

# Theorie A

# Quantenmechanik

Sommersemester 2011

KLAUS FREDENHAGEN

II. Institut für Theoretische Physik  
Universität Hamburg



## Überblick

Thema dieser Vorlesung ist die Quantenmechanik. Die Quantenmechanik ist eine physikalische Theorie, die unsere Auffassung von Realität sehr stark verändert hat. Trotz ihrer oft antiintuitiven Aussagen hat die Quantenmechanik in den letzten Jahrzehnten in zunehmendem Maße auch praktische Anwendungen gefunden. Ein wesentlicher Teil der modernen Technologie, insbesondere die Halbleiter und Laser, beruhen direkt auf Anwendungen der Quantenmechanik. Sogar die sogenannten Paradoxien der Quantenmechanik werden voraussichtlich in der nahen Zukunft im Rahmen der Quanteninformationstheorie praktisch benutzt werden.

Diese Vorlesung richtet sich an zukünftige Lehrer an Gymnasien. Diese stehen vor der schwierigen Aufgabe, Schülern ein Grundverständnis der Quantenmechanik zu vermitteln. Eine Hauptschwierigkeit dabei ist, dass die Quantenmechanik nur mit Hilfe angemessener mathematischer Begriffe richtig erfasst werden kann; diese mathematischen Begriffe stehen aber den Schülern noch nicht zur Verfügung.

Ziel dieser Vorlesung ist es, den Teilnehmern die Grundprinzipien der Quantenmechanik auf dem heutigen Stand der Forschung zu vermitteln. Dazu werden einige mathematische Voraussetzungen benötigt, die, soweit noch nicht vorhanden, in der Vorlesung bereitgestellt werden. Auf der Grundlage dieses eigenen Verständnisses sollen die Teilnehmer danach in der Lage sein, Schülern die Grundlagen der Quantenmechanik in einer vereinfachten, aber möglichst korrekten Form zu vermitteln.

Wir werden in dieser Vorlesung zunächst einige wichtige Phänomene besprechen, zu deren Erklärung Quantenmechanik wesentlich ist. Wir werden dann die formale Struktur der Theorie entwickeln und parallel dazu die benötigten mathematischen Begriffe einführen. Im dritten Teil der Vorlesung werden wir einige wichtige Anwendungen behandeln.

## Literatur

- Karl Schilcher: Theoretische Physik kompakt für das Lehramt, Oldenbourg 2010
- Klaus Fredenhagen: Quantenmechanik. Vorlesungsskript Hamburg 2008,  
<http://unith.desy.de/research/aqft>
- Peter Schmüser: Theoretische Physik für Studierende des Lehramts; Teil A: Quantenmechanik  
<http://www.schul-physik.de/downloads/TheorieA.pdf>
- Gernot Münster: Quantentheorie  
2. Auflage, De Gruyter 2010

- Torsten Fließbach: Quantenmechanik  
Lehrbuch zur Theoretischen Physik III  
5. Auflage, Spektrum 2008
- Franz Schwabl: Quantum Mechanics. Second revised edition.  
Springer 1995
- Claude Cohen-Tannoudji, Bernard Diu, Frank Laloe: Quan-  
tenmechanik 1 und 2. De Gruyter 1999

## Inhaltsverzeichnis

Kapitel I. Quantenphänomene	7
1. Hohlraumstrahlung	7
2. Photoelektrischer Effekt	8
3. Wasserstoffspektrum	9
4. Compton-Versuch	10
5. Materiewellen	11
6. Stern-Gerlach-Versuch	14
Kapitel II. Der mathematische Rahmen der Quantentheorie	15
1. Der Spin	15
2. Endlich dimensionale Quantensysteme	20
3. Kompatibilität und Unschärferelationen	24
4. Der Dirac-Formalismus	26
5. Schrödingerbild und Heisenbergbild; der Zeitentwicklungsoperator	27
6. Spinoperatoren im $\mathbb{C}^n$	28
7. Unabhängige Teilsysteme; das Tensorprodukt	30
8. Reine und gemischte Zustände; die Dichtematrix	34
9. Verschränkung (Entanglement), EPR Paradox und Bellsche Ungleichung	36
Kapitel III. Unendlich dimensionale Quantensysteme	41
1. Unendlich dimensionale Hilberträume	41
2. Teilchen im äußeren Feld	45
3. Der harmonische Oszillator	47
4. Das freie Teilchen in einer Dimension	50
5. Stufenpotentiale	50
6. Der Bahndrehimpuls	54
7. Die radiale Schrödingergleichung	56
8. Bindungszustände im Coulombpotential	57
9. Mehrteilchensysteme	58
10. Störungstheorie	61



## KAPITEL I

# Quantenphänomene

### 1. Hohlraumstrahlung

Die klassische Physik beruht auf den Gesetzen der Newtonschen Mechanik und der Maxwellschen Elektrodynamik. Diese haben einen weiten Anwendungsbereich, führen aber bei der Anwendung auf die Struktur der Materie auf große Probleme. Insbesondere die chemische Bindung und das Auftreten diskreter Spektrallinien lassen sich im Rahmen der klassischen Physik nicht erklären. Interessanterweise wurde der erste Quanteneffekt aber bei einem ganz anderen Problem gefunden, nämlich bei der Beschreibung der Wärmestrahlung.

Ein Körper mit Temperatur  $T$  gibt Energie in Form von elektromagnetischen Wellen ab. Die Intensitätsverteilung der verschiedenen Frequenzen nimmt für den ideal schwarzen Körper (am besten realisiert durch einen Hohlraum mit einem kleinen Loch) eine universelle Form an, die nicht mehr von der Art des Körpers, sondern nur noch von seiner Temperatur abhängt. Die genaue Kenntnis dieser Verteilung ermöglicht es, die Temperatur zu messen, ohne den Körper zu berühren; das ist z.B. bei der Verhüttung von Stahl wichtig. Eine moderne Anwendung ist die Thermographie, mit der man schlecht isolierte Stellen an Hauswänden finden kann. Nach den Prinzipien der statistischen Mechanik sollte jeder Freiheitsgrad die mittlere Energie  $\frac{1}{2}kT$  ( $k$  Boltzmannkonstante) haben. Da das elektromagnetische Feld unendlich viele Freiheitsgrade besitzt, würde das eine unendliche Energie bedeuten, was offenbar nicht der Fall ist. Die Beobachtungen zeigten, dass die Intensität einer ebenen Welle mit Frequenz  $\nu$  für kleine Frequenzen proportional zur Temperatur  $T$  ist; dies ist im Einklang mit dem Gleichverteilungssatz der statistischen Mechanik. Dieser Satz kann aber nicht für hohe Frequenzen gelten, da dann die Gesamtintensität der Strahlung divergieren würde. Tatsächlich hat Wien ein Gesetz gefunden, das für hohe Frequenzen gültig ist. Danach ist die Intensität proportional zu  $\nu e^{-\frac{c\nu}{T}}$ , mit  $c = \frac{h}{k}$  mit der Planckschen Konstante  $h$ .<sup>1</sup>

Planck vermutete dann, dass die Verteilung für alle Frequenzen durch die interpolierende Funktion

$$w(\nu) = h\nu(e^{\frac{h\nu}{kT}} - 1)^{-1} \quad (1.1)$$

---

<sup>1</sup>Diese Angaben beziehen sich auf eine ebene Welle mit vorgegebener Ausbreitungsrichtung. Berücksichtigt man alle Richtungen, so muss der Ausdruck noch mit dem Faktor  $4\pi\nu^2$ , der Oberfläche einer Kugel mit Radius  $\nu$ , multipliziert werden.

gegeben ist. Diese Funktion hat das richtige Verhalten für kleine und große Frequenzen. Auch das Verhalten im Zwischenbereich wird richtig wiedergegeben. Aber wie lässt sich dieses Verhalten theoretisch begründen?

Planck beobachtete zunächst, dass die Funktion, die mit dem Faktor  $h\nu$  multipliziert wird, die Summe einer geometrischen Reihe ist,

$$(e^{\frac{h\nu}{kT}} - 1)^{-1} = \sum_{n=1}^{\infty} (e^{-\frac{h\nu}{kT}})^n . \quad (1.2)$$

Er postulierte dann, dass elektromagnetische Energie nur in Form von Portionen ("Quanten") der Größe  $E = h\nu$  abgegeben werden kann. Die relative Wahrscheinlichkeit von  $n$  solchen Emissionen wird durch den Boltzmannfaktor  $e^{-n\frac{E}{kT}}$  gegeben. Summation über  $n$  ergibt den Normierungsfaktor  $1 - e^{-\frac{h\nu}{kT}}$ . Die mittlere Energie einer ebenen Welle mit Frequenz  $\nu$  und vorgegebener Ausbreitungsrichtung ist dann

$$E = \sum_{n=1}^{\infty} nh\nu e^{-\frac{nh\nu}{kT}} (1 - e^{-\frac{h\nu}{kT}}) . \quad (1.3)$$

Die Summation über  $n$  kann mit Hilfe der Formel

$$\sum_{n=1}^{\infty} nq^n = q \frac{d}{dq} \sum_{n=0}^{\infty} q^n = q \frac{d}{dq} (1 - q)^{-1} = q(1 - q)^{-2} , \quad (1.4)$$

die für  $|q| < 1$  gilt, ausgeführt werden. Mit  $q = e^{-\frac{h\nu}{kT}}$  ergibt sich die gewünschte Formel.

## 2. Photoelektrischer Effekt

Der nächste Hinweis auf die Quantentheorie kam aus der Einsteinschen Analyse des photoelektrischen Effekts. Man beobachtet, dass bei Bestrahlung einer Metalloberfläche mit Licht Elektronen austreten. Dieser Effekt setzt ein, wenn die Frequenz  $\nu$  des Lichtes einen gewissen, von der Metallsorte abhängigen Wert  $\nu_g$  überschreitet. Die maximale kinetische Energie der austretenden Elektronen wird durch das Anlegen einer kompensierenden Gegenspannung  $U$  gemessen. Man findet den Zusammenhang

$$h(\nu - \nu_g) = eU . \quad (2.1)$$

Einstein erklärte diesen Effekt damit, dass das Licht aus Photonen der Energie  $h\nu$  besteht. Die Grenzfrequenz ergibt sich aus der Austrittsarbeit  $W$ , die das Elektron beim Verlassen des Metalls leisten muss,  $W = h\nu_g$ . Die überschüssige Energie wird dann zur kinetischen Energie des Elektrons. Der Photostrom kommt zum Erliegen, wenn die kinetische Energie nicht ausreicht, die Gegenspannung zu überwinden.

### 3. Wasserstoffspektrum

Die Spektrallinien von Wasserstoff folgen der Formel

$$\nu = \nu_0 \left( \frac{1}{n^2} - \frac{1}{m^2} \right) \quad (3.1)$$

mit natürlichen Zahlen  $n < m$  (Balmer-Formel). Aus Streuexperimenten von Rutherford ( $\alpha$ -Strahlen auf Goldfolie) wusste man, dass Atome aus einem positiv geladenem Kern, der nahezu punktförmig ist und fast die gesamte Masse des Atoms enthält, und negativ geladenen Elektronen besteht. Man vermutete daher, dass die Atome ähnlich wie Planetensysteme aufgebaut sind. Bei Planeten ist aber die Umlaufzeit (und daher auch die Frequenz) kontinuierlich veränderbar, sodass dieses Modell die beobachteten diskreten Frequenzen nicht erklären kann. Zusätzlich sollte ein Elektron, dass sich mit kurzer Umlaufzeit um den Kern bewegt, wie ein Sender elektromagnetischer Wellen wirken. Die Abschätzung der Abstrahlung ergibt eine sehr kurze Lebensdauer von Atomen, im Widerspruch zur beobachteten Stabilität. Eine Erklärung der beobachteten Linien wurde dann von Bohr gefunden. Am leichtesten lässt sich die Bohrsche Theorie verstehen, wenn man die später von de Broglie stammende Hypothese aufstellt, dass auch Elektronen Welleneigenschaften besitzen, insbesondere eine Wellenlänge, die sich nach der Einsteinschen Formel für Lichtquanten aus dem Impuls  $p$  durch  $\lambda = \frac{h}{p}$  ergibt. Nach Bohr besitzt das Wasserstoffatom nur solche stationären Bahnen, bei denen der Umfang ein Vielfaches der Wellenlänge ist. Die Idee dabei ist, dass sich nur dann eine stehende Welle bilden kann, wenn diese Bedingung erfüllt ist. Auf einer Kreisbahn mit Radius  $r$  ist die Zentrifugalkraft  $F_Z = m\dot{\varphi}^2 r$  gleich der Anziehungskraft durch das elektrische Feld des Kerns,  $F = \frac{e^2}{4\pi\epsilon_0 r^2}$ . Für die Winkelgeschwindigkeit gilt also

$$\dot{\varphi}^2 = \frac{e^2}{4\pi\epsilon_0 r^3 m} . \quad (3.2)$$

Mit  $p = mr\dot{\varphi}$  folgt

$$\lambda = \frac{h}{mr\dot{\varphi}} = \frac{h\sqrt{4\pi\epsilon_0 mr^3}}{mre} = \frac{h\sqrt{4\pi\epsilon_0 r}}{e\sqrt{m}} . \quad (3.3)$$

Die Quantisierungsbedingung ist

$$2\pi r = n\lambda \quad (3.4)$$

für eine natürliche Zahl  $n$ . Die erlaubten Radien der Bahnen sind daher

$$r_n = \frac{n^2 \hbar^2 4\pi\epsilon_0}{e^2 m} , \quad \hbar = h/2\pi . \quad (3.5)$$

Also haben die stationären Bahnen die Energien

$$E_n = \left( \frac{m\dot{\varphi}^2 r^2}{2} - \frac{e^2}{4\pi\epsilon_0 r} \right) \Big|_{r=r_n} = -\frac{R}{n^2} \quad (3.6)$$

mit der Rydberg-Konstanten  $R = \frac{me^4}{2(4\pi\epsilon_0)^2\hbar^2} \approx 13,6\text{eV}$ . Die beobachteten Frequenzen ergeben sich dann als Differenzen zweier Energie-Niveaus.

Diese Interpretation der Spektren wird durch den Franck-Hertz-Versuch bestätigt. Bei diesem Versuch beobachtet man die Energieverluste durch Stöße in einer Ionisationsröhre und findet gerade die durch die Bohrsche Theorie vorhergesagten Energie-Niveaus.

Eine der Balmer-Formel analoge Formel wird bei den Röntgenspektren schwerer Atome beobachtet. Röntgenspektren bestehen aus einem kontinuierlichen Teil, dem Bremsspektrum (dieses entsteht bei Abbremsung der Elektronen in der Röntgenröhre) und einem sogenannten charakteristischen Spektrum, das aus scharfen Linien besteht, die für die Atome auf der Anode charakteristisch sind. Für die sogenannte K-Serie, den kurzwelligsten Bereich des charakteristischen Spektrums, findet man das Moseleysche Gesetz

$$h\nu = R(Z - 1)^2\left(1 - \frac{1}{n^2}\right), \quad n = 2, 3, \dots \quad (3.7)$$

mit der Kernladungszahl  $Z$ . Dies kann man so interpretieren, dass das auftreffende Elektron ein Elektron aus der innersten Schale herausgeschlagen hat. Ein Elektron aus der  $n$ -ten Schale springt in die Lücke und sendet dabei eine elektromagnetische Welle aus. Hierbei wird die Kernladung durch das in der innersten Schale verbliebene Elektron etwas abgeschirmt, sodass man ein Zentralpotential mit Ladung  $(Z - 1)e$  sieht. Ein ähnliches Gesetz beobachtet man für die nächste Serie, die L-Serie, mit einer effektiven Kernladungszahl von  $Z - 7, 4$ .

#### 4. Compton-Versuch

Die Teilcheninterpretation elektromagnetischer Wellen wurde eindrucksvoll durch den Compton-Versuch bestätigt. Bei diesem Versuch beobachtet man die Streuung von Röntgenstrahlen an Elektronen. Merkwürdiger Weise findet man dabei Änderungen der Wellenlänge. Ordnet man den Photonen Energie und Impuls nach den Einsteinschen Formeln zu,

$$E = h\nu, \quad |\mathbf{p}| = \frac{h}{\lambda} \quad (4.1)$$

und interpretiert die Streuung als elastische Stöße, so folgt aus Energie und Impulserhaltung für die Änderung der Wellenlänge bei der Streuung an ursprünglich ruhenden Elektronen

$$\Delta\lambda = \frac{h}{mc}(1 - \cos\theta) \quad (4.2)$$

mit dem Ablenkungswinkel  $\theta$ .  $\frac{h}{mc}$  nennt man die Comptonwellenlänge des Teilchens. Man interpretiert sie als Maß dafür, wie genau der Ort eines Teilchens der Masse  $m$  bestimmt ist.

## 5. Materiewellen

Nachdem bei elektromagnetischen Wellen teilchenartiges Verhalten beobachtet worden war, kam de Broglie auf die Idee, dass vielleicht auch die bekannten Teilchen (damals vor allem die Elektronen) Wellenphänomene zeigen könnten. Wir haben bereits gesehen, dass diese Vorstellung eine natürliche Erklärung der Bohrschen Bahnen als stehende Wellen erlaubt. Identifiziert man die (Kreis-)Frequenz  $\omega = 2\pi\nu$  mit der Energie,  $E = \hbar\omega$ , und den Wellenzahlvektor  $\mathbf{k}$  (die Richtung von  $\mathbf{k}$  beschreibt die Ausbreitungsrichtung der Welle, der Betrag von  $\mathbf{k}$  ergibt sich aus der Wellenlänge  $\lambda$  durch  $|\mathbf{k}| = 2\pi/\lambda$ ) mit dem Impuls,  $\mathbf{p} = \hbar\mathbf{k}$ , so erhält man für ein nichtrelativistisches kräftefreies Teilchen der Masse  $m$  mit Impuls  $\mathbf{p}$  und Energie  $E = \frac{|\mathbf{p}|^2}{2m}$  eine ebene Welle

$$\psi(t, \mathbf{x}) = e^{-it\omega + i\mathbf{k}\cdot\mathbf{x}} \quad (5.1)$$

mit der Dispersionsrelation

$$\omega(\mathbf{k}) = \hbar \frac{|\mathbf{k}|^2}{2m} . \quad (5.2)$$

Überlagert man solche Wellen, so erhält man Interferenzeffekte, die auch tatsächlich einige Jahre später von Davisson und Germer nachgewiesen werden konnten. Eine Schwierigkeit aber ist die Verträglichkeit dieser Effekte mit den Teilcheneigenschaften. Berühmt ist die Interferenz am Doppelspalt, die als Gedankenexperiment diente. Eigenartiger Weise verschwindet das Interferenzbild, wenn man feststellt, durch welchen Spalt das Teilchen geflogen ist. Dieses Experiment ist ein deutlicher Hinweis darauf, dass die Quantenmechanik unsere Vorstellung von der der Beobachtung zugrunde liegenden Realität verändert.

Eine Schwierigkeit bei der Interpretation der Materiewellen ist die Bedeutung der Werte der Wellenfunktion. Die bis heute gültige Interpretation geht auf Bohr zurück und postuliert, dass das Absolutquadrat  $|\psi(t, \mathbf{x})|^2$  der Wellenfunktion die Wahrscheinlichkeitsdichte dafür ist, dass sich das Teilchen zur Zeit  $t$  am Ort  $\mathbf{x}$  aufhält.

Diese Interpretation hat weitreichende Konsequenzen. Sie besagt nämlich, dass bei genügend häufiger Wiederholung eines Experiments die beobachteten Häufigkeiten den berechneten Wahrscheinlichkeiten entsprechen. Für das einzelne Experiment wird aber, solange die Wahrscheinlichkeit kleiner als 1 ist, keine Aussage gemacht. Bis heute gibt es eine lebhaft diskutierte Deutung der Quantentheorie ausreicht.

Ein anderes Problem ist, dass die ebenen Wellen, denen man eindeutig Energie und Impuls zuordnen kann, zu einer konstanten Aufenthaltswahrscheinlichkeit führen. Dies ist aber in einem unendlich ausgedehnten Raum nicht möglich und entspricht auch nicht dem beobachteten Verhalten von Teilchen. Man gibt daher die Bedingung auf, dass

Energie und Impuls scharfe Werte annehmen, und betrachtet Wellenpakete, das sind kontinuierliche Überlagerungen ebener Wellen,

$$\psi(t, \mathbf{x}) = \int d^3\mathbf{k} \varphi(\mathbf{k}) e^{-it\frac{\hbar|\mathbf{k}|^2}{2m} + i\mathbf{k}\cdot\mathbf{x}} \quad (5.3)$$

Die durch  $\psi$  beschriebene Aufenthaltswahrscheinlichkeit hängt von den Eigenschaften der Amplitudenfunktion  $\varphi$  ab. Grob gesagt gilt: je glatter  $\varphi$  ist, umso besser ist das Wellenpaket konzentriert, je besser  $\varphi$  konzentriert ist, umso glatter wird  $\psi$ . Dies sind allgemeine Eigenschaften der Fourier-Transformation. Auch die geradlinige Bahnkurve des kräftefreien Teilchens spiegelt sich in einer Eigenschaft des Wellenpakets. Hierzu nehmen wir an, dass  $\varphi$  eine unendlich oft differenzierbare (kurz: glatte) Funktion ist, die verschwindet, wenn  $\mathbf{k}$  von einem vorgegebenen Wellenzahlvektor  $\mathbf{k}_0$  zu stark abweicht,  $\varphi(\mathbf{k}) = 0$  für  $|\mathbf{k} - \mathbf{k}_0| < \epsilon$ . Für große Werte von  $t$  oszilliert der Integrand in (5.3) sehr schnell, sodass das Integral nahezu verschwindet. Das gilt aber nicht an den stationären Punkten des Exponenten, d.h. an den Punkten  $\mathbf{k}$ , an denen die Ableitung von  $\mathbf{k}\cdot\mathbf{x} - \frac{\hbar|\mathbf{k}|^2 t}{2m}$  nach  $\mathbf{k}$  verschwindet. Die Bedingung für stationäre Punkte ist

$$\mathbf{x} = \frac{\hbar\mathbf{k}}{m} t . \quad (5.4)$$

Diese Bedingung ist nur erfüllt, wenn  $\frac{m\mathbf{x}}{\hbar t}$  genügend dicht bei  $\mathbf{k}_0$  liegt. Wir schließen also, dass die Aufenthaltswahrscheinlichkeit des Teilchens zu späten Zeiten  $t$  im wesentlichen in der Nähe des Punktes  $\mathbf{x} = \frac{\hbar\mathbf{k}_0}{m} t$  konzentriert ist, das Wellenpaket beschreibt also näherungsweise die Bahnkurve eines Teilchens mit Geschwindigkeit  $\frac{\hbar\mathbf{k}_0}{m}$ . Allerdings wächst dabei die Unschärfe mit der Zeit an, sodass man aus dem asymptotischen Verhalten keine Rückschlüsse auf den Ort zur Zeit  $t = 0$  ziehen kann.

Wir wollen untersuchen, wie nahe die Materiewellenbeschreibung an der klassischen Beschreibung ist. In der klassischen Mechanik lässt sich die Bewegung eines Massenpunktes eindeutig vorhersagen, wenn zum Zeitpunkt  $t = 0$  Ort und Geschwindigkeit bekannt sind. Bei den Materiewellen sind Orts- und Geschwindigkeitsverteilung durch die Fouriertransformation verknüpft. Ist die Amplitudenfunktion  $\varphi$  bis auf  $\epsilon$  bei einem Wellenzahlvektor  $\mathbf{k}_0$  konzentriert, so ist die Ortsraumwellenfunktion zur Zeit  $t = 0$  in etwa bis auf  $1/\epsilon$  konzentriert. Der Ort zur Zeit  $t$  ist dann nach dem Argument der stationären Phase in einem Gebiet konzentriert, das mit den Geschwindigkeiten  $\mathbf{v} = \frac{\hbar\mathbf{k}}{m}$ ,  $|\mathbf{k} - \mathbf{k}_0| < \epsilon$  in der Zeit  $t$  erreicht werden kann. Der Radius dieses Gebietes ist

$$r(t) = \frac{1}{\epsilon} + |t| \frac{\epsilon\hbar}{m} . \quad (5.5)$$

Der Radius wird minimal, wenn  $1/\epsilon = \sqrt{|t|\hbar/m}$  gewählt wird. Die optimale Lokalisation zur Zeit  $t$  bei Festlegung von Ort und Geschwindigkeit zur Zeit 0 ist also

$$r(t) = 2\sqrt{\frac{\hbar|t|}{m}}. \quad (5.6)$$

Wir wollen dieses Resultat an einigen Beispielen illustrieren. Die Plancksche Konstante ist

$$\hbar = 1,055 \cdot 10^{-34} \text{Js}, \quad (5.7)$$

die Masse eines Elektrons ist

$$m_e = 9,11 \cdot 10^{-31} \text{kg}. \quad (5.8)$$

Nach einer Sekunde ist daher der Ort nur noch bis auf

$$r(t) = 2,1 \text{cm} \quad (5.9)$$

bestimmt. Die Periode bei einem gebundenen Elektron in einer Bohrschen Bahn ist

$$T = n^2 \frac{h}{R} = n^2 \cdot 3,04 \cdot 10^{-16} \text{s} \quad (5.10)$$

nach einer Periode ist also die Ortsunschärfe in etwa

$$r(T) = 3n \cdot 10^{-10} \text{m}. \quad (5.11)$$

Der Radius der Bohrschen Bahnen beträgt

$$r_n = n^2 \cdot 0,5 \cdot 10^{-10} \text{m}. \quad (5.12)$$

Wir schließen, dass für kleine  $n$  der Begriff der Bahn nicht mehr sinnvoll ist. Für große  $n$  aber (sogenannte Rydberg-Atome) kann näherungsweise von Bahnen gesprochen werden.

Ersetzen wir das Elektron durch ein Fulleren-Molekül ( $C_{60}$ ), so ist die Masse  $m = 1,2 \cdot 10^{-24} \text{kg}$ , und die Unschärfe nach einer Sekunde ist in etwa  $2 \cdot 10^{-5} \text{m}$ . Bei einem makroskopischen Körper mit der Masse 10g erhalten wir nur eine Abweichung von  $2 \cdot 10^{-16} \text{m}$ .

Zum Abschluss dieses Abschnitts wollen wir das Doppelspaltexperiment etwas genauer betrachten. Wir nehmen der Einfachheit halber an, dass zwei Kugelwellen gleicher Frequenz, die von zwei Punkten  $\mathbf{x}_1$  und  $\mathbf{x}_2$  ausgehen, superponiert werden. Die Wellenfunktion sei

$$\psi(t, \mathbf{x}) = e^{-i\omega t} \left( \frac{e^{ik|\mathbf{x}-\mathbf{x}_1|}}{|\mathbf{x}-\mathbf{x}_1|} + \frac{e^{ik|\mathbf{x}-\mathbf{x}_2|}}{|\mathbf{x}-\mathbf{x}_2|} \right). \quad (5.13)$$

Die Wahrscheinlichkeitsdichte ergibt sich zu

$$|\psi(t, \mathbf{x})|^2 = \frac{1}{r_1^2} + \frac{1}{r_2^2} + \frac{2}{r_1 r_2} \cos k(r_1 - r_2) \quad (5.14)$$

mit  $r_1 = |\mathbf{x} - \mathbf{x}_1|$  und  $r_2 = |\mathbf{x} - \mathbf{x}_2|$ .

Im Experiment schickt man ein einzelnes Teilchen durch den Doppelspalt. Auf einem Beobachtungsschirm wird das Teilchen an einem

Punkt nachgewiesen. Bei genügend häufiger Wiederholung des Experiments stellt sich eine Häufigkeitsverteilung ein, die durch die oben angegebene Verteilung gegeben ist. Modifiziert man das Experiment, sodass bekannt ist, durch welchen Spalt das Teilchen geflogen ist, so verschwindet der Interferenzterm.

## 6. Stern-Gerlach-Versuch

Im Stern-Gerlach-Versuch betrachtet man die Bewegung eines Strahls von Silberatomen unter dem Einfluss eines inhomogenen Magnetfelds. Die Atome besitzen ein magnetisches Moment. Daher wirkt auf die Atome eine Kraft, die von der Orientierung des Moments abhängt. Zusätzlich wirkt auf das magnetische Moment ein Drehmoment, das senkrecht auf dem magnetischen Moment und dem Magnetfeld steht und dadurch zu Drehungen um die Richtung des Magnetfeldes führt. Der zeitliche Mittelwert des Moments zeigt daher in die Richtung des Magnetfeldes, und je nach dem Wert des Mittelwerts sollte der Strahl durch die Inhomogenität des Feldes abgelenkt werden. Der überraschende Ausgang des Experiments ist, dass der Strahl in genau zwei Teilstrahlen aufspaltete; dies deutet darauf hin, dass das magnetische Moment nur 2 mögliche Orientierungen relativ zum Magnetfeld annehmen kann. Dies ist merkwürdig, da das Magnetfeld ganz beliebig gewählt werden kann. Das magnetische Moment kann danach in einem Magnetfeld nur die beiden Werte  $\pm\mu_B$  besitzen, mit dem Bohrschen Magneton  $\mu_B = \frac{\hbar e}{2m}$ .

## KAPITEL II

# Der mathematische Rahmen der Quantentheorie

### 1. Der Spin

Wir betrachten zunächst den Stern-Gerlach-Versuch. Wir benötigen einen Formalismus, der einerseits kontinuierliche Transformationen zulässt (hier die Veränderung der Orientierung des Magnetfeldes) und andererseits erklärt, warum nur zwei mögliche Messergebnisse auftreten. Die grundlegende Idee dazu ist Heisenberg zu verdanken. Er postulierte, dass die Observablen der Quantentheorie nicht durch zahlenwertige Funktionen auf den Zuständen des Systems gegeben sind, sondern durch die Gesamtheit der Übergangsamplituden zwischen den diskreten Zuständen. Gibt es wie im Stern-Gerlach-Versuch nur zwei Zustände, so stellt man die Observablen durch  $2 \times 2$ -Matrizen

$$A = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix} \quad (1.1)$$

dar. Beim Produkt zweier Observablen werden die Übergangsamplituden nach der Formel

$$(AB)_{jk} = \sum_{l=1}^2 A_{jl} B_{lk} \quad (1.2)$$

kombiniert; das ist nichts anderes als das Produkt der beiden Matrizen, eine mathematische Struktur, die Heisenberg damals noch nicht kannte. Die Spinkomponente in Richtung des Magnetfeldes kann dann als Diagonalmatrix dargestellt werden,

$$S_3 = \frac{\hbar}{2} \sigma_3, \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \quad (1.3)$$

Sie beschreibt die möglichen Werte  $\pm\hbar/2$  der  $z$ -Komponente des Spins.

Wie können jetzt die dazu senkrechten Komponenten beschrieben werden? Offenbar kann man ihnen nicht direkt Zahlenwerte zuordnen. Stattdessen postuliert man, dass sie Übergänge zwischen den beiden auftretenden Zuständen verursachen. Die einfachste Möglichkeit ist die Matrix

$$S_1 = \frac{\hbar}{2} \sigma_1, \sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad (1.4)$$

eine weitere die Matrix

$$S_2 = \frac{\hbar}{2} \sigma_2, \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}. \quad (1.5)$$

Die Matrizen  $\sigma_1, \sigma_2, \sigma_3$  werden die Pauli-Matrizen genannt. Offenbar kommt man nicht mit reellen Matrizen aus. Die Matrizen sind stattdessen hermitesch, d.h.  $A = A^*$ , wobei die adjungierte Matrix einer komplexen Matrix definiert ist durch Transposition und komplexe Konjugation der Matrixelemente,

$$(A^*)_{jk} = \overline{A_{kj}} . \quad (1.6)$$

Die Adjunktion erfüllt die folgenden Bedingungen

$$(A + B)^* = A^* + B^* \quad (1.7)$$

$$(\alpha A)^* = \overline{\alpha} A^* \quad (1.8)$$

$$(AB)^* = B^* A^* \quad (1.9)$$

$$(A^*)^* = A \quad (1.10)$$

Ihre Produkte sind

$$\sigma_1^2 = \sigma_2^2 = \sigma_3^2 = \mathbf{1} , \quad (1.11)$$

$$\sigma_1 \sigma_2 = i \sigma_3 = -\sigma_2 \sigma_1 , \quad \sigma_2 \sigma_3 = i \sigma_1 = -\sigma_3 \sigma_2 , \quad \sigma_3 \sigma_1 = i \sigma_2 = -\sigma_1 \sigma_3 . \quad (1.12)$$

Jede  $(2 \times 2)$ -Matrix mit komplexen Einträgen lässt sich in der Form

$$A = a^0 \mathbf{1} + a^1 \sigma_1 + a^2 \sigma_2 + a^3 \sigma_3 = \begin{pmatrix} a^0 + a^3 & a^1 - i a^2 \\ a^1 + i a^2 & a^0 - a^3 \end{pmatrix} \quad (1.13)$$

darstellen, mit eindeutig bestimmten komplexen Zahlen  $a^0, a^1, a^2, a^3$  und der Einheitsmatrix  $\mathbf{1}$ . Wir schreiben  $(a^1, a^2, a^3) = \mathbf{a}$  und  $a^1 \sigma_1 + a^2 \sigma_2 + a^3 \sigma_3 = \mathbf{a} \cdot \vec{\sigma}$ . Wir können das Produkt zweier beliebiger  $(2 \times 2)$ -Matrizen in der Form

$$(a^0 \mathbf{1} + \mathbf{a} \cdot \vec{\sigma})(b^0 \mathbf{1} + \mathbf{b} \cdot \vec{\sigma}) = a \cdot b \mathbf{1} + (a^0 \mathbf{b} + \mathbf{a} b^0 + i \mathbf{a} \times \mathbf{b}) \cdot \vec{\sigma} \quad (1.14)$$

darstellen. Wählen wir insbesondere ein Tripel paarweise orthogonaler Einheitsvektoren  $\mathbf{n}_1, \mathbf{n}_2, \mathbf{n}_3 \in \mathbb{R}^3$  mit  $\mathbf{n}_1 \times \mathbf{n}_2 = \mathbf{n}_3$  (Rechtssystem), so erfüllen die Matrizen

$$\sigma'_1 = \mathbf{n}_1 \cdot \vec{\sigma} , \quad \sigma'_2 = \mathbf{n}_2 \cdot \vec{\sigma} , \quad \sigma'_3 = \mathbf{n}_3 \cdot \vec{\sigma} \quad (1.15)$$

die gleichen algebraischen Relationen wie die ursprünglichen Pauli-Matrizen. Eine Drehung, die die Einheitsvektoren in die  $x, y$  und  $z$ -Richtung in die Einheitsvektoren  $\mathbf{n}_1, \mathbf{n}_2$  und  $\mathbf{n}_3$  überführt, wirkt auf eine beliebige  $(2 \times 2)$ -Matrix durch

$$a^0 \mathbf{1} + \mathbf{a} \cdot \vec{\sigma} \mapsto a^0 \mathbf{1} + \mathbf{a} \cdot \vec{\sigma}' . \quad (1.16)$$

Diese Transformation ist ein Isomorphismus der Algebra der  $(2 \times 2)$ -Matrizen.

Wir betrachten jetzt die durch das magnetische Feld verursachte Zeitentwicklung des Systems. Ein Magnetfeld  $\mathbf{B}$  übt auf ein magnetisches Moment  $\vec{\mu}$  ein Drehmoment  $\mathbf{N} = \vec{\mu} \times \mathbf{B}$  aus. Ist das magnetische Moment ein Vielfaches des Spins (wie beim für den Stern-Gerlach-Versuch entscheidenden äußeren Elektron),

$$\vec{\mu} = \gamma \mathbf{S} \quad (1.17)$$

mit dem gyromagnetischen Faktor des Elektrons  $\gamma = -g \frac{e}{2m}$ ,  $g \approx 2$ , so ergibt sich für die Zeitableitung des Spins

$$\frac{d}{dt} \mathbf{S} = \gamma \mathbf{B} \times \mathbf{S} . \quad (1.18)$$

Die Energie eines magnetischen Moments im Magnetfeld ist

$$H = -\vec{\mu} \cdot \mathbf{B} . \quad (1.19)$$

Wir identifizieren  $H$  mit der Matrix

$$H = -\gamma \frac{\hbar}{2} \mathbf{B} \cdot \vec{\sigma} \quad (1.20)$$

(„Hamiltonoperator“) und finden die folgende Beziehung (Heisenberg-Gleichung)

$$\frac{d}{dt} \mathbf{S} = \frac{i}{\hbar} [H, \mathbf{S}] \quad (1.21)$$

wobei die eckigen Klammern den Kommutator  $[A, B] = AB - BA$  zweier Matrizen bezeichnen.

Wir haben hiermit das physikalische System des Elektrons einschließlich seiner zeitlichen Entwicklung durch die Algebra der  $(2 \times 2)$ -Matrizen beschrieben. Doch wie ist dieses System den Beobachtungen zugänglich?

Dirac hat für die Observablen der Quantentheorie den einprägsamen Begriff „q-Zahlen“ geprägt, im Gegensatz zu den „c-Zahlen“, die für die klassische Physik verwendet werden. Um die c-Zahlen zu bestimmen, die einer q-Zahl entsprechen, betrachtet man die Gleichungen, die die q-Zahlen erfüllen. Im Fall des Spins erfüllen die Matrizen  $\mathbf{n} \cdot \vec{\sigma}$  mit  $|\mathbf{n}| = 1$  die Gleichung

$$(\mathbf{n} \cdot \vec{\sigma})^2 = \mathbf{1} . \quad (1.22)$$

Die beiden c-Zahl-Lösungen dieser Gleichung sind  $\pm 1$ . Diese entsprechen den beiden möglichen Werten, die eine Komponente des Spins annehmen kann.

Wir haben gesehen, dass beim Stern-Gerlach-Versuch die durch das Magnetfeld ausgezeichnete Komponente des Spins die beiden Werte  $\pm \hbar/2$  annehmen kann. Welcher Wert bei einem einzelnen Atom auftritt, ist offenbar zufällig, und bei einem unpräparierten Strahl ist die Wahrscheinlichkeit für beide Einstellungen gleich. Man kann jetzt bei einem der Teilstrahlen einen erneuten Stern-Gerlach-Versuch unternehmen. Das Resultat hängt dann von der Ausrichtung der beiden Magnetfelder ab. Das Ergebnis lässt sich wie folgt beschreiben:

Seien  $\mathbf{n}_1, \mathbf{n}_2$  die Richtungsvektoren des ersten, bzw. des zweiten Magnetfeldes. Für die Wahrscheinlichkeit, dass der Spin, der nach dem ersten Versuch in Richtung  $\mathbf{n}_1$  zeigt, beim zweiten Versuch in Richtung  $\mathbf{n}_2$  zeigt, findet man die Wahrscheinlichkeit

$$W(\mathbf{n}_2|\mathbf{n}_1) = \frac{1}{2}(1 + \mathbf{n}_2 \cdot \mathbf{n}_1) = \frac{1}{2}(1 + \cos \theta) = \cos^2 \frac{\theta}{2} \quad (1.23)$$

mit dem Winkel  $\theta$  zwischen den beiden Vektoren. Insbesondere ist die Wahrscheinlichkeit gleich 1, wenn die Richtungen übereinstimmen, und gleich  $\frac{1}{2}$ , wenn sie senkrecht aufeinander stehen.

Sei  $P(\mathbf{n}) = \frac{1}{2}(\mathbf{1} + \mathbf{n} \cdot \vec{\sigma})$  mit einem Einheitsvektor  $\mathbf{n} \in \mathbb{R}^3$ . Die Matrizen  $P(\mathbf{n})$  haben die Eigenschaften

$$P(\mathbf{n})^2 = P(\mathbf{n}) , \quad (1.24)$$

$$P(\mathbf{n})^* = P(\mathbf{n}) , \quad (1.25)$$

$$(\mathbf{n} \cdot \vec{\sigma})P(\mathbf{n}) = P(\mathbf{n})(\mathbf{n} \cdot \vec{\sigma}) = P(\mathbf{n}) . \quad (1.26)$$

Falls z.B.  $\mathbf{n}$  der Einheitsvektor in  $z$ -Richtung ist, so ergibt sich

$$P(\mathbf{n}) = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} . \quad (1.27)$$

Wir interpretieren  $P(\mathbf{n})$  als die Projektion auf den Zustand, in dem  $\mathbf{n} \cdot \vec{\sigma}$  den Wert 1 annimmt, im Einklang mit der Tatsache, dass der  $q$ -Zahl  $P(\mathbf{n})$  nach (1.24) die Zahlen 0 und 1 entsprechen. Die gemessenen Wahrscheinlichkeiten ergeben sich jetzt aus der Formel

$$\text{Tr}P(\mathbf{n}_2)P(\mathbf{n}_1) = \frac{1}{2}(1 + \mathbf{n}_2 \cdot \mathbf{n}_1) . \quad (1.28)$$

Hierbei bezeichnet  $\text{Tr}$  die Spur einer Matrix, d.h. die Summe ihrer Diagonalelemente.

Wir können die obige Formel auch noch etwas anders schreiben. Dazu nutzen wir aus, dass die Matrizen  $P(\mathbf{n})$  hermitesch sind und Determinante 0 haben,

$$\begin{aligned} \det P(\mathbf{n}) &= \det \frac{1}{2} \begin{pmatrix} 1 + n_3 & n_1 - in_2 \\ n_1 + in_2 & 1 - n_3 \end{pmatrix} \\ &= \frac{1}{4}((1 + n_3)(1 - n_3) - (n_1 - in_2)(n_1 + in_2)) \\ &= \frac{1}{4}(1 - |\mathbf{n}|^2) = 0 , \end{aligned} \quad (1.29)$$

und sich daher in der Form

$$P(\mathbf{n}) = \begin{pmatrix} a \\ b \end{pmatrix} \begin{pmatrix} \bar{a} & \bar{b} \end{pmatrix} \quad (1.30)$$

mit komplexen Zahlen  $a, b$  schreiben lassen. Eine Lösung für  $a$  und  $b$  ist

$$a = \frac{1}{\sqrt{2}}\sqrt{1 + n_3} , \quad b = \frac{1}{\sqrt{2}}\frac{n_1 + in_2}{\sqrt{1 + n_3}} . \quad (1.31)$$

Sei  $A$  eine  $(2 \times 2)$ -Matrix. Dann gilt

$$\text{Tr}AP(\mathbf{n}) = \begin{pmatrix} \bar{a} & \bar{b} \end{pmatrix} A \begin{pmatrix} a \\ b \end{pmatrix} \quad (1.32)$$

Wir fassen die rechte Seite dieser Gleichung jetzt in der folgenden Weise auf. Die Spaltenvektoren bilden einen zweidimensionalen komplexen Vektorraum

$$\mathfrak{H} = \mathbb{C}^2 . \quad (1.33)$$

Multipliziert man einen Spaltenvektor  $\Phi$  von links mit einer  $(2 \times 2)$ -Matrix  $A$ , so erhält man einen neuen Spaltenvektor  $A\Phi$ .  $A$  kann daher als eine Abbildung von  $\mathfrak{H}$  nach  $\mathfrak{H}$  angesehen werden.  $A$  ist linear, d.h.

$$A(\lambda\Phi + \mu\Psi) = \lambda A\Phi + \mu A\Psi , \quad \lambda, \mu \in \mathbb{C} , \quad \Phi, \Psi \in \mathfrak{H} . \quad (1.34)$$

Adjungiert man einen Spaltenvektor, so erhält man einen Zeilenvektor mit den komplex konjugierten Einträgen. Das Produkt eines Zeilenvektors mit einem Spaltenvektor ist eine  $(1 \times 1)$ -Matrix, also eine komplexe Zahl. Wir definieren jetzt auf  $\mathfrak{H}$  ein Skalarprodukt

$$\langle \Phi, \Psi \rangle = \Phi^* \Psi . \quad (1.35)$$

Ein Skalarprodukt auf einem komplexen Vektorraum ist eine Abbildung, die jedem Paar von Vektoren eine komplexe Zahl zuordnet, so dass die folgenden Bedingungen erfüllt sind:

$$\langle \Phi, \lambda\Psi_1 + \mu\Psi_2 \rangle = \lambda \langle \Phi, \Psi_1 \rangle + \mu \langle \Phi, \Psi_2 \rangle \quad (1.36)$$

$$\langle \Phi, \Psi \rangle = \overline{\langle \Psi, \Phi \rangle} \quad (1.37)$$

$$\langle \Phi, \Phi \rangle > 0 \text{ für } \Phi \neq 0 . \quad (1.38)$$

Die Wahrscheinlichkeit  $W(\mathbf{n}_2|\mathbf{n}_1)$  ergibt sich jetzt aus der Formel

$$W(\mathbf{n}_2|\mathbf{n}_1) = \langle \Phi, P(\mathbf{n}_2)\Phi \rangle \text{ mit } \Phi \text{ aus (1.31)} \quad (1.39)$$

Wir interpretieren den Vektor  $\Phi$  als den Zustand des Quantensystems, in dem die  $\mathbf{n}_1$ -Komponente des Spins den Wert  $\hbar/2$  hat. Der Vektor  $\Phi$  hat die Länge 1,

$$\langle \Phi, \Phi \rangle = |a|^2 + |b|^2 = \frac{1}{2} \left( 1 + n_3 + \frac{n_1^2 + n_2^2}{1 + n_3} \right) = 1 . \quad (1.40)$$

Mit der obigen Formel können wir jetzt alle Wahrscheinlichkeitsaussagen über die Beobachtungen in diesem Zustand ausdrücken.

Zum Beispiel können wir fragen, wie wahrscheinlich es ist, dass der Spin in  $\mathbf{n}_2$ -Richtung entweder den Wert  $\hbar/2$  oder den Wert  $-\hbar/2$  annimmt. Es gilt

$$W(\mathbf{n}_2|\mathbf{n}_1) + W(-\mathbf{n}_2|\mathbf{n}_1) = 1 . \quad (1.41)$$

Mit (1.39) ergibt sich

$$\langle \Phi, P(\mathbf{n}_2)\Phi \rangle + \langle \Phi, P(-\mathbf{n}_2)\Phi \rangle = \langle \Phi, (P(\mathbf{n}_2) + P(-\mathbf{n}_2))\Phi \rangle . \quad (1.42)$$

Die Formel  $P(\mathbf{n}) + P(-\mathbf{n}) = \mathbf{1}$  drückt, unabhängig von der Wahl des Zustands, die Tatsache aus, dass nur die beiden Werte vorkommen

können (die Observable  $\mathbf{1}$  nimmt per definitionem in jedem Zustand den Wert 1 an).

Wir können auch den Erwartungswert der  $\mathbf{n}_2$ -Komponente des Spins bestimmen. Dieser berechnet sich aus den Wahrscheinlichkeiten  $W(\pm\mathbf{n}_2|\mathbf{n}_1)$  zu

$$\langle \mathbf{n}_2 \cdot \mathbf{S} \rangle = \frac{\hbar}{2} (W(\mathbf{n}_2|\mathbf{n}_1) - W(-\mathbf{n}_2|\mathbf{n}_1)) . \quad (1.43)$$

Mit

$$\mathbf{n} \cdot \vec{\sigma} = P(\mathbf{n}) - P(-\mathbf{n}) \quad (1.44)$$

schreibt sich diese Relation als

$$\langle \mathbf{n}_2 \cdot \mathbf{S} \rangle = \langle \Phi, \mathbf{n}_2 \cdot \mathbf{S} \Phi \rangle \quad (1.45)$$

## 2. Endlich dimensionale Quantensysteme

Wir können die gefundenen Strukturen jetzt leicht auf Systeme verallgemeinern, in denen die Observablen höchstens  $n$  Werte annehmen können. Beispiele sind Spinsysteme mit Spin  $s = (n - 1)/2$ . Wir identifizieren die Observablen mit den hermiteschen  $(n \times n)$ -Matrizen. Die möglichen Messwerte sind die Eigenwerte der Matrix. Diese bestimmt man aus den Lösungen der charakteristischen Gleichung

$$\det(A - \lambda \mathbf{1}) = 0 \quad (2.1)$$

Seien  $\lambda_1, \dots, \lambda_k \in \mathbb{R}$ ,  $k \leq n$  die (verschiedenen) Eigenwerte der hermiteschen Matrix  $A$ . Es gilt dann

$$(A - \lambda_1 \mathbf{1}) \dots (A - \lambda_k \mathbf{1}) = 0 . \quad (2.2)$$

Wir wollen jetzt ebenso wie im Beispiel des vorherigen Abschnitts Projektoren finden, die den einzelnen Eigenwerten entsprechen. Wir suchen hermitesche Matrizen  $P_i$ ,  $i = 1, \dots, k$  mit den Eigenschaften

$$P_i P_j = \delta_{ij} P_j \quad (2.3)$$

$$A P_i = P_i A = \lambda_i P_i \quad (2.4)$$

$$\sum_{i=1}^k P_i = \mathbf{1} \quad (2.5)$$

Aus (2.3) und (2.5) folgen die Gleichungen

$$A^m = \sum_{i=1}^k \lambda_i^m P_i , \quad m = 0, \dots, k \quad (2.6)$$

(mit  $A^0 = \mathbf{1}$ ). Dieses Gleichungssystem ist genau dann eindeutig lösbar, wenn die Determinante der Koeffizientenmatrix nicht verschwindet. Es gilt

$$\det \begin{pmatrix} 1 & \dots & 1 \\ \lambda_1 & \dots & \lambda_k \\ \dots & \dots & \dots \\ \lambda_1^k & \dots & \lambda_k^k \end{pmatrix} = \prod_{1 \leq i < j \leq k} (\lambda_j - \lambda_i) \quad (2.7)$$

(Vandermond-Determinante). Da die Eigenwerte nach Voraussetzung verschieden sind, ist die Determinante nicht Null, und die Matrizen  $P_i$  sind eindeutig bestimmt. Man findet

$$P_i = \prod_{j \neq i} \frac{A - \lambda_j \mathbf{1}}{\lambda_i - \lambda_j}. \quad (2.8)$$

Aus (2.2) folgt  $(A - \lambda_i)P_i = 0$  und damit (2.8). Hieraus ergibt sich wiederum (2.3).

Die Projektoren  $P_i$  sind Polynome in  $A$ . Es gilt der allgemeine Satz:

**THEOREM 2.1.** Ist  $A$  eine quadratische  $n$ -reihige Matrix mit Eigenwerten  $\lambda_i$ ,  $i = 1, \dots, k$  und  $p$  ein Polynom, so sind die Eigenwerte der Matrix  $p(A)$  gegeben durch  $p(\lambda_i)$ ,  $i = 1, \dots, k$ .

**PROOF.** Die Behauptung ist trivial für  $p = 0$ . Sei  $p \neq 0$  und sei  $\mu \in \mathbb{C}$ . Dann lässt sich  $p(x) - \mu$  in ein Produkt linearer Faktoren zerlegen,

$$p(x) - \mu = a_0(x - a_1) \dots (x - a_m) \quad (2.9)$$

mit  $a_0 \neq 0$ . Diese Relation bleibt richtig, wenn wir für  $x$  die Matrix  $A$  einsetzen,

$$p(A) - \mu \mathbf{1} = a_0(A - a_1 \mathbf{1}) \dots (A - a_m \mathbf{1}). \quad (2.10)$$

Aufgrund des Multiplikationssatzes für Determinanten gilt

$$\det(p(A) - \mu \mathbf{1}) = a_0^n \det(A - a_1 \mathbf{1}) \dots \det(A - a_m \mathbf{1}). \quad (2.11)$$

Wir erkennen, dass  $\mu$  genau dann eine Nullstelle des charakteristischen Polynoms von  $p(A)$  ist, falls einer der Zahlen  $a_i$  eine Nullstelle des charakteristischen Polynoms von  $A$  ist. Dies beweist die Behauptung. ■

Man erkennt auch aus dem Beweis, dass die Vielfachheit eines Eigenwerts  $\mu$  von  $p(A)$  die Summe der Vielfachheiten der Eigenwerte  $\lambda_i$  von  $A$  mit  $p(\lambda_i) = \mu$  ist. Wir sehen also insbesondere, dass die Vielfachheit des Eigenwerts 1 von  $P_i(A)$  gleich der Vielfachheit des Eigenwerts  $\lambda_i$  von  $A$  ist. Das charakteristische Polynom von  $P_i$  ist also

$$\det(P_i - \lambda \mathbf{1}) = (-1)^n (\lambda - 1)^m \lambda^{n-m} = (-1)^n (\lambda^n - m\lambda^{n-1} + O(\lambda^{n-2})) \quad (2.12)$$

wenn  $m$  die Vielfachheit von  $\lambda_i$  ist. Wir erkennen, dass die Vielfachheit als Koeffizient des Terms  $\lambda^{n-1}$  auftritt. Dies ist aber gerade die Spur von  $P_i$ , also gilt

$$\text{Tr} P_i = m. \quad (2.13)$$

Ein Projektor  $P$  mit  $\text{Tr} P = 1$  ist minimal, im Sinne, dass er sich nicht als Summe zweier von Null verschiedener Projektoren schreiben lässt. Wir fassen wie im Stern-Gerlach-Versuch einen minimalen Projektor  $P$

als eine optimale Zustandspräparation auf. Messungen einer Observablen  $A$  mit Eigenprojektoren  $P_i$  zum Eigenwert  $\lambda_i$  ergeben in diesem Zustand den Messwert  $\lambda_i$  mit Wahrscheinlichkeit

$$W(A = \lambda_i | P) = \text{Tr} P_i P \quad (2.14)$$

Für  $A = P$  ist die Wahrscheinlichkeit also 1, dass die Messung den Wert 1 ergibt. Im allgemeinen aber sind die Wahrscheinlichkeiten  $< 1$ . Wir wollen uns davon überzeugen, dass sie immer zwischen 0 und 1 liegen. Seien  $P$  und  $Q$  zwei hermitesche Projektoren, d.h.  $P = P^2 = P^*$ ,  $Q = Q^2 = Q^*$ . Dann gilt, unter Verwendung der Formel  $\text{Tr} AB = \text{Tr} BA$  für die Spur

$$\text{Tr} QP = \text{Tr} Q^2 P^2 = \text{Tr} P Q^2 P = \text{Tr} P^* Q^* Q P = \text{Tr} (QP)^* Q P . \quad (2.15)$$

Aber die Diagonalelemente einer Matrix der Form  $A^* A$  sind immer positiv,

$$(A^* A)_{kk} = \sum_{j=1}^n (A^*)_{kj} A_{jk} = \sum_{j=1}^n |A_{jk}|^2 , \quad (2.16)$$

und damit auch ihre Spur.  $\text{Tr} QP$  kann aber auch nicht größer sein als  $\text{Tr} P$ , denn mit  $Q$  ist auch  $\mathbf{1} - Q$  ein Projektor, sodass folgt

$$0 \leq \text{Tr}(\mathbf{1} - Q)P = \text{Tr} P - \text{Tr} QP . \quad (2.17)$$

Mit  $\text{Tr} P = 1$  ergibt sich, dass (2.14) tatsächlich eine Wahrscheinlichkeitsverteilung definiert.

Wichtige Charakteristika einer Wahrscheinlichkeitsverteilung sind der Erwartungswert und die Varianz. Für den Erwartungswert ergibt sich

$$\langle A \rangle_P = \sum_{i=1}^k \lambda_i W(A = \lambda_i | P) = \text{Tr} A P \quad (2.18)$$

und für die Varianz

$$\text{var}(A)_P = \sum_{i=1}^k (\lambda_i - \langle A \rangle_P)^2 W(A = \lambda_i | P) = \text{Tr} A^2 P - (\text{Tr} A P)^2 . \quad (2.19)$$

Da die Eigenprojektoren von  $A$  Polynome in  $A$  vom Grade  $k - 1$  sind, kann die Wahrscheinlichkeitsverteilung aus den ersten  $k - 1$  Momenten, d.h. aus den Erwartungswerten der Potenzen von  $A$  mit Exponenten  $< k$  gewonnen werden.

Wir wollen jetzt wieder ausnutzen, dass ein minimaler Projektor  $P$  in der Form

$$P = \Phi \Phi^* \quad (2.20)$$

mit einem Spaltenvektor  $\Phi \in \mathbb{C}^n$  dargestellt werden kann. Wir fassen den Raum der Spaltenvektoren als einen Vektorraum auf

$$\mathfrak{H} = \mathbb{C}^n \quad (2.21)$$

mit einem Skalarprodukt

$$\langle \Phi, \Psi \rangle = \Phi^* \Psi . \quad (2.22)$$

Die positive Definitheit des Skalarprodukts erlaubt es, den Vektoren durch

$$\|\Phi\| := \sqrt{\langle \Phi, \Phi \rangle} \quad (2.23)$$

eine Norm ("Länge")  $\geq 0$  zuzuordnen. Endlich dimensionale komplexe Vektorräume mit einem Skalarprodukt nennt man endlich dimensionale Hilberträume. In der Quantenmechanik treten auch unendlich dimensionale Hilberträume auf. Darauf werden wir später eingehen.

Die Matrizen  $A$  wirken als lineare Abbildungen (auch Operatoren genannt) auf  $\mathfrak{H}$ , und die Formel für den Erwartungswert schreibt sich als

$$\langle A \rangle = \langle \Phi, A\Phi \rangle . \quad (2.24)$$

Die adjungierte Matrix lässt sich durch die Relation

$$\langle \Psi, A^*\Phi \rangle = \Psi^* A^* \Phi = (A\Psi)^* \Phi = \langle A\Psi, \Phi \rangle \quad (2.25)$$

charakterisieren. Ist  $\lambda$  ein Eigenwert von  $A$ , so gibt es einen nicht verschwindenden Vektor  $\Phi \in \mathfrak{H}$ , der die Eigenwertgleichung

$$A\Phi = \lambda\Phi \quad (2.26)$$

erfüllt.  $\Phi$  heißt Eigenvektor von  $A$  zum Eigenwert  $\lambda$ . Ist  $P_\lambda$  der zugehörige Eigenprojektor, so bilden die Eigenvektoren den Unterraum  $P_\lambda \mathfrak{H}$ . Die Dimension dieses Unterraums (des Eigenraums zum Eigenwert  $\lambda$ ) ist gleich der Vielfachheit von  $\lambda$ .

Man sieht jetzt auch sehr leicht, warum die Eigenwerte einer hermiteschen Matrix reell sein müssen. Es folgt nämlich aus (2.26) und  $A = A^*$

$$\lambda \langle \Phi, \Phi \rangle = \langle \Phi, A\Phi \rangle = \langle A\Phi, \Phi \rangle = \bar{\lambda} \langle \Phi, \Phi \rangle . \quad (2.27)$$

Eigenvektoren zu verschiedenen Eigenwerten einer hermiteschen Matrix sind orthogonal zueinander. Dies folgt aus der Orthogonalitätsrelation der Eigenprojektoren (2.3), ergibt sich aber auch leicht direkt aus der folgenden Rechnung. Sei  $A\Phi = \lambda\Phi$  und  $A\Psi = \mu\Psi$  mit  $\lambda, \mu \in \mathbb{R}$ . Dann gilt

$$\langle \Phi, A\Psi \rangle = \langle A\Phi, \Psi \rangle \implies (\lambda - \mu) \langle \Phi, \Psi \rangle = 0 . \quad (2.28)$$

Eine Orthonormalbasis des  $n$ -dimensionalen Hilbertraums  $\mathfrak{H}$  ist eine Familie von Vektoren  $(\Phi_i)_{i=1, \dots, n}$  mit der Eigenschaft

$$\langle \Phi_i, \Phi_j \rangle = \delta_{ij} . \quad (2.29)$$

Jeder Hilbertraum besitzt eine Orthonormalbasis. Es gilt der folgende Satz:

**THEOREM 2.2.** Eine hermitesche Matrix besitzt eine Orthonormalbasis von Eigenvektoren.

PROOF. Falls die Eigenwerte nicht entartet sind, wählt man einfach die Eigenvektoren mit Länge 1. Falls ein Eigenwert entartet ist, wählt man eine Orthonormalbasis im zugehörigen Eigenraum. Die Vereinigung dieser Orthonormalbasen ist dann eine Orthonormalbasis im ganzen Hilbertraum. ■

### 3. Kompatibilität und Unschärferelationen

Am Beispiel des Spins im Stern-Gerlach-Versuch haben wir gesehen, dass die verschiedenen Komponenten des Spins nicht gleichzeitig festgelegt werden können. Hat man durch einen ersten Stern-Gerlach-Versuch die 3-Komponente fixiert, so ist danach die 1-Komponente völlig unbestimmt. Man unterscheidet daher zwischen paarweise kompatiblen und inkompatiblen Observablen. Zwei Observable  $A, B$  heißen kompatibel, wenn es eine Familie  $(P_i)_{i=1, \dots, k}$  von (hermiteschen) Projektoren gibt, die paarweise orthogonal sind, deren Summe die Einheitsmatrix ist und die eine Darstellung von  $A$  und  $B$  der folgenden Form ermöglichen:

$$A = \sum_i a_i P_i, \quad B = \sum_i b_i P_i. \quad (3.1)$$

Hierbei sind die reellen Zahlen  $a_i, b_i$  die (nicht notwendig paarweise verschiedenen) Eigenwerte von  $A$ , bzw. von  $B$ . Kompatible Observable lassen sich gleichzeitig genau festlegen. Ist  $\Phi \in \mathfrak{H}$  und  $\Psi := P_i \Phi \neq 0$ , so ist im Zustand  $\Psi$  der Messwert von  $A$  mit Sicherheit  $a_i$  und der von  $B$  mit Sicherheit  $b_i$ . Für die Kompatibilität zweier Observablen gibt es ein einfaches Kriterium:

**THEOREM 3.1.** Zwei Observable sind dann und nur dann kompatibel, wenn ihr Kommutator verschwindet.

PROOF. Seien  $A$  und  $B$  kompatibel. Dann gilt

$$[A, B] = \sum_{i,j} a_i b_j [P_i, P_j] = 0, \quad (3.2)$$

da wegen der paarweisen Orthogonalität der Projektoren das Produkt  $P_i P_j$  für  $i \neq j$  verschwindet.

Sei nun umgekehrt  $[A, B] = 0$ , und seien  $P_j^A, j = 1, \dots, k, P_l^B, l = 1, \dots, k'$  die Eigenprojektoren von  $A$  bzw. von  $B$ . Aus der Formel (2.8) folgt dann, dass mit  $A$  und  $B$  auch die Eigenprojektoren von  $A$  und  $B$  miteinander kommutieren. Das Produkt zweier kommutierender Projektoren ist aber wieder ein Projektor. Man setzt jetzt

$$P_{j+k(l-1)} = P_j^A P_l^B \quad (3.3)$$

und erhält eine Familie von paarweise orthogonalen Projektoren  $P_i, i = 1, \dots, kk'$  mit den gewünschten Eigenschaften. Dabei können einige dieser Projektoren verschwinden. ■

Die Einschränkungen an die gemeinsame Festlegbarkeit inkompatibler Observablen nennt man Unschärferelationen. Die einfachste dieser Relationen liefert einen direkten Zusammenhang zwischen dem Produkt der Varianzen der jeweiligen Wahrscheinlichkeitsverteilungen und dem Kommutator.

Sei  $\Psi \in \mathfrak{H}$  mit  $\|\Psi\| = 1$  ein optimal präparierter Zustand. Die Wahrscheinlichkeitsverteilung der Messung einer Observablen  $A$  ist  $W(A = a_i | \Psi) = \langle \Psi, P_i^A \Psi \rangle$ . Die Varianz (quadratische Abweichung) ist

$$\text{var}(W) \equiv \Delta_\Psi(A)^2 = \langle \Psi, A^2 \Psi \rangle - \langle \Psi, A \Psi \rangle^2 . \quad (3.4)$$

Es gilt die folgende Ungleichung

$$\Delta_\Psi(A) \Delta_\Psi(B) \geq \frac{1}{2} | \langle \Psi, [A, B] \Psi \rangle | . \quad (3.5)$$

PROOF. Da die Einheitsmatrix mit allen Matrizen vertauscht, gilt für alle  $a, b \in \mathbb{R}$

$$[A, B] = [A - a\mathbf{1}, B - b\mathbf{1}] . \quad (3.6)$$

Nach der Schwarzischen Ungleichung gilt

$$| \langle \Psi, [A - a\mathbf{1}, B - b\mathbf{1}] \Psi \rangle | \leq 2 \| (A - a\mathbf{1}) \Psi \| \| (B - b\mathbf{1}) \Psi \| . \quad (3.7)$$

Weiter gilt für  $a = \langle \Psi, A \Psi \rangle$

$$\| (A - a\mathbf{1}) \Psi \|^2 = \langle \Psi, A^2 \Psi \rangle - \langle \Psi, A \Psi \rangle^2 = \Delta_\Psi(A)^2 , \quad (3.8)$$

und eine entsprechende Gleichung gilt für  $B$ . Fasst man diese Relationen zusammen, so erhält man die Behauptung. ■

Wendet man die Unschärferelation auf die Paulimatrizen an, so folgt aus  $[\sigma_1, \sigma_2] = 2i\sigma_3$ , dass in einem Zustand, in dem die beiden Messwerte von  $\sigma_3$  unterschiedliche Wahrscheinlichkeiten besitzen, weder die Schwankungen von  $\sigma_1$  noch von  $\sigma_2$  verschwinden können.

Ein anderes wichtiges Beispiel, an dem die Unschärferelation von Heisenberg zuerst entdeckt worden ist, bildet das Paar Ort-Impuls. Ort und Impuls werden durch Operatoren  $q$  und  $p$  auf einem unendlich dimensionalen Hilbertraum dargestellt und erfüllen die Vertauschungsrelation

$$[q, p] = i\hbar \mathbf{1} . \quad (3.9)$$

Da der Erwartungswert des Einheitsoperators immer 1 ist, erhält man die Unschärferelation

$$\Delta q \Delta p \geq \frac{\hbar}{2} . \quad (3.10)$$

Eine ähnliche Unschärferelation gibt es auch zwischen Energie und Zeit; allerdings ist die Darstellung der Zeitmessung durch Operatoren etwas komplizierter, sodass ein Beweis der Unschärferelation etwas anders geführt werden muss.

Interessant ist auch die Unschärferelation zwischen magnetischem und elektrischem Feld. Nach der Quantenelektrodynamik gilt für dese

eine Vertauschungsrelation ähnlich derjenigen zwischen Ort und Impuls. Als eine Konsequenz gibt es keinen Zustand, in dem elektrisches und magnetisches Feld gleichzeitig verschwinden.

#### 4. Der Dirac-Formalismus

Der Zusammenhang zwischen den Vektoren aus  $\mathfrak{H}$ , den Operatoren auf  $\mathfrak{H}$  und den Erwartungswerten lässt sich elegant in einem Formalismus beschreiben, der von Dirac entwickelt worden ist. Zunächst einmal können wir jedem Vektor  $\Phi$  einen Operator

$$|\Phi\rangle : \begin{cases} \mathbb{C} & \rightarrow & \mathfrak{H} \\ \lambda & \mapsto & \lambda\Phi \end{cases} \quad (4.1)$$

umkehrbar eindeutig zuordnen. Nach Dirac nennt man  $|\Phi\rangle$  einen Ket-Vektor. Umgekehrt kann man jedem Vektor  $\Psi \in \mathfrak{H}$  über das Skalarprodukt auch einen Operator

$$\langle\Phi| : \begin{cases} \mathfrak{H} & \rightarrow & \mathbb{C} \\ \Phi & \mapsto & \langle\Psi, \Phi\rangle \end{cases} \quad (4.2)$$

zuordnen.  $\langle\Psi|$  wird Bra-Vektor genannt. Das Skalarprodukt ("bracket") ist dann das Produkt eines Bra- und eines Ket-Vektors,

$$\langle\Psi, \Phi\rangle = \langle\Psi||\Phi\rangle \quad (4.3)$$

und die Formel für den Erwartungswert ergibt sich als Produkt dreier Operatoren

$$\langle\Phi, A\Phi\rangle = \langle\Phi|A|\Phi\rangle . \quad (4.4)$$

Der Formalismus liefert auch eine bequeme Möglichkeit, Observable durch Vektoren auszudrücken. Denn ein minimaler Projektor schreibt sich in der Form

$$P = |\Phi\rangle\langle\Phi| , \quad (4.5)$$

und jeder hermitesche Operator auf  $\mathfrak{H}$  lässt sich als Linearkombination solcher Terme darstellen. Hierzu wählen wir eine Orthonormalbasis von Eigenvektoren und erhalten

$$A = \sum_{i=1}^n \lambda_i |\Phi_i\rangle\langle\Phi_i| . \quad (4.6)$$

Hierbei sind die Eigenwerte nicht notwendig paarweise verschieden. Ein nützlicher Spezialfall dieser Formel ist die Zerlegung der Eins,

$$\mathbf{1} = \sum_{i=1}^n |\Phi_i\rangle\langle\Phi_i| . \quad (4.7)$$

### 5. Schrödingerbild und Heisenbergbild; der Zeitentwicklungsoperator

Wir haben bisher die Dynamik durch die Zeitabhängigkeit der Observablen beschrieben. Diese erfüllen die Heisenberggleichung. Die Zustände hingegen wurden als zeitunabhängig angenommen. Diese Formulierung der Dynamik nennt man das Heisenbergbild. Sei  $A(t)$  die Lösung der Heisenberggleichung mit der Anfangsbedingung  $A(0) = A$ . Mit Hilfe der Spektraldarstellung von  $H$ ,

$$H = \sum_j E_j P_j , \quad (5.1)$$

mit den Energieeigenwerten  $E_i$  und den zugehörigen Eigenprojektoren  $P_i$ , ergibt sich die Lösung der Heisenberggleichung zu

$$A(t) = \sum_{j,k} e^{it(E_j - E_k)} P_j A P_k . \quad (5.2)$$

Für den Erwartungswert von  $A(t)$  in einem Zustand, der durch einen Einheitsvektor  $\Psi$  angegeben wird, gilt

$$\langle \Psi, A(t) \Psi \rangle = i \sum_{jk} e^{it(E_j - E_k)} \langle P_j \Psi, A P_k \Psi \rangle = \langle \Psi(t), A \Psi(t) \rangle \quad (5.3)$$

mit

$$\Psi(t) = \sum_j e^{-itE_j} P_j \Psi . \quad (5.4)$$

Der zeitabhängige Vektor  $\Psi(t)$  erfüllt die Gleichung (Schrödingergleichung)

$$i \frac{d}{dt} \Psi(t) = H \Psi(t) \quad (5.5)$$

mit der Anfangsbedingung  $\Psi(0) = \Psi$ . In dieser Formulierung der Dynamik sieht man die Zustände als zeitabhängig und die Observablen als zeitunabhängig an. Diese Auffassung der Dynamik nennt man das Schrödingerbild. Da die Menge aller Erwartungswerte die vollständige Information über das System enthält, sind beide Auffassungen äquivalent.

Ein Vorteil des Schrödingerbildes ist, dass in der Lösung nur eine einfache Summe über die Eigenwerte auftaucht, im Gegensatz zur Doppelsumme im Heisenbergbild. Ist  $(\Phi_j)$  ein Orthonormalsystem von Eigenvektoren von  $H$  mit Eigenwerten  $E_j$ , so lautet die Lösung der Schrödingergleichung

$$|\Psi(t)\rangle = \sum_j e^{-itE_j} |\Phi_j\rangle \langle \Phi_j, \Psi \rangle . \quad (5.6)$$

Die Wahrscheinlichkeit  $W_t$ , dass der Zustand nach der Zeit  $t$  im Ausgangszustand ist, ist der Erwartungswert des Projektors  $|\Psi\rangle\langle\Psi|$ . Man findet

$$W_t = \langle \Psi(t), |\Psi\rangle\langle\Psi| \Psi(t) \rangle \quad (5.7)$$

Die Beziehungen zwischen den beiden Bildern lassen sich mit Hilfe des Zeitentwicklungsoperators  $U(t)$  übersichtlich schreiben.  $U(t)$  ist eine operatorwertige Lösung der Schrödingergleichung

$$i \frac{d}{dt} U(t) = H U(t) \quad (5.8)$$

mit der Anfangsbedingung  $U(0) = \mathbf{1}$ . Mit Hilfe der Spektraldarstellung von  $H$  ergibt sich

$$U(t) = \sum_j e^{-itE_j} P_j . \quad (5.9)$$

Die Operatoren  $U(t)$  sind unitär, da die Eigenwerte von  $H$  reell sind. Die Lösung der Schrödingergleichung für vektorwertige Funktionen  $\Psi(t)$  mit Anfangsbedingung  $\Psi(0) = \Psi$  ist dann

$$\Psi(t) = U(t)\Psi . \quad (5.10)$$

Die Lösung der Heisenberggleichung

$$\frac{d}{dt} A(t) = i[H, A(t)] \quad (5.11)$$

mit der Anfangsbedingung  $A(0) = A$  ist

$$A(t) = U(t)^* A U(t) . \quad (5.12)$$

Die Äquivalenz zwischen den beiden Bildern folgt aus der Identität der Erwartungswerte

$$\langle \Psi(t), A \Psi(t) \rangle = \langle U(t)\Psi, A U(t)\Psi(t) \rangle = \langle \Psi, U(t)^* A U(t)\Psi \rangle = \langle \Psi, A(t)\Psi \rangle . \quad (5.13)$$

## 6. Spinoperatoren im $\mathbb{C}^n$

Auch unter den  $(n \times n)$ -Matrizen lassen sich Systeme von 3 Matrizen  $\mathbf{S} = (S_1, S_2, S_3)$  finden, die sich als Spinobservable interpretieren lassen. Ausgangspunkt ist die Überlegung, dass sich der Vektor  $\mathbf{S}$  unter Drehungen um die Achse  $\mathbf{n}$  so verhalten soll, wie ein Vektor im  $\mathbb{R}^3$ ,

$$R(\mathbf{n}, \theta)\mathbf{S} = (\mathbf{n} \cdot \mathbf{S})\mathbf{n} + \cos \theta((\mathbf{S} - (\mathbf{n} \cdot \mathbf{S})\mathbf{n}) + \sin \theta(\mathbf{n} \times \mathbf{S})) \quad (6.1)$$

und dass die Drehungen durch ein angelegtes Magnetfeld wie im Spin- $\frac{1}{2}$ -Fall verursacht werden sollen. Aus der Heisenberggleichung folgt jetzt

$$\frac{d}{d\theta} R(\mathbf{n}, \theta)\mathbf{S}|_{\theta=0} = \mathbf{n} \times \mathbf{S} = i[\mathbf{n} \cdot \mathbf{S}, \mathbf{S}] \quad (6.2)$$

also

$$[S_j, S_k] = i\epsilon_{jkl} S_l , \quad j, k, l = 1, 2, 3 . \quad (6.3)$$

Diese Relationen definieren eine sogenannte Lie-Algebra, hier die Lie-Algebra der Drehgruppe  $SO(3)$ . Gesucht ist also eine Darstellung dieser Lie-Algebra durch hermitesche  $(n \times n)$ -Matrizen.

Wir beobachten zunächst dass der Operator

$$|\mathbf{S}|^2 = S_1^2 + S_2^2 + S_3^2 \quad (6.4)$$

mit allen Operatoren  $S_j, j = 1, 2, 3$  vertauscht<sup>1</sup>. Insbesondere gibt es also eine Orthonormalbasis gemeinsamer Eigenvektoren von  $|\mathbf{S}|^2$  und  $S_3$ .

Die Darstellungen der Spinoperatoren lassen sich am einfachsten finden, wenn man die Vertauschungsrelationen (6.3) in der folgenden Form schreibt

$$[S_3, S_{\pm}] = \pm S_{\pm}, \quad [S_+, S_-] = 2S_3 \quad \text{mit } S_{\pm} = S_1 \pm iS_2. \quad (6.5)$$

Ist jetzt  $\Phi$  ein gemeinsamer Eigenvektor von  $S_3$  zum Eigenwert  $m$  und von  $|\mathbf{S}|^2$  zum Eigenwert  $\mu$ , so gilt

$$S_3 S_{\pm} \Phi = [S_3, S_{\pm}] \Phi + S_{\pm} S_3 \Phi = \pm S_{\pm} \Phi + S_{\pm} m \Phi = (m \pm 1) \Phi, \quad (6.6)$$

also ist mit  $m$  auch  $m \pm 1$  Eigenwert von  $S_3$ , falls der Vektor  $S_{\pm} \Phi$  nicht verschwindet. Man nennt  $S_+$  den Aufsteige-Operator und  $S_-$  den Absteige-Operator. Wir berechnen

$$S_+^* S_+ = (S_1 - iS_2)(S_1 + iS_2) = S_1^2 + S_2^2 + i[S_1, S_2] = |\mathbf{S}|^2 - S_3^2 - S_3 \quad (6.7)$$

und entsprechend

$$S_-^* S_- = (S_1 + iS_2)(S_1 - iS_2) = S_1^2 + S_2^2 - i[S_1, S_2] = |\mathbf{S}|^2 - S_3^2 + S_3. \quad (6.8)$$

$\Phi$  ist also auch Eigenvektor dieser beiden Operatoren mit Eigenwerten

$$\mu - m(m + 1) \quad \text{bzw.} \quad \mu - m(m - 1) \quad (6.9)$$

Da Eigenwerte von Operatoren der Form  $A^* A$  nicht negativ sein können, kann  $m$  nur in dem Intervall

$$\frac{1}{2} - \sqrt{\frac{1}{4} + \mu} \leq m \leq -\frac{1}{2} + \sqrt{\frac{1}{4} + \mu} \quad (6.10)$$

liegen. Die Reihe der Eigenvektoren von  $S_3$ , die man durch Anwendung der Aufsteige- und AbsteigeOperatoren erhält, muss also abbrechen. Insbesondere gibt es einen Eigenvektor  $\Phi$  von  $S_3$  mit  $S_+ \Phi = 0$ . Dieser Vektor ist ein Eigenvektor von  $S_+^* S_+$  zum Eigenwert 0 und damit auch ein Eigenvektor von  $|\mathbf{S}|^2$  zum Eigenwert  $\mu = m(m + 1)$ .<sup>2</sup>

Sei  $\Phi$  ein solcher Vektor mit dem Eigenwert  $l$  für  $S_3$ . Wir wenden jetzt wiederholt den Absteige-Operator  $S_-$  auf  $\Phi$  an. Dabei ändert sich der Eigenwert von  $|\mathbf{S}|^2$  nicht. Auf diese Weise erhalten wir gemeinsame Eigenvektoren  $S_-^k \Phi$  von  $S_3$  und  $|\vec{S}|^2$  zu den Eigenwerten  $l - k$  bzw.  $l(l + 1)$  für natürliche Zahlen  $k$ . Die Reihe bricht ab, wenn  $S_-^{k+1} \Phi = 0$  ist. Nach (6.8) ist dies der Fall, wenn die Gleichung

$$l(l + 1) = (l - k)^2 - (l - k) \quad (6.11)$$

erfüllt ist. Sie muss abbrechen, da die Eigenwerte von  $S_-^* S_-$  nicht negativ sein können. Wie schließen, dass  $2l$  der maximale Wert von  $k$  ist.

<sup>1</sup>Einen solchen Operator nennt man einen Casimir-Operator der Lie-Algebra.

<sup>2</sup>In der Theorie der Lie-Algebren nennt man einen solchen Vektor einen Höchstgewichtsvektor.

Da  $k$  eine natürliche Zahl (einschließlich der Null) ist, kann  $l$  die Werte  $0, \frac{1}{2}, 1, \frac{3}{2}$  etc. annehmen.

Wir haben, ausgehend von einem höchsten Gewichtsvektor  $\Phi$ , Eigenvektoren von  $S_3$  mit den Eigenwerten  $-l, -l+1, \dots, l-1, l$  gefunden. Diese erzeugen einen Teilraum  $\mathfrak{H}_l$  von  $\mathfrak{H}$ , der unter der Anwendung der Spin-Operatoren invariant ist. Wir wollen jetzt annehmen, dass  $\mathfrak{H}_l = \mathfrak{H}$  ist. Eine solche Darstellung der Lie-Algebra nennt man irreduzibel. Wir können jetzt  $\mathfrak{H} = \mathbb{C}^{2l+1}$  setzen. Wir wählen  $S_3$  als Diagonalmatrix,

$$S_3 = \begin{pmatrix} l & 0 & \dots & 0 \\ 0 & l-1 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & -l \end{pmatrix}. \quad (6.12)$$

Der Aufsteige-Operator hat die folgende Form

$$S_+ = \begin{pmatrix} 0 & \sqrt{2l} & 0 & \dots & 0 \\ 0 & 0 & \sqrt{4l-2} & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 0 & \sqrt{2l} \\ 0 & 0 & \dots & \dots & 0 \end{pmatrix}. \quad (6.13)$$

Hierbei steht in der  $k$ -ten Zeile an der  $(k+1)$ -ten Spalte

$$\sqrt{l(l+1) - (l-k)(l-k+1)}, \quad (6.14)$$

während alle anderen Einträge verschwinden.

## 7. Unabhängige Teilsysteme; das Tensorprodukt

Grundlegend für die Physik ist die Möglichkeit, isolierte Teilsysteme zu betrachten. Für die Konsistenz einer solchen isolierten Betrachtung ist es erforderlich, dass man verschiedene Teilsysteme, die sich gegenseitig nicht beeinflussen, zu einem Gesamtsystem zusammenfassen kann, ohne dass sich dabei die Eigenschaften der Teilsysteme ändern. In der relativistischen Physik denkt man hierbei an Teilsysteme, die raumartig zueinander lokalisiert sind, so dass nach den Prinzipien der Relativitätstheorie keine Signale zwischen den Systemen ausgetauscht werden können.

Wir betrachten zwei Systeme, die durch die Matrixalgebren  $\mathfrak{A}_1 = M(n_1)$  und  $\mathfrak{A}_2 = M(n_2)$  beschrieben werden. Wir können jetzt den Matrizen aus  $\mathfrak{A}_1$  Matrizen aus  $M(n_1 n_2)$  zuordnen. Hierzu ersetzen wir in Blockschreibweise das Matrixelement  $A_{ik}$  durch die  $(n_2 \times n_2)$ -Matrix  $A_{ik} \mathbf{1}_{n_2}$ ,

$$A \mapsto A \otimes \mathbf{1}_{n_2} = \begin{pmatrix} A_{11} \mathbf{1}_{n_2} & \dots & \dots & A_{1n_1} \mathbf{1}_{n_2} \\ \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots \\ A_{n_1 1} \mathbf{1}_{n_2} & \dots & \dots & A_{n_1 n_1} \mathbf{1}_{n_2} \end{pmatrix}. \quad (7.1)$$

Für Blockmatrizen gelten dieselben Regeln für Addition und Multiplikation wie für Matrizen mit Zahleinträgen. Wir erkennen, dass die algebraischen Eigenschaften bei dieser Abbildung erhalten bleiben. Dies schließt insbesondere die Eigenwerte ein. Ihre Vielfachheit aber wird mit dem Faktor  $n_2$  multipliziert.

In ähnlicher Weise können wir auch den Matrizen aus  $\mathfrak{A}_2$  Matrizen aus  $M(n_1 n_2)$  durch die Vorschrift

$$A \mapsto \mathbf{1}_{n_1} \otimes A = \begin{pmatrix} A & 0 & \dots & 0 \\ 0 & A & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & \dots & \dots & A \end{pmatrix} \quad (7.2)$$

zuordnen. Auch diese Abbildung erhält alle algebraischen Relationen. Multipliziert man Matrizen aus den unabhängigen Teilsystemen, so findet man

$$(A \otimes \mathbf{1})(\mathbf{1} \otimes B) = \begin{pmatrix} A_{11}B & \dots & \dots & A_{1n_1}B \\ \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots \\ A_{n_1 1}B & \dots & \dots & A_{n_1 n_1}B \end{pmatrix} = (\mathbf{1} \otimes B)(A \otimes \mathbf{1}) . \quad (7.3)$$

Observable aus den unabhängigen Teilsystemen vertauschen miteinander, sind also immer kompatibel. Dies wird auch als Lokalisitätspostulat für unabhängige Teilsysteme bezeichnet. Man nennt (7.3) das Tensorprodukt von Matrizen, Notation  $A \otimes B$ . Es gelten die Regeln:

$$(A \otimes B)(C \otimes D) = AC \otimes BD \quad (7.4)$$

$$(A + B) \otimes C = A \otimes C + B \otimes C \quad (7.5)$$

$$A \otimes (B + C) = A \otimes B + A \otimes C \quad (7.6)$$

$$A \otimes (\lambda B) = (\lambda A) \otimes B = \lambda(A \otimes B) \quad (7.7)$$

$$(A \otimes B)^* = A^* \otimes B^* \quad (7.8)$$

Wir haben das Tensorprodukt zunächst für quadratische Matrizen definiert. Es kann aber in derselben Weise für beliebige Matrizen erklärt werden, insbesondere für Spaltenvektoren. Sei  $\mathfrak{H}_1 = \mathbb{C}^{n_1}$  und  $\mathfrak{H}_2 = \mathbb{C}^{n_2}$ . Dann kann die Matrix  $A \otimes B$  als Operator auf dem Raum  $\mathfrak{H} = \mathfrak{H}_1 \otimes \mathfrak{H}_2$  aufgefasst werden,

$$(A \otimes B)(\Phi \otimes \Psi) = A\Phi \otimes B\Psi . \quad (7.9)$$

Das Skalarprodukt auf dem Tensorproduktraum  $\mathfrak{H}$  hat die Produkteigenschaft

$$\begin{aligned} \langle \Phi_1 \otimes \Phi_2, \Psi_1 \otimes \Psi_2 \rangle &= (\Phi_1 \otimes \Phi_2)^*(\Psi_1 \otimes \Psi_2) \\ &= (\Phi_1^* \otimes \Phi_2^*)(\Psi_1 \otimes \Psi_2) \\ &= \Phi_1^* \Psi_1 \otimes \Phi_2^* \Psi_2 \\ &= \langle \Phi_1, \Psi_1 \rangle \langle \Phi_2, \Psi_2 \rangle . \end{aligned} \quad (7.10)$$

Hierbei haben wir im letzten Schritt benutzt, dass das Tensorprodukt von  $(1 \times 1)$ -Matrizen wieder eine  $(1 \times 1)$ -Matrix, also eine komplexe Zahl ist.

In der Physik wird das Tensorprodukt oft als direktes Produkt bezeichnet. Dies ist missverständlich, da der Ausdruck direktes Produkt in der Mathematik für eine andere Konstruktion verwendet wird. In der Mathematik wird das Tensorprodukt von Matrizen auch Kroneckerprodukt genannt.

Als eine Anwendung betrachten wir die Kopplung von Drehimpulsen. Sei  $n_1 = n_2 = 2$  und sei  $\mathbf{S} = \frac{1}{2}\vec{\sigma}$  der Spin. Wir definieren den Spin des ersten Systems als

$$\mathbf{S}_1 := \mathbf{S} \otimes \mathbf{1} \quad (7.11)$$

und den des zweiten als

$$\mathbf{S}_2 = \mathbf{1} \otimes \mathbf{S} \quad (7.12)$$

Der Spin des Gesamtsystems ist dann

$$\mathbf{S} = \mathbf{S}_1 + \mathbf{S}_2 \quad (7.13)$$

Für die  $x$ -Komponente z.B. ergibt sich die  $(4 \times 4)$ -Matrix

$$\begin{aligned} S_x &= \frac{1}{2}(\sigma_x \otimes \mathbf{1} + \mathbf{1} \otimes \sigma_x) \quad (7.14) \\ &= \frac{\hbar}{2} \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix} + \frac{1}{2} \begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix} = \frac{1}{2} \begin{pmatrix} 0 & 1 & 1 & 0 \\ 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 \end{pmatrix}. \end{aligned}$$

Die Eigenwerte der Komponenten des Gesamtspins ergeben sich als Summe der Eigenwerte der Komponenten der Spins der Teilsysteme, ihre Eigenvektoren als Tensorprodukte der Eigenvektoren in den Teilsystemen. Besonders einfach sieht man das am Beispiel der  $z$ -Komponente. Es gilt

$$\begin{aligned} S_z &= \frac{1}{2}(\sigma_z \otimes \mathbf{1} + \mathbf{1} \otimes \sigma_z) \quad (7.15) \\ &= \frac{1}{2} \begin{pmatrix} +1 & 0 & 0 & 0 \\ 0 & +1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} + \frac{1}{2} \begin{pmatrix} +1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & +1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} = \begin{pmatrix} +1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}. \end{aligned}$$

Der Eigenwert 0 hat die Vielfachheit 2, die Eigenwerte 1 und -1 sind nicht entartet. Die Aufsteige- und Absteige-Operatoren ergeben sich zu

$$S_+ = \begin{pmatrix} 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}, S_- = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 \end{pmatrix}. \quad (7.16)$$

Der Vektor

$$\Phi_{1,1} = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} \quad (7.17)$$

ist offenbar ein Höchstgewichtsvektor mit Eigenwerten  $s$  für  $S_z$  und  $s(s+1)$  für  $|S|^2$ ,  $s = 1$ . Anwendung von  $S_-$  ergibt

$$S_- \Phi_{1,1} = \begin{pmatrix} 0 \\ 1 \\ 1 \\ 0 \end{pmatrix} = \sqrt{2} \Phi_{1,0} \quad (7.18)$$

und

$$S_-^2 \Phi_{1,1} = \sqrt{2} S_- \Phi_{0,1} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 2 \end{pmatrix} = 2 \Phi_{1,-1} \quad (7.19)$$

Die Vektoren  $\Phi_{1,m}$ ,  $l = 1$ ,  $m = 1, 0, -1$  (eine andere gebräuchliche Notation ist  $\Phi_{l,m} = |l, m\rangle$ ) spannen einen 3-dimensionalen Unterraum  $\mathfrak{H}_1$  des  $\mathbb{C}^4$  auf. Das orthogonale Komplement ist ein 1-dimensionaler Unterraum, der von einem zu  $\Phi_{1,0}$  orthogonalen Eigenvektor  $\Phi_{0,0} = |0, 0\rangle$  von  $S_z$  zum Eigenwert 0 aufgespannt wird,

$$\Phi_{0,0} = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ 1 \\ -1 \\ 0 \end{pmatrix} \quad (7.20)$$

$\Phi_{0,0}$  ist ein weiterer Höchstgewichtsvektor, diesmal zum Eigenwert 0 von  $|\mathbf{S}|^2$ .

Wir erkennen, dass das Tensorprodukt zweier Spin- $\frac{1}{2}$ -Darstellungen in zwei Unterdarstellungen zerfällt, die den Spinquantenzahlen  $s = 1$  und  $s = 0$  entsprechen.

Die Kopplung zweier Spindarstellungen mit beliebigen Spins  $s_1$  und  $s_2$  lässt sich in analoger Weise durchführen. Man erhält eine Zerlegung des  $(2s_1 + 1)(2s_2 + 1)$  dimensionalen Produkt-Hilbertraums in eine direkte Summe von Unterräumen, die den Spinquantenzahlen  $s = s_1 + s_2, s_1 + s_2 - 1, \dots, s_1 - s_2$  entsprechen,

$$\mathfrak{H}_{s_1} \otimes \mathfrak{H}_{s_2} = \bigoplus_{s=|s_1-s_2|}^{s_1+s_2} \mathfrak{H}_s. \quad (7.21)$$

Die Projektoren  $P_s$  auf die Unterräume  $\mathfrak{H}_s$  lassen sich nach der Formel (2.8) explizit angeben. Im Fall  $s_1 = s_2 = \frac{1}{2}$  findet man

$$|\mathbf{S}|^2 = \begin{pmatrix} 2 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 2 \end{pmatrix} \quad (7.22)$$

und damit

$$P_1 = \frac{|\mathbf{S}|^2}{2} = \frac{1}{2} \begin{pmatrix} 2 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 2 \end{pmatrix} \quad (7.23)$$

und

$$P_0 = \frac{|\mathbf{S}|^2 - 2\mathbf{1}}{-2} = \frac{1}{2} \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & +1 & -1 & 0 \\ 0 & -1 & +1 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}. \quad (7.24)$$

## 8. Reine und gemischte Zustände; die Dichtematrix

Bisher haben wir optimal präparierte Zustände betrachtet. Diese konnten durch minimale Projektoren  $P$  oder, äquivalent, durch Einheitsvektoren  $\Psi$  mit  $P = |\Psi\rangle\langle\Psi|$  beschrieben werden. Die Wahrscheinlichkeitsverteilungen von Observablen in einem solchen Zustand werden durch die Erwartungswerte

$$\langle A \rangle = \langle \Psi, A\Psi \rangle = \text{Tr}AP \quad (8.1)$$

bestimmt.

Ist ein Zustand nicht optimal präpariert, wie z.B. beim einfachen Stern-Gerlach-Versuch, so ergibt sich die Wahrscheinlichkeitsverteilung aus der Mittelung von Erwartungswerten in verschiedenen optimal präparierten Zuständen. Im Stern-Gerlach-Versuch sind beide Spinorientierungen gleich wahrscheinlich, daher ist der Erwartungswert einer Observablen  $A$  gegeben durch

$$\langle A \rangle = \frac{1}{2}\text{Tr}AP_+ + \frac{1}{2}\text{Tr}AP_- = \text{Tr}A\left(\frac{1}{2}\mathbf{1}\right) \quad (8.2)$$

mit den Projektoren  $P_{\pm}$  auf die beiden Spinorientierungen.

Im allgemeinen müssen die beiden Spinorientierungen nicht gleich wahrscheinlich sein. Ein Zustand wird dann durch eine Matrix  $\rho$  beschrieben,

$$\langle A \rangle = \text{Tr}A\rho, \quad (8.3)$$

die die folgenden Eigenschaften besitzt:

$$\langle \Psi, \rho\Psi \rangle \geq 0 \forall \Psi. \quad (8.4)$$

$$\text{Tr}\rho = 1. \quad (8.5)$$

Man nennt  $\rho$  die Dichtematrix. Eine Dichtematrix ist hermitesch und hat Eigenwerte, die nicht negativ sein können. Wenn sich  $\rho$  in eine konvexe Kombination anderer Dichtematrizen zerlegen lässt,

$$\rho = \lambda\rho_1 + (1 - \lambda)\rho_2, \quad \lambda \neq 0, 1, \quad \rho_{1,2} \neq \rho, \quad (8.6)$$

nennt man den Zustand gemischt, anderenfalls rein. Ein Zustand ist genau dann rein, wenn die Dichtematrix ein minimaler Projektor ist.

Im Fall der  $(2 \times 2)$ -Matrizen lässt sich der Raum der Dichtematrizen gut geometrisch veranschaulichen. Als hermitesche Matrix kann  $\rho$  in der Form

$$\rho = a\mathbf{1} + \mathbf{b} \cdot \vec{\sigma}, \quad a \in \mathbb{R}, \quad \mathbf{b} \in \mathbb{R}^3 \quad (8.7)$$

geschrieben werden. Aus  $\text{Tr}\rho = 1$  folgt  $a = \frac{1}{2}$ , da die Spur der Pauli-Matrizen verschwindet. Die Positivitätsbedingung an die Eigenwerte ist genau dann erfüllt, wenn die Determinante nicht negativ ist. Man berechnet

$$\det \rho = a^2 - |\mathbf{b}|^2. \quad (8.8)$$

Wir schließen, dass die Dichtematrix  $\rho$  von der Form

$$\rho = \frac{1}{2}(\mathbf{1} + \mathbf{n} \cdot \vec{\sigma}) \quad (8.9)$$

ist, mit  $\mathbf{n} \in \mathbb{R}^3$ ,  $|\mathbf{n}| \leq 1$ . Die Dichtematrizen können also durch die Punkte der Einheitskugel im  $\mathbb{R}^3$  parametrisiert werden (Bloch-Kugel). Die konvexen Kombinationen zweier Dichtematrizen entsprechen dabei den Punkten auf der Verbindungslinie zwischen den beiden Punkten der Einheitskugel. Reine Zustände liegen auf der Oberfläche der Kugel.

Der Ausgangszustand des Stern-Gerlach-Versuchs wird durch den Mittelpunkt der Kugel gegeben. Dieser liegt auf jedem Durchmesser der Kugel, enthält also keinerlei Information über die Komponenten bei einer Zerlegung in reine Zustände. Diese Nichteindeutigkeit der Zerlegung ist ein entscheidender Unterschied zur klassischen Physik. Auch dort treten nicht optimal präparierte Zustände auf, z.B. bei Systemen mit endlicher Temperatur. Doch lassen diese sich immer in eindeutiger Weise als konvexe Kombination reiner Zustände auffassen. So bildet der klassische Zustandsraum bei  $n$  reinen Zuständen ein  $n$ -Simplex in  $n - 1$ -Dimensionen, also eine Strecke für  $n = 2$ , ein Dreieck für  $n = 3$  und einen Tetraeder für  $n = 4$ . Die reinen Zustände sind dabei die Ecken des Simplex.

Bei Quantensystemen mit  $(n \times n)$ -Matrizen mit  $n > 2$  ist die Geometrie des Zustandsraums komplizierter. Für große Werte von  $n$  lassen sich Ähnlichkeiten mit der klassischen Situation finden; auf diese Weise kann die klassische Physik als Grenzfall der Quantenphysik für große Systeme aufgefasst werden.

Man kann auch bei gemischten Zuständen das Schrödingerbild benutzen. Wir gehen wieder von der Erwartungswerten aus und erhalten

aus (8.3)

$$\langle A(t) \rangle_\rho = \text{Tr} A(t)\rho = \text{Tr} U(t)^* A U(t)\rho = \text{Tr} A U(t)\rho U(t)^* = \text{Tr} A \rho(t) \quad (8.10)$$

mit der zeitabhängigen Dichtematrix  $\rho(t) = U(t)\rho U(t)^*$ . Die Schrödingergleichung für die Dichtematrix stimmt bis auf ein Vorzeichen mit der Heisenberggleichung überein,

$$i \frac{d}{dt} \rho(t) = [H, \rho(t)] \quad (8.11)$$

(Liouville-Gleichung, nach der entsprechenden Gleichung in der klassischen statistischen Mechanik, oder auch von Neumann-Gleichung, nach von Neumann, der die wesentlichen Beiträge zur mathematischen Formulierung der Quantentheorie geleistet hat und auf den das Konzept der Dichtematrizen zurück geht.)

### 9. Verschränkung (Entanglement), EPR Paradox und Bellsche Ungleichung

Eine wesentliche Aussage der Quantenmechanik ist, dass auch in einem optimal präparierten Zustand nicht alle Observable scharfe Werte annehmen, sodass nur Wahrscheinlichkeitsaussagen möglich sind. Viele Physiker haben bezweifelt, dass diese Aussage die Wirklichkeit wiedergibt; von Einstein stammt der berühmte Satz: Gott würfelt nicht. Die Hoffnung war, dass es eine umfassendere Theorie gibt, in der dann wieder alle Observable definierte Werte annehmen, sofern der Zustand optimal präpariert worden ist. Diese Theorie sollte zusätzliche Parameter (hidden variables) enthalten, die mit den bis jetzt bekannten Methoden nicht beobachtet werden können. Hingegen haben maßgebliche Anhänger der Quantenmechanik, insbesondere Bohr und Heisenberg, immer die Vollständigkeit der Quantenmechanik behauptet.

Ein wichtiger Beitrag zu dieser Diskussion ist eine Arbeit von Einstein, Podolsky und Rosen aus dem Jahr 1935, in der sie argumentieren, dass die Quantenmechanik nicht vollständig sei. Dieses Papier wurde lange Zeit nicht richtig verstanden; erst in den 50er Jahren fand Bohm (auf den eine Version der Quantenmechanik mit verborgenen Parametern zurückgeht) eine Reformulierung des EPR-Arguments, die so einfach war, dass sie der weiteren Analyse zugänglich wurde. Man betrachtet ein Teilchen mit Spin 0, das in zwei Teilchen mit Spin  $\frac{1}{2}$  zerfällt, die sich in verschiedenen Richtungen bewegen. Betrachtet man nur die Spinfreiheitsgrade, so hat man ein Tensorprodukt zweier Spin- $\frac{1}{2}$ -Systeme im Zustand mit Gesamtspin 0. In diesem Zustand sind die Spinkomponenten in einer beliebigen Richtung  $\mathbf{n}$  genau entgegengesetzt. Misst man nun im System 1 die  $z$ -Komponente und findet den Wert  $\hbar/2$ , so ist die  $z$ -Komponente in System 2 mit Sicherheit  $-\hbar/2$ . Man kann nun an System 2 eine andere Komponente, z.B. die  $x$ -Komponente messen. Nach EPR ist der nicht gemessene (und nach

den Prinzipien der Quantentheorie auch nicht zusammen mit der  $x$ -Komponente messbare) Wert der  $z$ -Komponente ein Element der Realität, das durch die Quantentheorie nicht beschrieben wird; in diesem Sinne sei die Quantentheorie unvollständig.

Bell gelang es, aus den im EPR-Argument gemachten Annahmen eine Ungleichung herzuleiten, die dann experimentell überprüft werden konnte<sup>3</sup>. Die Annahme ist, dass auch ohne Messung die Werte der Spinkomponenten in jedem Einzelversuch festliegen. Sei  $M_1(\mathbf{a}) = \pm 1$  der Wert (in Einheiten von  $\hbar/2$ ) der Komponente des Spins in Richtung  $\mathbf{a}$  in System 1 und entsprechend  $M_2(\mathbf{a})$  der Wert der entsprechenden Komponente in System 2. Da der Gesamtspin Null ist, muss  $M_1(\mathbf{a}) + M_2(\mathbf{a}) = 0$  für alle Richtungen  $\mathbf{a}$  gelten.

Wir wählen jetzt 3 Richtungsvektoren  $\mathbf{a}, \mathbf{b}, \mathbf{c}$  und bemerken, dass die folgende Gleichung gilt:

$$|M_1(\mathbf{a})M_2(\mathbf{b}) - M_1(\mathbf{a})M_2(\mathbf{c})| = 1 + M_1(\mathbf{b})M_2(\mathbf{c}) . \quad (9.1)$$

PROOF. Ist  $M_2(\mathbf{b}) = M_2(\mathbf{c})$ , dann ist  $M_1(\mathbf{b}) = -M_2(\mathbf{c})$ , also verschwinden beide Seiten der Ungleichung. Ist  $M_2(\mathbf{b}) \neq M_2(\mathbf{c})$ , so ist  $M_1(\mathbf{b}) = M_2(\mathbf{c})$ , und beide Seiten sind gleich 2. ■

Wir wiederholen den Versuch jetzt mehrfach und bilden die Mittelwerte der Produkte  $M_1(\mathbf{e})M_2(\mathbf{f})$  für  $\mathbf{e}, \mathbf{f} = \mathbf{a}, \mathbf{b}, \mathbf{c}$ . Sei  $n(\alpha, \beta, \gamma)$  die relative Häufigkeit dafür, dass  $M_2(\mathbf{a}) = \alpha, M_2(\mathbf{b}) = \beta, M_2(\mathbf{c}) = \gamma$  ist. Dann gilt für die Mittelwerte die Abschätzung

$$|\langle M_1(\mathbf{a})M_2(\mathbf{b}) \rangle - \langle M_1(\mathbf{a})M_2(\mathbf{c}) \rangle| \quad (9.2)$$

$$= \left| \sum_{\alpha, \beta, \gamma = \pm 1} n(\alpha, \beta, \gamma)(-\alpha\beta + \alpha\gamma) \right| \quad (9.3)$$

$$\leq \sum_{\alpha, \beta, \gamma = \pm 1} n(\alpha, \beta, \gamma)|-\alpha\beta + \alpha\gamma| \quad (9.4)$$

$$= \sum_{\alpha, \beta, \gamma = \pm 1} n(\alpha, \beta, \gamma)(1 - \beta\gamma) \quad (9.5)$$

$$= 1 + \langle M_1(\mathbf{b})M_2(\mathbf{c}) \rangle . \quad (9.6)$$

Hierbei haben wir im Schritt von der 3. zur 4. Zeile die Gleichung (9.1) benutzt.

Wir betrachten nun die Vorhersage der Quantenmechanik. Der Zustand des Gesamtsystems wird durch den minimalen Projektor  $P_0$  aus

<sup>3</sup>Wir folgen hier der Darstellung aus dem zitierten Buch von Gernot Münster.

(7.24) auf den Spin-0-Unterraum beschrieben. Wir berechnen

$$\begin{aligned} \langle M_1(\mathbf{e})M_2(\mathbf{f}) \rangle &= \text{Tr}(\mathbf{e} \cdot \vec{\sigma} \otimes \mathbf{f} \cdot \vec{\sigma})P_0 = \\ \text{Tr} \left( \begin{array}{cccc} e_3 f_3 & e_3(f_1 - if_2) & (e_1 - ie_2)f_3 & (e_1 - ie_2)(f_1 - if_2) \\ e_3(f_1 + if_2) & -e_3 f_3 & (e_1 - ie_2)(f_1 + if_2) & -(e_1 - ie_2)f_3 \\ (e_1 + ie_2)f_3 & (e_1 + ie_2)(f_1 - if_2) & -e_3 f_3 & -e_3(f_1 - if_2) \\ (e_1 + ie_2)(f_1 + if_2) & -(e_1 + ie_2)f_3 & -e_3(f_1 + if_2) & e_3 f_3 \end{array} \right) P_0 \\ &= \frac{1}{2}(-e_3 f_3 - (e_1 - ie_2)(f_1 + if_2) - (e_1 + ie_2)(f_1 - if_2) - e_3 f_3) = -\mathbf{e} \cdot \mathbf{f} . \end{aligned} \quad (9.7)$$

Setzen wir dies in die Bellsche Ungleichung ein, so erhalten wir die Ungleichung

$$|\mathbf{a} \cdot (\mathbf{b} - \mathbf{c})| \leq 1 - \mathbf{b} \cdot \mathbf{c} . \quad (9.8)$$

Die Schwarzsche Ungleichung liefert für Einheitsvektoren  $\mathbf{a}, \mathbf{b}, \mathbf{c}$

$$|\mathbf{a} \cdot (\mathbf{b} - \mathbf{c})| \leq |\mathbf{b} - \mathbf{c}| = \sqrt{2(1 - \mathbf{b} \cdot \mathbf{c})} . \quad (9.9)$$

Falls  $\mathbf{a}$  parallel zu  $\mathbf{b} - \mathbf{c}$  ist, gilt das Gleichheitszeichen. Setzen wir  $\cos \theta = \mathbf{b} \cdot \mathbf{c}$  mit  $\theta \in [0, \pi]$ , so liefert die Quantenmechanik mit Hilfe der Formel  $(1 - \cos \theta) = 2 \sin^2 \frac{\theta}{2}$  die Schranke  $2 \sin \frac{\theta}{2}$ . Diese liegt für  $\theta \neq 0, \pi$  immer oberhalb der Schranke  $2 \sin^2 \frac{\theta}{2}$  aus der Bellschen Ungleichung.

Messungen haben ergeben, dass tatsächlich die Bellschen Ungleichungen verletzt sind, und dass die Vorhersagen der Quantentheorie zutreffen. Will man also nicht eine Wechselwirkung zwischen den beiden Systemen annehmen, so bleibt nur der Schluss, dass die physikalischen Observablen im allgemeinen vor einer Messung keinen definierten Wert besitzen.

Während auf der Seite der Observablen die Kopplung unabhängiger Teilsysteme mit Hilfe des Tensorproduktes möglich ist, ist die Situation auf der Seite der Zustände komplizierter. Betrachten wir zunächst reine Zustände. Sei  $\Phi \in \mathfrak{H}_1$  und  $\Psi \in \mathfrak{H}_2$ . Im Produktzustand  $\Phi \otimes \Psi$  faktorisieren alle Erwartungswerte, die Messungen an den beiden Teilsystemen sind unkorreliert. Bilden wir nun Linearkombinationen  $\Omega = \sum_{i=1}^n \Phi_i \otimes \Psi_i$  von Produktvektoren, so ergeben sich bei den Erwartungswerten von Tensorprodukten  $A \otimes B$  von Observablen der Teilsysteme

$$\langle \Omega, (A \otimes B)\Omega \rangle = \sum_{i,j=1}^n \langle \Phi_i, A\Phi_j \rangle \langle \Psi_i, B\Psi_j \rangle \quad (9.10)$$

neben der Summe der Diagonalterme ( $i = j$ ), die klassischen Korrelationen entspricht, auch die Beiträge der Nichtdiagonalterme ( $i \neq j$ ), für die es kein klassisches Analogon gibt.

Schrödinger hat solche Zustände verschränkt (englisch: entangled) genannt, als Reaktion auf das EPR-Argument. Bei gemischten Zuständen

unterscheidet man separable von verschränkten Zuständen. Ein Zustand mit Dichtematrix  $\rho$  heißt separabel, wenn  $\rho$  als konvexe Kombination von Tensorprodukten von Dichtematrizen der Teilsysteme dargestellt werden kann,

$$\rho = \sum \lambda_j \rho_j^{(1)} \otimes \rho_j^{(2)} \quad (9.11)$$

mit Dichtematrizen  $\rho_j^{(1)} \in \mathfrak{A}_1$  und  $\rho_j^{(2)} \in \mathfrak{A}_2$  und positiven Zahlen  $\lambda_j$  mit  $\sum_j \lambda_j = 1$ . In separablen Zuständen sind die Bellschen Ungleichungen erfüllt. Denn seien  $\rho_j^{(1)}, \rho_j^{(2)}, j = 1 \dots, n$  Dichtematrizen mit der Eigenschaft

$$\text{Tr}(\mathbf{n} \cdot \vec{\sigma}) \rho_j^{(1)} = -\text{Tr}(\mathbf{n} \cdot \vec{\sigma}) \rho_j^{(2)} \quad \forall \mathbf{n} \in \mathbb{R}^3. \quad (9.12)$$

Dann gilt für alle  $j$

$$|\text{Tr}(\mathbf{a} \cdot \vec{\sigma} \otimes \mathbf{b} \cdot \vec{\sigma})(\rho_j^{(1)} \otimes \rho_j^{(2)}) - \text{Tr}(\mathbf{a} \cdot \vec{\sigma} \otimes \mathbf{c} \cdot \vec{\sigma})(\rho_j^{(1)} \otimes \rho_j^{(2)})| \quad (9.13)$$

$$= |\text{Tr}(\mathbf{a} \cdot \vec{\sigma}) \rho_j^{(1)}| |\text{Tr}(\mathbf{b} \cdot \vec{\sigma}) \rho_j^{(2)} - \text{Tr}(\mathbf{c} \cdot \vec{\sigma}) \rho_j^{(2)}| \quad (9.14)$$

$$= |\text{Tr}(\mathbf{a} \cdot \vec{\sigma}) \rho_j^{(1)}| |\text{Tr}(\mathbf{b} \cdot \vec{\sigma}) \rho_j^{(1)} + \text{Tr}(\mathbf{c} \cdot \vec{\sigma}) \rho_j^{(2)}|. \quad (9.15)$$

Wir nutzen jetzt aus, dass der Erwartungswert  $\langle \mathbf{n} \rangle$  von  $\mathbf{n} \cdot \sigma$  für Einheitsvektoren  $\mathbf{n}$  immer zwischen -1 und 1 liegt. Wir finden

$$|\langle \mathbf{a} \rangle| |\langle \mathbf{b} \rangle + \langle \mathbf{c} \rangle| \leq |\langle \mathbf{b} \rangle + \langle \mathbf{c} \rangle| \leq 1 + \langle \mathbf{b} \rangle \langle \mathbf{c} \rangle \quad (9.16)$$

und damit die Bellsche Ungleichung für jedes  $j$ . Hierbei haben wir im letzten Schritt die Ungleichung

$$0 \leq (1 \pm \langle \mathbf{b} \rangle)(1 \pm \langle \mathbf{c} \rangle) = 1 + \langle \mathbf{b} \rangle \langle \mathbf{c} \rangle \pm (\langle \mathbf{b} \rangle + \langle \mathbf{c} \rangle) \quad (9.17)$$

verwendet. Die Ungleichung für den Zustand mit der Dichtematrix  $\rho$  folgt jetzt durch Multiplikation der Ungleichung für  $j$  mit der positiven Zahl  $\lambda_j$  und Summation über  $j$ . Beide Operationen erhalten die Ungleichung.

Ein Zustand heißt verschränkt, wenn er nicht separabel ist. Die Verletzung der Bellschen Ungleichungen durch geeignete Zustände zeigt also, dass nicht alle Zustände separabel sind. Verschränkte Zustände lassen sich nicht auf Zustände der Teilsysteme zurückführen. Dieser Aspekt der Quantentheorie wird als Quanten-Nichtlokalität bezeichnet.

Im allgemeinen ist es schwierig, die Menge der verschränkten Zustände explizit anzugeben. Die verschränkten Zustände spielen eine entscheidende Rolle in der Quanteninformationstheorie.



## KAPITEL III

### Unendlich dimensionale Quantensysteme

Die wesentlichen Züge der Quantentheorie zeigen sich bereits in endlich dimensionalen Systemen. Für die physikalischen Anwendungen spielen aber die Systeme, in denen die Observablen unendlich viele verschiedene Werte annehmen können, eine wichtige Rolle. Hier gibt es zwei verschiedene Fälle: das diskrete Spektrum, wie z.B. bei den Bindungsenergien von Atomen, und das kontinuierliche Spektrum wie bei den kinetischen Energien freier Teilchen oder bei Observablen wie den Ortskoordinaten. Während Observable mit diskretem Spektrum sich ähnlich wie Observable mit endlich vielen Werten verhalten, sind Observable mit kontinuierlichem Spektrum ein neues Phänomen, das nur in unendlich dimensionalen Hilberträumen auftreten kann. In vielen Fällen besitzen Observable kontinuierliches und diskretes Spektrum, wie z.B. der Hamiltonoperator des Wasserstoffatoms, mit den diskreten Bindungsenergien und dem kontinuierlichen Spektrum des ionisierten Atoms.

#### 1. Unendlich dimensionale Hilberträume

Die Mathematik unendlich dimensionaler Hilberträume ist komplizierter als im endlich dimensionalen Fall. Wir betrachten zunächst zwei Beispiele:

**1.1. Der Folgenraum  $l^2$ .** Eine direkte Verallgemeinerung des  $\mathbb{C}^n$  mit seinem natürlichen Skalarprodukt ist der Raum der Folgen komplexer Zahlen  $\Phi = (\Phi_n)_{n \in \mathbb{N}}$  mit dem Skalarprodukt

$$\langle \Phi, \Psi \rangle = \sum_{n=1}^{\infty} \overline{\Phi_n} \Psi_n . \quad (1.1)$$

Um zu garantieren, dass die Summen konvergieren, verlangt man, dass nur Folgen zugelassen werden, für die die Summe der Absolutquadrate konvergiert,

$$\|\Phi\|^2 \equiv \sum_{n=1}^{\infty} |\Phi_n|^2 < \infty . \quad (1.2)$$

Dann folgt die Konvergenz des Skalarprodukts aus der Schwarzischen Ungleichung. Den Raum der quadratsummablen Folgen bezeichnet man mit dem Symbol  $l^2$ .

Beispiele für quadratsummable Folgen sind

- (i) Sei  $\Phi_n = q^n$ ,  $q \in \mathbb{C}$ ,  $|q| \neq 1$ . Dann gilt nach der Summenformel für die geometrische Reihe

$$\sum_{n=1}^N |\Phi_n|^2 = \frac{|q|^2 - |q|^{2(N+1)}}{1 - |q|^2} . \quad (1.3)$$

Für  $|q| < 1$  strebt  $|q|^{2(N+1)}$  gegen Null. Daher ist in diesem Fall

$$\sum_{n=1}^{\infty} |\Phi_n|^2 = \frac{|q|^2}{1 - |q|^2} < \infty . \quad (1.4)$$

Für  $|q| \geq 1$  hingegen divergiert die Summe.

- (ii) Sei  $\Phi_n = \frac{1}{n}$ . Dann ist

$$\sum_{n=1}^N |\Phi_n|^2 \leq 1 + \sum_{n=2}^N \frac{1}{n(n-1)} . \quad (1.5)$$

Mit

$$\frac{1}{n-1} - \frac{1}{n} = \frac{1}{n(n-1)} \quad (1.6)$$

folgt

$$\sum_{n=2}^N \frac{1}{n(n-1)} = \left(1 - \frac{1}{2}\right) + \left(\frac{1}{2} - \frac{1}{3}\right) + \left(\frac{1}{3} - \frac{1}{4}\right) + \cdots + \left(\frac{1}{N-1} - \frac{1}{N}\right) = 1 - \frac{1}{N} . \quad (1.7)$$

Daher ist

$$\sum_{n=1}^{\infty} \frac{1}{n^2} < 2 < \infty . \quad (1.8)$$

Tatsächlich lässt sich diese unendliche Summe auch explizit berechnen, das Ergebnis ist  $\frac{\pi^2}{6}$ .

Ein Beispiel für eine Summe, die nicht quadratsummabel ist, ist  $\Phi_n = \frac{1}{\sqrt{n}}$ . Dies kann man in der folgenden Weise einsehen: Es gilt für  $k > 2$

$$\sum_{n=1}^{2^k} = 1 + \frac{1}{2} + \left(\frac{1}{3} + \frac{1}{4}\right) + \cdots + \left(\sum_{n=2^{k-1}+1}^{2^k} \frac{1}{n}\right) \quad (1.9)$$

Wir haben dabei die Summe in  $k+1$  Summanden zerlegt, von denen jeder  $\geq \frac{1}{2}$  ist. Daher ist die Summe  $\geq \frac{k+1}{2}$  und divergiert für  $k \rightarrow \infty$ .

**1.2. Der Funktionenraum  $L^2(\mathbb{R})$ .** Auf dem Raum der stetigen Funktionen einer reellen Variablen kann man durch

$$\langle \Phi, \Psi \rangle = \int_{-\infty}^{\infty} \overline{\Phi(x)} \Psi(x) dx \quad (1.10)$$

ein Skalarprodukt einführen. Um zu garantieren, dass dieses Integral konvergiert, lässt man nur Funktionen zu, die quadratintegabel sind,

$$\|\Phi\|^2 = \int |\Phi(x)|^2 dx < \infty . \quad (1.11)$$

Wieder folgt mittels der Schwarzischen Ungleichung, dass für solche Funktionen das Skalarprodukt wohl definiert ist. Beispiele für quadratintegable Funktionen:

(i) Sei  $\Phi(x) = \frac{1}{x+i}$ . Dann gilt für  $R > 1$

$$\int_{-R}^R |\Phi(x)|^2 dx = 2 \int_0^R \frac{dx}{x^2+1} \leq \int_0^1 dx + \int_1^R \frac{dx}{x^2} = 1 - \frac{1}{x} \Big|_1^R = 2 - \frac{1}{R} \leq 2 , \quad (1.12)$$

das Integral konvergiert also für  $R \rightarrow \infty$ . Tatsächlich kann man es auch explizit berechnen, indem man  $x = \tan \alpha$  substituiert. Man findet unter Benutzung der Formel  $\frac{d}{d\alpha} \tan \alpha = \tan^2 \alpha + 1$

$$\int_0^R \frac{dx}{x^2+1} = \int_0^{\arctan R} d\alpha = \arctan R \quad (1.13)$$

Mit  $\arctan R \rightarrow \frac{\pi}{2}$  für  $R \rightarrow \infty$  folgt

$$\int_{-\infty}^{\infty} \frac{dx}{x^2+1} = \pi . \quad (1.14)$$

(ii) Sei  $\Phi(x) = e^{-|x|}$ . Es gilt

$$\int_{-R}^R |\Phi(x)|^2 dx = 2 \int_0^R dx e^{-2x} = -e^{-2x} \Big|_0^R = 1 - e^{-2R} \rightarrow 1 \text{ für } R \rightarrow \infty . \quad (1.15)$$

Der Raum der quadratintegablen Funktionen  $L^2(\mathbb{R})$  enthält auch noch gewisse unstetige Funktionen, z.B. stückweise stetige Funktionen mit endlich vielen Sprungstellen.

Entsprechend definiert man die Räume  $L^2([a, b])$  für ein Intervall  $[a, b]$  der reellen Achse oder, allgemeiner, für Gebiete des  $\mathbb{R}^n$ .

**1.3. Orthonormalbasen.** Wie in endlich dimensionalen Hilberträumen gibt es auch in unendlich dimensionalen Hilberträumen Orthonormalbasen. Eine Orthonormalbasis ist dabei eine maximale Menge normierter und paarweise orthogonaler Vektoren.

Im  $l^2$  kann man die kanonische Basis  $\{\Phi^{(j)}, j = 1, \dots, \infty\}$  mit  $\Phi_n^{(j)} = \delta_{jn}$  wählen. Einen Hilbertraum mit einer abzählbaren Orthonormalbasis nennt man separabel. In der Quantenmechanik treten fast ausschließlich separable Hilberträume auf.

Sei  $(\Phi^{(j)})_{j=1, \dots, \infty}$  eine Orthonormalbasis. Dann kann jedem Vektor  $\Psi$  des Hilbertraums eine Folge komplexer Zahlen

$$\Psi_j = \langle \Phi^{(j)}, \Psi \rangle , \quad j = 1 \dots \infty \quad (1.16)$$

zugeordnet werden. Es gilt

$$\sum_{j=1}^{\infty} |\langle \Phi^{(j)}, \Psi \rangle|^2 = \|\Psi\|^2 . \quad (1.17)$$

Die Folge  $(\Psi_j)$  ist also quadratsummabel. Umgekehrt gibt es zu jeder quadratsummablen Folge komplexer Zahlen  $(\Psi_n)_{n \in \mathbb{N}}$  einen eindeutig bestimmten Vektor  $\Psi$  des Hilbertraums, der mit den Elementen der Orthonormalbasis die Relation (1.16) erfüllt. Wir schreiben

$$\Psi = \sum_{n=1}^{\infty} \Psi_n \Phi^{(n)} . \quad (1.18)$$

Die unendliche Summe bedeutet dabei, dass

$$\|\Psi - \sum_{n=1}^N \Psi_n \Phi^{(n)}\| \rightarrow 0 \text{ für } N \rightarrow \infty \quad (1.19)$$

gilt.

Wir schließen, dass jeder unendlich dimensionale separable Hilbertraum isomorph zu  $l^2$  ist.

**1.4. Operatoren.** Operatoren auf  $l^2$  lassen sich als Matrizen mit unendlich vielen Zeilen und Spalten auffassen,

$$(A\Phi)_n = \sum_{k=1}^{\infty} A_{nk} \Phi_k . \quad (1.20)$$

Bei der Multiplikation dieser Matrizen treten allerdings unendliche Summen auf, deren Konvergenz kontrolliert werden muss.

Auf  $L^2(\mathbb{R})$  sind Integraloperatoren

$$(A\Phi)(x) = \int dy A(x, y) \Phi(y) \quad (1.21)$$

kontinuierliche Verallgemeinerungen von Matrizen. Weitere Operatoren sind Multiplikationsoperatoren

$$\Phi \mapsto f\Phi , (f\Phi)(x) = f(x)\Phi(x) , \quad (1.22)$$

bei denen die Funktion  $\Phi$  mit einer Funktion  $f$  multipliziert wird, und Differentialoperatoren

$$\Phi \mapsto \frac{d}{dx} \Phi . \quad (1.23)$$

Eine Schwierigkeit bei Operatoren auf unendlich dimensionalen Hilberträumen ist, dass sie oft nur auf einem Teilraum definiert sind; z.B. ist der Differentialoperator  $\frac{d}{dx}$  nur auf differenzierbaren Funktionen erklärt, und man muss zusätzlich verlangen, dass die Ableitung quadratintegrabel ist.

## 2. Teilchen im äußeren Feld

Die Observablen für ein Teilchen sind Operatoren auf  $L^2(\mathbb{R}^3)$ . Hierbei sind die Multiplikationsoperatoren mit den Komponenten  $x_j$  des Ortsvektors die Ortsobservablen, und die partiellen Ableitungen nach den Koordinaten ergeben die Impulsoperatoren,

$$p_j = \frac{\hbar}{i} \frac{\partial}{\partial x_j} . \quad (2.1)$$

Orts- und Impulsoperatoren erfüllen die kanonischen Vertauschungsrelationen

$$[x_j, p_k] = i\hbar\delta_{jk} . \quad (2.2)$$

Hieraus folgen die Unschärferelationen

$$\Delta x_j \Delta p_k \geq \frac{\hbar}{2} \delta_{jk} . \quad (2.3)$$

Je genauer also ein Teilchen im Ort konzentriert ist, umso unbestimmter wird sein Impuls, und umgekehrt.

Orts- und Impulsoperatoren besitzen keine Eigenvektoren in  $L^2(\mathbb{R}^3)$ . Bei der Analyse dieses Problems beschränken wir uns der einfacheren Notation wegen auf Teilchen, die sich nur in einer Dimension bewegen können, ersetzen also  $L^2(\mathbb{R}^3)$  durch  $L^2(\mathbb{R})$ .

Statt der Eigenvektoren betrachtet man Folgen von Funktionen  $\Phi_n$ , die einen Eigenvektor zum Eigenwert  $a$  im folgenden Sinne approximieren:

$$\int dx |\Phi_n(x)|^2 = 1 , \quad \int dx |(x-a)^2 \Phi_n(x)|^2 \rightarrow 0 \text{ für } n \rightarrow \infty \quad (2.4)$$

Man nennt  $a$  einen verallgemeinerten Eigenwert. Die möglichen Messwerte der Observablen sind die verallgemeinerten Eigenwerte. Eigenprojektoren zu einem verallgemeinerten Eigenwert gibt es i.a. nicht. Es gibt aber Projektoren  $P(I)$ , die einem Intervall der reellen Achse zugeordnet sind; dies sind Multiplikationsoperatoren mit der Funktion, die 1 ist für  $x \in I$  und sonst verschwindet. Man nennt sie Spektralprojektoren; die Menge der verallgemeinerten Eigenwerte ist das Spektrum. Ist  $\Psi \in L^2(\mathbb{R})$  mit  $\|\Psi\| = 1$  der Zustandsvektor eines reinen Zustands, so ist

$$W(x \in [a, b] | \Psi) = \langle \Psi, P([a, b]) \Psi \rangle = \int_a^b |\Psi(x)|^2 dx \quad (2.5)$$

die Wahrscheinlichkeit dafür, dass bei einer Messung des Ortes des Teilchens der Messwert im Intervall  $[a, b]$  liegt.

Ganz ähnlich verhält sich der Impulsoperator  $p$ . Lösungen der Eigenwertgleichung für den Eigenwert  $k$  sind die Funktionen  $e^{ikx}$ . Diese sind allerdings nicht quadratintegabel. Wir multiplizieren diese Funktionen mit der quadratintegablen Funktion  $f_n(x) = (\frac{x}{n} + i)^{-1}$ . Wir

berechnen

$$\int dx |e^{ikx} f_n(x)|^2 = \int dx \frac{1}{\frac{x^2}{n^2} + 1} = n \int dy \frac{1}{y^2 + 1} = n\pi . \quad (2.6)$$

Wir setzen daher

$$\Phi_n(x) = \frac{1}{\sqrt{n\pi}} e^{ikx} \left(\frac{x}{n} + i\right)^{-1} \quad (2.7)$$

und erhalten eine Folge normierter Funktionen. Es gilt

$$((p-k)\Phi_n)(x) = \frac{1}{\sqrt{n\pi}} \frac{\hbar}{i} e^{ikx} (-1) \left(\frac{x}{n} + i\right)^{-2} \frac{1}{n} = \frac{i\hbar}{\sqrt{n^3\pi}} e^{ikx} \left(\frac{x}{n} + i\right)^{-2} \quad (2.8)$$

und damit

$$\int dx |(p-k)\Phi_n(x)|^2 = \frac{\hbar^2}{n^3\pi} \int dx \frac{1}{\left(\frac{x^2}{n^2} + 1\right)^2} = \frac{\hbar^2}{n^2\pi} \int dy \frac{1}{(y^2 + 1)^2} = \frac{\hbar^2}{2n^2} . \quad (2.9)$$

Diese Folge konvergiert für  $n \rightarrow \infty$  gegen 0, also ist jedes  $k \in \mathbb{R}$  ein verallgemeinerter Eigenwert des Impulses.

Die Spektralprojektoren des Impulses ergeben sich als Integraloperatoren der Form

$$(P([a, b])\Phi)(x) = \frac{1}{2\pi} \int_a^b dk \int dy e^{ik(x-y)} \Psi(y) \equiv \int_a^b dk |k\rangle \langle k| \Psi . \quad (2.10)$$

Die letzte Formel ist eine formale Analogie zum Fall der Entwicklung nach Eigenvektoren. Hier fasst man die Funktionen  $|k\rangle := \frac{1}{\sqrt{2\pi}} e^{ikx}$  als (uneigentliche) Eigenvektoren auf, die selbst nicht Elemente des Hilbertraums sind, und die eine Normierungsbedingung der Form

$$\langle k || k' \rangle = \delta(k - k') \quad (2.11)$$

erfüllen. Im Fall des Impulses ist  $\langle k | \Phi$  die Fouriertransformierte  $\hat{\Phi}$  von  $\Phi$  an der Stelle  $k$ , und die Normierungsbedingung ist äquivalent zur Parsevalschen Gleichung

$$\int dk |\hat{\Phi}(k)|^2 = \int dx |\Phi(x)|^2 . \quad (2.12)$$

Orts- und Impulsoperator sind hermitesch. Für den Ortsoperator folgt dies unmittelbar aus der Definition, für den Impulsoperator folgt es aus der folgenden Betrachtung. Seien  $\Phi$  und  $\Psi$  stetig differenzierbare Funktionen, die außerhalb eines endlichen Intervalls  $[a, b]$  verschwinden. Dann gilt

$$\langle p\Phi, \Psi \rangle - \langle \Phi, p\Psi \rangle = i\hbar \int dx \left( \overline{\frac{d}{dx} \Phi(x)} \Psi(x) + \overline{\Phi(x)} \frac{d}{dx} \Psi(x) \right) \quad (2.13)$$

$$= i\hbar \int_a^b dx \frac{d}{dx} \overline{\Phi(x)} \Psi(x) = i\hbar (\overline{\Phi(b)} \Psi(b) - \overline{\Phi(a)} \Psi(a)) = 0 . \quad (2.14)$$

Das Argument kann verallgemeinert werden auf Funktionen, die bei unendlich gegen Null konvergieren.

Die kinetische Energie ist als Operator auf  $L^2(\mathbb{R}^3)$  durch

$$T = \frac{|\mathbf{p}|^2}{2m} = -\frac{\hbar^2}{2m}\Delta \quad (2.15)$$

definiert, mit dem Laplaceoperator  $\Delta = \sum_{j=1}^3 \frac{\partial^2}{\partial x_j^2}$ . Auch dieser Operator ist hermitesch und besitzt als verallgemeinerte Eigenwerte alle reellen Zahlen  $\geq 0$ . Die uneigentlichen Impulseigenfunktionen sind auch uneigentliche Eigenfunktionen der kinetischen Energie.

Die potentielle Energie  $U$  in einem konservativem Kraftfeld wird als Multiplikationsoperator mit der klassischen potentiellen Energie dargestellt und meist mit demselben Symbol bezeichnet. Die verallgemeinerten Eigenwerte sind die Werte der klassischen potentiellen Energie.

Der Hamiltonoperator als Observable der Gesamtenergie ist definiert als die Summe der beiden Operatoren,

$$H = T + U . \quad (2.16)$$

Damit nimmt die Schrödingergleichung für ein Teilchen die Form an

$$i\hbar \frac{\partial}{\partial t} \psi(t, \mathbf{x}) = -\frac{\hbar^2}{2m} \Delta \psi(t, \mathbf{x}) + U(\mathbf{x}) \psi(t, \mathbf{x}) . \quad (2.17)$$

Wie im endlich dimensionalen Fall lösen wir die Schrödingergleichung, indem wir die Spektraldarstellung des Hamiltonoperators konstruieren. Hierzu suchen wir die Eigenwerte und Eigenvektoren (diese entsprechen den Bindungszuständen) sowie die verallgemeinerten Eigenwerte und die zugehörigen uneigentlichen Eigenfunktionen (diese können als Streuzustände interpretiert werden).

Die Eigenwertgleichung

$$H\psi = E\psi \quad (2.18)$$

für eine Ortsfunktion  $\psi$  wird als zeitunabhängige Schrödingergleichung bezeichnet.

### 3. Der harmonische Oszillator

Wir betrachten ein Teilchen der Masse  $m$ , das sich in einer Dimension unter dem Einfluss des Potentials  $U = \frac{k}{2}x^2$  bewegt. In der klassischen Physik führt ein solches Teilchen Schwingungen aus mit der Kreisfrequenz  $\omega = \sqrt{k/m}$ , und die Lösung der Bewegungsgleichung ist

$$x(t) = x(0) \cos \omega t + \frac{\dot{x}(0)}{\omega} \sin \omega t . \quad (3.1)$$

In der Hamiltonschen Form lauten die Bewegungsgleichungen

$$\dot{x} = \frac{p}{m} \quad (3.2)$$

$$\dot{p} = -kx . \quad (3.3)$$

Wir betrachten jetzt die Heisenbergsche Bewegungsgleichung für Ort- und Impulsoperator. Wir finden

$$\dot{x} = \frac{i}{\hbar}[H, x] = \frac{i}{\hbar}\left[\frac{p^2}{2m}, x\right] = \frac{p}{m} \quad (3.4)$$

$$\dot{p} = \frac{i}{\hbar}[H, p] = \frac{i}{\hbar}\left[\frac{k}{2}x^2, p\right] = -kx . \quad (3.5)$$

Wir beobachten, dass die quantenmechanischen Bewegungsgleichungen für dieses System mit den klassischen Bewegungsgleichungen identisch sind. Daher ist die Lösung (3.1) mit  $\dot{x} = \frac{p}{m}$  auch in der Quantenmechanik gültig, wenn Ort und Impuls als Operatoren aufgefasst werden. Wir identifizieren den Ortsoperator zur Zeit  $t = 0$  mit dem Operator der Multiplikation mit  $x$  und den Impulsoperator zur Zeit  $t = 0$  mit dem Differentialoperator  $p = \frac{\hbar}{i} \frac{d}{dx}$ . Die Lösung der Heisenberggleichung ist dann

$$x(t) = x \cos \omega t + \frac{p}{m\omega} \sin \omega t \quad (3.6)$$

$$p(t) = p \cos \omega t - m\omega x \sin \omega t . \quad (3.7)$$

Wir zerlegen jetzt die operatorwertige Funktion  $x(t)$  nach positiven und negativen Frequenzen. Wir erhalten

$$x(t) = x_+ e^{i\omega t} + x_- e^{-i\omega t} \quad (3.8)$$

mit

$$x_{\pm} = \frac{1}{2}\left(x \pm \frac{p}{im\omega}\right) . \quad (3.9)$$

Die Operatoren  $x_{\pm}$  erfüllen die Vertauschungsrelationen

$$[H, x_{\pm}] = \pm \hbar\omega x_{\pm} , \quad (3.10)$$

und

$$[x_-, x_+] = \frac{\hbar}{2m\omega} \quad (3.11)$$

Wir erkennen, dass die Anwendung des Operators  $x_+$  auf einen Eigenvektor von  $H$  den Eigenwert um  $\hbar\omega$  erhöht, während  $x_-$  ihn um  $\hbar\omega$  verringert. Da alle Erwartungswerte von  $H$  positiv sind, kann  $H$  keine negativen Eigenwerte haben. Wir suchen daher einen Eigenvektor  $\Phi_0$  mit dem kleinstmöglichen Eigenwert. Dieser muss die Gleichung

$$x_- \Phi_0 = 0 \quad (3.12)$$

erfüllen.

Einsetzen der Definition von  $x_-$  ergibt die Differentialgleichung

$$\frac{d}{dx} \Phi_0(x) = -\frac{m\omega x}{\hbar} \Phi_0(x) , \quad (3.13)$$

oder, äquivalent

$$\frac{d}{dx} \ln \Phi_0(x) = -\frac{m\omega x}{\hbar} . \quad (3.14)$$

Wir erhalten die Lösung

$$\Phi_0(x) = C e^{-\frac{m\omega x^2}{2\hbar}} \quad (3.15)$$

mit  $C \in \mathbb{C}$ ,  $C \neq 0$ .

$\Phi_0$  ist in der Tat ein Eigenvektor von  $H$ ,

$$H\Phi_0(x) = \frac{1}{2}\hbar\omega\Phi_0(x) . \quad (3.16)$$

In Analogie zur Analyse der Darstellungen der Drehimpulsvertauschungsrelation können wir  $\Phi_0$  einen Niedrigstgewichtsvektor nennen. Wie dort erhalten wir Eigenvektoren mit höheren Eigenwerten durch wiederholte Anwendung des Operators  $x_+$ . Im Gegensatz zum Drehimpuls endet diese Reihe nicht, da die Gleichung  $x_+\Phi = 0$  keine quadratintegrale von Null verschiedene Lösung besitzt.

Wir schließen also, dass  $H$  Eigenwerte der Form  $E_n = (n + \frac{1}{2})\hbar\omega$ ,  $n \in \mathbb{N}_0$  besitzt. Die zugehörigen normierten Eigenvektoren

$$\Phi_n = x_+^n \Phi_0 \|x_+^n \Phi_0\|^{-1} \quad (3.17)$$

bilden eine Orthonormalbasis von  $L^2(\mathbb{R})$ .

Es ist üblich, die Operatoren

$$a = \sqrt{\frac{2m\omega}{\hbar}} x_- \quad (3.18)$$

(den Vernichtungsoperator) und

$$a^* = \sqrt{\frac{2m\omega}{\hbar}} x_+ \quad (3.19)$$

(den Erzeugungsoperator) einzuführen. Diese erfüllen die Vertauschungsrelation

$$[a, a^*] = 1 . \quad (3.20)$$

Der Operator

$$N := a^* a \quad (3.21)$$

hat als Eigenwerte die natürlichen Zahlen (einschließlich der Null) und wird Teilchenzahloperator genannt. Man fasst dann den Grundzustand  $\Phi_0 =: |0\rangle$  als den Nullteilchenzustand (das Vakuum) auf. Der  $n$ -te angeregte Zustand  $\Phi_n =: |n\rangle$  wird als  $n$ -Teilchenzustand gedeutet. Der Zusammenhang zwischen  $N$  und  $H$  ist

$$H = \hbar\omega(N + \frac{1}{2}) . \quad (3.22)$$

#### 4. Das freie Teilchen in einer Dimension

Wirken keine Kräfte auf das Teilchen, so ist der Hamiltonoperator  $H$  gleich dem Operator  $T$  der kinetischen Energie

$$H = T = \frac{p^2}{2m} = -\frac{\hbar^2}{2m} \frac{d^2}{dx^2}. \quad (4.1)$$

(Im folgenden wird wieder  $\hbar = 1$  gesetzt.) Die Heisenberggleichung für  $x$  und  $p$  ist

$$\dot{x} = \frac{p}{m}, \quad \dot{p} = 0. \quad (4.2)$$

Wie beim harmonischen Oszillator stimmt sie mit der klassischen Bewegungsgleichung überein, und wir erhalten als Lösung die operatorwertige Funktion

$$x(t) = x + \frac{p}{m}t. \quad (4.3)$$

Im Gegensatz zum harmonischen Oszillator besitzt aber der Hamiltonoperator für das kräftefreie Teilchen keine Eigenvektoren in  $L^2(\mathbb{R})$ . Er besitzt stattdessen verallgemeinerte Eigenwerte  $E \geq 0$ . Zu jedem  $E > 0$  gibt es zwei verallgemeinerte Eigenwerte  $k = \pm\sqrt{2mE}$  des Impulses und auch zwei linear unabhängige uneigentliche Eigenfunktionen, nämlich die rechts- bzw. links laufenden ebenen Wellen  $e^{ikx}$ .

Der Erwartungswert des Ortes folgt genau der klassischen Bewegungsgleichung. Interessant ist das zeitliche Verhalten der Ortsunschärfe. Es gilt in einem Zustand  $\Psi \in L^2(\mathbb{R})$

$$\Delta(x(t))^2 = \langle x(t)^2 \rangle - \langle x(t) \rangle^2 = \Delta(x)^2 + 2\text{Cor}(p, x)\frac{t}{m} + \Delta(p)^2\frac{t^2}{m^2}. \quad (4.4)$$

mit der symmetrisierten Korrelation zweier Observabler

$$\text{Cor}(A, B) = \langle \frac{1}{2}(AB + BA) \rangle - \langle A \rangle \langle B \rangle. \quad (4.5)$$

Da es keine Zustände mit  $\Delta(p) = 0$  gibt, wächst die Ortsunschärfe bei der kräftefreien Bewegung für große Zeiten immer unbeschränkt an.

#### 5. Stufenpotentiale

Besonders drastisch zeigen sich die Effekte der Quantenmechanik bei Potentialen, die stückweise konstant sind. An den Stellen, an denen sich das Potential ändert, ist die Kraft unendlich groß, sodass die Heisenberggleichung für den Impuls ihren Sinn verliert. Wir betrachten zunächst den Fall einer einzigen Stufe,

$$U(x) = \begin{cases} U_0 & , \quad x > 0 \\ 0 & , \quad x \leq 0 \end{cases} \quad (5.1)$$

mit  $U_0 > 0$ .

Außerhalb des Nullpunktes ist die zeitunabhängige Schrödingergleichung gerade die Schwingungsgleichung mit Frequenz  $k(x) = \sqrt{2m(E - U(x))}$ .

Am Nullpunkt ersetzen wir die Differentialgleichung durch die Bedingung, dass die Lösung stetig differenzierbar ist. Damit erhalten wir als allgemeine Lösung

$$\psi(x) = \psi(0) \cos k(x)x + \psi'(0) \frac{\sin k(x)x}{k(x)}. \quad (5.2)$$

Als Konvention wählen wir im Fall  $E < U(x)$  die Wurzel  $k(x)$  mit positivem Imaginärteil, also

$$k(x) = \sqrt{2m(E - U(x))}, \quad E > U(x) \quad (5.3)$$

$$k(x) = i\sqrt{2m(U(x) - E)}, \quad E < U(x). \quad (5.4)$$

Weiter beachten wir den folgenden Zusammenhang zwischen den trigonometrischen und den hyperbolischen Funktionen,

$$\cos ix = \frac{1}{2}(e^{i^2x} + e^{-i^2x}) = \frac{1}{2}(e^{-x} + e^x) = \cosh x \quad (5.5)$$

$$\sin ix = \frac{1}{2i}(e^{i^2x} - e^{-i^2x}) = \frac{1}{2i}(e^{-x} - e^x) = i \sinh x. \quad (5.6)$$

Wir betrachten jetzt den Fall, dass eine Welle von links kommt und an der Potentialschwelle bei  $x = 0$  teilweise reflektiert und teilweise durchgelassen wird. Als eine Konsequenz findet man auf der rechten Halbachse nur eine Welle, die sich nach rechts ausbreitet ( $E > U_0$ ) oder exponentiell gedämpft wird ( $E < U_0$ ). Daher gilt mit  $k_+ := k(x), x > 0$

$$\psi'(0) = ik_+\psi(0). \quad (5.7)$$

Wir zerlegen nun die Wellenfunktion auf der linken Halbachse in eine Summe von Exponentialfunktionen, ( $k_- := k(x), x < 0$ )

$$\psi(x) = \psi(0) \left( \frac{1}{2} \left(1 + \frac{k_+}{k_-}\right) e^{ik_-x} + \frac{1}{2} \left(1 - \frac{k_+}{k_-}\right) e^{-ik_-x} \right) \quad (5.8)$$

Das Verhältnis der reflektierten zur einlaufenden Welle bei  $x = 0$ , der Reflexionskoeffizient, ergibt sich zu

$$R = \frac{1 - \frac{k_+}{k_-}}{1 + \frac{k_+}{k_-}} = \frac{k_- - k_+}{k_- + k_+}, \quad (5.9)$$

das Verhältnis der durchgelassenen zur einlaufenden Welle, der Transmissionskoeffizient, zu

$$T = \frac{2k_-}{k_- + k_+}. \quad (5.10)$$

Aus den gefundenen uneigentlichen Energieeigenfunktionen  $\psi_E$  bilden wir jetzt ein Wellenpaket,

$$\varphi = \int dE g(E) \psi_E, \quad (5.11)$$

mit einer glatten Funktion  $g$ , die um einen Energiewert  $E_0$  konzentriert ist. Wir unterscheiden zwei Fälle:

- (i)  $0 < E_0 < U_0$ : in diesem Fall wählen wir  $g$  mit Träger im offenen Intervall  $(0, U_0)$  mit  $g(E_0) = 1$ . Dann fällt  $\varphi$  auf der rechten Halbachse exponentiell ab. Auf der linken Halbachse gilt

$$x^2\varphi(x) = - \int dE g(E) \frac{d^2}{dk_-^2} \psi_E. \quad (5.12)$$

Mit  $k_- = \sqrt{2mE}$  folgt aus der Kettenregel

$$\frac{d}{dk_-} = \frac{dE}{dk_-} \frac{d}{dE} = \sqrt{\frac{2E}{m}} \frac{d}{dE} \quad (5.13)$$

und damit nach der Produktregel

$$\frac{d^2}{dk_-^2} = \sqrt{\frac{2E}{m}} \frac{d}{dE} \sqrt{\frac{2E}{m}} \frac{d}{dE} = \frac{2E}{m} \frac{d^2}{dE^2} + \frac{1}{m} \frac{d}{dE}. \quad (5.14)$$

Wir setzen diese Formel in die Gleichung (5.12) ein und integrieren partiell über  $E$ . Wegen der Annahmen an  $g$  verschwinden die Randterme, und wir erhalten

$$x^2\varphi(x) = \int dE \left( \frac{d^2}{dE^2} \frac{2Eg(E)}{m} - \frac{d}{dE} \frac{g(E)}{m} \right) \psi_E(x). \quad (5.15)$$

Die Lösungen  $\psi_E$  sind beschränkt durch eine Konstante, die auf dem Träger von  $g$  nicht von  $E$  abhängt. Da  $g$  als glatt vorausgesetzt war, schließen wir, dass auch  $x^2\varphi(x)$  für  $x \rightarrow -\infty$  beschränkt ist. Das aber bedeutet, dass  $\varphi(x)$  wie  $x^{-2}$  abfallen muss und daher quadratintegrabel ist.

Wir können jetzt zeigen, dass  $E_0$  ein verallgemeinerter Eigenwert von  $H$  ist. Dazu ersetzen wir  $g$  durch

$$g_n(E) = g(E_0 + n(E - E_0)) \quad (5.16)$$

und bilden die entsprechenden Wellenpakete  $\varphi_n$ . Die Folge

$$\|(H - E_0)\varphi_n\|/\|\varphi_n\| \quad (5.17)$$

konvergiert dann gegen Null (Beweis?).

Als nächstes betrachten wir die Zeitentwicklung des Wellenpakets  $\varphi$ . Es gilt mit  $k_- = k$

$$\begin{aligned} \varphi(t, x) &= \int dE g(E) e^{-itE} \psi_E(x) \\ &= \int_0^\infty dk \frac{k}{m} g\left(\frac{k^2}{2m}\right) e^{-i\frac{k^2 t}{2m}} \frac{1}{2} \left(1 + \frac{k_+}{k}\right) (e^{ikx} + R e^{-ikx}). \end{aligned} \quad (5.18)$$

Für große Werte von  $t$  und  $|x|$  oszilliert der Integrand stark mit Ausnahme der Punkte stationärer Phase, an denen die Ableitung des Exponenten nach  $k$  verschwindet. Für die einlaufende Welle ist das erfüllt für  $k = \frac{mx}{t}$ , für die reflektierte Welle bei  $k = -\frac{mx}{t}$ . Für sehr frühe Zeiten ( $t \ll 0$ ) ist  $\varphi(t, x)$  proportional zu  $g\left(\frac{mx^2}{2t^2}\right)$ , zu sehr späten Zeiten ( $t \gg 0$ )

zu  $g(\frac{mx^2}{2t^2})R$ .  $|R|^2$  ergibt also die Reflexionswahrscheinlichkeit. Im betrachteten Fall ist  $k_+$  imaginär und daher ist  $|R| = 1$ . Das Teilchen wird also in diesem Fall mit Wahrscheinlichkeit 1 reflektiert.

- (ii)  $E_0 > U_0$ . In diesem Fall ist  $k_+$  reell. Wie im ersten Fall bildet man Wellenpakete durch Integration mit einer glatten Funktion  $g$ , die jetzt Träger in einem endlichen Intervall  $I \subset (U_0, \infty)$  hat. Der Beweis, dass  $\varphi$  quadratintegrabel ist, muss jetzt auf beiden Halbachsen geführt werden, entspricht aber dem Beweis im Fall  $0 < E_0 < U_0$ . Ebenso sieht man, dass  $E_0$  ein verallgemeinerter Eigenwert ist. Die Menge der verallgemeinerten Eigenwerte ist abgeschlossen, deshalb sind auch 0 und  $U_0$  verallgemeinerte Eigenwerte.

Die zeitliche Entwicklung des Wellenpakets kann wieder mit Hilfe des Arguments der stationären Phase durchgeführt werden. Man findet, dass  $|R|^2$  die Reflexions- und  $\frac{k_+}{k_-}|T|^2$  die Transmissionswahrscheinlichkeit ist. Weiter sieht man, dass  $|R|^2 + \frac{k_+}{k_-}|T|^2 = 1$  gilt, das Teilchen wird also entweder reflektiert oder durchgelassen.

Wir beobachten zwei charakteristische Unterschiede zwischen dem quantenmechanischen und dem klassischen Verhalten des Teilchens: Auch im Fall  $E < U_0$  dringt das Teilchen in die rechte Halbachse ein, bevor es vollständig reflektiert wird. Umgekehrt ist im Fall  $E > U_0$  die Wahrscheinlichkeit, dass das Teilchen durchgelassen wird, kleiner als 1.

Wir wollen als nächstes den Tunneleffekt besprechen. Dazu betrachten wir ein Potential der Form  $U(x) = U_0 > 0$  für  $0 < x < a$  und  $U(x) = 0$  sonst. Sei  $0 < E < U_0$ . Die Lösungen der zeitunabhängigen Schrödingergleichung zur Energie  $E$  sind

$$\psi(x) = \psi(a) \cos k(x)(x - a) + \psi'(a) \frac{\sin k(x)(x - a)}{k(x)} \quad (5.19)$$

für  $x \geq 0$ . Für  $x \leq a$  ist die Lösung durch (5.2) gegeben. Offenbar gilt

$$\begin{pmatrix} \psi(0) \\ \psi'(0) \end{pmatrix} = \begin{pmatrix} \cos k(x)a & -\frac{\sin k(x)a}{k(x)} \\ k(x) \sin k(x)a & \cos k(x)a \end{pmatrix} \begin{pmatrix} \psi(a) \\ \psi'(a) \end{pmatrix}. \quad (5.20)$$

Sei  $k = \sqrt{2mE}$  und  $\kappa = \sqrt{2m(U_0 - E)}$ , also  $k(x) = i\kappa$  für  $0 < x < a$ . Wir betrachten wieder eine Welle, die von links einläuft. Daher ist für  $x > a$  nur eine auslaufende Welle vorhanden, also ist  $\psi'(a) = ik\psi(a)$ . Der Koeffizient  $A$  der einlaufenden Welle ist  $\frac{1}{2}(\psi(0) + \frac{\psi'(0)}{ik})$ . Wir erhalten

$$\begin{aligned} A &= \frac{1}{2} \left(1 + \frac{1}{ik}\right) \begin{pmatrix} \cosh \kappa a & -\frac{\sinh \kappa a}{\kappa} \\ -\kappa \sinh \kappa a & \cosh \kappa a \end{pmatrix} \begin{pmatrix} \psi(a) \\ ik\psi(a) \end{pmatrix} \\ &= \left( \cosh \kappa a + \frac{i}{2} \left( \frac{\kappa}{k} - \frac{k}{\kappa} \right) \sinh \kappa a \right) \psi(a). \end{aligned} \quad (5.21)$$

Die Tunnelwahrscheinlichkeit ergibt sich also mit  $\epsilon = \frac{1}{2} \left( \frac{\kappa}{k} - \frac{k}{\kappa} \right)$  zu

$$W = \frac{|\psi(a)|^2}{|A|^2} = (\cosh^2 \kappa a + \epsilon^2 \sinh^2 \kappa a)^{-1} = \frac{1}{1 + (1 + \epsilon^2) \sinh^2 \kappa a} . \quad (5.22)$$

Im Limes großer Werte von  $a$  verhält sich die Tunnelwahrscheinlichkeit wie

$$W \approx \frac{4}{1 + \epsilon^2} e^{-2\kappa a} . \quad (5.23)$$

Man erkennt, dass die Variation der Breite des Potentialwalls die Tunnelwahrscheinlichkeit stark beeinflusst. Mit diesem Effekt erklärt man die sehr unterschiedlichen Lebensdauern von Kernen, die durch  $\alpha$ -Strahlung zerfallen. Eine wichtige Anwendung des Tunneleffekts ist die Rastertunnelmikroskopie.

## 6. Der Bahndrehimpuls

Die Bewegung von Teilchen in 3 Dimensionen, die unter dem Einfluss eines rotationssymmetrischen Potentials stehen, kann auf die Bewegung in einer Dimension zurückgeführt werden, indem man die Rotationssymmetrie ausnutzt.

Auf eine Wellenfunktion  $\psi \in L^2(\mathbb{R}^3)$  wirken die Rotationen  $R$  in der folgenden Weise:

$$U(R)\psi(\mathbf{x}) = \psi(R^{-1}\mathbf{x}) . \quad (6.1)$$

Da die Rotationen das Volumenelement nicht ändern, ist  $U(R)$  ein unitärer Operator. Wir betrachten jetzt Drehungen mit dem Winkel  $\theta$  um eine Achse  $\mathbf{n}$ ,  $\mathbf{n} \in \mathbb{R}^3$ ,  $|\mathbf{n}| = 1$ . Diese wirken auf  $\mathbf{x}$  nach der Formel

$$R(\mathbf{n}, \theta)\mathbf{x} = (\mathbf{x} \cdot \mathbf{n})\mathbf{n} + \sin \theta \mathbf{n} \times \mathbf{x} + \cos \theta (\mathbf{x} - (\mathbf{x} \cdot \mathbf{n})\mathbf{n}) \quad (6.2)$$

Dann gilt an der Stelle  $\theta = 0$

$$\frac{d}{d\theta} U(R(\mathbf{n}, \theta))\psi(\mathbf{x}) = \frac{d}{d\theta} \psi(R(\mathbf{n}, -\theta)\mathbf{x}) = -(\mathbf{n} \times \mathbf{x}) \cdot \nabla \psi(\mathbf{x}) \quad (6.3)$$

Der Differentialoperator auf der rechten Seite ist

$$-(\mathbf{n} \times \mathbf{x}) \cdot \nabla = -\mathbf{n} \cdot (\mathbf{x} \times \nabla) = -i\mathbf{n} \cdot \mathbf{L} \quad (6.4)$$

mit dem *Bahndrehimpulsoperator*

$$\mathbf{L} = \mathbf{x} \times \mathbf{p} . \quad (6.5)$$

Man verifiziert leicht, dass die Komponenten des Bahndrehimpulses die aus der Theorie des Spins bekannten Vertauschungsrelationen erfüllen. Daher lassen sich die Eigenwerte und Eigenvektoren des Bahndrehimpulses mit der Methode aus Abschnitt 6 in Kapitel II bestimmen.

Wir führen Kugelkoordinaten  $(r, \theta, \phi)$  ein,

$$x = r \sin \theta \cos \phi, \quad y = r \sin \theta \sin \phi, \quad z = r \cos \theta, \quad r \geq 0, \quad 0 \leq \theta \leq \pi, \quad 0 \leq \phi < 2\pi . \quad (6.6)$$

Das Volumenelement in den neuen Koordinaten ist

$$dx dy dz = r^2 \sin \theta \, dr d\theta d\phi \quad (6.7)$$

Die Einheitsvektoren in den neuen Koordinatenrichtungen sind

$$\mathbf{e}_r = \begin{pmatrix} \sin \theta \cos \phi \\ \sin \theta \sin \phi \\ \cos \theta \end{pmatrix}, \quad \mathbf{e}_\theta = \begin{pmatrix} \cos \theta \cos \phi \\ \cos \theta \sin \phi \\ -\sin \theta \end{pmatrix}, \quad \mathbf{e}_\phi = \begin{pmatrix} -\sin \phi \\ \cos \phi \\ 0 \end{pmatrix} \quad (6.8)$$

Diese Vektoren bilden für jeden Wert der Winkel  $\theta$  und  $\phi$  eine Orthonormalbasis des  $\mathbb{R}^3$  und erfüllen die Gleichung

$$\mathbf{e}_r \times \mathbf{e}_\theta = \mathbf{e}_\phi, \quad (6.9)$$

und der Nabla-Operator schreibt sich in Kugelkoordinaten als

$$\nabla = \mathbf{e}_r \frac{\partial}{\partial r} + \frac{1}{r} \mathbf{e}_\theta \frac{\partial}{\partial \theta} + \frac{1}{r \sin \theta} \mathbf{e}_\phi \frac{\partial}{\partial \phi}. \quad (6.10)$$

Daraus ergibt sich der Bahndrehimpuls in Kugelkoordinaten zu

$$\mathbf{L} = \frac{1}{i} \mathbf{x} \times \nabla = \frac{1}{i} r \mathbf{e}_r \times \left( \mathbf{e}_r \frac{\partial}{\partial r} + \frac{1}{r} \mathbf{e}_\theta \frac{\partial}{\partial \theta} + \frac{1}{r \sin \theta} \mathbf{e}_\phi \frac{\partial}{\partial \phi} \right) = \frac{1}{i} \left( \mathbf{e}_\phi \frac{\partial}{\partial \theta} - \frac{1}{\sin \theta} \mathbf{e}_\theta \frac{\partial}{\partial \phi} \right). \quad (6.11)$$

Hieraus lesen wir ab

$$L_z = \frac{1}{i} \frac{\partial}{\partial \phi}, \quad L_\pm = e^{\pm i\phi} \left( i \cot \theta \frac{\partial}{\partial \phi} \pm \frac{\partial}{\partial \theta} \right). \quad (6.12)$$

Wir erkennen unmittelbar, dass die Eigenvektoren von  $L_z$  durch

$$\chi(r, \theta, \phi) = e^{im\phi} \chi(r, \theta, 0) \quad (6.13)$$

gegeben sind mit  $m \in \mathbb{Z}$ . Die halbzahligen Eigenwerte, die wir bei den Spindarstellungen beobachtet haben, treten beim Bahndrehimpuls also nicht auf. Wir suchen jetzt Höchstgewichtsvektoren, also Eigenvektoren  $\Phi$  von  $L_z$  zum Eigenwert  $l \geq 0$ , die die Gleichung

$$L_+ \Phi = 0 \quad (6.14)$$

erfüllen. Diese sind gegeben durch

$$\Phi(r, \theta, \phi) = (\sin \theta)^l e^{il\phi} g(r) \quad (6.15)$$

mit einer beliebigen Funktion  $g$ . Wir schließen, dass alle ganzzahligen Drehimpulsquantenzahlen auftreten. Durch wiederholte Anwendung von  $L_-$  gewinnen wir alle Eigenfunktionen von  $L_z$ . Wir erhalten auf diese Weise eine Orthonormalbasis  $Y_{lm}$  von Funktionen der Winkel bezüglich des Skalarproduktes mit dem Volumenelement  $\sin \theta d\theta d\phi$ . Dies lässt sich geometrisch interpretieren als eine Orthonormalbasis im Raum  $L^2(S^2)$  der quadratintegriblen Funktionen auf der Einheitskugel  $S^2$ . Die Funktionen  $Y_{lm}$  heißen Kugelflächenfunktionen (englisch: spherical harmonics).

### 7. Die radiale Schrödingergleichung

Der Operator der kinetischen Energie ist invariant unter Rotationen. In Kugelkoordinaten hat er die Form

$$T = -\frac{1}{2m} \left( \frac{1}{r^2} \frac{\partial}{\partial r} r^2 \frac{\partial}{\partial r} - \frac{|\mathbf{L}|^2}{r^2} \right). \quad (7.1)$$

Wenn das Teilchen sich in einem Potential bewegt, das nur vom Abstand von einem Kraftzentrum abhängt, dann ist der Hamiltonoperator invariant unter Drehungen um dieses Zentrum. Daher vertauscht auch der Bahndrehimpuls mit dem Hamiltonoperator. Wir können also gemeinsame Eigenfunktionen von  $H$ ,  $|L|^2$  und  $L_z$  wählen. Diese haben die Form

$$\varphi(r, \theta, \phi) = \Phi(r) Y_{lm}(\theta, \phi) \quad (7.2)$$

Ist  $\varphi$  eine Lösung der zeitunabhängigen Schrödingergleichung zum Energiewert  $E$ , so ist  $\Phi$  eine Lösung der *radialen Schrödingergleichung*

$$\left( -\frac{1}{2m} \left( \frac{d^2}{dr^2} + \frac{2}{r} \frac{d}{dr} - \frac{l(l+1)}{r^2} \right) + U(r) \right) \Phi(r) = E\Phi(r). \quad (7.3)$$

Man kann diese Gleichung noch etwas vereinfachen, indem man  $\Phi(r) = \frac{\chi(r)}{r}$  setzt. Dann ergibt sich

$$-\frac{1}{2m} \chi'' + \left( \frac{l(l+1)}{2mr^2} + U(r) \right) \chi = E\chi. \quad (7.4)$$

Diese Gleichung hat dieselbe Form wie die Schrödingergleichung für ein Teilchen in einer Dimension, das sich nur auf der rechten Halbachse bewegen kann, und auf das das *effektive Potential*

$$U_{\text{eff}}(r) = \frac{l(l+1)}{2mr^2} + U(r) \quad (7.5)$$

wirkt.

Zur Definition des Hamiltonoperators  $H$  auf einem unendlich dimensionalen Hilbertraum gehört die Angabe des Teilraums  $\mathfrak{D}(H)$ , auf dem er erklärt wird. Für den Operator der kinetischen Energie  $T$  geschieht das am einfachsten durch die folgende Bedingung an die Fouriertransformierte:

$$\mathfrak{D}(T) = \{ \varphi \in L^2(\mathbb{R}^3) \mid \int d^3\mathbf{k} |\hat{\varphi}(\mathbf{k})|^2 (|\mathbf{k}|^2 + 1)^2 < \infty \} \quad (7.6)$$

Sei  $\varphi \in \mathfrak{D}(T)$ . Dann ergibt sich der Wert von  $\varphi$  am Nullpunkt mittels der inversen Fouriertransformation zu

$$\begin{aligned} \varphi(0) &= (2\pi)^{-3/2} \int d^3\mathbf{k} \hat{\varphi}(\mathbf{k}) \\ &= (2\pi)^{-3/2} \int d^3\mathbf{k} \hat{\varphi}(\mathbf{k}) (|\mathbf{k}|^2 + 1) (|\mathbf{k}|^2 + 1)^{-1}. \end{aligned} \quad (7.7)$$

Anwendung der Cauchy-Schwarzschen Ungleichung liefert die Abschätzung

$$|\varphi(0)| \leq (2\pi)^{-3/2} \sqrt{ab} \quad (7.8)$$

mit

$$a = \int d^3\mathbf{k} |\hat{\varphi}(\mathbf{k})|^2 (|\mathbf{k}|^2 + 1)^2 = \|(2mT + \mathbf{1})\varphi\|^2 \quad (7.9)$$

und

$$b = \int d^3\mathbf{k} (|\mathbf{k}|^2 + 1)^{-2} = 4\pi \int_0^\infty dk \frac{k^2}{(k^2 + 1)^2} = 4\pi \int_0^{\pi/2} d\alpha \sin^2 \alpha = \pi^2. \quad (7.10)$$

Hierbei haben wir  $k = \tan \alpha$  substituiert.

Wir sehen, dass  $\varphi$  am Ursprung beschränkt ist; daher muss  $\chi$  bei Null verschwinden. Falls das Potential bei Null nicht zu schnell negativ wird (das anziehende Coulombpotential ist noch zulässig), gilt dieselbe Aussage auch für die Elemente des Definitionsbereichs des vollen Hamiltonoperators.

## 8. Bindungszustände im Coulombpotential

Wir betrachten effektive Potentiale der Form

$$U_{\text{eff}}(r) = -\frac{A}{r} + \frac{B}{r^2}$$

mit  $A > 0$  und  $B = \frac{l(l+1)}{2m}$ ,  $l \in \mathbb{N}_0$ . Der Bohrsche Radius  $a$  ist definiert als der Abstand  $r$ , bei dem die klassische Zentrifugalkraft für Drehimpuls 1 die Coulombkraft kompensiert,  $a = 1/mA$ .

Die radiale Schrödingergleichung lautet

$$-\frac{1}{2m}\chi''(r) + U_{\text{eff}}(r)\chi(r) = E\chi(r)$$

mit  $\chi(0) = 0$ . Für Bindungszustände sollten die Energien  $E$  negativ sein. Wir setzen  $\varepsilon = -2mE$ . Damit vereinfacht sich die Gleichung zu

$$-\chi''(r) + \frac{l(l+1)}{r^2}\chi(r) - \frac{2}{ar}\chi(r) = -\varepsilon\chi(r).$$

Wir suchen zunächst eine Näherungslösung für  $r \rightarrow \infty$ . Dazu vernachlässigen wir das Potential. Die Lösungen sind dann Linearkombinationen von

$$e^{\pm\sqrt{\varepsilon}r}.$$

Da wir normierte Lösungen suchen, berücksichtigen wir nur den exponentiell abfallenden Term.

Um das Verhalten bei  $r = 0$  zu studieren, vernachlässigen wir alle Terme außer den ersten beiden und lösen die verbleibende Gleichung durch den Ansatz

$$\chi(r) = r^s.$$

Wir finden die Bedingung

$$s(s-1) = l(l+1)$$

mit den beiden Lösungen  $s = l + 1$  und  $s = -l$ . Nur die erste Lösung ist mit der Randbedingung  $\chi(0) = 0$  verträglich. Wir machen daher für die volle Gleichung den Ansatz

$$\chi(r) = r^{l+1} H(r/a) e^{-\sqrt{\varepsilon} r} .$$

Durch Einsetzen erhalten wir die folgende Gleichung für  $H$ :

$$\frac{r}{a} H'' + 2(l+1 - \sqrt{\varepsilon} r) H' + (2 - 2(l+1)\sqrt{\varepsilon} a) H = 0 .$$

Diese Gleichung läßt sich durch einen Potenzreihenansatz lösen,  $H = \sum a_n \left(\frac{r}{a}\right)^n$ . Durch Vergleich der Koeffizienten ergibt sich die Rekursionsrelation

$$\frac{a_{n+1}}{a_n} = \frac{2\sqrt{\varepsilon} a(n+l+1) - 2}{(n+1)(n+2l+2)} .$$

Für große Werte von  $n$  verhalten sich aufeinanderfolgende Koeffizienten also wie  $\frac{2\sqrt{\varepsilon} a}{n+1}$ , die Potenzreihe wächst daher an wie  $e^{2\sqrt{\varepsilon} r}$ , sodass die Normierbarkeitsbedingung verletzt ist. Wir müssen daher verlangen, dass die Potenzreihe abbricht, d.h. dass für ein  $n \in \mathbb{N}_0$  gilt

$$\sqrt{\varepsilon} a(n+l+1) = 1 .$$

Wir erhalten daher als mögliche Energieeigenwerte

$$E_n = -\frac{1}{2ma^2(n+l+1)^2} .$$

Die zugehörigen Wellenfunktionen sind von der Form

$$\chi_n(r) = r^{l+1} H_n\left(\frac{r}{a}\right) e^{-\frac{r}{(n+l+1)a}}$$

mit Polynomen  $n$ -ter Ordnung  $H_n$ .

## 9. Mehrteilchensysteme

**9.1. Schwerpunkt- und Relativbewegung.** Zur Beschreibung von Mehrteilchensystemen verwendet man Tensorprodukte von Einteilchensystemen. Die reinen Zustände eines spinlosen Teilchens werden durch die Einheitsvektoren des Hilbertraums  $\mathfrak{H} = L^2(\mathbb{R}^3)$  beschrieben. Das Tensorprodukt von  $n$  solchen Hilberträumen ist

$$\mathfrak{H}^{\otimes n} = L^2(\mathbb{R}^{3n}) . \quad (9.1)$$

Die Schrödingergleichung nimmt dann die Form an

$$i \frac{\partial}{\partial t} \psi(t, \mathbf{x}_1, \dots, \mathbf{x}_n) = \left( - \sum_{j=1}^n \frac{1}{2m_j} \Delta_{\mathbf{x}_j} + U(\mathbf{x}_1, \dots, \mathbf{x}_n) \right) \psi(t, \mathbf{x}_1, \dots, \mathbf{x}_n) . \quad (9.2)$$

Hierbei ist  $\Delta_{\mathbf{x}_j}$  der Laplaceoperator zur Ortsvariablen  $\mathbf{x}_j$  des  $j$ -ten Teilchens. Die potentielle Energie ist im Fall von Zweiteilchenwechselwirkungen gegeben durch

$$U(\mathbf{x}_1, \dots, \mathbf{x}_n) = \sum_{1 \leq j < k \leq n} U_{jk}(\mathbf{x}_j - \mathbf{x}_k) \quad (9.3)$$

mit dem Wechselwirkungspotential  $U_{jk}$  zwischen dem  $j$ -ten und dem  $k$ -ten Teilchen. Für das Wasserstoffatom erhält man z.B.

$$U(\mathbf{x}_1, \mathbf{x}_2) = -\frac{e^2}{4\pi\epsilon_0|\mathbf{x}_1 - \mathbf{x}_2|} \quad (9.4)$$

mit der Elektronkoordinate  $\mathbf{x}_1$  und der Protonkoordinate  $\mathbf{x}_2$ .

Da das Potential in diesem Fall nur von den Differenzen der Ortskoordinaten abhängt, ist der Hamiltonoperator invariant unter Translationen des gesamten Systems. Um dies ausnutzen zu können, transformieren wir das System in Schwerpunkts- und Relativkoordinaten

$$\mathbf{X} := \frac{m_1\mathbf{x}_1 + m_2\mathbf{x}_2}{m_1 + m_2}, \quad \mathbf{x} = \mathbf{x}_1 - \mathbf{x}_2. \quad (9.5)$$

In der klassischen Physik kann man jetzt in der folgenden Weise vorgehen. Die Ortsvektoren der beiden Teilchen in den neuen Koordinaten sind

$$\mathbf{x}_1 = \mathbf{X} + \frac{m_2}{m_1 + m_2}\mathbf{x}, \quad \mathbf{x}_2 = \mathbf{X} - \frac{m_1}{m_1 + m_2}\mathbf{x}. \quad (9.6)$$

Eine entsprechende Formel gilt für die Geschwindigkeiten. Die kinetische Energie ergibt sich jetzt zu

$$T = \frac{m_1}{2}|\dot{\mathbf{x}}_1|^2 + \frac{m_2}{2}|\dot{\mathbf{x}}_2|^2 = \frac{M}{2}|\dot{\mathbf{X}}|^2 + \frac{m}{2}|\dot{\mathbf{x}}|^2 \quad (9.7)$$

mit der Gesamtmasse  $M = m_1 + m_2$  und der sogenannten reduzierten Masse

$$m = \frac{m_1 m_2}{m_1 + m_2}. \quad (9.8)$$

Definiert man den Gesamtimpuls  $\mathbf{P} = M\dot{\mathbf{X}}$  und den Relativimpuls  $\mathbf{p} = m\dot{\mathbf{x}}$ , so findet man die Formel

$$T = \frac{|\mathbf{P}|^2}{2M} + \frac{|\mathbf{p}|^2}{2m}. \quad (9.9)$$

Dieselbe Formel gilt auch in der Quantenmechanik. Hier definiert man den Gesamtimpuls und den Relativimpuls als Ableitungsoperatoren in den neuen Variablen. Die Umrechnung des Nabla-Operators in die neuen Koordinaten ergibt nach der Kettenregel

$$\nabla_{\mathbf{x}} = \frac{m_2}{m_1 + m_2}\nabla_{\mathbf{x}_1} - \frac{m_1}{m_1 + m_2}\nabla_{\mathbf{x}_2}, \quad \nabla_{\mathbf{X}} = \nabla_{\mathbf{x}_1} + \nabla_{\mathbf{x}_2}. \quad (9.10)$$

Damit folgt

$$\mathbf{P} = \frac{1}{i}\nabla_{\mathbf{X}} = \mathbf{p}_1 + \mathbf{p}_2, \quad \mathbf{p} = \frac{1}{i}\nabla_{\mathbf{x}} = \frac{m_2}{m_1 + m_2}\mathbf{p}_1 - \frac{m_1}{m_1 + m_2}\mathbf{p}_2. \quad (9.11)$$

Der Operator der kinetischen Energie nimmt dann in den neuen Koordinaten die Form an

$$T = -\frac{1}{2M}\Delta_{\mathbf{x}} - \frac{1}{2m}\Delta_{\mathbf{x}}. \quad (9.12)$$

Da die potentielle Energie nur von der Relativkoordinate abhängt, können die Schrödingergleichungen für die Schwerpunktkoordinate und die Relativkoordinate unabhängig voneinander gelöst werden. Für die Schwerpunktskoordinate erhält man die Gleichung für ein kräftefreies Teilchen der Masse  $M$ , für die Relativkoordinate die Gleichung eines Teilchens der Masse  $m$  in einem  $\mathbf{x}$ -abhängigen Potential. Im Falle des Wasserstoffatoms setzen wir also in der Formel für die Energieeigenwerte für die Masse nicht die Elektronmasse  $m_e$  ein, sondern die reduzierte Masse des Elektron-Proton-Systems

$$m = \frac{m_e m_p}{m_e + m_p}. \quad (9.13)$$

Daher gibt es kleine Unterschiede in den Energie-Niveaus für Kerne mit derselben Ladung, aber unterschiedlicher Masse (Isotopeneffekt). Aus demselben Grund gibt es im Sonnensystem kleine Abweichungen vom dritten Keplerschen Gesetz.

**9.2. Identische Teilchen; Bose- und Fermistatistik.** Falls die Teilchen identisch sind, gibt es keine Observablen, die die Nummerierung der Teilchen feststellen können. Im Falle von zwei identischen Teilchen bedeutet das, dass alle Observablen mit dem Permutationsoperator

$$U : \begin{cases} \mathfrak{H} \otimes \mathfrak{H} & \rightarrow \mathfrak{H} \otimes \mathfrak{H} \\ \Phi \otimes \Psi & \mapsto \Psi \otimes \Phi \end{cases} \quad (9.14)$$

kommutieren.  $U^2 = \mathbf{1}$  bedeutet, dass  $U$  die Eigenwerte  $\pm 1$  besitzt. Eine Observable, die mit allen Observablen vertauscht, nennt man eine Superauswahlregel. Eine Superauswahlregel kann gleichzeitig mit jeder beliebigen Observablen Eigenwerte annehmen.

In der relativistischen Quantenfeldtheorie kann man zeigen, dass Teilchen mit ganzzahligem Spin immer den Wert 1 annehmen, während Teilchen mit halbzahligem Spin immer den Wert -1 annehmen. Die Zustandsvektoren eines Mehrteilchensystems identischer Teilchen mit ganzzahligem Spin sind also symmetrisch unter Vertauschungen der Faktoren. Man spricht von Bose-Statistik und nennt solche Teilchen Bosonen. Dagegen sind die Zustandsvektoren eines Systems identischer Teilchen mit halbzahligem Spin total antisymmetrisch, d.h. sie ändern ihr Vorzeichen, sowie 2 Faktoren vertauscht werden. Dies nennt man Fermi-Statistik und nennt die Teilchen Fermionen. Beispiele für Bosonen sind Photonen, W- und Z-Bosonen, Pionen und andere Mesonen, aber auch Atome und Atomkerne, die aus einer geraden Anzahl von Fermionen bestehen. Beispiele für Fermionen sind Elektronen, Protonen,

Neutronen, Neutronen, und außerdem Atome, die aus einer ungeraden Anzahl von Fermionen bestehen.

Die wichtigste Konsequenz der Fermistatistik ist das Pauliprinzip. Dies besagt, dass Elektronen einen Einteilchenzustand nicht doppelt besetzen dürfen. Der Grund ist, dass ein total antisymmetrisiertes Tensorprodukt verschwindet, wenn die Faktoren linear abhängig sind.

Das Pauliprinzip erklärt den Aufbau der Atome. Vernachlässigt man die Coulombabstoßung der Elektronen, so ergibt sich das periodische System der Elemente dadurch, dass man die Wasserstoffeigenzustände (unter Berücksichtigung der Kernladungszahl  $Z$ ) der Reihe nach auffüllt. Sind  $\Phi_1, \dots, \Phi_Z$  die  $Z$  unteren Niveaus, so ergibt sich in dieser Näherung der Grundzustandsvektor zu

$$\Psi = \sum_{\sigma} \text{sign}(\sigma) \Phi_{\sigma(1)} \otimes \dots \otimes \Phi_{\sigma(Z)} . \quad (9.15)$$

Die Summe erstreckt sich über alle Permutationen  $\sigma$  der Zahlen  $1, \dots, Z$ . Bei der Formel ist zu berücksichtigen, dass Elektronen Spin 1/2-Teilchen sind. Daher tritt jede Lösung der Einteilchen-Schrödingergleichung für den Ort zweimal auf.

Bei Bosonen werden hingegen Mehrfachbesetzungen desselben Einteilchenzustands begünstigt. Das führt zur Existenz makroskopischer Wellen, wie z.B. bei der elektromagnetischen Strahlung. Eine spektakuläre Konsequenz ist die Bose-Einstein-Kondensation. Hier gelingt es, ein System mehrerer bosonischer Atome in einem Zustand zu präparieren, der das mehrfache Tensorprodukt eines Einteilchenzustands mit sich selbst ist.

## 10. Störungstheorie

Nur für wenige, zum Glück physikalisch wichtige, Hamiltonoperatoren kann das Eigenwertproblem explizit gelöst werden. Man hat aber gute Erfahrungen mit der folgenden Näherungsmethode gemacht. Sei

$$H = H_0 + H_1 \quad (10.1)$$

der Hamiltonoperator eines Systems. Wir nehmen an, dass wir einen Eigenvektor  $\Psi_0$  des Operators  $H_0$  kennen, mit Eigenwert  $E_0$  mit Vielfachheit 1. Wir untersuchen jetzt die interpolierenden Eigenwertgleichungen

$$(H_0 + \lambda H_1) \Psi_{\lambda} = E_{\lambda} \Psi_{\lambda} , \quad \lambda \in [0, 1] . \quad (10.2)$$

Um den frei wählbaren Faktor beim Eigenvektor festzulegen, setzen wir

$$\langle \Psi_0, \Psi_{\lambda} \rangle = 1 . \quad (10.3)$$

Wir bilden jetzt das Skalarprodukt beider Seiten der obigen Eigenwertgleichung mit  $\Psi_0$  und erhalten

$$E_0 + \lambda \langle \Psi_0, H_1 \Psi_{\lambda} \rangle = E_{\lambda} \quad (10.4)$$

Differentiation nach  $\lambda$  an der Stelle Null liefert

$$\frac{d}{d\lambda} E_\lambda|_{\lambda=0} = \langle \Psi_0, H_1 \Psi_0 \rangle . \quad (10.5)$$

In der Störungstheorie 1. Ordnung approximieren wir die Funktion  $E_\lambda$  durch ihre Tangente bei Null und erhalten als genäherten Eigenwert von  $H$  den Wert  $E_0 + \langle \Psi_0, H_1 \Psi_0 \rangle$ . Man kann jetzt auch die höheren Ableitungen von  $E_\lambda$  sowie die Ableitungen von  $\Psi_\lambda$  bei Null bestimmen und die gesuchten Eigenwerte und Eigenfunktionen durch Taylorpolynome approximieren. Dies liefert oft gute oder zumindest akzeptable Werte; in günstigen Fällen kann auch die Konvergenz des Verfahrens bewiesen werden.

Komplizierter wird die Methode, wenn der Eigenwert  $E_0$  von  $H_0$  endlich entartet ist. Sei  $\Phi_0^{(j)}, j = 1, \dots, n$  eine Orthonormalbasis des Eigenraums von  $H_0$  zum Eigenwert  $E_0$ . Wir müssen damit rechnen, dass bei Einschaltung einer Störung die Entartung aufgehoben wird. Wir suchen jetzt eine Familie von Vektoren  $\Phi_\lambda^{(j)}, j = 1, \dots, n$ , sodass der erzeugte Unterraum invariant unter  $H_0 + \lambda H_1$  ist, d.h. es soll gelten

$$(H_0 + \lambda H_1) \Phi_\lambda^{(j)} = \sum_{k=1}^n \Phi_\lambda^{(k)} E_{kj}(\lambda) \text{ mit } E_{kj} \in \mathbb{C} . \quad (10.6)$$

Wir können die Zahlen  $E_{kj}$  als die Einträge einer  $(n \times n)$ -Matrix auffassen. Ist  $a \in \mathbb{C}^n$  ein Eigenvektor dieser Matrix zum Eigenwert  $E$ , so ist

$$\Phi := \sum_{j=1}^n a_j \Phi_\lambda^{(j)} \quad (10.7)$$

ein Eigenvektor von  $H_0 + \lambda H_1$  zum selben Eigenwert. Die Gleichung (10.6) reduziert also das Eigenwertproblem für Hilbertraumoperatoren auf eines für endliche Matrizen. Als Normierungsbedingung an die Vektoren  $\Phi_\lambda^{(k)}$  wählen wir

$$\langle \Phi_0^{(j)}, \Phi_\lambda^{(k)} \rangle = \delta_{jk} . \quad (10.8)$$

Ähnlich wie im nicht entarteten Fall bilden wir das Skalarprodukt beider Seiten von (10.6) mit  $\Phi_0^{(l)}, l = 1, \dots, n$  und erhalten

$$E_0 \delta_{lj} + \lambda \langle \Phi_0^{(l)}, H_1 \Phi_\lambda^{(j)} \rangle = \sum_{k=1}^n \delta_{lk} E_{kj}(\lambda) = E_{lj}(\lambda) . \quad (10.9)$$

Differentiation nach  $\lambda$  bei Null ergibt

$$\frac{d}{d\lambda} E_{lj}(0) = \langle \Phi_0^{(l)}, H_1 \Phi_0^{(j)} \rangle . \quad (10.10)$$

Die Eigenwerte von  $H$  in der Nähe von  $E_0$  unterscheiden sich also in erster Ordnung um die Eigenwerte der  $(n \times n)$ -Matrix mit den Einträgen  $\langle \Phi_0^{(l)}, H_1 \Phi_0^{(j)} \rangle$  von dem Eigenwert  $E_0$  von  $H_0$ .