

Die Quantentheorie der einfachen Alternative

(Komplementarität und Logik II)

Von C. F. v. WEIZSÄCKER

Aus dem Philosophischen Seminar der Universität Hamburg

(Z. Naturforschg. 13 a, 245—253 [1958] ; eingegangen am 3. Februar 1958)

Die von J. v. NEUMANN gegebene Formulierung der Quantentheorie wird auf einen zweidimensionalen Zustandsraum spezialisiert. In 1. wird dieses Verfahren im Sinne der Arbeit I logisch interpretiert. Die in diesem Fall sich ergebende Mathematik der Zweierspinoren wird in 2. in ihrem Zusammenhang mit der Theorie eines drei- bzw. vierdimensionalen reellen Raumes noch einmal dargestellt. In 3. sind vier physikalisch verschiedene Beispiele erörtert: Elektronenspin, Isotopspin, Polarisierung des Lichts und Youngscher Versuch. Dieselbe abstrakte Mathematik erhält in jedem Fall eine andere physikalische Bedeutung. In 4. werden einige Grundbegriffe der v. NEUMANNschen Deutung spezialisiert: Projektionsoperatoren, SCHRÖDINGER-Gleichung, Zerlegung der Einheit und statistische Deutung. Ein für spätere Arbeiten wesentliches Resultat dieser Arbeit ist die allgemeingültige Gleichung (2, 17).

1. Alternativen in der Quantentheorie

Die Quantenmechanik beschreibt die Gesamtheit der möglichen Zustände eines physikalischen Systems als einen linearen komplexen Vektorraum. Die statistische Deutung benutzt eine Metrik, die in diesem Raum definiert ist. Der Raum hat im allgemeinen unendlich viele Dimensionen, ist also ein HILBERT-Raum. J. v. NEUMANN¹ hat wohl als erster den begrifflichen Aufbau der Quantentheorie unter dem mathematischen Aspekt, daß sie eine Theorie im HILBERT-Raum ist, zusammenhängend analysiert. Das Verständnis seiner Analyse ist vielleicht dadurch etwas erschwert worden, daß er zwei Probleme zugleich behandelt hat: die neuen physikalischen Gedanken der Quantentheorie und den mathematischen Übergang von Vektorräumen endlicher Dimensionszahl zum HILBERT-Raum. Die vorliegende Arbeit versucht einen Beitrag zur Erleichterung dieses Verständnisses zu geben, indem sie die beiden Probleme trennt und die physikalischen Gedanken der Quantentheorie am mathematisch wohl einfachsten denkbaren Beispiel des zweidimensionalen komplexen Vektorraumes erläutert.

Wir knüpfen hierzu an den Begriff der n -fachen Alternative an, der in einer vorangegangenen Arbeit über Komplementarität und Logik (weiterhin als I zitiert) entwickelt worden ist². Doch ist eine Kenntnis jener Arbeit hier nicht vorausgesetzt, und ihre ins Philosophische gehende Fragestellung wird hier nicht verfolgt. Hier soll lediglich ein einfaches Modell der NEUMANNschen Deutung der Quantentheorie

gegeben werden. Die Formeln, die sich dabei ergeben, werden ferner in einer nachfolgenden Arbeit über mehrfache Quantelung gebraucht werden.

Als n -fache Alternative wollen wir logisch eine Menge von n Aussagen a_k ($k=1, \dots, n$) bezeichnen, die den folgenden beiden Bedingungen genügen:

1. Wenn eine Aussage a_k wahr ist, sind alle übrigen a_l ($l \neq k$) falsch.
2. Wenn alle Aussagen a_l bis auf eine, a_k , falsch sind, so ist a_k wahr.

Physikalisch kann man statt dessen auch von n Zuständen eines Systems sprechen mit den Eigenschaften, daß das Vorliegen eines von ihnen (ψ_k) das Nichtvorliegen aller anderen impliziert und ebenso das Nichtvorliegen aller anderen das des einen ψ_k impliziert.

Es gibt in der Quantenmechanik n -dimensionale Zustandsräume mit beliebigen endlichen (ganzen und positiven) n . Ein Beispiel ist die Gesamtheit der Einstellungsmöglichkeiten eines Drehimpulses $j \cdot \hbar$; hier ist $n=2j+1$. n orthogonale Basisvektoren eines solchen n -dimensionalen Zustandsraumes bilden genau eine n -fache Alternative. Ist z. B. ein Drehimpuls mit $j=1$ vorgegeben und weiß man, daß die magnetische Quantenzahl m bezüglich einer vorgegebenen Richtung den Wert $+1$ hat, so ist sicher, daß sie nicht die Werte 0 oder -1 hat, weiß man umgekehrt, daß die Werte 0 und -1 im Zustand nicht vorkommen, so ist er ein Zustand mit dem scharfen Wert $+1$.

¹ J. v. NEUMANN, Mathematische Grundlagen der Quantenmechanik, Springer-Verlag, Berlin 1932.

² C. F. v. WEIZSÄCKER, Naturwiss. 42, Nr. 19/20, 548 [1955].



Im letzten Satz deuten die Worte „nicht vorkommen“ den Unterschied der Quantentheorie von der klassischen Theorie an. Nach klassischer Auffassung dürfte man sagen: „weiß man, daß m die Werte 0 und -1 nicht hat, so hat m den Wert 1“. Nach der Quantentheorie ist es möglich, daß m gar keinen bestimmten Wert hat; man muß daher die schärfere Forderung stellen, daß in der Entwicklung des Zustandes nach der durch die drei möglichen Werte von m definierten Basis die beiden Werte 0 und -1 nicht vorkommen. Das heißt die betreffenden Entwicklungskoeffizienten sollen verschwinden. Im Sinne der statistischen Deutung kann man sagen: es soll sicher sein, daß m die Werte 0 und -1 nicht hat. Nur wenn man das Wort „falsch“ in den Bedingungen 1. und 2. in diesem Sinne als „sicher falsch“ interpretiert, bleibt in der Quantentheorie der Begriff der n -fachen Alternative auf die Basis eines n -dimensionalen Zustandsraumes anwendbar.

Im Anschluß an v. NEUMANN und I werden wir im folgenden die Quantentheorie des n -dimensionalen Zustandsraumes gelegentlich als eine abgeänderte *Logik* der n -fachen Alternative bezeichnen. Da wir das philosophische Problem des Wesens der Logik hier beiseite lassen, wollen wir genau angeben, was mit dieser Terminologie in der vorliegenden Arbeit gemeint sein soll. Es handelt sich um zwei definierte Schritte:

- a) die Einführung statistischer Wahrheitswerte,
 - b) die Forderung der Allgemeingültigkeit.
- Beide Schritte sollen kurz erläutert werden.

a) Ist etwa im Falle $j=1$ der Drehimpuls gar nicht relativ zu der Achse orientiert, welche bei der Definition von m ausgezeichnet wurde, so ist die Aussage „ $m=1$ “ weder sicher wahr noch sicher falsch, sondern hat eine Wahrscheinlichkeit, sich bei Nachprüfung als wahr zu erweisen. Man kann der Aussage „ $m=1$ “ diese Wahrscheinlichkeit als „Wahrheitswert“ zuschreiben. Bekanntlich definiert aber auch die Angabe der Wahrscheinlichkeiten der drei Basiszustände $m=1, 0, -1$ den Zustand im allgemeinen noch nicht vollständig, sondern es müssen noch relative Phasen bekannt sein. Erst die drei Entwicklungskoeffizienten c_m in der Darstellung des Zustandes

$$\psi = \sum_{m=-1}^{+1} c_m \psi_m \quad (1.1)$$

definieren den Zustand ψ . In I wurde dafür der Ausdruck gebraucht, c_1, c_0, c_{-1} seien die „komple-

xen Wahrheitswerte“ der drei Aussagen „ $m=1$ “, „ $m=0$ “, „ $m=-1$ “ im Zustand ψ . Mit dieser Ausdruckweise soll hier keine Hypothese über die Möglichkeit oder Unmöglichkeit verborgener Parameter verbunden werden. Sie ist vielmehr rein deskriptiv gemeint: wenn man nicht mehr über den Zustand weiß, als die Quantentheorie angibt, so kann man bei gegebenem Zustand gewisse Aussagen zwar nicht als wahr oder falsch, wohl aber durch die ihnen zukommenden Entwicklungskoeffizienten bezüglich ihrer Verifizierbarkeit kennzeichnen.

b) Von einer Logik wird verlangt, daß sie für alle Aussagen unabhängig von deren Inhalt gilt. Nun ist in der Tat die mathematische Struktur des quantentheoretischen Zustandsraums der n -fachen Alternative eindeutig festgelegt, ohne jede Rücksicht darauf, was die n -Aussagen physikalisch bedeuten; sie müssen nur die Bedingungen 1. und 2. erfüllen. Die Quantentheorie ist demnach nicht eine Theorie über spezielle physikalische Gegenstände, sondern über alle Gegenstände, über die überhaupt Aussagen gemacht werden können, die sich zu Alternativen zusammenfassen lassen. In diesem Sinne soll die Zustandsraum-Theorie als eine Logik bezeichnet werden. Natürlich ist die Geltung dieser Logik auf den Geltungsbereich der Quantentheorie beschränkt, dessen Grenzen uns unbekannt sind. Wir erheben für eine empirisch begründete Theorie nicht den Anspruch einer Geltung a priori. Damit ist aber die mögliche Grenze der Geltung unserer „Logik“ ins Unbekannte verlegt. Uns genügt es hier, daß sie im bekannten Bereich so allgemein gilt wie die Logik: unsere positiven Kenntnisse weisen ihr keine Grenze an. Der einzige Einwand, der hiergegen heute wohl erhoben werden könnte, ist, daß die Mathematik, mit deren Hilfe wir sie definieren, selbst der klassischen Logik unterliege. Dieser Einwand ist in I besprochen und wird in den nachfolgenden Arbeiten über mehrfache Quantelung wieder aufgenommen. Hier lassen wir ihn beiseite.

Wir wenden uns nun der „einfachen Alternative“, d. h. dem Falle $n=2$ zu.

2. Die Mathematik der einfachen Alternative

Es seien zwei Aussagen a_1 und a_2 gegeben, die eine Alternative bilden. Im nächsten Abschnitt werden wir Beispiele betrachten; vorerst genüge der Hinweis, daß sie z. B. die beiden Einstellungsmöglichkeiten eines Elektronenspins, die beiden Polari-

sationsrichtungen eines Lichtquants, die beiden möglichen Ladungen eines Nukleons unterscheiden können. Es ist eine Eigentümlichkeit des Falles $n=2$, daß a_2 zur Negation von a_1 äquivalent ist und umgekehrt.

Nach der Quantentheorie kann jede der beiden Aussagen einen „komplexen Wahrheitswert“ haben, den wir mit u_1 bzw. u_2 bezeichnen wollen. Als Variable für den Index benutzen wir große lateinische Buchstaben: die Aussage a_A hat den Wahrheitswert u_A ($A=1,2$). Das komplexe Zahlenpaar u_1, u_2 fassen wir zum Vektor u zusammen, der auch Spinor (genauer: Zweierspinor) heißt. Die Gesamtheit aller Spinoren u ist ein Vektorraum. In ihm bilden die Vektoren $(1,0)$ und $(0,1)$, welche die Wahrheit von a_1 bzw. a_2 anzeigen, eine Basis. Diese Begriffe sind physikalisch sinnvoll wegen des quantentheoretischen Superpositionsprinzips: Bezeichnen die Spinoren u und v mögliche Zustände und sind α und β komplexe Zahlen, so bezeichnet auch

$$w = \alpha u + \beta v \quad (2.1)$$

einen möglichen Zustand. Dies gilt uneingeschränkt, wenn man auf Normierung verzichtet; sonst müssen α und β der Forderung

$$|\alpha|^2 + |\beta|^2 = 1 \quad (2.2)$$

unterworfen werden. Es wird sich im Lauf dieser und der folgenden Arbeiten als zweckmäßig herausstellen, die Normierung zunächst zu unterlassen und erst in einer späteren Stufe, in der sie sich als Ausdruck einer bestimmten Auffassung von selbst ergibt, einzuführen. In der jetzigen Betrachtungsstufe muß man dann alle Spinoren, die durch Multiplikation mit einer komplexen Zahl auseinander hervorgehen, demselben Zustand zuordnen.

Eine lineare Transformation

$$u'_A = s_A{}^B u_B \quad (2.3)$$

mit einer beliebigen festen Matrix $s_A{}^B$ läßt die linearen Beziehungen (2.1) invariant. Man kann die Transformation in bekannter Weise als den Übergang zur Darstellung desselben Zustands bezüglich einer neuen Basis oder als eine Abbildung des Zustandsraumes auf sich auffassen. Die Transformationen (2.3) bilden eine Gruppe, wenn die Matrix nicht singular ist. Verzichten wir auf Normierung, so brauchen wir $s_A{}^B$ keinen weiteren Einschränkungen zu unterwerfen. Eine für das folgende unwesentliche Einschränkung ist die Forderung, die Determinante von $s_A{}^B$ solle 1 sein:

$$s_1^1 s_2^2 - s_1^2 s_2^1 = 1. \quad (2.4)$$

Solche $s_A{}^B$ heißen *unimodular*. Die übliche Normierung ist definiert durch den Begriff des inneren Produkts

$$(u, v) = u_1^* v_1 + u_2^* v_2. \quad (2.5)$$

Die inneren Produkte bleiben invariant, wenn $s_A{}^B$ *unitär* ist:

$$s_1^1 = s_2^{2*}, \quad s_1^2 = -s_2^{1*}, \quad |s_1^1|^2 + |s_1^2|^2 = 1. \quad (2.6)$$

Die unimodularen Transformationen lassen die Fläche des von zwei Vektoren u, v aufgespannten Parallelogramms

$$f = u_1 v_2 - u_2 v_1 \quad (2.7)$$

invariant. Wir definieren bezüglich der durch f bestimmten „Metrik“ kontravariante Komponenten eines Spinors durch die Gleichungen

$$u^1 = -u_2, \quad u^2 = u_1. \quad (2.8)$$

Man kann dann schreiben

$$f = u^1 v_1 + u^2 v_2 = u^A v_A. \quad (2.9)$$

In diesem Sinne sollen auch die oberen Indizes der $s_A{}^B$ verstanden werden. Wir werden im folgenden über doppelte Spinorindizes nur summieren, wenn einer oben, einer unten steht.

Es ist bekannt, daß man einen Spinor durch einen reellen Dreier- bzw. Vierervektor repräsentieren kann und umgekehrt. Unter Benutzung der PAULI-Matrizen

$$\sigma^0 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad \sigma^1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma^2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma^3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \quad (2.10)$$

definieren wir zu jedem Spinor u einen vierkomponentigen reellen Vektor

$$k^\mu = u_A^* \sigma^{\mu A B} u_B, \quad (2.11)$$

explizit:

$$\begin{aligned} k^0 &= u_1^* u_1 + u_2^* u_2, \\ k^1 &= u_1^* u_2 + u_2^* u_1, \\ k^2 &= \frac{1}{i} (u_1^* u_2 - u_2^* u_1), \\ k^3 &= u_1^* u_1 - u_2^* u_2. \end{aligned} \quad (2.12)$$

Diese Beziehung ist nicht ein-eindeutig. Zwar ist ein Spinor durch zwei komplexe, also vier reelle Zahlen bestimmt. Aber zwischen den vier Zahlen k^μ besteht eine algebraische Beziehung:

$$-(k^0)^2 + (k^1)^2 + (k^2)^2 + (k^3)^2 = k^\mu k_\mu = 0. \quad (2.13)$$

Wir haben hier von der aus der Relativitätstheorie bekannten Schreibweise Gebrauch gemacht mit der

Definition

$$k_0 = -k^0, \quad k_l = k^l \quad (l = 1, 2, 3). \quad (2.14)$$

Also genügt der dreidimensionale reelle Vektor

$$\mathfrak{k} = \{k^1, k^2, k^3\}, \quad (2.15)$$

um alle vier k^μ zu bestimmen (das Vorzeichen von k^0 ist durch (2.12) willkürlich als positiv festgelegt). In der Tat definieren die k^μ den Spinor u nicht vollständig, denn wenn man u mit einem Phasenfaktor $e^{i\alpha}$ (α reell) multipliziert, ergeben sich dieselben k^μ . Da aber in der Quantentheorie ein Phasenfaktor physikalisch belanglos ist, bestimmen die k^μ gerade so viel von u als wir physikalisch wissen wollen.

Daß wir den Spinor nicht durch die drei Zahlen \mathfrak{k} , sondern durch die vier Zahlen k^μ beschreiben, bedeutet im Sprachgebrauch der Informationstheorie eine *redundante* Ausdrucksweise. Der mathematische Grund für die Wahl dieser vier Zahlen ist, daß einer linearen Transformation der Spinoren eine lineare Transformation der k^μ , aber im allgemeinen nicht der \mathfrak{k} entspricht. Die vier Matrizen σ^μ bilden nämlich eine Basis im vierdimensionalen Vektorraum der hermiteschen zweireihigen Matrizen. Daher müssen sich bei einer linearen Transformation der u die reellen k^μ (u') linear durch die k^μ (u) ausdrücken lassen. Dabei entsprechen den unimodularen Transformationen eigentliche LORENTZ-Transformationen der k^μ (ohne Raum- und Zeitspiegelung). Die unitäre Transformation läßt k^0 invariant, bedeutet also eine Drehung im dreidimensionalen \mathfrak{k} -Raum. All dies erlaubt eine anschauliche Deutung der u_A im reellen dreidimensionalen Raum. Führt man im \mathfrak{k} -Raum Polarkoordinaten um die k^3 -Achse ein, so ist

$$\begin{aligned} u_1 &= \sqrt{k^0} \cos(\vartheta/2) e^{-(i\varphi/2)} e^{(i\chi/2)}, \\ u_2 &= \sqrt{k^0} \sin(\vartheta/2) e^{(i\varphi/2)} e^{(i\chi/2)}. \end{aligned} \quad (2.16)$$

ϑ , φ und χ sind EULERSche Winkel. ϑ und φ definieren die Richtung des Vektors \mathfrak{k} , k^0 seinen Betrag. Der Phasenwinkel χ kann als eine Drehung um den Vektor \mathfrak{k} aufgefaßt werden.

Zwischen den k^μ und u_A gilt die „WEYLSche Gleichung“³

$$k_\mu \sigma^{\mu AB} u_B = 0. \quad (2.17)$$

Man kann diese Gleichung durch Einsetzen der Ausdrücke (2.11) für k^μ und explizites Ausrechnen be-

stätigen. Ihre geometrische Bedeutung besprechen wir im 4. Abschnitt. Explizit geschrieben ist (2.17) ein Gleichungspaar:

$$\begin{aligned} (k_0 + k_3) u_1 + (k_1 - i k_2) u_2 &= 0, \\ (k_1 + i k_2) u_1 + (k_0 - k_3) u_2 &= 0. \end{aligned} \quad (2.18)$$

Aus (2.17) bzw. (2.18) kann man (2.11) nur bis auf einen unbestimmt bleibenden Faktor zurückgewinnen. Zunächst ist das homogene System (2.18) nur lösbar, wenn (2.13) gilt. Unter dieser Bedingung folgt das Verhältnis u_1/u_2 , während der Absolutwert unbestimmt bleibt.

3. Physikalische Beispiele

a) Elektronenspin

Das Beispiel, in dem die soeben entwickelten Formeln ihre vertraute physikalische Bedeutung haben, ist der Spin eines Fermions. Wir beschränken uns hier auf die PAULISChe Näherung der Spintheorie. Die DIRACsche Theorie wird in den nachfolgenden Arbeiten über mehrfache Quantelung behandelt.

Ein Strahl von Atomen mit einem einzelnen Leuchtelektron wird im STERN-GERLACH-Versuch in genau zwei Strahlen aufgespalten. Man folgert daraus, daß das Elektron im Magnetfeld zwei mögliche Einstellungen hat. Wir unterscheiden sie durch die Ziffern 1 und 2. Quantentheoretisch muß der Zustand bezüglich dieser Alternative durch einen Spinor (u_1, u_2) dargestellt werden. Diesem Spinor entspricht ein Vektor \mathfrak{k} . Wir behaupten, daß die Richtung dieses Vektors identisch ist mit der Richtung des Elektronenspins im gewöhnlichen physikalischen Raum. Dies folgt sofort aus der bekannten Formel für den Operator der Wechselwirkungsenergie des Elektrons mit einem äußeren Magnetfeld

$$E = (\mathfrak{S}, \vec{\sigma}). \quad (3a.1)$$

Wir nehmen an, daß der Spin des Elektrons von seiner Bahn entkoppelt ist (nur dann hat das Reden von einer einfachen Alternative einen klaren Sinn). Dann ist seine Wellenfunktion

$$\psi(\mathfrak{r}, A) = \varphi(\mathfrak{r}) \cdot u_A \quad (3a.2)$$

und der Erwartungswert der Energie im Magnetfeld

$$\bar{E} = \sum_A \int \psi^* E \psi d\tau = (u_A^* \vec{\sigma}^{AB} u_B, \mathfrak{S}) = (\mathfrak{k}, \mathfrak{S}). \quad (3a.3)$$

Dieses vertraute Ergebnis ist eigentlich sehr wunderbar. Es ist nur möglich infolge einer „prästabili-

³ Diese Bezeichnung der Gleichung wird in der Arbeit über mehrfache Quantelung motiviert werden. Im dortigen Sinne ist (2.17) die „WEYLSche Gleichung erster Stufe.“

lierten Harmonie“ zwischen der Quantentheorie des Spins und der Geometrie des physikalischen Raumes, nämlich eben der im vorigen Abschnitt besprochenen Beziehung der Spinoren zu den dreidimensionalen reellen Vektoren. Der Spin wird ja in der Quantentheorie eigentlich in doppelter Weise eingeführt. Traditionell erscheint er als ein Drehimpuls im physikalischen Raum. Auf Grund der Spektren und des STERN-GERLACH-Versuchs behandelt man ihn jedoch als einfache Alternative. Die Quantentheorie der einfachen Alternative ist aber eindeutig festgelegt. Wäre der physikalische Raum nun z. B. vierdimensional oder wiche er sonst vom bekannten dreidimensional-euklidischen Verhalten ab, so wäre die Identifizierung beider Auffassungen vom Spin unmöglich: eine einfache Alternative würde nicht einfach einen räumlichen Vektor definieren. Hier können wir diesen Zusammenhang nur aussprechen. In den nachfolgenden Arbeiten werden wir versuchen, ihn durch den Prozeß der mehrfachen Quantelung zu erklären.

Zwei bezüglich der Metrik (2.5) orthogonale Spinoren entsprechen räumlichen Vektoren entgegengesetzter Richtung. Eine orthogonale Basis im Spinraum ist also einer Achse im Ortsraum zugeordnet. Einer unitären Transformation im Spinraum entspricht die in 2. genannte Drehung im Ortsraum. Man kann auch umgekehrt sagen: jeder Drehung im Ortsraum entspricht eine unitäre Transformation der Spinalternative.

Auf den ersten Blick überrascht es, daß der allgemeinen LORENTZ-Transformation keine unitäre Transformation im Spinraum entspricht. Zwei Spinzustände, die von einem LORENTZ-System aus gesehen orthogonal sind, können doch von einem anderen LORENTZ-System aus gesehen nicht auf einmal nicht-orthogonal sein; m. a. W., das Bezugssystem kann nicht aus zwei Zuständen, die, in einem anderen System betrachtet, die Übereinstimmungswahrscheinlichkeit Null haben, zwei Zustände mit endlicher Übereinstimmungswahrscheinlichkeit machen. Die Antwort ist, daß die Zustände im neuen System zwar natürlich quantentheoretisch orthogonal bleiben, aber räumlich nicht mehr entgegengesetzte Spinrichtungen haben.

Unter Verzicht auf die DIRACsche Theorie können wir dies nur qualitativ diskutieren. Die Spinrichtung ist als Richtung eines Drehimpulses nicht rein geometrisch, sondern durch einen Bewegungsvorgang definiert. Eine LORENTZ-Transformation kann aber

entgegengesetzte Bewegungsrichtungen in nicht-entgegengesetzte verwandeln. Am anschaulichsten ist das wohl bei Impulsrichtungen. Zwei gleiche Teilchen, die von ihrem gemeinsamen Schwerpunkt aus betrachtet in entgegengesetzter Richtung fliegen, fliegen von einem quer dazu bewegten System aus gesehen unter einem Winkel gegeneinander. Analoges gilt für Drehbewegungen um entgegengesetzte Achsen. Nun scheint freilich aus dem zu Anfang gegebenen Beweis zu folgen, daß quantentheoretisch orthogonale Spinzustände stets entgegengesetzt gerichtet sind. Aber die Formel (3 a. 1) gilt nur in Abwesenheit von elektrischen Feldern und Bahnbewegungen. Der STERN-GERLACH-Versuch zeichnet das Bezugssystem aus, in dem der Magnet ruht. In einem rasch bewegten Bezugssystem treten elektrische Felder auf und der Widerspruch verschwindet. Der Meßapparat ist dann ein solcher, der zwei nicht-entgegengesetzte Richtungen voneinander eindeutig trennt. In bezug auf diese Basis hat dann die metrische Fundamentalform eine andere Gestalt, die aus (2.5) durch die der LORENTZ-Transformation entsprechende unimodulare Transformation folgt.

b) Isotopenspin

HEISENBERG hat wohl als erster bewußten Gebrauch von der Allgemeingültigkeit der Quantentheorie der einfachen Alternative gemacht, als er den Spinformalismus auf die Alternative anwandte, ob ein Nukleon ein Proton oder ein Neutron ist. Der Isotopenspin-Raum ist das bekannteste Beispiel eines aus dieser Allgemeingültigkeit folgenden dreidimensionalen euklidischen Raumes, der mit „dem“ Raum der Physik nichts zu tun hat.

c) Polarisation des Lichts

Ist eine ebene Lichtwelle fester Frequenz und Fortschrittingsrichtung vorgegeben, so bleibt ihr bezüglich der Polarisation und Intensität noch eine dreidimensionale reelle Mannigfaltigkeit möglicher Zustände. Der allgemeinste Polarisationstyp ist ja die elliptische Polarisation. Die Ellipse, die z. B. der elektrische Vektor dabei beschreibt, hat drei reelle Parameter: etwa die Längen der beiden Achsen und den Winkel der großen Achse gegen eine gegebene Richtung in der Wellenebene. Diese drei Parameter hat STOKES bereits 1852 mit den Feldstärken durch einen Formalismus verknüpft, der essentiell der der Spinoren ist. Quantentheoretisch kann man um-

gekehrt das Auftreten der drei Parameter schon daraus ableiten, daß es Apparate (z. B. NICOLSche Prismen) gibt, die ein Lichtquant zwingen, zwischen genau zwei Polarisationsrichtungen zu wählen.

Wir wählen die Fortschrittrichtung des Lichts als positive z -Richtung und fassen die Feldstärken an einem festen Ort zu einem komplexen Vektor zusammen:

$$F_x = E_x + i H_y, \quad F_y = E_y - i H_x. \quad (3c.1)$$

Für in der x -Richtung linear polarisiertes Licht ist dann

$$F_x^{(1)} = F_0 e^{i\omega t}, \quad F_y^{(1)} = 0; \quad (3c.2)$$

für in der y -Richtung linear polarisiertes Licht ist

$$F_x^{(2)} = 0, \quad F_y^{(2)} = F_0 e^{i\omega t}. \quad (3c.3)$$

Superponieren wir beide Lösungen mit den Amplituden u_1 und u_2 , so ergibt sich die allgemeine Lösung

$$\begin{aligned} F_x &= u_1 F_0 e^{i\omega t} = F_0 k^0 \cos(\vartheta/2) e^{i(\omega t - \vartheta/2)}, \\ F_y &= u_2 F_0 e^{i\omega t} = F_0 k^0 \sin(\vartheta/2) e^{i(\omega t + \vartheta/2)}. \end{aligned} \quad (3c.4)$$

$\varphi = 0$ ist linear polarisiertes Licht; $\vartheta/2$ ist der Winkel der Polarisationssebene gegen die x -Achse. $\vartheta = \pi/2$ und $\varphi = \pm \pi/2$ sind die zwei Richtungen des zirkular polarisierten Lichtes. Im k -Raum sind diese Richtungen entgegengesetzt, also bilden auch diese beiden Richtungen eine quantentheoretische orthogonale Basis.

Auch in der Möglichkeit, in dieser Weise eine einfache Alternative in die MAXWELLSche Theorie einzubetten, steckt ein nichttrivialer Sachverhalt, über den im Zusammenhang der mehrfachen Quantelung noch einmal zu reden sein wird.

d) YOUNG'scher Versuch

Der Durchgang von Strahlung durch einen Schirm mit zwei Löchern ist das klassische Diskussionsbeispiel für einfache Alternativen in der Quantentheorie. Wir wollen hier zeigen, wie man aus der

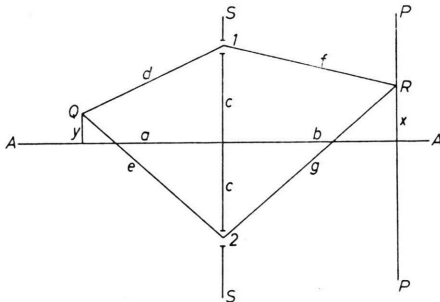


Abb. 1.

Interferenzfigur die drei reellen Parameter bestimmen kann.

Im Schirm S sind zwei um die Strecke $2c$ voneinander entfernte Löcher verschiedener Größe, 1 und 2, durch die von der Quelle Q Strahlung (z. B. Licht oder Elektronen) auf die zu S parallele photographische Platte P fällt. Q liegt in der Ebene, welche die Verbindungslinie 12 mit ihrer Mittelsenkrechten, der Figurenachse A, bildet, aber gegen A um y verschoben. Wir betrachten Strahlung, die auf einen in derselben Ebene liegenden Aufpunkt R auf P im Abstand x von A fällt. x , y und c seien klein gegen a und b . Die Durchmesser der Löcher seien nicht groß gegen die Wellenlänge der Strahlung.

Für die Abstände der Löcher von Q und R gilt streng

$$\begin{aligned} d^2 &= a^2 + (c - y)^2, & f^2 &= b^2 + (c - x)^2, \\ e^2 &= a^2 + (c + y)^2, & g^2 &= b^2 + (c + x)^2, \end{aligned} \quad (3d.1)$$

und für ihre Differenzen gilt genähert

$$L = e - d = \frac{2cy}{a}, \quad l = g - f = \frac{2cx}{b}. \quad (3d.2)$$

Im Teilchenbild kann nun ein von Q ausgegangenes und in R eingetroffenes Quant nur durch eins der beiden Löcher, aber nicht durch beide zugleich gegangen sein. Dies ist die einfache Alternative. Das Wellenbild bewertet diese beiden Aussagen mit den komplexen Zahlen u_1 und u_2 . Die Beträge von u_1 und u_2 ergeben sich dabei aus der Größe der Löcher (da $d \approx e$ angenommen ist), ihre Phasendifferenz aus der seitlichen Verschiebung y der Quelle. Die Wellenfunktion ψ sei so normiert, daß sie einen zeitlich konstanten Strom darstellt. Die mit ihr gebildeten Erwartungswerte beziehen sich dann auf die Zeiteinheit. ψ mag eine Elektronenwelle oder sein Realteil mag eine Vektorkomponente des elektromagnetischen Feldes bedeuten. Schwingt ψ im Umkreis einer Wellenlänge um Q wie

$$\psi_Q = \psi_0 e^{-i\omega t}, \quad (3d.3)$$

so erreicht es die Löcher in der Form

$$\psi_1 = (\lambda/d) \psi_0 e^{i(kd - \omega t)}, \quad \psi_2 = (\lambda/e) \psi_0 e^{i(ke - \omega t)}; \quad (3d.4)$$

$k = 2\pi/\lambda$ ist die Wellenzahl. Nach rechts strahlen die kleinen Löcher wie Punktquellen, deren Intensität durch die Lochgröße bestimmt ist. Bis auf einen gemeinsamen Phasenfaktor, der im Zeitnullpunkt untergeht, sind also die Schwingungen rechts in der Nachbarschaft der beiden Löcher

$$u_1 = s_1 e^{-(i\varphi/2) - i\omega t}, \quad u_2 = s_2 e^{(i\varphi/2) - i\omega t} \quad (3d.5)$$

$$\text{mit} \quad \varphi = kL \quad (3d.6)$$

und reellen s_1, s_2 . In R ist dann die Wellenfunktion

$$\psi = (\lambda/f) u_1 e^{ikl} + (\lambda/g) u_2 e^{ikg} \\ \approx \text{const} \cdot (s_1 e^{-i(\eta+\chi)/2} + s_2 e^{i(\eta+\chi)/2}) e^{-i\omega t} \quad (3 \text{ d. } 7)$$

$$\text{mit} \quad \chi = kl. \quad (3 \text{ d. } 8)$$

Aus (3 d. 7) ergibt sich die Interferenzfigur auf der Platte. Ein Schwärzungsmaximum tritt ein für $\varphi = -\chi$, z. B. für $x = x_0 = (b/a) y$. Ist x_0 gemessen, so kennt man also $\varphi = 2kx_0(c/b)$. Der zweite Winkel, ϑ , ist durch

$$\text{tg}(\vartheta/2) = s_2/s_1 \quad (3 \text{ d. } 9)$$

bestimmt, und das Intensitätsverhältnis zwischen Minimum und Maximum der Schwärzung ist genähert

$$\frac{I_{\min}}{I_{\max}} = \left(\frac{s_1 - s_2}{s_1 + s_2} \right)^2 = \left(\frac{\cos(\vartheta/2) - \sin(\vartheta/2)}{\cos(\vartheta/2) + \sin(\vartheta/2)} \right)^2. \quad (3 \text{ d. } 10)$$

Schließlich mißt $(k^0)^2 = s_1^2 + s_2^2$ die Gesamtintensität.

4. Lineare Operatoren

Als Abschluß illustrieren wir ein paar Begriffe des NEUMANNschen Buches an unserem zweidimensionalen Zustandsraum.

a) Projektionsoperatoren

Jeder Vektor u definiert eine Richtung im Zustandsraum, in der alle seine Vielfachen αu liegen. Die Projektion eines Vektors v auf die Richtung von u hat, sofern u auf 1 normiert ist, die Länge (u, v) und die Richtung u . Heißt der Projektionsoperator auf die Richtung von u : P_u , so ist

$$P_u v = (u, v) u \quad (4 \text{ a. } 1)$$

oder explizit

$$P_{u11} v_1 + P_{u12} v_2 = (u_1^* v_1 + u_2^* v_2) u_1, \\ P_{u21} v_1 + P_{u22} v_2 = (u_1^* v_1 + u_2^* v_2) u_2. \quad (4 \text{ a. } 2)$$

Eine Matrix P_u , die diese Gleichungen für alle v erfüllt, ist

$$P_u = \begin{pmatrix} u_1 u_1^* & u_1 u_2^* \\ u_2 u_1^* & u_2 u_2^* \end{pmatrix}, \quad (4 \text{ a. } 3)$$

$$\text{d. h.} \quad P_{uAB} = u_A u_B^*.$$

Man bestätigt leicht die Idempotenz $P_u^2 = P_u$.

Im zweidimensionalen Raum gibt es zu einer gegebenen Richtung nur eine orthogonale. Wir nennen den Projektionsoperator auf die zu u senkrechte Richtung Q_u . Jeder Vektor v ist die Summe seiner

Projektionen auf u und die zu u senkrechte Richtung:

$$P_u v + Q_u v = v \quad (4 \text{ a. } 5)$$

$$\text{oder} \quad P_u + Q_u = 1. \quad (4 \text{ a. } 6)$$

Mit der Normierungsbedingung $(u, u) = 1$ folgt

$$Q_u = - \begin{pmatrix} -u_1 u_2^* & u_1 u_1^* \\ u_2 u_1^* & -u_2 u_2^* \end{pmatrix} \quad (4 \text{ a. } 7) \\ = - \frac{1}{2} \begin{pmatrix} k_0 + k_3 & k_1 - i k_2 \\ k_1 + i k_2 & k_0 - k_3 \end{pmatrix} = - \frac{1}{2} k_\mu \sigma^\mu.$$

Die Projektion von u auf die zu u senkrechte Richtung verschwindet:

$$Q_u u = - \frac{1}{2} k_\mu \sigma^\mu u = 0. \quad (4 \text{ a. } 8)$$

Dies ist die WEYLSche Gleichung (2. 17), die also als die Lösung der Aufgabe angesehen werden kann, einen u annullierenden Projektionsoperator zu finden. Für P_u können wir nun auch schreiben

$$P_u = 1 + \frac{1}{2} k_\mu \sigma^\mu. \quad (4 \text{ a. } 9)$$

v. NEUMANN betrachtet die Projektionsoperatoren als quantenmechanische Analoga von *Aussagen*. Als idempotente Operatoren haben sie nur die Eigenwerte 0 und 1. $P_u + 1$ heißt: „ u liegt vor“ (explizit: $P_u v = v$ heißt: u und v sind derselbe Zustand; liegt also v vor, so liegt u vor); $P_u = 0$ heißt: „ u liegt vor“. Aber es ist möglich, daß P_u nicht in Diagonalform ist. Dies ist die „Mehrwertigkeit“ der quantentheoretischen Logik. Die Übertragung auf die Projektion nicht auf einzelne Vektoren, sondern auf Unterräume entspricht der Einführung des logischen „oder“: irgendeiner der im Unterraum liegenden Zustände liegt vor.

b) SCHRÖDINGER-Gleichung

Wir lassen den Zustand von der Zeit abhängen. Von einer individuellen „Bahn“ eines Zustands im Zustandsraum unterscheidet sich ein allgemeines Gesetz zeitlicher Veränderung: es ist durch eine zeitabhängige Abbildung des ganzen Zustandsraumes in sich bestimmt. Sie ist durch einen Operator $T(t)$ beschrieben. Man nennt

$$-i \dot{T} = H \quad (4 \text{ b. } 1)$$

den HAMILTON-Operator. Aus

$$u(t + dt) = T(dt) u(t) = (1 + \dot{T} dt) u(t) \quad (4 \text{ b. } 2) \\ = u(t) + \dot{u}(t) dt$$

folgt die SCHRÖDINGER-Gleichung

$$\dot{u} = i H u. \quad (4 \text{ b. } 3)$$

In dieser Allgemeinheit ist die SCHRÖDINGER-Gleichung also lediglich der Ausdruck der Forderung, daß sich der Zustand nach einem allgemeinen Gesetz stetig ändere; wir sagen, sie sei Ausdruck der Kausalität.

Man pflegt zu verlangen, daß T unitär, also H hermitesch sei. Nur dann bleibt die Metrik des Zustandsraumes zeitlich bewahrt. Nicht nötig ist dafür, daß T bzw. H von u unabhängig, also die SCHRÖDINGER-Gleichung linear sei. Die allgemeinste SCHRÖDINGER-Gleichung mit hermiteschem H können wir in unserem Fall sofort angeben. Die σ -Matrizen sind eine Basis der hermiteschen Matrizen, also ist

$$H = a_\mu \sigma^\mu \quad (4b. 4)$$

mit beliebigen reellen a_μ die allgemeinste hermitesche HAMILTON-Funktion.

Sind die a_μ von u und t unabhängig, so wird die SCHRÖDINGER-Gleichung durch den Ansatz

$$u = u_0 e^{i\lambda t} \quad (4b. 5)$$

gelöst, die zu dem Eigenwertproblem

$$H u_0 = \lambda u_0 \quad (4b. 6)$$

führt. Dieser Ansatz mag übrigens vielleicht eine Art „darwinistischer“ Rechtfertigung der Forderung geben, daß H hermitesch sei. Auch für nichthermitesche H wäre der Ansatz möglich, aber λ wäre nicht reell und der Zustand wäre nicht „lebenstüchtig“, sondern würde durch Riesen- oder Zwergwuchs verlorengelangen.

Mit konstanten a_μ ist H im Beispiel des Elektronenspins genau die Energie in einem Magnetfeld, vermehrt um eine beliebige additive Konstante a_0 . Dies ist also schon das allgemeinste mögliche „äußere Feld“ für eine einfache Alternative.

Wir können das Eigenwertproblem sofort explizit lösen. Wir schreiben die vier a_μ wie Quaternionen $\{a_0, \mathfrak{a}\}$, wobei der Betrag des Vektors \mathfrak{a} mit a bezeichnet werde. Nun hat das homogene lineare Gleichungssystem

$$(H - \lambda) u = 0 \quad (4b. 7)$$

(wir schreiben wieder u statt u_0) die Lösbarkeitsbedingung

$$|H - \lambda I| = \lambda^2 - S\lambda + D = 0. \quad (4b. 8)$$

S und D sind Spur und Determinante von H . Es ist

$$S = 2 a_0, \quad D = a_0^2 - a^2 \quad (4b. 9)$$

und folglich

$$\lambda_{\pm} = a_0 \pm a. \quad (4b. 10)$$

$$\text{Es folgt} \quad H - \lambda I = b_\mu \sigma^\mu \quad (4b. 11)$$

$$\text{mit} \quad b = \{\mp a, \mathfrak{a}\}. \quad (4b. 12)$$

Die zeitunabhängige SCHRÖDINGER-Gleichung hat die Gestalt

$$(H - \lambda) u = b_\mu \sigma^\mu u = 0, \quad (4b. 13)$$

die an die WEYLSche Gleichung (2. 17) erinnert. Sie ist in der Tat mit ihr identisch bis auf einen konstanten Faktor; denn bis auf einen solchen Faktor ist der u annullierende Operator Q_u eindeutig bestimmt. Da $a > 0$, $k_0 < 0$, ist mit irgendeinem bestimmten $\varepsilon > 0$

$$a = -\varepsilon k_0. \quad (4b. 14)$$

Also ist

$$\begin{aligned} H - \lambda_+ &= H - a_0 - a = +\varepsilon k_\mu^+ \sigma^\mu = b_\mu \sigma^\mu, \\ H - \lambda_- &= H - a_0 + a = -\varepsilon k_\mu^- \sigma^\mu = b_\mu \sigma^\mu. \end{aligned} \quad (4b. 15)$$

Für λ_+ ist somit $\mathfrak{a} = \varepsilon \mathfrak{f}$, für λ_- aber $\mathfrak{a} = -\varepsilon \mathfrak{f}$. Hier ist \mathfrak{a} ein fest gegebener Vektor. Dies bedeutet nichts anderes als daß \mathfrak{f}^+ entgegengesetzt gleich \mathfrak{f}^- und u_+ orthogonal auf u_- ist.

c) Zerlegung der Einheit

Diese, von NEUMANN in Abschnitt II, 7 seines Buches eingeführte Methode ist für endliche Dimensionszahl dem Eigenwertproblem äquivalent. Sein Ziel ist, solche mathematischen Größen einzuführen, die durch das Eigenwertproblem eindeutig bestimmt sind. Die Eigenwerte sind eindeutig, nicht aber die Eigenvektoren, die stets einen skalaren Faktor und bei Entartung sogar eine lineare Transformation zulassen. Eindeutig sind hingegen wieder die linearen Unterräume des Zustandsraums, in denen die zu den jeweiligen Eigenwerten gehörigen Eigenvektoren liegen. Diese linearen Unterräume charakterisiert man durch die Projektionsoperatoren, die auf sie projizieren. v. NEUMANN ordnet die Eigenwerte nach ihrer Größe. Jedem Eigenwert λ_τ entspricht der auf seinen Eigen-Unterraum projizierende Projektionsoperator F_τ . Dann wird eine Operatorfunktion $E(\lambda)$ der reellen Zahl λ definiert als die Summe aller der F_τ , deren Eigenwert λ_τ kleiner oder gleich λ ist. Ist λ kleiner als der kleinste Eigenwert, so setzt man $E(\lambda) = 0$; ist λ größer als der größte Eigenwert, so ist der Raum, auf den projiziert wird, der ganze Zustandsraum, also $E(\lambda) = 1$. $E(\lambda)$ heißt eine Zerlegung der Einheit.

Im zweidimensionalen Zustandsraum können wir den Fall der Entartung ausschließen, denn er würde

bedeuten, daß der ganze Zustandsraum aus Eigenvektoren von H bestünde, also ein Vielfaches der Einheitsmatrix wäre. Wir setzen also $\lambda_- < \lambda_+$. Zum Eigenvektor u_- gehört der Projektionsoperator P_- , zu u_+ gehört P_+ . Es ist

$$P_- + P_+ = 1, \quad (4c.1)$$

und die Zerlegung der Einheit ist

$$E(\lambda) = \begin{cases} 0 & \text{für } \lambda < \lambda_-, \\ P_- & \text{für } \lambda_- \leq \lambda < \lambda_+, \\ 1 & \text{für } \lambda_+ \leq \lambda. \end{cases} \quad (4c.2)$$

v. NEUMANN'S Darstellung des HAMILTON-Operators durch ein STIELTJES-Integral wird einfach

$$H = \int_{-\infty}^{+\infty} \lambda dE(\lambda) = \lambda_- P_- + \lambda_+ P_+. \quad (4c.3)$$

Wählt man u_- und u_+ als Basis, so wird

$$P_- = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \quad P_+ = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}, \quad H = \begin{pmatrix} \lambda_- & 0 \\ 0 & \lambda_+ \end{pmatrix}. \quad (4c.4)$$

d) Statistische Deutung

Als Grundlagen der statistischen Deutung wählt v. NEUMANN die Aussagen W. (S. 104) oder E. (S. 157). Mit angepaßten Bezeichnungen lautet:

W. Die Wahrscheinlichkeit w dafür, daß im Zustande v die Größe R einen Wert aus dem Intervall I annimmt, ist $\|E(I)v\|$. Dabei ist definiert $\|u\|$ als Länge des Vektors u und

$$E(I) = E(\lambda'') - E(\lambda'), \quad \text{wenn } I: \lambda' < \lambda < \lambda''. \quad (4d.1)$$

In unserem Falle ist

$$E(I) = \begin{cases} 0, & \text{wenn kein } \lambda_{\pm} \text{ in } I, \\ P_-, & \text{wenn nur } \lambda_- \text{ in } I, \\ P_+, & \text{wenn nur } \lambda_+ \text{ in } I, \\ 1, & \text{wenn beide } \lambda_{\pm} \text{ in } I. \end{cases} \quad (4d.2)$$

Bezeichnet in den beiden mittleren Fällen u den Eigenvektor des in I enthaltenen Eigenwerts der Energie und ist R die Energie, so folgt die „Übereinstimmungswahrscheinlichkeit von u und v “:

$$w = \|P_u v\| = (P_u v, P_u v) = (P_u v, v) = (v u) (u v) = |(u, v)|^2. \quad (4d.3)$$

Dasselbe folgt aus

E. Der Erwartungswert einer Größe R im Zustand v ist

$$\text{Erw}(R, v) = (R v, v) = \text{Spur}(P_v R). \quad (4d.4)$$

Hier muß man für R die „Eigenschaft des Zustandes, u zu sein“, d. h. P_u einsetzen und erhält

$$\text{Erw}(P_u, v) = (P_u v, v). \quad (4d.5)$$

Schreiben wir für einen beliebigen hermiteschen Operator

$$R = a_{\mu} \sigma^{\mu}, \quad (4d.6)$$

$$\text{so folgt} \quad (R v, v) = a_{\mu} l^{\mu} = a_0 + (a \mathbf{l}), \quad (4d.7)$$

wobei l^{μ} der v zugeordnete Vierervektor ist. Setzt man für R nach (4a.9) P_u ein, so folgt

$$w = \frac{1}{2} (1 + \mathbf{f} \mathbf{l}) = \cos^2(\Theta/2), \quad (4d.8)$$

wobei Θ der Winkel zwischen \mathbf{f} und \mathbf{l} ist.

In einem Gemenge ist nur mit Wahrscheinlichkeit bekannt, welcher Zustand vorliegt. Habe der Eigenvektor u_- die Wahrscheinlichkeit w_- und u_+ habe w_+ , mit

$$w_+ + w_- = 1, \quad (4d.9)$$

so definieren wir

$$W = w_+ - w_-. \quad (4d.10)$$

Im Sinne der Statistik sind Erwartungswerte in einander ausschließenden Zuständen additiv. Also ist

$$\begin{aligned} \text{Erw}(R) &= w_+ (R u_+, u_+) + w_- (R u_-, u_-) \\ &= \text{Spur}(U R) \end{aligned} \quad (4d.11)$$

$$\text{mit} \quad U = w_+ P_+ + w_- P_-. \quad (4d.12)$$

Wir definieren einen Operator N durch

$$N = P_+ - P_- = 1 + k_{\mu} \sigma^{\mu}. \quad (4d.13)$$

$$\text{Es folgt} \quad U = \frac{1}{2} (1 + w Q) \quad (4d.14)$$

$$\text{und} \quad \text{Erw}(R) = a_0 + W(a, \mathbf{f}). \quad (4d.15)$$