgt aus der Kettenregel

$$\frac{\partial \Phi}{\partial x'^\mu} = \frac{\partial \Phi}{\partial x^\nu} \frac{\partial x^\nu}{\partial x'^\mu} \overset{(4.8.15)}{=} \frac{\partial \Phi}{\partial x^\nu} \Lambda_\mu{}^\nu. \tag{4.8.16}$$

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:

$$\frac{\partial \Phi}{\partial x^\nu} = \partial_\nu \Phi. \tag{4.8.17}$$

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

$$\Phi'(\underline{x}') = \Phi(\underline{x}) = \Phi(\hat{\Lambda}^{-1}\underline{x}'). \tag{4.8.18}$$

Entsprechend gilt für ein **Vierervektorfeld**

$$A'^{\mu}(\underline{x}') = \Lambda^{\mu}{}_{\rho} A^{\rho}(\underline{x}) = \Lambda^{\mu}{}_{\rho} A^{\rho}(\hat{\Lambda}^{-1}\underline{x}'). \tag{4.8.19}$$

Demnach ist

$$T_{\mu}{}^{\nu} = \partial_{\mu} A^{\nu} \tag{4.8.20}$$

ein **Tensorfeld 2. Stufe**. Die Transformationvorschrift ergibt sich aus der Indexstellung. Das rechnet man auch leicht direkt nach:

$$T'^{\sigma}_{\rho} = \partial'_{\rho} A'^{\sigma} = \frac{\partial x^{\mu}}{\partial x'^{\rho}} \Lambda^{\sigma}{}_{\nu} \partial_{\mu} A^{\nu} = \Lambda_{\rho}{}^{\mu} \Lambda^{\sigma}{}_{\nu} T_{\mu}{}^{\nu}. \tag{4.8.21}$$

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

$$\Phi = \partial_{\mu} A^{\mu}. \tag{4.8.22}$$

Dies sollte nun natürlich ein Skalarfeld sein. In der Tat ergibt sich aus (4.8.21) (*Nachrechnen!*)

$$\Phi' = \partial'_{\rho} A'^{\rho} = \Lambda_{\rho}{}^{\mu} \Lambda^{\rho}{}_{\nu} \partial_{\mu} A^{\nu} \overset{(4.8.14)}{=} \delta^{\mu}{}_{\nu} \partial_{\mu} A^{\nu} = \partial_{\mu} A^{\mu} = \Phi. \tag{4.8.23}$$

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 $\partial_0 = \partial/\partial x^0 = (1/c)\partial_t$, folgt[6]

$$(\partial_{\mu} F^{\mu\nu}) = \begin{pmatrix} \partial_m F^{m0} \\ \partial_0 F^{01} + \partial_m F^{m1} \\ \partial_0 F^{02} + \partial_m F^{m2} \\ \partial_0 F^{03} + \partial_m F^{m3} \end{pmatrix} = \begin{pmatrix} \vec{\nabla} \cdot \vec{E}/c \\ -\partial_t E^1/c^2 + \partial_2 B^3 - \partial_3 B^2 \\ -\partial_t E^2/c^2 + \partial_3 B^1 - \partial_1 B^3 \\ -\partial_t E^3/c^2 + \partial_1 B^2 - \partial_2 B^3 \end{pmatrix} = \begin{pmatrix} \vec{\nabla} \cdot \vec{E}/c \\ -\partial_t E^1/c^2 + (\vec{\nabla} \times \vec{B})^1 \\ -\partial_t E^2/c^2 + (\vec{\nabla} \times \vec{B})^2 \\ -\partial_t E^3/c^2 + (\vec{\nabla} \times \vec{B})^3 \end{pmatrix}. \tag{4.8.24}$$

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

$$(\partial_{\mu} F^{\mu\nu}) = \begin{pmatrix} c\rho/(\epsilon_0 c^2) \\ \mu_0 j^1 \\ \mu_0 j^2 \\ \mu_0 j^3 \end{pmatrix} = \mu_0 \begin{pmatrix} c\rho \\ j^1 \\ j^2 \\ j^3 \end{pmatrix}. \tag{4.8.25}$$

Damit wird klar, dass offenbar

$$\partial_{\mu} F^{\mu\nu} = \mu_0 j^{\nu} \tag{4.8.26}$$

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

[6] Im Folgenden laufen definitionsgemäß mit lateinischen Buchstaben bezeichnete Indizes von 1 bis 3.

Komponenten des elektromagnetischen Feldes unter Lorentz-Transformationen transformieren. Aus (4.8.26) folgt dann zwingend, dass die Komponenten des **elektrischen Viererstroms** $\underline{j} = (j^\nu) = (c\rho, \vec{j})^{\mathrm{T}}$ eben Komponenten eines Vierervektorfelds 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

$$\epsilon_{\alpha\beta\gamma\delta} = \eta_{\alpha\mu}\eta_{\beta\nu}\eta_{\gamma\rho}\eta_{\delta\sigma}\epsilon^{\mu\nu\rho\sigma} = \det(\hat{\eta})\epsilon^{\alpha\beta\gamma\delta} = -\epsilon^{\alpha\beta\gamma\delta}. \tag{4.8.27}$$

Das Transformationsverhalten unter Lorentz-Transformationen ergibt sich entsprechend zu (*Nachrechnen!*)

$$\epsilon'^{\alpha\beta\gamma\delta} = \det(\hat{\Lambda})\epsilon^{\alpha\beta\gamma\delta}, \tag{4.8.28}$$

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 (4.1.41) 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:

$$[({}^\dagger F)_{\mu\nu}] = \frac{1}{2}(\epsilon_{\mu\nu\rho\sigma}F^{\rho\sigma}) = \begin{pmatrix} 0 & B^1 & B^2 & B^3 \\ -B^1 & 0 & E^3/c & -E^2/c \\ -B^2 & -E^3/c & 0 & E^1/c \\ -B^3 & E^2/c & -E^1/c & 0 \end{pmatrix}. \tag{4.8.29}$$

Daraus erhalten wir

$$[\partial^\mu({}^\dagger F)_{\mu\nu}] = \begin{pmatrix} \vec{\nabla}\cdot\vec{B} \\ \partial_t B^1/c - \partial_2 E^3/c + \partial_3 E^2/c \\ \partial_t B^2/c - \partial_1 E^3/c + \partial_3 E^1/c \\ \partial_t B^3/c + \partial_1 E^2/c - \partial_2 E^1/c \end{pmatrix} = \begin{pmatrix} \vec{\nabla}\cdot\vec{B} \\ \partial_t\vec{B}/c + \vec{\nabla}\times\vec{E}/c \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}. \tag{4.8.30}$$

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)

$$\partial_\mu F^{\mu\nu} = \mu_0 j^\nu, \tag{4.8.31}$$

$$\partial^\mu({}^\dagger F)_{\mu\nu} = 0 \tag{4.8.32}$$

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)

$$\partial_\nu\partial_\mu F^{\mu\nu} = \mu_0\partial_\nu j^\nu = 0. \tag{4.8.33}$$

Damit also die Maxwell-Gleichungen überhaupt eine Lösung haben können, muss für den elektrischen Viererstrom die Bedingung

$$\partial_\nu j^\nu = \partial_0(c\rho) + \vec{\nabla}\cdot\vec{j} = \partial_t\rho + \vec{\nabla}\cdot\vec{j} = 0 \tag{4.8.34}$$

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.

Wir leiten noch eine alternative Form der homogenen Maxwell-Gleichungen (4.8.32) her, die manchmal nützlich ist. Dazu setzen wir die Definition des dualen Feldstärke-Tensors (4.8.29) in (4.8.32) ein. Da $\epsilon^{\mu\nu\rho\sigma} = \text{const}$ gilt

$$\partial_\mu^\dagger F^{\mu\nu} = \epsilon^{\mu\nu\rho\sigma}\partial_\mu F_{\rho\sigma} = 0. \tag{4.8.35}$$

Es ist klar, dass nur Terme des Tensors $\partial_\mu F_{\nu\rho}$, die total antisymmetrisch unter Vertauschen der Indizes μ, ν und ρ sind, nichttriviale Beiträge zur Summe über diese Indizes in (4.8.35) liefern, d.h. dass der total antisymmetrische Anteil dieses Tensors eindeutig verschwindet. Da $F_{\rho\sigma} = -F_{\sigma\rho}$ bereits antisymmetrisch ist, ist dieser antisymmetrisierte Ausdruck durch

$$\partial_\mu F_{\rho\sigma} - \partial_\rho F_{\mu\sigma} - \partial_\sigma F_{\rho\mu} = \partial_\mu F_{\rho\sigma} + \partial_\rho F_{\sigma\mu} + \partial_\sigma F_{\mu\rho} \tag{4.8.36}$$

gegeben, wobei wir die Antisymmetrie des Feldstärketensors verwendet haben. Das Verschwinden dieser Summe über die „zyklischen Vertauschungen“ der Indizes ist also äquivalent zu den homogenen Maxwell-Gleichungen (4.8.32):

$$\partial_\mu F_{\rho\sigma} + \partial_\rho F_{\sigma\mu} + \partial_\sigma F_{\mu\rho} = 0. \tag{4.8.37}$$

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

$$\vec{E} = -\partial_t\vec{A} - \vec{\nabla}\phi, \quad \vec{B} = \vec{\nabla}\times\vec{A} \tag{4.8.38}$$

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

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu \tag{4.8.39}$$

macht.

Trennen wir dies in Zeit- und Raum-Komponenten, ergibt sich mit (4.7.3)[7]

$$F_{0n} = -F^{0n} = E^n/c = \partial_t A_n/c - \partial_n A_0 = -\partial_t A^n/c - \partial_n A_0, \tag{4.8.40}$$

$$F_{mn} = F^{mn} = -\epsilon^{lmn}B^l = -\epsilon^{lmn}(\vec{\nabla}\times\vec{A})^l = \partial_m A_n - \partial_n A_m = -(\partial_m A^n - \partial_n A^m). \tag{4.8.41}$$

In der Tat ergeben sich demnach die vier Komponenten des **Viererpotentials** $\underline{A}$ aus dem Skalar- und Dreiervektorpotential des Dreierformalismusses:

$$\underline{A} = (A^\mu) = \begin{pmatrix} \phi/c \\ \vec{A} \end{pmatrix}. \tag{4.8.42}$$

[7] Wir bemerken, dass für den Levi-Civita-Tensor im dreidimensionalen Euklidischen Raum nach wie vor $\epsilon_{lmn} = \epsilon^{lmn}$ mit $\epsilon_{123} = +1$ gilt.

Dabei ist $\underline{A}$ nur bis auf den Vierergradienten eines Viererskalarfeldes bestimmt, denn setzen wir

$$A'_\mu = A_\mu + \partial_\mu \chi, \tag{4.8.43}$$

folgt wegen der Vertauschbarkeit der partiellen Ableitungen

$$F'_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A'_\mu = (\partial_\mu A_\nu - \partial_\nu A_\mu) + (\partial_\mu \partial_\nu \chi - \partial_\nu \partial_\mu \chi) = F_{\mu\nu}, \tag{4.8.44}$$

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 (4.8.39) in die homogenen Maxwell-Gleichungen (4.8.32) ein, folgt

$$\partial^\mu F_{\mu\nu} = \partial^\mu (\partial_\mu A_\nu - \partial_\nu A_\mu) = \Box A_\nu - \partial_\nu \partial_\mu A^\mu = \mu_0 j_\nu. \tag{4.8.45}$$

Dabei ist

$$\Box = \partial_\mu \partial^\mu = \partial_0^2 - \vec{\nabla}^2 = \frac{1}{c^2}\partial_t^2 - \Delta \tag{4.8.46}$$

der bereits vom Dreierformalismus her bekannte **d' Alembert-Operator**.
Im Dreierformalismus hatten wir die Integration der Gleichungen (4.8.45) 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 (4.8.44), wenn wir die **Lorenz-Eichbedingung**

$$\partial_\mu A^\mu = 0 \Leftrightarrow \frac{1}{c^2}\partial_t \phi + \vec{\nabla} \cdot \vec{A} = 0. \tag{4.8.47}$$

erfüllen. Dann vereinfacht sich nämlich zum Einen (4.8.45) zu vier getrennten **Wellengleichungen** für die Komponenten des Viererpotentials,

$$(\Box A^\nu) = \begin{pmatrix} \Box \phi / c \\ \Box \vec{A} \end{pmatrix} = \mu_0 \begin{pmatrix} c\rho \\ \vec{j} \end{pmatrix} \tag{4.8.48}$$

bzw.

$$\Box \phi = \mu_0 c^2 \rho = \frac{1}{\epsilon_0}\rho, \quad \Box \vec{A} = \mu_0 \vec{j}. \tag{4.8.49}$$

Dies stimmt mit dem entsprechenden Resultat im Dreierformalismus überein (vgl. [Hee18c], Abschnitt 5.6). Zum Anderen ist der Vorteil der Lorenz-Eichbedingung (4.8.47), dass es sich um eine vom Bezugssystem unabhängige Größe handelt, denn offensichtlich is $\partial_\mu A^\mu$ ein Skalarfeld, d.h. gilt die Lorenz-Eichbedingung in einem Inertialsystem, so ist sie automatisch auch in jedem anderen Inertialsystem erfüllt.

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:

$$(F^{\mu\nu}) = \begin{pmatrix} 0 & -E^1/c & -E^2/c & -E^3/c \\ E^1/c & 0 & -B^3 & B^2 \\ E^2/c & B^3 & 0 & -B^1 \\ E^3/c & -B^2 & B^1 & 0 \end{pmatrix}. \tag{4.8.50}$$

Die Tensorkomponenten transformieren sich gemäß der entsprechenden Vorschrift analog zu (4.8.21), nur dass wir jetzt zwei kontravariante (obere) Indizes vorliegen haben:

$$F'^{\mu\nu}(\underline{x}') = \Lambda^{\mu}{}_{\rho}\Lambda^{\nu}{}_{\sigma}F^{\rho\sigma}(\underbrace{\hat{\Lambda}^{-1}\underline{x}'}_{=\underline{x}}). \tag{4.8.51}$$

Dabei ist $\hat{\Lambda} = (\Lambda^{\mu}{}_{\nu})$ die Lorentz-Boost-Matrix (4.1.41). In Matrix-Vektor-Schreibweise entspricht dies der Formel

$$\hat{F}' = \hat{\Lambda}\hat{F}\hat{\Lambda}^{\mathrm{T}}. \tag{4.8.52}$$

Führt man die Matrizenmultiplikationen aus (*Übung!*), kann man die Komponenten im Dreierformalismus ablesen:

$$\begin{aligned} &E'^1 = cF'^{10} = E^1, \quad E'^2 = cF'^{20} = \gamma(E^2 - \beta c B^3), \quad E'^3 = cF'^{30} = \gamma(E^3 + \beta c B^2), \\ &B'^1(\underline{x}') = F'^{32} = B^1, \quad B'^2 = F'^{13} = \gamma(B^2 + \beta E^3/c), \quad B'^3 = F'^{21} = \gamma(B^3 - \beta E^2/c). \end{aligned} \tag{4.8.53}$$

Wegen $\vec{\beta} = \beta\vec{e}_1$ können wir dies auch in 3D-kovarianter Form schreiben (*Nachrechnen!*). Dazu spalten wir $\vec{E}$ und $\vec{B}$ in die Anteile parallel und senkrecht zur Boost-Geschwindigkeit auf (*Nachrechnen!*):

$$\begin{aligned} &\vec{E}'_{\|} = \vec{E}_{\|}, \quad \vec{E}'_{\perp} = \gamma(\vec{E}_{\perp} + \vec{v} \times \vec{B}), \\ &\vec{B}'_{\|} = \vec{B}_{\|}, \quad \vec{B}'_{\perp} = \gamma(\vec{B}_{\perp} - \vec{\beta} \times \vec{E}/c). \end{aligned} \tag{4.8.54}$$

Dabei sind freilich auch die Raumzeit-Argumente x^{μ} auf der rechten Seite durch die x'^{μ} auszudrücken.
Die Umkehrtransformation ergibt sich daraus einfach durch Ersetzen von $\vec{v} \to -\vec{v}$:

$$\begin{aligned} &\vec{E}_{\|} = \vec{E}'_{\|}, \quad \vec{E}_{\perp} = \gamma(\vec{E}'_{\perp} - \vec{v} \times \vec{B}'), \\ &\vec{B}_{\|} = \vec{B}'_{\|}, \quad \vec{B}_{\perp} = \gamma(\vec{B}'_{\perp} + \vec{\beta} \times \vec{E}'/c). \end{aligned} \tag{4.8.55}$$

Manchmal ist es einfacher, zuerst das Viererpotential (4.8.42) zu transformieren:

$$\underline{A}'(\underline{x}') = \begin{pmatrix} \phi'/c \\ A'^1 \\ A'^2 \\ A'^3 \end{pmatrix} = \hat{\Lambda}\underline{A}(\underline{x}) = \begin{pmatrix} \gamma(A^0 - \beta A^1) \\ \gamma(A^1 - \beta A^0) \\ A^2 \\ A^3 \end{pmatrix} \tag{4.8.56}$$

und dann die entsprechenden Ableitungen des Dreierformalismusses zu verwenden

$$\vec{E}'(\underline{x}') = -\partial_{t'}\vec{A}'(\underline{x}') - \vec{\nabla}'\phi'(\underline{x}'), \quad \vec{B}'(\underline{x}') = \vec{\nabla}' \times \vec{A}'(\underline{x}'). \tag{4.8.57}$$

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 Koodinatensystems von Σ' an den Ort der Ladung. Dort ist das elektromagnetische Feld durch das **Coulomb-Feld** gegeben:

$$\vec{E}'(\underline{x}') = \frac{q}{4\pi\epsilon_0}\frac{\vec{x}'}{r'^3}, \quad \vec{B} = \vec{0}, \quad r' = |\vec{x}'|. \tag{4.8.58}$$

Nun benötigen wir die Umkehrtransformation von (4.8.53). Diese erhalten wir aus (4.8.53), indem wir dort überall β durch $-\beta$ ersetzen. Damit wird

$$\vec{E} = \begin{pmatrix} E'^1 \\ \gamma E'^2 \\ \gamma E'^3 \end{pmatrix}, \quad \vec{B} = \begin{pmatrix} 0 \\ -\gamma\beta E'^3/c \\ \gamma\beta E'^2/c \end{pmatrix} = \vec{\beta} \times \vec{E}/c. \tag{4.8.59}$$

Dabei müssen wir noch die Argumente der Felder durch die Raumzeit-Koordinaten $\underline{x}$ ausdrücken:

$$\underline{x}' = \begin{pmatrix} ct' \\ x'^1 \\ x'^2 \\ x'^3 \end{pmatrix} = \begin{pmatrix} \gamma(ct - \beta x^1) \\ \gamma(x^1 - \beta ct) \\ x^2 \\ x^3 \end{pmatrix}. \tag{4.8.60}$$

Demnach wird

$$r' = \sqrt{\gamma^2(x^1 - \beta ct)^2 + (x^2)^2 + (x^3)^2} \tag{4.8.61}$$

und somit

$$\begin{aligned} \vec{E} &= \frac{q\gamma}{4\pi\epsilon_0\sqrt{\gamma^2(x^1-\beta ct)^2+(x^2)^2+(x^3)^2}^3} \begin{pmatrix} x^1 - \beta ct \\ x^2 \\ x^3 \end{pmatrix}, \\ \vec{B} &= \frac{q\gamma\beta}{4\pi\epsilon_0 c\sqrt{\gamma^2(x^1-\beta ct)^2+(x^2)^2+(x^3)^2}^3} \begin{pmatrix} 0 \\ -x^3 \\ x^2 \end{pmatrix}. \end{aligned} \tag{4.8.62}$$

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. [Hee18c], Abschnitt 5.1)

$$\vec{E}(t,\vec{x}) \equiv \vec{E}(\underline{x}) = \vec{E}_0 \cos[k(ct - \vec{n}\cdot\vec{x})], \quad \vec{B}(\underline{x}) = \frac{1}{c}\vec{n} \times \vec{E}(\underline{x}). \tag{4.8.63}$$

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

$$\underline{k} = (k^\mu) = \begin{pmatrix} k \\ k\vec{n} \end{pmatrix} = \begin{pmatrix} \omega/c \\ k\vec{n} \end{pmatrix} \tag{4.8.64}$$

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 (4.8.64)

$$\vec{E}(\underline{x}) = \vec{E}_0 \cos(\underline{k}\cdot\underline{x}), \quad \vec{B}(\underline{x}) = \frac{1}{c}\vec{n} \times \vec{E}(\underline{x}). \tag{4.8.65}$$

Betrachten wir nun die Situation von Bobs Standpunkt aus, also im Ruhsystem der Lichtquelle. Er verwendet zur Beschreibung des elektromagnetischen Feldes die Komponenten

$$\vec{E}'(\underline{x}') = \vec{E}'_0 \cos(\underline{k}' \cdot \underline{x}'). \tag{4.8.66}$$

Ohne Beschränkung der Allgemeinheit können wir annehmen, dass er sein Koordinatensystem so wählt, dass $\vec{n}' = (\cos\alpha', \sin\alpha', 0)^{\mathrm{T}}$ und $\vec{E}'_0 = E'_{01}\vec{\epsilon}'_1 + E'_{02}\vec{\epsilon}_2{}'$ ist. Dabei sind

$$\vec{\epsilon}'_1 = \begin{pmatrix} -\sin\alpha' \\ \cos\alpha' \\ 0 \end{pmatrix}, \quad \vec{\epsilon}'_2 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \tag{4.8.67}$$

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

$$\vec{E}'(\underline{x}') = \vec{E}'_0 \cos(\underline{k}' \cdot \underline{x}'), \quad \vec{B}'(\underline{x}') = \vec{B}'_0 \cos(\underline{k}' \cdot \underline{x}'). \tag{4.8.68}$$

Da weiter $\underline{k}' \cdot \underline{x}' = \underline{k} \cdot \underline{x}$ ist, gilt gemäß (4.8.53)

$$\vec{E}(\underline{x}) = \vec{E}_0 \cos(\underline{k} \cdot \underline{x}), \quad \vec{B}(\underline{x}) = \vec{B}_0 \cos(\underline{k} \cdot \underline{x}) \tag{4.8.69}$$

mit

$$\vec{E}_0 = \begin{pmatrix} -E'_{01}\sin\alpha' \\ \gamma E'_{01}(\beta + \cos\alpha') \\ E'_{02}\gamma(1 + \beta\cos\alpha') \end{pmatrix} \quad \vec{B}_0 = \frac{1}{c}\begin{pmatrix} E'_{02}\sin\alpha' \\ -E'_{02}\gamma(\beta + \cos\alpha') \\ E'_{01}\gamma(1 + \beta\cos\alpha') \end{pmatrix}. \tag{4.8.70}$$

Nun ist

$$\underline{k} = \hat{\Lambda}^{-1}\underline{k}' = \hat{\Lambda}^{-1}\begin{pmatrix} k' \\ k'\cos\alpha' \\ k'\sin\alpha' \\ 0 \end{pmatrix} = \begin{pmatrix} \gamma k'(1 + \beta\cos\alpha') \\ \gamma k'(\beta + \cos\alpha') \\ k'\sin\alpha' \\ 0 \end{pmatrix} = \begin{pmatrix} k \\ k\vec{n} \end{pmatrix}. \tag{4.8.71}$$

Es ist also

$$k = \gamma k'(1 + \beta\cos\alpha'), \quad \vec{n} = \begin{pmatrix} (\beta + \cos\alpha')/(1 + \beta\cos\alpha') \\ \sqrt{1 - \beta^2}\sin\alpha'/(1 + \beta\cos\alpha') \\ 0 \end{pmatrix} = \begin{pmatrix} \cos\alpha \\ \sin\alpha \\ 0 \end{pmatrix}. \tag{4.8.72}$$

Die Felder (4.8.70) besitzen dann exakt die gleiche Form wie im System Σ', wobei

$$\vec{\epsilon}_1 = \begin{pmatrix} -\sin\alpha \\ \cos\alpha \\ 0 \end{pmatrix}, \quad \vec{\epsilon}_2 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}, \quad \vec{E} = (E_{01}\vec{\epsilon}_1 + E_{02}\vec{\epsilon}_2)\cos(\underline{k} \cdot \underline{x}), \quad \vec{B} = \frac{1}{c}\vec{n} \times \vec{E}. \tag{4.8.73}$$

Es sind also wieder ebene Wellen wie im Ruhsystem 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 Paramtern E'_{01}, E'_{02} und α'. Dies ist zu erwarten, weil die Forminvarianz der Maxwell-Gleichungen unter Lorentz-Transformationen bedeutet, dass die Lorentz-tranformierten 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 Ruhsystem der Lichtquelle Σ'.

So ändert sich die Frequenz gemäß (4.8.71) zu

$$\omega = ck = \gamma\omega'(1 + \beta\cos\alpha'). \tag{4.8.74}$$

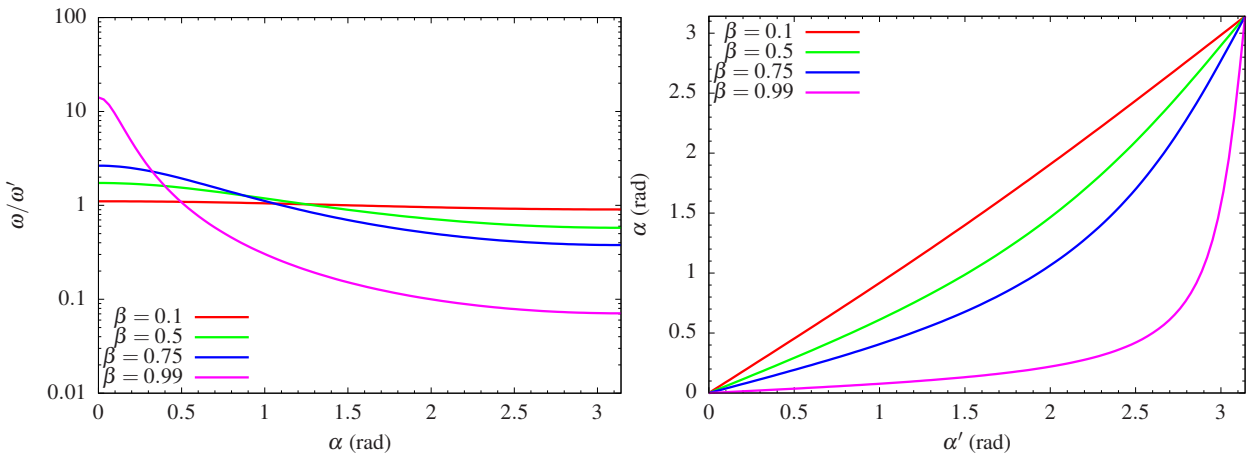


Abbildung 4.5: *Links: Dopplereffekt beim Licht gemäß (4.8.80); rechts: Aberration des Lichts gemäß (4.8.78).*

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 wegbewegt, entsprechend $\alpha = 0$ bzw. $\alpha = \pi$. Dann erhält man aus (4.8.72) auch $\alpha' = 0$ bzw. $\alpha' = \pi$ (*warum?*)

$$\omega_\pm = \gamma\omega'(1 \pm \beta) = \omega'\sqrt{\frac{1 \pm \beta}{1 \mp \beta}}. \tag{4.8.75}$$

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 Ausbereitungsrichtung des Lichtes bewegt. Das entspricht $\alpha = \pi/2$ bzw. $\cos\alpha = 0$ und gemäß (4.8.72) $\cos\alpha' = -\beta$ und $\sin\alpha' = \sqrt{1-\beta^2}$. Dann wird gemäß (4.8.74

$$\omega = \gamma\omega'(1-\beta^2) = \omega'\sqrt{1-\beta^2} \quad \text{für} \quad \alpha = \pi/2 \tag{4.8.76}$$

d.h. im Gegensatz zum analogen Sachverhalt bei Schallwellen, gibt es also auch einen **transversalen Doppler-Effekt**. Da hier $\omega < \omega'$ ist, handelt es sich dabei um eine **Rotverschiebung**.

Zu beachten ist, dass der transversale Doppler-Effekt zu einer **Blauverschiebung** führt, wenn man den Fall betrachtet, dass sich der Beobachter im Ruhsystem der Quelle bewegt und die Frequenz von Licht misst, das senkrecht zu seiner Bewegungsrichtung emittiert wird. Dann ist nämlich $\alpha' = \pi/2$ und gemäß (4.8.74)

$$\omega = \gamma\omega' = \frac{\omega'}{\sqrt{1-\beta^2}} \quad \text{für} \quad \alpha' = \pi/2. \tag{4.8.77}$$

Weiter zeigt sich, dass sich der Winkel α im Bezugssystem, in dem sich die Lichtquelle bewegt, gegenüber dem Winkel α' im Ruhsystem der Quelle ändert. Dies nennt man die **Aberrations des Lichts**. Da $\alpha, \alpha' \in [0, \pi]$, errechnet sich gemäß (4.8.72) der Winkel α eindeutig zu

$$\alpha = \arccos\left(\frac{\beta + \cos\alpha'}{1 + \beta\cos\alpha'}\right). \tag{4.8.78}$$

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 (4.8.73).

Natürlich können wir auch umgekehrt den Dopplereffekt direkt durch den Winkel α bzgl. Σ ausdrücken, denn offenbar gilt

$$\underline{k}' = \hat{\Lambda}\underline{k} = \hat{\Lambda}\begin{pmatrix} k \\ k\cos\alpha \\ k\sin\alpha \\ 0 \end{pmatrix} = \begin{pmatrix} \gamma k(1-\beta\cos\alpha) \\ \gamma k(-\beta+\cos\alpha) \\ k\sin\alpha \\ 0 \end{pmatrix}. \tag{4.8.79}$$

Wie zu erwarten, unterscheidet sich dies von (4.8.71) nur durch das Ersetzen der Größen $(k,\alpha) \longleftrightarrow (k',\alpha')$ und $\beta \to -\beta$. Aus der Zeitkomponente der Gl. (4.8.79) liest man ab (vgl. Abb. 4.5, rechts)

$$\omega' = \omega\frac{1-\beta\cos\alpha}{\sqrt{1-\beta^2}} \Rightarrow \omega = \omega'\frac{\sqrt{1-\beta^2}}{1-\beta\cos\alpha}. \tag{4.8.80}$$

Schließlich berechnen wir noch die Intensität (also die Energiedichte) der elektromagnetischen Welle in Alices Bezugssystem und vergleichen sie mit der Intensität im Ruhsystem der Quelle. Es ergibt sich nach einfacher Rechnung

$$\varepsilon = \frac{\epsilon_0}{2}\vec{E}_0^2 = \frac{\epsilon_0}{2}\vec{E}_0'^2\gamma^2(1+\beta\cos\alpha')^2 = \varepsilon'\left(\frac{\omega}{\omega'}\right)^2. \tag{4.8.81}$$

Die Intensität ändert sich also im Vergleich zum Ruhsystem der Quelle mit dem Quadrat des „Doppler-Faktors“ ω/ω'. Für das Intensitätsverhältnis zwischen der Intensität im System, wo sich die Lichtquelle bewegt, zur Intensität im Ruhsystem der Quelle gilt also, jeweils ausgedrückt durch Größen bzgl. des Ruhsystems der Quelle Σ' bzw. im System Σ, wo sich die Quelle mit der Geschwindigkeit $\vec{v}$ bewegt.

$$\frac{\epsilon}{\epsilon'} = \left(\frac{\omega}{\omega'}\right)^2 \overset{(4.8.74)}{=} \frac{(1+\beta\cos\alpha')^2}{1-\beta^2} \overset{(4.8.80)}{=} \frac{1-\beta^2}{(1-\beta\cos\alpha)^2}. \tag{4.8.82}$$

4.8.6 Der Energie-Impuls-Tensor

In der Vorlesung Theoretische Physik 2 haben wir die Ausdrücke für die **Energie- und Impulsdichte** des elektromagnetischen Feldes sowie die dazugehörigen **Energie- und Impulsstromdichten** anhand der kontinuumsmechanischen Formulierung der elektromagnetischen Bewgungsgleichung für Teilchen in einem gegebenen elektromagnetischen Feld aus der Erhaltung von Energie und Impuls des abgeschlossenen Systems aus Ladungen und dem elektromagnetischen Feld bestimmt. Wir wollen diese Rechnung hier in kovarianter Form wiederholen. Es ist ein Beispiel dafür, dass die manifest kovariante Formulierung der Elektrodynamik die Rechnungen wesentlich vereinfacht, freilich um den Preis, dass wir dazu zuerst die vierdimensionale Vektor- bzw. Tensoralgebra und -analysis entwickeln mussten.

Betrachten wir zunächst die Formulierung der Kontinummsmechanik einer Verteilung nichtwechselwirkender Teilchen. Es ist klar, dass wir dazu zunächst die entsprechenden Dichten für die (invariante!) Masse sowie die dazugehörige Massenstromdichte einführen müssen. Uns leitet dabei die Strategie, dass wir „lokale intrinsische Eigenschaften“ der Materie wie die Massen- bzw. Ladungsdichten als **skalare Felder** dadurch definieren können, dass wir diese im **momentanen Ruhsystem des betrachteten Fluid-Elements** festlegen. Eine Dichte liefert dann zusammen mit der dazugehörigen Stromdichte einen Vierervektor. Im lokalen Ruhsystem des Fluidelements, das wir mit einem * an den entsprechenden Vektor- bzw. Tensor-Komponenten bezeichnen, lautet dann der **Massen-Viererstromdichtevektor**

$$(j_{\mathrm{m}}^{*\mu}) = \mu^* c\underline{j}_m^* = \mu^*\begin{pmatrix} c \\ 0 \\ 0 \\ 0 \end{pmatrix} = \mu^* c\underline{u}^*. \tag{4.8.83}$$

Dabei bezeichnet $\underline{u}^*$ den **Geschwindigkeits-Vierervektor** des Fluidelements, entsprechend $c\underline{u} = \mathrm{d}\underline{x}/\mathrm{d}\tau$ (mit der Eigenzeit τ) eines Punktteilchens. Für diesen Vektor gilt also in jedem Inertialsystem $u_\mu u^\mu = \eta_{\mu\nu} u^\mu u^\nu = \underline{u} \cdot \underline{u} = 1$. Die Ruhemassedichte ist dabei ein Skalarfeld, das wir mit $\mu^*(\underline{x})$ bezeichnen. Dabei ist der Lorentz-Boost vom lokalen Ruhsystem der Fluidzelle zum Laborsystem durch die im Laborsystem gemessene Geschwindigkeit $\vec{v}(\underline{x}) = c\beta(\underline{x})$ des Fluidelements gegeben, d.h. ein beliebiger Vierervektor $\underline{V}$ transformiert sich gemäss

$$\underline{V} = \hat{\Lambda}(-\vec{v})\underline{V}^* \Rightarrow V^0 = \gamma(V^{*0} + \vec{\beta} \cdot \vec{x}^*), \quad \vec{V}_{\|} = \gamma(\vec{V}_{\|}^* + \vec{\beta} V^{*0}), \quad \vec{V}_{\perp} = \vec{V}_{\perp}^*, \tag{4.8.84}$$

wobei $\vec{V}_{\|}$ die Komponenten des räumlichen Anteils des Vierervektors in Richtung des Boosts und $\vec{V}_{\perp}$ diejenigen senkrecht dazu bezeichnen. Für $\underline{u}$ selbst folgt daraus wegen $\underline{u}^* = (1,0,0,0)^{\mathrm{T}}$

$$\underline{u} = \gamma \begin{pmatrix} 1 \\ \vec{\beta} \end{pmatrix}, \tag{4.8.85}$$

wie es sein muss.

Da nun (4.8.83) eine Gleichung zwischen Vierervektorkomponenten ist, gilt sie in einem beliebigen Bezugssystem, insbesondere im fest gewählten „Laborsystem“, d.h. es ist

$$\underline{j}_m = \mu^* c \underline{u} = \gamma \mu^* \begin{pmatrix} c \\ \vec{v} \end{pmatrix} = \mu \begin{pmatrix} c \\ \vec{v} \end{pmatrix}. \tag{4.8.86}$$

Daraus liest man ab, dass im Laborsystem die Massendichte $\mu = \gamma \mu^*$ ist. Dies trägt automatisch der Tatsache Rechnung, dass das Volumen $\mathrm{d}^3 x^*$ des Fluidelements im lokalen Ruhsystem, das der Definition der Massendichte im lokalen Ruhsystem $\mu^* = \mathrm{d}m/\mathrm{d}^3 x^*$ entspricht, für einen Beobachter im Laborsystem aufgrund der Längenkontraktion $\mathrm{d}^3 x = \mathrm{d}^3 x^*/\gamma$ beträgt. Die im Volumenelment enthaltene (invariante!) Masse $\mathrm{d}m$ ändert sich dabei nicht, denn es ist eine skalare Größe. Demnach gilt für die Massendichte im Laborsystem in der Tat $\mu = \mathrm{d}m/\mathrm{d}^3 x = \gamma \mathrm{d}m/\mathrm{d}^3 x^* = \gamma \mu^*$. Die Massenstromdichte ist im Laborsystem entsprechend $\vec{j}_m = \mu \vec{v} = \gamma \mu^* \vec{v} = \mu^* c \vec{u}$.

Betrachten wir nun eine beliebige auf das Fluid bezogene (skalare) Größe $\Phi(\underline{x})$, so ist die materielle zeitliche Änderung durch

$$\mathrm{D}_t \Phi(\vec{x}) = \partial_t \Phi(\underline{x}) + \vec{v} \cdot \Phi(\underline{x}) \tag{4.8.87}$$

gegeben, denn während einer kleinen Zeit $\mathrm{d}t$ hat sich das Fluidelement um $\mathrm{d}t\vec{v}$ weiterbewegt. Dies ist nun jedoch kein Skalar mehr. Betrachtet man aber stattdessen die materielle Ableitung nach der Eigenzeit des Fluid elements erhält man

$$\mathrm{D}_\tau \Phi(\vec{x}) = \gamma \mathrm{D}_t \Phi(\vec{x}) = \gamma \partial_t \Phi(\underline{x}) + \gamma \vec{v} \cdot \vec{\nabla} \Phi = c\gamma \partial_0 \Phi + c\gamma \vec{\beta} \cdot \vec{\nabla} \Phi = c u^\mu \partial_\mu \Phi, \tag{4.8.88}$$

und das ist ein Skalarfeld.

Die relativistische Form der Newtonsche Bewegungsgleichung lässt sich demnach entsprechend der obigen Betrachtung zur Punktteilchenmechanik in die Kontinuumsmechanik übertragen:

$$\mu^* c \mathrm{D}_\tau u^\mu = \mu^* c^2 u^\rho \partial_\rho u^\mu = c j_{\mathrm{m}}^\rho \partial_\rho u^\mu = k^\mu, \tag{4.8.89}$$

wobei k^μ die Dichte der **Lorentz-Kraft** ist. Für die Bewegung in einem äußeren elektromagnetischen Feld, repräsentiert durch den Feldstärketensor $F^{\mu\nu}$ ist

$$k^\mu = \rho^* c F^{\mu\nu} u_\nu = F^{\mu\nu} j_\nu, \tag{4.8.90}$$

wobei hier

$$\underline{j} = \rho^* c \underline{u} \tag{4.8.91}$$

die Viererstromdichte der elektrischen Ladung (und entsprechend ρ^* die Ruheladungsdichte) ist.

Weiter ist die lokale Form der Gesetzes von der **Erhaltung der invarianten Masse** und **elektrischen Ladung** durch die **Kontinuitätsgleichungen**

$$\partial_t \mu + \vec{\nabla} \cdot (\mu \vec{v}) = \partial_\mu j_{\mathrm{m}}^\mu = 0, \tag{4.8.92}$$

$$\partial_t \rho + \vec{\nabla} \cdot (\rho \vec{v}) = \partial_\mu j^\mu = 0, \tag{4.8.93}$$

also ebenfalls durch kovariante Gleichungen gegeben.

Die lokale Form der Energie- und Impulsbilanz wird nun ebenfalls durch solche Kontinuitätsgleichungen formuliert. Man benötigt für die Energiebilanz dazu eine **Energiedichte** und eine **Energiestromdichte** und für jede Komponente des Impulses ebenfalls eine Impulsdichte und eine Impulsstromdichte. Insgesamt lassen sich diese Größen in einen **Energie-Impuls-Vierertensor** zusammenfassen.

Für unseren Fall nichtwechselwirkender Materie haben wir im lokalen Ruhsystem der Fluidzelle nur eine Ruheenergiedichte $u^*_{\mathrm{mech}} = \mu c^2$ und sowohl Energiestrom- als auch die Impuls- und Impulsstromdichten verschwinden. Dies lässt sich in kovarianter Form durch den **Energie-Impuls-Tensor**

$$T^{\mu\nu}_{\mathrm{mech}} = \mu^* c^2 u^\mu u^\nu = c j_{\mathrm{m}}^\mu u^\nu \tag{4.8.94}$$

ausdrücken. Da $\underline{u}^* = (1,0,0,0)^{\mathrm{T}}$ ist, ist in der Tat im lokalen Ruhsystem nur $T^{*00}_{\mathrm{mech}} = \mu^* c^2$ von 0 verschieden.

Dass es sich tatsächlich um den Energie-Impulstensor handelt, erkennt man an der entsprechenden Bilanzgleichung. Diese erhalten wir einfach durch Anwenden von ∂_μ auf (4.8.94) unter Verwendung der Massenerhaltungsgleichung (4.8.92) und der Bewegungsgleichung (4.8.89):

$$\partial_\mu T^{\mu\nu}_{\mathrm{mech}} = c[(\partial_\mu j_{\mathrm{m}}^\mu) u^\nu + j_{\mathrm{m}}^\mu \partial_\mu u^\nu] = c j_{\mathrm{m}}^\mu \partial_\mu u^\nu = k^\nu = F^{\nu\rho} j_\rho. \tag{4.8.95}$$

Dies beschreibt die Übertragung von Energie und Impuls auf die Ladungsverteilung aufgrund des äußeren elektromagnetischen Feldes. Da wegen des **Noether-Theorems** wegen der Homogenität des Raumes und der Zeit die Gesamtenergie und der Gesamtimpuls eines abgeschlossenen Systems erhalten sein müssen, muss der entsprechende entgegengesetzte Übertrag auf das elektromagnetische Feld erfolgen, d.h. es ist

$$\partial_\mu T^{\mu\nu}_{\mathrm{em}} = -F^{\nu\rho} j_\rho. \tag{4.8.96}$$

Hier wird nun die Wichtigkeit des **Nahwirkungskonzepts** der feldtheoretischen Beschreibung der Kräfte klar: In der Newtonschen Mechanik wurden die Wechselwirkungen von weit voneinander entfernten Körpern (z.B. die Schwerkraft zwischen Himmelskörpern) durch **instantane Fernwirkungen** beschrieben. Nun kann aber die Lageänderung des einen Körpers nicht instantan die Änderung einer Kraft auf einen anderen weit entfernten Körper bewirken, da sich keine kausalen Wirkungen mit einer schnelleren Geschwindigkeit als der Lichtgeschwindigkeit ausbreiten können. Damit wäre in der Relativitätstheorie vermeintlich die Impulserhaltung verletzt. Beschreibt man allerdings die Wechselwirkungen zwischen Körpern über Felder, wo z.B. die elektromagnetische Kraft auf eine Ladung durch das elektromagnetische Feld am Ort dieser Ladung bestimmt ist, ergibt sich kein Paradox, denn das elektromagnetische Feld kann Energie und Impuls mit der Ladung austauschen. Das elektromagnetische Feld selbst wird von anderen evtl. weit entfernten Ladungs- und Stromverteilungen erzeugt. Die Lösungen der entsprechenden Maxwell-Gleichungen sind durch die **retardierten Potentiale** bzw. die **Jefimenko-Gleichungen** für das elektrische und magnetische Feld gegeben, die ebenfalls über die retardierte Greensche Funktion des d'Alembert-Operators $\Box = 1/c^2 \partial_t^2 - \Delta$ gegeben sind, und dies beschreibt genau eine Signalausbreitung mit Lichtgeschwindigkeit, was mit dem relativistischen Kausalitätsprinzip kompatibel ist.

Um nun den Energie-Impulstensor des elektromagnetischen Feldes $T^{\mu\nu}_{\mathrm{em}}$ zu finden, verwenden wir auf der rechten Seite von (4.8.96) die inhomogenen Maxwell-Gleichungen (4.8.31):

$$\partial_\mu T^{\mu\nu}_{\mathrm{em}} = -\frac{1}{\mu_0} F^{\nu\rho} \partial_\mu F^\mu{}_\rho = -\frac{1}{\mu_0} \left[\partial_\mu (F^{\nu\rho} F^\mu{}_\rho) - F^\mu{}_\rho \partial_\mu F^{\nu\rho} \right]. \tag{4.8.97}$$

Ziehen wir nun alle Indizes in den homogenen Maxwell-Gleichungen in der Form (4.8.37) nach oben, erhalten wir

$$\partial^\mu F^{\nu\rho} + \partial^\nu F^{\rho\mu} + \partial^\rho F^{\mu\nu} = 0. \tag{4.8.98}$$

Damit wird unter Berücksichtung der Antisymmetrie $F_{\mu\rho} = -F_{\rho\mu}$ des Feldstärketensors

$$\begin{aligned} F^\mu{}_\rho \partial_\mu F^{\nu\rho} &= \frac{1}{2} F_{\mu\rho}(\partial^\mu F^{\nu\rho} - \partial^\rho F^{\nu\mu}) \\ &= \frac{1}{2} F_{\mu\rho}(\partial^\mu F^{\nu\rho} + \partial^\rho F^{\mu\nu}) \overset{(4.8.98)}{=} -\frac{1}{2} F_{\mu\rho} \partial^\nu F^{\rho\mu} \\ &= \frac{1}{2} F_{\mu\rho} \partial^\nu F^{\mu\rho} = \frac{1}{4} \partial^\nu (F_{\rho\sigma} F^{\rho\sigma}). \end{aligned} \tag{4.8.99}$$

Dabei haben wir im letzten Schritt den Summationsindex μ in σ umbenannt. Damit können wir nun schreiben

$$F^\mu{}_\rho \partial_\mu F^{\nu\rho} = \frac{1}{4} \partial_\mu (F_{\rho\sigma} F^{\rho\sigma} \eta^{\mu\nu}). \tag{4.8.100}$$

Setzen wir dies in (4.8.97) ein, erhalten wir schließlich tatsächlich auf der rechten Seite eine Kontinuitätsgleichung, wie von Energie- und Impulserhaltungssatz verlangt:

$$\partial_\mu T^{\mu\nu}_{\text{em}} = \frac{1}{\mu_0} \partial_\mu \left(\frac{1}{4} F_{\rho\sigma} F^{\rho\sigma} \eta^{\mu\nu} - F^\mu{}_\rho F^{\nu\rho} \right). \tag{4.8.101}$$

Damit ergibt sich der **Energie-Impulstensor** des elektromagnetischen Feldes zu

$$T^{\mu\nu}_{\text{em}} = \frac{1}{\mu_0} \left(\frac{1}{4} F_{\rho\sigma} F^{\rho\sigma} \eta^{\mu\nu} - F^\mu{}_\rho F^{\nu\rho} \right) = \frac{1}{\mu_0} \left(\frac{1}{4} F_{\rho\sigma} F^{\rho\sigma} \eta^{\mu\nu} + F^\mu{}_\rho F^{\rho\nu} \right). \tag{4.8.102}$$

Rechnet man diese Matrix unter Verwendung von (4.7.3) und mit $\mu_0 \epsilon_0 = c^2$ aus, erhält man (*Nachrechnen!*) in Übereinstimmung mit der dreidimensionalen Rechnung in [Hee18c] die **Energiedichte** des elektromagnetischen Feldes

$$T^{00}_{\text{em}} = w^{(\text{em})} = \frac{\epsilon_0}{2} \vec{E}^2 + \frac{1}{2\mu_0} \vec{B}^2, \tag{4.8.103}$$

die Energiestromdichte bzw. Impulsdichte

$$c(T^{0n}_{\text{em}}) = c(T^{n0}_{\text{em}}) = \vec{J}^{W,\text{em}} = c^2 \vec{g}^{(\text{em})} = \frac{1}{\mu_0} \vec{E} \times \vec{B} = \vec{S} \Rightarrow \vec{g}^{(\text{em})} = \frac{1}{c^2} \vec{S} = \epsilon_0 \vec{E} \times \vec{B} \tag{4.8.104}$$

und die Impulsstromdichte bzw. den Maxwellschen Spannungstensor

$$(T^{mn}_{\text{em}}) = (J^{(p,\text{em})}_{mn}) = -(T^{(\text{Maxwell})}_{mn}) = -\epsilon_0 \left(E^m E^n - \frac{1}{2} \delta_{mn} \vec{E}^2 \right) - \frac{1}{\mu_0} \left(B^m B^n - \frac{1}{2} \delta_{mn} \vec{B}^2 \right). \tag{4.8.105}$$

Es ist also

$$(T^{\mu\nu}_{\text{em}}) = \begin{pmatrix} w^{(\text{em})} & c\vec{g}^{(\text{em})\text{T}} \\ c\vec{g}^{(\text{em})} & -T^{\text{Maxwell}}_{mn} \end{pmatrix}. \tag{4.8.106}$$

Bzgl. der diversen Faktoren c in (4.8.104) ist zu beachten, dass die Dimension der $T^{\mu\nu}$ Energie/Volumen = Energie/Länge^3 ist. Entsprechend ist Impuls/Volumen = Energie/(Volumen · Geschwindigkeit). Nun bedeuten aufgrund der Bewegungsgleichung (4.8.95) die $T^{0n} = T^{n0}$ bis auf entsprechende Faktoren c Energiestromdichte bzw. Impulsdichte, und es ergeben sich die besagten Faktoren c in (4.8.104). Bzgl. (4.8.105) hat eine Impulsstromdichte bzw. ein Spannungstensor die Dimension Impuls/(Fläche · Zeit) = Energie/(Fläche · Zeit · Geschwindigkeit) = Energie/Volumen, also wieder die Dimension einer Energiedichte. Dies ist zugleich auch die Dimension (Impuls/Zeit)/Fläche = Kraft/Fläche, also tatsächlich die Dimension einer mechanischen Spannung bzw. im Spezialfall idealer Fluide eines Drucks.

Kapitel 5

Elementarteilchen

Anhang A

Die Baker-Campbell-Hausdorff-Formel

A.1 Herleitung der BCH-Formel

Es seien **A** und **B** beliebige Operatoren. Dann gilt

$$\exp(\mathbf{A})\mathbf{B}\exp(-\mathbf{A}) = \sum_{n=0}^{\infty} \frac{1}{n!}[\mathbf{A},\mathbf{B}]_n \,, \tag{A.1.1}$$

wobei der „Multikommutator“ rekursiv durch

$$[\mathbf{A},\mathbf{B}]_0 = \mathbf{B}, \quad [\mathbf{A},\mathbf{B}]_n = \left[\mathbf{A},[\mathbf{A},\mathbf{B}]_{n-1}\right] \tag{A.1.2}$$

definiert ist.
Zum **Beweis** betrachten wir die Funktion

$$\mathbf{F}(z) = \exp(z\mathbf{A})\mathbf{B}\exp(-z\mathbf{A}) \tag{A.1.3}$$

und entwickeln diese Funktion nach Potenzen von z. Offenbar gilt

$$\mathbf{F}'(z) = \mathbf{A}\mathbf{F}(z) - \mathbf{F}(z)\mathbf{A} = [\mathbf{A},\mathbf{F}(z)], \tag{A.1.4}$$

und weiter durch Iteration

$$\mathbf{F}^{(n)}(z) = [\mathbf{A},\mathbf{F}(z)]_n \,. \tag{A.1.5}$$

Die Potenzreihenentwicklung lautet wegen $\mathbf{F}(0) = \mathbf{B}$ also

$$\mathbf{F}(z) = \sum_{n=0}^{\infty} \frac{z^n}{n!}[\mathbf{A},\mathbf{B}]_n \,, \tag{A.1.6}$$

und für $z = 1$ folgt die Behauptung. QED.

A.2 Anwendung auf Exponentialfunktion einer Summe

Als Anwendung der Baker-Campbell-Hausdorff-Formel (BCHF) beweisen wir noch folgenden Satz:
Seien **A** und **B** Operatoren, für die

$$[\mathbf{A},[\mathbf{A},\mathbf{B}]] = [[\mathbf{A},\mathbf{B}],\mathbf{B}] = 0 \tag{A.2.1}$$

gilt, so ist

$$\exp(\mathbf{A}+\mathbf{B}) = \exp\mathbf{A}\,\exp\mathbf{B}\,\exp\left(-\frac{1}{2}[\mathbf{A},\mathbf{B}]\right). \tag{A.2.2}$$

Zum Beweis wenden wir die BCHF (A.1.1) wie folgt an. Zunächst definieren wir

$$\mathbf{F}(z) = \exp[z(\mathbf{A}+\mathbf{B})]. \tag{A.2.3}$$

Wegen der Annahmen (A.2.1) bricht die BCHF bereits nach einem Term ab:

$$\begin{aligned} \mathbf{F}(z)\mathbf{A}\mathbf{F}^{-1}(z) &= \mathbf{A} + z\,[\mathbf{A}+\mathbf{B},\mathbf{A}] = \mathbf{A} - z\,[\mathbf{A},\mathbf{B}] \\ &\Rightarrow \mathbf{F}(z)\mathbf{A} = \mathbf{A}\mathbf{F}(z) - z\,[\mathbf{A},\mathbf{B}]\,\mathbf{F}(z). \end{aligned} \tag{A.2.4}$$

Ableiten nach von (A.2.3) nach z liefert hingegen mit Hilfe dieser Gleichung

$$\mathbf{F}'(z) = \mathbf{F}(z)(\mathbf{A}+\mathbf{B}) = \mathbf{A}\mathbf{F}(z) + \mathbf{F}(z)\mathbf{B} - z\,[\mathbf{A},\mathbf{B}]\,\mathbf{F}(z). \tag{A.2.5}$$

Offenbar wird diese Differentialgleichung durch

$$\mathbf{F}(z) = \exp(z\mathbf{A})\,\exp(z\mathbf{B})\,\exp\left(-\frac{z^2}{2}[\mathbf{A},\mathbf{B}]\right) \tag{A.2.6}$$

gelöst, und für $z = 1$ folgt (A.2.2). QED.

Anhang B

Spezielle Funktionen

In diesem Anhang stellen wir einige spezielle Funktionen zusammen, die uns im Haupttext aus Lösungen von Eigenwertgleichungen begegnen.

B.1 Hermite-Poloynome

B.1.1 Erzeugende Funktion

Wir haben die Hermite-Polynome im Zusammenhang mit den Energieeigenfunktionen des harmonischen Oszillators in (3.9.29) definiert:

$$\mathrm{H}_n(x) := \exp\left(\frac{x^2}{2}\right)\left(x - \frac{\mathrm{d}}{\mathrm{d}x}\right)^n \exp\left(-\frac{x^2}{2}\right). \tag{B.1.1}$$

Zunächst ist es leicht, eine Rekursionsformel herzuleiten. Dazu schreiben wir

$$\begin{aligned}\mathrm{H}_{n+1}(x) &= \exp(x^2/2)(x-\mathrm{d}_x)(x-\mathrm{d}_x)^n \exp(-x^2/2)\\ &= \exp(x^2/2)(x-\mathrm{d}_x)[\exp(x^2/2)\mathrm{H}_n(x)]\\ &= 2x\mathrm{H}_n(x) - \mathrm{d}_x H_n(x)\\ &= -\exp(x^2)\mathrm{d}_x[\exp(-x^2)\mathrm{H}_n(x)].\end{aligned} \tag{B.1.2}$$

Zusammen mit der Anfangsbedingung $\mathrm{H}_0(x) = 1$ folgt daraus durch Iteration die alternative Definition über die **Rodrigues-Formel**

$$\mathrm{H}_n(x) = (-1)^n \exp(x^2)\mathrm{d}_x^n \exp(-x^2). \tag{B.1.3}$$

Wir können nun einige nützliche Eigenschaften der Hermite-Polyonome herleiten, indem wir durch

$$F(x,z) = \sum_{n=0}^{\infty} \mathrm{H}_n(x)\frac{z^n}{n!} \tag{B.1.4}$$

eine sog. **erzeugende Funktion** definieren. Sie ist eine erzeugende Funktion in dem Sinne, dass aufgrund des Taylor-Satzes offenbar

$$\mathrm{H}_n(x) = \frac{\partial^n}{\partial z^n} F(x,z)\bigg|_{z=0} \tag{B.1.5}$$

gilt. Unser nächstes Ziel ist die explizite Berechnung der erzeugenden Funktion (B.1.2). Dazu setzen wir $z = x - y$ und verwenden (B.1.3):

$$\begin{aligned} F(x,z) &= \sum_{n=0}^{\infty} \frac{1}{n!} \mathrm{H}_n(x)(x-y)^n \\ &= \exp(x^2) \sum_{n=0}^{\infty} \frac{1}{n!} (-1)^n (x-y)^n \mathrm{d}_x^n \exp(-x^2) \\ &= \exp(x^2) \sum_{n=0}^{\infty} \frac{1}{n!} (y-x)^n \mathrm{d}_x^n \exp(-x^2) \\ &= \exp(x^2)\exp(-y^2) = \exp(x^2 - y^2). \end{aligned} \tag{B.1.6}$$

Dabei haben wir im letzten Schritt wieder den Taylor-Satz für die Entwicklung einer Funktion in eine Potenzreihe verwendet. Nun ist aber $y = x - z$ und damit

$$F(x,z) = \exp(2xz - z^2). \tag{B.1.7}$$

Wir können nun die Orthogonalität und das Normierungsintegral der Hermite-Polynome, das wir bereits auf andere Art bei der Berechnung der Energieeigenfunktionen des harmonischen Oszillators hergeleitet haben, jetzt leicht auch nochmals direkt berechnen. Dazu definieren wir

$$f(z_1, z_2) = \sum_{n_1, n_2 = 0}^{\infty} \frac{z_1^{n_1}}{n_1!} \frac{z_2^{n_2}}{n_1!} \int_{-\infty}^{\infty} \mathrm{d}x \exp(-x^2) \mathrm{H}_{n_1}(x) \mathrm{H}_{n_2}(x). \tag{B.1.8}$$

Mit (B.1.4) und (B.1.7) ist andererseits

$$f(z_1, z_2) = \int_{-\infty}^{\infty} \mathrm{d}x \exp[-x^2 + 2x(z_1 + z_2) - z_1^2 - z_2^2]. \tag{B.1.9}$$

Dieses Gaußsche Integral lässt sich mittels (C.1.22) geschlossen integrieren, was

$$f(z_1, z_2) = \sqrt{\pi} \exp(2z_1 z_2) \tag{B.1.10}$$

liefert. Mittels des Taylor-Satzes finden wir die fraglichen Integrale nun durch einfache Ableitungen dieser Funktion:

$$\int_{-\infty}^{\infty} \mathrm{d}x \exp(-x^2) \mathrm{H}_{n_1}(x) \mathrm{H}_{n_2}(x) = \partial_{z_1}^{n_1} \partial_{z_2}^{n_2} f(z_1, z_2) \Big|_{z_1 = z_2 = 0}. \tag{B.1.11}$$

Nun ist aber

$$\partial_{z_2}^{n_2} f(z_1, z_2) = \sqrt{\pi} (2z_1)^{n_2} \exp(2z_1 z_2). \tag{B.1.12}$$

Um dies wiederum n_1-mal nach z_1 abzuleiten, erinnern wir uns an die Leibnizsche Produktformel

$$\partial_{z_1}^{n_1} F(z_1) G(z_1) = \sum_{k=0}^{n_1} \binom{n_1}{k} [\partial_{x_1}^k F(z_1)][\partial_{x_1}^{n_1 - k} G(z_1)]. \tag{B.1.13}$$

Wenden wir dies auf (B.1.12) mit $F(z_1) = (2z_1)^{n_2}$ und $G(z_1) = \exp(2z_1 z_2)$ an, folgt

$$\partial_{z_1}^{n_1} \partial_{z_2}^{n_2} f(z_1, z_2) = \sqrt{\pi} 2^{n_2} \sum_{k=0}^{\infty} \binom{n_1}{k} (\partial_{z_1}^k z_1^{n_2}) [\partial_{z_1}^{n_1 - k} \exp(2z_1 z_2)]. \tag{B.1.14}$$

Gemäß (B.1.11) wollen wir dies bei $z_1 = z_2 = 0$ auswerten. Jede Ableitung von $\exp(2z_1 z_2)$ nach z_1 liefert einen Faktor $2z_2$, d.h. alle Beiträge in der Summe (B.1.14) außer derjenige für $k = n_1$ verschwindet. Es ist also

$$\partial_{z_1}^{n_1} \partial_{z_2}^{n_2} f(z_1, z_2) \Big|_{z_1 = z_2 = 0} = \sqrt{\pi}\, 2^{n_2} \partial_{z_1}^{n_1} z_1^{n_2} \Big|_{z_1 = 0}. \tag{B.1.15}$$

Nun verschwindet für $n_1 > n_2$ die Ableitung in (B.1.15) und für $n_1 < n_2$ wird die Ableitung $\propto z_1^{n_2-n_1}$ und verschwindet folglich für $z_1 = 0$. Also gilt

$$\partial_{z_1}^{n_1}\partial_{z_2}^{n_2} f(z_1,z_2)\Big|_{z_1=z_2=0} = \sqrt{\pi}2^{n_1}\delta_{n_1n_2}\partial_{z_1}^{n_1} z_1^{n_1} = \sqrt{\pi}2^{n_1}n_1!\delta_{n_1n_2}. \tag{B.1.16}$$

Wegen (4.1.28) liefert dies

$$\int_{-\infty}^{\infty} \mathrm{d}x \exp(-x^2)\mathrm{H}_{n_1}(x)\mathrm{H}_{n_2}(x) = \sqrt{\pi}2^{n_1}n_1!\delta_{n_1n_2}. \tag{B.1.17}$$

Dies stimmt damit überein, dass die Energieeigenfunktionen des harmonischen Oszillators (3.9.28) orthonormiert sind (*Nachrechnen!*), wie wir es mit Hilfe der Leiteroperatoren auf praktisch rein algebraischem Wege berechnet haben.

Eine nützliche Darstellung der Hermite-Polynome als Integral ergibt sich, indem man die Gaußfunktion in der folgenden Fourierdarstellung schreibt

$$\exp(-\xi^2) = \frac{1}{\sqrt{\pi}}\int_{\mathbb{R}} \mathrm{d}u \exp(-u^2 + 2\mathrm{i}u\xi) \tag{B.1.18}$$

und in die Rodrigues-Formel (B.1.3) einsetzt:

$$\mathrm{H}_n(\xi) = \frac{\exp(\xi^2)}{\sqrt{\pi}}\int_{\mathbb{R}} \mathrm{d}u(-2\mathrm{i}u)^n \exp(-u^2 + 2\mathrm{i}u\xi). \tag{B.1.19}$$

Wir können weiter mit Hilfe der erzeugenden Funktion (B.1.7) Rekursionsformeln für die Hermite-Polynome herleiten. Differenzieren wir die entsprechende Gleichung

$$\exp(2\xi z - z^2) = \sum_{n=0}^{\infty}\frac{1}{n!}\mathrm{H}_n(\xi)z^n. \tag{B.1.20}$$

nach z, finden wir

$$2(\xi - z)\exp(2\xi z - z^2) = \sum_{n=1}^{\infty}\frac{1}{(n-1)!}\mathrm{H}_n(\xi)z^{n-1} = \sum_{n=0}^{\infty}\frac{1}{n!}\mathrm{H}_{n+1}(\xi)z^n. \tag{B.1.21}$$

Multiplizieren wir andererseits (B.1.20) mit $2(\xi - z)$ und vergleichen die Koeffizienten von z^n, finden wir die Rekursionsformel

$$\mathrm{H}_1(\xi) = 2\xi\mathrm{H}_0(\xi), \quad \mathrm{H}_{n+1}(\xi) = 2\xi\mathrm{H}_n(\xi) - 2n\mathrm{H}_{n-1}(\xi), \tag{B.1.22}$$

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

Auf dieselbe Weise erhalten wir durch Ableiten von (B.1.20) nach ξ die Ableitungen der Hermite-Polynome

$$\mathrm{H}_n'(\xi) = \begin{cases} 0 & \text{für } n = 0, \\ 2n\mathrm{H}_{n-1}(\xi) & \text{für } n \geq 1. \end{cases} \tag{B.1.23}$$

B.1.2 Die Green-Funktion des harmonischen Oszillators

Als Anwendung berechnen wir die Green-Funktion des harmonischen Oszillators. Dabei verstehen wir unter der Green-Funktion $G(t,x,x')$ die Lösung der zeitabhängigen Schrödinger-Gleichung

$$\mathrm{i}\hbar\partial_t\psi(t,x) = \mathbf{H}\psi(t,x) \tag{B.1.24}$$

in der Form

$$\psi(t,x)=\int_{\mathbb{R}} \mathrm{d}x' G(t,x,x')\psi_0(x') \quad \text{mit} \quad \psi(t=0,x)=\psi_0(x). \tag{B.1.25}$$

Offenbar muss nun

$$\mathrm{i}\hbar\partial_t G(t,x,x')=\mathbf{H}_x G(t,x,x'), \quad G(0,x,x')=\delta(x-x') \tag{B.1.26}$$

erfüllen. Dabei bedeutet $\mathbf{H}_x$, dass der Hamilton-Operator, der auf die Abhängigkeit von $G(t,x,x')$ bzgl. x (und *nicht* x') wirken soll.

Wir wollen nun versuchen, diese Green-Funktion nach dem vollständigen Satz von Energie-Eigenfunktionen $u_n(x)$ zu entwickeln, d.h. in der Form

$$G(t,x,x')=\sum_{n=0}^{\infty} g_n(t,x')u_n(x). \tag{B.1.27}$$

Da $\mathbf{H}u_n=E_n u_n$ ist, folgt aus (B.1.19)

$$\mathrm{i}\hbar\partial_t g_n(t,x')=E_n g_n(t,x') \Rightarrow g_n(t,x')=C_n(x')\exp(-\mathrm{i}E_n t/\hbar). \tag{B.1.28}$$

Damit wird (B.1.24) zu

$$G(t,x,x')=\sum_{n=0}^{\infty} C_n(x')u_n(x)\exp(-\mathrm{i}E_n t/\hbar). \tag{B.1.29}$$

Um auch die Anfangsbedingung zu erfüllen, bemerken wir, dass die Vollständigkeit der orthonormierten Energie-Eigenfunktionen

$$\sum_{n=0}^{\infty} u_n^*(x')u_n(x)=\delta(x-x') \tag{B.1.30}$$

impliziert. Setzt man in (B.1.27) $t=0$, folgt aus $G(t=0,x,x')=\delta(x-x')$, dass $C_n(x')=u_n^*(x')$ sein muss, d.h.

$$G(t,x,x')=\sum_{n=0}^{\infty} u_n^*(x')u_n(x)\exp(-\mathrm{i}E_n t/\hbar). \tag{B.1.31}$$

Gemäß (3.9.28) sind die Energie-Eigenfunktionen durch

$$u_n(x)=\left(\frac{1}{\pi a^2}\right)^{1/4}\sqrt{\frac{1}{2^n n!}}\left[\mathrm{H}_n(\xi)\exp\left(-\frac{\xi^2}{2}\right)\right]_{\xi=\frac{x}{a}} \tag{B.1.32}$$

gegeben. Setzen wir dies in (B.1.31) ein, erhalten wir wegen $E_n=\hbar\omega(1/2+n)$

$$\begin{aligned} G(t;x,x')&=\frac{1}{\sqrt{\pi}a}\exp\left(-\frac{\mathrm{i}\omega t}{2}\right)\sum_{n=0}^{\infty}\frac{1}{2^n n!}\mathrm{H}_n(\xi)\mathrm{H}_n(\xi')\exp\left(-\frac{\xi^2+\xi'^2}{2}-\mathrm{i}n\omega t\right)\\ &\text{mit} \quad \xi=\frac{x}{a}, \quad \xi'=\frac{x'}{a}. \end{aligned} \tag{B.1.33}$$

Um die Summe explizit auszuführen, substituieren wir die Integraldarstellung der Hermite-Polynome (B.1.19)

$$G(t;x,x')=\sqrt{\frac{\alpha}{\pi^3}}\frac{1}{a}\exp\left(\frac{\xi^2+\xi'^2}{2}\right)\int_{\mathbb{R}}\mathrm{d}u\int_{\mathbb{R}}\mathrm{d}v\sum_{n=0}^{\infty}\frac{(-2\alpha uv)^n}{n!}\exp(-u^2-v^2+2\mathrm{i}\xi u+2\mathrm{i}\xi' v) \tag{B.1.34}$$

mit $\alpha=\exp(-\mathrm{i}\omega t)$. Die Summe stellt gerade die Exponentialreihe dar, so daß wir für sie sofort

$$G(t;x,x')=\sqrt{\frac{\alpha}{\pi^3}}\frac{1}{a}\exp\left(\frac{\xi^2+\xi'^2}{2}\right)\int_{\mathbb{R}^3}\mathrm{d}u\int_{\mathbb{R}^3}\mathrm{d}v\exp(-u^2-v^2+2\mathrm{i}\xi u+2\mathrm{i}\xi' v-2\alpha uv) \tag{B.1.35}$$

erhalten. Die beiden Gauß-Integrale lassen sich nacheinander ausführen, und das Ergebnis lautet (*nachrechnen!*)

$$G(t;x,x') = \frac{1}{\sqrt{\pi}a}\sqrt{\frac{\alpha}{1-\alpha^2}}\exp\left(-\frac{\xi^2+\xi'^2}{2}\frac{1+\alpha^2}{1-\alpha^2}+\frac{2\alpha}{1-\alpha^2}\xi\xi'\right). \tag{B.1.36}$$

Wegen

$$\begin{aligned}\frac{\alpha}{1-\alpha^2} &= \frac{1}{1/\alpha-\alpha} = \frac{1}{\exp(\mathrm{i}\omega t)-\exp(-\mathrm{i}\omega t)} = \frac{1}{2\mathrm{i}\sin(\omega t)},\\ \frac{1+\alpha^2}{1-\alpha^2} &= \frac{1/\alpha+\alpha}{1/\alpha-\alpha} = -\mathrm{i}\frac{\cos(\omega t)}{\sin(\omega t)}\end{aligned} \tag{B.1.37}$$

ergibt sich daraus

$$G(t;x,x') = \sqrt{\frac{1}{2\pi\mathrm{i}\sin(\omega t)}}\frac{1}{a}\exp\left\{\frac{\mathrm{i}}{2a^2\sin(\omega t)}[(x^2+x'^2)\cos(\omega t)-2xx']\right\}. \tag{B.1.38}$$

B.2 Sphärische Bessel-, Neumann- und Hankel-Funktionen

In diesem Anhang betrachten wir die Lösungen der Radialgleichung für die Energie-Drehimpuls-Eigenzustände (also die simultanen Eigenvektoren für $\mathbf{H}$, $\vec{\mathbf{L}}^2$ und $\mathbf{L}_3$) für das freie Teilchen genauer. In Abschnitt 3.12 haben wir für den Radialanteil $R_{E\ell}(r)$ der entsprechenden Eigenfunktionen (3.12.7) hergeleitet. Mit $r = k\rho$ lautet die entsprechende Differentialgleichung

$$\frac{1}{\rho}(\rho u_\ell)'' - \frac{\ell(\ell+1)}{\rho^2}u_\ell + u_\ell = 0 \tag{B.2.1}$$

mit der Drehimpulsquantenzahl $\ell \in \mathbb{N}_0$. Diese Gleichung ist von der Fuchsschen Klasse und lässt sich demnach mit dem verallgemeinerten Potenzreihenansatz

$$u_\ell(\rho) = \sum_{j=0}^{\infty} a_j \rho^{j+\lambda} \tag{B.2.2}$$

lösen. Dabei nehmen wir $a_0 \neq 0$ an, weil man anderernfalls nur λ umdefinieren müsst, um diese Bedingung zu erreichen. Einsetzen des Reihenansatzes in die Differentialgleichung (B.2.1) liefert die Rekursionsformeln

$$\begin{aligned}&a_0[\lambda(\lambda+1)-\ell(\ell+1)=0, \quad a_1[(\lambda+1)(\lambda+2)-\ell(\ell+1)]=0,\\ &a_j = -\frac{a_{j-2}}{(j+\lambda)(j+\lambda+1-\ell(\ell+1))}.\end{aligned} \tag{B.2.3}$$

Da $a_0 \neq 0$ sein soll, muss

$$\lambda = \ell \quad \text{quad} \quad \lambda = -\ell-1 \tag{B.2.4}$$

und demnach notwendig $a_1 = 0$ sein. Wir erhalten für diese beiden Werte von λ also zwei linear unabhängige Lösungen, von denen die erste mit $\lambda = \ell$ die in 0 analytische Lösung ist. Sie ist auch die einzige Lösung, die gleichzeitig die Randbedingung $rR_{k\ell}(r) \to 0$ für $r \to 0$ erfüllt.
Mit der Wahl

$$a_0 = \frac{2^\ell \ell!}{(2\ell+1)} \tag{B.2.5}$$

definiert dies die **sphärischen Besselfunktionen**

$$\mathrm{j}_\ell(\rho) = (2\rho)^\ell \sum_{k=0}^{\infty} (-1)^k \frac{(\ell+k)!}{k!(2k+2\ell+1)!} \rho^{2k}. \tag{B.2.6}$$

Mit dem Quotientenkriterium für Potenzreihen erkennt man sofort, dass der Konvergenzradius ∞ ist, und somit die j_ℓ überall analytische Funktionen sind. Diese bilden den vollständigen Satz von verallgemeinerten Radialfunktionen $R_\ell(r) = N_\ell \mathrm{j}_\ell(kr)$ für das freie Teilchen. Die weiteren Details werden in Abschnitt 3.12 ausgearbeitet.

Hier beschäftigen wir uns auch mit den irregulären Lösungen, die für $\rho \to 0$ mit der Ordnung $\mathcal{O}(r^{-\ell-1})$ divergieren. Einfacher als eine Berechnung der entsprechenden für diese zweite linear unabhängige Lösung von (B.2.1) mittels der verallgemeinerten Potenzreihe ist die folgende Betrachtung.

Als erstes wollen wir eine Rekursionsformel für allgemeine Lösungen $u_\ell(\rho)$ herleiten. Dazu beachten wir, dass man durch gliedweises Ableiten mittels der Rekursionsformel nach einigen einfachen Umformungen (*Übung*) für die sphärischen Besselfunktionen die Rekursionsformel

$$j_{\ell+1}(\rho) = -\rho^\ell \frac{\mathrm{d}}{\mathrm{d}\rho}\left[\frac{1}{\rho^\ell} j_\ell(\rho)\right] \tag{B.2.7}$$

gilt. Es ist nun durch direktes Einsetzen in (B.2.1) zu zeigen, dass für jede Lösung u_ℓ

$$u_{\ell+1}(\rho) = -\rho^\ell \frac{\mathrm{d}}{\mathrm{d}\rho}\left[\frac{1}{\rho^\ell} u_\ell\right] \tag{B.2.8}$$

eine Lösung dieser Differentialgleichung mit $\ell \to \ell+1$ ist. Dazu muss man nur den Ansatz (B.2.8) in die entsprechende DGL einsetzen und verwenden, dass für u_ℓ die Differentialgleichung gilt. Daraus folgt insbesondere, dass

$$u_\ell'''(\rho) = -\frac{2\ell(\ell+1)}{\rho^3} u_\ell(\rho) + \left[\frac{\ell(\ell+1)}{\rho^2} - 1\right] u_\ell'(\rho) - \frac{2}{\rho} u_\ell''(\rho) \tag{B.2.9}$$

ist. Verwendet man dies, um u_ℓ''' zu eliminieren und verwendet nochmals, dass u_ℓ (B.2.1) erfüllt, ergibt sich, dass $u_{\ell+1}$ eine Lösung für diese DGL mit $(\ell+1)$ statt ℓ ist. Die Rekursionsformel (B.2.8) liefert also für beliebige Lösungen

$$u_\ell(\rho) = (-1)^\ell \rho^\ell \left(\frac{1}{\rho}\frac{\mathrm{d}}{\mathrm{d}\rho}\right)^\ell u_0(\rho). \tag{B.2.10}$$

Offenbar erhält man für zwei linear unabhängige Lösungen der DGL (B.2.1) für $\ell = 0$ mit dieser Gleichung zwei linear unabhängige Lösungen für ℓ. Für $\ell = 0$ lautet (B.2.1) nach Multiplikation mit ρ

$$(\rho u_0)'' + \rho u_0 = 0. \tag{B.2.11}$$

Als die beiden linear unabhängigen Lösungen (*Nachrechnen*) wählt man die beiden **sphärischen Hankelfunktionen**

$$\mathrm{h}_0^{(1)}(\rho) = \frac{\exp(\mathrm{i}\rho)}{\mathrm{i}\rho}, \quad \mathrm{h}_0^{(2)}(\rho) = \frac{\exp(\mathrm{i}\rho)}{-\mathrm{i}\rho}, \tag{B.2.12}$$

wobei die Wahl des konstanten Faktors eine Konvention ist. Die in $\rho = 0$ reguläre Lösung, also die sphärische Bessel-Funktion j_0, ist offenbar durch

$$\mathrm{j}_0(\rho) = \mathrm{Re}\,\mathrm{h}_0^{(1)}(\rho) = \frac{\sin\rho}{\rho} \tag{B.2.13}$$

gegeben. Entsprechend definiert man die andere irreguläre reelle Lösung, die wie $\mathcal{O}(1/\rho)$ für $\rho \to 0$ geht, als die **sphärische Neumann-Funktion**

$$\mathrm{n}_0(\rho) = \operatorname{Im} \mathrm{h}_0^{(1)}(\rho) = -\frac{\cos\rho}{\rho}. \tag{B.2.14}$$

Die sphärischen Hankelfunktionen sind nun am bequemsten mittels der Rekursionsformel (B.2.10) zu berechnen. Wir benötigen auch nur $h_\ell^{(1)}$, da für $\rho \in \mathbb{R}$ offenbar $h_\ell^{(2)}(\rho) = h_\ell^{(1)*}(\rho)$ gilt. Offenbar können wir (B.2.10) mit (B.2.12) umschreiben zu (*Nachprüfen!*)

$$h_\ell^{(1)}(\rho) = -\mathrm{i}(-1)^\ell \rho^{\ell-1} \left(\frac{\mathrm{d}}{\mathrm{d}\rho}\frac{1}{\rho}\right)^\ell \exp(\mathrm{i}\rho). \tag{B.2.15}$$

In der Entwicklung nach Potenzen von ρ erhält man den führenden Term $\mathcal{O}(\rho^{-\ell-1})$ für $\rho \to 0$ offenbar, indem man die Ableitungen immer nur auf die Faktoren $1/\rho$ anwendet, d.h. für $\ell \geq 1$ (*Nachrechnen*)

$$h_\ell^{(1)}(\rho) = -\mathrm{i}(-1)^\ell \rho^{\ell-1} \exp(\mathrm{i}k\rho)\left[\left(\frac{\mathrm{d}}{\mathrm{d}\rho}\frac{1}{\rho}\right)^\ell + \mathcal{O}\left(\frac{1}{\rho^\ell}\right)\right] = -\mathrm{i}\frac{\exp(\mathrm{i}\rho)}{x^{\ell+1}}\left[\frac{(2\ell)!}{2^\ell l!} + \mathcal{O}(\rho)\right]. \tag{B.2.16}$$

Daraus wird klar, dass in dem Ansatz

$$h_\ell^{(1)}(\rho) = \frac{\exp(\mathrm{i}\rho)}{\rho^{\ell+1}}\eta_\ell(\rho) \tag{B.2.17}$$

die Funktion $\eta_\ell(\rho)$ ein Polynom vom Grad ρ sein muss, für das

$$\eta_\ell(0) = -\mathrm{i}\frac{(2\ell)!}{2^\ell \ell!} \tag{B.2.18}$$

ist. Geht man mit dem Ansatz (B.2.17) in die Differentialgleichung (B.2.1) ein, erhält man die DGL

$$\rho\eta_\ell''(\rho) + 2(\mathrm{i}\rho - \ell)\rho_\ell'(\rho) - 2\mathrm{i}\ell\eta_\ell(\rho) = 0. \tag{B.2.19}$$

Da wir schon wissen, dass η_ℓ ein Polynom vom Grade ℓ ist, können wir mit dem einfachen Potenzreihenansatz

$$\eta_\ell(\rho) = \sum_{j=0}^{\ell} a_j \rho^j \tag{B.2.20}$$

in diese DGL gehen, und erhalten als Rekursionsformel (*Nachrechnen!*)

$$a_{j+1} = -2\mathrm{i}\frac{(\ell - j)}{(j+1)(2\ell - j)}a_j. \tag{B.2.21}$$

Es ist klar, dass tatsächlich $a_j = 0$ für $j \geq \ell + 1$, wie zu erwarten. Mit (B.2.18) erhält man a_0, und die Rekursionsgleichung (B.2.21) liefert für $j \in \{0, 1, \ldots, \ell\}$ (*Nachrechnen!*)

$$a_j = \frac{2^{j-\ell}(-\mathrm{i})^{j+1}(2\ell - j)!}{j!(\ell - j)!}. \tag{B.2.22}$$

Damit ist

$$\mathrm{h}_\ell^{(1)}(\rho) = \frac{\exp(\mathrm{i}\rho)}{\rho^{\ell+1}}\sum_{j=0}^{\ell}\frac{(-\mathrm{i})^{j+1}2^{j-\ell}(2\ell - j)!}{j!(\ell - j)!}\rho^j. \tag{B.2.23}$$

Eine andere manchmal bequemere Darstellung erhält man, indem man $k = \ell - j$ als Summationsvariable einführt (*Nachrechnen!*)

$$\mathrm{h}_\ell^{(1)}(\rho) = \frac{(-\mathrm{i})^{\ell+1}\exp(\mathrm{i}\rho)}{\rho}\sum_{k=0}^{\ell}\frac{(k+\ell)!}{k!(\ell-k)!}\left(\frac{\mathrm{i}}{2\rho}\right)^j. \tag{B.2.24}$$

Aus (B.2.23) erhält man für die asymptotische Form für $\rho \to 0$ aus dem Term mit $j = 0$

$$\mathrm{h}_\ell^{(1)}(\rho) \underset{\rho\to 0}{\cong} -\mathrm{i}\frac{(2\ell)!}{2^\ell \ell!}\frac{1}{\rho^{\ell+1}}, \tag{B.2.25}$$

und aus (B.2.25) die asymptotische Formel für $\rho \to \infty$ aus dem Term für $k = 0$

$$\mathrm{h}_\ell^{(1)}(\rho) \underset{\rho\to\infty}{\cong} \frac{(-\mathrm{i})^{\ell+1}\exp(\mathrm{i}\rho)}{\rho} = -\frac{\mathrm{i}}{\rho}\exp[\mathrm{i}(\rho - \ell\pi/2)]. \tag{B.2.26}$$

Die entsprechenden asymptotischen Formeln für die sphärischen Bessel- und Neumann-Funktionen folgen aus (B.2.13), (B.2.14) und (B.2.10):

$$\mathrm{j}_\ell(\rho) = \mathrm{Re}\,\mathrm{h}_\ell^{(1)}(\rho) \underset{\rho\to\infty}{\cong} \frac{\sin(\rho - \ell\pi/2)}{\rho}, \quad \mathrm{n}_\ell(\rho) = \mathrm{Im}\,\mathrm{h}_\ell^{(1)}(\rho) \underset{\rho\to\infty}{\cong} -\frac{\cos(\rho - \ell\pi/2)}{\rho}. \tag{B.2.27}$$

Für $\rho \to 0$ erhalten wir das asymptotische Verhalten für j_ℓ offenbar am einfachsten aus der Reihe (B.2.6) mit $j = 0$:

$$\mathrm{j}_\ell(\rho) \underset{\rho\to 0}{\cong} \frac{2^\ell \ell!}{(2\ell+1)!}\rho^\ell. \tag{B.2.28}$$

Für die sphärische Neumann-Funktion kann man hingegen (B.2.25) verwenden:

$$\mathrm{n}_\ell(\rho) = \mathrm{Im}\,\mathrm{h}_\ell^{(1)}(\rho) \underset{\rho\to 0}{\cong} -\frac{(2\ell)!}{2^\ell \ell!}\frac{1}{\rho^{\ell+1}}. \tag{B.2.29}$$

Aus (B.2.8) folgt für die Ableitung beliebiger Lösungen von (B.2.1), also auch für die spärischen Bessel-, Neumann- und Hankel-Funktionen

$$u_\ell'(\rho) = \frac{\ell}{\rho}u_\ell(\rho) - u_\ell'(\rho). \tag{B.2.30}$$

Genau wie die Rekursionsformel (B.2.8) kann man für $\ell \geq 1$ auch die Formel

$$u_{\ell-1}(\rho) = \frac{1}{\rho^{\ell+1}}\frac{\mathrm{d}}{\mathrm{d}\rho}\left[\rho^{\ell+1}u_\ell(\rho)\right] \tag{B.2.31}$$

herleiten. Daraus folgt

$$u_\ell'(\rho) = u_\ell'(\rho) - \frac{\ell+1}{\rho}u_\ell(\rho). \tag{B.2.32}$$

Aus (B.2.31) und (B.2.32) folgt für $\ell \geq 1$ daraus die Rekursionsformel

$$u_{\ell+1}(\rho) = \frac{2\ell+1}{\rho}u_\ell(\rho) - u_{\ell-1}(\rho). \tag{B.2.33}$$

B.3 Kugelflächenfunktionen

Wir haben die Kugelflächenfunktionen in Abschnitt 3.11 unter Verwendung der Methode der Auf- und Absteigeoperatoren $\mathbf{l}_\pm$ berechnet. In diesem Anhang wollen wir noch einige interessante Eigenschaften herleiten. Dabei bedienen wir uns ausgiebig physikalisch motivierter Methoden.

B.3.1 Green-Funktion des Laplace-Operators

Als erstes betrachten wir die **Green-Funktion des Laplace-Operators**, die allerdings eher in der Elektrostatik (vgl. [Hee18c]) als in der Quantentheorie benötigt wird. Wir werden sehen, dass wir daraus eine elegante Formulierung der **Multipolentwicklung** elektrostatischer Potentiale bzw. Felder erhalten.
Die Green-Funktion des Laplaceoperators ist durch

$$-\Delta G(\vec{x},\vec{x}') = \delta^{(3)}(\vec{x}-\vec{x}') \tag{B.3.1}$$

definiert. Dabei beziehen sich partielle Ableitungen, also $\vec{\nabla}$ und Δ, auf die Koordinaten $\vec{x}$. Wollen wir nach Komponenten von $\vec{x}'$ ableiten, schreiben wir $\vec{\nabla}'$ bzw. Δ'.
Wir erinnern daran, dass in der **Elektrostatik** das Potential der **Poisson-Gleichung**

$$-\Delta\Phi(\vec{x}) = \frac{1}{\epsilon_0}\rho(\vec{x}) \tag{B.3.2}$$

genügt und die Lösung

$$\Phi(\vec{x}) = \frac{1}{\epsilon_0}\int_{\mathbb{R}^3} \mathrm{d}^3\vec{x}'\,\frac{\rho(\vec{x}')}{4\pi|\vec{x}-\vec{x}'|} = \frac{1}{\epsilon_0}\int_{\mathbb{R}^3} \mathrm{d}^3\vec{x}'\,G(\vec{x},\vec{x}')\rho(\vec{x}') \tag{B.3.3}$$

besitzt. Da dies für alle Ladungsverteilungen $\rho(\vec{x})$ gilt, folgern wir, dass

$$G(\vec{x},\vec{x}') = \frac{1}{4\pi|\vec{x}-\vec{x}'|} \tag{B.3.4}$$

gilt.
Wir leiten dieses Resultat noch einmal formal aus (B.3.1) her. Die rechte Seite von (B.3.1) legt den Ansatz

$$G(\vec{x},\vec{x}') = g(\vec{x}-\vec{x}') \tag{B.3.5}$$

nahe. Setzen wir also $\vec{r} = \vec{x}-\vec{x}'$, können wir (B.3.1) in der Form

$$-\Delta g(\vec{r}) = \delta^{(3)}(\vec{r}) \tag{B.3.6}$$

schreiben. Nun ist die rechte Seite auch unter beliebigen Drehungen invariant, d.h. wir dürfen erwarten, dass $g(\vec{r})$ tatsächlich nur von $r = |\vec{r}|$ abhängt. Außer für $\vec{r} = 0$ gilt

$$\Delta g(r) = \frac{1}{r}\partial_r^2[r\,g(r)] = 0. \tag{B.3.7}$$

Dies lässt sich durch zweimaliges Integrieren leicht lösen:

$$g(r) = \frac{A}{r} + B. \tag{B.3.8}$$

Dabei sind A und B Integrationskonstanten. Es ist klar, dass wir $g(r) \to 0$ für $r \to \infty$ haben wollen, d.h. es ist $B = 0$.
Um nun A zu bestimmen, integrieren wir (B.3.6) über eine Kugel um den Ursprung und wenden den Gaußschen Integralsatz an. Es folgt

$$-\int_{K_R} \mathrm{d}^3 r\,\Delta g(\vec{r}) = -\int_{\partial K_R} \mathrm{d}^2\vec{f}\cdot\vec{\nabla} g(\vec{r}) \overset{!}{=} \int_{K_R} \mathrm{d}^3 r\,\delta^{(3)}(\vec{r}) = 1. \tag{B.3.9}$$

Nun ist

$$\mathrm{d}^2\vec{f} = \mathrm{d}\vartheta\mathrm{d}\varphi\frac{\vec{r}}{r}R^2\sin\vartheta, \quad \nabla g(\vec{r}) = -\frac{A\vec{r}}{r^3} \tag{B.3.10}$$

und damit

$$-\int_{\partial K_r}\mathrm{d}^2\vec{f}\cdot\vec{\nabla}g(\vec{r}) = \int_0^\pi \mathrm{d}\vartheta\int_0^{2\pi}\mathrm{d}\varphi\sin\vartheta A = 4\pi A. \tag{B.3.11}$$

Mit (B.3.9) folgt also $A = 1/(4\pi)$, und damit (B.3.4).
Weiter wollen wir die Green-Funktion nach Kugelflächenfunktionen entwickeln, d.h. wir machen den Ansatz

$$G(\vec{x},\vec{x}') = \sum_{\ell=0}^{\infty}\sum_{m=-\ell}^{\ell} g_{\ell m}(r,\vec{r}')\mathrm{Y}_{\ell m}(\vartheta,\varphi). \tag{B.3.12}$$

Um die Funktionen $g_{\ell m}$ zu finden, bemerken wir, dass die $\mathrm{Y}_{\ell m}$ ein vollständiges orthonormiertes Funktionensystem für quadratintegrable Funktionen $f(\vartheta,\varphi)$, die auf der Einheitskugeloberfläche Ω definiert sind, bilden. Dabei ist das Skalarprodukt durch

$$\langle f_1|f_2\rangle_\Omega = \int_0^\pi \mathrm{d}\vartheta\int_0^{2\pi}\mathrm{d}\varphi\sin\vartheta f_1^*(\vartheta,\varphi)f_2(\vartheta,\varphi) \tag{B.3.13}$$

definiert (vgl. Abschnitt 3.11). Daraus folgt, dass man jede solche Funktion in der Form

$$f(\vartheta,\varphi) = \sum_{\ell=0}^{\infty}\sum_{m=-\ell}^{\ell} f_{\ell m}\mathrm{Y}_{\ell m}(\vartheta,\varphi) \tag{B.3.14}$$

schreiben kann und die entsprechenden Koeffizienten durch

$$f_{\ell m} = \langle \mathrm{Y}_{\ell m}|f\rangle_\Omega = \int_0^\pi \mathrm{d}\vartheta'\int_0^{2\pi}\mathrm{d}\varphi'\sin\vartheta\mathrm{Y}_{\ell m}^*(\vartheta',\varphi')f(\vartheta',\varphi') \tag{B.3.15}$$

gegeben sind. Setzt man das in (B.3.14) ein und vertauscht Integration und Summation, erhält man

$$f(\vartheta,\varphi) = \int_0^\pi \mathrm{d}\vartheta'\int_0^{2\pi}\mathrm{d}\varphi'\sin\vartheta' f(\vartheta',\varphi')\sum_{\ell=0}^{\infty}\sum_{m=-\ell}^{\ell}\mathrm{Y}_{\ell m}^*(\vartheta',\varphi')\mathrm{Y}_{\ell m}(\vartheta,\varphi). \tag{B.3.16}$$

Daraus folgt, dass

$$\sum_{\ell=0}^{\infty}\sum_{m=-\ell}^{\ell}\mathrm{Y}_{\ell m}^*(\vartheta',\varphi')\mathrm{Y}_{\ell m}(\vartheta,\varphi) = \frac{1}{\sin\vartheta}\delta(\vartheta-\vartheta')\delta(\varphi-\varphi') \tag{B.3.17}$$

ist.
Nun gilt mit dem Ansatz (B.3.12) wegen (3.11.27)

$$\begin{aligned}
-\Delta G(\vec{x},\vec{x}') &= -\left\{\frac{1}{r}\partial_r^2[rG(\vec{x},\vec{x}')] - \frac{\vec{\mathrm{l}}^2}{r^2}G(\vec{x},\vec{x}')\right\} \\
&= -\sum_{\ell=0}^{\infty}\sum_{m=-\ell}^{\ell}\left\{\frac{1}{r}\partial_r^2[r\,g_{\ell m}(r,\vec{r}')] - \frac{\ell(\ell+1)}{r^2}g_{\ell m}(r,\vec{r}')\right\}\mathrm{Y}_{\ell m}(\vartheta,\varphi) \\
&= \delta^{(3)}(\vec{x}-\vec{x}') = \frac{1}{r^2\sin\vartheta}\delta(r-r')\delta(\vartheta-\vartheta')\delta(\varphi-\varphi').
\end{aligned} \tag{B.3.18}$$

Dies legt wegen (B.3.17) den Ansatz

$$g_{\ell m}(r,\vec{r}\,') = h_{\ell m}(r,r')\mathrm{Y}^*_{\ell m}(\vartheta',\varphi') \tag{B.3.19}$$

nahe. Setzen wir dies nämlich in (B.3.18) ein, folgt

$$\begin{aligned} -\sum_{\ell=0}^{\infty}\sum_{m=-\ell}^{\ell}\left\{\frac{1}{r}\partial_r^2[r h_{\ell m}(r,r')] - \frac{\ell(\ell+1)}{r^2}h_{\ell m}(r,r')\right\}\mathrm{Y}^*_{\ell m}(\vartheta',\varphi')\mathrm{Y}_{\ell m}(\vartheta,\varphi) \\ = \frac{1}{r^2\sin\vartheta}\delta(r-r')\delta(\vartheta-\vartheta')\delta(\varphi-\varphi'). \end{aligned} \tag{B.3.20}$$

Der Vergleich mit (B.3.17) zeigt dann, dass wir diese Gleichung erfüllen, wenn wir verlangen, dass

$$-\left\{\frac{1}{r}\partial_r^2[r h_{\ell m}(r,r')] - \frac{\ell(\ell+1)}{r^2}h_{\ell m}(r,r')\right\} = \frac{1}{r^2}\delta(r-r') \tag{B.3.21}$$

gilt. Daraus folgt sofort, dass $h_{\ell m}$ nicht von m abhängt, d.h. wir können $h_{\ell m} = h_\ell$ setzen. Für $(r \neq r')$ gilt

$$\partial_r^2[r h_\ell(r,r')] = \frac{\ell(\ell+1)}{r}h_\ell(r,r'). \tag{B.3.22}$$

Setzen wir hierin $h_\ell(r,r') = \chi_\ell(r,r')/r$, folgt

$$\partial_r^2\chi_\ell = \frac{\ell(\ell+1)}{r^2}\chi_\ell. \tag{B.3.23}$$

Diese Gleichung lässt sich durch den Ansatz $\chi_\ell(r,r') = A(r')r^{\lambda+1}$ lösen. Setzen wir dies nämlich in (B.3.23 ein, folgt

$$\lambda(\lambda+1) = \ell(\ell+1) \Rightarrow \lambda_1 = \ell, \quad \lambda_2 = -(\ell+1). \tag{B.3.24}$$

Da nun gemäß (B.3.21) $h_\ell(r,r')$ bei $r = r'$ eine Singularität aufweisen muss und die Green-Funktion (B.3.4) für $\vec{x}' \neq 0$ in $\vec{x} = 0$ keine Singularität aufweist und für beliebiges festgehaltenes $\vec{x}'$ im Limes $\vec{x} \to \infty$ verschwindet, folgt, dass

$$h_\ell(r,r') = Ar^\ell\Theta(r'-r) + \frac{B}{r^{\ell+1}}\Theta(r-r') \tag{B.3.25}$$

sein muss. Wir müssen nun nur noch die Integrationskonstanten A und B aus (B.3.21) bestimmen. Da die δ-Distribution aus der höchsten vorkommenden Ableitung auf der linken Seite, also $\partial_r^2 h_{\ell m}(r,r')$ stammen muss, folgt, dass $\partial_r h_\ell(r,r')$ bei $r = h$ einen Sprung und $h_\ell(r,r')$ stetig bei $r = r'$ sein müssen, d.h. es muss

$$Ar'^\ell = \frac{B}{r'^{\ell+1}} \Rightarrow B = Ar'^{2\ell+1} \tag{B.3.26}$$

sein und folglich

$$h_\ell(r,r') = A\left[\Theta(r'-r)r^\ell + \Theta(r-r')\frac{r'^{2\ell+1}}{r^{\ell+1}}\right]. \tag{B.3.27}$$

Eine zweite Bedingung erhalten wir, indem wir (B.3.21) zuerst mit r^2 multiplizieren, dann bzgl. r über ein kleines Intervall $(r'-\epsilon, r'+\epsilon)$ integrieren und schließlich $\epsilon \to 0^+$ bilden. Daraus folgt

$$Ar'^l(2\ell+1) = \frac{1}{r'} \Rightarrow A = \frac{1}{(2\ell+1)r'^{\ell+1}}. \tag{B.3.28}$$

Setzen wir dies in (B.3.27) ein, können wir das Endergebnis in der kompakten Form

$$h_\ell(r,r')\frac{1}{2\ell+1}\frac{r_<^\ell}{r_>^{\ell+1}} \quad \text{mit} \quad r_< = \min(r,r'), \quad r_> = \max(r,r') \tag{B.3.29}$$

schreiben. Damit erhalten wir die sehr nützliche Formel

$$G(\vec{x},\vec{x}') = \frac{1}{4\pi|\vec{x}-\vec{x}'|} = \sum_{\ell=0}^{\infty}\sum_{m=-\ell}^{\ell}\frac{1}{2\ell+1}\frac{r_<^\ell}{r_>^{\ell+1}}\mathrm{Y}^*_{\ell m}(\vartheta',\varphi')\mathrm{Y}_{\ell m}(\vartheta,\varphi). \tag{B.3.30}$$

Daraus können wir nun das **Additionstheorem für Kugelflächenfunktionen** herleiten. Dazu denken wir uns ein neues Koordinatensystem so eingeführt, dass die 3-Achse in Richtung von $\vec{x}'$ weist, so dass in diesem Koordinatensystem $\vec{x}' = r'\vec{e}_3$ ist. Führen wir dann für $\vec{x}$ in diesem Koordinatensystem Polarkoordinaten (r,γ,δ) ein, hängt wegen

$$|\vec{x}-\vec{x}'| = \sqrt{(\vec{x}-\vec{x}')^2} = \sqrt{r^2+r'^2-2\vec{x}\cdot\vec{x}'} = \sqrt{r^2+r'^2-2r\,r'\cos\gamma} \tag{B.3.31}$$

die linke Seite der Gleichung nicht von δ ab. Demnach können in (B.3.30) (wo man in diesem Fall offenbar $\vartheta=\gamma$, $\varphi=\delta$, $\vartheta'=0$ zu setzen hat und φ' unbestimmt bleibt) nur Beiträge von $m=0$ vorkommen. Demnach ist

$$G(\vec{x},\vec{x}') = \sum_{\ell=0}^{\infty}\frac{1}{2\ell+1}\frac{r_<^\ell}{r_>^{\ell+1}}\mathrm{Y}^*_{\ell 0}(0)\mathrm{Y}_{\ell 0}(\gamma). \tag{B.3.32}$$

Dabei haben wir für $m=0$ berücksichtigt, dass die $\mathrm{Y}_{\ell 0}(\gamma,\delta)\equiv\mathrm{Y}_{\ell 0}(\gamma)$ nicht von δ abhängen. Mit (3.11.58) und $P_\ell(1)=1$ folgt

$$G(\vec{x},\vec{x}') = \frac{1}{4\pi}\sum_{\ell=0}^{\infty}\frac{r_<^\ell}{r_>^{\ell+1}}\mathrm{P}_\ell(\cos\gamma). \tag{B.3.33}$$

Der Vergleich mit (B.3.30) liefert dann das Additionstheorem für die Kugelflächenfunktionen

$$\sum_{m=-\ell}^{\ell}\frac{4\pi}{2\ell+1}\mathrm{Y}^*_{\ell m}(\vartheta',\varphi')\mathrm{Y}_{\ell m}(\vartheta,\varphi) = \mathrm{P}_\ell(\cos\gamma). \tag{B.3.34}$$

Dabei erhalten wir $\cos\gamma$ aufgrund der obigen Bedeutung des Winkels γ aus

$$\cos\gamma = \frac{\vec{x}\cdot\vec{x}'}{r\,r'} = \sin\vartheta\sin\vartheta'\cos(\varphi-\varphi')+\cos\vartheta\cos\vartheta'. \tag{B.3.35}$$

Verwenden wir schließlich (B.3.4) in (B.3.33) folgt

$$\frac{1}{|\vec{x}-\vec{x}'|} = \sum_{\ell=0}^{\infty}\frac{r_<^\ell}{r_>^{\ell+1}}\mathrm{P}_\ell(\cos\gamma). \tag{B.3.36}$$

Verwenden wir zur Abkürzung $u=\cos\gamma$ und setzen wir (B.3.31) auf der linken Seite von (B.3.36) ein, folgt mit $r'=1$ die erzeugende Funktion für die Legendre-Polynome

$$(1+r^2-2r\,u)^{-1/2} = \sum_{\ell=0}^{\infty} r^\ell\mathrm{P}_\ell(u). \tag{B.3.37}$$

Aus dem Taylorschen Lehrsatz folgt daraus die Formel

$$\mathrm{P}_\ell(u) = \frac{1}{\ell!}\partial_r^\ell(1+r^2-2r\,u)^{-1/2}. \tag{B.3.38}$$

B.3.2 Green-Funktion des Helmholtz-Operators

Für die Quantentheorie wichtiger ist die Green-Funktion für den **Helmholtz-Operator**

$$-(\Delta + k^2)G_k(\vec{x},\vec{x}') = \delta^{(3)}(\vec{x}-\vec{x}'). \tag{B.3.39}$$

Ähnlich wie im vorigen Abschnitt wollen wir mit Hilfe dieser Green-Funktion noch weitere wichtige Eigenschaften der Kugelflächenfunktionen sowie der sphärischen Bessel-Funktionen (s. den vorigen Abschnitt B.2) herleiten, insbesondere die Entwicklung ebener Wellen nach Kugelflächenfunktionen.

Zunächst können wir (B.3.39) analytisch lösen. Das Vorgehen ist dabei genau analog wie bei der Herleitung der Green-Funktion für den Laplace-Operator im vorigen Abschnitt. Zunächst ist G_k eine unter Drehungen des Relativvektors $\vec{r} = \vec{x} - \vec{x}'$ invariante Funktion, und für $\vec{r} \neq 0$ finden wir, indem wir $G_k(\vec{x},\vec{x}') = g_k(r) = \chi_k(r)/r$ in (B.3.39) einsetzen (*Nachrechnen!*)

$$\chi_k''(r) = -k^2\chi(r) \Rightarrow \chi_k^{(\pm)}(r) = C_k \exp(\pm \mathrm{i}kr). \tag{B.3.40}$$

Wir erhalten also zwei Lösungen für G_k:

$$G_k^{(\pm)}(\vec{x},\vec{x}') = \frac{C_k}{|\vec{x}-\vec{x}'|}\exp(\pm \mathrm{i}k|\vec{x}-\vec{x}'|). \tag{B.3.41}$$

Die physikalische Bedeutung wird klar, wenn wir bedenken, dass sich die Helmholtz-Gleichung aus der Schrödinger-Gleichung für ein freies Teilchen

$$\mathrm{i}\hbar\partial_t\psi(t,\vec{x}) = -\frac{\hbar^2}{2m}\Delta\psi(t,\vec{x}) \tag{B.3.42}$$

durch den Ansatz

$$\psi(t,\vec{x}) = \exp(-\mathrm{i}\omega t)u_\omega(\vec{x}) \tag{B.3.43}$$

ergibt. Setzt man diesen Ansatz nämlich in (B.3.42) ein, folgt

$$\hbar\omega u_\omega = -\frac{\hbar^2}{2m}\Delta u_\omega \Rightarrow -\Delta u_\omega = \frac{2m\omega}{\hbar}u_\omega. \tag{B.3.44}$$

Setzt man darin $2m\omega/\hbar = k^2$, erhält man

$$-(\Delta + k^2)u_\omega = 0, \tag{B.3.45}$$

also die Helmholtz-Gleichung. Dass hier nur $\omega \geq 0$ erlaubt ist, folgt daraus, dass u_ω die **Energieeigenfunktionen** sind, wobei die Energieigenwerte durch die de-Broglie-Beziehung $E = \hbar\omega$ gegeben ist und für ein freies Teilchen $E = \vec{p}^2/(2m) \geq 0$ mit den Impulseigenwerten $\vec{p} \in \mathbb{R}^3$ ist.

Somit sind

$$\psi(t,\vec{x}) = \exp(-\mathrm{i}\omega t)G_k^{(\pm)}(\vec{x},\vec{x}') \tag{B.3.46}$$

Lösungen der zeitabhängigen Schrödinger-Gleichung, wobei die Lösung mit $G_k^{(+)}$ einer von $\vec{x}'$ her auslaufenden und die Lösung mit $G_k^{(-)}$ einer zu $\vec{x}'$ hin einlaufenden **Kugelwelle** entspricht. In den meisten physikalischen Anwendungen benötigen wir auslaufende Kugelwellen. Daher beschäftigen wir uns im Folgenden mit $G_k^{(+)}$. Alle Eigenschaften lassen sich sofort auf $G_k^{(-)}$ übertragen, da ja einfach $G_k^{(-)} = G_k^{(+)*}$ ist.

Die Integrationskonstante in (B.3.41) folgt nun daraus, dass sich für $k = 0$ die Green-Funktion (B.3.4) des Laplace-Operators ergeben muss, d.h. es ist $C_k = 1/(4\pi)$ und somit

$$G_k^{(+)}(\vec{x},\vec{x}') = \frac{\exp(\mathrm{i}k|\vec{x}-\vec{x}'|)}{4\pi|\vec{x}-\vec{x}'|}. \tag{B.3.47}$$

Nun ist es klar, dass in vollständiger Analogie zur Betrachtung für die Green-Funktion des Laplace-Operators die Entwicklung dieser Funktion nach Kugelflächenfunktionen die Form

$$G_k^{(+)}(\vec{x},\vec{x}') = \sum_{\ell=0}^{\infty} g_\ell(r,r') \sum_{m=-\ell}^{\ell} \mathrm{Y}^*_{\ell m}(\hat{x}')\mathrm{Y}_{\ell m}(\hat{x}) \tag{B.3.48}$$

besitzen muss. Dabei bezeichnen wir mit $\hat{x} = \vec{x}/r$ und $\hat{x}' = \vec{x}'/r'$ die Einheitsvektoren in Richtung der Ortsvektoren $\vec{x}$ bzw. $\vec{x}'$.

Man rechnet leicht nach, dass g_ℓ sowohl bzgl. r als auch bzgl. r' die Differentialgleichung (B.2.1) für die sphärischen Bessel-, Neumann- und Hankel-Funktionen erfüllen, wobei $\rho = kr$ ist. Es liegt nun nahe, zu vermuten, dass wie bei der Green-Funktion des Laplace-Operators $g_\ell(r,r') = g_{\ell 1}(r)g_{\ell 2}(r')$ ist. Damit ist auch klar, dass $g_{\ell 1}$ und $g_{\ell 2}$ irgendwelche sphärischen Bessel-, Neumann- oder Hankel-Funktionen sein müssen.

Welche dies sein müssen, ergibt sich nun durch einige Betrachtungen zum asymptotischen Verhalten der Green-Funktion (B.3.47) für $r \to 0$ und $r \to \infty$ für festgehaltenes $\vec{r}' \neq 0$. Dann gilt

$$|\vec{x}-\vec{x}'| = r|\hat{x}-\vec{x}'/r| = r\sqrt{1-2\hat{r}\cdot\vec{x}'/r + \mathcal{O}(1/r^2)} = r - \hat{r}\cdot\vec{r}' + \mathcal{O}(r'^2/r). \tag{B.3.49}$$

Damit können wir (B.3.47) durch

$$G_k^{(+)}(\vec{x},\vec{x}') \underset{r\to\infty}{\cong} \frac{\exp(\mathrm{i}kr)}{4\pi r}\exp(-\mathrm{i}k\hat{r}\cdot\vec{r}') \tag{B.3.50}$$

nähern. Daraus folgt, dass dann $g_{\ell 1}(r) \propto \mathrm{h}_\ell^{(1)}(kr)$ sein muss, da diese sphärische Hankelfunktion gemäß (B.2.26) genau dieses asymptotische Verhalten aufweist.

Für $r \to 0$ und $\vec{r}' \neq 0$ bleibt offensichtlich $G_k^{(+)}$ endlich, und die einzige der sphärischen Bessel-Funktionen, die diese Randbedingung erfüllt, ist die sphärische Besselfunktion, d.h. $g_{\ell 1}^{(1)}(r) \propto \mathrm{j}_\ell(kr)$ für $r \to 0$.

Genauso kann man nun für $r' \to \infty$ und $r' \to 0$ und $g_\ell^{(2)}$ argumentieren. Insgesamt folgt, dass mit irgendwelchen noch unbestimmten Normierungskonstanten C_ℓ

$$g_\ell(r,r') = C_\ell \mathrm{j}_\ell(kr_<)\mathrm{h}_\ell^{(1)}(kr_>) \tag{B.3.51}$$

sein muss.

Diese Konstanten folgen wiederum daraus, dass (B.3.48) für $k \to 0$ in (B.3.30) übergehen muss, da $G_k^{(+)} \to G$ für $k \to 0$ gilt. Nun ist aber wegen (B.2.25) und (B.2.28) gemäß (B.3.30)

$$g_\ell(r,r') \underset{k\to 0}{\cong} C_\ell \frac{2^\ell \ell!}{(2\ell+1)!}(kr_<)^\ell(-\mathrm{i})\frac{(2\ell)!}{2^l \ell!}(kr_>)^{-\ell+1} = \frac{C_\ell}{2\ell+1}\frac{r_<^\ell}{r_>^{\ell+1}}\frac{1}{\mathrm{i}k} \overset{!}{=} \frac{1}{2\ell+1}\frac{r_<^\ell}{r_>^{\ell+1}} \Rightarrow C_\ell = \mathrm{i}k. \tag{B.3.52}$$

Damit erhalten wir schließlich für (B.3.48)

$$G_k^{(+)}(\vec{x},\vec{x}') = \frac{\exp(\mathrm{i}k|\vec{x}-\vec{x}'|)}{4\pi|\vec{x}-\vec{x}'|} = \mathrm{i}k\sum_{\ell=0}^{\infty}\mathrm{j}_\ell(kr_<)\mathrm{h}_\ell^{(1)}(kr_>)\sum_{m=-\ell}^{\ell}\mathrm{Y}^*_{\ell m}(\hat{x}')\mathrm{Y}_{\ell m}(\hat{x}). \tag{B.3.53}$$

Wir können daraus leicht die **Partialwellenentwicklung** der ebenen Welle

$$f_{\vec{k}}(\vec{x}) = \exp(\mathrm{i}\vec{k}\cdot\vec{x}) \tag{B.3.54}$$

erhalten. Anschaulich ist nämlich klar, dass für $r' \to \infty$ die Green-Funktion für endliche r näherungsweise wie eine ebene Welle aussehen muss. In der Tat erhalten wir durch Vertauschen von $\vec{x}$ mit $\vec{x}'$ in (B.3.49)

$$|\vec{x}-\vec{x}'| = r'\left|\frac{r}{r'}\hat{x} - \hat{x}'\right| = r' - \vec{x}\cdot\hat{x}' + \mathcal{O}\left(\frac{r^2}{r'}\right) \tag{B.3.55}$$

durch Einsetzen in (B.3.47)

$$G_k^{(+)}(\vec{x},\vec{x}') \underset{r'\to\infty}{\cong} \frac{\exp(\mathrm{i}kr')}{4\pi r'}\exp(-\mathrm{i}\vec{k}\cdot\vec{x}) + \mathcal{O}\left(\frac{1}{r'^2}\right), \quad \vec{k} = k\hat{x}'. \tag{B.3.56}$$

Verwenden wir in der Reihenentwicklung (B.3.53) die asymptotische Form für $\mathrm{h}_\ell^{(1)}$ (B.2.26), erhalten wir andererseits

$$\begin{aligned} G_k^{(+)}(\vec{x},\vec{x}') &\underset{r'\to\infty}{\cong} \mathrm{i}k\sum_{\ell=0}^{\infty}\mathrm{j}_\ell(kr)(-\mathrm{i})^{\ell+1}\frac{\exp(\mathrm{i}kr')}{kr'}\sum_{m=-\ell}^{\ell}\mathrm{Y}^*_{\ell m}(\hat{k})\mathrm{Y}_{\ell m}(\hat{x}) \\ &= \frac{\exp(\mathrm{i}kr')}{r'}\sum_{\ell=0}^{\infty}(-\mathrm{i})^\ell\mathrm{j}_\ell(kr)\sum_{m=-\ell}^{\ell}\mathrm{Y}^*_{\ell m}(\hat{k})\mathrm{Y}_{\ell m}(\hat{x}). \end{aligned} \tag{B.3.57}$$

Dabei haben wir verwendet, dass gemäß (B.3.56) $\hat{k} = \hat{x}'$ ist. Vergleichen wir nun (B.3.56) mit (B.3.57) folgt

$$\exp(-\mathrm{i}\vec{k}\cdot\vec{x}) = 4\pi\sum_{\ell=0}^{\infty}(-\mathrm{i})^\ell\mathrm{j}_\ell(kr)\sum_{m=-\ell}^{\ell}\mathrm{Y}^*_{\ell m}(\hat{k})\mathrm{Y}_{\ell m}(\hat{x}). \tag{B.3.58}$$

Bilden wir davon das konjugiert Komplexe und berücksichtigen, dass $\mathrm{j}_\ell(kr)\in\mathbb{R}$ ist, folgt schließlich

$$\exp(+\mathrm{i}\vec{k}\cdot\vec{x}) = 4\pi\sum_{\ell=0}^{\infty}\mathrm{i}^\ell\mathrm{j}_\ell(kr)\sum_{m=-\ell}^{\ell}\mathrm{Y}_{\ell m}(\hat{k})\mathrm{Y}^*_{\ell m}(\hat{x}). \tag{B.3.59}$$

Nun ist aufgrund des Additionstheorems der Kugelflächenfunktionen (B.3.34) wegen $P_\ell(\cos\gamma)\in\mathbb{R}$ die Summe über m reell, so dass wir auch das Konjugiert Komplexe für diese Summanden einsetzen können, d.h. es gilt auch

$$\exp(+\mathrm{i}\vec{k}\cdot\vec{x}) = 4\pi\sum_{\ell=0}^{\infty}\mathrm{i}^\ell\mathrm{j}_\ell(kr)\sum_{m=-\ell}^{\ell}\mathrm{Y}^*_{\ell m}(\hat{k})\mathrm{Y}_{\ell m}(\hat{x}). \tag{B.3.60}$$

Dies ist die gewünschte Partialwellenentwicklung der ebenen Welle.
Daraus können wir nun noch eine wichtige Formel für die sphärischen Bessel-Funktionen herleiten, die wir z.B. bei der Normierung der simultanen Eigenfunktionen für $\mathbf{H}$ (Energie), $\vec{\mathbf{L}}^2$ und $\mathbf{L}_3$ des freien Teilchens benötigen (s. Abschnitt 3.12). Dazu verwenden wir

$$\int_{\mathbb{R}^3}\mathrm{d}^3x\,\exp[\mathrm{i}\vec{x}\cdot(\vec{k}_1-\vec{k}_2)] = (2\pi)^3\delta^{(3)}(\vec{k}_1-\vec{k}_2). \tag{B.3.61}$$

Andererseits ist wegen (B.3.60)

$$\begin{aligned} \exp[\mathrm{i}\vec{x}\cdot(\vec{k}_1-\vec{k}_2)] &= \exp(\mathrm{i}\vec{x}\cdot\vec{k}_1)\exp(\mathrm{i}\vec{x}\cdot\vec{k}_2) \\ &= (4\pi)^2\sum_{\ell_1,m_1,\ell_2,m_2}\mathrm{i}^{\ell_1+\ell_2}\mathrm{j}_{\ell_1}(k_1r)\mathrm{j}_{\ell_2}(k_2r)\mathrm{Y}_{l_1m_1}(\hat{k}_1)\mathrm{Y}^*_{\ell_2m_2}(-\hat{k}_2)\mathrm{Y}^*_{\ell_1m_1}(\hat{x})\mathrm{Y}_{\ell_2m_2}(\hat{x}) \end{aligned} \tag{B.3.62}$$

Integrieren wir dies gemäß (B.3.61) über $\vec{x}$ und führen die Integration in Kugelkoordinaten aus, ergibt sich die Integration über die Winkel ϑ und φ aus der der Orthonormalität der Kugelflächenfunktionen

$$\int_0^\pi\mathrm{d}\vartheta\int_0^{2\pi}\mathrm{d}\varphi\,\sin\vartheta\,\mathrm{Y}^*_{\ell_1m_1}(\hat{x})\mathrm{Y}_{\ell_2m_2}(\hat{x}) = \delta_{\ell_1\ell_2}\delta_{m_1m_2}. \tag{B.3.63}$$

Setzt man dies in (B.3.62) ein, führt die Summe über ℓ_2 und m_2 aus und benennt die verbliebenen Summationsindizes ℓ_1 und m_1 in ℓ bzw. m um, folgt

$$(2\pi)^3\delta^{(3)}(\vec{k}_1-\vec{k}_2)=(4\pi)^2\sum_{\ell,m}(-1)^\ell\int_0^\infty \mathrm{d}r\, r^2 \mathrm{j}_\ell(k_1 r)\mathrm{j}_\ell(k_2 r)\mathrm{Y}_{\ell m}(\hat{k}_1)\mathrm{Y}^*_{\ell m}(-\hat{k}_2). \tag{B.3.64}$$

Wegen (3.11.53) wird dies zu

$$(2\pi)^3\delta^{(3)}(\vec{k}_1-\vec{k}_2)=(4\pi)^2\sum_{\ell,m}\int_0^\infty \mathrm{d}r\, r^2 \mathrm{j}_\ell(k_1 r)\mathrm{j}_\ell(k_2 r)\mathrm{Y}_{\ell m}(\hat{k}_1)\mathrm{Y}^*_{\ell m}(\hat{k}_2). \tag{B.3.65}$$

Die linke Seite lässt sich nun in Kugelkoordinaten umschreiben:

$$(2\pi)^3\delta^{(3)}(\vec{k}_1-\vec{k}_2)=(2\pi)^3\frac{1}{k_1^2\sin\vartheta_1}\delta(k_1-k_2)\delta(\vartheta_1-\vartheta_2)\delta(\varphi_1-\varphi_2). \tag{B.3.66}$$

Wegen der Vollständigkeit der Kugelflächenfunktionen ist

$$\frac{1}{\sin\vartheta_1}\delta(\vartheta_1-\vartheta_2)\delta(\varphi_1-\varphi_2)=\sum_{\ell,m}\mathrm{Y}_{\ell m}(\hat{k}_1)\mathrm{Y}^*_{\ell m}(\hat{k}_2). \tag{B.3.67}$$

Damit kann (B.3.65) nur korrekt sein, wenn

$$\int_0^\infty \mathrm{d}r\, r^2 \mathrm{j}_\ell(k_1 r)\mathrm{j}_\ell(k_2 r)=\frac{\pi}{2k_1^2}\delta(k_1-k_2) \tag{B.3.68}$$

gilt, und das ist das gewünschte Normierungsintegral für die sphärischen Bessel-Funktionen.

B.4 Laguerre- und assoziierte Laguerre-Polynome

Wir haben die **assoziierten Laguerre-Polynome** bei der Lösung des Energieeigenwertproblems des Wasserstoffatoms in (3.13.32) eingeführt. Mit $n,k\in\mathbb{N}_0$ lautet diese Definition (wir setzen in (3.13.32) $n_r=n$ und $2\ell+1=k$)

$$L_n^k(x)=\sum_{j=0}^n\frac{(-1)^j}{j!}\binom{n+k}{n-j}x^j. \tag{B.4.1}$$

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**:

$$\frac{\mathrm{d}^a}{\mathrm{d}x^a}[f(x)g(x)]=\sum_{j=0}^a\binom{a}{j}f^{(j)}(x)g^{(a-j)}(x) \tag{B.4.2}$$

mit den Binomialkoeffizienten

$$\binom{a}{j}=\frac{a!}{j!(a-j)!}. \tag{B.4.3}$$

Außerdem bedeutet $f^{(j)}$ die j-te Ableitung der Funktion f. Für $a=1$ erhält man die übliche Produktregel

$$\frac{\mathrm{d}}{\mathrm{d}x}[f(x)g(x)]=f'(x)g(x)+f(x)g'(x). \tag{B.4.4}$$

Für beliebige a folgt dann der Beweis von (B.4.2) durch vollständige Induktion. Sei also die Formel für a gültig. Dann ist

$$\begin{aligned}\frac{\mathrm{d}^{a+1}}{\mathrm{d}x^{a+1}}[f(x)g(x)] &= \frac{\mathrm{d}}{\mathrm{d}x}\sum_{j=0}^{a}\binom{a}{j}f^{(j)}(x)g^{(a-j)}(x)\\ &= \sum_{j=0}^{a}\binom{a}{j}\Big[f^{(j+1)}(x)g^{(a-j)}(x)+f^{(j)}(x)g^{(a-j+1)}(x)\Big].\end{aligned} \tag{B.4.5}$$

Dabei haben wir im letzten Schritt die übliche Produktformel für Ableitungen verwendet. Wir numerieren nun in der ersten Summe den Summationsindex zu $j' = j+1$ um. Dann erhält man

$$\sum_{j=0}^{a}\binom{a}{j}f^{(j+1)}(x)g^{(a-j)}(x) = \sum_{j'=1}^{a+1}\binom{a}{j'-1}f^{(j')}(x)g^{(a+1-j')}(x). \tag{B.4.6}$$

Benennen wir dann j' wieder in j um und setzten das Resultat in (B.4.5) ein, erhalten wir

$$\frac{\mathrm{d}^{a+1}}{\mathrm{d}x^{a+1}}[f(x)g(x)] = f(x)g^{(a+1)}(x)+\sum_{j=1}^{a}\left[\binom{a}{j-1}+\binom{a}{j}\right]f^{(j)}(x)g^{(a+1-j)}(x)+f^{(a+1)}(x)g(x). \tag{B.4.7}$$

Nun ist

$$\binom{a}{j-1}+\binom{a}{j} = \frac{a!}{(j-1)!(a+1-j)!}+\frac{a!}{j!(a-j)!} = a!\left[\frac{j+(a+1-j)}{j!(a+1-j)!}\right] = \frac{(a+1)!}{j!(a+1-j)!} = \binom{a+1}{j}. \tag{B.4.8}$$

Da weiter $\binom{a+1}{0} = \binom{a+1}{a+1} = 1$ ist also schließlich

$$\frac{\mathrm{d}^{a+1}}{\mathrm{d}x^{a+1}}[f(x)g(x)] = \sum_{j=0}^{a+1}\binom{a+1}{j}f^{(j)}(x)g^{(a+1-j)}(x), \tag{B.4.9}$$

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!*)

$$f^{(j)}(x) = (-1)^j\exp(-x), \quad g^{(a-j)}(x) = \frac{b!}{(b-a+j)!}x^{b-a+j} \tag{B.4.10}$$

und damit vermöge (B.4.2)

$$\frac{\mathrm{d}^a}{\mathrm{d}x^a}\Big[x^b\exp(-x)\Big] = \exp(-x)\sum_{j=0}^{a}\binom{a}{j}\frac{b!}{(b-a+j)!}(-1)^j x^{b-a+j}. \tag{B.4.11}$$

Setzt man hierin $a = n$ und $b = n+k$, erhalten wir

$$\begin{aligned}\frac{\mathrm{d}^n}{\mathrm{d}x^n}\Big[x^{n+k}\exp(-x)\Big] &= \exp(-x)\sum_{j=0}^{n}(-1)^j\binom{n}{j}\frac{(n+k)!}{(k+j)!}x^{k+j}\\ &= x^k\exp(-x)\sum_{j=0}^{n}(-1)^j\frac{n!(n+k)!}{j!(n-j)!(k+j)!}.\end{aligned} \tag{B.4.12}$$

Nun ist der in (B.4.1) vorkommende Faktor

$$\frac{1}{j!}\binom{n+k}{n-j} = \frac{(n+k)!}{j!(n-j)!(k+j)!} \tag{B.4.13}$$

Vergleicht man demnach (B.4.1) mit (B.4.12), ergibt sich die **Rodrigues-Formel** für die assoziierten Laguerre-Polynome

$$L_n^k(x) = \frac{1}{n!} x^{-k} \exp x \frac{\mathrm{d}^n}{\mathrm{d}x^n} \left[x^{n+k} \exp(-x) \right]. \tag{B.4.14}$$

Definieren wir nun ein Skalarprodukt auf dem Hilbertraum von Funktionen durch

$$\langle \psi \,|\, \phi \rangle_{\mathrm{L},k} = \int_0^\infty \mathrm{d}x x^k \exp(-x) \psi^*(x) \phi(x), \tag{B.4.15}$$

so bilden die assoziierten Laguerre-Polyonome 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 (B.4.14)

$$\left\langle L_n^k \middle| L_{n'}^k \right\rangle_{\mathrm{L},k} = \frac{1}{n!} \int_0^\infty \mathrm{d}x L_n^k(x) \frac{\mathrm{d}^{n'}}{\mathrm{d}x^{n'}} \left[x^{n'+k} \exp(-x) \right]. \tag{B.4.16}$$

Nun können wir n'-mal partiell integrieren, wobei alle „integralfreien“ Terme verschwinden, weil an der unteren Grenze stets mindestens ein Faktor x übrig bleibt. Es entsteht schließlich

$$\left\langle L_n^k \middle| L_{n'}^k \right\rangle_{\mathrm{L},k} = \frac{1}{n!} (-1)^{n'} \int_0^\infty \mathrm{d}x \frac{\mathrm{d}^{n'}}{\mathrm{d}x^{n'}} \left[L_n^k(x) \right] x^{n'+k} \exp(-x) = 0. \tag{B.4.17}$$

Letzteres folgt daraus, dass L_n^k ein Polynom n-ten Grades ist. Für $n' > n$ 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 (B.4.1) folgt, dass $L_n^k(x) = (-1)^n x^n / n! + \cdots$ und folglich $\mathrm{d}_x^n L_n^k(x) = (-1)^n$. Damit erhalten wir

$$\left\langle L_n^k \middle| L_n^k \right\rangle_{\mathrm{L},k} = \frac{1}{n!} (-1)^n \int_0^\infty \frac{\mathrm{d}^n}{\mathrm{d}x^n} \left[L_n^k(x) \right] x^{n+k} \exp(-x) = \frac{1}{n!} \int_0^\infty \mathrm{d}x x^{n+k} \exp(-x) = \frac{(n+k)!}{n!}. \tag{B.4.18}$$

Dabei ergibt sich das Integral im letzten Schritt mit Hilfe einer erzeugenden Funktion. Definieren wir

$$F(y) = \int_0^\infty \mathrm{d}x \exp(-yx) = -\frac{1}{y} \exp(-yx) \Big|_{x=0}^{\infty} = \frac{1}{y}, \tag{B.4.19}$$

folgt

$$\int_0^\infty \mathrm{d}x x^k \exp(-x) = (-1)^k F^{(k)}(1). \tag{B.4.20}$$

Nun ist

$$F'(y) = -\frac{1}{y^2}, \quad F''(y) = +\frac{2}{y^3}, \ldots, F^{(k)}(y) = (-1)^k \frac{k!}{y^{k+1}} \Rightarrow F^{(k)}(1) = (-1)^k k!. \tag{B.4.21}$$

Damit wird aber in der Tat

$$\int_0^\infty \mathrm{d}x x^k \exp(-x) = k!. \tag{B.4.22}$$

Anhang C

Gauß-Integrale

C.1 Grundlegende Formeln

Die **Gaußsche Normalverteilung**

$$f(x) = \frac{1}{\sqrt{\pi}} \exp(-x^2) \tag{C.1.1}$$

spielt als die sog. **Normalverteilung** eine wichtige Rolle in der Wahrscheinlichkeitsrechnung. Oft werden Integrale über diese Funktion benötigt. Die Stammfunktion ist (bis auf einen Faktor 2) die **Gaußsche Fehlerfunktion**

$$\operatorname{erf}(x) = \frac{2}{\sqrt{\pi}} \int_0^x \mathrm{d}x' \exp(-x^2). \tag{C.1.2}$$

Sie lässt sich nicht mit den üblichen elementaren Funktionen darstellen, ist jedoch bequem durch numerische Integration zu berechnen.

Das Integral über die ganze reelle Achse kann man aber mit einem Trick berechnen. Dazu definieren wir zunächst

$$I = \frac{1}{\sqrt{\pi}} \int_{\mathbb{R}} \mathrm{d}x \exp(-x^2). \tag{C.1.3}$$

Offensichtlich ist $I > 0$, und wir können auch I^2 berechnen. Dazu schreiben wir das Integral einfach zweimal hin, nennen aber die Integrationsvariable im zweiten Integral anders:

$$I^2 = \frac{1}{\pi} \int_{\mathbb{R}} \mathrm{d}x \exp(-x^2) \int_{\mathbb{R}} \mathrm{d}y \exp(-y^2). \tag{C.1.4}$$

Dieses Integral können wir auch als Integral über die gesamte (x, y)-Ebene betrachten und dann in Polarkoordinaten (R, φ) ausdrücken. Betrachten wir die ebenen Vektoren als (x_1, x_2)-Ebene im Raum, können wir schreiben

$$\underline{r} = \begin{pmatrix} x \\ y \\ 0 \end{pmatrix} = \begin{pmatrix} R\cos\varphi \\ R\sin\varphi \\ 0 \end{pmatrix}, \quad R \geq 0, \quad \varphi \in (0, 2\pi). \tag{C.1.5}$$

Die Koordinatenlinien $R = \text{const}$ sind offenbar Kreise in der (x, y)-Ebene und $\varphi = \text{const}$ beschreibt radiale Strahlen in der (x, y)-Ebene. Den Flächeninhalt der entsprechenden Flächenstücke können wir mit Hilfe des

Kreuzproduktes bestimmen. Der Flächennormalenvektor ist (*Nachrechnen!*):

$$\begin{aligned} \mathrm{d}^2\underline{f} &= \mathrm{d}R\,\mathrm{d}\varphi\,\frac{\partial \underline{r}}{\partial R} \times \frac{\partial \underline{r}}{\partial \varphi} \\ &= \mathrm{d}R\,\mathrm{d}\varphi \begin{pmatrix} \cos\varphi \\ \sin\varphi \\ 0 \end{pmatrix} \times \begin{pmatrix} -R\sin\varphi \\ R\cos\varphi \\ 0 \end{pmatrix} \\ &= \mathrm{d}R\,\mathrm{d}\varphi \begin{pmatrix} 0 \\ 0 \\ R \end{pmatrix}, \end{aligned} \tag{C.1.6}$$

und der Betrag

$$\mathrm{d}^2 f = |\mathrm{d}^2\underline{f}| = \mathrm{d}R\,\mathrm{d}\varphi\,R. \tag{C.1.7}$$

Dies in (C.1.4) eingesetzt, liefert

$$I^2 = \frac{1}{\pi}\int_0^\infty \mathrm{d}R \int_0^{2\pi} \mathrm{d}\varphi\,R\exp(-R^2) = 2\int_0^\infty \mathrm{d}R R\exp(-R^2) = -\exp(-R^2)\Big|_{R=0}^{R\to\infty} = 1. \tag{C.1.8}$$

Es ist also

$$\frac{1}{\sqrt{\pi}}\int_{\mathbb{R}} \mathrm{d}x\exp(-x^2) = 1. \tag{C.1.9}$$

Wichtig sind noch die **Momente der Gauß-Verteilung**, also

$$I_n = \frac{1}{\sqrt{\pi}}\int_{\mathbb{R}} \mathrm{d}x\,x^n\exp(-x^2). \tag{C.1.10}$$

Offenbar sind aus Symmetriegründen für ungerade n diese Integrale 0. Für alle anderen Integrale empfiehlt sich die Berechnung der **erzeugenden Funktion**

$$Z(y) = \frac{1}{\sqrt{\pi}}\int_{\mathbb{R}} \mathrm{d}x\exp(-x^2 + yx). \tag{C.1.11}$$

Dieses Integral läßt sich auf (C.1.9) zurückführen. Dazu schreiben wir für das Argument der Exponentialfunktion (*Nachrechnen!*)

$$-x^2 + yx = -[(x - y/2)^2 - y^2/4]. \tag{C.1.12}$$

Dann folgt

$$Z(y) = \frac{1}{\sqrt{\pi}}\exp(y^2/4)\int_{\mathbb{R}} \mathrm{d}x\exp[-(x-y)^2]. \tag{C.1.13}$$

Substituieren wir nun $x' = x - y$, $\mathrm{d}x' = \mathrm{d}x$ finden wir in der tat mit (C.1.9)

$$Z(y) = \frac{1}{\sqrt{\pi}}\exp(y^2/4)\int_{\mathbb{R}} \mathrm{d}x'\exp(-x'^2) = \exp(-y^2/4). \tag{C.1.14}$$

Nun lassen sich die Momente (C.1.11) offenbar durch die Ableitungen von Z berechnen, denn gemäß (C.1.11) ist

$$\begin{aligned} Z'(y)|_{y=0} &= \frac{1}{\sqrt{\pi}}\int_{\mathbb{R}} \mathrm{d}x\,x\exp(-x^2) = I_1, \\ Z''(y)|_{y=0} &\frac{1}{\sqrt{\pi}}\int_{\mathbb{R}} \mathrm{d}x\,x^2\exp(-x^2) = I_2, \\ &\dots \\ Z^{(n)}(y)|_{y=0} &= \frac{1}{\sqrt{\pi}}\int_{\mathbb{R}} \mathrm{d}x\,x^n\exp(-x^2) = I_n. \end{aligned} \tag{C.1.15}$$

Die ersten Ableitungen sind

$$Z'(y) = \frac{y}{2}\exp(y^2/4), \quad Z''(y) = \left(\frac{y^2}{4} + \frac{1}{2}\right)\exp(y^2/4), \quad Z^{(3)}(y) = \left(\frac{y^3}{8} + y\right)\exp(y^2/4), \ldots \tag{C.1.16}$$

Dies ist ein wenig mühsam. Wir benötigen aber nur die Werte der Ableitungen bei $y = 0$, die wir aus der bekannten Taylor-Reihe für die Exponentialfunktion

$$\exp z = \sum_{n=0}^{\infty} \frac{z^n}{n!} \Rightarrow \exp(y^2/4) = \sum_{n=0}^{\infty} \frac{1}{n!}\left(\frac{y^2}{4}\right)^n = \sum_{n=0}^{\infty} \frac{1}{n!}\left(\frac{1}{2}\right)^{2n} y^{2n} = \sum_{j=0}^{\infty} \frac{1}{j!} I_j y^j \tag{C.1.17}$$

unmittelbar ablesen können. Es ist demnach

$$I_j = \begin{cases} 0 & \text{für} \quad j = 2n+1, \\ \frac{(2n)!}{n!}\left(\frac{1}{2}\right)^{2n} & \text{für} \quad j = 2n. \end{cases} \tag{C.1.18}$$

Schließlich berechnen wir noch das etwas allgemeinere Gauß-Integral

$$G(a,b) = \int_{\mathbb{R}} \mathrm{d}x \exp(-ax^2 + bx) \tag{C.1.19}$$

Dieses Integral konvergiert offenbar für beliebige $a \in \mathbb{C}$ mit $\mathrm{Re}\, A > 0$. Wir führen es auf das Integral (C.1.9) zurück. Dazu formen wir zunächst das Argument der Exponentialfunktion ein wenig um:

$$-ax^2 + bx = -a\left(x^2 - \frac{b}{a}x\right) = -a\left[\left(x - \frac{b}{2a}\right)^2 - \frac{b^2}{4a^2}\right] \tag{C.1.20}$$

Setzen wir das in (C.1.19) ein, folgt

$$G(a,b) = \exp\left(\frac{b^2}{4a}\right)\int_{\mathbb{R}} \mathrm{d}x \exp\left[-a\left(x - \frac{b}{2a}\right)^2\right]. \tag{C.1.21}$$

Nehmen wir nun an, dass $a > 0$ und substituieren wir $y = \sqrt{a}[x - b/(2a)]$ folgt wegen $\mathrm{d}x = \mathrm{d}y/\sqrt{a}$

$$G(a,b) = \int_{\mathbb{R}} \mathrm{d}x \exp(-ax^2 + bx) = \frac{1}{\sqrt{a}}\exp\left(\frac{b^2}{4a}\right)\int_{\mathbb{R}} \mathrm{d}y \exp(-y^2) \overset{\text{(C.1.9)}}{=} \sqrt{\frac{\pi}{a}}\exp\left(\frac{b^2}{4a}\right). \tag{C.1.22}$$

Man kann mit Hilfe der Funktionentheorie zeigen, dass diese Formel auch für $a \in \mathbb{C}$ mit $\mathrm{Re}\, a > 0$ gilt.

C.2 Fourier-Transformation der Gauß-Verteilung

Wir können nun mit Hilfe des Integrale (C.1.9) auch die Fourier-Transformierte der Gauß-Verteilung berechnen. Wir wollen dies sogleich für eine etwas allgemeinere Funktion tun:

$$f(x) = \int_{\mathbb{R}} \frac{\mathrm{d}k}{2\pi}\tilde{f}(k)\exp(\mathrm{i}kx) = \int_{\mathbb{R}} \frac{\mathrm{d}k}{2\pi}\exp\left[-\frac{(k-k_0)^2}{2\sigma^2}\right]\exp(\mathrm{i}kx) = \int_{\mathbb{R}} \frac{\mathrm{d}k}{2\pi}\exp\left[-\frac{(k-k_0)^2}{2\sigma^2} + \mathrm{i}kx\right]. \tag{C.2.1}$$

Wir können wieder ähnlich wie bei der Berechnung von (C.1.13) zunächst das Argument der Exponentialfunktion umformen

$$-\left[\frac{(k-k_0)^2}{2\sigma^2} - \mathrm{i}kx\right] = -\left[\frac{k^2 + k_0^2 - 2k(k_0 + \mathrm{i}\sigma^2 x)}{2\sigma^2}\right] = -\left[\frac{(k - k_0 - \mathrm{i}\sigma^2 x)^2 + k_0^2 - (k_0 + \mathrm{i}\sigma^2 x)^2}{2\sigma^2}\right]. \tag{C.2.2}$$

Dies in (C.2.1) eingesetzt und noch etwas vereinfacht, ergibt

$$f(x) = \frac{1}{2\pi} \exp\left(-\frac{\sigma^2 x^2}{2} + \mathrm{i}k_0 x\right) \int_{\mathbb{R}} \mathrm{d}k \exp\left[-\frac{(k - k_0 - \mathrm{i}\sigma^2 x)^2}{2\sigma^2}\right]. \tag{C.2.3}$$

Nun substituieren wir

$$k - k_0 - \mathrm{i}\sigma^2 x = \sqrt{2}\sigma x \Rightarrow \mathrm{d}k = \sqrt{2}\sigma \mathrm{d}x. \tag{C.2.4}$$

Damit wird unter Zuhilfenahme von (C.1.9)

$$f(x) = \frac{\sigma}{\sqrt{2\pi}} \exp\left(-\frac{\sigma^2 x^2}{2} + \mathrm{i}k_0 x\right) \frac{1}{\sqrt{\pi}} \int_{\mathbb{R}} \mathrm{d}x \exp(-x^2) = \frac{\sigma}{\sqrt{2\pi}} \exp\left(-\frac{\sigma^2 x^2}{2} + \mathrm{i}k_0 x\right). \tag{C.2.5}$$

Definieren wir nun

$$\delta_\sigma(x) = \frac{1}{\sigma\sqrt{2\pi}} f(x) = \frac{1}{2\pi} \exp\left(-\frac{\sigma^2 x^2}{2} + \mathrm{i}k_0 x\right), \tag{C.2.6}$$

folgt

$$\lim_{\sigma \to 0^+} \delta_\sigma(x) = \frac{1}{2\pi} \exp(\mathrm{i}k_0 x). \tag{C.2.7}$$

Die Umkehrung der Fourier-Transformation[1] liefert

$$\tilde{\delta}_\sigma(k) = \int_{\mathbb{R}} \mathrm{d}x \, \delta_\sigma(x) \exp(-\mathrm{i}kx) = \frac{1}{\sigma} \tilde{f}(k) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left[-\frac{(k - k_0)^2}{2\sigma^2}\right]. \tag{C.2.8}$$

Im Sinne von verallgemeinerten Funktionen (Distributionen) können wir nun den Limes $\sigma \to 0^+$ mit der Integration vertauschen. Dann folgt aus (C.2.6)

$$\lim_{\sigma \to 0^+} \tilde{\delta}_\sigma(k) = \int_{\mathbb{R}} \mathrm{d}x \frac{1}{2\pi} \exp(\mathrm{i}k_0 x) \exp(-\mathrm{i}kx) = \int_{\mathbb{R}} \mathrm{d}x \frac{1}{2\pi} \exp[\mathrm{i}(k_0 - k)x] = \delta(k_0 - k). \tag{C.2.9}$$

Damit stellt (C.2.8) eine Regularisierung der Diracschen δ-Distribution dar, die manchmal nützlich ist.

[1] Genaueres zur Fourier-Transformation findet sich in [Hee18c].

Anhang D

Zur Ganzzahligkeit der Bahndrehimpulseigenwerte

In diesem Anhang wollen wir die Ganzzahligkeit der Bahndrehimpulseigenwerte noch auf eine andere rein algebraische Art beweisen. Dazu gehen wir von der Definition des Bahndrehimpulsoperators

$$\vec{\mathbf{l}} = -\mathrm{i}\vec{x} \times \vec{\nabla} \tag{D.0.1}$$

aus. Wie wir in Abschnitt 3.11 gezeigt haben, erfüllt dieser Operator die allgemeinen Drehimpulsvertauschungsregeln (3.10.2), und das simultane Eigenwertproblem zu $\vec{\mathbf{l}}^2$ und $\mathbf{l}_3$ wird durch das im Abschnitt 3.10 gezeigte Verfahren auch gelöst.

Es zeigt sich nun aber, dass aufgrund der speziellen Realisierung von Drehimpulsoperatoren als Bahndrehimpulsoperatoren in der Form $\vec{\mathbf{L}} = \vec{\mathbf{x}} \times \vec{\mathbf{p}}$ die Bahndrehimpulsbetragsquantenzahl ℓ **notwendig ganzzahlig** sein muss, d.h. es ist $\ell \in \{0,1,2,\ldots\} = \mathbb{N}_0$. Damit sind dann auch die Eigenwerte von $\mathbf{l}_3$ notwendig ganzzahlig, denn zu jedem ℓ ist $m \in \{-\ell, -\ell+1, \ldots, \ell-1, \ell\}$ (für halbzahlige j sind aufgrund der analogen Überlegung alle m notwendig auch halbzahlig).

Wir können nun, wieder ohne konkrete Rechnung mit Differentialoperatoren diese Ganzzahligkeit der $\mathbf{l}_3$-Eigenwerte beweisen. Dazu betrachten wir den zweidimensionalen isotropen harmonischen Oszillator und die entsprechenden beiden Vernichtungsoperatoren analog zum eindimensionalen Oszillator, also

$$\mathbf{a}_j = \sqrt{\frac{\mu\omega}{2\hbar}}\mathbf{x}_j + \frac{\mathrm{i}}{\sqrt{2\mu\omega\hbar}}\mathbf{p}_j, \quad j \in \{1,2\}. \tag{D.0.2}$$

Diese erfüllen offenbar die Kommutatorrelationen (*nachrechnen!*)

$$\left[\mathbf{a}_j, \mathbf{a}_k\right] = \left[\mathbf{a}_j^\dagger, \mathbf{a}_k^\dagger\right] = 0, \quad \left[\mathbf{a}_j, \mathbf{a}_k^\dagger\right] = \delta_{jk}\mathbb{1}. \tag{D.0.3}$$

Wie wir vom eindimensionalen Oszillator wissen, besitzen die selbstadjungierten Operatoren $\mathbf{N}_j = \mathbf{a}_j^\dagger \mathbf{a}_j$ beide die Eigenwerte $n_j \in \{0,1,2,\ldots\} = \mathbb{N}_0$, und es gibt einen vollständigen Satz gemeinsamer Eigenfunktionen u_{n_1,n_2}, da

$$[\mathbf{N}_1, \mathbf{N}_2] = 0 \tag{D.0.4}$$

gilt.

Andererseits können wir aber auch einen neuen Vernichtungsoperator

$$\mathbf{b} = \frac{1}{\sqrt{2}}(\mathbf{a}_1 + \mathrm{i}\mathbf{a}_2) \tag{D.0.5}$$

einführen. Dass dies wieder ein Vernichtungsoperator ist, ergibt sich wegen (D.0.3) aus (*Nachrechnen!*)

$$\left[\mathbf{b}, \mathbf{b}^\dagger\right] = \mathbf{1}. \tag{D.0.6}$$

D.h. der Operator $\mathbf{N}_b = \mathbf{b}^\dagger \mathbf{b}$ besitzt ebenfalls die Eigenwerte $n_b \in \mathbb{N}_0$. Weiter folgt

$$[\mathbf{N}_1 + \mathbf{N}_2, \mathbf{N}_b] = 0, \tag{D.0.7}$$

d.h. es gibt auch einen vollständigen Satz orthonormaler simultaner Eigenfunktionen $\tilde{u}_{n,n_b}$ mit Eigenwerten $n \in \mathbb{N}_0$ von $\mathbf{N} = \mathbf{N}_1 + \mathbf{N}_2$ und $n_b \in \mathbb{N}_0$ des Operators $\mathbf{N}_b$.
Rechnet man nun mit Hilfe von (D.0.2) und (D.0.5) $\mathbf{N}_b$ explizit aus, findet man nach einiger Rechnerei (*Übung!*)

$$\mathbf{N}_b = \mathbf{b}^\dagger \mathbf{b} = \frac{1}{2}\left[\mathbf{N} + \mathrm{i}(\mathbf{a}_1^\dagger \mathbf{a}_2 - \mathbf{a}_2^\dagger \mathbf{a}_1)\right] = \frac{1}{2}(\mathbf{N} - \mathbf{l}_3). \tag{D.0.8}$$

Damit folgt, dass

$$\mathbf{l}_3 = \mathbf{N} - 2\mathbf{N}_b \tag{D.0.9}$$

ist. Da $\mathbf{N} = \mathbf{N}_1 + \mathbf{N}_2$ offenbar die Eigenwerte $n \in \mathbb{N}_0$ besitzt und ebenso $\mathbf{N}_b$ die Eigenwerte $n_b \in \mathbb{N}_0$, sind die Eigenwerte m von $\mathbf{l}_3$ ganzzahlig, und man überlegt sich leicht, dass alle ganzzahligen Werte vorkommen, d.h. $m \in \mathbb{Z}$.

Anhang E

Teilchen im elektromagnetischen Feld und Eichinvarianz

In diesem Abschnitt wollen wir die Quantentheorie für ein Teilchen in einem elektromagnetischen Feld und den Zusammenhang mit der Eichinvarianz genauer analysieren.

E.1 Klassische Theorie

Wir analysieren die Bewegung eines Teilchens in einem vorgegebenen elektromagnetischen Feld im Rahmen des Lagrange- und Hamilton-Formalismus'. Das besondere Augenmerk soll dabei darauf gerichtet werden, dass sowohl die Lagrange- als auch die Hamilton-Funktion von den elektromagnetischen Potentialen Φ und $\vec{A}$ abhängen und demnach nicht direkt physikalisch observable Größen repräsentieren, da sie von der konkreten Wahl der Potentiale abhängen und diese nur bis auf eine **Eichtransformation** bestimmt sind. In der Tat sind die physikalisch observablen Felder die elektrische und magnetische Feldstärke, die wie folgt durch die Potentiale bestimmt sind:

$$\vec{E} = -\vec{\nabla}\Phi - \partial_t \vec{A}, \quad \vec{B} = \vec{\nabla} \times \vec{A}. \tag{E.1.1}$$

Diese Felder werden durch die Maxwellgleichungen durch gegebene Ladungs- und Stromverteilungen eindeutig bestimmt, aber die Potentiale sind nur bis auf eine Eichtransformation definiert, denn führt man neue Potentiale gemäß

$$\Phi' = \Phi + \partial_t \chi, \quad \vec{A}' = \vec{A} - \vec{\nabla}\chi \tag{E.1.2}$$

mit einem beliebigen Skalarfeld χ ein, so ergibt sich

$$\begin{aligned} -\vec{\nabla}\Phi' - \partial_t \vec{A}' &= -\vec{\nabla}\Phi - \partial_t \vec{A} - \partial_t \vec{\nabla}\chi + \partial_t \vec{\nabla}\chi = -\vec{\nabla}\Phi - \partial_t \vec{A} = \vec{E}, \\ \vec{\nabla} \times \vec{A}' &= \vec{\nabla} \times \vec{A} - \vec{\nabla} \times \vec{\nabla}\chi = \vec{\nabla} \times \vec{A} = \vec{B}. \end{aligned} \tag{E.1.3}$$

Das bedeutet, dass die mit einem beliebigen Skalarfeld χ gemäß (E.1.2) transformierten Potentiale Φ' und $\vec{A}'$ dieselbe physikalische Situation wie die ursprünglichen Potentiale Φ und $\vec{A}$ beschreiben.
Die Lagrange-Funktion für ein geladenes Teilchen mit der Ladung q lautet nun

$$L = \frac{m}{2}\dot{\vec{x}}^2 - q\Phi(t,\vec{x}) + q\dot{\vec{x}} \cdot \vec{A}(t,\vec{x}). \tag{E.1.4}$$

Um dies zu verifizieren, bestimmen wir die Euler-Lagrange-Gleichungen, die gemäß dem **Prinzip der kleinsten Wirkung** die Bewgungsgleichungen für das Teilchen ergeben. Berechnen wir also zuerst die **kanonischen Impulse**

$$\vec{p} = \frac{\partial L}{\partial \dot{\vec{x}}} = m\dot{\vec{x}} + q\vec{A}(t,\vec{x}). \tag{E.1.5}$$

Die Euler-Lagrange-Gleichungen lauten demnach (im Ricci-Indexkalkül geschrieben)

$$\dot{p}_k = m\ddot{x}_k + q[\partial_t A_k(t,\vec{x}) + q\dot{x}_j \cdot \partial_j A_k(t,\vec{x})] = \frac{\partial L}{\partial x_k} = -q\partial_k \Phi(t,\vec{x}) + q\dot{x}_j \partial_k A_j(t,\vec{x}). \tag{E.1.6}$$

Wir bringen dies nun in die Form einer Newtonschen Bewegungsgleichung, indem wir nach $m\ddot{x}_k$ auflösen:

$$m\ddot{x}_k = q[-\partial_t A_k(t,\vec{x}) - \partial_k \Phi(t,\vec{x})] + q\dot{x}_j[\partial_k A_j(t,\vec{x}) - \partial_j A_k(t,\vec{x})]. \tag{E.1.7}$$

Nun ist aber wegen (E.1.1)

$$\partial_k A_j - \partial_j A_k = \epsilon_{kjl}(\vec{\nabla} \times \vec{A})_l = \epsilon_{kjl} B_l \tag{E.1.8}$$

und folglich, wieder unter Verwendung von (E.1.1),

$$m\ddot{x}_k = qE_k(t,\vec{x}) + q\epsilon_{kjl}\dot{x}_j B_l(t,\vec{x}). \tag{E.1.9}$$

In der Vektorschreibweise lautet diese Gleichung

$$m\ddot{\vec{x}} = q[\vec{E}(t,\vec{x}) + \dot{\vec{x}} \times \vec{B}(t,\vec{x})]. \tag{E.1.10}$$

Auf der rechten Seite der Newtonschen Bewegungsgleichung steht also die **Lorentz-Kraft**. Somit liefert also die Lagrange-Funktion (E.1.4) tatsächlich die korrekten Bewegungsgleichungen, und diese sind auch unabhängig von der willkürlichen Wahl der Eichung der Potentiale, denn die Lorentz-Kraft hängt von den Feldern $\vec{E}$ und $\vec{B}$ ab, die eichinvariant sind.

Betrachten wir nun die Lagrange-Funktion mit den eichtransformierten Potentialen (E.1.2), erhalten wir

$$\begin{aligned} L' &= \frac{m}{2}\dot{\vec{x}}^2 - q\Phi'(t,\vec{x}) + q\dot{\vec{x}} \cdot \vec{A}'(t,\vec{x}) \\ &= \frac{m}{2}\dot{\vec{x}}^2 - q\Phi(t,\vec{x}) + q\dot{\vec{x}} \cdot \vec{A}(t,\vec{x}) - q[\partial_t \chi(t,\vec{x}) + \dot{\vec{x}} \cdot \vec{\nabla}\chi(t,\vec{x})] \\ &= L - q\frac{\mathrm{d}}{\mathrm{d}t}\chi(t,\vec{x}). \end{aligned} \tag{E.1.11}$$

Die Lagrangefunktion L' unterscheidet sich von L also nur durch eine totale Zeitableitung einer beliebigen Funktion $\Omega(t,\vec{x}) = -q\chi(t,\vec{x})$. Damit führt aber L' zu den gleichen Bewegungsgleichungen wie L (vgl. [Hee18b]), was man auch leicht durch Auswertung der Euler-Lagrange-Gleichungen direkt nachrechnet (*Übung!*).

Bestimmen wir nun die Hamilton-Funktionen zu L und L'

$$\begin{aligned} H &= \vec{p} \cdot \dot{\vec{x}} - L = \frac{m}{2}\dot{\vec{x}}^2 + q\Phi, \\ H' &= \vec{p}' \cdot \dot{\vec{x}} - L' = \frac{m}{2}\dot{\vec{x}}^2 + q\Phi' = \frac{m}{2}\dot{\vec{x}}^2 + q\Phi + q\partial_t \chi = H + q\partial_t \chi. \end{aligned} \tag{E.1.12}$$

Das bedeutet, dass sowohl die kanonischen Impulse als auch die Hamilton-Funktion *nicht* eichinvariant sind, also keine physikalischen Observablen darstellen.

Die Hamiltonschen kanonischen Gleichungen sind allerdings eichinvariant, denn die Eichtransformation ist durch die kanonische Transformation mit der Erzeugenden (vgl. wieder [Hee18b])

$$G(t,\vec{x},\vec{p}') = \vec{x} \cdot \vec{p}' + q\chi(t,\vec{x}). \tag{E.1.13}$$

In der Tat ist

$$\vec{x}' = \frac{\partial g}{\partial \vec{p}'} = \vec{x}, \quad \vec{p} = \frac{\partial G}{\partial \vec{x}} = \vec{p}' + q\vec{\nabla}\chi(t,\vec{x}), \quad H' = H + \partial_t G = H + q\partial_t \chi(t,\vec{x}) \tag{E.1.14}$$

in Übereinstimmung mit (E.1.12). Die Hamilton-Funktionen als Funktionen der jeweiligen generalisierten Koordinaten und generalisierten Impulse

$$\begin{aligned} H(\vec{x},\vec{p},t) &= \frac{m}{2}[\vec{p}-q\vec{A}(t,\vec{x})]^2 + q\Phi(t,\vec{x}), \\ H'(\vec{x}',\vec{p}',t) &= \frac{m}{2}[\vec{p}'-q\vec{A}'(t,\vec{x}')]^2 + q\Phi'(t,\vec{x}') \end{aligned} \tag{E.1.15}$$

liefern demnach auch dieselben Hamiltonschen kanonischen Gleichungen und also die eichinvarianten Bewegungsgleichungen (E.1.10).

Hinsichtlich der Energie liegt nun ein spezieller Fall vor, wenn sich das Teilchen in einem **statischen elektromagnetischen Feld** bewegt. Dann kann man offenbar zeitunabhängige elektromagnetische Potentiale $\Phi(\vec{x})$ und $\vec{A}(\vec{x})$ verwenden, und da dann die Lagrange-Funktion (E.1.4) nicht explizit von der Zeit abhängt, ist in diesem Fall die Energie

$$E = H = \frac{m}{2}\dot{\vec{x}}^2 + q\Phi(\vec{x}) \tag{E.1.16}$$

erhalten. Führt man eine Eichtransformation (E.1.2) durch, wird $H' = H + q\partial_t\chi \neq E$. Allerdings kommt E in diesem Fall eine spezielle Rolle zu, denn das Noether-Theorem für die Invarianz unter Zeittranslationen liefert E unabhängig von der Eichung als Erhaltungsgröße, und diese besitzt daher eine physikalische Bedeutung.

Für explizit zeitabhängige elektromagnetische Felder ist die Energie des Teilchens i.a. nicht erhalten, und die Hamilton-Funktion besitzt daher nicht den Charakter einer physikalisch beobachtbaren Größe, da sie eichabhängig ist.

Für ein allgemeines elektromagnetisches Feld kann man den elektrostatischen Anteil separieren, d.h. wir schreiben

$$\vec{E}(t,\vec{x}) = \vec{E}_{\text{stat}}(\vec{x}) + \vec{E}_2(t,\vec{x}) \tag{E.1.17}$$

und entsprechend

$$\vec{E}_{\text{stat}}(\vec{x}) = -\vec{\nabla}\Phi_{\text{stat}}(\vec{x}), \quad \vec{E}_2(\vec{x}) = -\vec{\nabla}\Phi_2(t,\vec{x}) - \partial_t\vec{A}(t,\vec{x}), \quad \vec{B}(t,\vec{x}) = \vec{\nabla}\times\vec{A}(t,\vec{x}). \tag{E.1.18}$$

Dann kann man die allgemeine Eichtransformation (E.1.2) für $\Phi = \Phi_{\text{stat}} + \Phi_2$ und $\vec{A}$ allein auf Φ_2 beziehen, d.h. man definiert die Eichtransformation durch

$$\Phi'_{\text{stat}}(\vec{x}) = \Phi_{\text{stat}}(\vec{x}), \quad \Phi'_2(t,\vec{x}) = \Phi_2(t,\vec{x}) + \partial_t\chi(t,\vec{x}), \quad \vec{A}'(t,\vec{x}) = \vec{A}(t,\vec{x}) - \vec{\nabla}\chi(t,\vec{x}). \tag{E.1.19}$$

Dann ist der Anteil der Hamilton-Funktion

$$H_{\text{stat}} = \frac{m}{2}\dot{\vec{x}}^2 + q\Phi_{\text{stat}}(\vec{x}) = \frac{m}{2}[\vec{p}-q\vec{A}(t,\vec{x})]^2 + q\Phi_{\text{stat}}(\vec{x}) \tag{E.1.20}$$

eichinvariant, während die Hamiltonsche Funktion

$$H = H_{\text{stat}} + q\Phi_2(t,\vec{x}) \tag{E.1.21}$$

sich gemäß (E.1.15) unter Eichtransformationen ändert.

Wie wir gleich sehen werden, wird diese für die klassische Mechnik irrelevante Abseparation des elektrostatischen Anteils der Hamiltonfunktion in der Quantenmechanik wichtig, da hier z.B. die Energieeigenwerte des Wasserstoffatoms (bzw. deren Differenzen) durch die Spektralanalyse beobachtbar sind und daher durch **eichinvariante Größen** beschrieben werden müssen.

E.2 Quantenmechanik und Eichtransformationen

In der Quantenmechanik spielt der Hamilton-Operator eine ähnliche Rolle wie die Hamilton-Funktion in der kanonischen Hamiltonschen Formulierung der klassischen Mechanik, denn wie dort bestimmt dieser vermöge der **Schrödinger-Gleichung** zunächst die Zeitentwicklung der Wellenfunktion, und es ist nicht notwendig, dass sie auch als eine observable Energie interpretierbar ist. Als erstes zeigen wir, dass die Uneindeutigkeit der elektromagnetischen Potentiale kein Problem für die Zeitentwicklung der Quantentheorie spielt, solange man nur eichunabhängige, physikalische Observablen betrachtet. Wie wir sehen werden, müssen dazu die **selbstadjungierten Operatoren**, die Observablen repräsentieren, ein bestimmtes Transformationsverhalten unter Eichtransformationen aufweisen. Außerdem wird sich zeigen, dass zwar die **Wellenfunktion eichabhängig** ist, nicht jedoch die mit ihr verknüpften physikalischen Inhalte, wie die Aufenthaltswahrscheinlichkeiten.

Wir betrachten zuerst ein geladenes Teilchen mit Spin 0 in einem beliebigen elektromagnetischen Feld, das durch die Potentiale Φ und $\vec{A}$ beschrieben wird. Gemäß der klassischen Hamilton-Funktion (E.1.15) postulieren wir, dass der Hamilton-Operator in der Quantentheorie durch

$$\mathbf{H} = \frac{1}{2m}[-\mathrm{i}\hbar\vec{\nabla} - q\vec{A}(t,\vec{x})]^2 + q\Phi(t,\vec{x}) \tag{E.2.1}$$

gegeben ist. Der Operator $\vec{\mathbf{p}} = -\mathrm{i}\hbar\vec{\nabla}$ repräsentiert demnach den **kanonischen Impuls** des Teilchens, der sich gemäß (E.1.5) vom **kinetischen Impuls** $\vec{\pi} = m\dot{\vec{v}} = \vec{p} - q\vec{A}$ unterscheidet.

Um zu sehen, dass sich der kinetische Impuls in der Quantenmechanik analog ergibt, bestimmen wir den Operator, der die Geschwindigkeit $\vec{v} = \dot{\vec{x}}$ repräsentiert. Gemäß (3.1.3) wird diese Observable in der Tat durch den Operator (*nachrechnen*)

$$\vec{\mathbf{v}} = \mathring{\vec{\mathbf{x}}} = \frac{1}{\mathrm{i}\hbar}[\vec{\mathbf{x}},\mathbf{H}] = \frac{1}{m}[-\mathrm{i}\hbar\vec{\nabla} - q\vec{A}(t,\vec{x})] \tag{E.2.2}$$

repräsentiert, so dass auch in der Quantenmechanik für die Operatoren die zum klassischen Hamilton-Formalismus analogen Beziehungen

$$\vec{\pi} = m\vec{\mathbf{v}} = \vec{\mathbf{p}} - q\vec{A}(t,\vec{\mathbf{x}}) \tag{E.2.3}$$

gelten.

Betrachten wir nun die Schrödinger-Gleichung mit dem Hamilton-Operator (E.2.1)

$$\mathrm{i}\hbar\partial_t\psi(t,\vec{x}) = \mathbf{H}\psi(t,\vec{x}) \tag{E.2.4}$$

und untersuchen, wie sich die entsprechende Wellenfunktion mit dem Hamilton-Operator für die gemäß (E.1.2) eichtransformierten elektrodynamischen Potentiale verhält. Dazu betrachten wir zuerst die Wirkung des kinetischen Impulsoperators auf eine beliebige Wellenfunktion. Mit den eichtransformierten Potentialen ist (*Nachrechnen!*)

$$\vec{\pi}'\psi' = (-\mathrm{i}\hbar\vec{\nabla} - q\vec{A} + q\vec{\nabla}\chi)\psi' = \exp\left(-\frac{\mathrm{i}q}{\hbar}\chi\right)\vec{\pi}\exp\left(\frac{\mathrm{i}q}{\hbar}\chi\right)\psi'. \tag{E.2.5}$$

Es folgt daraus sofort, dass

$$\begin{aligned}\mathbf{H}' &= \frac{1}{2m}\vec{\pi}'^2 + q\Phi' \\ &= \exp\left(-\frac{\mathrm{i}q}{\hbar}\chi\right)\left[\frac{1}{2m}\vec{\pi}^2 + q\Phi\right]\exp\left(\frac{\mathrm{i}q}{\hbar}\chi\right) + q\partial_t\chi \\ &= \exp\left(-\frac{\mathrm{i}q}{\hbar}\chi\right)(\mathbf{H} + q\partial_t\chi)\exp\left(\frac{\mathrm{i}q}{\hbar}\chi\right)\end{aligned} \tag{E.2.6}$$

gilt (*warum?*).
Ist dann $\psi(t,\vec{x})$ eine Lösung der Schrödingergleichung (E.2.5), so liegt es nahe, dass

$$\psi'(t,\vec{x}) = \exp\left(-\frac{\mathrm{i}q}{\hbar}\chi\right)\psi(t,\vec{x}) \tag{E.2.7}$$

eine Lösung der Schrödinger Gleichung mit $\mathbf{H}'$ als Hamilton-Operator ist. In der Tat folgt mit diesem Ansatz und (E.2.6)

$$\begin{aligned}
\mathrm{i}\hbar\partial_t\psi'(t,\vec{x}) &= \exp\left(-\frac{\mathrm{i}q}{\hbar}\chi\right)(\mathrm{i}\hbar\partial_t + q\partial_t\chi)\psi(t,\vec{x}) \\
&= \exp\left(-\frac{\mathrm{i}q}{\hbar}\chi\right)(\mathbf{H}\psi + q\partial_t\chi)\psi \\
&= \exp\left(-\frac{\mathrm{i}q}{\hbar}\chi\right)(\mathbf{H}\psi + q\partial_t\chi)\exp\left(\frac{\mathrm{i}q}{\hbar}\chi\right)\psi' = \mathbf{H}'\psi'.
\end{aligned} \tag{E.2.8}$$

Die Lösung der Schrödingerfunktion mit $\mathbf{H}'$ unterscheidet sich von einer Lösung mit $\mathbf{H}$ gemäß (E.2.7) also nur um einen Phasenfaktor. Für die entsprechenden Ortswahrscheinlichkeitsdichten gilt also

$$P'(t,\vec{x}) = |\psi'(t,\vec{x})|^2 = |\psi(t,\vec{x})|^2 = P(t,\vec{x}), \tag{E.2.9}$$

d.h. die Ortswahrscheinlichkeitsdichte ist **eichinvariant**, wie es für eine beobachtbare Größe sein muss.
Damit ein selbstadjungierter Operator $\mathbf{O}$ eine physikalisch beobachtbare Größe repräsentiert, muss sie sich demnach unter Eichtransformationen gemäß

$$\mathbf{O}' = \exp\left(-\frac{\mathrm{i}q}{\hbar}\chi\right)\mathbf{O}\exp\left(\frac{\mathrm{i}q}{\hbar}\chi\right) \tag{E.2.10}$$

eichkovariant transformieren, denn nur dann ergibt sich für die Erwartungswerte dieser Observablen

$$\left\langle\psi'\left|\mathbf{O}'\right|\psi'\right\rangle = \langle\exp(-\mathrm{i}q\chi/\hbar)\psi\,|\exp(-\mathrm{i}q\chi/\hbar)\mathbf{O}\exp(\mathrm{i}q\chi/\hbar)|\exp(-\mathrm{i}q\chi/\hbar)\psi\rangle = \langle\psi\,|\mathbf{O}|\,\psi\rangle\,, \tag{E.2.11}$$

also ein von der Eichung der elektromagnetischen Potentiale unabhängiges Ergebnis. Offenbar repräsentieren demnach $\vec{\mathbf{x}}$ und $\vec{\pi}$ eichinvariante Observablen und damit auch beliebige Funktionen $\mathbf{O} = O(\vec{x},\vec{\pi})$, während die kanonischen Impulse $\vec{\mathbf{p}} = -\mathrm{i}\hbar\vec{\nabla}$ und $\mathbf{H}$ sich *nicht* eichkovariant transformieren und somit keine physikalischen Observablen repräsentieren.
Damit ist klar, dass wir für elektrostatische Probleme das elektrostatische Potential analog zum klassischen Fall (E.2.1) als eichinvariant anzusehen haben. In der Tat ist dann der „Energieoperator“

$$\mathbf{E} = \frac{1}{2m}\vec{\pi}^2 + q\Phi_{\text{stat}}(\vec{x}) \tag{E.2.12}$$

eichkovariant, während der Hamiltonoperator für ein Teilchen in einem allgemeinen elektromagnetischen Feld

$$\mathbf{H} = \mathbf{E} + q\Phi_2(t,\vec{x}) \tag{E.2.13}$$

nicht eichkovariant ist und damit keine Observable repräsentiert, allerdings in der Schrödingergleichung die Zeitentwicklung der Wellenfunktion bestimmt.
Gemäß dieser Überlegung sind die Eigenwerte von $\mathbf{E}$ ebenfalls eichinvariant. Denn sind $u_E(\vec{x})$ ein beliebiges vollständiges Orthonormalsystem von Eigenfunktionen von $\mathbf{E}$, so bilden die Funktionen $u'_E(\vec{x}) = \exp(-\mathrm{i}q\chi/\hbar)u_E(\vec{x})$ offenbar ein vollständiges System von Eigenfunktionen von $\mathbf{E}'$, und für u'_E ist der Eigenwert von $\mathbf{E}'$ durch $E' = E$ gegeben, d.h. $\mathbf{E}$ und $\mathbf{E}'$ besitzen dieselben möglichen Messwerte, wie es für eine eichinvariante Observable sein muss.

E.3 Eichtheoretischer Zugang zur Quantentheorie geladener Teilchen

Die Eichsymmetrie der klassischen Elektrodynamik erweist sich für die gesamte moderne Physik der Elementarteilen als grundlegendes Paradigma. So beruht die Beschreibung der Elementarteilchen (Quarks und Leptonen) und ihrer Wechselwirkungen im Rahmen des Standardmodells der Elementarteilchen auf einer Verallgemeinerung dieser elektrodynamischen Eichtheorie.

Hier beschränken wir uns auf die entsprechende Betrachtung im Rahmen der nichtrelativistischen Quantentheorie eines einzelnen Teilchens in einem gegebenen äußeren klassischen elektromagnetischen Feldes (zunächst mit Spin 0, dann mit Spin 1/2). Hierbei erweist sich die Formulierung in der Form als „Wellenmechanik" als besonders zielführend. Die Idee ist, von der Schrödingergleichung für das freie Teilchen auszugehen:

$$\mathrm{i}\hbar\partial_t\psi(t,\vec{x}) = \hat{H}\psi(t,\vec{x}) = -\frac{\hbar^2}{2m}\Delta\psi(t,\vec{x}). \tag{E.3.1}$$

Die physikalische Bedeutung der Wellenfunktion ist nun, dass $|\psi(t,\vec{x})|^2$ die Wahrscheinlichkeitsdichte für das Finden des Teilchens am Ort $\vec{x}$ zur Zeit t ist. Damit repräsentiert die Wellenfunktion

$$\psi'(t,\vec{x}) = \exp(-\mathrm{i}q\chi/\hbar)\psi(t,\vec{x}), \quad \chi = \mathrm{const} \in \mathbb{R} \tag{E.3.2}$$

den gleichen Zustand. Es ist auch klar, dass ψ' die Schrödinger-Gleichung (E.3.1) erfüllt, wenn dies für ψ der Fall ist. Die gesamte Wellenmechanik ist also invariant unter der Transformation (E.3.1).

Die Idee der „Eichsymmetrie" ist nun, dass es möglich sein sollte die Wahl des Phasenfaktors „lokal" zu treffen, d.h. eigentlich sollte die Theorie immer noch unabhängig von χ sein, wenn man annimmt, dass es nicht einfach eine reelle Konstante sein soll sondern ein beliebiges skalares Feld sein darf: $\chi = \chi(t,\vec{x})$. Freilich ist unter solch einer **lokalen Eichtransformation** die Schrödinger-Gleichung des freien Teilchens nicht invariant. Z.B. ist für die Zeitableitung auf der linken Seite

$$\partial_t\psi'(t,\vec{x}) = \partial_t[\exp(-\mathrm{i}q\chi/\hbar)\psi(t,\vec{x}) = \exp(-\mathrm{i}q\chi/\hbar)[\partial_t\psi(t,\vec{x}) - \mathrm{i}q/\hbar\partial_t\chi\,\psi(t,\vec{x})]. \tag{E.3.3}$$

Auf der rechten Seite ergibt sich ein noch komplizierterer Ausdruck, da man ja sogar zweimal nach den Ortskoordinaten ableiten muss. Jedenfalls erfüllt ψ' nicht die Schrödinger-Gleichung, auch wenn dies für ψ der Fall ist. Für $\chi = \mathrm{const}$ ist das trivialerweise der Fall, weil sich auf beiden Seiten der Schrödingergleichung der konstante Phasenfaktor in (E.3.2) einfach wegdividieren lässt.

Die Idee der Eichinvarianz ist es nun, eine sogenannte **eichkovariante Ableitung** bzgl. der Zeit und der Ortskomponenten einzuführen:

$$\mathrm{D}_t\psi(t,\vec{x}) = [\partial_t + \mathrm{i}q/\hbar\Phi(t,\vec{x})]\psi(t,\vec{x}), \quad \mathrm{D}_j\psi(t,\vec{x}) = [\partial_j - \mathrm{i}q/\hbar A_j(t,\vec{x})]\psi(t,\vec{x}), \tag{E.3.4}$$

und zwar so, dass

$$\mathrm{D}'_t\psi'(t,\vec{x}) = \exp(-\mathrm{i}q\chi/\hbar)\mathrm{D}_t\psi(t,\vec{x}), \quad \mathrm{D}'_j\psi'(t,\vec{x}) = \exp(-\mathrm{i}q\chi/\hbar)\psi(t,\vec{x}). \tag{E.3.5}$$

Dabei ist

$$\mathrm{D}'_t = \partial_t + \mathrm{i}q/\hbar\Phi'(t,\vec{x}), \quad \mathrm{D}'_j = \partial_j - \mathrm{i}q/\hbar A'_j(t,\vec{x}). \tag{E.3.6}$$

Setzen wir (E.3.6) in (E.3.5) ein, ergibt sich die Transformationsvorschrift für Φ und $\vec{A}$ zu (*Nachrechnen!*)

$$\Phi'(t,\vec{x}) = \Phi + \partial_t\chi(t,\vec{x}), \quad \vec{A}'(t,\vec{x}) = \vec{A}(t,\vec{x}) - \vec{\nabla}\chi(t,\vec{x}). \tag{E.3.7}$$

Die Felder Φ und $\vec{A}$, die die eichkovarianten Ableitungen definieren, verhalten sich also wie das Skalar- und Vektorpotential des elektromagnetischen Feldes unter Eichtransformationen.

Eine unter Eichtransformationen invariante Schrödingergleichung erhält man aus der Schrödinger-Gleichung für das freie Teilchen (E.3.1), indem man für die partiellen Ableitungen ∂_t und ∂_j (bzw. $\vec{\nabla}$) die eichkovarianten Ableitungen einsetzt:

$$\mathrm{i}\hbar \mathrm{D}_t \psi(t,\vec{x}) = -\frac{\hbar^2}{2M}\mathrm{D}_j \mathrm{D}_j, \tag{E.3.8}$$

d.h.

$$(\mathrm{i}\hbar\partial_t - q\Phi)\psi = -\frac{\hbar^2}{2M}(\vec{\nabla} - q\vec{A})^2\psi \Rightarrow \mathrm{i}\hbar\partial_t\psi = -\frac{\hbar^2}{2M}(\vec{\nabla} - q\vec{A})^2\psi + q\Phi\psi. \tag{E.3.9}$$

Das ist die Schrödinger-Gleichung für ein Teilchen in einem elektromagnetischen Feld mit den elektromagnetischen Potentialen Φ und $\vec{A}$.

Mit einem Trick erhält man durch dieses heuristische Verfahren, aus unter globalen Eichtransformationen invarianten Gleichungen neue Gleichungen zu gewinnen, die unter lokalen Eichtransformationen invariant sind, die sog. **minimale Substitution** (also die Ersetzung der partiellen Ableitungen ∂_t und ∂_j nach der Zeit bzw. nach den x_j durch die eichkovarianten Ableitungen, auch die korrekte Pauli-Gleichung, einschließlich des Gyro-Faktors 2 für elementare Spin-1/2-Teilchen. Dazu führt man die minimale Substitution im freien Hamilton-Operator aus, den man aber zuvor unter Einführung der Pauli-Matrizen wie folgt ausdrückt

$$\hat{H} = -\frac{\hbar^2}{2M}\partial_j\partial_j = -\frac{\hbar^2}{2m}\hat{\sigma}_j\hat{\sigma}_k\partial_j\partial_k. \tag{E.3.10}$$

Dabei haben wir die Antikommutatorregeln (3.14.24) für die Pauli-Matrizen verwendet

$$\hat{\sigma}_k\hat{\sigma}_j\partial_j\partial_k = \frac{1}{2}\left\{\hat{\sigma}_j,\hat{\sigma}_k\right\}\partial_j\partial_k = \delta_{jk}\partial_j\partial_k = \partial_j\partial_j. \tag{E.3.11}$$

Führen wir die minimale Substitution $\partial_j \to \mathrm{D}_j$ in (E.3.10) aus, erhalten wir (*Nachrechnen*)

$$(-\mathrm{i}\hbar)^2\hat{\sigma}_j\hat{\sigma}_k\mathrm{D}_j\mathrm{D}_k\psi = -\hbar^2\Delta\psi + \mathrm{i}q\hbar\hat{\sigma}_j\hat{\sigma}_k(\partial_j A_k)\psi + \mathrm{i}q\hbar\left\{\hat{\sigma}_j,\hat{\sigma}_k\right\}A_k\partial_j\psi + q^2\hat{\sigma}_j\hat{\sigma}_k A_j A_k\psi. \tag{E.3.12}$$

Im zweiten Term verwenden wir nun

$$\hat{\sigma}_j\hat{\sigma}_k = \frac{1}{2}\left\{\hat{\sigma}_j,\hat{\sigma}_k\right\} + \frac{1}{2}\left[\hat{\sigma}_j,\hat{\sigma}_k\right] = \delta_{jk} + \mathrm{i}\epsilon_{jkl}\hat{\sigma}_l. \tag{E.3.13}$$

Dabei haben wir wieder die Antikommutatorregel für die Pauli-Matrizen sowie die Kommutatoren

$$\left[\hat{\sigma}_j,\hat{\sigma}_k\right] = 2\mathrm{i}\epsilon_{jkl}\hat{\sigma}_l, \tag{E.3.14}$$

die aus den Drehimpulskommutatorregeln für die $\hat{S}_j = \hbar\hat{\sigma}/2$ folgen, verwendet. Setzt man (E.3.13) in den zweiten Term von (E.3.12) ein, erhalten wir

$$\mathrm{i}q\hbar\hat{\sigma}_j\hat{\sigma}_k(\partial_j A_k)\psi = \mathrm{i}q\hbar(\vec{\nabla}\cdot\vec{A})\psi - q\hbar\hat{\vec{\sigma}}\cdot\vec{B}, \tag{E.3.15}$$

wobei $\vec{B} = \vec{\nabla}\times\vec{A}$ das Magnetfeld zum Vektorpotential $\vec{A}$ ist. Im 3. und 4. Term von (E.3.12) verwenden wir wieder die Antikommutatorregeln für die Pauli-Matrizen. Insgesamt ergibt sich dann die Pauli-Gleichung (*Nachrechnen*)

$$\mathrm{i}\hbar\mathrm{D}_t\psi = \mathrm{i}\hbar\partial_t\psi - q\Phi\psi = \frac{1}{2M}(-\mathrm{i}\hbar\vec{\nabla} - q\vec{A})^2\psi - \frac{\hbar q}{2M}\hat{\vec{\sigma}}\cdot\vec{B} = \frac{1}{2M}(-\mathrm{i}\hbar\vec{\nabla} - q\vec{A})^2\psi - \frac{q}{2M}2\hat{\vec{S}}\cdot\vec{B}\psi. \tag{E.3.16}$$

Das ist die korrekte Pauli-Gleichung für ein Elementarteilchen mit spin 1/2 im elektromagnetischen Feld. Die minimale Substitution in den freien Hamiltonoperator in der Form (E.3.10) liefert dabei sogar den korrekten Landé-Faktor $g_s = 2$.

Literaturverzeichnis

[BA58] D. Bohm, Y. Aharonov, Discussion of Experimental Proof of the Paradox of Einstein, Rosen, and Podolsky, Phys. Rev. **108**, 1070 (1958).
URL https://doi.org/10.1103/PhysRev.108.1070

[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).

[Bel64] J. S. Bell, On the Einstein-Podolsky-Rosen paradox, Physics **1**, 195 (1964).
URL https://doi.org/10.1103/PhysicsPhysiqueFizika.1.195

[Bel66] J. S. Bell, On the Problem of Hidden Variables in Quantum Mechanics, Rev. Mod. Phys. **38**, 447 (1966).
URL https://link.aps.org/doi/10.1103/RevModPhys.38.447

[Bel87] J. S. Bell, *Speakable and unspeakable in quantum mechanics*, Cambridge University Press, Cambridge, New York, Melbourne (1987).

[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

[Boh35] N. Bohr, Can quantum-mechanical description of physical reality be considered complete?, Physical Review **48**, 696 (1935).
URL https://link.aps.org/doi/10.1103/PhysRev.48.696

[Boh89] D. Bohm, *Quantum Theory*, Dover Publications, New York (1989).

[CH10] W. Cassing, H. v. Hees, *Mathematische Methoden für Physiker*, Universität Gießen (2010).
URL `https://itp.uni-frankfurt.de/~hees/publ/maphy.pdf`

[Com23] A. H. Compton, A quantum theory of the scattering of X-rays by light elements, Phys. Rev. **21**, 483 (1923).
URL `https://doi.org/10.1103/PhysRev.21.483`

[DL18] D. Dürr, D. Lazarovici, *Verständliche Quantenmechanik*, Springer Spektrum, Berlin (2018).
URL `https://doi.org/10.1007/978-3-662-55888-1`

[Ein05a] A. Einstein, Über einen die Erzeugung und Verwandlung des Lichtes betreffenden heuristischen Gesichtspunkt [AdP 17, 132 (1905)], Ann. d. Phys. **322**, 132 (1905).
URL `https://doi.org/10.1002/andp.19053220607`

[Ein05b] 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).

[Ein48] A. Einstein, Quantenmechanik und Wirklichkeit, Dialectica **2**, 320 (1948).
URL `https://doi.org/10.1111/j.1746-8361.1948.tb00704.x`

[Ein09a] A. Einstein, *Grundüge 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`

[EPR35] A. Einstein, B. Podolsky, N. Rosen, Can quantum-mechanical description of physical reality be considered complete?, Phys. Rev. **47**, 777 (1935).
URL `https://link.aps.org/doi/10.1103/PhysRev.47.777`

[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).

[FK06] H. Fischer, H. Kaul, *Mathematik für Physiker*, Bd. 3, 2. Aufl., Springer Verlag (2006).
URL `https://doi.org/10.1007/978-3-8351-9000-9`

[FK07] H. Fischer, H. Kaul, *Mathematik für Physiker*, Bd. 1, 6. Aufl., Springer Verlag (2007).
URL `https://doi.org/10.1007/978-3-8351-9207-2`

[FK08] H. Fischer, H. Kaul, *Mathematik für Physiker*, Bd. 2, 3. Aufl., Springer Verlag (2008).
URL `https://doi.org/10.1007/978-3-8351-9222-5`

[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).

[Gen05] M. Genovese, Research on hidden variable theories: A review of recent progresses, Physics Reports **413**, 319 (2005).
URL https://doi.org/10.1016/j.physrep.2005.03.003

[GK20] J. Grehn, J. Krause, *Metzler Physik SII*, 5. Aufl., Westermann Schulbuchverlag (2020).

[GP90] A. Galindo, P. Pascual, *Quantum Mechanics*, Springer Verlag, Heidelberg (1990), 2 Vols.

[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, *Mathematische Methoden für das Lehramt L3*, Goethe-Universität Frankfurt (2018).
URL https://itp.uni-frankfurt.de/~hees/publ/mameth-l3.pdf

[Hee18b] H. van Hees, *Theoretische Physik 1 für das Lehramt L3*, Goethe-Universität Frankfurt (2018).
URL https://itp.uni-frankfurt.de/~hees/publ/theo1-l3.pdf

[Hee18c] H. van Hees, *Theoretische Physik 2 für das Lehramt L3*, Goethe-Universität Frankfurt (2018).
URL https://itp.uni-frankfurt.de/~hees/publ/theo2-l3.pdf

[Hen17] K. Hentschel, *Lichtquanten*, Springer Spektrum, Berlin, Heidelberg (2017).
URL https://doi.org/10.1007/978-3-662-55273-5

[Kai11] D. Kaiser, *How the Hippies saved Physics*, W. W. Norton & Company, New York, London (2011).

[LL92] L. D. Landau, E. M. Lifschitz, *Lehrbuch der Theoretischen Physik 3, Quantenmechanik*, 9. Aufl., Harri Deutsch, Frankfurt/Main (1992).

[LL97] L. D. Landau, E. M. Lifschitz, *Lehrbuch der Theoretischen Physik, Bd. 1, Mechanik*, 14. Aufl., Verlag Harri Deutsch, Frankfurt/Main (1997).

[Mil16a] R. A. Millikan, A direct photoelectric determination of Planck's h, Phys. Rev. **7**, 355 (1916).
URL https://doi.org/10.1103/PhysRev.7.355

[Mil16b] R. A. Millikan, Einstein's photoelectric equation and contact electromotive force, Physical Review **7**, 18 (1916).
URL https://doi.org/10.1103/PhysRev.7.18

[Mil21] R. Millikan, The Distinction between Intrinsic and Spurious Contact EMFS and the Question of the Absorption of Radiation by Metals in Quanta, Phys. Rev. **18**, 236 (1921).
URL https://doi.org/10.1103/PhysRev.18.236

[Min09] H. Minkowski, Raum und Zeit, Jahresbericht der Deutschen Mathematiker-Vereinigung, vol. 18, p. 75-88 **18**, 75 (1909).

[RT76] J. Rudnick, D. Tannhauser, Concerning a widespread error in the description of the photoelectric effect, Am. J. Phys. **44**, 796 (1976).
URL `https://doi.org/10.1119/1.10130`

[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).

[TEM+89] A. Tonomura, et al., Demonstration of single-electron buildup of an interference pattern, Am. J. Phys. **57**, 117 (1989).
URL `https://doi.org/10.1119/1.16104`

[Wei15] S. Weinberg, *Lectures on Quantum Mechanics*, 2. Aufl., Cambridge University Press, Cambridge (2015).
URL `https://www.cambridge.org/9781107111660`