
Quantentheorie 1

© Carsten Timm 2018–2024

Inhaltsverzeichnis

1	Einführung	6
1.1	Ziele und Arbeitsweise der Quantentheorie	6
1.2	Überblick	8
1.3	Lehrbücher	8
2	Grenzen der klassischen Physik	11
2.1	Das goldene Zeitalter der klassischen Physik	11
2.2	Der schwarze Körper	12
2.3	Der Photoeffekt	15
2.4	Der Compton-Effekt	17
2.5	Welle oder Teilchen?	18
2.5.1	Welle-Teilchen-Dualität des Lichts	18
2.5.2	Welle-Teilchen-Dualität der Materie	19
2.6	Das Versagen des Rutherford'schen Atommodells	19
2.6.1	Stabilität der Atome	20
2.6.2	Atomspektren	20
2.6.3	Der Zeeman-Effekt	21
2.6.4	Der Franck-Hertz-Versuch	21
2.7	Der Stern-Gerlach-Versuch	22
3	Die ältere Quantentheorie	25
3.1	Die Bohrschen Postulate	25
3.2	Die Quantisierungsregeln von Bohr und Sommerfeld	25
3.2.1	Das Wasserstoffatom	26
3.2.2	Phasenraumquantisierung	28
3.2.3	Das Korrespondenzprinzip	31
3.3	Schwierigkeiten der älteren Quantentheorie	31
4	Ursprünge der Wellenmechanik	33
4.1	Materiewellen	33
4.2	Induktive Begründung der Schrödinger-Gleichung: Wellenpakete	33
4.3	Induktive Begründung der Schrödinger-Gleichung: Hamilton-Jacobi-Theorie	38
4.3.1	Die Eikonal-Gleichung	38
4.3.2	Erinnerung an die Hamilton-Jacobi-Theorie	39
4.3.3	Die Wellengleichung der Quantenmechanik	41
4.4	Weitere Bemerkungen zu Schrödinger-Gleichung	42
5	Die Schrödingersche Wellenmechanik	44
5.1	Operatoren	44
5.2	Die Schrödinger-Gleichung als partielle Differentialgleichung	47
5.3	Wahrscheinlichkeitswellen	47

5.3.1	Der Hilbert-Raum der quadratintegrablen Funktionen	49
5.3.2	Entwicklung nach Basisfunktionen	52
5.3.3	Uneigentliche (Dirac-) Zustände	53
5.3.4	Lokale Erhaltung der Wahrscheinlichkeit	55
5.3.5	Die Impulsdarstellung	55
5.4	Erwartungswerte und Schwankungen	56
5.4.1	Hermitizität	57
5.4.2	Schwankungen	58
5.5	Orts-Impuls-Unschärferelation	59
5.6	Messungen I	61
5.7	Die zeitunabhängige Schrödinger-Gleichung	64
6	Quantensysteme in einer Dimension	68
6.1	Allgemeine Eigenschaften	68
6.1.1	Klassisch verbotene und erlaubte Bereiche	69
6.1.2	Das Spektrum	70
6.2	Rechteckpotentiale	75
6.2.1	Kasten endlicher Tiefe	75
6.2.2	Kasten mit unendlich hohen Wänden	78
6.2.3	Rechteckige Potentialbarriere: Tunneleffekt	80
6.3	Der harmonische Oszillator I	84
6.4	Parität und Knotenzahl	88
6.4.1	Knotenzahl	90
6.4.2	Parität	91
7	Dirac-Formalismus	93
7.1	Zustände	93
7.1.1	Quantenmechanische Zustände	93
7.1.2	Der Hilbert-Raum	93
7.1.3	Produkträume	94
7.1.4	Der Dirac-Raum	95
7.2	Lineare Operatoren	95
7.2.1	Hermitesche und selbstadjungierte Operatoren	96
7.2.2	Unitäre Operatoren	100
7.3	Zeitentwicklung	101
7.3.1	Zeitentwicklung für zeitabhängige Hamiltonians	102
7.3.2	Heisenberg-Bild	105
7.3.3	Wechselwirkungs- oder Dirac-Bild	106
7.4	Messungen II	107
7.5	Postulate der Quantenmechanik	108
7.6	Verträgliche und nicht verträgliche Observable	109
7.6.1	Schwankungen	110
7.6.2	Allgemeine Unschärferelation	111
7.7	Ehrenfestsches Theorem	112
7.7.1	Energie-Zeit-Unschärferelation	113
7.8	Orts- und Impulsdarstellung	114
7.9	Der harmonische Oszillator II	114

8	Quantentheorie des Drehimpulses	118
8.1	Korrespondenzprinzip für den Bahndrehimpuls	118
8.2	Die Drehimpulsalgebra	121
8.2.1	Matrixdarstellung	124
8.2.2	Spin 1/2	124
8.2.3	Die Bloch-Kugel	125
8.3	Drehoperatoren	126
8.4	Addition von Drehimpulsen	128
8.4.1	Abzählung der Zustände	129
8.4.2	Clebsch-Gordan-Koeffizienten	131
9	Symmetrien	133
9.1	Unitäre Transformationen	133
9.1.1	Transformationsgruppen	134
9.1.2	Symmetrien	135
9.2	Zeitumkehr	136
9.2.1	Zeitumkehr, Drehimpulse und das Kramers-Theorem	139
10	Zentralpotentiale	141
10.1	Allgemeines Zentralpotential	141
10.2	Anwendung auf das Wasserstoff-Atom	145
11	Näherungsmethoden	150
11.1	Variationsverfahren	151
11.1.1	Das Variationsprinzip für Eigenvektoren	151
11.1.2	Ritzsches Variationsverfahren	152
11.2	Zeitunabhängige Störungstheorie	155
11.2.1	Störungstheorie für nichtentartete Energieniveaus	156
11.2.2	Störungstheorie für entartete Energieniveaus	159
11.3	Die quasiklassische Näherung	163
11.3.1	Entwicklung der Wirkung nach \hbar	163
11.3.2	Die WKB-Näherung	165
12	Der Dichteoperator	170
12.1	Projektionsoperatoren und Dichteoperatoren für reine Zustände	170
12.2	Gemischte Zustände	171
12.2.1	Messung und Projektionspostulat	175
12.2.2	Teilspur und Zustand in Faktorräumen	176
13	Konsequenzen und Deutungen der Quantenmechanik	178
13.1	Verschränkung und verborgene Variable	179
13.1.1	Das Einstein-Podolsky-Rosen-Paradoxon	179
13.1.2	Bellsches Theorem und Bellsche Ungleichung	181
13.2	Schrödingers Katze	184
13.2.1	Wigners Freund: Bewusstsein erzeugt Kollaps	185
13.2.2	Stern-Gerlach-Experimente als Prototyp des Messprozesses	185
13.3	Das No-Cloning-Theorem	187
13.4	Ensemble-Interpretation	188
13.5	Kopenhagener Deutung(en)	189
13.5.1	Bohrs Interpretation	189
13.5.2	Heisenbergs Interpretation	190
13.5.3	Zustandskollaps	190
13.6	Dekohärenz	192

13.7	Kein Kollaps: Viele-Welten-Interpretation	195
13.8	Weitere Interpretationen	199
13.8.1	De Broglie-Bohm-Führungswellen	199
13.8.2	GRW-Interpretation: Realistischer Kollaps	200

Kapitel 1

Einführung

Es ist sicherlich nicht übertrieben, die Vorlesung Quantentheorie 1 als zentralen Baustein des Physikstudiums zu betrachten. Hier wird zum ersten Mal systematisch die moderne quantentheoretische Beschreibung von Prozessen besprochen, und zwar zunächst angewandt auf mechanische Systeme. Diese Beschreibung unterscheidet sich fundamental von der aus der klassischen theoretischen Physik gewohnten. Das bezieht sich nicht in erster Linie auf die mathematische Methodik, sondern auf die Verknüpfung der formalen Objekte mit der realen Welt. Voraussagen in der Quantentheorie sind i. A. probabilistisch, d. h. Wahrscheinlichkeitsaussagen, im Gegensatz zu den deterministischen Aussagen der klassischen Physik. Das gilt offenbar auch dann, wenn wir maximale Kenntnis über den Zustand eines Systems haben. Ein bekanntes Beispiel ist der Zerfall instabiler Teilchen: Die Zerfallszeiten für mehrere Teilchen derselben Sorte sind nicht alle gleich, sondern folgen einer Zufallsverteilung. Dieser fundamentale Unterschied zur klassischen Physik provoziert Fragen nach der Deutung der Quantentheorie.

Ich möchte hier Jan Wüstemann, Bogusz Bujnowski und Georg Börner für die Erstellung verschiedener Versionen dieses Skriptes mit LaTeX sowie für die Erzeugung vieler Abbildungen danken. Außerdem danke ich allen, die mich auf Fehler und Unklarheiten aufmerksam gemacht haben, ganz besonders Herrn Dr. Jürgen Wagner.

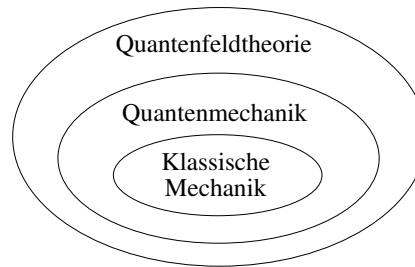
1.1 Ziele und Arbeitsweise der Quantentheorie

Die Quantentheorie ist sicherlich einer der weniger anschaulichen Zweige der Physik. Wir haben zunächst keine Intuition für Konzepte wie Wellenfunktionen, Operatoren und den Hilbert-Raum. In der klassischen Mechanik ist das anders, z. B. beim schrägen Wurf: Wir können einen geworfenen Ball fangen, ohne jemals etwas über Physik gelernt zu haben. Das Verständnis der Flugbahn ist intuitiv. Es ist aber nicht verwunderlich, dass wir für Prozesse, bei denen die Quantentheorie eine entscheidende Rolle spielt, keine solche Intuition besitzen. Im Laufe der Evolution haben wir geistige Fähigkeiten entwickelt, die für das Überleben in der gegebenen Umwelt nützlich waren. Quantenprozesse sind aber auf den unmittelbar beobachtbaren Längen- und Zeitskalen weitgehend unsichtbar. Es gab daher keinen Selektionsdruck, solche Prozesse in demselben Sinn voraussehen zu können, wie wir es für die Flugbahn eines makroskopischen Objekts können. Zur Beruhigung: Die Erfahrung zeigt, dass der Mensch in der Lage ist, auch für zunächst unanschauliche Konzepte wie die der Quantenmechanik eine Intuition zu entwickeln.

Das heißt natürlich nicht, dass Quantenprozesse für unser Leben unwichtig sind. In einem rein klassischen Universum ginge gar nichts, zum Beispiel gäbe es keine Energieproduktion in der Sonne. Selbst wenn die Sonne dennoch schiene, würde auf der Erde keine Photosynthese stattfinden, so dass die Strahlungsenergie nicht von Lebewesen genutzt werden könnte. Tatsächlich gäbe es gar keine Erde, geschweige denn Lebewesen, weil keine stabilen Atome existieren würden. Auch für technische Anwendungen ist die Quantenmechanik von überragender Bedeutung. Ohne Quantenmechanik kann man die Funktion von elektronischen Bauelementen wie Transistoren nicht verstehen, um nur ein Beispiel zu nennen. Außerdem ist die gesamte Chemie angewandte Quantenmechanik.

Die Quantenmechanik ist die Theorie der Dynamik von Teilchen. Sie betrifft damit im Prinzip dieselben Systeme wie die klassische Mechanik, ist aber die umfassendere Theorie – die Quantenmechanik enthält die klassische Mechanik als Grenzfall. Das ist der Inhalt des wichtigen *Korrespondenzprinzips*. Die Quantenmechanik

ist ihrerseits in der Quantenfeldtheorie enthalten.



In dieser Vorlesung werden wir uns ausschließlich mit der nichtrelativistischen Quantenmechanik beschäftigen. Diese beschreibt Teilchen mit (Relativ-) Geschwindigkeiten v , die klein im Vergleich zur Lichtgeschwindigkeit c sind. Die nichtrelativistische Quantenmechanik ist die Näherung für $v \ll c$ einer relativistischen Quantenmechanik, die in der relativistischen Quantenfeldtheorie enthalten ist. Auch diese Quantenfeldtheorie und ihre Anwendung auf elementare Teilchen und Felder, also das Standardmodell, liefert noch keine vollständige Beschreibung der Welt. Dazu muss zusätzlich die Gravitation beschrieben werden, was die Integration der Quantenfeldtheorie mit der (bisher klassischen) Allgemeinen Relativitätstheorie erfordert. Es ist noch nicht klar, wie dies gelingen kann. Die String-Theorie ist ein z. Zt. diskutierter Rahmen für eine solche Vereinheitlichung.

Ein Ziel der Quantenmechanik ist die quantitative Beschreibung von physikalischen Vorgängen. Die Arbeitsweise der Theoretischen Physik besteht in der Formulierung von *Theorien*, d. h. Beschreibungen von allgemeinen Gesetzmäßigkeiten. Aus einer brauchbaren Theorie lassen sich Voraussagen für Experimente herleiten, die es gestatten, die Theorie zu überprüfen. Wie der Naturphilosoph Sir Karl Popper sagte, kann man eine Theorie niemals beweisen, aber im Prinzip leicht widerlegen (falsifizieren). Experimente, die mit den Vorhersagen einer Theorie übereinstimmen, *stützen* diese, beweisen sie aber nicht. Für die Widerlegung reicht dagegen *eine* Beobachtung aus, die der Theorie widerspricht. Das Experiment ist immer die letzte Instanz in der Physik. Wir können uns also auf den Standpunkt stellen, dass die Quantenmechanik so ist wie sie ist, weil umfangreiche Experimente diese Theorie stützen und nicht irgendeine alternative Theorie. Manche Autoren belassen es dabei und halten weitergehende Frage nach der Deutung der Quantenmechanik für wenig produktiv. Diese Haltung wurde mit dem Aphorismus „shut up and calculate“ charakterisiert, der offenbar auf N. D. Mermin zurückgeht, der diese Haltung im Übrigen nicht teilt.

Einige weitere Bemerkungen in diesem Zusammenhang: 1. Man muss sich klarmachen, was Popper mit „beweisen“ meinte: Man kann eine Theorie über die reale Welt nicht *in mathematischer Strenge* beweisen, aber viele physikalische Theorien sind im „juristischen“ Sinne bewiesen, sie sind nämlich „nach menschlichem Ermessen“ wahr. Die englische Formulierung „without reasonable doubt“ ist noch treffender. 2. Die strikte Widerlegung einer Theorie durch ein Experiment im Sinne Poppers ist auch eine idealisierte Vorstellung, da man nie absolut sicher ist, dass ein Experiment wirklich zeigt, was man denkt, dass es zeigt. 3. Viele Theorien sind im Sinne Poppers falsifiziert. Wie schon erwähnt, versagt die nichtrelativistische Quantenmechanik bei hohen Teilchengeschwindigkeiten nahe c . Das bedeutet nicht, dass diese Theorie nutzlos oder nur von historischem Interesse wäre. Es ist gut verstanden, unter welchen Bedingungen sie präzise Voraussagen macht. In diesen Fällen wäre es unsinnig, die viel kompliziertere Quantenfeldtheorie zu verwenden.

Die Theoretische Physik formuliert die zu Grunde liegenden Gesetzmäßigkeiten in der Sprache der Mathematik, weil diese für die Beschreibung quantitativer Zusammenhänge am besten geeignet ist. Wir werden daher zahlreiche mathematische Methoden verwenden. Die Formulierung verwendet meist Begriffe der Analysis und der Linearen Algebra, nicht selten aber auch solche der Gruppentheorie und Geometrie. Die in dieser Vorlesung notwendigen mathematischen Hilfsmittel werden in der Vorlesung entwickelt oder wiederholt, soweit dies notwendig erscheint.

Die Quantentheorie, die in den letzten etwa hundert Jahren entwickelt wurde und sich gut bewährt hat, unterscheidet sich wie gesagt in der Natur ihrer Vorhersagen fundamental von der klassischen theoretischen Physik. Die Quantentheorie macht oft Wahrscheinlichkeitsaussagen der Art „Ein Atomkern eines bestimmten Isotops zerfällt mit einer Wahrscheinlichkeit von 83% innerhalb einer Sekunde“, selbst wenn der Ausgangszustand des Systems vollständig bekannt ist. Nach der klassischen Theorie ist der Ausgang in einem solchen Fall dagegen determiniert. Dieser fundamentale Unterschied ruft geradezu nach einer philosophischen Diskussion der Deutung

der Quantentheorie, die daher von Anfang an parallel zur methodologischen Entwicklung und Anwendung der Quantentheorie geführt wurde.

1.2 Überblick

Die Reihenfolge der Themen in dieser Vorlesung orientiert sich teilweise an der historischen Entwicklung. Dies gestattet uns, den Prozess der Konstruktion der Quantentheorie zu verfolgen. Uns interessiert nicht nur, was die moderne Quantentheorie aussagt, sondern auch, wie man darauf gekommen ist. Es ist ja nicht nur unsere Aufgabe in der Wissenschaft, existierende Theorien zu bewahren, anzuwenden und weiterzugeben, sondern auch, sie zu verbessern und letztlich zu überwinden. Die folgenden Themen sollen in der Vorlesung Quantentheorie 1 behandelt werden:

- Grenzen der klassischen Physik
- Die ältere Quantentheorie
- Die Schrödingersche Wellenmechanik
- Anwendungen auf eindimensionale Systeme
- Dirac-Formalismus
- Quantentheorie des Drehimpulses
- Zentralpotentiale
- Näherungsmethoden
- Der Dichteoperator
- Konsequenzen und Deutungen der Quantenmechanik

Als Ausblick seien hier die typischen Themen der Vorlesung Quantentheorie 2 genannt:

- Ununterscheidbare Teilchen und zweite Quantisierung
- Relativistische Quantentheorie
- Streutheorie

1.3 Lehrbücher

Es existiert eine große Auswahl von Lehrbüchern zur Quantenmechanik. Wegen der zentralen Bedeutung des Themas ist es eine gute Idee, zumindest ein Buch zu besitzen. Hier ist eine unvollständige Auswahl:

- W. Nolting, *Grundkurs Theoretische Physik, Band 5/1: Quantenmechanik – Grundlagen*, 7. Aufl. (Springer-Verlag, 2009) und *Band 5/2: Quantenmechanik – Methoden und Anwendungen*, 6. Aufl. (Springer-Verlag, 2006): Die gesamte Reihe von Lehrbüchern ist empfehlenswert. Beide Bände werden für diese Vorlesung benötigt. Nolting legt relativ großes Gewicht auf das Einüben der Formalismen und entsprechend weniger auf die ausführliche Diskussion des physikalischen Gehalts. Er führt Herleitungen oft im Detail vor, wo andere Autoren nur das Ergebnis angeben. Die Darstellung ist fast immer klar. Die Bücher enthalten viele gute Übungsaufgaben mit Lösungen und sehr hilfreiche Kontrollfragen.
- A. Messiah, *Quantenmechanik, Band 1*, 2. Aufl. (de Gruyter, 1991) und *Band 2*, 3. Aufl. (de Gruyter, 1990): Ein empfehlenswertes klassisches Lehrbuch in zwei Bänden. Umfasst den wesentlichen Stoff der Vorlesungen Quantentheorie 1 und 2 und mehr, das Material wird aber in anderer Reihenfolge diskutiert und tw. etwas verwirrend über verschiedene Kapitel verteilt. Mathematisch recht ausführlich. Deutlich mehr Diskussion als in Noltings Büchern, aber manchmal etwas wortreich und sprachlich nicht immer ganz exakt, evtl. aufgrund der Übersetzung ins Deutsche. Die Darstellung der Interpretation ist nicht ganz modern. Die Bände enthalten Übungsaufgaben ohne Lösungen.

- D. J. Griffiths und D. F. Schroeter, *Introduction to Quantum Mechanics*, 3. Aufl. (Cambridge University Press, 2018) [2. Aufl.: D. J. Griffiths (Pearson Prentice Hall, 2005)]: Ursprünglich amerikanisches Standardlehrbuch für *undergraduate quantum mechanics*. Eine deutsche Übersetzung der zweiten Auflage existiert ebenfalls bei Pearson Prentice Hall. Didaktisch und in Form und Sprache ansprechend (zumindest in der englischen Ausgabe). Eher knapp in der Stoffauswahl, niedrigeres Niveau als Noltings und Messiahs Bücher, deckt diese Vorlesung nicht vollständig ab und ist nicht sehr systematisch. Erwähnt den Dirac-Formalismus nur am Rande. Sorgfältige und klare Diskussion der Punkte, die überhaupt diskutiert werden. Viele Übungsaufgaben ohne Lösungen.
- R. L. Liboff, *Introductory Quantum Mechanics*, 4. Aufl. (Addison Wesley, 2003): Ein anderes amerikanisches *undergraduate* Lehrbuch. Hat Ähnlichkeiten im Stil mit Griffiths Buch, ist aber umfangreicher, enthält z. B. eine Darstellung der grundlegenden Experimente, eine kurze Einführung zum Pfadintegral und Kapitel zu Anwendungen in der Atom-, Molekül-, Kern- und Festkörperphysik. Deckt auch viel Material aus Quantentheorie 2 ab. Diskussionen sind eher knapper als bei Griffiths und vergleichbar mit Nolting. Enthält mehr zu modernen Themen wie Bellscher Ungleichung und Quanteninformation (ein ganzes Kapitel zu Quantum Computing!) als andere Bücher, aber teilweise veraltete Bemerkungen zur Deutung. Viele Übungsaufgaben, z. T. mit Lösungen.
- S. Weinberg, *Lectures on Quantum Mechanics*, 2. Aufl. (Cambridge University Press, 2015): Lehrbuch für *graduate students* (Master-Niveau) vom Nobelpreisträger Steven Weinberg. Führt extrem zügig in die Quantenmechanik ein, die Untersuchung des Zentralpotentials beginnt auf S. 32. Viele Themen werden nicht oder im Vorübergehen behandelt, dafür einige Themen aus der Quantentheorie 2. Keine Abbildungen! Die Stärke des Buches ist eine recht ausführliche Einführung in die Interpretationsdebatte. Einige wenige Übungsaufgaben ohne Lösungen.
- L. D. Landau und E. M. Lifschitz, *Lehrbuch der theoretischen Physik, Band 3: Quantenmechanik*, 9. Aufl. (Verlag Harri Deutsch, 1992): Teil der klassischen Reihe von russischen Lehrbüchern. Inzwischen altmodisch in der Stoffauswahl und Darstellung. Eher knapp und anspruchsvoll. Zwischenschritte werden selten angegeben. Kann zum Nachschlagen nützlich sein, wenn andere Bücher nicht weiter helfen. Die Bücher enthalten recht schwierige Übungsaufgaben ohne Lösungen.
- C. Cohen-Tannoudji, B. Diu und F. Laloë, *Quantenmechanik*, Band 1 und 2, 4. Aufl. (de Gruyter, 2010): Ein sehr gutes klassisches Lehrbuch, umfangreich in der Stoffauswahl, die über die Quantentheorie 1 hinaus geht, aber nicht die gesamte Quantentheorie 2 umfasst, und recht ausführlich hinsichtlich der Herleitungen. Für diese Vorlesung werden beiden Bände benötigt. Mit Übungsaufgaben.
- R. Shankar, *Principles of Quantum Mechanics*, 2. Aufl., 3. Nachdruck (Springer-Verlag, 2008): Amerikanisches Lehrbuch für fortgeschrittene *undergraduate* Kurse und für *graduate* Studenten. Im Umfang geht Shankars Buch deutlich über das von Griffiths hinaus und ähnelt eher Messiahs zwei Bänden, enthält aber zusätzlich eine Einführung zum Pfadintegral. Der Sprachstil ist informell, mehr noch als bei Griffiths, was man mögen muss. Enthält Übungsaufgaben.

Folgende Bücher enthalten moderne Darstellungen zur Interpretation der Quantenmechanik:

- G. Auletta, *Foundations and Interpretations of Quantum Mechanics* (World Scientific, 2001): Beginnt mit einer recht ausführlichen, lehrbuchmäßigen Einführung in die Quantenmechanik. Der Stil des gesamten Buches ähnelt eher mathematischen Lehrbüchern, mit Definitionen und Sätzen. Sehr umfangreiche Darstellung der informationstheoretischen Grundlagen, der Deutungsdebatte der Quantenmechanik und neuerer Ergebnisse (Bellsche Ungleichungen, Quanteninformation) mit zahlreichen Referenzen. Gut gegliedert, aber teilweise in überraschender Reihenfolge. Sprache und Gestaltung etwas unprofessionell (kein Lektorat?).
- C. Friebe *et al.*, *Philosophie der Quantenphysik* (Springer 2015): Erfolgreiches Ergebnis des Versuchs, ein Lehrbuch der Philosophie der Quantenphysik zu schaffen. Mehrere bekannte deutschsprachige Physikphilosophen haben Kapitel beigetragen. Gut verständlich. Die Kapitel sind relativ unabhängig voneinander. Die Konzepte der Quantentheorie, die philosophisch durchleuchtet werden sollen, werden zunächst sorgfältig eingeführt. Aber auch in diesen einführenden Teilen ist das Buch erstaunlich textlastig.

- M. Esfeld (Hrsg.), *Philosophie der Physik*, 3. Aufl. (Suhrkamp, 2012): ein weiteres Sammelwerk, mit noch mehr Kapiteln von mehr Autoren als das Buch von Friebe *et al.* Breiter in der Thematik, daher weniger ausführlich in Bezug auf die Quantentheorie.
- J. Bub, *Interpreting the Quantum Word* (Cambridge University Press, 1997): Kein Lehrbuch, sondern eine Monographie, die *einen* Vorschlag zur Interpretation vorstellt. Dieser scheint in der Tradition der de Broglie-Führungswellen zu stehen und moderne informationstheoretische Ideen zu verarbeiten. Interessant für die Vorlesung ist v. a. der Überblick über andere Interpretation aus Sicht eines nicht deutschen Physikphilosophen (in der Philosophie macht die nationale und sprachliche Tradition einen Unterschied). Deutlich mehr mathematisches Material und auch Abbildungen als in den deutschen Büchern.
- S. Carroll, *Something Deeply Hidden: Quantum Worlds and the Emergence of Spacetime* (Oneworld Publications, 2021) [deutsch *Was ist die Welt und wenn ja, wie viele: Wie die Quantenmechanik unser Weltbild verändert* (Klett-Cotta, 2021)]: Carroll ist ein Anhänger der Viele-Welten-Interpretation der Quantentheorie, die er in diesem Buch für eine breitere Öffentlichkeit vertritt. Anregend und lesenswert, aber entsprechend der Zielgruppe technisch nicht sehr tiefgehend.

Kapitel 2

Grenzen der klassischen Physik

In diesem Kapitel werden wir diskutieren, welche Experimente eine Erweiterung der Physik in der Richtung der Quantentheorie notwendig gemacht haben. Dazu werden wir zunächst den Stand der Physik gegen Ende des 19. Jahrhunderts umreißen.

2.1 Das goldene Zeitalter der klassischen Physik

Gegen Ende des 19. Jahrhunderts schien die Physik im Wesentlichen verstanden und vollständig. Es gab zwar noch offene Fragen, aber diese meinte man der *mathematischen* Schwierigkeit der Lösung der Gleichungen zuschreiben zu können. So riet der Münchner Physik-Professor Philipp von Jolly bekanntlich Max Planck 1874 davon ab, Physik zu studieren, weil nur noch Detailfragen zu klären wären.

Die damalige Physik unterschied zwei Kategorien von Objekten, nämlich Materie (Teilchen) und Strahlung (Felder):

Materie

- besteht aus Teilchen, die zu jeder Zeit durch ihren Ort \vec{r} und ihren Impuls \vec{p} charakterisiert sind
- gehorcht den Gesetzen der klassischen Mechanik (Newton, Lagrange, Hamilton)

Strahlung

- besteht nicht aus lokalisierten Teilchen, sondern zeigt wellenartiges Verhalten
- gehorcht den Maxwell'schen Gleichungen

Die Atomstruktur der Materie konnte zwar noch nicht direkt nachgewiesen werden, war aber indirekt sehr gut bestätigt. Da typische Körper offenbar aus sehr vielen Teilchen bestanden (1 mol aus $N_L = 6,022 \cdot 10^{23}$ Teilchen), konnte man die Bewegungsgleichungen der klassischen Mechanik nicht direkt lösen. Außerdem war die individuelle Bewegung der Teilchen gar nicht interessant, da im Allgemeinen nicht beobachtbar, sondern es galt, die gemittelten, makroskopischen Eigenschaften der Materie zu verstehen. Aus diesen Gründen entwickelten sich die *Thermodynamik* und die *Statistische Physik*, die aber konzeptionell die klassische Physik noch nicht in Frage stellten.

Es war natürlich gut bekannt, dass Teilchensysteme wellenartige Phänomene zeigen konnten, z. B. beim Schall oder bei Wasserwellen. Daher lag es nahe, auch für die elektromagnetischen Wellen einen materiellen Träger zu vermuten, der „Äther“ genannt wurde. Wenn diese Idee erfolgreich gewesen wäre, hätte sie die Strahlung in der obigen Tabelle auf die Materie zurückgeführt und somit eine „große Vereinheitlichung“ der klassischen Physik bewirkt. Die Idee wurde aber durch die Experimente von Michelson und Morley (1887) widerlegt, wonach die Lichtgeschwindigkeit unabhängig von der Richtung relativ zur Bewegung der Erde und damit des Labors ist. Das kann man nicht verstehen, wenn sich die Erde gegenüber einem Äther bewegt. Damit schien zunächst klar, dass Materie und elektromagnetische Strahlung zwei wesentlich verschiedene Kategorien waren. Aber mit einiger Verzögerung führten die Experimente auch zum ersten großen Umsturz der Physik am Anfang des 20. Jahrhunderts, nämlich zur Formulierung der Speziellen Relativitätstheorie. Wichtig für uns ist, dass auch die Spezielle

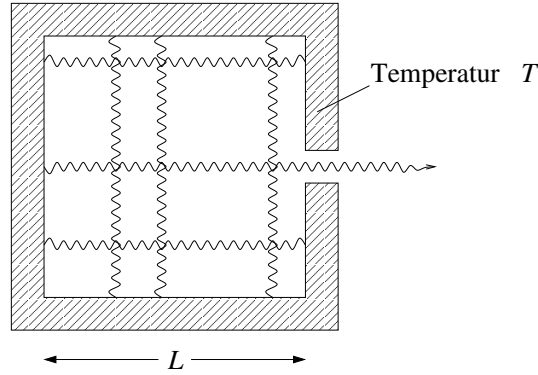
(und die Allgemeine) Relativitätstheorie nicht konzeptionell über das klassische Bild von genau lokalisierbaren Teilchen und von wellenartiger Strahlung hinausgeht.

2.2 Der schwarze Körper

Ein *schwarzer Körper* ist definiert als ein Körper, der die gesamte auftreffende Strahlung absorbiert. Es ist eigentlich eine Fehlbezeichnung, da der Körper im thermischen Gleichgewicht mit dem elektromagnetischen Feld genauso viel Energie emittiert wie absorbiert. Ein schwarzer Körper bei hohen Temperaturen ist also nicht schwarz, sondern er glüht.

Die Sonne verhält sich in guter Näherung wie ein schwarzer Körper, obwohl sie sicher nicht im thermischen Gleichgewicht mit dem Außenraum ist. Aber das Plasma, das wir tatsächlich sehen, die sogenannte Photosphäre der Sonne, ist näherungsweise im Gleichgewicht mit den tieferen Schichten, die als Wärmebad wirken. Das Emissionsspektrum entspricht daher näherungsweise dem eines schwarzen Strahlers.

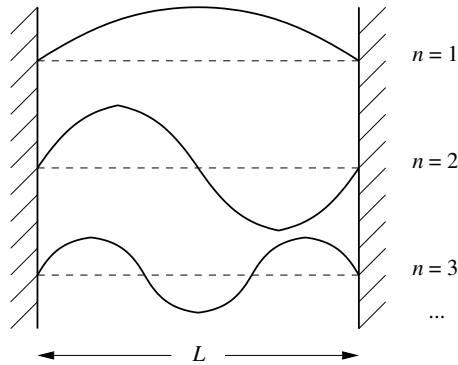
Eine sehr gute experimentelle Realisierung eines schwarzen Körpers ist ein Hohlraum mit einer kleinen Öffnung. Fast das gesamte von außen auf die Öffnung treffende Licht wird absorbiert und fast das gesamte aus der Öffnung austretende Licht wurde von den Innenwänden emittiert.



Im Rahmen der klassischen Physik lässt sich die Energieverteilung, die spektrale Energiedichte $w(\nu)$, des austretenden Lichts berechnen. Wir betrachten einen würfelförmigen, geerdeten, metallischen Hohlraum. Das elektromagnetische Feld im Innern lässt sich unter Vernachlässigung des Loches in stehende Wellen zerlegen. Es sind nur solche stehenden Wellen möglich, für die die parallelen Komponenten \vec{E}_{\parallel} des elektrischen Feldes und die Normalkomponente B_{\perp} des Induktionsfeldes auf der Oberfläche verschwinden.

In einer Dimension hätten wir die Bedingung

$$n \frac{\lambda}{2} = L, \quad n = 1, 2, 3, \dots \quad (2.1)$$



Also, mit der Wellenzahl $k = 2\pi/\lambda$,

$$k = n \frac{\pi}{L}, \quad n = 1, 2, 3, \dots \quad (2.2)$$

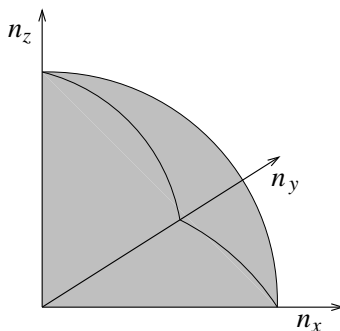
In einem würfelförmigen Hohlraum haben wir entsprechend für den Wellenvektor

$$\vec{k} = \begin{pmatrix} n_x \\ n_y \\ n_z \end{pmatrix} \frac{\pi}{L}, \quad n = 1, 2, 3, \dots \quad (2.3)$$

Die Frequenz der Welle ist

$$\nu = \frac{c}{\lambda} = \frac{c|\vec{k}|}{2\pi} = \frac{c}{2L} \sqrt{n_x^2 + n_y^2 + n_z^2} = \frac{c}{2L} |\vec{n}|. \quad (2.4)$$

Wie viele Feldmoden mit Frequenzen kleiner oder gleich ν existieren?



Für $\sqrt{n_x^2 + n_y^2 + n_z^2} \gg 1$ oder $L \gg \lambda$ können wir die Diskretheit der Punkte (n_x, n_y, n_z) vernachlässigen. Dann gibt es

$$N(\nu) = \underbrace{2}_{\substack{\text{Polarisationen} \\ \text{nur ein Oktant} \\ (n_x, n_y, n_z > 0)}} \cdot \underbrace{\frac{1}{8}}_{\substack{\text{Polarisationen} \\ \text{nur ein Oktant} \\ (n_x, n_y, n_z > 0)}} \cdot \frac{4\pi}{3} \left(\underbrace{\frac{2L}{c} \nu}_{=|\vec{n}|} \right)^3 = \frac{8\pi}{3} \frac{L^3}{c^3} \nu^3 \quad (2.5)$$

Feldmoden mit Frequenzen kleiner oder gleich ν . Die Anzahl der Moden mit Frequenzen im Intervall $[\nu, \nu + d\nu]$ ist $dN = (dN/d\nu) d\nu$ mit

$$\frac{dN}{d\nu} = 8\pi \frac{L^3}{c^3} \nu^2. \quad (2.6)$$

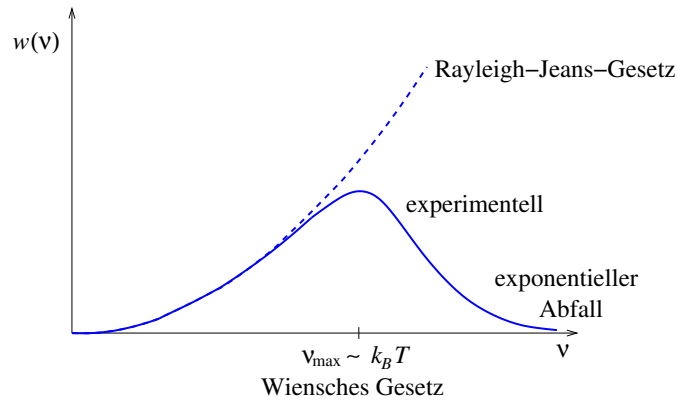
Nach dem Gleichverteilungssatz (siehe Vorlesung Thermodynamik und Statistische Physik) enthält im Gleichgewicht jede Mode die mittlere Energie $k_B T$, dazu tragen das \vec{E} - und das \vec{B} -Feld jeweils $k_B T/2$ bei. Die spektrale Energiedichte pro Volumen ist dann

$$w(\nu) = \frac{1}{L^3} \frac{dN}{d\nu} k_B T = 8\pi \frac{\nu^2}{c^3} k_B T. \quad (2.7)$$

Das ist das sogenannte *Rayleigh-Jeans-Gesetz*. Es ist im Rahmen der klassischen Physik exakt. Wir sehen aber sofort, dass sie nicht stimmen kann: die gesamte Energiedichte ist

$$w = \int_0^\infty d\nu w(\nu) = 8\pi \frac{k_B T}{c^3} \int_0^\infty d\nu \nu^2, \quad (2.8)$$

was bei großen Frequenzen divergiert (*Ultraviolett-Katastrophe*). Auch im Vergleich zum Experiment versagt die klassische Theorie:



Zur Lösung des Problems nahm Planck im Jahr 1900 an, dass die stehenden elektromagnetischen Wellen in Resonanz mit gewissen (nicht näher charakterisierten) Oszillatoren in der Innenwand des Hohlraums stünden und dass diese nur in Zuständen mit bestimmten *diskreten* Energien $E_n = n\varepsilon$, $n = 0, 1, 2, \dots$ existieren können. Das ist bemerkenswert, da es der Entwicklung der Quantentheorie um mehrere Jahre voraus ging; Niels Bohr veröffentlichte sein Atommodell erst 1913. Man zeigt in der Statistischen Physik, dass die Wahrscheinlichkeit für das Auftreten der Energie E_n proportional zu $e^{-E_n/k_B T} = e^{-n\varepsilon/k_B T}$ (genannt Boltzmann-Faktor) ist. Daher ist die mittlere Energie eines solchen Oszillators

$$\bar{E} = \frac{0 e^{-0/k_B T} + \varepsilon e^{-\varepsilon/k_B T} + 2\varepsilon e^{-2\varepsilon/k_B T} + \dots}{e^{-0/k_B T} + e^{-\varepsilon/k_B T} + e^{-2\varepsilon/k_B T} + \dots} = \frac{\sum_{n=0}^{\infty} n\varepsilon e^{-n\varepsilon/k_B T}}{\sum_{n=0}^{\infty} e^{-n\varepsilon/k_B T}}. \quad (2.9)$$

Dies können wir auch schreiben als

$$\bar{E} = \frac{\sum_{n=0}^{\infty} n\varepsilon e^{-n\beta\varepsilon}}{\sum_{n=0}^{\infty} e^{-n\beta\varepsilon}} \quad (2.10)$$

mit $\beta = 1/k_B T$. Die unendlichen Reihen lassen sich mittels der folgenden Umschreibung aufsummieren:

$$\begin{aligned} \bar{E} &= -\frac{d}{d\beta} \ln \sum_{n=0}^{\infty} e^{-n\beta\varepsilon} = -\frac{d}{d\beta} \ln \frac{1}{1 - e^{-\beta\varepsilon}} \quad \left| \text{(geometrische Reihe)} \right. \\ &= \frac{d}{d\beta} \ln(1 - e^{-\beta\varepsilon}) = \frac{\varepsilon e^{-\beta\varepsilon}}{1 - e^{-\beta\varepsilon}} \\ &= \frac{\varepsilon}{e^{\beta\varepsilon} - 1} = \frac{\varepsilon}{e^{\varepsilon/k_B T} - 1}. \end{aligned} \quad (2.11)$$

Die Experimente zeigten, dass ε proportional zur Frequenz ν sein muss. Die Proportionalitätskonstante nennen wir das *Plancksche Wirkungsquantum* h . Also schreiben wir

$$\varepsilon = h\nu. \quad (2.12)$$

Planck ersetzte dann die mittlere Energie $k_B T$ im Rayleigh-Jeans-Gesetz durch \bar{E} und erhielt so die *Plancksche Strahlungsformel*

$$w(\nu) = 8\pi \frac{\nu^3}{c^3} \frac{h}{e^{h\nu/k_B T} - 1}. \quad (2.13)$$

Sie geht für $h\nu \ll k_B T$ in die Rayleigh-Jeans-Formel über (da $e^{h\nu/k_B T} - 1 \cong h\nu/k_B T$) und sagt für $h\nu \gg k_B T$ einen exponentiellen Abfall voraus. Beachte, dass Planck diskrete Energieniveaus für die Oszillatoren in der Wand angenommen hat. Wir wissen heute, dass die Strahlungsformel tatsächlich von der diskreten Besetzung der Feldmoden, d. h. von der Teilchennatur des Lichts, herrührt. Die Wand hat nur die Funktion eines Wärmebades mit der Temperatur T .

2.3 Der Photoeffekt

Heinrich Hertz beobachtete 1887, dass bei Bestrahlung mit (UV-) Licht Elektronen aus Metalloberflächen austreten. Dieser sogenannte *Photoeffekt* oder *lichtelektrische Effekt* wurde von Wilhelm Hallwachs (1893–1922 Professor an der damaligen TH Dresden, 1921–1922 deren Rektor) im Detail experimentell untersucht. Hertz und Hallwachs machten folgende Beobachtungen:

1. Der Photoeffekt tritt nur auf, wenn die Lichtfrequenz ν eine materialabhängige Grenzfrequenz ν_g übersteigt.
2. Die (kinetische) Energie der austretenden *Photoelektronen* ist proportional zu $\nu - \nu_g$:

$$\frac{1}{2}mv^2 \sim \nu - \nu_g \text{ für } \nu > \nu_g \quad (2.14)$$

3. Sie ist unabhängig von der Lichtintensität, d.h. der transportierten Energie pro Zeit pro Fläche senkrecht zur Ausbreitungsrichtung,

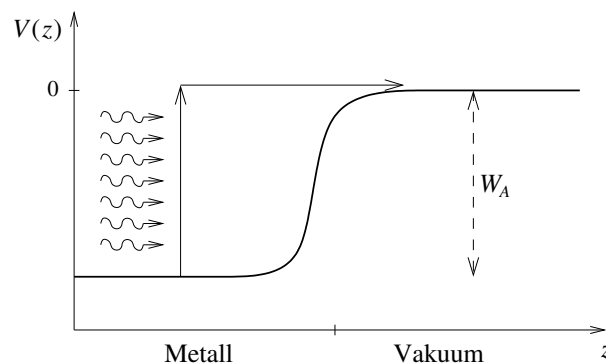
$$I = \frac{1}{2} c \varepsilon_0 E_0^2. \quad (2.15)$$

Hier ist E_0 ist die Amplitude des elektrischen Feldes der elektromagnetischen Lichtwelle.

4. Die Anzahl der pro Zeit emittierten Photoelektronen, also der Strom, ist proportional zur Intensität I .
5. Die Emission erfolgt ohne messbare Verzögerung, selbst bei sehr geringer Lichtintensität.

Aus der klassischen Elektrodynamik war bekannt, dass elektromagnetische Wellen Energie tragen, beschrieben durch die Intensität I . Es schien plausibel, dass diese Energie bei Absorption an einer Metalloberfläche auf Elektronen übertragen wird und diese so herauslösen kann. Was sagt die klassische Physik für die Photoelektronen voraus? Wir betrachten zwei Grenzfälle:

1. Falls die auf ein Elektron pro Periode τ der Lichtwelle übertragene Energie sehr klein im Vergleich zu seiner Bindungsenergie im Metall (der *Austrittsarbeit* W_A) ist, sollte es über viele Perioden immer mehr Energie aufnehmen, bis es schließlich genug Energie hat, um das Metall zu verlassen.

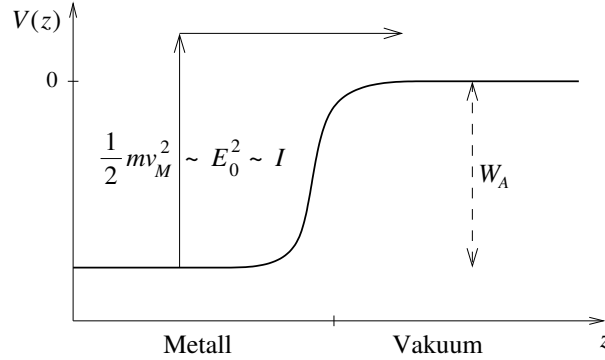


Die Energie der Photoelektronen sollte dann gering sein, denn sie haben ja gerade genug Energie, um das Metall zu verlassen. Insbesondere sollte die Elektronenenergie nicht von der Lichtfrequenz abhängen. Dies steht im Widerspruch zum Experiment. Außerdem sollte es bei geringer Lichtintensität eine Verzögerung bis zum Austritt der Photoelektronen geben, ebenfalls im Widerspruch zum Experiment.

2. Falls die Dauer der Energieübertragung bis zum Austritt klein im Vergleich zur Periode τ ist, sollte ein Kraftstoß proportional zur Feldamplitude E_0 auf die Elektronen wirken (die Kraft ist $\vec{F} = -e\vec{E}$). Waren sie vorher in Ruhe, sollten sie auf Geschwindigkeiten v_M (M für „Metall“) proportional zu E_0 beschleunigt werden. Ihre Energie relativ zum Grundzustand im Metall sollte dann

$$\frac{1}{2}mv_M^2 \sim E_0^2 \sim I \quad (2.16)$$

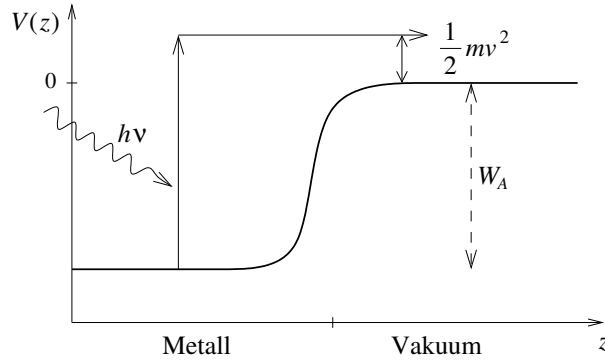
sein, also proportional zur Lichtintensität I , ebenfalls im Widerspruch zum Experiment. (Nach dem Austritt aus dem Metall sollte die kinetische Energie $(1/2)mv^2 = \alpha I - W_A$ mit einer Konstanten α betragen.)



Der Photoeffekt wurde 1905 korrekt von Albert Einstein erklärt, wofür er 1922 den Physik-Nobelpreis (für das Jahr 1921) erhielt. Einstein machte den Schritt, den Planck noch vermieden hatte, und postulierte, dass die elektromagnetische Strahlung selbst (und nicht gewisse Oszillatoren in festen Körpern wie bei Planck) aus Paketen mit der Energie $E = h\nu$ besteht. Deren Anzahl pro Zeit und Querschnittsfläche muss dann proportional zur Intensität sein. Nach Einstein kann Energie zwischen dem Licht und den Metallelektronen nur durch Absorption oder Emission ganzer Pakete ausgetauscht werden (*Lichtquantenhypothese*). Diese Pakete (Lichtquanten) verhalten sich also wie Teilchen, deren Anzahl aber im Unterschied z. B. zu Elektronen nicht erhalten ist. Diese Teilchen nennen wir heute *Photonen*.

Absorbiert ein Elektron mit der Bindungsenergie (Austrittsarbeit) W_A ein Photon der Energie $h\nu$, so bleibt ihm die kinetische Energie

$$\frac{1}{2}mv^2 = h\nu - W_A. \quad (2.17)$$



Damit $(1/2)mv^2 > 0$ ist, muss $h\nu > W_A$ gelten, was für die Grenzfrequenz

$$\nu_g = \frac{W_A}{h} \quad (2.18)$$

impliziert, in sehr guter Übereinstimmung mit dem Experiment. Weiter ist der Photoelektronenstrom proportional zur Zahl der absorbierten Photonen pro Zeit und damit zur Intensität. Alle Ergebnisse der Theorie sind experimentell sehr gut bestätigt.

Da die Photonen die Geschwindigkeit c haben, sind sie nach der Speziellen Relativitätstheorie (SRT) masselos. Die SRT ergibt dann für ihren Impuls p :

$$E^2 = p^2c^2 + m^2c^4 = p^2c^2 \quad \Rightarrow \quad p = \frac{E}{c} = \frac{h\nu}{c} = \frac{h}{\lambda}. \quad (2.19)$$

Mit der Kreisfrequenz $\omega = 2\pi\nu$ und der Wellenzahl $k = 2\pi/\lambda$ findet man die häufig verwendeten Beziehungen

$$E = \frac{h\omega}{2\pi} \quad \text{und} \quad p = \frac{hk}{2\pi}. \quad (2.20)$$

Man definiert die Abkürzung

$$\hbar := \frac{h}{2\pi} = 1,055 \cdot 10^{-34} \text{ Js} \quad (2.21)$$

und erhält

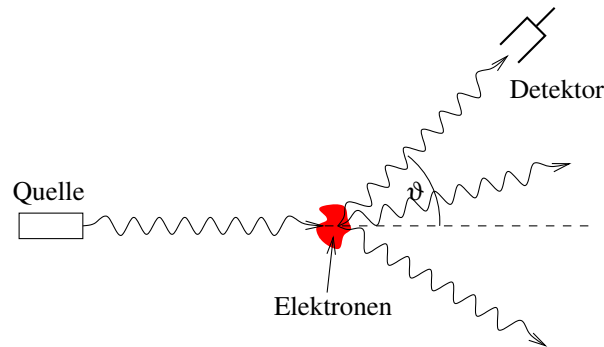
$$E = \hbar\omega \quad \text{und} \quad p = \hbar k. \quad (2.22)$$

2.4 Der Compton-Effekt

Der Compton-Effekt ist die elastische Streuung von Licht an freien oder zumindest schwach gebundenen Elektronen. Er konnte nicht mehr zum Umsturz der klassischen Physik beitragen, da er erst 1924 entdeckt wurde, hat aber Einsteins Teilchenbild der elektromagnetischen Strahlung bestätigt. Experimental fand Arthur H. Compton, dass sich die Wellenlänge von Licht bei der Streuung um

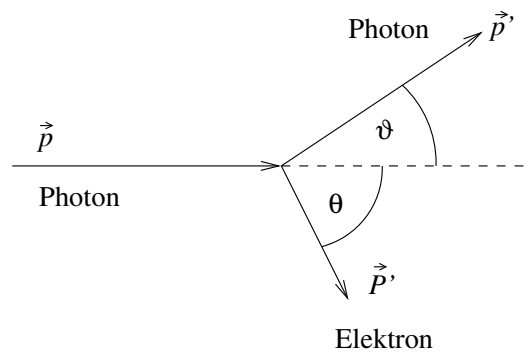
$$\Delta\lambda = 4\pi \frac{\hbar}{mc} \sin^2 \frac{\vartheta}{2} \quad (2.23)$$

verschiebt, wobei ϑ der Ablenkungswinkel (Streuwinkel) des Lichts ist.



$\Delta\lambda$ ist insbesondere unabhängig von der Wellenlänge λ der einfallenden Strahlung und damit von ihrer Frequenz. Klassisch würde man eine kontinuierliche Energieübertragung von Licht an die anfangs ruhenden Elektronen erwarten und damit eine Verteilung von $\Delta\lambda$ für gegebene Raumrichtung, nicht einen scharfen Wert für $\Delta\lambda$.

Im Rahmen des Photonenbildes kann man das Experiment dagegen als elastischen Stoß zwischen einem Photon und einem anfangs ruhenden Elektron verstehen. Beim elastischen Stoß müssen Impuls und Energie erhalten bleiben:



$$\vec{p} + 0 = \vec{p}' + \vec{P}', \quad (2.24)$$

$$pc + mc^2 = p'c + \underbrace{\sqrt{(P')^2 c^2 + m^2 c^4}}_{\text{(relativistische Energie-Impuls-Beziehung)}}. \quad (2.25)$$

Aus der Impulserhaltung folgt

$$\begin{aligned} \vec{P}' = \vec{p} - \vec{p}' &\Rightarrow (\vec{P}')^2 = p^2 + (p')^2 - 2\vec{p} \cdot \vec{p}' \\ &= p^2 + (p')^2 - 2pp' \cos \vartheta. \end{aligned} \quad (2.26)$$

Eingesetzt in die Energieerhaltung ergibt sich

$$pc + mc^2 = p'c + \sqrt{p^2 + (p')^2 - 2pp' \cos \vartheta + m^2 c^2} c \quad (2.27)$$

$$\Rightarrow p - p' + mc = \sqrt{p^2 + (p')^2 - 2pp' \cos \vartheta + m^2 c^2} \quad (2.28)$$

$$\Rightarrow p^2 + (p')^2 + m^2 c^2 - 2pp' + 2pmc - 2p'mc = p^2 + (p')^2 - 2pp' \cos \vartheta + m^2 c^2. \quad (2.29)$$

Also mit $p = h/\lambda$,

$$-\frac{2h^2}{\lambda\lambda'} + \frac{2mhc}{\lambda} - \frac{2mhc}{\lambda'} = -\frac{2h^2}{\lambda\lambda'} \cos \vartheta \quad (2.30)$$

$$\Rightarrow -1 + \frac{mc}{h} \lambda' - \frac{mc}{h} \lambda = -\cos \vartheta \quad (2.31)$$

$$\Rightarrow \Delta\lambda = \lambda' - \lambda = \frac{h}{mc} (1 - \cos \vartheta) = 2 \frac{h}{mc} \sin^2 \frac{\vartheta}{2} = 4\pi \frac{h}{mc} \sin^2 \frac{\vartheta}{2}. \quad (2.32)$$

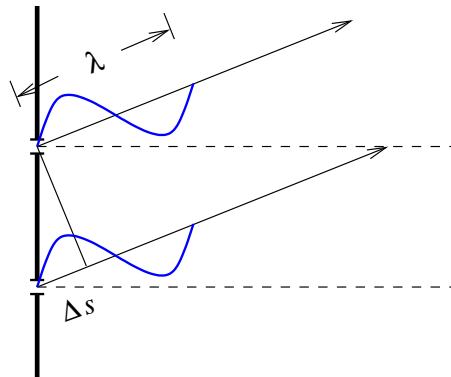
Das ist die beobachtete Verschiebung. $\lambda_C := h/mc = 2,426 \cdot 10^{-12} \text{ m}$ nennt man die *Compton-Wellenlänge* des Elektrons und $\lambda_C := h/mc$ die *reduzierte Compton-Wellenlänge*.

2.5 Welle oder Teilchen?

Die bisher besprochenen Experimente lassen sich auf natürliche Weise nur im Teilchenbild des Lichts erklären. Für Materie erschien das Teilchenbild ohnehin als gesichert. Es hat eine lange Geschichte, wichtig waren z. B. die Experimente von Thomson, Townsend und Wilson von 1896 zum elementaren Charakter des Elektrons und zu seiner Ladung und Masse. In diesem Abschnitt besprechen wir kurz Experimente, die zeigen, dass das Teilchenbild nicht ausreicht.

2.5.1 Welle-Teilchen-Dualität des Lichts

Die schon lange bekannten Beugungs- und Interferenzerscheinungen des Lichts werden sehr erfolgreich durch das Wellenbild erklärt. Z. B. finden wir beim Doppelspaltexperiment ungefähr äquidistante Maxima (und Minima), die auftreten, wenn der Wegunterschied Δs zwischen den beiden Spalten und einem Punkt auf dem weit entfernten Schirm $\Delta s = n\lambda$ (bzw. $\Delta s = (n + 1/2)\lambda$) mit $n = 0, \pm 1, \pm 2, \dots$ beträgt.



Man beobachtet dasselbe Interferenzmuster, wenn man die Lichtintensität so weit verringert, dass praktisch sicher immer nur ein *einzelnes* Photon zur Zeit den Doppelspalt passiert. Im klassischen Teilchenbild erwartet man, dass, wenn das Photon durch den rechten Spalt fliegt, die Position des linken Spalts keine Rolle spielt und umgekehrt. Das Bild auf dem Schirm sollte also einfach die Überlagerung der von den beiden Spalten erzeugten Bilder sein. Das beobachtet man aber offenbar nicht. Man ist also gezwungen anzunehmen, dass jedes einzelne Photon *beide* Spalte sieht, obwohl diese einen makroskopischen Abstand voneinander haben. Das ist mit klassischen Teilchenbahnen nicht vereinbar.

Die Schlussfolgerung ist, dass sich Licht in manchen Situationen wie eine Welle und in anderen wie eine Ansammlung von Teilchen verhält. Dies wurde *Welle-Teilchen-Dualität* genannt, aber zunächst nicht verstanden.

2.5.2 Welle-Teilchen-Dualität der Materie

In den Jahren 1927/28 wiesen mehrere Experimente unabhängig voneinander nach, dass auch Materiestrahlen Beugung und Interferenz zeigen. Dabei wurden die Reflexion und Transmission von Elektronenstrahlen an Kristallen untersucht. Die aus den Experimenten ermittelte Wellenlänge der Elektronenwellen betrug

$$\lambda = \frac{h}{p} = \frac{h}{mv} \quad (2.33)$$

(*de Broglie-Wellenlänge*). Beachte, dass diese immer größer als die Compton-Wellenlänge ist und sich ihr im relativistischen Limes annähert:

$$\lambda = \frac{h}{mv} \stackrel{v < c}{>} \frac{h}{mc} = \lambda_C. \quad (2.34)$$

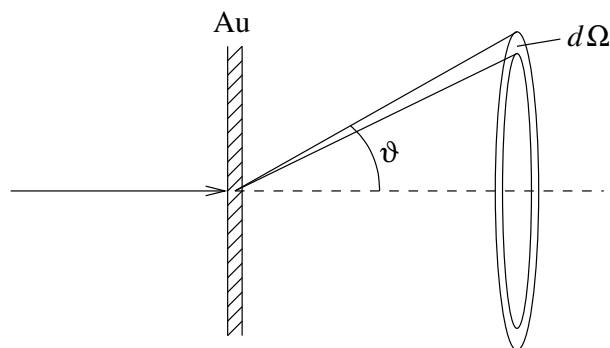
In diesem Fall ging die theoretische Voraussage (de Broglie 1923) dem Nachweis voraus.

Die Experimente zeigen also, dass sich auch Elektronen manchmal wie Wellen und manchmal wie Teilchen verhalten. Später wurde dies auch für andere Teilchen gezeigt, für Neutronen ist es z. B. inzwischen eine Standardtechnik für die Strukturbestimmung von Kristallen und die Bestimmung von magnetischer Ordnung.

2.6 Das Versagen des Rutherford'schen Atommodells

Ernest Rutherford untersuchte in den Jahren 1906 bis 1913 die Streuung von α -Teilchen (aus radioaktivem Zerfall) in Transmission durch dünne Goldfolien. Die gemessene Häufigkeit der Ablenkung in ein Raumwinkelement $d\Omega$ unter einem Winkel ϑ gegenüber der Einfallsrichtung lässt sich sehr gut klassisch beschreiben. Dabei muss man annehmen, dass fast die gesamte Atommasse in einem sehr kleinen Kern der Ladung $+Ze$ vereinigt ist. Dann ergibt sich die Rutherford'sche Streuformel

$$dN \sim N \frac{Z^2}{\sin^4 \frac{\vartheta}{2}} d\Omega. \quad (2.35)$$



Die Übereinstimmung mit der klassischen Streuformel mit dem Experiment scheint zunächst die klassische Physik (Mechanik und Elektrodynamik) glänzend zu bestätigen. (In diesem Fall ergibt eine quantenmechanische Betrachtung jedoch dieselbe Streuformel, siehe Vorlesung Quantentheorie 2.) Die Experimente untergraben die klassische Physik aber an anderer Stelle, nämlich beim Verständnis des Rutherford'schen Atommodells.

2.6.1 Stabilität der Atome

Nach Rutherford bestehen Atome aus Elektronen (leichten Teilchen der Ladung $-e$) und Kernen (schweren Teilchen der Ladung $+Ze$). Die Elektronen laufen gemäß der Gesetze der klassischen Mechanik auf Bahnen um den Kern. Da das Coulomb-Potential von der Form $V(\vec{r}) \sim -1/r$ und anziehend ist, handelt es sich um ein Kepler-Problem; die Bahnen müssen im klassischen Bild Ellipsen sein. Nun findet man in der klassischen Elektrodynamik, dass beschleunigte Ladungen elektromagnetische Wellen abstrahlen. Elektronen auf Ellipsenbahnen sind natürlich beschleunigt,

$$\ddot{\vec{r}} = \frac{\vec{F}}{m} = -\frac{e\vec{E}}{m} = -\frac{1}{m} \frac{Ze^2}{4\pi\epsilon_0 r^3} \vec{r} \neq 0. \quad (2.36)$$

Die Abstrahlung reduziert die Energie des Elektrons. Da seine kinetische Energie nicht negativ werden kann, muss die potentielle Energie immer kleiner werden, d. h. das Elektron fällt in den Potentialtopf des Kerns. Man kann die möglichen Elektronenbahnen im Rahmen der klassischen Physik berechnen. Man findet, dass ein Elektron innerhalb einer Zeit von der Größenordnung 10^{-11} s von einer Bahn mit einem typischen Atomradius (10^{-10} m) in den Kern fällt. (Was im Kern mit dem Elektron passiert, kann die klassische Physik nicht beantworten. Ist die Kernladung über das Kernvolumen ausgeschmiert, gibt es aber zumindest keine Divergenz der Elektronenenergie nach $-\infty$.)

Die klassische Physik sagt also voraus, dass das Rutherford-Atom auf einer typischen Zeit von 10^{-11} s kollabiert. Atome sollten unter Abstrahlung von elektromagnetischen Wellen in kleine, elektrisch neutrale Partikel (Kerne mit darin gefangenen Elektronen) übergehen, im Widerspruch zu allen Beobachtungen. In diesem Bild gäbe es nichts, was den Abstand der Kerne in Festkörpern und Flüssigkeiten bei etwa 10^{-10} m stabilisieren würde. Es gäbe damit insbesondere keine Erde.

2.6.2 Atomspektren

Sollte das Rutherford-Atom durch unbekannte Mechanismen den Kollaps vermeiden, so sollte es jedenfalls beliebige Beträge von Energie aufnehmen und abgeben können, zumindest bis zu einem gewissen Maximum. Beim klassischen Kepler-Problem gibt es nämlich keinen Grund, warum bestimmte Werte der Bindungsenergie vor anderen ausgezeichnet sein sollten.

Balmer beobachtete nun 1885, dass ein durch Stöße mit Elektronen angeregtes verdünntes Wasserstoffgas nur Licht mit bestimmten, diskreten Wellenlängen abstrahlte. Später wurden noch weitere Emissionslinien bei UV- und IR-Wellenlängen beobachtet. *Alle* beobachteten Wellenlängen gehorchten der empirischen *Rydberg-Formel*

$$\frac{\nu}{c} = \frac{1}{\lambda} = R \left(\frac{1}{n^2} - \frac{1}{m^2} \right) \quad (2.37)$$

mit natürlichen Zahlen $m > n$ und einem Faktor R von der Dimension einer inversen Länge. Außerdem fand man, dass offenbar alle dadurch vorhergesagten Linien im beobachtbarem Spektrum tatsächlich vorkamen. Für feste n ergeben sich *Serien* von Spektrallinien mit $m = n + 1, n + 2, n + 3, \dots$:

$n = 1$	Lyman-Serie
$n = 2$	Balmer-Serie
$n = 3$	Paschen-Serie
$n = 4$	Bracket-Serie
$n = 5$	Pfund-Serie
\dots	\dots

Dasselbe Verhalten, nur mit anderem Vorfaktor, erhält man für die wasserstoffähnlichen Ionen He^+ , Li^{2+} usw., die alle genau ein Elektron enthalten. Bei komplizierteren Atomen und Ionen findet man ebenfalls Serien, die aber nicht einer einfachen Formel gehorchen. Es gilt jedoch auch hier das *Ritzsche Kombinationsprinzip*: Die Summe und die Differenz zweier Frequenzen im Spektrum sind oft wieder Frequenzen im Spektrum.

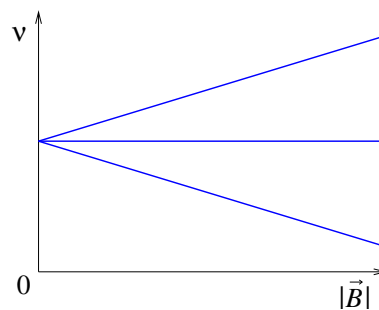
Wir kommen im nächsten Kapitel 3 auf die Deutung der Spektren im Rahmen der „älteren Quantentheorie“ (Bohr, Sommerfeld) zurück. Im Vorgriff darauf ist es nach Einsteins Lichtquantenhypothese naheliegend, das Spektrum durch Emission einzelner Photonen der Energie $h\nu$ zu interpretieren. Dann ist es natürlich,

$$h\nu = \underbrace{hcR}_{=: E_R} \left(\frac{1}{n^2} - \frac{1}{m^2} \right) = -\frac{E_R}{m^2} + \frac{E_R}{n^2} \quad (2.38)$$

unter Berücksichtigung der Energieerhaltung so zu deuten, dass das Wasserstoffatom bei der Emission einen Übergang zwischen zwei Zuständen mit den Energien $-E_R/n^2$ und $-E_R/m^2$ ausführt. $E_R = 13,6 \text{ eV}$ heißt *Rydberg-Energie* oder, als Einheit der Energie aufgefasst, *Rydberg*, d. h. $1 \text{ R} = 13,6 \text{ eV}$. Das Argument führt zu dem Schluss, dass das Wasserstoffatom nur *diskrete* Energiewerte $-E_R/n^2$, $n = 1, 2, \dots$ haben kann.

2.6.3 Der Zeeman-Effekt

Die Spektren von Atomen zeigen in einem angelegten Magnetfeld \vec{B} eine zu $|\vec{B}|$ proportionale Aufspaltung in eine ganze Zahl von Linien. Dies beobachtete Zeeman 1896.



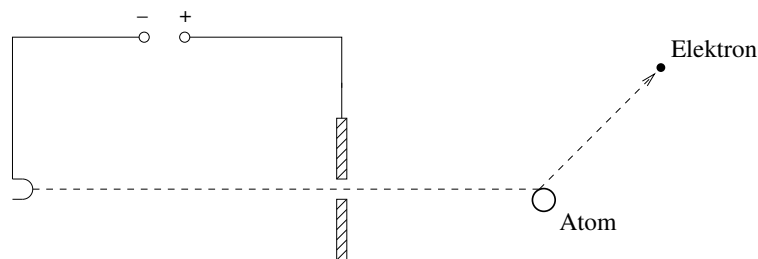
Der lineare Zusammenhang deutet auf die Energie $E_Z = -\vec{\mu} \cdot \vec{B}$ eines magnetischen Dipolmoments im \vec{B} -Feld hin. Im Rahmen des Rutherford'schen Atommodells ist im Prinzip verständlich, dass die Elektronenbahnen im Atom mit einem magnetischen Dipolmoment verbunden sind. Ein auf einer Kepler-Bahn umlaufendes Elektron realisiert einen Kreisstrom. Dieser erzeugt nach den Gesetzen der Elektrodynamik ein magnetisches Dipolmoment senkrecht zur Bahnebene. Im Photonenbild ist es naheliegend, die Verschiebung der Frequenz im Feld als

$$\Delta\nu = \frac{\Delta E_Z}{h} = -\frac{\Delta\mu_z B}{h} \quad (2.39)$$

mit $\vec{B} = B\hat{e}_z$ zu deuten. Die Diskretheit der Linien zeigt dann, dass sich die z -Komponente des Dipolmoments nur um diskrete Werte ändern kann, im Einklang mit dem späteren Stern-Gerlach-Experiment, siehe Abschnitt 2.7.

2.6.4 Der Franck-Hertz-Versuch

Die Vorstellung diskreter Energiezustände der Atome wurde 1914 durch Franck und Hertz bestätigt. Sie untersuchten die Streuung von monoenergetischen Elektronen an Atomen.



Die Messung der (kinetischen) Energie der Elektronen nach der Streuung ergab, dass die Energiedifferenz

$$\Delta E = E_{\text{vorher}} - E_{\text{nachher}} \quad (2.40)$$

nur diskrete Werte $0 < \Delta E_1 < \Delta E_2 < \dots$ annehmen konnte. War insbesondere $E_{\text{vorher}} < \Delta E_1$, so betrug die Energiedifferenz $\Delta E = 0$, da schon die kleinste nicht verschwindende Energieänderung $\Delta E = \Delta E_1$ wegen $E_{\text{nachher}} = E_{\text{vorher}} - \Delta E_1 < 0$ nicht möglich war. Es trat dann also nur elastische Streuung auf.

Die diskreten Energien $\Delta E_1, \dots$ nimmt offenbar das Atom auf. Es liegt nahe, sie als Differenzen zwischen diskreten atomaren Energiezuständen zu deuten. Man beobachtet dieselben Energien wie im optischen Spektrum. Für das Wasserstoffatom ist z. B.

$$\Delta E_1 = -\frac{E_R}{2^2} + \frac{E_R}{1^2} = \frac{3}{4} E_R, \quad (2.41)$$

$$\Delta E_2 = -\frac{E_R}{3^2} + \frac{E_R}{1^2} = \frac{8}{9} E_R \quad \text{usw.} \quad (2.42)$$

2.7 Der Stern-Gerlach-Versuch

Dauermagneten sind seit der Antike bekannt. Experimente von Oersted zeigten 1819, dass elektrische Ströme Magnetfelder hervorrufen. Es lag daher nahe, das Magnetfeld von Dauermagneten durch elektrische Ströme in ihrem Inneren zu erklären. Das Rutherfordsche Atommodell lieferte eine offensichtliche Erklärung für Ströme im Inneren der einzelnen Atome durch die Bahnbewegung geladener Elektronen. Diese Ströme sollten magnetische Dipolmomente erzeugen. Otto Stern und Walther Gerlach wollten 1921/22 das magnetische Dipolmoment einzelner Atome untersuchen und schickten dazu einen Strahl von Silberatomen durch ein inhomogenes Magnetfeld. Die Energie eines magnetischen Dipolmoments $\vec{\mu}$ in einem \vec{B} -Feld ist

$$E_{\text{Zeeman}} = -\vec{\mu} \cdot \vec{B}. \quad (2.43)$$

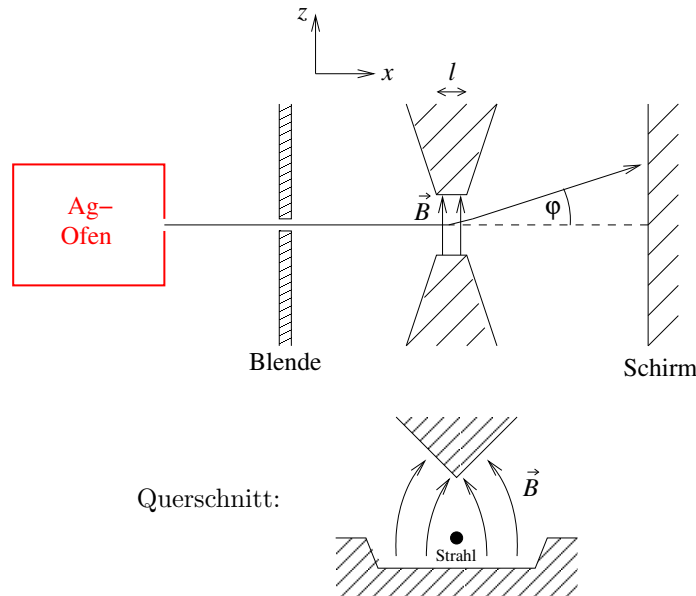
Wir legen die z -Achse entlang \vec{B} , dann ist

$$E_{\text{Zeeman}} = -\vec{\mu} \cdot \hat{z} B = -\mu_z B \quad (2.44)$$

(wir bezeichnen Einheitsvektoren durch ein Dach über dem Symbol). Aufgrund dieser potentiellen Energie wirkt auf den Dipol die Kraft

$$\vec{F} = -\vec{\nabla} E_{\text{Zeeman}} = \mu_z \vec{\nabla} B. \quad (2.45)$$

Diese Kraft ist offensichtlich nur in einem inhomogenen Feld von Null verschieden.



Wir berechnen die Ablenkung der Atome unter der Annahme, dass sie sich nur sehr kurz im Bereich des \vec{B} -Feldes aufhalten. Dann erfahren sie einen Kraftstoß

$$\Delta \vec{p} = \vec{F} \Delta t \quad \text{mit} \quad \Delta t = \frac{l}{v}. \quad (2.46)$$

Hier ist v die Geschwindigkeit des einlaufenden Atoms. Die Geschwindigkeitskomponente in z -Richtung ist vor dem \vec{B} -Feld-Bereich Null und dahinter

$$v_z = \frac{p_z}{m} = \frac{F_z}{m} \Delta t = \frac{\mu_z \frac{\partial B}{\partial z}}{m} \frac{l}{v}. \quad (2.47)$$

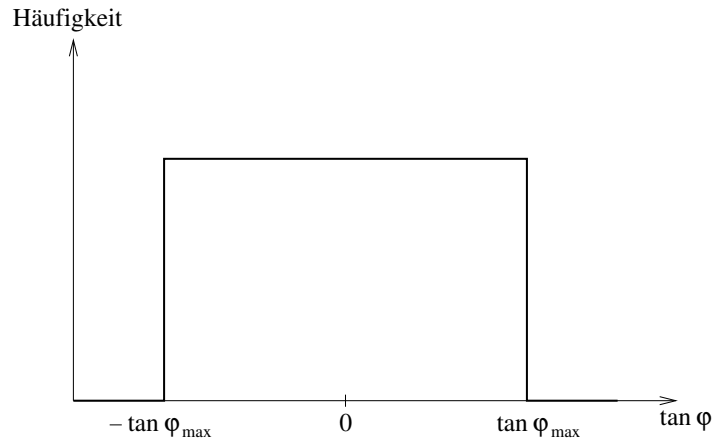
Der Ablenkwinkel φ erfüllt

$$\tan \varphi = \frac{v_z}{v_x} = \frac{v_z}{v} = \frac{l}{mv^2} \frac{\partial B}{\partial z} \mu_z. \quad (2.48)$$

Wir nehmen an, dass die Silberatome ein Dipolmoment vom Betrag $\mu > 0$ haben. Beim Eintritt in das \vec{B} -Feld sollte die Richtung des Dipolmoments der Atome beliebig verteilt (isotrop) sein, daher sollte die z -Komponente μ_z eine kontinuierliche Verteilung zwischen $-\mu$ und μ zeigen. Entsprechend sollte man auf dem Schirm eine breite Verteilung zwischen den Winkeln $-\varphi_{\max}$ und φ_{\max} mit

$$\tan \varphi_{\max} = \frac{l}{mv^2} \frac{\partial B}{\partial z} \mu \quad (2.49)$$

erwarten.

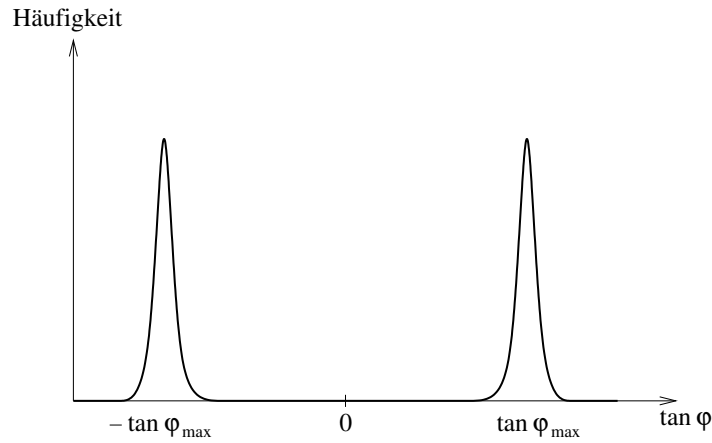


$\tan \varphi$ sollte sogar *gleichverteilt* sein, denn $\tan \varphi \sim \mu_z = \vec{\mu} \cdot \hat{z} = \mu \cos \theta$ (θ ist der Polarwinkel des Dipolmoments in Kugelkoordinaten) und die Verteilungsfunktion von $\cos \theta$ ist

$$\begin{aligned} P(\cos \theta) &= \frac{1}{4\pi} \int_0^{2\pi} d\phi' \int_0^\pi d\theta' \sin \theta' \delta(\cos \theta - \cos \theta') \\ &= \frac{1}{2} \int_0^\pi d\theta' \sin \theta' \delta(\cos \theta - \cos \theta') \quad \left| \text{Substitution } u' = \cos \theta' \right. \\ &= \frac{1}{2} \int_{-1}^1 du' \delta(\cos \theta - u') = \frac{1}{2} = \text{const} \end{aligned} \quad (2.50)$$

für $\cos \theta \in [-1, 1]$.

Das beobachtet man aber nicht. Stattdessen findet man zwei ziemlich schmale Maxima in der Häufigkeit, die die gleiche Höhe haben.



Die Verbreiterung beruht auf der endlichen Auflösung des Experiments und ist nicht fundamentaler Natur. Die zwei beobachteten Ablenkungswinkel entsprechen in guter Näherung

$$\mu_z = \pm \mu_B := \pm \frac{e\hbar}{2m_e} \quad (2.51)$$

mit der Elektronenmasse m_e (nicht der Masse der Silberatome). Die z -Komponente des Dipolmoments kann offenbar nur zwei diskrete Werte $\pm \mu_B$ (μ_B ist das Bohr-Magneton) annehmen. Das ist klassisch nicht verständlich. Wir werden die in den Versuchen von Stern und Gerlach sowie Zeeman beobachtete Diskretheit später aus den quantenmechanischen Eigenschaften des Drehimpulses erklären.

Für Ag^+ -Ionen beobachtet man keine Aufspaltung. Sie haben also kein magnetisches Dipolmoment. Man kann das beobachtete Moment demnach dem im neutralen Atom aber nicht im Ion vorhandenen (Leucht-) Elektron in der äußeren Schale zuschreiben.

Kapitel 3

Die ältere Quantentheorie

Die heute so genannte ältere Quantentheorie wurde von Niels Bohr und anderen ab 1913 entwickelt, um die offenbar diskreten Energiewerte von Atomen zu beschreiben und die Berechnung der Atomspektren zu ermöglichen. Sie beruht auf der Annahme, dass die klassische Physik die Dynamik von Teilchensystemen im Prinzip beschreibt, dass aber nur bestimmte Teilchenbahnen tatsächlich realisiert sind. Welche dies sind, ergibt sich aus *ad hoc* eingeführten *Quantisierungsregeln*. Wir werden sehen, dass die Bohrsche Quantentheorie nur etwas über geschlossene Bahnen aussagt.

3.1 Die Bohrschen Postulate

Die Annahme, dass nur bestimmte Bahnen vorkommen, widerspricht offensichtlich der klassischen Physik, die für geladene Teilchen eine kontinuierliche Energieabnahme durch Abstrahlung elektromagnetischer Wellen voraussagt. Diese Abstrahlung muss durch einen (von Bohr nicht erklärten) Mechanismus verhindert werden. Andererseits wurde ja eine Abstrahlung beobachtet, aber nur bei diskreten Frequenzen. Bohr formulierte seine Annahmen in der Form von zwei Postulaten:

1. Periodische Bewegungen können nur mit bestimmten diskreten Energien E_1, E_2, \dots erfolgen. Sie sind strahlungslos.
2. Übergänge zwischen erlaubten periodischen Bahnen der Energien E_n und E_m erfolgen unter Emission oder Absorption von elektromagnetischer Strahlung mit einer Frequenz der Form

$$\nu = \frac{E_n - E_m}{h}. \quad (3.1)$$

Hieraus folgt noch nicht, dass ein System mit periodischen Bahnen, z. B. ein Atom, einen *Grundzustand* niedrigster Energie haben muss. (Das folgt selbst dann nicht, wenn die Energie wie üblich nach unten beschränkt ist. Gegenbeispiel: $E_n = E_1/n, n = 1, 2, \dots$) In physikalisch relevanten Fällen existiert aber ein Grundzustand. Ist ein Atom im Grundzustand, so kann es keine Energie mehr abstrahlen und bleibt daher in diesem Zustand, bis es durch eine äußere Störung angeregt wird. Die Frequenz $\nu = (E_n - E_m)/h$ können wir im Teilchenbild des Lichts leicht deuten. Mit der Planck-Beziehung $\epsilon = h\nu$ ergibt sich, dass bei einem Übergang zwischen erlaubten Bahnen ein Photon emittiert oder absorbiert wird. Beachte, dass das Photon in moderner Sprechweise “virtuell” sein kann, d. h. es kann unmittelbar von einem anderen Teilchen emittiert oder absorbiert werden, z. B. von einem freien Elektron beim Franck-Hertz-Versuch.

3.2 Die Quantisierungsregeln von Bohr und Sommerfeld

Die Bohrschen Postulate sagen noch nichts darüber, *welche* der klassischen periodischen Bahnen erlaubt sind. Wir diskutieren dies, wie Bohr, zunächst für das Wasserstoffatom und dann allgemeiner.

3.2.1 Das Wasserstoffatom

In der klassischen Mechanik führt die anziehende Coulomb-Wechselwirkung zwischen Proton und Elektron auf ein Kepler-Problem. Für alle negativen Energien $E < 0$ findet man elliptische (im Grenzfall kreisförmige) Bahnen mit der Umlauffrequenz

$$\nu_{\text{Bahn}} = \frac{1}{T_{\text{Bahn}}} = \frac{1}{\pi} \frac{4\pi\epsilon_0}{e^2} \sqrt{\frac{2|E|^3}{m}}, \quad (3.2)$$

wobei $m \approx m_e$ die reduzierte Masse ist, und mit der großen Halbachse

$$a = \frac{1}{2} \frac{e^2}{4\pi\epsilon_0} \frac{1}{|E|}. \quad (3.3)$$

Wie in Kap. 1 diskutiert, gilt die klassische Physik auf makroskopischen Längenskalen sehr gut. Also erwarten wir, dass die Quantentheorie im Grenzfall großer Längenskalen, also hier großer a , mit der klassischen Physik übereinstimmt. Das ist die Anwendung des wichtigen *Korrespondenzprinzips* auf das Wasserstoffatom. Wir kommen darauf zurück. Große a entsprechen nach Glg. (3.2) und (3.3) betragsmäßig kleinen Energien $|E|$ und kleinen Frequenzen ν_{Bahn} .

Klassisch betrachtet, sollte ein Atom mit der gegebenen Energie E elektromagnetische Wellen mit der Frequenz ν_{Bahn} und höheren Harmonischen $2\nu_{\text{Bahn}}, 3\nu_{\text{Bahn}}, \dots$ abstrahlen. (Höhere Harmonische treten auf, weil das Elektron auf einer Ellipsenbahn keinen idealen schwingenden Dipol darstellt.) Die Bohrschen Postulate behaupten, dass das Atom Wellen mit Frequenzen der Form $(E_n - E_m)/h$ abstrahlt. Gemäß des Korrespondenzprinzips fordern wir nun, dass für kleine Energie $|E| = |E_n|$ die kleinste klassisch mögliche Frequenz ν_{Bahn} (die Grundfrequenz) gleich der kleinsten quantenmechanisch möglichen Frequenz ist,

$$\nu_{\text{Bahn}} = \nu = \frac{E_n - E_{n-1}}{h}. \quad (3.4)$$

Da E_n mit n anwächst – so zählen wir die erlaubten Energien – und E_n negativ ist, entsprechen kleine $|E_n|$ großen Quantenzahlen n . Wenn $E_n - E_{n-1} \ll |E_n|$ gilt, was wir nachträglich überprüfen sollten, können wir schreiben

$$\nu_{\text{Bahn}} = \frac{1}{h} \frac{dE_n}{dn}. \quad (3.5)$$

Diese Beziehung gilt gemäß des Korrespondenzprinzips für große n . Bohr forderte nun, dass sie für *alle* n gelten soll. Die Gleichung

$$\frac{1}{h} \frac{dE_n}{dn} = \nu_{\text{Bahn}}(E_n) \quad (3.6)$$

ist dann eine Differentialgleichung für E_n als Funktion von n . Trennung der Variablen ergibt

$$\frac{dE_n}{\nu_{\text{Bahn}}(E_n)} = h dn \quad (3.7)$$

und

$$\int_{-\infty}^{E_n} \frac{dE'_n}{\nu_{\text{Bahn}}(E'_n)} = (n + n_c) h, \quad (3.8)$$

wobei n_c eine (Integrations-) Konstante ist. Für das Wasserstoffatom erhalten wir

$$(n + n_c) h = \int_{-\infty}^{E_n} \frac{dE'_n}{\frac{1}{\pi} \frac{4\pi\epsilon_0}{e^2} \sqrt{\frac{2|E'_n|^3}{m}}} = \frac{e^2}{4\epsilon_0} \sqrt{\frac{m}{2}} \int_{-\infty}^{E_n} \frac{dE'_n}{(-E'_n)^{3/2}}. \quad (3.9)$$

Für $E_n < 0$ können wir das Integral auswerten,

$$(n + n_c) h = \frac{e^2}{4\epsilon_0} \sqrt{\frac{m}{2}} 2 \frac{1}{\sqrt{-E'_n}} \Big|_{-\infty}^{E_n} = \frac{e^2}{2\epsilon_0} \sqrt{\frac{m}{2}} \frac{1}{\sqrt{-E_n}}. \quad (3.10)$$

Dies ergibt

$$E_n = -\frac{e^4 m}{8\varepsilon_0^2 h^2} \frac{1}{(n + n_c)^2}. \quad (3.11)$$

Die Konstante n_c bekommt man nicht aus dem Korrespondenzprinzip. Vergleich mit dem beobachteten Spektrum ergibt $n_c = 0$, wenn $n = 1, 2, 3, \dots$. Dann ist

$$E_n = -\frac{E_R}{n^2} \quad (3.12)$$

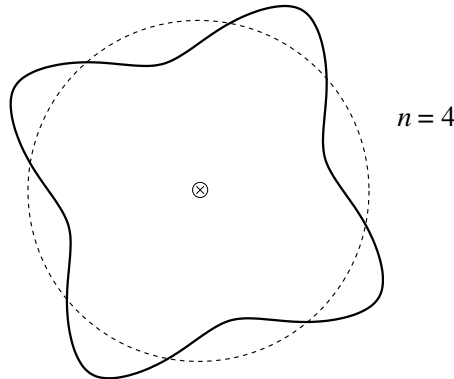
mit

$$E_R = \frac{e^4 m}{8\varepsilon_0^2 h^2} = 2\pi^2 \left(\frac{e^2}{4\pi\varepsilon_0} \right)^2 \frac{m}{h^2} = 13,6 \text{ eV}. \quad (3.13)$$

Die Differenzen dieser quantisierten Energien ergeben genau die Frequenzen der beobachteten Spektrallinien über $\nu = (E_n - E_m)/h$. Das war ein großer Erfolg des Bohrschen Atommodells.

De Broglie hat eine anschauliche Begründung für die diskreten Energiewerte formuliert. Die Idee ist, die auf Kreisbahnen umlaufenden Elektronen als Wellen aufzufassen. Eine stabile Welle sollte nur möglich sein, wenn der Bahnumfang ein ganzzahliges Vielfaches der (de Broglie-) Wellenlänge ist, denn sonst führt destruktive Interferenz zu einer mittleren Amplitude von Null. Also sollte gelten:

$$2\pi r = n\lambda = n \frac{h}{p}, \quad n = 1, 2, 3, \dots \quad (3.14)$$



Nun gilt

$$V = -\frac{e^2}{4\pi\varepsilon_0} \frac{1}{r} \quad (3.15)$$

und nach dem Virialsatz für die kinetische Energie

$$T = -\frac{V}{2} = +\frac{1}{2} \frac{e^2}{4\pi\varepsilon_0} \frac{1}{r}. \quad (3.16)$$

Damit ist der Impuls

$$p = \sqrt{2mT} = \sqrt{m \frac{e^2}{4\pi\varepsilon_0} \frac{1}{r}}. \quad (3.17)$$

Gleichung (3.14) ergibt also

$$r = n \frac{h}{p} = \frac{n\hbar}{\sqrt{m}} \sqrt{\frac{4\pi\varepsilon_0}{e^2}} \sqrt{r} \quad (3.18)$$

$$\Rightarrow \sqrt{r} = \frac{\hbar}{\sqrt{m}} \sqrt{\frac{4\pi\varepsilon_0}{e^2}} n \quad (3.19)$$

$$\Rightarrow r = \frac{\hbar^2}{m} \frac{4\pi\varepsilon_0}{e^2} n^2 \quad (3.20)$$

und für die Energie folgt

$$\begin{aligned}
E = T + V &= -\frac{1}{2} \frac{e^2}{4\pi\epsilon_0} \frac{1}{r} \\
&= -\frac{1}{2} \left(\frac{e^2}{4\pi\epsilon_0} \right)^2 \frac{m}{\hbar^2} \frac{1}{n^2} \\
&= -2\pi^2 \left(\frac{e^2}{4\pi\epsilon_0} \right)^2 \frac{m}{\hbar^2} \frac{1}{n^2}, \quad n = 1, 2, \dots
\end{aligned} \tag{3.21}$$

Wir finden also dieselben diskreten Energiewerte wie nach der Bohrschen Quantisierungsregel. Interessant ist auch Gl. (3.20): Der Bahnradius wird für große n viel größer als für den Grundzustand, in dem $r = r_1 \approx 1 \text{ \AA} = 10^{-10} \text{ m}$ gilt. Für große n sollten wir den klassischen Grenzfall erreichen, also sollte Gl. (3.20) dann eine vernünftige Abschätzung der Größe von Atomen ergeben. Solche vergleichsweise riesigen Atome mit großen n nennt man *Rydberg-Atome*. Da sie durch Stöße leicht abgeregt oder ionisiert werden, existieren sie für experimentell relevante Zeiten nur in stark verdünnten Gasen, entweder im Labor oder im interstellaren Raum.

3.2.2 Phasenraumquantisierung

Die bisherige Betrachtung schränkt zwar die große Halbachse a der Bahn und die Energie auf diskrete Werte ein, aber nicht die kleine Halbachse b . Demnach müsste es für jede Energie unendlich viele verschiedene Bahnen mit unterschiedlichen b geben. Das ist noch nicht richtig. Für die Lösung des Problems und die Behandlung anderer Potentiale benötigen wir eine systematische Theorie der Quantisierung von Phasenraumbahnen, die in diesem Abschnitt vorgestellt wird.

Wir formulieren nun die Ideen von Bohr als Quantisierung periodischer Bahnen $(\vec{q}, \vec{p})(t)$ im Phasenraum. Diese Formulierung geht v. a. auf Wilson und Sommerfeld zurück. Zunächst betrachten wir periodische Bewegungen in Systemen mit *einem* Freiheitsgrad. In diesem Fall ist der Phasenraum zweidimensional und wird durch q, p parametrisiert. Die Quantisierungsbedingung schreiben wir jetzt als

$$\int_{E_{\min}}^{E_n} \frac{dE'_n}{\nu_{\text{Bahn}}(E'_n)} = (n + n_c) h, \tag{3.22}$$

wobei die Grenze E_{\min} berücksichtigt, dass die klassische Energie für andere als das Coulomb-Potential nach unten beschränkt sein kann. Die linke Seite können wir als Integral im Phasenraum umschreiben (wir schreiben die Energie jetzt als Hamilton-Funktion H):

$$\int_{E_{\min}}^{E_n} \frac{dH}{\nu_{\text{Bahn}}(H)} = \int_{E_{\min}}^{E_n} dH T_{\text{Bahn}}(H) = \int_{E_{\min}}^{E_n} dH \int_0^{T_{\text{Bahn}}(H)} dt = \iint_{H(q,p) \leq E_n} dq dp \underbrace{\begin{vmatrix} \frac{\partial H}{\partial q} & \frac{\partial H}{\partial p} \\ \frac{\partial t}{\partial q} & \frac{\partial t}{\partial p} \end{vmatrix}}_{\text{Jacobi-Determinante}}. \tag{3.23}$$

Nach den kanonischen Gleichungen der Hamilton-Mechanik ist

$$\begin{vmatrix} \frac{\partial H}{\partial q} & \frac{\partial H}{\partial p} \\ \frac{\partial t}{\partial q} & \frac{\partial t}{\partial p} \end{vmatrix} = \begin{vmatrix} -\frac{dp}{dt} & \frac{dq}{dt} \\ \frac{\partial t}{\partial q} & \frac{\partial t}{\partial p} \end{vmatrix} = \begin{vmatrix} -\frac{dp}{dt} \frac{\partial t}{\partial p} - \frac{dq}{dt} \frac{\partial t}{\partial q} \end{vmatrix} = \begin{vmatrix} \frac{\partial t}{\partial q} \frac{dq}{dt} + \frac{\partial t}{\partial p} \frac{dp}{dt} \end{vmatrix}.$$

Wir betrachten nun $t = t(q, p)$ als Funktion von q und p . Wenn nun q und p ihrerseits von einer Variablen x abhängen, gilt nach der Kettenregel

$$\frac{dt}{dx} = \frac{\partial t}{\partial q} \frac{dq}{dx} + \frac{\partial t}{\partial p} \frac{dp}{dx}. \tag{3.24}$$

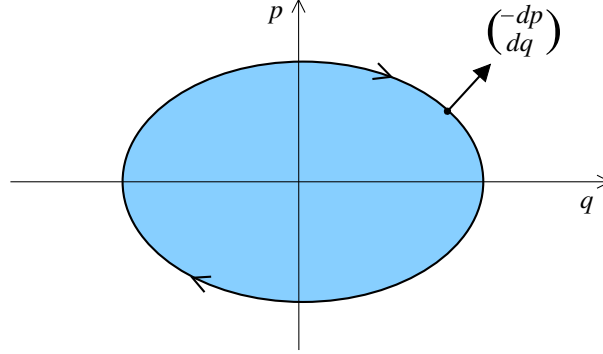
Hier setzen wir jetzt $x \equiv t$ und erhalten mit der vorigen Gleichung

$$\begin{vmatrix} \frac{\partial H}{\partial q} & \frac{\partial H}{\partial p} \\ \frac{\partial t}{\partial q} & \frac{\partial t}{\partial p} \end{vmatrix} = \left| \frac{dt}{dt} \right| = 1. \tag{3.25}$$

Damit folgt

$$\int_{E_{\min}}^{E_n} \frac{dH}{\nu_{\text{Bahn}}(H)} = \iint_{H(q,p) \leq E_n} dq dp. \quad (3.26)$$

Dies ist offenbar die Fläche im Phasenraum, die von der Bahn zur Energie $E = E_n$ eingeschlossen wird.



Diesen Ausdruck schreiben wir mit Hilfe des Gaußschen Satzes in zwei Dimensionen als Linienintegral um:

$$\begin{aligned} \iint_{H(p,q) \leq E_n} dq dp &= \frac{1}{2} \iint_{H \leq E_n} dq dp \left(\frac{\partial}{\partial q}, \frac{\partial}{\partial p} \right) \cdot \begin{pmatrix} q \\ p \end{pmatrix} \\ &\stackrel{\text{Gauß}}{=} \frac{1}{2} \oint_{H=E_n} (-dp, dq) \cdot \begin{pmatrix} q \\ p \end{pmatrix} = \frac{1}{2} \oint_{H=E_n} (p dq - q dp) \\ &= \frac{1}{2} \oint_{H=E_n} p dq - \frac{1}{2} \left[\underbrace{\oint_{H=E_n} d(qp)}_{= qp|_{(q_0, p_0)}^{(q_0, p_0)} = 0} - \oint_{H=E_n} p dq \right] = \oint_{H=E_n} p dq. \end{aligned} \quad (3.27)$$

Wir erhalten also eine kompakte Formulierung der Quantisierungsregel im Phasenraum:

$$\oint_{H(q,p)=E_n} p dq = (n + n_c) h, \quad n = 1, 2, \dots \quad (3.28)$$

Die von der Bahn eingeschlossene Fläche im Phasenraum, die auch für generalisierte Koordinaten immer die Dimension einer Wirkung (Energie mal Zeit) hat, muss nach Bohr und Sommerfeld bis auf die Korrektur $n_c h$, die die kleinst mögliche Bahn bestimmt, ein ganzzahliges Vielfaches des Planckschen Wirkungsquantums sein. Man erkennt auch sofort, dass die Regel nur für periodische Phasenraumbahnen anwendbar ist, also z. B. nicht für Streubahnen.

Als Beispiel betrachten wir den harmonischen Oszillator. Hier ist

$$H = \frac{p^2}{2m} + \frac{1}{2} m \omega_0^2 q^2. \quad (3.29)$$

Für $H = E = \text{const}$ ist das eine Ellipsengleichung mit den Halbachsen

$$q_{\max} = \sqrt{\frac{2E}{m\omega_0^2}}, \quad p_{\max} = \sqrt{2mE}. \quad (3.30)$$

Damit ist die Fläche der Ellipse

$$\oint_{H=E} p dq = \iint_{H \leq E} dq dp = \pi q_{\max} p_{\max} = \frac{2\pi E}{\omega_0} \stackrel{!}{=} (n + n_c) h. \quad (3.31)$$

Es folgt

$$E_n = \hbar\omega_0 (n + n_c). \quad (3.32)$$

Die Abstände der Energieniveaus kommen korrekt heraus, wie wir sehen werden. Die Konstante bekommen wir nicht aus der Bohr-Sommerfeld-Quantisierung. Die Schrödinger-Quantenmechanik zeigt, dass wir $n_c = -1/2$ für $n = 1, 2, 3, \dots$ wählen müssen, oder äquivalent (und üblicher) $n_c = +1/2$ für $n = 0, 1, 2, \dots$

Auf mehrere Freiheitsgrade lässt sich die Quantisierungsregel übertragen, wenn sich generalisierte Koordinaten (\vec{q}, \vec{p}) finden lassen, so dass die Bewegungsgleichungen für verschiedene Freiheitsgrade entkoppeln. Die Bewegung von (q_i, p_i) ist dann unabhängig von (q_j, p_j) für $i \neq j$. Dies ist insbesondere für sogenannte Wirkungs-Winkel-Variablen der Fall. Nicht alle mechanischen Systeme erlauben dies. Dafür ist notwendig und hinreichend, dass das System *integrabel* ist, siehe Vorlesung Theoretische Mechanik.

Ist diese Entkopplung möglich, muss zusätzlich die Phasenraumbahn in jedem Paar (q_i, p_i) periodisch sein, evtl. mit nicht identischen Umlaufzeiten bzw. Frequenzen. In diesem Fall lauten die Quantisierungsbedingungen, nach Wilson und Sommerfeld,

$$\oint p_i dq_i = (n_i + n_i^c) h, \quad n_i = 1, 2, 3, \dots \quad (3.33)$$

Wir wenden dies auf das Wasserstoffatom an, wobei wir zunächst die Korrekturen n_i^c außer Acht lassen.

Da die Kepler-Bahnen eben sind, handelt es sich um ein zweidimensionales Problem. Geeignete generalisierte Koordinaten und konjugierte Impulse sind die Polarkoordinaten r, φ und p_r sowie $p_\varphi = L$. Es ist

$$H = \underbrace{\frac{1}{2m} \left(p_r^2 + \frac{L^2}{r^2} \right)}_T - \underbrace{\frac{e^2}{4\pi\epsilon_0} \frac{1}{r}}_V. \quad (3.34)$$

Aufgrund der Erhaltung von Energie und Drehimpuls hängt, für gegebene Energie $H = E = \text{const}$ und gegebenen Drehimpuls $L = \text{const}$, p_r bis auf das Vorzeichen nur von r , aber nicht von φ ab:

$$\frac{1}{2m} \left(p_r^2 + \frac{L^2}{r^2} \right) - \frac{e^2}{4\pi\epsilon_0} \frac{1}{r} = E. \quad (3.35)$$

Andererseits hängt p_φ nur von φ ab (in trivialer Weise: p_φ ist eine konstante Funktion von φ):

$$p_\varphi = L = \text{const}. \quad (3.36)$$

Die beiden Freiheitsgrade sind also entkoppelt. Die letzte Beziehung führt auf die Quantisierung

$$\oint p_\varphi d\varphi = 2\pi L \stackrel{!}{=} n_\varphi h \quad (3.37)$$

$$\Rightarrow L = n_\varphi \hbar, \quad n_\varphi = 1, 2, 3, \dots \quad (3.38)$$

Gleichung (3.35) ist etwas komplizierter und führt nach einiger Rechnung auf

$$\oint p_r dr = 2\pi \frac{e^2}{4\pi\epsilon_0} \sqrt{-\frac{m}{2E}} - 2\pi L \stackrel{!}{=} n_r h, \quad n_r = 1, 2, \dots \quad (3.39)$$

Definieren wir die *Drehimpulsquantenzahl* $l := n_\varphi = 1, 2, \dots$ und die *Hauptquantenzahl* $n := n_\varphi + n_r = 2, 3, \dots$, so erhalten wir

$$L = \hbar l \quad \text{und} \quad E = -\frac{E_R}{n^2}. \quad (3.40)$$

Das stimmt insoweit mit dem Ergebnis der modernen Quantenmechanik überein, wie wir sehen werden, aber die Quantenzahlen durchlaufen in Wirklichkeit die Werte $n = 1, 2, 3, \dots$ und $l = 0, 1, \dots, n-1$. Wir wussten schon, dass Konstanten n_i^c nicht bestimmt werden können, also auch nicht der Startwert der ganzen Zahlen. Die Tatsache, dass die Drehimpulsquantenzahl l auch nach oben beschränkt ist, kommt aber unerwartet. Wieder kann die Bohr-Sommerfeld-Quantisierungsregel die erlaubten Quantenzahlen nicht bestimmen und macht *ad-hoc*-Annahmen notwendig.

Der Vollständigkeit halber geben wir die Halbachsen der klassischen Ellipsenbahnen an:

$$a = \frac{1}{2} \frac{e^2}{4\pi\epsilon_0} \frac{1}{|E|} = \frac{\hbar^2}{m} \frac{4\pi\epsilon_0}{e^2} n^2, \quad (3.41)$$

siehe Gl. (3.20), und

$$b = \frac{l}{n} a. \quad (3.42)$$

Beachte, dass für $l = 0$, d. h. Drehimpuls $L = 0$, die Ellipse zu einer linearen Schwingbahn durch den Kern entartet, was man wegen der Singularität von $V(\vec{r})$ bei $\vec{r} = 0$ nicht gut verstehen kann. Aber die Realität wird ohnehin nicht gut durch die klassischen Kepler-Bahnen beschrieben.

3.2.3 Das Korrespondenzprinzip

Als allgemeinste Formulierung der älteren Quantentheorie haben wir gefunden, dass sich die Wirkung

$$S = \sum_i \oint p_i dq_i \quad (3.43)$$

von mechanischen Systemen nur um ganzzahlige Vielfache eines *Wirkungsquantums* h ändern kann. Dies haben wir zumindest für periodische Bewegungen gesehen. Wir erwarten daher, dass sich das Verhalten eines mechanischen Systems deutlich vom klassischen Ergebnis unterscheidet, wenn seine Wirkung S von der Größenordnung von h ist. Andererseits sollte die Diskretheit der Wirkung keine großen Auswirkungen haben, wenn $S \gg h$ gilt. Das ist der Inhalt des *Korrespondenzprinzips*: Die Quantenmechanik enthält die klassische Mechanik als Grenzfall für $S \gg h$. Mathematisch sollten wir die Gesetze der klassischen Mechanik aus denen der Quantenmechanik erhalten, indem wir in geeigneter Weise den Grenzübergang $h \rightarrow 0$ ausführen. Dies sollte auf der Ebene von Observablen geschehen. Ein Beispiel ist die Plancksche Strahlungsformel aus 2.2:

$$w(\nu) = 8\pi \frac{\nu^3}{c^3} \frac{h}{e^{h\nu/k_B T} - 1} \stackrel{h \rightarrow 0}{\cong} 8\pi \frac{\nu^3}{c^3} \frac{h}{h\nu/k_B T} = 8\pi \frac{\nu^2}{c^3} k_B T. \quad (3.44)$$

Im klassischen Grenzfall ergibt sich, wie schon gesehen, das Rayleigh-Jeans-Gesetz.

3.3 Schwierigkeiten der älteren Quantentheorie

Die ältere Quantentheorie erlaubt, unter Hinzunahme eines empirischen Wertes für die kleinste erlaubte Bahn, eine quantitative Beschreibung der Spektren des Wasserstoffatoms und einer ganzen Reihe weiterer Systeme. Das Prinzip der Quantisierung der Phasenraumbahn ist so nützlich, dass es auch heute noch angewendet wird, um Quantenkorrekturen für (periodische) Bewegungen zu untersuchen, die sich klassisch viel leichter als quantenmechanisch berechnen lassen. Jedoch liefert die ältere Theorie keine befriedigende fundamentale Beschreibung und ist darüber hinaus unvollständig:

- Empirische Zusatzannahmen über die erlaubten Werte von Quantenzahlen sind erforderlich (Konstanten n_i^c und evtl. Verbot gewisser Werte, wie beim Drehimpuls).
- Die Beschreibung klassisch *ungebundener* Bewegungen, z. B. in Streuprozessen, ist nicht möglich.
- Die Beschreibung klassisch gebundener, aber *aperiodischer* Bewegungen ist ebenfalls nicht möglich.
- Die Beschreibung von Bewegungen durch *klassisch verbotene* Bereiche (Tunneleffekt) ist nicht möglich, da es in diesen Bereichen offensichtlich keine klassische Phasenraumbahn gibt, die quantisiert werden könnte.
- Der Spin und damit der von ihm abhängige anomale Zeeman-Effekt wird nicht beschrieben.

- Die Theorie macht inkonsistente Voraussagen. Zum Beispiel müsste beim Wasserstoffatom *jede* Drehimpulskomponente die Quantisierungsbedingung $L/\hbar = 0, 1, 2, \dots$ erfüllen, also

$$\frac{\vec{L} \cdot \hat{e}}{\hbar} = 0, 1, 2, \dots \quad (3.45)$$

für alle Einheitsvektoren \hat{e} . (Dies folgt aus der Isotropie des Raumes.) Dann muss aber $\vec{L} = 0$ gelten.

- Der wichtigste Kritikpunkt ist, dass die Theorie keine Begründung oder Rechtfertigung der *ad-hoc* Quantisierungsregel versucht. Die Theorie *erklärt* nichts. Im Gegenteil steht die Idee klassischer Bahnen im Konflikt zu un stetigen Übergängen. Die ältere Quantentheorie liefert keinen Hinweis darauf, was hier „wirklich passiert“.

Aufgrund der genannten Probleme ist die ältere Quantentheorie noch keine befriedigende Theorie von Quantenphänomenen. Die Probleme werden weitgehend durch die moderne Quantenmechanik nach Schrödinger, Heisenberg u.a. behoben.

Kapitel 4

Ursprünge der Wellenmechanik

Die moderne Quantenmechanik kann in verschiedener Weise formuliert werden. Für die Beschreibung räumlicher Bewegungen von Teilchen ist die Wellenmechanik besonders günstig. Ihre Grundgleichung ist die Schrödinger-Gleichung für die Wellenfunktion. In diesem Kapitel versuchen wir zunächst, die Form der Schrödinger-Gleichung plausibel zu machen. Im folgenden Kapitel formulieren wir die Wellenmechanik dann mathematisch sorgfältiger.

4.1 Materiewellen

Louis de Broglie war 1923 klar, dass das Licht sowohl Wellen-Charakter (Beugung) als auch Teilchen-Charakter (Photoeffekt) hat. Er schlug nun vor, dass die Konstituenten der Materie ebenfalls eine solche Welle-Teilchen-Dualität aufweisen. Diese gewagt erscheinende Idee wurde ab 1927 durch Beugungsexperimente mit Elektronenstrahlen direkt bestätigt. Sie gestattete eine sehr elegante Deutung der Quantisierung geschlossener Bahnen, z. B. im Wasserstoffatom, wie wir in Abschnitt 3.2.1 gesehen haben: Auf einer geschlossenen Bahn erhalten wir stationäre Wellen nur, wenn der Umfang ein ganzzahliges Vielfaches der Wellenlänge ist. Das ist ein allgemeines Konzept für Wellen oder allgemein für Felder und nicht beschränkt auf die Quantentheorie. Beispiele aus der klassischen Physik sind die mechanischen Eigenmoden einer Saite und die elektromagnetischen Eigenmoden eines Hohlraumresonators.

Es stellt sich nun die Frage, welche Bewegungsgleichung die Materiewellen beschreibt. Der wichtigste Anhaltspunkt bei ihrer Konstruktion ist wieder, dass sie im klassischen Grenzfall mit der klassischen Mechanik verträglich sein muss (Korrespondenzprinzip).

4.2 Induktive Begründung der Schrödinger-Gleichung: Wellenpakete

Wenn wir Materie durch Wellen beschreiben wollen, müssen wir uns klarmachen, wie wir das erfolgreiche Teilchenbild wieder gewinnen. Eine naheliegende Idee ist, Teilchen als *Wellenpakete* zu beschreiben, die im Ort und im Impuls möglichst schmal sind. Dabei können wir ein Wellenpaket nicht im Ort und im Impuls beliebig scharf lokalisieren: Sei $\psi(\vec{k}, t)$ ein Wellenpaket im Ortsraum, dann beschreibt die Fourier-Transformierte

$$\tilde{\psi}(\vec{k}, t) = \int d^3r e^{-i\vec{k}\cdot\vec{r}} \psi(\vec{r}, t) \quad (4.1)$$

die zugehörige Verteilung im Impulsraum (\vec{k} -Raum). Ist nun $\psi(\vec{r}, t)$ schmal, dann ist $\tilde{\psi}(\vec{k}, t)$ notwendigerweise breit, und ist $\tilde{\psi}$ schmal, dann ist ψ breit. Wir skizzieren die Begründung für den Fall, dass der mittlere Ort \vec{r} und der mittlere Wellenvektor \vec{k} verschwinden. Im Vorgriff auf die Schrödingersche Wellenmechanik nehmen wir als Gewichtsfunktionen $|\psi|^2$ und $|\tilde{\psi}|^2$ anstelle von ψ und $\tilde{\psi}$. Die mittleren quadratischen Abweichungen von Ort und Wellenvektor sind dann

$$\Delta r^2 = \frac{\int d^3r r^2 |\psi(\vec{r}, t)|^2}{\int d^3r |\psi(\vec{r}, t)|^2}, \quad (4.2)$$

$$\Delta k^2 = \frac{\int \frac{d^3 k}{(2\pi)^3} k^2 |\tilde{\psi}(\vec{k}, t)|^2}{\int \frac{d^3 k}{(2\pi)^3} |\tilde{\psi}(\vec{k}, t)|^2}. \quad (4.3)$$

Wir nehmen außerdem an, dass ψ normiert ist:

$$\int d^3 r |\psi(\vec{r}, t)|^2 = 1. \quad (4.4)$$

Dann gilt auch

$$\begin{aligned} \int \frac{d^3 k}{(2\pi)^3} |\tilde{\psi}(\vec{k}, t)|^2 &= \int \frac{d^3 k}{(2\pi)^3} \int d^3 r e^{i\vec{k} \cdot \vec{r}} \psi^*(\vec{r}, t) \int d^3 r' e^{-i\vec{k} \cdot \vec{r}'} \psi(\vec{r}', t) \\ &= \int d^3 r d^3 r' \delta(\vec{r} - \vec{r}') \psi^*(\vec{r}, t) \psi(\vec{r}', t) \\ &= \int d^3 r |\psi(\vec{r}, t)|^2 = 1. \end{aligned} \quad (4.5)$$

Damit erhalten wir

$$\Delta r^2 = \int d^3 r r^2 |\psi(\vec{r}, t)|^2, \quad (4.6)$$

$$\Delta k^2 = \int \frac{d^3 k}{(2\pi)^3} k^2 |\tilde{\psi}(\vec{k}, t)|^2. \quad (4.7)$$

Wir betrachten die Hilfsgröße

$$I(\lambda) := \int d^3 r |\vec{r} \psi(\vec{r}, t) + \lambda \vec{\nabla} \psi(\vec{r}, t)|^2, \quad (4.8)$$

die offensichtlich für alle λ nicht negativ ist. Wir formen das Integral um:

$$\begin{aligned} I(\lambda) &= \int d^3 r \left[\psi^* r^2 \psi + \psi^* \vec{r} \cdot \lambda \vec{\nabla} \psi + \lambda (\vec{\nabla} \psi^*) \cdot \vec{r} \psi + \lambda (\vec{\nabla} \psi^*) \cdot \lambda \vec{\nabla} \psi \right] \\ &\stackrel{\text{partiell}}{=} \int d^3 r \left[\psi^* r^2 \psi + \lambda \psi^* \vec{r} \cdot \vec{\nabla} \psi - \lambda \psi^* \vec{\nabla} \cdot (\vec{r} \psi) - \lambda^2 \psi^* \nabla^2 \psi \right] \\ &= \Delta r^2 - \lambda \int d^3 r \underbrace{\psi^* (\vec{\nabla} \cdot \vec{r}) \psi}_{=3} - \lambda^2 \int d^3 r \psi^* \nabla^2 \psi \\ &= \Delta r^2 - 3\lambda - \lambda^2 \int d^3 r \psi^* \nabla^2 \psi. \end{aligned} \quad (4.9)$$

Im letzten Term ist

$$\begin{aligned} \int d^3 r \psi^* \nabla^2 \psi &= \int d^3 r \int \frac{d^3 k d^3 k'}{(2\pi)^6} e^{-i\vec{k} \cdot \vec{r}} \tilde{\psi}^*(\vec{k}, t) \nabla^2 e^{i\vec{k}' \cdot \vec{r}} \tilde{\psi}(\vec{k}', t) \\ &= \int d^3 r \int \frac{d^3 k d^3 k'}{(2\pi)^6} e^{-i\vec{k} \cdot \vec{r}} \tilde{\psi}^*(\vec{k}, t) (-k'^2) e^{i\vec{k}' \cdot \vec{r}} \tilde{\psi}(\vec{k}', t) \\ &= \int \frac{d^3 k d^3 k'}{(2\pi)^6} (2\pi)^3 \delta(\vec{k}' - \vec{k}) \tilde{\psi}^*(\vec{k}, t) (-k'^2) \tilde{\psi}(\vec{k}', t) \\ &= - \int \frac{d^3 k}{(2\pi)^3} k^2 |\tilde{\psi}(\vec{k}, t)|^2 = -\Delta k^2. \end{aligned} \quad (4.10)$$

Es folgt

$$I(\lambda) = \Delta r^2 - 3\lambda + \lambda^2 \Delta k^2 \geq 0 \quad (4.11)$$

$$\Rightarrow \Delta r^2 \geq -\lambda^2 \Delta k^2 + 3\lambda \quad \forall \lambda. \quad (4.12)$$

Da dies für alle λ gilt, gilt es auch für das λ , das die rechte Seite maximiert. Diesen Wert von λ erhalten wir aus

$$\frac{d}{d\lambda} (-\lambda^2 \Delta k^2 + 3\lambda) = -2\lambda \Delta k^2 + 3 = 0 \quad (4.13)$$

$$\Rightarrow \lambda = \frac{3}{2 \Delta k^2}. \quad (4.14)$$

Wir folgern

$$\Delta r^2 \geq -\frac{9}{4} \frac{1}{\Delta k^2} + \frac{9}{2} \frac{1}{\Delta k^2} = \frac{9}{4} \frac{1}{\Delta k^2} \quad (4.15)$$

$$\Rightarrow \Delta r^2 \Delta k^2 \geq \frac{9}{4} \quad (4.16)$$

$$\Rightarrow \Delta r \Delta k \geq \frac{3}{2}. \quad (4.17)$$

Man kann sich überzeugen, dass der Faktor 3 im Zähler die drei Raumdimensionen widerspiegelt.

Wenn wir Teilchen durch Wellenpakete repräsentieren, liegt es nahe, den Ort des Teilchens mit dem Schwerpunkt oder mittleren Ort

$$\bar{\vec{r}} = \int d^3 r \vec{r} |\psi(\vec{r}, t)|^2 \quad (4.18)$$

des Wellenpakets zu identifizieren. Unsere Aufgabe besteht nun darin, eine Gleichung für die Dynamik der Welle $\psi(\vec{r}, t)$ zu finden, so dass der mittlere Ort $\bar{\vec{r}}$ im klassischen Grenzfall der klassischen Mechanik gehorcht.

Die Wellenlänge einer ebenen Materiewelle ist nach de Broglies später bestätigter Vorhersage, siehe Abschnitt 2.5, $\lambda = h/p$. Daraus folgt für die Wellenzahl

$$k = \frac{2\pi}{\lambda} = \frac{2\pi}{h} p \equiv \frac{p}{h}, \quad (4.19)$$

also

$$p = \hbar k. \quad (4.20)$$

Der natürliche Ansatz für die entsprechenden Vektoren ist dann

$$\vec{p} = \hbar \vec{k}. \quad (4.21)$$

Der mittlere Impuls eines Wellenpaketes ist daher

$$\begin{aligned} \bar{\vec{p}} &= \hbar \bar{\vec{k}} = \hbar \int \frac{d^3 k}{(2\pi)^3} \vec{k} |\tilde{\psi}(\vec{k}, t)|^2 = \hbar \int \frac{d^3 k}{(2\pi)^3} \tilde{\psi}^* \vec{k} \tilde{\psi} \\ &= \hbar \int \frac{d^3 k}{(2\pi)^3} \int d^3 r d^3 r' e^{i\vec{k} \cdot \vec{r}} \psi^*(\vec{r}, t) \vec{k} e^{-i\vec{k} \cdot \vec{r}'} \psi(\vec{r}', t) \\ &= \hbar \int \frac{d^3 k}{(2\pi)^3} \int d^3 r d^3 r' e^{i\vec{k} \cdot \vec{r}} \psi^*(\vec{r}, t) \left(i \vec{\nabla}' e^{-i\vec{k} \cdot \vec{r}'} \right) \psi(\vec{r}', t) \quad \left| \vec{\nabla}' \text{ wirkt nur auf } e^{-i\vec{k} \cdot \vec{r}'} \right. \\ &\stackrel{\text{partiell}}{=} \hbar \int \frac{d^3 k}{(2\pi)^3} \int d^3 r d^3 r' e^{i\vec{k} \cdot \vec{r}} \psi^*(\vec{r}, t) e^{-i\vec{k} \cdot \vec{r}'} (-i \vec{\nabla}') \psi(\vec{r}', t) \\ &= -i\hbar \int d^3 r d^3 r' \delta(\vec{r} - \vec{r}') \psi^*(\vec{r}, t) \vec{\nabla} \psi(\vec{r}', t) \\ &= -i\hbar \int d^3 r \psi^* \vec{\nabla} \psi \\ &= \int d^3 r \psi^* \frac{\hbar}{i} \vec{\nabla} \psi. \end{aligned} \quad (4.22)$$

Wir finden damit eine heuristische Regel: Der Impuls ist in der Ortsdarstellung der „Mittelwert“ des Differentialoperators $\frac{\hbar}{i} \vec{\nabla}$. Analoge Ergebnisse erhalten wir für gemittelte Funktionen des Impulses, z. B.

$$\overline{p^2} = \hbar^2 \overline{k^2} = \hbar^2 \int \frac{d^3 k}{(2\pi)^3} \tilde{\psi}^* k^2 \tilde{\psi} = -\hbar^2 \int d^3 r \psi^* \nabla^2 \psi = \int d^3 r \psi^* \left(\frac{\hbar}{i} \vec{\nabla} \right)^2 \psi. \quad (4.23)$$

Wir vermuten, dass wir überall in Integralen über den Ortsraum den Impuls durch $\frac{\hbar}{i} \vec{\nabla}$ ersetzen sollen. Dann ist z. B. die gemittelte Hamilton-Funktion

$$\overline{H(\vec{r}, \vec{p})} = \int d^3r \psi^* H\left(\vec{r}, \frac{\hbar}{i} \vec{\nabla}\right) \psi. \quad (4.24)$$

Unsere Aufgabe ist also, eine Bewegungsgleichung für $\psi(\vec{r}, t)$ so zu konstruieren, dass der mittlere Ort \vec{r} und der mittlere Impuls \vec{p} die klassischen Bewegungsgleichungen eines Teilchens erfüllen. Wir betrachten zunächst den dreidimensionalen harmonischen Oszillator. Die klassische Hamilton-Funktion lautet

$$H(\vec{r}, \vec{p}) = \frac{p^2}{2m} + \frac{1}{2} \kappa r^2. \quad (4.25)$$

Die kanonischen (Hamiltonschen) Bewegungsgleichungen sind

$$\dot{\vec{r}} = \frac{\partial H}{\partial \vec{p}} = \frac{\vec{p}}{m}, \quad (4.26)$$

$$\dot{\vec{p}} = -\frac{\partial H}{\partial \vec{r}} = -\kappa \vec{r}. \quad (4.27)$$

Es soll also gelten

$$\frac{d}{dt} \vec{r} = \int d^3r \left(\psi^* \vec{r} \dot{\psi} + \dot{\psi}^* \vec{r} \psi \right) = \frac{1}{m} \int d^3r \psi^* \frac{\hbar}{i} \vec{\nabla} \psi = \frac{\vec{p}}{m} \quad (4.28)$$

und

$$\frac{d}{dt} \vec{p} = \int d^3r \left(\dot{\psi}^* \frac{\hbar}{i} \vec{\nabla} \psi + \psi^* \frac{\hbar}{i} \vec{\nabla} \dot{\psi} \right) = -\kappa \int d^3r \psi^* \vec{r} \psi = -\kappa \vec{r}. \quad (4.29)$$

Wir schreiben beide Gleichungen um:

$$\int d^3r \left(\dot{\psi}^* \vec{r} \psi + \psi^* \vec{r} \dot{\psi} - \psi^* \frac{1}{m} \frac{\hbar}{i} \vec{\nabla} \psi \right) = 0, \quad (4.30)$$

$$\int d^3r \left(\dot{\psi}^* \frac{\hbar}{i} \vec{\nabla} \psi + \psi^* \frac{\hbar}{i} \vec{\nabla} \dot{\psi} + \psi^* \kappa \vec{r} \psi \right) = 0. \quad (4.31)$$

Wir suchen wie gesagt eine Bewegungsgleichung für ψ . Es ist sicherlich sinnvoll, eine möglichst einfache Gleichung zu konstruieren – wir können sie immer noch verwerfen, wenn sie sich nicht bewährt. Daher versuchen wir zunächst, eine *lineare* Gleichung *erster* Ordnung in der Zeit zu finden, also

$$\dot{\psi} = G\psi, \quad (4.32)$$

wobei G ein noch unbekannter Differentialoperator ist; G kann Funktionen von \vec{r} und t sowie Ableitungen nach \vec{r} (aber nicht t) enthalten. Einsetzen ergibt

$$\int d^3r \left[(G\psi)^* \vec{r} \psi + \psi^* \vec{r} G\psi - \psi^* \frac{1}{m} \frac{\hbar}{i} \vec{\nabla} \psi \right] = 0, \quad (4.33)$$

$$\int d^3r \left[(G\psi)^* \frac{\hbar}{i} \vec{\nabla} \psi + \psi^* \frac{\hbar}{i} \vec{\nabla} G\psi + \psi^* \kappa \vec{r} \psi \right] = 0. \quad (4.34)$$

Es gibt keinen systematischen Lösungsweg, man kommt also nicht ohne intelligentes Herumprobieren weiter. Wir machen den Ansatz, dass G eine Summe von zwei Termen ist, von denen einer nur vom Nabla-Operator $\vec{\nabla}$ und der andere nur von \vec{r} abhängt:

$$G = G_1(\vec{\nabla}) + G_2(\vec{r}). \quad (4.35)$$

Dann erhalten wir, mit partieller Integration für die G_1 -Terme,

$$\int d^3r \psi^* \left[G_1^*(-\vec{\nabla}) \vec{r} + G_2^*(\vec{r}) \vec{r} + \vec{r} G_1(\vec{\nabla}) + \vec{r} G_2(\vec{r}) - \frac{1}{m} \frac{\hbar}{i} \vec{\nabla} \right] \psi = 0, \quad (4.36)$$

$$\int d^3r \psi^* \left[G_1^*(-\vec{\nabla}) \frac{\hbar}{i} \vec{\nabla} + G_2^*(\vec{r}) \frac{\hbar}{i} \vec{\nabla} + \frac{\hbar}{i} \vec{\nabla} G_1(\vec{\nabla}) + \frac{\hbar}{i} \vec{\nabla} G_2(\vec{r}) + \kappa \vec{r} \right] \psi = 0. \quad (4.37)$$

Dies legt nahe, zu setzen

$$G_1^*(-\vec{\nabla}) = -G_1(\vec{\nabla}), \quad (4.38)$$

$$G_2^*(\vec{r}) = -G_2(\vec{r}), \quad (4.39)$$

denn dann folgt

$$\begin{aligned} & \int d^3r \psi^* \left[-G_1(\vec{\nabla}) \vec{r} - \cancel{G_2(\vec{r}) \vec{r}} + \vec{r} G_1(\vec{\nabla}) + \cancel{\vec{r} G_2(\vec{r})} - \frac{1}{m} \frac{\hbar}{i} \vec{\nabla} \right] \psi \\ &= \int d^3r \psi^* \left(-G_1 \vec{r} + \vec{r} G_1 - \frac{1}{m} \frac{\hbar}{i} \vec{\nabla} \right) \psi = 0, \end{aligned} \quad (4.40)$$

$$\begin{aligned} & \int d^3r \psi^* \left[\cancel{-G_1(\vec{\nabla}) \frac{\hbar}{i} \vec{\nabla}} - G_2(\vec{r}) \frac{\hbar}{i} \vec{\nabla} + \cancel{\frac{\hbar}{i} \vec{\nabla} G_1(\vec{\nabla})} + \frac{\hbar}{i} \vec{\nabla} G_2(\vec{r}) + \kappa \vec{r} \right] \psi \\ &= \int d^3r \psi^* \left(-G_2 \frac{\hbar}{i} \vec{\nabla} + \frac{\hbar}{i} \vec{\nabla} G_2 + \kappa \vec{r} \right) \psi = 0. \end{aligned} \quad (4.41)$$

Die übrig bleibenden Terme mit G_1 und G_2 verschwinden nicht, da $\vec{\nabla}$ und \vec{r} nicht aneinander vorbei gezogen werden können. Die weitere Strategie erläutern wir kurz anhand von G_2 :

1. Eine Konstante löst Gl. (4.41) offensichtlich nicht:

$$G_2 = ic, \quad c \in \mathbb{R} \quad \Rightarrow \quad \int d^3r \psi^* \kappa \vec{r} \psi = 0, \quad \text{falsch!} \quad (4.42)$$

2. Eine lineare Funktion hilft auch nicht:

$$G_2 = i\vec{c} \cdot \vec{r}, \quad \vec{c} \in \mathbb{R}^3 \quad (4.43)$$

$$\begin{aligned} & \Rightarrow \int d^3r \psi^* \left[-i(\vec{c} \cdot \vec{r}) \frac{\hbar}{i} \vec{\nabla} + \frac{\hbar}{i} \vec{\nabla} i(\vec{c} \cdot \vec{r}) + \kappa \vec{r} \right] \psi \\ &= \int d^3r \psi^* \left(\cancel{-\hbar(\vec{c} \cdot \vec{r}) \vec{\nabla}} + \hbar \vec{c} + \cancel{\hbar(\vec{c} \cdot \vec{r}) \vec{\nabla}} + \kappa \vec{r} \right) \psi = 0, \quad \text{falsch!} \end{aligned} \quad (4.44)$$

3. Eine quadratische Funktion

$$G_2 = icr^2, \quad c \in \mathbb{R} \quad (4.45)$$

löst das Problem:

$$\begin{aligned} \int d^3r \psi^* \left(-icr^2 \frac{\hbar}{i} \vec{\nabla} + \frac{\hbar}{i} \vec{\nabla} icr^2 + \kappa \vec{r} \right) \psi &= \int d^3r \psi^* \left(\cancel{-\hbar cr^2 \vec{\nabla}} + \hbar c \underbrace{2 \vec{r}}_{=\vec{r}} + \cancel{\hbar cr^2 \vec{\nabla}} + \kappa \vec{r} \right) \psi \\ &= \int d^3r \psi^* (2\hbar c + \kappa) \vec{r} \psi = 0. \end{aligned} \quad (4.46)$$

Dies erfordert $c = -\kappa/2\hbar$, also

$$G_2 = -\frac{i}{\hbar} \frac{\kappa}{2} r^2. \quad (4.47)$$

Analog findet man

$$G_1 = i\hbar \frac{1}{2m} \nabla^2 \quad (4.48)$$

und somit die Bewegungsgleichung

$$\dot{\psi} = \frac{i\hbar}{2m} \nabla^2 \psi - \frac{i}{\hbar} \frac{\kappa}{2} r^2 \psi. \quad (4.49)$$

Etwas umgeschrieben lautet sie

$$i\hbar \dot{\psi} = \left(-\frac{\hbar^2}{2m} \nabla^2 + \frac{\kappa}{2} r^2 \right) \psi. \quad (4.50)$$

Das ist die *Schrödinger-Gleichung* für den (dreidimensionalen) harmonischen Oszillator. Dieses Beispiel enthält offenbar auch das freie Teilchen als Spezialfall für $\kappa = 0$. Wir haben tatsächlich mehr geschafft, als zu erwarten war: \vec{r} und \vec{p} erfüllen für den harmonischen Oszillator *immer* die klassischen Bewegungsgleichungen. Unser Ziel war dagegen nur, dass alles im klassischen Grenzfall gelten sollte. Versucht man denselben Lösungsweg für irgendein anderes Potential, so findet man, wie wir noch sehen werden, keine Bewegungsgleichung für ψ , die immer die klassische Bewegung von \vec{r} und \vec{p} ergibt. Es ist sinnvoll, per Analogie für ein beliebiges Potential $V(\vec{r})$ die Schrödinger-Gleichung

$$i\hbar \dot{\psi} = \left(-\frac{\hbar^2}{2m} \nabla^2 + V(\vec{r}) \right) \psi \quad (4.51)$$

zu *postulieren*. Das ist sicherlich richtig, wenn die Bewegung in der Nähe eines normalen Minimums von $V(\vec{r})$ erfolgt, so dass man $V(\vec{r})$ bis zur zweiten Ordnung um dieses Minimum entwickeln kann. Gilt es aber auch im klassischen Grenzfall? Darunter verstehen wir hier den Fall eines schmalen Wellenpakets im Ort und im Wellenvektor. Wir können das Wellenpaket als schmal im Ort ansehen, wenn sich $V(\vec{r})$ (die einzige andere ortsabhängige Funktion im Problem) auf der Längenskala der Breite Δr des Pakets nur schwach ändert. Aber dann dürfen wir lokal $V(\vec{r})$ wieder bis zur zweiten Ordnung entwickeln und Gl. (4.51) ist gerechtfertigt. (Für die kinetische Energie finden wir analog, dass sie im klassischen Grenzfall annähernd quadratisch in \vec{p} sein soll, aber das hatten wir ohnehin angenommen.)

Wir haben die Schrödinger-Gleichung nicht hergeleitet, sondern bestenfalls plausibel gemacht. Eine Herleitung aus der klassischen Mechanik kann auch gar nicht möglich sein, weil die Quantenmechanik die allgemeinere Theorie ist, die die klassische Mechanik als Grenzfall einschließt.

4.3 Induktive Begründung der Schrödinger-Gleichung: Hamilton-Jacobi-Theorie

In diesem Abschnitt besprechen wir eine alternative induktive Begründung der Wellenmechanik. Dazu werden wir zunächst einige Konzepte aus der klassischen Elektrodynamik und Mechanik wiederholen.

4.3.1 Die Eikonal-Gleichung

Die Wellengleichung lautet bekanntlich

$$\nabla^2 F - \frac{1}{c^2} \frac{\partial^2 F}{\partial t^2} = 0, \quad (4.52)$$

wobei in der Elektrodynamik F z. B. eine Komponente des \mathbf{E} - oder \mathbf{B} -Feldes ist. Für konstante (Licht-) Geschwindigkeit c wird die Gleichung durch ebene Wellen

$$F(\vec{r}, t) = F_0 e^{i(\vec{k} \cdot \vec{r} - \omega t)} = F_0 e^{i(\vec{k} \cdot \vec{r} - kct)} \quad (4.53)$$

gelöst. Wenn die Geschwindigkeit nicht räumlich konstant ist, lautet die Wellengleichung

$$\nabla^2 F - \frac{n^2(\vec{r})}{c^2} \frac{\partial^2 F}{\partial t^2} = 0, \quad (4.54)$$

$n(\vec{r})$ ist der Brechungsindex. Uns interessiert hier der Fall, dass sich $n(\vec{r})$ räumlich langsam ändert, wobei „langsam“ nur im Vergleich zur Wellenlänge λ gemeint sein kann: $|\vec{\nabla} n|/n \ll 1/\lambda = k/2\pi$. Wir machen den Ansatz

$$F(\vec{r}, t) = F_0 \exp \left(i\omega \left[\frac{L(\vec{r})}{c} - t \right] \right). \quad (4.55)$$

Hier nennt man $L(\vec{r})$ den *Lichtweg* oder das *Eikonal*. Einsetzen in die Wellengleichung ergibt

$$\vec{\nabla} \cdot i \frac{\omega}{c} (\vec{\nabla} L) \exp \left(i\omega \left[\frac{L}{c} - t \right] \right) - \frac{n^2}{c^2} \frac{\partial}{\partial t} (-i\omega) \exp \left(i\omega \left[\frac{L}{c} - t \right] \right) = 0 \quad (4.56)$$

$$\Rightarrow i \frac{\omega}{c} (\nabla^2 L) \exp(\dots) - \frac{\omega^2}{c^2} (\vec{\nabla} L)^2 \exp(\dots) + \frac{n^2 \omega^2}{c^2} \exp(\dots) = 0 \quad (4.57)$$

$$\Rightarrow (\vec{\nabla} L)^2 = n^2 + \frac{ic}{\omega} \nabla^2 L. \quad (4.58)$$

Wir sehen, dass L i. A. komplex sein muss, was nicht mit der Interpretation als Lichtweg vereinbar ist. Das Problem resultiert daraus, dass wir die Amplitude F_0 als konstant angesetzt haben.

Wir haben aber noch nicht ausgenutzt, dass sich n räumlich nur langsam ändert. Für konstantes n wird Gl. (4.58) durch lineare Funktionen

$$L = L_0 + n \hat{e} \cdot \vec{r} \quad (4.59)$$

mit beliebigen, festen Einheitsvektoren \hat{e} gelöst. Als nicht lineare partielle Differentialgleichung wird Gl. (4.58) i. A. noch weitere Lösungen haben. Die linearen Funktionen sind aber für $n = \text{const}$ die physikalisch korrekten Lösungen, sie ergeben nämlich

$$F(\vec{r}, t) = F_0 \exp \left(i\omega \left[\frac{L_0 + n \hat{e} \cdot \vec{r}}{c} - t \right] \right) = F_0 \underbrace{e^{i\omega L_0/c}}_{=\text{const}} e^{i(\vec{k} \cdot \vec{r} - \omega t)} \quad (4.60)$$

mit $\vec{k} = (\omega n/c) \hat{e}$. Für diese Lösungen gilt

$$\nabla^2 L = \vec{\nabla} \cdot n \hat{e} = 0. \quad (4.61)$$

Für konstantes n verschwindet also der letzte Term in Gl. (4.58). Ist n schwach räumlich veränderlich, so sollten $\nabla^2 L$ und damit der letzte Term immerhin klein sein. Dann können wir diesen Term gegenüber n^2 vernachlässigen und erhalten die *Eikonal-Gleichung*

$$(\vec{\nabla} L)^2 = n^2. \quad (4.62)$$

Wie gesagt gilt sie im Grenzfall räumlich langsam veränderlicher Geschwindigkeit. Dies ist gerade der Gültigkeitsbereich der *geometrischen Optik* oder *Strahlenoptik* (sie lässt sich auch für *sprunghafte* Änderungen anwenden, was darauf hinaus läuft, Lösungen in Bereichen mit langsamen Änderungen an den Sprüngen mit Hilfe von Stetigkeitsbedingungen aneinander anzuschließen).

Gleichungen der Form $L(\vec{r}) = \text{const}$ beschreiben Flächen im Raum. Die Bedingung konstanter Phase, $\omega(L/c - t) = \text{const}$, beschreibt dagegen die Wellenfronten der elektromagnetischen Welle. Offenbar durchlaufen die Wellenfronten im Laufe der Zeit die Schar der Flächen $L(\vec{r}) = \text{const}$. Die *Strahlen* der geometrischen Optik stehen überall senkrecht auf den Flächen $L(\vec{r}) = \text{const}$ und damit auch auf den Wellenfronten.

4.3.2 Erinnerung an die Hamilton-Jacobi-Theorie

Die Grundidee der Hamilton-Jacobi-Theorie der klassischen Mechanik ist, eine Transformation von den gegebenen kanonischen Variablen (\vec{q}, \vec{p}) auf neue Variable (\vec{Q}, \vec{P}) zu finden, die die transformierte Hamilton-Funktion trivial macht. Man wählt dazu eine Erzeugende vom Typ $F_2(\vec{q}, \vec{P}, t)$ (siehe Vorlesung Theoretische Mechanik), die in diesem Zusammenhang mit dem Symbol S bezeichnet wird, also

$$p_i = \frac{\partial S}{\partial q_i}, \quad (4.63)$$

$$Q_i = \frac{\partial S}{\partial P_i}, \quad (4.64)$$

$$\mathcal{H} = H + \frac{\partial S}{\partial t}. \quad (4.65)$$

S soll so gewählt werden, dass alle neuen Koordinaten Q_i zyklisch sind, also in der transformierten Hamilton-Funktion \mathcal{H} gar nicht vorkommen. Dann sind wegen

$$\dot{P}_i = -\frac{\partial \mathcal{H}}{\partial Q_i} \quad (4.66)$$

alle neuen Impulse P_i erhalten. Da die Q_i in \mathcal{H} nicht vorkommen und die P_i in \mathcal{H} zwar i. A. vorkommen, aber erhalten sind, ist die einzige mögliche Zeitabhängigkeit von \mathcal{H} explizit. Gleichung (4.65) zeigt, dass wir zu irgendeinem S , das die Q_i zyklisch macht, eine nur von t abhängige Funktion addieren können, so dass \mathcal{H} nicht mehr von t abhängt und sogar identisch verschwindet: $\mathcal{H} \equiv 0$.

Die Bestimmungsgleichung für die Erzeugende mit den gewünschten Eigenschaften, also die *Hamilton-Jacobi-Gleichung*, erhalten wir aus $\mathcal{H} \equiv 0$ sowie Glg. (4.63) und (4.65):

$$H\left(q_1, \dots, q_s, \frac{\partial S}{\partial q_1}, \dots, \frac{\partial S}{\partial q_s}, t\right) + \frac{\partial S}{\partial t} = 0, \quad (4.67)$$

wobei s die Zahl der Freiheitsgrade ist. Es ist eine partielle Differentialgleichung erster Ordnung für S als Funktion von \vec{q} und t , die aber i. A. nicht linear in den $\partial S / \partial q_i$ ist. Die Lösungen für S nennt man *Wirkungswellen*, aus Gründen, die später klar werden. Da wir Materie als Feld beschreiben wollen (Materiewellen!), haben wir einen wichtigen ersten Schritt getan: Die Hamilton-Jacobi-Theorie ist eine *Feldtheorie* der klassischen Mechanik, nämlich für das Feld S .

Wie in der Vorlesung Theoretische Mechanik diskutiert, hat die Hamilton-Jacobi-Theorie ein Problem, das oft unter den Teppich gekehrt wird: Nach der Transformation haben wir s erhaltene Impulse P_i . Da dies kanonische Impulse sind, erfüllen sie die Relation

$$\{P_i, P_j\} = 0 \quad (4.68)$$

mit der Poisson-Klammer

$$\{A, B\} := \sum_i \left(\frac{\partial A}{\partial q_i} \frac{\partial B}{\partial p_i} - \frac{\partial A}{\partial p_i} \frac{\partial B}{\partial q_i} \right). \quad (4.69)$$

Man sagt, die P_i stehen in *Involution*. Also finden wir s Erhaltungsgrößen in Involution. Aber die Zahl der Erhaltungsgrößen in Involution ist unabhängig von der Wahl der kanonischen Variablen und ändert sich also insbesondere nicht bei der Transformation mit S . Weil nun aber nicht jedes mechanische System überhaupt s Erhaltungsgrößen in Involution besitzt, kann es nicht allgemein möglich sein, eine Erzeugende S zu finden, so dass $\mathcal{H} \equiv 0$ gilt. Systeme, die s Erhaltungsgrößen in Involution besitzen, nennt man *integrabel*.

Andererseits sieht die Hamilton-Jacobi-Gleichung ziemlich harmlos aus, unabhängig von der Integrabilität (die man der Gleichung nicht leicht ansehen kann). Wir können keinen Grund erkennen, warum sie nicht zumindest für fast alle (\vec{q}, \vec{P}, t) lösbar sein sollte. Die Auflösung des scheinbaren Paradoxons ist, dass man typischerweise für nicht integrable Systeme keine Lösungen finden kann, die zugleich überall eindeutig und stetig sind. Wir nehmen jetzt an, dass die Lösung zumindest im interessierenden Teilvolumen eindeutig und stetig ist.

Wir betrachten jetzt speziell den Fall einer nicht explizit zeitabhängigen Hamilton-Funktion H . In diesem Fall können wir die Zeitabhängigkeit von S abseparieren:

$$S(\vec{q}, \vec{P}, t) = W(\vec{q}, \vec{P}) - Et. \quad (4.70)$$

Einsetzen in die Hamilton-Jacobi-Gleichung ergibt

$$H\left(q_1, \dots, q_s, \frac{\partial W}{\partial q_1}, \dots, \frac{\partial W}{\partial q_s}\right) = E. \quad (4.71)$$

Da die P_i konstant sind, definieren Gleichungen der Form $W = \text{const}$ Flächen im Konfigurationsraum, also im Raum der \vec{q} . Flächen mit $S = \text{const}$, d. h. Wellenfronten der Wirkungswellen S , schieben sich mit der Zeit über die stationären Flächen $W = \text{const}$ hinweg. Die Geschwindigkeit der Wellenfronten erhalten wir aus

$$S = \text{const} \quad \Rightarrow \quad 0 = dS = \frac{\partial W}{\partial \vec{q}} \cdot d\vec{q} - E dt, \quad (4.72)$$

so dass gilt

$$\frac{\partial W}{\partial \vec{q}} \cdot \vec{u} = E \quad (4.73)$$

mit $\vec{u}(\vec{r}) := d\vec{q}/dt$. Gleichung (4.73) bestimmt \vec{u} nicht eindeutig – es ist nur eine skalare Gleichung, aber \vec{u} hat s Komponenten. Wir *definieren* \vec{u} orthogonal zu den Flächen $W = \text{const}$ und damit zu fester Zeit auch orthogonal

zu den Wellenfronten $S = \text{const}$, da Bewegungen parallel dazu keine realen Änderungen von S beschreiben. Dann ist die Wellengeschwindigkeit \vec{u} parallel zum Gradienten $\partial W / \partial \vec{q}$. Wegen der Parallelität folgt

$$\left(\frac{\partial W}{\partial \vec{q}} \right)^2 u^2 = E^2 \quad (4.74)$$

$$\Rightarrow \left(\frac{\partial W}{\partial \vec{q}} \right)^2 = \frac{E^2}{u^2}. \quad (4.75)$$

Dies ist nur dann eine nützliche Gleichung zur Bestimmung von W , wenn wir $E^2/u^2(\vec{q})$ explizit angeben können. Das ist nicht immer der Fall. Wir betrachten den Spezialfall

$$H = \frac{\vec{p} \cdot \vec{p}}{2m} + V(\vec{q}), \quad (4.76)$$

wobei H auch nicht explizit von der Zeit abhängen soll. Diese Hamilton-Funktion beschreibt beliebig viele Teilchen gleicher Masse und nicht relativistischer kinetischer Energie mit beliebigen Ein- und Mehrteilchenpotentialen. Für diese Hamilton-Funktion folgt

$$p^2 = 2m(H - V) = 2m(E - V). \quad (4.77)$$

Nun ist der Gradient $\partial W / \partial \vec{q}$ in Gl. (4.75) gerade der Vektor aller Impulse, \vec{p} . Also können wir diese Gleichung umschreiben als

$$\left(\frac{\partial W}{\partial \vec{q}} \right)^2 = 2m[E - V(\vec{q})]. \quad (4.78)$$

Diese Gleichung und Gl. (4.75) haben dieselbe Form wie die *Eikonal-Gleichung* (4.62).

Die klassischen Teilchenbahnen sind die *Strahlen* der Wirkungswellen. Dies sind die Kurven im Konfigurationsraum der \vec{q} , die überall senkrecht auf den Flächen $W = \text{const}$ stehen.

4.3.3 Die Wellengleichung der Quantenmechanik

Die Eikonal-Gleichung für elektromagnetische Wellen ergab sich im geometrisch-optischen Grenzfall aus der Wellengleichung. Die zentrale Idee ist nun, die formal analoge Eikonal-Gleichung (4.75) der klassischen Mechanik,

$$\left(\frac{\partial W}{\partial \vec{q}} \right)^2 = \frac{E^2}{u^2} \quad (4.79)$$

als geometrisch-optischen Grenzfall einer allgemeineren Mechanik, nämlich der Quantenmechanik, zu verstehen. Wie erwähnt ergeben sich die klassischen Teilchenbahnen als Strahlen im Rahmen dieses Grenzfalls. Die Aufgabe besteht also darin, eine Wellengleichung zu finden, die im geometrisch-optischen Grenzfall diese Eikonal-Gleichung ergibt.

In Analogie zum Ansatz

$$F(\vec{r}, t) = F_0 \exp \left(i\omega \left[\frac{L}{c} - t \right] \right) \quad (4.80)$$

für elektromagnetische Wellen setzen wir die *Wellenfunktion* für Materiewellen als

$$\psi(\vec{q}, t) = \psi_0 \exp \left(i \underbrace{\omega}_{= E/\hbar} \left[\frac{W}{E} - t \right] \right) = \psi_0 \exp \left(\frac{i}{\hbar} [W - Et] \right) = \psi_0 e^{iS/\hbar} \quad (4.81)$$

an. Die Konstante \hbar von der Dimension einer Wirkung wird benötigt, um die Phase dimensionslos zu machen. Aber nichts in der Eikonal-Gleichung (und der gesamten klassischen Physik) sagt uns, welchen Wert \hbar annimmt. Das ist analog dazu, dass die elektromagnetische Eikonal-Gleichung $(\vec{\nabla} L)^2 = n^2$ die Konstante c nicht enthält, die für die Wellengleichung aber essentiell ist.

Welche Wellengleichung wird nun von ψ erfüllt? Wir beachten, dass gilt

$$\frac{\partial}{\partial \vec{q}} \psi = \frac{i}{\hbar} \frac{\partial W}{\partial \vec{q}} \psi, \quad (4.82)$$

$$\frac{\partial}{\partial t} \psi = -\frac{i}{\hbar} E \psi. \quad (4.83)$$

Lassen wir nun eine operatorwertige Funktion $G(\vec{q}, \partial/\partial\vec{q})$ der Koordinaten und der Ableitungen nach den Koordinaten auf ψ wirken, so holt jede Ableitung $\partial/\partial q_i$ einen Faktor $(i/\hbar) \partial W/\partial q_i$ vor die Exponentialfunktion. Aber es passiert noch mehr, da die Ableitungen auf alles wirken, was dahinter steht: Produkte mit mehreren Faktoren $\partial/\partial q_i$ erzeugen höhere Ableitungen von W . Als Beispiel betrachten wir die Funktion

$$H\left(\vec{q}, \frac{\hbar}{i} \frac{\partial}{\partial \vec{q}}\right) = -\frac{\hbar^2}{2m} \frac{\partial}{\partial \vec{q}} \cdot \frac{\partial}{\partial \vec{q}} + V(\vec{q}). \quad (4.84)$$

Wir finden

$$H\left(\vec{q}, \frac{\hbar}{i} \frac{\partial}{\partial \vec{q}}\right) \psi = \frac{1}{2m} \frac{\partial W}{\partial \vec{q}} \cdot \frac{\partial W}{\partial \vec{q}} \psi - i \frac{\hbar}{2m} \underbrace{\left(\frac{\partial}{\partial \vec{q}} \cdot \frac{\partial}{\partial \vec{q}} W\right)}_{=\nabla_{\vec{q}}^2 W} \psi + V(\vec{q}) \psi. \quad (4.85)$$

Ändert sich nun $V(\vec{q})$ nur langsam oder sprunghaft als Funktion von \vec{q} , so können wir mit derselben Begründung wie bei der Herleitung der elektromagnetischen Eikonal-Gleichung die zweiten Ableitungen von W weglassen. Im geometrisch-optischen Grenzfall folgt also

$$H\left(\vec{q}, \frac{\hbar}{i} \frac{\partial}{\partial \vec{q}}\right) \psi = \frac{1}{2m} \left(\frac{\partial W}{\partial \vec{q}}\right)^2 \psi + V(\vec{q}) \psi. \quad (4.86)$$

Für die Hamilton-Funktion

$$H(\vec{q}, \vec{p}) = \frac{\vec{p} \cdot \vec{p}}{2m} + V(\vec{q}), \quad (4.87)$$

das ist *dieselbe* Funktion H , aber mit anderen Argumenten, lautet die Hamilton-Jacobi-Gleichung

$$H\left(\vec{q}, \frac{\partial W}{\partial \vec{q}}\right) = \frac{1}{2m} \left(\frac{\partial W}{\partial \vec{q}}\right)^2 + V(\vec{q}) = E. \quad (4.88)$$

Damit folgt im geometrisch-optischen Grenzfall

$$i\hbar \frac{\partial \psi}{\partial t} \stackrel{\text{Gl. (4.83)}}{=} E \psi \stackrel{\text{Gl. (4.88)}}{=} H\left(\vec{q}, \frac{\partial W}{\partial \vec{q}}\right) \psi \stackrel{\text{Gl. (4.86)}}{=} H\left(\vec{q}, \frac{\hbar}{i} \frac{\partial}{\partial \vec{q}}\right) \psi. \quad (4.89)$$

Wir haben damit die Schrödinger-Gleichung

$$i\hbar \frac{\partial \psi}{\partial t} = H\left(\vec{q}, \frac{\hbar}{i} \frac{\partial}{\partial \vec{q}}\right) \psi \quad (4.90)$$

erhalten. Offenbar wurden in der Hamilton-Funktion die Impulse p_i durch Ableitungen nach den Koordinaten ersetzt, genauer durch $(\hbar/i) \partial/\partial q_i$.

Wenn wir die Argumentation nun rückwärts lesen, erhalten wir die Eikonal-Gleichung der klassischen Mechanik, und damit im wesentlichen deren Hamilton-Jacobi-Formulierung, als geometrisch-optischen Grenzfall der durch die Schrödinger-Gleichung beschriebenen Dynamik von Materiewellen. Der induktive Schluss vom Speziellen (dem geometrisch-optischen Grenzfall) auf das Allgemeine ist natürlich nicht zwingend, sondern hat den Charakter eines Postulats. Schrödinger hat weiter postuliert, dass die Gleichung allgemein gilt, nicht nur für die Form von Hamilton-Funktionen, die wir betrachtet haben.

4.4 Weitere Bemerkungen zu Schrödinger-Gleichung

Für ein massives Teilchen mit nichtrelativistischer kinetischer Energie und in Abwesenheit eines Magnetfeldes lautet die Schrödinger-Gleichung

$$i\hbar \frac{\partial \psi}{\partial t} = -\frac{\hbar^2}{2m} \nabla^2 \psi + V(\vec{r}) \psi. \quad (4.91)$$

Wir werden uns ausführlich mit Lösungen dieser Gleichung beschäftigen. Es wurde schon erwähnt, dass die rechte Seite der Gleichung bis auf die Wellenfunktion ψ aus der klassischen Hamilton-Funktion hervorgeht, indem der Impuls \vec{p} durch $\frac{\hbar}{i} \vec{\nabla}$ ersetzt wird. Für ein konservatives System, wie hier angenommen, ist die Hamilton-Funktion gleich der Energie. Dies führt auf die Idee, auf der linken Seite der Schrödinger-Gleichung $i\hbar \frac{\partial}{\partial t}$ als Darstellung der Energie E aufzufassen. Die Ersetzungen

$$E \rightarrow i\hbar \frac{\partial}{\partial t}, \quad (4.92)$$

$$\vec{p} \rightarrow -i\hbar \vec{\nabla} \quad (4.93)$$

sind aus Sicht der Relativitätstheorie durchaus plausibel, denn sie können zu der kovarianten Regel

$$p^\mu \rightarrow i\hbar \partial^\mu \quad (4.94)$$

zusammengefasst werden. Dies ist kein sehr starkes Argument, da die Schrödinger-Theorie nicht kovariant ist. Auf jeden Fall muss man sich klar machen, dass Glg. (4.92) und (4.93) nur eine Zusammenfassung des Übergangs zur Quantenmechanik darstellen, die allein durch das Korrespondenzprinzip (der klassische Grenzfall ergibt sich korrekt) und Vergleich mit Experimenten gerechtfertigt wird.

Wir betrachten noch kurz den Fall mit verschwindenden Potential:

$$i\hbar \frac{\partial \psi}{\partial t} = -\frac{\hbar^2}{2m} \nabla^2 \psi, \quad (4.95)$$

Die Lösungen sind ebene Wellen:

$$\psi(\vec{r}, t) = \psi_0 e^{i(\vec{k} \cdot \vec{r} - \omega t)}. \quad (4.96)$$

Einsetzen ergibt

$$\hbar \omega \psi = \frac{\hbar^2 k^2}{2m} \psi, \quad (4.97)$$

also

$$\hbar \omega = \frac{\hbar^2 k^2}{2m}. \quad (4.98)$$

Dies liefert die *Dispersionsrelation*

$$\omega(\vec{k}) = \frac{\hbar k^2}{2m} \quad (4.99)$$

von freien Materiewellen. (Vergleiche die Dispersionsrelation freier elektromagnetischer Wellen: $\omega(\vec{k}) = kc$.) Unter der Annahme der de Broglie-Postulate $E = \hbar \omega$ und $\vec{p} = \hbar \vec{k}$ ist die Relation offensichtlich konsistent mit der klassischen kinetischen Energie

$$E = \frac{p^2}{2m}. \quad (4.100)$$

Schrödinger hat übrigens zuerst eine Gleichung der Form

$$\left(\frac{\hbar^2}{c^2} \frac{\partial^2}{\partial t^2} - \hbar^2 \nabla^2 + m^2 c^4 \right) \psi(\vec{r}, t) = 0 \quad (4.101)$$

hingeschrieben. Diese ergibt nicht die gewünschte, quadratische Dispersion, sondern

$$\omega^2(\vec{k}) = k^2 c^2 + \frac{m^2 c^4}{\hbar^2}. \quad (4.102)$$

Übersetzt für Energie und Impuls erhalten wir

$$E^2 = p^2 c^2 + m^2 c^4, \quad (4.103)$$

also die relativistische Energie-Impuls-Beziehung. Gleichung (4.101) ist die sogenannte *Klein-Gordon-Gleichung*, die in der Tat eine relativistische Quantentheorie definiert, vgl. Vorlesung Quantentheorie 2. Schrödinger hat sie verworfen, weil sie von zweiter Ordnung in der Zeit ist.

Kapitel 5

Die Schrödingersche Wellenmechanik

5.1 Operatoren

Die Schrödinger-Gleichung

$$i\hbar \frac{\partial}{\partial t} \psi(\vec{r}, t) = -\frac{\hbar^2}{2m} \nabla^2 \psi(\vec{r}, t) + V(\vec{r}) \psi(\vec{r}, t) \quad (5.1)$$

enthält Terme, die durch Ableitungen der *Wellenfunktion* $\psi(\vec{r}, t)$ oder durch ihre Multiplikation mit anderen Funktionen entstehen. Diese Terme können wir durch die Einführung des Begriffs des *Operators* vereinheitlichen. Ein Operator \hat{A} ist eine Abbildung von einem gewissen Raum R , hier einem *Funktionsraum*, in *denselben* Raum R , also

$$\hat{A}: \begin{array}{ccc} R & \rightarrow & R \\ \psi & \mapsto & \hat{A}\psi. \end{array} \quad (5.2)$$

Beachte, dass man meistens $\hat{A}\psi$ und nicht $\hat{A}(\psi)$ für das Bild schreibt. In dieser Vorlesung kennzeichnen wir Operatoren üblicherweise mit einem Dach (Zirkumflex). In der Quantentheorie interessieren uns ganz überwiegend *lineare* Operatoren. Diese sind durch die Eigenschaft

$$\hat{A}(\lambda_1 \psi_1 + \lambda_2 \psi_2) = \lambda_1 \hat{A}\psi_1 + \lambda_2 \hat{A}\psi_2 \quad \forall \psi_1, \psi_2 \in R \quad \text{und} \quad \lambda_1, \lambda_2 \in \mathbb{C} \quad (5.3)$$

definiert. Zwei Typen von linearen Operatoren sind in der Wellenmechanik besonders wichtig:

1. Differentialoperatoren, z. B.

$$\frac{\partial}{\partial x} : \psi(\vec{r}, t) \mapsto \frac{\partial}{\partial x} \psi(\vec{r}, t), \quad (5.4)$$

2. Multiplikationsoperatoren, z. B.

$$\hat{V}(\vec{r}) : \psi(\vec{r}, t) \mapsto V(\vec{r}) \psi(\vec{r}, t). \quad (5.5)$$

Operatoren können auch Vektorcharakter haben; ein Vektoroperator ist ein Vektor, dessen Komponenten Operatoren im obigen Sinne sind. Wichtige Beispiele sind

1. der Nabla-Operator

$$\vec{\nabla} := \begin{pmatrix} \partial/\partial x \\ \partial/\partial y \\ \partial/\partial z \end{pmatrix}; \quad \vec{\nabla} : \psi(\vec{r}, t) \mapsto \begin{pmatrix} \partial\psi/\partial x \\ \partial\psi/\partial y \\ \partial\psi/\partial z \end{pmatrix}, \quad (5.6)$$

2. die Multiplikation mit einem Vektor, z. B.

$$\hat{\vec{r}} : \psi(\vec{r}, t) \mapsto \vec{r} \psi(\vec{r}, t). \quad (5.7)$$

Man zeigt leicht, dass für lineare Operatoren \hat{A} , \hat{B} folgendes gilt:

1. $c\hat{A}$ mit einer Zahl $c \in \mathbb{C}$ ist ein linearer Operator,
2. $\hat{A} + \hat{B}$ ist ein linearer Operator,
3. $\hat{A}\hat{B}$ ist ein linearer Operator, wobei das Produkt als Hintereinanderausführung von rechts nach links definiert ist:

$$\hat{A}\hat{B}\psi(\vec{r}, t) := \hat{A}[\hat{B}\psi(\vec{r}, t)]. \quad (5.8)$$

Es ist wichtig zu beachten, dass das Operatorprodukt i. A. *nicht kommutativ* ist. Beispiel:

$$\hat{A} = \frac{\partial}{\partial x}, \quad \hat{B} = x, \quad \psi(x) = x^2. \quad (5.9)$$

Dann ist

$$\hat{A}\hat{B}\psi = \frac{\partial}{\partial x}(xx^2) = \frac{\partial}{\partial x}x^3 = 3x^2, \quad (5.10)$$

aber

$$\hat{B}\hat{A}\psi = x \frac{\partial}{\partial x}x^2 = x \cdot 2x = 2x^2. \quad (5.11)$$

Man definiert den *Kommutator*

$$[\hat{A}, \hat{B}] := \hat{A}\hat{B} - \hat{B}\hat{A}. \quad (5.12)$$

Der Kommutator $[\hat{A}, \hat{B}]$ ist ebenfalls ein Operator auf demselben Raum wie \hat{A} und \hat{B} . Er drückt offenbar das Maß der Nichtkommutativität aus. Aus der Definition folgen leicht eine Reihe von nützlichen Rechenregeln:

1. Der Kommutator ist antisymmetrisch unter Vertauschung der beiden Argumente:

$$[\hat{B}, \hat{A}] = \hat{B}\hat{A} - \hat{A}\hat{B} = -(\hat{A}\hat{B} - \hat{B}\hat{A}) = -[\hat{A}, \hat{B}]. \quad (5.13)$$

2. Der Kommutator ist additiv in jedem der beiden Argumente:

$$[\hat{A} + \hat{B}, \hat{C}] = (\hat{A} + \hat{B})\hat{C} - \hat{C}(\hat{A} + \hat{B}) = \hat{A}\hat{C} + \hat{B}\hat{C} - \hat{C}\hat{A} - \hat{C}\hat{B} = [\hat{A}, \hat{C}] + [\hat{B}, \hat{C}], \quad (5.14)$$

$$[\hat{A}, \hat{B} + \hat{C}] = \hat{A}(\hat{B} + \hat{C}) - (\hat{B} + \hat{C})\hat{A} = \hat{A}\hat{B} + \hat{A}\hat{C} - \hat{B}\hat{A} - \hat{C}\hat{A} = [\hat{A}, \hat{B}] + [\hat{A}, \hat{C}]. \quad (5.15)$$

3. Kommutatoren, die Produkte enthalten, erfüllen

$$\begin{aligned} [\hat{A}\hat{B}, \hat{C}] &= \hat{A}\hat{B}\hat{C} - \hat{C}\hat{A}\hat{B} \\ &= \hat{A}\hat{B}\hat{C} - \underbrace{\hat{A}\hat{C}\hat{B} + \hat{A}\hat{C}\hat{B}}_{=0} - \hat{C}\hat{A}\hat{B} \\ &= \hat{A}(\hat{B}\hat{C} - \hat{C}\hat{B}) + (\hat{A}\hat{C} - \hat{C}\hat{A})\hat{B} \\ &= \hat{A}[\hat{B}, \hat{C}] + [\hat{A}, \hat{C}]\hat{B}. \end{aligned} \quad (5.16)$$

Die Reihenfolge der Faktoren ist wichtig, da sie i. A. nicht vertauschen! Zusammen mit der Antisymmetrie des Kommutators erhalten wir auch

$$[\hat{A}, \hat{B}\hat{C}] = \hat{B}[\hat{A}, \hat{C}] + [\hat{A}, \hat{B}]\hat{C}. \quad (5.17)$$

Speziell, wenn einer der Operatoren die Multiplikation mit einer Zahl $\lambda \in \mathbb{C}$ ist, folgt

$$[\lambda\hat{A}, \hat{B}] = \lambda[\hat{A}, \hat{B}] + \underbrace{[\lambda, \hat{B}]}_{=0}\hat{A} = \lambda[\hat{A}, \hat{B}], \quad (5.18)$$

$$[\hat{A}, \lambda\hat{B}] = \lambda[\hat{A}, \hat{B}] + \underbrace{[\hat{A}, \lambda]}_{=0}\hat{B} = \lambda[\hat{A}, \hat{B}]. \quad (5.19)$$

Zusammen mit Punkt 2 folgt, dass der Kommutator in jedem seiner Argumente linear ist.

Beispiel: Für eine beliebige differenzierbare Funktion $\psi(x)$ gilt

$$\left[\frac{\partial}{\partial x}, x \right] \psi(x) = \frac{\partial}{\partial x} x \psi(x) - x \frac{\partial}{\partial x} \psi(x) = \psi(x) + x \frac{\partial}{\partial x} \psi(x) - x \frac{\partial}{\partial x} \psi(x) = \psi(x). \quad (5.20)$$

Da dies für beliebige differenzierbare ψ gilt, können wir das Ergebnis als *Operatoridentität*

$$\left[\frac{\partial}{\partial x}, x \right] = 1 \quad (5.21)$$

auf dem Raum der differenzierbaren Funktionen schreiben.

Die heuristische Begründung in 4.2 und 4.3 hatte die folgenden *Quantisierungsregeln* nahegelegt:

$$\vec{r} \rightarrow \hat{\vec{r}} = \vec{r}, \quad (5.22)$$

$$\vec{p} \rightarrow \hat{\vec{p}} = \frac{\hbar}{i} \vec{\nabla}. \quad (5.23)$$

Dies führt auf

$$[\hat{r}_i, \hat{p}_j] = \left[r_i, \frac{\hbar}{i} \frac{\partial}{\partial r_j} \right] = \frac{\hbar}{i} \left[r_i, \frac{\partial}{\partial r_j} \right] = -\frac{\hbar}{i} \underbrace{\left[\frac{\partial}{\partial r_j}, r_i \right]}_{\delta_{ij}} = i\hbar \delta_{ij}. \quad (5.24)$$

Diese *Orts-Impuls-Vertauschungsrelation* ist fundamental für die Quantentheorie.

Mittels der Quantisierungsregeln ergibt sich auch der *Hamilton-Operator* \hat{H} aus der Hamilton-Funktion:

$$H(\vec{r}, \vec{p}) \rightarrow \hat{H} = H(\hat{\vec{r}}, \hat{\vec{p}}). \quad (5.25)$$

Dies ist nicht immer eindeutig, weil Ort und Impuls nicht miteinander vertauschen. Enthält die Hamilton-Funktion z. B. einen Term der Form

$$H_1 = c \vec{r} \cdot \vec{p} = c \vec{p} \cdot \vec{r} \quad (5.26)$$

mit $c = \text{const}$, so ist nicht klar, welchen der folgenden Operatoren wir verwenden sollen:

$$\hat{H}_1 = \begin{cases} c \hat{\vec{r}} \cdot \hat{\vec{p}}, \\ c \hat{\vec{p}} \cdot \hat{\vec{r}}, \\ c \frac{\hat{\vec{r}} \cdot \hat{\vec{p}} + \hat{\vec{p}} \cdot \hat{\vec{r}}}{2}. \end{cases} \quad (5.27)$$

Oder ist der richtige Operator keiner der drei? Wir werden sehen, dass die letzte, symmetrisierte Form die richtige ist. Letztlich handelt es sich aber um ein Scheinproblem. Die Quantenmechanik ist die umfassendere Theorie und wir können daher nicht erwarten, sie eindeutig aus der klassischen Mechanik zu gewinnen.

Allgemeiner hängt die klassische Hamilton-Funktion von generalisierten Koordinaten q_j und zugehörigen Impulsen p_j und evtl. der Zeit ab:

$$H = H(q_1, \dots, p_1, \dots, t). \quad (5.28)$$

Nach 4.3 würden wir die Quantisierungsregeln gern verallgemeinern zu

$$\begin{aligned} q_j &\xrightarrow{?} \hat{q}_j = q_j, \\ p_j &\xrightarrow{?} \hat{p}_j = \frac{\hbar}{i} \frac{\partial}{\partial q_j}. \end{aligned} \quad (5.29)$$

Aber diese Regeln wären widersprüchlich. Z. B. erhalten wir nicht denselben Hamilton-Operator, wenn wir die Regeln für kartesische oder für sphärische Koordinaten verwenden. Vergleich mit dem Experiment und die Überlegung, dass die Koordinateneinheitsvektoren und damit die Quantisierungsregeln ortsunabhängig sein sollten, ergeben, dass man *kartesische* Koordinaten verwenden muss, wie Dirac betont hat. Die Quantenmechanik lässt sich für generalisierte Koordinaten verallgemeinern, die Quantisierungsregeln fußen aber immer auf der fundamentalen Orts-Impuls-Vertauschungsrelation in kartesischen Koordinaten. Die Quantenmechanik in Schrödingers Formulierung wirft uns auf das Niveau der Newton-Mechanik zurück, die ja auch primär mit kartesischen Koordinaten formuliert ist.

5.2 Die Schrödinger-Gleichung als partielle Differentialgleichung

Die allgemeine Schrödinger-Gleichung für ein Teilchen lautet

$$i\hbar \frac{\partial}{\partial t} \psi(\vec{r}, t) = \hat{H} \psi(\vec{r}, t), \quad (5.30)$$

wobei \hat{H} der Hamilton-Operator ist. \hat{H} enthält Ableitungen nach dem Ort, was die Schrödinger-Gleichung zu einer *partiellen* Differentialgleichung macht. Wir besprechen nun kurz wichtige Eigenschaften der Gleichung und deren Konsequenzen.

- Die Schrödinger-Gleichung ist *linear*. Daraus folgt das *Superpositionsprinzip*: Sind ψ_1 und ψ_2 zwei Lösungen der Gleichung, so ist

$$\psi = \lambda_1 \psi_1 + \lambda_2 \psi_2 \quad (5.31)$$

mit beliebigen Konstanten $\lambda_1, \lambda_2 \in \mathbb{C}$ ebenfalls eine Lösung. Das ist nicht nur sehr wichtig für die praktische Lösung der Gleichung für spezielle Systeme – viele Lösungsmethoden funktionieren nur für lineare Gleichungen – sondern auch von überragender fundamentaler Bedeutung. Wir kommen bei der Diskussion der Interpretation der Quantenmechanik darauf zurück.

- Die Schrödinger-Gleichung ist von *erster Ordnung in der Zeit*. Daher legt für gegebenes \hat{H} die Angabe von $\psi(\vec{r}, t_0)$ für alle \vec{r} und feste Zeit t_0 als Anfangsbedingung die Lösung für alle Zeiten $t > t_0$ (und auch $t < t_0$) eindeutig fest.
- Wenn \hat{H} zweite (oder höhere) Ableitungen nach dem Ort enthält, z. B. für

$$\hat{H} = \frac{\hat{p}^2}{2m} + V(\vec{r}), \quad (5.32)$$

dann ist die Schrödinger-Gleichung sicher nicht kovariant, d.h. sie steht im Widerspruch zur Speziellen Relativitätstheorie. Denn um kovariant zu sein, müsste sie Ableitungen in der Form des Vierergradienten

$$(\partial_\mu) = \left(\frac{1}{c} \frac{\partial}{\partial t}, \vec{\nabla} \right) \quad (5.33)$$

enthalten. Da sie von erster Ordnung in der Zeit ist, müsste sie insbesondere auch von erster Ordnung im Ort sein. Dirac ist es gelungen, eine solche kovariante Quantenmechanik zu konstruieren, vgl. die Vorlesung Quantentheorie 2.

5.3 Wahrscheinlichkeitswellen

Wir haben noch nicht diskutiert, was die Wellenfunktion $\psi(\vec{r}, t)$ eigentlich bedeutet. Nehmen wir ein Wellenpaket als Anfangsbedingung, so finden wir im Allgemeinen, dass seine Breite Δx mit der Zeit anwächst. Der harmonische Oszillator bildet hier eine Ausnahme. Für allgemeine Potentiale zerfließt das Wellenpaket. Der Teilchencharakter geht damit mehr und mehr verloren, wenn wir der semiklassischen Deutung von Wellenpaketen aus 4.2 folgen. Wir beobachten jedoch nicht, dass Teilchen zerfließen – Streuexperimente zeigen z. B. dass Elektronen punktförmig sind und bleiben. Auch bei der Beugung von Elektronen am Doppelspalt, bei der ja der Wellencharakter essentiell ist (Interferenz!), werden immer punktförmige Elektronen nachgewiesen, wenn sie auf den Schirm treffen. Auch nicht punktförmige zusammengesetzte Teilchen wie Atomkerne zerfließen nicht.

Wie können wir den scheinbaren Widerspruch zwischen der Existenz lokalisierter Teilchen und dem Zerfließen von Wellenpaketen auflösen? Einen Hinweis liefert die Feststellung, dass wir weder im Rahmen der Quantenmechanik noch auf andere Weise vorhersagen können, wo das nächste Teilchen beim Doppelspaltexperiment auftreffen wird. Ebenso wenig können wir beim Stern-Gerlach-Versuch vorhersagen, ob das nächste Atom nach oben oder nach unten abgelenkt werden wird. Dies legt eine *statistische Interpretation* der Wellenfunktion nahe: Die Wellenfunktion $\psi(\vec{r}, t)$ beschreibt demnach *nicht*, wie die Materie zur Zeit t im Raum verteilt ist, sondern, mit welcher

Wahrscheinlichkeit das – an sich exakt oder näherungsweise punktförmige – Teilchen zur Zeit t am Ort \vec{r} zu finden ist.

Wir müssen uns überlegen, wie $\psi(\vec{r}, t)$ im Detail diese Wahrscheinlichkeit beschreibt. Eine naheliegende Idee ist, dass $\psi(\vec{r}, t)$ die Wahrscheinlichkeitsdichte ist, dass also $\psi(\vec{r}, t) d^3r$ die Wahrscheinlichkeit dafür ist, ein Teilchen im Volumenelement d^3r zu finden. Aber das kann nicht stimmen, denn eine Wahrscheinlichkeitsdichte $\varrho(\vec{r}, t)$ muss

1. $\varrho(\vec{r}, t) \in \mathbb{R}$ und $\varrho(\vec{r}, t) \geq 0 \forall \vec{r}, t$ und
2. $\int d^3r \varrho(\vec{r}, t) = 1 \forall t$ (die Gesamtwahrscheinlichkeit ist eins)

erfüllen. Wenn wir zu einem Zeitpunkt, z. B. $t = 0$, eine Funktion $\psi(\vec{r}, t = 0)$ mit diesen Eigenschaften wählen, zeigt aber die Schrödinger-Gleichung, dass sie für spätere Zeiten i. A. nicht mehr gelten: Unter der Wirkung der Schrödinger-Gleichung ist weder $\int d^3r \psi(\vec{r}, t)$ erhalten, noch ist $\psi(\vec{r}, t)$ positiv semidefinit.

Andererseits führt der Ansatz

$$\varrho(\vec{r}, t) = |\psi(\vec{r}, t)|^2 = \psi^*(\vec{r}, t) \psi(\vec{r}, t) \quad (5.34)$$

nicht zu Problemen. Bedingung 1 ist automatisch erfüllt. Außerdem ist

$$\begin{aligned} \frac{d}{dt} \int d^3r |\psi(\vec{r}, t)|^2 &= \int d^3r \left[\frac{\partial \psi^*}{\partial t} \psi + \psi^* \frac{\partial \psi}{\partial t} \right] \\ &= \int d^3r \left[\left(\frac{1}{i\hbar} \hat{H} \psi \right)^* \psi + \psi^* \frac{1}{i\hbar} \hat{H} \psi \right] \\ &= \frac{1}{i\hbar} \int d^3r \left[-(\hat{H} \psi)^* \psi + \psi^* \hat{H} \psi \right]. \end{aligned} \quad (5.35)$$

Hierbei ist für den Fall eines Teilchens

$$\begin{aligned} (\hat{H} \psi)^* &= \left[-\frac{\hbar^2}{2m} \nabla^2 \psi(\vec{r}, t) + V(\vec{r}) \psi(\vec{r}, t) \right]^* \\ &= -\frac{\hbar^2}{2m} \nabla^2 \psi^*(\vec{r}, t) + V(\vec{r}) \psi^*(\vec{r}, t) \end{aligned} \quad (5.36)$$

und damit

$$\begin{aligned} \frac{d}{dt} \int d^3r |\psi(\vec{r}, t)|^2 &= \frac{1}{i\hbar} \int d^3r \left[\frac{\hbar^2}{2m} (\nabla^2 \psi^*) \psi - \cancel{V(\vec{r}) \psi^* \psi} - \frac{\hbar^2}{2m} \psi^* \nabla^2 \psi + \cancel{V(\vec{r}) \psi^* \psi} \right] \\ &\stackrel{\text{partiell}}{=} \underbrace{\text{Oberflächenterm}}_{=0} - \frac{1}{i\hbar} \frac{\hbar^2}{2m} \int d^3r [(\vec{\nabla} \psi^*) \cdot (\vec{\nabla} \psi) - (\vec{\nabla} \psi^*) \cdot (\vec{\nabla} \psi)] \\ &= 0. \end{aligned} \quad (5.37)$$

Die Oberflächenterme verschwinden hier, wenn wir annehmen, dass $\psi(\vec{r}, t)$ für $|\vec{r}| \rightarrow \infty$ hinreichend schnell abfällt. Es folgt, dass die Gesamtwahrscheinlichkeit $\int d^3r |\psi|^2$ eine Erhaltungsgröße ist. Fordern wir also $\int d^3r |\psi|^2 = 1$ im Anfangszustand, so gilt diese Normierung für alle Zeiten. (Der Ansatz $\rho = |\psi|$ würde zwar auch Bedingung 1 erfüllen, aber nicht Bedingung 2.)

Wir haben oben eine spezielle Form des Hamilton-Operators \hat{H} angenommen. Es ist nach Gl. (5.35) klar, dass Erhaltung der Gesamtwahrscheinlichkeit gilt, sofern allgemeiner

$$\int d^3r (\hat{H} \psi)^* \psi = \int d^3r \psi^* \hat{H} \psi \quad (5.38)$$

erfüllt ist. Diese Eigenschaft des Operators \hat{H} nennt man *Hermitizität*. \hat{H} muss also *hermitesch* sein, um $|\psi|^2$ als Wahrscheinlichkeitsdichte interpretieren zu können. $\psi(\vec{r}, t)$ selbst nennt man *Wahrscheinlichkeitsamplitude*.

Als Wellenfunktion für ein Teilchen kommen also solche Funktionen $\psi(\vec{r}, t)$ in Frage, die die Normierungsbedingung

$$\int d^3r |\psi(\vec{r}, t)|^2 = 1 \quad (5.39)$$

erfüllen. Etwas allgemeiner können wir auch nur fordern, dass

$$\int d^3r |\psi(\vec{r}, t)|^2 < \infty \quad (5.40)$$

ist, d. h., dass das Integral konvergiert. Dann können wir $\psi(\vec{r}, t)$ einfach durch Multiplikation mit einer Zahl normieren. Funktionen, die Gleichung (5.40) erfüllen, heißen *quadratintegrabel*.

5.3.1 Der Hilbert-Raum der quadratintegrablen Funktionen

Die Menge der quadratintegrablen Funktionen $f : M \rightarrow \mathbb{C}$ bezeichnen wir mit $L_2(M)$. In der Einteilchen-Quantenmechanik interessiert uns also insbesondere $L_2(\mathbb{R}^3)$. Diese Menge ausgestattet mit der gewöhnlichen Addition und der Multiplikation mit Zahlen aus \mathbb{C} ist ein *Vektorraum* (linearer Raum) über \mathbb{C} , denn mit $\psi_1(\vec{r})$ und $\psi_2(\vec{r})$ (wir unterdrücken das Zeitargument, das hier keine Rolle spielt) ist auch $\lambda_1 \psi_1(\vec{r}) + \lambda_2 \psi_2(\vec{r})$ mit beliebigen $\lambda_1, \lambda_2 \in \mathbb{C}$ quadratintegrabel. Beweis:

$$\begin{aligned} \int d^3r |\lambda_1 \psi_1(\vec{r}) + \lambda_2 \psi_2(\vec{r})|^2 &\leq \int d^3r (|\lambda_1 \psi_1(\vec{r})| + |\lambda_2 \psi_2(\vec{r})|)^2 \quad \left| \text{Dreiecksungleichung} \right. \\ &\leq \int d^3r [(|\lambda_1 \psi_1(\vec{r})| + |\lambda_2 \psi_2(\vec{r})|)^2 + (|\lambda_1 \psi_1(\vec{r})| - |\lambda_2 \psi_2(\vec{r})|)^2] \\ &= 2 \int d^3r [|\lambda_1|^2 |\psi_1(\vec{r})|^2 + |\lambda_2|^2 |\psi_2(\vec{r})|^2] \\ &= 2 |\lambda_1|^2 \int d^3r |\psi_1(\vec{r})|^2 + 2 |\lambda_2|^2 \int d^3r |\psi_2(\vec{r})|^2 \end{aligned} \quad (5.41)$$

und dies ist endlich, da ψ_1 und ψ_2 quadratintegrabel sind. Es sei darauf hingewiesen, dass wir in diesem Abschnitt den Integralbegriff nach *Lebesgue* benötigen, der allgemeiner ist als der nach Riemann. Nicht ganz rigoros formuliert ist eine Funktion f Lebesgue-integrierbar, wenn sie sich von einer Riemann- (d.h. im üblichen Sinne) integrierbaren Funktion \tilde{f} nur auf einer Menge vom Maß Null unterscheidet, und das Lebesgue-Integral von f ist dann gleich dem Riemann-Integral von \tilde{f} . Der Begriff ist erforderlich, um einige Beweise führen zu können, wir beachten den Unterschied aber nicht weiter.

Für die Quantenmechanik wichtig ist, dass $L_2(\mathbb{R}^3)$ nicht nur ein Vektorraum, sondern darüber hinaus ein *separabler Hilbert-Raum* ist. Hier müssen wir zunächst diesen Begriff definieren. Dazu benötigen wir einige Vorbereitungen.

1. Ein *Vektorraum* V über dem Körper K erfüllt folgende Axiome:

- (a) Es existiert eine Addition $+: V \otimes V \rightarrow V$, bezüglich der V eine abelsche Gruppe ist.
- (b) Es existiert eine Multiplikation mit einem Skalar $\cdot : K \otimes V \rightarrow V$ (das Verknüpfungssymbol wird nicht ausgeschrieben) mit den Eigenschaften

$$\lambda(x + y) = \lambda x + \lambda y \quad (\text{Distributivität}), \quad (5.42)$$

$$(\lambda + \mu)x = \lambda x + \mu x \quad (\text{Distributivität}), \quad (5.43)$$

$$(\lambda\mu)x = \lambda(\mu x) \quad (\text{Assoziativität}), \quad (5.44)$$

$$1x = x, \quad (5.45)$$

jeweils für alle möglichen λ, μ, x, y .

2. Ein *Skalarprodukt* zwischen Elementen eines Vektorraums V über dem Körper $K = \mathbb{R}$ oder \mathbb{C} ist eine Abbildung $\langle \bullet, \bullet \rangle : V \otimes V \rightarrow K$ mit den folgenden Eigenschaften:

- (a) $\langle x, x \rangle \in \mathbb{R}$ und $\langle x, x \rangle \geq 0$, wobei $\langle x, x \rangle = 0$ genau dann, wenn $x = 0$ (Nullelement von V).
- (b) $\langle x, y \rangle = \langle y, x \rangle^*$, wobei $*$ die komplexe Konjugation bezeichnet (überflüssig für $K = \mathbb{R}$).

(c)

$$\langle x, \lambda_1 y_1 + \lambda_2 y_2 \rangle = \lambda_1 \langle x, y_1 \rangle + \lambda_2 \langle x, y_2 \rangle, \quad (5.46)$$

d. h. das Skalarprodukt ist linear im zweiten Argument (in der reinen Mathematik wird es oft als linear im ersten Argument definiert). Daraus folgt mit (b), dass gilt

$$\langle \lambda_1 x_1 + \lambda_2 x_2, y \rangle = \lambda_1^* \langle x_1, y \rangle + \lambda_2^* \langle x_2, y \rangle, \quad (5.47)$$

d. h. das Skalarprodukt für $K = \mathbb{C}$ ist *antilinear* (für $K = \mathbb{R}$ linear) im ersten Argument.

Zwei Elemente $x, y \in V$, $x, y \neq 0$, heißen *orthogonal*, wenn gilt

$$\langle x, y \rangle = 0. \quad (5.48)$$

Mit Hilfe des Skalarprodukts kann man eine *Norm*

$$\|x\| := \sqrt{\langle x, x \rangle} \quad (5.49)$$

auf V definieren.

3. Ein mit einem Skalarprodukt ausgestatteter Vektorraum heißt *unitärer Raum*. (Ein mit einer Norm ausgestatteter Vektorraum heißt normierter Raum. Offenbar ist jeder unitäre Raum auch normiert.) $L_2(\mathbb{R}^3)$ mit dem Skalarprodukt

$$\langle \varphi, \psi \rangle = \int d^3r \varphi^*(\vec{r}) \psi(\vec{r}) \quad (5.50)$$

ist ein unitärer Raum. Zum Beweis ist zu zeigen, dass $\langle \varphi, \psi \rangle$ tatsächlich ein Skalarprodukt ist. Dies ist einfach, wobei man aber vereinbaren muss, dass zwei Funktionen als gleich anzusehen sind, wenn sie sich höchstens auf einer Menge vom Maß Null unterscheiden. Beispiel: Es sei

$$\psi(\vec{r}) = \begin{cases} 1 & \text{für } \vec{r} \in \mathbb{Z}^3, \\ 0 & \text{sonst.} \end{cases} \quad (5.51)$$

Es ist

$$\langle \psi, \psi \rangle = \int d^3r \psi^2(\vec{r}) \stackrel{\text{Lebesgue!}}{=} 0. \quad (5.52)$$

Das Axiom $\langle \psi, \psi \rangle = 0 \Leftrightarrow \psi \equiv 0$ ist nur erfüllt, wenn wir ψ als identisch mit der Funktion

$$\tilde{\psi}(\vec{r}) = 0 \quad \forall \vec{r} \quad (5.53)$$

ansehen.

Die nun folgenden Begriffe sind nur notwendig, weil der unitäre Raum $L_2(\mathbb{R}^3)$ keine endliche Dimension hat.

4. Eine Folge x_n von Elementen eines unitären (oder zumindest normierten) Vektorraums heißt *Cauchy-Folge*, wenn zu jedem $\epsilon > 0$ ein $N \in \mathbb{N}$ existiert, so dass gilt

$$\|x_m - x_n\| < \epsilon \quad \forall m, n > N. \quad (5.54)$$

Also konvergieren die Abstände gegen Null für große Indizes.

5. Ein unitärer (oder normierter) Vektorraum V heißt *vollständig*, wenn jede Cauchy-Folge $x_n \in V$ stark gegen ein Element $x \in V$ konvergiert, d. h. wenn zu jedem $\epsilon > 0$ ein $N \in \mathbb{N}$ existiert, so dass gilt

$$\|x_n - x\| < \epsilon \quad \forall n > N. \quad (5.55)$$

Diese Bedingung kann man auch schreiben als

$$\lim_{n \rightarrow \infty} \|x_n - x\| = 0. \quad (5.56)$$

Jede stark konvergierende Folge ist auch Cauchy-Folge, aber die Umkehrung gilt nicht automatisch und man führt daher den zusätzlichen Begriff der Vollständigkeit ein. Er bedeutet, dass V „keine Löcher hat“. Zum Beispiel ist die Menge \mathbb{Q} der rationalen Zahlen (\mathbb{Q} ist zwar ein Vektorraum über sich selbst, aber nicht über \mathbb{R} oder \mathbb{C}) nicht vollständig, denn die Cauchy-Folge

$$\left\{ \left(1 + \frac{1}{n} \right)^n \mid n \in \mathbb{N} \right\} = \left\{ 1, \left(\frac{3}{2} \right)^2, \left(\frac{4}{3} \right)^3, \left(\frac{5}{4} \right)^4, \dots \right\} \quad (5.57)$$

in \mathbb{Q} konvergiert nicht gegen ein Element von \mathbb{Q} . Die Folge konvergiert in \mathbb{R} , nämlich gegen e . Man kann zeigen, dass ein *endlichdimensionaler* unitärer Raum immer vollständig ist.

6. Ein vollständiger, unitärer Raum heißt *Hilbert-Raum*. (Ein vollständiger, normierter Raum heißt *Banach-Raum*. Das ist eine schwächere Definition, jeder Hilbert-Raum ist also auch ein Banach-Raum, aber nicht umgekehrt.)

$L_2(\mathbb{R}^3)$ ist ein Hilbert-Raum. Zum Beweis muss man offenbar noch die Vollständigkeit zeigen. Dies führen wir hier nicht durch. Der Lebesguesche Integralbegriff ist hier notwendig.

7. Für die Quantentheorie benötigen wir noch eine weitere Eigenschaft: Nicht nur soll jede Cauchy-Folge konvergieren (Vollständigkeit), sondern es soll auch eine Folge $F = \{x_n\}$ existieren, die jedem $x \in V$ beliebig nahe kommt. Man sagt, diese Folge ist *dicht* in V . Man kann zeigen, dass diese Eigenschaft äquivalent ist zu der Aussage, dass für jedes $x \in V$ eine Teilfolge von F existiert, die gegen x konvergiert. Existiert eine in V dichte Folge, so nennt man V *separabel*. Separabilität ist eine zusätzliche Forderung über die Hilbert-Raum-Eigenschaften hinaus. In Physik-Lehrbüchern wird jedoch manchmal ein Hilbert-Raum strenger als separabler, vollständiger, unitärer Raum definiert.

$L_2(\mathbb{R}^3)$ ist ein separabler Hilbert-Raum. Den Beweis der Separabilität übergehen wir wieder.

Die Separabilität des Hilbert-Raums V ist wichtig, wie wir jetzt diskutieren werden. Zunächst definieren wir ein *vollständiges Orthonormalsystem* in V als eine Menge $M \subset V$ mit den Eigenschaften

1. $\|x\| = 1 \quad \forall x \in M$ (Normiertheit),
2. $\langle x, y \rangle = 0 \quad \forall x, y \in M$ mit $x \neq y$ (Orthogonalität),
3. es gibt kein $v \in V$, $v \neq 0$, so dass $\langle v, x \rangle = 0 \quad \forall x \in M$. (Vollständigkeit von M – beachte, dass dies nicht derselbe Begriff von Vollständigkeit ist, wie beim Hilbert-Raum V).

Separabilität bedeutet nun, dass eine Folge $F = \{f_1, f_2, \dots\}$ existiert, die jedem $x \in M$ beliebig nahe kommt. F ist, wie jede Folge, abzählbar. Wir geben nun einen Beweis dafür an, dass das vollständige Orthonormalsystem M ebenfalls abzählbar ist.

Beweis (*reductio ad absurdum*): Wir nehmen an, dass M überzählbar ist. Für $x, y \in M$, $x \neq y$, ist

$$\begin{aligned} \|x - y\| &= \sqrt{\langle x - y, x - y \rangle} \\ &= \sqrt{\langle x, x \rangle + \langle y, y \rangle - \underbrace{\langle x, y \rangle}_{=0} - \underbrace{\langle y, x \rangle}_{=0}} \\ &= \sqrt{\|x\|^2 + \|y\|^2} = \sqrt{2}. \end{aligned} \quad (5.58)$$

Zwei verschiedene Elemente von M haben also immer den Abstand $\sqrt{2}$. Betrachte nun Kugeln

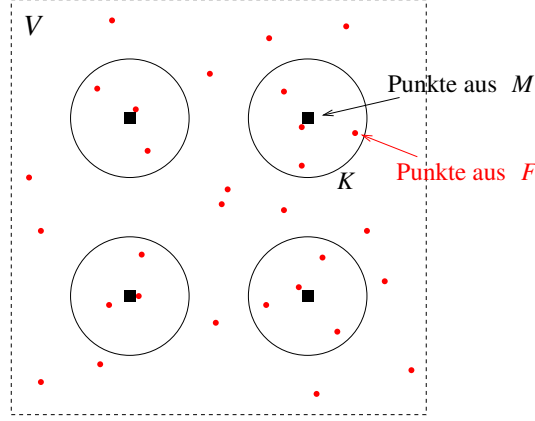
$$K\left(x, \frac{1}{2}\right) := \left\{ z \in V \mid \|z - x\| \leq \frac{1}{2} \right\}, \quad (5.59)$$

zentriert bei den Elementen $x \in M \subset V$. Da der Radius der Kugeln kleiner ist als der Abstand ihrer Mittelpunkte, enthält jede Kugel $K(x, 1/2)$ genau ein Element von M , nämlich x . Die Kugeln sind sogar disjunkt, denn gäbe es ein z mit $z \in K(x, 1/2)$ und $z \in K(y, 1/2)$ mit $x, y \in M$, $x \neq y$, so wäre der Abstand der Mittelpunkte x, y wegen der Dreiecksungleichung

$$\|x - y\| = \|(x - z) - (y - z)\|$$

$$\begin{aligned}
&\leq \|x - z\| + \|(y - z)\| = \|x - z\| + \|y - z\| \\
&= \frac{1}{2} + \frac{1}{2} = 1,
\end{aligned} \tag{5.60}$$

im Widerspruch zu Gl. (5.58). Anschaulich gesprochen: Der Radius der Kugeln ist kleiner als die Hälfte des Abstands der Mittelpunkte.



Da nun F dicht in V ist, enthält jede Kugel mindestens ein Element $f \in F$ (tatsächlich abzählbar unendlich viele). Es gibt genau eine Kugel für jedes $x \in M$ und diese sind disjunkt. Dann ist die Mächtigkeit von F nicht kleiner als die Mächtigkeit von M . M ist aber nach Voraussetzung überabzählbar. Dann ist F ebenfalls überabzählbar. Widerspruch!

Die Separabilität garantiert also die Existenz eines abzählbaren Orthonormalsystems. Es sei noch daran erinnert, dass bei *endlichdimensionalen* Hilbert-Räumen Vollständigkeit und Separabilität automatisch gegeben sind. Für endlichdimensionale Hilbert-Räume existiert auch immer ein endliches Orthonormalsystem.

5.3.2 Entwicklung nach Basisfunktionen

Wir haben gesehen, dass ein separabler Hilbert-Raum, insbesondere der Raum $L_2(\mathbb{R}^3)$, abzählbare, vollständige Orthonormalsysteme besitzt. Jedes davon kann als Orthonormalbasis B des Raumes dienen, d. h. man kann jede Wellenfunktion $\psi \in L_2(\mathbb{R}^3)$ nach den Basiselementen $\varphi_n \in B \subset L_2(\mathbb{R})$ entwickeln. Da ein separabler Hilbert-Raum eine abzählbare Basis besitzt, sagt man, seine Vektorraumdimension sei abzählbar.

Es ist nützlich, die Vollständigkeit der Basis durch die *Vollständigkeitsrelation*

$$\sum_n \varphi_n(\vec{r}) \varphi_n^*(\vec{r}') = \delta(\vec{r} - \vec{r}') \tag{5.61}$$

auszudrücken. Wir beweisen nun diese Relation.

Da δ eine Distribution ist, sind beide Seiten der Relation unter einem Integral zu verstehen. Die Behauptung lautet also: Für alle $f \in L_2(\mathbb{R}^3)$ gilt

$$\int d^3r' \sum_n \varphi_n(\vec{r}) \varphi_n^*(\vec{r}') f(\vec{r}') = \int d^3r' \delta(\vec{r} - \vec{r}') f(\vec{r}'). \tag{5.62}$$

Dies ist äquivalent zu

$$\sum_n \varphi_n(\vec{r}) \underbrace{\int d^3r' \varphi_n^*(\vec{r}') f(\vec{r}')}_{= \langle \varphi_n, f \rangle} = f(\vec{r}). \tag{5.63}$$

Wegen der Vollständigkeit von $L_2(\mathbb{R}^3)$ konvergiert die Summe (Reihe) auf der linken Seite. Es bleibt zu zeigen, dass die Reihe gegen $f(\vec{r})$ konvergiert. Das erfolgt wieder mittels *reductio ad absurdum*: Wir nehmen an, dass die Reihe gegen eine andere Funktion $g(\vec{r})$ konvergiert, also

$$\sum_n \varphi_n(\vec{r}) \langle \varphi_n, f \rangle - f(\vec{r}) =: g(\vec{r}) \neq 0. \tag{5.64}$$

Wegen der Abgeschlossenheit von $L_2(\mathbb{R}^3)$ unter Addition liegt g in $L_2(\mathbb{R}^3)$. Es folgt

$$\sum \underbrace{\langle \varphi_m, \varphi_n \rangle}_{=\delta_{m,n}} \langle \varphi_n, f \rangle - \langle \varphi_m, f \rangle = \langle \varphi_m, g \rangle \quad \forall \varphi_m \in B \quad (5.65)$$

$$\Rightarrow \langle \varphi_m, f \rangle - \langle \varphi_m, f \rangle = 0 = \langle \varphi_m, g \rangle \quad \forall \varphi_m \in B. \quad (5.66)$$

Aber dann ist g orthogonal zu allen φ_m , B ist also kein vollständiges Orthonormalsystem. Widerspruch!

Mit Hilfe der Vollständigkeitsrelation können wir nun leicht eine beliebige Funktion $\varphi \in L_2(\mathbb{R}^3)$ entwickeln. Die Methode, das „Einschieben der Eins“, ist sehr nützlich und wird häufiger vorkommen. Es funktioniert wie folgt:

$$\begin{aligned} \psi(\vec{r}) &= \int d^3r' \delta(\vec{r} - \vec{r}') \psi(\vec{r}') \\ &= \int d^3r' \sum_n \varphi_n(\vec{r}) \varphi_n^*(\vec{r}') \psi(\vec{r}') \\ &= \sum_n \varphi_n(\vec{r}) \underbrace{\int d^3r' \varphi_n^*(\vec{r}') \psi(\vec{r}')}_{=\langle \varphi_n, \psi \rangle} \\ &= \sum_n \langle \varphi_n, \psi \rangle \varphi_n(\vec{r}). \end{aligned} \quad (5.67)$$

Hier sind $\langle \varphi_n, \psi \rangle$ (i. A. komplexe) Zahlen. Sie sind die Entwicklungskoeffizienten von ψ in der Basis $B = \{\varphi_n | n = 1, 2, \dots\}$.

5.3.3 Uneigentliche (Dirac-) Zustände

Es ist oft wünschenswert, Funktionen zu betrachten, die nicht quadratintegrabel, also nicht Elemente von $L_2(\mathbb{R}^3)$, sind. So hat die Schrödingergleichung oft sinnvolle Lösungen, die nicht in $L_2(\mathbb{R}^3)$ liegen. Die freie Schrödingergleichung

$$i\hbar \frac{\partial \psi}{\partial t} = -\frac{\hbar^2}{2m} \nabla^2 \psi \quad (5.68)$$

wird z. B. von ebenen Wellen

$$\psi = \psi_0 e^{i(\vec{k} \cdot \vec{r} - \omega t)} \quad (5.69)$$

gelöst. Diese sind nicht quadratintegrabel:

$$\int d^3r |\psi|^2 = |\psi_0|^2 \int d^3r 1 = \infty. \quad (5.70)$$

Wenn wir den Funktionenraum um ebene Wellen und um alle möglichen Superpositionen erweitern, erhalten wir wieder einen Vektorraum. Aber dessen Elemente haben nun i. A. keine endliche Norm. Wir können daher keine Cauchy-Folgen definieren und der Raum kann daher nicht vollständig sein (also auch kein Hilbert-Raum) und auch nicht separabel. Daher ist die Existenz einer abzählbaren Basis nicht garantiert. In der Tat existiert keine: Die Funktionen $\psi = \psi_0 e^{i\vec{k} \cdot \vec{r}}$ sind für unterschiedliche \vec{k} linear unabhängig. Es gibt aber überabzählbar viele \vec{k} , nämlich alle $\vec{k} \in \mathbb{R}^3$. Daher kann der Funktionenraum keine abzählbare Basis haben.

Man erweitert den Raum der zulässigen „Funktionen“ sogar noch weiter, indem man Distributionen wie $\delta(\vec{r} - \vec{R})$, \vec{R} fest, hinzufügt. Welche mathematischen Objekte sinnvoll sind, wird durch die physikalische Fragestellung bestimmt. Den „sinnvoll“ erweiterten Raum nennt man den *Dirac-Raum*. Man kann aber den Standpunkt verteidigen, dass physikalische Zustände nur durch Wellenfunktionen in $L_2(\mathbb{R}^3)$ beschrieben werden und alle anderen Funktionen nur mathematische Hilfsmittel darstellen, die Herleitungen und Rechnungen erleichtern. Da der beobachtbare Teil des Universums endlich ist und seine quantentheoretische Beschreibung nicht vom unbeobachtbaren Teil abhängen sollte, kann man sogar davon ausgehen, dass der relevante Hilbert-Raum abzählbar

unendliche Dimension hat (siehe Abschnitt 6.2.2). Es wurde sogar argumentiert, seine Dimension sei endlich, wenn auch riesig.

Als physikalisch sinnvoll erweisen sich meist solche nicht quadratintegrablen Funktionen und Distributionen, die immerhin eine verallgemeinerte Orthonormalitätsbedingung und eine verallgemeinerte Vollständigkeitsrelation zulassen. Ein überabzählbares Funktionensystem $\{\varphi_\alpha(\vec{r})\}$ mit α aus einer überabzählbaren Menge, z. B. $\alpha \in \mathbb{R}$, nennen wir ein Orthonormalsystem, wenn gilt

$$\int d^3r \varphi_\alpha^*(\vec{r}) \varphi_\beta(\vec{r}) = \delta(\alpha - \beta). \quad (5.71)$$

Weiter nennen wir das System vollständig, wenn gilt

$$\int d\alpha \varphi_\alpha(\vec{r}) \varphi_\alpha^*(\vec{r}') = \delta(\vec{r} - \vec{r}'). \quad (5.72)$$

Diese Bedingungen werden von ebenen Wellen erfüllt: Sei $\varphi_{\vec{k}}(\vec{r}) = \frac{1}{(2\pi)^{3/2}} e^{i\vec{k} \cdot \vec{r}}$, dann ist

$$\int d^3r \varphi_{\vec{k}}^*(\vec{r}) \varphi_{\vec{k}'}(\vec{r}) = \frac{1}{(2\pi)^3} \underbrace{\int d^3r e^{-i\vec{k} \cdot \vec{r} + i\vec{k}' \cdot \vec{r}}}_{= (2\pi)^3 \delta(\vec{k} - \vec{k}')} = \delta(\vec{k} - \vec{k}'). \quad (5.73)$$

Die hier verwendete Beziehung für das Integral über eine ebene Welle beruht auf der Identität

$$\int_{-\infty}^{\infty} dx e^{ikx} = 2\pi \delta(k). \quad (5.74)$$

Sie gilt „unter einem Integral“, d. h. es ist gemeint, dass für jede integrierbare Testfunktion $\tilde{f}(k)$ gilt

$$\int_{-\infty}^{\infty} dk \tilde{f}(k) \int_{-\infty}^{\infty} dx e^{ikx} = 2\pi \int_{-\infty}^{\infty} dk \tilde{f}(k) \delta(k) = 2\pi \tilde{f}(k=0). \quad (5.75)$$

Diese Beziehung kann man leicht verstehen: Unter geeigneten Konvergenzannahmen ist die linke Seite

$$\begin{aligned} \int_{-\infty}^{\infty} dk \tilde{f}(k) \int_{-\infty}^{\infty} dx e^{ikx} &= 2\pi \int_{-\infty}^{\infty} dx \int_{-\infty}^{\infty} \frac{dk}{2\pi} e^{ikx} \tilde{f}(k) \\ &= 2\pi \int_{-\infty}^{\infty} dx f(x) = 2\pi \int_{-\infty}^{\infty} dx e^{-i0x} f(x) \\ &= 2\pi \tilde{f}(k=0), \end{aligned} \quad (5.76)$$

wobei $f(x)$ die Fourier-(Rück-)Transformierte von $\tilde{f}(k)$ ist.

Weiter gilt analog zu Gl. (5.73)

$$\int d^3k \varphi_{\vec{k}}(\vec{r}) \varphi_{\vec{k}'}^*(\vec{r}') = \frac{1}{(2\pi)^3} \underbrace{\int d^3k e^{i\vec{k} \cdot \vec{r} - i\vec{k}' \cdot \vec{r}'}}_{= (2\pi)^3 \delta(\vec{r} - \vec{r}')} = \delta(\vec{r} - \vec{r}'). \quad (5.77)$$

Die ebenen Wellen sind also in verallgemeinertem Sinn vollständig. Orthonormalität und Vollständigkeit gestatten die Entwicklung nach einem Funktionensystem. Dies ist für ebene Wellen die bekannte Fourier-Transformation.

Im Ortsraum lokalisierte δ -Distributionen bilden ebenfalls ein Orthonormalsystem: Sei $\varphi_{\vec{R}}(\vec{r}) = \delta(\vec{r} - \vec{R})$, dann ist

$$\int d^3r \varphi_{\vec{R}}^*(\vec{r}) \varphi_{\vec{R}'}(\vec{r}) = \int d^3r \delta(\vec{r} - \vec{R}) \delta(\vec{r} - \vec{R}') = \delta(\vec{R} - \vec{R}'). \quad (5.78)$$

Dieses Orthonormalsystem ist auch vollständig:

$$\int d^3R \varphi_{\vec{R}}(\vec{r}) \varphi_{\vec{R}'}^*(\vec{r}') = \int d^3R \delta(\vec{r} - \vec{R}) \delta(\vec{r}' - \vec{R}) = \delta(\vec{r} - \vec{r}'). \quad (5.79)$$

5.3.4 Lokale Erhaltung der Wahrscheinlichkeit

Wir hatten gefordert, dass die Gesamtwahrscheinlichkeit $\int d^3r |\psi|^2$ erhalten ist. Das ist ein globaler Erhaltungssatz. Die durch die Schrödinger-Gleichung beschriebene Dynamik führt jedoch auf eine noch strengere Erhaltung: Für ein Teilchen im Potential $V(\vec{r})$ ist

$$\begin{aligned}
\frac{\partial}{\partial t} |\psi(\vec{r}, t)|^2 &= \left(\frac{\partial}{\partial t} \psi^*(\vec{r}, t) \right) \psi(\vec{r}, t) + \psi^*(\vec{r}, t) \frac{\partial}{\partial t} \psi(\vec{r}, t) \\
&= -\frac{1}{i\hbar} (\hat{H}\psi)^* \psi + \frac{1}{i\hbar} \psi^* \hat{H}\psi \quad \left| \quad \text{da } \hat{H} \text{ hermitesch ist} \right. \\
&= -\frac{1}{i\hbar} \left(-\frac{\hbar^2}{2m} \nabla^2 \psi^* \right) \psi - \frac{1}{i\hbar} \overbrace{(V(\vec{r})\psi^*)} \psi + \frac{1}{i\hbar} \psi^* \left(-\frac{\hbar^2}{2m} \nabla^2 \psi \right) + \frac{1}{i\hbar} \cancel{\psi^* V(\vec{r}) \psi} \\
&= \frac{\hbar}{2mi} [(\nabla^2 \psi^*) \psi - \psi^* \nabla^2 \psi] \\
&= \frac{\hbar}{2mi} \left[\vec{\nabla} \cdot ((\vec{\nabla} \psi^*) \psi) - (\vec{\nabla} \psi^*) \cdot (\vec{\nabla} \psi) - \vec{\nabla} \cdot (\psi^* \vec{\nabla} \psi) + (\vec{\nabla} \psi^*) \cdot (\vec{\nabla} \psi) \right] \\
&= -\vec{\nabla} \cdot \frac{\hbar}{2mi} [\psi^* \vec{\nabla} \psi - (\vec{\nabla} \psi^*) \psi], \tag{5.80}
\end{aligned}$$

also

$$\frac{\partial}{\partial t} |\psi|^2 + \vec{\nabla} \cdot \frac{\hbar}{2mi} [\psi^* \vec{\nabla} \psi - (\vec{\nabla} \psi^*) \psi] = 0. \tag{5.81}$$

Da $\varrho = |\psi|^2$ eine *Dichte* ist, liegt es nahe, den zweiten Term als Divergenz einer *Stromdichte*

$$\vec{j}(\vec{r}, t) := \frac{\hbar}{2mi} [\psi^*(\vec{r}, t) \vec{\nabla} \psi(\vec{r}, t) - (\vec{\nabla} \psi^*(\vec{r}, t)) \psi(\vec{r}, t)] \tag{5.82}$$

zu schreiben. \vec{j} ist die *Wahrscheinlichkeits-Stromdichte*. Damit erhalten wir eine *Kontinuitätsgleichung*

$$\frac{\partial \varrho}{\partial t} + \vec{\nabla} \cdot \vec{j} = 0 \tag{5.83}$$

wie in der Elektrodynamik, aber nun für die Wahrscheinlichkeit anstelle der Ladung. Sie beschreibt die *lokale* Erhaltung der Wahrscheinlichkeit: ϱ kann sich nur dadurch ändern, dass Wahrscheinlichkeit zu- oder abfließt.

5.3.5 Die Impulsdarstellung

Für quadratintegrierbare Funktionen und für die physikalisch relevanten uneigentlichen Zustände können wir die Fourier-Transformierte definieren. Wir drücken sie in der Quantenmechanik oft als Funktion des Impulses $\vec{p} = \hbar \vec{k}$ statt des Wellenvektors \vec{k} aus:

$$\tilde{\psi}(\vec{p}, t) = \int d^3r e^{-i\vec{p} \cdot \vec{r} / \hbar} \psi(\vec{r}, t), \tag{5.84}$$

$$\psi(\vec{r}, t) = \int \frac{d^3p}{(2\pi\hbar)^3} e^{i\vec{p} \cdot \vec{r} / \hbar} \tilde{\psi}(\vec{p}, t) = \int \frac{d^3p}{h^3} e^{i\vec{p} \cdot \vec{r} / \hbar} \tilde{\psi}(\vec{p}, t) \tag{5.85}$$

Die asymmetrische Definition der Vor- und Rücktransformation ist in der Physik üblich, ψ und $\tilde{\psi}$ enthalten dieselbe Information. $\tilde{\psi}(\vec{p}, t)$ heißt *Wellenfunktion im Impulsraum*. Die Möglichkeit äquivalenter Darstellungen im Orts- und Impulsraum ist schon in der klassischen Hamilton-Mechanik angelegt, in der generalisierte Koordinaten und Impulse weitgehend gleichberechtigt auftreten. Ist $\psi(\vec{r}, t)$ normiert,

$$\int d^3r |\psi(\vec{r}, t)|^2 = 1, \tag{5.86}$$

so folgt (Satz von Plancherel)

$$\int \frac{d^3p}{(2\pi\hbar)^3} |\tilde{\psi}(\vec{p}, t)|^2 = \int \frac{d^3p}{(2\pi\hbar)^3} \tilde{\psi}^*(\vec{p}, t) \tilde{\psi}(\vec{p}, t)$$

$$\begin{aligned}
&= \int \frac{d^3 p}{(2\pi\hbar)^3} \int d^3 r e^{i\vec{p}\cdot\vec{r}/\hbar} \psi^*(\vec{r}, t) \int d^3 r' e^{-i\vec{p}\cdot\vec{r}'/\hbar} \psi(\vec{r}', t) \\
&= \int d^3 r d^3 r' \psi^*(\vec{r}, t) \psi(\vec{r}', t) \underbrace{\int \frac{d^3 p}{(2\pi\hbar)^3} e^{i\vec{p}\cdot(\vec{r}-\vec{r}')/\hbar}}_{= \frac{1}{(2\pi\hbar)^3} (2\pi)^3 \delta\left(\frac{\vec{r}-\vec{r}'}{\hbar}\right) = \delta(\vec{r}-\vec{r}')} \\
&= \int d^3 r |\psi(\vec{r}, t)|^2 = 1.
\end{aligned} \tag{5.87}$$

Die Normierung bleibt unter Fourier-Transformation also erhalten. Da $\tilde{\psi}$ normiert und $|\tilde{\psi}|^2$ natürlich nichtnegativ ist, liegt es nahe, $\tilde{\psi}$ als Wahrscheinlichkeitsdichte im Impulsraum zu interpretieren.

Wir können auch die Schrödinger-Gleichung in Impulsdarstellung schreiben:

$$i\hbar \frac{\partial}{\partial t} \tilde{\psi}(\vec{p}, t) = \int d^3 r e^{-i\vec{p}\cdot\vec{r}/\hbar} i\hbar \frac{\partial}{\partial t} \psi(\vec{r}, t) = \int d^3 r e^{-i\vec{p}\cdot\vec{r}/\hbar} \left[-\frac{\hbar^2}{2m} \nabla^2 \psi(\vec{r}, t) + V(\vec{r}) \psi(\vec{r}, t) \right]. \tag{5.88}$$

Durch zweimalige partielle Integration im Term der kinetischen Energie erhalten wir

$$\begin{aligned}
\ldots &= \int d^3 r \left[-\frac{\hbar^2}{2m} \left(-\frac{i\vec{p}}{\hbar} \right)^2 e^{-i\vec{p}\cdot\vec{r}/\hbar} \psi(\vec{r}, t) + V(\vec{r}) e^{-i\vec{p}\cdot\vec{r}/\hbar} \psi(\vec{r}, t) \right] \\
&= \int d^3 r \left[\frac{p^2}{2m} e^{-i\vec{p}\cdot\vec{r}/\hbar} \psi(\vec{r}, t) + V(i\hbar \vec{\nabla}_{\vec{p}}) e^{-i\vec{p}\cdot\vec{r}/\hbar} \psi(\vec{r}, t) \right].
\end{aligned} \tag{5.89}$$

Die letzte Identität verstehen wir am besten, wenn wir $V(\vec{r})$ in eine Potenzreihe entwickeln. Jede Potenz von $i\hbar \vec{\nabla}_{\vec{p}}$ bringt genau die gleiche Potenz von \vec{r} aus dem Exponenten herunter. Jetzt können wir die Integration ausführen, das Ergebnis lautet

$$i\hbar \frac{\partial}{\partial t} \tilde{\psi}(\vec{p}, t) = \frac{p^2}{2m} \tilde{\psi}(\vec{p}, t) + V(i\hbar \vec{\nabla}_{\vec{p}}) \tilde{\psi}(\vec{p}, t). \tag{5.90}$$

In der Impulsdarstellung ist also der Impulsoperator einfach die Multiplikation mit dem Zahlenvektor \vec{p} :

$$\hat{\vec{p}} = \vec{p}. \tag{5.91}$$

Der Ortsoperator ist dagegen nun komplizierter, nämlich i. W. der Gradient nach dem Impuls:

$$\hat{\vec{r}} = i\hbar \vec{\nabla}_{\vec{p}}. \tag{5.92}$$

Beachte die Analogie zum Impulsoperator in Ortsdarstellung, $\hat{p} = -i\hbar \vec{\nabla}$.

Die Impulsdarstellung vereinfacht einerseits die Lösung der Schrödinger-Gleichung für bestimmte Potentiale. Zum Beispiel wird die freie Gleichung zu einer *gewöhnlichen* (nicht partiellen) Differentialgleichung:

$$i\hbar \frac{d\tilde{\psi}}{dt} = \frac{p^2}{2m} \tilde{\psi}. \tag{5.93}$$

Auch für gitterperiodische Potentiale (Kristallgitter) ist die Impulsdarstellung sehr nützlich, was ihre breite Anwendung in der Festkörperphysik begründet. Andererseits erleichtert sie gewisse formale Herleitungen, wie auch der folgende Abschnitt zeigt.

5.4 Erwartungswerte und Schwankungen

Da wir $|\psi(\vec{r}, t)|^2$ als Wahrscheinlichkeitsdichte identifiziert haben, können wir sofort die Ausdrücke für die Erwartungswerte (Mittelwerte) von nur ortsabhängigen Größen $A(\vec{r})$ angeben. In der Quantenmechanik notieren wir Erwartungswerte mit spitzen Klammern:

$$\langle A(\vec{r}) \rangle = \int d^3 r |\psi(\vec{r}, t)|^2 A(\vec{r}) = \int d^3 r \psi^*(\vec{r}, t) A(\vec{r}) \psi(\vec{r}, t). \tag{5.94}$$

Für nur impulsabhängige Größen gehen wir zur Impulsdarstellung über:

$$\langle B(\vec{p}) \rangle = \int \frac{d^3 p}{(2\pi\hbar)^3} |\tilde{\psi}(\vec{p}, t)|^2 B(\vec{p}) = \int \frac{d^3 p}{(2\pi\hbar)^3} \tilde{\psi}^*(\vec{p}, t) B(\vec{p}) \tilde{\psi}(\vec{p}, t). \quad (5.95)$$

Den Erwartungswert $\langle \vec{p} \rangle$ hatten wir bereits in 4.2 ausgerechnet. In leicht veränderter Notation:

$$\begin{aligned} \langle \vec{p} \rangle &= \int \frac{d^3 p}{(2\pi\hbar)^3} \tilde{\psi}^*(\vec{p}, t) \vec{p} \tilde{\psi}(\vec{p}, t) \\ &= \int d^3 r d^3 r' \frac{d^3 p}{(2\pi\hbar)^3} e^{i\vec{p}\cdot\vec{r}/\hbar} \psi^*(\vec{r}, t) \vec{p} e^{-i\vec{p}\cdot\vec{r}'/\hbar} \psi(\vec{r}', t) \\ &= \int d^3 r d^3 r' \frac{d^3 p}{(2\pi\hbar)^3} e^{i\vec{p}\cdot\vec{r}/\hbar} \psi^*(\vec{r}, t) \left(-\frac{\hbar}{i} \right) \vec{\nabla}' e^{-i\vec{p}\cdot\vec{r}'/\hbar} \psi(\vec{r}', t). \end{aligned} \quad (5.96)$$

Die wesentliche Idee hier war, den Impuls \vec{p} durch den Gradienten der Exponentialfunktion darzustellen. Das wird klarer, wenn man die Gleichungskette von hinten nach vorne liest. Mittels partieller Integration in \vec{r}' übertragen wir nun den Gradienten auf den anderen von \vec{r}' abhängenden Faktor,

$$\begin{aligned} \langle \vec{p} \rangle &= \int d^3 r d^3 r' \frac{d^3 p}{(2\pi\hbar)^3} e^{i\vec{p}\cdot\vec{r}/\hbar} \psi^*(\vec{r}, t) e^{-i\vec{p}\cdot\vec{r}'/\hbar} \frac{\hbar}{i} \vec{\nabla}' \psi(\vec{r}', t) \\ &= \int d^3 r d^3 r' \psi^*(\vec{r}, t) \left[\frac{\hbar}{i} \vec{\nabla}' \psi(\vec{r}', t) \right] \underbrace{\int \frac{d^3 p}{(2\pi\hbar)^3} e^{i\vec{p}\cdot(\vec{r}-\vec{r}')/\hbar}}_{\delta(\vec{r}-\vec{r}')} \\ &= \int d^3 r \psi^*(\vec{r}, t) \frac{\hbar}{i} \vec{\nabla} \psi(\vec{r}, t) \\ &= \int d^3 r \psi^*(\vec{r}, t) \hat{\vec{p}} \psi(\vec{r}, t), \end{aligned} \quad (5.97)$$

mit dem Impulsoperator $\hat{\vec{p}} = -i\hbar \vec{\nabla}$ in Ortsdarstellung.

Analog findet man allgemeiner

$$\langle B(\vec{p}) \rangle = \int d^3 r \psi^*(\vec{r}, t) B\left(\frac{\hbar}{i} \vec{\nabla}\right) \psi(\vec{r}, t) = \int d^3 r \psi^*(\vec{r}, t) B(\hat{\vec{p}}) \psi(\vec{r}, t). \quad (5.98)$$

Ganz allgemein schreiben wir für Messgrößen $C(\vec{r}, \vec{p})$:

$$\langle C \rangle = \int d^3 r \psi^*(\vec{r}, t) C\left(\vec{r}, \frac{\hbar}{i} \vec{\nabla}\right) \psi(\vec{r}, t) \quad (5.99)$$

Beachte, dass in der Quantentheorie die Reihenfolge von Faktoren r_i und p_i in C relevant ist, im Unterschied zur klassischen Physik, siehe Abschnitt 5.1.

Mit der Impulsdarstellung erhalten wir auch

$$\langle C \rangle = \int \frac{d^3 p}{(2\pi\hbar)^3} \tilde{\psi}^*(\vec{p}, t) C(i\hbar \vec{\nabla}_{\vec{p}}, \vec{p}) \tilde{\psi}(\vec{p}, t). \quad (5.100)$$

5.4.1 Hermitizität

Da wir immer nur reelle Werte für Messgrößen messen – es ist gar nicht klar, was es bedeuten würde, eine komplexe Größe zu messen – ist es sinnvoll, zu fordern, dass die Erwartungswerte von Messgrößen ebenfalls reell sind. Welche Eigenschaft muss der eine Messgröße darstellende Operator \hat{C} haben, um dies zu garantieren? Es ist

$$\langle \hat{C} \rangle = \int d^3 r \psi^*(\vec{r}, t) \hat{C} \psi(\vec{r}, t), \quad (5.101)$$

woraus folgt

$$\langle \hat{C} \rangle^* = \int d^3r \psi(\vec{r}, t) [\hat{C} \psi(\vec{r}, t)]^* = \int d^3r [\hat{C} \psi(\vec{r}, t)]^* \psi(\vec{r}, t) \quad (5.102)$$

Die Forderung $\langle \hat{C} \rangle^* = \langle \hat{C} \rangle$ ergibt

$$\int d^3r [\hat{C} \psi(\vec{r}, t)]^* \psi(\vec{r}, t) = \int d^3r \psi^*(\vec{r}, t) \hat{C} \psi(\vec{r}, t). \quad (5.103)$$

Dies soll für alle quadratintegrablen $\psi(\vec{r}, t)$ gelten. Ist diese Bedingung erfüllt, so nennt man den Operator \hat{C} *hermitesch*, wie wir schon für den Spezialfall des Hamilton-Operators gesehen hatten.

Als Beispiel betrachten wir den Impulsoperator $\langle p \rangle = \frac{\hbar}{i} \vec{\nabla}$. Es ist

$$\begin{aligned} \int d^3r \left[\frac{\hbar}{i} \vec{\nabla} \psi(\vec{r}, t) \right]^* \psi(\vec{r}, t) &= \int d^3r \left(-\frac{\hbar}{i} \right) [\vec{\nabla} \psi^*(\vec{r}, t)] \psi(\vec{r}, t) \\ &\stackrel{\text{partiell}}{=} \int d^3r \frac{\hbar}{i} \psi^*(\vec{r}, t) \vec{\nabla} \psi(\vec{r}, t) \\ &= \int d^3r \psi^*(\vec{r}, t) \frac{\hbar}{i} \vec{\nabla} \psi(\vec{r}, t) \end{aligned} \quad (5.104)$$

Der Impulsoperator ist also hermitesch. Hermizität zeigt man auch leicht für den Hamilton-Operator $\hat{H} = \hat{p}^2/2m + V(\vec{r})$.

5.4.2 Schwankungen

Ebenso wie Mittelwerte können wir auch höhere Momente der Verteilung von Messwerten ausrechnen, insbesondere deren Schwankungen. Wir werden sehen, dass in der Quantenmechanik solche Schwankungen oder Streuungen von Messwerten aus *prinzipiellen* Gründen auftreten müssen, selbst für ideale Experimente, die keine zusätzlichen Messfehler einführen. Als *mittlere quadratische Schwankung* einer Größe $C(\vec{r}, \vec{p})$ definieren wir

$$\begin{aligned} \Delta C &:= \sqrt{\langle (C - \langle C \rangle)^2 \rangle} = \sqrt{\langle C^2 - 2C\langle C \rangle + \langle C \rangle^2 \rangle} \\ &= \sqrt{\langle C^2 \rangle - 2\langle C \rangle \langle C \rangle + \langle C \rangle^2} \\ &= \sqrt{\langle C^2 \rangle - \langle C \rangle^2}. \end{aligned} \quad (5.105)$$

ΔC^2 nennt man das *Schwankungsquadrat*. Es ist zu beachten, dass $\langle C^2 \rangle$ nicht dasselbe ist wie $\langle C \rangle^2$. Nimmt C z. B. mit gleichen Wahrscheinlichkeiten die Werte ± 1 an, so ist $\langle C^2 \rangle = \langle 1 \rangle = 1$, aber

$$\langle C \rangle^2 = \left(\frac{1}{2} \cdot 1 + \frac{1}{2} \cdot (-1) \right)^2 = 0^2 = 0. \quad (5.106)$$

Beispiel: Gaußsches Wellenpaket in einer Dimension. Wir betrachten die Wellenfunktion, in der Ortsdarstellung,

$$\psi(x, t=0) = \frac{1}{(2\pi)^{1/4} \sqrt{\sigma}} e^{ip_0 x/\hbar} \exp\left(-\frac{(x-x_0)^2}{4\sigma^2}\right). \quad (5.107)$$

Uns interessiert hier nicht die Zeitentwicklung, daher wählen wir eine feste Zeit, $t=0$. Es ist

$$|\psi(x)|^2 = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(x-x_0)^2}{2\sigma^2}\right), \quad (5.108)$$

die Wahrscheinlichkeitsdichte ist also eine auf eins normierte Gauß-Funktion um den Mittelwert x_0 und mit der Breite σ . Das Schwankungsquadrat des Ortes ist

$$\Delta x^2 = (\Delta x)^2 = \langle (x - \langle x \rangle)^2 \rangle = \langle (x - x_0)^2 \rangle$$

$$\begin{aligned}
&= \frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^{\infty} dx (x - x_0)^2 \exp\left(-\frac{(x - x_0)^2}{2\sigma^2}\right) \\
&\stackrel{u=x-x_0}{=} \frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^{\infty} du u^2 \exp\left(-\frac{u^2}{2\sigma^2}\right) = \sigma^2.
\end{aligned} \tag{5.109}$$

Die Schwankung des Impulses kann man mit Hilfe des Impulsoperators in Ortsdarstellung ausrechnen oder unter Verwendung der Wellenfunktion im Impulsraum. Wir wählen hier den letzten Weg:

$$\begin{aligned}
\tilde{\psi}(\vec{p}) &= \int dx e^{-ipx/\hbar} \psi(x) \\
&= \frac{1}{(2\pi)^{(1/4)}\sqrt{\sigma}} \int dx e^{-i(p-p_0)x/\hbar} \exp\left(-\frac{(x - x_0)^2}{4\sigma^2}\right) \\
&= 2^{3/4} \pi^{1/4} \sqrt{\sigma} e^{-i(p-p_0)x_0/\hbar} \exp\left(-\frac{\sigma^2 (p-p_0)^2}{\hbar^2}\right).
\end{aligned} \tag{5.110}$$

Die Fourier-Transformierte einer (normierten) Gauß-Funktion ist wieder eine (normierte) Gauß-Funktion. Der Mittelwert des Impulses ist offenbar p_0 . Das Schwankungsquadrat ist

$$\begin{aligned}
\Delta p^2 &= \langle (p - p_0)^2 \rangle \\
&= 2^{3/2} \sqrt{\pi} \sigma \int \frac{dp}{2\pi\hbar} (p - p_0)^2 \exp\left(-\frac{2\sigma^2 (p - p_0)^2}{\hbar^2}\right) \\
&\stackrel{p'=p-p_0}{=} \sqrt{\frac{2}{\pi}} \frac{\sigma}{\hbar} \int dp' (p')^2 \exp\left(-\frac{2\sigma^2 (p')^2}{\hbar^2}\right) = \frac{\hbar^2}{4\sigma^2}.
\end{aligned} \tag{5.111}$$

Also ist die Fourier-Transformierte $\tilde{\psi}(p)$ umso breiter, je schmaler die ursprüngliche Funktion $\psi(x)$ ist. Das ist eine allgemeine Eigenschaft der Fouriertransformation. Konkret erhalten wir

$$\Delta x \Delta p = \sigma \frac{\hbar}{2\sigma} = \frac{\hbar}{2}. \tag{5.112}$$

Der spezielle Wert $\hbar/2$ beruht auf der Wahl einer Gauß-Funktion für $\psi(x)$.

5.5 Orts-Impuls-Unschärferelation

Wir haben im letzten Abschnitt gesehen, dass für Gauß-Pakete $\Delta x \Delta p = \hbar/2$ gilt. Wir wollen dieses Ergebnis jetzt auf beliebige Wellenfunktionen verallgemeinern. Die Herleitung ist mathematisch etwas aufwendig und beginnt mit der Definition einer Hilfsgröße

$$\begin{aligned}
I(\lambda) &:= \int_{-\infty}^{\infty} dx |(x - \langle x \rangle)\psi(x) + i\lambda(\hat{p} - \langle \hat{p} \rangle)\psi(x)|^2 \\
&= \int_{-\infty}^{\infty} dx \left| (x - \langle x \rangle)\psi(x) + i\lambda \left(\frac{\hbar}{i} \frac{\partial}{\partial x} - \langle \hat{p} \rangle \right) \psi(x) \right|^2.
\end{aligned} \tag{5.113}$$

Aufgrund des Betragsquadrats im Integranden ist $I(\lambda) \geq 0 \forall \lambda$. Wir formen das Integral um:

$$\begin{aligned}
I(\lambda) &= \int_{-\infty}^{\infty} dx \psi^*(x) (x - \langle x \rangle)^2 \psi(x) \\
&\quad + \int_{-\infty}^{\infty} dx \psi^*(x) (x - \langle x \rangle) i\lambda \left(\frac{\hbar}{i} \frac{\partial}{\partial x} - \langle \hat{p} \rangle \right) \psi(x) \\
&\quad - \int_{-\infty}^{\infty} dx i\lambda \left[\left(-\frac{\hbar}{i} \frac{\partial}{\partial x} - \langle \hat{p} \rangle \right) \psi^*(x) \right] (x - \langle x \rangle) \psi(x)
\end{aligned}$$

$$\begin{aligned}
& + \int_{-\infty}^{\infty} dx \lambda \left[\left(-\frac{\hbar}{i} \frac{\partial}{\partial x} - \langle \hat{p} \rangle \right) \psi^*(x) \right] \lambda \left(\frac{\hbar}{i} \frac{\partial}{\partial x} - \langle \hat{p} \rangle \right) \psi(x) \\
& \stackrel{\text{partiell}}{=} \Delta x^2 + i\lambda \int_{-\infty}^{\infty} dx \psi^*(x) (x - \langle x \rangle) \left(\frac{\hbar}{i} \frac{\partial}{\partial x} - \langle \hat{p} \rangle \right) \psi(x) \\
& \quad - i\lambda \int_{-\infty}^{\infty} dx \psi^*(x) \left(\frac{\hbar}{i} \frac{\partial}{\partial x} - \langle \hat{p} \rangle \right) (x - \langle x \rangle) \psi(x) \\
& \quad + \lambda^2 \int_{-\infty}^{\infty} dx \psi^*(x) \left(\frac{\hbar}{i} \frac{\partial}{\partial x} - \langle \hat{p} \rangle \right)^2 \psi(x) \\
& = \Delta x^2 + i\lambda \int_{-\infty}^{\infty} dx \psi^*(x) (x - \langle x \rangle) \left(\frac{\hbar}{i} \frac{\partial}{\partial x} - \langle \hat{p} \rangle \right) \psi(x) \\
& \quad - i\lambda \int_{-\infty}^{\infty} dx \psi^*(x) \frac{\hbar}{i} \psi(x) \\
& \quad - i\lambda \int_{-\infty}^{\infty} dx \psi^*(x) (x - \langle x \rangle) \left(\frac{\hbar}{i} \frac{\partial}{\partial x} - \langle \hat{p} \rangle \right) \psi(x) \\
& \quad + \lambda^2 \Delta p^2 \\
& = \Delta x^2 - \hbar \lambda + \Delta p^2 \lambda^2.
\end{aligned} \tag{5.114}$$

Also ist

$$\Delta x^2 \geq -\Delta p^2 \lambda^2 + \hbar \lambda \quad \forall \lambda. \tag{5.115}$$

Da dies für alle λ gilt, gilt es auch für das λ , welches die rechte Seite maximiert. Diesen Wert von λ erhalten wir aus

$$\frac{d}{d\lambda} (-\Delta p^2 \lambda^2 + \hbar \lambda) = -2\Delta p^2 \lambda + \hbar = 0 \tag{5.116}$$

$$\Rightarrow \lambda = \frac{\hbar}{2\Delta p^2}. \tag{5.117}$$

Damit ist

$$\Delta x^2 \geq -\Delta p^2 \frac{\hbar^2}{4\Delta p^4} + \frac{\hbar^2}{2\Delta p^2} = \frac{\hbar^2}{4\Delta p^2} \tag{5.118}$$

und

$$\Delta x^2 \Delta p^2 \geq \frac{\hbar^2}{4} \tag{5.119}$$

und schließlich

$$\Delta x \Delta p \geq \frac{\hbar}{2}. \tag{5.120}$$

Dies ist die *Heisenbergsche Orts-Impuls-Unschärferelation*. Aus dieser Relation folgt, dass es prinzipiell unmöglich ist, den Ort und den Impuls eines Teilchens zugleich scharf zu messen. Denn dann wären $\Delta x = 0$ und $\Delta p = 0$, im Widerspruch zur Unschärferelation. An der Proportionalität von $\Delta x \Delta p$ zu \hbar erkennt man, dass es sich um eine Folge der Quantenmechanik handelt – der klassische Grenzfall entspricht ja $\hbar \rightarrow 0$.

Wenn wir die Herleitung betrachten, sehen wir, dass das Ergebnis auf der Anwesenheit des Terms $-\hbar \lambda$ in Gleichung (5.114) beruht. Dieser ergab sich aus der Differenz von

$$\int_{-\infty}^{\infty} dx \psi^*(x) (x - \langle x \rangle) (\hat{p} - \langle \hat{p} \rangle) \psi(x) \tag{5.121}$$

und

$$\int_{-\infty}^{\infty} dx \psi^*(x) (\hat{p} - \langle \hat{p} \rangle) (x - \langle x \rangle) \psi(x), \tag{5.122}$$

d. h. daraus, dass Ort und Impuls nicht vertauschen. Es ist nämlich, wie wir in 5.1 gesehen haben,

$$[\hat{x}, \hat{p}] = i\hbar. \tag{5.123}$$

Wir werden auf den Zusammenhang zwischen Kommutatoren und Unschärferelationen in Abschnitt 7.6.2 zurückkommen. Dort wird auch eine elegantere, operatorbasierte Herleitung gegeben, die keine Wellenfunktionen verwendet.

5.6 Messungen I

Die bisher entwickelte Schrödinger-Theorie ist *statistischer* Natur: Sie macht Aussagen über Wahrscheinlichkeitsverteilungen. Die Verteilungsfunktionen können nicht für alle Messgrößen beliebig scharf sein, wie die Orts-Impuls-Unschärferelation zeigt. Wenn wir z. B. den Ort eines Teilchens sicher wissen, haben wir keine Information über seinen Impuls. Der statistische Charakter ist also im Rahmen der Standard-Quantentheorie unvermeidbar. Es stellt sich die Frage, ob diese Unbestimmtheit eine Eigenschaft der realen Welt ist oder eine Schwäche der Standard-Quantentheorie. Mit anderen Worten, welche der folgenden Aussagen ist korrekt?

- Die Wellenfunktion beschreibt den Ort und den Impuls eines Teilchens so genau wie möglich. Das Teilchen hat also zumindest hinsichtlich Ort und Impuls keine Eigenschaften, die nicht durch die Wellenfunktion ausgedrückt werden. Es hat insbesondere wirklich nicht zugleich einen scharfen Ort und einen scharfen Impuls.
- Ein Teilchen hat in Wirklichkeit zu jeder Zeit einen scharfen Ort und einen scharfen Impuls. Die Wellenfunktion enthält also nur eine unvollständige Beschreibung des Teilchens. Da die Standard-Quantentheorie aus Prinzip nicht Teilchen mit scharfem Ort und Impuls beschreiben kann, die aber real existieren, ist sie unvollständig.

Diese Frage ist eng mit der Interpretation des *Messprozesses* verknüpft. Was passiert bei der Messung einer Observablen A an einem System mit gegebener Wellenfunktion ψ ? Nach der statistischen Interpretation ist $|\psi(\vec{r}, t)|^2$ die Wahrscheinlichkeitsdichte eines Teilchens im Ortsraum zur Zeit t . Angenommen, die Wellenfunktion vor der Messung sei nicht perfekt lokalisiert. Zur Zeit $t = 0$ messen wir nun z. B. den Ort des Teilchens mit einer idealen Apparatur. Unmittelbar vor der Messung ist die Wahrscheinlichkeitsdichte $|\psi(\vec{r}, 0)|^2$. Unmittelbar nach der Messung, zur Zeit $t = dt$, wissen wir, wo das Teilchen ist. Sagen wir, wir beobachten das Teilchen am Ort \vec{R} . Dann ist die Wahrscheinlichkeitsdichte zur Zeit $t = dt$ nicht mehr $|\psi(\vec{r}, 0)|^2$, denn $|\psi(\vec{r}, 0)|^2$ drückt aus, dass das Teilchen auch an Orten $\vec{r} \neq \vec{R}$ sein kann, was nicht zutrifft. Wir folgern, dass sich die Wahrscheinlichkeitsdichte und damit die Wellenfunktion $\psi(\vec{r}, t)$ bei der Messung sprunghaft ändern muss. Dies ist kein Artefakt der idealisierten scharfen Messung. Wenn wir $|\psi(\vec{r}, t)|^2$ als Wahrscheinlichkeitsdichte interpretieren, die unsere Kenntnis über das System ausdrückt, muss jeder Änderung der Kenntnis, insbesondere jede erfolgreiche Messung, zu einer Änderung der Wellenfunktion führen.

Diese Folgerung führt auf die eingangs gestellte Frage zurück: Hat die Messung dazu geführt, dass das vorher delokalisierte Teilchen am Ort \vec{R} lokalisiert wird, oder hat sie nur den schon vorher vorhandenen, aber noch unbekannten, Ort \vec{R} des Teilchens enthüllt? Die zweite Interpretation ist sicherlich die natürlichere aus Sicht der klassischen Physik. Sie ist aber nicht haltbar, wie wir noch sehen werden.

Wir wenden uns nun der Messung einer beliebigen Observablen \hat{A} zu. Wir schreiben einen Zirkumflex, da Observable durch Operatoren dargestellt werden. (Ortsmessungen sind für die Diskussion des Messprozesses nicht sehr günstig, weil Wellenfunktionen mit scharfem Ort δ -Distributionen und damit nicht quadratintegrabel sind.) Unmittelbar vor der Messung sei die Wellenfunktion $\psi(\vec{r}, 0)$. Bei der Messung von \hat{A} finden wir einen Messwert A (wir betrachten hier nur solche Messungen, die den Messwert eindeutig bestimmen, nicht nur einschränken). Durch die Messung muss sich die Wellenfunktion so geändert haben, dass die Observable \hat{A} nun den scharfen Wert A hat. Das bedeutet, dass der Mittelwert $\langle \hat{A} \rangle$ mit A übereinstimmen und die Schwankung ΔA verschwinden muss.

Der Mittelwert der Observablen ist (wir lassen das Zeitargument $t = dt$ weg)

$$\langle \hat{A} \rangle = \int d^3r \psi^*(\vec{r}) \hat{A} \psi(\vec{r}) \stackrel{!}{=} A. \quad (5.124)$$

Das Schwankungsquadrat ist

$$\Delta A^2 = \int d^3r \psi^*(\vec{r}) (\hat{A} - A)^2 \psi(\vec{r}). \quad (5.125)$$

Als Observable muss \hat{A} hermitesch sein. Dann ist auch $\hat{A} - A$ hermitesch, weil die Multiplikation mit einer reellen Konstanten A hermitesch ist. Also folgt

$$\begin{aligned}\Delta A^2 &= \int d^3r [(\hat{A} - A) \psi(\vec{r})]^* (\hat{A} - A) \psi(\vec{r}) \\ &= \int d^3r |(\hat{A} - A) \psi(\vec{r})|^2 \\ &= \|(\hat{A} - A) \psi(\vec{r})\|^2 \stackrel{!}{=} 0.\end{aligned}\tag{5.126}$$

Der letzte Ausdruck ist aber positiv semidefinit und Null genau dann, wenn gilt

$$(\hat{A} - A) \psi(\vec{r}) = 0 \quad \forall \vec{r},\tag{5.127}$$

also wenn

$$\hat{A} \psi(\vec{r}) = A \psi(\vec{r}) \quad \forall \vec{r}.\tag{5.128}$$

Die Wellenfunktion muss also unmittelbar nach einer Messung mit dem Messwert A diese Gleichung erfüllen.

Die hergeleitete Gleichung für schwankungsfreie Wellenfunktionen,

$$\hat{A} \psi_A(\vec{r}) = A \psi_A(\vec{r}),\tag{5.129}$$

ist eine *Eigenwertgleichung* für den linearen Operator \hat{A} auf $L_2(\mathbb{R}^3)$. A ist der *Eigenwert* zur *Eigenfunktion* $\psi_A(\vec{r})$. Diese Begriffe hängen eng mit den Eigenvektoren und Eigenwerten von Matrizen zusammen. Wir schreiben den Eigenwert A als Subskript von ψ_A , um anzudeuten, dass diese Gleichung eben nur für die Eigenfunktion zum Eigenwert A und nicht für beliebige Wellenfunktionen gilt. Als Ergebnis einer Messung können sich nur Werte ergeben, die verschwindende Schwankung ΔA zulassen, also Eigenwerte. Nach der Messung ist die Wellenfunktion dann eine Eigenfunktion von \hat{A} zum gemessenen Eigenwert.

Zusammengefasst haben wir folgendes gefunden:

- Bei der Messung einer Observablen \hat{A} geht die Wellenfunktion sprunghaft in eine Eigenfunktion von \hat{A} über („Kollaps der Wellenfunktion“).
- Der Messwert ist der zu dieser Eigenfunktion gehörende Eigenwert.

Diese Aussagen bilden das *von Neumannsche Projektionspostulat*. Es ist in der Tat ein zusätzliches Postulat, weil es nicht aus der durch die Schrödinger-Gleichung beschriebenen Zeitentwicklung folgt, wie wir sehen werden.

Da $L_2(\mathbb{R}^3)$ ein separabler Hilbert-Raum ist, existiert eine abzählbare Orthonormalbasis. Es existieren sogar unendlich viele solche Basen. Man kann nun zeigen, dass die Eigenfunktionen eines beliebigen hermiteschen Operators \hat{A} auf $L_2(\mathbb{R}^3)$ eine solche Basis bilden. (Wir betrachten zunächst nur den Fall, dass die Eigenwerte nicht *entartet* sind, d. h., dass es zu einem A nicht mehrere Eigenfunktionen ψ_A gibt.) Das hat wichtige Konsequenzen.

Betrachten wir wieder den Erwartungswert von \hat{A} :

$$\langle \hat{A} \rangle = \int d^3r \psi^*(\vec{r}) \hat{A} \psi(\vec{r}).\tag{5.130}$$

Die Vollständigkeit der Basis der Eigenfunktionen impliziert, dass gilt

$$\sum_A \psi_A(\vec{r}) \psi_A^*(\vec{r}') = \delta(\vec{r} - \vec{r}').\tag{5.131}$$

Einschieben dieser Vollständigkeitsrelation in Gl. (5.130) ergibt

$$\begin{aligned}\langle \hat{A} \rangle &= \int d^3r d^3r' \psi^*(\vec{r}) \hat{A} \delta(\vec{r} - \vec{r}') \psi(\vec{r}') \\ &= \int d^3r d^3r' \psi^*(\vec{r}) \hat{A} \sum_A \psi_A(\vec{r}) \psi_A^*(\vec{r}') \psi(\vec{r}') \quad \Bigg| \quad \hat{A} \text{ wirkt nur auf } \vec{r}'\end{aligned}$$

$$\begin{aligned}
&= \sum_A \int d^3r d^3r' \psi^*(\vec{r}) A \psi_A(\vec{r}) \psi_A^*(\vec{r}') \psi(\vec{r}') \quad \Big| \text{ mit Gl. (5.129)} \\
&= \sum_A A \int d^3r \psi^*(\vec{r}) \psi_A(\vec{r}) \int d^3r' \psi_A^*(\vec{r}') \psi(\vec{r}') \\
&= \sum_A A \langle \psi, \psi_A \rangle \langle \psi_A, \psi \rangle \\
&= \sum_A |\langle \psi_A, \psi \rangle|^2 A.
\end{aligned} \tag{5.132}$$

Der Erwartungswert ist also das gewichtete Mittel der Eigenwerte A mit Gewichten $|\langle \psi_A, \psi \rangle|^2$.

Was ist die Bedeutung des Erwartungswertes für Messungen? Da als Messwerte nur Eigenwerte auftreten können, der Erwartungswert aber kein Eigenwert sein muss (er liegt i. A. zwischen den diskreten Eigenwerten), kann er allgemein nicht das „erwartete Messergebnis“ beschreiben. Wenn wir die Interpretation von $|\psi|^2$ als Verteilungsfunktion ernst nehmen und wenn Gl. (5.130) für die Bestimmung von Erwartungswerten korrekt ist, sollte $\langle \hat{A} \rangle$ den Mittelwert der Messwerte für wiederholte Messungen an gleich präparierten Systemen darstellen. Also: Führen wir N Messungen von \hat{A} aus und erhalten jeweils die Messwerte A_n , so sollte gelten

$$\frac{A_1 + A_2 + \dots + A_N}{N} \rightarrow \langle \hat{A} \rangle \quad \text{für } N \rightarrow \infty \tag{5.133}$$

mit Wahrscheinlichkeit eins. Dann ist $\langle \hat{A} \rangle$ tatsächlich der Erwartungswert der Messgröße \hat{A} , wenn wir die Messung als Zufallsexperiment auffassen. Die Bezeichnung ist dann gerechtfertigt. Ganz allgemein können wir schreiben

$$\langle \hat{A} \rangle = \sum_A p_A A, \tag{5.134}$$

wobei p_A mit $\sum_A p_A = 1$ die Wahrscheinlichkeit des Auftretens des Messergebnisses A ist. Vergleich mit Gl. (5.132) zeigt sofort, dass gilt

$$p_A = |\langle \psi_A, \psi \rangle|^2. \tag{5.135}$$

Diese Beziehung nennt man die *Bornsche Regel*. Sie folgt aber weder aus der Schrödinger-Gleichung, noch aus dem von Neumannschen Projektionspostulat. Letzteres sagt uns nur, welche Messwerte auftreten können und in welchem Zustand ein System nach der Messung ist. Die Bornsche Regel ist damit ein weiteres Postulat der modernen Quantenmechanik. Konzeptionell sagt dieses Postulat aus, dass die Wahrscheinlichkeitsinterpretation zutrifft und wie man Wahrscheinlichkeiten für Ergebnisse von Messungen ausrechnen soll.

Kann man die Bornsche Regel experimentell nachprüfen? Sicherlich nicht für eine einzelne Messung, da es sich um eine statistische Aussage handelt. Aber zusammen mit der üblichen frequentistischen Interpretation der Wahrscheinlichkeit ergibt die Bornsche Regel eine Vorhersage für den Fall vieler Messungen: Angenommen, wir führen $N \gg 1$ Messungen von \hat{A} unter denselben Bedingungen durch und messen dabei N_A -mal den Wert A . Dann erwarten wir, dass der Anteil N_A/N von Messungen des Wertes A für große N gegen p_A konvergiert:

$$\lim_{N \rightarrow \infty} \frac{N_A}{N} = p_A. \tag{5.136}$$

Für die Interpretationsdebatte der Quantenmechanik ist aber wichtig, dass diese Aussagen auch nur „mit Wahrscheinlichkeit eins“ gilt. Es *kann* auch passieren, dass bei einer solchen Versuchsreihe N_A/N gegen einen anderen Wert strebt, nur wird dies für wachsende N immer seltener passieren.

Die Bornsche Regel ist im beschriebenen Sinne sehr gut bestätigt. Tatsächlich gibt es ein Experiment, dass mit bloßem Auge und ohne besondere Geräte funktioniert: Nach Born sind vier eng zusammenstehende Kerzen in 200 m Entfernung genauso hell wie eine Kerze in 100 m Entfernung. Im Detail: Wir können, wie man in der Feldtheorie sieht, das klassische elektromagnetische Feld als Wellenfunktion der Photonen interpretieren (genauer ist es das Vierervektorpotential (A^μ)). Das Analogon zur Schrödinger-Gleichung ist die Wellengleichung für das Feld. Für eine Punktquelle sagt sie voraus, dass die Amplitude des Feldes wie $1/r$ mit dem Abstand r abfällt. Aus klassischer Sicht muss das so sein, weil dann die Intensität mit $1/r^2$ abfällt. Die Oberfläche von gedachten

Kugeln des Radius r um die Punktquelle nimmt natürlich wie r^2 zu. Daher ist der Energiefluss pro Zeit durch solche Kugeloberflächen unabhängig vom Radius, wie es aufgrund der Energieerhaltung sein muss.

Wie sieht es nun quantentheoretisch aus? Das Sehen ist eigentlich eine Ortsmessung von Photonen. Wir nehmen Licht wahr, wenn ein Photon von einem Rezeptormolekül in der Netzhaut des Auges absorbiert wird. Das geschieht mit einer gewissen Wahrscheinlichkeit p . Nach Born ist dieser Wahrscheinlichkeit proportional zum Betragsquadrat des Feldes und damit für eine Punktquelle zu $1/r^2$. Würden wir eine andere Abhängigkeit (einschließlich der Möglichkeit, dass p gar nicht von r abhängt) beobachten, so wäre die Bornsche Regel experimentell widerlegt. Die Bornsche Regel sagt also voraus, dass die Wahrscheinlichkeit, Photonen einer Punktquelle wahrzunehmen, um den Faktor $1/4$ abnimmt, wenn die Quelle doppelt so weit entfernt ist. Die wesentliche Idee ist nun, dass mehrere, sagen wir n , gleichartige Lichtquellen (etwa Kerzenflammen oder gleiche LEDs), die so nah benachbart sind, dass das Auge sie nicht einzeln auflösen kann, wirken wie eine Lichtquelle, die n -mal so viele Photonen aussendet. Wenn wir nun eine Punktquelle im Abstand l betrachten und vergleichen mit vier nah benachbarte Punktequellen derselben Art im Abstand $2l$, so sollten nach Born über hinreichend lange Zeit ($N \gg 1$) im Mittel gleich viele Photonen der beiden Anordnungen vom Auge wahrgenommen werden, d. h., sie sollten gleich hell erscheinen. Dazu muss man nicht lange warten, weil sehr viele Photonen wahrgenommen werden.

5.7 Die zeitunabhängige Schrödinger-Gleichung

Wir beschränken uns hier auf den Fall, dass der Hamilton-Operator nicht explizit von der Zeit abhängt. Dann können wir die Schrödinger-Gleichung

$$i\hbar \frac{\partial}{\partial t} \psi(\vec{r}, t) = \hat{H} \psi(\vec{r}, t) \quad (5.137)$$

mittels eines *Separationsansatzes* in eine einfachere Form überführen. Wir machen den Ansatz

$$\psi(\vec{r}, t) = \psi(\vec{r}) f(t). \quad (5.138)$$

Dann folgt

$$i\hbar \psi(\vec{r}) \frac{df}{dt} = [\hat{H} \psi(\vec{r})] f(t). \quad (5.139)$$

Wir teilen durch $\psi(\vec{r}, t)$:

$$i\hbar \frac{\dot{f}(t)}{f(t)} = \frac{\hat{H} \psi(\vec{r})}{\psi(\vec{r})}. \quad (5.140)$$

Wir sollten nachträglich prüfen, was an Punkten mit $\psi(\vec{r}, t) = 0$ geschieht, an denen wir diese Division nicht ausführen können. Nun hängt die linke Seite der Gleichung nicht von \vec{r} ab und die rechte Seite nicht von t , aber beide Seiten sind für alle \vec{r} und t gleich. Daher hängt diese (gleiche) Größe weder von \vec{r} noch von t ab, ist also eine Konstante, die *Separationskonstante* genannt wird. Diese bezeichnen wir hier mit E . Dann gilt

$$i\hbar \frac{\dot{f}(t)}{f(t)} = E, \quad (5.141)$$

$$\frac{\hat{H} \psi(\vec{r})}{\psi(\vec{r})} = E. \quad (5.142)$$

Es folgen die Gleichungen

$$i\hbar \frac{df}{dt} = E f(t), \quad (5.143)$$

$$\hat{H} \psi(\vec{r}) = E \psi(\vec{r}). \quad (5.144)$$

E hat offenbar die Dimension einer Energie. Gleichung (5.143) hat die einzige linear unabhängige Lösung $e^{-iEt/\hbar}$ und entsprechend die allgemeine Lösung

$$f(t) = f_0 e^{-iEt/\hbar}. \quad (5.145)$$

Dies gilt für jede vorgegebene komplexe Zahl E . Aus physikalischen Gründen muss E aber reell sein, da $f(t)$ sonst einen exponentiell anwachsenden oder abfallenden Faktor enthielte, was mit der für alle Zeiten t geforderten Normierungsbedingung $\int d^3r |\psi|^2 = 1$ unvereinbar wäre. Abgesehen davon, schränkt Gleichung (5.143) die möglichen Werte für E nicht ein. Gleichung (5.144) ist die *zeitunabhängige Schrödinger-Gleichung*:

$$\hat{H} \psi_n(\vec{r}) = E_n \psi_n(\vec{r}). \quad (5.146)$$

Dies ist eine *Eigenwertgleichung* für die Eigenwerte (hier *Eigenenergien*) E_n der Hamilton-Operatoren \hat{H} und die zugehörigen Eigenfunktionen $\psi_n(\vec{r})$. Da die Energie reell sein muss – einerseits als Messgröße und andererseits, damit die Lösung der zeitabhängigen Schrödinger-Gleichung normierbar bleibt – ist es sinnvoll anzunehmen, dass der Hamilton-Operator \hat{H} hermitesch ist. (Nicht hermitesche Hamiltonians werden manchmal auch betrachtet, z. B. für offene Systeme.) Es kommt oft vor, dass Eigenwerte *entartet* sind, d. h., dass ein Wert E in der Folge E_1, E_2, \dots mehrfach vorkommt.

Wie im letzten Abschnitt besprochen, kann man aus den Eigenfunktionen eines hermiteschen Operators eine Orthonormalbasis bilden. Daher können wir jede quadratintegrierbare Funktion nach diesen Eigenfunktionen entwickeln. Das ist besonders nützlich für die Lösung der *zeitabhängigen* Schrödinger-Gleichung: Die speziellen Lösungen dieser Gleichung können wir aus den Lösungen der separierten Gleichungen für $\psi(\vec{r})$ und $f(t)$ zusammensetzen, sie lauten

$$\psi_n(\vec{r}, t) = \psi_n(\vec{r}) e^{-iE_n t/\hbar}. \quad (5.147)$$

Die allgemeine Lösung erhalten wir durch Superposition der speziellen Lösungen zu

$$\psi(\vec{r}, t) = \sum_n a_n \psi_n(\vec{r}) e^{-iE_n t/\hbar} \quad (5.148)$$

mit beliebigen Koeffizienten $a_n \in \mathbb{C}$ (wenn $\psi(\vec{r}, t)$ normiert sein soll, stellt dies eine Bedingung an die a_n).

Eine typische Aufgabenstellung besteht darin, für ein gegebenes System, also bekannten Hamilton-Operator \hat{H} , $\psi(\vec{r}, t)$ für $t > 0$ zu finden, wenn $\psi(\vec{r}, 0)$ als Anfangsbedingung vorgegeben ist. Das System soll also zu einer Zeit in einem bestimmten Zustand, beschrieben durch eine Wellenfunktion, präpariert werden und wir interessieren uns für die zeitliche Entwicklung zu späteren Zeiten. Zur Lösung eines solchen Anfangswertproblems zerlegen wir $\psi(\vec{r}, 0)$ in Eigenfunktionen von \hat{H} :

$$\psi(\vec{r}, 0) = \sum_n a_n \psi_n(\vec{r}). \quad (5.149)$$

Wir finden die Koeffizienten a_n unter Ausnutzung der Orthonormalität der Eigenfunktionen:

$$\int d^3r \psi_n^*(\vec{r}) \psi(\vec{r}, 0) = \sum_{n'} a_{n'} \int d^3r \psi_n^*(\vec{r}) \psi_{n'}(\vec{r}) = \sum_{n'} a_{n'} \delta_{nn'} = a_n, \quad (5.150)$$

also

$$a_n = \int d^3r \psi_n^*(\vec{r}) \psi(\vec{r}, 0) = \langle \psi_n, \psi(t=0) \rangle. \quad (5.151)$$

Die Lösung für beliebige Zeiten ist also

$$\psi(\vec{r}, t) = \sum_n \langle \psi_n, \psi(t=0) \rangle \psi_n(\vec{r}) e^{-iE_n t/\hbar}. \quad (5.152)$$

Beachte, dass dieser im Prinzip einfache Lösungsweg nur funktioniert, weil die Schrödinger-Gleichung linear ist.

Beispiel: Wir betrachten einen eindimensionalen Kasten der Länge L mit undurchdringlichen Wänden bei $x = 0$ und $x = L$. Es ist plausibel und wird später genauer begründet, dass dann $\psi(x, t)$ an den Rändern verschwinden muss. Das Potential $V(\vec{r})$ verschwindet in diesem Fall im Inneren des Kastens und der Hamilton-Operator hat die Form

$$\hat{H} = -\frac{\hbar^2}{2m} \frac{d^2}{dx^2}. \quad (5.153)$$

Zur Zeit $t = 0$ soll die Wellenfunktion

$$\psi(x, 0) = c \sin^3 \frac{\pi x}{L} \quad (5.154)$$

vorgegeben sein, wobei c eine Normierungskonstante ist (man berechnet $c = 4/\sqrt{5L}$). Die Funktion $\psi(x, 0)$ erfüllt offenbar die Randbedingungen bei $x = 0$ und $x = L$.

Schritt 1: Eigenfunktionen $\psi_n(\vec{r})$ und Eigenenergien E_n . Die zeitunabhängige Schrödinger-Gleichung lautet

$$-\frac{\hbar^2}{2m} \frac{d^2}{dx^2} \psi_n(x) = E_n \psi_n(x) \quad (5.155)$$

mit $\psi_n(0) = \psi_n(L) = 0$. Das können wir auch schreiben als

$$\psi_n''(x) = -\frac{2mE_n}{\hbar^2} \psi_n(x). \quad (5.156)$$

Ansatz:

$$\psi_n(x) = Ae^{ikx} + Be^{-ikx} \quad (5.157)$$

$$\Rightarrow \psi_n''(x) = -Ak^2 e^{ikx} - Bk^2 e^{-ikx} = -k^2 \psi_n(x) \quad (5.158)$$

$$\Rightarrow k^2 = \frac{2mE_n}{\hbar^2} \quad (5.159)$$

$$\Rightarrow E_n = \frac{\hbar^2 k^2}{2m}. \quad (5.160)$$

Randbedingungen:

$$\psi_n(0) = A + B = 0 \quad (5.161)$$

und

$$\psi_n(L) = Ae^{ikL} + Be^{-ikL} = 0 \quad (5.162)$$

$$\Rightarrow B = -A \quad \text{und} \quad \psi_n(L) = A(e^{ikL} - e^{-ikL}) = 2iA \sin kL = 0 \quad (5.163)$$

$$\Rightarrow kL = n\pi, \quad n \in \mathbb{N}. \quad (5.164)$$

Daher lauten die Eigenenergien

$$E_n = \frac{\pi^2 \hbar^2}{2mL^2} n^2, \quad n \in \mathbb{N} \quad (5.165)$$

und die dazugehörigen Eigenfunktionen

$$\psi_n(x) = c_n \sin \frac{n\pi x}{L}. \quad (5.166)$$

c_n sind Normierungskonstanten. Wir finden

$$\int_0^L dx |\psi_n(x)|^2 = |c_n|^2 \int_0^L dx \sin^2 \frac{n\pi x}{L} = |c_n|^2 \frac{L}{2} \stackrel{!}{=} 1 \quad (5.167)$$

$$\Rightarrow |c_n|^2 = \frac{2}{L}. \quad (5.168)$$

Wir wählen $c_n = \sqrt{2/L}$, also

$$\psi_n(x) = \sqrt{\frac{2}{L}} \sin \frac{n\pi x}{L}. \quad (5.169)$$

Schritt 2: Entwicklung der Anfangsbedingung.

$$\begin{aligned} a_n &= \int_0^L dx \psi_n^*(x) \psi(x, 0) \\ &= \int_0^L dx \sqrt{\frac{2}{L}} \sin \left(\frac{n\pi x}{L} \right) \frac{4}{\sqrt{5L}} \sin^3 \frac{\pi x}{L} \end{aligned}$$

$$\begin{aligned}
& \stackrel{\text{Tafel}}{=} \frac{4\sqrt{2}}{\sqrt{5}L} \cdot \begin{cases} \frac{3L}{8} & \text{für } n = 1 \\ -\frac{L}{8} & \text{für } n = 3 \\ 0 & \text{sonst} \end{cases} \\
& = \begin{cases} \frac{3}{\sqrt{10}} & \text{für } n = 1 \\ -\frac{1}{\sqrt{10}} & \text{für } n = 3 \\ 0 & \text{sonst.} \end{cases}
\end{aligned} \tag{5.170}$$

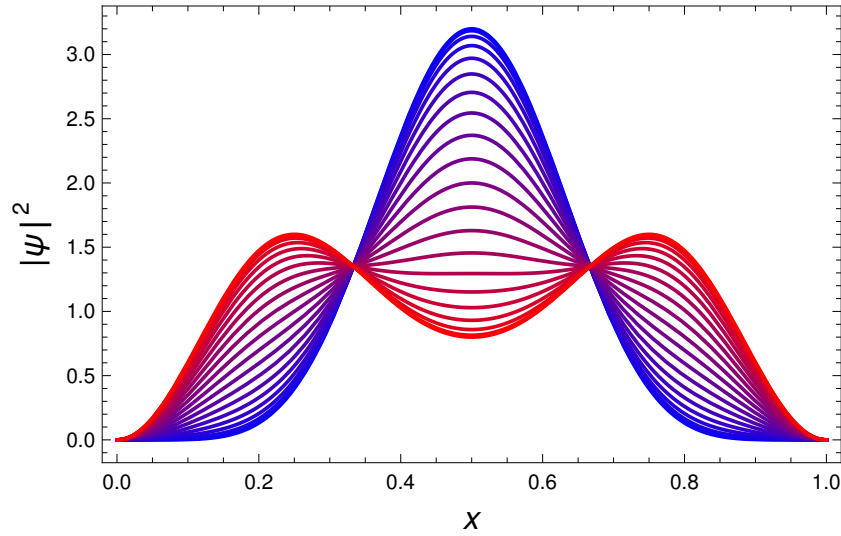
Also ist

$$\psi(x, 0) = \frac{3}{\sqrt{10}} \psi_1(x) - \frac{1}{\sqrt{10}} \psi_3(x). \tag{5.171}$$

Schritt 3: Zusammensetzen der Lösung. Für alle t folgt

$$\begin{aligned}
\psi(x, t) &= \frac{3}{\sqrt{10}} \psi_1(x) \exp\left(-i \frac{\pi^2 \hbar}{2mL^2} t\right) - \frac{1}{\sqrt{10}} \psi_3(x) \exp\left(-i \frac{9\pi^2 \hbar}{2mL^2} t\right) \\
&= \frac{3}{\sqrt{5}L} \sin \frac{\pi x}{L} \exp\left(-i \frac{\pi^2 \hbar}{2mL^2} t\right) - \frac{1}{\sqrt{5}L} \sin \frac{3\pi x}{L} \exp\left(-i \frac{9\pi^2 \hbar}{2mL^2} t\right).
\end{aligned} \tag{5.172}$$

Hier ist $|\psi(x, t)|^2$ für $L = 1$ dargestellt, wobei die Zeit t von 0 (blau) bis $m/4\pi\hbar = m/2h$ (rot) fortschreitet:



Kapitel 6

Quantensysteme in einer Dimension

In diesem Kapitel untersuchen wir die Bewegung eines Teilchens in einem äußeren Potential in einer Dimension. Dies ist auch für realistische Situationen in drei Dimensionen interessant, da sich viele (aber nicht alle) höherdimensionale Probleme auf eindimensionale reduzieren lassen und da viele der allgemeinen Einsichten unabhängig von der Dimension sind.

6.1 Allgemeine Eigenschaften

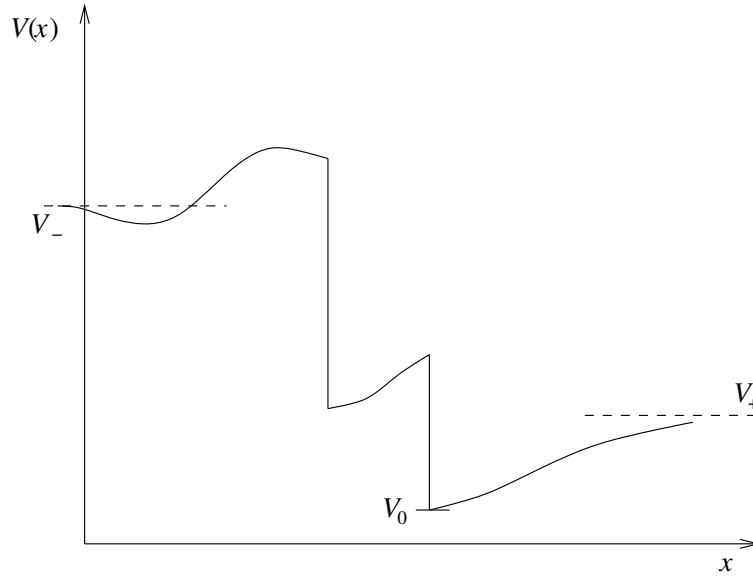
Wir beginnen mit der Diskussion allgemeiner Eigenschaften der eindimensionalen zeitunabhängigen Schrödinger-Gleichung für ein Teilchen in einem Potential,

$$-\frac{\hbar^2}{2m} \frac{d^2\psi}{dx^2} + V(x)\psi(x) = E\psi(x), \quad (6.1)$$

mit der Nebenbedingung, dass $|\psi(x)|^2$ für $x \in \mathbb{R}$ beschränkt ist. Diese Bedingung ist natürlich schwächer als Quadratintegrabilität, erlaubt aber die Behandlung von stationären Streuzuständen, die auf \mathbb{R} nicht quadratintegrabel sind. Da die hier untersuchte Schrödinger-Gleichung nur reelle Koeffizienten enthält, reicht es aus, reelle Lösungen zu betrachten; alle komplexen Lösungen ergeben sich bei Bedarf einfach durch Multiplikation mit beliebigen $c \in \mathbb{C}$. Allerdings ist es manchmal, gerade bei Streuzuständen, nützlich, komplexe Funktionen als Lösungen anzusetzen. (Die Schrödinger-Gleichung enthält nicht mehr nur reelle Koeffizienten, wenn sich das Teilchen in einem Magnetfeld bewegt.)

Die Resultate in diesem Abschnitt beruhen auf verschiedenen Sätzen aus der Theorie gewöhnlicher Differentialgleichungen, die wir hier nicht im Detail besprechen oder beweisen. Für das Potential $V(x)$ nehmen wir *zunächst* folgendes an:

1. $V(x)$ sei beschränkt von unten. Dies gestattet die Definition des *Infimums* $V_0 := \inf V(x)$ des Potentials.
2. $V(x)$ sei stückweise stetig und alle Unstetigkeiten seien Sprünge.
3. $\lim_{x \rightarrow \infty} V(x) =: V_+$ und $\lim_{x \rightarrow -\infty} V(x) =: V_-$ existieren als reelle Zahlen oder $+\infty$. V_+ und V_- müssen nicht gleich sein. Annahme 3 schließt z. B. oszillierende Potentiale der Art $V(x) = v \sin \kappa x$ aus.



Wir erwarten, dass $\psi(x)$ zumindest stetig ist. Da $V(x)$ stetig bis auf endliche Sprünge ist, gilt dies auch für

$$\psi''(x) = -\frac{2m}{\hbar^2} [E - V(x)] \psi(x). \quad (6.2)$$

Daher ist ψ'' integrierbar. Somit existiert ψ' und ist stetig. Schließlich existiert dann auch ψ und ist – wie angenommen – stetig. Wir finden also, dass die Wellenfunktion und ihre erste räumliche Ableitung unter den gegebenen Voraussetzungen stetig sein müssen, insbesondere auch an den Sprüngen von $V(x)$. Das gilt nicht mehr, wenn ein Sprung unendlich hoch wird, wie wir sehen werden.

6.1.1 Klassisch verbotene und erlaubte Bereiche

In der klassischen Mechanik kann sich ein Teilchen nur in Bereichen befinden, in denen $E \geq V(x)$ gilt. Denn für $E < V(x)$ wäre die kinetische Energie $T = E - V(x) < 0$, was klassisch unmöglich ist. Sind klassisch erlaubte Bereiche durch klassisch verbotene Bereiche getrennt, so kann ein Teilchen klassisch betrachtet niemals von einem erlaubten Bereich zum anderen gelangen.

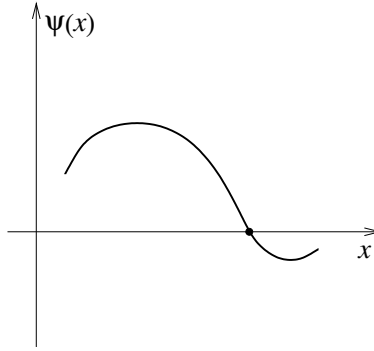
In der Quantenmechanik ist die Situation anders. Für $E > V(x)$ (klassisch erlaubter Bereich) haben wir

$$\psi''(x) = \underbrace{-\frac{2m}{\hbar^2} [E - V(x)]}_{< 0} \psi(x), \quad (6.3)$$

also $\psi''(x) = 0$ dann und nur dann, wenn $\psi(x) = 0$, und sonst

$$\frac{\psi''(x)}{\psi(x)} = -\frac{2m}{\hbar^2} [E - V(x)] < 0 \quad (6.4)$$

(wie gesagt können wir reelle Lösungen $\psi(x)$ annehmen). Damit ist der Graph von $\psi(x)$ immer zur x -Achse hin gekrümmt, typisch ist daher *oszillierendes* Verhalten.



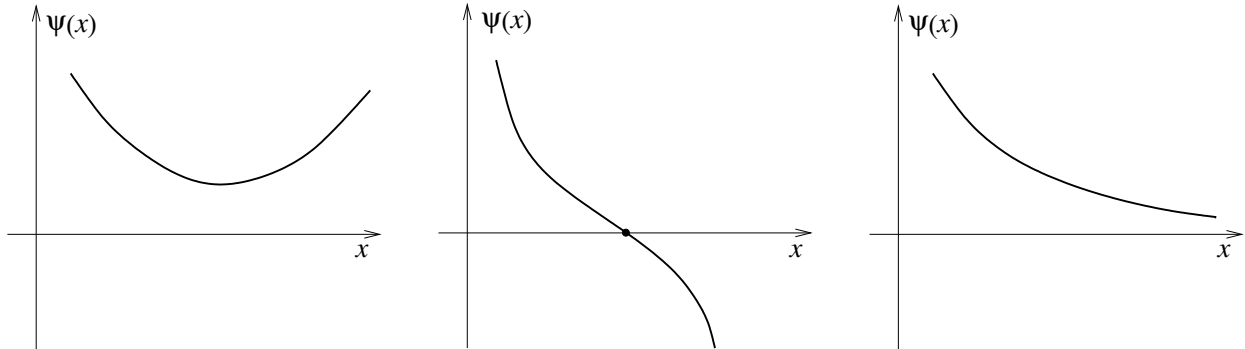
Für $E < V(x)$ (klassisch verbotener Bereich) haben wir stattdessen

$$\psi''(x) = - \underbrace{\frac{2m}{\hbar^2} [E - V(x)]}_{> 0} \psi(x), \quad (6.5)$$

also wieder $\psi''(x) = 0 \Leftrightarrow \psi(x) = 0$, aber andernfalls nun

$$\frac{\psi''(x)}{\psi(x)} = - \frac{2m}{\hbar^2} [E - V(x)] > 0. \quad (6.6)$$

Der Graph von $\psi(x)$ ist immer von der x -Achse weg gekrümmt. Typisch sind zumindest ein Maximum des Betrages $|\psi(x)|$ am Rand des Bereichs und, für ein unbeschränktes Intervall, asymptotisch verschwindendes $\psi(x)$.



Insbesondere erhalten wir i. A. auch in klassisch verbotenen Bereichen eine nicht verschwindende Wahrscheinlichkeitsdichte $|\psi(x)|^2$.

6.1.2 Das Spektrum

Wir untersuchen nun das *Spektrum* des Hamilton-Operators. Das Spektrum eines Operators ist definiert als (typischerweise geordnete) Menge aller seiner Eigenwerte.

Zunächst betrachten wir das Verhalten der Lösungen für $x \rightarrow -\infty$, wenn diese x im klassisch verbotenen Bereich liegen. Das ist der Fall für $E < V_-$. Wir zeigen jetzt, dass, wenn ψ beschränkt ist, sogar $\lim_{x \rightarrow \pm\infty} \psi(x) = 0$ gelten muss. Beweis: Sei $\tilde{V}_- := \frac{E+V_-}{2}$, so dass $E < \tilde{V}_- < V_-$. Es existiert ein x_- , so dass

$$V(x) > \tilde{V}_- > E \quad \forall x < x_-. \quad (6.7)$$

Es folgt

$$\frac{\psi''(x)}{\psi(x)} = - \frac{2m}{\hbar^2} [E - V(x)] = \frac{2m}{\hbar^2} [V(x) - E] > \frac{2m}{\hbar^2} (\tilde{V}_- - E) > 0 \quad \forall x < x_-. \quad (6.8)$$

Wir führen zur Abkürzung

$$c := \frac{2m}{\hbar^2} (\tilde{V}_- - E) > 0 \quad (6.9)$$

ein. Dann ist

$$\frac{\psi''(x)}{\psi(x)} > c > 0 \quad \forall x < x_-. \quad (6.10)$$

Es gilt

$$\psi''(x) = \frac{2m}{\hbar^2} [V(x) - E] \psi(x) \quad (6.11)$$

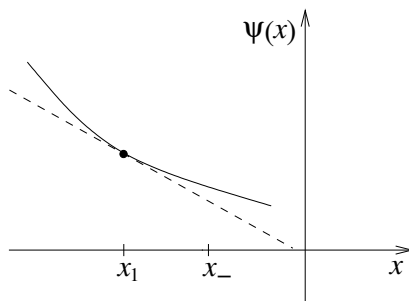
mit

$$\frac{2m}{\hbar^2} [V(x) - E] > c > 0 \quad \forall x < x_-. \quad (6.12)$$

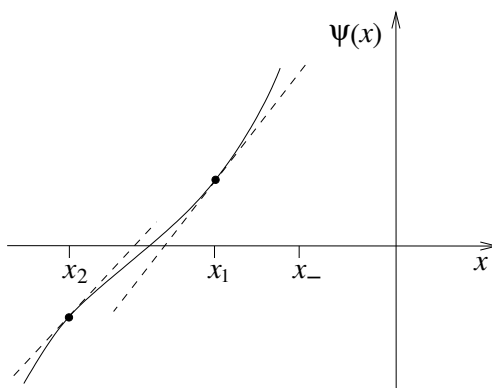
Falls $\psi(x) = 0 \quad \forall x < x_-$ gilt natürlich $\lim_{x \rightarrow -\infty} \psi(x) = 0$. Ansonsten existiert ein $x_1 < x_-$ mit $\psi(x_1) \neq 0$. Wir nehmen o. B. d. A. $\psi(x_1) > 0$ an. Wegen der Krümmung von $\psi(x)$ weg von der x -Achse muss gelten

$$\psi(x) > \psi(x_1) + \psi'(x_1)(x - x_1) \quad (6.13)$$

für $x < x_-$ und $x \neq x_1$, solange $\psi(x) > 0$ ist.



Ist nun $\psi'(x_1) < 0$, so divergiert $\psi(x)$ für $x \rightarrow -\infty$, im Widerspruch zum hier betrachteten Fall einer beschränkten Funktion $\psi(x)$. Für $\psi'(x_1) = 0$ divergiert $\psi(x)$ für $x \rightarrow -\infty$ ebenfalls, wegen $\psi''(x) > c\psi(x) > c\psi(x_1) > 0 \quad \forall x < x_1$. Es bleibt der Fall $\psi'(x_1) > 0$.



Angenommen, $\psi(x)$ wird für $x < x_1$ irgendwo negativ, d. h. es gibt ein $x_2 < x_1$ mit $\psi(x_2) < 0$. Es ist sicher $\psi'(x_2) > 0$, denn andernfalls könnte (nach einem analogen Argument wie oben), nicht $\psi(x_1) > 0$ sein. Aber dann folgt (wieder analog), dass $\psi(x)$ für $x \rightarrow -\infty$ nach $-\infty$ divergiert, im Widerspruch zur Annahme. Daher muss für beschränktes $\psi(x)$ gelten $\psi(x) \geq 0 \quad \forall x < x_1$ und damit $\psi''(x) \geq 0$. Es gilt dann auch $\psi'(x) \geq 0 \quad \forall x < x_1$, denn sonst wäre wieder $\lim_{x \rightarrow -\infty} \psi(x) = \infty$. Also ist ψ für $x < x_1$ monoton wachsend.

Da $\psi(x)$ von unten beschränkt und monoton wachsend ist, folgt mit der Vollständigkeit von \mathbb{R} , dass $\psi(x)$ für $x \rightarrow -\infty$ konvergiert. (Nämlich gegen das Infimum von $\psi(x)$ auf $x \in]-\infty, x_1[$.) Angenommen, $\psi_- := \lim_{x \rightarrow -\infty} \psi(x) > 0$. Dann existiert ein $x_3 < x_1$, so dass

$$\psi(x) > \frac{\psi_-}{2} \quad \forall x < x_3. \quad (6.14)$$

Es folgt

$$\psi''x > c\psi(x) > \frac{c\psi_-}{2} > 0 \quad \forall x < x_3. \quad (6.15)$$

Andererseits folgt aus der Konvergenz von $\psi(x)$, dass

$$\lim_{x \rightarrow -\infty} \psi''(x) = 0. \quad (6.16)$$

Widerspruch! Analog zeigt man die Aussage für $x \rightarrow +\infty$.

Nun können wir verschiedene Fälle unterscheiden:

1. Für Energien $E < V_0$ existieren keine beschränkten Lösungen. Denn für $E < V_0$ gilt

$$\frac{\psi''(x)}{\psi(x)} = -\frac{2m}{\hbar^2} [E - V(x)] = \frac{2m}{\hbar^2} [V(x) - E] \geq \frac{2m}{\hbar^2} (V_0 - E) > 0 \quad \forall x. \quad (6.17)$$

Nach dem eben bewiesenen Theorem ist

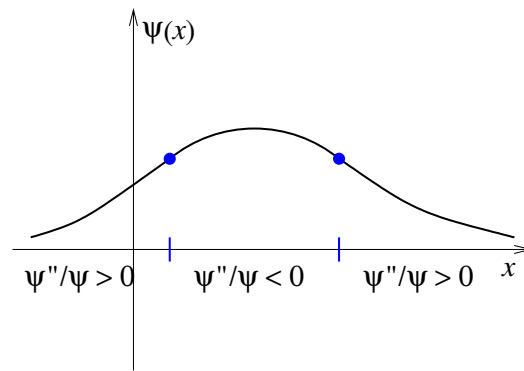
$$\lim_{x \rightarrow -\infty} \psi(x) = 0 \quad \wedge \quad \lim_{x \rightarrow +\infty} \psi(x) = 0. \quad (6.18)$$

Wenn nicht $\psi \equiv 0$ gilt, existiert ein x_1 mit $\psi(x_1) \neq 0$, o. B. d. A. $\psi(x_1) > 0$. Dann ist $\psi''(x_1) > 0$. $\psi(x)$ hat Wendepunkte nur an Nullstellen (falls vorhanden). Daher ist $\psi(x)$ für $x \geq x_1$ monoton wachsend oder für $x \leq x_1$ monoton fallend. Die Aussage (6.18) ist dann nicht erfüllbar. Die einzige Lösung ist $\psi \equiv 0$, so dass Gl. (6.5) durch $0 = 0$ trivial erfüllt ist und man die Ungleichung (6.17) nicht herleiten kann. Dies ist keine zulässige Wellenfunktion.

2. Für $V_0 < E < V_+, V_-$ müssen alle beschränkten Lösungen ebenfalls

$$\lim_{x \rightarrow -\infty} \psi(x) = 0 \quad \wedge \quad \lim_{x \rightarrow +\infty} \psi(x) = 0 \quad (6.19)$$

erfüllen. Nicht verschwindende Lösungen können auftreten, weil wegen $V_0 \leq E$ mindestens ein Bereich mit $\psi''(x)/\psi(x) < 0$ existiert.



Aber die *zwei* Randbedingungen (6.19) führen dazu, dass i. A. nur für eine diskrete Menge $\{E_n\}$ von Eigenenergien beschränkte Lösungen existieren. Dies wird durch das folgende, nicht rigorose, Argument untermauert: Die Schrödinger-Gleichung ist linear und zweiter Ordnung und hat daher zwei linear unabhängige spezielle Lösungen. Die allgemeine Lösung lautet also

$$\psi(x) = \lambda_1 \psi_1(x) + \lambda_2 \psi_2(x), \quad \lambda_1, \lambda_2 \in \mathbb{C}. \quad (6.20)$$

Für ψ_1 und ψ_2 können i. A. *unbeschränkte* Funktionen gewählt werden. Das ist plausibel, da die Graphen von ψ_1 und ψ_2 für hinreichend große $|x|$ von der x -Achse weg gekrümmt sind. Wir können die unbeschränkten Lösungen z. B. erhalten, indem wir für einen Punkt x_0 mit

$$V(x) > E \quad \forall x > x_0 \quad (6.21)$$

(existiert wegen $\lim_{x \rightarrow \infty} V(x) = V_+ > E$) Randbedingungen (a) $\psi_1(x_0) = \psi_0 > 0 \wedge \psi'_1(x_0) = \psi'_0 > 0$ sowie (b) $\psi_2(x) = \psi_0 > 0 \wedge \psi'_2(x_0) = -\psi'_0 < 0$ fordern. ψ_1 ist dann für $x \rightarrow \infty$ unbeschränkt. ψ_2 ist i. A. ebenfalls unbeschränkt und falls nicht, ersetzen wir ψ_2 durch $\psi_2 + \psi_1$, was unbeschränkt und weiterhin linear unabhängig ist. Für $x \rightarrow -\infty$ sind ψ_1 und ψ_2 generisch unbeschränkt und wenn nicht, können wir i. A. wieder durch Superposition beide unbeschränkt machen.

Die Randbedingung

$$\lim_{x \rightarrow -\infty} \psi(x) = 0 \quad (6.22)$$

erfordert, da ψ_1 und ψ_2 für $x \rightarrow -\infty$ unbeschränkt sind, dass

$$\lim_{x \rightarrow -\infty} \frac{\psi_1(x)}{\psi_2(x)} \quad (6.23)$$

existiert und

$$\frac{\lambda_2}{\lambda_1} = - \lim_{x \rightarrow -\infty} \frac{\psi_1(x)}{\psi_2(x)} \quad (6.24)$$

ist. Dies reduziert die Zahl der freien Parameter auf einen, sagen wir λ_1 :

$$\psi(x) = \lambda_1 \left[\psi_1(x) - \left(\lim_{x' \rightarrow -\infty} \frac{\psi_1(x')}{\psi_2(x')} \right) \psi_2(x) \right]. \quad (6.25)$$

λ_1 hat offenbar die Bedeutung eines Normierungsfaktors. Es muss $\lambda_1 \neq 0$ gelten, um eine nichttriviale Lösung zu bekommen. Die zweite Randbedingung

$$\lim_{x \rightarrow \infty} \psi(x) = 0 \quad (6.26)$$

ergibt

$$\lambda_1 \lim_{x \rightarrow \infty} \left[\psi_1(x) - \left(\lim_{x' \rightarrow -\infty} \frac{\psi_1(x')}{\psi_2(x')} \right) \psi_2(x) \right] = 0 \quad (6.27)$$

$$\stackrel{\lambda_1 \neq 0}{\Rightarrow} \lim_{x \rightarrow \infty} \left[\psi_1(x) - \left(\lim_{x' \rightarrow -\infty} \frac{\psi_1(x')}{\psi_2(x')} \right) \psi_2(x) \right] = 0. \quad (6.28)$$

Diese Bedingung hat keine Auswirkung auf λ_1 ! Also kann Gl. (6.26) nicht durch geeignete Wahl von λ_1 erfüllt werden. Der Grenzwert in der letzten Gleichung ist nun eine Funktion des Parameters E in der Schrödinger-Gleichung. Da es eine reelle Gleichung für eine reelle Größe E ist, erwarten wir generisch eine diskrete Lösungsmenge $\{E_n\}$.

Wir haben damit gefunden, dass für $V_0 < E < V_+, V_-$ generisch ein *diskretes Spektrum* existiert. Es kann auch leer sein. Die Eigenfunktionen gehen für $x \pm \infty$ gegen Null und beschreiben daher Zustände, die *gebunden* sind. (Man kann zeigen, dass die Lösungen sogar quadratintegrabel sind.)

3. Für $\min(V_+, V_-) < E < \max(V_+, V_-)$ gehen beschränkte Lösungen entweder für $x \rightarrow -\infty$ (falls $V_- > V_+$) oder für $x \rightarrow +\infty$ (falls $V_- < V_+$) gegen Null. Dies ergibt *eine* Randbedingung, die einen der Koeffizienten λ_1, λ_2 festlegt.

In der anderen Richtung, o. B. d. A. für $x \rightarrow \infty$, ist die Energie größer als das Potential $V(x)$ (klassisch erlaubter Bereich). Genauer existiert ein x_1 , so dass gilt

$$E > V(x) \quad \forall x > x_1. \quad (6.29)$$

Nun konvergiert $V(x)$ für $x \rightarrow \infty$ nach Voraussetzung gegen V_+ . Für „natürliche“ Potentiale $V(x)$ geht dann $V'(x)$ für $x \rightarrow \infty$ gegen Null und $V(x)$ ist für hinreichend große x langsam veränderlich. Dann sind die Voraussetzungen der geometrischen Optik erfüllt und $\psi(x)$ sollte für große x gut durch die Lösungen der mechanischen Eikonal-Gleichung genähert werden:

$$\psi(x, t) \approx \psi_0 \exp\left(\frac{i}{\hbar} [W(x) - Et]\right) \quad (6.30)$$

mit

$$\left(\frac{dW}{dx}\right)^2 = 2m[E - V(x)], \quad (6.31)$$

vgl. Abschnitt 4.3. Es folgt

$$\frac{dW}{dx} = \sqrt{2m[E - V(x)]} \quad (6.32)$$

$$\Rightarrow W(x) = \int_{x_1}^x dx' \sqrt{2m[E - V(x')]} \quad (6.33)$$

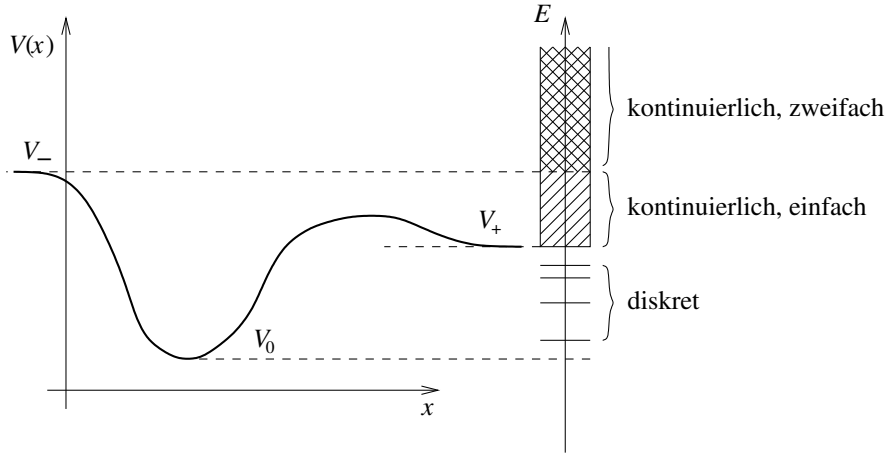
$$\Rightarrow \psi(x, t = 0) \approx \psi_0 \exp\left(\frac{i}{\hbar} \int_{x_1}^x dx' \sqrt{2m[E - V(x')]} \right). \quad (6.34)$$

Wir sind hier an reellen Lösungen interessiert. Diese erfüllen

$$\psi(x, t = 0) \approx |\psi_0| \sin\left(\alpha + \frac{1}{\hbar} \int_{x_1}^x dx' \sqrt{2m[E - V(x')]} \right). \quad (6.35)$$

mit einer konstanten Phase α . $\psi(x)$ ist also insbesondere beschränkt für $x \rightarrow \infty$ (aber nicht konvergent). Die Forderung der Beschränktheit liefert also keine weitere Bedingung, die die möglichen Eigenenergien einschränken könnte. Für jede Energie E mit $\min(V_+, V_-) < E < \max(V_+, V_-)$ existiert also *eine* linear unabhängige Lösung. Das Spektrum ist *kontinuierlich* und *nicht entartet*. Dieser Fall tritt natürlich nur auf, wenn $V_+ \neq V_-$ ist.

4. Für $E > \max(V_+, V_-)$ haben wir keine Randbedingungen für $\psi(x)$ für $x \rightarrow \pm\infty$. Analog zu 3. erwartet man, dass $\psi(x)$ für $x \rightarrow \pm\infty$ für gutartige $V(x)$ automatisch beschränkt ist. Wir erhalten also keine Einschränkung der freien Parameter in der allgemeinen Lösung. Es gibt daher für jede Energie E in diesem Bereich *zwei* linear unabhängige Lösungen. Das Spektrum ist *kontinuierlich* und *zweifach entartet*. Dieser Fall tritt natürlich nur auf, wenn $V_+, V_- < \infty$ sind.



Für Potentiale mit den angenommenen Eigenschaften ist das kontinuierliche Spektrum, falls es existiert, lückenlos. Es ist also das Intervall $[\min(V_+, V_-), +\infty[$. Das gilt nicht mehr, wenn $\lim_{x \rightarrow \pm\infty} V(x)$ nicht existiert, z. B. weil $V(x)$ periodisch ist. In diesem Fall kann es Lücken im kontinuierlichen Spektrum geben („Bandlücken“). Ein wichtiges Beispiel sind Elektronen im periodischen Potential der Atomkerne in Kristallen. In den Bandlücken können diskrete Eigenenergien existieren.

6.2 Rechteckpotentiale

Wir betrachten als Beispiel zunächst Potentiale, die bis auf Sprünge konstant sind, sogenannte *Rechteckpotentiale*. Für diese können wir die Schrödinger-Gleichung im Prinzip einfach lösen: in den einzelnen Bereichen $n = 1, 2, \dots$ mit konstantem Potential ist die Schrödinger-Gleichung von der Form

$$\psi''(x) = c_n \psi(x) \quad (6.36)$$

mit Konstanten c_n und an den Sprüngen sind ψ und ψ' stetig (siehe den Anfang von Abschnitt 6.1), also gelten die *Anschlussbedingungen*

$$\psi(x - \varepsilon) = \psi(x + \varepsilon) \quad \text{für } \varepsilon \rightarrow 0^+, \quad (6.37)$$

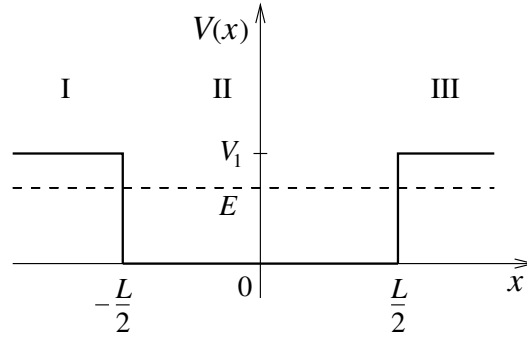
$$\psi'(x - \varepsilon) = \psi'(x + \varepsilon) \quad \text{für } \varepsilon \rightarrow 0^+. \quad (6.38)$$

6.2.1 Kasten endlicher Tiefe

Sei

$$V(x) = \begin{cases} 0 & \text{für } -\frac{L}{2} \leq x \leq \frac{L}{2} \\ V_1 > 0 & \text{sonst.} \end{cases} \quad (6.39)$$

Es ist $V_+ = V_- = V_1$ und $V_0 = 0$. Daher existiert nach Abschnitt 6.1.2 ein zweifach entartetes kontinuierliches Spektrum für $E > V_1$ und evtl. ein diskretes Spektrum im Bereich $0 < E < V_1$. Wir betrachten zunächst den letzteren Fall.



Es liegen drei Bereiche I, II, III vor, wobei Bereiche I und III klassisch verboten sind. In I und III gilt

$$\psi''(x) = -\underbrace{\frac{2m}{\hbar^2} [E - V_1]}_{> 0} \psi(x) =: \kappa^2 \psi(x), \quad (6.40)$$

wobei $\kappa = \sqrt{2m(V_1 - E)}/\hbar$ ist. Allgemeine Lösung:

$$\psi_{\text{I}}(x) = A_{\text{I}} e^{\kappa x} + B_{\text{I}} e^{-\kappa x}, \quad (6.41)$$

$$\psi_{\text{III}}(x) = A_{\text{III}} e^{\kappa x} + B_{\text{III}} e^{-\kappa x}. \quad (6.42)$$

Damit $\psi(x)$ beschränkt bleibt, muss

$$B_{\text{I}} = A_{\text{III}} = 0 \quad (6.43)$$

gelten.

In II (klassisch erlaubt) gilt

$$\psi''(x) = -\underbrace{\frac{2m}{\hbar^2} E}_{< 0} \psi(x) =: -k^2 \psi(x) \quad (6.44)$$

mit $k = \sqrt{2mE}/\hbar$. Allgemeine Lösung:

$$\psi_{\text{II}}(x) = C_{\text{II}} \cos kx + D_{\text{II}} \sin kx. \quad (6.45)$$

Die Anschlussbedingungen ergeben

$$\begin{aligned} \psi_{\text{I}}(-L/2) &= A_{\text{I}} e^{-\kappa L/2} \stackrel{!}{=} \psi_{\text{II}}(-L/2) \\ &= C_{\text{II}} \cos \frac{kL}{2} - D_{\text{II}} \sin \frac{kL}{2}, \end{aligned} \quad (6.46)$$

$$\begin{aligned} \psi'_{\text{I}}(-L/2) &= A_{\text{I}} \kappa e^{-\kappa L/2} \stackrel{!}{=} \psi'_{\text{II}}(-L/2) \\ &= C_{\text{II}} k \sin \frac{kL}{2} + D_{\text{II}} k \cos \frac{kL}{2}, \end{aligned} \quad (6.47)$$

$$\begin{aligned} \psi_{\text{III}}(L/2) &= B_{\text{III}} e^{-\kappa L/2} \stackrel{!}{=} \psi_{\text{II}}(L/2) \\ &= C_{\text{II}} \cos \frac{kL}{2} + D_{\text{II}} \sin \frac{kL}{2}, \end{aligned} \quad (6.48)$$

$$\begin{aligned} \psi'_{\text{III}}(L/2) &= -B_{\text{III}} \kappa e^{-\kappa L/2} \stackrel{!}{=} \psi'_{\text{II}}(L/2) \\ &= -C_{\text{II}} k \sin \frac{kL}{2} + D_{\text{II}} k \cos \frac{kL}{2}. \end{aligned} \quad (6.49)$$

Dies sind 4 lineare Gleichungen für die 4 Unbekannten $A_{\text{I}}, C_{\text{II}}, D_{\text{II}}, B_{\text{III}}$, jedoch sind sie nicht unabhängig. Wir finden durch Addition und Subtraktion

$$(A_{\text{I}} + B_{\text{III}}) e^{-\kappa L/2} = 2C_{\text{II}} \cos \frac{kL}{2}, \quad (6.50)$$

$$(A_{\text{I}} - B_{\text{III}}) e^{-\kappa L/2} = -2D_{\text{II}} \sin \frac{kL}{2}, \quad (6.51)$$

$$(A_{\text{I}} - B_{\text{III}}) \kappa e^{-\kappa L/2} = 2D_{\text{II}} k \cos \frac{kL}{2}, \quad (6.52)$$

$$(A_{\text{I}} + B_{\text{III}}) \kappa e^{-\kappa L/2} = 2C_{\text{II}} k \sin \frac{kL}{2}. \quad (6.53)$$

Es folgt

$$\cancel{2} C_{\text{II}} \kappa \cos \frac{kL}{2} = \cancel{2} C_{\text{II}} k \sin \frac{kL}{2} \quad (6.54)$$

$$-\cancel{2} D_{\text{II}} \kappa \sin \frac{kL}{2} = \cancel{2} D_{\text{II}} k \cos \frac{kL}{2}. \quad (6.55)$$

Nun kann nicht $C_{\text{II}} = D_{\text{II}} = 0$ gelten, da sonst auch $A_{\text{I}} = B_{\text{III}}$ folgen würde und damit $\psi \equiv 0$. Also muss zumindest einer der Parameter $C_{\text{II}}, D_{\text{II}}$ von Null verschieden sein.

Ist $C_{\text{II}} \neq 0$, so folgt

$$\frac{\sin \frac{kL}{2}}{\cos \frac{kL}{2}} = \tan \frac{kL}{2} = \frac{\kappa}{k} \quad (6.56)$$

und

$$-D_{\text{II}} \kappa = D_{\text{II}} k \frac{\cos \frac{kL}{2}}{\sin \frac{kL}{2}} = D_{\text{II}} k \frac{k}{\kappa} \quad (6.57)$$

$$\Rightarrow D_{\text{II}}(\kappa^2 + k^2) = 0 \quad (6.58)$$

$$\Rightarrow D_{\text{II}} = 0. \quad (6.59)$$

Ist dagegen $D_{II} \neq 0$, so folgt analog

$$\tan \frac{kL}{2} = -\frac{k}{\kappa} \quad (6.60)$$

und $C_{II} = 0$.

Somit existieren Lösungen für alle $E \in [0, V_1[$, die

$$\tan \frac{\sqrt{2mE}L}{2\hbar} = \sqrt{\frac{V_1 - E}{E}} \quad (6.61)$$

oder

$$\tan \frac{\sqrt{2mE}L}{2\hbar} = -\sqrt{\frac{E}{V_1 - E}} \quad (6.62)$$

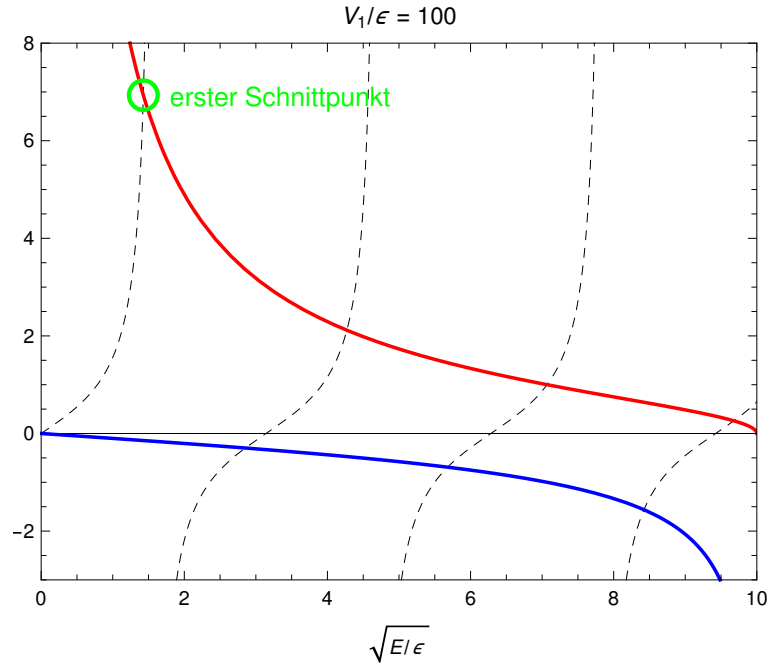
erfüllen. Diese Gleichungen sind nicht analytisch lösbar, sondern nur numerisch oder graphisch. Dazu führen wir die Energieeinheit

$$\varepsilon := \frac{2\hbar^2}{mL^2} \quad (6.63)$$

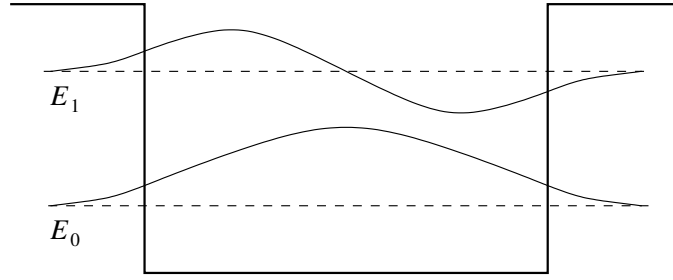
ein und schreiben die Gleichungen als

$$\tan \sqrt{\frac{E}{\varepsilon}} = \sqrt{\frac{V_1/\varepsilon - E/\varepsilon}{E/\varepsilon}} \quad \text{bzw.} \quad \tan \sqrt{\frac{E}{\varepsilon}} = -\sqrt{\frac{E/\varepsilon}{V_1/\varepsilon - E/\varepsilon}}. \quad (6.64)$$

Wir zeichnen beide Seiten der beiden Gleichungen als Funktionen von $\sqrt{E/\varepsilon}$ und suchen die Schnittpunkte. Dabei ist die mathematisch korrekte Lösung $E = 0$ der zweiten Gleichung physikalisch nicht sinnvoll, da sie zu $k = 0$, $\psi_{II} \equiv 0$ und damit schließlich zu $\psi \equiv 0$ führt.



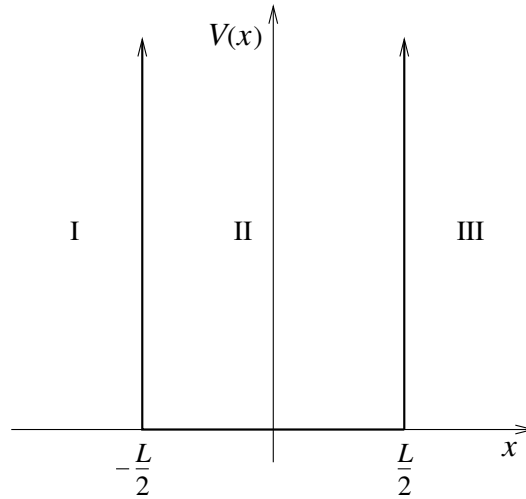
Wir erkennen, dass die niedrigste Eigenenergie eine Lösung der ersten Gleichung ist und für *alle* $V_1/\varepsilon > 0$ existiert. Es gibt also immer mindestens einen gebundenen Zustand. Existiert mehr als eine Eigenenergie, so wechseln sich Lösungen der ersten und zweiten Gleichung ab. Diese entsprechen geraden ($D_{II} = 0$) bzw. ungeraden ($C_{II} = 0$) Funktionen $\psi(x)$. Die Lösungen in Bereichen I und III, d. h. die Koeffizienten A_I und B_{III} , ergeben sich aus den Anschlussbedingungen. Die ersten beiden Eigenfunktionen sind hier skizziert:



Für $E \geq V_1$ existiert ein kontinuierliches Spektrum und die Eigenzustände sind ungebunden. Diese diskutieren wir unten in 6.2.3 im Zusammenhang mit dem Tunneleffekt.

6.2.2 Kasten mit unendlich hohen Wänden

Im Grenzfall $V_1 \rightarrow \infty$ lassen sich die Eigenenergien und Eigenfunktionen geschlossen ausrechnen. Dazu müssen wir zunächst verstehen, was in diesem Fall aus den Anschlussbedingungen wird.



Für beliebige endliche Energie $E \geq 0$ ist $\kappa = \sqrt{2m(V_1 - E)}/\hbar \rightarrow \infty$ und damit $\psi_I(x) = A_I e^{\kappa x} \rightarrow 0$ für $x \leq -L/2$ und $\psi_{III}(x) = B_{III} e^{-\kappa x} \rightarrow 0$ für $x \geq L/2$. Stetigkeit erfordert dann $\psi_{II}(\pm L/2) = 0$.

Über die Ableitung $\psi'_{II}(\pm L/2)$ können wir aber nichts schließen: Am Anfang des Kapitels hatten wir gesehen, dass aus

$$\psi''(x) = -\frac{2m}{\hbar^2} [E - V(x)] \psi(x) \quad (6.65)$$

für ein bis auf endliche Sprünge stetiges Potential $V(x)$ folgt, dass $\psi''(x)$ ebenfalls stetig bis auf endliche Sprünge ist. Aber jetzt divergieren die Sprünge in $V(x)$ für $V_1 \rightarrow \infty$. Zwar geht $\psi(x)$ an diesen Sprüngen gegen Null, wir können aber nicht ausschließen, dass $\psi''(x)$ divergierende Sprünge enthält. An diesen kann das Integral $\psi'(x)$ von $\psi''(x)$ unstetig sein.

Wir können das Problem demnach mathematisch wie folgt formulieren (siehe auch Abschnitt 5.7):

$$\psi''(x) = -\frac{2m}{\hbar^2} E \psi(x) \quad \text{für } x \in \left[-\frac{L}{2}, \frac{L}{2}\right] \quad (6.66)$$

mit den Randbedingungen $\psi_{II}(\pm L/2) = 0$.

Zur Lösung definieren wir zunächst wieder $k := \sqrt{2mE}/\hbar$. Damit ist die Differentialgleichung

$$\psi''(x) = -k^2 \psi(x) \quad (6.67)$$

zu lösen. Ansatz:

$$\psi(x) = C \cos kx + D \sin kx \quad (6.68)$$

$$\Rightarrow \psi''(x) = -Ck^2 \cos kx - Dk^2 \sin kx = -k^2 \psi(x), \quad (6.69)$$

der Ansatz erfüllt die Schrödinger-Gleichung für alle C, D . Die Randbedingungen lauten

$$\psi(-L/2) = C \cos \frac{kL}{2} - D \sin \frac{kL}{2} = 0, \quad (6.70)$$

$$\psi(L/2) = C \cos \frac{kL}{2} + D \sin \frac{kL}{2} = 0. \quad (6.71)$$

Daraus folgt

$$\Rightarrow C \cos \frac{kL}{2} = 0 \quad \wedge \quad D \sin \frac{kL}{2} = 0. \quad (6.72)$$

$C = D = 0$ ist keine physikalisch sinnvolle Lösung. Sei nun $C \neq 0$, dann folgt

$$\cos \frac{kL}{2} = 0 \quad \Rightarrow \quad \frac{kL}{2} = \left(j + \frac{1}{2}\right) \pi \quad \text{für } j = 0, 1, 2, \dots \quad (6.73)$$

$$\Rightarrow \sin \frac{kL}{2} \neq 0 \quad \Rightarrow \quad D = 0. \quad (6.74)$$

Ist dagegen $D \neq 0$, so folgt

$$\sin \frac{kL}{2} = 0 \quad \Rightarrow \quad \frac{kL}{2} = j\pi \quad \text{für } j = 1, 2, 3, \dots \quad (6.75)$$

$$\Rightarrow \cos \frac{kL}{2} \neq 0 \quad \Rightarrow \quad C = 0. \quad (6.76)$$

Wir finden also wieder zwei Klassen von Lösungen mit geraden bzw. ungeraden Eigenfunktionen. Eigenenergien sind alle E_n mit (beachte $k = \sqrt{2mE}/\hbar$)

$$\frac{\sqrt{2mE_n} L}{2\hbar} = \frac{n}{2} \pi \quad \text{mit } n = 1, 2, 3, \dots \quad (6.77)$$

$$\Rightarrow E_n = \frac{\pi^2 \hbar^2}{2mL^2} n^2 \quad \text{mit } n = 1, 2, 3, \dots \quad (6.78)$$

Die Eigenfunktionen sind:

(a) für ungerades n ($D = 0$)

$$\psi_n(x) = C \cos kx, \quad k = \sqrt{2mE_n}/\hbar. \quad (6.79)$$

Normierung:

$$\int_{-L/2}^{L/2} dx |\psi_n(x)|^2 = \int_{-L/2}^{L/2} dx C^2 \cos^2 kx = \frac{C^2 L}{2} \stackrel{!}{=} 1. \quad (6.80)$$

Wähle $C = \sqrt{2/L}$, also

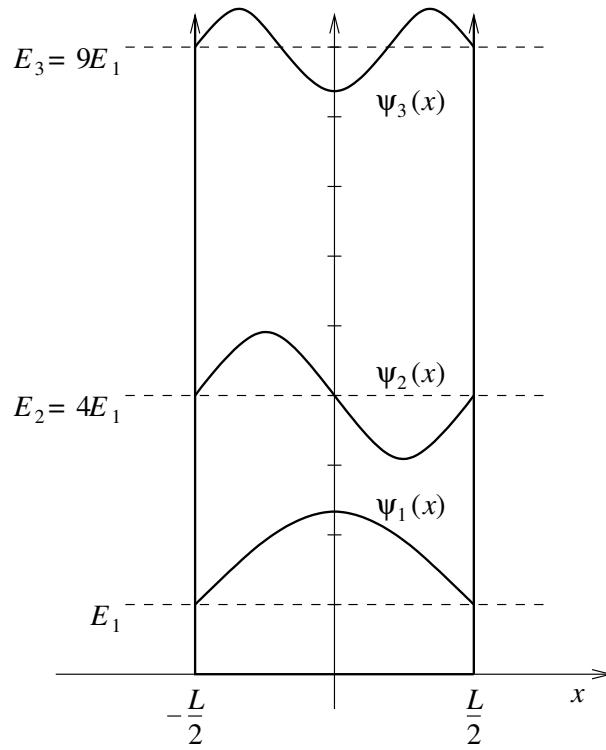
$$\psi_n(x) = \sqrt{\frac{2}{L}} \cos \frac{\sqrt{2mE_n} x}{\hbar} = \sqrt{\frac{2}{L}} \cos \frac{n\pi x}{L}. \quad (6.81)$$

(b) für gerades n ($C = 0$):

$$\psi_n(x) = D \sin kx, \quad k = \sqrt{2mE_n}/\hbar. \quad (6.82)$$

Analog:

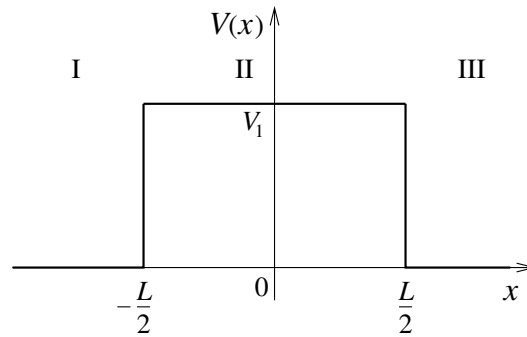
$$\psi_n(x) = \sqrt{\frac{2}{L}} \sin \frac{\sqrt{2mE_n} x}{\hbar} = \sqrt{\frac{2}{L}} \sin \frac{n\pi x}{L}. \quad (6.83)$$



6.2.3 Rechteckige Potentialbarriere: Tunneleffekt

Wir untersuchen nun die Potentialbarriere

$$V(x) = \begin{cases} 0 & \text{für } |x| \geq L/2 \\ V_1 > 0 & \text{für } |x| < L/2. \end{cases} \quad (6.84)$$



Es gilt $V_+ = V_- = V_0 = 0$. Das System hat daher ein zweifach entartetes kontinuierliches Spektrum, d. h. es existieren nur ungebundene Zustände. Wir betrachten o. B. d. A. den Fall eines von links einlaufenden Stroms von Teilchen der Energie $E > 0$. Die Lösung für einen von rechts einlaufenden Strom bei derselben Energie ist analog und ergibt den anderen der beiden entarteten Zustände.

Klassisch würden für $E < V_1$ alle Teilchen reflektiert und für $E > V_1$ alle transmittiert. Wie sieht das in der Quantenmechanik aus? Wir machen einen Ansatz durch ebene Wellen,

$$\psi_{\text{I}}(x) = A_{\text{I}} e^{ikx} + B_{\text{I}} e^{-ikx}, \quad (6.85)$$

$$\psi_{\text{II}}(x) = A_{\text{II}} e^{i\kappa x} + B_{\text{II}} e^{-i\kappa x}, \quad (6.86)$$

$$\psi_{\text{III}}(x) = A_{\text{III}} e^{ikx} + B_{\text{III}} e^{-ikx}, \quad (6.87)$$

mit

$$k = \sqrt{2mE}/\hbar \quad \text{und} \quad \kappa = \sqrt{2m(E - V_1)}/\hbar. \quad (6.88)$$

κ wird imaginär, falls $E < V_1$ ist. Dann besteht ψ_{II} aus exponentiell abfallenden und anwachsenden Anteilen. Wenn der Teilchenstrom von links einläuft, kommen keine Teilchen von rechts, also existiert im Bereich III keine linkslaufende Welle. Führen wir vorübergehend die Zeitabhängigkeit wieder ein,

$$\psi_{\text{III}}(x, t) = \underbrace{A_{\text{III}} e^{ikx - iEt/\hbar}}_{\text{rechtslaufend}} + \underbrace{B_{\text{III}} e^{-ikx - iEt/\hbar}}_{\text{linkslaufend}}, \quad (6.89)$$

so sehen wir, dass $B_{\text{III}} = 0$ sein muss.

Die Wellenfunktion ist nicht quadratintegrierbar, aber beschränkt. Die Wahl des Vorfaktors ist willkürlich, es ist aber praktisch, die Amplitude der einlaufenden Welle zu 1 zu wählen. Also setzen wir

$$\psi_{\text{I}}(x) = \underbrace{e^{ikx}}_{\text{einlaufend}} + \underbrace{r e^{-ikx}}_{\text{reflektiert}}, \quad (6.90)$$

$$\psi_{\text{II}}(x) = A e^{i\kappa x} + B e^{-i\kappa x}, \quad (6.91)$$

$$\psi_{\text{III}}(x) = \underbrace{t e^{ikx}}_{\text{transmittiert}} \quad (6.92)$$

mit noch unbekannten Koeffizienten r, t, A, B .

Der Ansatz erfüllt die Schrödinger-Gleichung. Wir müssen noch die Anschlussbedingungen berücksichtigen:

$$\begin{aligned} \psi_{\text{I}}(-L/2) &= e^{-ikL/2} + r e^{ikL/2} \\ &\stackrel{!}{=} \psi_{\text{II}}(-L/2) = A e^{-i\kappa L/2} + B e^{i\kappa L/2}, \end{aligned} \quad (6.93)$$

$$\begin{aligned} \psi'_{\text{I}}(-L/2) &= i k e^{-ikL/2} - i k r e^{ikL/2} \\ &\stackrel{!}{=} \psi'_{\text{II}}(-L/2) = i \kappa A e^{-i\kappa L/2} - i \kappa B e^{i\kappa L/2}, \end{aligned} \quad (6.94)$$

$$\begin{aligned} \psi_{\text{III}}(L/2) &= t e^{ikL/2} \\ &\stackrel{!}{=} \psi_{\text{II}}(L/2) = A e^{i\kappa L/2} + B e^{-i\kappa L/2}, \end{aligned} \quad (6.95)$$

$$\begin{aligned} \psi'_{\text{III}}(L/2) &= i k t e^{ikL/2} \\ &\stackrel{!}{=} \psi'_{\text{II}}(L/2) = i \kappa A e^{i\kappa L/2} - i \kappa B e^{-i\kappa L/2}. \end{aligned} \quad (6.96)$$

Dies ist ein inhomogenes lineares Gleichungssystem für r, t, A und B (inhomogen aufgrund des Terms von der einlaufenden Welle, der keine der Unbekannten enthält). Die Lösung ist elementar, aber etwas mühsam.

Wir betrachten nun die Wahrscheinlichkeitsstromdichte

$$\begin{aligned} j(x) &= \frac{\hbar}{2mi} \left[\psi^*(x) \frac{d}{dx} \psi(x) - \left(\frac{d}{dx} \psi^*(x) \right) \psi(x) \right] \\ &= \frac{\hbar}{2mi} [\psi^*(x) \psi'(x) - (\psi'(x))^* \psi(x)]. \end{aligned} \quad (6.97)$$

Man kann zeigen, dass $j(x)$ auf ganz \mathbb{R} räumlich konstant ist: Die Kontinuitätsgleichung für dieses eindimensionale System lautet

$$\frac{\partial \varrho}{\partial t} + \frac{\partial j}{\partial x} = 0. \quad (6.98)$$

Für die Eigenfunktion ψ_n ist jedoch

$$\varrho(x, t) = |\psi_n(x) e^{-iE_n t/\hbar}|^2 = |\psi_n(x)|^2 \quad (6.99)$$

und somit $\partial \varrho / \partial t = 0$. Es folgt $\partial j / \partial x = 0$.

Der Strom links der Barriere (im Bereich I) ist

$$\begin{aligned}
j_I(x) &= \frac{\hbar}{2mi} [(e^{-ikx} + r^* e^{ikx}) (ik e^{ikx} - rik e^{-ikx}) - (-ik e^{-ikx} + r^* ik e^{ikx}) (e^{ikx} + r e^{-ikx})] \\
&= \frac{\hbar k}{2m} [1 - \cancel{r e^{-2ikx}} + \cancel{r^* e^{2ikx}} - |r|^2 + 1 + \cancel{r e^{-2ikx}} - \cancel{r^* e^{2ikx}} - |r|^2] \\
&= \frac{\hbar k}{m} (1 - |r|^2).
\end{aligned} \tag{6.100}$$

Er besteht aus dem einlaufenden Strom

$$j_i(x) = \frac{\hbar k}{m} \tag{6.101}$$

und dem reflektierten Strom

$$j_r(x) = -\frac{\hbar k}{m} |r|^2, \tag{6.102}$$

der negativ ist, da er nach links fließt. Der Strom rechts der Barriere (im Bereich III) ist der transmittierte Strom

$$j_{III}(x) = j_t(x) = \frac{\hbar}{2mi} [t^* e^{-ikx} tik e^{ikx} + t^* ik e^{-ikx} t e^{ikx}] = \frac{\hbar k}{m} |t|^2. \tag{6.103}$$

Man definiert die *Reflexions-* und *Transmissionswahrscheinlichkeiten* (oder *-koeffizienten*) als

$$R := -\frac{j_r}{j_i}, \tag{6.104}$$

$$T := \frac{j_t}{j_i}. \tag{6.105}$$

Es folgt für die betrachtete Barriere

$$R = |r|^2, \tag{6.106}$$

$$T = |t|^2. \tag{6.107}$$

Es ist wichtig, die Koeffizienten über die Stromdichten zu bestimmen, nicht direkt über die Amplituden. Der Weg über die Stromdichten gewährleistet automatisch die Erhaltung der Wahrscheinlichkeit und funktioniert für beliebige Potentiale $V(x)$, sofern diese überhaupt ungebundene Zustände erlauben.

Da der Gesamtstrom konstant ist, muss $j_I(x) = j_{III}(x)$ gelten, also

$$1 - |r|^2 = |t|^2. \tag{6.108}$$

Die explizite Lösung für r , t , A und B erfüllt dies natürlich. Es folgt

$$1 - R = T \tag{6.109}$$

$$\Rightarrow R + T = 1. \tag{6.110}$$

Das ist vernünftig – jedes Teilchen wird sicher entweder reflektiert oder transmittiert, die Summe der beiden Wahrscheinlichkeiten ist eins.

Die Lösung des Gleichungssystems ergibt, hier ohne Details,

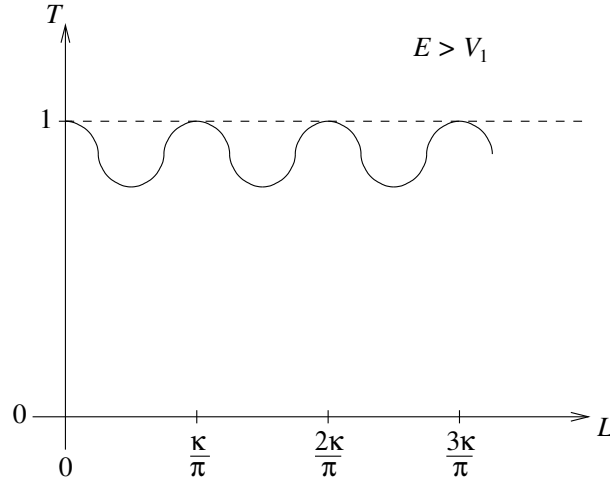
$$T = |t|^2 = \frac{16E(E - V_1)}{16E(E - V_1) - V_1^2(e^{i\kappa L} - e^{-i\kappa L})^2} \tag{6.111}$$

mit $\kappa = \sqrt{2m(E - V_1)}/\hbar$. Dieser Ausdruck gilt sowohl für $E > V_1$ als auch für $0 < E < V_1$.

Wir können auch schreiben

$$T = \frac{4E(E - V_1)}{4E(E - V_1) + V_1^2 \sin^2 \kappa L}, \tag{6.112}$$

was allgemein richtig, aber besonders für $E > V_1$, also für reelles κ , nützlich ist. Liegt die Energie oberhalb der Barriere, finden wir also, dass sich die Transmissionswahrscheinlichkeit T *periodisch* mit der Dicke L der Barriere ändert. Insbesondere ist $T = 1$ für $L = n\pi/\kappa$, $n = 0, 1, 2, \dots$. Diese nicht monotone Abhängigkeit der Transmission von L beruht auf der Interferenz zwischen den an den beiden Sprüngen reflektierten Anteilen der Welle. Beachte auch, dass wir, anders als im klassischen Fall, auch für $E > V_1$ im Allgemeinen keine vollständige Transmission erhalten.



Die Herleitung hat nie benutzt, dass $V_1 > 0$ ist. Wir erhalten für den Potentialtopf endlicher Tiefe, $V_1 < 0$, also dieselbe Formel. Damit könnten wir nun auch die Streuzustände des oben behandelten Kastens endlicher Tiefe angeben.

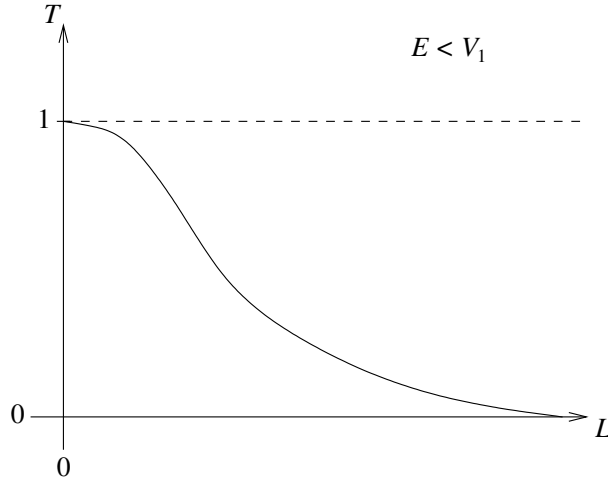
Wir kommen nun zum Fall $E < V_1$ (mit $V_1 > 0$). Gleichung (6.112) ist allgemein richtig. Für $E < V_1$ ist allerdings κ imaginär. Mit

$$\sin z = \frac{e^{iz} - e^{-iz}}{2i} = \frac{1}{i} \sinh iz = i \sinh(-iz) \quad (6.113)$$

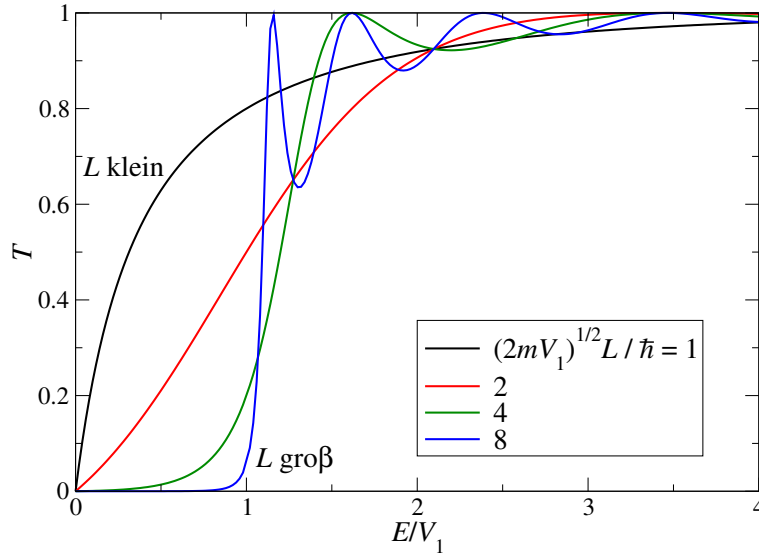
können wir dann schreiben

$$\begin{aligned} T &= \frac{4E(E - V_1)}{4E(E - V_1) - V_1^2 \sinh^2(-i\kappa L)} \\ &= \frac{4E(V_1 - E)}{4E(V_1 - E) + V_1^2 \sinh^2(-i\kappa L)} \\ &= \frac{4E(V_1 - E)}{4E(V_1 - E) + V_1^2 \sinh^2 \frac{\sqrt{2m(V_1 - E)} L}{\hbar}}. \end{aligned} \quad (6.114)$$

Auch dieser Ausdruck ist allgemein richtig. Er ist besonders nützlich für $E < V_1$. Liegt die Energie also unterhalb der Barrierenhöhe, finden wir quantenmechanisch dennoch eine nichtverschwindende Transmissionswahrscheinlichkeit. Das ist der quantenmechanische *Tunneleffekt*. Er ist zum Beispiel wichtig für α -Zerfälle von Atomkernen, für die Kernfusion und auch für die Funktion eines Rastertunnelmikroskops (scanning tunneling microscope, STM). Wir sehen, dass T für wachsende Dicke L monoton abnimmt.



Schließlich skizzieren wir noch T als Funktion der Energie E für verschiedene Dicken L :



Die Tatsache, dass $T > 0$ für $E/V_1 < 1$ gilt, ist der Tunneleffekt. Wir sehen außerdem, dass T für $E/V_1 > 1$ oszilliert, was wieder ein Interferenzeffekt ist. Für $E \rightarrow \infty$ finden wir $T \rightarrow 1$, also vollständige Transmission.

6.3 Der harmonische Oszillator I

Der harmonischer Oszillator, beschrieben durch den Hamilton-Operator

$$\hat{H} = -\frac{\hbar^2}{2m} \frac{d^2}{dx^2} + \frac{1}{2} m \omega^2 x^2, \quad (6.115)$$

ist eines der wichtigsten Modelle in der Physik. So kann man kompliziertere Potentiale $V(x)$ in der Nähe „generischer“ Minima in eine Taylor-Reihe entwickeln und durch den konstanten und den quadratischen Term annähern, was einen harmonischen Oszillator ergibt. Außerdem wird in der Quantenfeldtheorie jede Mode (vgl. den Abschnitt 2.2 zum Begriff der Mode) des Feldes durch einen harmonischer Oszillator beschrieben. Die zeitunabhängige Schrödinger-Gleichung

$$-\frac{\hbar^2}{2m} \psi''(x) + \frac{1}{2} m \omega^2 x^2 \psi(x) = E \psi(x) \quad (6.116)$$

lässt sich durch die Ersetzungen

$$\xi := \sqrt{\frac{m\omega}{\hbar}} x, \quad (6.117)$$

$$K := \frac{2E}{\hbar\omega} \quad (6.118)$$

vereinfachen zu

$$\psi''(\xi) = (\xi^2 - K) \psi(\xi). \quad (6.119)$$

Hier ist $\psi(\xi)$ nicht dieselbe Funktion wie $\psi(x)$, sondern es gilt $\psi(x) = \psi(\xi = \sqrt{m\omega/\hbar} x)$. Wie in der Physik üblich, verwenden wir dasselbe Symbol, da es sich um dieselbe *physikalische Größe* handelt. Da das Potential für $x \rightarrow \pm\infty$ ($\xi \rightarrow \pm\infty$) divergiert, $V_+ = V_- = \infty$, existiert nur ein diskretes Spektrum mit gebundenen Zuständen. Die Eigenfunktionen müssen also quadratintegrabel sein und für $\xi \rightarrow \pm\infty$ verschwinden.

Wir betrachten zunächst das asymptotische Verhalten der Lösungen für große $|\xi|$ – dies ist eine allgemein nützliche Methode für die Lösung von Differentialgleichungen. Dann lautet die Gleichung

$$\psi'' \approx \xi^2 \psi \quad (6.120)$$

mit der asymptotischen Lösung

$$\psi \approx A e^{-\xi^2/2} + B e^{\xi^2/2}. \quad (6.121)$$

Hier muss $B = 0$ sein, damit $\lim_{\xi \rightarrow \pm\infty} \psi = 0$ gilt. Es ist nützlich, die asymptotische Abhängigkeit abzuspalten,

$$\psi(\xi) = h(\xi) e^{-\xi^2/2}, \quad (6.122)$$

und Lösungen für $h(\xi)$ zu suchen. Mit

$$\psi' = (h' - \xi h) e^{-\xi^2/2}, \quad (6.123)$$

$$\psi'' = [h'' - 2\xi h' + (\xi^2 - 1)h] e^{-\xi^2/2} \quad (6.124)$$

finden wir

$$h'' - 2\xi h' + (K - 1)h = 0. \quad (6.125)$$

Diese Gleichung ist linear, homogen und von zweiter Ordnung, aber nicht einfach, weil der Koeffizient des Terms h' nicht konstant ist. Eine sinnvolle Methode für die Lösung ist die Darstellung von $h(\xi)$ als Potenzreihe (*Frobenius-Methode*):

$$h(\xi) = \sum_{j=0}^{\infty} a_j \xi^j \quad (6.126)$$

mit noch unbekannten a_j . Dann ist

$$h'(\xi) = \sum_{j=1}^{\infty} j a_j \xi^{j-1} = \sum_{j=0}^{\infty} (j+1) a_{j+1} \xi^j \quad (6.127)$$

und

$$h''(\xi) = \sum_{j=1}^{\infty} j(j+1) a_{j+1} \xi^{j-1} = \sum_{j=0}^{\infty} (j+1)(j+2) a_{j+2} \xi^j. \quad (6.128)$$

Einsetzen in die Gleichung ergibt

$$\sum_{j=0}^{\infty} [(j+1)(j+2) a_{j+2} - 2j a_j + (K-1) a_j] \xi^j = 0. \quad (6.129)$$

Die beiden Seiten müssen in jeder Ordnung in ξ gleich sein, weil die ξ^j linear unabhängig sind, also

$$(j+1)(j+2) a_{j+2} - 2j a_j + (K-1) a_j = 0 \quad \forall j. \quad (6.130)$$

Dies ergibt die Rekursionsformel

$$a_{j+2} = \frac{2j+1-K}{(j+1)(j+2)} a_j. \quad (6.131)$$

Die Lösung hat offenbar die Form

$$h(\xi) = h_{\text{gerade}}(\xi) + h_{\text{ungerade}}(\xi), \quad (6.132)$$

wobei

$$h_{\text{gerade}}(\xi) = a_0 + a_2 \xi^2 + a_4 \xi^4 + \dots \quad (6.133)$$

nur von a_0 abhängt und

$$h_{\text{ungerade}}(\xi) = a_1 \xi + a_3 \xi^3 + a_5 \xi^5 + \dots \quad (6.134)$$

nur von a_1 abhängt. Wir haben also zwei freie Parameter a_0 und a_1 in der Lösung, wie für eine Gleichung zweiter Ordnung auch zu erwarten war. Jedoch sind nicht alle so erhaltenen Lösungen quadratintegrabel. Es ist zunächst überraschend, dass es überhaupt beschränkte Lösungen gibt: Für große j wird Gleichung (6.131) nämlich zu

$$a_{j+2} \approx \frac{2}{j} a_j \quad (6.135)$$

mit der Lösung

$$a_j \approx \begin{cases} \frac{c_0}{(\frac{j}{2})!} & \text{für gerade } j, \\ \frac{c_1}{(\frac{j-1}{2})!} & \text{für ungerade } j \end{cases} \quad (6.136)$$

mit Konstanten c_0 und c_1 . Der Beweis ist einfach: Es folgt

$$\frac{a_{j+2}}{a_j} = \begin{cases} \frac{(\frac{j}{2})!}{(\frac{j}{2}+1)!} = \frac{1}{\frac{j}{2}+1} \approx \frac{2}{j} & \text{für gerade } j, \\ \frac{(\frac{j-1}{2})!}{(\frac{j+1}{2})!} = \frac{1}{\frac{j+1}{2}} \approx \frac{2}{j} & \text{für ungerade } j. \end{cases} \quad (6.137)$$

Dann ist

$$\begin{aligned} h(\xi) &\approx c_0 \sum_{j \text{ gerade}} \frac{1}{(\frac{j}{2})!} \xi^j + c_1 \sum_{j \text{ ungerade}} \frac{1}{(\frac{j-1}{2})!} \xi^j \\ &= c_0 \sum_{n=0}^{\infty} \frac{1}{n!} \xi^{2n} + c_1 \sum_{n=0}^{\infty} \frac{1}{n!} \xi^{2n+1} \\ &= c_0 e^{\xi^2} + c_1 \xi e^{\xi^2} = (c_0 + c_1 \xi) e^{\xi^2}. \end{aligned} \quad (6.138)$$

Aber damit wird

$$\psi(\xi) = h(\xi) e^{-\xi^2/2} \approx (c_0 + c_1 \xi) e^{\xi^2/2}, \quad (6.139)$$

was für $\xi \rightarrow \pm\infty$ divergiert. Der einzige Ausweg ist, dass die Iteration, Gl. (6.131), abbrechen kann. Dies geschieht für $K = 2n + 1$ mit $n = 0, 1, 2, \dots$, dann gilt nämlich

$$a_{n+2} = \frac{2n+1-K}{(n+1)(n+2)} a_n = \frac{0}{(n+1)(n+2)} a_n = 0 \quad (6.140)$$

unabhängig von a_n , und dann natürlich auch $a_{n+4} = a_{n+6} = \dots = 0$. Dann ist $h(\xi)$ ein Polynom und $h(\xi)e^{-\xi^2/2}$ verschwindet für $\xi \rightarrow \pm\infty$. Für $K = 2n + 1$ nimmt die Rekursionsformel die Form

$$a_{j+2} = \frac{2(j-n)}{(j+1)(j+2)} a_j \quad (6.141)$$

an. Die ersten paar Lösungen sind:

$n = 0$:

$$a_2 = \frac{2(0-0)}{1 \times 2} a_0 = 0, \quad (6.142)$$

wähle $a_1 = 0$, dann ist $a_3 = a_5 = \dots = 0$ (sonst würde die Folge a_n nicht abbrechen und die Lösungsfunktion wäre nicht quadratintegrabel) und es folgt

$$h_0(\xi) = a_0 \quad (6.143)$$

$$\Rightarrow \psi_0(\xi) = a_0 e^{-\xi^2/2} \quad (6.144)$$

(a_0 erhalten wir aus der Normierung).

$n = 1$:

$$a_3 = \frac{2(1-1)}{2 \times 3} a_1 = 0, \quad (6.145)$$

wähle $a_0 = 0$, dann ist $a_2 = a_4 = \dots = 0$ und es folgt

$$h_1(\xi) = a_1 \xi \quad (6.146)$$

$$\Rightarrow \psi_1(\xi) = a_1 \xi e^{-\xi^2/2}. \quad (6.147)$$

$n = 2$:

$$a_2 = \frac{2(0-2)}{1 \times 2} a_0 = -2a_0, \quad (6.148)$$

$$a_4 = \frac{2(2-2)}{3 \times 4} a_2 = 0, \quad (6.149)$$

wähle $a_1 = 0$, dann folgt

$$h_2(\xi) = a_0 - 2a_0 \xi^2 = a_0 (1 - 2\xi^2) \quad (6.150)$$

$$\Rightarrow \psi_2(\xi) = a_0 (1 - 2\xi^2) e^{-\xi^2/2}. \quad (6.151)$$

$n = 3$:

$$a_3 = \frac{2(1-3)}{2 \times 3} a_1 = -\frac{2}{3} a_1, \quad (6.152)$$

$$a_5 = \frac{2(3-3)}{4 \times 5} a_3 = 0, \quad (6.153)$$

wähle $a_0 = 0$, dann folgt

$$h_3(\xi) = a_1 \xi - \frac{2}{3} a_1 \xi^3 = a_1 \left(\xi - \frac{2}{3} \xi^3 \right) \quad (6.154)$$

$$\Rightarrow \psi_3(\xi) = a_1 \left(\xi - \frac{2}{3} \xi^3 \right) e^{-\xi^2/2}. \quad (6.155)$$

Ohne Beweis geben wir an, dass die allgemeine, normierte Lösung lautet

$$\begin{aligned} \psi_n(x) &= \left(\frac{m\omega}{\pi\hbar} \right)^{1/4} \frac{1}{\sqrt{2^n n!}} H_n(\xi) e^{-\xi^2/2} \quad \left| \text{mit } \xi = \sqrt{\frac{m\omega}{\hbar}} x \right. \\ &= \left(\frac{m\omega}{\pi\hbar} \right)^{1/4} \frac{1}{\sqrt{2^n n!}} H_n \left(\sqrt{\frac{m\omega}{\hbar}} x \right) \exp \left(-\frac{m\omega}{2\hbar} x^2 \right), \quad n = 0, 1, 2, \dots, \end{aligned} \quad (6.156)$$

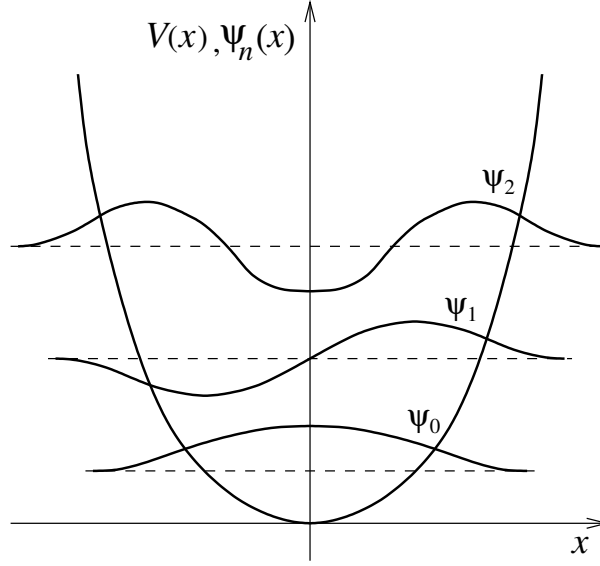
wobei $H_n(\xi)$ die *Hermite-Polynome* sind. Die ersten Hermite-Polynome lauten

$$H_0(\xi) = 1, \quad (6.157)$$

$$H_1(\xi) = 2\xi, \quad (6.158)$$

$$H_2(\xi) = 4\xi^2 - 2, \quad (6.159)$$

$$H_3(\xi) = 8\xi^3 - 12\xi, \quad \dots \quad (6.160)$$



Die Skizze legt die Vermutung nahe, dass n die Anzahl der Nullstellen von $\psi_n(x)$ angibt. Das ist tatsächlich für alle n korrekt.

Die Eigenenergien lauten

$$E_n = \frac{\hbar\omega}{2} K(n) = \frac{\hbar\omega}{2} (2n + 1) = \hbar\omega \left(n + \frac{1}{2} \right). \quad (6.161)$$

Insbesondere ist die Grundzustandsenergie $E_0 = \hbar\omega/2$ und damit größer als die minimale klassische Energie, $E_0^{\text{klassisch}} = \min V(x) = 0$. Die Differenz nennt man *Nullpunktsenergie* oder, im Zusammenhang mit der Quantenfeldtheorie, auch *Vakuumenergie*. Ihr Auftreten ist ein fundamental quantenmechanisches Phänomen. Wir werden im Rahmen des Dirac-Formalismus eine elegantere Beschreibung des harmonischen Oszillators kennenlernen.

6.4 Parität und Knotenzahl

Bei der Betrachtung der gebundenen Lösungen für den Kasten endlicher Tiefe und für den harmonischen Oszillator fallen mehrere Gemeinsamkeiten auf: Für beide Potentiale sind die Eigenfunktionen entweder gerade oder ungerade Funktionen von x und ihre *Parität* – die Eigenschaft, gerade oder ungerade zu sein – wechselt sich für aufsteigende Eigenenergie ab. Der Grundzustand ist gerade. Außerdem haben die Eigenfunktionen für aufsteigende Energie $0, 1, 2, \dots$ Nullstellen (Knoten). Hier wollen wir diese Eigenschaften für allgemeine Potentiale in einer Dimension untersuchen. Dafür ist die Definition der *Wronski-Determinante* von zwei Funktionen $\psi(x), \varphi(x)$ nützlich:

$$W(\psi, \varphi; x) := \begin{vmatrix} \psi(x) & \varphi(x) \\ \psi'(x) & \varphi'(x) \end{vmatrix} = \psi(x)\varphi'(x) - \varphi(x)\psi'(x). \quad (6.162)$$

Seien ψ_m und ψ_n zwei Eigenfunktionen des Hamiltonians \hat{H} zu den Eigenenergien E_m und E_n . Dann gilt:

$$\psi_m''(x) = \frac{2m}{\hbar^2} [V(x) - E_m] \psi_m(x), \quad (6.163)$$

$$\psi_n''(x) = \frac{2m}{\hbar^2} [V(x) - E_n] \psi_n(x), \quad (6.164)$$

Es folgt

$$\frac{d}{dx} [\psi_m'(x) \psi_n(x) - \psi_n'(x) \psi_m(x)] = \psi_m''(x) \psi_n(x) - \psi_n''(x) \psi_m(x)$$

$$\begin{aligned}
&= \frac{2m}{\hbar^2} [V(x) - E_m - V(x) + E_n] \psi_m(x) \psi_n(x) \\
&= \frac{2m}{\hbar^2} [E_n - E_m] \psi_m(x) \psi_n(x)
\end{aligned} \tag{6.165}$$

und daher

$$\begin{aligned}
W(\psi_m, \psi_n; x_2) - W(\psi_m, \psi_n; x_1) &= W(\psi_m, \psi_n; x) \Big|_{x_1}^{x_2} \\
&= [\psi_m(x) \psi'_n(x) - \psi_n(x) \psi'_m(x)]_{x_1}^{x_2} \\
&= -[\psi'_m(x) \psi_n(x) - \psi'_n(x) \psi_m(x)]_{x_1}^{x_2} \\
&= - \int_{x_1}^{x_2} dx \frac{d}{dx} [\psi'_m(x) \psi_n(x) - \psi'_n(x) \psi_m(x)] \\
&= - \int_{x_1}^{x_2} dx \frac{2m}{\hbar^2} (E_n - E_m) \psi_m(x) \psi_n(x) \\
&= \frac{2m}{\hbar^2} (E_m - E_n) \int_{x_1}^{x_2} dx \psi_m(x) \psi_n(x).
\end{aligned} \tag{6.166}$$

Für quadratintegrale ψ_q ($q = m, n$) gilt $\lim_{x \rightarrow \pm\infty} \psi_q(x) = 0$ und $\lim_{x \rightarrow \pm\infty} \psi'_q(x) = 0$ und damit $\lim_{x \rightarrow \pm\infty} W(\psi_m, \psi_n; x) = 0$. Wir erhalten dann

$$W(\psi_m, \psi_n; x) - \underbrace{W(\psi_m, \psi_n; -\infty)}_{=0} = \frac{2m}{\hbar^2} (E_m - E_n) \int_{-\infty}^x dx' \psi_m(x') \psi_n(x'). \tag{6.167}$$

Betrachten wir den Fall $E_m = E_n$. Dann folgt

$$W(\psi_m, \psi_n; x) = 0 \quad \forall x, \tag{6.168}$$

also

$$\psi_m(x) \psi'_n(x) - \psi_n(x) \psi'_m(x) = 0 \tag{6.169}$$

$$\Rightarrow \psi_m(x) \psi'_n(x) = \psi_n(x) \psi'_m(x). \tag{6.170}$$

Hieraus folgt für alle x , außer an isolierten Nullstellen von ψ_m und ψ_n ,

$$\frac{\psi'_n(x)}{\psi_n(x)} = \frac{\psi'_m(x)}{\psi_m(x)} \tag{6.171}$$

$$\Rightarrow \frac{d}{dx} \ln \psi_n(x) = \frac{d}{dx} \ln \psi_m(x) \tag{6.172}$$

$$\Rightarrow \ln \psi_n(x) = \ln \psi_m(x) + C \tag{6.173}$$

mit einer Konstanten C . Damit erhalten wir

$$\psi_n(x) = e^C \psi_m(x). \tag{6.174}$$

Also sind ψ_m und ψ_n linear abhängig. Es folgt, dass für eine gegebene Energie nur eine linear unabhängige Lösung existieren kann. Die Eigenenergien der gebundenen Zustände sind also *nicht entartet*.

Eine weitere Aussage ergibt sich für $E_m \neq E_n$. Wir schicken x nach $+\infty$ und erhalten aus Gl. (6.167)

$$0 = W(\psi_m, \psi_n; \infty) - W(\psi_m, \psi_n; -\infty) = \frac{2m}{\hbar^2} (E_m - E_n) \int_{-\infty}^{\infty} dx \psi_m(x) \psi_n(x) \tag{6.175}$$

$$\xrightarrow{E_m \neq E_n} 0 = \int_{-\infty}^{\infty} dx \psi_m(x) \psi_n(x) = \int_{-\infty}^{\infty} dx \psi_m^*(x) \psi_n(x) = \langle \psi_m, \psi_n \rangle. \tag{6.176}$$

Verschiedene gebundene Eigenzustände sind also orthogonal. Das hatten wir zwar schon behauptet, aber noch nicht gezeigt.

6.4.1 Knotenzahl

Wir beweisen nun zwei Aussagen über die Nullstellen (Knoten) der Eigenfunktionen für gebundene Zustände. Die zweite Aussage schließt die erste ein und ist etwas schwieriger zu zeigen. Wir ordnen die Eigenzustände $n = 0, 1, 2, \dots$ so, dass gilt $E_0 < E_1 < E_2 < \dots$. Die Zahl der Eigenwerte im diskreten Spektrum kann endlich oder (abzählbar) unendlich sein. Es ist hier wichtig, dass wir mit $n = 0$ beginnen, nicht mit $n = 1$.

Sei nun $n > m$, also $E_n > E_m$. Seien x_1 und x_2 , $x_1 < x_2$, zwei benachbarte Nullstellen von $\psi_m(x)$. Wir lassen auch Werte $x_1 = -\infty$ und $x_2 = \infty$ zu, mit offensichtlicher Bedeutung. O. B. d. A. nehmen wir

$$\psi_m(x) > 0 \quad \forall x > x_1 \wedge x < x_2 \quad (6.177)$$

an. Gleichung (6.166) ergibt dann

$$W(\psi_m, \psi_n; x_2) - W(\psi_m, \psi_n; x_1) = \frac{2m}{\hbar^2} (E_m - E_n) \int_{x_1}^{x_2} dx \psi_m(x) \psi_n(x) \quad (6.178)$$

$$\begin{aligned} \Rightarrow \frac{2m}{\hbar^2} (E_m - E_n) \int_{x_1}^{x_2} dx \psi_m(x) \psi_n(x) &= \underbrace{\psi_m(x_2)}_{=0} \psi'_n(x_2) - \psi_n(x_2) \psi'_m(x_2) - \underbrace{\psi_m(x_1)}_{=0} \psi'_n(x_1) + \psi_n(x_1) \psi'_m(x_1) \\ &= -\psi_n(x_2) \psi'_m(x_2) + \psi_n(x_1) \psi'_m(x_1) \end{aligned} \quad (6.179)$$

$$\Rightarrow \frac{2m}{\hbar^2} (E_n - E_m) \int_{x_1}^{x_2} dx \psi_m(x) \psi_n(x) = \psi_n(x_2) \psi'_m(x_2) - \psi_n(x_1) \psi'_m(x_1) \quad (6.180)$$

Nehmen wir nun an, dass $\psi_m(x)$ auf $]x_1, x_2[$ das Vorzeichen nicht wechselt. Sei o. B. d. A. $\psi_n(x) > 0 \quad \forall x > x_1 \wedge x < x_2$. Dann ist die linke Seite von Gl. (6.180) positiv. Die rechte Seite ist aber

$$\underbrace{\psi_n(x_2)}_{\geq 0} \underbrace{\psi'_m(x_2)}_{\leq 0} - \underbrace{\psi_n(x_1)}_{\geq 0} \underbrace{\psi'_m(x_1)}_{\geq 0} \leq 0. \quad (6.181)$$

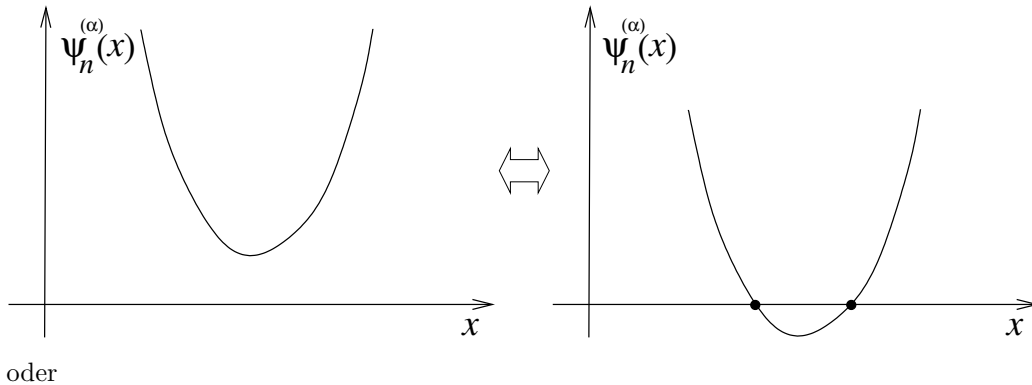
Widerspruch! Also muss $\psi_n(x)$ zwischen den Nullstellen von $\psi_m(x)$ – einschließlich der „asymptotischen“ Nullstellen für $x \rightarrow \pm\infty$ – mindestens einmal das Vorzeichen wechseln. Daraus folgt sofort, dass $\psi_n(x)$ für $E_n > E_m$ mehr Nullstellen hat als $\psi_m(x)$.

Wir begründen nun nicht rigoros, dass die Zahl der Nullstellen (für endliche x) gegeben ist durch n . Das ist die Aussage des *Knotensatzes*. Die Beweisidee ist konzeptionell ähnlich zu M. Moriconi, arXiv:quant-ph/0702260: Wir definieren eine Schar von Potentialfunktionen:

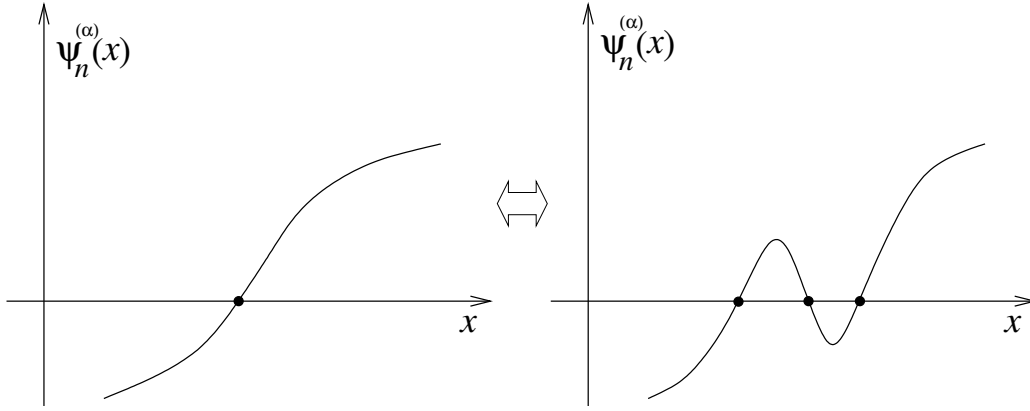
$$V_\alpha(x) = \alpha \frac{1}{2} m \omega_0^2 x^2 + (1 - \alpha) V(x). \quad (6.182)$$

Für $\alpha = 1$ erhalten wir den harmonischen Oszillator, für den wir die Gültigkeit des Knotensatzes bereits explizit gezeigt haben. Für $\alpha = 0$ ist V_α mit dem betrachteten Potential V identisch. Zumindest für hinreichend gutartige V hängen die Eigenwerte $E_n^{(\alpha)}$ und Eigenfunktionen $\psi_n^{(\alpha)}(x)$ stetig von $\alpha \in [0, 1]$ ab. Dann könnte sich die Zahl der Knoten von $\psi_n^{(\alpha)}$ nur dadurch ändern, dass für einen gewissen kritischen Wert $\alpha = \alpha_c$ entweder

- (a) ein Extremum von $x \mapsto \psi_n^{(\alpha)}(x)$ durch Null geschoben wird,



(b) eine Nullstelle in drei aufgespalten wird oder umgekehrt.



In beiden Fällen existiert ein x_c , so dass bei $\alpha = \alpha_c$ gilt

$$\psi_n^{(\alpha_c)}(x_c) = 0 \wedge \psi_n^{(\alpha_c)'}(x_c) = 0 \quad (6.183)$$

Da die Schrödinger-Gleichung linear und von zweiter Ordnung ist, ist die Lösung durch die Angabe der Funktion und ihrer Ableitung an einem Punkt eindeutig bestimmt. Da $\psi \equiv 0$ offensichtlich eine Lösung ist, ist dies dann auch die einzige Lösung. Aber dies ist keine Eigenfunktion. Also verschwindet die n -te Lösung aus dem diskreten Spektrum für $\alpha = \alpha_c$, im Widerspruch zur Erwartung, dass das Spektrum stetig von α abhängt. Es ist hier wichtig zu beachten, dass für $\alpha > 0$ gilt $\lim_{x \rightarrow \pm\infty} V_\alpha(x) = \infty$, so dass kein kontinuierliches Spektrum existiert. Das Verschwinden einer Lösung für $\alpha_c > 0$ kann also nicht darauf beruhen, dass eine diskrete Lösung ins kontinuierliche Spektrum eintaucht. Insgesamt finden wir für gutartige Potentiale, dass sich die Zahl der Knoten von $\psi_n^{(\alpha)}(x)$ als Funktion von $\alpha \in [0, 1]$ nicht ändert. Andererseits können sich Eigenwerte auch nicht schneiden, da Entartung bereits ausgeschlossen wurde. Damit bleiben die Eigenenergien $E_n^{(\alpha)}$ aufsteigend geordnet und wir finden, dass $\psi_n^{(\alpha)}(x)$ mit geordneten $E_n^{(\alpha)}$ dieselbe Zahl von Knoten hat wie für den harmonischen Oszillator, also n .

Eine Bemerkung ist noch notwendig: $V(x)$ muss nicht $\lim_{x \rightarrow \pm\infty} V(x) = \infty$ erfüllen und hat dann i. A. nicht unendlich viele diskrete Eigenenergien. Da die hier betrachteten Potentiale keine Lücken im kontinuierlichen Spektrum haben und da sich Eigenwerte für $\alpha > 0$ nicht schneiden, müssen dann alle Zustände mit $n \geq n_{\text{dis}}$, n_{dis} ist die Mächtigkeit des diskreten Spektrums, ins kontinuierliche Spektrum übergehen.

6.4.2 Parität

Der Knotensatz gilt unabhängig von der Symmetrie des Potentials $V(x)$. Die Beispiele des Kastenpotentials und des harmonischen Oszillators betrafen Potentiale mit der speziellen Eigenschaft

$$V(-x) = V(x). \quad (6.184)$$

Für diese geraden Funktionen finden wir

$$-\frac{\hbar^2}{2m} \frac{d^2 \psi_n}{dx^2} + V(x) \psi_n(x) = E_n \psi_n(x) \quad (6.185)$$

$$\stackrel{u \Rightarrow -x}{\Rightarrow} -\frac{\hbar^2}{2m} \left(-\frac{d}{du} \right) \left(-\frac{d}{du} \right) \psi_n(-u) + \underbrace{V(-u)}_{=V(u)} \psi_n(-u) = E_n \psi_n(-u) \quad (6.186)$$

$$\Rightarrow -\frac{\hbar^2}{2m} \frac{d^2}{du^2} \psi_n(-u) + V(u) \psi_n(-u) = E_n \psi_n(-u) \quad (6.187)$$

und mit Umbenennung $u \rightarrow x$:

$$-\frac{\hbar^2}{2m} \frac{d^2}{dx^2} \psi_n(-x) + V(x) \psi_n(-x) = E_n \psi_n(-x) \quad (6.188)$$

$\psi_n(-x)$ erfüllt also dieselbe Schrödinger-Gleichung wie $\psi_n(x)$. Da die diskreten Eigenenergien nicht entartet sind, folgt, dass sich $\psi_n(-x)$ und $\psi_n(x)$ nur um einen Faktor unterscheiden. Fordern wir Normiertheit und Realität, so kann nur gelten

$$\psi_n(-x) = \pm \psi_n(x). \quad (6.189)$$

Die Eigenfunktionen sind also entweder gerade oder ungerade.

Gerade (ungerade) Funktionen können nur eine gerade (ungerade) Zahl von Nullstellen haben. Mit dem Knotensatz folgt, dass die $\psi_n(x)$ für aufsteigend geordnete Eigenenergien E_n abwechselnd gerade und ungerade sind. Die Grundzustandswellenfunktion ($n = 0$) ist gerade.

Kapitel 7

Dirac-Formalismus

Wir haben im vorigen Kapitel gesehen, dass wir die Dynamik eines Teilchens äquivalent in der Orts- und Impulsdarstellung beschreiben können. Das legt nahe, dass dies nur zwei spezielle Darstellungen einer fundamentalen Theorie sind. Von welcher Art kann diese sein? Die Wellenmechanik führte bereits auf Operatoren und deren Eigenwerte sowie auf das Superpositionsprinzip. Wir können also eine lineare Algebra von Operatoren als fundamentale Theorie erwarten. In diesem Kapitel wird diese Theorie, P. Dirac folgend, axiomatisch aufgebaut. Es wird sich zeigen, dass sie sogar noch allgemeiner ist als gedacht, sie beschreibt nämlich auch Systeme, die sich in der Schrödingerschen Wellenmechanik nicht beschreiben lassen, z. B. Spins.

7.1 Zustände

Es ist naheliegend, den *Zustand* eines Systems durch Angabe eines minimalen Satzes von Größen zu beschreiben, der ausreicht, alle Eigenschaften festzulegen. In der klassischen Mechanik wird ein Zustand demnach durch die Angabe der Koordinaten q_i und Impulse p_i aller Teilchen charakterisiert. Wir können einen solchen Zustand präparieren, indem wir alle unabhängigen Größen q_i, p_i messen und das Experiment nur dann weiterführen, wenn sie die gewünschten Werte haben.

7.1.1 Quantenmechanische Zustände

Wie sieht das in der Quantenmechanik aus? Wir wollen im Prinzip ebenso vorgehen, wissen aber schon, dass gewisse Größen, wie z. B. Ort und Impuls eines Teilchens, nicht gleichzeitig scharf messbar sind. Wir sagen, diese Größen sind *nicht verträglich*. Wir wollen also zur Präparation eines Quantenzustands eine hinreichend große Zahl von verträglichen Größen messen. Einen so präparierten sogenannten „reinen Zustand“ bezeichnen wir abstrakt durch das Symbol $|\dots\rangle$, z. B. $|\psi\rangle$ oder $|n\rangle$, genannt *Ket-Vektor*. Die Wellenfunktion im Ortsraum, $\psi(\vec{r}, t)$, und im Impulsraum, $\tilde{\psi}(\vec{p}, t)$, sind zwei spezielle Darstellungen des Zustands $|\psi\rangle$ eines Einteilchensystems.

7.1.2 Der Hilbert-Raum

Da die quadratintegrierbaren Wellenfunktionen einen separablen Hilbert-Raum über \mathbb{C} bilden, postulieren wir, dass dies auch für die Zustandsvektoren $|\psi\rangle$ in der allgemeinen Formulierung gilt. Die Eigenschaften eines separablen Hilbert-Raums hatten wir bereits in 5.3.1 besprochen. So existiert für Elemente eines Hilbert-Raums \mathcal{H} ein Skalarprodukt, das wir jetzt als

$$\langle \bullet, \bullet \rangle : \mathcal{H} \otimes \mathcal{H} \rightarrow \mathbb{C} \quad (7.1)$$

notieren, und das folgende Eigenschaften hat:

1. $\langle \psi | \psi \rangle \in \mathbb{R}$ und $\langle \psi | \psi \rangle \geq 0$, wobei $\langle \psi | \psi \rangle = 0$ genau dann, wenn $|\psi\rangle = 0$. Wir schreiben das Nullelement von \mathcal{H} als 0 und niemals als $|0\rangle$, da $|0\rangle$ oft zur Bezeichnung des Grundzustandes, d. h. des Eigenzustandes des Hamilton-Operators mit der kleinsten Energie, verwendet wird.

$$2. \langle \psi | \varphi \rangle = \langle \varphi | \psi \rangle^*.$$

$$3. \langle \psi | \lambda_1 \varphi_1 + \lambda_2 \varphi_2 \rangle = \lambda_1 \langle \psi | \varphi_1 \rangle + \lambda_2 \langle \psi | \varphi_2 \rangle, \text{ woraus folgt } \langle \lambda_1 \psi_1 + \lambda_2 \psi_2 | \varphi \rangle = \lambda_1^* \langle \psi_1 | \varphi \rangle + \lambda_2^* \langle \psi_2 | \varphi \rangle.$$

Man kann das Skalarprodukt auch als Produkt eines Elements $\langle \psi |$ aus dem *Dualraum* \mathcal{H}^* zum Hilbertraum \mathcal{H} und eines Elements $|\psi\rangle \in \mathcal{H}$ auffassen. $\langle \psi | \in \mathcal{H}^*$ nennt man einen *Bra-Vektor*. Das Skalarprodukt $\langle \psi | \varphi \rangle$ ist dann ein *Bra(c)ket*. Die zusätzliche Nomenklatur mit Einführung des Dualraums ist jedoch nicht erforderlich, der Begriff des Skalarprodukts auf \mathcal{H} ist ausreichend.

Zwei Vektoren $|\psi\rangle, |\varphi\rangle \neq 0$ nennt man *orthogonal*, wenn $\langle \psi | \varphi \rangle = 0$ ist. Als *Norm* von $|\psi\rangle \in \mathcal{H}$ definieren wir

$$\|\psi\| := \sqrt{\langle \psi | \psi \rangle}. \quad (7.2)$$

Da \mathcal{H} ein separabler Hilbert-Raum ist, existiert eine endliche oder abzählbar unendliche Orthonormalbasis. Ist $\{|\varphi_n\rangle\}$ eine solche Orthonormalbasis, so können wir einen beliebigen Zustandsvektor $|\psi\rangle$ entwickeln:

$$|\psi\rangle = \sum_n c_n |\varphi_n\rangle. \quad (7.3)$$

Die Koeffizienten c_n erhalten wir aus

$$\langle \varphi_m | \psi \rangle = \sum_n c_n \underbrace{\langle \varphi_m | \varphi_n \rangle}_{=\delta_{mn}} = c_m. \quad (7.4)$$

Es folgt

$$|\psi\rangle = \sum_n \langle \varphi_n | \psi \rangle |\varphi_n\rangle = \sum_n |\varphi_n\rangle \langle \varphi_n | \psi \rangle. \quad (7.5)$$

Weil dies für alle $|\psi\rangle \in \mathcal{H}$ gilt, können wir als Operatoridentität schreiben

$$\sum_n |\varphi_n\rangle \langle \varphi_n| = \mathbb{1}, \quad (7.6)$$

wobei $\mathbb{1}$ der identische (Eins-) Operator ist. Dies ist die *Vollständigkeitsrelation* für die Basis $\{|\varphi_n\rangle\}$.

7.1.3 Produkträume

Oft interessieren uns zusammengesetzte Systeme, z. B. mehrere Teilchen. Was ist der Zustandsraum für zusammengesetzte Systeme? Angenommen, zwei Teilsysteme werden durch Zustände in den beiden Hilbert-Räumen \mathcal{H}_1 und \mathcal{H}_2 beschrieben. Der einfachste Fall ist, dass die möglichen Zustände für Teilsystem 1 durch das Vorhandensein von Teilsystem 2 weder eingeschränkt noch erweitert werden. Dies schließt *nicht* aus, dass die Teilsysteme wechselwirken, d. h., dass der Hamilton-Operator Terme enthält, die irreduzibel von beiden Teilsystemen abhängen. Dasselbe soll für Teilsystem 2 aufgrund von Teilsystem 1 gelten. Diese Annahmen stellen ein weiteres Postulat dar, das nicht immer erfüllt ist. Zum Beispiel können zwei ununterscheidbare Fermionen nicht in demselben Zustand sein (Pauli-Prinzip, siehe Vorlesung Quantentheorie 2). Gilt das Postulat, so ist der Zustandsraum des zusammengesetzten Systems das *Tensorprodukt* $\mathcal{H} = \mathcal{H}_1 \otimes \mathcal{H}_2$. Ist $\{|m\rangle_1\}$ eine Basis von \mathcal{H}_1 und $\{|n\rangle_2\}$ eine Basis von \mathcal{H}_2 , so bilden

$$|m, n\rangle := |m\rangle_1 |n\rangle_2 \quad (\text{man schreibt auch } |m\rangle_1 \otimes |n\rangle_2) \quad (7.7)$$

eine (Produkt-) Basis von \mathcal{H} . Die Dimension von \mathcal{H} ist demnach das Produkt der Dimensionen von \mathcal{H}_1 und \mathcal{H}_2 , wobei die Dimension von \mathcal{H} abzählbar unendlich ist, wenn die Dimension von \mathcal{H}_1 oder \mathcal{H}_2 abzählbar unendlich ist. Physikalisch ist die Produktbasis sinnvoll: Wir müssen die Quantenzahlen m und n für beide Teilsysteme angeben, um den Zustand vollständig zu beschreiben.

7.1.4 Der Dirac-Raum

Wie schon für Funktionenräume diskutiert, ist es oft praktisch, auch nicht normierbare Zustandsvektoren, genannt *uneigentliche* oder *Dirac-Zustände*, zu erlauben. Dadurch wird der Zustandsraum zum Dirac-Raum erweitert, der wegen der fehlenden Normierbarkeit kein Hilbert-Raum ist. Die Dimension des Dirac-Raums muss nicht endlich oder abzählbar unendlich sein und ist tatsächlich meist überabzählbar. Es stellt sich heraus, dass in physikalisch relevanten Fällen ein verallgemeinertes Orthonormalsystem existiert. Wegen der überabzählbaren Dimension können seine Elemente $|\varphi(\alpha)\rangle$ nur durch kontinuierliche Quantenzahlen α , o. B. d. A. aus \mathbb{R} , unterschieden werden. Es gilt eine verallgemeinerte Orthonormalitätsbedingung,

$$\langle\varphi(\alpha)|\varphi(\alpha')\rangle = \delta(\alpha - \alpha') \quad (7.8)$$

und eine verallgemeinerte Vollständigkeitsrelation

$$\int d\alpha |\varphi(\alpha)\rangle\langle\varphi(\alpha)| = \mathbb{1}. \quad (7.9)$$

Die Entwicklung in Dirac-Basisvektoren hat dann die Form

$$|\psi\rangle = \int d\alpha f(\alpha) |\varphi(\alpha)\rangle. \quad (7.10)$$

Wie in Abschnitt 5.3.3 erwähnt, ist die Fourier-Transformation von dieser Form.

Manchmal treten auch Orthonormalsysteme auf, die sowohl abzählbare als auch überabzählbare Anteile enthalten. Das ist z. B. der Fall für gebundene und ungebundene Eigenzustände eines Hamilton-Operators mit diskretem und kontinuierlichem Spektrum. Die obigen Beziehungen lauten dann:

- Orthonormalität:

$$\langle\varphi_n|\varphi_{n'}\rangle = \delta_{nn'}, \quad (7.11)$$

$$\langle\varphi(\alpha)|\varphi(\alpha')\rangle = \delta(\alpha - \alpha'), \quad (7.12)$$

$$\langle\varphi_n|\varphi(\alpha')\rangle = \langle\varphi(\alpha)|\varphi_{n'}\rangle = 0. \quad (7.13)$$

- Vollständigkeit:

$$\sum_n |\varphi_n\rangle\langle\varphi_n| + \int d\alpha |\varphi(\alpha)\rangle\langle\varphi(\alpha)| = \mathbb{1}. \quad (7.14)$$

- Entwicklung:

$$|\psi\rangle = \sum_n c_n |\varphi_n\rangle + \int d\alpha f(\alpha) |\varphi(\alpha)\rangle. \quad (7.15)$$

7.2 Lineare Operatoren

Wir hatten in der Wellenmechanik gesehen, dass Messgrößen („Observable“) durch *lineare* Operatoren auf dem Raum der Wellenfunktionen dargestellt werden. In der allgemeinen Formulierung haben wir es also mit linearen Operatoren auf dem Hilbert-Raum zu tun (wir gehen jetzt wieder von einem Hilbert-Raum, nicht einem Dirac-Raum, aus, um die entsprechenden mathematischen Eigenschaften zur Verfügung zu haben). Sei z. B. \hat{A} ein linearer Operator auf dem Raum \mathcal{H} :

$$\hat{A}: \begin{array}{l} D_{\hat{A}} \rightarrow \mathcal{H} \\ |\psi\rangle \mapsto \hat{A}|\psi\rangle, \end{array} \quad (7.16)$$

wobei $D_{\hat{A}} \subseteq \mathcal{H}$ der Definitionsbereich von \hat{A} ist. Wir betrachten, wie üblich in der Physik, den Fall, dass $D_{\hat{A}}$ *dicht* in \mathcal{H} ist. Zur Erinnerung: „dicht“ bedeutet, dass zu jedem $|\psi\rangle \in \mathcal{H}$ und $\epsilon \in \mathbb{R}$ ein $|\varphi\rangle \in D_{\hat{A}}$ existiert, so dass gilt $\|\psi - \varphi\| < \epsilon$. In Worten: Zu jedem $|\psi\rangle \in \mathcal{H}$ existieren beliebig nah benachbarte $|\varphi\rangle \in D_{\hat{A}}$. Das schließt die

Möglichkeit ein, dass $D_{\hat{A}} = \mathcal{H}$ ist. Wir erlauben hier aber $D_{\hat{A}} \subset \mathcal{H}$, um einen wichtigen mathematischen Punkt diskutieren zu können.

Linearität bedeutet

$$\hat{A}(\lambda_1|\psi_1\rangle + \lambda_2|\psi_2\rangle) = \lambda_1 \hat{A}|\psi_1\rangle + \lambda_2 \hat{A}|\psi_2\rangle \quad \forall \lambda_1, \lambda_2 \in \mathbb{C}, |\psi_1\rangle, |\psi_2\rangle \in D_{\hat{A}}. \quad (7.17)$$

Zu jedem linearen Operator \hat{A} definiert man den *adjungierten* oder *hermitesch konjugierten* Operator \hat{A}^\dagger (gesprochen „A Kreuz“ oder „A dagger“) gemäß

$$\langle \psi | \hat{A}^\dagger | \varphi \rangle := \langle \varphi | \hat{A} | \psi \rangle^* \quad (7.18)$$

für alle $|\psi\rangle \in D_{\hat{A}}$. Der Definitionsbereich $D_{\hat{A}^\dagger} \subseteq \mathcal{H}$ von \hat{A}^\dagger ist die Menge aller Vektoren $|\varphi\rangle$, für die die rechte Seite von Gl. (7.18) für alle $|\psi\rangle \in D_{\hat{A}}$ existiert. Offensichtlich ist dies äquivalent zu der Bedingung, dass $\langle \varphi | \hat{A} | \psi \rangle$ für alle $|\psi\rangle \in D_{\hat{A}}$ existiert.

Man kann zeigen, dass \hat{A}^\dagger eindeutig bestimmt ist, wenn $D_{\hat{A}}$ dicht in \mathcal{H} ist. Man kann außerdem zeigen, dass auf der Schnittmenge der jeweiligen Definitionsbereiche der beiden Seiten folgende Identitäten gelten:

$$(\hat{A}^\dagger)^\dagger = \hat{A}, \quad (7.19)$$

$$(\hat{A} + \hat{B})^\dagger = \hat{A}^\dagger + \hat{B}^\dagger, \quad (7.20)$$

$$(c\hat{A})^\dagger = c^* \hat{A}^\dagger \quad \text{für } c \in \mathbb{C}, \quad (7.21)$$

$$(\hat{A}\hat{B})^\dagger = \hat{B}^\dagger \hat{A}^\dagger. \quad (7.22)$$

Außerdem ist der zu $\hat{A}|\psi\rangle$ für $|\psi\rangle \in D_{\hat{A}}$ gehörende Bra-Vektor $\langle \psi | \hat{A}^\dagger$.

7.2.1 Hermitesche und selbstadjungierte Operatoren

Es gibt nun eine Reihe von Begriffen im Zusammenhang mit der Hermitizität von Operatoren, die in der Literatur nicht einheitlich definiert werden. Insbesondere unterscheidet sich die Verwendung zwischen der Mathematik, speziell der Funktionalanalysis, und der Physik. In der Physik werden die Unterschiede oft gänzlich verschwiegen.

Wir definieren zunächst den Begriff der Beschränktheit von linearen Operatoren: Ein linearer Operator \hat{A} ist *beschränkt*, wenn eine Zahl $M \in \mathbb{R}$ existiert, so dass gilt

$$\|\hat{A}\psi\| \leq M\|\psi\| \quad \forall |\psi\rangle \in D_{\hat{A}}. \quad (7.23)$$

Beispiele:

1. Die Multiplikation mit einer Konstanten $c \in \mathbb{C}$ ist beschränkt:

$$\|c\psi\| = \sqrt{\langle c\psi | c\psi \rangle} = \sqrt{c^* c \langle \psi | \psi \rangle} = |c| \|\psi\|, \quad (7.24)$$

wir können also $M = |c|$ wählen.

2. Die Multiplikation mit x ist auf $\mathcal{H} = L_2(\mathbb{R})$ gar kein linearer Operator, da für $\psi \in L_2(\mathbb{R})$ nicht notwendig $x\psi \in L_2(\mathbb{R})$ gilt. Gegenbeispiel:

$$\psi(x) = \frac{\sqrt{\xi}}{\sqrt{x^2 + \xi^2}}. \quad (7.25)$$

3. Die Multiplikation mit x ist für

$$D_{\hat{x}} = \{ \psi \in L_2(\mathbb{R}) \mid x\psi \in L_2(\mathbb{R}) \}, \quad (7.26)$$

also die Menge aller quadratintegrablen Funktionen, die mit x multipliziert ebenfalls quadratintegrabel sind, ein linearer Operator. Er ist aber nicht beschränkt. Denn z. B. hat

$$\psi(x) = \frac{\xi^{3/2}}{x^2 + \xi^2} \quad \text{mit } \xi > 0 \quad (7.27)$$

die Norm

$$\|\psi\| = \sqrt{\int_{-\infty}^{\infty} dx \psi^2(x)} = \sqrt{\int_{-\infty}^{\infty} dx \frac{\xi^3}{(x^2 + \xi^2)^2}} = \sqrt{\pi/2}. \quad (7.28)$$

Aber

$$\|x\psi\| = \sqrt{\int_{-\infty}^{\infty} dx x^2 \psi^2(x)} = \sqrt{\int_{-\infty}^{\infty} dx \frac{\xi^3 x^2}{(x^2 + \xi^2)^2}} = \sqrt{\pi/2} \xi \quad (7.29)$$

$$\Rightarrow \|x\psi\| = \xi \|\psi\|. \quad (7.30)$$

ξ kann beliebig groß sein, also existiert kein M , so dass $\|x\psi\| \leq M\|\psi\|$ für alle ξ .

Sei nun \hat{A} ein linearer Operator mit in \mathcal{H} dichtem Definitionsbereich $D_{\hat{A}}$ und sei \hat{A}^\dagger der zugehörige adjungierte Operator. Dann führt man folgende Begriffe ein:

1. Wenn gilt

$$\langle \psi | \hat{A} | \varphi \rangle = \langle \varphi | \hat{A} | \psi \rangle^* \quad \forall |\psi\rangle, |\varphi\rangle \in D_{\hat{A}}, \quad (7.31)$$

so heißt \hat{A} *hermitesch* (in der Mathematik oft *symmetrisch* genannt). Mit der Definition (7.18) gilt dann

$$\langle \psi | \hat{A} | \varphi \rangle = |\psi\rangle \hat{A}^\dagger |\varphi\rangle \quad \forall |\psi\rangle, |\varphi\rangle \in D_{\hat{A}}. \quad (7.32)$$

Das bedeutet aber nicht, dass \hat{A} und \hat{A}^\dagger als Operatoren gleich sind, denn Gl. (7.32) beinhaltet nur die Gleichheit der Matrixelemente auf der Teilmenge $D_{\hat{A}} \subseteq \mathcal{H}$. \hat{A} kann wirklich nur auf $|\varphi\rangle \in D_{\hat{A}}$ angewendet werden, aber in Gl. (7.32) wenden wir auch \hat{A}^\dagger nur auf $|\varphi\rangle \in D_{\hat{A}}$ an. Der Definitionsbereich $D_{\hat{A}^\dagger}$ von \hat{A}^\dagger kann jedoch *größer* sein als $D_{\hat{A}}$. $D_{\hat{A}^\dagger}$ ist die Menge aller Vektoren $|\varphi\rangle$, für die $\langle \varphi | \hat{A} | \psi \rangle$ für alle $|\psi\rangle \in D_{\hat{A}}$ existiert. Wir werden gleich ein Beispiel betrachten. $D_{\hat{A}^\dagger}$ kann aber nicht *kleiner* sein als $D_{\hat{A}}$, da die linke Seite von Gl. (7.32) nach Definition der Hermitizität existiert. Damit gilt

$$D_{\hat{A}} \subseteq D_{\hat{A}^\dagger}. \quad (7.33)$$

Mit der offensichtlichen Notation

$$|\hat{A}\psi\rangle \equiv \hat{A}|\psi\rangle \quad (7.34)$$

und $\langle \hat{A}\psi | = \langle \psi | \hat{A}^\dagger$ für den zugehörigen Bra-Vektor (s. o.) ist die Definition äquivalent zu

$$\langle \psi | \hat{A} | \varphi \rangle = \langle \varphi | \hat{A} | \psi \rangle^* = \langle \hat{A}\psi | \varphi \rangle \quad \forall |\psi\rangle, |\varphi\rangle \in D_{\hat{A}}. \quad (7.35)$$

2. Wenn gilt

$$\hat{A} = \hat{A}^\dagger, \quad (7.36)$$

dann heißt \hat{A} *selbstadjungiert*. Dies erfordert insbesondere, dass $D_{\hat{A}} = D_{\hat{A}^\dagger}$ ist. Offenbar folgt aus Selbstadjungiertheit Hermitizität, aber nicht umgekehrt.

Beispiel: Wir betrachten den Ortsoperator \hat{x} in einer Dimension. Der Definitionsbereich $D_{\hat{x}} = \{\psi \in L_2(\mathbb{R}) | x\psi \in L_2(\mathbb{R})\}$ ist *dicht* in $\mathcal{H} = L_2(\mathbb{R})$. Beweis: Es reicht hin, zu zeigen, dass für jedes $\psi \in L_2(\mathbb{R})$ eine Cauchy-Folge in $D_{\hat{x}}$ existiert, die gegen ψ konvergiert. Jedes $\psi \in L_2(\mathbb{R})$ lässt sich aber in Eigenfunktionen des harmonischen Oszillators entwickeln:

$$\psi(x) = \sum_{n=0}^{\infty} c_n \varphi_n(x). \quad (7.37)$$

Die Eigenfunktionen φ_n sind (Hermite-) Polynome multipliziert mit einer Gauß-Funktion und liegen daher in $D_{\hat{x}}$. Das gilt auch für jede endliche Teilsumme $\sum_{n=0}^N c_n \varphi_n(x)$. Die Teilsummen für verschiedene N bilden die gesuchte Cauchy-Folge. \hat{x} ist *hermitesch*: Es ist

$$\langle \psi | \hat{x} | \varphi \rangle = \int_{-\infty}^{\infty} dx \psi^*(x) x \varphi(x) = \left[\int_{-\infty}^{\infty} dx \varphi^*(x) x \psi(x) \right]^* = \langle \varphi | \hat{x} | \psi \rangle^*. \quad (7.38)$$

\hat{x} ist aber *nicht selbstadjungiert*: Der Definitionsbereich von \hat{x}^\dagger ist

$$\begin{aligned} D_{\hat{x}^\dagger} &= \left\{ \varphi \in L_2(\mathbb{R}) \left| \langle \varphi | \hat{x} | \psi \rangle^* \text{ existiert für alle } \psi \in D_{\hat{x}} \right. \right\} \\ &= \left\{ \varphi \in L_2(\mathbb{R}) \left| \langle \varphi | \hat{x} | \psi \rangle \text{ existiert für alle } \psi \in D_{\hat{x}} \right. \right\} \\ &= \left\{ \varphi \in L_2(\mathbb{R}) \left| \int_{-\infty}^{\infty} dx \varphi^*(x) x \psi(x) \text{ existiert für alle } \psi \in D_{\hat{x}} \right. \right\}. \end{aligned} \quad (7.39)$$

Aber $D_{\hat{x}^\dagger}$ ist echt größer als $D_{\hat{x}}$, denn die Funktion

$$\varphi(x) = \frac{\sqrt{\xi}}{\sqrt{x^2 + \xi^2}} \quad (7.40)$$

liegt nicht in $D_{\hat{x}}$, aber φ ist quadratintegrabel und es gilt

$$\int_{-\infty}^{\infty} dx \varphi^*(x) x \psi(x) = \int_{-\infty}^{\infty} dx \frac{\sqrt{\xi}}{\sqrt{x^2 + \xi^2}} x \psi(x) = \sqrt{\xi} \int_{-\infty}^{\infty} dx \frac{x}{\sqrt{x^2 + \xi^2}} \psi(x). \quad (7.41)$$

$\varphi(x)$ ist beschränkt, da Element von $L_2(\mathbb{R})$. Außerdem folgt aus $\psi \in D_{\hat{x}}$, dass

$$\int_{-\infty}^{\infty} dx x^2 \psi^2(x) \quad (7.42)$$

konvergiert. Daher fällt $x^2 \psi^2(x)$ für $x \rightarrow \pm\infty$ schneller als $1/|x|$ ab und $\psi(x)$ fällt schneller als $1/|x|^{3/2}$ ab. Damit konvergiert das Integral in Gl. (7.41). φ liegt also in $D_{\hat{x}^\dagger}$, aber nicht in $D_{\hat{x}}$. Damit ist

$$D_{\hat{x}^\dagger} \neq D_{\hat{x}} \quad (7.43)$$

(nämlich $D_{\hat{x}^\dagger} \supset D_{\hat{x}}$) und damit $\hat{x}^\dagger \neq \hat{x}$.

Mit ähnlichen Argumenten kann man zeigen, dass man für $\varphi(x)$ jede quadratintegrale Funktion wählen kann. Es gilt also sogar $D_{\hat{x}^\dagger} = \mathcal{H}$.

Für *beschränkte* Operatoren mit dichtem Definitionsbereich sind Hermitizität und Selbstadjungiertheit i. W. äquivalent. Man kann nämlich zeigen (Hellinger-Toeplitz-Theorem), dass man dann einen hermiteschen Operator \hat{A} auf eindeutige Weise so erweitern kann, dass $D_{\hat{A}} = \mathcal{H}$ ist. Dann muss $D_{\hat{A}^\dagger} \supseteq D_{\hat{A}}$ auch gleich \mathcal{H} sein.

Für *unbeschränkte* Operatoren ist die Unterscheidung aber wichtig. Die Existenz eines *vollständigen Orthornormalsystems* von Eigenzuständen ist nämlich nur für *selbstadjungierte* Operatoren gesichert. Beispiel: Der hermitesche, aber nicht selbstadjungierte Ortsoperator \hat{x} hat überhaupt keine Eigenfunktionen in seinem Definitionsbereich, denn für diese müsste gelten

$$0 = \Delta x^2 = \langle (x - \langle x \rangle)^2 \rangle = \int_{-\infty}^{\infty} dx (x - \langle x \rangle)^2 \psi^2(x). \quad (7.44)$$

O. B. d. A. sei $\langle x \rangle = 0$ (sonst führen wir eine Substitution $u = x - \langle x \rangle$ durch), dann folgt

$$0 = \int_{-\infty}^{\infty} dx x^2 \psi^2(x) \quad (7.45)$$

$$\Rightarrow \psi(x) = 0 \quad \forall x \neq 0 \quad (7.46)$$

$$\Rightarrow \int_{-\infty}^{\infty} dx \psi^2(x) = 0 \quad (\text{Lebesgue-Integral}). \quad (7.47)$$

Also kann $\psi(x)$ keine Eigenfunktion sein. Verallgemeinerte Eigenfunktionen von \hat{x} existieren, nämlich die δ -Distributionen

$$\psi_{x_0}(x) = \delta(x - x_0) \quad (7.48)$$

mit

$$\hat{x} \psi_{x_0}(x) = x \delta(x - x_0) = x_0 \delta(x - x_0) = x_0 \psi_{x_0}(x), \quad (7.49)$$

aber diese liegen nicht in $L_2(\mathbb{R})$ und damit auch sicher nicht in $D_{\hat{x}}$.

In der Physik spricht man meistens – korrekt – von hermiteschen Operatoren. Der Ortsoperator \hat{r} , der Impulsoperator \hat{p} und der Hamiltonian

$$\hat{H} = \frac{\hat{p}^2}{2m} + V(\hat{r}) \quad (7.50)$$

sind alle hermitesch. Aber \hat{r} und \hat{p} sind nicht selbstadjungiert. \hat{H} ist nur selbstadjungiert, wenn $\lim_{r \rightarrow \infty} V(r\hat{n}) = \infty$ für alle Einheitsvektoren \hat{n} gilt, dann hat \hat{H} nur ein diskretes Spektrum. Die Entwicklung nach einem abzählbaren Satz von Eigenzuständen funktioniert nur für selbstadjungierte Operatoren. Für hermitesche, aber nicht selbstadjungierte Operatoren ist meist eine Entwicklung nach einem überabzählbaren Satz von verallgemeinerten Eigenzuständen möglich. Diese spannen den Dirac-Raum auf. Diese Entwicklung steht auf mathematisch weniger festem Grund.

Oft findet man in Lehrbüchern der Quantentheorie für Hermitizität die Bedingung

$$\hat{A} = \hat{A}^\dagger, \quad (7.51)$$

die eigentlich Selbstadjungiertheit bedeutet. Hermitizität liegt vor, wenn die schwächere Bedingung

$$\hat{A} = \hat{A}^\dagger \quad \text{auf } D_{\hat{A}} \quad (7.52)$$

erfüllt ist. Es ist also üblich, den Gültigkeitsbereich zu unterdrücken.

Wichtig sind insbesondere die Eigenwerte und Eigenvektoren der Operatoren. $|a\rangle$ ist ein *Eigenvektor* (*Eigenzustand*) zum *Eigenwert* a des Operators \hat{A} , wenn die Eigenwertgleichung

$$\hat{A}|a\rangle = a|a\rangle \quad (7.53)$$

erfüllt ist. Wir gehen nun einige wichtige Eigenschaften hermitescher bzw. selbstadjungierter Operatoren an.

1. Für hermitesche Operatoren sind alle Eigenwerte reell. Beweis:

$$\langle a|\hat{A}|a\rangle = \langle a|a|a\rangle = a||a||^2 \quad (7.54)$$

und

$$\langle a|\hat{A}^\dagger|a\rangle = \langle a|a|a\rangle^* = a^*||a||^2. \quad (7.55)$$

Da $\hat{A} = \hat{A}^\dagger$ auf $D_{\hat{A}}$ gilt, folgt $a = a^*$, also ist $a \in \mathbb{R}$. Allgemeiner sind sogar alle Erwartungswerte reell: Für beliebiges $|\alpha\rangle \in D_{\hat{A}}$ ist

$$\langle \alpha|\hat{A}^\dagger|\alpha\rangle = \langle \alpha|\hat{A}|\alpha\rangle^* \quad (7.56)$$

und andererseits, wegen $\hat{A} = \hat{A}^\dagger$ auf $D_{\hat{A}}$,

$$\langle \alpha|\hat{A}^\dagger|\alpha\rangle = \langle \alpha|\hat{A}|\alpha\rangle, \quad (7.57)$$

also

$$\langle \alpha|\hat{A}|\alpha\rangle = \langle \alpha|\hat{A}|\alpha\rangle^*. \quad (7.58)$$

Da alle bekannten Messgrößen (Observable) nur *reelle* Messwerte zeigen, liegt es nahe, Observable nicht durch irgendwelche linearen Operatoren darzustellen, sondern durch hermitesche. Dann sind alle Erwartungswerte garantiert reell.

2. Ist \hat{A} sogar selbstadjungiert, so existiert ein *vollständiges Orthonormalsystem* (eine Orthonormalbasis) von Eigenzuständen. D. h. es existiert eine Basis $\{|a_i\rangle\}$ mit

$$\langle a_i|a_j\rangle = \delta_{ij} \quad (7.59)$$

und der Vollständigkeitsrelation

$$\sum_i |a_i\rangle\langle a_i| = \mathbb{1}. \quad (7.60)$$

Lassen wir \hat{A} auf einen Zustand $|\alpha\rangle$ wirken, so erhalten wir

$$\begin{aligned}\hat{A}|\alpha\rangle &= \sum_i \hat{A}|a_i\rangle\langle a_i|\alpha\rangle \\ &= \sum_i a_i|a_i\rangle\langle a_i|\alpha\rangle.\end{aligned}\tag{7.61}$$

Da das für alle $|\alpha\rangle \in \mathcal{H}$ gilt, folgt die *Spektraldarstellung* des Operators,

$$\hat{A} = \sum_i a_i|a_i\rangle\langle a_i|.\tag{7.62}$$

3. Zwei selbstadjungierte Operatoren \hat{A}, \hat{B} sind genau dann vertauschbar, $[\hat{A}, \hat{B}] = 0$, wenn sie ein *gemeinsames* vollständiges Orthonormalsystem von Eigenzuständen besitzen.

7.2.2 Unitäre Operatoren

Wir definieren zunächst den zu \hat{A} *inversen* Operator: Ist die Abbildung

$$|\alpha\rangle \mapsto \hat{A}|\alpha\rangle = |\beta\rangle\tag{7.63}$$

umkehrbar eindeutig, dann ist der zu \hat{A} inverse Operator \hat{A}^{-1} definiert durch

$$\hat{A}^{-1}|\beta\rangle = |\alpha\rangle.\tag{7.64}$$

Es gilt dann auf den jeweiligen Definitionsbereichen

$$\hat{A}^{-1}\hat{A} = \hat{A}\hat{A}^{-1} = \mathbb{1}\tag{7.65}$$

und

$$(\hat{A}^\dagger)^{-1} = (\hat{A}^{-1})^\dagger.\tag{7.66}$$

Schließlich definieren wir noch *unitäre* Operatoren \hat{U} durch

$$\hat{U}^\dagger\hat{U} = \hat{U}\hat{U}^\dagger = \mathbb{1}\tag{7.67}$$

oder äquivalent

$$\hat{U}^\dagger = \hat{U}^{-1}.\tag{7.68}$$

Auf ganz \mathcal{H} definierte unitäre Operatoren, $D_{\hat{U}} = D_{\hat{U}^\dagger} = \mathcal{H}$, haben, wie selbstadjungierte Operatoren, ein vollständiges Orthonormalsystem von Eigenzuständen. Ihre Eigenwerte sind jedoch komplex mit dem Betrag eins, denn für einen normierten Eigenzustand $|u_i\rangle$ gilt

$$\mathbb{1} = \langle u_i|u_i\rangle = \langle u_i|\hat{U}^\dagger\hat{U}|u_i\rangle = \langle u_i|\hat{U}^\dagger u_i|u_i\rangle = u_i\langle u_i|\hat{U}^\dagger|u_i\rangle = u_i(\langle u_i|\hat{U}|u_i\rangle)^* = u_i u_i^* \langle u_i|u_i\rangle = |u_i|^2.\tag{7.69}$$

Unter einer *unitären Transformation* versteht man die *gleichzeitige* Transformation von Zuständen gemäß

$$|\alpha\rangle \mapsto \hat{U}|\alpha\rangle, \quad \langle\alpha| \mapsto \langle\alpha|\hat{U}^\dagger\tag{7.70}$$

und von Operatoren gemäß

$$\hat{A} \mapsto \hat{U}\hat{A}\hat{U}^\dagger.\tag{7.71}$$

Die Relevanz der unitären Transformationen besteht darin, dass alle experimentell zugänglichen Größen bei solchen Transformationen unverändert bleiben. Genauer sind alle Skalarprodukte und Matrixelemente invariant unter unitären Transformationen:

$$\langle\alpha|\beta\rangle \mapsto \langle\alpha|\hat{U}^\dagger\hat{U}|\beta\rangle = \langle\alpha|\beta\rangle\tag{7.72}$$

und

$$\langle\alpha|\hat{A}|\beta\rangle \mapsto \langle\alpha|\hat{U}^\dagger\hat{U}\hat{A}\hat{U}^\dagger\hat{U}|\beta\rangle = \langle\alpha|\hat{A}|\beta\rangle.\tag{7.73}$$

7.3 Zeitentwicklung

Ein quantenmechanisches System kann in einem beliebigen Zustand $|\psi\rangle$ präpariert werden. Die eigentlich interessante Frage ist, wie es sich danach mit fortschreitender Zeit entwickelt. Wir interessieren uns also für die *Dynamik*. Die naheliegende Verallgemeinerung der Schrödinger-Gleichung aus der Wellenmechanik ist (Postulat!)

$$i\hbar \frac{d}{dt} |\psi\rangle = \hat{H}|\psi\rangle. \quad (7.74)$$

Gesucht ist $|\psi(t)\rangle$, $t > t_0$, für gegebenes $|\psi(t_0)\rangle$.

Wir schreiben

$$|\psi(t)\rangle = \hat{U}(t, t_0) |\psi(t_0)\rangle. \quad (7.75)$$

$\hat{U}(t, t_0)$ heißt *Zeitentwicklungsoperator*. Die Gleichung sagt zunächst nichts weiter aus, als dass der Zustand zur Zeit t durch den Zustand zur Zeit t_0 bestimmt ist. Wir haben das Problem natürlich nur auf die Bestimmung von $\hat{U}(t, t_0)$ abgewälzt. Der Zeitentwicklungsoperator muss folgende Eigenschaften haben:

1. $\hat{U}(t, t_0)$ ist *linear*. Da die Schrödinger-Gleichung linear ist, erfüllen ihre Lösungen das Superpositionsprinzip. Das ist aber nur sichergestellt, wenn $\hat{U}(t, t_0)$ linear ist.

2. $\hat{U}(t, t_0)$ muss die Norm erhalten:

$$\langle\psi(t)|\psi(t)\rangle = \langle\psi(t_0)|\psi(t_0)\rangle. \quad (7.76)$$

Es folgt

$$\langle\psi(t_0)|\hat{U}^\dagger\hat{U}|\psi(t_0)\rangle = \langle\psi(t_0)|\psi(t_0)\rangle. \quad (7.77)$$

Da das für alle $|\psi(t_0)\rangle$ gelten muss, folgt $\hat{U}^\dagger\hat{U} = \mathbb{1}$. \hat{U} ist also *unitär*.

3. Offensichtlich ist $\hat{U}(t_0, t_0) = \mathbb{1}$.

4. $\hat{U}(t, t_0) = \hat{U}(t, t') \hat{U}(t', t_0) \forall t' \in [t_0, t]$, dies folgt aus der Definition.

5. Ist \hat{H} zeitunabhängig, was wir hier zunächst annehmen, so ist die Wahl des Nullpunktes der Zeitmessung beliebig und $\hat{U}(t, t_0)$ kann daher nur von der Zeitdifferenz abhängen:

$$\hat{U}(t, t_0) = \hat{U}(t - t_0). \quad (7.78)$$

Aber wie sieht $\hat{U}(t, t_0)$ explizit aus? Einsetzen in die Schrödinger-Gleichung ergibt

$$i\hbar \frac{d}{dt} \hat{U}(t, t_0) |\psi(t_0)\rangle = \hat{H} \hat{U}(t, t_0) |\psi(t_0)\rangle \quad (7.79)$$

für alle $|\psi(t_0)\rangle$. Es folgt eine Differentialgleichung für eine operatorwertige Funktion,

$$i\hbar \frac{d}{dt} \hat{U}(t, t_0) = \hat{H} \hat{U}(t, t_0). \quad (7.80)$$

Die Anfangsbedingung lautet $\hat{U}(t_0, t_0) = \mathbb{1}$.

Wir machen denselben Ansatz wie für die entsprechende Gleichung für zahlenwertige Funktionen:

$$\hat{U}(t, t_0) = \exp\left(-\frac{i\hat{H}}{\hbar}(t - t_0)\right). \quad (7.81)$$

Probe:

$$i\hbar \frac{d}{dt} \hat{U}(t, t_0) = i\hbar \frac{d}{dt} \exp\left(-\frac{i\hat{H}}{\hbar}(t - t_0)\right) = \hat{H} \exp\left(-\frac{i\hat{H}}{\hbar}(t - t_0)\right) = \hat{H} \hat{U}(t, t_0). \quad (7.82)$$

Aber können wir mit Operatoren einfach so rechnen? Funktionen von Operatoren lassen sich sauber über Potenzreihenentwicklungen definieren, also z. B. für einen Operator \hat{A} :

$$e^{\hat{A}} := \sum_{n=0}^{\infty} \frac{1}{n!} \hat{A}^n, \quad (7.83)$$

was die Definition auf positiv-ganzzahlige Potenzen von Operatoren zurückführt. Diese sind aber wohldefiniert: $\hat{A}^0 = \mathbb{1}$, $\hat{A}^1 = \hat{A}$, $\hat{A}^2 = \hat{A}\hat{A}$, $\hat{A}^3 = \hat{A}\hat{A}\hat{A}$, ... Also behaupten wir eigentlich

$$\hat{U}(t, t_0) = \sum_{n=0}^{\infty} \frac{1}{n!} \left(-\frac{i}{\hbar}\right)^n \hat{H}^n (t - t_0)^n. \quad (7.84)$$

Es folgt

$$i\hbar \frac{d}{dt} \hat{U}(t, t_0) = \sum_{n=1}^{\infty} \frac{1}{n!} \left(-\frac{i}{\hbar}\right)^{n-1} \hat{H}^n n (t - t_0)^{n-1} \quad (7.85)$$

und mit $m = n - 1$

$$\begin{aligned} \dots &= \sum_{m=0}^{\infty} \frac{1}{(m+1)!} \left(-\frac{i}{\hbar}\right)^m \hat{H}^{m+1} (m+1) (t - t_0)^m = \hat{H} \sum_{m=0}^{\infty} \frac{1}{m!} \left(-\frac{i}{\hbar}\right)^m \hat{H}^m (t - t_0)^m \\ &= \hat{H} \hat{U}(t, t_0). \end{aligned} \quad (7.86)$$

Die Lösung ist also korrekt. Die Anfangsbedingung ist ebenfalls erfüllt:

$$\exp\left(-\frac{i\hat{H}}{\hbar} (t_0 - t_0)\right) = \exp(0\mathbb{1}) = \mathbb{1}. \quad (7.87)$$

Der Zeitentwicklungsoperator erfüllt die oben geforderten Bedingungen, insbesondere ist er unitär:

$$\begin{aligned} \hat{U}(t, t_0)^\dagger \hat{U}(t, t_0) &= \left[\exp\left(-\frac{i\hat{H}}{\hbar} (t - t_0)\right) \right]^\dagger \exp\left(-\frac{i\hat{H}}{\hbar} (t - t_0)\right) \\ &= \exp\left(+\frac{i\hat{H}^\dagger}{\hbar} (t - t_0)\right) \exp\left(-\frac{i\hat{H}(t - t_0)}{\hbar}\right) \stackrel{\hat{H}^\dagger = \hat{H} \text{ auf } D_{\hat{H}}}{=} \mathbb{1}, \end{aligned} \quad (7.88)$$

wobei wir die Identität

$$(e^{\hat{A}})^\dagger = \left(\sum_{n=0}^{\infty} \frac{1}{n!} \hat{A}^n\right)^\dagger = \sum_{n=0}^{\infty} \frac{1}{n!} (\hat{A}^\dagger)^n = e^{\hat{A}^\dagger} \quad (7.89)$$

und die Hermitizität von \hat{H} ausgenutzt haben. Nur für hermitesches \hat{H} ist der Zeitentwicklungsoperator unitär.

Nun können wir die Zeitabhängigkeit von uns interessierenden Größen ausrechnen, z. B. von Erwartungswerten wie

$$\langle \hat{A} \rangle(t) = \langle \psi(t) | \hat{A} | \psi(t) \rangle = \langle \psi(t_0) | \hat{U}(t, t_0)^\dagger \hat{A} \hat{U}(t, t_0) | \psi(t_0) \rangle. \quad (7.90)$$

Wir haben hier ohne weitere Diskussion die Zeitentwicklung den Zuständen zugeordnet, während die Observable keine Zeitabhängigkeit (es sei denn eine explizite) haben. Dies bezeichnet man als *Schrödinger-Bild*. Es ist jedoch nur eine von mehreren möglichen Betrachtungsweisen, wie wir sehen werden.

7.3.1 Zeitentwicklung für zeitabhängige Hamiltonians

Wir wenden uns nun dem Fall zu, dass der Hamilton-Operator $\hat{H}(t)$ explizit von der Zeit abhängt. Die hier entwickelte Methode ist in der Vielteilchentheorie und Feldtheorie von großer Bedeutung. Die Eigenschaften 1-4

des Zeitentwicklungsoperators $\hat{U}(t, t_0)$ sind davon nicht betroffen. Dagegen gilt i. A. nicht mehr $\hat{U}(t, t_0) = \hat{U}(t - t_0)$. Für den Zeitentwicklungsoperator haben wir nun die Differentialgleichung

$$i\hbar \frac{d}{dt} \hat{U}(t, t_0) = \hat{H}(t) \hat{U}(t, t_0) \quad (7.91)$$

mit der Anfangsbedingung $\hat{U}(t_0, t_0) = \mathbb{1}$.

Für zeitabhängiges $\hat{H}(t)$ ist eine andere Lösungsmethode als oben sinnvoll, nämlich *Iteration*. Wir integrieren die Differentialgleichung zunächst von t_0 bis t :

$$i\hbar \int_{t_0}^t dt_1 \frac{d}{dt_1} \hat{U}(t_1, t_0) = \int_{t_0}^t dt_1 \hat{H}(t_1) \hat{U}(t_1, t_0) \quad (7.92)$$

$$\Rightarrow i\hbar \left[\hat{U}(t, t_0) - \underbrace{\hat{U}(t_0, t_0)}_1 \right] = \int_{t_0}^t dt_1 \hat{H}(t_1) \hat{U}(t_1, t_0) \quad (7.93)$$

$$\Rightarrow \hat{U}(t, t_0) = \mathbb{1} - \frac{i}{\hbar} \int_{t_0}^t dt_1 \hat{H}(t_1) \hat{U}(t_1, t_0). \quad (7.94)$$

Diese Gleichung scheint nicht sehr hilfreich zu sein – wir haben die ursprüngliche Differentialgleichung für $\hat{U}(t, t_0)$ in eine Integralgleichung übersetzt. Der Nutzen besteht jedoch darin, dass wir diese Gleichung leicht iterieren können: wir setzen das Ergebnis für $\hat{U}(t, t_0)$ rechts wieder ein und erhalten

$$\begin{aligned} \hat{U}(t, t_0) &= \mathbb{1} - \frac{i}{\hbar} \int_{t_0}^t dt_1 \hat{H}(t_1) \left[\mathbb{1} - \frac{i}{\hbar} \int_{t_0}^{t_1} dt_2 \hat{H}(t_2) \hat{U}(t_2, t_0) \right] \\ &= \mathbb{1} - \frac{i}{\hbar} \int_{t_0}^t dt_1 \hat{H}(t_1) + \left(-\frac{i}{\hbar} \right)^2 \int_{t_0}^t dt_1 \int_{t_0}^{t_1} dt_2 \hat{H}(t_1) \hat{H}(t_2) \hat{U}(t_2, t_0). \end{aligned} \quad (7.95)$$

Wir wiederholen die Prozedur:

$$\begin{aligned} \hat{U}(t, t_0) &= \mathbb{1} - \frac{i}{\hbar} \int_{t_0}^t dt_1 \hat{H}(t_1) + \left(-\frac{i}{\hbar} \right)^2 \int_{t_0}^t dt_1 \int_{t_0}^{t_1} dt_2 \hat{H}(t_1) \hat{H}(t_2) \\ &\quad + \left(-\frac{i}{\hbar} \right)^3 \int_{t_0}^t dt_1 \int_{t_0}^{t_1} dt_2 \int_{t_0}^{t_2} dt_3 \hat{H}(t_1) \hat{H}(t_2) \hat{H}(t_3) \hat{U}(t_3, t_0) \end{aligned} \quad (7.96)$$

und erhalten nach weiterer Iteration die Lösung

$$\begin{aligned} \hat{U}(t, t_0) &= \mathbb{1} - \frac{i}{\hbar} \int_{t_0}^t dt_1 \hat{H}(t_1) + \left(-\frac{i}{\hbar} \right)^2 \int_{t_0}^t dt_1 \int_{t_0}^{t_1} dt_2 \hat{H}(t_1) \hat{H}(t_2) \\ &\quad + \left(-\frac{i}{\hbar} \right)^3 \int_{t_0}^t dt_1 \int_{t_0}^{t_1} dt_2 \int_{t_0}^{t_2} dt_3 \hat{H}(t_1) \hat{H}(t_2) \hat{H}(t_3) + \dots \end{aligned} \quad (7.97)$$

Beachte, dass die Integrationsgrenzen dafür sorgen, dass in jedem Term gilt

$$t_0 \leq \dots \leq t_3 \leq t_2 \leq t_1 \leq t. \quad (7.98)$$

Wir benennen die Integrationsvariablen um, was die zeitliche Ordnung klarer macht:

$$\begin{aligned} \hat{U}(t, t_0) &= \mathbb{1} - \frac{i}{\hbar} \int_{t_0}^t dt_1 \hat{H}(t_1) + \left(-\frac{i}{\hbar} \right)^2 \int_{t_0}^t dt_2 \int_{t_0}^{t_2} dt_1 \hat{H}(t_2) \hat{H}(t_1) \\ &\quad + \left(-\frac{i}{\hbar} \right)^3 \int_{t_0}^t dt_3 \int_{t_0}^{t_3} dt_2 \int_{t_0}^{t_2} dt_1 \hat{H}(t_3) \hat{H}(t_2) \hat{H}(t_1) + \dots, \end{aligned} \quad (7.99)$$

wobei nun gilt

$$t_0 \leq t_1 \leq t_2 \leq t_3 \leq \dots \leq t. \quad (7.100)$$

Wir sehen, dass die Argumente der Hamiltonians $\hat{H}(t_i)$ von rechts nach links anwachsen. Beachte, dass die $\hat{H}(t_i)$ zu verschiedenen Zeiten i. A. nicht kommutieren. Wir dürfen diese Reihenfolge daher nicht verändern. Die verschachtelten Integrationsgrenzen sind unpraktisch, es wäre günstiger, wenn alle Integrale von t_0 bis t liefen. Das können wir tatsächlich so schreiben, müssen dann aber für Mehrfachzählung korrigieren und für die korrekte Zeitordnung sorgen. Die Lösung lautet

$$\begin{aligned}\hat{U}(t, t_0) &= \mathbb{1} - \frac{i}{\hbar} \int_{t_0}^t dt_1 \hat{H}(t_1) \\ &+ \left(-\frac{i}{\hbar}\right)^2 \frac{1}{2!} \int_{t_0}^t dt_2 \int_{t_0}^{t_2} dt_1 \mathcal{T} \hat{H}(t_2) \hat{H}(t_1) \\ &+ \left(-\frac{i}{\hbar}\right)^3 \frac{1}{3!} \int_{t_0}^t dt_3 \int_{t_0}^{t_3} dt_2 \int_{t_0}^{t_2} dt_1 \mathcal{T} \hat{H}(t_3) \hat{H}(t_2) \hat{H}(t_1) + \dots,\end{aligned}\quad (7.101)$$

wobei \mathcal{T} der von F. Dyson eingeführte *Zeitordnungsoperator* ist. Er bringt Operatoren in zeitlich von rechts nach links aufsteigende Reihenfolge, also

$$\mathcal{T} \hat{A}(t) \hat{B}(t') := \begin{cases} \hat{A}(t) \hat{B}(t') & \text{für } t \geq t' \\ \hat{B}(t') \hat{A}(t) & \text{für } t < t' \end{cases} \quad (7.102)$$

und entsprechend für mehr Faktoren. Beachte, dass \mathcal{T} kein Operator auf dem Hilbert-Raum ist. Die demnach irreführende Bezeichnung als Zeitordnungsoperator ist aber üblich.

Wir zeigen die Korrektheit von Gl. (7.101) für den Term dritter Ordnung:

$$\begin{aligned}\frac{1}{3!} \int_{t_0}^t dt_3 \int_{t_0}^{t_3} dt_2 \int_{t_0}^{t_2} dt_1 \mathcal{T} \hat{H}(t_3) \hat{H}(t_2) \hat{H}(t_1) &= \frac{1}{3!} \left[\int_{t_0}^t dt_3 \int_{t_0}^{t_3} dt_2 \int_{t_0}^{t_2} dt_1 \hat{H}(t_3) \hat{H}(t_2) \hat{H}(t_1) \right. \\ &+ \int_{t_0}^t dt_3 \int_{t_0}^{t_3} dt_2 \int_{t_2}^{t_3} dt_1 \hat{H}(t_3) \hat{H}(t_1) \hat{H}(t_2) + \int_{t_0}^t dt_3 \int_{t_0}^{t_3} dt_2 \int_{t_3}^{t_2} dt_1 \hat{H}(t_1) \hat{H}(t_3) \hat{H}(t_2) \\ &+ \int_{t_0}^t dt_3 \int_{t_3}^{t_2} dt_2 \int_{t_0}^{t_3} dt_1 \hat{H}(t_2) \hat{H}(t_3) \hat{H}(t_1) + \int_{t_0}^t dt_3 \int_{t_3}^{t_2} dt_2 \int_{t_3}^{t_2} dt_1 \hat{H}(t_2) \hat{H}(t_1) \hat{H}(t_3) \\ &\left. + \int_{t_0}^t dt_1 \int_{t_3}^{t_2} dt_2 \int_{t_2}^{t_3} dt_1 \hat{H}(t_1) \hat{H}(t_2) \hat{H}(t_3) \right] \\ &= \frac{1}{3!} \left[\int_{t_0}^t dt_3 \int_{t_0}^{t_3} dt_2 \int_{t_0}^{t_2} dt_1 \hat{H}(t_3) \hat{H}(t_2) \hat{H}(t_1) + \int_{t_0}^t dt_3 \int_{t_0}^{t_3} dt_1 \int_{t_1}^{t_3} dt_2 \hat{H}(t_3) \hat{H}(t_2) \hat{H}(t_1) \right. \\ &+ \int_{t_0}^t dt_2 \int_{t_0}^{t_2} dt_1 \int_{t_2}^{t_3} dt_3 \hat{H}(t_3) \hat{H}(t_2) \hat{H}(t_1) + \int_{t_0}^t dt_2 \int_{t_2}^{t_3} dt_3 \int_{t_0}^{t_2} dt_1 \hat{H}(t_3) \hat{H}(t_2) \hat{H}(t_1) \\ &+ \int_{t_0}^t dt_1 \int_{t_1}^{t_3} dt_3 \int_{t_1}^{t_2} dt_2 \hat{H}(t_3) \hat{H}(t_2) \hat{H}(t_1) + \int_{t_0}^t dt_3 \int_{t_1}^{t_2} dt_2 \int_{t_2}^{t_3} dt_3 \hat{H}(t_3) \hat{H}(t_2) \hat{H}(t_1) \left. \right] \\ &= \frac{1}{3!} \left[\int_{t_0}^t dt_3 \int_{t_0}^{t_3} dt_2 \int_{t_0}^{t_2} dt_1 \hat{H}(t_3) \hat{H}(t_2) \hat{H}(t_1) + \int_{t_0}^t dt_3 \int_{t_0}^{t_3} dt_2 \int_{t_0}^{t_2} dt_1 \hat{H}(t_3) \hat{H}(t_2) \hat{H}(t_1) \right. \\ &\left. + \dots (\text{insgesamt 6 identische Terme}) \right] \\ &= \int_{t_0}^t dt_3 \int_{t_0}^{t_3} dt_2 \int_{t_0}^{t_2} dt_1 \hat{H}(t_3) \hat{H}(t_2) \hat{H}(t_1).\end{aligned}\quad (7.103)$$

Die Einführung von \mathcal{T} erlaubt es, die Reihendarstellung von $\hat{U}(t, t_0)$ formal aufzusummieren:

$$\hat{U}(t, t_0) = \sum_{n=0}^{\infty} \frac{1}{n!} \left(-\frac{i}{\hbar}\right)^n \mathcal{T} \int_{t_0}^t dt_n \hat{H}(t_n) \cdots \int_{t_0}^t dt_1 \hat{H}(t_1)$$

$$= \mathcal{T} \sum_{n=0}^{\infty} \frac{1}{n!} \left(-\frac{i}{\hbar}\right)^n \left[\int_{t_0}^t dt' \hat{H}(t') \right]^n = \mathcal{T} \exp \left(-\frac{i}{\hbar} \int_{t_0}^t dt' \hat{H}(t') \right) \quad (7.104)$$

Hier ist \mathcal{T} angewandt auf jeden Term in der Potenzreihe der Exponentialfunktion zu verstehen. $\mathcal{T} \exp(\bullet)$ nennt man *zeitgeordnete Exponentialfunktion*.

Offensichtlich ergibt sich das schon bekannte Ergebnis, wenn $\hat{H}(t)$ gar nicht zeitabhängig ist: Dann ist \mathcal{T} wirkungslos und kann weggelassen werden und es folgt

$$\hat{U}(t, t_0) = \exp \left(-\frac{i}{\hbar} \int_{t_0}^t dt' \hat{H}(t') \right) = \exp \left(-\frac{i\hat{H}}{\hbar} (t - t_0) \right). \quad (7.105)$$

7.3.2 Heisenberg-Bild

Die Idee hinter dem *Heisenberg-Bild* ist, die Zeitentwicklung den Observablen zuzuordnen und stattdessen die Zustände zeitunabhängig zu lassen. Anhand des Erwartungswertes $\langle A \rangle(t)$ lässt sich dies gut darstellen:

$$\begin{aligned} \langle \hat{A} \rangle(t) &= \langle \psi(t) | \hat{A} | \psi(t) \rangle \\ &= \langle \psi(t_0) | \hat{U}(t, t_0)^\dagger \hat{A} \hat{U}(t, t_0) | \psi(t_0) \rangle \\ &=: \langle \psi_H | \hat{A}_H(t) | \psi_H \rangle \end{aligned} \quad (7.106)$$

mit

$$|\psi_H\rangle = |\psi(t_0)\rangle = \hat{U}(t, t_0)^\dagger |\psi(t)\rangle, \quad (7.107)$$

$$\hat{A}_H(t) = \hat{U}(t, t_0)^\dagger \hat{A} \hat{U}(t, t_0). \quad (7.108)$$

Der Subskript H bezeichnet Größen im Heisenberg-Bild. Größen im Schrödinger-Bild lassen wir ohne Subskript. Der Übergang zwischen den beiden Bildern ist offenbar eine unitäre Transformation und ändert daher die beobachtbaren Größen nicht.

Im Heisenberg-Bild sind die Zustandsvektoren $|\psi_H\rangle$ konstant in der Zeit, es existiert also keine Schrödinger-Gleichung, die ihre Zeitentwicklung beschreiben würde. Andererseits benötigen wir jetzt eine Bewegungsgleichung für Observable $\hat{A}_H(t)$. Diese erhalten wir aus

$$\begin{aligned} \frac{d}{dt} \hat{A}_H(t) &= \frac{d}{dt} \left[\hat{U}(t, t_0)^\dagger \hat{A} \hat{U}(t, t_0) \right] \\ &= \left[\frac{d}{dt} \hat{U}(t, t_0) \right]^\dagger \hat{A} \hat{U}(t, t_0) + \hat{U}(t, t_0)^\dagger \underbrace{\frac{\partial \hat{A}}{\partial t}}_{\text{Ableitung nach expliziter Zeitabhängigkeit}} \hat{U}(t, t_0) + \hat{U}(t, t_0)^\dagger \hat{A} \frac{d}{dt} \hat{U}(t, t_0) \\ &= \left[-\frac{i}{\hbar} \hat{H} \hat{U}(t, t_0) \right]^\dagger \hat{A} \hat{U}(t, t_0) + \hat{U}(t, t_0)^\dagger \frac{\partial \hat{A}}{\partial t} \hat{U}(t, t_0) + \hat{U}(t, t_0)^\dagger \hat{A} \left(-\frac{i}{\hbar} \right) \hat{H} \hat{U}(t, t_0) \\ &= \frac{i}{\hbar} \hat{U}(t, t_0)^\dagger \hat{H} \hat{A} \hat{U}(t, t_0) - \frac{i}{\hbar} \hat{U}(t, t_0)^\dagger \hat{A} \hat{H} \hat{U}(t, t_0) + \hat{U}(t, t_0)^\dagger \frac{\partial \hat{A}}{\partial t} \hat{U}(t, t_0). \end{aligned} \quad (7.109)$$

Nun definieren wir

$$\frac{\partial \hat{A}_H}{\partial t} := \left(\frac{\partial \hat{A}}{\partial t} \right)_H = \hat{U}(t, t_0)^\dagger \frac{\partial \hat{A}}{\partial t} \hat{U}(t, t_0) \quad (7.110)$$

(hier wird das Symbol $\partial \bullet_H / \partial t$ definiert) und schieben in den anderen Termen eine eins ein:

$$\begin{aligned} \frac{d}{dt} \hat{A}_H(t) &= \frac{i}{\hbar} \hat{U}(t, t_0)^\dagger \hat{H} \hat{U}(t, t_0) \hat{U}(t, t_0)^\dagger \hat{A} \hat{U}(t, t_0) - \frac{i}{\hbar} \hat{U}(t, t_0)^\dagger \hat{A} \hat{U}(t, t_0) \hat{U}(t, t_0)^\dagger \hat{H} \hat{U}(t, t_0) + \frac{\partial \hat{A}_H}{\partial t} \\ &= \frac{i}{\hbar} (\hat{H}_H \hat{A}_H - \hat{A}_H \hat{H}_H) + \frac{\partial \hat{A}_H}{\partial t}. \end{aligned} \quad (7.111)$$

Es folgt

$$i\hbar \frac{d}{dt} \hat{A}_H(t) = [\hat{A}_H, \hat{H}_H] + i\hbar \frac{\partial \hat{A}_H}{\partial t}. \quad (7.112)$$

Dies ist die *Heisenberg-Gleichung*. Sie ist äquivalent zur Schrödinger-Gleichung in dem Sinn, dass sie zu denselben Vorhersagen für beobachtbare Größen führt. Beachte die Ähnlichkeit zur Bewegungsgleichung für eine beliebige Messgröße A in der klassischen Mechanik: $dA/dt = \{A, H\} + \partial A/\partial t$, wobei H hier die Hamilton-*Funktion* und $\{\bullet, \bullet\}$ die Poisson-Klammer ist. Diese Ähnlichkeit gibt Anlass zur Hoffnung, dass der klassische Grenzfall die korrekten klassischen Bewegungsgleichungen ergibt.

Für nicht explizit zeitabhängige Observable \hat{A} vereinfacht sich die Heisenberg-Gleichung zu

$$i\hbar \frac{d}{dt} \hat{A}_H(t) = [\hat{A}_H, \hat{H}_H]. \quad (7.113)$$

Wir sehen, dass eine nicht explizit zeitabhängige Observable genau dann $d\hat{A}_H/dt = 0$ erfüllt, wenn $[\hat{A}_H, \hat{H}_H] = 0$ gilt. Solche Observablen heißen *Erhaltungsgrößen*.

Wir merken noch an, dass für nicht explizit zeitabhängige Hamilton-Operatoren gilt

$$\begin{aligned} \hat{H}_H &= \hat{U}(t, t_0)^\dagger \hat{H} \hat{U}(t, t_0) \\ &= \exp\left(\frac{i\hat{H}}{\hbar}(t - t_0)\right) \hat{H} \exp\left(-\frac{i\hat{H}}{\hbar}(t - t_0)\right) \\ &= \underbrace{\exp\left(\frac{i\hat{H}}{\hbar}(t - t_0)\right) \exp\left(-\frac{i\hat{H}}{\hbar}(t - t_0)\right)}_{=1} \hat{H} = \hat{H}. \end{aligned} \quad (7.114)$$

(Die Faktoren kommutieren, weil alle Terme in ihren Potenzreihenentwicklungen nur Potenzen des einen Operators \hat{H} enthalten.) Also ist \hat{H} im Schrödinger- und im Heisenberg-Bild identisch.

7.3.3 Wechselwirkungs- oder Dirac-Bild

Es ist oft nützlich, nur einen Teil der Zeitabhängigkeit auf die Observablen zu übertragen, nämlich den einfacheren. Das ist u. a. in der Störungstheorie wichtig. Der Hamilton-Operator habe die Form

$$\hat{H} = \hat{H}_0 + \hat{H}_1, \quad (7.115)$$

wobei \hat{H}_0 einfach in dem Sinne sein soll, dass wir die zugehörige Zeitentwicklung leicht ermitteln können. \hat{H}_1 sei der „komplizierte“ Anteil. In der Störungstheorie wird \hat{H}_1 als klein angenommen, d. h. seine Matrixelemente in den relevanten Eigenzuständen von \hat{H}_0 sollen klein sein, aber in diesem Abschnitt müssen wir das nicht tun. Sei nun $\hat{U}_0(t, t_0)$ der zu \hat{H}_0 gehörige Zeitentwicklungsoperator. Dann definieren wir das *Wechselwirkungs-* oder *Dirac-Bild* durch die Transformationen

$$|\psi_D(t)\rangle := \hat{U}_0(t, t_0)^\dagger |\psi(t)\rangle, \quad (7.116)$$

$$\hat{A}_D(t) := \hat{U}_0(t, t_0)^\dagger \hat{A} \hat{U}_0(t, t_0). \quad (7.117)$$

Die Observablen erhalten also die durch \hat{H}_0 bestimmte Zeitentwicklung. Die Zeitentwicklung des Zustands ergibt sich aus

$$|\psi_D(t)\rangle = \hat{U}_0(t, t_0)^\dagger \hat{U}(t, t') |\psi(t')\rangle = \hat{U}_0(t, t_0)^\dagger \hat{U}(t, t') \hat{U}_0(t', t_0) |\psi_D(t')\rangle =: \hat{U}_D(t, t') |\psi_D(t')\rangle. \quad (7.118)$$

In $\hat{U}_D(t, t')$ taucht i. A. nicht nur \hat{H}_1 auf, weil \hat{H}_0 und \hat{H}_1 und daher auch \hat{U}_0 und \hat{U} nicht miteinander vertauschen. Die Bewegungsgleichungen für $\hat{A}_D(t)$ und $|\psi_D(t)\rangle$ erhalten wir wie folgt: Aus Gl. (7.117) folgt analog zum letzten Abschnitt

$$i\hbar \frac{d}{dt} \hat{A}_D(t) = [\hat{A}_D(t), \hat{H}_{0D}] + i\hbar \frac{\partial \hat{A}_D}{\partial t}. \quad (7.119)$$

Diese Gleichung sieht aus wie die Heisenberg-Gleichung, enthält aber nur \hat{H}_{0D} , nicht den vollen Hamilton-Operator. Ist \hat{H}_0 nicht explizit zeitabhängig, so gilt $\hat{H}_{0D} = \hat{H}_0$.

Für den Zustand folgt aus Gl. (7.116)

$$\begin{aligned}
i\hbar \frac{d}{dt} |\psi_D(t)\rangle &= i\hbar \left[\frac{d}{dt} \hat{U}_0(t, t_0) \right]^\dagger |\psi(t)\rangle + \hat{U}_0(t, t_0)^\dagger i\hbar \frac{d}{dt} |\psi(t)\rangle \\
&= i\hbar \left[-\frac{i}{\hbar} \hat{H}_0 \hat{U}_0(t, t_0) \right]^\dagger |\psi(t)\rangle + \hat{U}_0(t, t_0)^\dagger \hat{H} |\psi(t)\rangle \\
&= -\hat{U}_0(t, t_0)^\dagger \hat{H}_0 |\psi(t)\rangle + \hat{U}_0(t, t_0)^\dagger \hat{H} |\psi(t)\rangle \\
&= \hat{U}_0(t, t_0)^\dagger \hat{H}_1 |\psi(t)\rangle \\
&= \hat{U}_0(t, t_0)^\dagger \hat{H}_1 \hat{U}_0(t, t_0) |\psi_D(t)\rangle \\
&= \hat{H}_{1D}(t) |\psi_D(t)\rangle.
\end{aligned} \tag{7.120}$$

Diese Gleichung hat die Form der Schrödinger-Gleichung, enthält aber nur den „Störoperator“ \hat{H}_{1D} im Dirac-Bild. Wegen der Transformation mit \hat{U}_0 hängt \hat{H}_{1D} auch von \hat{H}_0 ab.

7.4 Messungen II

In diesem Abschnitt übertragen wir kurz das schon im Zusammenhang mit der Wellenmechanik zu Messungen gesagte in den allgemeineren Dirac-Formalismus. Das von Neumannsche Projektionspostulat sagt aus, dass als Konsequenz einer Messung das System in einem Eigenzustand $|\psi_A\rangle$ (d. h. schwankungsfrei) zur gemessenen Observablen \hat{A} ist, wobei der Messwert der zugehörige Eigenwert A ist. Es gilt also

$$\hat{A}|\psi_A\rangle = A|\psi_A\rangle. \tag{7.121}$$

Da das System i. A. vor der Messung nicht im Zustand $|\psi_A\rangle$ war, muss sich der Zustand bei der Messung sprunghaft ändern („kollabieren“). Diese Art von Zeitentwicklung wird *nicht* durch die Schrödinger-Gleichung beschrieben, d. h. sie ist *nicht* durch den Zeitentwicklungsoperator $\hat{U}(t, t_0)$ gegeben. Das sehen wir leicht wie folgt:

1. Die unitäre Zeitentwicklung mit $\hat{U}(t, t_0)$ ist stetig und sogar differenzierbar, der Kollaps ist *unstetig*.
2. Die unitäre Zeitentwicklung ist reversibel, da $\hat{U}(t_0, t) = \hat{U}(t, t_0)^{-1} = \hat{U}(t, t_0)^\dagger$ existiert. Man kann also immer auch vergangene Zustände eindeutig bestimmen. Der Kollaps ist dagegen *irreversibel*. Denn ist z. B. vor der Messung

$$|\psi(t=0^-)\rangle = \frac{c_1|A_1\rangle + c_2|A_2\rangle}{\sqrt{|c_1|^2 + |c_2|^2}} \tag{7.122}$$

mit

$$\hat{A}|A_n\rangle = A_n|A_n\rangle, \tag{7.123}$$

und nach der Messung von \hat{A}

$$|\psi(t=0^+)\rangle = |A_1\rangle, \tag{7.124}$$

so ist Information verloren gegangen, nämlich über c_1 und c_2 . Aus $|\psi(t=0^+)\rangle$ kann daher $|\psi(t=0^-)\rangle$ nicht mehr rekonstruiert werden.

Es ist sehr wichtig, sich klarzumachen, dass die Eigenwertgleichung $\hat{A}|\psi_A\rangle = A|\psi_A\rangle$ *nicht* den Messprozess beschreibt. Sie ist nur eine Bestimmungsgleichung für die möglichen Messwerte und die zugehörigen Eigenzustände. Der Prozess im Fall der Messung des Wertes A kann durch die Gleichung

$$|\psi(t=0^+)\rangle = \frac{|\psi_A\rangle \langle \psi_A | \psi(t=0^-) \rangle}{\| |\psi_A\rangle \langle \psi_A | \psi(t=0^-) \rangle \|} \tag{7.125}$$

beschrieben werden. Hier *projiziert* $|A\rangle\langle A|$ den Zustand in den durch $|A\rangle$ aufgespannten eindimensionalen Unterraum (Strahl). Der Nenner sorgt einfach für die korrekte Normierung. Die Verallgemeinerung für *entartete* Eigenwerte ist plausibel: Zu einem entarteten Eigenwert A gehören mehrere orthonormale Eigenvektoren $|A^{(1)}\rangle, |A^{(2)}\rangle, \dots$. Diese spannen einen höherdimensionalen Unterraum auf. Der Projektionsoperator auf diesen Unterraum ist

$$P_A = |A^{(1)}\rangle\langle A^{(1)}| + |A^{(2)}\rangle\langle A^{(2)}| + \dots \quad (7.126)$$

Durch die Messung des Eigenwertes A wird der Zustand in den zugehörigen Unterraum projiziert:

$$|\psi(t=0^+)\rangle = \frac{P_A |\psi(t=0^-)\rangle}{\|P_A |\psi(t=0^-)\rangle\|}. \quad (7.127)$$

Beachte, dass uns bisher nichts sagt, *welcher* Eigenwert A gemessen wird – mit der Einschränkung, dass kein Eigenwert gemessen werden kann, für den $\langle A|\psi(t=0^-)\rangle = 0$ gilt, denn dann ist $|A\rangle\langle A|\psi(t=0^-)\rangle = 0$ kein Zustand. Also: Ein Eigenwert tritt nicht als möglicher Messwert auf, wenn der zugehörige Eigenvektor orthogonal zum Anfangszustand ist.

Die über das von Neumannsche Postulat hinausgehende Wahrscheinlichkeitsdeutung sagt aus, dass der Messwert (mit der eben besprochenen Ausnahme) aus Prinzip nicht sicher vorhergesagt werden kann. Die quantitative Fassung ist die Bornsche Regel: Wie in 5.6 gezeigt, gilt

$$\langle \hat{A} \rangle = \sum_A A |\langle \psi_A | \psi \rangle|^2 \quad (7.128)$$

und Born postuliert, dass $\langle \hat{A} \rangle$ tatsächlich der Mittelwert der Messwerte für viele Messungen an identisch präparierten Systemen ist. Dann gilt für die Wahrscheinlichkeit p_A der Messung des Wertes A :

$$p_A = |\langle \psi_A | \psi \rangle|^2. \quad (7.129)$$

Die Wahrscheinlichkeitsdeutung des Messprozesses ist natürlich auch nicht mit der unitären Zeitentwicklung vereinbar:

3. Die unitäre Zeitentwicklung ist deterministisch, da $\hat{U}(t, t_0)$ eindeutig ist. Der Kollaps ist dagegen *stochastisch*.

Anscheinend enthält die Standard-Quantenmechanik zwei unterschiedliche Arten von Zeitentwicklung. Ein Problem ist sofort erkennbar: Wann ist ein Prozess eine Messung und führt zum Kollaps des Zustands und wann nicht?

7.5 Postulate der Quantenmechanik

Mit Hilfe des bisher eingeführten Formalismus können wir nun die der Quantenmechanik zugrundeliegenden Postulate in moderner Form ausdrücken. Wie immer müssen sich solche Postulate durch Vergleich mit dem Experiment bewähren.

1. Eine Observable wird durch einen *hermiteschen Operator* beschrieben. Die Observable ist dabei letztlich durch die Messapparatur und den Messprozess charakterisiert.
2. Ein reiner Zustand wird durch einen Vektor oder präziser durch einen *Strahl* im Hilbert-Raum (bzw. Dirac-Raum) beschrieben. Ein Strahl ist eine Menge $\{c|\psi\rangle \mid c \in \mathbb{C}\}$ für $|\psi\rangle \neq 0$, d. h. ein Zustandsvektor $|\psi\rangle$ ohne Beachtung der Normierung. Ein Strahl ist ein eindimensionaler Unter-Hilbert-Raum.
3. Die freie Zeitentwicklung von Zuständen im Schrödinger-Bild gehorcht der Schrödinger-Gleichung

$$i\hbar \frac{d}{dt} |\psi\rangle = \hat{H} |\psi\rangle, \quad (7.130)$$

wobei \hat{H} der Hamilton-Operator des Systems ist. Dies ist äquivalent zu einer unitären Zeitentwicklung

$$|\psi(t)\rangle = \hat{U}(t, t_0) |\psi(t_0)\rangle \quad (7.131)$$

mit $\hat{U}(t, t_0) = \exp(-i\hat{H}(t - t_0)/\hbar)$ für einen zeitunabhängigen Operator \hat{H} .

4. Die freie Zeitentwicklung gilt nicht für Messprozesse. Eine Messung ist eine Wechselwirkung des Systems mit einer Messapparatur, die durch die gemessene Observable \hat{A} charakterisiert wird. Nach der Messung wird das System durch einen Eigenzustand $|a_i\rangle$ von \hat{A} beschrieben, es gilt also

$$\hat{A}|a_i\rangle = a_i|a_i\rangle, \quad (7.132)$$

und der beobachtete Messwert ist der zugehörige Eigenwert a_i . Der Zustand ändert sich bei der Messung also i. A. sprunghaft. Dies ist der Inhalt des von Neumannschen Projektionspostulats.

5. Das Ergebnis einer Messung kann im Rahmen der Quantenmechanik i. A. nicht sicher vorhergesagt werden. Man kann jedoch die Wahrscheinlichkeit exakt angeben, am Anfangszustand $|\psi\rangle$ den Eigenwert a_i zu messen (hier unter Vernachlässigung der Entartung): Diese beträgt

$$p_{a_i} = |\langle a_i|\psi\rangle|^2 \quad (7.133)$$

(Bornsche Regel).

7.6 Verträgliche und nicht verträgliche Observable

Das Phänomen der Zustandsreduktion führt in Verbindung mit nicht verschwindenden Kommutatoren $[\hat{A}, \hat{B}] \neq 0$ zu Effekten, die kein klassisches Analogon haben. Wir hatten schon gesehen, dass nicht verträgliche Observable \hat{A} , \hat{B} kein gemeinsames vollständiges Orthonormalsystem von Eigenzuständen haben. Die Messung von \hat{B} und dann \hat{A} führt i. A. zu einem anderen Zustand als die Messung von \hat{A} und dann \hat{B} :

$$|\psi\rangle \begin{cases} \xrightarrow{\hat{B}} |b_i\rangle \text{ mit Wahrsch. } |\langle b_i|\psi\rangle|^2 & \xrightarrow{\hat{A}} |a_j\rangle \text{ mit Wahrsch. } |\langle a_j|b_i\rangle|^2 \\ \xrightarrow{\hat{A}} |a_j\rangle \text{ mit Wahrsch. } |\langle a_j|\psi\rangle|^2 & \xrightarrow{\hat{B}} |b_i\rangle \text{ mit Wahrsch. } |\langle b_i|a_j\rangle|^2 \end{cases} \quad (7.134)$$

Da $|a_j\rangle$ und $|b_i\rangle$ nicht Elemente desselben vollständigen Orthonormalsystems sind, sind sie i. A. weder identisch noch orthogonal. Die Messung von \hat{A} zerstört („löscht“) die Präparation eines Eigenzustands $|b_i\rangle$ von \hat{B} durch die vorherige Messung von \hat{B} . Sind \hat{A} und \hat{B} hingegen verträglich, so zerstört die Messung von \hat{A} den präparierten Eigenzustand nicht. In diesem Fall spielt die Reihenfolge der Messungen keine Rolle.

Allgemein definieren wir einen *vollständigen* Satz kommutierender Observabler $\hat{A}, \hat{B}, \hat{C}, \dots$ dadurch, dass sie *genau ein* gemeinsames vollständiges Orthonormalsystem von Eigenzuständen

$$|a_i, b_j, c_k, \dots\rangle \quad (7.135)$$

haben (bis auf Normierung). Dabei können wir solche Observable ignorieren, die nur trivial zusammenhängen, z. B. kommutiert $\hat{A}^2 := \hat{A}\hat{A}$ immer mit \hat{A} .

Die Messung eines vollständigen Satzes $\hat{A}, \hat{B}, \hat{C}, \dots$ liefert die maximal mögliche Information über einen Zustand. Sie präpariert zugleich einen Zustand charakterisiert durch die gemessenen Eigenwerte a_i, b_j, c_k, \dots . Es ist zu beachten, dass die Wahl des vollständigen Satzes $\hat{A}, \hat{B}, \hat{C}, \dots$ nicht eindeutig ist. Es gibt also verschiedene vollständige Sätze für dasselbe System. Beispiel: für ein Teilchen im dreidimensionalen Raum sind

$$\{\hat{x}_1, \hat{x}_2, \hat{x}_3\}, \{\hat{p}_1, \hat{p}_2, \hat{p}_3\}, \{\hat{x}_1, \hat{x}_2, \hat{p}_3\}, \{\hat{x}_1 + \hat{x}_3, \hat{p}_2, \hat{x}_1 - \hat{x}_3\}, \dots \quad (7.136)$$

vollständige Sätze kommutierender Observabler, aber nicht

$$\{\hat{x}_1, \hat{x}_2, \hat{x}_3, \hat{p}_1\}, \{\hat{x}_1, \hat{x}_2, \hat{p}_2\}, \{\hat{x}_2, \hat{p}_2\}, \dots \quad (7.137)$$

und auch nicht

$$\{\hat{x}_1, \hat{x}_2\}, \dots, \quad (7.138)$$

da \hat{x}_1 und \hat{x}_2 zwar kommutieren, aber keinen vollständigen Satz bilden. Weitere Beispiele werden wir bei der Untersuchung des Drehimpulses sehen.

7.6.1 Schwankungen

Die Schwankung

$$\Delta A := \sqrt{\langle \psi | \hat{A}^2 | \psi \rangle - \langle \psi | \hat{A} | \psi \rangle^2} = \sqrt{\langle \psi | (\hat{A} - \langle \psi | \hat{A} | \psi \rangle)^2 | \psi \rangle} \quad (7.139)$$

der Observablen \hat{A} im Zustand $|\psi\rangle$ hat einen wichtigen Zusammenhang mit den Eigenzuständen von \hat{A} : Es gilt $\Delta A = 0$ *genau dann*, wenn $|\psi\rangle$ Eigenzustand von \hat{A} ist. Beweis:

1. Ist $|\psi\rangle = |a_i\rangle$ ein normierter Eigenzustand zum Eigenwert a_i , so folgt

$$\begin{aligned} \Delta A^2 &= \langle \psi | \hat{A}^2 | \psi \rangle - \langle \psi | \hat{A} | \psi \rangle^2 = \langle a_i | \hat{A} a_i | a_i \rangle - \langle a_i | a_i | a_i \rangle^2 = a_i \langle a_i | \hat{A} | a_i \rangle - a_i^2 \langle a_i | a_i \rangle^2 \\ &= a_i^2 (\langle a_i | a_i \rangle - \langle a_i | a_i \rangle) = 0. \end{aligned} \quad (7.140)$$

2. Ist umgekehrt $\Delta A^2 = 0$, so folgt (mit $|\psi\rangle$ o. B. d. A. normiert)

$$\begin{aligned} 0 &= \langle \psi | \hat{A}^2 | \psi \rangle - \langle \psi | \hat{A} | \psi \rangle^2 = \langle \psi | \left(\hat{A} - \langle \psi | \hat{A} | \psi \rangle \right)^2 | \psi \rangle = \langle \psi | \underbrace{\sum_j |a_j\rangle \langle a_j|}_{=1} \left(\hat{A} - \langle \psi | \hat{A} | \psi \rangle \right)^2 | \psi \rangle \\ &= \langle \psi | \sum_j |a_j\rangle \langle a_j| \left(a_j - \langle \psi | \hat{A} | \psi \rangle \right)^2 | \psi \rangle = \sum_j |\langle \psi | a_j \rangle|^2 (a_j - \langle \psi | \hat{A} | \psi \rangle)^2. \end{aligned} \quad (7.141)$$

Hier wurde verwendet, dass $\langle a_j | \hat{A} = \langle a_j | a_j$ gilt, was aus Gl. (7.53) folgt. Die Summanden sind alle nicht-negativ. Daher müssen sie alle verschwinden:

$$|\langle \psi | a_j \rangle|^2 \left(a_j - \langle \psi | \hat{A} | \psi \rangle \right)^2 = 0 \quad \forall j. \quad (7.142)$$

Da die $|a_j\rangle$ eine Basis bilden, muss ein j existieren mit $\langle \psi | a_j \rangle \neq 0$. Es folgt

$$a_j - \langle \psi | \hat{A} | \psi \rangle = 0 \quad (7.143)$$

$$\Rightarrow a_j = \langle \psi | \hat{A} | \psi \rangle. \quad (7.144)$$

Damit folgt aber für alle $a_k \neq a_j$, dass gilt

$$a_k - \langle \psi | \hat{A} | \psi \rangle \neq 0 \quad (7.145)$$

$$\Rightarrow \langle \psi | a_k \rangle = 0. \quad (7.146)$$

$|\psi\rangle$ ist also eine Linearkombination von höchstens solchen $|a_k\rangle$, für die $a_k = a_j$ gilt, die also mit $|a_j\rangle$ entartet sind. Damit ist

$$|\psi\rangle = \sum_{\substack{k \\ a_k = a_j}} c_k |a_k\rangle \quad (7.147)$$

$$\Rightarrow \hat{A}|\psi\rangle = \sum_{\substack{k \\ a_k = a_j}} c_k a_j |a_k\rangle = a_j |\psi\rangle \quad (7.148)$$

und $|\psi\rangle$ ist also Eigenzustand von \hat{A} .

Die bewiesene Aussage bedeutet, dass von allen Zuständen genau die Eigenzustände eine verschwindende Schwankung haben, also *scharf messbar* sind. Ist das System nicht in einem Eigenzustand von \hat{A} , z. B. aufgrund einer vorhergehenden Messung einer unverträglichen Observable \hat{B} mit $[\hat{A}, \hat{B}] \neq 0$, so ist \hat{A} nicht scharf messbar. Wir finden also einen Zusammenhang zwischen der Unverträglichkeit von Observablen und deren Schwankungen. Dieser wird im Folgenden exakt formuliert.

7.6.2 Allgemeine Unschärferelation

Wir betrachten zwei Observable (hermitesche Operatoren) \hat{A} , \hat{B} und einen normierten Zustandsvektor $|\psi\rangle$. Hilfsweise definieren wir die beiden Vektoren

$$|\alpha\rangle := \hat{a} |\psi\rangle := (\hat{A} - \langle\psi|\hat{A}|\psi\rangle) |\psi\rangle, \quad (7.149)$$

$$|\beta\rangle := \hat{b} |\psi\rangle := (\hat{B} - \langle\psi|\hat{B}|\psi\rangle) |\psi\rangle \quad (7.150)$$

(beachte, dass hier \hat{a} und \hat{b} Operatoren sind). Die Schwarzsche Ungleichung für $|\alpha\rangle$, $|\beta\rangle$ lautet

$$\|\alpha\|^2 \|\beta\|^2 \geq |\langle\alpha|\beta\rangle|^2 = \langle\alpha|\beta\rangle \langle\beta|\alpha\rangle. \quad (7.151)$$

Da \hat{a} und \hat{b} wie \hat{A} und \hat{B} hermitesche Operatoren sind, gilt

$$\|\alpha\|^2 = \langle\alpha|\alpha\rangle = \langle\psi| \underbrace{\hat{a}^2}_{\text{hermitesch}} |\psi\rangle = \langle\psi|(\hat{A} - \langle\psi|\hat{A}|\psi\rangle)^2 |\psi\rangle = \Delta A^2, \quad (7.152)$$

$$\|\beta\|^2 = \langle\beta|\beta\rangle = \langle\psi| \underbrace{\hat{b}^2}_{\text{hermitesch}} |\psi\rangle = \langle\psi|(\hat{B} - \langle\psi|\hat{B}|\psi\rangle)^2 |\psi\rangle = \Delta B^2. \quad (7.153)$$

Die Skalarprodukte lassen sich schreiben als

$$\langle\alpha|\beta\rangle = \langle\psi|\hat{a}\hat{b}|\psi\rangle = \langle\psi| \left(\frac{\hat{a}\hat{b} + \hat{b}\hat{a}}{2} + \frac{\hat{a}\hat{b} - \hat{b}\hat{a}}{2} \right) |\psi\rangle = \frac{1}{2} \langle\psi|(\hat{a}\hat{b} + \hat{b}\hat{a})|\psi\rangle + \frac{1}{2} \langle\psi|[\hat{a}, \hat{b}]|\psi\rangle, \quad (7.154)$$

$$\langle\beta|\alpha\rangle = \langle\psi|\hat{b}\hat{a}|\psi\rangle = \langle\psi| \left(\frac{\hat{a}\hat{b} + \hat{b}\hat{a}}{2} - \frac{\hat{a}\hat{b} - \hat{b}\hat{a}}{2} \right) |\psi\rangle = \frac{1}{2} \langle\psi|(\hat{a}\hat{b} + \hat{b}\hat{a})|\psi\rangle - \frac{1}{2} \langle\psi|[\hat{a}, \hat{b}]|\psi\rangle. \quad (7.155)$$

Es folgt

$$\begin{aligned} \langle\alpha|\beta\rangle \langle\beta|\alpha\rangle &= \left(\frac{1}{2} \langle\psi|(\hat{a}\hat{b} + \hat{b}\hat{a})|\psi\rangle + \frac{1}{2} \langle\psi|[\hat{a}, \hat{b}]|\psi\rangle \right) \left(\frac{1}{2} \langle\psi|(\hat{a}\hat{b} + \hat{b}\hat{a})|\psi\rangle - \frac{1}{2} \langle\psi|[\hat{a}, \hat{b}]|\psi\rangle \right) \\ &= \frac{1}{4} (\langle\psi|(\hat{a}\hat{b} + \hat{b}\hat{a})|\psi\rangle)^2 - \frac{1}{4} (\langle\psi|[\hat{a}, \hat{b}]|\psi\rangle)^2 \\ &= \frac{1}{4} (\langle\psi|(\hat{a}\hat{b} + \hat{b}\hat{a})|\psi\rangle)^2 + \frac{1}{4} (\langle\psi|i[\hat{a}, \hat{b}]|\psi\rangle)^2. \end{aligned} \quad (7.156)$$

Nun sind $\hat{a}\hat{b} + \hat{b}\hat{a}$ und $i[\hat{a}, \hat{b}]$ hermitesch:

$$(\hat{a}\hat{b} + \hat{b}\hat{a})^\dagger = \hat{b}^\dagger \hat{a}^\dagger + \hat{a}^\dagger \hat{b}^\dagger = \hat{b}\hat{a} + \hat{a}\hat{b} = \hat{a}\hat{b} + \hat{b}\hat{a} \quad (7.157)$$

und

$$(i[\hat{a}, \hat{b}])^\dagger = -i(\hat{a}\hat{b} - \hat{b}\hat{a})^\dagger = -i(\hat{b}^\dagger \hat{a}^\dagger - \hat{a}^\dagger \hat{b}^\dagger) = -i(\hat{b}\hat{a} - \hat{a}\hat{b}) = i(\hat{a}\hat{b} - \hat{b}\hat{a}) = i[\hat{a}, \hat{b}] \quad (7.158)$$

(beachte, dass der Kommutator $[\hat{a}, \hat{b}]$ selbst nicht hermitesch ist, sondern antihermitesch: $[\hat{a}, \hat{b}]^\dagger = -[\hat{a}, \hat{b}]$). Die Erwartungswerte hermitescher Operatoren sind reell und deren Quadrate sind dann ebenfalls reell und nicht-negativ. Es folgt

$$\begin{aligned} \langle\alpha|\beta\rangle \langle\beta|\alpha\rangle &= \frac{1}{4} (\langle\psi|(\hat{a}\hat{b} + \hat{b}\hat{a})|\psi\rangle)^2 + \frac{1}{4} (\langle\psi|i[\hat{a}, \hat{b}]|\psi\rangle)^2 \\ &\geq \frac{1}{4} (\langle\psi|i[\hat{a}, \hat{b}]|\psi\rangle)^2 \\ &= \frac{1}{4} |\langle\psi|i[\hat{a}, \hat{b}]|\psi\rangle|^2 \quad \left| \quad \text{da Erwartungswert reell} \right. \\ &= \frac{1}{4} |\langle\psi|[\hat{a}, \hat{b}]|\psi\rangle|^2 = \frac{1}{4} |\langle\psi|[\hat{A}, \hat{B}]|\psi\rangle|^2. \end{aligned} \quad (7.159)$$

Wir haben hier noch benutzt, dass die Zahlen $\langle \psi | \hat{A} | \psi \rangle$ und $\langle \psi | \hat{B} | \psi \rangle$ mit jedem Operator kommutieren. Also folgt insgesamt aus Glg. (7.151)–(7.153)

$$\Delta A^2 \Delta B^2 \geq \frac{1}{4} \left| \langle \psi | [\hat{A}, \hat{B}] | \psi \rangle \right|^2 \quad (7.160)$$

und schließlich

$$\Delta A \Delta B \geq \frac{1}{2} \left| \langle \psi | [\hat{A}, \hat{B}] | \psi \rangle \right|. \quad (7.161)$$

Dies ist die *allgemeine Heisenbergsche Unschärferelation*. Offensichtlich erhalten wir wegen $[\hat{x}_i, \hat{p}_i] = i\hbar$ die bekannte Orts-Impuls-Unschärferelation als Spezialfall. (Beachte, dass $|\psi\rangle$ ein normierbarer Zustandsvektor sein sollte, wir können hier also nicht $|\psi\rangle$ als Eigenzustand von \hat{x}_i oder von \hat{p}_i wählen und folgern, dass die rechte Seite der Ungleichung Null ist.) Andererseits können wir jetzt für zwei beliebige Observable die Unschärferelation aufstellen. Insbesondere erhalten wir für verträgliche Observable, d. h. für $[\hat{A}, \hat{B}] = 0$, keine Einschränkung von ΔA und ΔB . Es ist daher im Prinzip möglich, solche Observable scharf zu messen, ohne dass sich die Messungen stören.

7.7 Ehrenfest'sches Theorem

Es wäre praktisch, Bewegungsgleichungen direkt für Größen mit intuitiver physikalischer Bedeutung, insbesondere für Erwartungswerte, zu formulieren, anstatt für abstrakte Zustände und Operatoren. Bewegungsgleichungen für beobachtbare Größen sind in allen Bildern (Schrödinger, Heisenberg, Dirac) identisch, da es sich um äquivalente Formulierungen derselben Theorie handelt. Insbesondere sind Erwartungswerte $\langle \hat{A} \rangle$ bildunabhängig. Wir betrachten die Zeitableitung, im Schrödinger-Bild,

$$\begin{aligned} i\hbar \frac{d}{dt} \langle \hat{A} \rangle &= i\hbar \frac{d}{dt} \langle \psi(t) | \hat{A} | \psi(t) \rangle \\ &= i\hbar \left[\frac{d}{dt} \langle \psi(t) | \right] \hat{A} | \psi(t) \rangle + i\hbar \langle \psi(t) | \frac{\partial \hat{A}}{\partial t} | \psi(t) \rangle + \langle \psi(t) | \hat{A} i\hbar \frac{d}{dt} | \psi(t) \rangle, \end{aligned} \quad (7.162)$$

wobei $\partial \hat{A} / \partial t$ die Ableitung nach der expliziten Zeitabhängigkeit von \hat{A} ist. Aus der Schrödinger-Gleichung

$$i\hbar \frac{d}{dt} | \psi(t) \rangle = \hat{H} | \psi(t) \rangle \quad (7.163)$$

folgt die duale Gleichung (beachte, dass der zu $\hat{B} | \psi \rangle$ duale Bra-Vektor $\langle \psi | \hat{B}^\dagger$ ist)

$$-i\hbar \frac{d}{dt} \langle \psi(t) | = \langle \psi(t) | \hat{H}^\dagger = \langle \psi(t) | \hat{H}. \quad (7.164)$$

Es folgt

$$\begin{aligned} i\hbar \frac{d}{dt} \langle \hat{A} \rangle &= -\langle \psi(t) | \hat{H} \hat{A} | \psi(t) \rangle + i\hbar \langle \psi(t) | \frac{\partial \hat{A}}{\partial t} | \psi(t) \rangle + \langle \psi(t) | \hat{A} \hat{H} | \psi(t) \rangle \\ &= \langle \psi(t) | [\hat{A}, \hat{H}] | \psi(t) \rangle + i\hbar \langle \psi(t) | \frac{\partial \hat{A}}{\partial t} | \psi(t) \rangle \\ &= \langle [\hat{A}, \hat{H}] \rangle + i\hbar \left\langle \frac{\partial \hat{A}}{\partial t} \right\rangle. \end{aligned} \quad (7.165)$$

(Im Heisenberg-Bild erhalten wir die analoge Beziehung sofort aus der Heisenberg-Gleichung.) Diese Beziehung heißt *Ehrenfest'sches Theorem*. Sie erlaubt den Vergleich der Dynamik der Erwartungswerte und der entsprechenden klassischen Größen. Dies sieht man am besten am Beispiel eines Teilchens in einer Dimension: Sei

$$\hat{H} = \frac{\hat{p}^2}{2m} + \hat{V}(x). \quad (7.166)$$

Dann gilt

$$\frac{d}{dt}\langle\hat{x}\rangle = \frac{1}{i\hbar}\langle[\hat{x}, \hat{H}]\rangle = \frac{1}{m}\langle\hat{p}\rangle \quad (7.167)$$

und

$$\frac{d}{dt}\langle\hat{p}\rangle = \frac{1}{i\hbar}\langle[\hat{p}, \hat{H}]\rangle = -\left\langle\frac{d\hat{V}}{dx}\right\rangle. \quad (7.168)$$

Beachte die Ähnlichkeit mit den Hamiltonschen Gleichungen. Mit der Kraft $\hat{F}(x) := -d\hat{V}/dx$ erhalten wir

$$\frac{d^2}{dt^2}\langle\hat{x}\rangle = \frac{1}{m}\frac{d}{dt}\langle\hat{p}\rangle = \frac{1}{m}\langle\hat{F}(x)\rangle. \quad (7.169)$$

Dies sieht natürlich der Newtonschen Bewegungsgleichung sehr ähnlich, ist aber nicht dasselbe, da i. A.

$$\langle\hat{F}(x)\rangle \neq \hat{F}(\langle\hat{x}\rangle) \quad (7.170)$$

gilt. Wir erhalten also nicht die klassische Dynamik für den Erwartungswert $\langle\hat{x}\rangle$. Wir können uns fragen, wann doch $\langle\hat{F}(x)\rangle = \hat{F}(\langle\hat{x}\rangle)$ gilt. Dies ist genau dann der Fall, wenn \hat{F} eine *lineare* Funktion von \hat{x} ist,

$$\hat{F}(\hat{x}) = -k\hat{x} + F_0. \quad (7.171)$$

Dies schließt insbesondere den *harmonischen Oszillator* und das *freie Teilchen* ein. Für diese Fälle bewegt sich der mittlere Ort $\langle\hat{x}\rangle$ also wie der Ort eines klassischen Massenpunkts.

7.7.1 Energie-Zeit-Unschärferelation

Die Energie-Zeit-Unschärferelation, die oft zusammen mit der Orts-Impuls-Unschärferelation genannt wird, hat eine davon verschiedene Interpretation. Das sieht man schon daran, dass die Zeit t in der Quantenmechanik keine Observable ist – es existiert kein „Zeit-Operator“. Sie ist vielmehr ein Parameter, der durch eine reelle Zahl dargestellt wird und innerhalb der Quantenmechanik nicht weiter begründet werden kann. Dennoch werden wir im Folgenden eine Abschätzung für typische Zeitauern und typische Energien finden.

Seien \hat{A} eine Observable und \hat{H} der Hamilton-Operator. Außerdem möge \hat{A} nicht mit \hat{H} kommutieren. Dann lautet die allgemeine Unschärferelation, siehe Abschnitt 7.6.2,

$$\Delta A \Delta H \geq \frac{1}{2} \left| \langle[\hat{A}, \hat{H}]\rangle \right|. \quad (7.172)$$

Falls \hat{A} mit \hat{H} kommutiert, ist diese Ungleichung auch richtig, aber trivial. Sei \hat{A} nun nicht explizit zeitabhängig. Dann folgt aus dem Ehrenfestschen Theorem

$$\Delta A \Delta H \geq \frac{\hbar}{2} \left| \frac{d}{dt}\langle\hat{A}\rangle \right|. \quad (7.173)$$

Wir definieren Δt_A als die charakteristische Zeit für die Dynamik von \hat{A} , nämlich als die mit der charakteristischen Änderung ΔA von \hat{A} verknüpfte Zeit:

$$\frac{\Delta A}{\Delta t_A} := \left| \frac{d}{dt}\langle\hat{A}\rangle \right|. \quad (7.174)$$

Dies ist die natürliche Art, aus von \hat{A} abgeleiteten Größen eine Größe mit der Dimension einer Zeit zu konstruieren. Es folgt

$$\Delta A \Delta H \geq \frac{\hbar}{2} \frac{\Delta A}{\Delta t_A} \quad (7.175)$$

und, falls \hat{A} im betrachteten Zustand nicht scharf messbar ist, also eine nicht verschwindende Schwankung ΔA hat,

$$\Delta H \Delta t_A \geq \frac{\hbar}{2}. \quad (7.176)$$

Die muss aber für alle Observable gelten. Wir können daher den Subskript A weglassen und erhalten die *Energie-Zeit-Unschärferelation*

$$\Delta H \Delta t \geq \frac{\hbar}{2}. \quad (7.177)$$

Eine wichtige Konsistenzprüfung ist folgende: Ist das System in einem Eigenzustand zum Hamilton-Operator, so ist $\Delta H = 0$ und es folgt $\Delta t = \infty$. Das ist vernünftig, weil solche Zustände ja gerade stationär sind, so dass alle Messwerte zeitunabhängig werden. Die Relation ist z. B. beim Teilchenzerfall wichtig, wo sie die Zerfallszeit mit der Unschärfe der Ruheenergie mc^2 der Teilchen in Beziehung setzt.

7.8 Orts- und Impulsdarstellung

Wir betrachten noch kurz, wie sich die Orts- und Impulsdarstellung der Schrödingerschen Wellenmechanik aus dem allgemeinen Formalismus ergibt. Dazu schreiben wir die Eigenwertgleichungen für Ort und Impuls auf:

$$\hat{\vec{r}}|\vec{r}\rangle = \vec{r}|\vec{r}\rangle, \quad (7.178)$$

$$\hat{\vec{p}}|\vec{p}\rangle = \vec{p}|\vec{p}\rangle. \quad (7.179)$$

Die Eigenwerte \vec{r} und \vec{p} sind kontinuierliche Variable, d. h. die Operatoren $\hat{\vec{r}}$ und $\hat{\vec{p}}$ haben kontinuierliche, überabzählbare Spektren. Daher sind $|\vec{r}\rangle$ und $|\vec{p}\rangle$ uneigentliche (Dirac-) Zustände. Sie bilden verallgemeinerte vollständige Orthonormalsysteme, so dass wir jeden Zustand nach ihnen entwickeln können:

$$|\psi\rangle = \int d^3r |\vec{r}\rangle \langle \vec{r}|\psi\rangle = \int d^3r \langle \vec{r}|\psi\rangle |\vec{r}\rangle, \quad (7.180)$$

$$|\psi\rangle = \int d^3p |\vec{p}\rangle \langle \vec{p}|\psi\rangle = \int d^3p \langle \vec{p}|\psi\rangle |\vec{p}\rangle. \quad (7.181)$$

Die Entwicklungskoeffizienten sind komplexe Funktionen von \vec{r} und \vec{p} , die wir als *Wellenfunktionen* im Orts- bzw. Impulsraum bezeichnen:

$$\psi(\vec{r}, t) := \langle \vec{r}|\psi(t)\rangle, \quad (7.182)$$

$$\tilde{\psi}(\vec{p}, t) := \langle \vec{p}|\psi(t)\rangle. \quad (7.183)$$

Dies sind dieselben Wellenfunktionen, die wir in der Wellenmechanik kennengelernt hatten.

Wie sehen die Eigenzustände $|\vec{r}\rangle$ und $|\vec{p}\rangle$ in Orts- bzw. Impulsdarstellung konkret aus? Es gilt für den Ortseigenzustand $|\vec{r}'\rangle$:

$$\langle \vec{r}|\vec{r}'\rangle = \delta(\vec{r} - \vec{r}'), \quad (7.184)$$

$$\langle \vec{p}|\vec{r}'\rangle = \frac{1}{(2\pi\hbar)^{3/2}} e^{-i\vec{p}\cdot\vec{r}'/\hbar}, \quad (7.185)$$

und für den Impulseigenzustand $|\vec{p}'\rangle$:

$$\langle \vec{r}|\vec{p}'\rangle = \frac{1}{(2\pi\hbar)^{3/2}} e^{i\vec{p}'\cdot\vec{r}/\hbar}, \quad (7.186)$$

$$\langle \vec{p}|\vec{p}'\rangle = \delta(\vec{p} - \vec{p}'). \quad (7.187)$$

Also ist die Wellenfunktion für den Ortseigenzustand in der Ortsdarstellung eine δ -Funktion. Die Wellenfunktion für den Impulseigenzustand ist in der Ortsdarstellung eine ebene Welle, wie wir schon wussten.

7.9 Der harmonische Oszillator II

In diesem Abschnitt untersuchen wir den harmonischen Oszillator – wie erwähnt vielleicht das wichtigste Modell der theoretischen Physik – im Rahmen des Dirac-Formalismus. Der Hamilton-Operator lautet

$$\hat{H} = \frac{\hat{p}^2}{2m} + \frac{1}{2} m\omega^2 \hat{x}^2. \quad (7.188)$$

Das Problem besteht darin, dass \hat{x} und \hat{p} nicht vertauschen und daher \hat{H} die Summe zweier nicht vertauschbarer Operatoren $\hat{p}^2/2m$ und $(1/2)m\omega\hat{x}^2$ ist. Wir wollen \hat{H} in eine einfachere Form bringen. Dazu führen wir neue Operatoren ein: Den *Absteigeoperator*

$$\hat{a} := \frac{1}{\sqrt{2\hbar}} \left(\sqrt{m\omega} \hat{x} + \frac{i}{\sqrt{m\omega}} \hat{p} \right) \quad (7.189)$$

und den *Aufsteigeoperator*

$$\hat{a}^\dagger := \frac{1}{\sqrt{2\hbar}} \left(\sqrt{m\omega} \hat{x} - \frac{i}{\sqrt{m\omega}} \hat{p} \right). \quad (7.190)$$

Da \hat{x} und \hat{p} hermitesch sind, ist \hat{a}^\dagger offenbar tatsächlich der adjungierte Operator zu \hat{a} . \hat{a} und \hat{a}^\dagger sind nicht hermitesch. Der Kommutator von \hat{a} und \hat{a}^\dagger ist

$$\begin{aligned} [\hat{a}, \hat{a}^\dagger] &= \frac{1}{2\hbar} \left[\sqrt{m\omega} \hat{x} + \frac{i}{\sqrt{m\omega}} \hat{p}, \sqrt{m\omega} \hat{x} - \frac{i}{\sqrt{m\omega}} \hat{p} \right] \\ &= \frac{1}{2\hbar} \left(m\omega [\hat{x}, \hat{x}] - i [\hat{x}, \hat{p}] + i [\hat{p}, \hat{x}] + \frac{1}{m\omega} [\hat{p}, \hat{p}] \right) \\ &= \frac{1}{2\hbar} (-i\hbar + i(-i\hbar)) = \frac{1}{2\hbar} (\hbar + \hbar) = 1. \end{aligned} \quad (7.191)$$

Das ist sicherlich der einfachste mögliche Kommutator abgesehen von Null.

Die Auflösung der Definitionsgleichungen nach \hat{x} und \hat{p} ergibt

$$\hat{x} = \sqrt{\frac{\hbar}{2m\omega}} (\hat{a} + \hat{a}^\dagger), \quad (7.192)$$

$$\hat{p} = -i \sqrt{\frac{\hbar m\omega}{2}} (\hat{a} - \hat{a}^\dagger). \quad (7.193)$$

Durch Einsetzen in den Hamilton-Operator erhalten wir

$$\begin{aligned} \hat{H} &= \frac{1}{2m} (-1) \frac{\hbar m\omega}{2} (\hat{a} - \hat{a}^\dagger)^2 + \frac{1}{2} m\omega^2 \frac{\hbar}{2m\omega} (\hat{a} + \hat{a}^\dagger)^2 \\ &= \frac{\hbar\omega}{4} (\cancel{\hat{a}^2} + \hat{a}\hat{a}^\dagger + \hat{a}^\dagger\hat{a} \cancel{\hat{a}^{\dagger 2}} + \hat{a}\hat{a}^\dagger + \hat{a}^\dagger\hat{a} \cancel{\hat{a}^{\dagger 2}}) \\ &= \frac{\hbar\omega}{2} (\hat{a}^\dagger\hat{a} + \hat{a}\hat{a}^\dagger) = \frac{\hbar\omega}{2} (\hat{a}^\dagger\hat{a} + \underbrace{\hat{a}\hat{a}^\dagger - \hat{a}^\dagger\hat{a}}_{=[\hat{a}, \hat{a}^\dagger]=1} + \hat{a}^\dagger\hat{a}) \\ &= \frac{\hbar\omega}{2} (2\hat{a}^\dagger\hat{a} + 1) = \hbar\omega \left(\hat{a}^\dagger\hat{a} + \frac{1}{2} \right). \end{aligned} \quad (7.194)$$

Wir haben das Problem also auf die Bestimmung der Eigenzustände und Eigenwerte von $\hat{a}^\dagger\hat{a}$ zurückgeführt. Wir definieren zunächst den neuen Operator

$$\hat{n} = \hat{a}^\dagger\hat{a}. \quad (7.195)$$

Wir schreiben die Eigenwertgleichung als

$$\hat{n}|n\rangle = n|n\rangle. \quad (7.196)$$

\hat{n} ist hermitesch per Konstruktion: $(\hat{a}^\dagger\hat{a})^\dagger = \hat{a}^\dagger\hat{a}^{\dagger\dagger} = \hat{a}^\dagger\hat{a}$. Die Eigenwerte n sind also reell. Wegen

$$\langle n|\hat{n}|n\rangle = \langle n|\hat{a}^\dagger\hat{a}|n\rangle = ||\hat{a}|n\rangle||^2 \geq 0 \quad (7.197)$$

sind die Eigenwerte n nicht-negativ.

Wir untersuchen nun die Eigenschaften von \hat{n} , \hat{a} und \hat{a}^\dagger . Wir werden nur einen Teil der Behauptungen beweisen.

1. Ist $|n\rangle$ Eigenzustand von \hat{n} zum Eigenwert n , so ist $\hat{a}^\dagger|n\rangle$ Eigenzustand zum Eigenwert $n+1$ und $\hat{a}|n\rangle$ Eigenzustand zum Eigenwert $n-1$, falls $\hat{a}^\dagger|n\rangle \neq 0$ bzw. $\hat{a}|n\rangle \neq 0$.

Beweis:

$$\hat{n}\hat{a}^\dagger|n\rangle = \hat{a}^\dagger\hat{a}\hat{a}^\dagger|n\rangle = \hat{a}^\dagger(\mathbb{1} + \hat{a}^\dagger\hat{a})|n\rangle = \hat{a}^\dagger|n\rangle + \hat{a}^\dagger n|n\rangle = (n+1)\hat{a}^\dagger|n\rangle \quad (7.198)$$

und

$$\hat{n}\hat{a}|n\rangle = \hat{a}^\dagger\hat{a}\hat{a}|n\rangle = (-\mathbb{1} + \hat{a}\hat{a}^\dagger)\hat{a}|n\rangle = -\hat{a}|n\rangle + \hat{a}n|n\rangle = (n-1)\hat{a}|n\rangle. \quad (7.199)$$

2. Die Eigenwerte von \hat{n} sind nicht entartet (ohne Beweis). Damit folgt, dass Zahlen c_{n+1} und d_{n-1} existieren mit

$$\hat{a}^\dagger|n\rangle = c_{n+1}|n+1\rangle, \quad (7.200)$$

$$\hat{a}|n\rangle = d_{n-1}|n-1\rangle. \quad (7.201)$$

Die Koeffizienten erhalten wir aus der Normierung:

$$|c_{n+1}|^2 = \langle n+1|c_{n+1}^*c_{n+1}|n+1\rangle = \langle n|\hat{a}\hat{a}^\dagger|n\rangle = \langle n|(\mathbb{1} + \hat{a}^\dagger\hat{a})|n\rangle = n+1, \quad (7.202)$$

$$|d_{n-1}|^2 = \langle n-1|d_{n-1}^*d_{n-1}|n-1\rangle = \langle n|\hat{a}^\dagger\hat{a}|n\rangle = n, \quad (7.203)$$

Die Phasen von c_{n+1} und d_{n-1} sind noch nicht bestimmt. Als Ansatz wählen wir c_{n+1} und d_{n-1} reell und nicht negativ, also $c_{n+1} = \sqrt{n+1}$ und $d_{n-1} = \sqrt{n}$. Dann ist

$$\hat{a}^\dagger|n\rangle = \sqrt{n+1}|n+1\rangle, \quad (7.204)$$

$$\hat{a}|n\rangle = \sqrt{n}|n-1\rangle. \quad (7.205)$$

Die Wahl der Phasen ist konsistent, da

$$\hat{a}^\dagger\hat{a}|n\rangle = \hat{a}^\dagger\sqrt{n}|n-1\rangle = \sqrt{n}\hat{a}^\dagger|n-1\rangle = n|n\rangle. \quad (7.206)$$

3. Der kleinste Eigenwert von \hat{n} ist Null.

Beweis: Die Eigenwertgleichung für den kleinsten Eigenwert sei

$$\hat{n}|n_{\min}\rangle = n_{\min}|n_{\min}\rangle. \quad (7.207)$$

Wäre $n_{\min} > 0$, so würde aus 2. folgen, dass

$$|n_{\min}-1\rangle = \frac{1}{\sqrt{n_{\min}}}\hat{a}|n_{\min}\rangle \quad (7.208)$$

ein normierter Eigenzustand zum Eigenwert $n_{\min}-1$ wäre. n_{\min} wäre also nicht der kleinste Eigenwert im Widerspruch zur Voraussetzung.

Da $n_{\min} < 0$ ausgeschlossen ist, weil alle Eigenwerte nicht-negativ sind, folgt $n_{\min} = 0$,

Man nennt $|0\rangle$ den *Vakuumzustand*, er ist wegen $\hat{H} = \hbar\omega(\hat{n} + 1/2)$ zugleich der Grundzustand des harmonischen Oszillators. Beachte, dass $|0\rangle \equiv |n=0\rangle$ nicht der Null-Vektor 0 ist. Es gilt

$$\hat{a}|0\rangle = 0, \quad (7.209)$$

denn sonst wäre $\hat{a}|0\rangle$ Eigenzustand von \hat{n} zum Eigenwert -1 .

4. Das Eigenwertspektrum von \hat{n} ist nach oben unbeschränkt. Der Beweis ist ähnlich wie zu 3.

Es folgt, dass (zumindest) alle nicht-negativen ganzen Zahlen Eigenwerte von \hat{n} sind und dass die zugehörigen normierten Eigenzustände

$$|n\rangle = \frac{1}{\sqrt{n!}}(\hat{a}^\dagger)^n|0\rangle \quad (7.210)$$

sind.

5. Es existieren keine nicht-ganzzahligen Eigenwerte von \hat{n} . Gäbe es nämlich einen solchen, käme man durch hinreichend häufige Anwendung von \hat{a} zu einem Eigenzustand mit negativem Eigenwert.

Damit erhalten wir das Spektrum des harmonischen Oszillators:

$$E_n = \hbar\omega \left(n + \frac{1}{2} \right), \quad n = 0, 1, 2, \dots \quad (7.211)$$

Dies hatten wir schon in Abschnitt 6.3 gesehen. Die algebraische Herleitung in diesem Abschnitt wird sich in der Vielteilchentheorie als nützlich herausstellen, siehe Vorlesung Quantentheorie 2. Dort werden wir \hat{a}^\dagger und \hat{a} als Erzeugungs- bzw. Vernichtungsoperatoren von *Bosonen* interpretieren. \hat{n} ist dann der Teilchenzahloperator der Bosonen.

Die Eigenfunktionen in Ortsdarstellung sind

$$\psi_n(x) = \langle x | n \rangle. \quad (7.212)$$

Konkret finden wir diese wie folgt: Wir substituieren zunächst

$$\xi = \sqrt{\frac{m\omega}{\hbar}} x, \quad (7.213)$$

dann vereinfachen sich \hat{a} und \hat{a}^\dagger in Ortsdarstellung zu

$$\hat{a} = \frac{1}{\sqrt{2}} \left(\xi + \frac{d}{d\xi} \right), \quad \hat{a}^\dagger = \frac{1}{\sqrt{2}} \left(\xi - \frac{d}{d\xi} \right) \quad (7.214)$$

oder korrekter in Spektraldarstellung

$$\hat{a} = \int d\xi |\xi\rangle \frac{1}{\sqrt{2}} \left(\xi + \frac{d}{d\xi} \right) \langle \xi|, \quad \hat{a}^\dagger = \int d\xi |\xi\rangle \frac{1}{\sqrt{2}} \left(\xi - \frac{d}{d\xi} \right) \langle \xi|. \quad (7.215)$$

Der Grundzustand $|0\rangle$ erfüllt $\hat{a}|0\rangle = 0$, also

$$\langle \xi | \hat{a} | 0 \rangle = \int d\xi' \underbrace{\langle \xi | \xi' \rangle}_{=\delta(\xi-\xi')} \frac{1}{\sqrt{2}} \left(\xi + \frac{d}{d\xi} \right) \langle \xi' | 0 \rangle = \frac{1}{\sqrt{2}} \left(\xi + \frac{d}{d\xi} \right) \psi_0(\xi) = 0 \quad (7.216)$$

mit der normierten Lösung

$$\psi_0(\xi) = \left(\frac{m\omega}{\pi\hbar} \right)^{1/4} e^{-\xi^2/2}. \quad (7.217)$$

Die anderen Eigenfunktionen ergeben sich jetzt zu

$$\begin{aligned} \psi_n(\xi) &= \langle \xi | n \rangle = \langle \xi | \frac{1}{\sqrt{n!}} (\hat{a}^\dagger)^n | 0 \rangle \\ &= \frac{1}{\sqrt{2^n n!}} \left(\xi - \frac{d}{d\xi} \right)^n \psi_0(\xi) \\ &= \frac{1}{\sqrt{2^n n!}} \left(\frac{m\omega}{\pi\hbar} \right)^{1/4} \left(\xi - \frac{d}{d\xi} \right)^n e^{-\xi^2/2}. \end{aligned} \quad (7.218)$$

Dies ergibt dieselben Funktion wie in Abschnitt 6.3. Man *definiert* die Hermite-Polynome auch über diese Iteration,

$$H_n(\xi) := e^{\xi^2/2} \left(\xi - \frac{d}{d\xi} \right)^n e^{-\xi^2/2}. \quad (7.219)$$

Damit erhalten wir wieder

$$\psi_n(x) = \frac{1}{\sqrt{2^n n!}} \left(\frac{m\omega}{\pi\hbar} \right)^{1/4} e^{-\xi^2/2} H_n(\xi) \quad \text{mit } \xi = \sqrt{\frac{m\omega}{\hbar}} x. \quad (7.220)$$

Kapitel 8

Quantentheorie des Drehimpulses

In diesem Kapitel untersuchen wir die Quantenmechanik von Drehimpulsen. Dieses Thema hat ein eigenes Kapitel verdient, weil neue Konzepte auftreten, die über die Quantenmechanik der Teilchenbewegung in einer Dimension hinausgehen. Außerdem sind die Überlegungen in diesem Kapitel wichtig für das Verständnis des Wasserstoff-Atoms.

8.1 Korrespondenzprinzip für den Bahndrehimpuls

Der Bahndrehimpuls eines Teilchens, $\vec{L} = \vec{r} \times \vec{p}$, sollte nach den Quantisierungsregeln von Schrödinger in den Operator

$$\hat{\vec{L}} = \hat{\vec{r}} \times \hat{\vec{p}} \quad (8.1)$$

mit den Komponenten

$$\hat{L}_i = \sum_{jk} \varepsilon_{ijk} \hat{x}_j \hat{p}_k \quad (8.2)$$

übergehen, wobei die Vertauschungsrelationen

$$[\hat{x}_j, \hat{p}_k] = i\hbar \delta_{jk}, \quad (8.3)$$

$$[\hat{x}_j, \hat{x}_k] = [\hat{p}_j, \hat{p}_k] = 0 \quad (8.4)$$

gelten. \hat{L}_i enthält nur Produkte *unterschiedlicher* Orts- und Impulskomponenten, die also vertauschen. Daher tritt keine Mehrdeutigkeit bei der Quantisierung nach Schrödinger auf. Außerdem folgt daraus, dass gilt

$$\hat{L}_i^\dagger = \sum_{jk} \varepsilon_{ijk} (\hat{x}_j \hat{p}_k)^\dagger = \sum_{jk} \varepsilon_{ijk} \hat{p}_k \hat{x}_j = \sum_{jk} \varepsilon_{ijk} \hat{x}_j \hat{p}_k = \hat{L}_i, \quad (8.5)$$

\hat{L}_i ist somit hermitesch.

Die Vertauschungsrelationen der Drehimpulskomponenten \hat{L}_i ergeben sich aus den Gleichungen (8.3) und (8.4):

$$\begin{aligned} [\hat{L}_i, \hat{L}_j] &= \sum_{klmn} [\varepsilon_{ikl} \hat{x}_k \hat{p}_l, \varepsilon_{jmn} \hat{x}_m \hat{p}_n] \\ &= \sum_{klmn} \varepsilon_{ikl} \varepsilon_{jmn} [\hat{x}_k \hat{p}_l, \hat{x}_m \hat{p}_n] \\ &= \sum_{klmn} \varepsilon_{ikl} \varepsilon_{jmn} (\hat{x}_k [\hat{p}_l, \hat{x}_m \hat{p}_n] + [\hat{x}_k, \hat{x}_m \hat{p}_n] \hat{p}_l) \\ &= \sum_{klmn} \varepsilon_{ikl} \varepsilon_{jmn} (\underbrace{\hat{x}_k \hat{x}_m [\hat{p}_l, \hat{p}_n]}_{=0} + \underbrace{\hat{x}_k [\hat{p}_l, \hat{x}_m]}_{=-i\hbar \delta_{lm}} \hat{p}_n + \underbrace{\hat{x}_m [\hat{x}_k, \hat{p}_n]}_{=i\hbar \delta_{kn}} \hat{p}_l + \underbrace{[\hat{x}_k, \hat{x}_m] \hat{p}_n \hat{p}_l}_{=0}) \end{aligned}$$

$$\begin{aligned}
&= i\hbar \sum_{klmn} \varepsilon_{ikl} \varepsilon_{jmn} (\delta_{kn} \hat{x}_m \hat{p}_l - \delta_{lm} \hat{x}_k \hat{p}_n) \\
&= i\hbar \left(\sum_{klm} \underbrace{\varepsilon_{ikl} \varepsilon_{jmk}}_{=\varepsilon_{kli} \varepsilon_{kjm}} \hat{x}_m \hat{p}_l - \sum_{kln} \underbrace{\varepsilon_{ikl} \varepsilon_{jln}}_{=\varepsilon_{lik} \varepsilon_{lnj}} \hat{x}_k \hat{p}_n \right).
\end{aligned} \tag{8.6}$$

Nun gilt

$$\sum_a \varepsilon_{abc} \varepsilon_{ade} = \delta_{bd} \delta_{ce} - \delta_{be} \delta_{cd}, \tag{8.7}$$

woraus folgt

$$\begin{aligned}
[\hat{L}_i, \hat{L}_j] &= i\hbar \left(\sum_{lm} (\delta_{lj} \delta_{im} - \delta_{lm} \delta_{ij}) \hat{x}_m \hat{p}_l - \sum_{kn} (\delta_{in} \delta_{kj} - \delta_{ij} \delta_{kn}) \hat{x}_k \hat{p}_n \right) \\
&= i\hbar \left(\hat{x}_i \hat{p}_j - \cancel{\delta_{ij} \sum_l \hat{x}_l \hat{p}_l} - \hat{x}_j \hat{p}_i + \cancel{\delta_{ij} \sum_k \hat{x}_k \hat{p}_k} \right) \\
&= i\hbar (\hat{x}_i \hat{p}_j - \hat{x}_j \hat{p}_i).
\end{aligned} \tag{8.8}$$

Andererseits ist

$$\begin{aligned}
\sum_k \varepsilon_{ijk} \hat{L}_k &= \sum_{klm} \underbrace{\varepsilon_{ijk} \varepsilon_{klm}}_{=\varepsilon_{kij} \varepsilon_{klm}} \hat{x}_l \hat{p}_m \\
&= \sum_{lm} (\delta_{il} \delta_{jm} - \delta_{im} \delta_{jl}) \hat{x}_l \hat{p}_m \\
&= \hat{x}_i \hat{p}_j - \hat{x}_j \hat{p}_i.
\end{aligned} \tag{8.9}$$

Es folgt, dass gilt

$$[\hat{L}_i, \hat{L}_j] = i\hbar \sum_k \varepsilon_{ijk} \hat{L}_k. \tag{8.10}$$

Also gilt insbesondere $[\hat{L}_x, \hat{L}_y] = i\hbar \hat{L}_z$ usw., zyklisch.

Alternativ und einfacher können wir die Vertauschungsrelation für \hat{L}_x und \hat{L}_y zeigen:

$$\begin{aligned}
[\hat{L}_x, \hat{L}_y] &= [\hat{y} \hat{p}_z - \hat{z} \hat{p}_y, \hat{z} \hat{p}_x - \hat{x} \hat{p}_z] \\
&= [\hat{y} \hat{p}_z, \hat{z} \hat{p}_x] - [\hat{y} \hat{p}_z, \hat{x} \hat{p}_z] - [\hat{z} \hat{p}_y, \hat{z} \hat{p}_x] + [\hat{z} \hat{p}_y, \hat{x} \hat{p}_z] \\
&= \hat{y} \underbrace{[\hat{p}_z, \hat{z}]}_{=-i\hbar} \hat{p}_x - 0 - 0 + \hat{x} \underbrace{[\hat{z}, \hat{p}_z]}_{=i\hbar} \hat{p}_y \\
&= i\hbar (\hat{x} \hat{p}_y - \hat{y} \hat{p}_x) = i\hbar \hat{L}_z.
\end{aligned} \tag{8.11}$$

Da die Benennung der Koordinatenachsen beliebig ist, folgt die allgemeine Relation (8.10).

Wir definieren außerdem das *Betragsquadrat* des Drehimpulses,

$$\hat{\vec{L}}^2 := \hat{L}_x^2 + \hat{L}_y^2 + \hat{L}_z^2. \tag{8.12}$$

Es ist

$$\begin{aligned}
[\hat{\vec{L}}^2, \hat{L}_x] &= [\hat{L}_x^2, \hat{L}_x] + [\hat{L}_y^2, \hat{L}_x] + [\hat{L}_z^2, \hat{L}_x] \\
&= \underbrace{[\hat{L}_x^2, \hat{L}_x]}_{=0} + \hat{L}_y \underbrace{[\hat{L}_y, \hat{L}_x]}_{=-i\hbar \hat{L}_z} + \underbrace{[\hat{L}_y, \hat{L}_x]}_{=-i\hbar \hat{L}_z} \hat{L}_y + \hat{L}_z \underbrace{[\hat{L}_z, \hat{L}_x]}_{=i\hbar \hat{L}_y} + \underbrace{[\hat{L}_z, \hat{L}_x]}_{=i\hbar \hat{L}_y} \hat{L}_z
\end{aligned}$$

$$= -i\hbar \hat{L}_y \hat{L}_z - i\hbar \hat{L}_z \hat{L}_y + i\hbar \hat{L}_z \hat{L}_y + i\hbar \hat{L}_y \hat{L}_z = 0 \quad (8.13)$$

und analog für die übrigen Komponenten, also

$$[\hat{L}^2, \hat{L}_i] = 0 \quad \text{für } i = x, y, z. \quad (8.14)$$

Also kommutiert das Betragsquadrat mit allen Komponenten des Drehimpulses, aber nicht diese untereinander. Damit sind das Betragsquadrat und *eine* beliebige Komponente verträgliche Observable, aber nicht (das Betragsquadrat und) mehr als eine Komponente.

Wir können auch feststellen, welche Eigenwerte $\hat{L}_3 \equiv \hat{L}_z$ haben kann. (Die Wahl der z -Komponente ist völlig beliebig; wegen der Isotropie des Raumes müssen wir dieselben Eigenwerte für \hat{L}_x und \hat{L}_y und auch für $\hat{e} \cdot \hat{L}$ mit einem beliebigen Einheitsvektor \hat{e} finden.) Beachte dazu

$$\hat{L}_z = \hat{x}\hat{p}_y - \hat{y}\hat{p}_x \quad (8.15)$$

in Ortsdarstellung:

$$\hat{L}_z = \frac{\hbar}{i} \left(x \frac{\partial}{\partial y} - y \frac{\partial}{\partial x} \right). \quad (8.16)$$

Nun ist es nützlich, Kugelkoordinaten zu betrachten. Es ist

$$x = r \sin \theta \cos \varphi, \quad (8.17)$$

$$y = r \sin \theta \sin \varphi, \quad (8.18)$$

$$z = r \cos \theta \quad (8.19)$$

und daher

$$\begin{aligned} \frac{\partial}{\partial \varphi} &= \frac{\partial x}{\partial \varphi} \frac{\partial}{\partial x} + \frac{\partial y}{\partial \varphi} \frac{\partial}{\partial y} + \frac{\partial z}{\partial \varphi} \frac{\partial}{\partial z} = -r \sin \theta \sin \varphi \frac{\partial}{\partial x} + r \sin \theta \cos \varphi \frac{\partial}{\partial y} + 0 \\ &= -y \frac{\partial}{\partial x} + x \frac{\partial}{\partial y} = x \frac{\partial}{\partial y} - y \frac{\partial}{\partial x}. \end{aligned} \quad (8.20)$$

Also ist

$$\hat{L}_z = \frac{\hbar}{i} \frac{\partial}{\partial \varphi} \quad (8.21)$$

in Ortsdarstellung. Dieser einfache Zusammenhang spiegelt wider, dass L_z und φ in der klassischen Mechanik kanonisch konjugierte Variablen sind. Die möglichen Eigenwerte $\hbar m$ erhalten wir aus der Gleichung

$$\frac{\hbar}{i} \frac{\partial}{\partial \varphi} \psi(r, \theta, \varphi) = \hbar m \psi(r, \theta, \varphi). \quad (8.22)$$

Es ist üblich, einen Faktor \hbar abzuspalten, um m einheitenlos zu machen. Diese Gleichung können wir mittels des (Separations-) Ansatzes

$$\psi(r, \theta, \varphi) = f(r, \theta) g(\varphi) \quad (8.23)$$

lösen:

$$\frac{1}{i} f \frac{dg}{d\varphi} = m f g \quad (8.24)$$

$$\Rightarrow \frac{dg}{d\varphi} = i m g \quad (8.25)$$

mit der Lösung

$$g(\varphi) = e^{i m \varphi}. \quad (8.26)$$

Nun muss die Wellenfunktion ψ zweimal stetig differenzierbar also erst recht selbst stetig sein. Dies erfordert

$$g(2\pi) = g(0) \Leftrightarrow e^{2\pi i m} = e^0 = 1 \Leftrightarrow m \in \mathbb{Z}. \quad (8.27)$$

Die Eigenwerte der z -Komponente des Bahndrehimpulses sind also ganzzahlige Vielfache von \hbar .

8.2 Die Drehimpulsalgebra

In diesem Abschnitt werden wir die algebraischen Eigenschaften des Drehimpulsoperators untersuchen. Es erweist sich als möglich und günstig, dies allein ausgehend von den Kommutatorrelationen zu tun, ohne die konkrete Herleitung in Abschnitt 8.1 zu beachten. Wir werden sehen, dass die mathematische Struktur zusätzliche Lösungen erlaubt, die sich nicht aus der klassischen Mechanik mittels des Korrespondenzprinzips ergeben. Es ist bemerkenswert, dass diese zusätzlichen Lösungen in der Natur tatsächlich realisiert sind, nämlich als Spin.

Wir bezeichnen jetzt jeden Vektoroperator $\hat{\vec{J}}$, der die Vertauschungsrelationen

$$[\hat{J}_j, \hat{J}_k] = i\hbar \sum_l \varepsilon_{jkl} \hat{J}_l, \quad (8.28)$$

$$[\hat{J}^2, \hat{J}_j] = 0 \quad (8.29)$$

erfüllt, als Drehimpuls. Wir definieren noch die *Leiteroperatoren*

$$\hat{J}_{\pm} := \hat{J}_x \pm i\hat{J}_y, \quad (8.30)$$

die im Folgenden nützlich sein werden. Diese Definitionen und auch die folgenden Herleitungen zeichnen eine Drehimpulskomponente, nämlich \hat{J}_z , vor den anderen aus. Das ist üblich, aber völlig beliebig. Wir könnten die gesamte Diskussion z. B. auch für \hat{J}_x durchführen und würden dann $\hat{J}_{\pm} := \hat{J}_y \pm i\hat{J}_z$ definieren.

Offenbar gilt $\hat{J}_{-} = \hat{J}_{+}^{\dagger}$. Die Leiteroperatoren erfüllen außerdem die Vertauschungsrelationen

$$[\hat{J}_{+}, \hat{J}_{-}] = 2\hbar\hat{J}_z, \quad (8.31)$$

$$[\hat{J}_z, \hat{J}_{\pm}] = \pm\hbar\hat{J}_{\pm}, \quad (8.32)$$

$$[\hat{J}^2, \hat{J}_{\pm}] = 0. \quad (8.33)$$

Wir betrachten nun das Eigenwertproblem für Drehimpulse. Wir hatten in Gl. (8.29) gefordert, dass \hat{J}^2 und eine Komponente, z. B. \hat{J}_z , verträgliche Observable sind. Daher besitzen sie ein gemeinsames vollständiges Orthonormalsystem von Eigenvektoren (Eigenzuständen) $|\alpha, m\rangle$ mit

$$\hat{J}^2 |\alpha, m\rangle = \hbar^2 \alpha |\alpha, m\rangle, \quad (8.34)$$

$$\hat{J}_z |\alpha, m\rangle = \hbar m |\alpha, m\rangle. \quad (8.35)$$

Da Drehimpulse dieselbe Dimension haben wie \hbar , sind die Zahlen α und m dimensionslos. Aufgrund der Vollständigkeit spannen die $|\alpha, m\rangle$ den Hilbert-Raum eines Drehimpulses auf.

Die Bestimmung der möglichen Eigenwerte $\hbar^2 \alpha$ und $\hbar m$ ähnelt der algebraischen Lösung des harmonischen Oszillators in Abschnitt 7.9. Wir können folgendes zeigen:

1. Mit $|\alpha, m\rangle$ sind auch $\hat{J}_{\pm} |\alpha, m\rangle$ Eigenzustände zu \hat{J}^2 mit demselben Eigenwert $\hbar^2 \alpha$ und zu \hat{J}_z mit den Eigenwerten $\hbar(m \pm 1)$, falls $\hat{J}_{\pm} |\alpha, m\rangle \neq 0$.

Beweis:

$$\hat{J}^2 \hat{J}_{\pm} |\alpha, m\rangle = \hat{J}_{\pm} \hat{J}^2 |\alpha, m\rangle = \hbar^2 \alpha \hat{J}_{\pm} |\alpha, m\rangle \quad (8.36)$$

und

$$\hat{J}_z \hat{J}_{\pm} |\alpha, m\rangle = \underbrace{(\hat{J}_z \hat{J}_{\pm} - \hat{J}_{\pm} \hat{J}_z)}_{=\pm\hbar\hat{J}_{\pm}} |\alpha, m\rangle = (\pm\hbar\hat{J}_{\pm} + \hat{J}_{\pm} \hbar m) |\alpha, m\rangle = \hbar(m \pm 1) \hat{J}_{\pm} |\alpha, m\rangle. \quad (8.37)$$

\hat{J}_{\pm} erhöht bzw. erniedrigt den Eigenwert von \hat{J}_z um eins, lässt den Eigenwert von \hat{J}^2 aber unverändert.

2. Es gilt $-\sqrt{\alpha} \leq m \leq \sqrt{\alpha}$.

Beweis:

$$\langle \alpha, m | (\hat{J}^2 - \hat{J}_z^2) | \alpha, m \rangle = \langle \alpha, m | \hbar^2 (\alpha - m^2) | \alpha, m \rangle = \hbar^2 (\alpha - m^2) \quad (8.38)$$

und andererseits

$$\langle \alpha, m | (\hat{J}^2 - \hat{J}_z^2) | \alpha, m \rangle = \langle \alpha, m | (\hat{J}_x^2 + \hat{J}_y^2) | \alpha, m \rangle \geq 0, \quad (8.39)$$

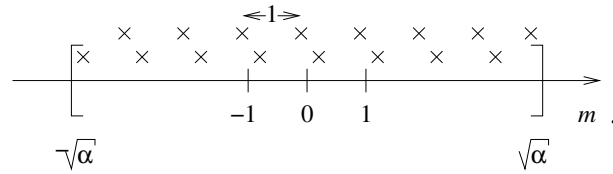
da alle Erwartungswerte von $\hat{J}_{x,y}$ reell und daher die von $\hat{J}_{x,y}^2$ reell und nichtnegativ sind. Es folgt

$$\hbar^2 (\alpha - m^2) \geq 0 \quad (8.40)$$

$$\Rightarrow m^2 \leq \alpha \quad (8.41)$$

$$\Rightarrow -\sqrt{\alpha} \leq m \leq \sqrt{\alpha}. \quad (8.42)$$

3. Wir haben gesehen, dass die Eigenzustände $|\alpha, m\rangle$ für festes α Leitern mit m -Werten im Abstand 1 im Intervall $[-\sqrt{\alpha}, \sqrt{\alpha}]$ bilden. Soviel wir bisher wissen, könnte es aber mehr als eine solche Leiter geben. In dieser Skizze z. B. zwei:



Weiter muss für $m > \sqrt{\alpha} - 1$ gelten

$$\hat{J}_+ |\alpha, m\rangle = 0, \quad (8.43)$$

denn wäre der resultierende Vektor nicht Null, so müsste er wegen Punkt 1 proportional zu $|\alpha, m+1\rangle$ sein. Hierin wäre aber $m+1 > \sqrt{\alpha}$, im Widerspruch zu Punkt 2. Entsprechend gilt

$$\hat{J}_- |\alpha, m\rangle = 0 \quad (8.44)$$

für $m < -\sqrt{\alpha} + 1$.

Nun gilt wegen 1., dass Zahlen $a_{\pm}(\alpha, m)$ existieren, die

$$\hat{J}_{\pm} |\alpha, m\rangle = a_{\pm}(\alpha, m) |\alpha, m \pm 1\rangle \quad (8.45)$$

erfüllen und die wir reell wählen. Es folgt

$$\langle \alpha, m | \hat{J}_- \hat{J}_+ | \alpha, m \rangle = \langle \alpha, m | \hat{J}_+^{\dagger} \hat{J}_+ | \alpha, m \rangle = \langle \alpha, m+1 | a_+^2(\alpha, m) | \alpha, m+1 \rangle = a_+^2(\alpha, m). \quad (8.46)$$

Andererseits zeigt man leicht, dass gilt

$$\hat{J}_+ \hat{J}_- = \hat{J}^2 - \hat{J}_z^2 + \hbar \hat{J}_z, \quad (8.47)$$

$$\hat{J}_- \hat{J}_+ = \hat{J}^2 - \hat{J}_z^2 - \hbar \hat{J}_z. \quad (8.48)$$

Also ist

$$a_+^2(\alpha, m) = \langle \alpha, m | (\hat{J}^2 - \hat{J}_z^2 - \hbar \hat{J}_z) | \alpha, m \rangle = \hbar^2 (\alpha - m^2 - m). \quad (8.49)$$

Analog findet man

$$a_-^2(\alpha, m) = \hbar^2 (\alpha - m^2 + m). \quad (8.50)$$

Also können wir

$$a_{\pm} = \hbar \sqrt{\alpha - m(m \pm 1)} \quad (8.51)$$

wählen, d. h.

$$\hat{J}_{\pm} |\alpha, m\rangle = \hbar \sqrt{\alpha - m(m \pm 1)} |\alpha, m \pm 1\rangle. \quad (8.52)$$

Wie oben gezeigt, muss für $\sqrt{\alpha} - 1 < m \leq \sqrt{\alpha}$ gelten, dass $a_+(\alpha, m) = 0$ ist. Dies ist der Fall, wenn gilt

$$\alpha = m(m+1) \quad (8.53)$$

$$\Leftrightarrow m = -\frac{1}{2} + \sqrt{\frac{1}{4} + \alpha} \quad \text{oder} \quad m = -\frac{1}{2} - \sqrt{\frac{1}{4} + \alpha}. \quad (8.54)$$

Die zweite Lösung liegt nicht im Intervall $]\sqrt{\alpha} - 1, \sqrt{\alpha}]$ und ist daher irrelevant. Es existiert also nur ein einziger Eigenwert $m = -1/2 + \sqrt{1/4 + \alpha}$ mit der Eigenschaft $\hat{J}_+ |\alpha, m\rangle = 0$. Analog findet man, dass nur ein einziger Eigenwert $m = 1/2 - \sqrt{1/4 + \alpha}$ mit $\hat{J}_- |\alpha, m\rangle = 0$ existiert. Wir folgern, dass es nur eine einzige Leiter von Zuständen mit m im Abstand 1 gibt. Weiter erfüllen das maximale bzw. minimale m die Gleichungen

$$\alpha = m_{\max}(m_{\max} + 1) = m_{\min}(m_{\min} - 1). \quad (8.55)$$

Wir nennen jetzt $m_{\max} = j$ und entsprechend $\alpha = j(j+1)$. Dann ist

$$m_{\min} = \frac{1}{2} - \sqrt{\frac{1}{4} + j(j+1)} = \frac{1}{2} - \sqrt{j^2 + j + \frac{1}{4}} = \frac{1}{2} - \left(j + \frac{1}{2}\right) = -j. \quad (8.56)$$

4. Da $|\alpha, m_{\max}\rangle$ und $|\alpha, m_{\min}\rangle$ zu derselben Leiter gehören, müssen sich $m_{\max} = j$ und $m_{\min} = -j$ um eine nichtnegative ganze Zahl unterscheiden:

$$m_{\max} - m_{\min} = 0, 1, 2, \dots, \quad (8.57)$$

woraus folgt

$$2j = 0, 1, 2, \dots \quad (8.58)$$

Damit ist j *ganz- oder halbzahlig* und $j \geq 0$. In dieser Vorlesung nennen wir j die *Länge* des Drehimpulses.

Wir folgen der üblichen Konvention und bezeichnen die Eigenzustände von nun an mit $|j, m\rangle$ oder auch $|jm\rangle$. Die bisherigen Überlegungen ergaben, dass j, m folgende Werte annehmen können:

$$j = 0, \frac{1}{2}, 1, \frac{3}{2}, 2, \dots, \quad (8.59)$$

$$m = -j, -j+1, \dots, j-1, j. \quad (8.60)$$

Letzteres sind $2j+1$ verschiedene Werte für m . Sowohl das Betragsquadrat als auch die z - oder eine andere Komponente des Drehimpulses können in der Quantenmechanik also nur diskrete Eigenwerte $\hbar^2 j(j+1)$ bzw. $\hbar m$ annehmen. Dies unterscheidet den Drehimpuls z. B. vom linearen Impuls \hat{p} , der ein kontinuierliches Spektrum hat.

Das zweite bemerkenswerte Ergebnis ist, dass die Algebra, festgelegt durch die Kommutatorrelation $[\hat{J}_j, \hat{J}_k] = i\hbar \sum_l \varepsilon_{ijk} \hat{J}_l$, auch Lösungen mit halbzahligem Eigenwerten j und m zulässt. Wir hatten oben gesehen, dass der Eigenwert m und damit auch j für einen *Bahndrehimpuls* nur ganzzahlig sein kann. Dies wurde zuerst von Lucy Mensing allgemein gezeigt [Z. Phys. **36**, 814 (1926)]. Es stellt sich die Frage, ob die halbzahligem Lösungen überhaupt physikalische Relevanz haben. Es zeigt sich, dass das tatsächlich der Fall ist. Zum Beispiel zeigt das Stern-Gerlach-Experiment, dass Silber-Atome, und letztlich Elektronen, einen Drehimpuls von $j = 1/2$ tragen, der offensichtlich kein Bahndrehimpuls sein kann. Er wird als *Eigendrehimpuls* oder *Spin* \hat{S} bezeichnet und kann bei verschiedenen Teilchen halb- oder ganzzahlig sein.

Teilchen mit halbzahligem Spin nennt man *Fermionen*, Teilchen mit ganzzahligem Spin *Bosonen*. Beispiele für elementare Fermionen sind die Leptonen (wie das Elektron) und Quarks und für Bosonen die Quanten der Wechselwirkungsfelder (Photonen, W^{\pm} , Z, Gluonen, evtl. Gravitonen) sowie das Higgs-Boson.

8.2.1 Matrixdarstellung

Wir können jetzt die Matrixelemente der relevanten Operatoren in der Basis $\{|j, m\rangle\}$ bestimmen. Dies ist nützlich für praktische Rechnungen. Zunächst gilt trivialerweise

$$\langle j, m | \hat{J}^2 | j', m' \rangle = \hbar^2 j(j+1) \delta_{jj'} \delta_{mm'}, \quad (8.61)$$

$$\langle j, m | \hat{J}_z | j', m' \rangle = \hbar m \delta_{jj'} \delta_{mm'}. \quad (8.62)$$

Aus Gl. (8.52) erhalten wir

$$\langle j, m | \hat{J}_\pm | j', m' \rangle = \hbar \sqrt{j'(j'+1) - m'(m' \pm 1)} \delta_{jj'} \delta_{m, m' \pm 1} \quad (8.63)$$

Aus $\hat{J}_\pm = \hat{J}_x \pm i\hat{J}_y$ folgt

$$\hat{J}_x = \frac{\hat{J}_+ + \hat{J}_-}{2}, \quad (8.64)$$

$$\hat{J}_y = \frac{\hat{J}_+ - \hat{J}_-}{2i}. \quad (8.65)$$

Daher ist

$$\langle j, m | \hat{J}_x | j', m' \rangle = \frac{\hbar}{2} [\sqrt{j'(j'+1) - m'(m'+1)} \delta_{m, m'+1} + \sqrt{j'(j'+1) - m'(m'-1)} \delta_{m, m'-1}] \delta_{jj'}, \quad (8.66)$$

$$\langle j, m | \hat{J}_y | j', m' \rangle = \frac{\hbar}{2i} [\sqrt{j'(j'+1) - m'(m'+1)} \delta_{m, m'+1} - \sqrt{j'(j'+1) - m'(m'-1)} \delta_{m, m'-1}] \delta_{jj'}. \quad (8.67)$$

Als Beispiel schreiben wir die Matrixdarstellung für $j = j' = 3/2$ explizit auf, wobei wir die Basisvektoren in der Reihenfolge $|\frac{3}{2}, \frac{3}{2}\rangle, |\frac{3}{2}, \frac{1}{2}\rangle, |\frac{3}{2}, -\frac{1}{2}\rangle, |\frac{3}{2}, -\frac{3}{2}\rangle$ nehmen:

$$\hat{J}_x = \hbar \begin{pmatrix} 0 & \sqrt{3}/2 & 0 & 0 \\ \sqrt{3}/2 & 0 & 1 & 0 \\ 0 & 1 & 0 & \sqrt{3}/2 \\ 0 & 0 & \sqrt{3}/2 & 0 \end{pmatrix}, \quad (8.68)$$

$$\hat{J}_y = \hbar \begin{pmatrix} 0 & -i\sqrt{3}/2 & 0 & 0 \\ i\sqrt{3}/2 & 0 & -i & 0 \\ 0 & i & 0 & -i\sqrt{3}/2 \\ 0 & 0 & i\sqrt{3}/2 & 0 \end{pmatrix}, \quad (8.69)$$

$$\hat{J}_z = \hbar \begin{pmatrix} 3/2 & 0 & 0 & 0 \\ 0 & 1/2 & 0 & 0 \\ 0 & 0 & -1/2 & 0 \\ 0 & 0 & 0 & -3/2 \end{pmatrix}, \quad (8.70)$$

$$\hat{J}^2 = \hbar^2 \begin{pmatrix} 15/4 & 0 & 0 & 0 \\ 0 & 15/4 & 0 & 0 \\ 0 & 0 & 15/4 & 0 \\ 0 & 0 & 0 & 15/4 \end{pmatrix}. \quad (8.71)$$

8.2.2 Spin 1/2

In der Vorlesung Quantentheorie 2 wird gezeigt, dass Teilchen mit dem Spin 1/2 die einfachste Möglichkeit darstellen, wenn man folgendes fordert:

- Die Dynamik soll durch eine lineare Gleichung von *erster Ordnung* in der Zeit beschrieben werden.
- Die Beschreibung soll *kovariant* sein, d. h. mit der Speziellen Relativitätstheorie verträglich.

Man findet auch, dass dieser Spin mit einem magnetischen Moment verknüpft ist. In diesem Abschnitt betrachten wir den Fall $j = s = 1/2$ genauer. Er beschreibt nicht nur die wichtigsten fermionischen Elementarteilchen, sondern ergibt auch das einfachste nichttriviale Quantensystem. Für $s = 1/2$ existieren nämlich nur zwei linear unabhängige Zustände $|s, m\rangle$:

$$|\tfrac{1}{2}, \tfrac{1}{2}\rangle =: |\uparrow\rangle \quad (8.72)$$

und

$$|\tfrac{1}{2}, -\tfrac{1}{2}\rangle =: |\downarrow\rangle. \quad (8.73)$$

Der Hilbert-Raum ist also *zweidimensional*. Zweikomponentige Vektoren nennt man *Spinoren*. Alle Operatoren können als 2×2 -Matrizen dargestellt werden. Durch Auswertung der Gleichungen (8.62), (8.66) und (8.67) findet man

$$\hat{S}_i = \frac{\hbar}{2} \sigma_i, \quad (8.74)$$

wobei σ_1, σ_2 und σ_3 die sogenannten *Pauli-Matrizen* sind:

$$\sigma_1 = \sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad (8.75)$$

$$\sigma_2 = \sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad (8.76)$$

$$\sigma_3 = \sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \quad (8.77)$$

Wir finden

$$\hat{S}^2 = \frac{\hbar^2}{4} (\sigma_1^2 + \sigma_2^2 + \sigma_3^2) = \frac{3\hbar^2}{4} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \hbar^2 \frac{1}{2} \left(\frac{1}{2} + 1 \right) \mathbb{1}, \quad (8.78)$$

was den korrekten (zweifach entarteten) Eigenwert $\hbar^2 s(s+1)$ mit $s = 1/2$ hat. Der Hamilton-Operator für einen Spin $1/2$ in einem Magnetfeld in z -Richtung lautet nun zum Beispiel

$$\hat{H} = -g\mu_B \vec{B} \cdot \hat{\vec{S}} = -g\mu_B B \hat{S}_z = -g\mu_B \frac{\hbar}{2} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \quad (8.79)$$

wobei $g \approx 2$ der g -Faktor und μ_B das Bohrsche Magneton sind. Ein Wert von $g = 2$ ergibt sich in der relativistischen Dirac-Quantenmechanik, siehe Vorlesung Quantentheorie 2.

8.2.3 Die Bloch-Kugel

Es stellt sich die Frage, wie der Zustandsvektor $|\psi\rangle \in \mathcal{H} = \mathbb{C}^2$ eines Spins der Länge $1/2$ mit der Richtung des Spins im Realraum \mathbb{R}^3 zusammenhängt. Wir können Spins eine Richtung im Realraum zuordnen, weil sie an das \vec{B} -Feld koppeln. Die Richtung kann durch einen Einheitsvektor \hat{n} (kein Operator!) beschrieben werden, den wir durch sphärische Winkel parametrisieren: In kartesischen Koordinaten sei

$$\hat{n} = \begin{pmatrix} \sin \vartheta \cos \varphi \\ \sin \vartheta \sin \varphi \\ \cos \vartheta \end{pmatrix}. \quad (8.80)$$

Alle \hat{n} bilden natürlich eine Einheitskugel im \mathbb{R}^3 , die *Bloch-Kugel*.

Wir betrachten nun Zustände

$$|\vartheta, \varphi\rangle := \cos \frac{\vartheta}{2} |\uparrow\rangle + e^{i\varphi} \sin \frac{\vartheta}{2} |\downarrow\rangle. \quad (8.81)$$

Der Erwartungswert des Spin-Vektors in solchen Zuständen ist

$$\langle \hat{\vec{S}} \rangle = \langle \vartheta, \varphi | \hat{\vec{S}} | \vartheta, \varphi \rangle = \frac{\hbar}{2} \langle \vartheta, \varphi | \begin{pmatrix} \sigma_x \\ \sigma_y \\ \sigma_z \end{pmatrix} | \vartheta, \varphi \rangle \quad (8.82)$$

mit den Komponenten

$$\begin{aligned}
\langle \hat{S}_x \rangle &= \frac{\hbar}{2} \left(\cos \frac{\vartheta}{2} \langle \uparrow | + e^{-i\varphi} \sin \frac{\vartheta}{2} \langle \downarrow | \right) \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \left(\cos \frac{\vartheta}{2} | \uparrow \rangle + e^{i\varphi} \sin \frac{\vartheta}{2} | \downarrow \rangle \right) \\
&= \frac{\hbar}{2} \left(\cos \frac{\vartheta}{2} e^{i\varphi} \sin \frac{\vartheta}{2} + e^{-i\varphi} \sin \frac{\vartheta}{2} \cos \frac{\vartheta}{2} \right) \\
&= \hbar \cos \frac{\vartheta}{2} \sin \frac{\vartheta}{2} \cos \varphi = \frac{\hbar}{2} \sin \vartheta \cos \varphi,
\end{aligned} \tag{8.83}$$

$$\begin{aligned}
\langle \hat{S}_y \rangle &= \frac{\hbar}{2} \left(\cos \frac{\vartheta}{2} \langle \uparrow | + e^{-i\varphi} \sin \frac{\vartheta}{2} \langle \downarrow | \right) \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \left(\cos \frac{\vartheta}{2} | \uparrow \rangle + e^{i\varphi} \sin \frac{\vartheta}{2} | \downarrow \rangle \right) \\
&= \frac{\hbar}{2} \left(-i \cos \frac{\vartheta}{2} e^{i\varphi} \sin \frac{\vartheta}{2} + i e^{-i\varphi} \sin \frac{\vartheta}{2} \cos \frac{\vartheta}{2} \right) \\
&= \hbar \cos \frac{\vartheta}{2} \sin \frac{\vartheta}{2} \sin \varphi = \frac{\hbar}{2} \sin \vartheta \sin \varphi,
\end{aligned} \tag{8.84}$$

$$\begin{aligned}
\langle \hat{S}_z \rangle &= \frac{\hbar}{2} \left(\cos \frac{\vartheta}{2} \langle \uparrow | + e^{-i\varphi} \sin \frac{\vartheta}{2} \langle \downarrow | \right) \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \left(\cos \frac{\vartheta}{2} | \uparrow \rangle + e^{i\varphi} \sin \frac{\vartheta}{2} | \downarrow \rangle \right) \\
&= \frac{\hbar}{2} \left(\cos^2 \frac{\vartheta}{2} - \sin^2 \frac{\vartheta}{2} \right) = \frac{\hbar}{2} \cos \vartheta.
\end{aligned} \tag{8.85}$$

Also erhalten wir

$$\langle \hat{\vec{S}} \rangle = \frac{\hbar}{2} \hat{n}. \tag{8.86}$$

Der Zustand $|\vartheta, \varphi\rangle$ beschreibt daher einen in der Richtung von \hat{n} ausgerichteten Spin 1/2. Die Verallgemeinerung auf Drehimpulse beliebiger Länge j ist möglich.

8.3 Drehoperatoren

Für einen Einheitsvektor \hat{n} ist $\hat{\vec{J}} \cdot \hat{n}$ die Drehimpulskomponente entlang \hat{n} und hermitesch. Daher ist

$$\hat{D}_\alpha(\hat{n}) := \exp \left(-\frac{i}{\hbar} \hat{\vec{J}} \cdot \hat{n} \alpha \right) \quad \text{mit } \alpha \in \mathbb{R} \tag{8.87}$$

unitär. $\hat{D}_\alpha(\hat{n})$ ist ein *Drehoperator*, der einen Zustand um den Winkel α um die Drehachse parallel zu \hat{n} dreht. Wir zeigen dies hier nicht allgemein, sondern überzeugen uns anhand des Falles $j = s = 1/2$, dass sich $\hat{D}_\alpha(\hat{n})$ wie erwartet verhält.

Für Drehungen um die z -Achse gilt

$$\hat{D}_\alpha(\hat{z}) = \exp \left(-\frac{i}{\hbar} \hat{S}_z \alpha \right) = \exp \left(-\frac{i}{2} \hat{\sigma}_z \alpha \right) = \begin{pmatrix} e^{-i\alpha/2} & 0 \\ 0 & e^{i\alpha/2} \end{pmatrix}. \tag{8.88}$$

Die letzte Gleichheit gilt, da σ_z diagonal ist. Man kann sich leicht überzeugen, dass alle Matricelemente gleich sind. Alternativ zeigt man mit Hilfe der Potenzreihenentwicklung

$$\begin{aligned}
\exp \left(-\frac{i\alpha}{2} \hat{\sigma}_z \right) &= \sum_{n=0}^{\infty} \frac{1}{n!} \left(-\frac{i\alpha}{2} \right)^n \sigma_z^n = \sum_{n=0}^{\infty} \frac{1}{n!} \left(-\frac{i\alpha}{2} \right)^n \begin{pmatrix} 1^n & 0 \\ 0 & (-1)^n \end{pmatrix} \\
&= \begin{pmatrix} \sum_{n=0}^{\infty} \frac{1}{n!} \left(-\frac{i\alpha}{2} \right)^n & 0 \\ 0 & \sum_{n=0}^{\infty} \frac{1}{n!} \left(\frac{i\alpha}{2} \right)^n \end{pmatrix}
\end{aligned}$$

$$= \begin{pmatrix} e^{-i\alpha/2} & 0 \\ 0 & e^{i\alpha/2} \end{pmatrix}. \quad (8.89)$$

Wir finden also

$$\hat{D}_\alpha(\hat{z}) = \begin{pmatrix} \cos \frac{\alpha}{2} - i \sin \frac{\alpha}{2} & 0 \\ 0 & \cos \frac{\alpha}{2} + i \sin \frac{\alpha}{2} \end{pmatrix} = \cos \frac{\alpha}{2} \mathbb{1} - i \sin \frac{\alpha}{2} \sigma_z. \quad (8.90)$$

Aber die z -Achse ist nicht ausgezeichnet. Wir erhalten daher analoge Ausdrücke für beliebige Richtungen \hat{n} :

$$\hat{D}_\alpha(\hat{n}) = \exp \left(-\frac{i}{\hbar} \hat{\vec{S}} \cdot \hat{n} \alpha \right) = \cos \frac{\alpha}{2} \mathbb{1} - i \sin \frac{\alpha}{2} \vec{\sigma} \cdot \hat{n}, \quad (8.91)$$

wobei

$$\vec{\sigma} := \begin{pmatrix} \sigma_x \\ \sigma_y \\ \sigma_z \end{pmatrix} \quad (8.92)$$

der oft verwendete Vektor der Pauli-Matrizen ist.

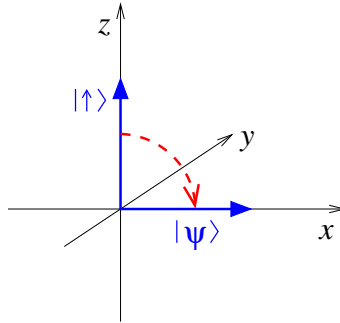
Wir betrachten nun zunächst die Drehung von $|\uparrow\rangle$ um 90° um die y -Achse:

$$\begin{aligned} |\psi\rangle &= \hat{D}_{\pi/2}(\hat{y})|\uparrow\rangle = \left(\cos \frac{\pi}{4} \mathbb{1} - i \sin \frac{\pi}{4} \sigma_y \right) |\uparrow\rangle = \left(\frac{1}{\sqrt{2}} \mathbb{1} - \frac{i}{\sqrt{2}} \sigma_y \right) |\uparrow\rangle \\ &= \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix}. \end{aligned} \quad (8.93)$$

Nun ist

$$\hat{S}_x |\psi\rangle = \frac{\hbar}{2} \sigma_x |\psi\rangle = \frac{\hbar}{2} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \frac{\hbar}{2} \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \frac{\hbar}{2} |\psi\rangle. \quad (8.94)$$

$|\psi\rangle$ ist also ein Eigenzustand von \hat{S}_x mit dem (maximalen) Eigenwert $\hbar/2$, d. h. der Spin im Zustand $|\psi\rangle$ zeigt in die x -Richtung. Das erwarten wir auch, wenn wir $|\uparrow\rangle$ (zeigt in die z -Richtung) um 90° um die y -Achse drehen.



Was passiert, wenn wir $|\uparrow\rangle$ um 180° drehen? Dann erhalten wir

$$\begin{aligned} |\psi'\rangle &= \hat{D}_\pi(\hat{y})|\uparrow\rangle = \left(\cos \frac{\pi}{2} \mathbb{1} - i \sin \frac{\pi}{2} \sigma_y \right) |\uparrow\rangle \\ &= -i \sigma_y |\uparrow\rangle = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \end{pmatrix} = |\downarrow\rangle. \end{aligned} \quad (8.95)$$

Der Spin zeigt in die negative z -Richtung, auch wie erwartet.

Eine Drehung um 360° ergibt

$$|\psi''\rangle = \hat{D}_{2\pi}(\hat{y})|\uparrow\rangle = \left(\cos \pi \mathbb{1} - i \sin \pi \sigma_y \right) |\uparrow\rangle = -|\uparrow\rangle. \quad (8.96)$$

Das ist derselbe Zustand wie $|\uparrow\rangle$, der Spin zeigt also wie erwartet wieder in die z -Richtung. Aber in Systemen mit mehr als einem Spin sind Phasenfaktoren relevant und speziell der Faktor -1 unter Rotationen um 360°

ist wesentlich. Er tritt allgemein bei halbzahligen Spins (also bei Fermionen) auf. Halbzahlige Spins muss man *zweimal* umdrehen, um wieder zum ursprünglichen Vektor zurück zu kommen. Mathematisch hängt dies damit zusammen, dass die Abbildung zwischen der Symmetriegruppe $SU(2)$ eines Quanten-Spins zur Symmetriegruppe $SO(3)$ für Rotationen im Realraum $2 : 1$ und nicht $1 : 1$ (bijektiv) ist. Physikalisch ergeben sich daraus, zusammen mit der Spin-Bahnkopplung, zusätzliche Symmetrieeigenschaften. Ein Beispiel sind vierfache Entartungen von Energiebändern in kubischen Kristallen.

Die Kommutatorrelationen für Drehoperatoren sind auch interessant. Für das Beispiel $j = s = 1/2$ finden wir

$$\begin{aligned}
[\hat{D}_\alpha(\hat{x}), \hat{D}_\beta(\hat{y})] &= \left[\exp\left(-\frac{i}{\hbar} \hat{S}_x \alpha\right), \exp\left(-\frac{i}{\hbar} \hat{S}_y \beta\right) \right] \\
&= \left[\cos \frac{\alpha}{2} \mathbb{1} - i \sin \frac{\alpha}{2} \sigma_x, \cos \frac{\beta}{2} \mathbb{1} - i \sin \frac{\beta}{2} \sigma_y \right] \\
&\quad (\mathbb{1} \text{ kommutiert mit allen Matrizen}) \\
&= -\sin \frac{\alpha}{2} \sin \frac{\beta}{2} \underbrace{[\sigma_x, \sigma_y]}_{= \sigma_x \sigma_y - \sigma_y \sigma_x = 2i\sigma_z} \\
&= -2i \sin \frac{\alpha}{2} \sin \frac{\beta}{2} \sigma_z \\
&= 2 \sin \frac{\alpha}{2} \sin \frac{\beta}{2} \hat{D}_\pi(\hat{z}).
\end{aligned} \tag{8.97}$$

Also speziell für Drehungen um 90° :

$$[\hat{D}_{\pi/2}(\hat{x}), \hat{D}_{\pi/2}(\hat{y})] = \hat{D}_\pi(\hat{z}). \tag{8.98}$$

Interessant hieran ist, dass diese Vertauschungsrelationen im Unterschied zu $[\hat{J}_x, \hat{J}_y] = i\hbar \hat{J}_z$ (und auch zu $[\hat{x}_i, \hat{p}_j] = i\hbar \delta_{ij}$) keinen Faktor \hbar enthalten. Das resultiert daraus, dass die Nichtvertauschbarkeit von Drehungen kein genuin quantenmechanisches Phänomen ist, sondern ein geometrisches. Schon klassisch hängt das Ergebnis von zwei Drehungen um unterschiedliche Achsen von ihrer Reihenfolge ab, vgl. Vorlesung Theoretische Mechanik.

Es sei noch angemerkt, dass die Quantentheorie des Drehimpulses auch umgekehrt aufgezogen werden kann: Man startet von den Drehoperatoren und ihrer geometrischen Nichtvertauschbarkeit. Dann konstruiert man die Drehimpulskomponenten als Generatoren der Drehungen im Sinne von $\hat{D}_\alpha(\hat{n}) = \exp(-\frac{i}{\hbar} \hat{\vec{J}} \cdot \hat{n} \alpha)$.

8.4 Addition von Drehimpulsen

Oft untersucht man Systeme mit mehreren gekoppelten Drehimpulsen. Dies können Spin und Bahndrehimpuls desselben Teilchens sein oder auch Drehimpulse verschiedener Teilchen. Ist das System insgesamt rotationsinvariant, so ist der Gesamtdrehimpuls $\hat{\vec{J}}$ erhalten, also

$$[\hat{\vec{J}}, \hat{H}] = 0. \tag{8.99}$$

Die einzelnen Drehimpulse $\hat{\vec{J}}_i$ von Teilsystemen sind i. A. nicht erhalten. Der Zustandsraum für mehrere Drehimpulse ist nach 7.1.3 das Tensorprodukt der Hilbert-Räume der einzelnen Drehimpulse $\hat{\vec{J}}_i$:

$$\mathcal{H} = \mathcal{H}_1 \otimes \mathcal{H}_2 \otimes \dots \tag{8.100}$$

Es ist nützlich, die Zustände des Gesamtdrehimpulses mit denen der einzelnen Drehimpulse in Beziehung zu setzen. Dabei werden wir auch herausfinden, welche Gesamtdrehimpulse sich überhaupt ergeben können. Wir beschränken uns hier auf den Fall der Kombination zweier Drehimpulse. Mehr als zwei kann man im Prinzip kombinieren, indem man konsekutiv jeweils einen hinzufügt.

8.4.1 Abzählung der Zustände

Wir betrachten zwei Drehimpulse der Längen j_1 und j_2 , d. h.

$$\hat{j}_1^2 = \hbar^2 j_1(j_1 + 1) \mathbb{1}_{2j_1+1}, \quad (8.101)$$

$$\hat{j}_2^2 = \hbar^2 j_2(j_2 + 1) \mathbb{1}_{2j_2+1}, \quad (8.102)$$

wobei $\mathbb{1}_n$ die $n \times n$ -Einheitsmatrix ist. Die Hilbert-Räume \mathcal{H}_1 und \mathcal{H}_2 haben die Dimensionen $2j_1 + 1$ bzw. $2j_2 + 1$. Der Hilbert-Raum \mathcal{H} des Gesamtsystems hat daher die Dimension $(2j_1 + 1)(2j_2 + 1)$. Als Basen von \mathcal{H}_n , $n = 1, 2$, nehmen wir die Eigenzustände zu (\hat{j}_n^2) und \hat{j}_{nz} , also $\{|j_n, m_n\rangle \mid m_n = -j_n, -j_n + 1, \dots, j_n\}$. Eine Basis von \mathcal{H} ist die Produktbasis bestehend aus den Vektoren

$$|j_1 j_2; m_1 m_2\rangle := |j_1, m_1\rangle |j_2, m_2\rangle. \quad (8.103)$$

Für den Gesamtdrehimpuls schreibt man meist

$$\hat{J} = \hat{j}_1 + \hat{j}_2. \quad (8.104)$$

Aber diese Schreibweise ist nicht ganz sauber, weil \hat{j}_1 auf \mathcal{H}_1 , \hat{j}_2 auf \mathcal{H}_2 , \hat{J} aber auf $\mathcal{H} = \mathcal{H}_1 \otimes \mathcal{H}_2$ wirkt. Korrekter müssen wir \hat{j}_1 und \hat{j}_2 auf \mathcal{H} erweitern. Dabei soll \hat{j}_1 mit Drehimpuls 2 nichts tun und umgekehrt. Daher schreiben wir korrekter

$$\hat{J} = \hat{j}_1 \otimes \mathbb{1}^{(2)} + \mathbb{1}^{(1)} \otimes \hat{j}_2, \quad (8.105)$$

wobei $\mathbb{1}^{(n)}$ jeweils der Einsoperator auf dem anderen Faktorraum ist. Im folgenden verwenden wir aber die abgekürzte Schreibweise aus Gl. (8.104). Man kann nachrechnen, dass gilt

$$[\hat{J}_i, \hat{J}_j] = i\hbar \sum_k \epsilon_{ijk} \hat{J}_k, \quad (8.106)$$

$$[\hat{J}^2, \hat{J}_i] = 0 \quad (8.107)$$

für $i, j = x, y, z$. \hat{J} ist also wirklich ein Drehimpuls. (Das ist nicht trivial, da die Vertauschungsrelationen nicht linear sind. Zum Beispiel ist das negative $-\hat{J}$ eines Drehimpulses \hat{J} kein Drehimpuls.) Daher haben \hat{J}^2 und \hat{J}_z Eigenwerte der üblichen Art. Weiter zeigt man

$$[\hat{j}_n^2, \hat{J}^2] = 0, \quad (8.108)$$

$$[\hat{j}_n^2, \hat{J}_z] = 0 \quad (8.109)$$

für $n = 1, 2$. Es gibt also ein gemeinsames vollständiges Orthonormalsystem von Eigenzuständen zu \hat{j}_1^2 , \hat{j}_2^2 , \hat{J}^2 und \hat{J}_z . Diese Eigenzustände bezeichnen wir mit $|j_1 j_2, JM\rangle$, wobei die Eigenwertgleichungen lauten

$$\hat{j}_1^2 |j_1 j_2, JM\rangle = \hbar^2 j_1(j_1 + 1) |j_1 j_2, JM\rangle, \quad (8.110)$$

$$\hat{j}_2^2 |j_1 j_2, JM\rangle = \hbar^2 j_2(j_2 + 1) |j_1 j_2, JM\rangle, \quad (8.111)$$

$$\hat{J}^2 |j_1 j_2, JM\rangle = \hbar^2 J(J + 1) |j_1 j_2, JM\rangle, \quad (8.112)$$

$$\hat{J}_z |j_1 j_2, JM\rangle = \hbar M |j_1 j_2, JM\rangle. \quad (8.113)$$

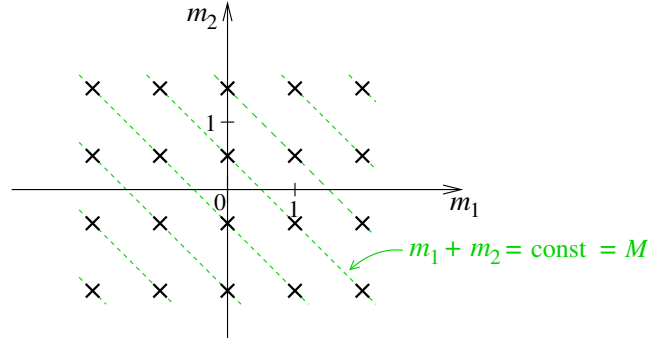
Wir wollen die Eigenzustände $|j_1 j_2, JM\rangle$ mit den Produktzuständen $|j_1 j_2; m_1 m_2\rangle$ in Beziehung setzen. Wir haben noch nicht gezeigt, dass \hat{j}_1^2 , \hat{j}_2^2 , \hat{J}^2 und \hat{J}_z ein maximaler Satz verträglicher Observabler ist. Im Prinzip könnte es also sein, dass wir in $|j_1 j_2, JM\rangle$ noch weitere Quantenzahlen brauchen, um ein vollständiges Orthonormalsystem zu erhalten. Mit anderen Worten, es könnte mehr als einen linear unabhängigen Zustand mit denselben Werten von J und M (und trivialerweise j_1, j_2) geben. Wir bezeichnen diesen Entartungsgrad mit $g(J, M)$. Aus einem

Zustand mit Quantenzahlen J, M können wir durch Anwendung von \hat{J}_\pm immer eine ganze Leiter von $2J + 1$ Zuständen mit $M = -J, \dots, J$ erzeugen. Daher hängt der Entartungsgrad nur von J ab: $g(J, M) = g(J)$.

Nun ist $|j_1 j_2; m_1 m_2\rangle$ schon Eigenzustand zu \hat{J}_z , denn

$$\begin{aligned}\hat{J}_z |j_1 j_2; m_1 m_2\rangle &= (\hat{j}_{1z} + \hat{j}_{2z}) |j_1 j_2; m_1 m_2\rangle \\ &= (\hat{j}_{1z} |j_1, m_1\rangle) |j_2, m_2\rangle + |j_1, m_1\rangle \hat{j}_{2z} |j_2, m_2\rangle \\ &= \hbar(m_1 + m_2) |j_1 j_2; m_1 m_2\rangle.\end{aligned}\tag{8.114}$$

Der zugehörige Eigenwert ist offenbar $M = m_1 + m_2$. Also ist $|j_1 j_2, JM\rangle$ eine Linearkombination nur von solchen $|j_1 j_2; m_1 m_2\rangle$, die $m_1 + m_2 = M$ erfüllen. Die Zahl $n(M)$ der Paare (m_1, m_2) , die dies erfüllen – und damit den Entartungsgrad des Eigenwerts M – erhält man mittels des folgenden Diagramms (hier für $j_1 = 2$ und $j_2 = 3/2$):



Wir finden

$$n(M) = \begin{cases} 0 & \text{für } |M| > j_1 + j_2, \\ j_1 + j_2 + 1 - |M| & \text{für } |j_1 - j_2| \leq |M| \leq j_1 + j_2, \\ 2 \min(j_1, j_2) + 1 & \text{für } 0 \leq |M| < |j_1 - j_2|. \end{cases}\tag{8.115}$$

Andererseits erhalten wir für jeden erlaubten Wert von $J \geq |M|$ gerade $g(J)$ linear unabhängige Zustände mit der Quantenzahl M . Also gilt

$$n(M) = n(|M|) = \sum_{J \geq |M|} g(J),\tag{8.116}$$

wobei wir $g(J) = 0$ setzen, wenn J gar nicht vorkommt. Es folgt

$$n(|M|) - n(|M| + 1) = \sum_{J \geq |M|} g(J) - \sum_{J \geq |M|+1} g(J) = g(|M|).\tag{8.117}$$

Das Argument $|M|$ können wir umbenennen:

$$g(J) = n(J) - n(J + 1).\tag{8.118}$$

Daraus folgt, mit Gl. (8.115),

$$g(J) = \begin{cases} 0 & \text{für } J > j_1 + j_2, \\ 1 & \text{für } |j_1 - j_2| \leq J \leq j_1 + j_2, \\ 0 & \text{für } J < |j_1 - j_2|. \end{cases}\tag{8.119}$$

Also sind nur die Werte

$$J = |j_1 - j_2|, |j_1 - j_2| + 1, \dots, j_1 + j_2\tag{8.120}$$

erlaubt und es gibt keine weitere Entartung. Zu jedem J erhalten wir natürlich

$$M = -J, -J + 1, \dots, J.\tag{8.121}$$

Beispiel: Für die Kombination zweier Spins $j_1 = j_2 = 1/2$ erhalten wir die Zustände

$$\begin{array}{l} \left| \frac{1}{2} \frac{1}{2}, 0 0 \right\rangle \\ \left| \frac{1}{2} \frac{1}{2}, 1 -1 \right\rangle \\ \left| \frac{1}{2} \frac{1}{2}, 1 0 \right\rangle \\ \left| \frac{1}{2} \frac{1}{2}, 1 1 \right\rangle \end{array} \left. \begin{array}{l} \} \text{ Singulett,} \\ \\ \\ \} \text{ Triplett.} \end{array} \right.$$

Die drei Zustände des Triplets sind entartet, wenn das System rotationsinvariant ist. Wird die Entartung aufgehoben, z. B. durch ein Magnetfeld, so spalten die Eigenenergien in drei Werte auf. Spektroskopisch erkennt man dies typischerweise an der Aufspaltung einer Spektrallinie in *drei* Linien im Magnetfeld, daher die Bezeichnung „Triplett“.

8.4.2 Clebsch-Gordan-Koeffizienten

Wir haben gesehen, dass die Zustände $|j_1 j_2, JM\rangle$ eindeutig sind und sich als Linearkombinationen von $|j_1 j_2; m_1 m_2\rangle$ mit $m_1 + m_2 = M$ schreiben lassen müssen. Jetzt bestimmen wir die Koeffizienten in dieser Linearkombination. Es gilt

$$\begin{aligned} |j_1 j_2, JM\rangle &= \sum_{j'_1, j'_2, m_1, m_2} |j'_1 j'_2; m_1 m_2\rangle \langle j'_1 j'_2; m_1 m_2 | j_1 j_2, JM\rangle \\ &= \sum_{\substack{m_1, m_2 \\ m_1 + m_2 = M}} |j_1 j_2; m_1 m_2\rangle \langle j_1 j_2; m_1 m_2 | j_1 j_2, JM\rangle \\ &= \sum_{\substack{m_1, m_2 \\ m_1 + m_2 = M}} \langle j_1 j_2; m_1 m_2 | j_1 j_2, JM\rangle |j_1 j_2; m_1 m_2\rangle. \end{aligned} \quad (8.122)$$

Die Koeffizienten nennt man *Clebsch-Gordan-Koeffizienten* und führt das Symbol

$$\langle j_1 j_2; m_1 m_2 | JM\rangle := \langle j_1 j_2; m_1 m_2 | j_1 j_2, JM\rangle \quad (8.123)$$

ein. Sie hängen per Konstruktion nur von den Größen j_1, j_2, m_1, m_2 und J ab ($M = m_1 + m_2$ ist redundant), also z. B. nicht davon, ob die beiden einzelnen Drehimpulse Spins oder Bahndrehimpulse oder selbst auch zusammengesetzte Drehimpulse darstellen. Mit dieser Abkürzung schreiben wir

$$|j_1 j_2, JM\rangle = \sum_{m_1, m_2} \langle j_1 j_2; m_1 m_2 | JM\rangle |j_1 j_2; m_1 m_2\rangle, \quad (8.124)$$

wobei wir die Summe über alle m_1, m_2 laufen lassen; für nicht vorkommende Terme ist einfach $\langle j_1 j_2; m_1 m_2 | JM\rangle = 0$. Aus dem letzten Abschnitt wissen wir, dass gilt

$$\langle j_1 j_2; m_1 m_2 | JM\rangle = 0 \quad \text{für } m_1 + m_2 \neq M, \quad (8.125)$$

$$\langle j_1 j_2; m_1 m_2 | JM\rangle = 0 \quad \text{für } J < |j_1 - j_2| \text{ oder } J > j_1 + j_2. \quad (8.126)$$

Um die Clebsch-Gordan-Koeffizienten eindeutig festzulegen, müssen wir die Phasen von $|j_1 j_2, JM\rangle$ (relativ zu denen von $|j_1 j_2; m_1 m_2\rangle$) wählen. Wir vereinbaren

- $\langle j_1 j_2; j_1 (J - j_1) | JJ\rangle \in \mathbb{R}^+$,
- für alle anderen Koeffizienten ergibt sich die Phase durch Anwendung der Auf- und Absteigeoperatoren \hat{J}_\pm des Gesamtdrehimpulses mittels Gl. (8.52).

Mann kann zeigen, dass dann alle Koeffizienten reell sind.

Da Gl. (8.124) eine unitäre Transformation darstellt, sind die $\langle j_1 j_2; m_1 m_2 | JM \rangle$ für feste j_1, j_2 Komponenten von unitären Matrizen. Da sie alle reell sind, sind diese Matrizen auch orthogonal (d. h. sie erfüllen $O^T O = O O^T = \mathbb{1}$). Aus der Unitarität bzw. Orthogonalität folgt

$$\sum_{m_1 m_2} \langle j_1 j_2; m_1 m_2 | JM \rangle \langle j_1 j_2; m_1 m_2 | J' M' \rangle = \delta_{JJ'} \delta_{MM'}, \quad (8.127)$$

$$\sum_{m_1 m_2} \langle j_1 j_2; m_1 m_2 | JM \rangle \langle j_1 j_2; m'_1 m'_2 | JM \rangle = \delta_{m_1 m'_1} \delta_{m_2 m'_2}. \quad (8.128)$$

Die Koeffizienten sind gleich eins, wenn die Linearkombination nur einen Term enthält. Das ist der Fall, wenn einer der Drehimpulse die Länge Null hat,

$$\langle j \ 0; m \ 0 | jm \rangle = \langle 0 \ j; 0 \ m | jm \rangle = 1, \quad (8.129)$$

und für zwei parallel zur gemeinsamen Quantisierungsachse ausgerichtete Drehimpulse,

$$\langle j_1 j_2; j_1 j_2 | (j_1 + j_2) \ (j_1 + j_2) \rangle = \langle j_1 j_2; (-j_1) \ (-j_2) | (j_1 + j_2) \ (-j_1 - j_2) \rangle = 1. \quad (8.130)$$

Außerdem kann man unter Verwendung von \hat{J}_\pm folgende Rekursionsformeln beweisen:

$$\begin{aligned} \sqrt{J(J+1) - M(M+1)} \langle j_1 j_2; m_1 m_2 | JM \rangle &= \sqrt{j_1(j_1+1) - m_1(m_1+1)} \langle j_1 j_2; (m_1+1) \ m_2 | J \ (M+1) \rangle \\ &\quad + \sqrt{j_2(j_2+1) - m_2(m_2+1)} \langle j_1 j_2; m_1 \ (m_2+1) | J \ (M+1) \rangle, \end{aligned} \quad (8.131)$$

$$\begin{aligned} \sqrt{J(J+1) - M(M-1)} \langle j_1 j_2; m_1 m_2 | JM \rangle &= \sqrt{j_1(j_1+1) - m_1(m_1-1)} \langle j_1 j_2; (m_1-1) \ m_2 | J \ (M-1) \rangle \\ &\quad + \sqrt{j_2(j_2+1) - m_2(m_2-1)} \langle j_1 j_2; m_1 \ (m_2-1) | J \ (M-1) \rangle. \end{aligned} \quad (8.132)$$

Mit den angegebenen Beziehungen können alle Koeffizienten bestimmt werden. In der Praxis entnimmt man sie Tabellen. Beispiel: Für zwei Spins $1/2$ erhält man

$$|\tfrac{1}{2} \ \tfrac{1}{2}, 0 \ 0\rangle = \frac{|\tfrac{1}{2} \ \tfrac{1}{2}; \tfrac{1}{2} \ -\tfrac{1}{2}\rangle - |\tfrac{1}{2} \ \tfrac{1}{2}; -\tfrac{1}{2} \ \tfrac{1}{2}\rangle}{\sqrt{2}} = \frac{|\tfrac{1}{2}, \tfrac{1}{2}\rangle |\tfrac{1}{2}, -\tfrac{1}{2}\rangle - |\tfrac{1}{2}, -\tfrac{1}{2}\rangle |\tfrac{1}{2}, \tfrac{1}{2}\rangle}{\sqrt{2}} \equiv \frac{|\uparrow\rangle |\downarrow\rangle - |\downarrow\rangle |\uparrow\rangle}{\sqrt{2}}, \quad (8.133)$$

$$|\tfrac{1}{2} \ \tfrac{1}{2}, 1 \ -1\rangle = |\downarrow\rangle |\downarrow\rangle, \quad (8.134)$$

$$|\tfrac{1}{2} \ \tfrac{1}{2}, 1 \ 0\rangle = \frac{|\uparrow\rangle |\downarrow\rangle + |\downarrow\rangle |\uparrow\rangle}{\sqrt{2}}, \quad (8.135)$$

$$|\tfrac{1}{2} \ \tfrac{1}{2}, 1 \ 1\rangle = |\uparrow\rangle |\uparrow\rangle. \quad (8.136)$$

Kapitel 9

Symmetrien

In diesem Kapitel werden wir kurz einige Konzepte besprechen, die bei der Betrachtung von Symmetrien quantenmechanischer Systeme wichtig sind. Aus den Symmetrien eines Systems kann man – wie schon in der klassischen Mechanik – viel über seine Eigenschaften lernen. In der Quantenmechanik erhält man z. B. Informationen über die Entartungen von Eigenwerten und über strikt verschwindende Matrixelemente, die zu Auswahlregeln für Übergänge gestörter Systeme führen. Wir können hier nur die Grundlagen besprechen, auf die in weiterführenden Vorlesungen, z. B. zur Gruppentheorie in der Physik, aufgebaut wird.

9.1 Unitäre Transformationen

Wir hatten in 7.2.2 gesehen, dass beobachtbare Größen invariant sind unter unitären Transformationen

$$\mathcal{U} : \begin{cases} |\alpha\rangle & \mapsto \hat{U}|\alpha\rangle, \\ \langle\alpha| & \mapsto \langle\alpha|\hat{U}^\dagger, \\ \hat{A} & \mapsto \hat{U}\hat{A}\hat{U}^\dagger. \end{cases} \quad (9.1)$$

Der Grund ist, dass \mathcal{U} beliebige Matrixelemente $\langle\alpha|\hat{A}|\beta\rangle$ abbildet auf

$$\mathcal{U}(\langle\alpha|\hat{A}|\beta\rangle) = \langle\alpha|\hat{U}^\dagger\hat{U}\hat{A}\hat{U}^\dagger\hat{U}|\beta\rangle = \langle\alpha|\hat{A}|\beta\rangle. \quad (9.2)$$

Die Transformation \mathcal{U} wird offenbar durch den unitären Operator \hat{U} charakterisiert, wobei \hat{U} bis auf einen irrelevanten Phasenfaktor eindeutig ist. (Ein Phasenfaktor ist irrelevant für Zustände und hebt sich in der Transformation der Operatoren \hat{A} heraus.) Die Eigenschaften von \hat{U} ergeben sich oft aus der physikalischen Bedeutung der jeweiligen Transformation. Als Beispiel betrachten wir die Raumspiegelung (Inversion) \mathcal{P} , die durch den *Paritätsoperator* $\hat{U} = \hat{P}$ vermittelt wird. Inversion kehrt alle Orts- und Impulskomponenten um:

$$\hat{P}\hat{r}_i\hat{P}^\dagger = -\hat{r}_i, \quad (9.3)$$

$$\hat{P}\hat{p}_i\hat{P}^\dagger = -\hat{p}_i. \quad (9.4)$$

Daraus folgt für den Bahndrehimpuls

$$\hat{P}\hat{L}_i\hat{P}^\dagger = +\hat{L}_i. \quad (9.5)$$

Wir fordern hier, dass auch der Spin gerade unter Inversion ist; im Rahmen der relativistischen Quantenmechanik kann man dies zeigen:

$$\hat{P}\hat{S}_i\hat{P}^\dagger = +\hat{S}_i. \quad (9.6)$$

Wie wirkt \hat{P} auf Zustände? Zustände im Produktraum für Spin und räumliche Bewegung lassen sich als zweikomponentige (Spinor-) Wellenfunktionen

$$\psi(\vec{r}) = \begin{pmatrix} \psi_\uparrow(\vec{r}) \\ \psi_\downarrow(\vec{r}) \end{pmatrix} = \begin{pmatrix} \langle\vec{r}|\psi_\uparrow\rangle \\ \langle\vec{r}|\psi_\downarrow\rangle \end{pmatrix} \quad (9.7)$$

darstellen. Der zugehörige „Zeilenspinor“ ist

$$(\psi_{\uparrow}^*(\vec{r}), \psi_{\downarrow}^*(\vec{r})). \quad (9.8)$$

Ein beliebiges Matricelement von \hat{r} lautet in Ortsdarstellung

$$\begin{aligned} \langle \psi | \hat{r} | \varphi \rangle &= \int d^3r \psi^\dagger(\vec{r}) \vec{r} \varphi(\vec{r}) \\ &= \int d^3r \psi^\dagger(\vec{r}) \hat{P}^\dagger \underbrace{\hat{P} \vec{r} \hat{P}^\dagger}_{= -\vec{r}} \hat{P} \varphi(\vec{r}) \\ &= - \int d^3r \psi^\dagger(\vec{r}) \hat{P}^\dagger \vec{r} \hat{P} \varphi(\vec{r}). \end{aligned} \quad (9.9)$$

Analog lauten die Matricelemente des Spins

$$\begin{aligned} \langle \psi | \hat{S} | \varphi \rangle &= \int d^3r \psi^\dagger(\vec{r}) \frac{\hbar}{2} \vec{\sigma} \varphi(\vec{r}) \\ &= \int d^3r \psi^\dagger(\vec{r}) \hat{P}^\dagger \hat{P} \frac{\hbar}{2} \vec{\sigma} \hat{P}^\dagger \hat{P} \varphi(\vec{r}) \\ &= \int d^3r \psi^\dagger(\vec{r}) \hat{P}^\dagger \frac{\hbar}{2} \vec{\sigma} \hat{P} \varphi(\vec{r}). \end{aligned} \quad (9.10)$$

Damit die Matricelemente unter der unitären Transformation mit \hat{P} invariant sind, muss gelten

$$\hat{P} \varphi(\vec{r}) = \varphi(-\vec{r}) \quad (9.11)$$

und

$$\psi^\dagger(\vec{r}) \hat{P}^\dagger = \psi^\dagger(-\vec{r}), \quad (9.12)$$

denn dann ergibt sich

$$\langle \psi | \hat{r} | \varphi \rangle = - \int d^3r \psi^\dagger(-\vec{r}) \vec{r} \varphi(-\vec{r}) \xrightarrow{\vec{r} \rightarrow -\vec{r}} - \int d^3r \psi^\dagger(\vec{r}) (-\vec{r}) \varphi(\vec{r}) = + \int d^3r \psi^\dagger(\vec{r}) \vec{r} \varphi(\vec{r}) \quad (9.13)$$

und

$$\langle \psi | \hat{S} | \varphi \rangle = \int d^3r \psi^\dagger(-\vec{r}) \frac{\hbar}{2} \vec{\sigma} \varphi(-\vec{r}) \xrightarrow{\vec{r} \rightarrow -\vec{r}} \int d^3r \psi^\dagger(\vec{r}) \frac{\hbar}{2} \vec{\sigma} \varphi(\vec{r}). \quad (9.14)$$

\hat{P} kehrt also, wie erwartet, das Ortsargument der Wellenfunktion um, ändert aber nichts am Spinor. Insbesondere wird der Spinor nicht invertiert. \hat{P} ist im Spin-Raum proportional zur 2×2 -Einheitsmatrix $\mathbb{1}$.

9.1.1 Transformationsgruppen

Die unitären Transformationen auf einem Hilbert-Raum bilden eine *Gruppe*. Diese Aussage beruht auf folgenden Erkenntnissen:

1. Das (Gruppen-) Produkt zweier Transformationen mit den unitären Operatoren \hat{U}_1 und \hat{U}_2 ist wieder eine unitäre Transformation mit dem Operator $\hat{U}_1 \hat{U}_2$. Insbesondere ist $\hat{U}_1 \hat{U}_2$ unitär. Diese Gruppenmultiplikation ist assoziativ, da die Multiplikation von Operatoren assoziativ ist.
2. Die identische Transformation mit $\hat{U} = \mathbb{1}$ ist das neutrale Element der Gruppe.
3. Das Inverse einer Transformation ist durch den inversen Operator $\hat{U}^{-1} = \hat{U}^\dagger$ charakterisiert, der wegen der Unitarität von \hat{U} immer existiert.

Die Gruppe ist i. A. nicht kommutativ, da verschiedene unitäre Operatoren nicht kommutieren müssen.

Für einen endlichdimensionalen Hilbert-Raum der Dimension N ist die vollständige Transformationsgruppe isomorph zur Gruppe $U(N)$ aller unitären $N \times N$ -Matrizen. Da ein skalarer Phasenfaktor $e^{i\alpha}$ keine Konsequenzen hat, kann man ihn festhalten, z. B. durch die Forderung $\det \hat{U} = 1 \forall \hat{U}$. Dies führt auf die *Spezielle unitäre Gruppe* $SU(N)$ aller unitären $N \times N$ -Matrizen mit Determinante 1. Oft interessiert man sich für Untergruppen der vollen Transformationsgruppe.

9.1.2 Symmetrien

Man sagt, eine Transformation lässt eine Observable \hat{A} invariant, wenn gilt

$$\hat{U}\hat{A}\hat{U}^\dagger = \hat{A}. \quad (9.15)$$

Beachte, dass diese Gleichung äquivalent ist zu

$$\hat{U}\hat{A} = \hat{A}\hat{U} \Leftrightarrow \hat{A}\hat{U} - \hat{U}\hat{A} = 0 \Leftrightarrow [\hat{A}, \hat{U}] = 0. \quad (9.16)$$

Eine Transformation lässt eine Observable also genau dann invariant, wenn der zugehörige unitäre Operator mit der Observablen kommutiert.

Man kann nun leicht folgendes zeigen:

1. Lassen \hat{U}_1 und \hat{U}_2 eine Observable invariant, so gilt dies auch für $\hat{U}_1\hat{U}_2$.
2. Die Identität $\mathbb{1}$ lässt jede Observable invariant (das ist natürlich trivial).
3. Lässt \hat{U} eine Observable invariant, so gilt dies auch für $\hat{U}^{-1} = \hat{U}^\dagger$.

Es folgt, dass die Transformationen (bzw. die unitären Operatoren), die eine Observable \hat{A} invariant lassen, eine Gruppe bilden. Diese nennen wir *Symmetriegruppe* von \hat{A} . Sie ist offenbar eine Untergruppe der vollen Transformationsgruppe. Beispiel: Eine Komponente des Drehimpulses, o. B. d. A. \hat{J}_z , kommutiert mit Drehoperatoren um die z -Achse, aber nicht mit Drehoperatoren um andere Achsen und auch nicht mit anderen unitären Operatoren. Die Symmetriegruppe von \hat{J}_z ist daher $\{\hat{D}_\alpha(\hat{z}) \mid \alpha \in [0, 2\pi[\}$. Der Bereich von α ist eingeschränkt, da $\alpha = 2\pi$ dasselbe ergibt wie $\alpha = 0$. Die Gruppe ist isomorph zu

$$U(1) = \{e^{i\alpha} \mid \alpha \in [0, 2\pi[\}. \quad (9.17)$$

(Für halbzahlige Spins ergibt sich eine Subtilität aus $\hat{D}_{2\pi}(\hat{z}) = -\mathbb{1}$, auf die wir hier nicht eingehen.)

Besonders wichtig ist die Symmetriegruppe des Hamiltonians \hat{H} , meist Symmetriegruppe des Systems oder einfach Symmetriegruppe genannt. Sie umfasst alle Transformationen bzw. unitären Operatoren \hat{U} mit

$$\hat{U}\hat{H}\hat{U}^\dagger = \hat{H} \quad (9.18)$$

oder äquivalent mit

$$[\hat{H}, \hat{U}] = 0. \quad (9.19)$$

Nun kann man jeden unitären Operator schreiben als

$$\hat{U} = e^{i\hat{B}}, \quad (9.20)$$

wobei \hat{B} hermitesch ist. Für Gl. (9.19) ist hinreichend, dass

$$[\hat{H}, \hat{B}] = 0 \quad (9.21)$$

gilt; ist \hat{B} also eine Erhaltungsgröße, so folgt eine unitäre Symmetrie. (Manchmal bezeichnet man auch unitäre Operatoren, die mit \hat{H} kommutieren, als Erhaltungsgrößen, obwohl sie keine Observablen darstellen.) Die Umkehrung gilt nicht allgemein. Zum Beispiel ist *jedes* System mit Drehimpuls ganzzahliger Länge j trivialerweise invariant unter

$$\hat{D}_{2\pi}(\hat{n}) = \exp\left(-\frac{2\pi i}{\hbar} \hat{J} \cdot \hat{n}\right) = \mathbb{1}, \quad (9.22)$$

aber daraus folgt nicht, dass jede Drehimpulskomponente $\hat{J} \cdot \hat{n}$ eine Erhaltungsgröße ist. Dies folgt aber, wenn das System invariant unter *allen* Drehungen um die Achse \hat{n} ist. Denn dann ist es insbesondere invariant für beliebig kleine Drehwinkel und wir können entwickeln:

$$\hat{D}_\alpha(\hat{n}) = \exp\left(-\frac{i}{\hbar} \hat{J} \cdot \hat{n} \alpha\right) \cong \mathbb{1} - \frac{i}{\hbar} \hat{J} \cdot \hat{n} \alpha \quad (9.23)$$

und $[\hat{H}, \hat{D}_\alpha(\hat{n})] = 0$ impliziert dann $[\hat{H}, \hat{\vec{J}} \cdot \hat{n}] = 0$.

Die Wichtigkeit der Symmetriegruppe für Entartungen von Eigenenergien kann man leicht einsehen: Sei \hat{U} der unitäre Operator eines Elements der Symmetriegruppe und sei $|\psi\rangle$ ein Eigenzustand von \hat{H} zur Energie E . Dann gilt

$$\hat{H}|\psi\rangle = E|\psi\rangle \quad (9.24)$$

$$\Rightarrow \hat{U}\hat{H}|\psi\rangle = \hat{U}E|\psi\rangle = E\hat{U}|\psi\rangle \quad (9.25)$$

$$\Rightarrow \hat{U}\hat{H}\hat{U}^\dagger\hat{U}|\psi\rangle = E\hat{U}|\psi\rangle. \quad (9.26)$$

Wegen der Symmetrie ist $\hat{U}\hat{H}\hat{U}^\dagger = \hat{H}$, also folgt

$$\hat{H}\hat{U}|\psi\rangle = E\hat{U}|\psi\rangle. \quad (9.27)$$

Damit ist $\hat{U}|\psi\rangle$ ebenfalls ein Eigenzustand von \hat{H} zur Energie E . Unterscheidet sich $\hat{U}|\psi\rangle$ von $|\psi\rangle$ nur um einen Phasenfaktor, so beschreiben $\hat{U}|\psi\rangle$ und $|\psi\rangle$ denselben Zustand. Andernfalls sind $\hat{U}|\psi\rangle$ und $|\psi\rangle$ linear unabhängig und beschreiben unterschiedliche Zustände. Der Eigenwert E ist dann (mindestens zweifach) entartet. Das gilt natürlich für alle \hat{U} aus der Symmetriegruppe. Ist der Eigenwert E n -fach entartet, so muss der entsprechende n -dimensionale Unterraum des Hilbert-Raums invariant unter allen \hat{U} aus der Symmetriegruppe sein. Die Existenz und Dimension solcher invarianter Unterräume untersucht man mittels der *Darstellungen* von Gruppen.

Beispiel: Der Drehimpuls $\hat{\vec{J}}$ eines gewissen Atoms in einem Kristall sei durch den Hamilton-Operator

$$\hat{H} = -\kappa \hat{J}_z^2 \quad (9.28)$$

beschrieben. Dies nennt man eine magnetokristalline Anisotropie, konkret für $\kappa > 0$ eine „leichte Achse“, weil die Energie durch Ausrichtung des Drehimpulses entlang $\pm \hat{z}$ minimiert wird. Die Symmetriegruppe enthält beliebige Drehungen um die z -Achse, $\hat{D}_\alpha(\hat{z}) = \exp(-\frac{i}{\hbar} \hat{J}_z \alpha)$, und äquivalent damit kommutiert \hat{H} mit \hat{J}_z (und auch mit \hat{J}^2). Die Zustände $|j, m\rangle$ aus der Standardbasis sind also Eigenzustände von \hat{H} . Zusätzlich enthält die Symmetriegruppe aber auch Drehungen um 180° bzgl. beliebiger Achsen in der xy -Ebene,

$$\hat{D}_\pi(\hat{n}) = \exp\left(-\frac{i\pi}{\hbar} \hat{\vec{J}} \cdot \hat{n}\right), \quad \hat{n} \perp \hat{z}. \quad (9.29)$$

Wir zeigen dies für $\hat{n} = \hat{x}$:

$$\begin{aligned} \hat{D}_\pi(\hat{x}) \hat{H} \hat{D}_\pi^\dagger(\hat{x}) &= -\kappa \hat{D}_\pi(\hat{x}) \hat{J}_z^2 \hat{D}_\pi^\dagger(\hat{x}) \\ &= -\kappa \sum_{mm'} \hat{D}_\pi(\hat{x}) |j, m\rangle \langle j, m| \hat{J}_z^2 |j, m'\rangle \langle j, m'| \hat{D}_\pi^\dagger(\hat{x}) \\ &= -\kappa \sum_{mm'} |j, -m\rangle \langle j, m| \hbar^2 m^2 \delta_{mm'} |j, m'\rangle \langle j, -m'| \quad \left| \text{da } \hat{D}_\pi(\hat{x}) |j, m\rangle = |j, -m\rangle \right. \\ &= -\kappa \sum_m \hbar^2 m^2 |j, -m\rangle \langle j, -m| = -\kappa \sum_m \hbar^2 m^2 |j, m\rangle \langle j, m| \\ &= \hat{H}. \end{aligned} \quad (9.30)$$

Da $|j, m\rangle$ Eigenzustand von \hat{H} zum Eigenwert $E_m = -\kappa \hbar^2 m^2$ ist, so folgt, dass $\hat{D}_\pi(\hat{x}) |j, m\rangle = |j, -m\rangle$ ebenfalls Eigenzustand zu derselben Energie ist. Die beiden Zustände sind linear unabhängig genau dann, wenn $m \neq 0$ ist. Daher sind die Energien E_m für $m \neq 0$ zweifach entartet und E_0 (existiert nur für ganzzahliges j) ist nicht entartet. Bei diesem einfachen Beispiel kann man dies natürlich direkt sehen.

9.2 Zeitumkehr

In der klassischen Physik diskutiert man die Zeitumkehr: Kann ein physikalischer Prozess auch rückwärts in der Zeit ablaufen? Oder exakter ausgedrückt: Wenn $t \mapsto (\vec{q}(t), \vec{p}(t))$ ein mit der Hamilton-Funktion konsistente

Zeitentwicklung der kanonischen Variablen (\vec{q}, \vec{p}) ist, gilt dies auch für $t \mapsto (\vec{q}(-t), \vec{p}(-t))$? In diesem Abschnitt besprechen wir kurz die Zeitumkehr in der Quantenmechanik.

Wir schreiben die Zeitumkehrtransformation analog zu Abschnitt 9.1 als

$$\mathcal{T} : \begin{cases} |\alpha\rangle & \mapsto \hat{T} |\alpha\rangle, \\ \langle\alpha| & \mapsto \langle\alpha| \hat{T}^\dagger, \\ \hat{A} & \mapsto \hat{T} \hat{A} \hat{T}^\dagger. \end{cases} \quad (9.31)$$

Aus dem Korrespondenzprinzip erhalten wir

$$\hat{T} \hat{r}_i \hat{T}^\dagger = \hat{r}_i, \quad (9.32)$$

$$\hat{T} \hat{p}_i \hat{T}^\dagger = -\hat{p}_i. \quad (9.33)$$

Daraus folgt

$$\hat{T} \hat{L}_i \hat{T}^\dagger = -\hat{L}_i \quad (9.34)$$

und wir fordern (oder entnehmen der relativistischen Quantenmechanik), dass analog für den Spin gilt

$$\hat{T} \hat{S}_i \hat{T}^\dagger = -\hat{S}_i. \quad (9.35)$$

Aber für den Kommutator konjugierter Orts- und Impulskomponenten finden wir dann

$$[\hat{r}_i, \hat{p}_i] = i\hbar \quad (9.36)$$

$$\Rightarrow \hat{T} [\hat{r}_i, \hat{p}_i] \hat{T}^\dagger = \hat{T} i\hbar \hat{T}^\dagger \quad (9.37)$$

$$\begin{aligned} \Rightarrow \hat{T} (\hat{r}_i \hat{p}_i - \hat{p}_i \hat{r}_i) \hat{T}^\dagger &= \hat{T} \hat{r}_i \hat{T}^\dagger \hat{T} \hat{p}_i \hat{T}^\dagger - \hat{T} \hat{p}_i \hat{T}^\dagger \hat{T} \hat{r}_i \hat{T}^\dagger \\ &= -\hat{r}_i \hat{p}_i + \hat{p}_i \hat{r}_i \\ &= -\underbrace{[\hat{r}_i, \hat{p}_i]}_{= i\hbar} \stackrel{!}{=} \hat{T} i\hbar \hat{T}^\dagger \end{aligned} \quad (9.38)$$

$$\Rightarrow -i\hbar = \hat{T} i\hbar \hat{T}^\dagger. \quad (9.39)$$

Hier tritt ein unerwarteter Vorzeichenwechsel auf. Da ein negativer Wert von \hbar keinen Sinn hat, folgt

$$\hat{T} i\hat{T}^\dagger = -i. \quad (9.40)$$

Aber unitäre Operatoren erfüllen dies nicht – für sie gilt $\hat{U} i \hat{U}^\dagger = i$. \hat{T} ist damit nicht unitär, und, da der Operator die Komplexkonjugation enthält, noch nicht einmal linear. Ein Operator \hat{X} mit

$$\hat{X} (\lambda_1 |\psi_1\rangle + \lambda_2 |\psi_2\rangle) = \lambda_1^* \hat{X} |\psi_1\rangle + \lambda_2^* \hat{X} |\psi_2\rangle \quad (9.41)$$

heißt *antilinear*. Den zu \hat{X} adjungierten Operator definiert man wie üblich durch

$$\langle\psi| \hat{X}^\dagger |\varphi\rangle := \langle\hat{X} \psi|\varphi\rangle = \langle\varphi|\hat{X} \psi\rangle^* \equiv (\langle\varphi|\hat{X}|\psi\rangle)^*. \quad (9.42)$$

Erfüllt ein antilinearer Operator

$$\hat{X} \hat{X}^\dagger = \hat{X}^\dagger \hat{X} = \mathbb{1}, \quad (9.43)$$

so heißt er *antiunitär*. Die obigen Eigenschaften von \hat{T} legen nahe, dass \hat{T} antiunitär ist. Das ist tatsächlich der Fall.

Man kann auch an der Schrödinger-Gleichung in Ortsdarstellung sehen, dass Zeitumkehr etwas mit Komplexkonjugation zu tun hat: Aus

$$i\hbar \frac{\partial}{\partial t} \psi(\vec{r}, t) = \hat{H} \psi(\vec{r}, t) \quad (9.44)$$

folgt durch Umbenennung $t \rightarrow -t$, dass gilt

$$i\hbar \frac{\partial}{\partial(-t)} \psi(\vec{r}, -t) = \hat{H} \psi(\vec{r}, -t), \quad (9.45)$$

wobei \hat{H} zeitunabhängig angenommen wurde. Es folgt

$$-i\hbar \frac{\partial}{\partial t} \psi(\vec{r}, -t) = \hat{H} \psi(\vec{r}, -t) \quad \left| \text{komplex konjugieren} \right. \quad (9.46)$$

$$\Rightarrow i\hbar \frac{\partial}{\partial t} \psi^*(\vec{r}, -t) = \hat{H}^* \psi^*(\vec{r}, -t). \quad (9.47)$$

Also ist $\psi^*(\vec{r}, -t)$ eine Lösung der Schrödinger-Gleichung zu \hat{H}^* . Ist \hat{H} reell, so ist $\psi^*(\vec{r}, -t)$ Lösung der Schrödinger-Gleichung zu \hat{H} , genau wie $\psi(\vec{r}, t)$. In diesem Fall nennt man das System *zeitumkehrinvariant*. Nun ist in Ortsdarstellung $\hat{r}_i^* = \hat{r}_i$ und

$$\hat{p}_i^* = \left(\frac{\hbar}{i} \frac{\partial}{\partial r_i} \right)^* = -\frac{\hbar}{i} \frac{\partial}{\partial r_i} = -\hat{p}_i. \quad (9.48)$$

Komplexbildung in Ortsdarstellung tut also genau das, was wir von der Zeitumkehrtransformation erwarten.

Beispiel: Der Hamiltonian eines Teilchens mit der Ladung q im Induktionsfeld $\vec{B} = \vec{\nabla} \times \vec{A}$ lautet (minimale Kopplung)

$$\hat{H} = \frac{1}{2m} \left(\hat{\vec{p}} - \frac{q}{c} \vec{A}(\vec{r}) \right)^2, \quad (9.49)$$

hier in Gaußschen Einheiten. Ein homogenes \vec{B} -Feld in z -Richtung wird z. B. durch

$$\vec{A} = \hat{y} B x \quad (9.50)$$

realisiert:

$$\vec{B} = \vec{\nabla} \times \vec{A} = \begin{pmatrix} \frac{\partial}{\partial y} A_z - \frac{\partial}{\partial z} A_y \\ \frac{\partial}{\partial z} A_x - \frac{\partial}{\partial x} A_z \\ \frac{\partial}{\partial x} A_y - \frac{\partial}{\partial y} A_x \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ B \end{pmatrix} = \hat{z} B. \quad (9.51)$$

Hier sind \hat{y} und \hat{z} Koordinateneinheitsvektoren, keine Operatoren. Dann ist

$$\hat{H} = \frac{1}{2m} \left(\hat{\vec{p}} - \frac{qB}{c} x \hat{y} \right)^2. \quad (9.52)$$

Zeitumkehr liefert

$$\hat{T} \hat{H} \hat{T}^\dagger = \frac{1}{2m} \left(-\hat{\vec{p}} - \frac{qB}{c} x \hat{y} \right)^2 = \frac{1}{2m} \left(\hat{\vec{p}} + \frac{qB}{c} x \hat{y} \right)^2. \quad (9.53)$$

Das System ist also nicht zeitumkehrinvariant, es sei denn, es ist $B = 0$ (oder $q = 0$). Ganz allgemein bricht ein äußeres Magnetfeld die Zeitumkehrsymmetrie.

Man kann folgendes zeigen, die Beweise sind meist elementar:

1. Sei $\mathcal{B} = \{|\psi_1\rangle, |\psi_2\rangle, \dots\}$ eine Orthonormalbasis des Hilbert-Raums. Die Komplexkonjugation \hat{K} bezüglich der Basis \mathcal{B} ist definiert durch

$$\hat{K} \sum_n c_n |\psi_n\rangle = \sum_n c_n^* \hat{K} |\psi_n\rangle = \sum_n c_n^* |\psi_n\rangle. \quad (9.54)$$

\hat{K} konjugiert also Koeffizienten und lässt die Basisvektoren unverändert. \hat{K} hängt von der Wahl der Basis ab, da die Transformation von einer Basis zu einer anderen komplexe Koeffizienten enthalten kann. \hat{K} ist offensichtlich antiunitär und erfüllt $\hat{K}^2 = \mathbb{1}$, also $\hat{K}^{-1} = \hat{K}^\dagger = \hat{K}$. Im letzten Beispiel wurde die Komplexkonjugation bezüglich der Ortsbasis betrachtet.

2. Jeder antiunitäre Operator \hat{T} lässt sich schreiben als

$$\hat{T} = \hat{K} \hat{U} = \hat{V} \hat{K} \quad (9.55)$$

mit unitären Operatoren \hat{U}, \hat{V} . (Da \hat{K} von der Basis abhängt, gilt dies auch für \hat{U}, \hat{V} .)

3. Ganz allgemein gilt: Ist \hat{U} unitär und \hat{T} antiunitär, so sind $\hat{U} \hat{T}$ und $\hat{T} \hat{U}$ antiunitär.

4. Für zwei antiunitäre Operatoren \hat{S}, \hat{T} gilt: $\hat{S}\hat{T}$ und $\hat{T}\hat{S}$ sind unitär. (Im Allgemeinen ist $\hat{S}\hat{T} \neq \hat{T}\hat{S}$.) Aus

$$\hat{U} := \hat{S}\hat{T} \quad (\hat{U} \text{ unitär}) \quad (9.56)$$

folgt

$$\hat{S} = \hat{U}\hat{T}^\dagger. \quad (9.57)$$

Zwei antiunitäre Operatoren unterscheiden sich also immer um einen unitären Faktor. Daher ist es ausreichend, *eine einzige* antiunitäre Transformation zu betrachten. Oft wählt man die Zeitumkehr. Alle anderen möglichen antiunitären Transformationen sind Produkte der ausgewählten antiunitären Transformation und unitärer Transformationen.

5. Das Quadrat des Zeitumkehroperators ist

$$\hat{T}^2 = \mathbb{1} \quad \text{oder} \quad \hat{T}^2 = -\mathbb{1}. \quad (9.58)$$

Beweis: Zweimalige Zeitumkehr führt auf den Ausgangszustand zurück, daher ist \hat{T}^2 ein reiner Phasenfaktor:

$$\hat{T}^2 = e^{i\alpha} \mathbb{1}. \quad (9.59)$$

Nun existiert ein unitärer Operator \hat{U}_T , so dass gilt $\hat{T} = \hat{U}_T \hat{K}$. Es folgt

$$\hat{U}_T \hat{K} \hat{U}_T \hat{K} = \hat{U}_T \hat{U}_T^* = e^{i\alpha} \mathbb{1}. \quad (9.60)$$

Daraus folgt sofort

$$\hat{U}_T \hat{U}_T^* \hat{U}_T \hat{U}_T^* = e^{i\alpha} e^{i\alpha} \mathbb{1} = e^{2i\alpha} \mathbb{1}. \quad (9.61)$$

Andererseits folgt mit komplexer Konjugation

$$\hat{U}_T^* \hat{U}_T = e^{-i\alpha} \mathbb{1} \quad (9.62)$$

und daher

$$\hat{U}_T \hat{U}_T^* \hat{U}_T \hat{U}_T^* = \hat{U}_T e^{i\alpha} \hat{U}_T^* = e^{i\alpha} \hat{U}_T \hat{U}_T^* = e^{-i\alpha} e^{i\alpha} \mathbb{1} = \mathbb{1}. \quad (9.63)$$

Wir erhalten also

$$e^{2i\alpha} = 1 \quad (9.64)$$

$$\Rightarrow e^{i\alpha} = \pm 1. \quad (9.65)$$

Also folgt $\hat{T}^2 = \pm \mathbb{1}$, was zu zeigen war.

9.2.1 Zeitumkehr, Drehimpulse und das Kramers-Theorem

Wie wirkt Zeitumkehr auf Drehimpulsoperatoren in der Standardbasis $\{|j, m\rangle\}$? Wir wissen schon, dass gilt

$$\hat{T} \hat{J}_i \hat{T}^\dagger = -\hat{J}_i, \quad i = x, y, z. \quad (9.66)$$

Mit Hilfe der Basis $\{|j, m\rangle\}$ können wir \hat{J}_i schreiben als

$$\hat{J}_i = \sum_{j,m} \sum_{j',m'} |j, m\rangle \langle j, m | \hat{J}_i | j', m' \rangle \langle j', m' | = \sum_j \sum_{m,m'} |j, m\rangle \langle j, m | \hat{J}_i | j, m' \rangle \langle j, m' |, \quad (9.67)$$

wobei $\langle j, m | \hat{J}_i | j, m' \rangle$ die Matrixelemente von \hat{J}_i aus 8.2.1 sind. Der Zeitumkehroperator lässt sich schreiben als

$$\hat{T} = \hat{U}_T \hat{K}, \quad (9.68)$$

wobei \hat{K} die Komplexkonjugation bezüglich der Basis $\{|j, m\rangle\}$ ist. Es folgt

$$\hat{J}_i = -\hat{T} \hat{J}_i \hat{T}^\dagger = -\hat{U}_T \hat{K} \hat{J}_i \hat{K}^{-1} \hat{U}_T^\dagger = -\hat{U}_T \hat{K} \hat{J}_i \hat{K} \hat{U}_T^\dagger$$

$$\begin{aligned}
&= - \sum_j \sum_{mm'} \hat{U}_T \hat{K} |j, m\rangle \underbrace{\langle j, m | \hat{J}_i | j, m' \rangle}_{\text{Zahl!}} \langle j, m' | \hat{K} \hat{U}_T^\dagger \\
&= - \sum_j \sum_{mm'} \hat{U}_T |j, m\rangle (\langle j, m | \hat{J}_i | j, m' \rangle)^* \langle j, m' | \hat{U}_T^\dagger.
\end{aligned} \tag{9.69}$$

Nun haben wir in 8.2.1 gesehen, dass die Matrixelemente von \hat{J}_x und \hat{J}_z reell und die von \hat{J}_y imaginär sind. Es folgt

$$\begin{aligned}
\hat{J}_i &= \sum_j \sum_{mm'} \hat{U}_T |j, m\rangle \langle j, m | \hat{J}_i | j, m' \rangle \langle j, m' | \hat{U}_T^\dagger \times \begin{cases} -1 & \text{für } i = x, z, \\ 1 & \text{für } i = y. \end{cases} \\
&= \hat{U}_T \hat{J}_i \hat{U}_T^\dagger \times \begin{cases} -1 & \text{für } i = x, z, \\ 1 & \text{für } i = y. \end{cases}
\end{aligned} \tag{9.70}$$

Wir suchen also einen unitären Operator \hat{U}_T , der die x - und z -Komponenten des Drehimpulses umkehrt und die y -Komponente unverändert lässt. Dies erfüllt gerade die Drehung um die y -Achse um 180° . Also ist

$$\hat{U}_T = \hat{D}_\pi(\hat{y}) = \exp\left(-\frac{i\pi}{\hbar} \hat{J}_y\right). \tag{9.71}$$

Der Zeitumkehroperator lässt sich auf dem Hilbert-Raum eines Drehimpulses also schreiben als

$$\hat{T} = \exp\left(-\frac{i\pi}{\hbar} \hat{J}_y\right) \hat{K}. \tag{9.72}$$

Speziell für einen Spin $1/2$ ist

$$\begin{aligned}
\hat{T} &= \exp\left(-\frac{i\pi}{\hbar} \hat{S}_y\right) \hat{K} = \exp\left(-\frac{i\pi}{2} \sigma_y\right) \hat{K} = \left(\cos \frac{\pi}{2} \mathbb{1} - i \sin \frac{\pi}{2} \sigma_y\right) \hat{K} \\
&= (-i\sigma_y) \hat{K} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \hat{K}.
\end{aligned} \tag{9.73}$$

Es folgt für den Spin $1/2$ und analog für alle halbzahligen Drehimpulse

$$\hat{T}^2 = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \hat{K} \underbrace{\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}}_{\text{reell}} \hat{K} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix} = -\mathbb{1}. \tag{9.74}$$

Für ganzzahlige Drehimpulse findet man dagegen $\hat{T}^2 = +\mathbb{1}$.

Die Identität $\hat{T}^2 = -\mathbb{1}$ für halbzahlige Drehimpulse hat eine wichtige Konsequenz: Sei der Hamiltonian \hat{H} eines Systems mit halbzahligem Drehimpuls zeitumkehrsymmetrisch und sei $|\psi\rangle$ ein Eigenzustand von \hat{H} . Dann gilt:

$$\hat{H}|\psi\rangle = E|\psi\rangle \tag{9.75}$$

$$\Rightarrow \hat{T}\hat{H}|\psi\rangle = \hat{T}E|\psi\rangle \tag{9.76}$$

$$\Rightarrow \hat{T}\hat{H}\hat{T}^\dagger \hat{T}|\psi\rangle = E \hat{T}|\psi\rangle \tag{9.77}$$

$$\Rightarrow \hat{H} \hat{T}|\psi\rangle = E \hat{T}|\psi\rangle. \tag{9.78}$$

Also ist $\hat{T}|\psi\rangle$ ebenfalls Eigenzustand von \hat{H} zu derselben Eigenenergie. Nun ist aber $\hat{T}|\psi\rangle$ linear unabhängig von $|\psi\rangle$. Beweis: Angenommen, $\hat{T}|\psi\rangle$ und $|\psi\rangle$ sind linear abhängig. Dann gibt es ein $c \in \mathbb{C}$, $c \neq 0$ mit $\hat{T}|\psi\rangle = c|\psi\rangle \Rightarrow \hat{T}^2|\psi\rangle = \hat{T}c|\psi\rangle = c^* \hat{T}|\psi\rangle = c^* c |\psi\rangle = |c|^2 |\psi\rangle$. Andererseits ist $\hat{T}^2|\psi\rangle = -|\psi\rangle$. Widerspruch!

Es folgt, dass für Systeme mit halbzahligem Gesamtdrehimpuls und Zeitumkehrsymmetrie jeder Eigenwert von \hat{H} mindestens zweifach und allgemein geradzahlig entartet ist. Diese durch zu -1 quadrierende Zeitumkehrsymmetrie erzwungene Entartung nennt man *Kramers-Entartung*. Die Aussage ist Inhalt des *Kramers-Theorems*. Es ist z. B. wichtig für die Festkörperphysik: Wenn man die Elektronen in guter Näherung als nicht wechselwirkende Quasiteilchen beschreiben kann, ist ein Ein-Teilchen-Bild anwendbar. Das Theorem sagt dann für den Fall mit Zeitumkehrsymmetrie eine zweifache Entartung aller Eigenenergien voraus, unabhängig von anderen Symmetrien.

Kapitel 10

Zentralpotentiale

In diesem Kapitel beschäftigen wir uns mit der Quantenmechanik eines Teilchens in einem dreidimensionalen Zentralpotential $V(\vec{r}) = V(r)$. Als wichtigste Anwendung besprechen wir dann das Wasserstoff-Atom.

10.1 Allgemeines Zentralpotential

Wir betrachten ein Teilchen der Masse M im Zentralpotential $V(r)$. $V(r)$ ist zunächst eine beliebige (hinreichend gutartige) Funktion. Der Hamiltonoperator in Ortsdarstellung lautet

$$\hat{H} = \frac{\hat{p}^2}{2M} + V(\vec{r}) = -\frac{\hbar^2}{2M} \nabla^2 + V(r). \quad (10.1)$$

Aufgrund der Rotationssymmetrie des Potentials V ist es günstig, zu Kugelkoordinaten überzugehen. Dann ist (siehe Skript zur Theoretischen Mechanik)

$$\nabla^2 = \frac{1}{r^2} \frac{\partial}{\partial r} \left(r^2 \frac{\partial}{\partial r} \right) + \frac{1}{r^2 \sin \theta} \frac{\partial}{\partial \theta} \left(\sin \theta \frac{\partial}{\partial \theta} \right) + \frac{1}{r^2 \sin^2 \theta} \frac{\partial^2}{\partial \varphi^2}. \quad (10.2)$$

Der erste Term führt auf die kinetische Energie der Radialbewegung, die beiden anderen auf die Energie der Tangentialbewegung. Es liegt daher nahe, zu vermuten, dass letztere mit dem Bahndrehimpuls zusammenhängen. Diesen Zusammenhang leiten wir nun her. Der Nabla-Operator lautet in Kugelkoordinaten

$$\vec{\nabla} = \hat{r} \frac{\partial}{\partial r} + \frac{\hat{\theta}}{r} \frac{\partial}{\partial \theta} + \frac{\hat{\varphi}}{r \sin \theta} \frac{\partial}{\partial \varphi}, \quad (10.3)$$

wobei \hat{r} , $\hat{\theta}$ und $\hat{\varphi}$ Koordinateneinheitsvektoren und keine Operatoren sind. Der Ortsvektor erfüllt trivialerweise

$$\vec{r} = r \hat{r}. \quad (10.4)$$

Daraus folgt, in Ortsdarstellung,

$$\begin{aligned} \hat{L} &= \vec{r} \times \frac{\hbar}{i} \vec{\nabla} \\ &= \frac{\hbar}{i} r \hat{r} \times \left(\hat{r} \frac{\partial}{\partial r} + \frac{\hat{\theta}}{r} \frac{\partial}{\partial \theta} + \frac{\hat{\varphi}}{r \sin \theta} \frac{\partial}{\partial \varphi} \right) \\ &= \frac{\hbar}{i} \cancel{r} \left(0 + \frac{\hat{r} \times \hat{\theta}}{\cancel{r}} \frac{\partial}{\partial \theta} + \frac{\hat{r} \times \hat{\varphi}}{\cancel{r} \sin \theta} \frac{\partial}{\partial \varphi} \right) \\ &= \frac{\hbar}{i} \left(\hat{\varphi} \frac{\partial}{\partial \theta} - \frac{\hat{\theta}}{\sin \theta} \frac{\partial}{\partial \varphi} \right). \end{aligned} \quad (10.5)$$

Damit ist

$$\hat{L}^2 = \hat{\vec{L}} \cdot \hat{\vec{L}} = -\hbar^2 \left(\hat{\varphi} \frac{\partial}{\partial \theta} \cdot \hat{\varphi} \frac{\partial}{\partial \theta} - \hat{\varphi} \frac{\partial}{\partial \theta} \cdot \frac{\hat{\theta}}{\sin \theta} \frac{\partial}{\partial \varphi} - \frac{\hat{\theta}}{\sin \theta} \frac{\partial}{\partial \varphi} \cdot \hat{\varphi} \frac{\partial}{\partial \theta} + \frac{\hat{\theta}}{\sin \theta} \frac{\partial}{\partial \varphi} \cdot \frac{\hat{\theta}}{\sin \theta} \frac{\partial}{\partial \varphi} \right). \quad (10.6)$$

Hier tritt als Komplikation auf, dass die Koordinateneinheitsvektoren $\hat{r}, \hat{\theta}, \hat{\varphi}$ von den Koordinaten r, θ, φ abhängen (anders als für kartesische Koordinaten). Es ist nämlich, ausgedrückt in kartesischen Komponenten,

$$\hat{r} = \begin{pmatrix} \sin \theta \cos \varphi \\ \sin \theta \sin \varphi \\ \cos \theta \end{pmatrix}, \quad (10.7)$$

$$\hat{\theta} = \begin{pmatrix} \cos \theta \cos \varphi \\ \cos \theta \sin \varphi \\ -\sin \theta \end{pmatrix}, \quad (10.8)$$

$$\hat{\varphi} = \begin{pmatrix} -\sin \varphi \\ \cos \varphi \\ 0 \end{pmatrix}. \quad (10.9)$$

Es folgt

$$\frac{\partial \hat{\theta}}{\partial \theta} = \begin{pmatrix} -\sin \theta \cos \varphi \\ -\sin \theta \sin \varphi \\ -\cos \theta \end{pmatrix} = -\hat{r}, \quad (10.10)$$

$$\frac{\partial \hat{\theta}}{\partial \varphi} = \begin{pmatrix} -\cos \theta \sin \varphi \\ \cos \theta \cos \varphi \\ 0 \end{pmatrix} = \cos \theta \hat{\varphi}, \quad (10.11)$$

$$\frac{\partial \hat{\varphi}}{\partial \theta} = 0, \quad (10.12)$$

$$\frac{\partial \hat{\varphi}}{\partial \varphi} = \begin{pmatrix} -\cos \varphi \\ -\sin \varphi \\ 0 \end{pmatrix} = \hat{z} \times \hat{\varphi}. \quad (10.13)$$

Daraus folgt für das Drehimpulsquadrat

$$\begin{aligned} \hat{L}^2 &= -\hbar^2 \left(\frac{\partial^2}{\partial \theta^2} + \underbrace{\hat{\varphi} \cdot \frac{\hat{r}}{\sin \theta}}_{=0} \frac{\partial}{\partial \varphi} - \underbrace{\hat{\varphi} \cdot \hat{\theta}}_{=0} \frac{\partial}{\partial \theta} \frac{1}{\sin \theta} \frac{\partial}{\partial \varphi} \right. \\ &\quad \left. - \frac{\hat{\theta}}{\sin \theta} \cdot (\hat{z} \times \hat{\varphi}) \frac{\partial}{\partial \theta} + \underbrace{\frac{\hat{\theta}}{\sin \theta} \cdot \frac{\cos \theta \hat{\varphi}}{\sin \theta}}_{=0} \frac{\partial}{\partial \varphi} + \frac{1}{\sin^2 \theta} \frac{\partial^2}{\partial \varphi^2} \right) \\ &= -\hbar^2 \left(\frac{\partial^2}{\partial \theta^2} + \frac{\cos \theta}{\sin \theta} \frac{\partial}{\partial \theta} + \frac{1}{\sin^2 \theta} \frac{\partial^2}{\partial \varphi^2} \right). \end{aligned} \quad (10.14)$$

Mit

$$\frac{1}{\sin \theta} \frac{\partial}{\partial \theta} \sin \theta \frac{\partial}{\partial \theta} = \frac{1}{\sin \theta} \cos \theta \frac{\partial}{\partial \theta} + \frac{\partial^2}{\partial \theta^2}, \quad (10.15)$$

ergibt sich schließlich

$$\hat{L}^2 = -\hbar^2 \left(\frac{1}{\sin \theta} \frac{\partial}{\partial \theta} \sin \theta \frac{\partial}{\partial \theta} + \frac{1}{\sin^2 \theta} \frac{\partial^2}{\partial \varphi^2} \right). \quad (10.16)$$

Damit finden wir, dass gilt

$$\nabla^2 = \frac{1}{r^2} \frac{\partial}{\partial r} \left(r^2 \frac{\partial}{\partial r} \right) - \frac{1}{\hbar^2 r^2} \hat{L}^2, \quad (10.17)$$

also

$$\hat{H} = -\frac{\hbar^2}{2M} \frac{1}{r^2} \frac{\partial}{\partial r} \left(r^2 \frac{\partial}{\partial r} \right) + \frac{1}{2Mr^2} \hat{L}^2 + V(r). \quad (10.18)$$

Da \hat{L} nur auf die Winkel wirkt, folgt

$$[\hat{H}, \hat{L}_z] = \frac{1}{2Mr^2} [\hat{L}^2, \hat{L}_z] = 0, \quad (10.19)$$

$$[\hat{H}, \hat{L}^2] = \frac{1}{2Mr^2} [\hat{L}^2, \hat{L}^2] = 0. \quad (10.20)$$

Da natürlich auch $[\hat{L}^2, \hat{L}_z] = 0$ ist, bilden \hat{H} , \hat{L}^2 , \hat{L}_z einen Satz verträglicher Observabler und haben ein gemeinsames vollständiges Orthonormalsystem von Eigenzuständen $|E, l, m\rangle$, die wir nun bestimmen werden. Wir haben bisher nicht gezeigt, dass es auch ein *vollständiger* Satz verträglicher Observabler ist, so dass keine weiteren Quantenzahlen zur eindeutigen Angabe der Eigenzustände $|E, l, m\rangle$ erforderlich sind, vgl. Abschnitt 7.6. Wir werden sehen aber sehen, dass das tatsächlich der Fall ist – für gegebene E , l und m finden wir höchstens eine Lösung.

Die zeitunabhängige Schrödinger-Gleichung lautet

$$\left[-\frac{\hbar^2}{2M} \frac{1}{r^2} \frac{\partial}{\partial r} \left(r^2 \frac{\partial}{\partial r} \right) + \frac{1}{2Mr^2} \hat{L}^2 + V(r) \right] \psi(\vec{r}) = E \psi(\vec{r}). \quad (10.21)$$

Wir suchen gebundene Zustände. Für diese fordern wir wieder, dass $\psi(\vec{r})$ quadratintegrabel ist. Da \hat{L}^2 nur Ableitungen nach den Winkeln θ , φ enthält, ist der Separationsansatz

$$\psi(\vec{r}) = R(r) Y(\theta, \varphi) \quad (10.22)$$

vielversprechend. Es folgt

$$-Y(\theta, \varphi) \frac{\hbar^2}{2Mr^2} \frac{\partial}{\partial r} \left(r^2 \frac{\partial R}{\partial r} \right) + R(r) \frac{1}{2Mr^2} \hat{L}^2 Y(\theta, \varphi) + V(r) R(r) Y(\theta, \varphi) = E R(r) Y(\theta, \varphi), \quad (10.23)$$

also

$$-\frac{\hbar^2}{2Mr^2} \frac{1}{R(r)} \frac{\partial}{\partial r} \left(r^2 \frac{\partial R}{\partial r} \right) + \frac{1}{2Mr^2} \frac{1}{Y(\theta, \varphi)} \hat{L}^2 Y(\theta, \varphi) + V(r) = E. \quad (10.24)$$

Wir multiplizieren mit $2Mr^2$ und bringen alle winkelabhängigen Terme auf eine Seite,

$$\frac{1}{Y(\theta, \varphi)} \hat{L}^2 Y(\theta, \varphi) = 2Mr^2 [E - V(r)] + \hbar^2 \frac{1}{R(r)} \frac{\partial}{\partial r} \left(r^2 \frac{\partial R}{\partial r} \right). \quad (10.25)$$

Die linke Seite hängt nicht von r ab, die rechte nicht von θ , φ . Da die Gleichung für alle r , θ , φ gelten soll, folgt, dass beide Seiten gleich einer Konstanten sind. Die Eigenwerte des Bahndrehimpulsquadrats \hat{L}^2 kennen wir aber schon: \hat{L}^2 und \hat{L}_z haben gemeinsame Eigenfunktionen zu den Eigenwerten $\hbar^2 l(l+1)$ und $\hbar m$ mit $l = 0, 1, 2, \dots$ und $m = -l, \dots, l$. Diese Eigenfunktionen nennen wir jetzt $Y_{lm}(\theta, \varphi)$. Es bietet sich also an, die Separationskonstante als $\hbar^2 l(l+1)$ zu schreiben. Dann gilt für den Winkelanteil

$$\hat{L}^2 Y_{lm}(\theta, \varphi) = \hbar^2 l(l+1) Y_{lm}(\theta, \varphi), \quad l = 0, 1, 2, \dots, \quad m = -l, \dots, l. \quad (10.26)$$

Um die Funktionen $Y_{lm}(\theta, \varphi)$ zu bestimmen, müssen wir die Differentialgleichung

$$-\left[\frac{1}{\sin \theta} \frac{\partial}{\partial \theta} \left(\sin \theta \frac{\partial}{\partial \theta} \right) + \frac{1}{\sin^2 \theta} \frac{\partial^2}{\partial \varphi^2} \right] Y_{lm}(\theta, \varphi) = l(l+1) Y_{lm}(\theta, \varphi) \quad (10.27)$$

explizit lösen. Als Randbedingung gilt dabei, dass die Funktion $Y_{lm}(\theta, \varphi)$ auf der Kugeloberfläche stetig sein muss. Die Lösung erfolgt durch weitere Separation der Abhängigkeiten von θ und φ . Die φ -Abhängigkeit ist einfach eine

komplexe Exponentialfunktion, da die Koeffizienten nicht von φ abhängen. Wir führen die Lösung hier nicht durch, sondern geben nur die ersten paar Lösungen (mit konventioneller Normierung) für kleine l an:

$$Y_{00}(\theta, \varphi) = \frac{1}{\sqrt{4\pi}}, \quad (10.28)$$

$$Y_{10}(\theta, \varphi) = \sqrt{\frac{3}{4\pi}} \cos \theta, \quad (10.29)$$

$$Y_{1,\pm 1}(\theta, \varphi) = \mp \sqrt{\frac{3}{8\pi}} \sin \theta e^{\pm i\varphi}, \quad (10.30)$$

$$Y_{20}(\theta, \varphi) = \sqrt{\frac{5}{16\pi}} (3 \cos^2 \theta - 1), \quad (10.31)$$

$$Y_{2,\pm 1}(\theta, \varphi) = \mp \sqrt{\frac{15}{8\pi}} \sin \theta \cos \theta e^{\pm i\varphi}, \quad (10.32)$$

$$Y_{2,\pm 2}(\theta, \varphi) = \sqrt{\frac{15}{32\pi}} \sin^2 \theta e^{\pm 2i\varphi}, \quad (10.33)$$

...

Die Y_{lm} heißen *Kugelflächenfunktionen*. Allgemein lassen sie sich durch die assoziierten Legendre-Polynome $P_l^m(z)$ ausdrücken:

$$Y_{lm}(\theta, \varphi) = \sqrt{\frac{2l+1}{4\pi} \frac{(l-m)!}{(l+m)!}} P_l^m(\cos \theta) e^{im\varphi}. \quad (10.34)$$

Die Polynome sind gegeben durch

$$\begin{aligned} P_l^m(z) &= (-1)^m \frac{(1-z^2)^{m/2}}{2^l l!} \left(\frac{d}{dz} \right)^{l+m} (z^2 - 1)^l \\ &= (-1)^{l+m} \frac{(1-z^2)^{m/2}}{2^l l!} \left(\frac{d}{dz} \right)^{l+m} (1-z^2)^l. \end{aligned} \quad (10.35)$$

Die *Radialgleichung*, d. h. die Gleichung für $R(r)$, lautet nun

$$2Mr^2 [E - V(r)] + \hbar^2 \frac{1}{R(r)} \frac{d}{dr} \left(r^2 \frac{dR}{dr} \right) = \hbar^2 l(l+1) \quad (10.36)$$

$$\Rightarrow -\frac{\hbar^2}{2Mr^2} \frac{d}{dr} \left(r^2 \frac{dR}{dr} \right) + \frac{\hbar^2 l(l+1)}{2Mr^2} R(r) + V(r)R(r) = ER(r). \quad (10.37)$$

Wie sehen die Randbedingungen aus? Für $r \rightarrow \infty$ muss $R(r)$ schneller als $1/r^{3/2}$ abfallen, so dass $\psi(\vec{r})$ quadratintegrabel ist. Das Normquadrat lautet nämlich

$$||\psi||^2 = \int d^3r |\psi(\vec{r})|^2 = \int dr d\theta d\varphi r^2 \sin \theta |\psi(\vec{r})|^2. \quad (10.38)$$

Für $\psi(\vec{r}) \sim R(r) \sim 1/r^{3/2}$ ist das radiale Integral

$$\int_0^\infty dr \frac{r^2}{r^3} = \int_0^\infty \frac{dr}{r}, \quad (10.39)$$

was für große r logarithmisch divergiert.

Für $r \rightarrow 0$ fordern wir, dass $\psi(\vec{r})$ und damit $R(r)$ beschränkt bleibt. Ein Pol der Form $1/r^\alpha$ mit $\alpha < 3/2$ wäre in drei Dimensionen zwar quadratintegrabel, aber die Schrödinger-Gleichung lässt sich damit nicht erfüllen, i. W. weil die Ableitung ∇^2 den Pol „verschlimmert“. Es folgt

$$\lim_{r \rightarrow 0} rR(r) = 0 \quad (10.40)$$

als zweite Randbedingung.

Es ist nützlich, eine neue Funktion

$$u(r) := rR(r) \quad (10.41)$$

einzuführen. Es gilt

$$\frac{d^2 u}{dr^2} = \frac{d}{dr} \frac{du}{dr} = \frac{d}{dr} (R + rR') = 2R' + rR'' \quad (10.42)$$

und andererseits

$$\frac{d}{dr} \left(r^2 \frac{dR}{dr} \right) = 2rR' + r^2 R'' = r \frac{d^2 u}{dr^2}. \quad (10.43)$$

In der Radialgleichung erhalten wir

$$-\frac{\hbar^2}{2Mr} \frac{d^2 u}{dr^2} + \frac{\hbar^2 l(l+1)}{2Mr^3} u(r) + \frac{V(r)}{r} u(r) = \frac{E}{r} u(r) \quad (10.44)$$

$$\Rightarrow -\frac{\hbar^2}{2M} \frac{d^2 u}{dr^2} + \frac{\hbar^2 l(l+1)}{2Mr^2} u(r) + V(r) u(r) = Eu(r). \quad (10.45)$$

Die Randbedingungen lauten nun $u(0) = 0$ und $u(r) \rightarrow 0$ für $r \rightarrow \infty$, schneller als $1/\sqrt{r}$.

Wir sehen, dass die Radialgleichung (10.45) dieselbe Struktur hat wie die Schrödinger-Gleichung in einer Dimension mit dem effektiven Potential

$$V_{\text{eff}}(r) = V(r) + \frac{\hbar^2 l(l+1)}{2Mr^2}. \quad (10.46)$$

Der zweite Term ist das aus der klassischen Mechanik bekannte *Zentrifugalpotential*. Wir nehmen realistischerweise an, dass $V(r)$ für $r \rightarrow 0$ nicht wie $1/r^2$ oder noch stärker divergiert. Dann dominiert für $r \rightarrow 0$ der abstoßende Zentrifugalterm, außer für $l = 0$.

Die Radialgleichung hat i. A. gebundene Lösungen $u_n(r) = rR_n(r)$, die zum diskreten Spektrum von Eigenenergien E_n gehören. Gilt $\lim_{r \rightarrow \infty} V(r) = 0$, so sind die Energien der gebundenen Zustände negativ, $E_n < 0$, siehe Abschnitt 6.1. Da die Radialgleichung l (aber nicht m) als Parameter enthält, werden die Eigenenergien und Radialfunktionen i. A. von l abhängen. Wir schreiben also die Lösungen der Radialgleichung als $u_{nl}(r)$ bzw. $R_{nl}(r)$ zu Eigenenergien E_{nl} . n zählt nun die Lösungen zu demselben l ab, falls es mehr als eine gibt. Die vollen Eigenfunktionen erhalten wir, indem wir Radial- und Winkelanteil wieder zusammensetzen:

$$\psi_{nlm}(r) = R_{nl}(r) Y_{lm}(\theta, \varphi) \quad (10.47)$$

zu den Eigenenergien E_{nl} . Beachte, dass der Winkelanteil $Y_{lm}(\theta, \varphi)$ unabhängig von der spezifischen Form des Zentralpotentials $V(r)$ ist. Zur Bestimmung von $R_{nl}(r)$ und E_{nl} benötigen wir aber diese Form.

10.2 Anwendung auf das Wasserstoff-Atom

Wenn wir die Überlegungen aus dem vorigen Abschnitt auf das Wasserstoff-Atom anwenden wollen, bemerken wir zunächst, dass dieses gar kein Ein-Teilchen-, sondern ein Zwei-Teilchen-System ist. Wir können jedoch wie in der klassischen Mechanik auf Schwerpunkts- und Relativkoordinaten transformieren. Für die Relativkoordinaten erhalten wir ein effektives Ein-Teilchen-Problem mit der reduzierten Masse

$$M = \frac{m_e m_p}{m_e + m_p} \approx m_e. \quad (10.48)$$

Dieses Ein-Teilchen-Problem wollen wir nun lösen.

Das Potential $V(r)$ ist in diesem Fall das Coulomb-Potential des Kerns (Protons),

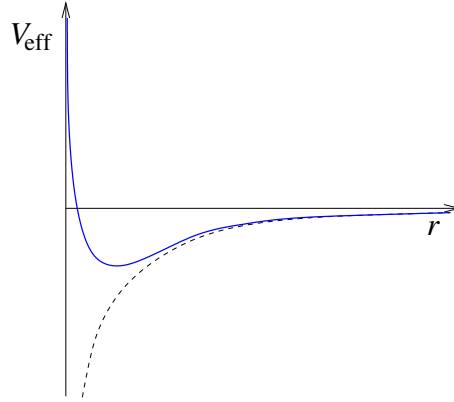
$$V(r) = -\frac{1}{4\pi\epsilon_0} \frac{e^2}{r}. \quad (10.49)$$

Wegen $\lim_{r \rightarrow \infty} V(r) = 0$ haben gebundene Zustände negative Energien $E_{nl} < 0$. Wir verallgemeinern das Problem ohne zusätzlichen Aufwand auf die wasserstoffähnlichen Ionen mit nur einem Elektron (He^+ , Li^{++} , ...), indem wir

$$V(r) = -\frac{1}{4\pi\epsilon_0} \frac{Ze^2}{r} \quad (10.50)$$

schreiben. Z ist die Kernladungszahl. Die Radialgleichung lautet nun

$$\left[-\frac{\hbar^2}{2M} \frac{d^2}{dr^2} - \underbrace{\frac{Ze^2}{4\pi\epsilon_0 r} + \frac{\hbar^2 l(l+1)}{2Mr^2}}_{= V_{\text{eff}}(r)} - E \right] u(r) = 0. \quad (10.51)$$



Effektives Potential für $l > 0$.

Durch die Reskalierung der Längen- und Energieeinheiten erhalten wir eine übersichtlichere Form:

$$\rho := Z \frac{r}{a_B}, \quad (10.52)$$

$$\eta^2 := -\frac{1}{Z^2} \frac{E}{E_R} > 0 \quad (10.53)$$

mit dem *Bohr-Radius*

$$a_B := \frac{4\pi\epsilon_0}{e^2} \frac{\hbar^2}{M} = 0,529 \text{ \AA} \quad (10.54)$$

und der *Rydberg-Energie*

$$E_R := \frac{\hbar^2}{2Ma_B^2} = 13,605 \text{ eV}, \quad (10.55)$$

vgl. Abschnitt 3.2.1. Damit wird die Radialgleichung

$$\left[\frac{d^2}{d\rho^2} + \frac{2}{\rho} - \frac{l(l+1)}{\rho^2} - \eta^2 \right] u(\rho) = 0 \quad (10.56)$$

mit den Randbedingungen

$$u(0) = 0, \quad (10.57)$$

$$u(\rho) \rightarrow 0 \quad \text{für } \rho \rightarrow \infty \text{ schneller als } 1/\sqrt{\rho}. \quad (10.58)$$

Die Lösung dieses mathematischen Problems erfolgt ähnlich wie die des harmonischen Oszillators in Abschnitt 6.3: Zunächst wird das asymptotische Verhalten bestimmt und als Faktor abgespalten. Dann wird der übrigbleibende Faktor in eine Potenzreihe in ρ entwickelt und gezeigt, dass diese abbrechen muss, um eine normierbare Lösung zu erhalten. Wir geben hier nur einige wichtige Schritte an.

1. Asymptotisches Verhalten für $\rho \rightarrow 0$: Hier können wir zunächst den Term $-\eta^2$ gegenüber $2/\rho$ vernachlässigen. Für $l = 0$ erhalten wir

$$\left(\frac{d^2}{d\rho^2} + \frac{2}{\rho}\right) u(\rho) = 0. \quad (10.59)$$

Die mit $u(0) = 0$ verträgliche Lösung ist eine Bessel-Funktion,

$$u(\rho) = \sqrt{2\rho} J_1(2\sqrt{2\rho}), \quad (10.60)$$

die sich für kleine ρ verhält wie

$$u(\rho) \cong \sqrt{2\rho}\sqrt{2\rho} = 2\rho. \quad (10.61)$$

Also ist das asymptotische Verhalten $u \sim \rho$ für $l = 0$.

Für $l \geq 1$ können wir auch den Term $2/\rho$ gegenüber $-l(l+1)/\rho^2$ vernachlässigen. Die resultierende Gleichung

$$\left(\frac{d^2}{d\rho^2} - \frac{l(l+1)}{\rho^2}\right) u(\rho) = 0 \quad (10.62)$$

hat die mit $u(0) = 0$ verträgliche Lösung

$$u(\rho) = \rho^{l+1}. \quad (10.63)$$

Also erhalten wir für *alle* $l = 0, 1, 2, \dots$ die asymptotische Form

$$u \sim \rho^{l+1}. \quad (10.64)$$

2. Asymptotisches Verhalten für $\rho \rightarrow \infty$: Hier können wir $2/\rho$ und $-l(l+1)/\rho^2$ gegenüber $-\eta^2$ vernachlässigen und erhalten die Gleichung

$$\left(\frac{d^2}{d\rho^2} - \eta^2\right) u(\rho) = 0. \quad (10.65)$$

Die normierbare Lösung lautet

$$u(\rho) = e^{-\eta\rho}. \quad (10.66)$$

3. Wir schreiben also

$$u(\rho) = \rho^{l+1} e^{-\eta\rho} P(\rho) \quad (10.67)$$

mit einer noch zu bestimmenden Funktion $P(\rho)$. Einsetzen in die Gleichung (10.56) ergibt

$$P''(\rho) + 2\left(\frac{l+1}{\rho} - \eta\right) P'(\rho) + 2\frac{1 - \eta(l+1)}{\rho} P(\rho) = 0. \quad (10.68)$$

Der Potenzreihenansatz

$$P(\rho) = \sum_{\mu=0}^{\infty} \alpha_{\mu} \rho^{\mu} \quad (10.69)$$

führt nur dann zu einer für $\rho \rightarrow \infty$ abfallenden Lösung, wenn die Reihe abbricht (ohne Beweis, aber ähnlich zu Abschnitt 6.3). Also muss ein $\mu_0 = 0, 1, 2, \dots$ existieren, so dass gilt

$$\alpha_{\mu_0} \neq 0 \quad \text{und} \quad \alpha_{\mu_0+1} = \alpha_{\mu_0+2} = \dots = 0. \quad (10.70)$$

Dann ist $P(\rho)$ ein Polynom vom Grad μ_0 . Einsetzen des Potenzreihenansatzes in die Differentialgleichung (10.68) ergibt

$$\sum_{\mu} \alpha_{\mu} \mu (\mu - 1) \rho^{\mu-2} + 2 \left(\frac{l+1}{\rho} - \eta \right) \sum_{\mu} \alpha_{\mu} \mu \rho^{\mu-1} + 2 \frac{1 - \eta(l+1)}{\rho} \sum_{\mu} \alpha_{\mu} \rho^{\mu} = 0. \quad (10.71)$$

Mit Umbenennung des Summationsindex μ in einem Teil der Terme erhalten wir

$$\begin{aligned} & \sum_{\mu} [\alpha_{\mu+1}(\mu+1)\mu + 2(l+1)\alpha_{\mu+1}(\mu+1) - 2\eta\alpha_{\mu}\mu + 2(1 - \eta(l+1))\alpha_{\mu}] \rho^{\mu-1} \\ &= \sum_{\mu} [\alpha_{\mu+1}(\mu+1)(\mu + 2(l+1)) + 2\alpha_{\mu}(1 - \eta(\mu + l + 1))] \rho^{\mu-1} = 0. \end{aligned} \quad (10.72)$$

Da die Potenzfunktionen $\rho^{\mu-1}$ linear unabhängig sind, müssen ihre Koeffizienten einzeln verschwinden. Daraus folgt die Rekursionsformel

$$\alpha_{\mu+1} = 2 \frac{\eta(\mu + l + 1) - 1}{(\mu + 1)(\mu + 2(l + 1))} \alpha_{\mu}. \quad (10.73)$$

Die Potenzreihe bricht nach dem Term der Ordnung $\mu = \mu_0$ ab, wenn gilt

$$\eta(\mu_0 + l + 1) - 1 = 0 \quad (10.74)$$

$$\Leftrightarrow \eta = \frac{1}{\mu_0 + l + 1}. \quad (10.75)$$

Die Quantenzahlen $\mu_0 = 0, 1, 2, \dots$, $l = 0, 1, 2, \dots$ und $m = -l, \dots, l$ zählen nun die diskreten, gebundenen Eigenzustände ab. Es ist aber üblich, die *Hauptquantenzahl*

$$n := \mu_0 + l + 1 = 1, 2, 3, \dots \quad (10.76)$$

zu definieren und die Zustände durch n , l , m abzuzählen. Wegen $l = n - \mu_0 - 1$ und $\mu_0 = 0, 1, 2, \dots$ ist l beschränkt auf die Werte

$$l = 0, 1, \dots, n - 1. \quad (10.77)$$

Also ist z. B. für $n = 1$ nur $l = 0$ möglich. Die Eigenenergien sind nun

$$E = -Z^2 E_R \eta^2 = -\frac{Z^2 E_R}{(\mu_0 + l + 1)^2} = -\frac{Z^2 E_R}{n^2} =: E_n. \quad (10.78)$$

Sie hängen offenbar nur von n , aber nicht von l ab, obwohl auch l als Parameter in der Radialgleichung auftritt. (Von m können die Eigenenergien ohnehin nicht abhängen, weil m in der Gleichung nicht vorkommt. Das liegt an der Kugelsymmetrie des Problems: Das Auftreten von m würde eine Drehimpulskomponente, nämlich \hat{L}_z , auszeichnen.) Die Energieniveaus des Wasserstoff-Atoms haben also eine höhere Entartung, nämlich

$$\sum_{l=0}^{n-1} (2l + 1) = n^2, \quad (10.79)$$

als aufgrund der Rotationssymmetrie zu erwarten (nämlich $2l + 1$). Dies gilt nur für das Coulomb-Potential, nicht für andere Zentralpotentiale. Das Ergebnis für E_n ergab sich schon aus der älteren Quantentheorie nach Bohr und Sommerfeld.

Die zu den Quantenzahlen n und l gehörende Lösungsfunktion $P(\rho)$ ist ein sogenanntes *Laguerre-Polynom*,

$$P(\rho) = L_{n+l}^{2l+1}(2\eta\rho), \quad (10.80)$$

mit der Definition

$$L_p^k(x) = (-1)^k \sum_{\mu=0}^{p-k} (-1)^\mu \frac{(p!)^2}{(p-k-\mu)!(k+\mu)!\mu!} x^\mu. \quad (10.81)$$

Hieraus erhalten wir $u(\rho)$ und schließlich die ursprüngliche Radialfunktion

$$R_{nl}(r) = \underbrace{\left(\frac{Z}{a_B}\right)^{3/2} \frac{2}{n^2(n+l)!} \sqrt{\frac{(n-l-1)!}{(n+l)!}}}_{\text{Normierungsfaktor}} (2\kappa r)^l e^{-\kappa r} L_{n+l}^{2l+1}(2\kappa r) \quad (10.82)$$

mit

$$\kappa := \frac{Z}{na_B}. \quad (10.83)$$

Die ersten paar Lösungen lauten

$$R_{10}(r) = 2 \left(\frac{Z}{a_B}\right)^{3/2} e^{-Zr/a_B}, \quad (10.84)$$

$$R_{20}(r) = 2 \left(\frac{Z}{a_B}\right)^{3/2} \left(1 - \frac{Zr}{2a_B}\right) e^{-Zr/2a_B}, \quad (10.85)$$

$$R_{21}(r) = \frac{1}{\sqrt{3}} \left(\frac{Z}{a_B}\right)^{3/2} \frac{Zr}{a_B} e^{-Zr/2a_B}, \quad (10.86)$$

...

Die Radialfunktionen haben $\mu_0 = n - l - 1 \geq 0$ Nullstellen für $r > 0$. Dies ergibt sich aus dem Knotensatz, siehe Abschnitt 6.4.1. Sie haben eine weitere Nullstelle bei $r = 0$ genau dann, wenn $l > 0$ ist. Wie aus der Atomphysik bekannt, nennen wir die Zustände mit Drehimpulsquantenzahl $l = 0, 1, 2, \dots$ die *s*, *p*, *d*, *f*, *g*, ...-Orbitale des Atoms.

Die vollen Eigenfunktionen sind $\psi_{nlm}(\vec{r}) = R_{nl}(r) Y_{lm}(\theta, \varphi)$. Abbildungen davon finden sich in allen Quantenmechanik-Lehrbüchern. Es sei darauf hingewiesen, dass relativistische Effekte, die Wechselwirkung mit dem Kernspin und die kleine aber endliche Ausdehnung des Kerns die Ergebnisse geringfügig ändern und insbesondere dafür sorgen, dass die Eigenenergien E_{nl} auch schwach von l abhängen.

Kapitel 11

Näherungsmethoden

In den vorausgegangenen Kapiteln haben wir eine ganze Reihe von Modellen exakt gelöst, d. h. wir haben geschlossene Ausdrücke für die Eigenwerte und Eigenzustände des Hamilton-Operators gefunden. Dies vermittelt den Eindruck, quantenmechanische Probleme seien typischerweise oder zumindest oft exakt lösbar. Das ist keineswegs der Fall. Die Kenntnis exakter Lösungen ist die große Ausnahme, v. a. für Systeme aus mehreren wechselwirkenden Teilchen. Wenn sich keine exakte Lösung finden lässt, bleiben i. W. zwei konzeptionell verschiedene Zugänge:

- numerische Lösung,
- Näherungsmethoden.

Numerische Lösungen können ausnutzen, dass man quantenmechanische Probleme oft in verschiedenen mathematischen Sprachen formulieren kann. Die zeitunabhängige Schrödinger-Gleichung für Teilchen im Ortsraum lässt sich einerseits als partielle Differentialgleichung oder System partieller Differentialgleichungen auffassen und andererseits als Eigenwertproblem des linearen Hamilton-Operators auf dem Hilbert-Raum. Das Problem bei der exakten numerischen Lösung ist oft, dass die Dimension des Systems sehr groß ist. Als Beispiel betrachten wir ein Heisenberg-Modell von Spins der Länge $1/2$ auf einem einfach kubischen Gitter. Der Hamiltonian lautet

$$\hat{H} = -J \sum_{\langle ij \rangle} \hat{S}_i \cdot \hat{S}_j. \quad (11.1)$$

Hier ist J die Stärke der Wechselwirkung (hier ist J/\hbar^2 die „Austauschenergie“, die in der Vorlesung Quantentheorie 2 besprochen wird) und die \hat{S}_i sind Spins an Gitterplätzen i des kubischen Gitters. Die Summe läuft über alle Nächste-Nachbar-Bindungen. Für eine numerische Lösung muss man die Dimension des Hilbert-Raums endlich halten. Dazu beschränkt man die Summe auf einen Würfel der Kantenlänge L (wir wählen die Gitterkonstante zu 1), üblicherweise mit periodischen Randbedingungen. Der Hilbert-Raum eines Spins ist zweidimensional. Der Gesamt-Hilbert-Raum ist ein Produktraum der Dimension 2^{L^3} für die L^3 Spins auf dem Gitter. Damit die Ergebnisse aussagekräftig für den eigentlich interessierenden thermodynamischen Limes $L \rightarrow \infty$ sind, muss $L \gg 1$ gelten. Nehmen wir als Beispiel $L = 10$, was noch nicht sehr groß ist. Dann ist die Dimension des Hilbert-Raums

$$2^{L^3} = 2^{1000} \approx 1,07 \cdot 10^{301}. \quad (11.2)$$

Diese Zahl ist weitaus größer als die Zahl der Elementarteilchen im beobachtbaren Universum. Man hat also keine Chance, einen Eigenvektor abzuspeichern, geschweige denn, ihn auszurechnen. Numerische Methoden sind aber für niedrigdimensionale Probleme sinnvoll, z. B. für ein Teilchen in einem komplizierten Potential $V(\vec{r})$.

Andererseits haben Näherungsmethoden den offensichtlichen Nachteil, schon vom Prinzip her keine exakten Ergebnisse zu liefern. Nützlich ist, wenn man die Näherung systematisch verbessern kann und ihren Fehler abschätzen kann, dies sind Charakteristika für *kontrollierte Näherungen*. Ein Beispiel ist die Annäherung einer Funktion durch die Partialsummen ihrer Taylor-Reihe.

Näherungsmethoden können auch Vorteile haben, die darüber hinaus gehen, dass nichts anderes möglich ist. Man gewinnt durch Näherungen Einsicht in die Physik eines Modells, die numerische Verfahren zunächst nicht

bieten, weil (a) Näherungsmethoden funktionale Abhängigkeiten von Parametern liefern können, nicht nur Zahlen für bestimmte Parametersätze, und (b) der Erfolg bestimmter Näherungsmethoden etwas über den Charakter des Systems aussagt. Wir werden in diesem Kapitel verschiedene Näherungsverfahren besprechen.

11.1 Variationsverfahren

Die Gleichungen der klassischen Mechanik kann man elegant aus Variationsprinzipien herleiten, insbesondere aus dem Hamiltonschen Prinzip. Solche Prinzipien kann man auch in der Quantentheorie formulieren. Sie liefern nicht nur eine alternative Betrachtung von Eigenwertgleichungen, z. B. der zeitunabhängigen Schrödinger-Gleichung, sondern führen auch auf eine nützliche Näherungsmethode.

11.1.1 Das Variationsprinzip für Eigenvektoren

Sei \hat{A} eine Observable. Ihr Erwartungswert in einem Zustand $|\psi\rangle$ ist

$$\langle \hat{A} \rangle_\psi = \frac{\langle \psi | \hat{A} | \psi \rangle}{\langle \psi | \psi \rangle}. \quad (11.3)$$

Dies können wir als *Funktional* von $|\psi\rangle$ auffassen. In diesem Abschnitt lassen wir zu, dass $|\psi\rangle$ nicht normiert (aber normierbar) ist. Daher müssen wir explizit durch das Normquadrat $\|\psi\|^2 = \langle \psi | \psi \rangle$ teilen.

Die Variation von $\langle \hat{A} \rangle_\psi$ – in demselben Sinne wie in der klassischen Mechanik – ist

$$\begin{aligned} \delta \langle \hat{A} \rangle_\psi &= \frac{(\delta \langle \psi | \hat{A} | \psi \rangle) \langle \psi | \psi \rangle - \langle \psi | \hat{A} | \psi \rangle \delta \langle \psi | \psi \rangle}{\langle \psi | \psi \rangle^2} \\ &= \frac{\delta \langle \psi | \hat{A} | \psi \rangle}{\langle \psi | \psi \rangle} - \frac{\langle \psi | \hat{A} | \psi \rangle \delta \langle \psi | \psi \rangle}{\langle \psi | \psi \rangle^2} \\ &= \frac{\langle \delta \psi | \hat{A} | \psi \rangle + \langle \psi | \hat{A} | \delta \psi \rangle}{\langle \psi | \psi \rangle} - \frac{\langle \hat{A} \rangle_\psi (\langle \delta \psi | \psi \rangle + \langle \psi | \delta \psi \rangle)}{\langle \psi | \psi \rangle} \\ &= \frac{\langle \delta \psi | (\hat{A} - \langle \hat{A} \rangle_\psi) | \psi \rangle}{\langle \psi | \psi \rangle} + \frac{\langle \psi | (\hat{A} - \langle \hat{A} \rangle_\psi) | \delta \psi \rangle}{\langle \psi | \psi \rangle} \\ &= \frac{\langle \delta \psi | (\hat{A} - \langle \hat{A} \rangle_\psi) | \psi \rangle}{\langle \psi | \psi \rangle} + \frac{[\langle \delta \psi | (\hat{A} - \langle \hat{A} \rangle_\psi) | \psi \rangle]^*}{\langle \psi | \psi \rangle} \\ &= \frac{2 \operatorname{Re} \langle \delta \psi | (\hat{A} - \langle \hat{A} \rangle_\psi) | \psi \rangle}{\langle \psi | \psi \rangle}. \end{aligned} \quad (11.4)$$

Wir haben verwendet, dass \hat{A} hermitesch ist (Observable!) und $\langle \hat{A} \rangle_\psi$ damit reell. Nun formulieren und beweisen wir das Extremalprinzip für Eigenvektoren:

Ein Vektor $|\psi\rangle \neq 0$ aus dem Hilbert-Raum ist Eigenvektor zur Observablen \hat{A} genau dann, wenn gilt

$$\delta \langle \hat{A} \rangle_\psi = 0. \quad (11.5)$$

Der zugehörige Eigenwert ist dann $\langle \hat{A} \rangle_\psi$.

Beweis: $\delta \langle \hat{A} \rangle_\psi = 0$ ist äquivalent zu

$$\frac{2 \operatorname{Re} \langle \delta \psi | (\hat{A} - \langle \hat{A} \rangle_\psi) | \psi \rangle}{\langle \psi | \psi \rangle} = 0 \quad (11.6)$$

und damit zu

$$\operatorname{Re} \langle \delta \psi | (\hat{A} - \langle \hat{A} \rangle_\psi) | \psi \rangle = 0, \quad (11.7)$$

für alle infinitesimalen Variationen $\langle \delta\psi |$ von Bra-Vektoren. Da mit $\langle \delta\psi |$ auch $-i\langle \delta\psi |$ eine mögliche Variation ist, schließt Gl. (11.7) die Aussage

$$\operatorname{Re}(-i\langle \delta\psi |(\hat{A} - \langle \hat{A} \rangle_\psi)|\psi\rangle) = \operatorname{Im}\langle \delta\psi |(\hat{A} - \langle \hat{A} \rangle_\psi)|\psi\rangle = 0 \quad (11.8)$$

mit ein. Also ist $\delta\langle \hat{A} \rangle_\psi = 0$ äquivalent zu

$$\langle \delta\psi |(\hat{A} - \langle \hat{A} \rangle_\psi)|\psi\rangle = 0 \quad (11.9)$$

für alle Variationen $\langle \delta\psi |$. Dies ist wiederum äquivalent zu

$$(\hat{A} - \langle \hat{A} \rangle_\psi)|\psi\rangle = 0 \quad \Leftrightarrow \quad \hat{A}|\psi\rangle = \langle \hat{A} \rangle_\psi|\psi\rangle. \quad (11.10)$$

Das ist gerade die Eigenwertgleichung für den Eigenvektor $|\psi\rangle$ zum Eigenwert $\langle \hat{A} \rangle_\psi$. \square

Wir haben damit gezeigt, dass der Erwartungswert $\langle \hat{A} \rangle_\psi$ genau für die Eigenvektoren von \hat{A} *stationär* wird. Dieses Variationsprinzip unterscheidet nicht zwischen verschiedenen Eigenvektoren. Es folgt aber sofort, dass Eigenvektoren zum *minimalen* Eigenwert A_0 den Erwartungswert $\langle \hat{A} \rangle_\psi$ *minimieren*: Das Minimum des Funktionals $\langle \hat{A} \rangle_\psi$ von $|\psi\rangle$ ist natürlich ein spezieller stationärer Punkt, d. h. für Grundzustandsvektoren ψ_0 gilt

$$\delta\langle \hat{A} \rangle_{\psi_0} = 0. \quad (11.11)$$

Es muss noch gezeigt werden, dass $\langle \hat{A} \rangle_\psi$ nicht kleiner werden kann als

$$\langle \hat{A} \rangle_{\psi_0} = \frac{\langle \psi_0 | \hat{A} | \psi_0 \rangle}{\langle \psi_0 | \psi_0 \rangle} = A_0. \quad (11.12)$$

Beweis: Die Eigenwertgleichung für \hat{A} laute $\hat{A}|\psi_n\rangle = A_n|\psi_n\rangle$ mit $A_n \geq A_0 \forall n$. Es folgt

$$\begin{aligned} \langle \hat{A} \rangle_\psi &= \frac{\langle \psi | \hat{A} | \psi \rangle}{\langle \psi | \psi \rangle} = \frac{\sum_n \langle \psi | \hat{A} | \psi_n \rangle \langle \psi_n | \psi \rangle}{\langle \psi | \psi \rangle} \\ &= \frac{\sum_n \langle \psi | A_n | \psi_n \rangle \langle \psi_n | \psi \rangle}{\langle \psi | \psi \rangle} \\ &= \frac{\sum_n A_n |\langle \psi_n | \psi \rangle|^2}{\langle \psi | \psi \rangle} \\ &\geq \frac{\sum_n A_0 |\langle \psi_n | \psi \rangle|^2}{\langle \psi | \psi \rangle} \\ &= A_0 \frac{\sum_n \langle \psi | \psi_n \rangle \langle \psi_n | \psi \rangle}{\langle \psi | \psi \rangle} \\ &= A_0 \frac{\langle \psi | \mathbb{1} | \psi \rangle}{\langle \psi | \psi \rangle} = A_0. \end{aligned} \quad (11.13)$$

\square

11.1.2 Ritzsches Variationsverfahren

Die Näherungsmethode besteht nun darin, $\delta\langle \hat{A} \rangle_\psi = 0$ nicht für $|\psi\rangle$ aus dem gesamten Hilbert-Raum auszuwerten, sondern für eine Teilmenge von *Testzuständen*. Sinnvollerweise wählt man eine Teilmenge, auf der sich die Erwartungswerte $\langle \hat{A} \rangle_\psi$ leicht berechnen lassen. Die Lösungen von $\delta\langle \hat{A} \rangle_\psi = 0$ auf der Teilmenge ergeben dann Näherungen für die wahren Eigenzustände und Eigenwerte. Oft interessiert man sich für den kleinsten Eigenwert oder einige wenige niedrig liegende Eigenwerte. Wendet man diese Ideen auf den Hamilton-Operator \hat{H} an, so spricht man vom *Ritzschen Verfahren*.

Grundzustände des Hamiltonians minimieren den Erwartungswert $\langle \hat{H} \rangle_\psi$, wie eben gesehen. Sei $M \subseteq \mathcal{H}$ eine Menge von Testzuständen (\mathcal{H} ist der Hilbert-Raum) und $|\psi\rangle \in M$. Dann gilt offensichtlich

$$\langle \hat{H} \rangle_\psi \equiv \frac{\langle \psi | \hat{H} | \psi \rangle}{\langle \psi | \psi \rangle} \geq E_0, \quad (11.14)$$

wo E_0 die exakte Grundzustandsenergie ist. Die Minimierung von $\langle \hat{H} \rangle_\psi$ auf M liefert damit eine obere Schranke von E_0 und zwar die optimale für alle $|\psi\rangle \in M$. Wir nehmen an, dass $|\psi\rangle$ von o. B. d. A. reellen Variationsparametern $\alpha_1, \dots, \alpha_n$ abhängt. Dann erhalten wir alle stationären Punkte durch Lösung von

$$\frac{\partial}{\partial \alpha_i} \langle \hat{H} \rangle_\psi = 0 \quad \forall i = 1, \dots, n. \quad (11.15)$$

Gibt es mehrere Lösungen, nehmen wir diejenige mit dem kleinsten $\langle \hat{H} \rangle_\psi$. Diese entspricht dann dem globalen Minimum von $\langle \hat{H} \rangle_\psi$ auf M .

In der Praxis wählt man $|\psi(\alpha_1, \dots, \alpha_n)\rangle$ natürlich so, dass

- man $\langle \hat{H} \rangle_\psi$ effizient berechnen kann und
- die Testzustände berücksichtigen, was man schon über die Lösung weiß, z. B. aus Symmetrieüberlegungen oder dem Knotensatz.

Beispiel: Wir tun so, als hätten wir vergessen, dass die Grundzustandswellenfunktion des harmonischen Oszillators eine Gauß-Funktion ist. Da es eine gerade Funktion ohne Nullstellen sein sollte, setzen wir eine Lorentz-Funktion an:

$$\psi_\alpha(x) = \frac{1}{x^2 + \alpha^2}. \quad (11.16)$$

Der Ansatz enthält einen reellen Parameter α . Es ist

$$\langle \hat{H} \rangle_\psi = \frac{\int_{-\infty}^{\infty} dx \psi_\alpha^*(x) H \psi_\alpha(x)}{\int_{-\infty}^{\infty} dx \psi_\alpha^*(x) \psi_\alpha(x)} \quad (11.17)$$

mit

$$\int_{-\infty}^{\infty} dx \psi_\alpha^*(x) \psi_\alpha(x) = \int_{-\infty}^{\infty} \frac{dx}{(x^2 + \alpha^2)^2} \stackrel{\text{Tafel}}{=} \frac{\pi}{2\alpha^3} \quad (11.18)$$

und

$$\begin{aligned} \int_{-\infty}^{\infty} dx \psi_\alpha^*(x) H \psi_\alpha(x) &= -\frac{\hbar^2}{2m} \int_{-\infty}^{\infty} dx \frac{1}{x^2 + \alpha^2} \frac{d^2}{dx^2} \frac{1}{x^2 + \alpha^2} + \frac{1}{2} m \omega^2 \int_{-\infty}^{\infty} dx \frac{x^2}{(x^2 + \alpha^2)^2} \\ &\stackrel{\text{part.}}{=} +\frac{\hbar^2}{2m} \int_{-\infty}^{\infty} dx \left(\frac{d}{dx} \frac{1}{x^2 + \alpha^2} \right) \frac{d}{dx} \frac{1}{x^2 + \alpha^2} + \frac{1}{2} m \omega^2 \int_{-\infty}^{\infty} dx \frac{x^2}{(x^2 + \alpha^2)^2} \\ &= \frac{\hbar^2}{2m} \int_{-\infty}^{\infty} dx \left(\frac{2x}{(x^2 + \alpha^2)^2} \right)^2 + \frac{1}{2} m \omega^2 \int_{-\infty}^{\infty} dx \frac{x^2}{(x^2 + \alpha^2)^2} \\ &\stackrel{\text{Tafel}}{=} \frac{2\hbar^2}{m} \frac{\pi}{16\alpha^5} + \frac{1}{2} m \omega^2 \frac{\pi}{2\alpha} \\ &= \frac{\pi}{8} \frac{\hbar^2}{m\alpha^5} + \frac{\pi}{4} \frac{m\omega^2}{\alpha}. \end{aligned} \quad (11.19)$$

Also ist

$$\langle \hat{H} \rangle_\psi = \frac{\frac{\pi}{8} \frac{\hbar^2}{m\alpha^5} + \frac{\pi}{4} \frac{m\omega^2}{\alpha}}{\frac{\pi}{2\alpha^3}} = \frac{1}{4} \frac{\hbar^2}{m\alpha^2} + \frac{1}{2} m \omega^2 \alpha^2. \quad (11.20)$$

Das Minimum bestimmen wir mittels

$$0 = \frac{d}{d\alpha} \langle \hat{H} \rangle_\psi = -\frac{\hbar^2}{2m\alpha^3} + m\omega^2 \alpha \quad (11.21)$$

$$\Rightarrow -\frac{\hbar^2}{2m^2\omega^2} + \alpha^4 = 0 \quad (11.22)$$

$$\Rightarrow \alpha^2 = \sqrt{\frac{\hbar^2}{2m^2\omega^2}} = \frac{1}{\sqrt{2}} \frac{\hbar}{m\omega} \quad \left| \alpha \in \mathbb{R} \text{ erfordert } \alpha^2 \geq 0 \right. \quad (11.23)$$

$$\Rightarrow \alpha = \pm \frac{1}{2^{1/4}} \sqrt{\frac{\hbar}{m\omega}}. \quad (11.24)$$

Die Näherung für die Grundzustandsenergie ist damit

$$\langle \hat{H} \rangle_\psi = \frac{\hbar\omega}{\sqrt{2}} = \sqrt{2} \frac{\hbar\omega}{2}. \quad (11.25)$$

Dieser Wert ist offensichtlich größer als die exakte Grundzustandsenergie $\hbar\omega/2$, aber immerhin von derselben Größenordnung. Die optimierte Näherung für die zugehörige Wellenfunktion, die Lorentz-Funktion, ist dagegen für große $|x|$ beliebig verschieden von der exakten Gauß-Funktion.

Die Abhängigkeit von $|\psi\rangle$ von den Parametern α_i kann im Prinzip beliebig sein, das Variationsproblem nimmt aber eine besonders einfache Form an, wenn die Testzustände linear von den α_i abhängen:

$$|\psi\rangle = \sum_{i=1}^n \alpha_i |\varphi_i\rangle, \quad (11.26)$$

wobei $|\varphi_i\rangle$ feste Hilbert-Raum-Vektoren sind, die nicht unbedingt normiert oder orthogonal zueinander sein müssen. Dann ist

$$\begin{aligned} \langle \psi | \psi \rangle &= \sum_{jk} \underbrace{\alpha_j}_{\text{reell}} \alpha_k \langle \varphi_j | \varphi_k \rangle \\ \Rightarrow \frac{\partial}{\partial \alpha_i} \langle \psi | \psi \rangle &= \sum_{jk} (\delta_{ij} \alpha_k + \alpha_j \delta_{ik}) \langle \varphi_j | \varphi_k \rangle \\ &= \sum_k \alpha_k \langle \varphi_i | \varphi_k \rangle + \sum_j \alpha_j \langle \varphi_j | \varphi_i \rangle \\ &= 2 \sum_j \alpha_j \text{Re} \langle \varphi_j | \varphi_i \rangle \end{aligned} \quad (11.27)$$

und analog

$$\langle \psi | \hat{H} | \psi \rangle = \sum_{jk} \alpha_j \alpha_k \langle \varphi_j | \hat{H} | \varphi_k \rangle \quad (11.29)$$

$$\Rightarrow \frac{\partial}{\partial \alpha_i} \langle \psi | \hat{H} | \psi \rangle = 2 \sum_j \alpha_j \text{Re} \langle \varphi_j | \hat{H} | \varphi_i \rangle. \quad (11.30)$$

Die Bedingungen für stationäre Lösungen lautet dann

$$\begin{aligned} 0 &= \frac{\partial}{\partial \alpha_i} \langle \hat{H} \rangle_\psi = \frac{\partial}{\partial \alpha_i} \frac{\langle \psi | \hat{H} | \psi \rangle}{\langle \psi | \psi \rangle} \\ &= \frac{1}{\langle \psi | \psi \rangle} \frac{\partial}{\partial \alpha_i} \langle \psi | \hat{H} | \psi \rangle - \frac{\langle \psi | \hat{H} | \psi \rangle}{\langle \psi | \psi \rangle^2} \frac{\partial}{\partial \alpha_i} \langle \psi | \psi \rangle \\ &= \frac{1}{\langle \psi | \psi \rangle} 2 \sum_j \alpha_j \text{Re} \langle \varphi_j | \hat{H} | \varphi_i \rangle - \frac{\langle \hat{H} \rangle_\psi}{\langle \psi | \psi \rangle} 2 \sum_j \alpha_j \text{Re} \langle \varphi_j | \varphi_i \rangle \end{aligned} \quad (11.31)$$

$$\Rightarrow \sum_j \alpha_j (\text{Re} \langle \varphi_j | \hat{H} | \varphi_i \rangle - \langle \hat{H} \rangle_\psi \text{Re} \langle \varphi_j | \varphi_i \rangle) = 0 \quad (11.32)$$

$$\Rightarrow \sum_j \operatorname{Re} \langle \varphi_j | \hat{H} | \varphi_i \rangle \alpha_j = \langle \hat{H} \rangle_\psi \sum_j \operatorname{Re} \langle \varphi_j | \varphi_i \rangle \alpha_j. \quad (11.33)$$

Mit den Definitionen

$$\tilde{H}_{ij} := \operatorname{Re} \langle \varphi_j | \hat{H} | \varphi_i \rangle, \quad (11.34)$$

$$\tilde{I}_{ij} := \operatorname{Re} \langle \varphi_j | \varphi_i \rangle, \quad (11.35)$$

$$E_{\text{var}} := \langle \hat{H} \rangle_\psi, \quad (11.36)$$

$$\vec{\alpha} := \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix} \quad (11.37)$$

können wir dies als Matrixgleichung

$$\tilde{H} \vec{\alpha} = E_{\text{var}} \tilde{I} \vec{\alpha} \quad (11.38)$$

schreiben. Das ist eine *verallgemeinerte Eigenwertgleichung* (verallgemeinert wegen des Auftretens der Matrix \tilde{I} , die nicht die Einheitsmatrix sein muss), wobei der Näherungswert der Grundzustandsenergie, E_{var} , als Eigenwert und der Vektor der Koeffizienten α_i als Eigenvektor auftritt. Wir können Gl. (11.38) als effektive Schrödinger-Gleichung für die Näherungslösung ansehen.

Lösungen mit $\vec{\alpha} \neq 0$ existieren für

$$\det(\tilde{H} - E_{\text{var}} \tilde{I}) = 0. \quad (11.39)$$

Wenn wir für $\{|\varphi_i\rangle\}$ ein vollständiges Orthonormalsystem nehmen, ist \tilde{H} die Matrixdarstellung des Hamiltonians in dieser Basis und $\tilde{I} = \mathbb{1}$. Dann erhalten wir die zeitunabhängige Schrödinger-Gleichung zurück. Aber die Motivation für die Näherungsmethode ist natürlich, dass wir die volle Schrödinger-Gleichung nicht lösen können.

Das Ritz-Verfahren kann auch zur Bestimmung angeregter Zustände verwendet werden. Dazu nutzt man aus, dass Eigenzustände orthogonal sind. Um den ersten angeregten Zustand zu finden, kann man also die Menge M der Testzustände auf solche beschränken, die orthogonal zum (vermutlich auch genäherten) Grundzustand $|\psi_0\rangle$ sind. Alternativ implementiert man die Nebenbedingung $\langle \psi | \psi_0 \rangle = 0$ mittels eines Lagrange-Multiplikators bei der Minimierung. Die Erweiterung auf höhere angeregte Zustände ist offensichtlich.

Abschließend einige Bemerkungen zur Beurteilung des Ritz-Verfahrens:

- Es ist immer durchführbar und liefert garantiert eine obere Schranke für die Grundzustandsenergie.
- Es stellt keine kontrollierte Näherung dar; zwar kann man die Näherung durch Vergrößerung der Menge der Testzustände potentiell verbessern, es gibt aber keine systematische Methode dafür. Auch ist der Fehler nicht systematisch abschätzbar.
- Meist ist die Näherung für die Eigenzustände deutlich schlechter als die für die Eigenwerte.

11.2 Zeitunabhängige Störungstheorie

Für viele – aber längst nicht alle – physikalischen Systeme kann man den Hamilton-Operator in zwei Teile zerlegen,

$$\hat{H} = \hat{H}_0 + \hat{H}_1, \quad (11.40)$$

wobei sich die Eigenenergien und Eigenzustände von \hat{H}_0 exakt bestimmen lassen und \hat{H}_1 eine „kleine Störung“ darstellt. Da \hat{H}_1 ein Operator ist, müssen wir sagen, was wir mit „klein“ meinen. Das geschieht letztlich *a posteriori*: Wir können \hat{H}_1 als kleine Störung betrachten, wenn die Störungstheorie vernünftige Resultate liefert. Das ist typischerweise der Fall, wenn für die relevanten Eigenzustände von \hat{H}_0 (meist diejenigen zu kleinen Energien) die Matrixelemente von \hat{H}_1 betragsmäßig klein im Vergleich zur Energieskala von \hat{H}_0 sind. Die zeitunabhängige Störungstheorie ist sehr wichtig in der Physik. Zum Beispiel sind die präzisen Rechnungen für die g -Faktoren von Elektron und Myon störungstheoretischer Natur. Die störungstheoretische Methode ist, ähnlich wie Variationsverfahren, nicht auf Eigenzustände des Hamilton-Operators beschränkt, diese stellen aber die wichtigste Anwendung dar. Wir entwickeln die Methode zunächst für den Fall nichtentartete Zustände von \hat{H}_0 .

11.2.1 Störungstheorie für nichtentartete Energieniveaus

Wir suchen einen gewissen Eigenzustand $|\psi_n\rangle$ und den zugehörigen Eigenwert E_n des Hamilton-Operators \hat{H} . Die Eigenwertgleichung

$$\hat{H}|\psi_n\rangle = E_n|\psi_n\rangle \quad (11.41)$$

können wir i. A. nicht exakt lösen. Wir können aber

$$\hat{H} = \hat{H}_0 + \hat{H}_1 \quad (11.42)$$

schreiben und die Eigenwertgleichung

$$\hat{H}_0|\psi_n^{(0)}\rangle = E_n^{(0)}|\psi_n^{(0)}\rangle \quad (11.43)$$

für den *ungestörten* Hamiltonian \hat{H}_0 lösen. Wir sortieren die Eigenwerte so, dass gilt

$$E_m \leq E_n \quad \text{für } m < n, \quad (11.44)$$

$$E_m^{(0)} \leq E_n^{(0)} \quad \text{für } m < n. \quad (11.45)$$

In diesem Abschnitt betrachten wir den Fall, dass $E_n^{(0)}$ für den untersuchten Zustand n nicht entartet ist. Die übrigen Eigenwerte von \hat{H}_0 dürfen entartet sein; wir zählen sie trotzdem nur mit einem Index m ab – bei Entartung sind dann manche $E_m^{(0)}$ gleich. Die Eigenzustände $|\psi_m^{(0)}\rangle$ zu \hat{H}_0 wählen wir orthonormal:

$$\langle \psi_m^{(0)} | \psi_n^{(0)} \rangle = \delta_{mn}. \quad (11.46)$$

Natürlich können wir für die Eigenzustände $|\psi_m\rangle$ zu \hat{H} ebenfalls Orthonormalität fordern, es wird sich aber als nützlich erweisen, den gesuchten Zustand $|\psi_n\rangle$ stattdessen durch

$$\langle \psi_n^{(0)} | \psi_n \rangle = 1 \quad (11.47)$$

zu normieren. (Das funktioniert offenbar nicht, wenn $|\psi_n\rangle$ zu $|\psi_n^{(0)}\rangle$ orthogonal ist. Dann ist aber die Störungstheorie ohnehin nicht anwendbar.) $|\psi_n\rangle$ muss dann ggf. am Ende noch normiert werden. Wir versehen nun die Störung \hat{H}_1 mit einem Vorfaktor λ :

$$\hat{H} = \hat{H}_0 + \lambda \hat{H}_1. \quad (11.48)$$

Natürlich ist am Ende $\lambda = 1$ zu setzen. Die Einführung dieses Parameters gestattet uns aber, alle Größen nach Ordnungen von \hat{H}_1 zu entwickeln, da dies dieselben Terme sind wie in der Entwicklung nach λ . Wir vermeiden damit, Potenzreihen-Entwicklungen nach Operatoren untersuchen zu müssen. Wir schreiben

$$E_n = E_n^{(0)} + \lambda E_n^{(1)} + \lambda^2 E_n^{(2)} + \dots, \quad (11.49)$$

$$|\psi_n\rangle = |\psi_n^{(0)}\rangle + \lambda |\psi_n^{(1)}\rangle + \lambda^2 |\psi_n^{(2)}\rangle + \dots \quad (11.50)$$

Der Superskript in Klammern gibt also immer die Ordnung in λ bzw. \hat{H}_1 an. Multiplikation der letzten Gleichung mit $\langle \psi_n^{(0)} |$ von links ergibt

$$\langle \psi_n^{(0)} | \psi_n \rangle = \langle \psi_n^{(0)} | \psi_n^{(0)} \rangle + \lambda \langle \psi_n^{(0)} | \psi_n^{(1)} \rangle + \lambda^2 \langle \psi_n^{(0)} | \psi_n^{(2)} \rangle + \dots \quad (11.51)$$

$$\Rightarrow 1 = 1 + \lambda \langle \psi_n^{(0)} | \psi_n^{(1)} \rangle + \lambda^2 \langle \psi_n^{(0)} | \psi_n^{(2)} \rangle + \dots \quad (11.52)$$

$$\Rightarrow \lambda \langle \psi_n^{(0)} | \psi_n^{(1)} \rangle + \lambda^2 \langle \psi_n^{(0)} | \psi_n^{(2)} \rangle + \dots = 0. \quad (11.53)$$

Da die Potenzfunktionen λ^j linear unabhängig sind, folgt

$$\langle \psi_n^{(0)} | \psi_n^{(j)} \rangle = 0 \quad \text{für } j \geq 1 \quad (11.54)$$

und daher

$$\langle \psi_n^{(0)} | \psi_n^{(j)} \rangle = \delta_{0j}. \quad (11.55)$$

Einsetzen von Glg. (11.49) und (11.50) in die volle Schrödinger-Gleichung ergibt

$$\begin{aligned}
\hat{H}|\psi_n\rangle &= (\hat{H}_0 + \lambda\hat{H}_1)(|\psi_n^{(0)}\rangle + \lambda|\psi_n^{(1)}\rangle + \lambda^2|\psi_n^{(2)}\rangle + \dots) \\
&= \hat{H}_0|\psi_n^{(0)}\rangle + \sum_{k=1}^{\infty} \lambda^k (\hat{H}_0|\psi_n^{(k)}\rangle + \hat{H}_1|\psi_n^{(k-1)}\rangle) \\
&\stackrel{!}{=} E_n|\psi_n\rangle \\
&= (E_n^{(0)} + \lambda E_n^{(1)} + \lambda^2 E_n^{(2)} + \dots)(|\psi_n^{(0)}\rangle + \lambda|\psi_n^{(1)}\rangle + \lambda^2|\psi_n^{(2)}\rangle + \dots) \\
&= E_n^{(0)}|\psi_n^{(0)}\rangle + \sum_{k=1}^{\infty} \lambda^k \sum_{l=0}^k E_n^{(l)}|\psi_n^{(k-l)}\rangle.
\end{aligned} \tag{11.56}$$

Da die Potenzfunktionen λ^k linear unabhängig sind, gilt die Gleichheit für jede Ordnung separat. Es folgt für die Ordnung $k = 0$:

$$\hat{H}_0|\psi_n^{(0)}\rangle = E_n^{(0)}|\psi_n^{(0)}\rangle, \tag{11.57}$$

also nicht Neues, und für $k \geq 1$:

$$\hat{H}_0|\psi_n^{(k)}\rangle + \hat{H}_1|\psi_n^{(k-1)}\rangle = \sum_{l=0}^k E_n^{(l)}|\psi_n^{(k-l)}\rangle. \tag{11.58}$$

Multiplikation von links mit $\langle\psi_n^{(0)}|$ ergibt mit Gl. (11.55)

$$\begin{aligned}
\underbrace{\langle\psi_n^{(0)}|\hat{H}_0|\psi_n^{(k)}\rangle}_{= \langle\psi_n^{(0)}|E_n^{(0)}|\psi_n^{(k)}\rangle = 0} + \langle\psi_n^{(0)}|\hat{H}_1|\psi_n^{(k-1)}\rangle &= \sum_{l=0}^k E_n^{(l)} \underbrace{\langle\psi_n^{(0)}|\psi_n^{(k-l)}\rangle}_{= \delta_{0,k-l}} \\
\Rightarrow E_n^{(k)} &= \langle\psi_n^{(0)}|\hat{H}_1|\psi_n^{(k-1)}\rangle.
\end{aligned} \tag{11.59}$$

$$\Rightarrow E_n^{(k)} = \langle\psi_n^{(0)}|\hat{H}_1|\psi_n^{(k-1)}\rangle. \tag{11.60}$$

Damit erhalten wir die Korrektur der Ordnung k zur Energie aus der Korrektur der Ordnung $k-1$ zum Zustand. Besonders einfach ist die Energiekorrektur erster Ordnung:

$$E_n^{(1)} = \langle\psi_n^{(0)}|\hat{H}_1|\psi_n^{(0)}\rangle. \tag{11.61}$$

Sie ist also der Erwartungswert des Störungsoperators im ungestörten Zustand.

Die Zustandskorrekturen mit $k \geq 1$ erhalten wir aus Gl. (11.58) durch Multiplikation von links mit $\langle\psi_m^{(0)}|$ für $m \neq n$:

$$\begin{aligned}
\underbrace{\langle\psi_m^{(0)}|\hat{H}_0|\psi_n^{(k)}\rangle}_{= E_m^{(0)}\langle\psi_m^{(0)}|\psi_n^{(k)}\rangle} + \langle\psi_m^{(0)}|\hat{H}_1|\psi_n^{(k-1)}\rangle &= \sum_{l=0}^k E_n^{(l)} \langle\psi_m^{(0)}|\psi_n^{(k-l)}\rangle \\
&= \sum_{l=0}^{k-1} E_n^{(l)} \langle\psi_m^{(0)}|\psi_n^{(k-l)}\rangle \quad \left| \text{wegen } \langle\psi_m^{(0)}|\psi_n^{(0)}\rangle = 0 \right.
\end{aligned} \tag{11.62}$$

$$\Rightarrow (E_m^{(0)} - E_n^{(0)})\langle\psi_m^{(0)}|\psi_n^{(k)}\rangle = -\langle\psi_m^{(0)}|\hat{H}_1|\psi_n^{(k-1)}\rangle + \sum_{l=1}^{k-1} E_n^{(l)} \langle\psi_m^{(0)}|\psi_n^{(k-l)}\rangle, \tag{11.63}$$

wobei die l -Summe für $k-1 = 0$ wegzulassen ist. Da $E_n^{(0)}$ nicht entartet ist, gilt $E_m^{(0)} \neq E_n^{(0)}$ und wir können durch $E_m^{(0)} - E_n^{(0)}$ teilen. Es folgt

$$\langle\psi_m^{(0)}|\psi_n^{(k)}\rangle = + \frac{\langle\psi_m^{(0)}|\hat{H}_1|\psi_n^{(k-1)}\rangle}{E_n^{(0)} - E_m^{(0)}} - \sum_{l=1}^{k-1} E_n^{(l)} \frac{\langle\psi_m^{(0)}|\psi_n^{(k-l)}\rangle}{E_n^{(0)} - E_m^{(0)}}. \tag{11.64}$$

Wir multiplizieren mit $|\psi_m^{(0)}\rangle$, summieren über m und nutzen die Vollständigkeit von $\{|\psi_m^{(0)}\rangle\}$ aus:

$$\begin{aligned} |\psi_n^{(k)}\rangle &= \sum_m |\psi_m^{(0)}\rangle \langle \psi_m^{(0)} | \psi_n^{(k)} \rangle \stackrel{\text{Gl. (11.55)}}{=} \sum_{m \neq n} |\psi_m^{(0)}\rangle \langle \psi_m^{(0)} | \psi_n^{(k)} \rangle \\ &= \sum_{m \neq n} |\psi_m^{(0)}\rangle \frac{\langle \psi_m^{(0)} | \hat{H}_1 | \psi_n^{(k-1)} \rangle}{E_n^{(0)} - E_m^{(0)}} - \sum_{l=1}^{k-1} E_n^{(l)} \sum_{m \neq n} |\psi_m^{(0)}\rangle \frac{\langle \psi_m^{(0)} | \psi_n^{(k-l)} \rangle}{E_n^{(0)} - E_m^{(0)}}. \end{aligned} \quad (11.65)$$

Zusammen mit Gl. (11.60) haben wir gekoppelte Rekursionsgleichungen für die Energie- und Zustandskorrekturen gefunden. Damit können wir im Prinzip alle Ordnungen ausrechnen.

Störungstheorie erster Ordnung liefert für die Energie, wie schon erwähnt,

$$E_n^{(1)} = \langle \psi_n^{(0)} | \hat{H}_1 | \psi_n^{(0)} \rangle, \quad (11.66)$$

und für den Zustand

$$|\psi_n^{(1)}\rangle = \sum_{m \neq n} |\psi_m^{(0)}\rangle \frac{\langle \psi_m^{(0)} | \hat{H}_1 | \psi_n^{(0)} \rangle}{E_n^{(0)} - E_m^{(0)}}. \quad (11.67)$$

Damit erhält man für die zweite Ordnung

$$\begin{aligned} E_n^{(2)} &= \langle \psi_n^{(0)} | \hat{H}_1 | \psi_n^{(1)} \rangle \\ &= \sum_{m \neq n} \frac{\langle \psi_n^{(0)} | \hat{H}_1 | \psi_m^{(0)} \rangle \langle \psi_m^{(0)} | \hat{H}_1 | \psi_n^{(0)} \rangle}{E_n^{(0)} - E_m^{(0)}} \\ &= \sum_{m \neq n} \frac{|\langle \psi_n^{(0)} | \hat{H}_1 | \psi_m^{(0)} \rangle|^2}{E_n^{(0)} - E_m^{(0)}} \end{aligned} \quad (11.68)$$

und

$$\begin{aligned} |\psi_n^{(2)}\rangle &= \sum_{m \neq n} |\psi_m^{(0)}\rangle \frac{\langle \psi_m^{(0)} | \hat{H}_1 | \psi_n^{(1)} \rangle}{E_n^{(0)} - E_m^{(0)}} - E_n^{(1)} \sum_{m \neq n} |\psi_m^{(0)}\rangle \frac{\langle \psi_m^{(0)} | \psi_n^{(1)} \rangle}{E_n^{(0)} - E_m^{(0)}} \\ &= \sum_{m, q \neq n} |\psi_m^{(0)}\rangle \frac{\langle \psi_m^{(0)} | \hat{H}_1 | \psi_q^{(0)} \rangle \langle \psi_q^{(0)} | \hat{H}_1 | \psi_n^{(0)} \rangle}{(E_n^{(0)} - E_m^{(0)})(E_n^{(0)} - E_q^{(0)})} - \langle \psi_n^{(0)} | \hat{H}_1 | \psi_n^{(0)} \rangle \sum_{m \neq n} |\psi_m^{(0)}\rangle \frac{\langle \psi_m^{(0)} | \hat{H}_1 | \psi_n^{(0)} \rangle}{(E_n^{(0)} - E_m^{(0)})^2}. \end{aligned} \quad (11.69)$$

Ganz allgemein lassen sich alle Korrekturterme durch die ungestörten Zustände und Energien sowie Matrixelemente der Störung \hat{H}_1 bzgl. der ungestörten Zustände ausdrücken.

Beispiel: Wir betrachten ein eindimensionales Kastenpotential mit unendlich hohen Wänden und als Störung ein quadratisches Potential: In Ortsdarstellung ist

$$H_0 = -\frac{\hbar^2}{2m} \frac{d^2}{dx^2}, \quad (11.70)$$

$$H_1 = \frac{1}{2} m \omega^2 x^2, \quad (11.71)$$

$$x \in \left[-\frac{L}{2}, \frac{L}{2} \right]. \quad (11.72)$$

Die Randbedingungen lauten

$$\psi\left(-\frac{L}{2}\right) = \psi\left(\frac{L}{2}\right) = 0. \quad (11.73)$$

Die ungestörten Lösungen kennen wir aus 6.2.2:

$$\psi_n^{(0)}(x) = \begin{cases} \sqrt{\frac{2}{L}} \cos \frac{n\pi x}{L} & \text{für } n \text{ ungerade,} \\ \sqrt{\frac{2}{L}} \sin \frac{n\pi x}{L} & \text{für } n \text{ gerade,} \end{cases} \quad (11.74)$$

$$E_n^{(0)} = \frac{\pi^2 \hbar^2}{2mL^2} n^2, \quad n = 1, 2, 3, \dots \quad (11.75)$$

Die Energiekorrekturen erster Ordnung sind

$$\begin{aligned} E_n^{(1)} &= \int_{-L/2}^{L/2} dx \psi_n^{(0)}(x)^* H_1 \psi_n^{(0)}(x) \\ &= \frac{2}{L} \frac{1}{2} m\omega^2 \int_{-L/2}^{L/2} dx x^2 \times \begin{cases} \cos^2 \frac{n\pi x}{L} & \text{für } n \text{ ungerade,} \\ \sin^2 \frac{n\pi x}{L} & \text{für } n \text{ gerade} \end{cases} \\ &\stackrel{\text{Tafel}}{=} \frac{m\omega^2 L^2}{24\pi^2} \frac{n^2\pi^2 - 6}{n^2} \quad | \text{ für alle } n = 1, 2, \dots \\ &= \frac{m\omega^2 L^2}{24} \left(1 - \frac{6}{n^2\pi^2} \right). \end{aligned} \quad (11.76)$$

Also ist speziell für den Grundzustand (hier $n = 1$):

$$E_1^{(1)} = \frac{m\omega^2 L^2}{24} \left(1 - \frac{6}{\pi^2} \right) \quad (11.77)$$

$$\Rightarrow E_1^{(0)} + E_1^{(1)} = \frac{\pi^2 \hbar^2}{2mL^2} + \frac{m\omega^2 L^2}{24} \left(1 - \frac{6}{\pi^2} \right). \quad (11.78)$$

Wir betrachten die Energiekorrekturen noch etwas genauer. Die Näherung für die Energie bis zur ersten Ordnung ist

$$E_n \cong E_n^{(0)} + E_n^{(1)} = \langle \psi_n^{(0)} | (\hat{H}_0 + \hat{H}_1) | \psi_n^{(0)} \rangle = \langle \psi_n^{(0)} | \hat{H} | \psi_n^{(0)} \rangle. \quad (11.79)$$

Speziell für den Grundzustand $n = 0$ hatten wir in 11.1 gesehen, dass $\langle \hat{H} \rangle_\psi$ niemals kleiner als die exakte Grundzustandsenergie sein kann. Damit ist $E_0^{(0)} + E_0^{(1)}$ eine obere Schranke für die exakte Grundzustandsenergie. Die Korrektur zweiter Ordnung zur Grundzustandsenergie ist

$$E_0^{(2)} = \sum_{m>0} \frac{|\langle \psi_0^{(0)} | \hat{H}_1 | \psi_m^{(0)} \rangle|^2}{E_0^{(0)} - E_m^{(0)}} = - \sum_{m>0} \frac{|\langle \psi_0^{(0)} | \hat{H}_1 | \psi_m^{(0)} \rangle|^2}{E_m^{(0)} - E_0^{(0)}}. \quad (11.80)$$

Dieser Ausdruck ist nichtpositiv und ist echt *negativ*, wenn nicht alle Matrixelemente $\langle \psi_0^{(0)} | \hat{H}_1 | \psi_m^{(0)} \rangle$ für $m > 0$ verschwinden. Da die Näherung erster Ordnung eine obere Schranke liefert, besteht damit die Chance, die Näherung durch die zweite Ordnung zu verbessern. Allerdings stellen die Näherungen ab einschließlich der zweiten Ordnung keine oberen Schranken der Grundzustandsenergie dar.

Die hier betrachtete Störungstheorie wird auch als *Schrödingersche Störungstheorie* bezeichnet. Für sie sind die Differenzen der ungestörten Energien $E_n^{(0)}$ im Nenner charakteristisch. Man kann zeigen, dass eine ähnliche Störungsentwicklung, bei der für den betrachteten Zustand die *volle* Eigenenergie E_n im Nenner erscheint, ebenfalls gegen die exakte Lösung konvergiert. Zum Beispiel lautet dann die Korrektur erster Ordnung zum Zustand

$$|\psi_n^{(1)}\rangle = \sum_{m \neq n} |\psi_m^{(0)}\rangle \frac{\langle \psi_m^{(0)} | \hat{H}_1 | \psi_n^{(0)} \rangle}{E_n - E_m^{(0)}}. \quad (11.81)$$

Diese alternative Entwicklung nennt man *Brillouin-Wignersche Störungstheorie*. Sie kann oft Probleme mit Entartungen und daraus resultierenden verschwindenden Nennern vermeiden, erfordert für konkrete Rechnungen aber i. A. Iteration, da die vollen Energien E_n ja *a priori* nicht bekannt sind. Welcher Zugang günstiger ist, hängt vom konkreten Fall ab.

11.2.2 Störungstheorie für entartete Energieniveaus

Wir hatten im vorigen Abschnitt vorausgesetzt, dass die ungestörte Energie $E_n^{(0)}$ nicht entartet ist. Die hergeleiteten Rekursionsgleichungen sind in der Tat nicht anwendbar, wenn $E_n^{(0)}$ entartet ist, da sie Energienenner der Form $E_n^{(0)} - E_m^{(0)}$ für $m \neq n$ enthalten. Jetzt wollen wir Entartung von $E_n^{(0)}$ zulassen.

Der Entartungsgrad der Energie $E_n^{(0)}$ sei g_n . Wir nummerieren nun die Eigenzustände mit n und $\alpha = 1, \dots, g_n$. Es gilt also

$$\hat{H}_0 |\psi_{n\alpha}^{(0)}\rangle = E_n^{(0)} |\psi_{n\alpha}^{(0)}\rangle, \quad (11.82)$$

$$\hat{H} |\psi_{n\alpha}\rangle = E_{n\alpha} |\psi_{n\alpha}\rangle. \quad (11.83)$$

Die vollen Eigenenergien $E_{n\alpha}$ müssen natürlich nicht entartet sein, nur weil die ungestörten Energien $E_n^{(0)}$ es sind. Wir nehmen nun an, dass die $E_{n\alpha}$ tatsächlich nicht entartet sind. Dann sind die zugehörigen Eigenvektoren $|\psi_{n\alpha}\rangle$ bis auf Zahlenfaktoren eindeutig bestimmt. Andererseits ist die Wahl der (orthonormierten) Eigenvektoren $|\psi_{n\alpha}^{(0)}\rangle$ zu \hat{H}_0 im entarteten Unterraum beliebig; alle nicht verschwindenden Superpositionen

$$\sum_{\alpha=1}^{g_n} c_\alpha |\psi_{n\alpha}^{(0)}\rangle \quad (11.84)$$

sind ebenfalls Eigenvektoren. Daher ist irgendein $|\psi_{n\alpha}\rangle$ i. A. nicht nah an einem $|\psi_{n\alpha}^{(0)}\rangle$ aus einem gegebenen Orthonormalsystem im entarteten Unterraum. Wir müssen also herausfinden, welches Orthonormalsystem in dieser Hinsicht das „richtige“ ist.

Dazu fordern wir, dass die richtigen Eigenvektoren $|\tilde{\psi}_{n\alpha}^{(0)}\rangle$ so gewählt sind, dass weiterhin Gl. (11.58) zur ersten Ordnung ($k = 1$) gilt:

$$\hat{H}_0 |\psi_{n\alpha}^{(1)}\rangle + \hat{H}_1 |\tilde{\psi}_{n\alpha}^{(0)}\rangle = E_n^{(0)} |\psi_{n\alpha}^{(1)}\rangle + E_{n\alpha}^{(1)} |\tilde{\psi}_{n\alpha}^{(0)}\rangle. \quad (11.85)$$

Das ist der Beitrag erster Ordnung in λ (und \hat{H}_1) zur Schrödinger-Gleichung. Wir schreiben $|\tilde{\psi}_{n\alpha}^{(0)}\rangle$ als Linearkombination der Vektoren aus einer beliebigen Orthonormalbasis $\{|\psi_{n\beta}^{(0)}\rangle\}$ des entarteten Unterraums:

$$|\tilde{\psi}_{n\alpha}^{(0)}\rangle = \sum_{\beta} u_{\alpha\beta} |\psi_{n\beta}^{(0)}\rangle. \quad (11.86)$$

Multiplikation der Gleichung (11.85) von links mit $\langle \psi_{n\gamma}^{(0)} |$ ergibt

$$\cancel{E_n^{(0)} \langle \psi_{n\gamma}^{(0)} | \psi_{n\alpha}^{(1)} \rangle} + \sum_{\beta} u_{\alpha\beta} \langle \psi_{n\gamma}^{(0)} | \hat{H}_1 | \psi_{n\beta}^{(0)} \rangle = \cancel{E_n^{(0)} \langle \psi_{n\gamma}^{(0)} | \psi_{n\alpha}^{(1)} \rangle} + E_{n\alpha}^{(1)} \sum_{\beta} u_{\alpha\beta} \underbrace{\langle \psi_{n\gamma}^{(0)} | \psi_{n\beta}^{(0)} \rangle}_{=\delta_{\gamma\beta}} \quad (11.87)$$

$$\Rightarrow \sum_{\beta} \langle \psi_{n\gamma}^{(0)} | \hat{H}_1 | \psi_{n\beta}^{(0)} \rangle u_{\alpha\beta} = E_{n\alpha}^{(1)} u_{\alpha\gamma}. \quad (11.88)$$

Sei H_{1n} die $g_n \times g_n$ -Matrix, genannt *Störmatrix*, mit den Komponenten

$$H_{1n,\alpha\beta} = \langle \psi_{n\alpha}^{(0)} | \hat{H}_1 | \psi_{n\beta}^{(0)} \rangle, \quad (11.89)$$

also die Matrixdarstellung des Störoperators auf dem entarteten Unterraum bzgl. der Basis $\{|\psi_{n\alpha}^{(0)}\rangle\}$. Sei außerdem \vec{u}_α der Spaltenvektor mit den Komponenten $u_{\alpha\beta}$. Dann lässt sich Gl. (11.89) schreiben als

$$H_{1n} \vec{u}_\alpha = E_{n\alpha}^{(1)} \vec{u}_\alpha. \quad (11.90)$$

Die Koeffizienten $u_{\alpha\beta}$ in $\{|\tilde{\psi}_{n\alpha}^{(0)}\rangle\}$ ergeben sich also als Komponenten des Eigenvektor \vec{u}_α von H_{1n} zum Eigenwert $E_{n\alpha}^{(1)}$. Die Lösung des Eigenwertproblems für H_{1n} liefert zugleich die richtige Basis des entarteten Unterraums und die Korrekturen erster Ordnung zur Energie.

Falls die Eigenwerte der Störmatrix H_{1n} nicht entartet sind, so spaltet die Energie E_n in erster Ordnung in g_n Niveaus auf. In der neuen Basis $\{|\tilde{\psi}_{n\alpha}^{(0)}\rangle\}$ gilt

$$\langle \tilde{\psi}_{n\alpha}^{(0)} | \hat{H}_1 | \tilde{\psi}_{n\beta}^{(0)} \rangle = E_{n\alpha}^{(1)} \delta_{\alpha\beta}, \quad (11.91)$$

insbesondere verschwinden alle Matrixelemente für $\alpha \neq \beta$. Die Korrekturen höherer Ordnung lassen sich dann herleiten. Die Herleitung ist nur teilweise analog zum nichtentarteten Fall. Wir betrachten hier nur die Zustandskorrekturen erster Ordnung und die Energiekorrekturen zweiter Ordnung. Wir lassen die Tilde weg und schreiben die neue Basis als $\{|\psi_{n\alpha}^{(0)}\rangle\}$.

Aus der Forderung $\langle \psi_{n\alpha}^{(0)} | \psi_{n\alpha} \rangle = 1$ folgt wie oben

$$\langle \psi_{n\alpha}^{(0)} | \psi_{n\alpha}^{(j)} \rangle = \delta_{0j}. \quad (11.92)$$

Aus der Entwicklung nach λ erhalten wir, für $k \geq 1$,

$$\hat{H}_0 |\psi_{n\alpha}^{(k)}\rangle + \hat{H}_1 |\psi_{n\alpha}^{(k-1)}\rangle = \sum_{l=0}^k E_{n\alpha}^{(l)} |\psi_{n\alpha}^{(k-l)}\rangle. \quad (11.93)$$

Multiplikation von links mit $\langle \psi_{n\alpha}^{(0)} |$ ergibt

$$E_{n\alpha}^{(k)} = \langle \psi_{n\alpha}^{(0)} | \hat{H}_1 | \psi_{n\alpha}^{(k-1)} \rangle. \quad (11.94)$$

Für $k = 1$ hatten wir dies schon vorher gefunden: Es ist $E_{n\alpha}^{(1)} = \langle \psi_{n\alpha}^{(0)} | \hat{H}_1 | \psi_{n\alpha}^{(0)} \rangle$ in der neuen Basis.

Da die ungestörten Eigenzustände ein vollständiges Orthonormalsystem bilden, können wir schreiben

$$|\psi_{n\alpha}^{(1)}\rangle = \sum_m \sum_{\beta=1}^{g_m} |\psi_{m\beta}^{(0)}\rangle \langle \psi_{m\beta}^{(0)} | \psi_{n\alpha}^{(1)} \rangle = \sum_{m \neq n} \sum_{\beta=1}^{g_m} |\psi_{m\beta}^{(0)}\rangle \langle \psi_{m\beta}^{(0)} | \psi_{n\alpha}^{(1)} \rangle + \sum_{\substack{\beta=1 \\ \beta \neq \alpha}}^{g_n} |\psi_{n\beta}^{(0)}\rangle \langle \psi_{n\beta}^{(0)} | \psi_{n\alpha}^{(1)} \rangle, \quad (11.95)$$

wobei in der letzten Summe der Term mit $\beta = \alpha$ wegen Gl. (11.92) nicht auftritt. Die erste Summe (über $m \neq n$) erhalten wir wie folgt. Wir multiplizieren zunächst Gl. (11.93) von links mit $\langle \psi_{m\beta}^{(0)} |$, $m \neq n$, für $k = 1$:

$$E_m^{(0)} \langle \psi_{m\beta}^{(0)} | \psi_{n\alpha}^{(1)} \rangle + \langle \psi_{m\beta}^{(0)} | \hat{H}_1 | \psi_{n\alpha}^{(0)} \rangle = E_n^{(0)} \langle \psi_{m\beta}^{(0)} | \psi_{n\alpha}^{(1)} \rangle + E_{n\alpha}^{(1)} \underbrace{\langle \psi_{m\beta}^{(0)} | \psi_{n\alpha}^{(0)} \rangle}_{=0} \quad (11.96)$$

$$\Rightarrow (E_m^{(0)} - E_n^{(0)}) \langle \psi_{m\beta}^{(0)} | \psi_{n\alpha}^{(1)} \rangle = -\langle \psi_{m\beta}^{(0)} | \hat{H}_1 | \psi_{n\alpha}^{(0)} \rangle. \quad (11.97)$$

Wegen $m \neq n$ ist $E_m^{(0)} - E_n^{(0)} \neq 0$ und wir erhalten

$$\langle \psi_{m\beta}^{(0)} | \psi_{n\alpha}^{(1)} \rangle = + \frac{\langle \psi_{m\beta}^{(0)} | \hat{H}_1 | \psi_{n\alpha}^{(0)} \rangle}{E_n^{(0)} - E_m^{(0)}}. \quad (11.98)$$

Es folgt

$$\sum_{m,\beta} |\psi_{m\beta}^{(0)}\rangle \langle \psi_{m\beta}^{(0)} | \psi_{n\alpha}^{(1)} \rangle = \sum_{m,\beta} |\psi_{m\beta}^{(0)}\rangle \frac{\langle \psi_{m\beta}^{(0)} | \hat{H}_1 | \psi_{n\alpha}^{(0)} \rangle}{E_n^{(0)} - E_m^{(0)}}. \quad (11.99)$$

Für die zweite Summe multiplizieren wir Gl. (11.93) für $k = 2$ von links mit $\langle \psi_{n\beta}^{(0)} |$:

$$\cancel{E_n^{(0)} \langle \psi_{n\beta}^{(0)} | \psi_{n\alpha}^{(2)} \rangle} + \langle \psi_{n\beta}^{(0)} | \hat{H}_1 | \psi_{n\alpha}^{(1)} \rangle = \cancel{E_n^{(0)} \langle \psi_{n\beta}^{(0)} | \psi_{n\alpha}^{(2)} \rangle} + E_{n\alpha}^{(1)} \langle \psi_{n\beta}^{(0)} | \psi_{n\alpha}^{(1)} \rangle + E_{n\alpha}^{(2)} \underbrace{\langle \psi_{n\beta}^{(0)} | \psi_{n\alpha}^{(0)} \rangle}_{= \delta_{\beta\alpha}}. \quad (11.100)$$

Für $|\psi_{n\alpha}^{(1)}\rangle$ setzen wir Gl. (11.95) ein:

$$\sum_{\substack{m,\gamma \\ m \neq n}} \langle \psi_{n\beta}^{(0)} | \hat{H}_1 | \psi_{m\gamma}^{(0)} \rangle \langle \psi_{m\gamma}^{(0)} | \psi_{n\alpha}^{(1)} \rangle + \sum_{\gamma \neq \alpha} \underbrace{\langle \psi_{n\beta}^{(0)} | \hat{H}_1 | \psi_{n\gamma}^{(0)} \rangle \langle \psi_{n\gamma}^{(0)} | \psi_{n\alpha}^{(1)} \rangle}_{= E_{n\beta}^{(1)} \delta_{\beta\gamma}} = E_{n\alpha}^{(1)} \sum_{\gamma \neq \alpha} \underbrace{\langle \psi_{n\beta}^{(0)} | \psi_{n\gamma}^{(0)} \rangle \langle \psi_{n\gamma}^{(0)} | \psi_{n\alpha}^{(1)} \rangle}_{= \delta_{\beta\gamma}} + E_{n\alpha}^{(2)} \delta_{\beta\alpha} \quad (11.101)$$

$$\Rightarrow (E_{n\beta}^{(1)} - E_{n\alpha}^{(1)}) \underbrace{(1 - \delta_{\alpha\beta})}_{\text{redundant}} \langle \psi_{n\beta}^{(0)} | \psi_{n\alpha}^{(1)} \rangle = E_{n\alpha}^{(1)} \delta_{\alpha\beta} - \sum_{\substack{m,\gamma \\ m \neq n}} \langle \psi_{n\beta}^{(0)} | \hat{H}_1 | \psi_{m\gamma}^{(0)} \rangle \langle \psi_{m\gamma}^{(0)} | \psi_{n\alpha}^{(1)} \rangle. \quad (11.102)$$

Für $\alpha = \beta$ verschwindet die linke Seite und wir erhalten die Energiekorrekturen zweiter Ordnung:

$$\begin{aligned} E_{n\alpha}^{(2)} &= \sum_{\substack{m,\beta \\ m \neq n}} \langle \psi_{n\alpha}^{(0)} | \hat{H}_1 | \psi_{m\beta}^{(0)} \rangle \langle \psi_{m\beta}^{(0)} | \psi_{n\alpha}^{(1)} \rangle \\ &\stackrel{\text{Gl. (11.98)}}{=} \sum_{\substack{m,\beta \\ m \neq n}} \frac{|\langle \psi_{n\alpha}^{(0)} | \hat{H}_1 | \psi_{m\beta}^{(0)} \rangle|^2}{E_n^{(0)} - E_m^{(0)}}. \end{aligned} \quad (11.103)$$

Die Summe läuft über alle Zustände, die mit dem betrachteten in nullter Ordnung nicht entartet sind. Andererseits, für $\alpha \neq \beta$ verschwindet der Term mit $E_{n\alpha}^{(2)}$ und wir erhalten

$$(E_{n\beta}^{(1)} - E_{n\alpha}^{(1)}) \langle \psi_{n\beta}^{(0)} | \psi_{n\alpha}^{(1)} \rangle = - \sum_{\substack{m,\gamma \\ m \neq n}} \langle \psi_{n\beta}^{(0)} | \hat{H}_1 | \psi_{m\gamma}^{(0)} \rangle \langle \psi_{m\gamma}^{(0)} | \psi_{n\alpha}^{(1)} \rangle. \quad (11.104)$$

$$\begin{aligned} \Rightarrow \langle \psi_{n\beta}^{(0)} | \psi_{n\alpha}^{(1)} \rangle &= + \sum_{\substack{m,\gamma \\ m \neq n}} \frac{\langle \psi_{n\beta}^{(0)} | \hat{H}_1 | \psi_{m\gamma}^{(0)} \rangle \langle \psi_{m\gamma}^{(0)} | \psi_{n\alpha}^{(1)} \rangle}{E_{n\alpha}^{(1)} - E_{n\beta}^{(1)}} \\ &\stackrel{\text{Gl. (11.98)}}{=} \sum_{\substack{m,\gamma \\ m \neq n}} \frac{\langle \psi_{n\beta}^{(0)} | \hat{H}_1 | \psi_{m\gamma}^{(0)} \rangle \langle \psi_{m\gamma}^{(0)} | \hat{H}_1 | \psi_{n\alpha}^{(1)} \rangle}{(E_{n\alpha}^{(1)} - E_{n\beta}^{(1)}) (E_n^{(0)} - E_m^{(0)})}. \end{aligned} \quad (11.105)$$

Wir haben ausgenutzt, dass \hat{H}_1 die Entartung *vollständig* aufhebt. Wenn wir alle Ergebnisse einsetzen, erhalten wir für die Zustandskorrektur nun aus Gl. (11.95)

$$|\psi_{n\alpha}^{(1)}\rangle = \sum_{\substack{m,\beta \\ m \neq n}} |\psi_{m\beta}^{(0)}\rangle \frac{\langle \psi_{m\beta}^{(0)} | \hat{H}_1 | \psi_{n\alpha}^{(0)} \rangle}{E_n^{(0)} - E_m^{(0)}} + \sum_{\beta \neq \alpha} \sum_{\substack{m,\gamma \\ m \neq n}} |\psi_{n\beta}^{(0)}\rangle \frac{\langle \psi_{n\beta}^{(0)} | \hat{H}_1 | \psi_{m\gamma}^{(0)} \rangle \langle \psi_{m\gamma}^{(0)} | \hat{H}_1 | \psi_{n\alpha}^{(1)} \rangle}{(E_{n\alpha}^{(1)} - E_{n\beta}^{(1)}) (E_n^{(0)} - E_m^{(0)})}. \quad (11.106)$$

Falls die Entartung in erster Ordnung *nicht* vollständig aufgehoben wird, geht die Rechnung zwar bis zur Bestimmung der Energien zur zweiten Ordnung durch, aber das Ergebnis für die Zustandskorrektur zur ersten Ordnung stimmt nicht mehr. Hier tritt ja eine Division durch $E_{n\alpha}^{(1)} - E_{n\beta}^{(1)}$ auf, was nun für gewisse $\alpha \neq \beta$ Null ist. In diesem Fall ist die Betrachtung der Störung zweiter oder noch höherer Ordnung notwendig, um zu entscheiden, welches die „richtige“ Basis im entarteten Unterraum ist. Die Korrekturen müssen mühsam hergeleitet werden. Es gibt aber eine elegante Formulierung der Störungstheorie, die i. W. die *Resolvente* $\hat{G} := (E - \hat{H})^{-1}$ der Schrödinger-Gleichung für kleine Störungen \hat{H}_1 entwickelt. Für diese Formulierung sei auf Lehrbücher verwiesen.

Was sind die wesentlichen Merkmale der Störungstheorie?

- Sie ist kontrolliert in dem Sinne, dass man systematisch zu höheren Ordnungen gehen kann. Der Aufwand vergrößert sich allerdings beträchtlich in höheren Ordnungen.
- Eine Abschätzung des Fehlers ist aber nicht streng möglich.
- Die Störungsentwicklung versagt, wenn die Störung nicht klein ist. Es gibt aber Fälle, in denen die Potenzreihen in λ nicht konvergieren, die ersten paar Terme aber trotzdem gute Näherungen darstellen.
- Es bringt oft keinen Vorteil, zu höheren Ordnungen zu gehen: Ist die Störung klein, so reicht die führende Ordnung oder einige wenige niedrige Ordnungen. Ist sie dagegen groß, versagt die Störungstheorie. Die einzige Situation, in der man hohe Ordnungen explizit berechnet, ist, wenn man an sehr präzisen Ergebnissen für einen Fall mit kleiner Störung interessiert ist, z. B. bei der Berechnung des magnetischen Moments des Elektrons.

11.3 Die quasiklassische Näherung

Wir hatten schon diskutiert, dass sich die klassische Mechanik im geeignet definierten Limes $\hbar \rightarrow 0$ aus der Quantenmechanik ergeben muss. Dies legt als weitere Näherungsmethode eine Entwicklung in \hbar als kleinem Parameter nahe. Die nullte Ordnung entspricht dann dem klassischen Grenzfall und die Terme ab erster Ordnung beschreiben die quantenmechanischen Korrekturen. Konkret hatten wir in 4.3.3 gesehen, dass sich die klassische Mechanik in der Hamilton-Jacobi-Formulierung als klassischer Grenzfall der Schrödingerschen Wellenmechanik ergibt (in 4.3.3 sind wir den umgekehrten Weg gegangen, um die Schrödinger-Gleichung induktiv zu begründen). Wir werden zunächst die dort angestellten Überlegungen zu einer mathematischen Theorie der Entwicklung nach \hbar ausbauen.

11.3.1 Entwicklung der Wirkung nach \hbar

Wie in 4.3.3 setzen wir die zeitabhängige Wellenfunktion eines Teilchens als

$$\psi(\vec{r}, t) = \psi_0 \exp\left(\frac{i}{\hbar} S(\vec{r}, t)\right) \quad (11.107)$$

an. (Die Theorie kann ohne Mühe für allgemeine Koordinaten $\vec{q} = (q_1, q_2, \dots)$ formuliert werden. Wir schreiben sie hier aber für den gewohnten Fall eines Teilchens in drei Raumdimensionen.) $S(\vec{r}, t)$ ist die i. A. komplexe Wirkungsfunktion. Einsetzen in die Schrödinger-Gleichung

$$i\hbar \frac{\partial \psi}{\partial t} = -\frac{\hbar^2}{2m} \nabla^2 \psi + V(\vec{r}, t) \psi, \quad (11.108)$$

in der das Potential völlig beliebig ist und sogar von der Zeit abhängen darf, ergibt

$$-\frac{\partial S}{\partial t} e^{iS/\hbar} = \frac{1}{2m} (\vec{\nabla} S)^2 e^{iS/\hbar} - \frac{i\hbar}{2m} (\nabla^2 S) e^{iS/\hbar} + V e^{iS/\hbar} \quad (11.109)$$

$$\Rightarrow -\frac{\partial S}{\partial t} = \frac{1}{2m} (\vec{\nabla} S)^2 - \frac{i\hbar}{2m} \nabla^2 S + V. \quad (11.110)$$

Das ist offensichtlich eine nichtlineare partielle Differentialgleichung für $S(\vec{r}, t)$. Nun entwickeln wir S in Potenzen von \hbar . Da \hbar in der Schrödinger-Gleichung immer zusammen mit einem Faktor i erscheint, schreiben wir

$$S(\vec{r}, t) = \sum_{n=0}^{\infty} (i\hbar)^n S_n(\vec{r}, t). \quad (11.111)$$

Einsetzen ergibt

$$\begin{aligned} -\left(\frac{\partial S_0}{\partial t} + i\hbar \frac{\partial S_1}{\partial t} - \hbar^2 \frac{\partial S_2}{\partial t} + \dots\right) &= \frac{1}{2m} \left[(\vec{\nabla} S_0)^2 - \hbar^2 (\vec{\nabla} S_1)^2 + \hbar^4 (\vec{\nabla} S_2)^2 + \dots \right. \\ &\quad \left. + 2i\hbar \vec{\nabla} S_0 \cdot \vec{\nabla} S_1 - 2\hbar^2 \vec{\nabla} S_0 \cdot \vec{\nabla} S_2 - 2i\hbar^3 \vec{\nabla} S_1 \cdot \vec{\nabla} S_2 + \dots \right] \\ &\quad - \frac{i\hbar}{2m} (\nabla^2 S_0 + i\hbar \nabla^2 S_1 - \hbar^2 \nabla^2 S_2 + \dots) + V. \end{aligned} \quad (11.112)$$

Die Reihenentwicklungen auf beiden Seiten der Gleichung sollen dieselbe Funktion von \hbar ergeben. Wegen der linearen Unabhängigkeit von \hbar^n für verschiedene n folgt die Gleichheit der Koeffizienten. In den Ordnungen \hbar^0 , \hbar^1 und \hbar^2 erhalten wir

$$-\frac{\partial S_0}{\partial t} = \frac{1}{2m} (\vec{\nabla} S_0)^2 + V, \quad (11.113)$$

$$-\frac{\partial S_1}{\partial t} = \frac{1}{m} \vec{\nabla} S_0 \cdot \vec{\nabla} S_1 - \frac{1}{2m} \nabla^2 S_0, \quad (11.114)$$

$$-\frac{\partial S_2}{\partial t} = \frac{1}{2m} (\vec{\nabla} S_1)^2 + \frac{1}{m} \vec{\nabla} S_0 \cdot \vec{\nabla} S_2 - \frac{1}{2m} \nabla^2 S_1. \quad (11.115)$$

Beachte, dass alle Koeffizienten in den Gleichungen reell sind. Wir könnten daher reelle Lösungen $S_n(\vec{r}, t)$ erwarten. In der Tat sind im *klassisch erlaubten Bereich* $[E \geq V(\vec{r})]$ alle S_n reell. Dann tragen die geraden Ordnungen zur Phase und die ungeraden zum Betrag von $\psi(\vec{r}, t)$ bei. Die Gleichung nullter Ordnung lässt sich umschreiben zu

$$H(\vec{r}, \vec{\nabla} S_0, t) = -\frac{\partial S_0}{\partial t} \quad (11.116)$$

mit

$$H(\vec{r}, \vec{p}, t) = \frac{p^2}{2m} + V(\vec{r}, t). \quad (11.117)$$

Gleichung (11.116) ist die Hamilton-Jacobi-Gleichung für die klassische Hamilton-Funktion H . Der Limes $\hbar \rightarrow 0$ S_0 beschreibt also wirklich den klassischen Grenzfall des Quantensystems.

Die Korrektur erster Ordnung hat auch einen klaren Ursprung, sie folgt nämlich aus der Kontinuitätsgleichung, wie wir nun für den klassisch erlaubten Bereich zeigen. Die Orts- und Zeitabhängigkeit von S_0 erzwingt einen orts- und zeitabhängigen Betrag von ψ . Die Wahrscheinlichkeitsdichte ist, in führender Ordnung in \hbar ,

$$\begin{aligned} \rho &= \psi^* \psi \\ &= |\psi_0|^2 \exp\left(-\frac{i}{\hbar}(S_0 - i\hbar S_1 - \hbar^2 S_2 + i\hbar^3 S_3 + \dots)\right) \exp\left(\frac{i}{\hbar}(S_0 + i\hbar S_1 - \hbar^2 S_2 - i\hbar^3 S_3 + \dots)\right) \\ &= |\psi_0|^2 \exp(-2S_1 + 2\hbar^2 S_3 + \dots) \\ &\cong |\psi_0|^2 e^{-2S_1}. \end{aligned} \quad (11.118)$$

Es folgt

$$\frac{\partial \rho}{\partial t} \cong -2 \frac{\partial S_1}{\partial t} \rho. \quad (11.119)$$

Die Wahrscheinlichkeitsstromdichte ist, zur gleichen Ordnung,

$$\begin{aligned} \vec{j} &= \frac{\hbar}{2mi} [\psi^* \vec{\nabla} \psi - (\vec{\nabla} \psi^*) \psi] \\ &= \frac{\hbar}{2mi} |\psi_0|^2 \left[\exp\left(-\frac{i}{\hbar}(S_0 - i\hbar S_1 - \hbar^2 S_2 + \dots)\right) \vec{\nabla} \exp\left(\frac{i}{\hbar}(S_0 + i\hbar S_1 - \hbar^2 S_2 + \dots)\right) \right. \\ &\quad \left. - \left(\vec{\nabla} \exp\left(-\frac{i}{\hbar}(S_0 - i\hbar S_1 - \hbar^2 S_2 + \dots)\right) \right) \exp\left(\frac{i}{\hbar}(S_0 + i\hbar S_1 - \hbar^2 S_2 + \dots)\right) \right] \\ &= \frac{1}{2m} |\psi_0|^2 \left[\exp\left(-\frac{i}{\hbar}(S_0 - i\hbar S_1 - \hbar^2 S_2 + \dots)\right) \right. \\ &\quad \times (\vec{\nabla} S_0 + i\hbar \vec{\nabla} S_1 - \hbar^2 \vec{\nabla} S_2 + \dots) \exp\left(\frac{i}{\hbar}(S_0 + i\hbar S_1 - \hbar^2 S_2 + \dots)\right) \\ &\quad \left. + (\vec{\nabla} S_0 - i\hbar \vec{\nabla} S_1 - \hbar^2 \vec{\nabla} S_2 + \dots) \right. \\ &\quad \left. \times \exp\left(-\frac{i}{\hbar}(S_0 - i\hbar S_1 - \hbar^2 S_2 + \dots)\right) \exp\left(\frac{i}{\hbar}(S_0 + i\hbar S_1 - \hbar^2 S_2 + \dots)\right) \right] \\ &= \frac{1}{m} |\psi_0|^2 (\vec{\nabla} S_0 - \hbar^2 \vec{\nabla} S_2 + \dots) \exp(-2S_1 + \dots) \\ &\cong \frac{|\psi_0|^2}{m} (\vec{\nabla} S_0) e^{-2S_1}. \end{aligned} \quad (11.120)$$

Daraus folgt

$$\begin{aligned} \vec{\nabla} \cdot \vec{j} &\cong \frac{|\psi_0|^2}{m} (\nabla^2 S_0) e^{-2S_1} - 2 \frac{|\psi_0|^2}{m} (\vec{\nabla} S_0 \cdot \vec{\nabla} S_1) e^{-2S_1} \\ &\cong \frac{1}{m} (\nabla^2 S_0) \rho - \frac{2}{m} (\vec{\nabla} S_0 \cdot \vec{\nabla} S_1) \rho \end{aligned} \quad (11.121)$$

und schließlich aus der Kontinuitätsgleichung, zur Ordnung \hbar^0 ,

$$0 = \frac{\partial \rho}{\partial t} + \vec{\nabla} \cdot \vec{j} \cong -2 \frac{\partial S_1}{\partial t} \rho + \frac{1}{m} (\vec{\nabla}^2 S_0) \rho - \frac{2}{m} (\vec{\nabla} S_0 \cdot \vec{\nabla} S_1) \rho \quad (11.122)$$

$$\Rightarrow \frac{\partial S_1}{\partial t} - \frac{1}{2m} \nabla^2 S_0 + \frac{1}{m} \vec{\nabla} S_0 \cdot \vec{\nabla} S_1 = 0. \quad (11.123)$$

Das ist gerade die Gleichung für die Korrektur erster Ordnung.

Die Gleichungen für die Funktionen $S_n(\vec{r}, t)$ sind offensichtlich rekursiv. Im Prinzip können wir nun diese Funktionen iterativ bestimmen. In höheren Ordnungen wird dies praktisch schwierig sein, da (a) die Gleichungen immer komplizierter werden und (b) die daher i. A. notwendige numerische Lösung zur Akkumulation von Fehlern mit wachsender Ordnung führt. Daher ist in der Praxis v. a. die erste Ordnung relevant.

11.3.2 Die WKB-Näherung

Die nach Wentzel, Kramers und Brillouin (1926) benannte *WKB-Näherung* besteht darin, die Entwicklung nach \hbar nach der ersten Ordnung abzuberechnen. Sie bezieht sich auf zeitunabhängige, effektiv eindimensionale Systeme. Effektiv eindimensionale Probleme können sich natürlich durch Separationsansätze aus höherdimensionalen ergeben, z. B. die Radialgleichung für Zentralpotentiale. Für zeitunabhängige Hamiltonians kann die Zeitentwicklung absepariert werden, was bekanntlich auf die zeitunabhängige Schrödinger-Gleichung führt. Wir betrachten hier konkret die Gleichung

$$-\frac{\hbar^2}{2m} \frac{d^2 \psi}{dx^2} + V(x) \psi(x) = E \psi(x). \quad (11.124)$$

Wir schreiben diese in der Form

$$\psi''(x) + k^2(x) \psi(x) = 0, \quad (11.125)$$

mit

$$k^2(x) := \frac{2m}{\hbar^2} [E - V(x)]. \quad (11.126)$$

Die *zeitabhängige* Wellenfunktion schreiben wir wie im vorigen Abschnitt als

$$\psi(x, t) = \psi_0 e^{iS(x, t)/\hbar}, \quad (11.127)$$

wobei für zeitunabhängige Systeme gilt

$$S(x, t) = W(x) - Et, \quad (11.128)$$

vgl. Abschnitt 4.3.2 und die entsprechende Beziehung in der Hamilton-Jacobi-Theorie. Damit wird

$$\psi(x, t) = \psi_0 e^{iW(x)/\hbar} e^{-Et/\hbar} \equiv \psi(x) e^{-Et/\hbar}, \quad (11.129)$$

wobei $\psi(x)$ die zeitunabhängige Wellenfunktion (Eigenfunktion zur Energie E) ist. Einsetzen in Gl. (11.125) ergibt

$$-\frac{\psi_0}{\hbar^2} (W'(x))^2 e^{iW(x)/\hbar} + \frac{i\psi_0}{\hbar} W''(x) e^{iW(x)/\hbar} + k^2(x) \psi_0 e^{iW(x)/\hbar} = 0 \quad (11.130)$$

$$\Rightarrow (W'(x))^2 - i\hbar W''(x) = \hbar^2 k^2(x) = 2m [E - V(x)]. \quad (11.131)$$

Jetzt entwickeln wir W nach $i\hbar$:

$$W(x) = \sum_{n=0}^{\infty} (i\hbar)^n W_n(x). \quad (11.132)$$

Einsetzen ergibt

$$(W'_0(x) + i\hbar W'_1(x) - \hbar^2 W'_2(x) + \dots)^2 - i\hbar (W''_0(x) + i\hbar W''_1(x) - \hbar^2 W''_2(x) + \dots) = \hbar^2 k^2(x). \quad (11.133)$$

Die niedrigsten Ordnungen in \hbar können wir direkt ablesen, wobei wir beachten müssen, dass $\hbar^2 k^2 = 2m(E - V)$ ein Term *nullter* und nicht zweiter Ordnung ist:

$$(W'_0(x))^2 = \hbar^2 k^2(x), \quad (11.134)$$

$$2W_0'(x)W_1'(x) - W_0''(x) = 0, \quad (11.135)$$

$$(W_1'(x))^2 + 2W_0'(x)W_2'(x) - W_1''(x) = 0, \quad (11.136)$$

...

Die Gleichung nullter Ordnung ist gerade die Eikonal-Gleichung der Klassischen Mechanik, siehe Abschnitt 4.3.2. Sie lässt sich explizit lösen: Zunächst folgt

$$W_0'(x) = \pm \hbar k(x) \quad (11.137)$$

$$\begin{aligned} \Rightarrow W_0(x) &= \pm \hbar \int_{x_{\pm}}^x dx' k(x') \\ &= \pm \hbar \int_{x_{\pm}}^x dx' \sqrt{2m[E - V(x')]}, \end{aligned} \quad (11.138)$$

wobei x_{\pm} unbestimmt ist; es ist der erwartete freie Parameter in der Lösung einer Differentialgleichung erster Ordnung. Soweit ist die Rechnung rein klassisch.

Die Gleichung erster Ordnung kann leicht für $W_1(x)$ gelöst werden:

$$W_1'(x) = \frac{1}{2} \frac{W_0''(x)}{W_0'(x)} = \frac{1}{2} \frac{k'(x)}{k(x)} = \frac{1}{2} \frac{d}{dx} \ln k(x) = \frac{d}{dx} \ln \sqrt{k(x)} \quad (11.139)$$

$$\Rightarrow W_1(x) = \ln \sqrt{k(x)} + C \quad (11.140)$$

mit einer Integrationskonstanten C . In der WKB-Näherung werden die höheren Ordnungen $W_2(x)$, $W_3(x)$, ... vernachlässigt. Für die Wellenfunktion erhalten wir nun

$$\begin{aligned} \psi(x) &\cong \psi_0 \exp\left(\frac{i}{\hbar}(W_0 + i\hbar W_1)\right) = \psi_0 \exp\left(\pm i \int_{x_{\pm}}^x dx' k(x')\right) \exp\left(-\ln \sqrt{k(x)} - C\right) \\ &= \psi_0 e^{-C} \frac{1}{\sqrt{k(x)}} \exp\left(\pm i \int_{x_{\pm}}^x dx' k(x')\right). \end{aligned} \quad (11.141)$$

Der Faktor e^{-C} kann offensichtlich in den Normierungsfaktor ψ_0 absorbiert werden. Die (genäherte) allgemeine Lösung der Schrödinger-Gleichung ergibt sich als Superposition der beiden Lösungen mit Vorzeichen \pm :

$$\psi(x) \cong \frac{\psi_+}{\sqrt{k(x)}} \exp\left(i \int_{x_+}^x dx' k(x')\right) + \frac{\psi_-}{\sqrt{k(x)}} \exp\left(-i \int_{x_-}^x dx' k(x')\right), \quad (11.142)$$

wobei ψ_{\pm} und x_{\pm} nicht unabhängig sind; es gibt nur zwei unabhängige Parameter, da die Schrödinger-Gleichung zweiter Ordnung ist. Beachte, dass der Grenzfall eines konstanten Potentials exakt beschrieben wird: In diesem Fall ist $\hbar k = \sqrt{2m(E - V)}$ der konstante Impuls, die Integrale ergeben $k(x - x_{\pm})$ und die Lösung $\psi(x)$ beschreibt nach rechts und links laufende ebene Wellen. Es liegt nahe zu vermuten, dass die Näherung insbesondere dann gut ist, wenn sich $V(x)$ schwach räumlich ändert.

Nichts in der Herleitung verbietet uns die Auswertung von $\psi(x)$ im klassisch verbotenen Bereich, wo $E < V(x)$ gilt. Hier wird $k(x)$ imaginär. Wir schreiben

$$k(x) = i\kappa(x) \quad (11.143)$$

mit

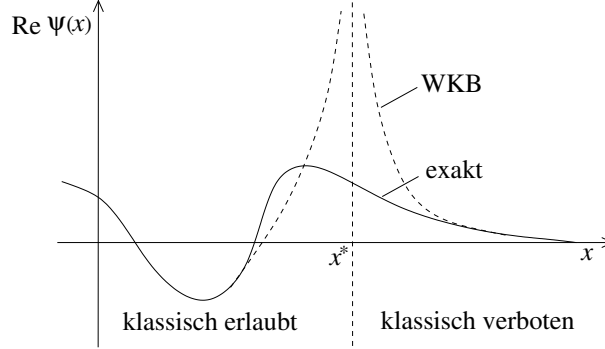
$$\kappa(x) = \frac{1}{\hbar} \sqrt{2m[V(x) - E]} \in \mathbb{R}. \quad (11.144)$$

Die WKB-Näherung für $\psi(x)$ lautet dann

$$\psi(x) \cong \frac{\psi_+}{\sqrt{\kappa(x)}} \exp\left(-\int_{x_+}^x dx' \kappa(x')\right) + \frac{\psi_-}{\sqrt{\kappa(x)}} \exp\left(\int_{x_-}^x dx' \kappa(x')\right), \quad (11.145)$$

wobei komplexe Faktoren in ψ_{\pm} absorbiert wurden. Wenn sich ein klassisch verbotener Bereich bis $\pm\infty$ erstreckt, schließt die Forderung der Quadratintegrabilität oder zumindest Beschränktheit oft eine der beiden unabhängigen Lösungen aus, weil sie exponentiell anwächst.

An den klassischen Umkehrpunkten x^* gilt $E = V(x^*)$ und damit $k = 0$ und $\kappa = 0$. Dann versagt die WKB-Näherung offensichtlich, da die Vorfaktoren $1/\sqrt{k(x^*)}$ und $1/\sqrt{\kappa(x^*)}$ divergieren.



Auch in einer Umgebung der Umkehrpunkte ist die Wellenzahl k i. A. klein und $\psi(x)$ daher betragsmäßig zu groß. Unter welchen Bedingungen ist die WKB-Näherung gültig? Um dies zu beantworten, untersuchen wir, für welches Potential die genäherte Wellenfunktion exakt wäre. Es ist

$$\begin{aligned}
 \frac{d^2}{dx^2} \frac{\psi_{\pm}}{\sqrt{k(x)}} \exp\left(\pm i \int_{x_{\pm}}^x dx' k(x')\right) &= \psi_{\pm} \frac{d}{dx} \left[-\frac{k'(x)}{2k^{3/2}(x)} \exp\left(\pm i \int_{x_{\pm}}^x dx' k(x')\right) \right. \\
 &\quad \left. \pm i \sqrt{k(x)} \exp\left(\pm i \int_{x_{\pm}}^x dx' k(x')\right) \right] \\
 &= \psi_{\pm} \left[-\frac{k''(x)}{2k^{3/2}(x)} + \frac{3(k'(x))^2}{4k^{5/2}(x)} \mp i \frac{k'(x)}{2\sqrt{k(x)}} \pm i \frac{k'(x)}{2\sqrt{k(x)}} - k^{3/2}(x) \right] \\
 &\quad \times \exp\left(\pm i \int_{x_{\pm}}^x dx' k(x')\right) \\
 &= \left[-\frac{k''(x)}{2k(x)} + \frac{3(k'(x))^2}{4k^2(x)} - k^2(x) \right] \frac{\psi_{\pm}}{\sqrt{k(x)}} \exp\left(\pm i \int_{x_{\pm}}^x dx' k(x')\right).
 \end{aligned} \tag{11.146}$$

Die WKB-Näherung für $\psi(x)$ erfüllt also die Gleichung

$$\psi''(x) + \left[k^2(x) + \frac{1}{2} \frac{k''(x)}{k(x)} - \frac{3}{4} \frac{(k'(x))^2}{k^2(x)} \right] \psi(x) = 0, \tag{11.147}$$

anstelle der eigentlich interessierenden (Schrödinger-) Gleichung

$$\psi''(x) + k^2(x) \psi(x) = 0. \tag{11.148}$$

Die Näherung ist gut, wenn die Zusatzterme in der ersten Gleichung klein im Vergleich zu $k^2\psi$ sind, d. h. für

$$\left| \frac{1}{2} \frac{k''(x)}{k^3(x)} - \frac{3}{4} \frac{(k'(x))^2}{k^4(x)} \right| \ll 1. \tag{11.149}$$

Generisch, d. h. ohne spezielle Wahl von $V(x)$ und damit $k(x)$, ist dies nur erfüllt, wenn beide Terme jeweils für sich klein sind. Dann ist

$$\left| \frac{(k'(x))^2}{k^4(x)} \right| \ll 1, \tag{11.150}$$

was äquivalent ist zu

$$\left| \frac{k'(x)}{k^2(x)} \right| = \left| \frac{d}{dx} \frac{1}{k(x)} \right| \ll 1 \quad (11.151)$$

und auch zu

$$\left| \frac{d}{dx} \frac{2\pi}{k(x)} \right| \equiv \left| \frac{d\lambda}{dx} \right| \ll 1 \quad (11.152)$$

mit der de Broglie-Wellenlänge $\lambda = 2\pi/k$. Die WKB-Näherung erfordert also, dass sich die de Broglie-Wellenlänge räumlich nur langsam ändert, also wie erwartet schwach veränderliche Potentiale. An den klassischen Umkehrpunkten wird aber $k = 0$ und λ divergiert. In ihrer Umgebung ist die Bedingung $|d\lambda/dx| \ll 1$ daher sicher nicht erfüllt. Wegen der Divergenz von $\psi(x)$ in der WKB-Näherung an Umkehrpunkten kann man dort auch nicht Lösungen in den beiden Bereichen (erlaubt, verboten) stetig oder differenzierbar aneinander anschließen und so einen Teil der freien Parameter festlegen. Beachte, dass kein solches Problem auftritt, wenn das Potential $V(x)$ bei x^* einen *Sprung* von einem Wert weit unterhalb der Energie E auf einen Wert weit oberhalb von E hat. Dann bleiben die Lösungen für den erlaubten und den verbotenen Bereich bis zum Punkt $x = x^*$ gültig und können aneinander angeschlossen werden.

Ein Ausweg für den Fall mit stetigem $V(x)$ besteht darin, in der Umgebung des Umkehrpunktes x^* das Potential zu linearisieren:

$$V(x) \cong V(x^*) + V'(x^*) (x - x^*). \quad (11.153)$$

Damit kann die Schrödinger-Gleichung *exakt* gelöst werden. Man erhält Lösungen in den Bereichen:

- WKB-Näherung im erlaubten Bereich, hinreichend weit weg vom Umkehrpunkt x^* ,
- exakte Lösung für linearisiertes Potential in einer Umgebung von x^* ,
- WKB-Näherung im verbotenen Bereich, hinreichend weit weg von x^* .

Diese Lösungen können stetig differenzierbar aneinander angeschlossen werden. Eine Weiterentwicklung dieser Idee geht auf Rudolf E. Langer (1937) zurück. Die *Langer-Korrektur* besteht darin, die WKB-Gleichungen so abzuändern, dass sie weit weg vom Umkehrpunkt gegen die ursprünglichen WKB-Gleichungen gehen, aber für $x \rightarrow x^*$ zur vollen (nicht quasiklassischen) und exakt lösbaren Schrödinger-Gleichung für linearisiertes Potential äquivalent sind. Damit konnte Langer die freien Parameter der Lösungen im klassisch erlaubten und verbotenen Bereich direkt miteinander in Verbindung setzen. Der Bereich in der Nähe der Umkehrpunkte und die Abänderung der WKB-Gleichungen treten dann in der Praxis gar nicht mehr auf. Wir geben hier nur das Ergebnis an: Die WKB-Näherung lautet

$$\psi(x) = \begin{cases} \frac{\psi_0}{\sqrt{k(x)}} \cos\left(\alpha(x) - \frac{\pi}{4} + \varphi\right) & \text{im klassisch erlaubten Bereich,} \\ \frac{\psi_0}{\sqrt{|k(x)|}} \left(\frac{1}{2} \cos \varphi e^{-|\alpha(x)|} + \sin \varphi e^{|\alpha(x)|} \right) & \text{im klassisch verbotenen Bereich.} \end{cases} \quad (11.154)$$

Beachte, dass nur vor dem Kosinus-Term ein Faktor $1/2$ steht. Hier ist ψ_0 eine Normierungskonstante und φ ist eine konstante Phase, die sich aus Normierbarkeits- oder Randbedingungen ergibt. ψ_0 und φ nehmen im erlaubten und im verbotenen Bereich *dieselben* Werte an. Weiter ist

$$\alpha(x) := \begin{cases} \int_{x^*}^x dx' k(x') & \text{für } x^* \text{ linksseitig,} \\ \int_x^{x^*} dx' k(x') & \text{für } x^* \text{ rechtsseitig.} \end{cases} \quad (11.155)$$

Linksseitig (rechtsseitig) bedeutet, dass der klassisch verbotene Bereich links (rechts) von x^* liegt, d. h. bei $x < x^*$ ($x > x^*$). Mit der Langer-Methode können nun z. B. Reflexions- und Transmissionskoeffizienten für Tunnelbarrieren bestimmt werden. Die WKB-Näherung für die Wellenfunktion ist in der Nähe der Umkehrpunkte natürlich immer noch divergent und daher unbrauchbar.

Wir machen noch einige abschließende Bemerkungen zur WKB-Näherung:

- Sie gilt nur für langsam räumlich veränderliche Potentiale und für Sprünge.
- Sie ist im Prinzip eine kontrollierte Näherung, da man zu höheren Ordnungen in \hbar gehen kann. In der Praxis geschieht dies selten.
- Andererseits existiert wie in der Störungstheorie keine rigorose Fehlerabschätzung.
- Sie versagt in der Nähe der klassischen Umkehrpunkte, jedenfalls für stetige Potentiale $V(x)$. Dies kann durch hybride Methoden wie die Langer-Korrektur tw. behoben werden, die keine konsistenten Entwicklungen in \hbar darstellen.

Kapitel 12

Der Dichteoperator

Das Ergebnis einer Messung ist in der Quantenmechanik i. A. nicht sicher vorhersagbar, selbst wenn der Zustand des Systems vollständig bekannt ist. Die Bornsche Regel liefert die Wahrscheinlichkeiten für verschiedene Ausgänge der Messung: Für einen nicht entarteten Eigenwert A einer Observable \hat{A} mit zugehörigem Eigenzustand $|\psi_A\rangle$ ist die Wahrscheinlichkeit der Messung von A im Zustand $|\psi\rangle$ gegeben durch $W_A = |\langle\psi_A|\psi\rangle|^2$. Im Fall der Entartung von A mit mehreren Eigenzuständen $|\psi_{An}\rangle$, ist die entsprechende Wahrscheinlichkeit

$$W_A = \sum_n |\langle\psi_{An}|\psi\rangle|^2. \quad (12.1)$$

Diese Unvorhersehbarkeit ist intrinsisch für die Quantentheorie. Oft tritt aber noch eine andere Form von Unbestimmtheit auf, wenn wir nämlich gar nicht sicher wissen, in welchem Zustand ein System vor der Messung ist. Diese Unbestimmtheit hat zunächst nichts mit Quantentheorie zu tun, sie tritt auch in der klassischen Physik auf. In diesem kurzen Kapitel besprechen wir, wie beide Formen von Stochastik in einem einheitlichen Formalismus beschrieben werden können.

12.1 Projektionsoperatoren und Dichteoperatoren für reine Zustände

Wir definieren zunächst den Begriff des Projektionsoperators: Ein Operator \hat{P} ist ein *Projektionsoperator*, wenn \hat{P} (a) selbstadjungiert und (b) idempotent ist, d. h. wenn gilt $\hat{P}^2 = \hat{P}$.

Wenn $|\psi\rangle$ ein normierter Hilbert-Raum-Vektor ist, dann ist

$$\hat{P}_\psi = |\psi\rangle\langle\psi| \quad (12.2)$$

ein Projektionsoperator, denn

$$\hat{P}_\psi^\dagger = (|\psi\rangle\langle\psi|)^\dagger = |\psi\rangle\langle\psi| = \hat{P}_\psi, \quad (12.3)$$

$$\hat{P}_\psi^2 = |\psi\rangle \underbrace{\langle\psi|\psi\rangle}_{=1} \langle\psi| = \hat{P}_\psi. \quad (12.4)$$

Jedoch sind nicht alle Projektionsoperatoren von dieser Form.

Ein Projektionsoperator \hat{P} kann nur die Eigenwerte 0 und 1 haben. Der Beweis ist einfach: Unter Verwendung der Eigenbasis von \hat{P} sehen wir, dass die Eigenwerte $p_i^2 = p_i$ erfüllen müssen, was nur für $p_i = 0, 1$ möglich ist.

Was wir bisher einfach als Zustand bezeichnet haben, nennen wir jetzt einen *reinen Zustand*. Für einen reinen Zustand, der durch den normierten Hilbert-Raum-Vektor $|\psi\rangle$ beschrieben wird, definieren wir den *Dichteoperator* (oder *Statistischen Operator*) durch

$$\hat{\rho} := \hat{P}_\psi = |\psi\rangle\langle\psi| \quad (12.5)$$

Er ist also gerade der Projektionsoperator auf $|\psi\rangle$. Seine Spur ist

$$\text{Sp } \hat{\rho} = \sum_n \langle \psi_n | \psi \rangle \langle \psi | \psi_n \rangle = \langle \psi | \underbrace{\sum_n |\psi_n\rangle \langle \psi_n|}_{=1} \psi \rangle = \langle \psi | \psi \rangle = 1 \quad (12.6)$$

mit einer beliebigen Orthonormalbasis $\{|\psi_n\rangle\}$.

Man kann im Prinzip die gesamte Quantenmechanik unter Verwendung von Dichteoperatoren anstelle von Zustandsvektoren formulieren. Gleichung (12.5) legt eine Bijektion zwischen Zustandsvektoren und Projektionsoperatoren auf diese Zustände nahe. Diese gilt auch fast; nur ein Phasenfaktor in $|\psi\rangle$ ist in $\hat{\rho}$ nicht enthalten, weil $\langle \psi |$ den komplex konjugierten Faktor enthält. Aber da die globale Phase ja keine physikalischen Konsequenzen hat, ist es als Vorzug des Dichteoperator-Formalismus anzusehen, dass sie darin gar nicht vorkommt.

Erwartungswerte sind z. B.

$$\langle \psi | \hat{A} | \psi \rangle = \sum_n \langle \psi | \hat{A} | \psi_n \rangle \langle \psi_n | \psi \rangle = \sum_n \langle \psi_n | \psi \rangle \langle \psi | \hat{A} | \psi_n \rangle = \sum_n \langle \psi_n | \hat{\rho} \hat{A} | \psi_n \rangle = \text{Sp } \hat{\rho} \hat{A}. \quad (12.7)$$

Die Zeitentwicklung von $\hat{\rho}$ folgt natürlich aus der Schrödinger-Gleichung für $|\psi\rangle$:

$$\begin{aligned} \frac{d}{dt} \hat{\rho} &= \frac{d}{dt} |\psi\rangle \langle \psi| \\ &= \left(\frac{d}{dt} |\psi\rangle \right) \langle \psi| + |\psi\rangle \frac{d}{dt} \langle \psi| \\ &= \frac{1}{i\hbar} \hat{H} |\psi\rangle \langle \psi| - |\psi\rangle \frac{1}{i\hbar} \langle \psi| \hat{H} \\ &= \frac{1}{i\hbar} [\hat{H}, \hat{\rho}] = -\frac{1}{i\hbar} [\hat{\rho}, \hat{H}] \end{aligned} \quad (12.8)$$

$$\Rightarrow i\hbar \frac{d\hat{\rho}}{dt} = -[\hat{\rho}, \hat{H}]. \quad (12.9)$$

Dies ist die *von Neumann-Gleichung* für den Dichteoperator. Sie hat dieselbe Form wie die Heisenberg-Gleichung bis auf ein Vorzeichen, aber wir arbeiten hier im *Schrödinger*-, nicht im *Heisenberg*-Bild. Die Zeitentwicklung wird daher von den Zuständen getragen, die nun durch den Dichteoperator repräsentiert werden. Im Heisenberg-Bild ist $\hat{\rho}$ dagegen zeitunabhängig.

Für reine Zustände ist der Dichteoperator-Formalismus äquivalent zum Dirac-Formalismus und hat keine besonderen Vorteile. Höchstens hat er die ästhetisch ansprechende Eigenschaft, dass nur eine Klasse von mathematischen Objekten – nämlich Operatoren – betrachtet wird, und nicht zwei – Vektoren und Operatoren. Allerdings wirkt es etwas unnatürlich, Operatoren auf einem Hilbert-Raum zu untersuchen, ohne nach der physikalischen Bedeutung der Vektoren in diesem Hilbert-Raum zu fragen.

12.2 Gemischte Zustände

Die Stärke des Formalismus zeigt sich, wenn man kein vollständiges Wissen über den Zustand eines Systems hat. Wir nehmen an, ein System sei mit bestimmten Wahrscheinlichkeiten p_i in Zuständen $|\varphi_i\rangle$. Dabei sollen die Vektoren $|\varphi_i\rangle$ normiert sein, aber verschiedene $|\varphi_i\rangle$ müssen nicht orthogonal und noch nicht einmal linear unabhängig sein. Die Gesamtheit der p_i und $|\varphi_i\rangle$ beschreibt einen *gemischten Zustand* des Systems. Die Wahrscheinlichkeiten p_i haben wie gesagt nichts mit Quantenmechanik zu tun, sie beschreiben einfach unsere unvollständige Kenntnis des Systems. Sie ergeben sich daher auch nicht aus der Bornschen Regel für irgendeinen Zustandsvektor. Der übliche Dirac-Formalismus der Quantenmechanik auf Basis von Zustandsvektoren wird also nicht ausreichen, um die Unkenntnis auszudrücken.

Beispiel: Im zweidimensionalen Hilbert-Raum eines Spins der Länge 1/2 existiert kein normierter Vektor $|\psi\rangle$, der vollständige Unkenntnis des Zustandes beschreibt. Ein solcher müsste nämlich

$$\langle \psi | \hat{S}_x | \psi \rangle = \langle \psi | \hat{S}_y | \psi \rangle = \langle \psi | \hat{S}_z | \psi \rangle = 0 \quad (12.10)$$

erfüllen. Sei nun

$$|\psi\rangle = \alpha|\uparrow\rangle + \beta|\downarrow\rangle \quad (12.11)$$

mit $|\alpha|^2 + |\beta|^2 = 1$. Dann folgt

$$\langle\psi|\hat{S}_x|\psi\rangle = \frac{\hbar}{2} (\alpha^*\langle\uparrow| + \beta^*\langle\downarrow|) \sigma_x (\alpha|\uparrow\rangle + \beta|\downarrow\rangle) = \frac{\hbar}{2} (\alpha^*\beta + \beta^*\alpha) = 0, \quad (12.12)$$

$$\langle\psi|\hat{S}_y|\psi\rangle = \frac{\hbar}{2} (\alpha^*\langle\uparrow| + \beta^*\langle\downarrow|) \sigma_y (\alpha|\uparrow\rangle + \beta|\downarrow\rangle) = \frac{\hbar}{2} (-i\alpha^*\beta + i\beta^*\alpha) = 0, \quad (12.13)$$

$$\langle\psi|\hat{S}_z|\psi\rangle = \frac{\hbar}{2} (\alpha^*\langle\uparrow| + \beta^*\langle\downarrow|) \sigma_z (\alpha|\uparrow\rangle + \beta|\downarrow\rangle) = \frac{\hbar}{2} (\alpha^*\alpha - \beta^*\beta) = 0. \quad (12.14)$$

Aus der letzten Gleichung folgt $|\alpha|^2 = |\beta|^2$ und mit der Normierung

$$|\alpha|^2 = |\beta|^2 = \frac{1}{2}. \quad (12.15)$$

Aus den anderen beiden Gleichungen folgt

$$\operatorname{Re} \alpha^*\beta = \operatorname{Im} \alpha^*\beta = 0 \quad (12.16)$$

und daraus

$$\alpha^*\beta = 0 \quad \Rightarrow \quad \alpha\beta^*\alpha^*\beta = 0 \quad \Rightarrow \quad |\alpha|^2|\beta|^2 = 0. \quad (12.17)$$

Wir finden einen Widerspruch zu Gl. (12.15). Man kann auch mit Hilfe der Bloch-Kugel argumentieren: Jeder Vektor im Spin-1/2-Hilbert-Raum beschreibt einen Spin der Länge 1/2, der in eine bestimmte Richtung zeigt.

Die Unkenntnis kann aber mittels des Dichteoperators beschrieben werden, wie wir nun plausibel machen werden. Der Erwartungswert einer Observablen \hat{A} sollte nach den üblichen Regeln der Stochastik gegeben sein durch

$$\langle\hat{A}\rangle = \sum_i p_i \langle\hat{A}\rangle_{\varphi_i} \quad (12.18)$$

und die Quantenmechanik liefert dann

$$\langle\hat{A}\rangle = \sum_i p_i \langle\varphi_i|\hat{A}|\varphi_i\rangle. \quad (12.19)$$

Einschieben einer Orthonormalbasis $\{|\psi_n\rangle\}$ ergibt

$$\begin{aligned} \langle\hat{A}\rangle &= \sum_i \sum_n p_i \langle\varphi_i|\hat{A}|\psi_n\rangle \langle\psi_n|\varphi_i\rangle \\ &= \sum_n \langle\psi_n| \sum_i p_i |\varphi_i\rangle \langle\varphi_i|\hat{A}|\psi_n\rangle \\ &= \operatorname{Sp} \left(\sum_i p_i |\varphi_i\rangle \langle\varphi_i|\hat{A} \right). \end{aligned} \quad (12.20)$$

Dies ist von derselben Form wie $\langle\hat{A}\rangle = \operatorname{Sp} \hat{\rho} \hat{A}$ für einen reinen Zustand. Wir definieren daher nun den Dichteoperator für einen allgemeinen, gemischten Zustand durch

$$\hat{\rho} := \sum_i p_i |\varphi_i\rangle \langle\varphi_i|. \quad (12.21)$$

Für einen reinen Zustand haben wir volle Kenntnis, also enthält die Summe über i nur einen Term und dieser hat die Wahrscheinlichkeit $p = 1$. Damit erhalten wir wieder die Form aus 12.1, $\hat{\rho} = |\varphi\rangle \langle\varphi|$, als Spezialfall.

Wir untersuchen nun die Eigenschaften des Dichteoperators.

- $\hat{\rho}$ ist selbstadjungiert:

$$\hat{\rho}^\dagger = \sum_i p_i (|\varphi_i\rangle \langle\varphi_i|)^\dagger = \sum_i p_i |\varphi_i\rangle \langle\varphi_i| = \hat{\rho}. \quad (12.22)$$

- $\hat{\rho}$ hat die Spur 1:

$$\begin{aligned}
\text{Sp } \hat{\rho} &= \sum_n \langle \psi_n | \sum_i p_i |\varphi_i\rangle \langle \varphi_i | \psi_n \rangle \\
&= \sum_i p_i \langle \varphi_i | \underbrace{\sum_n |\psi_n\rangle \langle \psi_n |}_{=1} \varphi_i \rangle \\
&= \sum_i p_i = 1,
\end{aligned} \tag{12.23}$$

da sich die Wahrscheinlichkeiten bei Summation über alle Alternativen natürlich zu 1 addieren müssen.

- $\hat{\rho}$ ist positiv, d. h. $\hat{\rho}$ ist selbstadjungiert und hat keine negativen Eigenwerte. Beweis: Angenommen, $\hat{\rho}$ habe einen negativen Eigenwert $r < 0$ mit dem Eigenvektor $|r\rangle$. Dann folgt

$$0 > r = \langle r | \hat{\rho} | r \rangle = \sum_i p_i \langle r | \varphi_i \rangle \langle \varphi_i | r \rangle = \sum_i p_i |\langle r | \varphi_i \rangle|^2 \geq 0, \tag{12.24}$$

da Wahrscheinlichkeiten nichtnegativ sind. Widerspruch!

- Wie schon gesehen, gilt

$$\langle \hat{A} \rangle = \text{Sp } \hat{\rho} \hat{A}. \tag{12.25}$$

- Die Zeitentwicklung von $\hat{\rho}$ gehorcht der von Neumann-Gleichung

$$i\hbar \frac{d\hat{\rho}}{dt} = -[\hat{\rho}, \hat{H}]. \tag{12.26}$$

Beweis:

$$\begin{aligned}
i\hbar \frac{d\hat{\rho}}{dt} &= \sum_i p_i i\hbar \frac{d}{dt} |\varphi_i\rangle \langle \varphi_i| \\
&= \sum_i p_i \left[\left(i\hbar \frac{d}{dt} |\varphi_i\rangle \right) \langle \varphi_i| + |\varphi_i\rangle i\hbar \frac{d}{dt} \langle \varphi_i| \right] \\
&= \sum_i p_i \left(\hat{H} |\varphi_i\rangle \langle \varphi_i| - |\varphi_i\rangle \langle \varphi_i| \hat{H} \right) \\
&= [\hat{H}, \hat{\rho}] = -[\hat{\rho}, \hat{H}].
\end{aligned} \tag{12.27}$$

Diese Zeitentwicklung erhält die Spur:

$$\begin{aligned}
\frac{d}{dt} \text{Sp } \hat{\rho} &= \text{Sp } \frac{d\hat{\rho}}{dt} = -\frac{1}{i\hbar} \text{Sp } [\hat{\rho}, \hat{H}] \\
&= -\frac{1}{i\hbar} (\text{Sp } \hat{\rho} \hat{H} - \underbrace{\text{Sp } \hat{H} \hat{\rho}}_{= \text{Sp } \hat{\rho} \hat{H}}) = 0.
\end{aligned} \tag{12.28}$$

- Ein Zustand ist *rein* genau dann, wenn gilt $\text{Sp } \hat{\rho}^2 = 1$. Das ist ein nützliches Kriterium für die Reinheit von Zuständen, da diese der Matrixdarstellung von $\hat{\rho}$ in irgendeiner Basis oft nicht leicht anzusehen ist.

Beweis: 1. Ist $\hat{\rho}$ ein reiner Zustand, so existiert ein normierter Vektor $|\varphi\rangle$ mit

$$\hat{\rho} = |\varphi\rangle \langle \varphi| \tag{12.29}$$

$$\Rightarrow \hat{\rho}^2 = |\varphi\rangle \langle \varphi| \varphi\rangle \langle \varphi| = |\varphi\rangle \langle \varphi| = \hat{\rho} \quad \Big| \quad \text{idempotent!} \tag{12.30}$$

$$\Rightarrow \text{Sp } \hat{\rho}^2 = \text{Sp } \hat{\rho} = 1. \tag{12.31}$$

2. Es gelte $\text{Sp } \hat{\rho}^2 = 1$. Angenommen, $\hat{\rho}$ sei kein reiner Zustand, Dann existieren mindestens zwei linear unabhängige normierte Vektoren $|\varphi_i\rangle$ und Zahlen $p_i \in]0, 1[$ mit $\sum_i p_i = 1$, so dass

$$\hat{\rho} = \sum_i p_i |\varphi_i\rangle \langle \varphi_i|. \quad (12.32)$$

Es folgt

$$\hat{\rho}^2 = \sum_{ij} p_i p_j |\varphi_i\rangle \langle \varphi_i | \varphi_j\rangle \langle \varphi_j|. \quad (12.33)$$

Achtung: die $|\varphi_i\rangle$ sind i. A. nicht orthogonal. Nach Voraussetzung ist

$$\begin{aligned} 1 &= \text{Sp } \hat{\rho}^2 \\ &= \sum_n \sum_{ij} p_i p_j \langle \psi_n | \varphi_i\rangle \langle \varphi_i | \varphi_j\rangle \langle \varphi_j | \psi_n\rangle \\ &= \sum_{ij} p_i p_j \langle \varphi_i | \varphi_j\rangle \langle \varphi_j | \underbrace{\sum_n |\psi_n\rangle \langle \psi_n|}_{=1} \varphi_i\rangle \\ &= \sum_{ij} p_i p_j |\langle \varphi_i | \varphi_j\rangle|^2. \end{aligned} \quad (12.34)$$

Nach der Cauchy-Schwarzschen Ungleichung ist $|\langle \varphi_i | \varphi_j\rangle|^2 \leq \langle \varphi_i | \varphi_i\rangle \langle \varphi_j | \varphi_j\rangle$ und, da die $|\varphi_i\rangle$ linear unabhängig sind, sogar

$$|\langle \varphi_i | \varphi_j\rangle|^2 < \langle \varphi_i | \varphi_i\rangle \langle \varphi_j | \varphi_j\rangle = 1 \quad (12.35)$$

$$\Rightarrow 1 = \text{Sp } \hat{\rho}^2 = \sum_{ij} p_i p_j |\langle \varphi_i | \varphi_j\rangle|^2 < \sum_{ij} p_i p_j = \left(\sum_i p_i \right) \left(\sum_j p_j \right) = 1. \quad (12.36)$$

Widerspruch!

Beispiele: 1. Wir untersuchen

$$\hat{\rho} = \begin{pmatrix} 1/2 & 1/2 \\ 1/2 & 1/2 \end{pmatrix} \quad (12.37)$$

auf dem Hilbert-Raum eines Spins $1/2$. $\hat{\rho}$ ist selbstadjungiert, hat die Spur 1 und ist positiv. Beschreibt $\hat{\rho}$ einen reinen Zustand? Wir prüfen

$$\text{Sp } \hat{\rho}^2 = \text{Sp } \begin{pmatrix} 1/2 & 1/2 \\ 1/2 & 1/2 \end{pmatrix} \begin{pmatrix} 1/2 & 1/2 \\ 1/2 & 1/2 \end{pmatrix} = \text{Sp } \begin{pmatrix} 1/2 & 1/2 \\ 1/2 & 1/2 \end{pmatrix} = 1, \quad (12.38)$$

also ist $\hat{\rho}$ rein. (Man müsste die Außerdiagonalkomponenten von $\hat{\rho}^2$ nicht berechnen, um dies auszusagen.) Tatsächlich lässt sich $\hat{\rho}$ schreiben als $\hat{\rho} = |\varphi\rangle \langle \varphi|$ mit

$$|\varphi\rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix} \equiv |\rightarrow\rangle, \quad (12.39)$$

das ist der Eigenvektor zu \hat{S}_x zum Eigenwert $\hbar/2$.

2. Ist

$$\hat{\rho} = \begin{pmatrix} 1/2 & 0 \\ 0 & 1/2 \end{pmatrix} \quad (12.40)$$

rein? Nein, denn

$$\text{Sp } \hat{\rho}^2 = \text{Sp } \begin{pmatrix} 1/4 & 0 \\ 0 & 1/4 \end{pmatrix} = \frac{1}{2} < 1. \quad (12.41)$$

Explizit ist

$$\hat{\rho} = \frac{1}{2} |\uparrow\rangle \langle \uparrow| + \frac{1}{2} |\downarrow\rangle \langle \downarrow|. \quad (12.42)$$

Dieser Dichtoperator ist proportional zur Einheitsmatrix $\mathbb{1}$ und daher invariant unter allen Drehungen. Er drückt maximale Unkenntnis aus.

- Die formale Lösung der von Neumann-Gleichung erhält man aus der formalen Lösung der Schrödinger-Gleichung: mit dem Zeitentwicklungsoperator $\hat{U}(t, t_0)$ ist

$$|\psi(t)\rangle = \hat{U}(t, t_0)|\psi(t_0)\rangle. \quad (12.43)$$

Daraus folgt

$$\begin{aligned} \hat{\rho}(t) &= \sum_i p_i |\varphi_i(t)\rangle \langle \varphi_i(t)| \\ &= \hat{U}(t, t_0) \sum_i p_i |\varphi_i(t_0)\rangle \langle \varphi_i(t_0)| \hat{U}^\dagger(t, t_0) \\ &= \hat{U}(t, t_0) \hat{\rho}(t_0) \hat{U}^\dagger(t, t_0). \end{aligned} \quad (12.44)$$

Damit sieht man leicht, dass gilt

$$\text{Sp } \hat{\rho}(t) = \text{Sp } \hat{U}(t, t_0) \hat{\rho}(t_0) \hat{U}^\dagger(t, t_0) = \text{Sp } \underbrace{\hat{U}^\dagger(t, t_0) \hat{U}(t, t_0)}_{=1} \hat{\rho}(t_0) = \text{Sp } \hat{\rho}(t_0), \quad (12.45)$$

wie bekannt, und

$$\text{Sp } \hat{\rho}^2(t) = \text{Sp } \hat{U}(t, t_0) \hat{\rho}(t_0) \underbrace{\hat{U}^\dagger(t, t_0) \hat{U}(t, t_0)}_{=1} \hat{\rho}(t_0) \hat{U}^\dagger(t, t_0) = \text{Sp } \hat{\rho}^2(t_0). \quad (12.46)$$

Die Reinheit des Zustands ist also ebenfalls erhalten.

- Da $\hat{\rho}$ selbstadjungiert ist, existiert ein Orthonormalsystem von Eigenvektoren $|\xi_n\rangle$ und $\hat{\rho}$ lässt sich schreiben als

$$\hat{\rho} = \sum_n \pi_n |\chi_n\rangle \langle \chi_n|. \quad (12.47)$$

Das ist die Spektraldarstellung von $\hat{\rho}$, siehe 7.2.1. Es gilt offensichtlich

$$1 = \text{Sp } \hat{\rho} = \sum_n \pi_n \quad (12.48)$$

und wegen der Positivität von $\hat{\rho}$

$$\pi_n \geq 0 \quad \forall n. \quad (12.49)$$

Die π_n lassen sich als Wahrscheinlichkeiten der – nun orthonormalen – Zustände $|\xi_n\rangle$ verstehen.

Für gegebene Wahrscheinlichkeiten p_i bzw. π_n können wir die Erwartungswerte beliebiger Observabler ausrechnen, auch als Funktion der Zeit. Die Bestimmung der Wahrscheinlichkeiten selbst ist Gegenstand der Statistischen Physik, genauer der Quantenstatistik.

12.2.1 Messung und Projektionspostulat

Wie stellt sich nun der quantenmechanische Messprozess im Dichteoperator-Formalismus dar? Wir betrachten der Einfachheit halber die Messung eines nicht entarteten Eigenwertes.

Wird bei der Messung der Observable \hat{A} der (nicht entartete) Eigenwert A zum Eigenzustand $|\psi_A\rangle$ beobachtet, so geht nach dem von Neumannschen Projektionspostulat ein reiner Anfangszustand $|\varphi\rangle$ instantan in den Zustand $|\psi_A\rangle$ über. Nach der Bornschen Regel ist die Wahrscheinlichkeit dafür

$$p_A = |\langle \psi_A | \varphi \rangle|^2. \quad (12.50)$$

Wir schreiben jetzt den Dichteoperator nach der Messung (o. B. d. A. zur Zeit $t = 0$) in der Eigenbasis $\{|\psi_A\rangle\}$ von \hat{A} :

$$\hat{\rho}(t = 0^+) = \sum_A p_A |\psi_A\rangle \langle \psi_A|$$

$$\begin{aligned}
&= \sum_A |\langle \psi_A | \varphi \rangle|^2 |\psi_A\rangle\langle\psi_A| \\
&= \sum_A |\psi_A\rangle\langle\psi_A| \varphi \rangle \langle \varphi | \psi_A\rangle \langle\psi_A| \\
&= \sum_A \hat{P}_A \hat{\rho}(t=0^-) \hat{P}_A,
\end{aligned} \tag{12.51}$$

wobei $\hat{P}_A = |\psi_A\rangle\langle\psi_A|$ der Projektionsoperator auf den Zustand $|\psi_A\rangle$ und $\hat{\rho}(t=0^-)$ der Dichteoperator unmittelbar vor der Messung ist.

War der Zustand vor der Messung schon gemischt, so sollten wir das Ergebnis entsprechend mitteln:

$$\hat{\rho}(t=0^-) = \sum_i q_i |\varphi_i\rangle\langle\varphi_i|. \tag{12.52}$$

Dann ist

$$\hat{\rho}(t=0^+) = \sum_i q_i \sum_A |\psi_A\rangle\langle\psi_A| \varphi_i \rangle \langle \varphi_i | \psi_A\rangle \langle\psi_A| = \sum_A \hat{P}_A \hat{\rho}(t=0^-) \hat{P}_A, \tag{12.53}$$

also dieselbe Beziehung wie für einen reinen Anfangszustand. Hier sieht man übrigens gut, wieso das Projektionspostulat diesen Namen trägt.

Schreiben wir den Dichteoperator *vor* der Messung auch in der Eigenbasis von \hat{A} , so erhalten wir

$$\hat{\rho}(0^-) = \mathbb{1} \hat{\rho}(0^-) \mathbb{1} = \sum_{A,A'} |\psi_A\rangle\langle\psi_A| \hat{\rho}(0^-) |\psi_{A'}\rangle\langle\psi_{A'}|. \tag{12.54}$$

In Matrixschreibweise sind die Komponenten

$$\rho_{AA'}(0^-) = \langle\psi_A| \hat{\rho}(0^-) |\psi_{A'}\rangle. \tag{12.55}$$

Nach der Messung ist, wie gesehen,

$$\hat{\rho}(0^+) = \sum_A |\psi_A\rangle\langle\psi_A| \hat{\rho}(0^-) |\psi_A\rangle\langle\psi_A|. \tag{12.56}$$

In Matrixschreibweise sind die Komponenten nun

$$\rho_{AA'}(0^+) = \delta_{AA'} \rho_{AA}(0^-). \tag{12.57}$$

Die Messung führt also dazu, dass alle Außerdiagonalkomponenten von $\hat{\rho}$ in der Eigenbasis der gemessenen Observablen auf Null gesetzt werden (diese Komponenten nennt man auch *Kohärenzen*). Die Diagonalkomponenten bleiben dagegen unangetastet. Beispiel: Die Messung von \hat{S}_z am (reinen) Zustand

$$\hat{\rho}(0^-) = \begin{pmatrix} 1/2 & 1/2 \\ 1/2 & 1/2 \end{pmatrix} \tag{12.58}$$

führt zum Kollaps in den gemischten Zustand

$$\hat{\rho}(0^+) = \begin{pmatrix} 1/2 & 0 \\ 0 & 1/2 \end{pmatrix}. \tag{12.59}$$

12.2.2 Teilspur und Zustand in Faktorräumen

Nehmen wir an, der Hilbert-Raum eines zusammengesetzten Systems sei der Produktraum

$$\mathcal{H} = \mathcal{H}_1 \otimes \mathcal{H}_2. \tag{12.60}$$

Das System sei im reinen Zustand $|\psi\rangle \in \mathcal{H}$. Man interessiert sich oft für den Zustand in einem Teilsystem. Letztlich ist das immer der Fall, wenn wir ein kleines System (Spin, Atom ...) getrennt vom Rest des Universums

untersuchen. Der formale Weg, den Zustand in einem Teilsystem zu erhalten, ist die Bildung der Spur über eine Orthonormalbasis für den Rest des Universums. Seien also $\{|\psi_m^1\rangle\}$ und $\{|\psi_n^2\rangle\}$ Orthonormalbasen von \mathcal{H}_1 bzw. \mathcal{H}_2 . Wir hatten in 7.1.3 gesehen, dass dann $\{|\psi_m^1\rangle|\psi_n^2\rangle\}$ eine Basis von \mathcal{H} ist. Die *Teilspuren* über Teilsysteme 1 und 2 sind gegeben durch

$$\text{Sp}_1 \bullet = \sum_m \langle \psi_m^1 | \bullet | \psi_m^1 \rangle, \quad (12.61)$$

$$\text{Sp}_2 \bullet = \sum_n \langle \psi_n^2 | \bullet | \psi_n^2 \rangle. \quad (12.62)$$

Das Ergebnis ist jeweils ein Operator auf dem übrigbleibenden Raum \mathcal{H}_2 bzw. \mathcal{H}_1 .

Die Spurbildung ist nur für Operatoren definiert, also müssen wir den reinen Zustand $|\psi\rangle$ durch den Dichteoperator $\hat{\rho} = |\psi\rangle\langle\psi|$ darstellen. Der Zustand im Teilsystem 1 wird nun beschrieben durch den *reduzierten Dichteoperator*

$$\hat{\rho}_1 = \text{Sp}_2 \hat{\rho} = \sum_n \langle \psi_n^2 | \psi \rangle \langle \psi | \psi_n^2 \rangle. \quad (12.63)$$

Wir betrachten zunächst einen Produktzustand, o. B. d. A.

$$|\psi\rangle = |\psi_1^1\rangle|\psi_1^2\rangle. \quad (12.64)$$

Dann ist der reduzierte Dichteoperator

$$\hat{\rho}_1 = \sum_n \langle \psi_n^2 | \psi_1^1 \rangle |\psi_1^1\rangle \langle \psi_1^1 | \psi_n^2 \rangle = \sum_n \underbrace{\langle \psi_n^2 | \psi_1^1 \rangle}_{=\delta_{n1}} |\psi_1^1\rangle \underbrace{\langle \psi_1^1 | \psi_n^2 \rangle}_{=\delta_{n1}} = |\psi_1^1\rangle \langle \psi_1^1|. \quad (12.65)$$

Dies ist offensichtlich ein reiner Zustand. Bisher ergab sich nichts Überraschendes. Nun sei $|\psi\rangle$ kein Produktzustand – dann nennt man $|\psi\rangle$ *verschränkt*. Als Beispiel betrachten wir

$$|\psi\rangle = \frac{|\psi_1^1\rangle|\psi_1^2\rangle + |\psi_2^1\rangle|\psi_2^2\rangle}{\sqrt{2}}, \quad (12.66)$$

ein Zustandsvektor dieser Art ist $(|\uparrow\uparrow\rangle + |\downarrow\downarrow\rangle)/\sqrt{2}$. Dann ist

$$\begin{aligned} \hat{\rho}_1 &= \frac{1}{2} \left(\sum_n \langle \psi_n^2 | \psi_1^1 \rangle |\psi_1^1\rangle \langle \psi_1^1 | \psi_n^2 \rangle + \underbrace{\sum_n \langle \psi_n^2 | \psi_1^1 \rangle |\psi_1^1\rangle \langle \psi_2^2 | \psi_n^2 \rangle}_{=0} \right. \\ &\quad \left. + \underbrace{\sum_n \langle \psi_n^2 | \psi_2^1 \rangle |\psi_2^1\rangle \langle \psi_1^1 | \psi_n^2 \rangle}_{=0} + \sum_n \langle \psi_n^2 | \psi_2^1 \rangle |\psi_2^1\rangle \langle \psi_2^1 | \psi_n^2 \rangle \right) \\ &= \frac{1}{2} (|\psi_1^1\rangle \langle \psi_1^1| + |\psi_2^1\rangle \langle \psi_2^1|) \end{aligned} \quad (12.67)$$

$$\begin{aligned} \Rightarrow (\hat{\rho}_1)^2 &= \frac{1}{4} (|\psi_1^1\rangle \underbrace{\langle \psi_1^1 | \psi_1^1 \rangle}_{=1} \langle \psi_1^1| + |\psi_1^1\rangle \underbrace{\langle \psi_1^1 | \psi_2^1 \rangle \langle \psi_2^1 | \psi_1^1 \rangle}_{=0} + |\psi_2^1\rangle \underbrace{\langle \psi_2^1 | \psi_1^1 \rangle \langle \psi_1^1 | \psi_2^1 \rangle}_{=0} + |\psi_2^1\rangle \underbrace{\langle \psi_2^1 | \psi_2^1 \rangle}_{=1} \langle \psi_2^1|) \\ &= \frac{1}{4} (|\psi_1^1\rangle \langle \psi_1^1| + |\psi_2^1\rangle \langle \psi_2^1|) \end{aligned} \quad (12.68)$$

$$\Rightarrow \text{Sp} (\hat{\rho}_1)^2 = \frac{1}{2} < 1. \quad (12.69)$$

Der Zustand des Teilsystems ist also kein reiner Zustand, obwohl das Gesamtsystem in einem reinen Zustand ist! Man kann zeigen, dass das für verschränkte Zustände immer gilt. Wann immer man Teilsysteme betrachten will, benötigt man also im Prinzip den Dichteoperator- (Ensemble-) Formalismus. Nun hatten wir gesehen, dass gemischte Zustände unvollständige Kenntnis anzeigen. Im vorliegenden Fall ist durch die Bildung der Teilspur Information verloren gegangen. Wir können den Zustand im Teilsystem nicht allein im Teilsystem vollständig beschreiben. Im Gegensatz zum vorigen Abschnitt liegt hier aber kein Zustandskollaps vor – der wahre Zustand ist weiterhin $|\psi\rangle \in \mathcal{H}$, also rein.

Kapitel 13

Konsequenzen und Deutungen der Quantenmechanik

Die Quantentheorie und speziell die Quantenmechanik sind zweifellos sehr erfolgreich in der Beschreibung der realen Welt. Es liegt nahe, zu vermuten, dass eine Theorie die Realität nicht zufällig gut beschreibt, sondern weil sie auf einem i. W. korrekten Bild der Welt beruht. Wenn das stimmt, können wir hoffen, durch Analyse der Quantentheorie etwas darüber zu lernen, wie die Welt wirklich beschaffen ist. Dieses Programm erfordert, die Elemente der Theorie mit Elementen der realen Welt in Beziehung zu bringen, d. h. es erfordert die *Interpretation* oder *Deutung* der Theorie.

Nun kommen wir in der Physik nie ganz ohne Deutung aus – sonst würden wir gar keine Beziehung zwischen Theorie und Welt herstellen und damit nicht Physik betreiben, sondern Mathematik. So enthält das Postulat, dass als Messwerte nur Eigenwerte hermitescher Operatoren auftreten, schon Deutung. Das für die praktische Anwendung notwendige Minimum an Deutung bezeichnet man als *Minimalinterpretation*. Obwohl Einführungen in die Quantenmechanik, z. B. auch diese Vorlesung, so formuliert sind, als würde die Theorie über Eigenschaften mikroskopischer Systeme sprechen, ist dies aus Sicht der Minimalinterpretation nicht der Fall. Stattdessen beschreibt sie letztlich nur, was man mit bestimmten makroskopischen Versuchsanordnungen an makroskopischen Messgeräten (repräsentiert durch „Zeigerstellungen“) abliest. Vertreter der Minimalinterpretation haben z. B. nicht unbedingt die Existenz von Atomen oder Elementarteilchen akzeptiert oder ihnen zumindest nicht dieselbe Existenzform wie makroskopischen Objekten zugestanden.

Man kann sich natürlich damit zufrieden geben, dass die Quantenmechanik mit der Minimalinterpretation gute Vorhersagen für Experimente macht, und nicht nach dem physikalischen Gehalt der übrigen Elemente der Theorie fragen. Dann sind z. B. Zustände nur Hilfsgrößen, die bei der Vorhersage von Messergebnissen auftreten. Diese Sichtweise wurde von N. David Mermin mit „shut up and calculate“ charakterisiert (er teilt sie nicht). Dafür spricht der Anspruch der theoretischen Physik, (nur?) experimentell überprüfbare Voraussagen zu machen. Einige zentrale Debatten über die Deutung der Quantenmechanik drehen sich um unterschiedliche Interpretationen, zwischen denen prinzipiell nicht experimentell entschieden werden kann, zumindest aus heutiger Sicht. Demnach sind diese Debatten eher Teil der Naturphilosophie als der Physik. Die Beschäftigung damit erscheint aber auch im Rahmen der Physikausbildung aus mehreren Gründen relevant:

- Wie schon gesagt: Wir vermuten, dass die Quantentheorie die Realität gut beschreibt, weil sie auf einem vernünftigen Bild der Welt beruht. Das legt nahe, den Elementen der Theorie physikalische Bedeutungen zuzuschreiben. Dann können wir aus der Analyse der Theorie etwas über die Welt lernen.
- Ein solches tieferes Verständnis der Theorie ermöglicht ihre Weiterentwicklung.
- Was nach heutiger Kenntnis nicht experimentell zugänglich ist, kann es durch neue Erkenntnisse in Zukunft werden.

In der Deutungsdebatte ist das Verständnis des Messprozesses und des (scheinbaren?) Zustandskollapses von Anfang an eine zentrale, vielleicht die wichtigste, Fragestellung gewesen. Ein Grund hierfür ist die (scheinba-

re?) Dichotomie von einerseits unitärer Zeitentwicklung und andererseits instantanem Zustandskollaps bei der Messung, die von vielen als Mangel der Theorie gesehen wurde.

In diesem Kapitel diskutieren wir zunächst wichtige Ergebnisse, die Konsequenzen der Quantenmechanik aufzeigen, die der Alltagserwartung widersprechen. Danach skizzieren wir einige repräsentative Deutungen speziell des Messprozesses.

13.1 Verschränkung und verborgene Variable

Wir besprechen nun eine Reihe von theoretischen und experimentellen Ergebnissen, die mit der *Verschränkung* von Teilsystemen zu tun haben. Zur Erinnerung: Wenn ein System aus zwei Teilsysteme besteht und die Hilbert-Räume der beiden Teilsystemen und des Gesamtsystems \mathcal{H}_1 , \mathcal{H}_2 und $\mathcal{H} = \mathcal{H}_1 \otimes \mathcal{H}_2$ sind, so nennt man einen Zustand $|\Psi\rangle \in \mathcal{H}$ Produktzustand, wenn er sich schreiben lässt als $|\Psi\rangle = |\psi_1\rangle|\psi_2\rangle$ mit $|\psi_1\rangle \in \mathcal{H}_1$, $|\psi_2\rangle \in \mathcal{H}_2$. Ist $|\Psi\rangle$ kein Produktzustand, so ist $|\Psi\rangle$ *verschränkt*.

13.1.1 Das Einstein-Podolsky-Rosen-Paradoxon

Albert Einstein, Boris Podolsky und Nathan Rosen haben 1935 ein Gedankenexperiment beschrieben, das allgemein als Einstein-Podolsky-Rosen- (EPR-) Paradoxon bekannt ist. Dieser Name wird weiterhin verwendet, obwohl das Gedankenexperiment heute nicht mehr paradox erscheint. Einstein *et al.* wollten zeigen, dass die Quantenmechanik *unvollständig* ist, d. h. dass die Realität Eigenschaften hat, die in der Theorie keine Entsprechung haben. Aus heutiger Sicht sind die Konsequenzen aber schwächer: Das Gedankenexperiment zeigt nur, dass die Quantenmechanik nicht zugleich *vollständig* und *lokal realistisch* sein kann. Hier bedeutet lokal realistisch, dass alle Eigenschaften eines Systems, insbesondere Korrelationen zwischen Teilsystemen, lokal den einzelnen Teilsystemen (z. B. Teilchen) zugeordnet werden können.

Wir geben hier die Formulierung des Paradoxons durch Bohm (1951) wieder (Einstein *et al.* hatten verschränkte Zustände im Hilbert-Raum der räumlichen Bewegung statt im Spin-Hilbert-Raum betrachtet): Diese Formulierung bezieht sich auf ein System aus zwei Teilchen mit Spins der Länge $1/2$, dessen Gesamt-Spin Null ist. Die Teilchen bilden also ein Spin-Singulett beschrieben durch den Spin-Zustand

$$|\psi_0\rangle = \frac{|\uparrow\rangle|\downarrow\rangle - |\downarrow\rangle|\uparrow\rangle}{\sqrt{2}} \equiv \frac{|\uparrow\downarrow\rangle - |\downarrow\uparrow\rangle}{\sqrt{2}}. \quad (13.1)$$

Das Singulett ist offensichtlich verschränkt. Zur Zeit $t = 0$ befinden sich die Teilchen an demselben Ort $\vec{r} = 0$. (Eine solche Situation tritt z. B. auf, wenn ein spinloses Teilchen in zwei Leptonen zerfällt: $\pi^0 \rightarrow e^+ + e^-$ oder $H \rightarrow \tau^+ + \tau^-$.) Für $t > 0$ fliegen die beiden Teilchen auseinander und wir nehmen an, dass sie nicht mehr miteinander wechselwirken. Ihr Spin-Zustand bleibt dann der Singulett-Zustand $|\psi_0\rangle$.

Nun werde die z -Komponente \hat{S}_z^1 des Spins von Teilchen 1 gemessen. Die beiden Terme in Gl. (13.1) sind Eigenzustände zu \hat{S}_z^1 mit den Eigenwerten $\pm\hbar/2$. Wenn wir den Wert $+\hbar/2$ ($-\hbar/2$) messen, muss nach dem von Neumannschen Projektionspostulat der Zustand in $|\uparrow\downarrow\rangle$ ($|\downarrow\uparrow\rangle$) kollabieren. Aber dann wissen wir sofort, was eine folgende Messung von \hat{S}_z^2 an Teilchen 2 ergeben muss.

Die z -Komponente ist nicht ausgezeichnet, da $|\psi_0\rangle$ als Spin-0-Zustand rotationsinvariant ist. In der Tat ist $|\psi_0\rangle$ Eigenzustand zu $(\hat{S}^1 \cdot \hat{n}) \otimes (\hat{S}^2 \cdot \hat{n})$ zum Eigenwert $-\hbar^2/4$ für einen beliebigen Einheitsvektor $\hat{n} = (n_x, n_y, n_z)$. Das bedeutet, dass bezüglich jeder Achse die Komponenten der beiden Spins entgegengesetzt sind. Zum Beweis verwenden wir die Basis $\{|\uparrow\uparrow\rangle, |\uparrow\downarrow\rangle, |\downarrow\uparrow\rangle, |\downarrow\downarrow\rangle\}$. In dieser Basis ist

$$|\psi_0\rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ 1 \\ -1 \\ 0 \end{pmatrix} \quad (13.2)$$

und

$$(\hat{S}^1 \cdot \hat{n}) \otimes (\hat{S}^2 \cdot \hat{n}) = \frac{\hbar^2}{4} \sum_{ij} \sigma_i n_i \otimes \sigma_j n_j$$

$$\begin{aligned}
&= \frac{\hbar^2}{4} \begin{pmatrix} n_z & n_x - in_y \\ n_x + in_y & -n_z \end{pmatrix} \otimes \begin{pmatrix} n_z & n_x - in_y \\ n_x + in_y & -n_z \end{pmatrix} \\
&= \frac{\hbar^2}{4} \begin{pmatrix} n_z^2 & n_z(n_x - in_y) & (n_x - in_y)n_z & (n_x - in_y)^2 \\ n_z(n_x + in_y) & -n_z^2 & n_x^2 + n_y^2 & -(n_x - in_y)n_z \\ (n_x + in_y)n_z & n_x^2 + n_y^2 & -n_z^2 & -n_z(n_x - in_y) \\ (n_x + in_y)^2 & -(n_x + in_y)n_z & -n_z(n_x + in_y) & n_z^2 \end{pmatrix}. \quad (13.3)
\end{aligned}$$

Damit ist

$$(\hat{S}^1 \cdot \hat{n}) \otimes (\hat{S}^2 \cdot \hat{n}) |\psi_0\rangle = \frac{\hbar^2}{4} \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ -n_1^2 - n_2^2 - n_3^2 \\ n_1^2 + n_2^2 + n_3^2 \\ 0 \end{pmatrix} = -\frac{\hbar^2}{4} |\psi_0\rangle. \quad (13.4)$$

□

Durch die Messung einer Spin-Komponente von Teilchen 1 wissen wir also sofort, was eine Messung derselben Spin-Komponente von Teilchen 2 ergeben wird. Nun kann die Messung an Teilchen 1 nach Einstein *et al.* keine Änderung des Zustands von Teilchen 2 bewirken, da die beiden Teilchen nicht mehr miteinander wechselwirken. Schon gar nicht kann Information über die Messung an Teilchen 1 instantan zu Teilchen 2 übertragen werden; dies steht offensichtlich im Widerspruch zur Speziellen Relativitätstheorie. Daher müsste schon vor der Messung an Teilchen 1 bestimmt gewesen sein, was bei der Messung irgendeiner Spin-Komponente an Teilchen 2 herauskommt. Aber eine solche Information ist im Zustand $|\psi_0\rangle$ nicht enthalten, denn $|\psi_0\rangle$ ist ja rotations-symmetrisch. Schlimmer noch, die Quantenmechanik kann gar keine Zustände von Teilchen 2 beschreiben, in denen alle Spin-Komponenten feststehen, weil unterschiedliche Komponenten unverträglich sind und daher kein vollständiger Satz gemeinsamer Eigenzustände existiert. Einstein *et al.* schlossen daraus, dass die Quantenmechanik unvollständig ist. Sie äußerten in ihrem Artikel aber die Vermutung, dass eine grundlegendere Theorie formuliert werden kann, die diese Unvollständigkeit beseitigt. Eigenschaften, die in der Realität vorliegen, aber nicht durch den Zustandsvektor beschrieben werden, nennt man *verborgene Variable*. Einstein *et al.* vermuteten also, dass verborgene Variable existieren.

Niels Bohr reagierte kurz darauf mit der Vermutung, dass man die behauptete Unvollständigkeit nicht durch eine fundamentalere Theorie beheben könnte. Nach Bohr ist die Quantenmechanik vollständig in dem Sinne, dass ein Quantensystem keine weiteren Eigenschaften hat, die nicht im Zustandsvektor enthalten sind. Sein Versuch, das Paradoxon auszuräumen, beruht auf der Forderung, dass die Diskussion von Aussagen der Quantenmechanik immer die Beschreibung der makroskopischen Messapparatur beinhalten muss. Das geht offenbar in die Richtung der Minimalinterpretation.

Erwin Schrödinger reagierte ebenfalls sehr schnell auf Einstein *et al.* Er wies zu Recht darauf hin, dass der Zustand des einzelnen Teilchens 1 oder 2 eben nicht durch einen Zustandsvektor beschrieben werden kann. Wie in Kapitel 12 besprochen, ist ein Teilsystem bei einem verschränkten reinen Zustand in einem *gemischten* Zustand. Daher ist die Beschreibung durch einen Dichteoperator möglich, aber nicht durch einen Hilbert-Raum-Vektor.

Die moderne Diskussion des EPR-Paradoxons bezieht sich eher auf den lokalen Realismus. Sie folgt Schrödingers Argument. Es ist nicht möglich, den Zustand eines verschränkten Systems auf lokale Eigenschaften von Teilsystemen, z. B. Teilchen, zu reduzieren. Die Quantenmechanik enthält im Gegenteil Korrelationen, die nur durch den Zustand des Gesamtsystems ausgedrückt werden, die also nicht lokal sind. Am Beispiel des Spin-Singulets: Der Dichteoperator des Spins von Teilchen 2 ist

$$\hat{\rho}_2 = \text{Sp}_1 |\psi_0\rangle\langle\psi_0| = \frac{1}{2} |\uparrow\rangle\langle\uparrow| + \frac{1}{2} |\downarrow\rangle\langle\downarrow|. \quad (13.5)$$

Dieser Dichteoperator ist weiterhin rotationsinvariant, denn er lautet in *jeder* Spin-Basis

$$\hat{\rho}_2 = \frac{1}{2} \mathbb{1}. \quad (13.6)$$

Er enthält offensichtlich keine Information darüber, was an Teilchen 1 gemessen wird. Die Korrelation von Messergebnissen ist nur in $|\psi_0\rangle$ enthalten. Dies führt aber auch zu einer für Einstein inakzeptablen Folgerung: Wird an Teilchen 1 der Eigenwert von \hat{S}_z^1 gemessen, kollabiert nach von Neumann der Zustand sofort und global.

Beschreibt der Zustand ein reales Element der Welt, so entspricht der Kollaps offenbar einer instantanen Fernwirkung. Beschreibt er dagegen nur unsere Kenntnis, so lässt man offenbar zu, dass das System Eigenschaften hat, die nicht im Zustandsvektor enthalten sind, und wird so wieder auf die Idee von verborgenen Variablen geführt.

13.1.2 Bellsches Theorem und Bellsche Ungleichung

Das EPR-Gedankenexperiment motivierte eine wichtige Weiterentwicklung durch John S. Bell (1964). Bell konnte zeigen, dass *jede* lokal realistische Theorie den Voraussagen der Quantenmechanik widerspricht – selbst wenn die Theorie verborgene Variablen zulässt. Mit anderen Worten, verborgene Variablen können den lokalen Realismus nicht retten.

Bell beginnt mit einem modifizierten Gedankenexperiment: Es werden wieder Spin-Komponenten an zwei Spin-1/2-Teilchen im Spin-Singulett-Zustand

$$|\psi_0\rangle = \frac{|\uparrow\downarrow\rangle - |\downarrow\uparrow\rangle}{\sqrt{2}} \quad (13.7)$$

gemessen, aber nun nicht unbedingt dieselbe Komponente. Stattdessen wird am ersten Teilchen $\hat{S}^1 \cdot \hat{n}_1$ und am zweiten $\hat{S}^2 \cdot \hat{n}_2$ gemessen, wobei \hat{n}_1 und \hat{n}_2 Einheitsvektoren sind. Ist $\hat{n}_2 = \hat{n}_1$, so erhalten wir den EPR-Fall. Die zentrale Größe ist

$$P(\hat{n}_1, \hat{n}_2) := \left\langle \left(\hat{S}^1 \cdot \hat{n}_1 \right) \otimes \left(\hat{S}^2 \cdot \hat{n}_2 \right) \right\rangle. \quad (13.8)$$

(Man könnte hier das Tensorprodukt-Symbol „ \otimes “ weglassen, ohne Verständnisprobleme zu erzeugen. Es drückt aus, dass \hat{S}^1 auf \mathcal{H}_1 und \hat{S}^2 auf \mathcal{H}_2 wirkt.) Im Spin-Singulett sind gleiche Komponenten der beiden Spins entgegengesetzt gerichtet, also gilt im EPR-Fall

$$P(\hat{n}_1, \hat{n}_1) = -\frac{\hbar^2}{4}. \quad (13.9)$$

Um $P(\hat{n}_1, \hat{n}_2)$ für beliebige Einheitsvektoren auszurechnen, beachten wir zunächst

$$\begin{aligned} \langle \hat{S}_i^1 \otimes \hat{S}_j^2 \rangle &= \frac{1}{2} (\langle \uparrow\downarrow | - \langle \downarrow\uparrow |) \hat{S}_i^1 \otimes \hat{S}_j^2 (|\uparrow\downarrow\rangle - |\downarrow\uparrow\rangle) \\ &= \frac{1}{2} \left(\langle \uparrow | \hat{S}_i | \uparrow \rangle \langle \downarrow | \hat{S}_j | \downarrow \rangle - \langle \uparrow | \hat{S}_i | \downarrow \rangle \langle \downarrow | \hat{S}_j | \uparrow \rangle - \langle \downarrow | \hat{S}_i | \uparrow \rangle \langle \uparrow | \hat{S}_j | \downarrow \rangle + \langle \downarrow | \hat{S}_i | \downarrow \rangle \langle \uparrow | \hat{S}_j | \uparrow \rangle \right) \\ &= \frac{\hbar^2}{8} [\delta_{iz}(-1)\delta_{jz} - (\delta_{ix} - i\delta_{iy})(\delta_{jx} + i\delta_{jy}) - (\delta_{ix} + i\delta_{iy})(\delta_{jx} - i\delta_{jy}) - \delta_{iz}\delta_{jz}] \\ &= \frac{\hbar^2}{8} (-\delta_{iz}\delta_{jz} - \delta_{ix}\delta_{jx} - \cancel{i\delta_{ix}\delta_{jy}} + \cancel{i\delta_{iy}\delta_{jx}} - \delta_{iy}\delta_{jy} - \delta_{ix}\delta_{jx} + \cancel{i\delta_{ix}\delta_{jy}} - \cancel{i\delta_{iy}\delta_{jx}} - \delta_{iy}\delta_{jy} - \delta_{iz}\delta_{jz}) \\ &= -\frac{\hbar^2}{4} (\delta_{ix}\delta_{jx} + \delta_{iy}\delta_{jy} + \delta_{iz}\delta_{jz}) \\ &= -\frac{\hbar^2}{4} \delta_{ij}. \end{aligned} \quad (13.10)$$

Damit ergibt die Standard-Quantenmechanik

$$\begin{aligned} P(\hat{n}_1, \hat{n}_2) &= \left\langle \left(\hat{S}^1 \cdot \hat{n}_1 \right) \otimes \left(\hat{S}^2 \cdot \hat{n}_2 \right) \right\rangle = \sum_{ij} \left\langle \hat{S}_i^1 n_{1i} \otimes \hat{S}_j^2 n_{2j} \right\rangle \\ &= -\frac{\hbar^2}{4} \sum_{ij} \delta_{ij} n_{1i} n_{2j} = -\frac{\hbar^2}{4} \hat{n}_1 \cdot \hat{n}_2. \end{aligned} \quad (13.11)$$

Nun zu Bells Theorem: Die Voraussagen der Standard-Quantenmechanik sind unvereinbar mit einer lokal realistischen Theorie, sogar wenn diese die Existenz verborgener Variabler annimmt.

Beweis: Der reale Zustand des Spin-Paars werde durch einen Satz verborgener Variabler λ und den Zustandsvektor $|\psi_0\rangle$ vollständig beschrieben. Wir nehmen an, dass das Ergebnis der Messung an Teilchen 1 unabhängig

von \hat{n}_2 ist, d. h. von der Einstellung des Detektors für Teilchen 2, und umgekehrt. Das ist die Annahme lokaler Realität. (Sie wirkt vernünftig, da \hat{n}_2 ja zu einem Raumzeitpunkt gewählt werden kann, der vom Raumzeitpunkt der Messung an Teilchen 1 raumartig getrennt ist, so dass keine Information über \hat{n}_2 zum ersten Teilchen und Detektor kommen kann.) Seien s_1 und s_2 die Messergebnisse der Spin-Komponenten $\hat{S}^1 \cdot \hat{n}_1$ und $\hat{S}^2 \cdot \hat{n}_2$ der beiden Teilchen. Die Vollständigkeit der Beschreibung impliziert, dass diese Messwerte eindeutig vom Zustand des Systems beschrieben werden:

$$s_1 = s_1(\hat{n}_1, \lambda), \quad (13.12)$$

$$s_2 = s_2(\hat{n}_2, \lambda). \quad (13.13)$$

Die Abhängigkeit von $|\psi_0\rangle$ haben wir unterdrückt. Das Produkt der Messwerte ist natürlich

$$P(\hat{n}_1, \hat{n}_2, \lambda) = s_1(\hat{n}_1, \lambda) s_2(\hat{n}_2, \lambda). \quad (13.14)$$

Eine lokal realistische Theorie mit verborgenen Variablen muss die beobachtete Streuung von P durch eine Zufallsverteilung von λ erklären, da \hat{n}_1 und \hat{n}_2 ja fest vorgegeben sind. Der Erwartungswert von P über viele Messungen ist dann

$$P(\hat{n}_1, \hat{n}_2) = \overline{P(\hat{n}_1, \hat{n}_2, \lambda)} = \int d\lambda w(\lambda) s_1(\hat{n}_1, \lambda) s_2(\hat{n}_2, \lambda), \quad (13.15)$$

wobei $w(\lambda)$ die Verteilungsfunktion von λ ist. Drehimpulserhaltung erfordert, dass für gleichen Achsen $\hat{n}_2 = \hat{n}_1$ gilt

$$s_2(\hat{n}_1, \lambda) = -s_1(\hat{n}_1, \lambda) \quad \forall \hat{n}_1, \lambda. \quad (13.16)$$

Damit können wir schreiben

$$P(\hat{n}_1, \hat{n}_2) = - \int d\lambda w(\lambda) s_1(\hat{n}_1, \lambda) s_1(\hat{n}_2, \lambda). \quad (13.17)$$

Sei nun \hat{n}_3 ein weiterer Einheitsvektor. Dann ist

$$\begin{aligned} P(\hat{n}_1, \hat{n}_2) - P(\hat{n}_1, \hat{n}_3) &= - \int d\lambda w(\lambda) [s_1(\hat{n}_1, \lambda) s_1(\hat{n}_2, \lambda) - s_1(\hat{n}_1, \lambda) s_1(\hat{n}_3, \lambda)] \\ &= - \int d\lambda w(\lambda) s_1(\hat{n}_1, \lambda) [s_1(\hat{n}_2, \lambda) - s_1(\hat{n}_3, \lambda)]. \end{aligned} \quad (13.18)$$

Die gesuchte lokal realistische Theorie soll dieselben Voraussagen machen, wie die Standard-Quantenmechanik. Also darf s_1 nur die Werte $\pm\hbar/2$ annehmen. Es ist also $s_1^2 = \hbar^2/4$ und wir können schreiben

$$\begin{aligned} \dots &= - \frac{4}{\hbar^2} \int d\lambda w(\lambda) s_1(\hat{n}_1, \lambda) s_1(\hat{n}_2, \lambda)^2 [s_1(\hat{n}_2, \lambda) - s_1(\hat{n}_3, \lambda)] \\ &= - \frac{4}{\hbar^2} \int d\lambda w(\lambda) s_1(\hat{n}_1, \lambda) s_1(\hat{n}_2, \lambda) \left[\frac{\hbar^2}{4} - s_1(\hat{n}_2, \lambda) s_1(\hat{n}_3, \lambda) \right] \\ &= - \int d\lambda w(\lambda) s_1(\hat{n}_1, \lambda) s_1(\hat{n}_2, \lambda) \left[1 - \frac{4}{\hbar^2} s_1(\hat{n}_2, \lambda) s_1(\hat{n}_3, \lambda) \right]. \end{aligned} \quad (13.19)$$

Nun ist

$$s_1(\hat{n}_1, \lambda) s_1(\hat{n}_2, \lambda) = \pm \frac{\hbar^2}{4} \quad (13.20)$$

und natürlich auch

$$s_1(\hat{n}_2, \lambda) s_1(\hat{n}_3, \lambda) = \pm \frac{\hbar^2}{4} \quad (13.21)$$

$$\Rightarrow 1 - \frac{4}{\hbar^2} s_1(\hat{n}_2, \lambda) s_1(\hat{n}_3, \lambda) \in \{0, 2\} \quad (13.22)$$

$$\Rightarrow w(\lambda) \left[1 - \frac{4}{\hbar^2} s_1(\hat{n}_2, \lambda) s_1(\hat{n}_3, \lambda) \right] \geq 0. \quad (13.23)$$

Daraus folgt

$$\begin{aligned}
|P(\hat{n}_1, \hat{n}_2) - P(\hat{n}_1, \hat{n}_3)| &= \left| \int d\lambda s_1(\hat{n}_1, \lambda) s_1(\hat{n}_2, \lambda) w(\lambda) \left[1 - \frac{4}{\hbar^2} s_1(\hat{n}_2, \lambda) s_1(\hat{n}_3, \lambda) \right] \right| \\
&\downarrow \quad \text{mit } s_1(\hat{n}_1, \lambda) s_1(\hat{n}_2, \lambda) = \pm \frac{\hbar^2}{4} \\
&\leq \frac{\hbar^2}{4} \int d\lambda w(\lambda) \left[1 - \frac{4}{\hbar^2} s_1(\hat{n}_2, \lambda) s_1(\hat{n}_3, \lambda) \right] \\
&= \int d\lambda w(\lambda) \left[\frac{\hbar^2}{4} - s_1(\hat{n}_2, \lambda) s_1(\hat{n}_3, \lambda) \right] \\
&= \frac{\hbar^2}{4} + P(\hat{n}_2, \hat{n}_3).
\end{aligned} \tag{13.24}$$

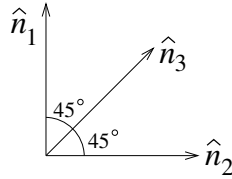
Wir haben damit die *Bellsche Ungleichung*

$$|P(\hat{n}_1, \hat{n}_2) - P(\hat{n}_1, \hat{n}_3)| \leq \frac{\hbar^2}{4} + P(\hat{n}_2, \hat{n}_3) \tag{13.25}$$

erhalten.

Betrachten wir nun als Beispiel die drei (koplanaren) Einheitsvektoren

$$\hat{n}_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \quad \hat{n}_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \quad \hat{n}_3 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}. \tag{13.26}$$



Die Standard-Quantenmechanik sagt voraus

$$P(\hat{n}_1, \hat{n}_2) = 0, \tag{13.27}$$

$$P(\hat{n}_1, \hat{n}_3) = P(\hat{n}_2, \hat{n}_3) = -\frac{1}{\sqrt{2}} \frac{\hbar^2}{4}. \tag{13.28}$$

Einsetzen in die Bellsche Ungleichung ergibt

$$\frac{1}{\sqrt{2}} \leq 1 - \frac{1}{\sqrt{2}} \tag{13.29}$$

$$\Rightarrow 1 \geq \frac{2}{\sqrt{2}} = \sqrt{2}. \tag{13.30}$$

Das ist offensichtlich falsch. Es ergibt sich also ein Widerspruch. \square

Das Ergebnis sagt aus, dass man auch durch Annahme beliebiger verborgener Variabler – wir haben nichts darüber angenommen, aus welcher Menge λ kommt oder wie die Verteilungsfunktion $w(\lambda)$ aussieht – *keine* lokal realistische Theorie konstruieren kann, die dieselben Voraussagen macht, wie die Standard-Quantenmechanik. Die Hoffnung von Einstein *et al.*, die vermeintlich unvollständige Quantenmechanik könne durch eine vollständige, lokal realistische Theorie ersetzt werden, die dieselbe Physik beschreibt, hat sich also zerschlagen. Es ist wichtig, sich klarzumachen, dass Bells Theorem *nicht* die Existenz verborgener Variabler verbietet. Es kann durchaus verborgene Variablen geben, nur kann man durch ihre Einführung die Theorie nicht lokal machen.

Soweit sagt Bells Theorem natürlich nichts darüber aus, ob die Welt überhaupt durch die Standard-Quantenmechanik beschrieben wird. Sie könnte ja in Wirklichkeit durch eine lokal realistische Theorie beschrieben werden, die dann eben nicht dieselben Voraussagen macht. Aber die Bellsche Ungleichung stellt auch hier

einen großen Fortschritt dar. Man kann ja das beschriebene Experiment wirklich ausführen und damit $P(\hat{n}_1, \hat{n}_2)$, $P(\hat{n}_1, \hat{n}_3)$ und $P(\hat{n}_2, \hat{n}_3)$ experimentell bestimmen. Dies wurde vielfach getan und die Ergebnisse stimmen mit der Standard-Quantenmechanik überein und verletzen somit die Bellsche Ungleichung. (Die Interpretation dieser Experimente lässt manchmal Schlupflöcher, wie sie unter – meist wenig plausiblen – Annahmen doch noch mit dem lokalen Realismus in Einklang gebracht werden können. Großer Aufwand ist sowohl experimentell als auch theoretisch getrieben worden, um diese Schlupflöcher zu schließen.) Es ist sehr naheliegend, aus dem Versagen jeglicher lokal realistischer Beschreibung zu schließen, dass die Welt selbst fundamental nichtlokal ist.

Die experimentelle Bestätigung der notwendig nichtlokalen Beschreibung, z. B. durch den verschränkten Zustand $|\psi_0\rangle$, provoziert erneut Einsteins Frage: Führt der Kollaps des Zustands bei der Messung an Teilchen 1 zu einer realen Fernwirkung? Die Antwort ist *nein*, wenn damit eine *kausale* Wirkung gemeint ist. Man kann zeigen, dass die Nichtlokalität nicht für die instantane oder auch überlichtschnelle Übertragung von Information genutzt werden kann. Wesentlich dafür ist, dass Experimentatorin Alice mit Teilchen 1 zwar die Richtung \hat{n}_1 wählen kann, aber nicht bestimmen, welcher Messwert $s_1(\hat{n}_1)$ herauskommt. Experimentator Bob kann aber durch keine Messung an Teilchen 2 feststellen, welche Richtung \hat{n}_1 Alice gewählt hat. Die Nichtlokalität zeigt sich erst, wenn Alice und Bob später ihre Datensätze für viele Messungen an Teilchen-Paaren zusammenführen und finden, dass die Bellsche Ungleichung verletzt ist.

13.2 Schrödingers Katze

Das bekannteste Gedankenexperiment zur Quantenmechanik stammt von Erwin Schrödinger (1935). Es beruht auf Schrödingers Beschäftigung mit dem EPR-Paradoxon. Schrödinger wollte mit seinem Gedankenexperiment die scheinbar paradoxen Konsequenzen aufzeigen, die sich ergeben, wenn ein mikroskopisches und ein makroskopisches Objekt in einem verschränkten Zustand sind. Letztlich kann man die paradoxen Konsequenzen der durch die Quantenmechanik vorhergesagten unitären (es reicht auch linearen) Zeitentwicklung und dem daraus folgenden Superpositionsprinzip zuschreiben.

Schrödinger betrachtete einen geschlossenen Kasten mit einer Katze und einer Killermaschine. Die Maschine enthält ein Atom mit radioaktivem Kern. Wenn der Kern zerfällt, wird ein Mechanismus ausgelöst, der die Katze tötet. Schrödinger ging davon aus, dass die beobachtete (Exponential-) Verteilung von Zerfallszeitpunkten intrinsisch quantenmechanischer Natur ist. Sie beruht dann nicht auf unserer Unkenntnis über einen hypothetischen inneren Zustand des Kerns (d. h. es läuft keine verborgene Uhr im Kern). Der Zustand des, evtl. zerfallenen, Kerns zur Zeit t lässt sich dann schreiben als

$$|\text{Kern}(t)\rangle = \sqrt{e^{-t/\tau}} |\text{heil}(t)\rangle + \sqrt{1 - e^{-t/\tau}} |\text{zerfallen}(t)\rangle. \quad (13.31)$$

Hier ist τ die mittlere Zerfallszeit, $|\text{heil}(t)\rangle$ der Zustand des nicht zerfallenen Kerns und $|\text{zerfallen}(t)\rangle$ der dazu orthonormierte (und komplizierte!) Zustand der Zerfallsprodukte des zerfallenen Kerns. Wir lassen jetzt das Zeitargument der Zustände weg. Der Vektor $|\text{Kern}\rangle$ ist normiert:

$$\langle \text{Kern} | \text{Kern} \rangle = e^{-t/\tau} \underbrace{\langle \text{heil} | \text{heil} \rangle}_{=1} + (1 - e^{-t/\tau}) \underbrace{\langle \text{zerfallen} | \text{zerfallen} \rangle}_{=1} = 1. \quad (13.32)$$

Da der Apparat die Katze kausal an den Kern koppelt, ist der Zustand des Gesamtsystems im Kasten von derselben Form:

$$|\text{Kern} + \text{Katze}\rangle = \sqrt{e^{-t/\tau}} |\text{heil}\rangle |\text{lebendig}\rangle + \sqrt{1 - e^{-t/\tau}} |\text{zerfallen}\rangle |\text{tot}\rangle. \quad (13.33)$$

Wenn ein Beobachter den Kasten öffnet, findet er entweder eine lebende oder eine tote Katze vor. Er findet nie eine Superposition von lebender und toter Katze (wie würde das überhaupt aussehen?). Dies ist die Motivation, einen Zustandskollaps zu postulieren: Nach dem Projektionspostulat kollabiert bei Messungen der Zustand instantan. Im vorliegenden Fall wird die Beobachtung des Zustandes der Katze als Messung aufgefasst. Die Wahrscheinlichkeit dafür, eine lebende Katze zu finden, ist nach der Bornschen Regel

$$W_{\text{lebendig}} = |\langle \text{lebendig} | \langle \text{heil} | |\text{Kern} + \text{Katze}\rangle|^2 = e^{-t/\tau}. \quad (13.34)$$

Vor der Beobachtung ist das makroskopische System im Kasten also in einer Superposition von deutlich verschiedenen Zuständen – sie enthalten eine lebende oder eine tote Katze. Dies hielt Schrödinger für absurd. Er sah

sein Gedankenexperiment als Argument gegen die Kopenhagener Deutung der Quantentheorie, auf die wir noch kommen werden. Aber das Paradoxon greift nicht wirklich: Die Kopenhagener Deutung neigt dazu, den Zustand als Ausdruck der Kenntnis eines Beobachters aufzufassen, und dann zeigt die Superposition nur an, dass der Beobachter eben nicht weiß, ob die Katze lebendig oder tot ist.

Man bezeichnet allgemein Superpositionen makroskopisch unterschiedlicher Zustände als *Schrödinger-Katzen-Zustände*. Dabei ist es keineswegs einfach, zu definieren, was mit „makroskopisch unterschiedlich“ gemeint ist, und es gibt verschiedene Ansichten dazu.

Schrödingers Katze ist ein Prüfstein für jeden Versuch der Interpretation der Quantenmechanik. Die Minimalinterpretation ist allerdings damit kaum angreifbar. In ihr wird ja gerade dem unbeobachteten Zustand der Katze im geschlossenen Kasten keine physikalische Bedeutung beigemessen. Ein Vertreter von „shut up and calculate“ würde also weder versuchen, der makroskopischen Superposition einen physikalischen Sinn zu geben, noch sie durch Änderung der Theorie zu vermeiden, sondern würde die Frage als gegenstandslos und damit unsinnig zurückweisen.

Das Gedankenexperiment wirft auch die Frage auf, wann genau der Zustand kollabiert. Gemäß dem Projektionspostulat und der Kopenhagener Deutung nach Heisenberg kollabiert der Zustand, wenn der Beobachter den Kasten öffnet und den Zustand der Katze überprüft. Aber ist nicht die Katze auch ein Beobachter? Dann würde der Zustand kollabieren, sobald der Kern zerfällt – hier könnten wir auch den Fall betrachten, dass die Apparatur den Kernzerfall für die Katze sichtbar anzeigt, diese aber nicht tötet. Die Frage nach der Rolle des Beobachters steht auch hinter dem folgenden Gedankenexperiment.

13.2.1 Wigners Freund: Bewusstsein erzeugt Kollaps

Eine Fortsetzung des Gedankenexperiments über Schrödingers Katze stammt von E. Wigner (1961): Dabei führt ein Freund von Wigner das Experiment mit der Katze aus, während Wigner nicht im Raum ist. Wigners Freund öffnet den Kasten und überprüft den Zustand der Katze. Nach der Kopenhagener Deutung ist aus Sicht von Wigner, der all dies nicht beobachtet, das System Kern+Katze+Freund nun in einem Zustand der Form

$$|\text{Kern}+\text{Katze}+\text{Freund}\rangle = \sqrt{e^{-t/\tau}} |\text{heil}\rangle |\text{lebendig}\rangle |\text{glücklich}\rangle + \sqrt{1 - e^{-t/\tau}} |\text{zerfallen}\rangle |\text{tot}\rangle |\text{traurig}\rangle. \quad (13.35)$$

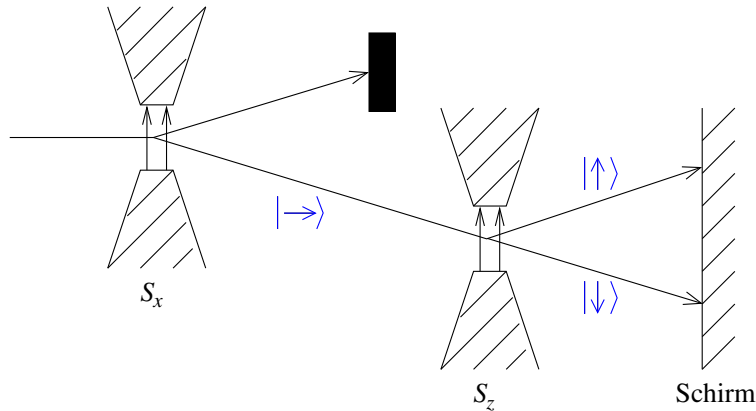
Dieser Zustand kollabiert demnach erst, wenn Wigner den Raum wieder betritt und seinen Freund oder die Katze untersucht. Das würde bedeuten, dass der Kollaps für Wigner und seinen Freund und evtl. die Katze zu unterschiedlichen Zeitpunkten stattfindet. Die Kopenhagener Deutung nach Heisenberg interpretiert den Zustand als Ausdruck des Wissens über ein System, nicht als Eigenschaft des Systems selbst. In dieser Sichtweise ist es natürlich kein Problem, wenn der Kollaps für die beteiligten Beobachter zu unterschiedlichen Zeitpunkten stattfindet, sondern dies ist zu erwarten.

Will man dem Kollaps aber realen Charakter zuschreiben, also sagen, dass sich über die Kenntnisse von Beobachtern hinaus etwas in der Welt ändert, sind die unterschiedlichen Kollapszeitpunkte nur schwer zu verteidigen. Wigner vertrat eine solche realistische Interpretation des Kollapses. Er war der Auffassung, dass ein *bewusster* Beobachter notwendig ist, um einen Zustand zu kollabieren. Dann findet der Kollaps eindeutig genau dann statt, wenn Wigners Freund den Kasten öffnet, und nicht, wenn Wigner den Raum betritt. Wigner billigte der Katze also kein Bewusstsein zu. Aber hier liegt ein Problem: Bewusstsein ist eine graduelle Eigenschaft und kann somit nicht leicht eine zweiwertige Konsequenz erklären, nämlich Kollaps oder kein Kollaps. Ab welchem Alter kann z. B. ein Kind einen Zustand kollabieren? Kann ein Schimpanse oder ein Delphin einen Zustand kollabieren? Wigners dualistische – und esoterisch anmutende – Auffassung von Materie vs. Bewusstsein ist auch kaum zu halten.

13.2.2 Stern-Gerlach-Experimente als Prototyp des Messprozesses

Varianten des in 2.7 besprochenen Stern-Gerlach-Versuchs sind für die Diskussion des Messprozesses besonders geeignet, weil sie eine minimale Realisierung darstellen, die dennoch alle relevanten Aspekte enthält. Wir beschreiben zunächst eine Variante des Versuchs, die etwas vom historischen Experiment abweicht. Die einlaufenden Teilchen sollen nämlich in einem wohldefinierten Spin-Zustand $|\psi\rangle$ sein, während sie bei Stern und Gerlach aus einem Ofen kamen und daher zufällig verteilte Spin-Zustände hatten. Beim historischen Experiment kann der einlaufende Zustand also korrekt nur mittels eines Dichteoperators beschrieben werden. Ein bestimmter (reiner)

Spin-Zustand kann im Prinzip leicht durch eine vorgeschaltete Stern-Gerlach-Apparatur realisiert werden, die den Strahl abhängig von einer einstellbaren Spin-Komponente $\hat{S} \cdot \hat{n}$ aufspaltet (\hat{n} ist ein Einheitsvektor). Hier ist dies für das Beispiel \hat{S}_x skizziert:



Was geschieht nun, wenn die zweite Stern-Gerlach-Apparatur die Komponente \hat{S}_z misst? Zunächst können wir den einlaufenden Zustand schreiben als

$$|\psi\rangle = |\uparrow\rangle\langle\uparrow|\psi\rangle + |\downarrow\rangle\langle\downarrow|\psi\rangle = \langle\uparrow|\psi\rangle |\uparrow\rangle + \langle\downarrow|\psi\rangle |\downarrow\rangle. \quad (13.36)$$

Dies ist eine mathematische Operation, die natürlich immer möglich ist, unabhängig davon, ob und was wir messen. Zur Abkürzung schreiben wir

$$|\psi\rangle = c_\uparrow |\uparrow\rangle + c_\downarrow |\downarrow\rangle. \quad (13.37)$$

Wir verfolgen ein Teilchen jetzt im mitbewegten Bezugssystem. Zunächst wirkt kein Magnetfeld und der Hamilton-Operator ist $\hat{H} = \hat{p}^2/2m$. Im zweiten Stern-Gerlach-Apparat wirkt ein inhomogenes Magnetfeld, das eine Kraft auf das Teilchen ausübt, die für die Eigenzustände von \hat{S}_z entgegengesetzt ist. Eine räumlich konstante Kraft entspricht einem linearen Potential. Im Bereich der zweiten Apparatur ist der Hamiltonian daher

$$\hat{H} = \frac{\hat{p}^2}{2m} + b x \hat{S}_z \quad (13.38)$$

mit einer Konstanten b . An dieser Stelle erweist es sich als nützlich, dass wir $|\psi\rangle$ nach Eigenzuständen von \hat{S}_z entwickelt haben und nicht in irgendeine andere Basis. \hat{H} ist nämlich in der Eigenbasis von \hat{S}_z (block-) diagonal:

$$\hat{H} = \begin{pmatrix} \frac{\hat{p}^2}{2m} + \frac{b\hbar}{2} x & 0 \\ 0 & \frac{\hat{p}^2}{2m} - \frac{b\hbar}{2} x \end{pmatrix}. \quad (13.39)$$

Beim Durchlaufen des inhomogenen Magnetfeldes erhalten die Teilchen im Spin-Zustand $|\uparrow\rangle$ bzw. $|\downarrow\rangle$ entgegengesetzte Impulse senkrecht zum einlaufenden Strahl. Hinter der Apparatur können wir den Zustand also schreiben als

$$|\psi\rangle = c_\uparrow |\uparrow, p_z > 0\rangle + c_\downarrow |\downarrow, p_z < 0\rangle. \quad (13.40)$$

Die Zustände $|\uparrow\rangle, |\downarrow\rangle$ im Spin-Faktorraum sind nun mit den Zuständen $|p_z > 0\rangle, |p_z < 0\rangle$ im Faktorraum der räumlichen Bewegung verschränkt. Dies würde man noch nicht als Schrödinger-Katzen-Zustand bezeichnen, weil $|\uparrow, p_z > 0\rangle$ und $|\downarrow, p_z < 0\rangle$ mikroskopische Teilchen an näherungsweise demselben Ort beschreiben.

Aber nun müssen wir nur abwarten. Wegen der entgegengesetzten Transversalimpulse laufen die Teilchen abhängig von $m = \uparrow, \downarrow$ räumlich auseinander und treffen an makroskopisch getrennten Orten auf den Schirm oder Detektor. Der Zustand hat dann die Form

$$|\psi\rangle = c_\uparrow |\uparrow, z = z_0\rangle + c_\downarrow |\downarrow, z = -z_0\rangle. \quad (13.41)$$

Ob das jetzt ein Schrödinger-Katzen-Zustand ist, hängt von der genauen Definition ab. Im nächsten Schritt wird das Auftreffen des Teilchens detektiert. Und schließlich erhält ein Beobachter die Nachricht, wo das Teilchen aufgetroffen ist. (Der Detektor spielt hier keine fundamentale Rolle, denn es ist für einen menschlichen Beobachter zwar schwierig, aber im Prinzip möglich, ein einzelnes Teilchen sinnlich wahrzunehmen.)

Ist die Zeitentwicklung weiterhin unitär, so ist der gemeinsame Zustand des Teilchens und der Beobachterin Alice nun von der Form

$$|\Psi\rangle = c_{\uparrow} |\uparrow, A_{\uparrow}\rangle + c_{\downarrow} |\downarrow, A_{\downarrow}\rangle, \quad (13.42)$$

wobei A_m bedeutet „Alice hat Spin m gemessen“. Das wäre nach jeder sinnvollen Definition ein Schrödinger-Katzen-Zustand.

Das Projektionspostulat sagt dagegen, dass der Zustand bei der Messung kollabiert. Dann ist der Zustand nach der Messung *entweder* $|\uparrow, A_{\uparrow}\rangle$ *oder* $|\downarrow, A_{\downarrow}\rangle$ mit Wahrscheinlichkeiten $|c_m|^2 = |\langle m|\psi\rangle|^2$. Wann findet dieser Kollaps statt? Oder, in anderen Worten, was genau macht hier die Messung aus? Die Minimalinterpretation macht dazu keine Aussage. Interpretiert man den Zustand als Ausdruck des Wissens über ein System, erfolgt der Kollaps offenbar dann, wenn das Messergebnis Alice bekannt wird. Dasselbe folgt aus Wigners Vermutung, wonach ein bewusster Beobachter notwendig ist.

13.3 Das No-Cloning-Theorem

Eine für die experimentelle Untersuchung von Zuständen und für auf Quantenmechanik beruhenden Technologien wesentliche Einschränkung besteht darin, dass Zustände von Quantensystemen prinzipiell nicht perfekt kopiert werden können. Dies ist die Aussage des *No-Cloning-Theorems* (Wootters und Zurek, 1982), das wir nun formulieren und beweisen werden: Wir betrachten zwei gleichartige Teilsysteme, insbesondere mit Hilbert-Räumen gleicher Dimension. Anfangs sei das Gesamtsystem in einem beliebigen Produktzustand $|\psi\rangle_1 |\nu\rangle_2$. Behauptung: Es existiert kein unitärer Prozess, der für alle $|\psi\rangle_1$ diesen Zustand in $|\psi\rangle_1 |\psi\rangle_2$ überführt, d. h. der den Zustand $|\psi\rangle_1$ von Teilsystem 1 unverändert lässt und in Teilsystem 2 kopiert.

Beweis: Angenommen, es gäbe einen solchen Prozess. Seien $|\psi\rangle_1$ und $|\phi\rangle_1$ zwei beliebige Zustände von Teilsystem 1. Alle vorkommenden Zustände seien normiert. Die Zeitentwicklung werde durch den unitären Zeitentwicklungsoperator \hat{U} beschrieben. Dann ist

$$\hat{U} |\psi\rangle_1 |\nu\rangle_2 = |\psi\rangle_1 |\psi\rangle_2, \quad (13.43)$$

$$\hat{U} |\phi\rangle_1 |\nu\rangle_2 = |\phi\rangle_1 |\phi\rangle_2. \quad (13.44)$$

Es folgt

$${}_2\langle\phi| {}_1\langle\phi| |\psi\rangle_1 |\psi\rangle_2 = {}_2\langle\nu| {}_1\langle\phi| \hat{U}^\dagger \hat{U} |\psi\rangle_1 |\nu\rangle_2 = {}_2\langle\nu| {}_1\langle\phi| |\psi\rangle_1 |\nu\rangle_2 = \underbrace{{}_2\langle\nu| \nu\rangle}_1 {}_1\langle\phi| \psi\rangle_1 = {}_1\langle\phi| \psi\rangle_1. \quad (13.45)$$

Andererseits gilt

$${}_2\langle\phi| {}_1\langle\phi| |\psi\rangle_1 |\psi\rangle_2 = {}_2\langle\phi| \psi\rangle_2 {}_1\langle\phi| \psi\rangle_1. \quad (13.46)$$

Die Teilraum-Indizes 1, 2 sind für die Berechnung der Skalarprodukte irrelevant, da die Zustände $|\psi\rangle$ und $|\phi\rangle$ in Teilraum 2 Kopien der Zustände in Teilraum 1 sein und sich also gleich verhalten sollen, also folgt aus Glg. (13.45) und (13.46)

$$(\langle\phi|\psi\rangle)^2 = \langle\phi|\psi\rangle. \quad (13.47)$$

Dafür gibt es nur zwei Lösungen:

$$\langle\phi|\psi\rangle \in \{0, 1\}. \quad (13.48)$$

Also sind die Zustände $|\psi\rangle$ und $|\phi\rangle$ entweder identisch oder orthogonal. Die Voraussetzung war aber, dass $|\psi\rangle$ und $|\phi\rangle$ beliebig sind, sie können insbesondere $\langle\phi|\psi\rangle \notin \{0, 1\}$ erfüllen. Widerspruch!

Der Kontext des No-Cloning-Theorems ist eine Sichtweise, wonach die gesamte Zeitentwicklung in der Quantenmechanik unitär ist, also kein Zustandskollaps stattfindet – wir kommen darauf zurück. Aber auch, wenn echter Kollaps zugelassen wird, bleibt die Aussage richtig. Das sieht man z. B. daran, dass Kollaps, also Projektion von

Zuständen, notwendigerweise gewisse Basen auszeichnet, nämlich die Eigenbasen der Projektionsoperatoren. Daher kann damit erst recht kein Prozess konstruiert werden, der alle möglichen Zustände kopieren kann.

Noch zwei Bemerkungen:

- Es ist möglich, den Zustand eines Teilsystems in ein anderes zu kopieren. Das No-Cloning-Theorem erzwingt nur, dass dabei der Zustand des ursprünglichen Teilsystems geändert wird. Es ist also insbesondere im Prinzip möglich, einen Quantenzustand über makroskopische Entfernungen zu verschicken, man spricht dann von *Teleportation* des Zustands.
- Es ist auch möglich, beliebig viele Teilsysteme in identischen Zuständen zu präparieren, z. B. mehrere Spins im Zustand $|\uparrow\rangle$. Das Theorem verbietet nur die Kopie vorher unbekannter Zustände.

Das Theorem hat wichtige Konsequenzen. Zunächst zeigt es, dass wir die Unmöglichkeit der gleichzeitigen Messung unverträglicher Observabler nicht umgehen können: Wäre es möglich, Zustände zu kopieren, so könnten wir z. B. den Zustand $|\psi\rangle$ eines Spins der Länge $1/2$ kopieren und dann \hat{S}_x und \hat{S}_y messen, ohne dass sich die Messungen gegenseitig stören. Das Argument lässt sich umdrehen: Es ist unmöglich, den Zustand $|\psi\rangle$ eines Quantensystems vollständig zu messen. Denn wäre es möglich, könnten wir danach beliebig viele Subsysteme in diesem Zustand $|\psi\rangle$ präparieren, was äquivalent zur Anfertigung von Kopien ist und damit durch das No-Cloning-Theorem verboten.

Damit im Zusammenhang steht die Erkenntnis von N. Herbert (1982), dass die Möglichkeit der Kopie von beliebigen Zuständen überlichtschnelle Informationsübertragung erlauben würde. Dies wird durch das No-Cloning-Theorem ausgeschlossen.

13.4 Ensemble-Interpretation

Wir haben gesehen, dass die Quantenmechanik eine stochastische Theorie ist – wenn bei einem Stern-Gerlach-Experiment der einlaufende Zustand der Eigenzustand $|\rightarrow\rangle$ zu \hat{S}_x ist und man \hat{S}_z misst, kann man wirklich nicht vorhersagen, ob ein Teilchen nach oben oder unten abgelenkt wird. Eine wesentliche Frage ist, ob dieses Element des Zufalls eine vermeidbare Unvollständigkeit der theoretischen Beschreibung ist (Einstein) oder intrinsischer Natur (Bohr). Unabhängig von der Haltung zu dieser Frage kann man sich nun auf den Standpunkt stellen, dass die Quantenmechanik nur als statistische Theorie von *Ensembles* zu verstehen sei, aber nicht von Einzelsystemen.

Im Rahmen dieser Sichtweise ist der Formalismus des Dichteoperators aus Kapitel 12 naheliegend, da er gerade Ensembles beschreibt. „Reine Zustände“ werden dann ebenfalls als Ensembles verstanden, bei denen man aber maximale Information über die Einzelsysteme hat. Für die Ensemble-Interpretation und den Dichteoperator-Formalismus spricht, dass man bei der Betrachtung von Teilsystemen generisch eine statistische Beschreibung benötigt, selbst wenn das Gesamtsystem in einem reinen Zustand ist, wie wir in 12.2.2 gesehen haben. Da man in der Praxis nicht den Zustand des gesamten Universums betrachtet, ist die Beschreibung mittels des Dichteoperators als primär und die Betrachtung reiner Zustände von Teilsystemen mittels der Schrödinger-Gleichung als Näherung anzusehen.

Außerdem kann man im Dichteoperator-Formalismus den Zustandskollaps recht elegant beschreiben: Die Außerdiagonalkomponenten des Dichteoperators in der Eigenbasis der gemessenen Observable werden auf Null gesetzt,

$$\begin{pmatrix} \rho_{11} & \rho_{12} & \cdots \\ \rho_{21} & \rho_{22} & \cdots \\ \vdots & \vdots & \ddots \end{pmatrix} \rightarrow \begin{pmatrix} \rho_{11} & 0 & \cdots \\ 0 & \rho_{22} & \cdots \\ \vdots & \vdots & \ddots \end{pmatrix}. \quad (13.49)$$

Erklärt wird der Kollaps damit aber sicher nicht.

Den Wahrscheinlichkeiten aus der Bornschen Regel wird im Rahmen der Ensemble-Interpretation eine Bedeutung in der Realität (also eine *ontologische* Bedeutung) zugeschrieben: Sie sind relative Häufigkeiten von Messergebnissen in Ensembles. Wenn man von Ensembles von Messungen spricht, geht die Ensemble-Interpretation nicht über die Minimalinterpretation hinaus. Wenn man dagegen an Ensembles von Mikrosystemen denkt – was natürlicher scheint – dann schon. Die Interpretation von Wahrscheinlichkeiten als relative Häufigkeiten wird philosophisch auch ohne Bezug zur Quantentheorie als problematisch angesehen. Ein Punkt dabei ist, dass zwar relative Häufigkeiten für Ensembles aus N Einzelsystemen für $N \rightarrow \infty$ mit Wahrscheinlichkeit eins gegen die

Wahrscheinlichkeiten streben, aber eben nur *mit Wahrscheinlichkeit eins* – auch für beliebig große N kann eine gemessene relative Häufigkeit beliebig weit von der tatsächlichen Wahrscheinlichkeit abweichen. Das Konzept der Wahrscheinlichkeit und damit die stochastische Natur der Theorie wird man in der Ensemble-Interpretation jedenfalls nicht los, was zumindest für einige ihrer Vertreter das Ziel war.

Auch aus anderen Gründen kann die Ensemble-Interpretation nicht befriedigen. Die Quantenmechanik, so wie sie in dieser Vorlesung entwickelt wurde, stellt sich als Theorie für die Dynamik von (reinen) Zuständen von Einzelsystemen dar. Der statistische Dichteoperator-Formalismus folgt aus dieser Schrödinger-Heisenberg-Dirac-Quantenmechanik und elementaren Aussagen der mathematischen Stochastik. Die statistische Deutung bricht i. W. die Quantenmechanik einzelner Zustände weg. Man kann dann nicht recht verstehen, wie eine Theorie für Ensembles begründet werden soll, wenn man keine fundamentale Theorie für Einzelsysteme hat. Insbesondere wird nicht klar, wieso die Wahrscheinlichkeiten gerade den von Born postulierten Wert haben.

Zum anderen kann man ja einzelne Quanten-Experimente durchführen. Es wäre unbefriedigend, wenn man theoretisch darüber nichts aussagen könnte. In der Tat macht die Quantenmechanik erfolgreiche Aussagen über Einzelsysteme, z. B. wenn man zweimal hintereinander dieselbe Observable misst. Wird bei einem Stern-Gerlach-Versuch zunächst der Eigenwert $m = \hbar/2$ für \hat{S}_z gemessen und dann die Messung von \hat{S}_z wiederholt, kommt mit Sicherheit wieder $m = \hbar/2$ heraus. Es liegt nahe, zu sagen, dass ein einzelnes Teilchen zwischen den Messungen im Zustand $|m\rangle$ ist. Dagegen erscheint es unnatürlich, nur die Aussage zuzulassen, dass die Teilsysteme eines Ensembles mit der Wahrscheinlichkeit 100% im Zustand $|m\rangle$ sind, d. h. dass der Dichteoperator in der Standardbasis

$$\hat{\rho} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \quad (13.50)$$

ist. Wenn man hier aber eine solche Aussage über das Einzelsystem zulässt, kann man nur schwer die Frage zurückweisen, was denn mit dem Einzelsystem im Zustand $|m\rangle$ passiert, wenn man dann \hat{S}_y statt \hat{S}_z misst.

Insgesamt können wir feststellen, dass die Ensemble-Interpretation dem Verständnis des Messprozesses nicht näher kommt. Sie verharret in der „shut up and calculate“-Haltung, indem sie den Anwendungsbereich der Quantenmechanik so einzuschränken versucht, dass schwierig zu deutende Aspekte herausfallen. Diese defensive Strategie ist aber noch nicht einmal erfolgreich. All dies bedeutet nicht, dass die Beschreibung von Ensembles mittels des Dichteoperator-Formalismus abzulehnen wäre. Sie ist im Gegenteil ein sehr erfolgreicher Teil der Standardformulierung. Problematisch ist nur, die Quantenmechanik als *ausschließlich* für Ensembles sinnvolle Theorie interpretieren zu wollen.

13.5 Kopenhagener Deutung(en)

Während gemäß der Minimalinterpretation die Quantenmechanik makroskopische Messapparaturen beschreibt, bezieht sie sich nach der Ensemble-Interpretation immerhin auf mikroskopische Quantensysteme, aber nur als Ensemble. Erst die Kopenhagener Deutung nach Niels Bohr und Werner Heisenberg geht davon aus, dass die Quantenmechanik wirklich *einzelne Quantensysteme* beschreibt. Sie sieht darüber hinaus die Quantenmechanik als *vollständig* an.

Die Kopenhagener Deutung galt für lange Zeit als Standardinterpretation und wird in Lehrbüchern oft immer noch so dargestellt. Allerdings wird nicht immer dasselbe gemeint, wenn von der Kopenhagener Deutung die Rede ist. Schon Bohr und Heisenberg haben nicht dieselben Standpunkte vertreten. Außerdem blieb v. a. Bohr recht vage in seiner Interpretation. Heisenbergs Sichtweise wird oft als Kopenhagener Deutung im engeren Sinne angesehen, nicht zuletzt, weil Heisenberg selbst diese Bezeichnung in einem Artikel von 1959 („Die Kopenhagener Deutung der Quantentheorie“) eingeführt hat.

13.5.1 Bohrs Interpretation

Niels Bohr vertrat drei Thesen:

- Bei der Messung existiert notwendig eine Verknüpfung zwischen dem Quantensystem und der Messapparatur.

- Experimente, insbesondere der Aufbau von Messapparaturen, müssen in der Sprache der klassischen Physik beschrieben werden.
- *Komplementarität*: Verschiedene Beschreibungen der Realität können jeweils für sich korrekt, aber unvereinbar sein. Zum Beispiel kann man einem Quantensysteme scharfe Werte von einer von zwei unverträglichen Observablen zuschreiben, aber nicht beiden gleichzeitig. Die Welle-Teilchen-Dualität ist auch von dieser Art – sie beruht ja auf der Unverträglichkeit von Ort und Impuls.

Damit ist laut Bohr das Quantensystem zwar real, aber es ist nicht unabhängig von der makroskopischen Apparatur. Quantensysteme haben Eigenschaften nur bezogen auf die zu ihrer Messung dienende Apparatur, die klassisch beschrieben wird. Dies legt nahe, dass Quantensysteme vielleicht zwar real existieren, aber ihre (oder ein Teil ihrer) Eigenschaften erst durch die Kopplung an eine Messapparatur erhalten. Andererseits besteht die Messapparatur aus Elementarteilchen, die man quantenmechanisch beschreiben möchte. Wenn aber diese Teilchen ihre Eigenschaften nicht ohne Bezug auf einen Messapparat haben, ergibt sich die Gefahr einer zirkulären Beschreibung.

Weiter ließ sich Bohr nicht darauf ein, zu sagen, was beim Messprozess wirklich passiert. Er akzeptierte jedenfalls nicht den Zustandskollaps als fundamentales Element der Quantenmechanik. Aber es fragt sich, wie dann bestimmte Messwerte zu Stande kommen. Damit blieben die wesentlichen Fragen der Deutung des Messprozesses unbeantwortet.

13.5.2 Heisenbergs Interpretation

Werner Heisenberg geht über Bohrs vorsichtige Interpretation hinaus und schreibt dem einzelnen Quantensystem Realität zu. Insbesondere hat nach Heisenberg ein Quantensystem in einem Eigenzustand $|\psi_A\rangle$ eines hermiteschen Operators \hat{A} wirklich eine durch den zugehörigen Eigenwert A charakterisierte Eigenschaft. Heisenberg akzeptiert die Postulate von von Neumann und Born und damit die Existenz zweier unterschiedlicher Dynamiken:

- die deterministische, stetige und reversible *unitäre Zeitentwicklung* gemäß der Schrödinger-Gleichung und
- den indeterministischen, unstetigen und irreversible *Kollaps* des Zustands bei einer Messung.

Nach Heisenberg beschreibt aber keine der beiden Dynamiken einen realen Prozess in der Welt. Denn er schreibt zwar Eigenwerten hermitescher Operatoren Realität zu, aber nicht dem Zustandsvektor bzw. der Wellenfunktion. Er deutet den Zustand als Ausdruck der Kenntnis des Beobachters über das System: Er schreibt nämlich 1959 über den Zustandskollaps: „hier handelt es sich um die unstetige Änderung unserer Kenntnis im Moment der Registrierung.“ Der Zustand hat demnach *epistemischen* Charakter (er ist Teil der Beschreibung) und nicht *ontologischen* (Teil der Welt).

13.5.3 Zustandskollaps

Wenn man die Existenz zweier Zeitentwicklungen annimmt, muss offenbar entschieden werden, wann welche Dynamik stattfindet.

1. Falls das wesentliche Kriterium ist, ob eine *Messung* stattfindet, stellt sich die Frage, wann ein Prozess eine Messung ist. Nimmt man an, dass der Zustand real ist in dem Sinne, dass er zu jedem Zeitpunkt unabhängig von irgendeinem Beobachter ist, kommt man schnell in Schwierigkeiten. Wenn z. B. ein Teilchen „zufällig“ durch ein starkes und inhomogenes Magnetfeld fliegt, sagen wir in der Nähe der Sonne, so wird es abhängig von seinem Spin abgelenkt. Die Situation unterscheidet sich physikalisch nicht von einer Stern-Gerlach-Apparatur. Ist der Vorgang deshalb eine Messung? Wie schwach muss die Inhomogenität des Feldes und damit die Ablenkung bei einer Stern-Gerlach-Anordnung im Labor oder in der Natur sein, damit es keine Messung mehr ist? Falls sie beliebig schwach sein kann, existieren dann überhaupt „Nicht-Messungen“? Doch wohl ja, denn sonst existierte gar keine unitäre Zeitentwicklung und die Schrödinger-Gleichung wäre niemals anwendbar. Falls es andererseits eine Schwelle für die Inhomogenität gibt, unterhalb derer man es nicht mehr mit Messungen zu tun hat, was bestimmt dann die Schwelle?

Wir schließen, dass kein *physikalisches* Kriterium dafür existiert, was eine Messung ausmacht. Das stimmt natürlich mit dem üblichen Verständnis von Messungen überein, das einen Beobachter einschließt. Aber

der Versuch, die Notwendigkeit eines Beobachters mit einem realen Kollaps zu vereinbaren, führt direkt auf Wigners These, dass *Bewusstsein* notwendig ist, um Zustände zu kollabieren. Wie in 13.2.1 diskutiert, hat man dann das Problem zu entscheiden, wann ein Beobachter „bewusst genug“ ist.

2. Diese Schwierigkeiten ergeben sich nicht, wenn der Zustand im Hilbert-Raum – wie in der Kopenhagener Deutung – als Ausdruck der Kenntnis eines Beobachters über das System interpretiert wird und nicht als Eigenschaft des Systems. Dies wirft sofort die Frage auf, ob und wie wir das System selbst beschreiben können und nicht nur unsere Kenntnis davon. Die Vermutung, dass man die Kenntnis von einer Sache beschreiben kann, aber aus Prinzip nicht die Sache selbst, ist recht nah an “shut up and calculate”. Andererseits ist jede Theorie, in der die Beschreibung des Systems zusätzliche Information erlaubt, eine Theorie mit verborgenen Variablen.

Die unmittelbare Folge dieser „subjektivistischen“ Interpretation ist, dass für verschiedene Beobachter dasselbe Quantensystem durch unterschiedliche Zustände beschrieben werden kann, weil sie unterschiedliche Kenntnis über das System haben (z.B. Wigner und sein Freund). Ein Zustandskollaps zeigt dann einfach an, dass ein Beobachter zusätzliche Information gewonnen hat. Die Frage, ob Schrödingers Katze einen Zustand kollabieren kann, reduziert sich auf eine Frage nach den geistigen Fähigkeiten von Katzen und wird irrelevant für die Vorhersage und Interpretation von Quantenexperimenten durch menschliche Beobachter.

Was können verschiedene Beobachter über dasselbe System „wissen“? Angenommen, zwei Beobachter Alice und Bob haben dieselbe Kenntnis über einen Spin der Länge $1/2$, weil sie beide die Apparatur kennen, die den Spin präpariert. Sie beschreiben den Spin dann mit demselben Zustand, dieser sei $|\psi_A\rangle = |\psi_B\rangle = |\rightarrow\rangle$. Nun misst Alice in Abwesenheit von Bob \hat{S}_z und erhält den Messwert $\hbar/2$. Dann wird für Alice $|\psi_A\rangle = |\uparrow\rangle$. Aber Bob weiß nichts vom Messergebnis und daher bleibt für ihn $|\psi_B\rangle = |\rightarrow\rangle$. Als nächstes misst Alice \hat{S}_x und erhält den Messwert $-\hbar/2$. Diese Abfolge von Messergebnissen ist im Rahmen der Quantenmechanik natürlich erlaubt; sie tritt mit einer Wahrscheinlichkeit von 25% auf. Nun ist $|\psi_A\rangle = |\leftarrow\rangle$ und immer noch $|\psi_B\rangle = |\rightarrow\rangle$. Nun sagen beide Beobachter das Ergebnis einer weiteren Messung von \hat{S}_x vorher:

- Alice: mit 100% Wahrscheinlichkeit $-\hbar/2$,
- Bob: mit 100% Wahrscheinlichkeit $\hbar/2$.

Wenn sie nun beide die Messung von \hat{S}_x beobachten, ist das Ergebnis mit Sicherheit im Widerspruch zur Vorhersage von einem von beiden. Diese Schlussfolgerung bezieht sich auf die Vorhersagekraft der Quantenmechanik und ändert sich also nicht, wenn wir annehmen, dass ein wahrer Zustand des Spins existiert, der bestimmt, was bei der letzten Messung herauskommen kann. Unter der Annahme, dass der Zustand nur die Kenntnis über das System beschreibt, sind die Vorhersagen der Quantenmechanik also zumindest für manche Beobachter falsch.

3. Ein physikalisch potentiell besser fassbares Kriterium besteht in der Unterscheidung zwischen Mikro- und Makrosystemen. Es beinhaltet jedenfalls keinen Bezug zu einem Beobachter und verhindert daher auch nicht eine realistische Deutung des Zustands. Demnach würde die Schrödinger-Gleichung nur mikroskopische Systeme beschreiben. Makroskopische Systeme würden einer anderen (noch zu bestimmenden, aber vermutlich klassischen) Dynamik gehorchen. Koppelt man ein Mikro- an ein Makrosystem, kollabiert der Zustand sofort. Das schließt offensichtlich Messprozesse mit ein, da, nach Bohr, Messapparaturen ja klassisch beschreibbar und damit sicher makroskopisch sein müssen. Messungen werden hier nicht mehr als wesentlich verschieden von anderen Wechselwirkungen mit makroskopischen Systemen betrachtet.

Aber wir haben wieder die Schwierigkeit, dass eine Ja-Nein-Entscheidung von einer (quasi-) kontinuierlichen Variablen, nämlich der Systemgröße, abhängen soll. Wie groß muss ein System sein, damit der Zustand kollabiert? Eine konkret ausgearbeitete Interpretation in dieser Richtung stammt von Ghirardi, Rimini und Weber und wird in 13.8.2 kurz besprochen. Dabei wird postuliert, dass die typische Zeit bis zum Kollaps proportional zur Teilchenzahl in einem System ist. Die Proportionalitätskonstante wird dann so gewählt, dass diese Zeit für wenige Teilchen sehr lang, aber für 10^{23} Teilchen sehr kurz ist.

13.6 Dekohärenz

Die von Heisenberg angenommene zweifache Dynamik (unitär vs. Kollaps) wurde als wesentlicher Mangel der Kopenhagener Deutung angesehen. Daher gab es verschiedene Versuche, ohne Kollaps auszukommen. Dies war eine Motivation für die Entwicklung der Theorie der *Dekohärenz* ab 1970. Diese steht konzeptionell eher in der Tradition der Ensemble-Interpretation und kann, letztlich aus diesem Grund, das Messproblem nicht lösen. Wir hatten in 13.4 gesehen, dass bei einer Messung die Außerdiagonalkomponenten, genannt *Kohärenzen*, des Dichteoperators in der Eigenbasis der gemessenen Observablen auf Null gesetzt werden. Hier geht es um die Frage, *wie* dies genau geschieht. Das Ziel ist, dabei ohne expliziten Zustandskollaps auszukommen. Die Zeitentwicklung des Gesamtsystems soll also immer unitär sein.

Wir betrachten als Beispiel einen Spin der Länge 1/2, an dem \hat{S}_z gemessen wird, so dass als Messwert $\hbar/2$ oder $-\hbar/2$ herauskommen kann. Die Zahl der möglichen Messwerte ist für die Argumentation unerheblich. Ein Messgerät sollte auf jeden Fall *eindeutig* den *richtigen* Messwert anzeigen, wenn der Spin schon vorher in einem der entsprechenden Eigenzustände $|\uparrow\rangle$ oder $|\downarrow\rangle$ präpariert wurde. Vor der Messung sei das Messgerät im Ausgangszustand $|M_0\rangle$ und vom zu messenden Quantensystem getrennt. Wenn wir nun annehmen, dass die Messung durch den unitären Zeitentwicklungsoperator \hat{U} beschrieben wird, so gilt

$$\hat{U} |\uparrow\rangle |M_0\rangle = |\uparrow\rangle |M_\uparrow\rangle, \quad (13.51)$$

$$\hat{U} |\downarrow\rangle |M_0\rangle = |\downarrow\rangle |M_\downarrow\rangle. \quad (13.52)$$

Hier sind $|M_\uparrow\rangle$ und $|M_\downarrow\rangle$ makroskopisch verschiedene Zustände des Messgerätes, die jeweils den Messwert $\hbar/2$ bzw. $-\hbar/2$ anzeigen. Aber wegen der Linearität der Schrödinger-Gleichung gilt dann notwendig für eine Superposition

$$\hat{U} (c_\uparrow |\uparrow\rangle + c_\downarrow |\downarrow\rangle) |M_0\rangle = c_\uparrow |\uparrow\rangle |M_\uparrow\rangle + c_\downarrow |\downarrow\rangle |M_\downarrow\rangle. \quad (13.53)$$

Da $|M_\uparrow\rangle$ und $|M_\downarrow\rangle$ makroskopisch verschieden sind, ist dies ein Schrödinger-Katzen-Zustand. Das ist (scheinbar?) nicht, was man jemals beobachtet. Mit der Dekohärenz-Theorie möchte man dies *ohne* Kollaps (Projektionspostulat) verstehen.

Die Lösung besteht *nicht* darin, das Quantensystem und das Messgerät als Teilsysteme anzusehen und die reduzierten Dichteoperatoren auszurechnen. Wenn wir dies trotzdem tun, erhalten wir

$$\hat{\rho}_{\text{Spin}} = |c_\uparrow|^2 |\uparrow\rangle\langle\uparrow| + |c_\downarrow|^2 |\downarrow\rangle\langle\downarrow|, \quad (13.54)$$

$$\hat{\rho}_M = |c_\uparrow|^2 |M_\uparrow\rangle\langle M_\uparrow| + |c_\downarrow|^2 |M_\downarrow\rangle\langle M_\downarrow|. \quad (13.55)$$

Aber hier haben wir relevante Information weggeworfen. Denn nichts in Glg. (13.54) und (13.55) verhindert, dass der Spin im Zustand $|\uparrow\rangle$ ist, das Messgerät aber im Zustand M_\downarrow , so dass es $-\hbar/2$ anzeigt. Das schließt der Zustand des Gesamtsystems aber aus. Also macht die Beschreibung mittels der reduzierten Dichteoperatoren falsche Voraussagen und ist daher ungeeignet, um irgendetwas zu verstehen.

Die wesentliche Idee ist nun, das Quantensystem und das Messgerät zusammen als nicht abgeschlossen anzunehmen. Stattdessen steht es, so das Argument, immer in Wechselwirkung mit seiner Umgebung, d. h. mit dem Rest des Universums. Es ist für die Argumentation nützlich, als Gedankenexperiment anzunehmen, dass die Wechselwirkung zwischen Quantensystem und Messgerät einerseits und der Umgebung andererseits erst zu einem späteren Zeitpunkt nach der Messung eingeschaltet wird (das erinnert stark an Schrödingers Katze!). Dann ist nach der Messung, aber vor der Kopplung an die Umgebung der Zustand von der Produktform

$$|\Psi\rangle = (c_\uparrow |\uparrow\rangle |M_\uparrow\rangle + c_\downarrow |\downarrow\rangle |M_\downarrow\rangle) |U_0\rangle. \quad (13.56)$$

Den reduzierten Dichteoperator von Quantensystem und das Messgerät erhalten wir durch Bildung der Spur über die Umgebung (Sp_U). Dadurch fällt einfach der Umgebungszustand U_0 heraus:

$$\begin{aligned} \hat{\rho}_{Q+M} &= \text{Sp}_U |\Psi\rangle\langle\Psi| \\ &= (c_\uparrow |\uparrow\rangle |M_\uparrow\rangle + c_\downarrow |\downarrow\rangle |M_\downarrow\rangle) (c_\uparrow^* \langle M_\uparrow| \langle\uparrow| + c_\downarrow^* \langle M_\downarrow| \langle\downarrow|) \\ &= |c_\uparrow|^2 |\uparrow\rangle |M_\uparrow\rangle\langle M_\uparrow| \langle\uparrow| + c_\uparrow c_\downarrow^* |\uparrow\rangle |M_\uparrow\rangle\langle M_\downarrow| \langle\downarrow| + c_\downarrow c_\uparrow^* |\downarrow\rangle |M_\downarrow\rangle\langle M_\uparrow| \langle\uparrow| + |c_\downarrow|^2 |\downarrow\rangle |M_\downarrow\rangle\langle M_\downarrow| \langle\downarrow| \end{aligned}$$

$$= \begin{pmatrix} |c_\uparrow|^2 & c_\uparrow c_\downarrow^* \\ c_\downarrow c_\uparrow^* & |c_\downarrow|^2 \end{pmatrix}. \quad (13.57)$$

Das ist natürlich ein reiner Zustand, weil Quantensystem und Messgerät nicht mit der Umgebung verschränkt waren.

Nun werde die Kopplung an die Umgebung eingeschaltet. Durch diese Wechselwirkung ändert sich die Umgebung in unterschiedlicher Weise für die beiden Spin-Einstellungen. Wenn wir idealisierend annehmen, dass das Messergebnis dadurch nicht verfälscht wird (wir können an ein Stern-Gerlach-Experiment denken; das Vorzeichen der Ablenkung der Teilchen ist vermutlich sehr robust gegenüber der Wechselwirkung mit der Umgebung), dann wird der Gesamtzustand schließlich von der Form

$$|\Psi\rangle = c_\uparrow |\uparrow\rangle |M_\uparrow\rangle |U_\uparrow\rangle + c_\downarrow |\downarrow\rangle |M_\downarrow\rangle |U_\downarrow\rangle \quad (13.58)$$

sein, wo $|U_m\rangle$ nun den Zustand der Umgebung repräsentiert. Die Umgebung ist kein Messgerät und insbesondere nicht so konstruiert worden, dass der Zustand der Umgebung *eindeutig* den Zustand des Quantensystems widerspiegelt. Daher ist nicht garantiert, dass $|U_\uparrow\rangle$ und $|U_\downarrow\rangle$ orthogonal sind (wir nehmen aber beide als normiert an). Das müssen wir beachten, wenn wir nun den reduzierten Dichteoperator des aus Quantensystem und Messgerät bestehenden Teilsystems ausrechnen:

$$\begin{aligned} \hat{\rho}_{Q+M} &= \text{Sp}_U \hat{\rho} = \text{Sp}_U |\Psi\rangle\langle\Psi| \\ &= \sum_n \langle u_n | (c_\uparrow |\uparrow\rangle |M_\uparrow\rangle |U_\uparrow\rangle + c_\downarrow |\downarrow\rangle |M_\downarrow\rangle |U_\downarrow\rangle) (c_\uparrow^* \langle U_\uparrow | \langle M_\uparrow | \langle \uparrow | + c_\downarrow^* \langle U_\downarrow | \langle M_\downarrow | \langle \downarrow |) | u_n \rangle \\ &= |c_\uparrow|^2 \langle U_\uparrow | \underbrace{\sum_n |u_n\rangle\langle u_n|}_{=1_U} |U_\uparrow\rangle \langle \uparrow | |M_\uparrow\rangle \langle M_\uparrow | \langle \uparrow | + c_\uparrow c_\downarrow^* \langle U_\uparrow | \underbrace{\sum_n |u_n\rangle\langle u_n|}_{=1_U} |U_\downarrow\rangle \langle \uparrow | |M_\uparrow\rangle \langle M_\downarrow | \langle \downarrow | \\ &\quad + c_\downarrow c_\uparrow^* \langle U_\downarrow | \underbrace{\sum_n |u_n\rangle\langle u_n|}_{=1_U} |U_\uparrow\rangle \langle \downarrow | |M_\downarrow\rangle \langle M_\uparrow | \langle \uparrow | + |c_\downarrow|^2 \langle U_\downarrow | \underbrace{\sum_n |u_n\rangle\langle u_n|}_{=1_U} |U_\downarrow\rangle \langle \downarrow | |M_\downarrow\rangle \langle M_\downarrow | \langle \downarrow | \\ &= |c_\uparrow|^2 |\uparrow\rangle |M_\uparrow\rangle \langle M_\uparrow | \langle \uparrow | + c_\uparrow c_\downarrow^* \langle U_\uparrow | U_\downarrow \rangle |\uparrow\rangle |M_\uparrow\rangle \langle M_\downarrow | \langle \downarrow | \\ &\quad + c_\downarrow c_\uparrow^* \langle U_\downarrow | U_\uparrow \rangle |\downarrow\rangle |M_\downarrow\rangle \langle M_\uparrow | \langle \uparrow | + |c_\downarrow|^2 |\downarrow\rangle |M_\downarrow\rangle \langle M_\downarrow | \langle \downarrow | \end{aligned} \quad (13.59)$$

wobei $\{|u_n\rangle\}$ eine Orthonormalbasis der Umgebung ist. In der Basis $\{|\uparrow\rangle |M_\uparrow\rangle, |\downarrow\rangle |M_\downarrow\rangle\}$ ist dieser Dichteoperator

$$\hat{\rho}_{Q+M} = \begin{pmatrix} |c_\uparrow|^2 & c_\uparrow c_\downarrow^* \langle U_\uparrow | U_\downarrow \rangle \\ c_\downarrow c_\uparrow^* \langle U_\downarrow | U_\uparrow \rangle & |c_\downarrow|^2 \end{pmatrix}. \quad (13.60)$$

Er ist offenbar nicht diagonal in der Eigenbasis der Messgröße \hat{S}_z , weil $|U_\uparrow\rangle$ und $|U_\downarrow\rangle$ i. A. nicht orthogonal sind. Aber dies sind Zustände des Restes des Universums, d. h. Zustände in einem Hilbert-Raum von unendlicher oder jedenfalls sehr hoher Dimension. Wir erwarten, dass eine unkontrollierte Wechselwirkung für die Zustände $|\uparrow\rangle |M_\uparrow\rangle$ und $|\downarrow\rangle |M_\downarrow\rangle$ von Spin und Messapparat den Umgebungszustand in unterschiedlicher Weise zufällig ändert. Dann sind $|U_\uparrow\rangle$ und $|U_\downarrow\rangle$ Ergebnisse unabhängiger Zufallsbewegungen in einem hochdimensionalen Vektorraum und daher ist ihr Skalarprodukt sehr klein.

Zur Veranschaulichung: Für zwei normierte Zufallsvektoren im \mathbb{R}^n ist der Erwartungswert des Quadrates ihres Skalarproduktes

$$\overline{(\vec{v} \cdot \vec{w})^2} = \frac{1}{n} \quad (13.61)$$

Beweis: Wir können das Koordinatensystem so wählen, dass die x_m -Achse, $m = 1, \dots, n$ parallel zu \vec{v} ist. Dann ist

$$\overline{(\vec{v} \cdot \vec{w})^2} = \overline{w_m^2}. \quad (13.62)$$

Andererseits gilt

$$\sum_m \overline{w_m^2} = \overline{\sum_m w_m^2} = \overline{1} = 1 \quad (13.63)$$

Aber nun ist $\overline{w_m^2}$ für alle m wegen der Isotropie des Raumes gleich. Es folgt

$$\overline{w_m^2} = \frac{1}{n} \quad (13.64)$$

und daraus die Behauptung. \square

Wird das Skalarprodukt $\langle U_\uparrow | U_\downarrow \rangle$ Null oder vernachlässigbar, so erhalten wir offensichtlich

$$\hat{\rho}_{Q+M} = |c_\uparrow|^2 |\uparrow, M_\uparrow\rangle\langle\uparrow, M_\uparrow| + |c_\downarrow|^2 |\downarrow, M_\downarrow\rangle\langle\downarrow, M_\downarrow| = \begin{pmatrix} |c_\uparrow|^2 & 0 \\ 0 & |c_\downarrow|^2 \end{pmatrix}. \quad (13.65)$$

Dieser reduzierte Dichteoperator drückt aus, dass

- das Messgerät *entweder* $\hbar/2$ oder $-\hbar/2$ anzeigt und
- das Messgerät genau den Zustand anzeigt, den der Spin wirklich hat.

Man kann ohne Probleme den Beobachter mit in $|M_m\rangle$ einschließen und sagen, dass nun der Beobachter entweder $\hbar/2$ oder $-\hbar/2$ abliest und dies dem tatsächlichen Spin-Zustand entspricht.

Die Außerdiagonalkomponenten (Kohärenzen) müssen stetig von i. A. großen Werten mit dem Betrag $|c_\uparrow c_\downarrow|$ in Eq. (13.57) auf Null in der letzten Gleichung abfallen. Diesen Prozess nennt man *Dekohärenz*. Die Grundlage des Arguments ist, dass dies *unvermeidbar* ist und wegen der riesigen Größe der Umgebung *sehr schnell* passiert. Der Zerfall der Kohärenzen ist eine *unitäre* Zeitentwicklung des Gesamtsystems, die aber für das Quantensystem und das Messgerät nicht von einem instantanen Kollaps zu unterscheiden ist.

Diese Interpretation klingt vernünftig. Hier ist aber noch ein Problem verborgen: Der Bezug auf Außerdiagonalkomponenten impliziert eine *Basisabhängigkeit*. Der reduzierte Dichteoperator $\hat{\rho}_{Q+M}$ ist i. A. in manchen Basen diagonal, in anderen aber nicht. (Dies gilt nur dann nicht, wenn der Dichteoperator proportional zum Identitätsoperator ist.) Aussagen der Quantenmechanik sollten üblicherweise basisunabhängig sein, was also zeichnet hier eine Basis aus? Die *scheinbar* offensichtliche Antwort ist, dass die Eigenbasis der gemessenen Observablen ausgezeichnet ist, weil das Messgerät eben so konstruiert ist, dass es diese Observable misst und nicht irgendeine andere. Wir würden in unserem Beispiel gern wie folgt argumentieren: Weil der Apparat \hat{S}_z misst und nicht z. B. \hat{S}_x , hat der Zustand nach der Messung aber vor der Wechselwirkung mit der Umgebung die Form

$$c_\uparrow |\uparrow\rangle |M_\uparrow\rangle + c_\downarrow |\downarrow\rangle |M_\downarrow\rangle \quad (13.66)$$

und nicht

$$c_{\rightarrow} |\rightarrow\rangle |M_{\rightarrow}\rangle + c_{\leftarrow} |\leftarrow\rangle |M_{\leftarrow}\rangle. \quad (13.67)$$

Hier sind $|\rightarrow\rangle$, $|\leftarrow\rangle$ Eigenzustände von \hat{S}_x zu den Eigenwerten $\pm\hbar/2$. Aber das ist *falsch*. Der Zustand in Gl. (13.66) lässt sich umschreiben als

$$\begin{aligned} c_\uparrow |\uparrow\rangle |M_\uparrow\rangle + c_\downarrow |\downarrow\rangle |M_\downarrow\rangle &= c_\uparrow \frac{|\rightarrow\rangle + |\leftarrow\rangle}{\sqrt{2}} |M_\uparrow\rangle + c_\downarrow \frac{|\rightarrow\rangle - |\leftarrow\rangle}{\sqrt{2}} |M_\downarrow\rangle \\ &= |\rightarrow\rangle \frac{c_\uparrow |M_\uparrow\rangle + c_\downarrow |M_\downarrow\rangle}{\sqrt{2}} + |\leftarrow\rangle \frac{c_\uparrow |M_\uparrow\rangle - c_\downarrow |M_\downarrow\rangle}{\sqrt{2}}. \end{aligned} \quad (13.68)$$

Dies hat genau dieselbe Form wie Gl. (13.67)! (Das Ergebnis hängt nicht davon ab, dass der Spin-Hilbert-Raum im Beispiel zweidimensional ist. Man kann jeden Zustand eines zusammengesetzten Systems in Produktzustände mit einer beliebig gewählten Basis *eines* Teilsystems, hier des Spins, entwickeln.) Der verschränkte Zustand von Quantensystem und Messgerät lässt sich auf unendlich viele verschiedene Weisen in eine Summe von Produktzuständen zerlegen und zwar unabhängig davon, was das Messgerät messen soll.

Es gibt aber einen entscheidenden Unterschied zwischen den verschiedenen Zerlegungen desselben Zustandes in Gl. (13.68): $|M_\uparrow\rangle$ und $|M_\downarrow\rangle$ sind makroskopisch verschiedene Zustände des Messgerätes, in denen es unterschiedliche Messwerte $\pm\hbar/2$ für \hat{S}_z anzeigt. Dagegen ist

$$|M_{\rightarrow}\rangle \sim c_\uparrow |M_\uparrow\rangle + c_\downarrow |M_\downarrow\rangle \quad (13.69)$$

nicht ein Zustand eines Messgerätes, das \hat{S}_x misst. Es ist ja immer noch derselbe Apparat. Vielmehr ist $|M_{\rightarrow}\rangle$ eine *Superposition* der beiden makroskopisch verschiedenen Zustände für unterschiedliche Messwerte von \hat{S}_z . $|M_{\rightarrow}\rangle$ ist also ein Schrödinger-Katzen-Zustand. Dekohärenz muss offenbar so stattfinden, dass Schrödinger-Katzen-Zustände sehr schnell zerfallen.

Es ist tatsächlich zu erwarten, dass Kohärenzen zwischen makroskopisch verschiedenen Zuständen schnell zerfallen. Zur Begründung kommen wir auf die Diskussion nach Gl. (13.60) zurück: Die Außerdiagonalkomponenten des reduzierten Dichteoperators in der Eigenbasis von \hat{S}_z , Gl. (13.60), fallen auf Null, wenn die Umgebungszustände $|U_{\uparrow}\rangle$ und $|U_{\downarrow}\rangle$ orthogonal werden. Die unkontrollierte Wechselwirkung der Umgebung mit den *makroskopisch verschiedenen* Zuständen $|M_{\uparrow}\rangle$ und $|M_{\downarrow}\rangle$ erzeugt deutlich unterschiedliche Zufallsbewegungen von $|U_{\uparrow}\rangle$ und $|U_{\downarrow}\rangle$ in einem hochdimensionalen Vektorraum und daher wird ihr Skalarprodukt schnell klein. Im Vergleich zur Diskussion nach Gl. (13.60) ist ein Punkt geschärft worden: Die Dekohärenz wird durch die makroskopisch verschiedenen Zustände des Messgerätes bewirkt. Rechnungen für Modelle unterstützen die Erwartung, dass Schrödinger-Katzen-Zustände schnell dekohärieren.

Zusammengefasst ist die Dekohärenz-Theorie recht erfolgreich darin, den scheinbaren Zustandskollaps im Dichteoperator-Formalismus durch eine unitäre Zeitentwicklung unter Einbeziehung der Umgebung zu erklären. Dies löst aber nicht die Aufgabe, den Messprozess auf Basis reiner Zustände zu erklären. Wir erwarten, dass das Universum in einem reinen Zustand ist und es ist daher relevant zu fragen, was mit diesem bei einem Stern-Gerlach-Experiment geschieht. Wenn aber, wie die Dekohärenz-Theorie nahelegt, die Zeitentwicklung von Dichteoperatoren allein durch die unitäre Zeitentwicklung des Gesamtsystems beschrieben werden kann und kein zusätzlicher Kollaps notwendig ist, liegt es nahe, auch einen reinen Zustand so zu beschreiben. Ansonsten müsste ein Kollaps so stattfinden, dass er mit der rein unitären Zeitentwicklung des Dichteoperators verträglich ist.

13.7 Kein Kollaps: Viele-Welten-Interpretation

T. Maudlin (1995) hat das Messproblem als Trilemma formuliert, d. h. von drei Aussagen über die Quantenmechanik muss mindestens eine falsch sein:

1. Die Quantenmechanik ist *vollständig*, d. h. der Zustandsvektor $|\psi\rangle$ beschreibt sämtliche Eigenschaften eines Systems.
2. Die Zustände gehorchen einer *linearen* (und reversiblen) Dynamik. Konkreter nimmt die Quantenmechanik die durch die Schrödinger-Gleichung beschriebene unitäre Dynamik an, dies ist hier aber nicht erforderlich.
3. Messungen liefern *definitive Messwerte*, die Eigenwerte von hermiteschen Operatoren sind, und sind wiederholbar.

Verschiedene Interpretationen der Quantenmechanik unterscheiden sich u. a. darin, welche dieser Aussagen sie für falsch halten. Vollständigkeit und Linearität erscheinen als natürliche und wünschenswerte Eigenschaften der Theorie. Fordert man sie, so muss die Existenz definitiver Messwerte verletzt sein. In diesem Abschnitt wollen wir die Konsequenzen der Annahme untersuchen, dass die unitäre Quantenmechanik – ohne Kollaps – eine vollständige Beschreibung darstellt. Die Strategie ist hier also, soweit möglich die Interpretation aus dem Formalismus abzuleiten. Dieses Programm wurde zunächst von Hugh Everett III. (1957) verfolgt.

Wir betrachten als Beispiel wieder eine Stern-Gerlach-Apparatur. Ein Spin-1/2-Teilchen soll auf eine Apparatur treffen, die Teilchen entsprechend der z -Komponente ihres Spins nach oben oder unten ablenkt. Wir können den einlaufenden Zustand schreiben als

$$c_{\uparrow} |\uparrow\rangle + c_{\downarrow} |\downarrow\rangle. \quad (13.70)$$

Bei unitärer Zeitentwicklung ist der Zustand nach der Messung dann

$$|\psi\rangle = c_{\uparrow} |\uparrow, z = z_0\rangle + c_{\downarrow} |\downarrow, z = -z_0\rangle, \quad (13.71)$$

in dem Teilchen mit $m = \uparrow, \downarrow$ bei $z = \pm z_0$ auf dem Schirm auftreffen. Der Beobachter ist nun auch Teil des Universums und kann (muss?) daher von der nach Voraussetzung vollständigen Quantentheorie mit beschrieben werden

(im Widerspruch zu Wigners „Kollaps durch Bewusstsein“!). Nach der Beobachtung sind das Quantensystem und Beobachterin Alice dann in einem Schrödinger-Katzen-Zustand der Form

$$|\Psi\rangle = c_{\uparrow} |\uparrow, A_{\uparrow}\rangle + c_{\downarrow} |\downarrow, A_{\downarrow}\rangle. \quad (13.72)$$

Diesen Zustand fanden Einstein, Schrödinger, Heisenberg usw. inakzeptabel und versuchten daher, ihn entweder als bloß epistemisch (Teil der Beschreibung, nicht der Realität) zu interpretieren oder seinen sehr schnellen Zerfall zu begründen.

Everett hat die Superposition dagegen akzeptiert und interpretiert, was sie eigentlich bedeutet. Seine Vorgänger (z. B. Schrödinger in seiner Diskussion des Katzen-Paradoxons) und auch viele Nachfolger haben argumentiert, dass wir nie Superpositionen von makroskopisch verschiedenen Zuständen beobachten und dass daher eine Theorie/Interpretation, die sie vorhersagt, abzulehnen oder durch Kollapsprozesse zu ergänzen sei. Everett hat dagegen gezeigt, dass die Standard-Quantenmechanik zwar solche Superpositionen vorhersagt, aber nicht, dass wir sie beobachten können. Im Gegenteil, seine Argumente erklären, warum wir sie nicht beobachten können.

$|\Psi\rangle$ beschreibt offenbar die Superposition von einerseits Spin- \uparrow und Alice beobachtet \uparrow und andererseits Spin- \downarrow und Alice beobachtet \downarrow . Nun ist die Zeitentwicklung linear, d. h. es gilt

$$|\Psi(t)\rangle = \hat{U}(t, t_0) (c_{\uparrow} |\uparrow, A_{\uparrow}(t_0)\rangle + c_{\downarrow} |\downarrow, A_{\downarrow}(t_0)\rangle) = c_{\uparrow} \hat{U}(t, t_0) |\uparrow, A_{\uparrow}(t_0)\rangle + c_{\downarrow} \hat{U}(t, t_0) |\downarrow, A_{\downarrow}(t_0)\rangle. \quad (13.73)$$

Die Zeitentwicklung des Zustandes mit „Alice beobachtet \uparrow “ ist daher unabhängig davon, ob er in Superposition zu einem Zustand mit „Alice beobachtet \downarrow “ steht oder nicht und wie sich dieser ggf. zeitlich entwickelt. Wegen der Linearität wechselwirken die beiden Vektoren $|\uparrow, A_{\uparrow}\rangle$ und $|\downarrow, A_{\downarrow}\rangle$ nie mehr miteinander. Der Gesamtzustand $|\Psi\rangle$ enthält jetzt zwei Instanzen von Alice in makroskopisch unterschiedlichen Zuständen (beobachten \uparrow bzw. \downarrow), die jeweils in ihrem Teilzustand gefangen sind.

Für Alice, die \uparrow beobachtet, scheint es so, als ob der einlaufende Zustand zu $|\uparrow\rangle$ kollabiert wäre. Wir können annehmen, dass Alice den einlaufenden Zustand kennt, weil sie weiß, durch welchen Prozess er präpariert wurde. Analog scheint es Alice, die \downarrow beobachtet, natürlich so, als ob der einlaufende Zustand zu $|\downarrow\rangle$ kollabiert wäre. Das ist Everetts Lösung für das Maudlinsche Trilemma: In Wirklichkeit existieren alle überhaupt möglichen Messwerte in Superposition, aber für jede Instanz der Beobachterin scheint es so, als ob ein bestimmter Wert gemessen wurde und der Spin-Zustand in den zugehörigen Eigenzustand kollabiert ist. Der beobachtete Kollaps ist also nur scheinbar, aber der Anschein ist eine perfekte Täuschung: Für Alice ist nicht unterscheidbar, ob der Zustand kollabiert ist, also nur noch der Teilzustand existiert, in dem sie steckt, oder ob noch die gesamte Superposition vorhanden ist.

Da nach der Messung eine Superposition von makroskopisch verschiedenen Zuständen vorhanden ist, die nie mehr miteinander wechselwirken, wird diese Sichtweise als *Viele-Welten-Interpretation* bezeichnet. Diese Bezeichnung stammt nicht von Everett, sondern von DeWitt und Graham (1973). Natürlich findet eine Aufspaltung in makroskopisch verschiedene Zustände nicht nur bei dezidierten Messungen statt, sondern bei allen Prozessen, die eben Schrödinger-Katzen-Zustände erzeugen. Einige Stärken der Interpretation sind die folgenden:

1. kein Kollaps, d. h. nur unitäre Zeitentwicklung gemäß der Schrödinger-Gleichung,
2. keine Sonderbehandlung von Messungen und daher auch keine Sonderrolle eines Beobachters,
3. keine verborgenen Variablen, zur Quantenmechanik wird nichts hinzugefügt.

Es gibt auch gewichtige Kritik an der Viele-Welten-Interpretation:

1. Die erste Kritik ergibt sich aus der (philosophischen) Forderung nach Sparsamkeit in der Theoriebildung. Diese lässt sich so ausdrücken, dass man nicht mehr Entitäten in die Theorie einführen sollte, als für deren Abgeschlossenheit und für korrekte Vorhersagen erforderlich. In diesem Sinne, so die Kritik, ist die Viele-Welten-Interpretation extrem verschwenderisch, denn sie behauptet die Existenz von sehr vielen Instanzen des Universums, die alle bis auf eine aus Prinzip nicht beobachtbar sind. Die Entgegnung ist folgende: Die Kritik ist nur überzeugend, wenn man die räumliche Struktur des Universums und seinen Gehalt an Teilchen und Feldern als fundamental ansieht. Sieht man dagegen den Zustandsvektor als fundamental an, so läuft die Kritik ins Leere – es gibt nur einen Zustand – und die Viele-Welten-Interpretation ist im Gegenteil sparsamer als die Kopenhagener Deutung, weil sie die Einführung einer zweiten (Kollaps-) Dynamik vermeidet.

2. Eine zweite Kritiklinie betrifft die Mehrdeutigkeit der Zerlegung von $|\Psi\rangle$ in eine Superposition von Produktzuständen. Ein ähnliches Problem tritt auch in der Beschreibung der Dekohärenz auf. Man kann den verschränkten Zustand nämlich umschreiben gemäß

$$\begin{aligned}
 |\Psi\rangle &= c_{\uparrow} |\uparrow, A_{\uparrow}\rangle + c_{\downarrow} |\downarrow, A_{\downarrow}\rangle \equiv c_{\uparrow} |\uparrow\rangle |A_{\uparrow}\rangle + c_{\downarrow} |\downarrow\rangle |A_{\downarrow}\rangle \\
 &= c_{\uparrow} \frac{|\rightarrow\rangle + |\leftarrow\rangle}{\sqrt{2}} |A_{\uparrow}\rangle + c_{\downarrow} \frac{|\rightarrow\rangle - |\leftarrow\rangle}{\sqrt{2}} |A_{\downarrow}\rangle \\
 &= |\rightarrow\rangle \frac{c_{\uparrow} |A_{\uparrow}\rangle + c_{\downarrow} |A_{\downarrow}\rangle}{\sqrt{2}} + |\leftarrow\rangle \frac{c_{\uparrow} |A_{\uparrow}\rangle - c_{\downarrow} |A_{\downarrow}\rangle}{\sqrt{2}} \\
 &=: c_{\rightarrow} |\rightarrow\rangle |A_{\rightarrow}\rangle + c_{\leftarrow} |\leftarrow\rangle |A_{\leftarrow}\rangle.
 \end{aligned} \tag{13.74}$$

Hier beschreiben $|A_{\rightarrow}\rangle$ und $|A_{\leftarrow}\rangle$ *nicht* Alice, die die Spin-Einstellung \rightarrow bzw. \leftarrow beobachtet, denn die Messapparatur kann nur \hat{S}_z messen. Vielmehr sind dies selbst Schrödinger-Katzen-Zustände von Alice in Superposition von „beobachtet \uparrow “ und „beobachtet \downarrow “. Daher kann man $|A_{\rightarrow}\rangle$ und $|A_{\leftarrow}\rangle$ nicht als Instanzen von Alice auffassen und die formal korrekte Form $c_{\rightarrow} |\rightarrow\rangle |A_{\rightarrow}\rangle + c_{\leftarrow} |\leftarrow\rangle |A_{\leftarrow}\rangle$ des Zustandes beschreibt keine Aufspaltung in zwei Welten. Wir sehen, dass eine bevorzugte Basis existiert, nämlich hier die Eigenbasis von \hat{S}_z . Diese Basisabhängigkeit erscheint aber unproblematisch, da der Apparat eben so konstruiert ist, dass er \hat{S}_z misst, und daher in der Lage ist, Alice abhängig von der z -Komponente in einen Schrödinger-Katzen-Zustand zu versetzen.

3. Der dritte und evtl. schwerwiegendste Einwand bezieht sich auf die Wahrscheinlichkeitsinterpretation. Schon in der Minimalinterpretation sind die Wahrscheinlichkeiten für Messergebnisse durch die Bornsche Regel gegeben. Im obigen Beispiel liefert diese die Wahrscheinlichkeiten $|c_{\uparrow}|^2$ und $|c_{\downarrow}|^2$ für die Messwerte $\pm\hbar/2$ bzw. die Beobachtungen \uparrow und \downarrow . Die Kritiker fordern, dass die Viele-Welten-Interpretation auch erklären können muss, wieso Alice den Beobachtungen \uparrow und \downarrow diese (Bornschen) Wahrscheinlichkeiten zuschreiben kann. Dies erweist sich als schwierig. In der Viele-Welten-Interpretation hat es keinen Sinn, wenn Alice vor der Messung fragt „mit welcher Wahrscheinlichkeit werde ich \uparrow bzw. \downarrow beobachten?“ Dies ist nicht die richtige Frage, weil ja *alle* überhaupt möglichen Ergebnisse auch eintreten.

Hier ist es wichtig festzustellen, dass der Messprozess für irgendeine Instanz von Alice ununterscheidbar von einem stochastischen Kollaps ist. Alice, die \uparrow beobachtet, erinnert sich nach der Messung eben daran, dass sie vor der Messung nicht wusste, was herauskommen würde, und dass sie dann \uparrow beobachtet hat. Da der Prozess ununterscheidbar von einem Zufallsexperiment ist, ist es im Prinzip sinnvoll, den möglichen Ausgängen Wahrscheinlichkeiten zuzuschreiben.

Was ist also die richtige Frage, wenn nicht „Mit welcher Wahrscheinlichkeit wird Alice \uparrow bzw. \downarrow beobachten?“ Wir betrachten die Situation, dass das Experiment bereits stattgefunden hat, aber Alice das Ergebnis noch nicht bekannt ist. Der Gesamtzustand befindet sich bereits in einer Superposition von zwei Zuständen mit dem Spin-Anteil $|\uparrow\rangle$ bzw. $|\downarrow\rangle$. Es ist zu diesem Zeitpunkt bereits sicher, dass sich diese beiden Zustände nach der Wechselwirkung mit Alice in Zustände $|\uparrow, A_{\uparrow}\rangle$ bzw. $|\downarrow, A_{\downarrow}\rangle$ entwickeln. Alice ist noch nicht bekannt, welcher Zustand es sein wird, dies steht aber bereits fest. Dafür wurde der Begriff *self-localization uncertainty* geprägt. Das unterscheidet sich konzeptionell nicht von einem Zufallsexperiment, in dem Alice würfelt, aber zunächst nicht nachsieht, welche Zahl gefallen ist. Den möglichen Ausgängen können daher konventionelle Wahrscheinlichkeiten zugeordnet werden.

Da die Zeitentwicklung linear ist und daher das Superpositionsprinzip gilt, kann der bereits aufgespaltene Teilzustand, in dem Alice \uparrow (oder \downarrow) beobachten wird, mittels des Zeitentwicklungsoperators auch zu Zeiten *vor* der Messung zurück entwickelt werden. In den dadurch entstehenden Zuständen ist Alice natürlich identisch. Trotzdem ist es eine sinnvolle Frage, in welchem Teilzustand sich Alice befindet. Was man gern zeigen würde ist, dass aus dem Formalismus der Standard-Quantenmechanik nicht nur die Beobachtung scheinbar zufälliger Messergebnisse folgt, sondern auch ihre fast immer beobachtete relative Häufigkeit. Vermutlich äquivalent dazu möchte man zeigen, dass es für uns vernünftig ist, auf den Ausgang von Messungen Wetten abzuschließen und dabei von den Wahrscheinlichkeiten nach Born auszugehen.

Man kann leicht einsehen, dass die Bornsche Regel für die Wahrscheinlichkeiten von Messergebnissen ein zulässiges Wahrscheinlichkeitsmaß darstellt. Die Bornsche Regel sagt folgendes aus, siehe Gl. (5.135): Der

Zustand eines Systems sei durch den normierten Hilbert-Raum-Vektor $|\phi\rangle$ beschrieben. Die Wahrscheinlichkeit, das System nach Messung eines vollständigen Satzes verträglicher Observabler in einem Zustand beschrieben durch den normierten Hilbert-Raum-Vektor $|\psi_i\rangle$ zu finden, ist

$$p_i = |\langle\psi_i|\phi\rangle|^2. \quad (13.75)$$

Dieses Wahrscheinlichkeitsmaß ist per Konstruktion nichtnegativ, es ergibt 0 (1) für ausgeschlossene (sichere) Ausgänge. Außerdem ist die Summe der Wahrscheinlichkeiten aller Zustände einer beliebigen Orthonormalbasis eins:

$$\sum_i p_i = \sum_i |\langle\psi_i|\phi\rangle|^2 = \sum_i \langle\phi|\psi_i\rangle\langle\psi_i|\phi\rangle = \langle\phi|\mathbb{1}|\phi\rangle = 1. \quad (13.76)$$

Es ist auch das *einzigste* Wahrscheinlichkeitsmaß, das nur vom aktuellen Zustand des Systems, $|\phi\rangle$, und vom Endzustand $|\psi_i\rangle$ abhängt. Dies ist der Inhalt von Gleasons Theorem (Andrew M. Gleason, 1957). Das Theorem wurde für allgemeine Ensemble-Zustände im Dichteoperator-Formalismus formuliert, gilt aber natürlich auch für den Spezialfall reiner Zustände.

Sei $\{|\psi_i\rangle\}$ eine Orthonormalbasis eines Hilbert-Raums \mathcal{H} der Dimension drei oder größer. Sei $\hat{P}_i := |\psi_i\rangle\langle\psi_i|$ der Projektionsoperator auf den Vektor $|\psi_i\rangle$. (Dann gilt $\sum_i \hat{P}_i = \mathbb{1}$.) Sei $f(\hat{P}_i)$ eine Abbildung von Projektionsoperatoren auf \mathcal{H} in das Intervall $[0, 1]$ mit der Eigenschaft

$$\sum_i f(\hat{P}_i) = 1. \quad (13.77)$$

Dann existiert ein positiver (hermitescher) Operator $\hat{\rho}$ auf \mathcal{H} mit Spur eins, so dass gilt

$$f(\hat{P}_i) = \text{Tr } \hat{\rho} \hat{P}_i. \quad (13.78)$$

Beachte, dass $\hat{\rho}$ genau die Eigenschaften eines Dichteoperators hat. Für einen reinen Zustand $|\phi\rangle$ ist $\hat{\rho} = |\phi\rangle\langle\phi|$ und damit

$$f(\hat{P}_i) = \text{Tr } |\phi\rangle\langle\phi| \hat{P}_i = |\langle\psi_i|\phi\rangle|^2. \quad (13.79)$$

Bemerkung: Es wird gefordert, dass $f(\hat{P}_i)$ nur von dem einen Zustand $|\psi_i\rangle$ abhängt, nicht davon, wie wir den Rest der Orthonormalbasis aufbauen, und auch nicht vom Messprotokoll. Die Messung muss nur geeignet sein, festzustellen, ob das System im Zustand $|\psi_i\rangle$ ist. Der Beweis ist erstaunlich kompliziert. Für einen zweidimensionalen Hilbert-Raum, d. h. für einen Spin der Länge 1/2 oder für ein Qubit, findet man auch Wahrscheinlichkeitsmaße, die sich nicht auf diese Weise durch einen Dichteoperator ausdrücken lassen.

Man kann den Standpunkt vertreten, dass durch die beiden Aussagen (a) „Für eine Beobachterin Alice ist ihre Zugehörigkeit zu einem der Zweige des Zustands eine Zufallsvariable (*self-localization uncertainty*).“ und (b) „Die Bornsche Regel liefert das einzige mögliche Wahrscheinlichkeitsmaß.“ die Bornsche Regel überzeugend begründet ist. Aber ganz zufrieden sind wir damit vielleicht nicht. Die Viele-Welten-Interpretation erklärt in sehr eleganter Weise, wieso uns die deterministische unitäre Zeitentwicklung der Quantentheorie als stochastischer Prozess erscheint. Es wäre wünschenswert, wenn sie ebenso elegant die zugehörigen Wahrscheinlichkeiten vorhersagen könnte. Stattdessen beruht letzteres auf einem mathematisch aufwendigen Beweis.

Auch in unserer Erinnerung finden wir, dass die Bornsche Regel erfüllt ist. Hier kann man einwenden, dass die meisten Menschen nicht viele Stern-Gerlach-Experimente tatsächlich durchgeführt haben. Wo im Alltagsleben bemerken wir die Gültigkeit der Bornschen Regel? Ein Beispiel ist das Sehen: Eine Oberfläche, die wir sehen, sendet Licht nach dem Huygensschen Prinzip i. A. als auslaufende Kugelwellen aus. Das ist unabhängig davon, ob es eine aktive Lichtquelle ist (Stern, Flamme, Glühfaden) oder eine diffus streuende beleuchtete Oberfläche. Laser fallen hier natürlich heraus. Quantenmechanisch können wir diese Kugelwellen in guter Näherung als Wellenfunktionen einzelner Photonen beschreiben. Die Photon-Photon-Wechselwirkung in Vakuum oder Luft ist sehr schwach, daher ist die Ein-Teilchen-Quantenmechanik ausreichend. Wenn wir ein Photon mit der Netzhaut registrieren, kollabiert nach der Kopenhagener Deutung diese Wellenfunktion zu einer stark lokalisierten Funktion, i. W. einer Ortseigenfunktion. Dies geschieht nach der Bornschen

Regel mit einer Wahrscheinlichkeit, die zum Betragsquadrat der Wellenfunktion am Ort der Registrierung proportional ist. Das Betragsquadrat entspricht der *Intensität* des Lichts. Wegen der Erhaltung der Wahrscheinlichkeit muss die Intensität einer Kugelwelle mit dem Abstand r von der Quelle wie $1/r^2$ abfallen. (Die Wellenfunktion muss also wie $1/r$ abfallen, vgl. die Einführung in die Streutheorie in der Vorlesung Quantentheorie 2.) Die Quantenmechanik sagt also voraus, dass die beobachtete und auch die objektiv gemessene Helligkeit einer Punktquelle proportional zu $1/r^2$ abfällt. Eine signifikante Abweichung davon, also von der Bornschen Regel, würden wir unmittelbar sehen. Eine kleinere, sogar sehr kleine, Abweichung wäre experimentell beobachtet worden. Im Ergebnis haben wir also sehr oft die Gültigkeit der Bornschen Regel überprüft.

Nach der Viele-Welten-Interpretation existieren aber Instanzen von uns, die jede überhaupt mögliche Abfolge von Beobachtungen gemacht haben. Das enthält zahlreiche völlig unplausible Situationen, z. B. Beobachter, die konsistent weiter entfernte Objekte als heller als nähere Objekte wahrnehmen. Diese tragen in der Superposition eine extrem kleine Amplitude, aber eben nicht Null. Es stellt sich also die Frage, wieso wir nicht solche Beobachter sind, die sich an ganz merkwürdige Beobachtungen erinnern. Man würde gern antworten, dass die Entscheidung, welche Instanz von uns wir sind, eine zufällige Auswahl, gewichtet mit dem Amplitudenquadrat, ist. Dann ist unsere Erinnerung zwar mit hoher Wahrscheinlichkeit mit der Bornschen Regel vereinbar, aber das Problem wurde nur verlagert: Wir müssen erklären, wieso die richtigen Gewichte denn die Amplitudenquadrate sind.

13.8 Weitere Interpretationen

In diesem Abschnitt erwähnen wir sehr knapp noch zwei weitere Interpretationen, nämlich die Idee der Führungswellen und den realistischen Kollaps von Ghirardi, Rimini und Weber. Diese beiden Interpretationen zeichnen den Ortsraum aus, haben aber sonst nicht viel miteinander gemein.

13.8.1 De Broglie-Bohm-Führungswellen

De Broglie (1927) und unabhängig davon Bohm (1952) haben eine Interpretation der Quantenmechanik vorgeschlagen, in der neben der Wellenfunktion auch die klassisch gedachten Trajektorien von Teilchen auftreten. Wir betrachten hier der Einfachheit halber ein System aus N Teilchen gleicher Art, insbesondere mit gleichen Massen. Seine Dynamik wird beschrieben durch einerseits die Schrödinger-Gleichung

$$i\hbar \frac{\partial \Psi}{\partial t} = -\frac{\hbar^2}{2m} \sum_{i=1}^{3N} \frac{\partial^2 \Psi}{\partial x_i^2} + V(x_1, \dots, x_{3N}) \Psi \quad (13.80)$$

für die Wellenfunktion $\Psi(x_1, \dots, x_{3N}, t)$ und die *Führungsgleichung* für die Teilchen,

$$\frac{dx_i}{dt} = \frac{1}{m} \frac{\partial S}{\partial x_i} \quad (13.81)$$

mit der Phase der Wellenfunktion, $S(x_1, \dots, x_{3N}, t)$, die definiert ist über

$$\Psi(x_1, \dots, x_{3N}, t) = R(x_1, \dots, x_{3N}, t) \exp\left(\frac{i}{\hbar} S(x_1, \dots, x_{3N}, t)\right). \quad (13.82)$$

Wir sehen, dass S die Rolle der *Wirkung* für die klassische Bewegung der Teilchen spielt: Die kanonischen Impulse sind gemäß 4.3.2 gegeben durch

$$p_i = \frac{\partial S}{\partial x_i}. \quad (13.83)$$

Gleichung (13.81) ist also einfach die Beziehung $v_i = p_i/m$ für nichtrelativistische Teilchen. Die Bewegung der Teilchen wird also von der Wellenfunktion determiniert, daher die Bezeichnung *Führungswelle*. Die Interpretation wird manchmal als Beispiel für eine Theorie mit verborgenen Variablen angeführt, wobei die Teilchenorte diese Variablen seien. Aber diese Orte sind von völlig anderem Charakter als die verborgenen Variablen im Sinne

von Einstein *et al.* oder Bell. Bei de Broglie und Bohm wird die Zeitentwicklung des Zustands allein durch die Schrödinger-Gleichung für Ψ beschrieben, die insbesondere nicht von den Teilchenorten abhängt. Der Zustandsvektor Ψ determiniert dann die Bewegung der Teilchen.

Die Orte der Teilchen mögen nun einer gewissen Verteilung $\rho(x_1, \dots, x_{3N}, t)$ im Konfigurationsraum \mathbb{R}^{3N} gehorchen – damit ist man notwendig bei einer statistischen Deutung. Dann kann man folgendes zeigen: Wenn zu einem Zeitpunkt t_0 gilt $\rho = |\Psi|^2$, so gilt diese Identität für alle Zeiten, insbesondere für alle $t > t_0$. Dann erhält man für alle Observablen, die von den Orten der Teilchen abhängen – das schließt auch Impulse ein, die sich aus $p_i = m dx_i/dt$ ergeben, und damit kinetische Energien – dieselben Ergebnisse aus den klassischen Trajektorien $x_i(t)$ wie aus der Standard-Quantenmechanik. Die De Broglie-Bohm-Führungswellen-Theorie macht also dieselben Voraussagen wie die Standard-Quantenmechanik. Man kann nicht experimentell zwischen den beiden unterscheiden und es geht insofern „nur“ um Interpretation.

Die wesentliche Voraussetzung für die Äquivalenz der Beschreibung ist, dass wirklich zu einem Zeitpunkt $\rho = |\Psi|^2$ gilt. Dies nennt man die *Quantengleichgewichtshypothese*. Sie wirkt *ad hoc* und es stellt sich die Frage, wie sie begründet werden kann.

Weitere Kritikpunkte sollen kurz genannt werden: Die De Broglie-Bohm-Interpretation ist nicht sparsam, da sie dieselbe Vorhersagekraft wie die Standard-Quantenmechanik hat, aber als zusätzliche Entitäten die Teilchen und ihre Trajektorien enthält. Sie beschreibt in der ursprünglichen Form nur Bewegung im Raum und der Spin der Teilchen findet daher keinen natürlichen Platz. Die relativistische Formulierung erweist sich ebenfalls als schwierig – vermutlich mit dem vorigen Punkt zusammenhängend, weil der Spin (von Fermionen) durch die relativistische Quantentheorie erzwungen wird. Und schließlich: Die de Broglie-Bohm-Interpretation steht in der Tradition der Ensemble-Interpretation, da sie von der Verteilungsfunktion ρ der Teilchenorte spricht, und hat dieselben Schwierigkeiten wie diese, Einzelprozesse zu verstehen.

13.8.2 GRW-Interpretation: Realistischer Kollaps

Eine von der Kopenhagener und der Viele-Welten-Interpretation radikal abweichende Sichtweise wurde von Ghirardi, Rimini und Weber (GRW) im Jahr 1986 vertreten: Sie interpretieren den Zustandskollaps nicht als scheinbaren Effekt einer in Wahrheit unitären Zeitentwicklung und auch nicht als bloßen Ausdruck veränderter Kenntnis eines Beobachters, sondern als real. Nach Ghirardi *et al.* ist die Zeitentwicklung also wirklich nicht unitär. Stattdessen postulieren sie eine Dynamik nah an der Kopenhagener Deutung, aber nun als real verstanden. Für ein System aus N Teilchen postulieren Ghirardi *et al.* folgende Dynamik:

- Der Zustand entwickelt sich unitär gemäß der Schrödinger-Gleichung, aber
- zu *zufälligen* Zeiten geht der Zustand eines der Teilchen in ein Gauß-Paket

$$\varphi(\vec{r}) = \left(\frac{\alpha}{\pi}\right)^{3/4} \exp\left(-\frac{\alpha}{2}(\vec{r} - \vec{r}_0)^2\right) \quad (13.84)$$

mit gemäß der Wahrscheinlichkeitsdichte $|\psi(\vec{r}, t)|^2$ *zufällig* gewähltem Zentrum \vec{r}_0 über; dies geschieht mit einer Rate (Übergangswahrscheinlichkeit pro Zeiteinheit) $1/\tau$ pro Teilchen, für N Teilchen ist die Gesamtrate daher N/τ .

Der reale Kollaps ist hier also eine *stochastische* Ergänzung zur unitären Dynamik. Dafür werden zwei neue Naturkonstanten benötigt, nämlich die Kollapsrate

$$\frac{1}{\tau} \approx 10^{-15} \text{ s} \quad (13.85)$$

und die Breite des Gauß-Pakets,

$$\alpha \approx 10^{-7} \text{ m}. \quad (13.86)$$

Die Vertreter dieser Idee argumentieren, dass wir aufgrund der Kleinheit von $1/\tau$ für mikroskopische Systeme praktisch nie einen Kollaps beobachten. Die charakteristische Zeit ist sehr lang: $\tau \approx 3,2 \times 10^7$ Jahre. Für makroskopische Systeme aus von der Größenordnung 10^{23} Teilchen findet hingegen sehr schnell ein Kollaps statt: $\tau/N \approx 10^{-8}$ s. Da die Teilchen gekoppelt (verschränkt?) sind, führt der Kollaps des Zustandes eines Teilchens zum Kollaps des Gesamtzustandes.

Der wichtigste Kritikpunkt an der GRW-Interpretation ist, dass sie *ad hoc* einen Mechanismus und zwei neue Naturkonstanten einführt, die nur dazu dienen, das Messproblem zu lösen. Kritisch ist auch, dass sie den Ortsraum und damit die Ortsdarstellung gegenüber allen anderen Basen auszeichnet – der Kollaps erfolgt ausdrücklich in ein Gauß-Paket im Ortsraum. Der Spin von Teilchen wird nicht auf natürliche Weise beschrieben. Die Gauß-Pakete haben $\langle \vec{r} \rangle = \vec{r}_0$, aber $\langle \vec{p} \rangle = 0$. Sie sind also im Mittel in Ruhe. Dies führt offensichtlich auch zu einem Konflikt mit dem Relativitätsprinzip, also der Äquivalenz von Inertialsystemen, schon bei Newton, nicht erst bei Einstein. Weiter postulieren Ghirardi *et al.* die Existenz eines realen Zufalls in den fundamentalen Naturgesetzen. Das ist schwer zu deuten. Und schließlich würden die Vertreter einer Viele-Welten-Interpretation sagen, dass Ghirardi *et al.* ein Problem zu lösen versuchen, das gar nicht existiert, nämlich dass wir nie makroskopische Superpositionen beobachten.