
Eugen Wigner

Gruppentheorie und ihre Anwendung auf die Quantenmechanik der Atomspektren

Springer Fachmedien Wiesbaden GmbH

Eugen Wigner

**Gruppentheorie und ihre
Anwendung auf die Quantenmechanik
der Atomspektren**

Gruppentheorie und ihre Anwendung auf die Quanten- mechanik der Atomspektren

Von

Eugen Wigner

Mit 12 Abbildungen

Springer Fachmedien Wiesbaden GmbH

1977

Alle Rechte vorbehalten

Unveränderter Nachdruck der 1931 bei Springer Fachmedien Wiesbaden

Ursprünglich erschienen bei Friedr. Vieweg & Sohn Akt.-Ges., 1931

Braunschweig, als Band 85 der Sammlung „Die Wissenschaft“ erschienenen Ausgabe.

Die Vervielfältigung und Übertragung einzelner Textabschnitte, Zeichnungen oder Bilder, auch für Zwecke der Unterrichtsgestaltung, gestattet das Urheberrecht nur, wenn sie mit dem Verlag vorher vereinbart wurden. Im Einzelfall muß über die Zahlung einer Gebühr für die Nutzung fremden geistigen Eigentums entschieden werden. Das gilt für die Vervielfältigung durch alle Verfahren einschließlich Speicherung und jede Übertragung auf Papier, Transparente, Filme, Bänder, Platten und andere Medien.

Buchbinder: Junghans, Darmstadt

ISBN 978-3-663-00642-8 ISBN 978-3-663-02555-9 (eBook)

DOI 10.1007/978-3-663-02555-9

Vorwort

Dieses Buch ist aus dem Wunsche entstanden, die Anwendung gruppentheoretischer Methoden in der Quantenmechanik einem weiteren Leserkreis zugänglich zu machen. Die wirkliche Lösung der quantenmechanischen Differentialgleichungen stößt im allgemeinen auf so große Schwierigkeiten, daß man durch direkte Rechnung zumeist nur eine grobe Annäherung zu erreichen vermag. Um so erfreulicher ist es, daß ein so großer Teil der quantenmechanischen Resultate schon durch reine Symmetrievergleichungen erhalten werden kann.

Man hat gegen die gruppentheoretische Behandlung der Schrödinger-Gleichung oft den Einwand erhoben, daß sie „nicht physikalisch“ sei. Es scheint mir aber, daß die bewußte Ausnutzung elementarer Symmetrieeigenschaften dem physikalischen Gefühl eher entsprechen muß, als die mehr rechnerische Behandlung. Der erwähnte Einwand dürfte darauf zurückzuführen sein, daß die Gruppentheorie einen wesentlich anderen Charakter hat, als die dem Physiker hauptsächlich geläufigen Teile der Mathematik, so daß es immerhin einige Zeit erfordert, bis man sich mit ihr befreundet hat. Sie besteht nämlich aus einer großen Zahl unscheinbarer Schlußfolgerungen, die zwar einzeln betrachtet trivial, in ihrer Gesamtheit aber doch nicht so leicht zu überblicken sind.

Deshalb hielt ich es schon im Interesse des ungeübten Lesers für ratsam, alle Zwischenüberlegungen nach Möglichkeit explizite auszuführen. Um das richtige „sich zu Hause fühlen“ im Gegenstand zu erleichtern, wurde vom mathematischen Teil wesentlich mehr gebracht, als für die physikalischen

Anwendungen notwendig gewesen wäre. In den physikalischen Anwendungen wollte ich mich auf einen Gegenstand beschränken und habe als diesen die Theorie der Atomspektren gewählt. So wurde der Inhalt zweier Arbeiten des Verfassers (Zeitschr. f. Phys. **40**, 883, 1926 und **43**, 624, 1927) aufgenommen, die im Anschluß an die Heisenberg-Diracschen Untersuchungen über Resonanz entstanden. Die Methode wurde bald von F. Hund (Zeitschr. f. Phys. **43**, 788, 1927), W. Heitler (ebenda **46**, 47, 1927), F. London (ebenda **46**, 455, 1927) und H. Weyl (Vorlesungen in Zürich, Wintersemester 1927/28) aufgegriffen und angewandt. Weiterhin wurde nur noch die gemeinsame Arbeit von J. v. Neumann und dem Verfasser und die neuere Slatersche Arbeit ausführlich behandelt. Dagegen konnten wegen des Umfanges die physikalisch besonders wichtigen Untersuchungen von Hund, London und Heitler über Moleküle nicht aufgenommen werden, obwohl sie die Anwendbarkeit der gruppentheoretischen Methoden auch auf andere Gebiete lebhaft bezeugen. Es schien mir aber wichtiger, ein Gebiet möglichst vollständig zu behandeln, als der verlockenden Idee nachzugeben, aus allen Gebieten die schönsten Resultate zusammenzustellen.

Zum Schluß ist es mir ein Bedürfnis, den Herren R. Becker, W. Bloch, E. Brody, H. Kallmann und W. Westphal herzlichst zu danken für viele wertvolle Ratschläge, Durchsicht von Teilen des Manuskripts und der Korrektur.

Berlin, im Februar 1931

Eugen Wigner

Inhaltsverzeichnis

	Seite
I. Vektoren und Matrizen	1
Lineare Abhangigkeit von Vektoren	11
II. Verallgemeinerungen	13
III. Hauptachsentransformation	22
Spezielle Matrizen	25
Unitare Matrizen und das skalare Produkt	27
Hauptachsentransformation unitarer und hermiteischer Ma- trizen	29
Reelle orthogonale und symmetrische Matrizen	33
IV. Grundlagen der Quantenmechanik	34
V. Storungstheorie	44
VI. Transformationstheorie und Grundlinien der statisti- schen Deutung der Quantenmechanik	51
VII. Abstrakte Gruppentheorie	63
Satze uber endliche Gruppen	65
Beispiele von Gruppen	67
Konjugierte Elemente und Klassen	72
VIII. Normalteiler	73
Holomorphic und Isomorphie	76
IX. Allgemeine Darstellungstheorie	79
X. Kontinuierliche Gruppen	97
XI. Darstellungen und Eigenfunktionen	110
XII. Algebra der Darstellungstheorie	120
XIII. Die symmetrische Gruppe	133
XIV. Die Drehgruppen	152
XV. Die Darstellungen der dreidimensionalen reinen Dreh- gruppe	164
Kugelfunktionen	164
Isomorphie der zweidimensionalen unitaren Gruppe zur Dreh- gruppe	168
Die Darstellungen der unitaren Gruppe	173
Die Darstellungen der dreidimensionalen reinen Drehgruppe	179
XVI. Die Darstellungen des direkten Produktes	184
XVII. Die Grundzuge der Atomspektren	190
Das Vektoradditionsmodell	198

	Seite
XVIII. Auswahlregeln und die Aufspaltung der Spektrallinien	209
XIX. Teilweise Bestimmung der Eigenfunktionen aus ihren Transformationseigenschaften	226
XX. Das Drehelektron	236
Die physikalischen Grundlagen der Paulischen Theorie	236
Die Invarianz der Beschreibung räumlichen Drehungen gegenüber	240
Zusammenhang mit der Darstellungstheorie	244
XXI. Die Gesamtquantenzahl	254
XXII. Die Feinstruktur der Spektrallinien	270
XXIII. Auswahl- und Intensitätsregeln bei Mitberücksichtigung des Spins	286
Die Hönl-Kronigschen Intensitätsformeln	297
Die Landésche <i>g</i> -Formel	300
Die Intervallregel	302
XXIV. Das Aufbauprinzip	306
Formelsammlung	325
Sachregister	328

Anmerkung

In diesem Buche wurde die Schrödinger-Gleichung in der Gestalt

$$\frac{h}{2\pi i} \frac{\partial \varphi}{\partial t} = \mathbf{H} \varphi$$

angenommen, während sonst

$$-\frac{h}{2\pi i} \frac{\partial \varphi}{\partial t} = \mathbf{H} \varphi$$

üblich ist. Dieses Umkehren des Vorzeichens der imaginären Einheit hat zur Wirkung, daß auch der Operator des linearen Impulses $-\frac{h}{2\pi i} \frac{\partial}{\partial x}$ anstatt $\frac{h}{2\pi i} \frac{\partial}{\partial x}$ wird.

I. Vektoren und Matrizen

Den Inbegriff von n Zahlen v_1, v_2, \dots, v_n nennt man bekanntlich einen n -dimensionalen Vektor v , oder einen Vektor im n -dimensionalen Raum, die n Zahlen selber v_1, v_2, \dots, v_n sind die Komponenten dieses Vektors. Man kann auch die Koordinaten eines Punktes im n -dimensionalen Raum als einen Vektor auffassen, denjenigen, der den Nullpunkt des Achsenkreuzes mit dem betrachteten Punkte verbindet. Die Vektoren als solche werden nach Möglichkeit mit gotischen Buchstaben bezeichnet. Zur Kennzeichnung der Komponenten des Vektors fügt man einen (lateinischen) Index an, den Namen der betreffenden Koordinatenachse. So ist v_k eine Zahl, v ein Vektor, eine Gesamtheit von n Zahlen.

Zwei Vektoren sind gleich bzw. benachbart (nahezu gleich), wenn ihre entsprechenden (d. h. sich auf dieselbe Koordinatenachse beziehenden) Komponenten gleich bzw. benachbart sind.

$$v = w \quad (1)$$

ist mit den n Gleichungen

$$v_1 = w_1; \quad v_2 = w_2; \quad \dots, \quad v_n = w_n$$

äquivalent. Ein Vektor ist Null, wenn alle seine Komponenten verschwinden. Das Produkt $c v$ einer Zahl c mit einem Vektor v ist ein Vektor, dessen Komponenten die c -fachen der Komponenten von v sind: $(cv)_k = cv_k$.

Man definiert eine Addition für Vektoren, indem man festsetzt, daß die Komponenten der Summe gleich der Summe der entsprechenden Komponenten seien, in Formel

$$(v + w)_k = v_k + w_k. \quad (2)$$

In sehr vielen mathematischen Problemen ist es zweckmäßig, an Stelle der zuerst verwendeten Variablen neue Variable ein-

zuführen. Im einfachsten Falle sind die neuen Variablen x'_1, x'_2, \dots, x'_n lineare Funktionen der alten x_1, x_2, \dots, x_n , etwa

$$\left. \begin{aligned} x'_1 &= a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n \\ x'_2 &= a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n \\ &\vdots \\ x'_n &= a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n \end{aligned} \right\} \quad (3)$$

oder

$$x'_i = \sum_{k=1}^n a_{ik} x_k.$$

Diese Art von Einführung neuer Variablen bezeichnet man als eine lineare Transformation. Sie ist vollkommen durch die Koeffizienten a_{11}, \dots, a_{nn} bestimmt, den Inbegriff dieser n^2 Zahlen — in ein quadratisches Schema angeordnet — nennt man die Matrix der linearen Transformation (3):

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix}, \quad (4)$$

wofür wir auch kürzer (a_{ik}) oder auch α schreiben wollen. Damit (3) wirklich eine Einführung neuer Variablen sei, ist es noch notwendig, daß man nicht nur die x' durch die x , sondern auch umgekehrt die x durch die x' ausdrücken kann. Mit anderen Worten, wenn wir (3) als die n Gleichungen zur Bestimmung der n Unbekannten x_1, \dots, x_n auffassen, wobei jetzt die x' als bekannt anzusehen sind, so müssen diese Gleichungen eindeutig auflösbar sein. Hierzu ist es notwendig und hinreichend, daß die aus den Koeffizienten a_{ik} gebildete Determinante

$$\begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix} \neq 0 \quad (4a)$$

nicht verschwinden soll. In diesem Falle spricht man von einer eigentlichen Transformation. Bei dem Begriff der Matrix wollen wir diese Bedingung jedoch nicht hinzunehmen und das Koeffizientenschema (4) immer als eine Matrix bezeichnen, ganz unabhängig davon, ob seine Determinante verschwindet oder nicht. Eine eigentliche Transformation wird aber dann nur von einer Matrix α

mit nicht verschwindender Determinante $|\alpha| \neq 0$ induziert. Für Matrizen wollen wir fette Buchstaben verwenden, zur Kennzeichnung der Koeffizienten α_{ik} fügen wir noch die entsprechenden Achsen (nicht mehr fettgedruckt) als Indizes zu. So ist α eine Matrix, eine Gesamtheit von n^2 Zahlen, α_{ik} eine Zahl.

Zwei Matrizen sind gleich, wenn ihre entsprechenden Koeffizienten gleich sind.

$$\alpha = \beta \quad (5)$$

ist mit den n^2 Gleichungen

$$\alpha_{ik} = \beta_{ik} \quad (i = 1, 2, \dots, n; k = 1, 2, \dots, n)$$

äquivalent.

Man kann der Gleichung

$$\xi'_i = \sum_{k=1}^n \alpha_{ik} \xi_k \quad (3a)$$

noch eine andere Bedeutung geben, indem man die ξ'_i nicht als die Komponenten desselben Vektors in einem anderen Achsenkreuz auffaßt, sondern als die Komponenten eines neuen Vektors in demselben Achsenkreuz. Man sagt, die Matrix α führt den Vektor ξ in den Vektor ξ' über, oder auch α angewendet auf ξ gibt ξ' :

$$\xi' = \alpha \xi \quad (3b)$$

gleichbedeutend mit (3a). (Eigentlich ist dieser Sprachgebrauch vom Standpunkt der Bezeichnung der logischere.)

Die n -dimensionale Matrix ist ein linearer Operator für n -dimensionale Vektoren. Ein Operator deshalb, weil sie einen Vektor in einen anderen Vektor überführt; linear, weil für beliebige Zahlen a und b

$$\alpha(a\xi + b\eta) = a\alpha\xi + b\alpha\eta \quad (6)$$

gilt. Die beiden Vektoren auf der linken und auf der rechten Seite von (6) sind nämlich einander gleich; die k -Komponente von $a\xi + b\eta$ ist $a\xi_k + b\eta_k$, die i -Komponente des Vektors der linken Seite von (6) ist also

$$\sum_{k=1}^n \alpha_{ik} (a\xi_k + b\eta_k).$$

Die i -Komponente der rechten Seite ist ebenfalls

$$a \sum_{k=1}^n \alpha_{ik} \xi_k + b \sum_{k=1}^n \alpha_{ik} \eta_k = \sum_{k=1}^n \alpha_{ik} (a\xi_k + b\eta_k).$$

Die n -dimensionale Matrix ist sogar der allgemeinste lineare Operator im n -dimensionalen Vektorraum, d. h. jeder lineare Operator in diesem Raum ist einer Matrix äquivalent. In der Tat: führt der beliebige lineare Operator O den Vektor $e_1 = (1, 0, 0, \dots, 0)$ in den Vektor $\xi_{.1}$, den Vektor $e_2 = (0, 1, 0, \dots, 0)$ in $\xi_{.2}$ usw., schließlich den Vektor $e_n = (0, 0, 0, \dots, 1)$ in $\xi_{.n}$ über, wobei die Komponenten von $\xi_{.i}$ der Reihe nach $\xi_{1i}, \xi_{2i}, \dots, \xi_{ni}$ seien, so ist O der Matrix (ξ_{ik}) äquivalent. Die Matrix (ξ_{ik}) führt nämlich zunächst jeden der n Vektoren e_1, e_2, \dots, e_n in dieselben Vektoren $\xi_{.1}, \xi_{.2}, \dots, \xi_{.n}$ über, in die sie O überführt. Ein beliebiger Vektor a ist aber eine Linearkombination der Vektoren e_1, e_2, \dots, e_n , nämlich $a = a_1 e_1 + a_2 e_2 + \dots + a_n e_n$. Da O und (ξ_{ik}) linear sind, wird also a durch beide in $a_1 \xi_{.1} + a_2 \xi_{.2} + \dots + a_n \xi_{.n}$ überführt, die Matrix (ξ_{ik}) ist O äquivalent.

Die linearen Transformationen haben eine sehr wichtige Eigenschaft, der sie ihre Bedeutung verdanken: ihre Zusammensetzbarkeit. Wenn wir nach (3) an Stelle der ursprünglichen Variablen x durch eine lineare Transformation die x' eingeführt haben und an Stelle dieser dann durch eine weitere lineare Transformation

$$\left. \begin{aligned} x'_1 &= \beta_{11} x'_1 + \beta_{12} x'_2 + \dots + \beta_{1n} x'_n \\ x'_2 &= \beta_{21} x'_1 + \beta_{22} x'_2 + \dots + \beta_{2n} x'_n \\ &\dots \\ x'_n &= \beta_{n1} x'_1 + \beta_{n2} x'_2 + \dots + \beta_{nn} x'_n \end{aligned} \right\} \quad (7)$$

die Variablen x'' einführen, so können wir diese beiden Schritte zu einem einzigen vereinigen und an Stelle der x sofort durch eine lineare Transformation die x'' einführen. Es ist wegen (7) und (3)

$$\left. \begin{aligned} x''_1 &= \beta_{11} (\alpha_{11} x_1 + \dots + \alpha_{1n} x_n) + \dots + \beta_{1n} (\alpha_{n1} x_1 + \dots + \alpha_{nn} x_n) \\ x''_2 &= \beta_{21} (\alpha_{11} x_1 + \dots + \alpha_{1n} x_n) + \dots + \beta_{2n} (\alpha_{n1} x_1 + \dots + \alpha_{nn} x_n) \\ &\dots \\ x''_n &= \beta_{n1} (\alpha_{11} x_1 + \dots + \alpha_{1n} x_n) + \dots + \beta_{nn} (\alpha_{n1} x_1 + \dots + \alpha_{nn} x_n) \end{aligned} \right\} \quad (8)$$

also sind die x'' lineare Funktionen der x . Wir können (8) etwas einfacher schreiben, indem wir für (3), (7) und (8)

$$x'_j = \sum_{k=1}^n \alpha_{jk} x_k \quad (j = 1, 2, \dots, n) \quad (3c)$$

$$x''_i = \sum_{j=1}^n \beta_{ij} x'_j \quad (i = 1, 2, \dots, n) \quad (7a)$$

und

$$x''_i = \sum_{j=1}^n \sum_{k=1}^n \beta_{ij} \alpha_{jk} x_k \quad (8a)$$

schreiben und weiter

$$\gamma_{ik} = \sum_{j=1}^n \beta_{ij} a_{jk} \quad (9)$$

setzen. Wir erhalten so

$$x_i'' = \sum_{k=1}^n \gamma_{ik} x_k. \quad (8b)$$

Die Zusammensetzung der beiden Transformationen (7) und (3) mit den Matrizen (β_{ik}) und (a_{ik}) ergibt wieder eine lineare Transformation. Man sagt, die Matrix (γ_{ik}) ist das Produkt der Matrizen (β_{ik}) und (a_{ik}) . Führt α den Vektor x in $x' = \alpha x$ und β den Vektor x' in $x'' = \beta x'$ über, so führt das Produkt $\beta \alpha$ definitionsgemäß x direkt in $(\beta \alpha)x = x''$ über. Unter dem Produkt (γ_{ik}) der Matrizen (β_{ik}) und (a_{ik}) verstehen wir also diejenige Matrix, deren Koeffizientenschema nach (9) aus den Koeffizientenschemata der beiden Matrizen β und α hervorgeht. Es ist zu beachten, daß der Koeffizient in der i -Zeile und k -Kolonne, das (i, k) -Element des Produkts, nur von den Koeffizienten der i -Zeile des ersten und k -Spalte des zweiten Faktors abhängt. Die soeben definierte Zusammensetzung von Transformationen, die sogenannte **Matrizenmultiplikation**, hat eine Reihe einfacher Eigenschaften, die wir nun betrachten wollen.

1. Vor allem bemerken wir, daß die Matrizenmultiplikationsregel formal dieselbe ist, wie eine Determinantenmultiplikationsregel. Die Determinante des Produkts zweier Matrizen ist also gleich dem Produkt der Determinanten der beiden Faktoren.

2. Es gilt sodann nicht notwendig

$$\alpha \beta = \beta \alpha. \quad (*)$$

Ein Beispiel, wo (*) nicht gilt, bilden die Matrizen

$$\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \text{ und } \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}$$

Es ist nämlich

$$\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix} \text{ und } \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix}.$$

Das Produkt zweier Matrizen hängt sogar in der Regel von der Reihenfolge ab, und es müssen schon ganz bestimmte Beziehungen vorhanden sein, damit (*) gelten soll. In diesem Falle nennen wir die Matrizen α und β vertauschbar.

3. Im Gegensatz zum kommutativen Gesetz gilt für Matrizen das assoziative Gesetz der Multiplikation. Es ist

$$\gamma(\beta\alpha) = (\gamma\beta)\alpha, \quad (10)$$

d. h. es ist gleichgültig, ob man γ mit dem Produkt von β und α oder das Produkt von γ und β mit α multipliziert. Bezeichnet man den (i, k) -Koeffizienten der Matrix auf der linken Seite von (10) mit ε_{ik} , so ist

$$\varepsilon_{ik} = \sum_{j=1}^n \gamma_{ij}(\beta\alpha)_{jk} = \sum_{j=1}^n \gamma_{ij} \sum_{l=1}^n \beta_{jl} \alpha_{lk}. \quad (10a)$$

Das (i, k) -Element ε'_{ik} der rechten Seite von (7) ist

$$\varepsilon_{ik} = \sum_{l=1}^n (\gamma\beta)_{il} \alpha_{lk} = \sum_{l=1}^n \sum_{j=1}^n \gamma_{ij} \beta_{jl} \alpha_{lk}. \quad (10b)$$

Da man die Summenzeichen an der rechten Seite von (10a) beide herausziehen kann, ist $\varepsilon_{ik} = \varepsilon'_{ik}$, womit (10) bewiesen ist. Man darf für beide Seiten von (10) einfacher $\gamma\beta\alpha$ schreiben. Daß man dies tun darf, wird auch unmittelbar klar, wenn man die Matrizen als lineare Operatoren auffaßt. Es führt nämlich α den Vektor ξ in $\alpha\xi = \xi'$, β den Vektor ξ' in $\beta\xi' = \xi''$ und γ den Vektor ξ'' in $\gamma\xi'' = \xi'''$ über. Dann bedeutet das Zusammenfassen zweier Matrizen zu einer einzigen, die Matrizenmultiplikation, nur das Zusammenfassen von zwei Operationen: $\beta\alpha$ führt ξ direkt in ξ'' und $\gamma\beta$ den Vektor ξ' in ξ''' über. So führt sowohl $\gamma(\beta\alpha)$ wie auch $(\gamma\beta)\alpha$ den Vektor ξ in ξ''' über: die beiden Operationen sind äquivalent.

Während also die Reihenfolge der Faktoren in (10) nicht geändert werden darf, kommt es auf die Klammern nicht an, sie können einfach weggelassen werden.

4. So wie die Zahl 1 in der gewöhnlichen Zahlenmultiplikation, so spielt hier die Matrix

$$1 = \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1 \end{pmatrix} \quad (11)$$

eine besondere Rolle. Es gilt für alle Matrizen α

$$1\alpha = \alpha 1 = \alpha.$$

1 ist mit allen Matrizen vertauschbar, und das Produkt ist die ursprüngliche Matrix. Die Koeffizienten der Matrix **1** bezeichnet man mit dem Symbol δ_{ik} , wobei also

$$\begin{aligned}\delta_{ik} &= 0 \quad \text{für } i \neq k, \\ \delta_{ik} &= 1 \quad \text{für } i = k.\end{aligned}\quad (12)$$

Das so definierte δ_{ik} nennt man das Weierstrasssche Delta-symbol. Die Matrix $(\delta_{ik}) = 1$ induziert auch eine Transformation, die identische, wobei die neuen Variablen gleich den alten sind.

Gibt es zur Matrix **a** eine Matrix **β** derart, daß

$$\beta a = 1, \quad (13)$$

so nennen wir **β** die Reziproke der Matrix **a**. Gleichung (13) bedeutet, daß eine Transformation mit der Matrix **β** existiert, die mit **a** zusammengesetzt die identische Transformation ergibt. Ist die Determinante von **a** nicht gleich Null, $|a_{ik}| \neq 0$, so existiert, wie wir schon S. 2 besprochen haben, immer eine solche Transformation. Um dies noch im einzelnen auszuführen, schreiben wir die n^2 Gleichungen von (13) etwas mehr explizite hin. Sie lauten

$$\sum_{j=1}^n \beta_{ij} a_{jk} = \delta_{ik} \quad (i = 1, 2, \dots, n; k = 1, 2, \dots, n). \quad (14)$$

Betrachten wir unter diesen etwa diejenigen n Gleichungen, bei denen i denselben Wert, etwa den Wert l hat. Diese bilden n lineare Gleichungen für die n Unbekannten $\beta_{l1}, \beta_{l2}, \dots, \beta_{ln}$, sie haben also eine und nur eine Auflösung, wenn die Determinante $|a_{ik}|$ nicht verschwindet. Dasselbe gilt für die anderen $n - 1$ Gleichungssysteme.

5. Ist also die Determinante $|a_{ik}| \neq 0$, so existiert eine und nur eine Matrix **β**, so daß $\beta a = 1$. Der Wert der Determinante dieser Matrix **β** ist der reziproke der von **a**, da die Determinante von **1** gleich 1 ist, und nach Satz 1

$$|\beta_{ik}| \cdot |a_{ik}| = 1 \quad (15)$$

sein muß. Hieraus folgt auch sofort, daß **a** keine Reziproke haben kann, wenn $|a_{ik}| = 0$, und daß **β**, die Reziproke von **a**, selber auch eine Reziproke haben muß.

Wir wollen nun zeigen, daß, wenn (13) gilt, auch

$$a \beta = 1 \quad (16)$$

ist, d. h. die Reziproke von β wieder α ist. Am einfachsten sehen wir dies ein, indem wir (13) von rechts mit β multiplizieren:

$$\beta \alpha \beta = \beta \quad (17)$$

und dies von links mit der Reziproken von β , die wir mit γ bezeichnen. Es folgt

$$\gamma \beta \alpha \beta = \gamma \beta$$

und, da $\gamma \beta = 1$ ist, ist dies mit (16) identisch. Umgekehrt folgt natürlich auch (13) aus (16).

6. Die Reziproke von α bezeichnet man mit α^{-1} . Ist α^{-1} die Reziproke von α , so ist auch α die von α^{-1} . Reziproke Matrizen sind miteinander vertauschbar.

Regel: Die Reziproke eines Produkts $\alpha \beta \gamma \delta \dots$ erhält man, wenn man die Reziproken der einzelnen Faktoren in umgekehrter Reihenfolge miteinander multipliziert, d. h. $\dots \delta^{-1} \gamma^{-1} \beta^{-1} \alpha^{-1}$ bildet. In der Tat ist (wegen des Assoziativgesetzes)

$$\dots \delta^{-1} \gamma^{-1} \beta^{-1} \alpha^{-1} \alpha \beta \gamma \delta \dots = 1.$$

7. Eine weitere sehr wichtige Matrix ist noch die Nullmatrix

$$\mathbf{0} = \begin{pmatrix} 0 & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \dots & 0 \end{pmatrix}, \quad (18)$$

deren Koeffizienten alle 0 sind. Es gilt offenbar

$$\alpha \mathbf{0} = \mathbf{0} \alpha = \mathbf{0}. \quad (18a)$$

Sie spielt bei einer anderen Verknüpfung der Matrizen, bei der Addition, eine wichtige Rolle. Die Summe γ der Matrizen α und β ist diejenige Matrix, deren Koeffizienten

$$\gamma_{ik} = \alpha_{ik} + \beta_{ik} \quad (i = 1, 2, \dots, n; k = 1, 2, \dots, n) \quad (19)$$

sind: die n^2 Gleichungen (19) sind mit der Gleichung

$$\gamma = \alpha + \beta \quad \text{oder} \quad \gamma - \alpha - \beta = \mathbf{0} \quad (19a)$$

äquivalent.

Es gilt offenbar

$$\alpha + \beta = \beta + \alpha \quad (20)$$

und auch

$$\left. \begin{aligned} \gamma(\alpha + \beta) &= \gamma \alpha + \gamma \beta, \\ (\alpha + \beta)\gamma &= \alpha \gamma + \beta \gamma. \end{aligned} \right\} \quad (20a)$$

Weiter versteht man unter dem Produkt einer Zahl a mit einer Matrix α diejenige Matrix γ , für die gilt

$$\gamma_{ik} = a \alpha_{ik}. \quad (21)$$

Die Formeln

$$(ab) \cdot \alpha = a \cdot (b \alpha); \quad \alpha \alpha \beta = \alpha \alpha \beta; \quad a(\alpha + \beta) = a\alpha + a\beta \quad (21a)$$

folgen dann leicht.

Da die ganzzahligen Potenzen einer Matrix α durch sukzessive Multiplikation leicht definiert werden können

$$\left. \begin{array}{l} \alpha^2 = \alpha \cdot \alpha; \quad \alpha^3 = \alpha \cdot \alpha \cdot \alpha; \dots \\ \alpha^{-2} = \alpha^{-1} \cdot \alpha^{-1}; \quad \alpha^{-3} = \alpha^{-1} \cdot \alpha^{-1} \cdot \alpha^{-1}; \dots \end{array} \right\} \quad (22)$$

so kann man auch jedes Polynom mit positiven und negativen ganzzahligen Exponenten etwa

$$\left. \begin{array}{l} \dots + a_{-3} \alpha^{-3} + a_{-2} \alpha^{-2} + a_{-1} \alpha^{-1} + a_0 \mathbf{1} + a_1 \alpha \\ \quad + a_2 \alpha^2 + a_3 \alpha^3 + \dots \end{array} \right\} \quad (23)$$

definieren. Die Koeffizienten a sollen dabei keine Matrizen, sondern Zahlen sein. Jede solche Funktion (23) von α ist dann mit jeder anderen Funktion von α (also unter anderem auch mit α selber) vertauschbar.

8. Eine sehr häufig vorkommende Art von Matrizen bilden die sogenannten Diagonalmatrizen, deren Koeffizienten außerhalb der Hauptdiagonalen alle 0 sind.

$$\begin{pmatrix} D_1 & 0 & 0 & \dots & 0 \\ 0 & D_2 & 0 & \dots & 0 \\ 0 & 0 & D_3 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & D_n \end{pmatrix} \quad (24)$$

Das allgemeine Element dieser Diagonalmatrix kann

$$D_{ik} = D_i \delta_{ik} \quad (25)$$

geschrieben werden.

Alle Diagonalmatrizen sind miteinander vertauschbar, und das Produkt ist wieder eine Diagonalmatrix.

Es ist nämlich

$$(\mathbf{D} \mathbf{D}')_{ik} = \sum_j \mathbf{D}_{ij} \mathbf{D}'_{jk} = \sum_j D_i \delta_{ij} D'_j \delta_{jk} = D_i D'_i \delta_{ik}. \quad (26)$$

Ist umgekehrt eine Matrix α mit einer Diagonalmatrix D vertauschbar, deren Diagonalelemente alle voneinander verschieden sind, so ist α selber eine Diagonalmatrix. Es ist nämlich

$$(\alpha D)_{ik} = \alpha_{ik} D_k = (D \alpha)_{ik} = D_i \alpha_{ik}, \quad (27)$$

woraus

$$(D_i - D_k) \alpha_{ik} = 0 \quad (27a)$$

folgt. Da aber für $i \neq k$ auch $D_i \neq D_k$ ist, muß dann $\alpha_{ik} = 0$ sein, d. h. α ist eine Diagonalmatrix.

Die Summe der Diagonalelemente einer Matrix, die Diagonalsumme, in der mathematischen Literatur zumeist Spur genannt, ist durch die Gleichung

$$\text{Spur } \alpha = \sum_i \alpha_{ii} = \alpha_{11} + \alpha_{22} + \cdots + \alpha_{nn} \quad (28)$$

definiert.

9. Die Spur des Produkts von zwei Matrizen ist unabhängig von der Reihenfolge der Faktoren.

$$\text{Spur } \alpha \beta = \text{Spur } \beta \alpha. \quad (29)$$

In der Tat, es ist

$$\text{Spur } \alpha \beta = \sum_i (\alpha \beta)_{ii} = \sum_i \sum_j \alpha_{ij} \beta_{ji} = \text{Spur } \beta \alpha. \quad (30)$$

Ihre wichtigste Anwendung findet diese Regel bei der sogenannten Ähnlichkeitstransformation von Matrizen. Diese bedeutet eine Multiplikation der zu transformierenden Matrix α mit der transformierenden Matrix β von hinten und ihrer Reziproken von vorne. Sie führt also α in $\beta^{-1} \alpha \beta$ über.

Bei einer Ähnlichkeitstransformation bleibt die Spur ungeändert. Es hat nämlich $\beta^{-1} \alpha \cdot \beta$ nach Satz 9 dieselbe Spur wie $\beta \cdot \beta^{-1} \alpha = \alpha$.

10. Die Wichtigkeit der Ähnlichkeitstransformationen beruht auf der Tatsache, daß alle Matrizengleichungen gültig bleiben, wenn man jede darin vorkommende Matrix ein und derselben Ähnlichkeitstransformation unterwirft. Ist z. B. $\alpha \beta = \gamma$, so ist auch $\sigma^{-1} \alpha \sigma \cdot \sigma^{-1} \beta \sigma = \sigma^{-1} \gamma \sigma$ und mit $\alpha \beta = 1$ auch $\sigma^{-1} \alpha \sigma \cdot \sigma^{-1} \beta \sigma = \sigma^{-1} 1 \sigma = 1$.

Auch auf die allgemeineren Matrizenoperationen, die auf S. 9 eingeführt sind, bezieht sich dieses. In der Tat folgt aus $\gamma = \alpha + \beta$ auch $\sigma^{-1} \gamma = \sigma^{-1} (\alpha + \beta) = \sigma^{-1} \alpha + \sigma^{-1} \beta$ und $\sigma^{-1} \gamma \sigma$

$= (\sigma^{-1} \alpha + \sigma^{-1} \beta) \sigma = \sigma^{-1} \alpha \sigma + \sigma^{-1} \beta \sigma$. Ebenso folgt aus $\beta = \alpha \alpha$ auch $\sigma^{-1} \beta \sigma = \sigma^{-1} \alpha \alpha \sigma = \alpha \sigma^{-1} \alpha \sigma$.

Hieraus folgt, daß Satz 10 für jede Matrizengleichung gilt, die nur die Operationen: Multiplikation einer Matrix mit einer Zahl oder anderen Matrix, Potenzierung einer Matrix mit positiven oder negativen, ganzzahligen Exponenten, Addition von Matrizen enthält.

Diese zehn einfachen Sätze des Matrizenkalküls werden zum großen Teil schon vielen Lesern bekannt sein. Ich wollte sie hier doch noch einmal wiederholen, da wir später öfters gezwungen sein werden, sie sogar „zwischen den Zeilen“ zu benutzen¹⁾, und es ist für das Spätere — wie überhaupt für jede quantenmechanische Rechnung — die sichere Beherrschung dieser Grundregeln unerlässlich. Sie kommen schon in der ersten Publikation über Quantenmechanik von M. Born und P. Jordan²⁾ vor.

Lineare Abhängigkeit von Vektoren

Man nennt die Vektoren v_1, v_2, \dots, v_t voneinander linear unabhängig, wenn unter ihnen keine Beziehung

$$a_1 v_1 + a_2 v_2 + \dots + a_k v_t = 0 \quad (31)$$

besteht, in der nicht alle a_1, a_2, \dots, a_k Null wären. Mit anderen Worten bedeutet dies, daß keiner der Vektoren v_1, v_2, \dots, v_t durch die übrigen linear ausgedrückt werden kann. Im Falle, daß ein Vektor, etwa $v_1 = 0$ (alle seine Komponenten Null) ist, sind die Vektoren nicht mehr linear unabhängig, da die Beziehung

$$1 \cdot v_1 + 0 \cdot v_2 + 0 \cdot v_3 + \dots + 0 \cdot v_t = 0$$

unter ihnen dann sicher besteht, obwohl $a_1 = 1 \neq 0$ ist.

Z. B. sind die vierdimensionalen Vektoren $v_1(1, 2, -1, 3); v_2(0, -2, 1, -1); v_3(2, 2, -1, 5)$ voneinander linear abhängig, da

$$2v_1 + v_2 - v_3 = 0, \quad v_1 = -\frac{1}{2}v_2 + \frac{1}{2}v_3$$

ist. Dagegen sind v_1 und v_2 noch linear unabhängig.

Sind k Vektoren v_1, v_2, \dots, v_t voneinander nicht linear unabhängig, so kann man unter ihnen $k' < k$ Vektoren finden, die

¹⁾ So ist z. B. Satz 3, das Assoziativgesetz der Multiplikation, bei der Ableitung von Satz 6 (die Vertauschbarkeit von Reziproken) implizite bereits dreimal benutzt. (Man überzeuge sich hiervon durch Ausschreiben aller Klammern!)

²⁾ Zeitschr. f. Phys. **34**, 858, 1925.

linear unabhängig sind und durch die man alle k linear ausdrücken kann. Um diese k' Vektoren herauszufinden, lassen wir zunächst alle Vektoren $v_i = 0$, deren Komponenten alle verschwinden, weg, diese kann man ja schon ohne die anderen ausdrücken, wie die Gleichung $v_i = 0$ zeigt. Dann geht man die übriggebliebenen Vektoren der Reihe nach durch, nimmt aber nur jene auf, die durch die vorangehenden nicht linear ausgedrückt werden können. Die so erhaltenen k' Vektoren sind voneinander linear unabhängig und alle anderen lassen sich durch diese ausdrücken. Wurde ein Vektor nicht aufgenommen, so konnte er schon durch die vorangehenden linear ausgedrückt werden. Unter den aufgenommenen kann dagegen keine lineare Beziehung bestehen. Ist nämlich etwa v_j der zuletzt aufgenommene Vektor, der in der linearen Beziehung vorkommt (mit dem Koeffizienten $a_j \neq 0$), so kann man v_j durch die vorangehenden linear ausdrücken und er hätte nicht aufgenommen werden dürfen.

Sind k Vektoren v_1, v_2, \dots, v_k linear unabhängig oder nicht, so gilt dies auch für die Vektoren $\alpha v_1, \alpha v_2, \dots, \alpha v_k$, wenn α eine eigentliche Transformation ist. Wäre nämlich

$$a_1 \alpha v_1 + a_2 \alpha v_2 + \cdots + a_k \alpha v_k = 0, \quad (32)$$

so folgte, wenn man auf beide Seiten α^{-1} anwendet, wegen der Linearität von α^{-1}

$$a_1 v_1 + a_2 v_2 + \cdots + a_k v_k = 0, \quad (32a)$$

und umgekehrt folgt aus (32a) auch (32). Besteht also zwischen den αv_i eine lineare Abhängigkeit (32), so besteht diese auch zwischen den v_i und umgekehrt.

Mehr als n n -dimensionale Vektoren $v_1, v_2, \dots, v_n, v_{n+1}$ können nicht linear unabhängig sein.

Die Gleichung

$$a_1 v_1 + a_2 v_2 + \cdots + a_n v_n + a_{n+1} v_{n+1} = 0 \quad (33)$$

ist nämlich den n linearen homogenen Gleichungen für die Komponenten

$$\left. \begin{aligned} a_1 v_{11} + a_2 v_{21} + \cdots + a_n v_{n1} + a_{n+1} v_{n+1,1} &= 0, \\ a_1 v_{12} + a_2 v_{22} + \cdots + a_n v_{n2} + a_{n+1} v_{n+1,2} &= 0, \\ a_1 v_{13} + a_2 v_{23} + \cdots + a_n v_{n3} + a_{n+1} v_{n+1,3} &= 0, \\ \vdots &\quad \vdots \end{aligned} \right\} (33a)$$

äquivalent. In diesen kann man die $a_1, a_2, \dots, a_n, a_{n+1}$ als Unbekannte auffassen: n homogene lineare Gleichungen für $n+1$ Unbekannte haben aber immer von Null verschiedene Lösungen, die in (33) eingesetzt die lineare Abhängigkeit ergeben. Hieraus folgt, daß durch n linear unabhängige n -dimensionale Vektoren v_1, v_2, \dots, v_n jeder n -dimensionale Vektor w linear ausgedrückt werden kann. Es muß nämlich eine Beziehung

$$a_1 v_1 + a_2 v_2 + \cdots + a_n v_n + a w = 0$$

unter ihnen mit $a \neq 0$ existieren (wäre $a = 0$, so wären schon die v_1, v_2, \dots, v_n nicht linear unabhängig), woraus man w berechnen kann. Die Vektoren v_1, v_2, \dots, v_n bilden ein vollständiges Vektorensystem.

Man kann eine Spalte oder auch eine Zeile einer n -dimensionalen Matrix (a_{it}) als einen Vektor auffassen. Die Komponenten des Vektors $a_{.t}$, der die t -Spalte bildet, sind $a_{1t}, a_{2t}, \dots, a_{nt}$, die Komponenten des Vektors $a_{i.}$, der die i -Zeile bildet, sind $a_{i1}, a_{i2}, \dots, a_{in}$. Die Existenz einer linearen Abhängigkeit zwischen den Vektoren $a_{.1}, a_{.2}, \dots, a_{.n}$

$$a_1 a_{.1} + a_2 a_{.2} + \cdots + a_n a_{.n} = 0$$

ist gleichbedeutend mit der Lösbarkeit der linearen homogenen Gleichungen für die a_1, a_2, \dots, a_n

$$a_1 a_{11} + a_2 a_{12} + \cdots + a_n a_{1n} = 0,$$

$$a_1 a_{21} + a_2 a_{22} + \cdots + a_n a_{2n} = 0,$$

.....

$$a_1 a_{n1} + a_2 a_{n2} + \cdots + a_n a_{nn} = 0,$$

ihre notwendige und hinreichende Bedingung ist das Verschwinden der Determinante $|a_{it}|$.

Ist die Determinante $|a_{it}| \neq 0$, so sind die Vektoren $a_{.1}, a_{.2}, \dots, a_{.n}$ linear unabhängig, sie bilden ein vollständiges Vektorsystem. Sind umgekehrt die Vektoren $v_{.1}, v_{.2}, \dots, v_{.n}$ linear unabhängig, so ist die Determinante der Matrix, deren Spalten diese Vektoren bilden, ungleich Null. Dasselbe gilt von den Zeilen.

II. Verallgemeinerungen

1. Wir wollen jetzt die Resultate des vorangehenden Kapitels — zunächst formal, dann aber auch dem Inhalte nach — verallgemeinern. Zur Kennzeichnung der Komponenten eines Vektors

bzw. der Koeffizienten einer Matrix haben wir die Nummer der Koordinatenachse, auf welche sie sich bezieht, als Index hinzugefügt. Die Koordinatenachsen waren mit den Zahlen $1, 2, \dots, n$ nummeriert. Hier von wollen wir jetzt abgehen und die Koordinatenachsen nach den Elementen einer beliebigen Gesamtheit benennen. Ist G eine Gesamtheit der Dinge g, h, i, \dots , so ist ein Vektor v im Raum der Gesamtheit G der Inbegriff der Zahlen v_g, v_h, v_i, \dots . Man kann dann natürlich nur Vektoren, die in demselben Raum definiert sind, miteinander vergleichen (oder auch addieren usw.), da nur diese entsprechende, d. h. gleichbenannte Komponenten haben.

Dasselbe wollen wir auch bei Matrizen einführen. Damit eine Matrix α auf den Vektor v mit den Komponenten v_g, v_h, v_i, \dots anwendbar sein soll, müssen die Spalten von α nach den Elementen derselben Gesamtheit G benannt sein, nach denen die Komponenten von v benannt sind. Im einfachsten Falle sind dann auch die Zeilen nach den Elementen g, h, i, \dots dieser Gesamtheit benannt und α führt den Vektor v im Raum der Gesamtheit G in den Vektor αv in demselben Raum über. Es ist

$$v_j = \sum_{l \in G} \alpha_{jl} v_l, \quad (1)$$

wo j ein Element der Gesamtheit G ist und l alle Elemente dieser Gesamtheit durchlaufen muß.

Z. B. kann man die Koordinatenachsen nach drei Buchstaben x, y, z benennen. Dann ist z. B. v mit den Komponenten $v_x = 1, v_y' = 0, v_z = -2$ ein Vektor, α

$$\alpha = \begin{pmatrix} x & y & z \\ 1 & 2 & 3 \\ 0 & 5 & -1 \\ -4 & -2 & 4 \end{pmatrix} \begin{matrix} x \\ y \\ z \end{matrix}$$

eine Matrix (die Zeichen der entsprechenden Zeilen und Spalten sind angefügt). Es ist z. B. $\alpha_{xz} = 1, \alpha_{xy} = 2, \alpha_{zy} = -2$. Die x -Komponente von $v' = \alpha v$ berechnet sich nach (1)

$$\begin{aligned} v'_x &= \alpha_{xz} v_x + \alpha_{xy} v_y + \alpha_{xz} v_z \\ &= 1 \cdot 1 + 2 \cdot 0 + 3 \cdot -2 = -5 \text{ usw.} \end{aligned}$$

Die Verallgemeinerung, die wir jetzt vorgenommen haben, ist offenbar rein formaler Natur, sie bezieht sich auf eine andere Benennung der Koordinatenachsen und der Komponenten der Vektoren und Matrizen. Man kann zwei Matrizen, die auf Vektoren

im gleichen Raum einwirken, miteinander ebenso multiplizieren, wie die Matrizen des vorangehenden Kapitels.

$$\gamma = \beta \alpha \quad (2)$$

ist gleichbedeutend mit

$$\gamma_{jk} = \sum_{l \in G} \beta_{jl} \alpha_{lk},$$

wo j und k zwei Elemente der Gesamtheit G sind und l alle Elemente dieser Gesamtheit durchläuft.

2. Eine weitere, weniger formale Verallgemeinerung ist die, daß man die Zeilen und Spalten der Matrizen nach den Elementen verschiedener Gesamtheiten F und G benennt. Dann wird aus (1)

$$w_j = \sum_{l \in G} \alpha_{jl} v_l, \quad (1a)$$

wo j ein Element der Gesamtheit F und l alle Elemente der Gesamtheit G durchläuft. Eine solche Matrix, deren Zeilen und Spalten in verschiedener Weise benannt sind, nennt man eine nichtquadratische Matrix (im Gegensatz zu den quadratischen Matrizen des vorangehenden Kapitels), sie führt einen Vektor v im Raum des G in einen Vektor w im Raum des F über. Im allgemeinen muß die Gesamtheit F gar nicht genau so viel Elemente umfassen, wie die Gesamtheit G . Umfassen sie gleich viele Elemente, so hat die Matrix gleich viel Zeilen und Spalten, wir nennen sie „quadratisch im weiteren Sinne“.

Die Gesamtheit G bestehe aus den Zeichen $*$, Δ , \square , die Gesamtheit F aus den Zahlen 1, 2. Dann ist

$$\alpha = \begin{pmatrix} * & \Delta & \square \\ 5 & 7 & 3 \\ 0 & -1 & -2 \end{pmatrix} \begin{matrix} 1 \\ 1 \\ 2 \end{matrix}$$

eine nichtquadratische Matrix (die Namen der Zeilen und Spalten sind wieder angefügt). Sie führt den Vektor $v_* = 1$, $v_\Delta = 0$, $v_\square = -2$ in den Vektor

$$w = \alpha v$$

über. Die Komponenten w_1 und w_2 berechnen sich

$$w_1 = \alpha_{1*} v_* + \alpha_{1\Delta} v_\Delta + \alpha_{1\square} v_\square = 5 \cdot 1 + 7 \cdot 0 + 3 \cdot -2 = -1,$$

$$w_2 = \alpha_{2*} v_* + \alpha_{2\Delta} v_\Delta + \alpha_{2\square} v_\square = 0 \cdot 1 + -1 \cdot 0 + -2 \cdot -2 = 4.$$

Zwei nichtquadratische Matrizen β und α kann man miteinander nur dann multiplizieren, wenn die Spalten des ersten Faktors und die Zeilen des zweiten Faktors nach der-

selben Gesamtheit F benannt sind, wenn die Zeilen des zweiten Faktors auf die Spalten des ersten „passen“. Die Zeilen des ersten und Spalten des zweiten Faktors dagegen können nach den Elementen verschiedener Gesamtheiten E und G benannt sein. Es ist dann

$$\gamma = \beta \alpha \quad (2a)$$

gleichbedeutend mit

$$\gamma_{jk} = \sum_{l < F} \beta_{jl} \alpha_{lk},$$

wo j ein Element von E , k eines von G ist und l alle Elemente von F durchläuft: Es führt dann α einen Vektor im Raume von G in einen im Raume von F über, β führt dann diesen Vektor in einen im Raume von E über. Daher führt γ einen Vektor im Raume von G in einen im Raume von E über.

Sei etwa G wieder $*$, Δ , \square , dagegen umfasse F die Buchstaben x, y und E die Zahlen 1, 2. Ist

$$\beta = \begin{pmatrix} x & y \\ 7 & 8 \\ 9 & 3 \end{pmatrix} \begin{matrix} 1 \\ 2 \end{matrix} \text{ und } \alpha = \begin{pmatrix} * & \Delta & \square \\ 2 & 3 & 4 \\ 5 & 6 & 7 \end{pmatrix} \begin{matrix} x \\ y \end{matrix},$$

so ist

$$\gamma_{1*} = \beta_{1x} \alpha_{x*} + \beta_{1y} \alpha_{y*} = 7 \cdot 2 + 8 \cdot 5 = 54$$

oder

$$\gamma_{2\Delta} = \beta_{2x} \alpha_{x\Delta} + \beta_{2y} \alpha_{y\Delta} = 9 \cdot 3 + 3 \cdot 6 = 45$$

und

$$\gamma = \begin{pmatrix} * & \Delta & \square \\ 5 & 4 & 6 & 9 & 8 & 4 \\ 3 & 3 & 4 & 5 & 5 & 7 \end{pmatrix} \begin{matrix} 1 \\ 2 \end{matrix}.$$

3. Wir müssen noch untersuchen, wie sich die 10 Gesetze der Matrizenrechnung, die wir im ersten Kapitel abgeleitet haben, für nichtquadratische Matrizen modifizieren. Daß sie auch für die allgemeineren quadratischen Matrizen, mit denen wir uns am Anfang dieses Kapitels befaßt haben, noch genau gelten, sieht man sofort, wenn man bedenkt, daß die Zahlennatur der Indizes im ersten Kapitel nirgends benutzt worden ist.

Die Addition zweier nichtquadratischer Matrizen — wie auch diejenige zweier Vektoren — setzt voraus, daß sie in denselben Achsenkreuzen definiert sind, d. h. daß die Zeilen auf Zeilen und Spalten auf Spalten passen. In der Gleichung

$$\alpha + \beta = \gamma$$

muß die Zeilenbenennung der drei Matrizen α , β , γ dieselbe sein, ebenso ihre Spaltenbenennung. Ganz anders ist es bei der Multiplikation, wo die Spalten des ersten Faktors auf die Zeilen des anderen Faktors passen mußten und das Produkt nur in diesem Falle und in diesem Falle immer gebildet werden konnte: es hatte die Zeilenbenennung des ersten und Spaltenbenennung des zweiten Faktors.

1. Von einer Determinante nichtquadratischer Matrizen kann noch gesprochen werden, wenn sie gleich viele Zeilen und Spalten haben, wenn diese auch verschieden benannt sind. Für diese „quadratischen Matrizen im weiteren Sinne“ gilt die Regel, daß die Determinante des Produkts gleich dem Produkte der Determinanten ist.

2. 3. Für die Multiplikation von nichtquadratischen Matrizen gilt ebenfalls das Assoziativgesetz:

$$(\alpha \beta) \gamma = \alpha (\beta \gamma), \quad (3)$$

und zwar lassen sich alle Multiplikationen auf der rechten Seite wirklich ausführen, wenn dies auf der linken Seite der Fall ist, und umgekehrt.

4. 5. 6. Die Matrix 1 soll immer als eine quadratische Matrix mit gleicher Zeilen- und Spaltenbenennung angesehen werden. Die Multiplikation mit ihr kann immer weggelassen werden.

Die quadratischen Matrizen im weiteren Sinne mit nichtverschwindender Determinante haben eine Reziproke. Für eigentliche nichtquadratische Matrizen wird keine Reziproke definiert. Ist α eine quadratische Matrix im weiteren Sinne, so müssen, damit die Gleichung

$$\beta \alpha = 1$$

sinnvoll sei, die Spalten von β auf die Zeilen von α passen. Weiter muß auch 1 die Zeilenbenennung von α haben, und da es im engeren Sinne quadratisch ist, auch die Spaltenbenennung. Folglich muß auch die Spaltenbenennung von β dieselbe sein.

Die Zeilen der Reziproken β einer Matrix α sind nach den Elementen derselben Gesamtheit benannt, nach dessen Elementen die Spalten von α benannt sind, ihre Spalten nach denselben Elementen, nach denen die Zeilen von α benannt sind. Es existiert zu jeder quadratischen Matrix im

weiteren Sinne α mit nichtverschwindender Determinante eine Reziproke β , so daß

$$\beta \alpha = 1; \quad (4)$$

es ist dann auch

$$\alpha \beta = 1, \quad (4a)$$

wobei aber zu bemerken ist, daß die Zeilen und Spalten der 1 in (4) anders benannt sind als die der 1 von (4a).

7. Für die Nullmatrix und die Addition von nichtquadratischen Matrizen gilt dasselbe, was für quadratische Matrizen gilt, ebenso von der Multiplikation mit einer Zahl. Dagegen kann man die Potenzen, z. B. von nichtquadratischen Matrizen, nicht bilden, weil die Multiplikation von α mit α voraussetzt, daß die Spalten von α auf die Zeilen von α passen, d. h. daß α quadratisch ist.

8. 9. 10. Diagonalmatrizen und Spur sind keine für nichtquadratische Matrizen sinnvolle Begriffe, auch wird keine Ähnlichkeitstransformation für diese definiert. Betrachten wir nämlich die Gleichung

$$\sigma \alpha \sigma^{-1} = \beta,$$

so ist die Zeilenbenennung von β und σ gleich. Diese ist aber auch die Spaltenbenennung von σ^{-1} , also auch von β . Die Matrix β ist im engeren Sinne quadratisch, ebenso α , dessen Zeilen auf die Spalten von σ und Spalten auf die Zeilen von σ^{-1} passen müssen.

Dagegen kann σ eine im weiteren Sinne quadratische Matrix sein: dann ist die Zeilen- und Spaltenbenennung von α und β verschieden. Solche Ähnlichkeitstransformationen, die die Benennung der Zeilen und Spalten ändern, haben sogar eine wesentliche Bedeutung. Zum Beispiel in der sogenannten Transformationstheorie der Quantenmechanik.

Die Einführung der nichtquadratischen Matrizen ist trotz der scheinbar mit ihnen verbundenen Komplikation sehr vorteilhaft, da man mit ihrer Hilfe die Ausdrucksweise wesentlich vereinfachen kann. Das soeben Beschriebene soll aber nicht etwa als ein starres Schema verwendet werden, sondern eher an das Denken mit diesen Größen gewöhnen. Wo später solche komplizierteren Matrizen gebraucht werden, werden sie immer separat erläutert. Zumeist ist die Benennung der Zeilen und Spalten so sehr durch die Gestalt

und Definition der Koeffizienten nahegelegt, daß man sie kaum weiter anzugeben braucht.

4. Am häufigsten kommt es vor, daß man die Zeilen nicht mit einer Zahl, sondern mit zwei oder mehreren Zahlen benennt, z. B.

$$\gamma = \begin{pmatrix} a_1 b_1 c_1 d_1, & a_1 b_1 c_1 d_2, & a_1 b_1 c_2 d_1, & a_1 b_1 c_2 d_2, \\ a_1 b_2 c_1 d_1, & a_1 b_2 c_1 d_2, & a_1 b_2 c_2 d_1, & a_1 b_2 c_2 d_2, \\ a_2 b_1 c_1 d_1, & a_2 b_1 c_1 d_2, & a_2 b_1 c_2 d_1, & a_2 b_1 c_2 d_2, \\ a_2 b_2 c_1 d_1, & a_2 b_2 c_1 d_2, & a_2 b_2 c_2 d_1, & a_2 b_2 c_2 d_2, \end{pmatrix} \quad (\dagger)$$

Die erste Zeile wird man „1, 1-Zeile“, die zweite „1, 2-Zeile“, die dritte „2, 1-Zeile“, die vierte „2, 2-Zeile“ nennen und entsprechend die Spalten. Die Koeffizienten von (\dagger) sind

$$\gamma_{ij;kl} = a_i b_j c_k d_l,$$

wobei der besseren Übersicht halber den Namen der Zeile (i, j) vom Namen der Spalte (k, l) ein Semikolon (;) trennt.

Eine wichtige Rolle unter diesen Matrizen spielt das direkte Produkt γ zweier Matrizen (α_{ik}) und (β_{jl}), wir bezeichnen sie nach H. Weyl¹⁾

$$\gamma = \alpha \times \beta. \quad (5)$$

(5) ist gleichbedeutend mit

$$\gamma_{ij;kl} = \alpha_{ik} \beta_{jl}. \quad (6)$$

Die Zeilenzahl von α sei n_1 , die Spaltenzahl n_2 ; die entsprechenden Zahlen für β seien n'_1 und n'_2 , dann hat γ genau $n_1 n'_1$ Zeilen und $n_2 n'_2$ Spalten. Ist insbesondere sowohl α wie β eine quadratische Matrix, so ist es auch $\alpha \times \beta$.

Satz 1. Ist $\alpha \bar{\alpha} = \bar{\alpha}$ und $\beta \bar{\beta} = \bar{\beta}$, und ist $\alpha \times \beta = \gamma$ und $\bar{\alpha} \times \bar{\beta} = \bar{\gamma}$, so ist auch $\gamma \bar{\gamma} = \bar{\alpha} \times \bar{\beta}$

$$(\alpha \times \beta) (\bar{\alpha} \times \bar{\beta}) = \alpha \bar{\alpha} \times \beta \bar{\beta}. \quad (7)$$

¹⁾ Die Faktoren α und $\bar{\alpha}$ der Matrizenmultiplikation schreiben wir einfach nebeneinander $\alpha \bar{\alpha}$. Die Matrix (\dagger) ist das direkte Produkt der beiden Matrizen:

$$\begin{pmatrix} a_1 c_1 & a_1 c_2 \\ a_2 c_1 & a_2 c_2 \end{pmatrix} \times \begin{pmatrix} b_1 d_1 & b_1 d_2 \\ b_2 d_1 & b_2 d_2 \end{pmatrix} = \gamma.$$

Das Matrizenprodukt zweier direkter Produkte ist das direkte Produkt der beiden Matrizenprodukte. In der Tat ist

$$(\alpha \times \beta)_{ik; i'k'} = \alpha_{ii'} \beta_{kk'}; (\bar{\alpha} \times \bar{\beta})_{i'k'; i''k''} = \bar{\alpha}_{i'i''} \bar{\beta}_{k'k''}$$

und

$$(\alpha \times \beta) \cdot (\bar{\alpha} \times \bar{\beta})_{ik; i'k'} = \sum_{i'k'} \alpha_{ii'} \beta_{kk'} \bar{\alpha}_{i'i''} \bar{\beta}_{k'k''}. \quad (8)$$

Ebenso ist

$$(\alpha \bar{\alpha})_{ii''} = \sum_{i'} \alpha_{ii'} \bar{\alpha}_{i'i''}; (\beta \bar{\beta})_{kk''} = \sum_{k'} \beta_{kk'} \bar{\beta}_{k'k''}$$

und

$$(\alpha \bar{\alpha} \times \beta \bar{\beta})_{ik; i''k''} = \sum_{i'} \alpha_{ii'} \alpha_{i'i''} \cdot \sum_{k'} \beta_{kk'} \bar{\beta}_{k'k''}, \quad (9)$$

aus (8) und (9) folgt aber

$$(\alpha \times \beta) (\bar{\alpha} \times \bar{\beta}) = \alpha \bar{\alpha} \times \beta \bar{\beta}. \quad (7)$$

Satz 2. Das direkte Produkt zweier Diagonalmatrizen ist wieder eine Diagonalmatrix, das direkte Produkt zweier Einheitsmatrizen ist wieder eine Einheitsmatrix. Man überzeugt sich hier von leicht an Hand der Definition des direkten Produkts.

Bei formalem Rechnen mit Matrizen muß man sich stets überzeugen, ob die angedeutete Multiplikation auch ausführbar ist. Im ersten Kapitel, wo wir durchwegs quadratische Matrizen mit n Zeilen und Spalten hatten, war dieses natürlich immer der Fall, im allgemeinen ist aber darauf zu achten, daß die Spalten des ersten Faktors der Matrizenmultiplikation auf die Zeilen des zweiten Faktors passen, d. h. daß die beiden mit dem gleichen Namen bzw. Symbol benannt seien. Das direkte Produkt zweier Matrizen läßt sich nach (6) immer bilden.

Eine etwas allgemeinere Art von Matrizen mit mehreren Indizes bezeichnen M. Born und P. Jordan als Übermatrizen. Sie fassen nämlich etwa die Matrix $(\alpha_{ij; kl})$ als eine Matrix (A_{ik}) auf, deren Koeffizienten A_{ik} selber Matrizen sind: A_{ik} ist diejenige Matrix, bei der in der Zeile j und Spalte l die Zahl $\alpha_{ij; kl}$ steht.

$$(A_{ij; kl}) = \alpha = (A_{ik}), \text{ wenn } (A_{ik})_{jl} = \alpha_{ij; kl}. \quad (10)$$

Satz 3. Ist $\alpha = (A_{iv})$ und $\beta = (B_{i'v'})$, so ist $\alpha \beta = \gamma = (C_{iv'})$, wobei

$$C_{iv'} = \sum_{i'} A_{iv} B_{i'v'}, \quad (11)$$

wobei rechts in (11) eine Summe von Produkten nach der Matrizenmultiplikation steht. Es ist

$$(\alpha \beta)_{ik; i'' k''} = \sum_{i', k'} \alpha_{ik; i' k'} \beta_{i' k'; i'' k''},$$

andererseits

$$\gamma_{ik; i'' k''} = (C_{i' i''})_{kk''} = \sum_{i'} (A_{i i'} B_{i' i''})_{kk''}$$

und

$$(A_{i i'} B_{i' i''})_{kk''} = \sum_{k'} (A_{i i'})_{kk'} (B_{i' i''})_{k' k''} = \sum_{k'} \alpha_{ik; i' k'} \beta_{i' k'; i'' k''},$$

also

$$(\alpha \beta)_{ik; i'' k''} = \gamma_{ik; i'' k''},$$

womit Satz 3 bewiesen ist. Natürlich muß man in der rechten Seite von (11) auf die Reihenfolge der Faktoren achten, was in der entsprechenden Gleichung bei der Multiplikation von einfachen Matrizen nicht notwendig war. Mit dieser Einschränkung kann man Übermatrizen nach der für einfache Matrizen gültigen Regel multiplizieren.

Im einfachsten Fall haben wir zwei quadratische Matrizen, etwa

$$\left(\begin{array}{cc|cc} \alpha_{11} & \alpha_{12} & \alpha_{13} & \alpha_{14} & \alpha_{15} \\ \alpha_{21} & \alpha_{22} & \alpha_{23} & \alpha_{24} & \alpha_{25} \\ \hline \alpha_{31} & \alpha_{32} & \alpha_{33} & \alpha_{34} & \alpha_{35} \\ \alpha_{41} & \alpha_{42} & \alpha_{43} & \alpha_{44} & \alpha_{45} \\ \alpha_{51} & \alpha_{52} & \alpha_{53} & \alpha_{54} & \alpha_{55} \end{array} \right) \quad \text{und} \quad \left(\begin{array}{cc|cc} \beta_{11} & \beta_{12} & \beta_{13} & \beta_{14} & \beta_{15} \\ \beta_{21} & \beta_{22} & \beta_{23} & \beta_{24} & \beta_{25} \\ \hline \beta_{31} & \beta_{32} & \beta_{33} & \beta_{34} & \beta_{35} \\ \beta_{41} & \beta_{42} & \beta_{43} & \beta_{44} & \beta_{45} \\ \beta_{51} & \beta_{52} & \beta_{53} & \beta_{54} & \beta_{55} \end{array} \right) \quad \Theta$$

wir können sie längs der punktierten Linien in Untermatrizen einteilen, aber so, daß die Spalteneinteilung der ersten (2:3) mit der Zeileneinteilung der zweiten übereinstimmt. Wir setzen dann für die beiden Matrizen Θ die Schemata

$$\left(\begin{array}{cc} A_{11} & A_{12} \\ A_{21} & A_{22} \end{array} \right) \quad \text{und} \quad \left(\begin{array}{cc} B_{11} & B_{12} \\ B_{21} & B_{22} \end{array} \right).$$

Das Produkt der beiden Matrizen Θ läßt sich

$$\left(\begin{array}{cc|cc} A_{11} B_{11} + A_{12} B_{21} & A_{11} B_{12} + A_{12} B_{22} \\ A_{21} B_{11} + A_{22} B_{21} & A_{21} B_{12} + A_{22} B_{22} \end{array} \right) = \left(\begin{array}{cc} C_{11} & C_{12} \\ C_{21} & C_{22} \end{array} \right)$$

schreiben.

Dagegen ist das Schema

$$\begin{pmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{pmatrix} \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix} = \begin{pmatrix} B_{11}A_{11} + B_{12}A_{21} & B_{11}A_{12} + B_{12}A_{22} \\ B_{21}A_{11} + B_{22}A_{21} & B_{21}A_{12} + B_{22}A_{22} \end{pmatrix}$$

sinnlos, weil die Zeilen von z. B. B_{11} nicht auf die Spalten von A_{11} passen.

III. Hauptachsentransformation

Im ersten Kapitel haben wir eine wichtige Eigenschaft der Ähnlichkeitstransformation kennengelernt: sie läßt die Spur einer Matrix¹⁾ ungeändert, α hat dieselbe Spur wie $\sigma^{-1}\alpha\sigma$. Ist die Spur einer Matrix ihre einzige Invariante gegenüber Ähnlichkeitstransformationen? Offenbar nicht, denn z. B. die Determinante von $|\sigma^{-1}\alpha\sigma|$ ist ja auch gleich der Determinante von $|\alpha|$. Um weitere Invarianten zu erhalten, betrachten wir die Determinantengleichung n -ten Grades für λ :

$$\left| \begin{array}{cccc} a_{11} - \lambda & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} - \lambda & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} - \lambda \end{array} \right| = 0, \quad (1)$$

oder anders geschrieben,

$$|\alpha - \lambda 1| = 0, \quad (2)$$

wir nennen sie die Säkulargleichung von α . Die Säkulargleichung von $\beta = \sigma^{-1}\alpha\sigma$ ist

$$|\beta - \lambda 1| = |\sigma^{-1}\alpha\sigma - \lambda 1| = 0. \quad (3)$$

Man kann hierfür offenbar auch $|\sigma(\alpha - \lambda 1)\sigma^{-1}| = 0$ oder

$$|\sigma^{-1}| |\alpha - \lambda 1| |\sigma| = 0 \quad (4)$$

schreiben. Gleichung (4) zeigt, daß die n Wurzeln der Säkulargleichung $|\beta - \lambda 1| = 0$ mit den n Wurzeln der Säkulargleichung $|\alpha - \lambda 1| = 0$ übereinstimmen²⁾. Die Wurzeln der Säkulargleichung, die Eigenwerte einer Matrix, sind Invarianten gegenüber Ähnlichkeitstransformationen. Wir werden später sehen,

¹⁾ Die Matrix, die einer Ähnlichkeitstransformation unterworfen wird, muß immer eine quadratische sein. Da kein anderer Grund vorliegt, benennen wir ihre Zeilen und Spalten wieder mit den Zahlen 1, 2, ..., n .

²⁾ $|\sigma^{-1}|$ und $|\sigma|$ sind Zahlen!

daß eine Matrix im allgemeinen keine weiteren Invarianten hat, jetzt bemerken wir nur, daß die Spur die Summe und die Determinante das Produkt der Eigenwerte ist und ihre Invarianz in dem soeben ausgesprochenen Satz enthalten ist.

Betrachten wir einen Eigenwert λ_1 . Die Determinante der Matrix $(\alpha - \lambda_1 I)$ ist Null, so daß die linearen homogenen Gleichungen

$$\left. \begin{aligned} \alpha_{11} x_1 + \alpha_{12} x_2 + \cdots + \alpha_{1n} x_n &= \lambda_1 x_1, \\ \alpha_{21} x_1 + \alpha_{22} x_2 + \cdots + \alpha_{2n} x_n &= \lambda_1 x_2, \\ \vdots &\quad \vdots \\ \alpha_{n1} x_1 + \alpha_{n2} x_2 + \cdots + \alpha_{nn} x_n &= \lambda_1 x_n \end{aligned} \right\} \quad (5)$$

lösbar sind. Es läßt sich zu allen n Eigenwerten λ_i ein lineares homogenes Gleichungssystem nach (5) aufstellen, die Lösungen dieses Systems, die nur bis auf einen konstanten Faktor bestimmt sind, bezeichnen wir mit $x_{1i}, x_{2i}, \dots, x_{ni}$, so daß

$$\sum_j \alpha_{ij} x_{ji} = \lambda_i x_{ii} \quad (5a)$$

gilt. Die Gesamtheit der n Zahlen $x_{1i}, x_{2i}, \dots, x_{ni}$ nennen wir einen Eigenvektor x_i der Matrix α , der Eigenvektor x_i gehört zum Eigenwert λ_i . Gleichung (5a) läßt sich auch

$$\alpha x_i = \lambda_i x_i \quad (5b)$$

schreiben: die Matrix führt ihre Eigenvektoren in Vektoren über, die sich vom Eigenvektor nur durch einen konstanten Faktor, den Eigenwert unterscheiden.

Man kann die Eigenvektoren $x_{1i}, x_{2i}, \dots, x_{ni}$ zu einer Matrix x zusammenfassen, so daß x_i die i -Spalte dieser Matrix bildet:

$$x_{ii} = (x_i)_i = x_{ii}$$

Dann steht in (5a) links der (i,j) -Koeffizient von αx . Auch die rechte Seite kann man als den (i,i) -Koeffizienten einer Matrix auffassen, nämlich als den (i,i) -Koeffizienten von $x A$, wo A eine Diagonalmatrix mit den Diagonalelementen $\lambda_1, \lambda_2, \dots, \lambda_n$

$$A_{jk} = \delta_{jk} \lambda_k$$

ist. Dann lautet (5a)

$$(\alpha x)_{ik} = \sum_j x_{ij} \delta_{jk} \lambda_k = (x A)_{ik},$$

und man kann die n^2 Gleichungen (5a) in

$$\alpha x = x A \quad (6)$$

oder, wenn \mathbf{z} eine Reziproke hat, in

$$\mathbf{z}^{-1} \mathbf{a} \mathbf{z} = \mathbf{A} \quad (6a)$$

zusammenfassen.

Eine Ähnlichkeitstransformation mit einer Matrix, deren Spalten die n Eigenvektoren sind, bringt die Matrix auf die Diagonalform, die Diagonalelemente sind die Eigenwerte der Matrix. Zwei Matrizen, die dieselben Eigenwerte haben, lassen sich ineinander transformieren, es lassen sich nämlich beide in dieselbe Matrix \mathbf{A} überführen. Die Eigenwerte sind die einzigen Invarianten gegenüber Ähnlichkeitstransformationen.

Dies gilt allerdings nur, wenn \mathbf{z} eine Reziproke hat, d. h. wenn die n Vektoren $\mathbf{x}_{.1}, \mathbf{x}_{.2}, \dots, \mathbf{x}_{.n}$ linear unabhängig sind. Das ist zwar der allgemeine Fall und tritt z. B. immer ein, wenn die Eigenwerte alle verschieden sind, doch kommen auch Ausnahmen vor, wie z. B. die Matrizen

$$\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \quad \text{oder} \quad \begin{pmatrix} 1 & i \\ i & -1 \end{pmatrix}$$

zeigen, die sich durch eine Ähnlichkeitstransformation überhaupt nicht auf die Diagonalform bringen lassen. Von solchen Matrizen handelt die Elementarteilertheorie. Wir brauchen indessen auf diese nicht einzugehen, weil wir immer mit solchen Matrizen zu tun haben werden (unitären bzw. hermiteischen Matrizen), die sich nach (6a) auf Diagonalform bringen lassen.

Die Vertauschbarkeitsverhältnisse zweier Matrizen lassen sich von dem soeben gewonnenen Standpunkt aus sehr gut übersehen. Lassen sich zwei Matrizen mit derselben Transformation auf die Diagonalform bringen, haben sie dieselben Eigenvektoren, so sind sie vertauschbar¹⁾. Sie sind es nämlich als Diagonalmatrizen sicher nach der Ähnlichkeitstransformation, also waren sie es auch schon vorher.

Im ersten Kapitel wurden die rationalen Funktionen

$$\begin{aligned} f(\mathbf{a}) = & \cdots a_{-3} \mathbf{a}^{-3} + a_{-2} \mathbf{a}^{-2} + a_{-1} \mathbf{a}^{-1} + a_0 \mathbf{1} + a_1 \mathbf{a} \\ & + a_2 \mathbf{a}^2 + a_3 \mathbf{a}^3 + \cdots \end{aligned}$$

¹⁾ Dagegen können die Eigenwerte beliebig verschieden sein.

einer Matrix α definiert. Um $f(\alpha)$ auf die Diagonalform zu bringen, genügt es, α auf die Diagonalform $\Lambda = \sigma^{-1} \alpha \sigma$ zu transformieren. Dann ist nach Satz 10, Kapitel I

$$\begin{aligned}\sigma^{-1} f(\alpha) \sigma &= \sigma^{-1} (\dots a_{-2} \alpha^{-2} + a_{-1} \alpha^{-1} + a_0 1 + a_1 \alpha + a_2 \alpha^2 + \dots) \sigma, \\ &= \dots a_{-2} \Lambda^{-2} + a_{-1} \Lambda^{-1} + a_0 1 + a_1 \Lambda + a_2 \Lambda^2 + \dots = f(\Lambda)\end{aligned}$$

und diese ist selber eine Diagonalmatrix. Ist λ_k das k -te Diagonalelement von $\Lambda = (\Lambda_{ik}) = (\delta_{ik} \lambda_k)$, so ist es λ_k^e in Λ^e und

$$\dots a_{-2} \lambda_k^{-2} + a_{-1} \lambda_k^{-1} + a_0 + a_1 \lambda_k + a_2 \lambda_k^2 + \dots = f(\lambda_k)$$

in $f(\Lambda)$.

Eine rationale Funktion $f(\alpha)$ einer Matrix α kann durch dieselbe Transformation auf Diagonalform gebracht werden, durch die α selber auf Diagonalform gebracht wird. Die Diagonalelemente, die Eigenwerte von $f(\alpha)$ sind die entsprechenden Funktionen $f(\lambda_1)$, $f(\lambda_2)$, $f(\lambda_3)$, ... der Diagonalelemente λ_1 , λ_2 , λ_3 , ... von α . Man nimmt an, daß dieses Gesetz nicht nur für rationale, sondern für beliebige Funktionen $F(\alpha)$ von α gilt, und betrachtet dies als die Definition der allgemeinen Matrixfunktionen.

Spezielle Matrizen

Es liegt nahe, aus einer quadratischen Matrix α dadurch eine neue α' zu bilden, daß man die Rolle der Zeilen und Spalten vertauscht. Die so entstandene Matrix α' nennt man die Transponierte zu α , die Transposition deutet man mit einem Strich an. Es gilt also

$$\alpha'_{ik} = \alpha_{ki}. \quad (7)$$

Regel. Die Transponierte eines Produkts $\alpha \beta \gamma \delta \dots$ ist das Produkt der Transponierten in umgekehrter Reihenfolge genommen, also

$$(\alpha \beta \gamma \dots \varepsilon)' = \varepsilon' \dots \gamma' \beta' \alpha'. \quad (7a)$$

Es ist nämlich

$$(\alpha \beta \gamma \dots \varepsilon)'_{ki} = (\alpha \beta \gamma \dots \varepsilon)_{ik} = \sum_{x \lambda \mu \dots \zeta} \alpha_{ix} \beta_{x\lambda} \gamma_{\lambda\mu} \dots \varepsilon_{\zeta k}.$$

Andererseits ist

$$(\varepsilon' \dots \gamma' \beta' \alpha')_{ki} = \sum_{\zeta \dots \mu \lambda x} \varepsilon'_{k\zeta} \dots \gamma'_{\mu\lambda} \beta'_{\lambda x} \alpha'_{xi} = \sum_{\zeta \dots \mu \lambda x} \varepsilon_{\zeta k} \dots \gamma_{\lambda\mu} \beta_{\lambda x} \alpha_{xi},$$

womit (7a) verifiziert ist.

Die Matrix, die entsteht, wenn man an Stelle eines jeden der n^2 Koeffizienten von α die konjugiert komplexe Zahl setzt, nennt man α^* , die konjugiert Komplexe zu α . Ist $\alpha = \alpha^*$, so sind alle Koeffizienten reell.

Vertauscht man Zeilen mit Spalten, und geht man außerdem zur konjugiert Komplexen über, bildet man also aus α die Matrix $\alpha^{*'} = \alpha'^*$, so erhält man die Adjungierte α^\dagger zu α :

$$\alpha^{*'} = \alpha^\dagger = \alpha'^*. \quad (8)$$

Die konjugiert Komplexe eines Produkts ist offenbar das Produkt der konjugiert Komplexen:

$$(\alpha \beta \gamma \dots \varepsilon)^* = \alpha^* \beta^* \gamma^* \dots \varepsilon^*;$$

beim Übergang zur Adjungierten muß man die Reihenfolge umkehren:

$$\begin{aligned} (\alpha \beta \gamma \dots \varepsilon)^\dagger &= (\alpha \beta \gamma \dots \varepsilon)^{*'} = (\alpha^* \beta^* \gamma^* \dots \varepsilon^*)' \\ &= \varepsilon^{*'} \dots \gamma^{*'} \beta^{*'} \alpha^{*'} = \varepsilon^\dagger \dots \gamma^\dagger \beta^\dagger \alpha^\dagger. \end{aligned} \quad (8a)$$

Man kann verschiedene Beziehungen zwischen einer Matrix α , ihrer Adjungierten, Transponierten und Reziproken annehmen und erhält dadurch spezielle Arten von Matrizen. Da ihre Namen in der Literatur häufig vorkommen, wollen wir sie der Vollständigkeit halber alle anführen, im folgenden werden außer den unitären und hermiteischen nur die reell orthogonalen gebraucht werden.

Ist $\alpha = \alpha^*$, also $\alpha_{ik} = \alpha_{ik}^*$, so heißt die Matrix reell, alle n^2 Koeffizienten α_{ik} sind reell. Ist $\alpha = -\alpha^*$, also $\alpha_{ik} = -\alpha_{ik}^*$, so ist die Matrix rein imaginär.

Ist $S = S'$, also $S_{ik} = S_{ki}$, so ist die Matrix symmetrisch. Ist $S_{ik} = -S_{ki}$, also $S = -S'$, so nennt man sie schief.

Ist $H = H^\dagger$, also $H_{ik} = H_{ki}^*$, so heißt die Matrix hermiteisch. Ist $\alpha = -\alpha^\dagger$, so nennt man sie hermiteisch schief.

Ist α sowohl reell als auch symmetrisch, so ist α natürlich auch hermiteisch usw.

Ist $O' = O^{-1}$, so ist O komplex orthogonal. Die Matrix U , für die $U^\dagger = U^{-1}$ gilt, heißt unitär. Ist $R^\dagger = R^{-1}$ und $R = R^*$ reell, so ist auch $R' = R^{*'} = R^\dagger = R^{-1}$, also auch $R' = R^{-1}$. Man nennt R reell orthogonal, oder schlechtweg orthogonal.

Unitäre Matrizen und das skalare Produkt

Bevor wir zu den unitären Matrizen übergehen, müssen wir eine letzte Art von Begriffsbildung vornehmen.

Schon im ersten Kapitel haben wir die Summe zweier Vektoren und konstante Vielfache eines Vektors definiert. Bereits in der elementaren Vektorrechnung gebraucht man daneben das skalare Produkt zweier Vektoren. Das skalare Produkt des Vektors a mit dem Vektor b ist eine Zahl, und zwar unterscheiden wir das hermiteische skalare Produkt

$$a_1^* b_1 + a_2^* b_2 + \cdots + a_n^* b_n = (a, b) \quad (\dagger)$$

vom einfachen skalaren Produkt

$$a_1 b_1 + a_2 b_2 + \cdots + a_n b_n = ((a, b)). \quad (\ddagger\ddagger)$$

Wenn wir nichts anderes bemerken, denken wir stets an das hermiteische skalare Produkt und nicht an das einfache.

Sind die Vektorkomponenten a_1, a_2, \dots, a_n reell, so fallen beide Produkte zusammen.

Ist $(a, b) = 0 = (b, a)$, so sagt man, daß a und b senkrecht aufeinander stehen, oder auch, daß sie orthogonal sind. Ist $(a, a) = 1$, so sagt man, daß a ein Einheitsvektor ist, oder daß er normiert ist. Sonst ist (a, a) immer reell und positiv und nur dann 0, wenn alle Komponenten von a verschwinden. Dies gilt nur für das hermiteische skalare Produkt, dagegen nicht für das einfache: ist a z. B. der zweidimensionale Vektor $1, i$, so ist $((a, a)) = 0$, aber $(a, a) = 2$. Aus $(a, a) = 0$ folgt $a = 0$, aus $((a, a)) = 0$ dagegen nicht.

Einfache Regeln. 1. Vertauschen der Vektoren

$$(a, b) = (b, a)^*, \quad (9)$$

dagegen $((a, b)) = ((b, a))$.

2. Ist c eine Zahl, so gilt

$$(a, c b) = c (a, b); \quad ((a, c b)) = c ((a, b)). \quad (9a)$$

Dagegen ist

$$(c a, b) = c^* (a, b); \quad ((c a, b)) = c (a, b). \quad (9b)$$

3. Das skalare Produkt ist linear, da noch

$$(a, b + c) = (a, b) + (a, c) \quad (9c)$$

4. Es gilt weiter die wichtige Regel

$(\mathbf{a}, \mathbf{a} \mathbf{b}) = (\mathbf{a}^\dagger \mathbf{a}, \mathbf{b})$ oder $(\boldsymbol{\beta} \mathbf{a}, \mathbf{b}) = (\mathbf{a}, \boldsymbol{\beta}^\dagger \mathbf{b})$ (9d)
für beliebige Vektoren \mathbf{a} und \mathbf{b} und jede Matrix \mathbf{a} . Es ist nämlich

$$(\mathbf{a}, \mathbf{a} \mathbf{b}) = \sum_{k=1}^n \mathbf{a}_k^* (\mathbf{a} \mathbf{b})_k = \sum_{k=1}^n \mathbf{a}_k^* \sum_{\lambda=1}^n \cdot \mathbf{a}_{k\lambda} \mathbf{b}_\lambda$$

und

$$(\mathbf{a}^\dagger \mathbf{a}, \mathbf{b}) = \sum_{\lambda=1}^n (\mathbf{a}^\dagger \mathbf{a})_\lambda^* \mathbf{b}_\lambda = \sum_{\lambda=1}^n \sum_{k=1}^n (\mathbf{a}_{k\lambda}^* \mathbf{a}_k)^* \mathbf{b}_\lambda = \sum_{\lambda=1}^n \sum_{k=1}^n \mathbf{a}_{k\lambda} \mathbf{a}_k^* \mathbf{b}_\lambda.$$

Statt einer Matrix \mathbf{a} auf den einen Faktor kann man auch ihre Adjungierte \mathbf{a}^\dagger auf den anderen Faktor des skalaren Produkts anwenden.

Bei dem einfachen skalaren Produkt gilt dasselbe von der transponierten Matrix: es ist

$$((\mathbf{a}, \mathbf{a} \mathbf{b})) = ((\mathbf{a}' \mathbf{a}, \mathbf{b})).$$

5. Schreiben wir jetzt die Bedingung $\mathbf{U}^\dagger = \mathbf{U}^{-1}$ für die Unitarität der Matrix \mathbf{U} etwas mehr explizit hin! Es ist $\mathbf{U}^\dagger \mathbf{U} = \mathbf{1}$, also

$$\sum_{j=1}^n (\mathbf{U}^\dagger)_{ij} \mathbf{U}_{jk} = \sum_{j=1}^n \mathbf{U}_{ji}^* \mathbf{U}_{jk} = \delta_{ik}; (\mathbf{U}_{\cdot i}, \mathbf{U}_{\cdot i}) = \delta_{ii}. \quad (10)$$

Faßt man die n Spalten einer unitären Matrix als Vektoren auf, so bilden sie n zueinander senkrechte Einheitsvektoren.

Aus $\mathbf{U} \mathbf{U}^\dagger = \mathbf{1}$ folgt

$$\sum_j \mathbf{U}_{ij} \mathbf{U}_{kj}^* = \delta_{ik}; (\mathbf{U}_{\cdot i}, \mathbf{U}_{\cdot i}) = \delta_{ii}. \quad (10a)$$

Auch die n Zeilen einer unitären Matrix sind n zueinander senkrechte Einheitsvektoren.

6. Die unitäre Transformation läßt das hermiteische innere Produkt unverändert, d. h. es gilt für beliebige Vektoren \mathbf{a} und \mathbf{b}

$$(\mathbf{U} \mathbf{a}, \mathbf{U} \mathbf{b}) = (\mathbf{a}, \mathbf{U}^\dagger \mathbf{U} \mathbf{b}) = (\mathbf{a}, \mathbf{b}). \quad (11)$$

Gilt umgekehrt (11) für die Matrix \mathbf{U} in Verbindung mit jedem beliebigen Vektor \mathbf{a} und \mathbf{b} , so ist \mathbf{U} unitär. Dann gilt nämlich (11) auch für $\mathbf{a} = \mathbf{e}_i$ und $\mathbf{b} = \mathbf{e}_l$ (mit $\mathbf{e}_{il} = \delta_{il}$) und aus (11) wird

$$\begin{aligned} \delta_{il} &= (\mathbf{e}_i, \mathbf{e}_l) = (\mathbf{U} \mathbf{e}_i, \mathbf{U} \mathbf{e}_l) = \sum_j (\mathbf{U} \mathbf{e}_i)_j^* (\mathbf{U} \mathbf{e}_l)_j \\ &= \sum_j (\sum_l \mathbf{U}_{jl} \delta_{ik})^* \cdot \sum_l \mathbf{U}_{jl} \delta_{il} = \sum_j \mathbf{U}_{ji}^* \mathbf{U}_{jl}, \end{aligned}$$

also (10): (11) ist hinreichend und notwendig für die Unitarität von \mathbf{U} .

Dasselbe gilt für komplexorthogonale Matrizen in Verbindung mit dem einfachen skalaren Produkt.

7. Das Produkt $\mathbf{U} \mathbf{V}$ von zwei unitären Matrizen \mathbf{U} und \mathbf{V} ist unitär.

$$(\mathbf{U} \mathbf{V})^\dagger = \mathbf{V}^\dagger \mathbf{U}^\dagger = \mathbf{V}^{-1} \mathbf{U}^{-1} = (\mathbf{U} \mathbf{V})^{-1}. \quad (12)$$

Dasselbe gilt von der Reziproken \mathbf{U}^{-1} einer unitären Matrix:

$$(\mathbf{U}^{-1})^\dagger = (\mathbf{U}^\dagger)^\dagger = \mathbf{U} = (\mathbf{U}^{-1})^{-1}. \quad (13)$$

Hauptachsentransformation unitärer und hermiteischer Matrizen

Man kann jede unitäre und jede hermiteische Matrix durch eine Ähnlichkeitstransformation mit einer unitären Matrix auf die Diagonalform bringen. Die auf S. 24 erwähnten Ausnahmefälle können bei diesen Matrizen nicht vorkommen. Zunächst bemerken wir, daß eine Matrix unitär bzw. hermiteisch bleibt, wenn man sie unitär transformiert. Als Produkt von drei unitären Matrizen ist $\mathbf{U}^{-1} \mathbf{V} \mathbf{U}$ selber unitär. Aber auch $\mathbf{U}^{-1} \mathbf{H} \mathbf{U}$ bleibt hermiteisch, wenn \mathbf{H} hermiteisch war. In der Tat ist wegen (13)

$$(\mathbf{U}^{-1} \mathbf{H} \mathbf{U})^\dagger = \mathbf{U}^\dagger \mathbf{H} \mathbf{U}^{-1\dagger} = \mathbf{U}^\dagger \mathbf{H} \mathbf{U}. \quad (14)$$

Um \mathbf{V} bzw. \mathbf{H} auf die Diagonalform zu bringen, bestimmen wir einen Eigenwert von \mathbf{V} bzw. \mathbf{H} . Dieser sei λ_1 , der zugehörige Eigenvektor, der nur bis auf einen konstanten Faktor bestimmt ist, $\mathbf{U}_{.1}(\mathbf{U}_{11}, \mathbf{U}_{21}, \dots, \mathbf{U}_{n1})$. Wir nehmen den konstanten Faktor nun so groß an, daß

$$(\mathbf{U}_{.1}, \mathbf{U}_{.1}) = 1$$

sei, was immer möglich ist, da $(\mathbf{U}_{.1}, \mathbf{U}_{.1})$ ja nicht verschwinden kann. Jetzt bilden wir eine unitäre Matrix \mathbf{U} , deren erste Spalte $\mathbf{U}_{.1}$ ist¹⁾. Mit dieser unitären Matrix \mathbf{U} transformieren wir nun \mathbf{V} bzw. \mathbf{H} zu $\mathbf{U}^{-1} \mathbf{V} \mathbf{U}$ bzw. $\mathbf{U}^{-1} \mathbf{H} \mathbf{U}$. In die erste Spalte kommt dabei im ersten Falle

$$\begin{aligned} X_{r1} = (\mathbf{U}^{-1} \mathbf{V} \mathbf{U})_{r1} &= (\mathbf{U}^\dagger \mathbf{V} \mathbf{U})_{r1} = \sum_v \mathbf{U}_{vr}^* \sum_\mu V_{v\mu} \mathbf{U}_{\mu 1} \\ &= \sum_v \mathbf{U}_{vr}^* \lambda_1 \mathbf{U}_{v1} = \delta_{r1} \lambda_1, \end{aligned}$$

da $\mathbf{U}_{.1}$ ja ein Eigenvektor von \mathbf{V} ist.

In die erste Zeile der ersten Spalte kommt λ_1 , sonst ist die erste Spalte leer.

¹⁾ Siehe Hilfssatz am Ende des Beweises.

30 III. Hauptachsentransformation unitärer und hermiteischer Matrizen

Dies gilt offenbar nicht nur für $\mathbf{U}^{-1} \mathbf{V} \mathbf{U}$, sondern entsprechend auch für $\mathbf{U}^{-1} \mathbf{H} \mathbf{U}$. In diesem ist sogar — da $\mathbf{U}^{-1} \mathbf{H} \mathbf{U}$ hermiteisch ist — auch die erste Zeile bis auf die erste Stelle leer, $\mathbf{U}^{-1} \mathbf{H} \mathbf{U}$ hat die Gestalt

$$\begin{pmatrix} \lambda_1 & 0 & 0 & \dots & 0 \\ 0 & & & & \\ 0 & & & & \\ \vdots & & & & \\ 0 & & & & \end{pmatrix} \quad (\Theta)$$

Genau so muß aber $\mathbf{U}^{-1} \mathbf{V} \mathbf{U} = \mathbf{X}$ aussehen! Da \mathbf{X} eine unitäre Matrix ist, ist seine erste Spalte $\mathbf{X}_{\cdot 1}$ ein Einheitsvektor, und daraus folgt, daß

$$|\mathbf{X}_{11}|^2 + |\mathbf{X}_{21}|^2 + \dots + |\mathbf{X}_{n1}|^2 = |\lambda_1|^2 = 1 \quad (*)$$

ist. Dasselbe muß aber von der ersten Zeile \mathbf{X}_1 von \mathbf{X} gelten,

$$|\mathbf{X}_{11}|^2 + |\mathbf{X}_{12}|^2 + \dots + |\mathbf{X}_{1n}|^2 = |\lambda_1|^2 + |\mathbf{X}_{12}|^2 + |\mathbf{X}_{13}|^2 + \dots + |\mathbf{X}_{1n}|^2 = 1,$$

so daß die Summe der Absolutwertquadrate von $\mathbf{X}_{12}, \mathbf{X}_{13}, \dots, \mathbf{X}_{1n}$ verschwindet, woraus auch das Verschwinden von $\mathbf{X}_{12}, \mathbf{X}_{13}, \dots, \mathbf{X}_{1n}$ folgt.

Jede unitäre bzw. hermiteische Matrix läßt sich unitär auf die Gestalt \oplus transformieren!

Nun ist \oplus noch keine Diagonalmatrix, was sie ja auch nicht sein kann, da wir ja nur die Existenz eines Eigenwertes benutzt haben. Sie ist aber einer Diagonalmatrix schon ähnlicher als die ursprüngliche Matrix \mathbf{V} oder \mathbf{H} . Es liegt nahe, \oplus als Übermatrix zu schreiben, etwa

$$\begin{pmatrix} \lambda_1 & \mathbf{0} \\ \mathbf{0} & \mathbf{V}_1 \end{pmatrix} \text{ bzw. } \begin{pmatrix} \lambda_1 & \mathbf{0} \\ \mathbf{0} & \mathbf{H}_1 \end{pmatrix}, \quad \oplus \oplus$$

wo die Matrix \mathbf{V}_1 bzw. \mathbf{H}_1 nur noch $n - 1$ Zeilen und Spalten hat. Dann können wir $\oplus \oplus$ durch eine weitere unitäre Matrix $\mathbf{U}_{(1)}$ von der Gestalt

$$\begin{pmatrix} 1 & \mathbf{0} \\ \mathbf{0} & \mathbf{U}_1 \end{pmatrix}$$

transformieren, wobei \mathbf{U}_1 nur $n - 1$ Zeilen und Spalten hat. Dabei geht $\oplus \oplus$ in

$$\begin{pmatrix} \lambda_1 & \mathbf{0} \\ \mathbf{0} & \mathbf{U}_1^\dagger \mathbf{V}_1 \mathbf{U}_1 \end{pmatrix} \text{ bzw. } \begin{pmatrix} \lambda_1 & \mathbf{0} \\ \mathbf{0} & \mathbf{U}_1^\dagger \mathbf{H}_1 \mathbf{U}_1 \end{pmatrix} \quad \oplus \oplus \oplus$$

über. Darin kann man aber nach dem Vorangehenden \mathbf{U}_1 so wählen, daß $\mathbf{U}_1^\dagger \mathbf{V}_1 \mathbf{U}_1$ bzw. $\mathbf{U}_1^\dagger \mathbf{H}_1 \mathbf{U}_1$ die Gestalt

$$\begin{pmatrix} \lambda_2 & 0 \\ 0 & \mathbf{V}_2 \end{pmatrix} \text{ bzw. } \begin{pmatrix} \lambda_2 & 0 \\ 0 & \mathbf{H}_2 \end{pmatrix}$$

hat, wo \mathbf{V}_2 bzw. \mathbf{H}_2 nur noch $n - 2$ -dimensional ist. Jetzt hat $\mathbf{U}_{(1)}^\dagger \mathbf{U}^\dagger \mathbf{V} \mathbf{U} \mathbf{U}_1$ die Gestalt

$$\begin{pmatrix} \Delta_1 & 0 \\ 0 & \mathbf{V}_2 \end{pmatrix} \text{ mit } \Delta_1 = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}.$$

Durch Fortsetzen dieses Verfahrens gelingt es offenbar, \mathbf{V} oder \mathbf{H} auch ganz auf die Diagonalform zu bringen.

Dieser Satz gilt nicht für symmetrische bzw. komplex orthogonale Matrizen, wie schon das zweite Beispiel auf S. 24 zeigt (die zweite Matrix ist symmetrisch und komplex orthogonal). Dagegen gilt er für reelle symmetrische oder reelle orthogonale Matrizen, die ja nur einen Spezialfall der hermiteischen bzw. unitären bilden.

Hilfssatz. Ist $(\mathbf{u}_{.1}, \mathbf{u}_{.1}) = 1$, so läßt sich (auf mannigfach verschiedene Weise) eine unitäre Matrix bilden, deren erste Spalte $\mathbf{u}_{.1} (\mathbf{u}_{11}, \mathbf{u}_{21}, \mathbf{u}_{31}, \dots, \mathbf{u}_{n1})$ ist.

Bilden wir nämlich zuerst überhaupt eine Matrix, deren erste Spalte $\mathbf{u}_{.1}$ ist und deren Determinante nicht verschwindet. Die zweite Spalte dieser Matrix sei $\mathbf{v}_{.2} (\mathbf{v}_{12}, \mathbf{v}_{22}, \mathbf{v}_{32}, \dots, \mathbf{v}_{n2})$, die dritte $\mathbf{v}_{.3}$ usw.

$$\begin{pmatrix} \mathbf{u}_{11} & \mathbf{v}_{12} & \mathbf{v}_{13} & \dots & \mathbf{v}_{1n} \\ \mathbf{u}_{21} & \mathbf{v}_{22} & \mathbf{v}_{23} & \dots & \mathbf{v}_{2n} \\ \mathbf{u}_{31} & \mathbf{v}_{32} & \mathbf{v}_{33} & \dots & \mathbf{v}_{3n} \\ \vdots & \vdots & \vdots & & \vdots \\ \mathbf{u}_{n1} & \mathbf{v}_{n2} & \mathbf{v}_{n3} & \dots & \mathbf{v}_{nn} \end{pmatrix}$$

Die Vektoren $\mathbf{u}_{.1}, \mathbf{v}_{.2}, \mathbf{v}_{.3}, \dots$ sind dann voneinander linear unabhängig. Wollen wir noch, daß sie gemäß (9) zueinander senkrecht stehen sollen, so müssen wir sie „orthogonalisieren“. Wir ersetzen zunächst $\mathbf{v}_{.2}$ durch $\mathbf{u}_{.2} = a_{21} \mathbf{u}_{.1} + \mathbf{v}_{.2}$. Dann setzen wir (Schmidtsches Orthogonalisierungsverfahren)

$(\mathbf{u}_{.1}, \mathbf{u}_{.2}) = 0 = a_{21}(\mathbf{u}_{.1}, \mathbf{u}_{.1}) + (\mathbf{u}_{.1}, \mathbf{v}_{.2}) = a_{21} + (\mathbf{u}_{.1}, \mathbf{v}_{.2})$ und bestimmen daraus a_{21} . Dann setzen wir $\mathbf{u}_{.3}$ an Stelle von $\mathbf{v}_{.3}$ mit $\mathbf{u}_{.3} = a_{31} \mathbf{u}_{.1} + a_{32} \mathbf{u}_{.2} + \mathbf{v}_{.3}$ und bestimmen a_{31} und a_{32} so, daß

$$0 = (\mathbf{u}_{.1}, \mathbf{u}_{.3}) = a_{31}(\mathbf{u}_{.1}, \mathbf{u}_{.1}) + (\mathbf{u}_{.1}, \mathbf{v}_{.3}),$$

$$0 = (\mathbf{u}_{.2}, \mathbf{u}_{.3}) = a_{32}(\mathbf{u}_{.2}, \mathbf{u}_{.2}) + (\mathbf{u}_{.2}, \mathbf{v}_{.3}),$$

so gehen wir weiter und setzen schließlich $u_{\cdot n}$ an Stelle von $v_{\cdot n}$ mit $u_{\cdot n} = a_{n1} u_{\cdot 1} + a_{n2} u_{\cdot 2} + a_{n3} u_{\cdot 3} + \dots + v_{\cdot n}$ und bestimmen $a_{n1}, a_{n2}, a_{n3}, \dots, a_{n(n-1)}$, so daß

$$\begin{aligned} 0 &= (u_{\cdot 1}, u_{\cdot n}) = a_{n1} (u_{\cdot 1} u_{\cdot 1}) + (u_{\cdot 1}, v_{\cdot n}), \\ 0 &= (u_{\cdot 1}, u_{\cdot n}) = a_{n2} (u_{\cdot 2} u_{\cdot 2}) + (u_{\cdot 2}, v_{\cdot n}), \\ &\quad \dots \\ 0 &= (u_{\cdot (n-1)}, u_{\cdot n}) = a_{n(n-1)} (u_{\cdot (n-1)} u_{\cdot (n-1)}) + (u_{\cdot (n-1)}, v_{\cdot n}). \end{aligned}$$

So haben wir mit Hilfe der $\frac{1}{2}n(n-1)$ Zahlen a erreicht, daß an Stelle der Vektoren v Vektoren u treten, die aufeinander senkrecht stehen und dabei wegen der linearen Unabhängigkeit der v nicht verschwinden können. In der Tat würde z. B. aus $u_{\cdot n} = 0$

$$a_{n1} u_{\cdot 1} + a_{n2} u_{\cdot 2} + \dots + a_{n(n-1)} u_{\cdot (n-1)} + v_{\cdot n} = 0$$

folgen, und da die $u_{\cdot 1} \dots u_{\cdot (n-1)}$ Linearkombinationen der $u_{\cdot 1}, u_{\cdot 2}, \dots, u_{\cdot (n-1)}$ sind, könnte man $v_{\cdot n}$ durch diese ausdrücken, was natürlich nicht sein kann.

Wir können noch die $u_{\cdot 2}, u_{\cdot 3}, \dots, u_{\cdot n}$ normieren, dann bilden sie eine unitäre Matrix, deren erste Spalte $u_{\cdot 1}$ ist.

Dieses sogenannte Schmidtsche Orthogonalisierungsverfahren lehrt also, aus einer Schar linear unabhängiger Vektoren normierte, zueinander senkrechte, zu bilden, wobei der k -te Einheitsvektor eine Linearkombination nur der ersten k Vektoren ist.

Geht man von n n -dimensionalen Vektoren aus einem vollständigen Vektorensystem aus, so erhält man ein vollständiges Orthogonalsystem.

Ist V bzw. H auf diese Weise auf Diagonalform gebracht, so ist die entstandene Matrix Λ_v bzw. Λ_h auch weiter unitär bzw. hermiteisch. Es folgt

$$\Lambda_v \Lambda_v^\dagger = 1, \quad \Lambda_h = \Lambda_h^\dagger. \quad (15)$$

Der Absolutwert der Eigenwerte einer unitären Matrix¹⁾ ist 1, die Eigenwerte einer hermiteischen Matrix sind reell. In der Tat steht in (15), daß für die Eigenwerte λ_v der unitären Matrix $\lambda_v \lambda_v^* = 1$, für die einer hermiteischen $\lambda_h = \lambda_h^*$ gilt.

¹⁾ Dies zeigt schon (*).

Reelle orthogonale und symmetrische Matrizen

Zuletzt untersuchen wir noch, was es bedeuten würde, wenn V bzw. H nicht nur unitär bzw. hermiteisch, sondern außerdem noch komplex orthogonal bzw. symmetrisch, also beide rein reell wären.

Aus $U^\dagger V U = \Lambda_v$ erhalten wir durch Übergang zum konjugiert komplexen $U^{*\dagger} V^* U^* = (U^*)^\dagger V U^* = \Lambda_v^*$; die Diagonalform Λ_v kann auch Λ_v^* geschrieben werden, die Zahlen $\lambda_1, \lambda_2, \dots, \lambda_n$ sind dieselben wie die Zahlen $\lambda_1^*, \lambda_2^*, \dots, \lambda_n^*$, weil die Eigenwerte als Wurzeln der Säkulargleichung unabhängig davon sind, wie (durch U oder U^*) man die Matrix auf Diagonalform bringt.

Die komplexen Eigenwerte einer reellen orthogonalen Matrix sind paarweise konjugiert komplex. Außerdem haben alle den Absolutwert 1, die reellen sind also ± 1 . Bei ungerader Dimensionszahl muß wenigstens ein Eigenwert reell sein.

Ist v der Eigenvektor zum Eigenwert λ , so ist v^* der Eigenvektor des konjugiert komplexen Eigenwerts λ^* . Es sei nämlich $Vv = \lambda v$, so ist auch $V^*v^* = Vv^* = \lambda^*v^*$. Ist λ von λ^* verschieden, so ist $(v^*, v) = 0 = ((v, v))$. Das einfache skalare Produkt eines Eigenvektors mit sich selber verschwindet, wenn der zugehörige Eigenwert nicht reell (also ± 1) ist. Zu den Eigenwerten ± 1 gehören dagegen reelle Eigenvektoren. Weiter sei v der Eigenvektor zu λ_1 , v^* zu λ_1^* und z zu λ_2 . Ist dann $\lambda_1^* \neq \lambda_2$, so ist $0 = (v^*, z) = ((v, z))$.

Das einfache skalare Produkt von zwei Eigenvektoren einer reellen orthogonalen Matrix ist immer Null, wenn die zugehörigen Eigenwerte nicht konjugiert komplex sind, in diesem Falle sind sie auch selber konjugiert komplex zueinander.

Die Determinante einer orthogonalen Matrix ist ± 1 . Aus $VV' = 1$ folgt, daß die Determinante von V mit der von V' multipliziert 1 ergeben muß, die Determinante von V' ist aber gleich der von V und beide müssen ± 1 sein.

Ist H reell, so ist die Gleichung (5) rein reell, da auch die λ reell sind: die Eigenvektoren einer reellen hermiteischen Matrix können reell angenommen werden (sie können, da sie nur bis auf eine Konstante bestimmt sind, mit einem komplexen Faktor multipliziert werden); die unitäre Matrix U in $U^{-1}H\bar{U} = \Lambda_h$ kann reell angenommen werden.

IV. Grundlagen der Quantenmechanik

1. Die Entwicklung der Quantentheorie in den letzten Jahren, die man unter dem Namen „Quantenmechanik“ zusammenzufassen pflegt, war zuerst auf die Bestimmung der Energie der stationären Zustände, die Berechnung der Energieniveaus, gerichtet. In der älteren Theorie von Epstein-Schwarzschild, der sogenannten Separationstheorie, war nur eine Vorschrift für die Bestimmung der Energieniveaus der Terme solcher Systeme gegeben, deren klassisch-mechanische Bewegungen gewisse sehr spezielle Eigenschaften hatten, streng periodisch oder doch quasiperiodisch waren.

Diesem Übel half eine Idee von W. Heisenberg ab, die eine Präzisierung des Bohrschen Korrespondenzprinzips zum Ziele hatte. Sie wurde von M. Born und P. Jordan einerseits, von P. A. M. Dirac andererseits ausgebaut. Den Gedanken kann man kurz dahin zusammenfassen, daß man von vornherein nur solche Bewegungen betrachtet und nur mit solchen rechnet, die später dann als quantentheoretisch erlaubte Bewegungen angesehen werden. Die Durchführung dieser Idee führte dann die erwähnten Verfasser auf die Einführung von Matrizen mit unendlich vielen Zeilen und Spalten, durch die man formal Lagen und Impulskoordinaten repräsentieren konnte, bzw. auf das formale Rechnen mit „*q*-Zahlen“, die nur dem Assoziativgesetz, nicht aber dem Kommutativgesetz gehorchen. So lautet z. B. die Gleichung für die Energie ***H*** des linearen Oszillators¹⁾

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

man erhält sie, indem man im klassischen Ausdruck für die Energie, in der sogenannten Hamiltonschen Funktion, *p* und *q* — die Impuls- und Lagenkoordinaten — formal durch die Matrizen ***p*** und ***q*** ersetzt. Es wird verlangt, daß ***H*** eine Diagonalmatrix sei. Die Diagonalglieder ***H***_{*n,n*} bedeuten dann die möglichen Energiewerte, die Terme des Systems. Das Matrixelement ***q***_{*n,k*} von ***q*** bedeutet dagegen (bis auf eine Proportionalitätskonstante), zum Quadrat erhoben, dem Absolutwert nach die spon-

1) *m* ist die Masse des oszillierenden Teilchens, *x*² die Proportionalitätskonstante der elastischen Kraft, ***q*** und ***p*** sind Lagen- und Impulskoordinaten.

tane Übergangswahrscheinlichkeit vom Zustand mit der Energie H_{nn} zum Zustand mit der Energie H_{kk} , also die Intensität der Linie mit der Frequenz $\frac{H_{nn} - H_{kk}}{\hbar}$. Alles dies folgte aus denselben Überlegungen, die die Einführung von Matrizen für p und q nahegelegt haben.

Um das Problem bestimmt zu machen, mußte noch eine „Vertauschungsrelation“ zwischen p und q eingeführt werden, diese wurde zu

$$pq - qp = \frac{\hbar}{2\pi i} 1 \quad (2)$$

angenommen, wobei gleichzeitig die Plancksche Konstante \hbar eingeführt werden konnte.

Das, oft recht mühselige, Rechnen mit diesen Größen führte sehr bald zu überraschend schönen und wichtigen Resultaten. So konnten die sogenannten Auswahlregeln für die Drehimpulse und gewisse „Summensätze“, die die relative Intensität von Zeeman-Komponenten einer Linie bestimmten, übereinstimmend mit der Erfahrung berechnet werden, während die Separationstheorie für diese Regeln keine hinreichende Erklärung geben konnte.

Ganz unabhängig von der Heisenbergschen Richtung kam E. Schrödinger zu mathematisch vollkommen äquivalenten Resultaten, seine Gedankengänge zeigen eher mit denen von L. de Broglie weitgehende Ähnlichkeit.

Betrachten wir einen mehrdimensionalen Raum mit so viel Koordinaten, wie das ins Auge gefaßte System Lagenkoordinaten hat! Jeder Lage, jeder momentanen Konfiguration des Systems entspricht ein Punkt in diesem mehrdimensionalen „Konfigurationsraume“. Dieser Punkt wird sich im Laufe der Zeit bewegen, eine Bahn beschreiben, durch die die Bewegung des Systems klassisch vollkommen beschrieben werden kann. Es besteht nun eine weitgehende Analogie zwischen der klassischen Bewegung dieses Punktes, des Systempunktes im Konfigurationsraum und der Bewegung einer ebenfalls in den Konfigurationsraum hineingedachten Wellengruppe,

wenn man den Brechungsindex für diese Wellen zu $\frac{\sqrt{2m(E-V)}}{E}$ annimmt¹⁾). Diese Analogie besteht darin, daß die Bahn des

¹⁾ E ist die Gesamtenergie des Systems, V die potentielle Energie als Funktion der Konfiguration.

Systempunktes im Konfigurationsraum durch die Wellengruppe um so genauer wiedergegeben werden kann, je kleiner die Wellenlänge im Verhältnis zur Bahnkrümmung ist. Wenn dagegen die Wellenlänge der die Wellengruppe bildenden Wellen vergleichbar mit der Krümmung der klassischen Bahn des Konfigurationspunktes wird, so kommen — durch Interferenzerscheinungen der Wellen — wesentliche Abweichungen der beiden Bewegungen zum Vorschein.

Schrödinger nimmt nun an, daß die wirkliche Bewegung der Konfigurationswellen und nicht der klassisch berechneten Bewegung des Konfigurationspunktes entspricht.

Die Wellengleichung für diese Wellen lautet zunächst, wenn man die skalare Amplitude mit ψ bezeichnet,

$$\frac{E - V}{E^2} \frac{\partial^2 \psi}{\partial t^2} = \frac{1}{2m_1} \frac{\partial^2 \psi}{\partial x_1^2} + \frac{1}{2m_2} \frac{\partial^2 \psi}{\partial x_2^2} + \cdots + \frac{1}{2m_f} \frac{\partial^2 \psi}{\partial x_f^2}, \quad (3)$$

wo x_1, x_2, \dots, x_f die Lagenkoordinaten der im System vorhandenen Teilchen, m_1, m_2, \dots, m_f die zugehörigen Massen und $V(x_1, x_2, \dots, x_f)$ die potentielle Energie der Lage bedeutet, in der die Koordinaten der einzelnen Teilchen x_1, x_2, \dots, x_f sind. In (3) kommt noch die Gesamtenergie des Systems explizite vor. Über die Frequenz, die zeitliche Periode der Wellen, kann dagegen noch frei verfügt werden. Schrödinger nimmt nun an, daß die Frequenz ν der Wellen, die mit einer Bewegung des Systems, mit der Gesamtenergie E , verknüpft sind, durch $\hbar\nu = E$ gegeben ist. Er führt daher in (3)

$$\psi = \psi_E e^{\frac{2\pi i}{\hbar} Et} \quad (4)$$

ein, wo ψ_E unabhängig von t ist. So erhält er die Eigenwertgleichung

$$\begin{aligned} \Delta \psi_E &= \frac{1}{2m_1} \frac{\partial^2 \psi_E}{\partial x_1^2} + \frac{1}{2m_2} \frac{\partial^2 \psi_E}{\partial x_2^2} + \cdots + \frac{1}{2m_f} \frac{\partial^2 \psi_E}{\partial x_f^2} \\ &= -\frac{4\pi^2}{\hbar^2} (E - V) \psi_E, \end{aligned} \quad \left. \right\} \quad (5)$$

wo ψ_E eine Funktion der Variablen x_1, x_2, \dots, x_f , der Lagenkoordinaten der das System bildenden Teilchen ist. Es wird verlangt, daß ψ_E quadratintegrierbar sei, d. h. daß das Integral

$$\int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} |\psi_E(x_1, x_2, \dots, x_f)|^2 dx_1 dx_2 \cdots dx_f,$$

erstreckt über den ganzen Konfigurationsraum, konvergiere. (Insbesondere muß also ψ im Unendlichen verschwinden.) Die Werte E , bei denen eine solche Bestimmung der Funktion ψ_E möglich ist, nennt man die Eigenwerte von (5), sie ergeben die möglichen Energiewerte des Systems. Die dazugehörige quadratintegrierbare Lösung von (5) nennt man die zu E gehörige Eigenfunktion.

Man pflegt (5) auch in der Gestalt

$$\mathbf{H} \psi_E = E \psi_E \quad (5a)$$

zu schreiben, wo \mathbf{H} ein linearer Operator (der Hamiltonsche oder Energieoperator):

$$\mathbf{H} = -\frac{\hbar^2}{4\pi^2} \left(\frac{1}{2m_1} \frac{\partial^2}{\partial x_1^2} + \frac{1}{2m_2} \frac{\partial^2}{\partial x_2^2} + \cdots + \frac{1}{2m_f} \frac{\partial^2}{\partial x_f^2} \right) + V(x_1 x_2 \dots x_f) \quad (5b)$$

ist. Er führt eine Funktion von x_1, x_2, \dots, x_f in eine andere Funktion über.

Die Funktion ψ von (4) erfüllt die Relation

$$\frac{\hbar}{2\pi i} \frac{\partial \psi}{\partial t} = \mathbf{H} \psi, \quad (6)$$

in der die Gesamtenergie des Systems nicht mehr explizite auftritt und die daher für alle Bewegungen — gleichgültig, welche Energie das System dabei hat — allgemein gilt. Man nennt (6) die zeitabhängige Schrödingergleichung.

Die beiden Gleichungen (5) [bzw. (5a), (5b)] und (6) sind die Grundgleichungen der Quantenmechanik. Die letztere gibt an, wie sich eine Konfigurationswelle — der man, wie wir sehen werden, eine weitgehende physikalische Realität zuschreibt — im Laufe der Zeit verändert, (5) bzw. (5a), (5b) ist die Gleichung für die Frequenz $\nu = E/\hbar$, die Energie E und die Wellenfunktion ψ_E der zeitlich periodischen Vorgänge. In der Tat erhält man (5a), (5b), wenn man für ψ in (6) den Ansatz

$$\psi = \psi_E e^{\frac{2\pi i}{\hbar} Et}$$

macht.

Wir wollen hier noch die wichtigsten Eigenschaften der Eigenfunktionen und Eigenwerte von (5b) zusammenfassen.

2. Zu diesem Zwecke definieren wir zuerst das skalare Produkt zweier Funktionen. Es sei

$$(\varphi, g) = \int_{-\infty}^{\infty} \dots \int \varphi(x_1 \dots x_f)^* g(x_1 \dots x_f) dx_1 \dots d x_f = \int \varphi^* g. \quad (7)$$

Dann gelten für dieses skalare Produkt alle einfachen Rechenregeln des dritten Kapitels, also (a, b Konstanten)

$$(\varphi, a_1 g_1 + a_2 g_2) = a_1 (\varphi, g_1) + a_2 (\varphi, g_2),$$

$$(\varphi, g) = (g, \varphi)^*.$$

$(\varphi, \varphi) = (\varphi, \varphi)^*$ reell positiv verschwindet nur für $\varphi = 0$. Ist $(\varphi, \varphi) = 1$, so sagt man, daß φ normiert ist. Ist das Integral

$$(\varphi, \varphi) = \int_{-\infty}^{\infty} \dots \int |\varphi(x_1 \dots x_f)|^2 dx_1 \dots d x_f = c^2$$

endlich, so läßt sich φ immer durch Multiplikation mit einer Konstanten (im vorliegenden Falle $1/c$) normieren. In der Tat ist $(\varphi/c, \varphi/c) = 1$. Zwei Funktionen sind orthogonal, wenn ihr skalares Produkt Null ist.

Man wird auf die in (7) gegebene Definition des skalaren Produkts geführt, wenn man die Funktionen $\varphi(x_1 \dots x_f)$, $g(x_1 \dots x_f)$ von $x_1 \dots x_f$ als Vektoren auffaßt, deren Komponenten nach den f kontinuierlich veränderlichen Indizes $x_1 \dots x_f$ benannt sind. Der Funktionenvektor $\varphi(x_1 \dots x_f)$ ist in einem f -fach unendlich viele-dimensionalen Raum definiert: Jedem Wertesystem von $x_1 \dots x_f$, jeder Konfiguration entspricht eine Dimension. Das skalare Produkt von φ und g ist dann in der Vektorsprache

$$(\varphi, g) = \sum_{x_1 \dots x_f} \varphi(x_1 \dots x_f)^* g(x_1 \dots x_f),$$

das durch das Integral (7) ersetzt worden ist.

Im Einklang hiermit steht auch die Definition der linearen Abhängigkeit bzw. Unabhängigkeit von Funktionen.

Es besteht zwischen $\varphi_1, \varphi_2, \dots, \varphi_k$ eine lineare Beziehung

$$a_1 \varphi_1 + a_2 \varphi_2 + \dots + a_k \varphi_k = 0,$$

wenn diese Gleichung bei konstantem a_1, a_2, \dots, a_k für alle Komponenten, alle Wertesysteme der x_1, x_2, \dots, x_f gilt. Weiter heiße ein Operator H linear, wenn für alle Funktionen φ und g die Gleichungen

$$H(a \varphi + b g) = a H \varphi + b H g \quad (8)$$

gelten. Wir werden uns überhaupt nur mit linearen Operatoren beschäftigen. Die linearen Operatoren der Funktionenvektoren entsprechen den Matrizen

der gewöhnlichen Vektoren. Beide führen Vektoren, auf die sie angewendet werden können, in andere Vektoren über. Die Linearitätsbedingung (8) gilt für alle Matrizen. Jeder Operator, der auf endlich-vieldimensionale Vektoren angewendet werden kann, ist einer Matrix äquivalent. Die unendlich viel-dimensionalen Operatoren haben auch eine Matrizenform, diese ist aber vielfach stark singulär.

Zum Beispiel sind die Koeffizienten der Matrix q_1 , die dem Operator „Multiplizieren mit x_1 “ entspricht,

$$(q_1)_{x_1 x_2 \dots x_f; x'_1 x'_2 \dots x'_f} = x_1 \delta_{x_1 x'_1} \delta_{x_2 x'_2} \dots \delta_{x_f x'_f}. \quad (*)$$

Er führt den Vektor ψ in den Vektor $q_1 \psi$ mit der $x_1 x_2 \dots x_f$ Komponente

$$\begin{aligned} q_1 \psi (x_1 x_2 \dots x_f) &= \sum_{x'_1 \dots x'_f} (q_1)_{x_1 \dots x_f; x'_1 \dots x'_f} \psi (x'_1 \dots x'_f) \\ &= \sum_{x'_1 \dots x'_f} x_1 \delta_{x_1 x'_1} \delta_{x_2 x'_2} \dots \delta_{x_f x'_f} \psi (x'_1 \dots x'_f) = x_1 \psi (x_1 \dots x_f) \end{aligned}$$

über. Dieser Vektor entspricht in der Tat der Funktion $x_1 \psi$, in die ψ durch den Operator „Multiplizieren mit x_1 “ übergeführt wird.

Die Matrix, die dem Operator „Differentiieren nach x_1 “ entspricht, bezeichnen wir mit $\frac{2\pi i}{h} p_1$:

$$\left. \left(\frac{2\pi i}{h} p_1 \right)_{x_1 \dots x_f; x'_1 \dots x'_f} \right\} = \lim_{\Delta=0} \frac{1}{\Delta} (\delta_{x_1 + \frac{1}{2}\Delta, x'_1} - \delta_{x_1 - \frac{1}{2}\Delta, x'_1}) \delta_{x_2 x'_2} \dots \delta_{x_f x'_f}, \quad (**)$$

er führt den Vektor ψ in

$$\begin{aligned} \sum_{x'_1 x'_f} \lim_{\Delta=0} \frac{1}{\Delta} (\delta_{x_1 + \frac{1}{2}\Delta, x'_1} - \delta_{x_1 - \frac{1}{2}\Delta, x'_1}) \delta_{x_2 x'_2} \dots \delta_{x_f x'_f} \psi (x'_1, x'_2, \dots, x'_f) \\ = \lim_{\Delta=0} \frac{1}{\Delta} (\psi (x_1 + \frac{1}{2}\Delta, x_2, \dots, x_f) - \psi (x_1 - \frac{1}{2}\Delta, x_2, \dots, x_f)) \end{aligned}$$

über, und das ist tatsächlich der Differentialquotient von ψ nach x_1 .

Wir nannten eine Matrix H hermiteisch, wenn $H = H^\dagger$, also wenn für beliebige Vektoren v, w

$$(v, Hw) = (H^\dagger v, w) = (Hv, w)$$

galt, d. h. wenn sie im skalaren Produkt vom einen auf den anderen Faktor abwählbar war. Entsprechend definieren wir auch die hermiteischen Operatoren.

Ein Operator H ist hermiteisch, wenn für jede (entsprechenden Nebenbedingungen genügende, im vorliegenden Fall quadrat-integrierbare, also im Unendlichen verschwindende) Funktion φ, g

$$(\varphi, Hg) = (H\varphi, g) \quad (9)$$

gilt. Die Summe, reelle Vielfache von hermiteischen Operatoren sind wieder linear hermiteisch, dasselbe gilt von ihren Potenzen, Reziproken usw.

Der Hamiltonsche Operator (5b) ist hermiteisch. Zunächst ist das „Multiplizieren mit einer reellen Funktion $V(x_1 \dots x_f)$ “ hermiteisch:

$$\begin{aligned} (\varphi, Vg) &= \int_{-\infty}^{\infty} \dots \int \varphi(x_1 \dots x_f)^* V(x_1 \dots x_f) g(x_1 \dots x_f) dx_1 \dots dx_f \\ &= \int_{-\infty}^{\infty} \dots \int (V(x_1 \dots x_f) \varphi(x_1 \dots x_f))^* g(x_1 \dots x_f) dx_1 \dots dx_f \\ &\quad = (V\varphi, g). \end{aligned} \quad \left. \right\} (9a)$$

Dann ist aber auch der Operator $\frac{\hbar}{2\pi i} \frac{\partial}{\partial x_k}$ hermiteisch. Durch partielle Integration findet man

$$\begin{aligned} \left(\varphi, \frac{\hbar}{2\pi i} \frac{\partial}{\partial x_k} g \right) &= \int_{-\infty}^{\infty} \dots \int \varphi(x_1 \dots x_f)^* \frac{\hbar}{2\pi i} \frac{\partial}{\partial x_k} g(x_1 \dots x_f) dx_1 \dots dx_f \\ &= \int_{-\infty}^{\infty} \dots \int -\frac{\hbar}{2\pi i} \left(\frac{\partial}{\partial x_k} \varphi(x_1 \dots x_f) \right)^* g(x_1 \dots x_f) dx_1 \dots dx_f \\ &\quad = \left(\frac{\hbar}{2\pi i} \frac{\partial}{\partial x_k} \varphi, g \right), \end{aligned} \quad \left. \right\} (10)$$

weil φ für $x_k = \pm \infty$ verschwindet und $i^* = -i$ ist.

Daher ist auch das Quadrat $-\frac{\hbar^2}{4\pi^2} \frac{\partial^2}{\partial x_k^2}$ hermiteisch, wie man durch zweimalige partielle Integration auch direkt zeigt. Da alle Summanden von \mathbf{H} hermiteisch sind, ist es auch \mathbf{H} selber.

Die Gleichung $\mathbf{H}\psi = E\psi$

für ψ hat bekanntlich nicht für jedes E eine nichtverschwindende, quadratintegrierbare Lösung. Die Werte E , bei denen eine solche Lösung existiert, nennt man die Eigenwerte, ihre Gesamtheit das Spektrum von \mathbf{H} .

Die Eigenwerte eines hermiteischen Operators sind alle reell: ist $\mathbf{H}\psi_E = E\psi_E$, so findet man, wenn man das skalare Produkt mit ψ_E bildet,

$$(\psi_E, \mathbf{H}\psi_E) = (\psi_E, E\psi_E) = E(\psi_E, \psi_E). \quad (11)$$

Nun ist in (11) sowohl $(\psi_E, \mathbf{H} \psi_E) = (\mathbf{H} \psi_E, \psi_E) = (\psi_E, \mathbf{H} \psi_E)^*$, wie auch (ψ_E, ψ_E) reell, so daß auch E reell sein muß.

Ein hermiteischer Operator hat im allgemeinen ein diskretes und ein kontinuierliches Spektrum.

Die Eigenwerte des diskreten Spektrums sind diskrete (endlich oder unendlich, aber jedenfalls abzählbar unendlich viele) Zahlen, die zugehörigen Eigenfunktionen lassen sich normieren [in unserem Falle bedeutet das, daß das Quadratintegral (ψ_E, ψ_E) konvergiert], und wir wollen sie fortan immer normiert annehmen. Man unterscheidet sie voneinander durch angehängte Indizes: $\psi_E, \psi_F \dots$. Die diskreten Eigenwerte bilden gewöhnlich den interessanteren Teil des Spektrums und wo bisher einfach über „Eigenwert“ gesprochen wurde, war ein diskreter Eigenwert gemeint.

Die zu einem Eigenwert E des kontinuierlichen Spektrums gehörige Lösung der Eigenwertgleichung $\psi(x_1, x_2, \dots, x_f; E)$ hat kein endliches Quadratintegral. Man könnte daher meinen, daß er gar nicht zum Spektrum gehört. Bildet man aber ein sogenanntes Eigendifferential:

$$\int_E^{E+\Delta} \psi(x_1, x_2, \dots, x_f; E) dE = \psi(x_1, x_2, \dots, x_f; E, E + \Delta), \quad (*)$$

so sieht man, daß dieses nunmehr quadratintegrable ist, so daß man es normieren kann. Dies wäre nicht der Fall, wenn E gar nicht zum Spektrum gehören würde. Das Eigendifferential (*) gehört zum Gebiet zwischen E und $E + \Delta$. Hierdurch ist es schon gegeben, daß das kontinuierliche Spektrum nicht aus Punkten, sondern aus stetigen Bereichen besteht. Die Lösungen $\psi(x_1, x_2, \dots, x_f; E)$ der Eigenwertgleichung nennt man — obwohl sie nicht normiert werden können — die Eigenfunktionen des kontinuierlichen Spektrums. Sie hängen vom Eigenwert E in stetiger Weise ab, man pflegt E als Variable an Stelle eines Index zur Unterscheidung der verschiedenen kontinuierlichen Eigenfunktionen einzuführen. Teilt man das kontinuierliche Spektrum in lauter kleine Gebiete (von der Länge Δ) ein, so kann man zu jedem ein Eigendifferential definieren, die — nachdem man sie normiert hat — den Eigenfunktionen des diskreten Spektrums um so genauer entsprechen, je kleiner Δ ist.

Die zu verschiedenen Eigenwerten des diskreten Spektrums gehörigen Eigenfunktionen sind zueinander orthogonal.

In der Tat folgt aus $\mathbf{H}\psi_E = E\psi_E$ durch skalare Produktbildung mit ψ_F

$$(\psi_F, \mathbf{H}\psi_E) = (\psi_F, E\psi_E); \quad (\mathbf{H}\psi_F, \psi_E) = E(\psi_F, \psi_E)$$

und ähnlich aus $\mathbf{H}\psi_F = F\psi_F$

$$(\mathbf{H}\psi_F, \psi_E) = (F\psi_F, \psi_E) = F^*(\psi_F, \psi_E) = F(\psi_F, \psi_E),$$

also muß für $E \neq F$ das $(\psi_F, \psi_E) = 0$ sein.

Ebenso sind die diskreten Eigenfunktionen orthogonal auf alle Eigendifferentiale. Auch die Eigendifferentiale sind orthogonal zueinander, wenn die zugehörigen Gebiete sich nicht überdecken.

Es können zu einem Eigenwert — etwa des diskreten Spektrums — auch mehrere voneinander linear unabhängige Eigenfunktionen gehören („Entartung“). Dann ist auch jede ihrer Linearkombinationen Eigenfunktion zu diesem Eigenwert. Aus dieser linearen Schar von Eigenfunktionen kann man zuerst voneinander linear unabhängige herauswählen, durch die man alle Eigenfunktionen dieses Eigenwertes linear ausdrücken kann. Aus diesen kann man dann — etwa nach dem Schmidtschen Orthogonalisierungsverfahren — zueinander orthogonale bilden. Natürlich kann diese Auswahl nicht ohne Willkür geschehen — bei dem Schmidtschen Verfahren äußert sich dies z. B. darin, daß man ein anderes Orthogonalsystem erhält, wenn man die Eigenfunktionen in anderer Reihenfolge durchläuft —, dies braucht uns aber im folgenden nicht zu stören. Wir wollen fortan immer annehmen, daß die Eigenfunktionen bei Entartungen schon in dieser Weise gewählt sind. Die Eigenfunktionen und die vorher definierten Eigendifferentiale bilden dann ein Orthogonalsystem: sind ψ und ψ' zwei beliebige, voneinander verschiedene Funktionen dieses Systems, so ist

$$(\psi, \psi') = 0. \tag{12}$$

Außerdem soll

$$(\psi, \psi) = 1 \tag{12a}$$

gelten.

Dieses Orthogonalsystem ist auch — bei genügend feiner Unterteilung des kontinuierlichen Spektrums, bei genügend kleinem Δ

— vollständig, d. h. man kann jede Funktion $\varphi(x_1, \dots, x_f)$ für die das Integral (φ, φ) konvergiert, in eine Reihe

$$\varphi = \sum_{\kappa} g_{\kappa} \psi_{\kappa} + \sum_E g(E, \Delta) \psi(E, E + \Delta) \quad (13)$$

entwickeln, wo der Index κ über alle diskreten Eigenfunktionen, E über die unteren Grenzen aller Eigendifferentielle läuft. Die Reihenentwicklung gilt eigentlich nur bei unendlich kleinem Δ , und die zweite Summe geht dann in ein Integral über:

$$\varphi = \sum_{\kappa} g_{\kappa} \psi_{\kappa} + \int g(E) \psi(E) dE, \quad (13a)$$

wo die Integration über das ganze Gebiet des kontinuierlichen Spektrums zu erstrecken ist. Gehören mehrere linear unabhängige Eigenfunktionen zu den Eigenwerten des kontinuierlichen Spektrums, so müssen in (13a) auch entsprechend mehrere Integrale, oder auch — wenn die Anzahl der Eigenfunktionen unendlich ist — ein oder mehrere Doppel- oder Mehrfachintegrale stehen. Hat das betrachtete Problem dagegen kein kontinuierliches Spektrum, so fällt das zweite Glied in (13) und das Integral in (13a) natürlich ganz weg. Durch skalare Produktbildung mit ψ_{κ} erhält man etwa aus (13) für g_{κ}

$$(\psi_{\kappa}, \varphi) = g_{\kappa}. \quad (14)$$

Und ähnlich für

$$(\psi(E, E + \Delta), \varphi) = g(E, \Delta). \quad (14a)$$

In formalen Rechnungen unterdrückt man häufig das kontinuierliche Spektrum und rechnet, als wenn nur ein diskretes existierte. Man kann sich jedoch gewöhnlich leicht überzeugen, was für Änderungen die Existenz des kontinuierlichen Spektrums nach sich zieht: zumeist treten zu den Summen noch Glieder mit Integralen hinzu.

Die Entwicklungen dieses Kapitels — namentlich sofern sie sich auf das kontinuierliche Spektrum beziehen — sind keineswegs als streng anzusehen. Die strenge Eigenwerttheorie beliebiger hermiteischer Operatoren wurde erst in jüngster Zeit erledigt¹⁾, und hier wurde nur ein Teil der Resultate zusammengefaßt. Man wird aber auch ohne Kenntnis der allgemeinen Theorie Fehler vermeiden können, namentlich wenn man sich auf das Rechnen mit Eigendifferentialen beschränkt.

¹⁾ J. v. Neumann, Mathem. Ann. 102, 49, 1929.

V. Störungstheorie

1. Es kommt häufig vor, daß man die Eigenwerte und Eigenfunktionen eines bestimmten Problems kennt und sich für die Eigenwerte und Eigenfunktionen eines benachbarten Problems interessiert, dessen Energieoperator aus dem Energieoperator dieses Problems durch eine verhältnismäßig geringe Änderung, durch eine „Störung“, hervorgeht. Mit Lösungsmethoden derartiger Aufgaben befaßt sich die Störungstheorie. Eine Störungstheorie haben schon M. Born, W. Heisenberg und P. Jordan mit Hilfe der Matrizentheorie gegeben, im folgenden schließen wir uns der Rayleigh-Schrödingerischen Methode an.

Wir rechnen so, wie wenn das Ausgangssystem kein kontinuierliches Spektrum hätte, und setzen voraus, daß auch das gestörte System ein reines Punktspektrum besitzt. Die durch das Vorhandensein bzw. das Entstehen des kontinuierlichen Spektrums bedingten Komplikationen sollen erst am Schlusse erörtert werden, zuerst sei die Theorie in ihrer einfachsten Form erläutert.

Wir haben einen hermiteischen Operator \mathbf{H} mit den Eigenwerten E_1, E_2, \dots und Eigenfunktionen ψ_1, ψ_2, \dots

$$\mathbf{H} \psi_k = E_k \psi_k. \quad (1)$$

Gesucht sind Eigenwerte F und Eigenfunktionen φ des Operators $\mathbf{H} + \lambda \mathbf{V}$, wo \mathbf{V} ebenfalls hermiteisch und λ eine kleine Zahl ist:

$$(\mathbf{H} + \lambda \mathbf{V}) \varphi_k = F_k \varphi_k. \quad (2)$$

Wir setzen zunächst für F und φ eine Potenzreihe nach Potenzen von λ an, die wir bei dem zweiten Gliede abbrechen:

$$F_k = E_k + \lambda E'_k + \lambda^2 E''_k \dots \quad (3a)$$

$$\varphi_k = \psi_k + \lambda \psi'_k + \lambda^2 \psi''_k \dots = \psi_k + \lambda \sum_l a_{kl} \psi_l + \lambda^2 \sum_l b_{kl} \psi_l \dots \quad (3b)$$

In (3a) und (3b) ist angenommen, daß F_k bzw. φ_k für $\lambda = 0$ in E_k bzw. ψ_k übergehen, außerdem ist ψ'_k und ψ''_k nach den Resultaten des vorangehenden Kapitels in eine Reihe nach den ψ entwickelt, die Entwicklungskoeffizienten sind mit a_{kl} bzw. b_{kl} bezeichnet.

Wir setzen (3a) und (3b) in (2) ein und erhalten:

$$\left. \begin{aligned} & \mathbf{H} [\psi_k + \lambda \sum_l a_{kl} \psi_l + \lambda^2 \sum_l b_{kl} \psi_l] + \lambda \mathbf{V} [\psi_k + \lambda \sum_l a_{kl} \psi_l] \\ &= (E_k + \lambda E'_k + \lambda^2 E''_k) (\psi_k + \lambda \sum_l a_{kl} \psi_l + \lambda^2 \sum_l b_{kl} \psi_l). \end{aligned} \right\} \quad (4)$$

Die Koeffizienten gleich hoher Potenzen von λ müssen einander gleich sein. Die von λ freien Glieder heben sich wegen (1) weg, der Vergleich der Koeffizienten von λ bzw. λ^2 ergibt

$$\sum_l a_{kl} E_l \psi_l + \mathbf{V} \psi_k = E'_k \psi_k + E_k \sum_l a_{kl} \psi_l, \quad (5a)$$

$$\sum_l b_{kl} E_l \psi_l + \sum_l a_{kl} \mathbf{V} \psi_l = E''_k \psi_k + E'_k \sum_l a_{kl} \psi_l + E_k \sum_l b_{kl} \psi_l. \quad (5b)$$

(5a) gestattet die Bestimmung von E'_k und a_{kl} für $l \neq k$: man erhält, indem man das skalare Produkt mit ψ_k bzw. ψ_l bildet, wegen der Orthogonalitätsrelationen

$$a_{kk} E_k + (\psi_k, \mathbf{V} \psi_k) = E'_k + a_{kk} E_k, \quad (6)$$

$$a_{kl} E_l + (\psi_l, \mathbf{V} \psi_k) = E_k a_{kl} \text{ (für } l \neq k\text{).} \quad (7)$$

Wenn man hier zur Abkürzung

$$V_{\alpha\beta} = (\psi_\alpha, \mathbf{V} \psi_\beta) = (\mathbf{V} \psi_\alpha, \psi_\beta) = (\psi_\beta, \mathbf{V} \psi_\alpha)^* = V_{\beta\alpha}^* \quad (8)$$

einführt, lautet dies

$$E'_k = (\psi_k, \mathbf{V} \psi_k) = V_{kk}, \quad (6a)$$

$$a_{kl} = \frac{(\psi_l, \mathbf{V} \psi_k)}{E_k - E_l} = \frac{V_{lk}}{E_k - E_l} \text{ (für } l \neq k\text{).} \quad (7a)$$

Ähnlich erhält man aus (5b), indem man mit ψ_k^* multipliziert und über den ganzen Konfigurationsraum integriert,

$$b_{kk} E_k + \sum_l a_{kl} (\psi_k, \mathbf{V} \psi_l) = E''_k + E'_k a_{kk} + E_k b_{kk}. \quad (9)$$

Auf der linken Seite zerlegen wir die Summe über l in zwei Teile, indem wir das Glied mit $l = k$ separat hinschreiben. Außerdem setzen wir für E'_k den Wert aus (6a) und für a_{kl} aus (7a) ein. Wir erhalten so

$$E''_k = \sum_{l \neq k} \frac{(\psi_l, \mathbf{V} \psi_k) (\psi_k, \mathbf{V} \psi_l)}{E_k - E_l} = \sum_{l \neq k} \frac{|V_{lk}|^2}{E_k - E_l}.$$

Dies ergibt den neuen Eigenwert F_k bis zu Gliedern von der Ordnung λ^2

$$F_k = E_k + \lambda V_{kk} + \lambda^2 \sum_{l \neq k} \frac{|V_{lk}|^2}{E_k - E_l}. \quad (10)$$

Die neue Eigenfunktion φ_k ist bis zu Gliedern von der Ordnung λ

$$\varphi_k = \psi_k + \lambda \sum_{l \neq k} \frac{V_{lk}}{E_k - E_l} \psi_l + \lambda a_{kk} \psi_k.$$

Aus den bisherigen Gleichungen fiel a_{kk} immer heraus, was dem Umstand entspricht, daß in φ_k der Normierungsfaktor frei ist. Setzen wir $(\varphi_k, \varphi_k) = 1$, so erhalten wir $a_{kk} = 0$, und

$$\varphi_k = \psi_k + \lambda \sum_{l \neq k} \frac{V_{lk}}{E_l - E_k} \psi_l \quad (11)$$

ist bis auf Glieder mit λ^2 normiert.

Über die Konvergenz dieses Störungsverfahrens ist folgendes zu sagen: In den Summen in (10) und (11) können — wenn die Eigenwerte E_l und E_k zweier Eigenfunktionen ψ_l und ψ_k des Ausgangsproblems zufällig einander gleich sind — unendlich große Glieder auftreten. Diese kann man aber leicht eliminieren, und ihr Vorkommen bildet — wie wir sogleich sehen werden — keine ernsten Konvergenzschwierigkeiten. Nachdem man sie eliminiert hat, lassen sich die Summationen in den meisten praktisch vorkommenden Fällen ausführen. Damit ist aber über die Konvergenz des ganzen Verfahrens — über die der Reihe (3a), (3b) — noch nichts ausgesagt. Diese kann noch sehr wohl divergieren, in vielen Beispielen ist schon das dritte Glied in (3a) allein unendlich groß! Außerdem ist es bekannt, daß ein diskreter Eigenwert — namentlich, wenn er schon im Ausgangsproblem vom kontinuierlichen Spektrum überdeckt war — sich bei der Störung „auflösen“, d. h. gänzlich im kontinuierlichen Spektrum untergehen kann.

Aber selbst in diesen Fällen hat (11) einen wohlbestimmten Sinn: er gibt einen Zustand wieder, der — bei kleinem λ — während sehr langer Zeiten, wenn auch nicht absolut, so doch fast stationär ist. Das F_k in (10) gibt die ungefähre Energie und durch \hbar dividiert die ungefähre Frequenz dieser Bewegung an. Wenn man mit Hilfe von (10) und (11) $a = (\mathbf{H} + \lambda \mathbf{V} - F_k) \varphi_k$ bildet, so findet man, daß es vom zweiten Grade in λ ist. Nimmt man also an, daß die Wellenfunktion $\varphi(t)$ des Systems zur Zeit $t = 0$ mit φ_k übereinstimmt: $\varphi(0) = \varphi_k$, so kann man allenfalls

$$\varphi(t) = \varphi_k e^{\frac{2\pi i F_k}{\hbar} t} + \chi(t) \quad (12)$$

ansetzen. Setzt man dies in die zeitabhängige Schrödingergleichung ein:

$$\begin{aligned} \frac{\hbar}{2\pi i} \frac{\partial \varphi}{\partial t} &= F_k \varphi_k e^{\frac{2\pi i F_k}{\hbar} t} + \frac{\hbar}{2\pi i} \frac{\partial \chi}{\partial t} \\ &= F_k \varphi_k e^{\frac{2\pi i F_k}{\hbar} t} + a e^{\frac{2\pi i F_k}{\hbar} t} + (\mathbf{H} + \lambda \mathbf{V}) \chi, \end{aligned}$$

so erhält man

$$\frac{h}{2\pi i} \frac{\partial \chi}{\partial t} = a e^{\frac{2\pi i F_k}{h} t} + (\mathbf{H} + \lambda \mathbf{V}) \chi \quad (13)$$

und daraus, indem man $\frac{\partial}{\partial t}(\chi, \chi)$ berechnet,

$$\frac{\partial}{\partial t}(\chi, \chi) = \frac{2\pi i}{h} \left(e^{\frac{2\pi i F_k}{h} t} (\chi, a) - e^{-\frac{2\pi i F_k}{h} t} (a, \chi) \right).$$

Hieraus folgt mit Hilfe der Schwarzschen Ungleichung
 $|(\chi, a)|^3 \leq (\chi, \chi)(a, a)$

$$\frac{\partial}{\partial t}(\chi, \chi) \leq \frac{4\pi}{h} \sqrt{(\chi, \chi)} \cdot \sqrt{(a, a)} \quad (14)$$

oder¹⁾

$$\frac{\partial \sqrt{(\chi, \chi)}}{\partial t} \leq \frac{2\pi}{h} \sqrt{(a, a)}; \quad \sqrt{(\chi, \chi)} \leq \frac{2\pi}{h} \sqrt{(a, a)} t + C.$$

Da für $t = 0$ auch $\chi = 0$ ist, ist $C = 0$ und

$$(\chi, \chi) \leq (a, a) \frac{4\pi^3}{h^2} t^3$$

d. h. die Abweichung von $\varphi(t)$ von $\varphi_k e^{\frac{2\pi i F_k}{h} t}$ ist für Zeiten, die klein gegen $h/2\pi \cdot \sqrt{(a, a)}$ sind, sicher sehr klein. Da aber (a, a) proportional λ^4 ist, ändert sich bei kleinem λ das φ_k während verhältnismäßig langer Zeiten so, als wenn es eine wirkliche Eigenfunktion wäre.

2. Einer Korrektion bedürfen, wie schon erwähnt, diese Entwicklungen in dem Falle, wenn die Eigenwerte zweier oder mehrerer Eigenfunktionen des Ausgangsproblems gleich sind, d. h. wenn mehrere linear unabhängige Eigenfunktionen zum selben Eigenwert des Ausgangsproblems gehören. Da die Summation in (10) und (11) über alle Eigenfunktionen zu erstrecken ist und in den Summen (10) und (11) auch der Index l jener Eigenfunktionen vorkommt, deren Eigenwert E_l gleich E_k ist, kann man diese Summen nur bilden, wenn für jene l , für welche $E_k = E_l$ mit $l \neq k$ verschwindet, auch $(\psi_l, \mathbf{V} \psi_k)$ Null ist.

Stellen wir uns vor, es seien die Eigenfunktionen $\psi_{k1}, \dots, \psi_{ks}$, die zum Eigenwert E_k gehören. Wir nehmen an (was auch bisher geschah), daß auch diese orthogonal aufeinander sind. Es

¹⁾ a ist unabhängig von t .

besteht dann eine gewisse Willkür in der Wahl der Eigenfunktionen für unser Näherungsverfahren, indem man als Ausgangsfunktionen an Stelle der $\psi_{k1}, \dots, \psi_{ks}$ andere, etwa

$$\left. \begin{aligned} \psi'_{k1} &= a_{11} \psi_{k1} + a_{12} \psi_{k2} + \dots + a_{1s} \psi_{ks} \\ \psi'_{k2} &= a_{21} \psi_{k1} + a_{22} \psi_{k2} + \dots + a_{2s} \psi_{ks} \\ \vdots &\quad \ddots \quad \ddots \quad \ddots \quad \ddots \\ \psi'_{ks} &= a_{s1} \psi_{k1} + a_{s2} \psi_{k2} + \dots + a_{ss} \psi_{ks} \end{aligned} \right\} \quad (15)$$

nehmen könnte, und man ist nicht mehr berechtigt, anzunehmen, daß die erste Näherung für φ_k eben ψ_k ist. Ist $(a_{\mu\mu})$ eine unitäre Matrix, so sind auch die ψ'_k , aufeinander orthogonal (natürlich auch orthogonal auf die übrigen Eigenfunktionen, die nicht zum Eigenwert E_k gehören),

$$\left. \begin{aligned} (\psi'_{kv}, \psi'_{k\mu}) &= (\sum_{v'} a_{vv'} \psi_{kv'}, \sum_{\mu'} a_{\mu\mu'} \psi_{k\mu'}) \\ &= \sum_{v'\mu'} a_{v'v}^* a_{\mu\mu'} (\psi_{kv'}, \psi_{k\mu'}) = \sum_{v'\mu'} a_{v'v}^* a_{\mu\mu'} \delta_{v'\mu'} = \delta_{v\mu}, \end{aligned} \right\} \quad (16)$$

und sind als Ausgangssystem für das Näherungsverfahren ebenso geeignet wie das ursprüngliche System der ψ_{kv} .

Es erhebt sich die Frage, ob man durch zweckmäßige Bestimmung der Matrix a nicht alle $(\psi'_{kv}, V \psi'_{k\mu})$ mit $v \neq \mu$ zum Verschwinden bringen könnte. In der Tat gelingt dies, es ist

$$(\psi'_{kv}, V \psi'_{k\mu}) = \sum_{v', \mu'=1}^s a_{v'v}^* a_{\mu\mu'} (\psi_{kv'}, V \psi_{k\mu'}). \quad (17)$$

Bezeichnet man also die aus den Größen $(\psi_{kv'}, V \psi_{k\mu'}) = V_{kv'; k\mu'} = v_{v'\mu'}$ gebildete hermitische Matrix mit s Zeilen und Spalten mit v , so muß die Matrix a so bestimmt werden, daß $a^* v a'$ eine Diagonalmatrix sei. Ist a in dieser Weise gewählt, so verschwinden für $v \neq \mu$ nach (17) die $(\psi'_{kv}, V \psi'_{k\mu})$, und in den Formeln (10) und (11) treten — wenn man die mit den ψ_{kv} vollkommen gleichberechtigten ψ'_{kv} als Ausgangssystem für das Störungsverfahren benutzt — keine Glieder mit verschwindendem Nenner auf.

Nun ist a , also auch a^* eine unitäre Matrix, $a^* = a'^{-1}$, und die Aufgabe besteht nach dem Vorangehenden darin, diese so zu wählen, daß sie v auf Diagonalform transformiere. Die Gleichung zur Bestimmung der $a_{\mu\mu'}$ lautet dann

$$\sum_{\mu'} v_{v'\mu'} a_{\mu\mu'} = a_{\mu}, v'_{\mu}, \quad (18)$$

wo die v'_{μ} die Eigenwerte der Matrix v sind.

Wir wollen auch in diesem Falle den Eigenwert bis zu Gliedern mit λ^2 und die Eigenfunktion bis zu Gliedern mit der ersten Potenz von λ berechnen. Die Eigenfunktionen $\psi_{k_1}, \psi_{k_2}, \dots, \psi_{k_s}$ des Eigenwertes E_k , dessen Verschiebung wir berechnen wollen, nehmen wir — was wir ja tun dürfen und was nach dem Vorangehenden zweckmäßig erscheint — von vornherein so an, daß

$$(\psi_{kv}, \mathbf{V} \psi_{k\mu}) = V_{kv; k\mu} = V_{kv; kv} \delta_{v\mu} = v'_v \delta_{v\mu} \quad (19)$$

sei, d. h. wir gehen sofort von den ψ' aus. Die übrigen Eigenfunktionen können wir durchnumerieren: ψ_l gehöre zu E_l . Den Eigenwert des Operators $\mathbf{H} + \lambda \mathbf{V}$, der zur Eigenfunktion φ_{kv} gehört, bezeichnen wir mit F_{kv} , es wird sich nämlich zeigen, daß der s -fache Eigenwert¹⁾ E_k aufspalten wird und im allgemeinen s neue Eigenwerte ergibt.

Es sei

$$F_{kv} = E_k + \lambda E'_{kv} + \lambda^2 E''_{kv} \dots \quad (20)$$

und

$$\begin{aligned} \varphi_{kv} &= \psi_{kv} + \lambda \sum_{\mu=1}^s \beta_{kv; k\mu} \psi_{k\mu} + \lambda \sum_{l \neq k} a_{kv; l} \psi_l \\ &\quad + \lambda^2 \sum_{\mu} \gamma_{kv; k\mu} \psi_{k\mu} + \lambda^2 \sum_{l \neq k} b_{kv; l} \psi_l. \end{aligned} \quad (20a)$$

Setzt man (20), (20a) in die Gleichung $(\mathbf{H} + \lambda \mathbf{V}) \varphi_{kv} = F_{kv} \varphi_{kv}$, ein und vergleicht wiederum die Koeffizienten gleich hoher Potenzen von λ , so sieht man, daß die Glieder mit der 0-ten Potenz sich gegenseitig wegheben, während der Vergleich der ersten und zweiten Potenz die beiden Gleichungen

$$\begin{aligned} \sum_{\mu} E_k \beta_{kv; k\mu} \psi_{k\mu} + \sum_{l \neq k} E_l a_{kv; l} \psi_l + \mathbf{V} \psi_{kv} \\ = \sum_{\mu} E_k \beta_{kv; k\mu} \psi_{k\mu} + \sum_{l \neq k} E_k a_{kv; l} \psi_l + E'_{kv} \psi_{kv}, \end{aligned} \quad (21)$$

$$\left. \begin{aligned} \sum_{\mu} E_k \gamma_{kv; k\mu} \psi_{k\mu} + \sum_{l \neq k} E_l b_{kv; l} \psi_l + \sum_{\mu} \beta_{kv; k\mu} \mathbf{V} \psi_{k\mu} \\ + \sum_{l \neq k} a_{kv; l} \mathbf{V} \psi_l = \sum_{\mu} E_k \gamma_{kv; k\mu} \psi_{k\mu} + \sum_{l \neq k} E_k b_{kv; l} \psi_l \\ + \sum_{\mu} E'_{kv} \beta_{kv; k\mu} \psi_{k\mu} + \sum_{l \neq k} E'_{kv} a_{kv; l} \psi_l + E''_{kv} \psi_{kv} \end{aligned} \right\} \quad (21a)$$

¹⁾ Man nennt ihn so, weil zu ihm s linear unabhängige Eigenfunktionen gehören.

ergibt. Aus diesen Gleichungen kann man, ebenso wie vorher, die Unbekannten $E'_{k\nu}$, $E''_{k\nu}$, $a_{k\nu;l}$, $\beta_{k\nu;k\mu}$ bestimmen. Für die Energie $F_{k\nu}$ ergibt sich

$$F_{k\nu} = E_k + \lambda V_{k\nu;k\nu} + \lambda^2 \sum_{l \neq k} \frac{|V_{l;k\nu}|^2}{E_k - E_l} \quad (22)$$

und die zugehörige Eigenfunktion ist

$$\left. \begin{aligned} \varphi_{k\nu} &= \psi_{k\nu} + \lambda \sum_{\mu \neq \nu} \sum_{l \neq k} \frac{V_{k\mu;l} V_{l;k\nu}}{(E_k - E_l)(V_{k\nu;k\nu} - V_{k\mu;k\mu})} \psi_{k\mu} \\ &\quad + \lambda \sum_{l \neq k} \frac{V_{l;k\nu}}{E_k - E_l} \psi_l. \end{aligned} \right\} \quad (23)$$

Dabei ist vorausgesetzt, daß die $\psi_{k1}, \psi_{k2}, \dots, \psi_{ks}$ schon so gewählt sind, daß $V_{k\nu;k\mu} = 0$ für $\nu \neq \mu$ ist.

Sind die $V_{k\nu;k\nu} = v'_\nu$, für $\nu = 1, 2, \dots, s$ alle verschieden voneinander, so spaltet der Eigenwert E_k schon in erster Näherung in s neue Eigenwerte auf. Dann läßt sich auch $\varphi_{k\nu}$ nach (23) ohne weiteres bilden, da in (23) keine verschwindenden Nenner vorkommen.

Sind aber unter den $v'_\nu = V_{k\nu;k\nu}$ gleiche vorhanden, so sind die zugehörigen Eigenfunktionen wieder nur bis auf eine unitäre Transformation bestimmt (weil die zum gleichen Eigenwert von v gehörigen Eigenvektoren nur bis auf eine solche festgelegt sind). Wir müssen sie dann so wählen, daß die hermitische Matrix w

$$w_{\mu\nu} = \sum_{l \neq k} \frac{V_{k\mu;l} V_{l;k\nu}}{E_k - E_l} \quad (24)$$

die Diagonalform haben soll. Dann kann man auch (23) bilden. All dies tritt automatisch ein, wenn man die richtigen Eigenfunktionen erster Näherung (15) aus anderen Überlegungen kennt und von vornherein mit diesen rechnet.

Mit diesen Modifikationen ist also das Störungsverfahren, auch wenn mehrere, allerdings nur endlich viele Eigenfunktionen zum selben Punkteigenwert gehören, noch ausführbar. Dieser Fall wird uns im folgenden viel beschäftigen, und die Entwicklungen dieses Kapitels bilden die Grundlage der meisten quantenmechanischen Rechnungen. Meistens beschränkt man sich dabei in (22) auf das Glied mit der ersten Potenz von λ , d. h. auf $V_{k\nu;k\nu} = v'_\nu$. Dieses kann man schon durch Auflösen der Säkulargleichung von (18) berechnen oder auch auf eine einfache Quadratur

zurückführen, wenn man die sogenannten „richtigen Linearkombinationen“ (15), für die

$$v_{\nu\mu} = (\psi_{k\nu}, \mathbf{V} \psi_{k\mu}) = 0 \text{ für } \nu \neq \mu$$

und

$$w_{\nu\mu} = 0 \text{ für } \nu \neq \mu \text{ und } v_{\nu\nu} = v_{\mu\mu}$$

ist, kennt. Diese kann man häufig, ohne die Säkulargleichung von (18) aufzulösen, aus gruppentheoretischen Überlegungen bestimmen. Dies ergibt eine wichtige Anwendung der Gruppentheorie auf quantenmechanische Probleme.

VI. Transformationstheorie und Grundlinien der statistischen Deutung der Quantenmechanik

1. Während anfangs das Bestreben der Quantenmechanik sich nur auf die Bestimmung der Energiewerte, spontanen Übergangswahrscheinlichkeiten usw. richtete, ging man später immer mehr zu prinzipiellen Fragen über und versuchte, sich unter den Matrizen, Operatoren und Eigenfunktionen auch etwas vorzustellen. So entstand die statistische Deutung der Quantenmechanik, bei deren Entwicklung M. Born, P. A. M. Dirac, P. Jordan, W. Pauli jr. und W. Heisenberg in erster Reihe beteiligt waren.

Während man in der klassischen Mechanik zur Beschreibung eines Systems mit f Freiheitsgraden $2f$ Zahlen: f Lagenkoordinaten und f Geschwindigkeiten, notwendig hatte, beschreibt die Quantenmechanik einen Zustand durch eine normierte Wellenfunktion $\varphi(x_1, x_2, \dots, x_r)$ [es ist $(\varphi, \varphi) = 1$], dessen Argumente die Lagenkoordinaten sind. Ebenso wie in der klassischen Theorie zu beliebigen $2f$ Zahlen ein Zustand zugeordnet war, ist jetzt zu jeder Wellenfunktion, die der Bedingung

$$\int_{-\infty}^{\infty} \cdots \int | \varphi(x_1, x_2, \dots, x_r) |^2 dx_1 \dots dx_r = 1$$

genügt (nicht etwa nur den Eigenfunktionen der Schrödinger-Gleichung, sondern z. B. auch Linearkombinationen von diesen), ein Zustand zugeordnet. Die Mannigfaltigkeit der Zustände ist also in der Quantenmechanik viel größer als in der klassischen Theorie.

Die zeitliche Fortentwicklung des Systems ist in der klassischen Mechanik durch die Newtonschen Bewegungsgleichungen bestimmt, in der Quantenmechanik durch die zeitabhängige Schrödinger-Gleichung

$$\frac{\hbar}{2\pi i} \frac{\partial \varphi}{\partial t} = \mathbf{H}\varphi, \quad (1)$$

wo \mathbf{H} der Hamiltonsche Operator ist: im einfachsten Falle

$$\mathbf{H} = - \sum_{k=1}^f \frac{1}{2m_k} \frac{\hbar^2}{4\pi^2} \frac{\partial^2}{\partial x_k^2} + V(x_1, x_2, \dots, x_f). \quad (2)$$

Seine genaue Bestimmung ist wohl die wichtigste Aufgabe der Quantenmechanik.

In der klassischen Mechanik gaben die $2f$ Zahlen, die zur Beschreibung des Zustandes dienen, die Koordinaten und Geschwindigkeiten der einzelnen Teilchen direkt an und man konnte aus ihnen auch mühelos beliebige Funktionen dieser Größen berechnen. In der Quantenmechanik hat die Frage nach der Lage eines Teilchens im allgemeinen gar keinen Sinn: sinnvoll ist nur die Frage nach der Wahrscheinlichkeit, mit der sich ein Teilchen an einer bestimmten Stelle befindet. Dasselbe gilt von Impulsen und von Funktionen dieser Größen, wie z. B. Energie usw.

Den einzelnen physikalischen Größen ordnet die Quantenmechanik hermitische Operatoren zu. So z. B. ist der Operator, der der x_k -Koordinate zugeordnet ist: „Multiplizieren mit x_k “, dem entsprechenden Impuls: $\frac{\hbar}{2\pi i} \frac{\partial}{\partial x_k}$, der Energie das \mathbf{H} aus (2) usw. Dieser letztere Operator ist übrigens der einzige, der eine besondere Rolle spielt: er tritt in der zeitabhängigen Schrödinger-Gleichung auf.

Im allgemeinen erhält man diesen Operator, wenn man die fragliche Größe klassisch mit Hilfe von Lagen- und Impulskoordinaten ausdrückt und dann die Lagenkoordinate x_k durch den Operator „Multiplizieren mit x_k “, die Impulskoordinate p_k durch den Operator $\frac{\hbar}{2\pi i} \frac{\partial}{\partial x_k}$ ersetzt. Z. B. ist bei einem harmonischen Oszillator die Energie

$$\frac{1}{2m} (p_1^2 + p_2^2 + p_3^2) + \frac{1}{2} \omega^2 (x_1^2 + x_2^2 + x_3^2).$$

Sie wird durch den Operator

$$\frac{1}{2m} \left[\left(\frac{\hbar}{2\pi i} \frac{\partial}{\partial x_1} \right)^2 + \left(\frac{\hbar}{2\pi i} \frac{\partial}{\partial x_2} \right)^2 + \left(\frac{\hbar}{2\pi i} \frac{\partial}{\partial x_3} \right)^2 \right] + x^3 (x_1 \cdot x_1 + x_2 \cdot x_2 + x_3 \cdot x_3)$$

ersetzt. Das ist genau das, was in (2) angegeben ist.

Die Messung einer Größe (Koordinate, Energie) kann überhaupt nur jene Werte ergeben, die als Eigenwerte des zugeordneten Operators vorkommen. So z. B. sind die möglichen Energiestufen die Eigenwerte von \mathbf{H} . Wie groß ist die Wahrscheinlichkeit, wenn sich das System im Zustand $\varphi(x_1 \dots x_f)$ befindet, daß die Größe mit dem Operator \mathbf{G} den Wert λ_k hat? Diese Wahrscheinlichkeit ist sicher Null, wenn λ_k kein Eigenwert von \mathbf{G} ist, ist er dagegen ein Eigenwert und ψ_k die zugehörige normierte Eigenfunktion, so gibt

$$|(\varphi, \psi_k)|^2 = |(\psi_k, \varphi)|^2 \quad (3)$$

die gefragte Wahrscheinlichkeit an.

Im Sinne der statistischen Deutung sollen nur diese Wahrscheinlichkeiten einen physikalischen Sinn haben, nur diese Folgerung aus der Theorie ist durch Experimente prüfbar.

Entwickelt man φ nach dem vollständigen Orthogonalsystem der Eigenfunktionen von \mathbf{G}

$$\varphi = a_1 \psi_1 + a_2 \psi_2 + a_3 \psi_3 + \dots, \quad (4)$$

so ist die Wahrscheinlichkeit für λ_k , wenn

$$\mathbf{G} \psi_k = \lambda_k \psi_k \quad (4a)$$

ist, nach (3) das Absolutwertquadrat $|a_k|^2$ von

$$(\psi_k, \varphi) = a_k. \quad (5)$$

Natürlich muß die Summe der Wahrscheinlichkeiten, daß der Wert der betrachteten Größe irgendeinen der Werte $\lambda_1, \lambda_2, \lambda_3 \dots$ annimmt,

$$|a_1|^2 + |a_2|^2 + |a_3|^2 + \dots = 1$$

sein. In der Tat folgt aus

$$\begin{aligned} (\varphi, \varphi) &= \left(\sum_k a_k \psi_k, \sum_l a_l \psi_l \right) = \sum_{k,l} a_k^* a_l (\psi_k, \psi_l) \\ &\Rightarrow \sum_{k,l} a_k^* a_l \delta_{kl} = \sum_k |a_k|^2 = 1. \end{aligned}$$

Die Wellenfunktion $c\varphi$ (mit $|c| = 1$) entspricht demselben Zustand, dem die Wellenfunktion φ entspricht, die Wellenfunktion ist nur bis auf einen Faktor vom Absolutwert Eins durch den physikalischen Zustand bestimmt. In der Tat sind alle

Wahrscheinlichkeiten, die man mit Hilfe der Wellenfunktion φ einerseits, mit der Wellenfunktion $c\varphi$ andererseits berechnet, einander gleich:

$$|\langle \psi_k, c\varphi \rangle|^2 = |c\langle \psi_k, \varphi \rangle|^2 = |c|^2 |\langle \psi_k, \varphi \rangle|^2 = |\langle \psi_k, \varphi \rangle|^2,$$

und diese Wahrscheinlichkeiten sind das einzige physikalisch Reale am Zustand.

Gehören zu einem Eigenwert λ_k mehrere linear unabhängige Eigenfunktionen $\psi_{k1}, \psi_{k2}, \psi_{k3}, \dots$ (die wir zueinander orthogonal annehmen wollen), so ist die Wahrscheinlichkeit für λ_k gleich der Summe der Quadrate der Entwicklungskoeffizienten

$$|\langle \psi_{k1}, \varphi \rangle|^2 + |\langle \psi_{k2}, \varphi \rangle|^2 + |\langle \psi_{k3}, \varphi \rangle|^2 + \dots$$

Dies alles gilt zunächst natürlich nur für die Wahrscheinlichkeiten der diskreten Eigenwerte. Die Wahrscheinlichkeit für einen ganz bestimmten Wert des kontinuierlichen Spektrums ist ja immer Null, weil im kontinuierlichen Spektrum nur endliche Bereiche eine endliche Wahrscheinlichkeit haben können. Diese ist — wenn das Bereich hinreichend klein ist — gleich dem Absolutwertquadrat des Entwicklungskoeffizienten des zu diesem Bereich gehörigen normierten Eigendifferentials.

2. Nur in einem Fall arten die Wahrscheinlichkeitsaussagen der Quantenmechanik in bestimmte Voraussagen aus: wenn die Zustandsfunktion φ eine Eigenfunktion des zu messenden Operators \mathbf{G} , also $\mathbf{G}\varphi = \lambda_k \varphi$ ist. Dann ist nämlich φ auf alle nicht zu λ_k gehörigen Eigenfunktionen von \mathbf{G} orthogonal und die Wahrscheinlichkeit für diese Eigenwerte ist Null, für λ_k also 1. In diesem Fall ergibt die Messung mit Sicherheit den Wert λ_k .

Wenn wir eine Größe gemessen und dabei einen bestimmten Wert gefunden haben, so müssen wir, wenn wir diese Messung nur genügend rasch wiederholen, denselben Wert finden. Sonst wäre die durch die Messung gewonnene Aussage, daß die betreffende Größe diesen oder jenen Wert hat, sinnlos. Die Wahrscheinlichkeiten für eine nochmalige Messung, also auch die Wellenfunktion, die ja nur zur Berechnung der Wahrscheinlichkeiten dienen soll, ändern sich¹⁾ durch eine Messung. Und zwar

¹⁾ Die Wellenfunktion ändert sich also auf zwei sehr verschiedene Weisen. Erstens im Laufe der Zeit, kontinuierlich, nach der Differentialgleichung (1) und zweitens bei den Messungen, die man am System vornimmt diskontinuierlich nach Wahrscheinlichkeitsgesetzen (s. weiter unten).

sehen wir, daß die Wellenfunktion nach der Messung, die für \mathbf{G} den Wert λ_k ergeben hat, eine zu λ_k gehörige Eigenfunktion von \mathbf{G} sein muß. Nur so ist es nämlich gewährleistet, daß eine nochmalige Messung von \mathbf{G} wiederum den Wert λ_k ergibt. Durch die Messung von \mathbf{G} ist die Wellenfunktion gezwungen worden, in eine der Eigenfunktionen von \mathbf{G} überzugehen. Und zwar ist sie in ψ_k übergegangen, wenn die Messung das Resultat λ_k ergab. In welche der Eigenfunktionen von \mathbf{G} die Zustandsfunktion des Systems übergehen wird, kann man im allgemeinen nicht mit Sicherheit voraussagen, die Quantenmechanik gibt nur die Wahrscheinlichkeiten

$$|(\psi_k, \varphi)|^2$$

für die einzelnen (Werte λ_k , also) Eigenfunktionen ψ_k an. Da man den Ausdruck $|(\psi_k, \varphi)|^2$, die Häufigkeit des Überganges der Wellenfunktion φ in ψ_k bei der Ausführung der Messung \mathbf{G} schon aus den beiden Funktionen φ und ψ_k berechnen kann, nennt man ihn einfacher die (durch einen Versuch bedingte) Übergangswahrscheinlichkeit vom Zustand φ in den Zustand ψ_k . Kennt man die Übergangswahrscheinlichkeiten einer Wellenfunktion in jede Funktion, so kann man die Wahrscheinlichkeiten der verschiedenen Resultate aller denkbaren Versuche berechnen.

Nach dem Vorangehenden ist es kaum noch nötig zu bemerken, daß die Übergangswahrscheinlichkeiten einen physikalischen Sinn haben und daß sie daher in zwei Beschreibungen desselben Systems, die miteinander äquivalent sein sollen, denselben Wert haben müssen.

3. Übergang zu einem anderen „Koordinatensystem“. Sind $\mathbf{G}, \mathbf{G}', \mathbf{G}''$, ... Operatoren, die verschiedenen physikalischen Größen, wie Energie, Koordinate, Impuls usw. zugeordnet sind, und $\varphi_1, \varphi_2, \varphi_3, \dots$ Wellenfunktionen verschiedener Zustände, so erhält man dieselben Resultate wie mit diesem System von Operatoren und Wellenfunktionen, wenn man die ersten durch

$$\bar{\mathbf{G}} = \mathbf{U} \mathbf{G} \mathbf{U}^{-1}; \quad \bar{\mathbf{G}'} = \mathbf{U} \mathbf{G}' \mathbf{U}^{-1}; \quad \bar{\mathbf{G}''} = \mathbf{U} \mathbf{G}'' \mathbf{U}^{-1}; \dots,$$

die Wellenfunktionen dagegen durch

$$\bar{\varphi}_1 = \mathbf{U} \varphi_1; \quad \bar{\varphi}_2 = \mathbf{U} \varphi_2; \quad \bar{\varphi}_3 = \mathbf{U} \varphi_3; \quad \dots$$

ersetzt, wobei \mathbf{U} ein beliebiger unitärer¹⁾ Operator ist. Zunächst sind die Eigenwerte, die überhaupt möglichen Versuchsresultate der \mathbf{G} und der $\bar{\mathbf{G}} = \mathbf{U} \mathbf{G} \mathbf{U}^{-1}$ dieselben, da sich die Eigenwerte durch eine Ähnlichkeitstransformation nicht ändern: ist λ_k Eigenwert von \mathbf{G} und ψ_k die zugehörige Eigenfunktion, so ist λ_k auch Eigenwert von $\bar{\mathbf{G}} = \mathbf{U} \mathbf{G} \mathbf{U}^{-1}$ und die zugehörige Eigenfunktion ist $\mathbf{U} \psi_k$, aus $\mathbf{G} \psi_k = \lambda_k \psi_k$ folgt

$$\bar{\mathbf{G}} \mathbf{U} \psi_k = \mathbf{U} \mathbf{G} \mathbf{U}^{-1} \mathbf{U} \psi_k = \mathbf{U} \mathbf{G} \psi_k = \mathbf{U} \lambda_k \psi_k = \lambda_k \mathbf{U} \psi_k.$$

Aber auch die Wahrscheinlichkeiten dieser Eigenwerte — etwa des Resultats λ_k der Größe, der im „ersten Koordinatensystem“ \mathbf{G} , im zweiten $\bar{\mathbf{G}}$ zugeordnet ist — ergeben sich beidemal gleich groß. Im ersten Falle haben wir

$$|(\psi_k, \varphi)|^2.$$

Im zweiten Falle ist φ durch $\mathbf{U} \varphi$ und ψ_k durch die zu λ_k gehörige Eigenfunktion von $\bar{\mathbf{G}}$, also durch $\mathbf{U} \psi_k$ zu ersetzen. Wir erhalten so für die Wahrscheinlichkeit im „zweiten Koordinatensystem“

$$|(\mathbf{U} \psi_k, \mathbf{U} \varphi)|^2$$

also wegen der Unitarität von \mathbf{U} dasselbe, was wir vorher erhalten haben. Natürlich sind auch die Übergangswahrscheinlichkeiten zwischen zwei einander entsprechenden Zuständen φ_1, φ_2 bzw. $\mathbf{U} \varphi_1, \mathbf{U} \varphi_2$ in beiden Koordinatensystemen dieselben: wegen

$$(\mathbf{U} \varphi_1, \mathbf{U} \varphi_2) = (\varphi_1, \varphi_2) \text{ ist } |(\mathbf{U} \varphi_1, \mathbf{U} \varphi_2)|^2 = |(\varphi_1, \varphi_2)|^2.$$

Einen solchen Übergang zu einem anderen Koordinatensystem durch eine Ähnlichkeitstransformation der Operatoren und gleichzeitiges Ersetzen der Wellenfunktion φ durch $\mathbf{U} \varphi$ bezeichnet man vielfach als eine kanonische Transformation. Zwei Beschreibungsweisen, die durch eine kanonische Transformation auseinander hervorgehen, sind einander äquivalent. Umgekehrt läßt sich zeigen, daß zwei quantenmechanische Beschreibungsweisen, die einander äquivalent sind, durch eine kanonische Transformation ineinander überführt werden können.

¹⁾ Die Unitarität eines Operators \mathbf{U} wird analog zur Hermitizität definiert: es wird verlangt, daß für beliebige zwei Funktionen f und g

$$(f, g) = (\mathbf{U} f, \mathbf{U} g)$$

gelte. Sind f und g Vektoren, \mathbf{U} also eine Matrix, so ist dies die notwendige und hinreichende Bedingung für ihre Unitarität.

4. Wir wollen ein Beispiel für die Anwendung der Transformationstheorie und der statistischen Deutung durchrechnen und wählen hierzu die von Schrödinger gegebene Begründung für die Bedeutung der Absolutwertquadrate der Matrixelemente ($N = \frac{1}{8}f$ = Anzahl der Elektronen, x_1, x_2, \dots, x_N ihre X-Koordinaten)

$$(x_1 + x_2 + \dots + x_N)_{FE} = (\psi_F, (x_1 + x_2 + \dots + x_N) \psi_E) = X_{FE}. \quad (6)$$

Diese soll nach der Matrizentheorie für die durch in der X-Achse polarisiertes Licht bedingte Übergangswahrscheinlichkeit vom stationären Zustand ψ_E (mit der Energie E) in den stationären Zustand ψ_F (mit der Energie F)

$$\mathbf{H} \psi_E = E \psi_E; \quad \mathbf{H} \psi_F = F \psi_F \quad (7)$$

verantwortlich sein.

Dieser Begriff der durch Strahlung bedingten Übergangswahrscheinlichkeit hat mit dem soeben erörterten Begriff der durch Versuche bedingten Übergangswahrscheinlichkeit nichts zu tun. Der letztere ist aus der Gedankenbildung der statischen Deutung entstanden und gibt die etwas paradox klingende Wahrscheinlichkeit für das Vorhandensein des Zustandes φ' an, wenn der Zustand φ ist. Sie ist eine dimensionslose Größe. Der jetzt wichtige Begriff gibt die Wahrscheinlichkeit an, daß in der nächsten Sekunde das Atom vom Zustand ψ_E unter Absorption des Lichtquantes $h\nu = F - E$ in den Zustand ψ_F übergeht. Sie hat die Dimension Sec^{-1} . Sie ist nur für Übergänge zwischen zwei stationären Zuständen (Eigenfunktionen des Hamiltonschen Operators \mathbf{H}) sinnvoll, während die erstere für beliebige Zustände φ, φ' definiert war. Da sie sich auf einen im Laufe der Zeit abspielenden Vorgang bezieht, muß sie durch die zeitabhängige Schrödingergleichung erfaßt werden können.

In allen Einzelheiten ist dies zwar nicht der Fall, weil die zeitabhängige Schrödingergleichung die spontane Emission nicht zu erklären vermag: nach ihr sind die Atome auch in angeregten Zuständen (wie etwa ψ_F) beliebig lange Zeit stabil, $\varphi = \psi_F \cdot e^{\frac{2\pi i F}{\hbar} t}$ ist eine Lösung der Gleichung (1). Doch enthält die Schrödingergleichung die Absorption (wie auch die Einsteinsche negative Einstrahlung), so daß wir jedenfalls zu richtigen Resultaten kommen müssen, solange die spontane Emission keine wesentliche Rolle spielt, solange sich das Atom fast im Normalzustand ψ_E befindet.

58 VI. Behandl. d. Lichtabsorption m. d. zeitabhängigen Schrödinger-Gleichung

Diese Annahme wird sich später auch aus rechnerischen Gründen notwendig erweisen, sie ist gerechtfertigt, wenn das Atom anfangs im untersten Zustande ψ_E war, man sich auf verhältnismäßig kurze Zeiten beschränkt und die Intensität der auffallenden Lichtwelle nicht extrem groß ist (was auch praktisch kaum erreicht werden könnte).

Unser Weg zur Behandlung des Absorptionsprozesses ist nun mehr vorgezeichnet: Wir nehmen an, daß zur Zeit $t = 0$ der Zustand des Systems $\varphi(0) = \psi_E$ war und sich dann nach der Gleichung

$$\frac{\hbar}{2\pi i} \frac{\partial \varphi}{\partial t} = \mathbf{H}\varphi = \mathbf{H}_0\varphi + \mathbf{H}_1\varphi \quad (8)$$

andert, wo \mathbf{H}_0 der Hamiltonsche Operator wäre, wenn kein Licht auf das Atom fallen würde, und \mathbf{H}_1 der durch das Licht bedingte Zusatzoperator ist. Das Licht bedeutet nämlich ein elektrisches Wechselfeld

$$\mathfrak{E}_x = P \sin 2\pi\nu t; \quad \mathfrak{E}_y = 0; \quad \mathfrak{E}_z = 0 \quad (9)$$

(die Abhängigkeit der Feldstärke von der Koordinate können wir wegen der Kleinheit der Atome vernachlässigen), wodurch die potentielle Energie V in (2) durch

$$V + \mathbf{H}_1 = V + Pe(x_1 + x_2 + \dots + x_N) \sin 2\pi\nu t \quad (8a)$$

ersetzt wird. Durch dieses Zusatzpotential, das als Störung behandelt wird, wird der Zeitablauf der Wellenfunktion, der bei $P = 0$

$$\varphi = \psi_E e^{\frac{2\pi i E}{\hbar} t} \quad (10)$$

wäre, modifiziert und wir haben die Gleichung

$$\frac{\hbar}{2\pi i} \frac{\partial \varphi}{\partial t} = \mathbf{H}_0\varphi + Pe \sin 2\pi\nu t (x_1 + x_2 + \dots + x_N)\varphi. \quad (11)$$

Um diese Gleichung zu lösen, entwickeln wir φ nach dem vollständigen System von Eigenfunktionen von \mathbf{H}_0

$$\varphi(t) = a_E(t)\psi_E + a_F(t)\psi_F + a_G(t)\psi_G + \dots, \quad (12)$$

wo die a_E, a_F, a_G nicht mehr von den Koordinaten x_1, x_2, \dots , die $\psi_E, \psi_F, \psi_G, \dots$ dagegen nicht von der Zeit t abhängen. Dann können wir einen Zustand anstatt mit seiner Wellenfunktion φ mit den Entwicklungskoeffizienten a_E, a_F, a_G, \dots dieser charakterisieren.

VI. Behandl. d. Lichtabsorption m. d. zeitabhängigen Schrödinger-Gleichung 59

Die Absolutwertquadrate $|\alpha_E|^2, |\alpha_F|^2, |\alpha_G|^2, \dots$ dieser Größen geben dann die Wahrscheinlichkeiten der verschiedenen Anregungsstufen des Atoms an. Wird das Atom durch keine Lichtwelle getroffen, so bleiben diese Wahrscheinlichkeiten zeitlich konstant, und da anfangs nur $|\alpha_E|^2 = 1$ von Null verschieden war, wird dies für alle Zeiten gelten. Fällt dagegen eine Lichtwelle auf das Atom, so müssen nach einer bestimmten Zeit auch die höheren Zustände angeregt erscheinen. Diese Anregungsstärken wollen wir jetzt berechnen.

Dabei nehmen wir an, daß zur Zeit $t = 0$

$$\alpha_E(0) = 1, \quad \alpha_F(0) = 0, \quad \alpha_G(0) = 0, \dots$$

ist und daß die Frequenz ν des Lichtes näherungsweise mit der Sprungfrequenz

$$\frac{1}{\hbar} (F - E) \approx \nu \quad (*)$$

übereinstimmt.

Setzen wir den Ausdruck (12) für φ in (11) ein, so erhalten wir eine Differentialgleichung für die Zeitabhängigkeit der $\alpha_E, \alpha_F, \alpha_G, \dots$. Da uns in erster Linie $\alpha_F(t)$ interessiert, bilden wir das skalare Produkt dieser Gleichung mit ψ_F , es bleibt dann links wegen der Orthogonalitätsrelationen der Eigenfunktionen von H_0 [(12) Kap. IV] nur das Glied mit α_F stehen und man erhält

$$\frac{\hbar}{2\pi i} \frac{\partial \alpha_F(t)}{\partial t} = F \alpha_F + P \sin 2\pi\nu t (\mathbf{X}_{FE} \alpha_E + \mathbf{X}_{FF} \alpha_F + \mathbf{X}_{FG} \alpha_G + \dots), \quad (13)$$

wo nach (6) für das Integral

$$(\psi_F, (x_1 + x_2 + \dots + x_N) \psi_E) = \mathbf{X}_{FE} \quad (6)$$

usw. eingesetzt ist.

Die Glieder rechts in (13) sind sehr verschiedener Größenordnung. Die Energie E hat die Größenordnung einiger Volt. Dagegen ist es schon ein sehr intensiver monochromatischer Lichtstrahl, bei dem P die Amplitude des elektrischen Vektors 10^{-2} Volt/cm ist. Die Matrixelemente von X sind etwa 10^{-8} cm, so daß $P e X \sim 10^{-10}$ Volt ist. Wir können daher im zweiten Glied rechts in (13), das schon ohnehin klein ist,

$$\alpha_E = e^{2\pi i \frac{E}{\hbar} t}; \quad \alpha_F = 0; \quad \alpha_G = 0; \dots$$

also die Näherungswerte (10) einsetzen, die diese Größen haben würden, wenn wir das zweite Glied rechts in (13) überhaupt vernachlässigen würden. Wir erhalten so

$$\frac{h}{2\pi i} \frac{\partial a_F(t)}{\partial t} = F a_F(t) + P e X_{FE} \sin 2\pi\nu t e^{2\pi i \frac{E}{\hbar} t}. \quad (14)$$

Um diese Gleichung zu integrieren, setzen wir

$$a_F(t) = b(t) e^{2\pi i \frac{F}{\hbar} t}$$

dann ist

$$e^{2\pi i \frac{F}{\hbar} t} \frac{h}{2\pi i} \frac{\partial b(t)}{\partial t} = \frac{P e}{2i} X_{FE} \left(e^{2\pi i \left(\frac{E}{\hbar} + \nu\right)t} - e^{2\pi i \left(\frac{E}{\hbar} - \nu\right)t} \right)$$

und daraus ergibt sich mit $e^{-2\pi i \frac{F}{\hbar} t}$ erweitert und integriert

$$\frac{h}{2\pi i} b(t) = \frac{P e}{2i} X_{FE} \left(\frac{e^{2\pi i [\nu - (F-E)/\hbar]t}}{2\pi i [\nu - (F-E)/\hbar]} - \frac{e^{2\pi i [-\nu - (F-E)/\hbar]t}}{2\pi i [-\nu - (F-E)/\hbar]} + C \right).$$

Die Integrationskonstante C bestimmt sich aus der Bedingung $b(0) = 0$, wir können sie in zwei Teile zerlegen und nach Multiplikation mit $2\pi i/\hbar$

$$b(t) = \frac{P e}{2i} X_{FE} \left(\frac{e^{2\pi i [\nu - (F-E)/\hbar]t} - 1}{h\nu - F + E} + \frac{e^{2\pi i [-\nu - (F-E)/\hbar]t} - 1}{h\nu + F - E} \right) \quad (15)$$

schreiben. Dieser Ausdruck verschwindet ja für $t = 0$ und wir sehen, daß $b(t)$ eine Summe von zwei periodischen Funktionen ist.

Wenn wir die Intensität P^2 des Lichtes konstant lassen und seine Frequenz ν ändern, so sehen wir, daß das erste Glied von (15) sehr stark ansteigt, wenn $h\nu$ nahezu $F - E$ wird. Nur in diesem Falle gelingt es überhaupt, eine einigermaßen merkliche Anregung zu bekommen und wir haben hiermit eine Erklärung für die Bohrsche Frequenzbedingung gefunden: die Frequenz des Lichtes, das einen beträchtlichen Übergang vom Zustand mit der Energie E in den Zustand F erzwingen soll, muß der Bedingung (*) genügen.

Wenn wir dies fortan annehmen, können wir in (15) das zweite Glied neben dem ersten vernachlässigen. Für die Wahrscheinlichkeit $|a_F(t)|^2 = |b(t)|^2$ des Zustandes F erhalten wir dann

$$|b(t)|^2 = \frac{P^2 e^2}{2} |X_{FE}|^2 \frac{1 - \cos 2\pi [\nu - (F-E)/\hbar]t}{(h\nu - F + E)^2}. \quad (15a)$$

5. Wir haben bisher vorausgesetzt, daß die Lichtwelle, die das Atom zur Zeit $t = 0$ trifft, die Gestalt einer reinen Sinusschwingung hat. Gewöhnlich ist dies nicht der Fall und das Licht besteht aus einer Superposition von vielen Sinusschwingungen, deren Frequenzen ein ganzes Intervall um $\nu = (F - E)/\hbar$ ungefähr gleichmäßig bedecken und deren Phasen statistisch unabhängig verteilt sind. Wegen dieses letzteren Umstandes kann man annehmen, daß sich die Wirkungen dieser Schwingungen additiv zusammensetzen und die Gesamtwahrscheinlichkeit, daß das Atom zur Zeit t im Zustand F sei,

$$|\mathbf{b}(t)|^2 = \sum_{\nu} |\mathbf{X}_{FE}|^2 \frac{P_{\nu}^2 e^2}{2} \frac{1 - \cos 2\pi [\nu - (F - E)/\hbar] t}{(\hbar\nu - F + E)^2} \quad (16)$$

ist, wo ν die Frequenzen aller im Licht auftretenden Sinusschwingungen durchläuft und P_{ν} die Amplitude der Schwingung mit der Frequenz ν ist.

Sind die Frequenzen der einzelnen Sinusschwingungen sehr nahe zueinander und bedecken sie ein gewisses um $(F - E)/\hbar$ gelegenes Gebiet — deren untere und obere Grenze ν_1 bzw. ν_2 sei — gleichmäßig und sehr dicht, so kann man für $P_{\nu}^2 = 8\pi J d\nu$ schreiben, wo J die Intensität (Energiedichte) des Lichtes pro Frequenzeinheit und $d\nu$ der Abstand zweier aufeinanderfolgender Frequenzen ist. Aus (16) wird dann in bekannter Weise das Integral

$$|\mathbf{b}(t)|^2 = 4\pi e^2 J |\mathbf{X}_{FE}|^2 \int_{\nu_1}^{\nu_2} \frac{1 - \cos 2\pi t [\nu - (F - E)/\hbar]}{(\hbar\nu - F + E)^2} d\nu \quad (16a)$$

oder wenn man an Stelle von ν als Integrationsvariable

$$x = 2\pi t [\nu - (F - E)/\hbar]$$

einführt:

$$|\mathbf{b}(t)|^2 = \frac{8\pi^2}{\hbar^2} e^2 J t |\mathbf{X}_{FE}|^2 \int_{x_1}^{x_2} \frac{1 - \cos x}{x^2} dx, \quad (16b)$$

wo die Integrationsgrenzen

$$x_1 = 2\pi t [\nu_1 - (F - E)/\hbar], \quad x_2 = 2\pi t [\nu_2 - (F - E)/\hbar] \quad (\dagger)$$

sein sollten. Da aber der Integrand von (16b) nur in einem schmalen Gebiet um $x = 0$ wesentliche Werte annimmt, kann man

die Integration von $-\infty$ bis ∞ erstrecken und erhält für die Wahrscheinlichkeit des Zustandes mit der Energie F

$$|\psi(t)|^2 = \frac{8\pi^3 e^3}{h^2} J t |\mathbf{X}_{FE}|^2. \quad (17)$$

Diese Verschiebung der Integrationsgrenzen ist eigentlich nur erlaubt wenn x_2 und $-x_1$ groß sind, was nach (†) darauf hinausläuft, daß die eingestrahlte Linie sich nach beiden Seiten von $(F-E)/h$ über ein Frequenzgebiet erstrecken muß, das groß gegen $1/2\pi t$ ist. Da unsere Rechnungen andererseits nur für Zeiten, die kurz sind gegen die Lebensdauer τ des Zustandes F , Gültigkeit beanspruchen können, besagt dies, daß die Breite der eingestrahlten Linie allenfalls sehr groß gegen die „natürliche Linienbreite“ $1/\tau$ angenommen werden muß.

Die Wahrscheinlichkeit, daß das Atom im Zustande mit der Energie F sei, ist nach (17) proportional zur Intensität J des eingestrahlten Lichtes, zum Quadrat des Matrixelements $|\mathbf{X}_{FE}|^2$, womit wir die matrizentheoretische Aussage tatsächlich bestätigen konnten — und zur Einwirkungsdauer t der Lichtwelle — wie erwartet werden mußte. Allerdings gilt (17) nur für Zeiten, die kurz gegen die Lebensdauer des angeregten Zustandes und lang gegen die reziproke Breite der eingestrahlten Linie sind.

Trotz dieses Umstandes und ihren nur näherungsweisen Gültigkeit liefert (17) eine sehr schöne Bestätigung der Annahme, daß $|\alpha_F|^2$ die Anregungsstärke des Zustandes mit der Energie F ist, sie ist — zusammen mit der Vorstellung des Wahrscheinlichkeitspaketes im Konfigurationsraum — eine der stärksten Stützen der statistischen Deutung der Quantenmechanik.

Wir haben (17) aber auch noch aus einem anderen Grunde abgeleitet, sie zeigt uns nämlich, daß

$$|\mathbf{X}_{FE}|^2 = |(\psi_F, (x_1 + x_2 + \dots + x_N) \psi_E)|^2$$

proportional der durch zur X -Achse parallel polarisiertes Licht erzwungenen Übergangswahrscheinlichkeit vom stationären Zustand ψ_E in den stationären Zustand ψ_F ist. Diese Tatsache — sie wurde verschiedentlich weit exakter begründet, als wir es hier taten — wird in der Folge die Grundlage der Berechnung der Intensitäten bzw. Intensitätsverhältnisse der Spektrallinien bilden.

VII. Abstrakte Gruppentheorie

Betrachten wir die sechs Matrizen

$$\left. \begin{array}{c} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \quad \begin{pmatrix} -\frac{1}{2} & \frac{1}{2}\sqrt{3} \\ \frac{1}{2}\sqrt{3} & \frac{1}{2} \end{pmatrix}, \quad \begin{pmatrix} -\frac{1}{2} & -\frac{1}{2}\sqrt{3} \\ -\frac{1}{2}\sqrt{3} & \frac{1}{2} \end{pmatrix}, \\ \mathbf{E} \qquad \mathbf{A} \qquad \mathbf{B} \qquad \mathbf{C} \\ \begin{pmatrix} -\frac{1}{2} & \frac{1}{2}\sqrt{3} \\ -\frac{1}{2}\sqrt{3} & -\frac{1}{2} \end{pmatrix}, \quad \begin{pmatrix} -\frac{1}{2} & -\frac{1}{2}\sqrt{3} \\ \frac{1}{2}\sqrt{3} & -\frac{1}{2} \end{pmatrix} \\ \mathbf{D} \qquad \mathbf{F} \end{array} \right\} (\dagger)$$

und bilden wir nach den Regeln der Matrizenmultiplikation die 36 Produkte, welche entstehen, wenn man jede der sechs Matrizen (\dagger) mit jeder Matrix (\dagger) multipliziert! Wir sehen, daß alle 36 so entstehenden Matrizen mit einer schon in (\dagger) vorkommenden Matrix identisch sind. Ein solches System bezeichnet man als eine Gruppe. Wir können diese Eigenschaft dieser Matrizen in einer Tabelle, der Gruppentafel, zusammenfassen:

	E	A	B	C	D	F
E	E	A	B	C	D	F
A	A	E	D	F	B	C
B	B	F	E	D	C	A
C	C	D	F	E	A	B
D	D	C	A	B	F	E
F	F	B	C	A	E	D

In der ersten Spalte steht der erste Faktor, in der ersten Zeile der zweite Faktor, und in der Tabelle selbst steht das Produkt. In dieser Tabelle sind alle Multiplikationsregeln der Matrizen (\dagger) vereinigt.

Die strenge Definition der Gruppe ist folgende:

Unter Gruppe versteht man eine Gesamtheit von Objekten, den Elementen der Gruppe, zwischen denen eine Art Verknüpfung, Multiplikation genannt, eindeutig definiert ist. Diese Multiplikation definiert zu je zwei Elementen der Gruppe (Faktoren) ein drittes Element der Gruppe, das Produkt¹⁾. Diese Gruppen-

¹⁾ Man denke im folgenden an ein System von n -zeiligen Matrizen.

multiplikation, die als eine den Gruppenelementen innewohnende Eigenschaft angesehen werden soll, muß ferner folgende Merkmale aufweisen.

1. Das Assoziativgesetz muß gelten, d. h. wenn $A B = F$ und $B C = G$ ist, so ist $F C = A G$. Sind die Gruppenelemente Matrizen und verstehen wir unter Gruppenmultiplikation Matrizenmultiplikation, so ist das Assoziativgesetz immer erfüllt. (Nach Satz 3 des I. Kapitels.) Gruppen, in denen auch das kommutative Gesetz der Multiplikation gilt, d. h. $A B = B A$ ist, nennt man Abelsche Gruppen.

2. Unter den Elementen ist eines (und nur eines) vorhanden, Einheit E genannt, das die Eigenschaft hat, daß sein Produkt mit einem beliebigen anderen Element das andere Element ergibt, d. h. $A E = E A = A$ gilt.

3. Jedes Element hat eine Reziproke, d. h. zu jedem A gibt es ein B , so daß $B A = E$. Dann können wir, wie in Kapitel I, Satz 6, zeigen, daß auch $A B = E$ ist. Aus $B A = E$ folgt $B A B = B$; ist C die Reziproke von B , so folgt weiter $C B A B = C B$, d. h. $A B = E$. Die Reziproke von A bezeichnet man mit A^{-1} .

Diese drei Eigenschaften der Gruppenelemente und der Gruppenmultiplikation dienen zur Definition der Gruppe, man nennt sie auch — so, oder etwas anders formuliert — Axiome der Gruppe oder Gruppenpostulate.

Regel. Die Reziproke eines Produkts $A B C D \dots$ bildet man — ebenso, wie bei Matrizen —, indem man die Reziproken der einzelnen Faktoren in umgekehrter Reihenfolge miteinander multipliziert. Es ist also

$$(A B C D \dots)^{-1} = \dots D^{-1} C^{-1} B^{-1} A^{-1},$$

in der Tat ist

$$\dots D^{-1} C^{-1} B^{-1} A^{-1} A B C D \dots = E.$$

Es ist zu beachten, daß etwa aus $A X = B$ und $A Y = B$ auch $X = Y$ folgt. In der Tat sind beide $X = Y = A^{-1} B$. Auch aus $X A = B$ und $Y A = B$ folgt $X = Y = B A^{-1}$.

Hat die Gruppe nur eine endliche Anzahl h von Elementen, so nennt man die Gruppe eine endliche und bezeichnet h als die Ordnung der Gruppe.

Sätze über endliche Gruppen¹⁾

Fassen wir ein Element X ins Auge. Dann können wir daraus der Reihe nach die Elemente

$$E = X^0, X = X^1, X^2, X^3, \dots \quad (\dagger\dagger)$$

usw. bilden. Da die Elemente von $(\dagger\dagger)$ alle Gruppenelemente sind, die Anzahl aller Elemente aber endlich ist, so muß nach einer gewissen Anzahl von Potenzen eine schon in der Reihe $(\dagger\dagger)$ vorgekommene wieder vorkommen. Sei das erste solche Element $X^n = X^k$ (mit $k < n$), so muß $k = 0$ und $X^n = E$ sein, denn sonst wäre schon $X^{n-1} = X^{k-1}$ in der Reihe vorgekommen und X^n wäre nicht das erste Element, das in der Reihe $(\dagger\dagger)$ zum zweiten Male vorkommt. Ist n die kleinste Zahl, für die $X^n = E$ gilt, so nennt man n die Ordnung von X . Die Reihe

$$E, X, X^2, X^3, \dots, X^{n-1} \quad (\dagger\dagger\dagger)$$

nennt man die Periode von X . Die Periode von D in der Gruppe (\dagger) ist z. B. $E, D, D^2 = F$ ($D^3 = FD = E$), die Ordnung von D ist also 3. Die Periode von F ist $E, F, F^2 = D$, ($F^3 = DF = E$), die Ordnung von F ist also ebenfalls 3. Die Ordnung von A dagegen ist 2, da bereits $A^2 = E$ gilt.

Die Periode von X $(\dagger\dagger\dagger)$ bildet selber eine Gruppe (und zwar eine Abelsche Gruppe). Eine Gesamtheit von Elementen einer Gruppe, die selber eine Gruppe bildet, nennt man eine Untergruppe. $(\dagger\dagger\dagger)$ ist eine Abelsche Untergruppe.

Satz 1. Ist \mathfrak{H} eine Gruppe von der Ordnung h mit den Elementen E, A_2, A_3, \dots, A_h , und ist A_k ein beliebiges Element dieser Gruppe, so kommt in der Reihe $E A_k = A_k, A_2 A_k, A_3 A_k, \dots, A_h A_k$ jedes Element einmal und nur einmal vor. In der Tat sei X ein Element und $X A_k^{-1} = A_r$, so ist $A_r A_k = X$, d. h. X kommt in der Reihe vor. Zweimal kann X dagegen nicht vorkommen, da aus $A_r A_k = X$ und $A_s A_k = X$ auch $A_r = A_s$ folgt.

Dasselbe gilt natürlich von der Reihe $A_k E, A_k A_2, A_k A_3, \dots, A_k A_h$. Satz 1 bringt zum Ausdruck, daß in der Gruppentafel in jeder Spalte (und ebenso in jeder Zeile) jedes Element einmal und nur einmal vorkommt. Die einfachste und wichtigste Anwendung dieses Satzes ist die folgende: Sind $J_E, J_{A_2}, J_{A_3}, \dots, J_{A_h}$ Zahlen,

¹⁾ Man verifiziere alle Sätze an der Gruppe (\dagger) . Man bediene sich hierzu der Gruppentafel!

so daß jedem Gruppenelement X eine Zahl J_X zugeordnet ist, dann ist

$$\sum_{v=1}^h J_{A_v} = \sum_{v=1}^h J_{A_v X} = \sum_{v=1}^h J_{X A_v}. \quad (1)$$

In den Summen rechts kommen nämlich genau dieselben Zahlen wie links vor, und jede einmal, nur in anderer Reihenfolge.

\mathfrak{B} sei eine Untergruppe der Gruppe \mathfrak{H} mit den Elementen E, B_1, B_2, \dots, B_g . Man nennt die Gesamtheit der g Elemente $EX, B_1 X, B_2 X, \dots, B_g X$ eine rechtsseitige Nebengruppe $\mathfrak{B}X$, wenn X nicht in der Untergruppe \mathfrak{B} vorkommt¹⁾. (In diesem Falle wären nämlich die Elemente von $\mathfrak{B}X$ nach Satz 1 angewandt auf \mathfrak{B} genau diejenigen von \mathfrak{B} .) Eine Nebengruppe ist sicher keine Gruppe, sie enthält ja die Einheit E sicher nicht und auch kein anderes Element von \mathfrak{B} . Wäre nämlich etwa $B_k X = B_l$, so wäre $X = B_k^{-1} B_l$, d. h. X in der Untergruppe \mathfrak{B} enthalten und $\mathfrak{B}X$ wäre \mathfrak{B} selber. Entsprechend bilden die Elemente $XE = X, XB_1, XB_2, \dots, XB_g$ eine linksseitige Nebengruppe von \mathfrak{B} .

Satz 2. Zwei rechtsseitige Nebengruppen einer Untergruppe \mathfrak{B} enthalten entweder beide dieselben Elemente, oder sie haben kein gemeinsames Element. Die eine Nebengruppe sei $\mathfrak{B}X$, die andere $\mathfrak{B}Y$. Ist etwa $B_k X = B_l Y$, so ist $YX^{-1} = B_l^{-1} B_k$, d. h. YX^{-1} wäre in \mathfrak{B} enthalten. Dann würde nach Satz 1, angewandt auf die Untergruppe \mathfrak{B} , die Reihe $EYX^{-1}, B_1 YX^{-1}, B_2 YX^{-1}, \dots, B_g YX^{-1}$ bis auf die Reihenfolge mit E, B_1, B_2, \dots, B_g identisch. Also wäre auch $EYX^{-1}X, B_1 YX^{-1}X, B_2 YX^{-1}X, \dots, B_g YX^{-1}X$ bis auf die Reihenfolge mit $EX, B_1 X, B_2 X, \dots, B_g X$ identisch. Die erstere Reihe ist aber nichts anderes als $EY, B_1 Y, B_2 Y, \dots, B_g Y = \mathfrak{B}Y$. Die Elemente von $\mathfrak{B}Y$ stimmen also mit den Elementen von $\mathfrak{B}X$ überein, falls nur ein einziges Element übereinstimmt. Das Kriterium hierfür ist, daß YX^{-1} in \mathfrak{B} enthalten sei.

Eine Untergruppe von (\dagger) bildet z. B. die Periode von A , d. h. die beiden Elemente E und A . Eine rechtsseitige Nebengruppe dieser Gruppe erhalten wir, wenn wir jedes Element rechts mit einem anderen Element, etwa B multiplizieren. Wir erhalten so die Nebengruppe $EB = B$,

¹⁾ Dagegen natürlich ein Element von \mathfrak{H} ist.

$A B = D$. Die anderen Nebengruppen erhalten wir, wenn wir die Elemente E, A mit den anderen Elementen C, D, F multiplizieren.

Die Nebengruppe mit B ist $B, D,$
 C " $C, F,$
 D " $D, B,$
 F " $F, C.$

In diesem Falle also sind die Nebengruppen, die wir durch Multiplikation mit B und D (bzw. C und F) erhalten, identisch. In der Tat ist $B D^{-1} = B F = A$ (bzw. $C F^{-1} = C D = A$) in der Untergruppe E, A enthalten.

Betrachten wir nun alle voneinander verschiedenen Nebengruppen von \mathfrak{B} ! Diese seien $\mathfrak{B} X_2, \mathfrak{B} X_3, \mathfrak{B} X_4, \dots, \mathfrak{B} X_l$. Jedes Element von \mathfrak{H} kommt entweder in \mathfrak{B} oder in einer dieser $l - 1$ Nebengruppen vor. So erhalten wir im ganzen lg Elemente. Da aber kein Element auf diese Weise doppelt gezählt ist, muß $lg = h$ sein.

Satz 3. Die Ordnung g einer Untergruppe ist Teiler der Ordnung h der ganzen Gruppe. Den Quotienten $\frac{h}{g} = l$ bezeichnet man als den Index der Untergruppe \mathfrak{B} unter der Gruppe \mathfrak{H} .

Da die Periode eines jeden Elements eine Untergruppe mit so viel Elementen, als seine Ordnung beträgt, ist, folgt, daß die Ordnung eines jeden Elements Teiler der Ordnung der Gruppe ist.

Kriterium für Untergruppen. Enthält ein Komplex von Gruppenelementen alle Produkte $A B$ der in ihm enthaltenen Elemente A und B , so bildet es schon eine Gruppe, ist also eine Untergruppe der ursprünglichen Gruppe. Das Assoziativgesetz der Multiplikation gilt für alle Elemente der Gruppe, also auch für die Elemente des betrachteten Systems. Da mit jedem Element A alle seine Potenzen im System vorkommen, kommt auch die Einheit E vor. Ist n die Ordnung von A , so ist $A^n = E$. Da $A^{n-1} = A^{-1}$ ist, kommt auch die Reziproke eines jeden Elements im betrachteten System vor. Hiermit sind aber alle drei Gruppenpostulate erfüllt.

Beispiele von Gruppen

1. Die Gruppe, die nur ein Element enthält, besteht aus der Einheit E allein.

2. Die Gruppe von der Ordnung zwei hat die Gruppentafel:

	E	A
E	E	A
A	A	E

Sie ist eine abelsche Gruppe. Wir nennen sie Spiegelungsgruppe, weil die Einheit zusammen mit der Transformation der Spiegelung $x' = -x$ diese Gruppe bildet.

3. Die Gruppe von der Ordnung drei kann außer der Einheit nur Elemente enthalten, deren Ordnung 3 ist, da dies der einzige Teiler von 3 (außer 1) ist. Sie besteht aus einer einzigen Periode. Ihre Elemente sind:

$$E, A, A^2 (A^3 = E),$$

sie ist also abelsch.

Genau dasselbe gilt für jede Gruppe, deren Ordnung eine Primzahl p ist. Ihre Elemente sind

$$E, A, A^2, \dots, A^{p-1}.$$

Gruppen von dieser Form, auch wenn p keine Primzahl ist, heißen zyklische Gruppen. Wenn ω eine n -te primitive Einheitswurzel ist (also ω^n die niedrigste Potenz von ω , die gleich 1 ist,

etwa $\omega = \cos \frac{2\pi}{n} + i \sin \frac{2\pi}{n}$), so bilden die Zahlen

$$1, \omega, \omega^2, \dots, \omega^{n-1} \tag{*}$$

eine zyklische Gruppe von der Ordnung n , wenn man unter Gruppenmultiplikation gewöhnliche Zahlemultiplikation versteht. Die zyklischen Gruppen sind alle abelsch. Die „gleiche Gruppe“ wie (*) bilden auch die Zahlen

$$0, 1, 2, \dots, n-1, \tag{**}$$

wenn man unter Gruppenmultiplikation Addition und Restnehmen nach n versteht. (Ist z. B. $n = 7$, so ist $5 \cdot 4 = 2$, da $5 + 4 = 9 = 7 + 2 = n + 2$ ist.) Die Elemente der Gruppe (**) lassen sich den Elementen von (*) ein-eindeutig zuordnen, indem man k zu ω^k zuordnet. Diese Zuordnung hat die Eigenschaft, daß dabei „Produkt in Produkt“ übergeht, d. h. mit

$k_1 \cdot k_2 = k_3$ ist auch $\omega^{k_1} \cdot \omega^{k_2} = \omega^{k_3}$. Solche zwei Gruppen bezeichnet man als holomorphe Gruppen¹⁾.

Zwei Gruppen sind holomorph, wenn die Elemente A der einen Gruppe sich so zu den Elementen \bar{A} der anderen Gruppe eindeutig umkehrbar zuordnen lassen, daß aus $A \cdot B = C$ auch $\bar{A} \cdot \bar{B} = \bar{C}$ gefolgert werden kann, d. h. $\bar{A} \cdot \bar{B} = \bar{A} \cdot \bar{B}$ gilt. Holomorphe Gruppen sind wesensgleich, die einzelnen Elemente sind nur verschieden benannt.

4. Es existieren zwei Gruppen von der Ordnung vier, d. h. zwei Gruppen, von denen keine mit der anderen holomorph wäre. Alle anderen Gruppen dagegen sind mit einer dieser Gruppen holomorph. Die erste Gruppe ist die zyklische Gruppe: etwa $1, i, -1, -i$ mit Gruppenmultiplikation = Zahlenmultiplikation. Die zweite Gruppe ist die sogenannte Vierergruppe. Ihre Gruppentafel ist:

	E	A	B	C
E	E	A	B	C
A	A	E	C	B
B	B	C	E	A
C	C	B	A	E

alle ihre Elemente (außer E) sind von der Ordnung 2, sie ist noch abelsch.

5. Die Vierergruppe ist die erste Repräsentantin einer umfassenden Art von Gruppen, der Symmetriegruppen.

Betrachten wir ein reguläres n -Eck in der XY-Ebene. Die Koordinaten der n Eckpunkte seien $x_k = r \cos \frac{2\pi k}{n}$, $y = r \sin \frac{2\pi k}{n}$ ($k = 0, 1, 2, 3, \dots, n-1$). Betrachten wir alle linearen Substitutionen,

$$x' = \alpha x + \beta y; \quad y' = \gamma x + \delta y,$$

¹⁾ Um zu zeigen, daß die bisher abgeleiteten Sätze keineswegs mehr trivial sind, sei folgendes angeführt: Ist $n+1$ eine Primzahl, so bilden die Zahlen $1, 2, 3, \dots, n$ noch auf eine andere Weise eine Gruppe, wenn wir nämlich unter Gruppenmultiplikation: Zahlenmultiplikation und Restnehmen nach $n+1$ verstehen. Ist etwa $n+1 = 7$, so ist $3 \cdot 5 = 1$, da $3 \cdot 5 = 15 = 2 \cdot 7 + 1$ ist. Die Einheit ist dann die 1. Die Periode eines jeden Elements ist dann Teiler von der Ordnung der Gruppe, von n . Es ist also sicher $A^n = 1$, wenn A ein Element der Gruppe ist. Dies bedeutet aber so viel, daß $a^n \equiv 1 \pmod{n+1}$ ist, wenn a eine der Zahlen $1, 2, 3, \dots, n$ ist. Dies ist ein — wie man zugeben wird — keineswegs trivialer Spezialfall des Fermatschen Satzes.

die das reguläre n -Eck „in sich selbst überführen“, d. h. für die die neuen Koordinaten der Ecken x'_x, y'_x ebenfalls in der Gestalt

$$x'_x = r \cos \frac{2\pi x}{n}, \quad y'_x = r \sin \frac{2\pi x}{n}$$

($x = 0, 1, 2, 3, \dots, n-1$) geschrieben werden können. Die Matrizen dieser linearen Substitutionen werden eine Gruppe bilden, da mit zwei Substitutionen auch ihr Produkt, die Reziproke einer jeden, sowie offenbar die Einheit $E = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ diese Eigenschaften haben.

Diese Substitutionen, die das n -Eck in sich selbst überführen, sind:

1. Drehungen der Ebene mit dem Winkel $\frac{2\pi r}{n}$ ($r = 0, 1, 2, 3, \dots, n-1$); die entsprechenden Matrizen sind

$$\begin{pmatrix} \cos \frac{2\pi r}{n}, & \sin \frac{2\pi r}{n} \\ -\sin \frac{2\pi r}{n}, & \cos \frac{2\pi r}{n} \end{pmatrix} = D_r$$

Diese bilden eine zyklische Gruppe. 2. Umklappung der Ebene und darauf folgende Drehung um einen der Winkel $\frac{2\pi r}{n}$. Die entsprechenden Matrizen sind

$$\begin{pmatrix} -\cos \frac{2\pi r}{n}, & \sin \frac{2\pi r}{n} \\ \sin \frac{2\pi r}{n}, & \cos \frac{2\pi r}{n} \end{pmatrix} = U_r$$

Diese $2n$ Matrizen bilden eine Gruppe von der Ordnung $2n$, man nennt sie Diedergruppe. Die Matrizen von 1. bilden eine Untergruppe dieser Gruppe, die von 2. eine Nebengruppe hierzu. Die Vierergruppe ist das einfachste Beispiel mit $n = 2$ für eine solche Gruppe, das n -Eck artet in ein Zweieck, eine gerade Linie aus. Während die Vierergruppe noch abelsch ist, sind die anderen Diedergruppen nicht mehr abelsch. Die Gruppe (\dagger) ist die Diedergruppe des regulären Dreiecks, sie ist die erste nichtabelsche Gruppe, die Elemente E, F, D gehören in die Untergruppe 1; A, B, C in die Nebengruppe. Die Substitutionen, die reguläre Körper in sich überführen, sind wichtige und interessante Gruppen, man nennt sie Symmetriegruppen. Man bezeichnet sie gewöhnlich nach denjenigen regulären Körpern, den sie in sich überführen. So existiert eine Tetraedergruppe, Oktaedergruppe, Ikosaedergruppe usw. Sie spielen in der Kristallphysik eine wichtige Rolle.

6. Sehr wichtige Gruppen sind endlich die Permutationsgruppen. Betrachten wir die Zahlen von 1 bis n : 1, 2, 3, ..., n . Jede Reihenfolge $\alpha_1, \alpha_2, \alpha_3, \dots, \alpha_n$ dieser n Zahlen bildet eine

Permutation. Es gibt also insgesamt $n!$ Permutationen von n Dingen, man bezeichnet sie gewöhnlich

$$\begin{pmatrix} 1, & 2, & 3, & \dots, & n \\ \alpha_1, & \alpha_2, & \alpha_3, & \dots, & \alpha_n \end{pmatrix},$$

indem man die zu permutierenden Dinge in die obere Zeile in natürlicher Reihenfolge, in die untere Zeile in derjenigen Reihenfolge schreibt, die aus der erstenen durch die zu kennzeichnende Permutation entsteht. Die Multiplikation von zwei Permutationen P_1 und P_2 geschieht so, daß man in der Reihenfolge P_1 diejenigen Veränderungen vornimmt, die P_2 in der normalen Reihenfolge bewirken würde. So ist z. B.

$$P_1 = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix}, \quad P_2 = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix};$$

$$P_1 P_2 = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix} \cdot \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix}.$$

Es führt nämlich P_2 die 1 in die 3 über, folglich steht in $P_1 P_2$ dort, wo in P_1 die 1 steht, die 3. Ebenso führt P_2 die 2 in 1 über, wo in P_1 die 2 steht, steht in $P_1 P_2$ die 1, usw.

Führt P_1 die k in α_k , P_2 dann α_k in β_k und P_3 weiter β_k in γ_k über, so führt $P_1 P_2$ die k in β_k , $P_2 P_3$ aber α_k in γ_k über. Es führt also sowohl $(P_1 P_2) \cdot P_3$ wie auch $P_1 (P_2 P_3)$ die k in γ_k über, die Permutationenmultiplikation ist assoziativ.

Sämtliche $n!$ Permutationen von n Dingen bilden eine Gruppe mit der Einheit

$$\begin{pmatrix} 1, & 2, & 3, & \dots, & n \\ 1, & 2, & 3, & \dots, & n \end{pmatrix},$$

man nennt sie die symmetrische Gruppe¹⁾. Die symmetrische Gruppe dritten Grades ist von der Ordnung 6, sie ist holomorph zur Gruppe (\dagger), also zur Diedergruppe mit $n = 3$. Die Zuordnung ist folgende:

$$\begin{array}{ccccccc} \begin{pmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \end{pmatrix}, & \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix}, & \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix}, & \begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix}, & \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix}, & \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix}. \\ E & A & B & C & D & F \end{array}$$

Die symmetrischen Gruppen spielen auch in der Quantenmechanik eine wichtige Rolle.

¹⁾ Vielfach auch Permutationsgruppe, keine Symmetriegruppe.

Konjugierte Elemente und Klassen

Das Element XAX^{-1} bezeichnet man als ein zu A konjugiertes Element. Sind zwei Elemente A und B zu einem dritten C konjugiert, so sind sie es auch untereinander. Aus $A = XCX^{-1}$ und $B = YCY^{-1}$ folgt $X^{-1}AX = C$ und $B = YX^{-1}AXY^{-1} = YX^{-1}A(YX^{-1})^{-1}$. Die zueinander konjugierten Elemente einer Gruppe bilden zusammen eine Klasse. Eine Klasse ist durch Vorgabe eines einzigen A ihrer Elemente bestimmt: Man erhält die ganze Klasse, indem man die Reihe

$$EAE^{-1} = A, A_2 A A_2^{-1}, A_3 A A_3^{-1}, \dots, A_h A A_h^{-1}$$

bildet. Alle Elemente dieser Reihe sind zu A , also auch untereinander konjugiert und es kommen alle Elemente, die zu A (also auch alle, die zu einem beliebigen anderen Element der Reihe) konjugiert sind, in der Reihe auch wirklich (sogar mehrfach) vor.

Man kann daher die Elemente einer Gruppe in Klassen einteilen, wobei jedes Element in einer und nur in einer Klasse vorkommt.

Die Einheit der Gruppe bildet eine Klasse für sich, da sie mit keinem anderen Element konjugiert ist: $XE X^{-1} = E$ gilt für alle X . Abgesehen von der Klasse, die nur aus dem Element E besteht, ist keine Klasse eine Untergruppe, da sonst keine die Einheit E enthält.

Alle Elemente einer Klasse haben die gleiche Ordnung. Ist nämlich $A^n = E$, so ist auch

$$\begin{aligned}(XAX^{-1})^n &= XAX^{-1} \cdot XAX^{-1} \cdot XAX^{-1} \dots XAX^{-1} \\ &= XA^nX^{-1} = XEX^{-1} = E.\end{aligned}$$

In einer Substitutionsgruppe haben alle Matrizen, die in dieselbe Klasse gehören, dieselbe Spur. Gehören nämlich die Matrizen α und β in dieselbe Klasse, so existiert ein Gruppenelement, eine Matrix γ , so daß

$$\beta = \gamma \alpha \gamma^{-1}.$$

Folglich ist die Spur $\beta = \text{Spur } \gamma \alpha \gamma^{-1} = \text{Spur } \alpha$.

Bilden wir z. B. die Klasse von C in der Gruppe (\dagger). Wir haben

$$\begin{aligned}ECE^{-1} &= C, \quad ACA^{-1} = B, \quad BCB^{-1} = A, \quad CCC^{-1} = C, \\ DCD^{-1} &= A, \quad FCF^{-1} = B,\end{aligned}$$

die Klasse von C besteht also aus den Elementen A, B, C , dies ist auch die Klasse von A oder B . Alle drei Elemente A, B, C sind von der

Ordnung 2, die Spur in ihrer Matrixdarstellung (\dagger) ist bei allen dreien 0. Die Klasse von D ist

$$EDE^{-1} = D, \quad ADA^{-1} = F, \quad BDB^{-1} = F, \quad CDC^{-1} = F,$$

$$DDD^{-1} = D, \quad FDF^{-1} = D,$$

die Klasse von D (oder F) besteht aus den beiden Elementen D, F . In einer Abelschen Gruppe besteht jede Klasse aus einem einzigen Element, da $XAX^{-1} = A$ ist.

VIII. Normalteiler

Besteht eine Untergruppe aus lauter ganzen Klassen der ursprünglichen Gruppe, so nennt man sie einen Normalteiler. Ist $\mathfrak{N} = E, N_2, N_3, \dots, N_n$ ein Normalteiler, so ist in ihm mit N_i und N_j auch $N_i N_j$ enthalten, weil er eine Gruppe ist. Außerdem ist aber auch $X N_i X^{-1}$ in ihm enthalten, wo X ein beliebiges Element der ganzen Gruppe ist, weil der Normalteiler alle Elemente $X N_i X^{-1}$ einer Klasse enthält, wenn er eines ihrer Elemente, N_i enthält. Gewöhnliche Untergruppen müssen $X N_i X^{-1}$ natürlich nur dann enthalten, wenn auch X eines ihrer Elemente ist.

Bei gewöhnlichen Untergruppen (wie bei jeder Gruppe) sind die Elemente

$$N_j E = N_j, \quad N_j N_2, \quad N_j N_3, \quad \dots, \quad N_j N_n \quad (a)$$

mit den Elementen der Untergruppe bis auf die Reihenfolge identisch. Dasselbe gilt von der Reihe

$$EN_j^{-1} = N_j^{-1}, \quad N_2 N_j^{-1}, \quad N_3 N_j^{-1}, \quad \dots, \quad N_n N_j^{-1}, \quad (b)$$

also auch von der Reihe

$$N_j EN_j^{-1} = E, \quad N_j N_2 N_j^{-1}, \quad N_j N_3 N_j^{-1}, \quad \dots, \quad N_j N_n N_j^{-1}, \quad (c)$$

die ja aus (b) entsteht, wenn man die ersten Faktoren E, N_2, N_3, \dots, N_n durch die Reihe (a), die mit ihnen bis auf die Reihenfolge identisch ist, ersetzt. Dabei ist N_j natürlich ein Element der Untergruppe.

Ist aber E, N_2, N_3, \dots, N_n ein Normalteiler, so ist, auch wenn X ein beliebiges Element der ganzen Gruppe ist, die Reihe

$$XE X^{-1} = E, \quad XN_2 X^{-1}, \quad XN_3 X^{-1}, \quad \dots, \quad XN_n X^{-1} \quad (d)$$

bis auf die Reihenfolge mit den Elementen des Normalteilers E, N_2, N_3, \dots, N_n identisch. Einerseits sind alle Elemente von (d) im Normalteiler enthalten — da sie durch Konjugation aus den Elementen dieses entstehen —, andererseits kommen alle Elemente des Normalteilers in (d) vor. Um N_k in der Reihe (d) zu finden, müssen wir nur $X^{-1} N_k X$ bilden. Da dies auch ein Element des Normalteilers ist, kommt es in der Reihe E, N_2, N_3, \dots, N_n auch vor. Es sei etwa N_i . Dann ist $N_k = X N_i X^{-1}$, so daß N_k in (d) an der i -ten Stelle steht.

Alle Untergruppen einer Abelschen Gruppe sind gleichzeitig Normalteiler. Da nämlich jedes Element eine ganze Klasse ist, besteht auch jede Untergruppe aus lauter ganzen Klassen. Die symmetrischen Gruppen haben einen und im allgemeinen nur einen Normalteiler, der aus allen geraden Permutationen gebildet wird. Diese bilden nämlich erstens eine Untergruppe, da das Produkt von zwei geraden Permutationen wieder eine gerade Permutation ist. Zweitens ist aber das konjugierte Element zu einer geraden Permutation wieder eine gerade Permutation, also unter allen geraden Permutationen enthalten. Vgl. auch Kap. XII.

In Beispiel (†) bilden die Elemente E, D, F diesen Normalteiler. Verifizieren wir an diesem die folgenden Sätze!

Die Aufsuchung der Normalteiler ist für die Konstitution der Gruppe sehr wichtig. Gruppen, die keinen Normalteiler haben, heißen einfach.

Betrachten wir nun die Nebengruppen des Normalteilers \mathfrak{N} . Eine rechtsseitige Nebengruppe mit U , die Elemente $U, N_2 U, N_3 U, \dots, N_n U$, bilden gleichzeitig die linksseitige Nebengruppe mit U , da $U = U \cdot U^{-1} E U, N_2 U = U \cdot U^{-1} N_2 U, N_3 U = U \cdot U^{-1} N_3 U, \dots, N_n U = U \cdot U^{-1} N_n U$ ist und die Elemente $E = U^{-1} E U, N'_2 = U^{-1} N_2 U, N'_3 = U^{-1} N_3 U, \dots, N'_n = U^{-1} N_n U$ mit den Elementen E, N_2, N_3, \dots, N_n bis auf die Reihenfolge identisch sind. Der Komplex $\mathfrak{N} U$ stimmt mit dem Komplex $U \mathfrak{N}$ überein. Man kann also schlechtweg von einer Nebengruppe des Normalteilers sprechen, ohne anzugeben, ob sie rechts- oder linksseitig ist¹⁾. Multiplizieren wir alle Elemente einer Neben-

¹⁾ Rechnerisch läßt sich das folgendermaßen einsehen: Daß $U V^{-1}$ in \mathfrak{N} sei, ist die Bedingung dafür, daß U und V in derselben rechtsseitigen Nebengruppe seien, daß $V^{-1} U$ in \mathfrak{N} sei, die Bedingung, daß sie in derselben linksseitigen Nebengruppe seien. Ist $U V^{-1}$ in \mathfrak{N} , so ist auch $V^{-1} \cdot U V^{-1} \cdot V = V^{-1} U$ in \mathfrak{N} , also sind zwei Elemente in derselben linksseitigen Nebengruppe, wenn sie in derselben rechtsseitigen Nebengruppe sind und umgekehrt.

gruppe $\mathfrak{N} U$ mit allen Elementen einer anderen Nebengruppe $\mathfrak{N} V$! Es ist $N_j U N_l V = N_j U N_l U^{-1} \cdot UV = N_k UV$, da sowohl N_j als auch $U N_l U^{-1}$, also auch ihr Produkt N_k , in \mathfrak{N} enthalten ist. Wir erhalten also die Elemente einer einzigen Nebengruppe, derjenigen, die mit UV gebildet werden kann.

Faßt man die Nebengruppen eines Normalteilers als selbständige Größen auf und bezeichnet man als Produkt zweier Nebengruppen diejenige Nebengruppe, deren Elemente man durch Multiplizieren der Elemente der einen Nebengruppe mit den Elementen der anderen Nebengruppe erhält, so bilden die Nebengruppen mit Hilfe dieses Multiplikationsgesetzes selber eine Gruppe, die Faktorgruppe des Normalteilers. Die Einheit in dieser Gruppe ist der Normalteiler selber: jedes Element $N_j U$ einer Nebengruppe $\mathfrak{N} U$ mit einem Element N_l von \mathfrak{N} von vorne oder auch von hinten multipliziert, ergibt ein Element der Nebengruppe $\mathfrak{N} U$. In der Tat ist $N_l \cdot N_j U = N_l N_j \cdot U$ und $N_j U \cdot N_l = N_j \cdot U N_l U^{-1} \cdot U = N_k U$. Außerdem existiert eine Reziproke zu jeder Nebengruppe $\mathfrak{N} U$, das ist die Nebengruppe $\mathfrak{N} U^{-1}$. Es ist nämlich

$$N_j U \cdot N_l U^{-1} = N_j \cdot U N_l U^{-1} = N_k,$$

also im Normalteiler selber enthalten. Das Produkt von $\mathfrak{N} U$ und $\mathfrak{N} U^{-1}$ ergibt also \mathfrak{N} , die Einheit.

Die Ordnung der Faktorgruppe von \mathfrak{N} ist gleich der Anzahl der Nebengruppen von \mathfrak{N} , also gleich seinem Index. Man hüte sich, die Faktorgruppe mit einer Untergruppe zu verwechseln, die Elemente der Untergruppe sind die Elemente der Gruppe, die Elemente der Faktorgruppe sind selber Nebengruppen.

Die Gesetze, die wir bisher abgeleitet haben, kann man formal etwas einfacher mit Hilfe einer symbolischen Methode gewinnen, indem man eine Gesamtheit von Elementen, einen Komplex, mit einem einzigen Buchstaben, etwa \mathfrak{C} bezeichnet. Das Produkt eines Komplexes \mathfrak{C} mit einem Element A ist wieder ein Komplex $\mathfrak{C} A$, dessen Elemente man erhält, wenn man alle Elemente von \mathfrak{C} rechts bzw. für $A \mathfrak{C}$ links mit A multipliziert. Das Produkt zweier Komplexe \mathfrak{C} und \mathfrak{D} ist ein Komplex $\mathfrak{C} \mathfrak{D}$, dessen Elemente man erhält, wenn man alle Elemente von \mathfrak{C} rechts mit allen Elementen von \mathfrak{D} multipliziert. Man überzeugt sich leicht, daß für diese Art Multiplikation das Assoziativgesetz ebenfalls gilt. Enthalten \mathfrak{C} und \mathfrak{D}

genau n bzw. n' Elemente, so enthält $\mathfrak{C} \mathfrak{D}$ höchstens $n n'$ Elemente. Es enthält aber im allgemeinen weniger Elemente, weil unter den $n n'$ Produkten der Elemente von \mathfrak{C} und \mathfrak{D} gleiche vorkommen können.

Die Bedingung, daß \mathfrak{C} eine Untergruppe sei, ist $\mathfrak{C} \cdot \mathfrak{C} = \mathfrak{C}^2 = \mathfrak{C}$. Diese Untergruppe ist ein Normalteiler, wenn für jedes Element U die Gleichung $U^{-1} \mathfrak{C} U = \mathfrak{C}$ gilt. Die rechtsseitigen Nebengruppen von \mathfrak{C} sind alle voneinander verschiedenen Komplexe $\mathfrak{C} U$. Ist \mathfrak{C} ein Normalteiler, so ist $U^{-1} \mathfrak{C} U = \mathfrak{C}$, also $\mathfrak{C} U = U \mathfrak{C}$, die rechtsseitigen Nebengruppen sind auch linksseitige. Die Elemente der Faktorgruppe sind die voneinander verschiedenen Komplexe $\mathfrak{C} U$. Das Produkt zweier Komplexe $\mathfrak{C} U$ und $\mathfrak{C} V$ im Sinne der Gruppenmultiplikation der Faktorgruppe ist gleich dem Produkte im Sinne der Multiplikation der Komplexe:

$$\mathfrak{C} U \cdot \mathfrak{C} V = \mathfrak{C} \cdot U \mathfrak{C} \cdot V = \mathfrak{C} \cdot \mathfrak{C} U \cdot V = \mathfrak{C}^2 U V = \mathfrak{C} \cdot U V.$$

Holomorphie und Isomorphie

Wir haben im vorangehenden Kapitel den Begriff der Holomorphie von zwei Gruppen kennengelernt. Zwei Gruppen sind holomorph, wenn eine ein-eindeutige Zuordnung zwischen ihren Elementen existiert, so daß Produkt in Produkt übergeht: entspricht dem A bzw. B der einen Gruppe das \bar{A} bzw. \bar{B} der anderen, holomorphen Gruppe, so entspricht dem Produkt $A B$ auch das Produkt $\bar{A} \bar{B}$ der holomorphen Gruppe. Holomorphe Gruppen haben natürlich gleiche Ordnung.

Eine weniger scharfe Beziehung zwischen zwei Gruppen ist die einfache Isomorphie. Eine Gruppe \mathfrak{G} ist zu einer anderen \mathfrak{H} isomorph, wenn zu jedem Element A bzw. B von \mathfrak{G} ein Element \bar{A} bzw. \bar{B} von \mathfrak{H} zugeordnet werden kann, so daß dem Produkt $A B$ das Produkt $\bar{A} \bar{B}$ zugeordnet ist und \mathfrak{G} zu jedem Element \bar{C} von \mathfrak{H} wenigstens ein Element C hat, so daß \bar{C} zu C zugeordnet ist. Die Zuordnung ist — im Gegensatz zu der Zuordnung der Holomorphie — im allgemeinen nicht eindeutig umkehrbar. Es ist vielmehr möglich, daß zu mehreren verschiedenen Elementen etwa A und A' von \mathfrak{G} dasselbe Element \bar{A} von \mathfrak{H} zugeordnet ist. Demgemäß ist auch die Isomorphie kein umkehrbarer Begriff; ist \mathfrak{G} isomorph zu \mathfrak{H} , so ist \mathfrak{H} im allgemeinen keineswegs isomorph zu \mathfrak{G} . Die Anzahl

der Elemente von \mathfrak{G} ist gleich oder größer als die Anzahl der Elemente von \mathfrak{H} , im ersteren Falle artet die Isomorphie in eine Holomorphie aus, die dann auch umkehrbar ist.

Der Einheit E von \mathfrak{G} entspricht die Einheit \bar{E} von \mathfrak{H} , da aus $E \cdot E = E$ auch $\bar{E} \cdot \bar{E} = \bar{E}$ folgt, was nur für die Einheit der Gruppe gilt. Reziproken Elementen von \mathfrak{G} entsprechen reziproke Elemente von \mathfrak{H} .

Betrachten wir alle diejenigen Elemente E, E_2, E_3, \dots, E_n von \mathfrak{G} , zu denen die Einheit \bar{E} von \mathfrak{H} zugeordnet ist. Bezeichnen wir diesen Komplex mit \mathfrak{E} . Da zum Produkt $E_k E_l$ das Produkt $\bar{E} \cdot \bar{E} = \bar{E}$ zugeordnet ist, ist auch $E_k E_l$ in \mathfrak{E} enthalten, \mathfrak{E} bildet eine Gruppe. Einem zu E_k konjugierten Element $U^{-1} E_k U$ entspricht auch $\bar{U}^{-1} \bar{E} \bar{U} = \bar{U}^{-1} \bar{E} \bar{U} = \bar{E}$: die konjugierten Elemente der E, E_2, E_3, \dots, E_n sind ebenfalls in \mathfrak{E} enthalten, \mathfrak{E} bildet einen Normalteiler von \mathfrak{G} . Die Elemente des Komplexes \mathfrak{A} , zu denen ein und dasselbe Element \bar{A} von \mathfrak{H} zugeordnet ist, bilden dagegen eine Nebengruppe von \mathfrak{E} . Sind nämlich A_j und A_l zwei Elemente von \mathfrak{A} , also $\bar{A}_j = \bar{A}_l = \bar{A}$, so entspricht dem Element $A_j A_l^{-1}$ das Element $\bar{A}_j \bar{A}_l^{-1} \bar{A}_j \bar{A}_l^{-1} = \bar{E}$: $A_j A_l^{-1}$ ist in \mathfrak{E} enthalten, was die Bedingung dafür ist, daß A_j und A_l in ein und derselben Nebengruppe von \mathfrak{E} seien. Den Nebengruppen von \mathfrak{E} sind in ein-eindeutiger Weise die Elemente von \mathfrak{H} zugeordnet, wobei dem Produkt zweier Nebengruppen $\mathfrak{E} U$ und $\mathfrak{E} V$ auch das Produkt $\bar{U} \bar{V}$ der beiden Elemente \bar{U} und \bar{V} entspricht. Die Nebengruppen bilden aber die Elemente der Faktorgruppe von \mathfrak{E} , diese ist also holomorph zu \mathfrak{H} .

Ist \mathfrak{G} isomorph zu \mathfrak{H} , so ist eine Faktorgruppe von \mathfrak{G} holomorph zu \mathfrak{H} . Die Ordnung von \mathfrak{G} ist ein ganzzahliges Vielfaches der Ordnung von \mathfrak{H} .

Ist die Isomorphie von vornherein eine Holomorphie, so artet der Normalteiler \mathfrak{E} von \mathfrak{G} in das einzige Element E aus.

Hieraus ergibt sich nach entsprechender Umnumerierung der Elemente folgendes Bild für die Zuordnung der Elemente von \mathfrak{G} und \mathfrak{H} zueinander:

$$\underbrace{E, G_2, G_3, \dots, G_n}_{\bar{E}}, \quad \underbrace{G_{n+1}, G_{n+2}, \dots, G_{2n}}_{H_2}, \quad \dots, \quad \underbrace{G_{(h-1)n}, \dots, G_{hn}}_{H_h}.$$

Wir sehen, daß zu jedem Element von \mathfrak{G} eindeutig ein Element von \mathfrak{H} zugeordnet ist. Auch umgekehrt sind jedem Element von \mathfrak{H} Elemente von \mathfrak{G} zugeordnet. Diese Zuordnung ist aber nicht eindeutig, sondern n -deutig, indem jedem Element von \mathfrak{H} genau n Elemente von \mathfrak{G} zugeordnet sind. Diese Zahl ist für alle Elemente von \mathfrak{H} dieselbe. Die Elemente E, G_1, G_2, \dots, G_n bilden einen Normalteiler (wir bezeichneten sie vorher mit E, E_1, E_2, \dots, E_n), die anderen durch Klammern zusammengefaßten, ein und demselben Element von \mathfrak{H} zugeordneten Elemente (z. B. $G_{n+1}, G_{n+2}, \dots, G_{2n}$) seine Nebengruppen.

Multipliziert man ein zu H_i zugeordnetes Element von \mathfrak{G} mit einem zu H_j zugeordneten, so erhält man ein zu $H_i H_j$ zugeordnetes. Multipliziert man alle n zu H_i zugeordneten Elemente von \mathfrak{G} mit allen n zu H_j zugeordneten, so erhält man die n Elemente, die zu $H_i H_j$ zugeordnet sind, und zwar alle n -mal.

Die Gruppe \mathfrak{H} ist eigentlich nichts anderes, als die zum Normalteiler E, G_1, G_2, \dots, G_n gehörige Faktorgruppe. Sie ist mit dieser holomorph.

Offenbar ist jede Gruppe mit sich selber holomorph. Weiter ist jede Gruppe mit der Gruppe isomorph, die nur aus dem Einheitselement \bar{E} besteht. Zu A ist eben \bar{E} , zu B ebenfalls \bar{E} und zu $A B$ auch $\bar{E} \cdot \bar{E} = \bar{E}$ zugeordnet. In diesem Falle artet der Normalteiler in die ganze Gruppe aus.

Zu jeder Substitutionsgruppe ist eine Abelsche Gruppe isomorph. Man stellt die Isomorphie her, indem man jeder Substitution den Wert ihrer Determinante zuordnet. Was ergibt dies im Falle des Beispiels (+) des vorangehenden Kapitels?

Hiermit schließen wir die abstrakte Theorie der endlichen Gruppen und gehen auf ihre Darstellungstheorie über. Mit den kontinuierlichen Gruppen werden wir uns noch einmal im X. Kapitel beschäftigen. Es sei bemerkt, daß hier nur die Anfänge dieser in der Einfachheit der Gedankenführung so sehr eindrucksvollen Theorie besprochen wurden. Eine ausführliche Darstellung findet man bei A. Speiser, Theorie der Gruppen von endlicher Ordnung, sowie in Webers bekannter Algebra. Wir wollen aber hier nicht länger bei diesem Thema verweilen und wollten außer dem, was für das Folgende notwendig ist, nur das besprechen, was zur Erlangung einer gewissen Sicherheit bei dem Arbeiten mit Gruppen notwendig erscheint.

IX. Allgemeine Darstellungstheorie

Die Darstellung¹⁾ einer Gruppe ist eine zu ihr isomorphe Substitutionsgruppe [Matrizengruppe mit quadratischen²⁾ Matrizen], also eigentlich eine Zuordnung je einer Matrix $D(A)$ (oder A) zu jedem Gruppenelement A , und zwar so, daß

$$D(A) D(B) = D(AB) \quad (1)$$

sei. Wenn alle Matrizen, die zu verschiedenen Gruppenelementen zugeordnet sind, verschieden sind, so ist die Matrizengruppe zur dargestellten Gruppe holomorph, die Darstellung treu. Sonst müssen — wie im vorangehenden Kapitel auseinandergesetzt — die Gruppenelemente, denen dieselbe Matrix entspricht, die auch der Gruppeneinheit zugeordnet ist, einen Normalteiler bilden und die Darstellung ist eigentlich die (treue) Darstellung der Faktorgruppe dieses Normalteilers.

Umgekehrt kann man aus jeder Darstellung einer Faktorgruppe eine derartige Darstellung der ganzen Gruppe bilden. Die Elemente der Faktorgruppe sind die Nebengruppen eines Normalteilers, und man kann allen Elementen der Gruppe, die diese Nebengruppe bilden, dieselbe Matrix, und zwar diejenige zuordnen, die der Nebengruppe als einem Element der Faktorgruppe in der Darstellung dieser zugeordnet war.

Natürlich ist jede Matrizengruppe ihre eigene (treue) Darstellung. Weiter kann man offenbar jedem Gruppenelement die Matrix (1) zuordnen, und wir haben es mit der trivialen Isomorphie einer jeden Gruppe mit der Gruppe zu tun, die lediglich die Einheit enthält. Im Beispiel (†), Kap. VII, haben wir eine (treue) Darstellung der symmetrischen Gruppe von drei Dingen, eine weitere (nicht treue) Darstellung erhält man z. B., indem man jedem Gruppenelement die darunterstehende Matrix zuordnet.

E	A	B	C	D	F	(††)
(1)	(-1)	(-1)	(-1)	(1)	(1)	

¹⁾ Genauer sagt man: „Darstellung durch lineare Substitutionen“.

²⁾ Die Zeilen und Spalten der quadratischen Matrizen, die die Darstellung bilden, sollen immer in der gleichen Weise numeriert sein. Außerdem gelte dieselbe Numerierung für alle Matrizen derselben Darstellung, damit die Matrizenmultiplikationen immer eindeutig definiert sein sollen. Wir wollen hieran bei allen Darstellungsmatrizen festhalten.

Sie ist eigentlich eine (treue) Darstellung der Faktorgruppe, die zum Normalteiler E, D, F gehört. Die Faktorgruppe hat zwei Elemente, den Normalteiler E, D, F und ihre Nebengruppe A, B, C . Dem ersten Element der Faktorgruppe entspricht die Matrix (1), dem zweiten die Matrix (-1) .

Die Anzahl der Zeilen und Spalten in den Darstellungsmatrizen nennt man die Dimension der Darstellung. Man kann aus einer Darstellung einer Gruppe dadurch eine neue Darstellung machen, daß man jede Matrix der Darstellung ein und derselben Ähnlichkeitstransformation unterwirft. Das ändert an den Multiplikationseigenschaften der Matrizen gar nichts und der ganze Charakter der Darstellung bleibt dabei erhalten. Zwei Darstellungen, die so auseinander hervorgehen, oder, was damit gleichbedeutend ist, die man ineinander transformieren kann, nennt man äquivalent und sieht sie als nicht wesentlich voneinander verschieden an.

Aus zwei Darstellungen kann man auf mehrere Art und Weise eine neue Darstellung bilden. Die einfachste ist wohl die, daß man die beiden Darstellungen einfach aneinanderreihet. Besteht die erste Darstellung aus den Matrizen $D(A_1), D(A_2), \dots, D(A_h)$, die zweite aus den Matrizen $D'(A_1), D'(A_2), \dots, D'(A_h)$, so besteht die neue aus den Übermatrizen

$$\begin{pmatrix} D(A_1) & 0 \\ 0 & D'(A_1) \end{pmatrix}, \begin{pmatrix} D(A_2) & 0 \\ 0 & D'(A_2) \end{pmatrix}, \dots, \begin{pmatrix} D(A_h) & 0 \\ 0 & D'(A_h) \end{pmatrix}. (*)$$

Unterwirft man diese neue Darstellung noch einer Ähnlichkeitstransformation, so sieht man ihr die Entstehung aus zwei Darstellungen nicht mehr ohne weiteres an. Darstellungen, die auf diese Weise aus anderen Darstellungen entstehen, nennt man reduzibel. Reduzible Darstellungen lassen sich durch eine Ähnlichkeitstransformation immer auf die Gestalt (*) bringen, sie sind mit einer Darstellung der Form (*) äquivalent. Darstellungen, für die dies nicht möglich ist, heißen irreduzibel. Natürlich ist jede Darstellung reduzibel, die durch bloße gleichzeitige, für alle Matrizen gemeinsame Umnumerierung der Zeilen und Spalten in die Form (*) gebracht werden kann. Eine solche Umnummerierung läßt sich in der Tat durch eine Ähnlichkeitstransformation erreichen. Will man die \bar{j} -Zeile und -Spalte zur j -Zeile und -Spalte machen, so nehme man zu S die Matrix $S_{ki} = \delta_{k\bar{i}}$, dann ist $(S^{-1})_{jm} = \delta_{j\bar{m}}$, da

$$\sum_i S_{ki} (S^{-1})_{ij} = \sum_i \delta_{k\bar{i}} \delta_{\bar{i}j} = \delta_{kj}$$

ist. Weiter wird

$$\bar{\mathbf{A}} = \mathbf{S}^{-1} \mathbf{A} \mathbf{S}; \quad \bar{\mathbf{A}}_{ji} = \sum_{mk} \delta_{jm} \mathbf{A}_{mk} \delta_{ki} = \mathbf{A}_{j\bar{i}}.$$

Jedes Matrzensystem ist also reduzibel bzw. schon ausreduziert, bei der Zeilen und Spalten so in „ungestrichene“ und „gestrichene“ eingeteilt werden können, daß an den Kreuzungspunkten von „ungestrichenen“ Zeilen und „gestrichenen“ Spalten oder umgekehrt lauter Nullen stehen. Man kann dann die „ungestrichenen“ Zeilen und Spalten zu den ersten machen und erhält die Form (*) für die Darstellung.

Im folgenden wollen wir uns auf Darstellungen mit Matrizen mit nichtverschwindender Determinante beschränken. Dann haben alle Matrizen $\mathbf{D}(A)$ eine Reziproke. Jedes Element A der Gruppe mit der Einheit E der Gruppe multipliziert, ergibt das Element A wieder: jede Matrix $\mathbf{D}(A)$ der Darstellung mit der zur Einheit zugeordneten Matrix $\mathbf{D}(E)$ multipliziert, ergibt $\mathbf{D}(A)$ und es folgt

$$\mathbf{D}(A) \cdot \mathbf{D}(E) = \mathbf{D}(A); \quad \mathbf{D}(E) = \mathbf{I}. \quad (2)$$

Der Einheit der Gruppe ist die Einheitsmatrix zugeordnet. Das Produkt reziproken Elementen zugeordneter Matrizen $\mathbf{D}(A)$ und $\mathbf{D}(A^{-1})$ ist $\mathbf{D}(E) = \mathbf{I}$. Daher ist

$$\mathbf{D}(A) \cdot \mathbf{D}(A^{-1}) = \mathbf{I}; \quad \mathbf{D}(A^{-1}) = [\mathbf{D}(A)]^{-1} \quad (3)$$

und es ist

$$\mathbf{D}(A^{-1}) = \mathbf{D}(A)^\dagger, \quad (3a)$$

wenn die Matrizen der Darstellung unitär sind.

Satz 1. Jede Darstellung mit Matrizen mit nichtverschwindender Determinante läßt sich durch eine Ähnlichkeitstransformation in eine Darstellung transformieren, deren Matrizen alle unitär sind.

Die Matrizen der Darstellung der Gruppe der Ordnung h seien $\mathbf{A}_1, \mathbf{A}_2, \mathbf{A}_3, \dots, \mathbf{A}_h$. (Unter den $\mathbf{A}_1, \dots, \mathbf{A}_h$ können auch gleiche vorkommen, wenn die Darstellung nicht treu ist.) Dann bilde man die hermitische Matrix \mathbf{H} durch Summation über alle Gruppenelemente

$$\sum_{s=1}^h \mathbf{A}_s \mathbf{A}_s^\dagger = \mathbf{H}. \quad (4)$$

Die hermitische Matrix \mathbf{H} lasse sich durch die unitäre Matrix \mathbf{U} auf die Diagonalform \mathbf{d} bringen

$$\left. \begin{aligned} \mathbf{d} &= \mathbf{U}^{-1} \mathbf{H} \mathbf{U} = \sum_{\mathbf{x}} \mathbf{U}^{-1} \mathbf{A}_{\mathbf{x}} \mathbf{A}_{\mathbf{x}}^{\dagger} \mathbf{U} \\ &= \sum_{\mathbf{x}} \mathbf{U}^{-1} \mathbf{A}_{\mathbf{x}} \mathbf{U} (\mathbf{U}^{-1} \mathbf{A}_{\mathbf{x}} \mathbf{U})^{\dagger} = \sum_{\mathbf{x}} \overline{\mathbf{A}}_{\mathbf{x}} \overline{\mathbf{A}}_{\mathbf{x}}^{\dagger}. \end{aligned} \right\} \quad (5)$$

Die Diagonalmatrix \mathbf{d} hat lauter reelle positive Diagonalelemente, z. B. ist

$$d_{kk} = \sum_{\mathbf{x}} \sum_j (\overline{\mathbf{A}}_{\mathbf{x}})_{kj} (\overline{\mathbf{A}}_{\mathbf{x}})^*_{kj} = \sum_{\mathbf{x}} \sum_j |(\overline{\mathbf{A}}_{\mathbf{x}})_{kj}|^2,$$

das nur Null sein könnte, wenn bei diesem k für alle j (und \mathbf{x}) $(\overline{\mathbf{A}}_{\mathbf{x}})_{kj} = 0$ wäre. Dann wäre aber eine ganze Zeile von $\mathbf{A}_{\mathbf{x}}$ Null, also auch seine Determinante, also auch die von $\mathbf{A}_{\mathbf{x}}$, entgegen der Voraussetzung. Man kann also aus \mathbf{d} in eindeutiger Weise $\mathbf{d}^{1/2}$ und $\mathbf{d}^{-1/2}$ bilden, indem man aus den Diagonalgliedern die Quadratwurzel zieht bzw. die $-1/2$ -Potenz bildet und den positiven Wert davon nimmt. Dann sind $\mathbf{d}^{1/2}$ und $\mathbf{d}^{-1/2}$ reelle Diagonalmatrizen: $\mathbf{d}^{1/2\dagger} = \mathbf{d}^{1/2}$; $\mathbf{d}^{-1/2\dagger} = \mathbf{d}^{-1/2}$. Aus (5) erhält man

$$\mathbf{I} = \mathbf{d}^{-1/2} \sum_{\mathbf{x}} \overline{\mathbf{A}}_{\mathbf{x}} \overline{\mathbf{A}}_{\mathbf{x}}^{\dagger} \mathbf{d}^{-1/2}$$

und für jedes $\overline{\mathbf{A}}_{\lambda}$ mit $\overline{\mathbf{A}}_{\lambda} = \mathbf{d}^{-1/2} \overline{\mathbf{A}}_{\lambda} \mathbf{d}^{1/2}$

$$\left. \begin{aligned} \overline{\mathbf{A}}_{\lambda} \overline{\mathbf{A}}_{\lambda}^{\dagger} &= \mathbf{d}^{-1/2} \overline{\mathbf{A}}_{\lambda} \mathbf{d}^{1/2} \cdot (\mathbf{d}^{-1/2} \sum_{\mathbf{x}} \overline{\mathbf{A}}_{\mathbf{x}} \overline{\mathbf{A}}_{\mathbf{x}}^{\dagger} \mathbf{d}^{-1/2}) \mathbf{d}^{1/2} \overline{\mathbf{A}}_{\lambda}^{\dagger} \mathbf{d}^{-1/2} \\ &= \mathbf{d}^{-1/2} \sum_{\mathbf{x}} \overline{\mathbf{A}}_{\lambda} \overline{\mathbf{A}}_{\mathbf{x}} \overline{\mathbf{A}}_{\mathbf{x}}^{\dagger} \overline{\mathbf{A}}_{\lambda}^{\dagger} \mathbf{d}^{-1/2}. \end{aligned} \right\} \quad (6)$$

Nun ist wegen der Gruppeneigenschaft der $\overline{\mathbf{A}}_{\mathbf{x}}$, da $\overline{\mathbf{A}}_{\lambda} \overline{\mathbf{A}}_{\mathbf{x}}$ für $\mathbf{x} = 1, 2, \dots, h$ zwar in anderer Reihenfolge, aber ebenfalls alle Matrizen $\overline{\mathbf{A}}_1, \overline{\mathbf{A}}_2, \dots, \overline{\mathbf{A}}_h$ durchläuft (S. 66),

$$\sum_{\mathbf{x}} \overline{\mathbf{A}}_{\lambda} \overline{\mathbf{A}}_{\mathbf{x}} (\overline{\mathbf{A}}_{\lambda} \overline{\mathbf{A}}_{\mathbf{x}})^{\dagger} = \sum_{\mathbf{x}} \overline{\mathbf{A}}_{\mathbf{x}} \overline{\mathbf{A}}_{\mathbf{x}}^{\dagger}$$

und demgemäß

$$\overline{\mathbf{A}}_{\lambda} \overline{\mathbf{A}}_{\lambda}^{\dagger} = \mathbf{d}^{-1/2} \sum_{\mathbf{x}} \overline{\mathbf{A}}_{\mathbf{x}} \overline{\mathbf{A}}_{\mathbf{x}}^{\dagger} \mathbf{d}^{-1/2} = \mathbf{I}, \quad (7)$$

also ist

$$\overline{\mathbf{A}}_{\lambda} = \mathbf{d}^{-1/2} \mathbf{U}^{-1} \mathbf{A}_{\lambda} \mathbf{U} \mathbf{d}^{1/2}$$

unitär.

Satz 2. Eine mit allen Matrizen einer irreduziblen Darstellung vertauschbare Matrix ist eine konstante Matrix, d. h. ein Vielfaches der Einheitsmatrix.

Wieder können wir annehmen, daß die Darstellung in unitärer Form vorliegt, da ja eine Ähnlichkeitstransformation die Einheitsmatrix und ihre Vielfache unverändert läßt. Dann sei \mathbf{M} vertauschbar mit allen $\mathbf{A}_1, \mathbf{A}_2, \dots, \mathbf{A}_h$ also

$$\mathbf{A}_x \mathbf{M} = \mathbf{M} \mathbf{A}_x \quad (x = 1, 2, \dots, h). \quad (8)$$

Gehen wir zur adjungierten Gleichung über, so erhalten wir

$$\mathbf{M}^\dagger \mathbf{A}_x^\dagger = \mathbf{A}_x^\dagger \mathbf{M}^\dagger$$

und vorne sowie hinten mit \mathbf{A}_x multipliziert, wegen $\mathbf{A}_x \mathbf{A}_x^\dagger = \mathbf{A}_x^\dagger \mathbf{A}_x = \mathbf{I}$

$$\mathbf{A}_x \mathbf{M}^\dagger = \mathbf{M}^\dagger \mathbf{A}_x \quad (x = 1, 2, \dots, h), \quad (9)$$

es ist also nicht nur \mathbf{M} , sondern auch \mathbf{M}^\dagger ist mit allen \mathbf{A} vertauschbar, also auch das hermiteische $\mathbf{H}_1 = \mathbf{M} + \mathbf{M}^\dagger$ und $\mathbf{H}_2 = i(\mathbf{M} - \mathbf{M}^\dagger)$. Es genügt nachzuweisen, daß jede mit allen \mathbf{A}_x vertauschbare hermiteische Matrix eine konstante Matrix ist, dann muß nämlich \mathbf{H}_1 und \mathbf{H}_2 ein Vielfaches der Einheit sein, also auch $2\mathbf{M} = \mathbf{H}_1 - i\mathbf{H}_2$.

Ist aber die Matrix \mathbf{M} in (8) hermiteisch, so kann man sie mit einer unitären Matrix \mathbf{V} auf Diagonalf orm $\mathbf{d} = \mathbf{V}^{-1} \mathbf{M} \mathbf{V}$ bringen und erhält, wenn man noch $\bar{\mathbf{A}}_x = \mathbf{V}^{-1} \mathbf{A}_x \mathbf{V}$ setzt (die $\bar{\mathbf{A}}_x$ bleiben dabei unitär),

$$\bar{\mathbf{A}}_x \mathbf{d} = \mathbf{d} \bar{\mathbf{A}}_x \quad (x = 1, 2, \dots, h). \quad (10)$$

Würde die Diagonalmatrix \mathbf{d} nicht lauter gleiche Diagonalelemente haben, so müßten an den Kreuzungspunkten verschiedener Diagonalelemente entsprechenden Punkten in allen $\bar{\mathbf{A}}_x$ lauter Nullen stehen. Aus

$$(\bar{\mathbf{A}}_x)_{kj} \mathbf{d}_{jj} = \mathbf{d}_{kk} (\bar{\mathbf{A}}_x)_{kj}$$

folgt für $\mathbf{d}_{jj} \neq \mathbf{d}_{kk}$ auch $(\bar{\mathbf{A}}_x)_{kj} = 0$, so daß die Darstellung im Sinne der Ausführungen auf S. 81 reduzibel wäre. Da dies nicht der Fall ist, sind alle \mathbf{d}_{kk} gleich, \mathbf{d} , also auch $\mathbf{V} \mathbf{d} \mathbf{V}^{-1} = \mathbf{M}$ ist eine konstante, mit jeder Matrix vertauschbare Matrix.

Die soeben gegebene Ableitung des Satzes 2 lehrt uns nicht nur, daß die Darstellung reduzibel sein muß, wenn eine nichtkonstante

Matrix mit der ganzen Darstellung vertauschbar ist, sie zeigt uns auch, wie man die Darstellung ausreduzieren, in die Form (*) bringen kann. Dies leistet dieselbe Ähnlichkeitstransformation, die die „vertauschbare Matrix“ auf die Diagonalform bringt.

Ist umgekehrt die Darstellung reduzibel, so existieren sicher nichtkonstante Matrizen, die mit allen Matrizen der Darstellung vertauschbar sind. In diesem Fall kann man nämlich die Darstellung durch Ähnlichkeitstransformation mit einer zweckmäßig gewählten Matrix S auf die Form (*) bringen. Mit Matrizen der Form (*) sind aber alle Matrizen

$$\mathbf{M} = \begin{pmatrix} a'1 & 0 \\ 0 & a'1 \end{pmatrix}$$

bei beliebigem a, a' vertauschbar. Wenn man \mathbf{M} mit S^{-1} transformiert, wird es mit den ursprünglichen Matrizen der Darstellung, die ebenfalls durch Transformation mit S^{-1} aus den Matrizen der Form (*) hervorgehen, vertauschbar.

Existiert eine nichtkonstante, mit allen Matrizen der Darstellung vertauschbare Matrix, so ist die Darstellung reduzibel, existiert keine, so ist sie irreduzibel.

Satz 3. Sind zwei Darstellungen $\mathbf{D}^{(1)}(A_1), \mathbf{D}^{(1)}(A_2), \dots, \mathbf{D}^{(1)}(A_h)$ und $\mathbf{D}^{(2)}(A_1), \mathbf{D}^{(2)}(A_2), \dots, \mathbf{D}^{(2)}(A_h)$ derselben Gruppe der Dimension l_1 bzw. l_2 irreduzibel und existiert eine Matrix \mathbf{M} mit l_1 Spalten und l_2 Zeilen, so daß

$$\mathbf{M} \mathbf{D}^{(1)}(A_x) = \mathbf{D}^{(2)}(A_x) \mathbf{M} \quad (x = 1, 2, \dots, h), \quad (11)$$

so muß bei $l_1 \neq l_2$ die Matrix \mathbf{M} eine Nullmatrix sein. Ist $l_1 = l_2$, so ist \mathbf{M} entweder eine Nullmatrix, oder eine Matrix mit nicht verschwindender Determinante. Im letzteren Falle hat \mathbf{M} eine Reziproke und die beiden irreduziblen Darstellungen sind miteinander äquivalent.

Zunächst können wir annehmen, daß die Darstellungen bereits in unitärer Form vorliegen. Wäre das nämlich nicht der Fall, so könnten wir sie durch S bzw. R in diese Form transformieren und aus (11) würde

$$\left. \begin{aligned} \mathbf{R}^{-1} \mathbf{M} \mathbf{S} \cdot \mathbf{S}^{-1} \mathbf{D}^{(1)}(A_x) \mathbf{S} &= \mathbf{R}^{-1} \mathbf{D}^{(2)}(A_x) \mathbf{R} \cdot \mathbf{R}^{-1} \mathbf{M} \mathbf{S} \\ \mathbf{R}^{-1} \mathbf{M} \mathbf{S} \cdot \bar{\mathbf{D}}^{(1)}(A_x) &= \bar{\mathbf{D}}^{(2)}(A_x) \cdot \mathbf{R}^{-1} \mathbf{M} \mathbf{S} \end{aligned} \right\} \quad (12)$$

folgen, wo wir für $\mathbf{R}^{-1} \mathbf{M} \mathbf{S}$ wieder \mathbf{M} setzen könnten.

Nehmen wir weiter an, daß $l_1 \leqq l_2$ ist. Wäre $l_1 > l_2$, so gingen wir zur transponierten Gleichung von (11) über, am folgenden würde sich dadurch nichts ändern. Schreiben wir die Adjungierte von (11) auf, so ist wegen der Unitarität $\mathbf{D}^{(1)}(A_x)^\dagger = \mathbf{D}^{(1)}(A_x)^{-1} = \mathbf{D}^{(1)}(A_x^{-1})$ und $\mathbf{D}^{(2)}(A_x)^\dagger = \mathbf{D}^{(2)}(A_x^{-1})$

$$\mathbf{D}^{(1)}(A_x^{-1}) \mathbf{M}^\dagger = \mathbf{M}^\dagger \mathbf{D}^{(2)}(A_x^{-1}). \quad (13)$$

Da (11) für alle Gruppenelemente, also auch für A_x^{-1} gilt, folgt aus (13) links mit \mathbf{M} multipliziert

$$\mathbf{M} \mathbf{D}^{(1)}(A_x^{-1}) \mathbf{M}^\dagger = \mathbf{M} \mathbf{M}^\dagger \mathbf{D}^{(2)}(A_x^{-1}), \quad (14)$$

$$\mathbf{D}^{(2)}(A_x^{-1}) \mathbf{M} \mathbf{M}^\dagger = \mathbf{M} \mathbf{M}^\dagger \mathbf{D}^{(2)}(A_x^{-1}), \quad (15)$$

d. h. das hermitische $\mathbf{M} \mathbf{M}^\dagger$ ist mit allen Darstellungsmatrizen $\mathbf{D}^{(2)}(A_1), \mathbf{D}^{(2)}(A_2), \dots, \mathbf{D}^{(2)}(A_h)$ vertauschbar. Es muß daher ein Vielfaches der Einheitsmatrix sein:

$$\mathbf{M} \mathbf{M}^\dagger = c \mathbf{1}. \quad (16)$$

Ist jetzt die Dimension der beiden Darstellungen $\mathbf{D}^{(1)}$ und $\mathbf{D}^{(2)}$ gleich, $l_1 = l_2$, so ist entweder $c \neq 0$, dann ist die Determinante $|c \mathbf{1}| = c^{l_1}$ nicht Null und daher verschwindet auch die von \mathbf{M} nicht, \mathbf{M} hat eine Reziproke. Ist aber $c = 0$, so ist $\mathbf{M} \mathbf{M}^\dagger = \mathbf{0}$ die Nullmatrix, was in den Koeffizienten ausgeschrieben

$$\sum_k \mathbf{M}_{ik} \mathbf{M}_{jk}^* = 0 \quad (17)$$

lautet. Setzen wir $i = j$, so ergibt dies

$$\sum_k |\mathbf{M}_{ik}|^2 = 0, \quad (18)$$

woraus auch $\mathbf{M}_{ik} = 0$ folgt, da kein $|\mathbf{M}_{ik}|^2$ negativ sein kann, also auch keines positiv sein darf.

Ist dagegen die Dimension der beiden Darstellungen nicht dieselbe, so ist auch \mathbf{M} nicht quadratisch, sondern, da $l_1 < l_2$, nach unten gestreckt. Wir können sie aber mit lauter Nullen zu einer quadratischen machen:

$$\mathbf{N} = \begin{pmatrix} \mathbf{M}_{11} & \mathbf{M}_{12} & \dots & \mathbf{M}_{1a} & 0 & \dots & 0 \\ \mathbf{M}_{21} & \mathbf{M}_{22} & \dots & \mathbf{M}_{2a} & 0 & \dots & 0 \\ \vdots & \vdots & & \vdots & \vdots & & \vdots \\ \mathbf{M}_{b1} & \mathbf{M}_{b2} & \dots & \mathbf{M}_{ba} & 0 & \dots & 0 \end{pmatrix} \quad (19)$$

Dabei bleibt $\mathbf{N} \mathbf{N}^\dagger = \mathbf{M} \mathbf{M}^\dagger$. Da aber die Determinante von \mathbf{N} offenbar Null ist, muß auch die von $\mathbf{N} \mathbf{N}^\dagger = \mathbf{M} \mathbf{M}^\dagger$, also auch das c in (16) verschwinden, so daß wieder (17) und (18) gilt.

Damit haben wir Satz 3 vollkommen bewiesen.

Satz 1a. Wenn zwei beliebige Darstellungen A_1, A_2, \dots, A_h und B_1, B_2, \dots, B_h derselben Gruppe in unitärer Form vorliegen und miteinander äquivalent sind, d. h. wenn eine beliebige Matrix M existiert, so daß

$$M A_x M^{-1} = B_x \quad (x = 1, 2, \dots, h), \quad (20)$$

dann kann man die beiden Darstellungen auch unitär ineinander transformieren. Es existiert dann auch eine unitäre Matrix U , für die

$$U A_x U^{-1} = B_x \quad (x = 1, 2, \dots, h) \quad (21)$$

gilt.

Aus (20) folgt

$$M A_x = B_x M \quad (x = 1, 2, \dots, h) \quad (22)$$

und daraus, ebenso wie vorher, daß $M M^\dagger$ mit allen Matrizen der zweiten Darstellung vertauschbar ist,

$$B_x M M^\dagger = M M^\dagger B_x \quad (x = 1, 2, \dots, h). \quad (22a)$$

Eine Ähnlichkeitstransformation mit ihr läßt die zweite Darstellung überhaupt ungeändert. Dies legt den Gedanken nahe, die erste Darstellung an Stelle von M mit KM zu transformieren, wo K das $M M^\dagger$ oder etwas Ähnliches ist, was mit allen Matrizen B_x vertauschbar ist, so daß KM in (20) an Stelle von M gesetzt werden kann. Natürlich werden wir dieses K so bestimmen, daß KM unitär werde, d. h.

$$M^\dagger K^\dagger KM = 1; \quad K^\dagger K = M^{\dagger -1} M^{-1} = (MM^\dagger)^{-1} \quad (23)$$

ist. Daher muß nicht K selber, sondern — wenn es hermiteisch ist — seine -2 Potenz MM^\dagger sein. Es handelt sich also wieder darum, aus MM^\dagger gewissermaßen die $-1/2$ Potenz zu bilden. Da MM^\dagger hermiteisch ist, läßt es sich mit einer unitären Matrix V auf die Diagonalform bringen: es sei

$$V^{-1} M M^\dagger V = d; \quad MM^\dagger = VdV^{-1}, \quad (24)$$

und die Diagonalelemente von d sind, wie bei Satz 1, alle positiv, so daß man ebenso wie dort $d^{-1/2}$ bilden kann. Diese ist eine Diagonalmatrix mit lauter reellen positiven Diagonagliedern. Nunmehr ist zu erwarten, daß

$$K = Vd^{-1/2}V^{-1} \quad (25)$$

erstens mit allen Matrizen B_1, B_2, \dots, B_h vertauschbar ist, und zweitens $U = KM$ unitär ist. Dann würde aus (20) wegen $K B_x K^{-1} = B_x$ auch folgen:

$$U A_x U^{-1} = K M A_x M^{-1} K^{-1} = K B_x K^{-1} = B_x.$$

In der Tat ist nun wegen (22a) und (24)

$$B_x V d V^{-1} = V d V^{-1} B_x; \quad V^{-1} B_x V d = d V^{-1} B_x V, \quad (26)$$

so daß die Diagonalmatrix d mit allen $V^{-1} B_x V$ vertauschbar ist. Daher stehen an den Kreuzungspunkten solcher Zeilen und Spalten, deren

Diagonalelemente in \mathbf{d} verschieden sind, in allen $V^{-1} \mathbf{B}_x V$ Nullen. Dann sind diese Matrizen aber auch mit $d^{-1/2}$ vertauschbar, weil in $d^{-1/2}$ nur jene Diagonalglieder verschieden sind, die schon in \mathbf{d} verschieden waren. Daher ist

$$V^{-1} \mathbf{B}_x V d^{-1/2} = d^{-1/2} V^{-1} \mathbf{B}_x V; \quad \mathbf{B}_x \mathbf{K} = \mathbf{K} \mathbf{B}_x. \quad (27)$$

\mathbf{K} ist tatsächlich mit allen Darstellungsmatrizen $\mathbf{B}_1, \mathbf{B}_2, \dots, \mathbf{B}_h$ vertauschbar. Zweitens ist

$$\mathbf{U} \mathbf{U}^\dagger = \mathbf{K} \mathbf{M} \mathbf{M}^\dagger \mathbf{K}^\dagger = V d^{-1/2} V^{-1} \mathbf{M} \mathbf{M}^\dagger V^{-1\dagger} d^{-1/2\dagger} V^\dagger \quad (28)$$

und wegen (24) und weil V unitär, $d^{-1/2}$ aber hermiteisch ist (reelle Diagonalmatrix):

$$\mathbf{U} \mathbf{U}^\dagger = V d^{-1/2} d d^{-1/2} V^\dagger = V V^\dagger = \mathbf{1}, \quad (29)$$

also \mathbf{U} unitär.

Die Bedeutung dieses Satzes liegt in der Tatsache, daß man sich überhaupt auf unitäre Ähnlichkeitstransformationen beschränken kann, solange man nur mit unitären Darstellungen operiert.

Satz 4. Der vierte, praktisch von allen wichtigste Satz ist die Orthogonalitätsrelation der Koeffizienten irreduzibler Darstellungen.

Sind

$$\mathbf{D}^{(1)}(E), \mathbf{D}^{(1)}(A_2), \dots, \mathbf{D}^{(1)}(A_h)$$

und

$$\mathbf{D}^{(2)}(E), \mathbf{D}^{(2)}(A_2), \dots, \mathbf{D}^{(2)}(A_h)$$

zwei nicht äquivalente irreduzible Darstellungen derselben Gruppe, die beide in unitärer Form vorliegen, so gilt für alle Koeffizienten $\mu \nu$ bzw. $\alpha \beta$

$$\sum_R \mathbf{D}^{(1)}(R)_\mu^* \mathbf{D}^{(2)}(R)_{\alpha \beta} = 0, \quad (30)$$

wo die Summation, wie angedeutet, über alle h Gruppenelemente E, A_2, \dots, A_h zu erstrecken ist¹⁾.

Für die Koeffizienten einer unitären, irreduziblen Darstellung gilt

$$\sum_R \mathbf{D}^{(1)}(R)_\mu^* \mathbf{D}^{(1)}(R)_{\mu' \nu'} = \frac{h}{l_1} \delta_{\mu \mu'} \delta_{\nu \nu'}, \quad (31)$$

wo h die Ordnung der Gruppe, l_1 die Dimension der Darstellung $\mathbf{D}^{(1)}(R)$ ist.

Satz 4 entspringt der Tatsache, daß man wegen der Gruppen-eigenschaft der Darstellung scheinbar leicht viele Matrizen \mathbf{M}

¹⁾ R (und später auch S) ist im folgenden immer eines der Gruppenelemente $E = A_1, A_2, \dots, A_h$.

konstruieren kann, die die Gleichung (11) oder (8) befriedigen; unsere Gleichungen (30) und (31) drücken dann die Tatsache aus, daß eine Matrix, die (11) befriedigt, die Nullmatrix sein muß, eine, die (8) befriedigt, eine Vielfache der Einheitsmatrix ist.

Eine Matrix, die (11) befriedigt, ist bei beliebigem l_1 -spaltigem, l_2 -zeiligem \mathbf{X} die Matrix

$$\mathbf{M} = \sum_R \mathbf{D}^{(2)}(R) \mathbf{X} \mathbf{D}^{(1)}(R)^{-1}.$$

Es ist nämlich

$$\begin{aligned} \mathbf{D}^{(2)}(S) \mathbf{M} &= \sum_R \mathbf{D}^{(2)}(S) \mathbf{D}^{(2)}(R) \mathbf{X} \mathbf{D}^{(1)}(R)^{-1} \\ &= \sum_R \mathbf{D}^{(2)}(SR) \mathbf{X} \mathbf{D}^{(1)}(SR)^{-1} \mathbf{D}^{(1)}(S). \end{aligned}$$

Nun ist wegen der Gruppeneigenschaft

$$\sum_R \mathbf{D}^{(2)}(SR) \mathbf{X} \mathbf{D}^{(1)}(SR)^{-1} = \sum_R \mathbf{D}^{(2)}(R) \mathbf{X} \mathbf{D}^{(1)}(R)^{-1} = \mathbf{M}.$$

Es kommen ja links und rechts dieselben Matrizen, nur in verschiedener Reihenfolge vor. Daher ist

$$\mathbf{D}^{(2)}(S) \mathbf{M} = \mathbf{M} \mathbf{D}^{(1)}(S), \quad (11a)$$

woraus nach Satz 3 folgt, daß \mathbf{M} die Nullmatrix sein muß. Es ist also bei beliebigen $\mathbf{X}_{\alpha\lambda}$

$$\mathbf{M}_{\alpha\mu} = \sum_{\alpha\lambda} \sum_R \mathbf{D}^{(2)}(R)_{\alpha\lambda} \mathbf{X}_{\alpha\lambda} \mathbf{D}^{(1)}(R^{-1})_{\lambda\mu} = 0.$$

Setzt man hierin alle $\mathbf{X}_{\alpha\lambda} = 0$, nur ein hervorgehobenes $\mathbf{X}_{\beta\gamma} = 1$, so erhält man die etwas verallgemeinerte Gleichung (30)

$$\sum_R \mathbf{D}^{(2)}(R)_{\alpha\beta} \mathbf{D}^{(1)}(R^{-1})_{\gamma\mu} = 0, \quad (30a)$$

wo $\mathbf{D}^{(2)}(R)$ und $\mathbf{D}^{(1)}(R)$ zwar irreduzibel, aber nicht in unitärer Form sein müssen. Sind die Matrizen $\mathbf{D}^{(1)}(R)$ unitär, so ist

$$\mathbf{D}^{(1)}(R^{-1}) = \mathbf{D}^{(1)}(R)^{-1} = \mathbf{D}^{(1)}(R)^\dagger$$

und (30a) geht in (30) über.

Um noch (31) zu beweisen, nehmen wir bei beliebigem \mathbf{X}

$$\mathbf{M} = \sum_R \mathbf{D}^{(1)}(R) \mathbf{X} \mathbf{D}^{(1)}(R)^{-1}$$

an. Dieses \mathbf{M} ist mit allen $\mathbf{D}^{(1)}(S)$ vertauschbar.

$$\begin{aligned} \mathbf{D}^{(1)}(S) \mathbf{M} &= \sum_R \mathbf{D}^{(1)}(S) \mathbf{D}^{(1)}(R) \mathbf{X} \mathbf{D}^{(1)}(R)^{-1} \\ &= \sum_R \mathbf{D}^{(1)}(SR) \mathbf{X} \mathbf{D}^{(1)}((SR)^{-1}) \mathbf{D}^{(1)}(S) = \mathbf{M} \mathbf{D}^{(1)}(S), \end{aligned}$$

so daß \mathbf{M} ein Vielfaches der Einheitsmatrix sein muß. Es ist

$$\sum_{\lambda} \sum_R \mathbf{D}^{(1)}(R)_{\mu\lambda} \mathbf{X}_{\lambda} \mathbf{D}^{(1)}(R^{-1})_{\lambda\mu'} = c \delta_{\mu\mu'},$$

wo c von μ und μ' unabhängig ist, von den \mathbf{X}_{λ} aber noch abhängen kann. Wählen wir wieder ein $\mathbf{X}_{\nu\nu'} = 1$, lassen aber alle anderen verschwinden, so erhalten wir, wenn wir das diesem System von \mathbf{X}_{λ} entsprechende c mit $c_{\nu\nu'}$ bezeichnen:

$$\sum_R \mathbf{D}^{(1)}(R)_{\mu\nu} \mathbf{D}^{(1)}(R^{-1})_{\nu'\mu'} = c_{\nu\nu'} \delta_{\mu\mu'}.$$

Zur Bestimmung der $c_{\nu\nu'}$ setzen wir $\mu = \mu'$ und summieren über μ von 1 bis l_1 . Wir erhalten wegen der Multiplikations-eigenschaft der Darstellungen, sowie wegen $\mathbf{D}(E)_{\nu\nu} = \delta_{\nu\nu}$,

$$\begin{aligned} \sum_u \sum_R \mathbf{D}^{(1)}(R^{-1})_{\nu'\mu} \mathbf{D}^{(1)}(R)_{\mu\nu} &= \sum_R \mathbf{D}^{(1)}(E)_{\nu\nu} = h \delta_{\nu\nu}, \\ &= \sum_{\mu} c_{\nu\nu'} \delta_{\mu\mu} = c_{\nu\nu'} l_1, \end{aligned}$$

so daß $c_{\nu\nu'} = \delta_{\nu\nu'} h/l_1$ ist. Wir haben wieder das etwas verallgemeinerte (31)

$$\sum_R \mathbf{D}^{(1)}(R)_{\mu\nu} \mathbf{D}^{(1)}(R^{-1})_{\nu'\mu'} = \frac{h}{l_1} \delta_{\mu\mu'} \delta_{\nu\nu'}, \quad (31a)$$

die sich für eine unitäre Darstellung in der Form (31) schreiben läßt.

Die Zahlen

$$\mathbf{D}^{(1)}(A_1)_{\mu\nu} = v_{A_1}^{(\mu\nu)}; \mathbf{D}^{(1)}(A_2)_{\mu\nu} = v_{A_2}^{(\mu\nu)}; \dots; \mathbf{D}^{(1)}(A_h)_{\mu\nu} = v_{A_h}^{(\mu\nu)}$$

kann man als die Komponenten eines h dimensionalen Vektors v auffassen, dessen Komponenten nach den Gruppenelementen benannt sind. Dann bedeutet (31), daß die hermitische Länge dieses Vektors $\sqrt{h/l_1}$ ist, und daß je zwei verschiedene dieser l_1^2 -Vektoren aufeinander orthogonal sind. Ebenfalls orthogonal sind nach (30) diese Vektoren auf alle Vektoren w [mit $w_{A_1}^{(\alpha\beta)} = \mathbf{D}^{(2)}(A_1)_{\alpha\beta}$, $w_{A_2}^{(\alpha\beta)} = \mathbf{D}^{(2)}(A_2)_{\alpha\beta}$, ..., $w_{A_h}^{(\alpha\beta)} = \mathbf{D}^{(2)}(A_h)_{\alpha\beta}$], die aus einer anderen irreduziblen Darstellung auf ähnliche Weise hervorgehen.

Die schon öfters betrachtete Darstellung der symmetrischen Gruppe von drei Dingen:

$$\left. \begin{aligned} \mathbf{D}(E) &= \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}; & \mathbf{D}(A) &= \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}; & \mathbf{D}(B) &= \begin{pmatrix} -\frac{1}{2} & \frac{1}{2}\sqrt{3} \\ \frac{1}{2}\sqrt{3} & \frac{1}{2} \end{pmatrix}, \\ \mathbf{D}(C) &= \begin{pmatrix} -\frac{1}{2} & -\frac{1}{2}\sqrt{3} \\ -\frac{1}{2}\sqrt{3} & \frac{1}{2} \end{pmatrix}; & \mathbf{D}(D) &= \begin{pmatrix} -\frac{1}{2} & \frac{1}{2}\sqrt{3} \\ -\frac{1}{2}\sqrt{3} & -\frac{1}{2} \end{pmatrix}, \\ \mathbf{D}(F) &= \begin{pmatrix} -\frac{1}{2} & -\frac{1}{2}\sqrt{3} \\ \frac{1}{2}\sqrt{3} & -\frac{1}{2} \end{pmatrix} \end{aligned} \right\} +$$

ist irreduzibel. Wäre sie nämlich reduzibel, d. h. könnte man alle Matrizen mit derselben Ähnlichkeitstransformation auf die Diagonalf orm bringen, so müßten alle Matrizen vertauschbar sein, da sie es in der Diagonalf orm wären. Dies ist jedoch nicht der Fall, wie man z. B. aus

$$\mathbf{D}(A)\mathbf{D}(B) = \mathbf{D}(D), \quad \mathbf{D}(B)\mathbf{D}(A) = \mathbf{D}(F) \neq \mathbf{D}(D)$$

ersieht. Nach Satz 2 darf nur ein Vielfaches der Einheit mit allen Matrizen (+) vertauschbar sein. Mit $\mathbf{D}(A)$ ist nur eine Diagonalmatrix vertauschbar, mit $\mathbf{D}(B)$ ist aber diese Diagonalmatrix nur vertauschbar, wenn die beiden Diagonalelemente gleich sind. Also sind schon mit $\mathbf{D}(A)$ und $\mathbf{D}(B)$ nur Vielfache der Einheitsmatrix vertauschbar. Nach (31) müssen die vier Vektoren $v_E^{(11)}$, $v_A^{(12)}$, $v_B^{(21)}$, $v_D^{(22)}$

$$v_E^{(11)} = 1; \quad v_A^{(11)} = 1; \quad v_B^{(11)} = -\frac{1}{2}; \quad v_C^{(11)} = -\frac{1}{2};$$

$$v_D^{(11)} = -\frac{1}{2}; \quad v_F^{(11)} = -\frac{1}{2}$$

$$v_E^{(12)} = 0; \quad v_A^{(12)} = 0; \quad v_B^{(12)} = \frac{1}{2}\sqrt{3}; \quad v_C^{(12)} = -\frac{1}{2}\sqrt{3};$$

$$v_D^{(12)} = \frac{1}{2}\sqrt{3}; \quad v_F^{(12)} = -\frac{1}{2}\sqrt{3}$$

$$v_E^{(21)} = 0; \quad v_A^{(21)} = 0; \quad v_B^{(21)} = \frac{1}{2}\sqrt{3}; \quad v_C^{(21)} = -\frac{1}{2}\sqrt{3};$$

$$v_D^{(21)} = -\frac{1}{2}\sqrt{3}; \quad v_F^{(21)} = \frac{1}{2}\sqrt{3}$$

$$v_E^{(22)} = 1; \quad v_A^{(22)} = -1; \quad v_B^{(22)} = \frac{1}{2}; \quad v_C^{(22)} = \frac{1}{2};$$

$$v_D^{(22)} = -\frac{1}{2}; \quad v_F^{(22)} = -\frac{1}{2}$$

aufeinander paarweise orthogonal sein. Z. B. ist

$$(v^{(11)}, v^{(12)}) = 1 \cdot 0 + 1 \cdot 0 + -\frac{1}{2} \cdot \frac{1}{2} \sqrt{3} + -\frac{1}{2} \cdot -\frac{1}{2} \sqrt{3} + -\frac{1}{2} \cdot \frac{1}{2} \sqrt{3} + -\frac{1}{2} \cdot -\frac{1}{2} \sqrt{3} = 0.$$

Ihre Länge muß $\sqrt{h^2} = \sqrt{6/2} = \sqrt{3}$ betragen. Z. B. ist

$$(v^{(21)}, v^{(22)}) = 0^2 + 0^2 + \frac{3}{4} + \frac{3}{4} + \frac{3}{4} = 3.$$

Um noch Beispiele für (30) zu sehen, betrachten wir die offenbar irreduzible Darstellung derselben Gruppe von S :

$$\begin{aligned} \overline{\mathbf{D}}(E) &= (1); \quad \overline{\mathbf{D}}(A) = (-1); \quad \overline{\mathbf{D}}(B) = (-1); \quad \overline{\mathbf{D}}(C) = (-1); \\ \overline{\mathbf{D}}(D) &= (1); \quad \overline{\mathbf{D}}(F) = (1) \end{aligned} \quad \left. \right\} \text{ (++)}$$

und die triviale Darstellung durch lauter (1):

$$\begin{aligned} \overline{\overline{\mathbf{D}}}(E) &= (1); \quad \overline{\overline{\mathbf{D}}}(A) = (1); \quad \overline{\overline{\mathbf{D}}}(B) = (1); \quad \overline{\overline{\mathbf{D}}}(C) = (1); \\ \overline{\overline{\mathbf{D}}}(D) &= (1); \quad \overline{\overline{\mathbf{D}}}(F) = (1). \end{aligned} \quad \left. \right\} \text{ (++++)}$$

Alle vier Vektoren v müssen sowohl auf den Vektor $w_R = \overline{\mathbf{D}}(R)_{11}$ wie auch auf den Vektor $j_R = \overline{\overline{\mathbf{D}}}(R)_{11} = 1$ senkrecht stehen. Z. B. ist

$$(v^{(22)}, w) = 1 \cdot 1 + -1 \cdot -1 + \frac{1}{2} \cdot -1 + \frac{1}{2} \cdot -1 + -\frac{1}{2} \cdot 1 + -\frac{1}{2} \cdot 1 = 0.$$

Betrachten wir nunmehr alle irreduziblen Darstellungen einer Gruppe: $\mathbf{D}^{(1)}(R)$ habe die Dimension l_1 ; $\mathbf{D}^{(2)}(R)$ die Dimension l_2 ; $\mathbf{D}^{(3)}(R)$ die Dimension l_3 ; usw. ... $\mathbf{D}^{(c)}(R)$ die Dimension l_c . Alle sollen in unitärer Form vorliegen. Dann läßt sich (30) und (31) zu

$$\left. \sum_R \mathbf{D}^{(j)}(R)_{\mu\nu} \sqrt{\frac{l_j}{h}} \mathbf{D}^{(j')}(\bar{R})_{\mu'\nu'}^* \sqrt{\frac{l_{j'}}{h}} = \delta_{jj'} \delta_{\mu\mu'} \delta_{\nu\nu'} \right\} \quad (32)$$

$(\mu, \nu = 1, 2, \dots, l_j; \mu' \nu' = 1, 2, \dots, l_{j'}; j, j' = 1, 2, \dots, c)$

zusammenfassen. Die $l_1^3 + l_2^3 + \dots + l_c^3$ h -dimensionalen Vektoren

$$v_R^{(j), \mu, \nu} = \mathbf{D}^{(j)}(R)_{\mu\nu}$$

im Raume der Gruppenelemente stehen im hermiteischen Sinne orthogonal aufeinander.

Da in einem Raume von h -Dimensionen höchstens h -Vektoren aufeinander paarweise orthogonal stehen können, folgt, daß $l_1^3 + l_2^3 + \dots + l_c^3$, die Summe der Quadrate der Dimensionen aller miteinander nichtäquivalenten irreduziblen Darstellungen höchstens gleich der Ordnung der dargestellten Gruppe ist.

In Wirklichkeit läßt sich zeigen, daß die Summe der Quadrate $l_1^3 + l_2^3 + \dots + l_c^3 = h$ genau gleich der Ordnung der Gruppe ist. Auf den Beweis dieses Satzes soll aber hier verzichtet werden.

Wir formen die Gleichung (32) noch etwas um. Bezeichnen wir die Diagonalsumme der Matrix $\mathbf{D}^{(j)}(R)$ mit $\chi^{(j)}(R)$, also

$$\chi^{(j)}(R) = \sum_{\mu=1}^{l_j} \mathbf{D}^{(j)}(R)_{\mu\mu}.$$

Das Zahlensystem, die h -Zahlen $\chi^{(j)}(E), \chi^{(j)}(A_2), \dots, \chi^{(j)}(A_h)$ nennt man den Charakter der Darstellung $\mathbf{D}^{(j)}(R)$. Das Charakterisieren einer Darstellung mit dem Charakter hat den Vorteil, daß sie Ähnlichkeitstransformationen gegenüber invariant ist. Nach (32) ist

$$\sum_R \mathbf{D}^{(j)}(R)_{\mu\mu} \mathbf{D}^{(j')}(R)_{\mu'\mu'}^* = \frac{h}{l_j} \delta_{jj'} \delta_{\mu\mu'},$$

was über μ von 1 bis l_j und über μ' von 1 bis $l_{j'}$ summiert

$$\sum_R \chi^{(j)}(R) \chi^{(j')}(R)^* = \frac{h}{l_j} \delta_{jj'} \sum_{\mu=1}^{l_j} \sum_{\mu'} \delta_{\mu\mu'} = \frac{h}{l_j} \delta_{jj'} \sum_{\mu=1}^{l_j} 1 = h \delta_{jj'} \quad (33)$$

ergibt. Auch die Charaktere $\chi^{(j)}(R)$ bilden ein orthogonales Vektorensystem im Raume der Gruppenelemente. Hieraus folgt auch, daß zwei nichtäquivalente irreduzible Darstellungen nicht dasselbe Charakterensystem haben können, irreduzible Darstellungen mit gleichen Charakteren sind äquivalent.

Die Gleichung (33) läßt sich noch etwas weiter umformen, wenn man die Charaktere $\chi^{(j)}(R)$ und $\chi^{(j)}(S)$ vergleicht, die zu zwei Elementen R und S derselben Klasse gehören. Es gibt dann ein Gruppenelement T , das R in S transformiert. Ist aber $T^{-1}RT = S$, so ist auch $D^{(j)}(T)^{-1}D^{(j)}(R)D^{(j)}(T) = D^{(j)}(S)$, so daß auch $D^{(j)}(R)$ in $D^{(j)}(S)$ transformiert werden kann. Daraus folgt aber, daß die Spur $\chi^{(j)}(R)$ der Matrix $D^{(j)}(R)$ gleich der Spur $\chi^{(j)}(S)$ der Matrix $D^{(j)}(S)$ ist, zu Elementen derselben Klasse gehören gleiche Charaktere. Es genügt also, um das Charakterensystem anzugeben, den Charakter je eines Elementes jeder Klasse der Gruppe anzugeben; den anderen Elementen derselben Klasse entspricht derselbe Charakter. Man kann daher diese Zahl als den Charakter der Klasse bezeichnen. Besteht die ganze Gruppe, um deren Darstellung es sich handelt, aus k -Klassen C_1, C_2, \dots, C_k mit je g_1, g_2, \dots, g_k Elementen ($g_1 + g_2 + \dots + g_k = h$), so ist der Charakter einer Darstellung bereits durch die k -Zahlen $\chi^{(j)}(C_1); \chi^{(j)}(C_2), \dots, \chi^{(j)}(C_k)$ gegeben. Wir können diese Zahlen an Stelle der $\chi^{(j)}(R)$ in (33) einführen und erhalten, indem wir bei der Summation über die h -Gruppenelemente zuerst über die g_ϱ -Elemente derselben Klasse (die entsprechenden g_ϱ -Glieder sind alle gleich), dann über alle k -Klassen summieren:

$$\sum_{\varrho=1}^k \chi^{(j)}(C_\varrho) \chi^{(j')} (C_\varrho)^* g_\varrho = h \delta_{jj'},$$

oder

$$\sum_{\varrho=1}^k \chi^{(j)}(C_\varrho) \sqrt{\frac{g_\varrho}{h}} \cdot \chi^{(j')} (C_\varrho)^* \sqrt{\frac{g_\varrho}{h}} = \delta_{jj'}. \quad (34)$$

Die normierten Charaktere $\chi^{(j)}(C_\varrho) \sqrt{\frac{g_\varrho}{h}}$ bilden schon im k -dimensionalen Raume der Klassen ein im hermiteischen Sinne orthogonales Vektorensystem.

Die Gleichungen (30), (31), (33), (34) sind die wichtigsten Gleichungen der Darstellungstheorie, auf diese werden wir immer wieder zurückgreifen müssen.

Die Charaktere der Darstellungen (\dagger), (\ddagger), ($\ddot{\dagger}$) sind

$$\begin{aligned}\chi^{(E)} &= 2; \quad \chi^{(A)} = 0; \quad \chi^{(B)} = 0; \quad \chi^{(C)} = 0; \quad \chi^{(D)} = -1; \quad \chi^{(F)} = -1 \\ \bar{\chi}^{(E)} &= 1; \quad \bar{\chi}^{(A)} = -1; \quad \bar{\chi}^{(B)} = -1; \quad \bar{\chi}^{(C)} = -1; \quad \bar{\chi}^{(D)} = 1; \quad \bar{\chi}^{(F)} = 1 \\ \bar{\bar{\chi}}^{(E)} &= 1; \quad \bar{\bar{\chi}}^{(A)} = 1; \quad \bar{\bar{\chi}}^{(B)} = 1; \quad \bar{\bar{\chi}}^{(C)} = 1; \quad \bar{\bar{\chi}}^{(D)} = 1; \quad \bar{\bar{\chi}}^{(F)} = 1\end{aligned}$$

D, F und A, B, C sind in derselben Klasse, ihnen entspricht derselbe Charakter. Man kann daher die Charaktere zusammenfassen:

$$\begin{aligned}\chi^{(E)} &= 2; \quad \chi^{(A, B, C)} = 0; \quad \chi^{(D, F)} = -1; \\ \bar{\chi}^{(E)} &= 1; \quad \bar{\chi}^{(A, B, C)} = -1; \quad \bar{\chi}^{(D, F)} = 1; \\ \bar{\bar{\chi}}^{(E)} &= 1; \quad \bar{\bar{\chi}}^{(A, B, C)} = 1; \quad \bar{\bar{\chi}}^{(D, F)} = 1.\end{aligned}$$

Die normierten Charaktere $\sqrt{\frac{g_\rho}{h}} \chi^{(l)}(C_g)$ sind orthogonal zueinander. Z. B. ist für χ und $\bar{\chi}$

$$\sqrt{\frac{1}{6}} 2 \sqrt{\frac{1}{6}} 1 + \sqrt{\frac{5}{6}} 0 \sqrt{\frac{2}{6}} \cdot (-1) + \sqrt{\frac{2}{6}} \cdot (-1) \sqrt{\frac{2}{6}} 1 = 0.$$

Da es nicht mehr als k orthogonal-normierte k -dimensionale Vektoren geben kann, folgt, daß die Anzahl c der miteinander nicht-äquivalenten irreduziblen Darstellungen höchstens gleich der Anzahl k der Klassen der dargestellten Gruppe ist. In Wirklichkeit ist diese Zahl immer erreicht, es läßt sich zeigen, daß die Anzahl der nichtäquivalenten irreduziblen Darstellungen einer Gruppe gleich der Anzahl der Klassen dieser Gruppe $c = k$ ist.

Ein Beispiel haben wir hierfür schon in den auf S. 89—90 gegebenen drei Darstellungen der symmetrischen Gruppe dritten Grades gesehen; die Gruppe besteht aus den drei Klassen $E; A, B, C; D, F$ und kann also außer den angeführten keine weiteren irreduziblen Darstellungen haben. Die Dimensionen der Darstellungen sind $2, 1, 1$, in der Tat ist $2^2 + 1^2 + 1^2 = 6$ die Ordnung der Gruppe.

Ausreduzieren einer Darstellung. Bisher hatten wir uns ziemlich regellos mit reduziblen und irreduziblen Darstellungen befaßt. Die Sätze 1 und 1a bezogen sich auf beliebige, die Sätze 2, 3 und 4 [unter anderen die Gleichungen (30), (31), (33), (34)] auf irreduzible Darstellungen.

Die Wichtigkeit der irreduziblen Darstellungen beruht darauf, daß jede Darstellung sich in eindeutiger Weise in irreduzible zerlegen, d. h. durch Ähnlichkeitstransformation mit einer zweckmäßig gewählten Matrix „ausreduzieren“, in die Gestalt

$$\left(\begin{array}{cccc} \mathbf{D}^{(1)}(R) & \mathbf{0} & \dots & \mathbf{0} \\ \mathbf{0} & \mathbf{D}^{(2)}(R) & \dots & \mathbf{0} \\ \vdots & \vdots & & \vdots \\ \mathbf{0} & \mathbf{0} & \dots & \mathbf{D}^{(s)}(R) \end{array} \right) \quad (**)$$

bringen läßt, wo die $\mathbf{D}^{(j)}(R)$ nunmehr irreduzible Darstellungen, die irreduziblen Bestandteile der ursprünglichen Darstellung sind.

Ist eine Darstellung $\mathbf{D}(R)$ nicht schon selber irreduzibel, so läßt sie sich (d. h. alle ihre Matrizen lassen sich) auf die Gestalt (*)

$$\bar{\mathbf{D}}(A_x) = \begin{pmatrix} \mathbf{D}'(A_x) & \mathbf{0} \\ \mathbf{0} & \mathbf{D}''(A_x) \end{pmatrix} \quad (*)$$

transformieren. Nunmehr sind entweder beide Teile $\mathbf{D}'(A_x)$ und $\mathbf{D}''(A_x)$ irreduzibel, oder ist z. B. $\mathbf{D}''(A_x)$ reduzibel. Dann kann man $\bar{\mathbf{D}}(A_x)$ einer weiteren Ähnlichkeitstransformation mit

$$\begin{pmatrix} \mathbf{S} & \mathbf{0} \\ \mathbf{0} & \mathbf{T} \end{pmatrix}$$

unterwerfen. So entsteht aus ihr

$$\bar{\bar{\mathbf{D}}}(A_x) = \begin{pmatrix} \mathbf{S}^{-1} \mathbf{D}'(A_x) \mathbf{S} & \mathbf{0} \\ \mathbf{0} & \mathbf{T}^{-1} \mathbf{D}''(A_x) \mathbf{T} \end{pmatrix}.$$

War etwa $\mathbf{D}''(A_x)$ noch reduzibel, so kann man \mathbf{T} so wählen, daß $\mathbf{T}^{-1} \mathbf{D}''(A_x) \mathbf{T}$ die Form (*) hat. Dann hat $\bar{\bar{\mathbf{D}}}(A_x)$ die Form

$$\begin{pmatrix} \mathbf{D}'(A_x) & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{D}'''(A_x) & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{D}''''(A_x) \end{pmatrix}$$

die man, sind noch immer nicht alle drei Darstellungen $\mathbf{D}', \mathbf{D}''', \mathbf{D}''''$ irreduzibel, weiter reduzieren kann. Da die Darstellung von endlicher Dimension ist, muß man sie so schließlich auf die Form (**) bringen können, wo alle Darstellungen $\mathbf{D}^{(1)}, \mathbf{D}^{(2)}, \dots, \mathbf{D}^{(s)}$ nunmehr irreduzibel sind. Da mehrere aufeinanderfolgende Ähnlichkeitstransformationen immer durch eine ersetzt werden können, kann man die vorgelegte Darstellung auch durch eine Transformation sofort auf die Form (**) bringen. Man nennt diesen Prozeß Ausreduktion, (**) die ausreduzierte Form.

Führt man diese Ausreduktion aus, so wäre es möglich, daß die irreduziblen Teile $\mathbf{D}^{(1)}(R), \mathbf{D}^{(2)}(R), \dots, \mathbf{D}^{(s)}(R)$ nicht eindeutig bis auf Ähnlichkeitstransformationen bestimmt sind, daß man $\mathbf{D}(R)$ vielmehr auf mehrere Arten ausreduzieren kann. Wir können aber zeigen, daß dies nicht der Fall ist. Ebenso, wie man eine ganze rationale Zahl nur auf eine Weise in ein Produkt von Primzahlen zerlegen kann, sind auch hier die irreduziblen Bestandteile —

natürlich bis auf die Reihenfolge — eindeutig bestimmt. Für den Charakter $\chi(R)$ einer beliebigen reduziblen Darstellung gilt

$$\chi(R) = \sum_{i=1}^c a_i \chi^{(i)}(R) \quad (R = E, A_2, A_3, \dots, A_h), \quad (35)$$

wenn es bei einer Ausreduktion in a_1 irreduzible Darstellungen $D^{(1)}(R)$ [mit dem Charakter $\chi^{(1)}(R)$], in a_2 Darstellungen $D^{(2)}(R)$ [mit dem Charakter $\chi^{(2)}(R)$] ... usw. zerfällt.

Nun bestimmen aber die h -Gleichungen (35) die Zahlen a_1, a_2, \dots, a_c vollkommen. Bildet man nämlich das skalare Produkt von (35) mit $\chi^{(j')}(R)$, d. h. multipliziert man mit $\chi^{(j')}(R)^*$ und summiert über alle Gruppenelemente, so erhält man wegen (33)

$$\sum_R \chi(R) \chi^{(j')}(R)^* = \sum_R \sum_j a_j \chi^{(j)}(R) \chi^{(j')}(R)^* = h a_{j'}, \quad (36)$$

so daß die ganze Zahl $a_{j'}$ eindeutig zu

$$a_{j'} = \frac{1}{h} \sum_R \chi(R) \chi^{(j')}(R)^* \quad (37)$$

bestimmt ist. Wie oft eine irreduzible Darstellung in der ausreduzierten Form einer Darstellung enthalten ist, ist nach (37) schon durch den Charakter der Darstellung bestimmt. Insbesondere sind also die irreduziblen Bestandteile unabhängig davon, wie man die Ausreduktion vornimmt. Außerdem sehen wir aber, daß zwei beliebige Darstellungen äquivalent sind, wenn sie denselben Charakter haben. Man kann sie nämlich ineinander transformieren, indem man sie beide ausreduziert. Dann können sie sich noch höchstens in der Form der einzelnen irreduziblen Bestandteile unterscheiden, was man durch eine Ähnlichkeitstransformation ausgleichen kann, und in der Reihenfolge der irreduziblen Bestandteile in der Form (**), die man durch eine bloße gleichzeitige Umbenennung der Zeilen und Spalten, also auch durch eine Ähnlichkeitstransformation beseitigen kann.

Da die Gleichheit der Charaktere andererseits notwendig für die Äquivalenz zweier Darstellungen ist, ist sie die notwendige und hinreichende Bedingung für ihre Äquivalenz, für ihre Ineinandertransformierbarkeit.

Für die Ineinandertransformierbarkeit zweier einzelner Matrizen würde dies, daß die Summe der Eigenwerte übereinstimmt, keineswegs genügen, die Eigenwerte müßten auch einzeln übereinstimmen. Bei zwei Darstellungen

folgt aber daraus, daß bei allen h entsprechenden Matrizenpaaren die Summe der Eigenwerte dieselbe ist, auch, daß alle entsprechenden Eigenwerte einzeln übereinstimmen. Es genügt sogar etwas weniger. Da die Charaktere aller Gruppenelemente derselben Klasse in jeder Darstellung gleich sind, genügt schon die Gleichheit der k -Zahlen $\chi(C_1) = \chi'(C_1)$, $\chi(C_2) = \chi'(C_2)$, ..., $\chi(C_k) = \chi'(C_k)$ für die Äquivalenz zweier Darstellungen mit den Charakteren χ und χ' .

Wir wollen noch eine Formel über die Anzahl der in einer Darstellung enthaltenen irreduziblen Bestandteile ableiten. Nehmen wir das skalare Produkt von (35) mit sich selber, so erhalten wir

$$\begin{aligned} \sum_R |\chi(R)|^2 &= \sum_R \sum_j a_j \chi^{(j)}(R) \sum_{j'} a_{j'} \chi^{(j')}(R)^* \\ &= \sum_j \sum_{j'} h \delta_{jj'} a_j a_{j'} = h \sum_j a_j^2. \end{aligned} \quad (38)$$

Die Summe der Absolutwertquadrat der Charaktere einer Darstellung ist gleich der Ordnung h der Gruppe, multipliziert mit der Summe der Quadrate der Anzahlen a_j , wie oft die einzelnen irreduziblen Darstellungen in dieser Darstellung enthalten sind. Für eine irreduzible Darstellung hat also die Summe

$$\sum_R |\chi(R)|^2 = h \quad (38a)$$

den kleinstmöglichen Wert h , und es folgt umgekehrt aus dem Bestehen von (38a), daß die Darstellung mit der Spur $\chi(R)$ irreduzibel ist, da sie wegen (38) nach der Ausreduktion nur einen Bestandteil enthält.

In vielen Fällen genügen schon die allgemeinen Sätze zur expliziten Angabe der irreduziblen Darstellungen. Besonders die beiden von uns nur teilweise bewiesenen Sätze über die Anzahl der nichtäquivalenten Darstellungen (gleich Anzahl der Klassen) und über die Summe der Quadrate ihrer Dimensionen (gleich der Ordnung der Gruppe) leisten dabei gute Dienste. In den meisten Fällen allerdings sind dazu noch weitergehende, spezielle Untersuchungen notwendig.

Zunächst bei Abelschen Gruppen bildet jedes Element eine Klasse für sich, die Gruppe hat so viele Darstellungen wie Elemente. Da die Summe der Quadrate aller Darstellungen ebenfalls gleich ihrer Ordnung ist, ist jede irreduzible Darstellung von der Dimension 1.

Weiter ist die schon am Anfang betonte Tatsache zu beachten, daß jede Darstellung einer Faktorgruppe auch eine Darstellung der ganzen Gruppe ist. So hat z. B. die schon oft besprochene symmetrische Gruppe von drei Elementen einen Normalteiler E, D, F , die Faktorgruppe ist von der Ordnung 2, ist also abelsch und hat zwei Darstellungen von der Dimension 1. Da die ganze Gruppe nur drei Klassen hat, hat die Gruppe nur noch eine irreduzible Darstellung, und diese muß zweidimensional sein, damit $1^2 + 1^2 + 2^2 = 6$ sei.

Die verschiedenen irreduziblen Darstellungen spielen in der Quantenmechanik eine besonders wichtige Rolle, weil sie zur Charakterisierung der verschiedenen Termsysteme dienen, die Terme gleicher Auswahlregeln usw.

zusammenfassen. Aber auch von rein mathematischem Standpunkt aus kann die soeben gegebene Theorie — sie stammt in erster Reihe von S. Frobenius, H. Burnside und J. Schur — als einer der schönsten Teile der Algebra betrachtet werden. Die Charaktere der Darstellungen enthalten auch einige interessante Ganzahligkeitsbeziehungen.

X. Kontinuierliche Gruppen

1. Bisher hatten wir es nur mit Gruppen mit endlich vielen Gruppenelementen, mit endlichen Gruppen zu tun. Unsere drei Gruppenpostulate (Assoziativgesetz, Einheit, Reziproke) lassen sich aber auch auf eine unendliche Mannigfaltigkeit von Dingen, auf unendliche Gruppen anwenden. Z. B. bilden die dreidimensionalen reellen orthogonalen Matrizen, die Drehungen im Raum, ein System von Dingen, die den Gruppenpostulaten genügen, wenn man die Gruppenmultiplikation der Matrixmultiplikation gleichsetzt, unter dem Produkt zweier Drehungen ihre Zusammensetzung versteht. Eine ähnliche Gruppe bilden alle dreidimensionalen Matrizen mit der Determinante 1, oder mit der Determinante ± 1 usw. Alle diese Gruppen bezeichnet man als unendliche Gruppen im Gegensatz zu den bisher betrachteten endlichen Gruppen.

Der Begriff der unendlichen Gruppe wäre, wenn man nur die erwähnten Eigenschaften von den Gruppen verlangen würde, für unsere Zwecke etwas zu weit gefaßt. Man könnte z. B. alle zweidimensionalen Matrizen mit der Determinante 1, deren Koeffizienten alle vier rational sind, zu einer Gruppe zusammenfassen. Ein solches System würde viele Stetigkeitseigenschaften entbehren, die wir voraussetzen wollen. Deshalb empfiehlt es sich, den Begriff der Gruppe von dieser Seite etwas einzuzengen. Unter einer kontinuierlichen Gruppe versteht man ein System von Dingen, Gruppenelemente genannt, die durch in einem gewissen Bereich stetig sich ändernde Parameter gekennzeichnet werden können. Es gehört zu jedem innerhalb des betreffenden Bereiches gelegenen Wertesystem der Parameter ein Gruppenelement und umgekehrt zu jedem Gruppenelement ein im Variabilitätsbereich der Parameter liegendes Wertesystem dieser.

Gruppenelemente, deren Parameter sich nur wenig unterscheiden, nennt man „benachbart“. Wenn man die Parameter stetig ändert, so sagt man, daß sich das Gruppenelement stetig ändert. Die anderen drei Gruppenpostulate bleiben erhalten und werden noch

durch Stetigkeitsforderungen verschärft, indem verlangt wird, daß das Produkt benachbarter Faktoren und die Reziproken benachbarter Elemente wieder benachbart seien.

Wir werden sogar annehmen, daß die Parameter $p_1(RS)$, $p_2(RS)$, ..., $p_n(RS)$ des Produkts zweier Elemente stetig differentiierbare Funktionen der Parameter $p_1(R)$, $p_2(R)$, ..., $p_n(R)$ und $p_1(S)$, $p_2(S)$, ..., $p_n(S)$ der beiden Faktoren R und S sind. Dasselbe werde von der Abhängigkeit der Parameter $p_1(R^{-1})$, $p_2(R^{-1})$, ..., $p_n(R^{-1})$ von den Parametern von R verlangt.

Gruppen, deren Elemente durch n Parameter gekennzeichnet werden können, nennt man eine n -parametrische Gruppe. Der Variabilitätsbereich der Parameter kann einfach oder mehrfach zusammenhängend sein, oder auch in mehrere getrennte Gebiete zerfallen. Im letzteren Fall spricht man von einer gemischtkontinuierlichen Gruppe im Gegensatz zu einfachkontinuierlichen Gruppen, bei denen der Variabilitätsbereich der Parameter zusammenhängend ist.

Betrachten wir z. B. die eingangs erwähnte Gruppe der Drehungen im dreidimensionalen Raum, die dreidimensionale Drehgruppe. Ein Element der Gruppe, eine reelle dreidimensionale orthogonale Matrix, läßt sich natürlich auch durch Angabe ihrer neun Koeffizienten charakterisieren. Diese können aber nicht als Parameter angesehen werden, da sie nicht unabhängig voneinander variierbar sind, weil unter ihnen noch Beziehungen bestehen. Nimmt man dagegen zur Charakterisierung einer Drehung etwa Azimut Φ und Polabstand ϑ der Drehachse und den Drehwinkel φ , so entspricht jedem in gewissen Bereichen liegenden Wertetripel ($0 \leq \Phi < 2\pi$, $0 \leq \vartheta < \pi$, $0 \leq \varphi \leq \pi$) dieser Zahlen eine Drehung. Umgekehrt entspricht auch jeder Drehung ein Wertesystem dieser Parameter¹⁾.

Eine Ausnahme bildet hier von nur die Drehung mit dem Drehwinkel $\varphi = 0$, also die Drehung, die eigentlich keine Drehung ist, sondern der Ruhe entspricht, die Einheit der Gruppe. Ihr entspricht nämlich jedes Parametertripel Φ , ϑ , 0 , die Zuordnung der Parameter zu diesem Gruppen-element ist nicht eindeutig. Man könnte daran denken, dem so abzuhelfen, daß man für $\varphi = 0$ den Variabilitätsbereich der beiden anderen Parameter Φ , ϑ auf den Wert 0 beschränkt. Dann wäre die Zuordnung von Drehungen zu Parametertripeln tatsächlich ein-eindeutig. Betrachten wir aber die stetige

¹⁾ Ein Punkt R des Parameterraumes entspricht nicht der Operation einer Drehung, sondern ihrem Resultat. Die Drehung in diesem Sinne ist vollkommen durch Anfangs- und Endlage der Kugel bestimmt. Will man die Operation, d. h. den Weg, über den die Drehung erfolgt ist, beschreiben, so muß man alle Zwischenlagen der Kugel, also eine Linie im Parameterraume oder eine stetige Folge $R(t)$ von „Drehungen“ angeben, die für $t = 0$ von $R(0) = E$ ausgeht und etwa für $t = 1$ in $R(1) = R$ der gewünschten Drehung endet.

Reihe der Drehungen um eine beliebige Drehachse um immer kleinere und kleinere Winkel, also die Drehungen mit den Parametern $\Phi = \Phi_0$, $\vartheta = \vartheta_0$, $\varphi = t\varphi_0$, wo t stetig abnimmt. Für $t = 0$ müssten auch Φ und ϑ gleich 0 werden, also einen Sprung erleiden, obwohl die Reihe der Drehungen durchaus stetig verändert wurde. Daher ist die zunächst naheliegende Beschränkung $\Phi = 0$, $\vartheta = 0$ für $\varphi = 0$ nicht zulässig und man muß sich anderswie zu helfen versuchen.

Am besten benutzt man als Parameter die kartesischen Koordinaten ξ , η , ζ desjenigen Punktes, dessen Polarkoordinaten q/π , ϑ , Φ sind, also $\xi = q/\pi \sin \vartheta \sin \Phi$, $\eta = q/\pi \sin \vartheta \cos \Phi$, $\zeta = q/\pi \cos \vartheta$. Der Ruhe, der Einheit der Gruppe, der bisher die Parameter Φ , ϑ , 0 (Φ und θ beliebig) entsprachen, entspricht jetzt der einzige Punkt $\xi = \eta = \zeta = 0$, der Koordinatenaufangspunkt. Die Zuordnung der Punkte des Parameterraumes zu den Drehungen wird durch Abb. 1 veranschaulicht. Jeder Drehung entspricht ein Punkt in der Einheitskugel im ξ , η , ζ -Raum.

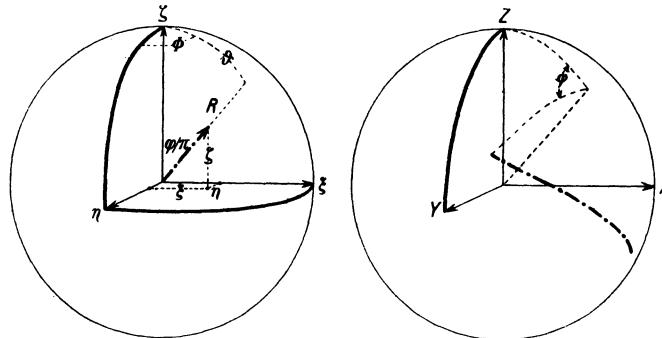


Abb. 1. Dem Punkte R der Figur links entspricht die Drehung, die auf der Figur rechts den ausgezogenen in den strichpunktieren Bogen überführt

Auch bei dieser Zuordnung existiert zunächst eine Ausnahme, wo die Zuordnung von Gruppenelementen zu Parametertripeln nicht ein-eindeutig ist: den beiden Antipoden der Kugeloberfläche entspricht dieselbe Drehung, da sich zwei Drehungen um entgegengerichtete Achsen mit dem Winkel π gar nicht unterscheiden. Die Antipoden der Kugeloberfläche muß man daher identifizieren und den Übergang zum Antipoden auf der Kugeloberfläche im ξ , η , ζ -Raum nicht als eine Unterbrechung einer stetigen Linie betrachten.

Hat man aber so den Parameterraum zu einem mehrfach zusammenhängenden gemacht, so ist die Zuordnung der Gruppenelemente zu den Punkten dieses Raumes wirklich ein-eindeutig, und diese Parameter sind daher für prinzipielle Betrachtungen über die Drehgruppe besonders geeignet.

Dies schließt natürlich keineswegs aus, daß man bei formelmäßigem Rechnen andere Parameter bevorzugt (die Eulerschen Winkel, Abb. 2), bei denen die Zuordnung zwar nicht ein-eindeutig ist, aber die meisten expliziten Formeln übersichtlicher als für die ξ , η , ζ ausfallen.

2. Den Teil der Gruppe, der der Einheit benachbart ist, bezeichnet man als **Infinitesimalgruppe**. Mit Infinitesimalgruppen von Transformationsgruppen beschäftigen sich die grundlegenden Arbeiten von Sophus Lie¹⁾. Auf diese leider so wenig bekannten Untersuchungen können wir nur andeutungsweise eingehen, indem wir uns auf die Grundtatsachen beschränken und alle Existenz- und Konvergenzbeweise weglassen. Bei den uns interessierenden Gruppen kann man sie durch wirkliches Vorzeigen der betreffenden Gruppenelemente usw. leicht nachholen.

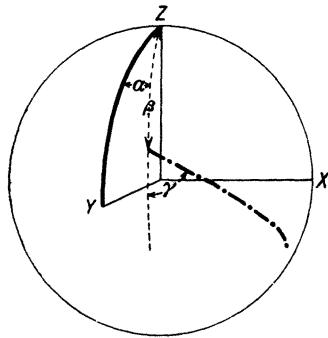


Abb. 2. Die Drehung der Abb. 1, die den ausgezogenen in den strichpunktierten Bogen überführt, hat die Eulerschen Winkel α, β, γ

Am besten führt man gleich eine unendlich kleine Zahl h ein. Die Parameter der Einheit der Gruppe seien $\pi_1, \pi_2, \dots, \pi_n$. Dann betrachten wir die n Elemente F_1, F_2, \dots, F_n , die Parameter von F_k seien $\pi_1, \pi_2, \dots, \pi_k + h, \dots, \pi_n$. Die Elemente $E, F_k, F_k^2, F_k^3, \dots$ liegen alle sehr nahe zueinander, ist h genügend klein, so bilden sie beinahe eine stetige Folge von Gruppenelementen. Man muß nur, wenn man einigermaßen weit von der Einheit kommen will, zu sehr hohen Potenzen von F_k vordringen. Eine solche einparametrische Schar von (miteinander vertauschbaren!) Gruppenelementen entspricht den sogenannten Perioden der Elemente endlicher Gruppen.

Eine einfachkontinuierliche einparametrische Gruppe ist immer abelsch, da sie aus einer einzigen Periode besteht.

Bei der dreidimensionalen Drehgruppe waren die Parameter der Einheit $\xi = \eta = \zeta = 0$, die drei Infinitesimalelemente $h, 0, 0; 0, h, 0; 0, 0, h$ sind die drei unendlich kleinen Drehungen um die drei Koordinatenachsen.

Betrachten wir nun die n -parametrische Schar $F_1^{p_1} F_2^{p_2} \dots F_n^{p_n}$. Beschränken wir uns auf solche Werte von h und p_1, p_2, \dots, p_n , bei denen die Elemente der Schar noch alle in der Nähe der Einheit sind, so wird sie, da sie n -parametrisch ist, die ganze Infinitesimalgruppe umfassen. Es ist zweckmäßig, wenigstens für die In-

¹⁾ S. Lie, Vorlesungen über kontinuierliche Gruppen. Leipzig 1893.

infinitesimalgruppe die p_1, p_2, \dots, p_n als Parameter einzuführen. Die n Parameter der Einheit sind alle 0, die Parameter der Elemente der Infinitesimalgruppe noch alle sehr klein.

Die Elemente der Infinitesimalgruppe sind vertauschbar: sind die Parameter zweier Faktoren R und S in der Größenordnung einer kleinen Zahl ε , so ist die Differenz der Parameter von RS und SR in der Größenordnung ε^2 . Betrachten wir die Elemente $R(t)$ und $S(t')$ mit den Parametern tr_1, tr_2, \dots, tr_n und $t's_1, t's_2, \dots, t's_n$, wo t und t' als stetig veränderliche Variable angesehen werden sollen. Es ist $E = R(0) = S(0)$ und t und t' liegen in der Größenordnung von ε . Man kann annehmen, daß man die Parameter von $R(t)S(t')$ und $S(t')R(t)$ in eine MacLaurinsche Reihe nach t und t' entwickeln kann. Die Reihen sollen für die n Parameter von $R(t)S(t')$ bzw. von $S(t')R(t)$ folgendermaßen lauten:

$$u_1 + tv_1 + t'w_1 + \dots; u_2 + tv_2 + t'w_2 + \dots; \dots; u_n + tv_n + t'w_n + \dots \\ \text{bzw.}$$

$$\bar{u}_1 + t\bar{v}_1 + t'\bar{w}_1 + \dots; \bar{u}_2 + t\bar{v}_2 + t'\bar{w}_2 + \dots; \dots; \bar{u}_n + t\bar{v}_n + t'\bar{w}_n + \dots$$

Um die u und \bar{u} zu bestimmen, setzen wir $t = t' = 0$. Dann sind beide Produkte $R(0)S(0)$ und $S(0)R(0)$ gleich E , die

$$u_1 = u_2 = \dots = u_n = \bar{u}_1 = \bar{u}_2 = \dots = \bar{u}_n = 0$$

alle Null. Um die v und \bar{v} zu bestimmen, setzen wir nunmehr $t = 0$, dann ist $R(t)S(0) = R(t)E = R(t)$ und $S(0)R(t) = E R(t) = R(t)$ und daher ist $v_1 = \bar{v}_1 = r_1; v_2 = \bar{v}_2 = r_2; \dots; v_n = \bar{v}_n = r_n$. Ebenso erhalten wir, indem wir $t = 0$ setzen,

$$w_1 = \bar{w}_1 = s_1; w_2 = \bar{w}_2 = s_2; \dots; w_n = \bar{w}_n = s_n.$$

Bei den in Betracht gezogenen Gliedern unterscheiden sich die Parameter der beiden Produkte noch nicht: ein Unterschied kann erst bei den Gliedern mit t^2, tt' , oder t'^2 eintreten, die aber schon alle in der Größenordnung von ε^2 liegen.

Die Vertauschbarkeit der Infinitesimalelemente bis auf quadratische Glieder beruht darauf, daß sie sowohl bei beliebigem S und $R = E$, wie auch bei beliebigem R und $S = E$ genau richtig ist. Weicht R nur wenig von E ab, so gilt sie noch näherungsweise, ebenso, wenn $S \sim E$ ist. Ist aber sowohl $R \sim E$, wie auch $S \sim E$, so gilt die Vertauschbarkeit besonders gut.

Sehr einfach stellt sich dieses Gesetz für Matrizengruppen dar: die Elemente in der Nähe der Einheit haben die Gestalt $1 + \varepsilon a$. Nun ist

$$(1 + \varepsilon a)(1 + \varepsilon b) = 1 + \varepsilon(a + b) + \varepsilon^2 ab$$

und

$$(1 + \varepsilon b)(1 + \varepsilon a) = 1 + \varepsilon(b + a) + \varepsilon^2 ba,$$

sie unterscheiden sich nur in Gliedern mit ε^2 .

Die vorher besprochene Parameterwahl, die dem Element $F_1^{p_1} F_2^{p_2} \dots F_n^{p_n}$ die Parameter p_1, p_2, \dots, p_n zuordnet, hat die besondere Eigenschaft, daß man die Parameter des Produktes zweier Elemente der Infinitesimalgruppe bis auf Glieder höherer Ordnung einfach durch Addition der Parameter der einzelnen Faktoren erhalten kann.

3. Bei einer gemischtkontinuierlichen Gruppe bilden die Elemente, die von der Einheit E aus stetig zu erreichen sind, eine Untergruppe, die einfachkontinuierlich ist. Daß ein Element stetig von der Einheit zu erreichen ist, bedeutet so viel, daß eine stetige Mannigfaltigkeit $R(t)$ von Elementen existiert, die von $R(0) = E$ ausgeht und — etwa für $t = 1$ — in $R(1) = R$ mündet. Sind R und S zwei derartige Elemente und $R(t)$ und $S(t)$ zwei zugehörige Wege, so ist ihr Produkt $R(1)S(1) = RS$ auch stetig von der Einheit aus erreichbar: ein zugehöriger Weg ist etwa $R(t)S(t)$. Dasselbe gilt von der Reziproken von R : der zugehörige Weg ist $R(t)^{-1}$.

Die stetig von der Einheit aus erreichbaren Elemente bilden daher eine Untergruppe der ganzen Gruppe, die, da der zugehörige Parameterbereich zusammenhängend ist, einfachkontinuierlich ist.

Die stetig von der Einheit erreichbaren Elemente bilden sogar einen Normalteiler der ganzen gemischtkontinuierlichen Gruppe. Mit R ist nämlich auch $X^{-1}RX$ stetig — etwa auf dem Wege $X^{-1}R(t)X$ — von der Einheit aus erreichbar. Die Nebengruppen des Normalteilers sind die anderen voneinander getrennten Teile des Parameterraumes. Die Faktorgruppe kann also als endlich, von der Ordnung der Anzahl der nichtzusammenhängenden Parameterbereiche angenommen werden.

Für das Folgende führen wir noch die Bezeichnung $\{r_1, r_2, \dots, r_n\}$ für das Gruppenelement mit den Parametern r_1, r_2, \dots, r_n ein. Es gilt dann natürlich identisch $p_k(\{r_1, r_2, \dots, r_n\}) = r_k$ und $\{p_1(R), p_2(R), \dots, p_n(R)\} = R$.

4. Die Darstellung einer kontinuierlichen Gruppe ist genau so wie die einer endlichen Gruppe definiert: es soll jedem Element R der Gruppe eine Matrix $D(R)$ zugeordnet werden, so daß stets

$\mathbf{D}(R) \mathbf{D}(S) = \mathbf{D}(RS)$ sei. Es wird nur noch die Stetigkeit der Darstellung verlangt, d. h. wenn R und S benachbarte Gruppenelemente sind, so dürfen sich alle l^2 Koeffizienten $\mathbf{D}(R)_{\alpha\lambda}$ von den entsprechenden Koeffizienten $\mathbf{D}(S)_{\alpha\lambda}$ nur unendlich wenig unterscheiden. Auch hier werden wir uns auf Darstellungen mit nicht-verschwindender Determinante beschränken.

Wollen wir die Gesetze für die Darstellungen endlicher Gruppen auf die Darstellungen kontinuierlicher Gruppen übertragen, so können wir davon ausgehen, daß wir für die ersten vier Sätze, also insbesondere auch für die Orthogonalitätsrelationen (30), (31) nur eine Gruppeneigenschaft benutzt haben, nämlich die, daß man gewisse Summen $\sum_R J_R$ bilden kann, für die, wenn man die Summation über alle Gruppenelemente erstreckt,

$$\sum_R J_R = \sum_R J_{SR} \quad (1)$$

gilt. Dabei sind die J_R ganz beliebige Zahlen (oder Matrizen), die zu je einem Gruppenelement R zugeordnet sind, und (1) gilt für jedes Gruppenelement S . In beiden Summen in (1) kommen ja dieselben Glieder, nur in verschiedener Reihenfolge vor. Mit Hilfe einer solchen Summe $\sum_R \mathbf{D}(R) \mathbf{D}(R)^\dagger$ haben wir den Satz 1, die Unitarisierbarkeit der Darstellungen, bewiesen, auch die Orthogonalitätsrelationen (Satz 4) erhielten wir durch Betrachtung einer solchen Summe $\sum_R \mathbf{D}^{(j)}(R) \mathbf{X} \mathbf{D}^{(j')}(R^{-1})$. Die Sätze 2 und 3 waren eher matrizentheoretischer Natur, zu ihrer Abteilung haben wir keine neuen Gruppeneigenschaften mehr benutzt.

Könnten wir auch für kontinuierliche Gruppen etwas Ähnliches, wie die Summe $\sum_R J_R$ definieren (es wird sich hier natürlich ein über den ganzen Variabilitätsbereich der Parameter erstrecktes Integral ergeben), so könnten wir die vier Sätze der Darstellungstheorie endlicher Gruppen auf kontinuierliche Gruppen übertragen.

5. Bei endlichen Gruppen beruht (1) darauf, daß die Reihe SE, SA_2, \dots, SA_h bei beliebigem S bis auf die Reihenfolge mit der Reihe E, A_1, \dots, A_h identisch ist. Dies wird durch die Abb. 3 veranschaulicht, in der jedem Element der Gruppe der

Vertauschungen von drei Objekten [(†), Kap. VII] ein Punkt entspricht. Die Pfeile deuten an, in welches Gruppenelement ein bestimmtes durch vordere Multiplikation mit F übergeht. Wir sehen, daß das Gesamtbild vollkommen erhalten bleibt, wenn man Pfeilbeginn durch Pfeilspitze ersetzt.

Wollen wir ein entsprechendes Bild für eine (n -parametrische) kontinuierliche Gruppe konstruieren, so müssen wir den ganzen

Variabilitätsbereich des n -dimensionalen Parameterraumes sehr dicht mit Punkten anfüllen. Die vordere Multiplikation bedeutet dann auch hier das Ansetzen eines Pfeiles an jeden Punkt. Es wird sich nun zeigen, daß es zwar im allgemeinen nicht möglich ist, die erwähnten Punkte so zu legen, daß bei dem Ersetzen von Pfeilbeginn durch Pfeilspitze das ganze Bild ungeändert

bleibe¹⁾), daß es aber sehr wohl möglich ist, sie so zu wählen, daß nach der erwähnten Operation die Dichte der Punkte an jeder Stelle des Parameterraumes ebenso groß sei, wie sie vor der Operation an derselben Stelle des Parameterraumes war. Ist $J(R)$ eine im Parameterraume stetige Funktion, so ist dann

$$\sum_R J(R) = \sum_{SR} J(SR), \quad (2)$$

wenn R beidemal alle diese Punkte durchläuft.

Bezeichnen wir die Dichte der in der Summe (2) vorkommenden Punkte im Parameterraum in der Nähe des Punktes $Q = \{q_1, q_2, \dots, q_n\}$ mit $g(Q) = g(\{q_1, q_2, \dots, q_n\})$, so bedeutet dies, daß in (2) $g(\{q_1, q_2, \dots, q_n\}) \Delta_1 \Delta_2 \dots \Delta_n$ Punkte vorkommen, deren erste Parameter zwischen q_1 und $q_1 + \Delta_1$, deren zweite Parameter zwischen q_2 und $q_2 + \Delta_2$ und deren n -te Parameter zwischen q_n und $q_n + \Delta_n$ liegen. Dieses parallelepipedische Gebiet bezeichnen wir mit U .

Die Dichte der Punkte soll sich bei dem Ersetzen eines jeden Punktes R durch SR an keiner Stelle ändern. Wählen wir insbesondere $S = Q^{-1}$, so gehen die Punkte des Gebietes U alle

¹⁾ Dies ist nur bei Abelschen Gruppen möglich. Auch dies zeigt, wie verschieden kontinuierliche von endlichen Gruppen sind und daß erstere nicht etwa durch Grenzübergang aus letzteren zu gewinnen sind.

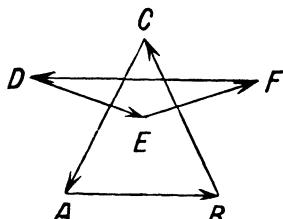


Abb. 3

in ein Gebiet U_0 über, das in der Nähe der Einheit liegt. Die Größe des Gebietes U_0 ist bis auf höhere Potenzen der Δ

$$\gamma_0 = \left| p_1(Q^{-1} \cdot \{q_1 + \Delta_1, q_2, \dots, q_n\}) \dots p_n(Q^{-1} \cdot \{q_1 + \Delta_1, q_2, \dots, q_n\}) \right| \\ \vdots \\ \left| p_1(Q^{-1} \cdot \{q_1, q_2, \dots, q_n + \Delta_n\}) \dots p_n(Q^{-1} \cdot \{q_1, q_2, \dots, q_n + \Delta_n\}) \right| \\ - \Delta_1 \Delta_2 \dots \Delta_n \frac{\partial [p_1(Q^{-1} \cdot \{q_1, \dots, q_n\}), p_2(Q^{-1} \cdot \{q_1, \dots, q_n\}), \dots, p_n(Q^{-1} \cdot \{q_1, \dots, q_n\})]}{\partial [q_1, q_2, \dots, q_n]} \\ [\text{für } ^1) q_1 = p_1(Q) \dots q_n = p_n(Q)].$$

In das Gebiet U_0 fallen nach dem Ersetzen von R durch $Q^{-1}R$ diejenigen $g(Q)\Delta_1\Delta_2\dots\Delta_n$ Punkte, die vorher im Gebiet U waren. Ursprünglich enthielt U_0 aber $g(E)V_0$ Punkte, da die Dichte darin $g(E)$ war und seine Größe V_0 ist. Es muß also

$$g(Q) = g(E) \frac{\partial [p_1(Q^{-1} \cdot \{q_1, \dots, q_n\}), \dots, p_n(Q^{-1} \cdot \{q_1, \dots, q_n\})]}{\partial [q_1, \dots, q_n]} \quad (3) \\ [\text{für } ^1) q_1 = p_1(Q) \dots q_n = p_n(Q)]$$

sein. Gilt (3), so ändert sich bei dem Ersetzen von R durch $Q^{-1}R$ die Dichte der Summationspunkte in (2) in der Nähe der Einheit nicht. Umgekehrt nimmt sie

bei dem Ersetzen von R durch QR in der Nähe von $Q = \{q_1, \dots, q_n\}$ den Wert (3) an, wenn vorher die Dichte bei der Einheit $g(E)$ war.

Aus der Gültigkeit von (3) für alle Punkte Q haben wir bisher die Invarianz der Belegungsdichte allgemein nur für die Einheit $E = Q^{-1}Q$ bewiesen und wir müssen sie noch für den beliebigen Punkt $T = SQ$ zeigen. Zu diesem Zweck zerlegen wir das Ersetzen von R durch SR in

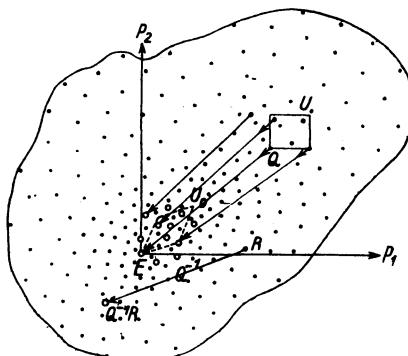


Abb. 4. Die Punkte sind die Summationspunkte R in (2), die Stellen, in die einige dieser Punkte durch vordere Multiplikation mit Q^{-1} übergeführt werden, sind durch kleine Kreise angegeben. Bei der Einheit ist die Dichte der Kreise und Punkte gleich

¹⁾ Die Werte $p_k(Q)$ für die q_k sind erst nach Ausführung der Differentiation einzusetzen: (3) ist eine Bildung der Art „ $\partial f(xy)/\partial x$ für $x=y^*$. Ist z. B. $f(xy)=x^3y^3$, so ist $\partial f(xy)/\partial x$ für $x=y$ gleich $2y^4$.

zwei Schritte: zuerst ersetzen wir R durch $Q^{-1}R$ und dann die so entstandenen Punkte R durch $SQR = TR^1$), dann haben wir im Ganzen R durch $SQ \cdot Q^{-1}R = SR$ ersetzt. War die Dichte der Summationspunkte anfangs bei Q durch (3) gegeben, so ist sie nach dem ersten Schritt bei der Einheit gleich $g(E)$. Ist dies aber

vor dem zweiten Schritt der Fall, so ist die Dichte nach dem zweiten Schritt bei $SQ = T$ ebenfalls durch (3) gegeben, womit die Invarianz der Belegungsdichte (3) bewiesen ist.

Weiterhin ist es klar, daß (2) für ein nirgends (d. h. für kein R) negatives $J(R)$ nur verschwinden kann, wenn dieses überall Null ist — eine Tatsache, die bei der Ableitung des ersten Satzes des vorangehenden Kapitels notwendig ist.

Für analytische Zwecke ist es natürlich bequemer, an Stelle der Summe (2) ein Integral zu benutzen. Anstatt aus dem Bereich U (Abb. 4) $g(Q) \Delta_1, \Delta_2, \dots, \Delta_n$ Summanden zu nehmen, kann man einen Summanden nehmen und diesen mit dem Gewicht $g(Q) \Delta_1, \Delta_2, \dots, \Delta_n$ multiplizieren. Da alle $J(R)$ stetig sein sollen, ändert dies den Wert der Summe (2) nicht, und dieses geht in das Hurwitzsche Integral

$$\int J(R) dR = \int \dots \int J(\{r_1, r_2, \dots, r_n\}) g(\{r_1, r_2, \dots, r_n\}) dr_1 dr_2 \dots dr_n \quad (4)$$

über. Die Integration muß in (4) über den ganzen Variabilitätsbereich der Parameter erstreckt werden. Wir wollen noch einmal direkt verifizieren, daß

$$\begin{aligned} \int J(R) dR &= \int \dots \int J(\{r_1, r_2, \dots, r_n\}) g(\{r_1, r_2, \dots, r_n\}) dr_1 \dots dr_n \\ &= \int J(SR) dR = \int \dots \int J(S \cdot \{r_1, r_2, \dots, r_n\}) g(\{r_1, r_2, \dots, r_n\}) dr_1 \dots dr_n \end{aligned} \quad (5)$$

für alle Elemente S gilt, wobei das $g(R)$ aus (3) zu entnehmen ist.

¹⁾ Die Möglichkeit dieser Zerlegung beruht wesentlich auf der Gültigkeit des Assoziativgesetzes.

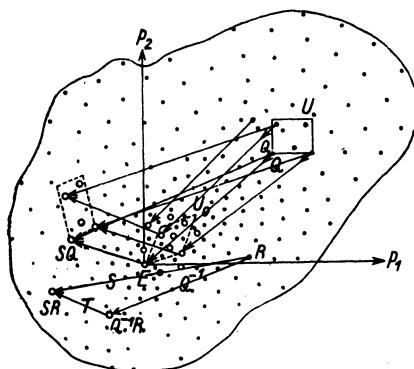


Abb. 5

Führen wir im bestimmten Integral rechts in (5) neue Variable, nämlich die Parameter x_1, x_2, \dots, x_n des Produkts

$X = \{x_1, x_2, \dots, x_n\} = SR = S \cdot \{r_1, r_2, \dots, r_n\}$
ein:

$$x_k = p_k(S \cdot \{r_1, r_2, \dots, r_n\}), \quad (6)$$

$$r_k = p_k(S^{-1} \cdot \{x_1, x_2, \dots, x_n\}), \quad (6a)$$

so ändert sich das Integrationsgebiet nicht, da SR mit R die ganze Gruppe durchstreicht. Wir erhalten

$$\begin{aligned} \int J(SR) dR &= \int \dots \int J(S \{r_1, \dots, r_n\}) g(\{r_1, \dots, r_n\}) dr_1 \dots dr_n \\ &= \int \dots \int J(\{x_1, \dots, x_n\}) g(R) \frac{\partial(r_1, \dots, r_n)}{\partial(x_1, \dots, x_n)} dx_1, \dots, dx_n, \end{aligned}$$

wo R und die r nach (6a) als Funktionen der x zu betrachten sind. Entnehmen wir noch $g(R)$ aus (3), so können wir das Produkt der beiden letzten Faktoren des Integranden nach dem Satz von der Funktionaldeterminante impliziter Funktionen umformen¹⁾

$$\left. \begin{aligned} g(E) \frac{\partial [\dots p_k(R^{-1} \{r_1, \dots, r_n\}) \dots]}{\partial [\dots r_k \dots]} \frac{\partial(r_1, \dots, r_n)}{\partial(x_1, \dots, x_n)} \\ = g(E) \frac{\partial [\dots p_k(R^{-1} \{r_1, \dots, r_n\}) \dots]}{\partial [\dots x_k \dots]} \end{aligned} \right\} (7)$$

[für $r_1 = p_1(R) \dots r_n = p_n(R)$],

wo bei der Differentiation nach x das R als konstant [wie in (3)], die r dagegen als die Funktionen (6a) der x anzusehen sind. Setzen wir daher $\{r_1, \dots, r_n\} = S^{-1} \{x_1, \dots, x_n\}$ in (7) ein, so geht dies wegen $R^{-1} S^{-1} = X^{-1}$ in

$$g(E) \frac{\partial [\dots p_k(X^{-1} \{x_1, \dots, x_n\}) \dots]}{\partial [\dots x_k \dots]} = g(\{x_1, \dots, x_n\})$$

[für²⁾ $x_1 = p_1(SR) = p_1(X) \dots x_n = p_n(SR) = p_n(X)$]

über, so daß die rechte Seite von (5) tatsächlich — bis auf die Benennung der Integrationsvariablen — mit der linken gleichlautend wird.

Wir können noch (3) etwas umformen, indem wir für die q , die wir für den Augenblick als frei veränderlich ansehen (wo-

¹⁾ Zunächst gilt (7) für beliebige Werte der r_k , also auch für die Werte $r_k = p_k(R)$, die wir für (3) brauchen.

²⁾ Die Werte $p_k(X)$ für die x_k sind auch hier nach Ausführung der Differentiation einzusetzen.

gegen Q als konstantes Gruppenelement betrachtet werde), neue Variable e mit Hilfe der Gleichung $Q^{-1}\{q_1, \dots, q_n\} = \{e_1, \dots, e_n\}$ einführen. Dann ist nach dem bereits benutzten Satze über die Funktionaldeterminante impliziter Funktionen

$$\begin{aligned} & \frac{\partial [\dots p_k(\{e_1 \dots e_n\}) \dots]}{\partial [\dots e_k \dots]} \\ &= \frac{\partial [\dots p_k(Q^{-1}\{q_1 \dots q_n\}) \dots]}{\partial [\dots q_k \dots]} \frac{\partial (\dots q_k \dots)}{\partial (\dots e_k \dots)}. \end{aligned}$$

Da aber $p_k(\{e_1 \dots e_n\}) = e_k$ ist, ist die linke Seite gleich 1 und wir erhalten, wenn wir die q_k mit Hilfe der e_k ausdrücken,

$$\left. \begin{aligned} & \frac{\partial [\dots p_k(Q^{-1}\{q_1 \dots q_n\}) \dots]}{\partial [\dots q_k \dots]} = \left[\frac{\partial (\dots q_k \dots)}{\partial (\dots e_k \dots)} \right]^{-1} \\ &= \left[\frac{\partial [\dots p_k(Q\{e_1 \dots e_n\}) \dots]}{\partial [\dots e_k \dots]} \right]^{-1}. \end{aligned} \right\} \quad (8)$$

Setzen wir jetzt links $q_k = p_k(Q)$, so müssen wir rechts $e_k = p_k(E)$ setzen und wir können an Stelle von (3)

$$g(Q) = g(E) \left. \begin{aligned} & \left[\frac{\partial [p_1(Q\{e_1 \dots e_n\}), \dots, p_n(Q\{e_1, \dots, e_n\})]}{\partial [e_1, \dots, e_n]} \right]^{-1} \\ & \text{[für }^1) e_1 = p_1(E) \dots e_n = p_n(E)] \end{aligned} \right\} \quad (3a)$$

schreiben.

6. Bei gemischtkontinuierlichen Gruppen kann man das Hurwitzsche Integral noch etwas umformen. Bezeichnen wir das mit der Einheit zusammenhängende Gebiet mit \mathfrak{G}_1 , die anderen zusammenhängenden Gebiete mit $\mathfrak{G}_2, \mathfrak{G}_3, \dots, \mathfrak{G}_\varrho$. Die Elemente von \mathfrak{G}_1 bilden, wie wir in 3. gesehen haben, eine Untergruppe (sogar einen Normalteiler), die der \mathfrak{G}_i die zugehörigen Nebengruppen. Zeichnet man aus jedem der Gebiete \mathfrak{G}_i ein beliebiges Element, etwa A_i aus, so kann man durch Multiplikation mit A_i das \mathfrak{G}_1 auf \mathfrak{G}_i abbilden. Da die Gewichte von Gebieten, die man aufeinander abbilden kann, gleich sind, kann man erwarten, daß das über \mathfrak{G}_1 erstreckte Integral

$$\left. \begin{aligned} & \int_{\mathfrak{G}_1} [J(R) + J(A_2 R) + \dots + J(A_\varrho R)] dR \\ &= \int_{\mathfrak{G}_1} [J(SR) + J(SA_2 R) + \dots + J(SA_\varrho R)] dR \end{aligned} \right\} \quad (9)$$

¹⁾ Ist auch erst nach der Differentiation einzusetzen.

gegenüber linksseitiger Multiplikation mit einem beliebigen S invariant, also das Hurwitzsche Integral für die ganze Gruppe ist, wenn mit $\int_{\mathfrak{G}_1} \dots dR$ das Hurwitzsche Integral für die einfach-kontinuierliche Gruppe \mathfrak{G}_1 bezeichnet ist.

Durchläuft R die Untergruppe \mathfrak{G}_1 , so durchläuft A, R die Elemente der Nebengruppe \mathfrak{G}_μ , links in (9) steht für jede Nebengruppe von \mathfrak{G}_1 ein Glied. Aber auch rechts in (9) steht für jede Nebengruppe ein Integral: für $\mathfrak{G}_\mu = A$, \mathfrak{G}_1 z. B. das Glied mit $SA_\mu R$, wenn $\mathfrak{G}_\mu = A_\mu \mathfrak{G}_1$ diejenige Nebengruppe ist, die $S^{-1}A$, enthält. Wenn wir also hierfür

$$\int_{\mathfrak{G}_1} J(A, R) dR = \int_{\mathfrak{G}_1} J(SA_\mu R) dR \quad (10)$$

zeigen könnten, so wäre damit gezeigt, daß die Glieder der beiden Seiten von (9) paarweise gleich sind.

Die Nebengruppe $A_\mu \mathfrak{G}_1$ soll $S^{-1}A$, enthalten, es sei etwa $A_\mu T = S^{-1}A$, $A_\mu = S^{-1}A, T^{-1}$, wo T in der Untergruppe \mathfrak{G}_1 enthalten ist. Führen wir diesen Ausdruck für A_μ in (10) ein, so geht es in

$$\int_{\mathfrak{G}_1} J(A, R) dR = \int_{\mathfrak{G}_1} J(S \cdot S^{-1}A, T^{-1}R) dR = \int_{\mathfrak{G}_1} J(A, T^{-1}R) dR$$

über. Diese Gleichung ist aber sicher richtig, weil sich die rechte Seite nur dadurch von der linken unterscheidet, daß in ihr $T^{-1}R$ an Stelle von R steht. Dieser Substitution gegenüber ist jedoch das Integral $\int_{\mathfrak{G}_1} \dots dR$ voraussetzungsgemäß invariant, da ja auch T^{-1} ein Element von \mathfrak{G}_1 ist.

Damit ist die Äquivalenz des Ausdrückes (9) mit dem Hurwitzschen Integral für die ganze gemischtkontinuierliche Gruppe nachgewiesen und dieses zugleich auf ein Integral, das nur über den mit der Einheit zusammenhängenden Teil der Gruppe zu erstrecken ist, zurückgeführt. (Diese ganze Betrachtung gilt natürlich nicht nur für \mathfrak{G}_1 , sondern für jede Untergruppe mit endlichem Index.)

7. Aus (5) folgt, genau wie bei endlichen Gruppen, daß jede Darstellung zu einer unitären gemacht werden kann, wenigstens dann, wenn die Integrale

$$\int \mathbf{D}(R)_{\kappa\lambda} \mathbf{D}(R)_{\mu\lambda}^* dR$$

konvergieren. Dies ist immer der Fall, wenn — wie etwa bei der Drehgruppe — das Grundgebiet für die Integration endlich ist. Die Orthogonalitätsrelationen der Darstellungskoeffizienten nehmen die Form

$$\int \mathbf{D}^{(\nu)}(R)_{\kappa\lambda}^* \mathbf{D}^{(\nu')}(\mathbf{R})_{\kappa'\lambda'} dR = \frac{\delta_{\nu\nu'} \delta_{\kappa\kappa'} \delta_{\lambda\lambda'}}{l_\nu} \int dR \quad (11)$$

[l_ν die Dimension der Darstellung $\mathbf{D}^{(\nu)}(R)$] an. Entsprechend lautet die Gleichung für die Charaktere

$$\int \chi^{(\nu)}(R)^* \chi^{(\nu')}(R) dR = \delta_{\nu\nu'} \int dR. \quad (12)$$

XI. Darstellungen und Eigenfunktionen

1. Betrachten wir die Schrödingergleichung $\mathbf{H}\psi = E\psi$ eines Systems, das aus zwei gleichen Teilchen aufgebaut ist. Der Einfachheit halber nehmen wir an, daß beide Teilchen nur je einen Freiheitsgrad haben, die entsprechenden Koordinaten seien x und y

$$\mathbf{H}\psi = -\frac{\hbar^2}{8\pi^2 m} \left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} \right) \psi(x, y) + V(x, y) \psi(x, y) = E\psi(x, y), \quad (1)$$

wo m die Masse der Teilchen ist. Die potentielle Energie muß gleich groß sein, wenn das erste Teilchen die Koordinate a und das zweite die Koordinate b , oder das erste die Koordinate b und das zweite die Koordinate a hat. Es muß daher für alle Werte von a und b

$$V(a, b) = V(b, a) \quad (2)$$

gelten. Nehmen wir weiter an, daß (1) ein diskretes Spektrum hat und $\psi_x(x, y)$ zum diskreten Eigenwert E_x gehört. Zu diesem Eigenwert gehöre keine andere, linear von $\psi_x(x, y)$ unabhängige Eigenfunktion: die allgemeine Lösung der Differentialgleichung für ψ_x

$$\mathbf{H}\psi_x = E_x \psi_x, \quad (3)$$

die für $x = \pm\infty$ und $y = \pm\infty$ verschwindet, sei ein konstantes Vielfaches von $\psi_x(x, y)$.

Betrachten wir die Funktion $\mathbf{P}\psi_x = \bar{\psi}_x$, die so definiert ist, daß für alle Werte von a und b

$$\mathbf{P}\psi_x(a, b) = \bar{\psi}_x(a, b) = \psi_x(b, a) \quad (4)$$

gilt. Wir wollen zeigen, daß $\bar{\psi}_x(x, y)$ die Differentialgleichung (3) auch befriedigt. Bezeichnen wir für den Augenblick die Ableitung

einer Funktion $f(x, y)$ nach der Variablen, die an erster Stelle steht, mit $f^{(1)}(x, y)$, nach der Variablen, die an zweiter Stelle steht, mit $f^{(2)}(x, y)$, also etwa

$$\left. \begin{aligned} \frac{\partial f(x, y)}{\partial x} &= f^{(1)}(x, y); & \frac{\partial f(x, y)}{\partial y} &= f^{(2)}(x, y), \\ \frac{\partial^2 f(x, y)}{\partial x^2} &= \frac{\partial f^{(1)}(x, y)}{\partial x} = f^{(1)(1)}(x, y) \end{aligned} \right\} \quad (5)$$

usw., oder auch

$$\frac{\partial f(y, x)}{\partial x} = f^{(2)}(y, x); \quad \frac{\partial f(y, x)}{\partial y} = f^{(1)}(y, x),$$

so ergibt sich aus (4) durch Differentiation nach a bzw. b

$$\left. \begin{aligned} \bar{\psi}_x^{(1)}(a, b) &= \psi_x^{(2)}(b, a); & \bar{\psi}_x^{(1)(1)}(a, b) &= \psi_x^{(2)(2)}(b, a), \\ \bar{\psi}_x^{(2)}(a, b) &= \psi_x^{(1)}(b, a); & \bar{\psi}_x^{(2)(2)}(a, b) &= \psi_x^{(1)(1)}(b, a). \end{aligned} \right\} \quad (6)$$

Berechnen wir jetzt

$$\left. \begin{aligned} \mathbf{H} \bar{\psi}_x(x, y) &= -\frac{\hbar^2}{8\pi^2 m} \left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} \right) \bar{\psi}_x(x, y) + V(x, y) \bar{\psi}_x(x, y) \\ &= -\frac{\hbar^2}{8\pi^2 m} (\bar{\psi}_x^{(1)(1)}(x, y) + \bar{\psi}_x^{(2)(2)}(x, y)) + V(x, y) \bar{\psi}_x(x, y), \end{aligned} \right\} \quad (7)$$

so ergibt sich dafür wegen (6), (4), (2) und (3) die Gleichung

$$\left. \begin{aligned} \mathbf{H} \bar{\psi}_x(x, y) &= -\frac{\hbar^2}{8\pi^2 m} (\psi_x^{(2)(2)}(y, x) + \psi_x^{(1)(1)}(y, x)) + V(y, x) \psi_x(y, x) \\ &= E_x \psi_x(y, x) = E_x \bar{\psi}_x(x, y), \end{aligned} \right\} \quad (8)$$

d. h., daß $\mathbf{P} \psi_x(x, y) = \bar{\psi}_x(x, y)$ ebenfalls eine Lösung der Differentialgleichung (3) ist. Sonach muß es ein konstantes Vielfaches von $\psi_x(x, y)$

$$\bar{\psi}_x(x, y) = c \psi_x(x, y) \quad (9)$$

sein, da es auch die Randbedingungen befriedigt. Zur Bestimmung von c beachten wir, daß wegen (4) für alle Wertepaare von a und b

$$c \psi_x(a, b) = \bar{\psi}_x(a, b) = \psi_x(b, a) \quad (10)$$

gelten muß. Wenden wir (10) erstens auf das Wertepaar y, x und dann auf das Wertepaar x, y an

$$c \psi_x(y, x) = \psi_x(x, y); \quad c \psi_x(x, y) = \psi_x(y, x). \quad (11)$$

Daraus ergibt sich

$$c^2 \psi_x(x, y) = \psi_x(x, y)$$

und, da $\psi_x(x, y)$ nicht identisch verschwinden soll, $c^2 = 1$; $c = \pm 1$. Wir haben dann ebenfalls identisch in x und y

$$\bar{\psi}_x(y, x) = \psi_x(x, y) = \pm \psi_x(y, x). \quad (12)$$

Die Eigenfunktion $\psi_x(x, y)$ ist an der Stelle x, y entweder gleich groß, wie an der Stelle y, x , oder entgegengesetzt gleich groß. Ob der erste oder der zweite Fall eintritt, kann man durch allgemeine Überlegungen nicht entscheiden, für eine Funktion $\psi_x(x, y)$, d. h. für einen Eigenwert, der den gestellten Bedingungen genügt, kann nur der eine der beiden Fälle eintreten. Eigenwerte bzw. Eigenfunktionen, für die (12) mit dem oberen Vorzeichen gilt, nennt man bekanntlich symmetrische Eigenwerte bzw. Eigenfunktionen, für die (12) mit dem unteren Vorzeichen gilt, antisymmetrische. Es ergibt sich also ein qualitatives Unterscheidungsmerkmal für die Eigenwerte und Eigenfunktionen der Schrödingergleichung (1); je nachdem die Eigenfunktion die Gleichung (12) mit dem oberen oder mit dem unteren Vorzeichen befriedigt, gehören sie in die erste oder zweite Klasse. Eine ganz analoge, vielleicht noch einfachere Betrachtung läßt sich mit einer Eigenwertgleichung

$$-\frac{\hbar^2}{8\pi^2 m} \frac{\partial^2 \psi(x)}{\partial x^2} + V(x) \psi(x) = E \psi(x) \quad (13)$$

anstellen, mit

$$V(x) = V(-x), \quad (14)$$

wenn man

$$\mathbf{P} \psi(x) = \bar{\psi}(x) = \psi(-x) \quad (15)$$

setzt. Diesmal ergibt sich an Stelle von (12)

$$\psi(x) = \pm \psi(-x), \quad (16)$$

d. h. die wohlbekannte Tatsache, daß alle Eigenfunktionen entweder gerade, oder ungerade Funktionen von x sind.

2. Wir wollen diese Überlegung, bei der wir uns auf einen einfachen Eigenwert beschränkt haben, noch in dem etwas allgemeineren Fall eines diskreten Eigenwerts mit mehreren, allerdings nur endlich vielen linear unabhängigen Eigenfunktionen ausführen¹⁾. Dabei wollen wir die Rechnungen nach Möglichkeit

¹⁾ Die Betonung liegt hier auf „endlich vielen“, nicht auf „diskreten“, die ganze Theorie läßt sich z. B. fast unverändert auf die „diskreten komplexen Eigenwerte“ anwenden, auf die die Gamowsche Theorie des Kernzerfallen geführt hat, obwohl sie in unserem Sinne natürlich nicht diskret sind, weil die Quadratintegrale der zugehörigen Eigenfunktionen divergieren.

durch begriffliche Überlegungen ersetzen: es ist ja klar, daß die spezielle Gestalt des Hamiltonschen Operators in (1) und (13) unwesentlich ist und es das erstmal nur auf die Äquivalenz der beiden Teilchen, das zweitemal nur auf die der beiden Richtungen X und $-X$ ankam. Wir werden Gleichungen erhalten, die (12) bzw. (16) analog sind und ebenso wie diese Alternativen enthalten, so daß für verschiedene Gruppen von Eigenfunktionen zum Teil verschiedene Beziehungen resultieren. Die Eigenfunktionen der verschiedenen Gruppen gehören zu Termen verschiedener Eigenschaften, deren Unterschiede die Grundlage der sogenannten „Termozoologie“ bilden.

Die Betrachtung, die uns zu (12) geführt hat, beruht darauf, daß (1) der Transformation

$$x' = y; \quad y' = x \quad (17)$$

gegenüber invariant ist. Dies soll bedeuten, daß die Funktion¹⁾ $\mathbf{P}\psi_x$, die durch die in a und b identische Gleichung

$$\mathbf{P}\psi_x(a, b) = \psi_x(b, a)$$

definiert ist, eine Lösung der Gleichung $\mathbf{H}\psi = E\psi$ für ψ ist, wenn dies für ψ_x gilt.

Sei allgemeiner \mathbf{R} eine reelle orthogonale Transformation

$$\left. \begin{aligned} x'_1 &= R_{11}x_1 + R_{12}x_2 + \cdots + R_{1n}x_n \\ x'_2 &= R_{21}x_1 + R_{22}x_2 + \cdots + R_{2n}x_n \\ &\vdots \\ x'_n &= R_{n1}x_1 + R_{n2}x_2 + \cdots + R_{nn}x_n \end{aligned} \right\} \quad (18a)$$

Wir verstehen unter $\mathbf{P}_R f$ diejenige Funktion, für die

$$\mathbf{P}_R f(x'_1, x'_2, \dots, x'_n) = f(x_1, x_2, \dots, x_n) \quad (19)$$

entweder identisch in x_1, x_2, \dots, x_n gilt, wobei für die x'_1, x'_2, \dots, x'_n die Ausdrücke aus (18a) eingesetzt gedacht werden müssen, oder identisch in x'_1, x'_2, \dots, x'_n gilt und für die x_i :

$$x_i = \sum_{j=1}^n R_{ji} x'_j \quad (18b)$$

eingesetzt gedacht werden muß: \mathbf{P}_R ist ein Operator, der gewissermaßen die x'_1, x'_2, \dots, x'_n durch die x_1, x_2, \dots, x_n ersetzt. Wir

¹⁾ $\mathbf{P}\psi$ ist das Zeichen für eine Funktion ebenso, wie etwa f oder g solche Zeichen zu sein pflegen, $\mathbf{P}\psi(x, y)$ ist der Wert dieser Funktion an der Stelle x, y . So bedeutet z. B. (19), daß $\mathbf{P}_R f$ an der Stelle x'_1, \dots, x'_n denselben Wert hat, den die Funktion f an der Stelle x_1, \dots, x_n aufnimmt.

werden die in (18), (19) definierte formale Symbolik mit dem \mathbf{P}_R dieser mehr anschaulichen Redeweise vorziehen, weil bei dieser erfahrungsgemäß sehr leicht Fehler dadurch entstehen, daß man etwa die x durch die x' , anstatt umgekehrt, ersetzt¹⁾.

Sind jetzt die beiden Punkte x_1, x_2, \dots, x_n und x'_1, x'_2, \dots, x'_n des Konfigurationsraumes, die durch eine bestimmte Transformation \mathbf{R} ineinander übergeführt werden, physikalisch gleichwertig (dies ist z. B. der Fall, wenn sie sich nur dadurch unterscheiden, daß zwei gleiche Teilchen ihre Plätze vertauscht haben), so sind auch die beiden Funktionen ψ und $\mathbf{P}_R\psi$ gleichwertig (in $\mathbf{P}_R\psi$ spielt nur das zweite Teilchen die Rolle, die in ψ das erste gespielt hat und umgekehrt). Ist ψ ein stationärer Zustand, so gilt dies auch für $\mathbf{P}_R\psi$ und beide haben dieselbe Energie: aus $\mathbf{H}\psi = E\psi$ folgt $\mathbf{H}\mathbf{P}_R\psi = E\mathbf{P}_R\psi$; \mathbf{H} ist dem Operator \mathbf{P}_R gegenüber invariant.

Die Transformationen \mathbf{R} , die einander gleichwertige Stellen ineinander überführen, bilden eine Gruppe, die „Gruppe der Schrödingergleichung“, da die Zusammensetzungen, wie auch die Reziproken solcher Transformationen diese Eigenschaft ebenfalls aufweisen. Die Einheit der Gruppe ist die identische Transformation, die jede Stelle in sich überführt. Die Gruppe selber ist die Symmetriegruppe des Konfigurationsraumes.

Entsprechendes gilt von den Operatoren \mathbf{P}_R . Man sieht ja leicht, daß $\mathbf{P}_S\mathbf{P}_R = \mathbf{P}_{SR}$ ist. Führt nämlich \mathbf{R} die x in x' über, so ist $\mathbf{P}_R f(x'_i) = f(x_i)$, und führt \mathbf{S} die x' in x'' über, so ist $\mathbf{P}_S g(x''_i) = g(x'_i)$, also für $g(x) = \mathbf{P}_R f(x)$

$$\mathbf{P}_S\mathbf{P}_R f(x''_i) = \mathbf{P}_R f(x'_i) = f(x).$$

Nun führt aber \mathbf{SR} die x direkt in die x'' über, so daß

$$\mathbf{P}_{SR} f(x''_i) = f(x_i)$$

die Definitionsgleichung von $\mathbf{P}_{SR}f$ ist. Es folgt, daß $\mathbf{P}_{SR}f(x) = \mathbf{P}_S\mathbf{P}_R f(x)$, und da f eine beliebige Funktion ist, gilt

$$\mathbf{P}_{SR} = \mathbf{P}_S\mathbf{P}_R. \quad (20)$$

Die Gruppe der \mathbf{P}_R ist holomorph zur Gruppe der \mathbf{R} .

3. Die Operatoren \mathbf{P} sind linear. Will man in einer Summe die x' durch die x ersetzen, so kann man es in den einzelnen

¹⁾ Nur wenn \mathbf{R} identisch mit seiner eigenen Reziproken, also $\mathbf{R}^2 = 1$ ist, braucht man hierauf nicht zu achten.

Summanden tun, und will man es in einem konstanten Vielfachen einer Funktion tun, so kann man das Ersetzen von x' durch x zuerst vornehmen und dann mit der Konstanten multiplizieren. In Formeln

$$\mathbf{P}(af + bg) = a\mathbf{P}f + b\mathbf{P}g. \quad (21)$$

Da der Operator \mathbf{P} lediglich einen Übergang zu einem neuen, ebenfalls orthogonalen Koordinatensystem im Konfigurationsraum bedeutet, muß für zwei beliebige Funktionen f, g das skalare Produkt $(f, g) = (\mathbf{P}f, \mathbf{P}g)$, d. h. \mathbf{P} unitär sein. Übrigens ist \mathbf{P} unter viel allgemeineren Voraussetzungen, als wir sie hier gemacht haben, ein linear-unitärer Operator.

In unserem Fall hat \mathbf{P} noch die Eigenschaft

$$\mathbf{P}fg = \mathbf{P}f \cdot \mathbf{P}g, \quad (22)$$

wie ebenfalls unmittelbar aus der Definition folgt. Diese Eigenschaft der \mathbf{P} ist nicht so allgemeiner Natur, wie ihr linear-unitärer Charakter.

4. Zumeist kann man die Gruppe einer Schrödinger-Gleichung auf allgemeine physikalische Gründe zurückführen.

Hat man ein System mit n Elektronen und bezeichnet die Koordinaten des k -ten Elektrons mit¹⁾ x_k, y_k, z_k , so ist die Schrödinger-Gleichung invariant gegenüber folgenden beiden Arten von Transformationen:

$$\left. \begin{array}{l} x_1 = x_{\alpha_1}; \quad y_1 = y_{\alpha_1}; \quad z_1 = z_{\alpha_1} \\ x_2 = x_{\alpha_2}; \quad y_2 = y_{\alpha_2}; \quad z_2 = z_{\alpha_2} \\ \dots \quad \dots \quad \dots \\ x_n = x_{\alpha_n}; \quad y_n = y_{\alpha_n}; \quad z_n = z_{\alpha_n} \end{array} \right\} \quad (A)$$

(„Vertauschung von Elektronen“), wo $\alpha_1, \alpha_2, \dots, \alpha_n$ eine beliebige Permutation der Zahlen 1, 2, ..., n ist. Es sind ja alle Elektronen gleichberechtigt. Dasselbe gilt natürlich für Protonen, He-Kerne usw.

$$\left. \begin{array}{l} x'_1 = \beta_{11}x_1 + \beta_{12}y_1 + \beta_{13}z_1; \quad y'_1 = \beta_{21}x_1 + \beta_{22}y_1 + \beta_{23}z_1; \\ x'_2 = \beta_{11}x_2 + \beta_{12}y_2 + \beta_{13}z_2; \quad y'_2 = \beta_{21}x_2 + \beta_{22}y_2 + \beta_{23}z_2; \\ \dots \quad \dots \quad \dots \\ x'_n = \beta_{11}x_n + \beta_{12}y_n + \beta_{13}z_n; \quad y'_n = \beta_{21}x_n + \beta_{22}y_n + \beta_{23}z_n; \\ z'_1 = \beta_{31}x_1 + \beta_{32}y_1 + \beta_{33}z_1 \\ z'_2 = \beta_{31}x_2 + \beta_{32}y_2 + \beta_{33}z_2 \\ \dots \quad \dots \quad \dots \\ z'_n = \beta_{31}x_n + \beta_{32}y_n + \beta_{33}z_n \end{array} \right\} \quad (B)$$

¹⁾ Wir haben also fortan $3n$ Variable $x_1, y_1, z_1, x_2, y_2, z_2, \dots, x_n, y_n, z_n$ an Stelle der n Variablen x_1, x_2, \dots, x_n .

(„Ausführung einer Drehung“), wo (β_{ik}) eine reelle orthogonale Matrix ist; (B) bedeutet ja nur einen Übergang zu einem anderen Achsenkreuz und — solange wenigstens keine äußeren elektrischen oder magnetischen Felder da sind — sind alle Raumrichtungen gleichberechtigt.

Weiter ist die Schrödinger-Gleichung invariant gegenüber Transformationen, die aus (A) und (B) zusammengesetzt sind. Die Transformationen (A) bilden eine zur Gruppe der Vertauschungen von n Dingen (symmetrische Gruppe), die von (B) der dreidimensionalen Drehgruppe holomorphe Gruppe.

Operatoren, die — wie der Hamiltonsche Operator im vorher betrachteten Fall — invariant sind gegenüber gewissen Transformationen, nennt man häufig diesen Transformationen gegenüber symmetrische Operatoren.

5. Die Feststellung, die uns zu (12) bzw. (16) geführt hat, war die, daß die Funktion $\mathbf{P}\psi$ — als auch zum Eigenwert von ψ gehörig — ein konstantes Vielfaches von ψ sein muß. Jetzt, wo wir einen Eigenwert, zu dem l linear unabhängige Eigenfunktionen $\psi_1, \psi_2, \dots, \psi_l$ gehören, vor uns haben, können wir nicht mehr so schließen. Wir können nur sagen, daß die $\mathbf{P}_R\psi_1, \mathbf{P}_R\psi_2, \dots, \mathbf{P}_R\psi_l$ sich alle als Linearkombinationen der $\psi_1, \psi_2, \dots, \psi_l$ schreiben lassen: jede Eigenfunktion zu diesem Eigenwert hat ja diese Eigenschaft. Die Koeffizienten bezeichnen wir mit $\mathbf{D}(R)_{\kappa\nu}$, so daß

$$\mathbf{P}_R\psi_\nu(x_1, y_1, z_1, \dots, x_n, y_n, z_n) = \sum_{\kappa=1}^l \mathbf{D}(R)_{\kappa\nu} \psi_\kappa(x_1, y_1, z_1, \dots, x_n, y_n, z_n) \quad (23)$$

wird¹⁾. Gehört S ebenfalls in die Gruppe der Schrödinger-Gleichung, so ist auch

$$\mathbf{P}_S\psi_\kappa = \sum_{\lambda=1}^l \mathbf{D}(S)_{\lambda\kappa} \psi_\lambda$$

und man erhält, wenn man in (23) die Variablen an beiden Seiten der Transformation S unterwirft, auf beide Seiten \mathbf{P}_S anwendet (\mathbf{P}_S ist linear, die $\mathbf{D}(R)_{\kappa\nu}$ Konstanten!),

$$\left. \begin{aligned} \mathbf{P}_S \mathbf{P}_R \psi_\nu &= \mathbf{P}_S \sum_{\kappa=1}^l \mathbf{D}(R)_{\kappa\nu} \psi_\kappa = \sum_{\kappa=1}^l \mathbf{D}(R)_{\kappa\nu} \mathbf{P}_S \psi_\kappa \\ &= \sum_{\kappa=1}^l \mathbf{D}(R)_{\kappa\nu} \sum_{\lambda=1}^l \mathbf{D}(S)_{\lambda\kappa} \psi_\lambda = \sum_{\lambda=1}^l \sum_{\kappa=1}^l \mathbf{D}(S)_{\lambda\kappa} \mathbf{D}(R)_{\kappa\nu} \psi_\lambda. \end{aligned} \right\} (24)$$

¹⁾ Die Indizes $\kappa\nu$ stehen hier in dieser Reihenfolge, damit [vgl. (25)] $\mathbf{D}(R)$ selber und nicht seine Transponierte $\mathbf{D}(R)'$ eine Darstellung sei.

Andererseits ist $\mathbf{P}_S \mathbf{P}_R \psi_r = \mathbf{P}_{SR} \psi_r$, also

$$\mathbf{P}_S \mathbf{P}_R \psi_r = \mathbf{P}_{SR} \psi_r = \sum_{\lambda=1}^l \mathbf{D}(SR)_{\lambda r} \psi_\lambda$$

und hieraus folgt durch Koeffizientenvergleich mit (24)

$$\mathbf{D}(SR)_{\lambda r} = \sum_{x=1}^l \mathbf{D}(S)_{\lambda x} \mathbf{D}(R)_{xr}. \quad (25)$$

Die aus den Koeffizienten von (23) gebildeten l -dimensionalen Matrizen $\mathbf{D}(R)$, die nach (23) die Eigenfunktionen ψ_r eines Eigenwerts in die transformierten Eigenfunktionen $\mathbf{P}_R \psi_r$ überführen, genügen den Gleichungen $\mathbf{D}(SR) = \mathbf{D}(S) \mathbf{D}(R)$, sie bilden eine Darstellung der Gruppe der Transformationen, denen gegenüber $\mathbf{H} \psi = E \psi$ invariant ist. Die Dimension der Darstellung ist gleich der Anzahl der linear unabhängigen, zu dem betreffenden Eigenwert gehörigen Eigenfunktionen $\psi_1, \psi_2, \dots, \psi_l$.

Welche Darstellung es ist, deren Koeffizienten in (23) vorkommen, lässt sich mit Hilfe von allgemeinen Überlegungen ebenso wenig entscheiden, wie die Frage, ob in (12) das obere oder das untere Vorzeichen gilt: (23) gibt ebenso eine Alternative (allerdings unter mehreren Möglichkeiten) wie (12). Für verschiedene Eigenwerte werden auch hier im allgemeinen verschiedene Fälle der Alternative zutreffen.

Kombinieren wir noch (23) mit der Definitionsgleichung von \mathbf{P}_R

$$\mathbf{P}_R \psi_r(x'_1, y'_1, z'_1, \dots, x'_n, y'_n, z'_n) = \psi_r(x_1, y_1, z_1, \dots, x_n, y_n, z_n). \quad (18c)$$

Indem wir auf beide Seiten $\mathbf{P}_R^{-1} = \mathbf{P}_{R^{-1}}$ anwenden, erhalten wir mit Hilfe von (23)

$$\left. \begin{aligned} \psi_r(x'_1, y'_1, z'_1, \dots, x'_n, y'_n, z'_n) &= \mathbf{P}_{R^{-1}} \psi_r(x_1, y_1, z_1, \dots, x_n, y_n, z_n) \\ &= \sum_{x=1}^l \mathbf{D}(R^{-1})_{xr} \psi_x(x_1, y_1, z_1, \dots, x_n, y_n, z_n), \end{aligned} \right\} \quad (26)$$

die Transformationsformel für die ψ_r . Sie verbindet die Werte der Eigenfunktionen an gleichwertigen Stellen.

Die Gruppe, der gegenüber (1) invariant ist, besteht aus der Einheit und der Vertauschung R von x und y . Der Eigenwert war voraussetzungsgemäß einfach, wir müssten also eine eindimensionale Darstellung der Spiegelungsgruppe erhalten. Da \mathbf{P}_E der Einheitsoperator ist, ist

$$\mathbf{P}_E \psi = \psi = 1 \cdot \psi.$$

Der Einheit der Gruppe entspricht die Matrix (1). Weiter bedeutet (12)

$$\mathbf{P}_R \psi = \bar{\psi} = \pm \psi = \pm 1 \cdot \psi. \quad (12a)$$

Für einige Eigenwerte gilt das obere Vorzeichen, bei diesen entspricht auch dem Element R der Vertauschung der beiden Teilchen die Matrix (1), bei anderen Eigenwerten gilt das untere Vorzeichen, bei diesen entspricht dem Gruppenelement R die Matrix (-1) . Wir erhalten also zwei ein-dimensionale Darstellungen der symmetrischen Gruppe von zwei Elementen: die eine ist durch die Zuordnung $\mathbf{D}(E) = (1)$, $\mathbf{D}(R) = (1)$, die andere durch die Zuordnung $\mathbf{D}(E) = (1)$, $\mathbf{D}(R) = (-1)$ gegeben.

6. Wenn man zu einer anderen Wahl der linear unabhängigen Eigenfunktionen übergeht und an Stelle der $\psi_1, \psi_2, \dots, \psi_l$ aus den $\psi'_1, \psi'_2, \dots, \psi'_l$

$$\psi'_\mu = \sum_{v=1}^l a_{v\mu} \psi_v \quad (27)$$

ausgeht, so erhält man in (23) auch eine andere Darstellung der Gruppe der Operatoren \mathbf{P} , und es fragt sich, wie sich diese beiden Darstellungen zueinander verhalten.

Es sei

$$\psi_x = \sum \beta_{\lambda x} \psi'_\lambda \quad (28)$$

β die zu α reziproke Matrix. Dann ist wegen der Linearität von \mathbf{P}_R

$$\begin{aligned} \mathbf{P}_R \psi'_\mu &= \sum_v a_{v\mu} \mathbf{P}_R \psi_v = \sum_v \sum_x a_{v\mu} \mathbf{D}(R)_{xv} \psi_x \\ &= \sum_v \sum_x \sum_\lambda a_{v\mu} \mathbf{D}(R)_{xv} \beta_{\lambda x} \psi'_\lambda = \sum_\lambda \left(\sum_{xv} a_{v\mu} \mathbf{D}(R)_{xv} a_{x\lambda} \right) \psi'_\lambda. \end{aligned} \quad (29)$$

Die Matrix $\bar{\mathbf{D}}(R)$, die die ψ' in die $\mathbf{P}_R \psi'$ überführt, entsteht aus $\mathbf{D}(R)$ durch Ähnlichkeitstransformation mit α

$$\bar{\mathbf{D}}(R) = \alpha^{-1} \mathbf{D}(R) \alpha. \quad (30)$$

Durch eine andere Wahl der linear unabhängigen Eigenfunktionen, die zu einem bestimmten Eigenwert gehören, erleidet die zugehörige Darstellung nur eine Ähnlichkeitstransformation: zu jedem Eigenwert gehört eine bis auf eine Ähnlichkeitstransformation eindeutig bestimmte Darstellung der Gruppe der Schrödingergleichung.

Wollen wir demnach solche Eigenfunktionen haben, die sich nicht nach $\mathbf{D}(R)$, sondern nach einer ihr äquivalenten Darstellung transformieren, so müssen wir durch diejenige Matrix neue Linearkombinationen der Eigenfunktionen bilden, durch welche die Darstellung in die gewünschte Form übergeht.

Die bis auf eine Ähnlichkeitstransformation eindeutig bestimmte Darstellung bildet das qualitative Merkmal, durch das sich die verschiedenen Termarten unterscheiden. Es wird zu Singulett-*S*-Termen eine andere Darstellung gehören, als etwa zu Triplet-S-Termen oder auch zu Singulett-*D*-Termen, während alle Darstellungen — die zu Singulett-*S*-Termen gehören — äquivalent, d. h. da sie wegen der willkürlichen Wählbarkeit der linear unabhängigen Eigenfunktionen nur bis auf eine Ähnlichkeitstransformation bestimmt sind, einander gleich sind. Diese Darstellungen werden praktisch immer irreduzibel sein, was die Wichtigkeit der irreduziblen Darstellungen erklärt.

7. Sind die *l*-Eigenfunktionen $\psi_1, \psi_2, \dots, \psi_l$ des Eigenwerts *E* — was wir immer annehmen wollen — paarweise orthogonal: $(\psi_x, \psi_y) = \delta_{xy}$, so hat auch die zugehörige Darstellung die unitäre Form. Aus dem unitären Charakter des Operators P_R folgt nämlich, daß auch die *l*-Funktionen $P_R \psi_1, P_R \psi_2, \dots, P_R \psi_l$ paarweise orthogonal sind:

$$(P_R \psi_x, P_R \psi_y) = (\psi_x, \psi_y) = \delta_{xy}, \quad (31)$$

oder mit Hilfe von (23):

$$\left. \begin{aligned} \delta_{xy} &= (P_R \psi_x, P_R \psi_y) = \left(\sum_{\lambda} D(R)_{\lambda x} \psi_{\lambda}, \sum_{\mu} D(R)_{\mu y} \psi_{\mu} \right) \\ &= \sum_{\lambda} \sum_{\mu} D(R)_{\lambda x}^* D(R)_{\mu y} (\psi_{\lambda}, \psi_{\mu}) = \sum_{\lambda} D(R)_{\lambda x}^* D(R)_{\lambda y}, \\ &\quad 1 = D(R)^{\dagger} D(R), \end{aligned} \right\} \quad (32)$$

d. h. $D(R)$ ist eine unitäre Matrix¹⁾. Demzufolge gelten für die $D(R)$, wenn sie schon irreduzible Darstellungen sind, die Orthogonalitätsrelationen des IX. Kapitels ohne weiteres.

Die Gleichung (26) läßt sich nunmehr auch

$$\left. \begin{aligned} \psi_y(x'_1, y'_1, z'_1, \dots, x'_n, y'_n, z'_n) \\ &= \sum_x D(R^{-1})_{x y} \psi_x(x_1, y_1, z_1, \dots, x_n, y_n, z_n) \\ &= \sum_x D(R)_{y x}^* \psi_x(x_1, y_1, z_1, \dots, x_n, y_n, z_n) \end{aligned} \right\} \quad (26a)$$

schreiben. Auch sehen wir, daß für uns nur Darstellungen mit Matrizen mit nichtverschwindender Determinante in Frage kommen.

¹⁾ Da die Eigenfunktionen immer paarweise orthogonal gewählt werden können, ist dies ein besonders einfacher Beweis für die Unitarisierbarkeit der Darstellungen.

XII. Algebra der Darstellungstheorie

Wir wollen noch einige algebraische Betrachtungen an die Resultate des vorangehenden Kapitels knüpfen. Zunächst seien einige rein mathematische Sätze zusammengestellt.

1. Es sei $\mathbf{D}^{(j)}(R)$ eine irreduzible unitäre Darstellung der Dimension l_j der Gruppe der Operatoren \mathbf{P}_R , und $f_1^{(j)}, f_2^{(j)}, \dots, f_{l_j}^{(j)}$ solche l_j Funktionen, für die

$$\mathbf{P}_R f_\mu^{(j)} = \sum_{\lambda=1}^{l_j} \mathbf{D}^{(j)}(R)_{\lambda\mu} f_\lambda^{(j)} \quad (\mu = 1, 2, \dots, l_j) \quad (1)$$

für alle \mathbf{P}_R gilt. Wir nennen eine Funktion $f_x^{(j)}$, zu der sich solche $l_j - 1$ Funktionen $f_1^{(j)}, f_2^{(j)}, \dots, f_{x-1}^{(j)}, f_{x+1}^{(j)}, \dots, f_{l_j}^{(j)}$, „Partner“ finden lassen, daß für ihre Gesamtheit (1) gilt, zur x -Zeile der irreduziblen Darstellung $\mathbf{D}^{(j)}(R)$ gehörig. Diese Aussage ist nur sinnvoll, wenn die Darstellung $\mathbf{D}^{(j)}(R)$ ganz (nicht nur bis auf eine Ähnlichkeitstransformation) vorgegeben ist.

Aus (1) folgt für $\mu = x$, wenn man mit $\mathbf{D}^{(j)}(R)_{\lambda'x}^*$ multipliziert und über die ganze Gruppe summiert (bzw. bei kontinuierlichen Gruppen integriert),

$$\left. \begin{aligned} \sum_R \mathbf{D}^{(j)}(R)_{\lambda'x}^* \mathbf{P}_R f_x^{(j)} &= \sum_R \sum_\lambda \mathbf{D}^{(j)}(R)_{\lambda'x}^* \mathbf{D}^{(j)}(R)_{\lambda x} f_\lambda^{(j)} \\ &= \sum_\lambda \frac{h}{l_j} \delta_{j,j'} \delta_{\lambda\lambda'} \delta_{xx'} f_\lambda^{(j)} = \frac{h}{l_j} \delta_{j,j'} \delta_{xx'} f_\lambda^{(j)}. \end{aligned} \right\} \quad (2)$$

Hieraus ersehen wir zunächst, daß

$$\sum_R \mathbf{D}^{(j)}(R)_{xx}^* \mathbf{P}_R f_x^{(j)} = \frac{h}{l_j} f_x^{(j)} \quad (3)$$

für jede Funktion gilt, die zur x -Zeile der irreduziblen Darstellung $\mathbf{D}^{(j)}(R)$ gehört. Umgekehrt kann man auch zu jedem $f_x^{(j)}$, für das (3) gilt, solche $f_1^{(j)}, f_2^{(j)}, \dots, f_{x-1}^{(j)}, f_{x+1}^{(j)}, \dots, f_{l_j}^{(j)}$ Partner zuordnen, daß für ihre Gesamtheit (1) gelten soll: (3) ist die notwendige und hinreichende Bedingung dafür, daß $f_x^{(j)}$ zur x -Zeile der irreduziblen Darstellung $\mathbf{D}^{(j)}(R)$ gehören.

Aus (2) ergibt sich nämlich zunächst, daß, wenn $f_x^{(j)}$ überhaupt Partner hat, diese durch

$$f_\lambda^{(j)} = \frac{l_j}{h} \sum_S \mathbf{D}^{(j)}(S)_{\lambda x}^* \mathbf{P}_S f_x^{(j)} \quad (3a)$$

gegeben sind, dies sei also die Definition von $f_1^{(j)}, \dots, f_{x-1}^{(j)}, f_{x+1}^{(j)}, \dots, f_{l_j}^{(j)}$. Außerdem soll aber (3a) voraussetzungsgemäß auch für $\kappa = \lambda$ gelten, es gilt also für alle $f_\lambda^{(j)}$. Wir wollen jetzt zeigen, daß für diese $f_\lambda^{(j)}$ auch (1) gilt. Zu diesem Zwecke setzen wir für die $f_\mu^{(j)}, f_\lambda^{(j)}$ in (1) die rechten Seiten von (3a) ein, die Gültigkeit von (1) ist dann gleichbedeutend mit der Gültigkeit von

$$\mathbf{P}_R \frac{l_j}{h} \sum_S \mathbf{D}^{(j)}(S)_{\mu\kappa}^* \mathbf{P}_S f_\kappa^{(j)} = \sum_\lambda \mathbf{D}^{(j)}(R)_{\lambda\mu} \frac{l_j}{h} \sum_S \mathbf{D}^{(j)}(S)_{\lambda\kappa}^* \mathbf{P}_S f_\kappa^{(j)}.$$

Dies ist aber eine Identität, wie man sofort einsieht, wenn man auf beide Seiten $\mathbf{P}_{R^{-1}}$ anwendet und $\mathbf{D}^{(j)}(R)_{\lambda\mu} = \mathbf{D}^{(j)}(R^{-1})_{\mu\lambda}^*$ einsetzt. Wegen der Darstellungseigenschaft der $\mathbf{D}^{(j)}$ ergibt sich dann

$$\begin{aligned} \sum_S \mathbf{D}^{(j)}(S)_{\mu\kappa}^* \mathbf{P}_S f_\kappa^{(j)} &= \sum_S \sum_\lambda \mathbf{P}_{R^{-1}} \mathbf{P}_S \mathbf{D}^{(j)}(R^{-1})_{\mu\lambda}^* \mathbf{D}^{(j)}(S)_{\lambda\kappa}^* f_\kappa^{(j)} \\ &= \sum_S \mathbf{P}_{R^{-1}S} \mathbf{D}^{(j)}(R^{-1}S)_{\mu\kappa}^* f_\kappa^{(j)}. \end{aligned}$$

Jetzt steht auf beiden Seiten dasselbe, weil man die Summation rechts über $R^{-1}S$ statt über S erstrecken darf: zu einem $f_\kappa^{(j)}$, das (3) befriedigt, konnten wir in der Tat in (3a) Partner definieren, so daß für ihre Gesamtheit (1) erfüllt ist.

2. Eine Linearkombination $a f_x^{(j)} + b g_x^{(j)}$ von Funktionen $f_x^{(j)}$ und $g_x^{(j)}$, die zur κ -Zeile der Darstellung $\mathbf{D}^{(j)}$ gehören, hat wieder diese Eigenschaft. Es folgt dies einfach aus der Linearität von (3) oder aus der Definition (1).

3. Sind $\mathbf{D}^{(1)}(R), \mathbf{D}^{(2)}(R), \dots, \mathbf{D}^{(c)}(R)$ die sämtlichen irreduziblen Darstellungen der Gruppe der Operatoren \mathbf{P}_R , so läßt sich jede Funktion F , auf die die \mathbf{P}_R angewendet werden können, so in eine Summe zerlegen

$$F = \sum_{j=1}^c \sum_{\kappa=1}^{l_j} f_\kappa^{(j)}, \quad (4)$$

dass $f_\kappa^{(j)}$ zur κ -Zeile der Darstellung $\mathbf{D}^{(j)}(R)$ gehöre.

Zum Beweise betrachten wir die h Funktionen $F = \mathbf{P}_E F, \mathbf{P}_{A_2} F, \mathbf{P}_{A_3} F, \dots, \mathbf{P}_{A_h} F$, die entstehen, wenn man auf F die h Operationen der Gruppe der \mathbf{P}_R anwendet. Sollten diese Funktionen voneinander nicht linear unabhängig sein, so lassen wir so viele weg, daß unter den übriggebliebenen $F, F_2, \dots, F_{h'}$ keine linearen Beziehungen mehr bestehen sollen. Diese h' Funktionen spannen eine Darstellung

der Gruppe der \mathbf{P}_R auf. Wenden wir nämlich auf diese Funktionen einen der Operatoren \mathbf{P}_R an, so lassen sich die entstehenden Funktionen durch die F, F_2, \dots, F_n linear ausdrücken. Sei z. B. $F_k = \mathbf{P}_T F$, so ist $\mathbf{P}_R \mathbf{P}_T F = \mathbf{P}_{RT} F$ also entweder schon selber eines der F_i oder doch durch diese linear ausdrückbar:

$$\mathbf{P}_R F_k = \sum_{i=1}^n \Delta(R)_{ik} F_i. \quad (5)$$

Die Matrizen $\Delta(R)$ bilden nun eine Darstellung der Gruppe der \mathbf{P}_R . Dies entspricht genau der Darstellungseigenschaft der Eigenfunktionen des vorangegangenen Kapitels, es ist

$$\begin{aligned} \sum_n \Delta(SR)_{nk} F_n &= \mathbf{P}_{SR} F_k = \mathbf{P}_S \mathbf{P}_R F_k = \mathbf{P}_S \sum_i \Delta(R)_{ik} F_i \\ &= \sum_i \sum_n \Delta(R)_{ik} \Delta(S)_{ni} F_n \end{aligned}$$

also wegen der linearen Unabhängigkeit

$$\Delta(S) \cdot \Delta(R) = \Delta(SR).$$

Diese Methode der Erzeugung von Darstellungen wird z. B. auch bei der expliziten Erzeugung irreduzibler Darstellungen der symmetrischen Gruppen eine wichtige Rolle spielen. Durch besondere Wahl der ursprünglichen Funktion F kann man nämlich sehr vielerlei Darstellungen erhalten, die zur Bestimmung von irreduziblen dienen werden.

Ist die Darstellung in (5) noch nicht irreduzibel, so kann man sie doch mit einer Ähnlichkeitstransformation ausreduzieren, d. h. die Matrizen $\Delta(R)$ alle gleichzeitig in die Form

$$\begin{pmatrix} \mathbf{D}^{(1)}(R) & \mathbf{0} & \dots \\ \mathbf{0} & \mathbf{D}^{(2)}(R) & \dots \\ \vdots & \vdots & \ddots \end{pmatrix} = \alpha^{-1} \Delta(R) \alpha \quad (*)$$

bringen, wo die \mathbf{D} alle unitären irreduzible Darstellungen sind. Nach 6, Kap. XI kann man aus den F_k mit Hilfe von α solche Linearkombinationen bilden, die sich bei der Anwendung der \mathbf{P}_R nach (*) transformieren, also zu den verschiedenen Zeilen der irreduziblen Darstellungen $\mathbf{D}^{(1)}, \mathbf{D}^{(2)}, \dots$ gehören. Da α keine singuläre Matrix ist, kann man umgekehrt mit Hilfe dieser Linearkombinationen die F_k , insbesondere auch F linear ausdrücken. Damit ist die Möglichkeit der Summendarstellung (4) der beliebigen Funktion F bewiesen.

Um in (4) noch die $f_x^{(j)}$ explizite zu berechnen, wenden wir darauf \mathbf{P}_R an, multiplizieren mit $\mathbf{D}^{(j)}(R)_{xx}^*$ und summieren über alle R . Wir erhalten

$$\sum_R \mathbf{D}^{(j)}(R)_{xx}^* \mathbf{P}_R F = \sum_{r'} \sum_{x'} \sum_R \mathbf{D}^{(j)}(R)_{xx}^* \mathbf{P}_R f_{x'}^{(j')} = \frac{h}{l_j} f_x^{(j)}, \quad (6)$$

das letzte wegen (2).

Diese Gleichung zeigt, daß $\sum_R \mathbf{D}^{(j)}(R)_{xx}^* \mathbf{P}_R F$ bei ganz beliebigem F zur x -Zeile der Darstellung $\mathbf{D}^{(j)}(R)$ gehört, eine Tatsache, die man mit Hilfe von (3) auch direkt verifizieren kann. Es ist ja, wenn man $SR = T$ setzt und anstatt über S über T summiert,

$$\begin{aligned} \sum_S \mathbf{D}^{(j)}(S)_{xx}^* \mathbf{P}_S \cdot \left(\sum_R \mathbf{D}^{(j)}(R)_{xx}^* \mathbf{P}_R F \right) &= \sum_{T, R} \mathbf{D}^{(j)}(TR^{-1})_{xx}^* \mathbf{P}_T \mathbf{D}^{(j)}(R)_{xx}^* F, \\ \sum_{T, R} \sum_{\lambda} \mathbf{D}^{(j)}(T)_{x\lambda}^* \mathbf{D}^{(j)}(R^{-1})_{\lambda x}^* \mathbf{D}^{(j)}(R)_{xx}^* \mathbf{P}_T F &= \frac{h}{l_j} \left(\sum_T \mathbf{D}^{(j)}(T)_{xx}^* \mathbf{P}_T F \right), \end{aligned}$$

da man die Summation über R nach (31a), Kap. IX ausführen kann.

Die Identität

$$F = \sum_j \sum_x \sum_R \frac{l_j}{h} \mathbf{D}^{(j)}(R)_{xx}^* \mathbf{P}_R F \quad (6a)$$

zeigt, da F ganz beliebig sein kann und mit den $\mathbf{P}_R F$ in keinerlei Zusammenhang stehen muß, daß

$$\sum_{j=1}^c \sum_{x=1}^{l_j} \frac{l_j}{h} \mathbf{D}^{(j)}(R)_{xx}^* \begin{cases} = 1 & \text{für } R = E \\ = 0 & \text{sonst} \end{cases}$$

ist. Für $R = E$ lautet dies, da $\mathbf{D}^{(j)}(E)_{xx} = 1$ ist,

$$\sum_{j=1}^c \sum_{x=1}^{l_j} \frac{l_j}{h} = \sum_{j=1}^c \frac{l_j^2}{h} = 1,$$

d. h. die Summe der Quadrate der Dimensionen aller irreduziblen Darstellungen ist gleich der Ordnung der dargestellten Gruppe.

4. Zwei Funktionen, die zu verschiedenen irreduziblen Darstellungen gehören, oder zwar zu derselben Darstellung, aber zu verschiedenen Zeilen, sind orthogonal. Zu $f_x^{(j)}$ wie auch zu $g_x^{(j')}$ kann man definitionsgemäß solche $f_1^{(j)}, f_2^{(j)}, f_3^{(j)}, \dots$ bzw. $g_1^{(j')}, g_2^{(j')}, g_3^{(j')}, \dots$ finden, daß

$$\mathbf{P}_R f_x^{(j)} = \sum_{\lambda} \mathbf{D}^{(j)}(R)_{\lambda x} f_{\lambda}^{(j)}, \quad \mathbf{P}_R g_x^{(j')} = \sum_{\lambda'} \mathbf{D}^{(j')}(R)_{\lambda' x} g_{\lambda'}^{(j')}$$

sei. Da \mathbf{P}_R unitär ist, gilt

$$\begin{aligned} (f_{\kappa}^{(j)}, g_{\kappa'}^{(j')}) &= (\mathbf{P}_R f_{\kappa}^{(j)}, \mathbf{P}_R g_{\kappa'}^{(j')}) \\ &= \sum_{\lambda} \sum_{\lambda'} \mathbf{D}^{(j)}(R)_{\lambda \kappa}^* \mathbf{D}^{(j')}(R)_{\lambda' \kappa'} (f_{\lambda}^{(j)}, g_{\lambda'}^{(j')}). \end{aligned} \quad (7)$$

Summieren wir diese Gleichung über alle Operationen \mathbf{P}_R der Gruppe, so ergibt sich

$$h(f_{\kappa}^{(j)}, g_{\kappa'}^{(j')}) = \frac{h}{l_j} \delta_{jj'} \delta_{\kappa\kappa'} \sum_{\lambda} (f_{\lambda}^{(j)}, g_{\lambda}^{(j')}). \quad (8)$$

Daraus folgt erstens der vorher ausgesprochene fundamentale Satz, daß $(f_{\kappa}^{(j)}, g_{\kappa'}^{(j')})$ für $j \neq j'$ oder $\kappa \neq \kappa'$ verschwindet, und zweitens, daß $(f_{\kappa}^{(j)}, g_{\kappa}^{(j)})$ unabhängig von κ für alle Partner gleich ist.

5. Schon im vorangehenden Kapitel haben wir von den \mathbf{P}_R gegenüber symmetrischen Operatoren gesprochen. So war der Hamiltonsche Operator \mathbf{H} den Operationen (A) und (B) gegenüber symmetrisch. Dies bedeutet so viel, daß \mathbf{P}_R in der Funktion, auf die es angewendet wird, nur solche Veränderungen hervorruft, die vom Standpunkt des \mathbf{H} irrelevant sind (sie vertauscht etwa die Rollen zweier Elektronen).

Wir wollen hier diesen Begriff etwas genauer präzisieren. Die Operatoren, die wir als symmetrische bezeichnen, sind immer hermiteische, sie sind physikalischen Größen (z. B. Energie) zugeordnet. Die \mathbf{P}_R , denen gegenüber ein Operator symmetrisch sein kann, sind zwar selber auch Operatoren, aber unitäre, die nicht physikalischen Größen zugeordnet sind, sondern eine Wellenfunktion, einen Zustand in einen anderen überführen. Man nennt \mathbf{S} symmetrisch, wenn sich ihm gegenüber alle $\mathbf{P}_R \varphi$ ebenso wie φ verhalten. (Wir werden sogleich sehen, daß sich dieser Begriff mit dem des vorangehenden Kapitels deckt.)

Ist \mathbf{S} ein den \mathbf{P}_R gegenüber symmetrischer Operator und ψ eine seiner Eigenfunktionen $\mathbf{S} \psi = s \psi$, so bedeutet dies, daß im Zustand ψ die Messung der Größe, zu der \mathbf{S} zugeordnet ist, mit Sicherheit den Wert s ergibt. Dies muß dann auch für $\mathbf{P}_R \psi$ gelten, d. h. diese muß auch eine zum Eigenwert s gehörige Eigenfunktion von \mathbf{S} sein.

Aus $\mathbf{S} \psi = s \psi$ folgt durch beiderseitige Anwendung von \mathbf{P}_R , daß $\mathbf{P}_R \mathbf{S} \psi = \mathbf{P}_R s \psi = s \mathbf{P}_R \psi$. Da auch $\mathbf{S} \mathbf{P}_R \psi = s \mathbf{P}_R \psi$ gilt, ist auch $\mathbf{P}_R \mathbf{S} \psi = \mathbf{S} \mathbf{P}_R \psi$, und diese Gleichung gilt für jede

Eigenfunktion von \mathbf{S} , da der Eigenwert darin nicht mehr enthalten ist. Da sie linear ist, gilt sie für jede Linearkombination der Eigenfunktionen, also für jede Funktion überhaupt. Daraus folgt aber $\mathbf{P}_R \mathbf{S} = \mathbf{S} \mathbf{P}_R$: ein in bezug auf die \mathbf{P}_R symmetrischer Operator ist mit allen \mathbf{P}_R vertauschbar, es ist gleichgültig, ob man auf eine Funktion zuerst das \mathbf{P}_R und dann das \mathbf{S} oder umgekehrt anwendet („ \mathbf{S} ist dem \mathbf{P}_R gegenüber invariant“).

Wendet man auf (1) \mathbf{S} an, so sieht man, daß auch $\mathbf{S} f_x^{(j)}$ zur x -Zeile von $\mathbf{D}^{(j)}$ gehört, wenn dies für $f_x^{(j)}$ gilt. Daraus folgt nach (8), daß

$$(f_x^{(j)}, \mathbf{S} g_{x'}^{(j')}) = \delta_{jj'} \delta_{xx'} (f_x^{(j)}, \mathbf{S} g_x^{(j)}) \quad (8a)$$

für $j \neq j'$ oder $x \neq x'$ verschwindet und für $j = j'; x = x'$ von x unabhängig ist.

Obwohl diese Gesetze recht allgemeiner Natur sind, sind sie doch nur für die einfachsten Gruppen von Operatoren allgemein bekannt. Eine Gruppe, für die diese Gesetze geläufig sind, besteht aus dem Einheitsoperator \mathbf{P}_E und dem von (15) des vorangehenden Kapitels

$$\mathbf{P}_R f(x) = f(-x).$$

Es ist $\mathbf{P}_R^2 = \mathbf{P}_E$, die Gruppe \mathbf{P}_E , \mathbf{P}_R ist die Spiegelungsgruppe. Sie hat zwei irreduzible Darstellungen, beide sind eindimensional:

$$\mathbf{D}^{(1)}(E) = (1), \mathbf{D}^{(1)}(R) = (1) \text{ und } \mathbf{D}^{(2)}(E) = (1), \mathbf{D}^{(2)}(R) = (-1).$$

Für Funktionen, die zur ersten Darstellung (sie hat nur eine Zeile) gehören, lautet (1)

$$\mathbf{P}_R f^{(1)}(x) = f^{(1)}(-x) = 1 \cdot f^{(1)}(x),$$

sie sind die geraden Funktionen von x . Für Funktionen, die zur zweiten Darstellung gehören, ist (1)

$$\mathbf{P}_R f^{(2)}(x) = f^{(2)}(-x) = -1 f^{(2)}(x),$$

das sind die ungeraden Funktionen. Die Gleichung (3) lautet für $f^{(1)}(x)$

$$\mathbf{D}^{(1)}(E) \mathbf{P}_E f^{(1)}(x) + \mathbf{D}^{(1)}(R) \mathbf{P}_R f^{(1)}(x) = 1 f^{(1)}(x) + 1 f^{(1)}(-x) = \frac{2}{1} f^{(1)}(x)$$

und für $f^{(2)}(x)$

$$\mathbf{D}^{(2)}(E) \mathbf{P}_E f^{(2)}(x) + \mathbf{D}^{(2)}(R) \mathbf{P}_R f^{(2)}(x) = 1 f^{(2)}(x) - 1 f^{(2)}(-x) = \frac{2}{1} f^{(2)}(x),$$

und umgekehrt folgt aus diesen Gleichungen, daß $f^{(1)}(x)$ gerade und $f^{(2)}(x)$ ungerade ist. Daß sich jede Funktion in einen geraden und einen ungeraden Teil zerlegen läßt, ist eine bekannte Tatsache, ebenso wie daß jede gerade Funktion auf jede ungerade orthogonal steht.

6. Bisher haben wir die Darstellungen irgendwie willkürlich bestimmt annehmen müssen und eine Funktion, die zur κ -Zeile einer irreduziblen Darstellung gehört, gehört nicht mehr zur κ -Zeile einer äquivalenten Darstellung.

Unabhängig von dieser speziellen Festlegung der Darstellung sind die nun folgenden Sätze.

Nach (2) gilt für jede Funktion $f_{\kappa}^{(j)}$, die zur κ -Zeile der irreduziblen Darstellung $D^{(j)}(R)$ gehört,

$$\sum_R D^{(j)}(R)_{\lambda\lambda}^* \mathbf{P}_R f_{\kappa}^{(j)} = \frac{h}{l_j} \delta_{\kappa\lambda} f_{\kappa}^{(j)}, \quad (2a)$$

was über λ von 1 bis l_j summiert

$$\sum_R \chi^{(j)}(R)^* \mathbf{P}_R f_{\kappa}^{(j)} = \frac{h}{l_j} f_{\kappa}^{(j)} \quad (9)$$

(für $\kappa = 1, 2, \dots, l_j$)

ergibt. Da in (9) κ nicht mehr vorkommt, genügen ihm alle Funktionen, die zu einer beliebigen Zeile der Darstellung $D^{(j)}(R)$ gehören, und auch beliebige Linearkombinationen solcher Funktionen. Man sagt von einer Funktion, die (9) befriedigt, sie gehöre zur Darstellung $D^{(j)}(R)$. Diese Tatsache ist — wie der Charakter — unabhängig von der speziellen Form der Darstellung. Umgekehrt ist jede Funktion, die (9) genügt, eine Linearkombination von Funktionen, die je zu einer Zeile der Darstellung $D^{(j)}(R)$ gehören. Nach (9) ist nämlich

$$\frac{h}{l_j} f^{(j)} = \sum_R \chi^{(j)}(R)^* \mathbf{P}_R f^{(j)} = \sum_{\lambda} \sum_R D^{(j)}(R)_{\lambda\lambda}^* \mathbf{P}_R f^{(j)}. \quad (10)$$

Jede Funktion der Gestalt $\sum_R D^{(j)}(R)_{\lambda\lambda}^* \mathbf{P}_R F$ gehört aber nach 3 zur λ -Zeile der Darstellung $D^{(j)}(R)$.

Daraus folgt auch, daß Funktionen, die zu nichtäquivalenten irreduziblen Darstellungen gehören, orthogonal zueinander sind. Weiter läßt sich jede Funktion F in eine Summe

$$F = \sum_{j=1}^c f^{(j)} \quad (11)$$

zerlegen, wo $f^{(j)}$ zur Darstellung $D^{(j)}(R)$ gehört: man braucht dazu ja nur die Gleichung (4)

$$\left. \begin{aligned} F &= \sum_{j=1}^c f^{(j)} \\ f^{(j)} &= \sum_{x=1}^{l_j} f_x^{(j)} \end{aligned} \right\} \quad (4a)$$

umzuschreiben.

Die Funktionen, die zu einer bestimmten irreduziblen Darstellung gehören, haben also ganz analoge Eigenschaften, wie die zu einer Zeile einer irreduziblen Darstellung gehörenden: eine Linearkombination von Funktionen irgendeiner Art ist wieder eine Funktion dieser Art, man kann jede beliebige Funktion als eine Summe von Funktionen je einer Art schreiben, zwei Funktionen verschiedener Art sind jeweils orthogonal aufeinander, ein Operator S , der den P gegenüber invariant ist, führt eine Funktion einer Art in eine andere Funktion derselben Art über.

Die hier besprochenen allgemeinen Sätze über Funktionen kann man so zusammenfassen, daß man sagt, Funktionen verschiedener Art (sowohl im ersten wie im zweiten Sinne) gehören zu verschiedenen Eigenwerten eines hermiteischen Operators, der — wie alle P und ihre Funktionen — mit allen invarianten Operatoren S vertauschbar ist.

Der Operator O_{jx} , der F in

$$O_{jx} F = \sum_R D^{(j)}(R)_{xx}^* P_R F \quad (12)$$

bzw. im zweiten Fall in

$$O_j F = \sum_R \chi^{(j)}(R)^* P_R F \quad (12a)$$

überführt, hat die beiden Eigenwerte 0 und h/l_j . Zum letzteren Eigenwert gehören alle Funktionen, die zur x -Zeile der Darstellung $D^{(j)}(R)$ bzw. einfach zu dieser Darstellung gehören. Zum Eigenwert 0 dagegen gehören diejenigen Funktionen, die zu anderen Darstellungen gehören, bzw. bei (12) auch die, die zu anderen Zeilen (nicht zur x -Zeile) von $D^{(j)}(R)$ gehören.

Die soeben besprochenen Sätze sind dann nichts anderes als die Orthogonalitäts- und Vollständigkeitsrelationen der Eigenfunktionen der Operatoren (12), (12a). Das Besondere gegenüber den gewöhnlichen hermiteischen Operatoren besteht lediglich darin, daß (12), (12a) unendlich vielfach entartete Operatoren sind, da zu beiden Eigenwerten unendlich viele linear unabhängige Eigenfunktionen gehören.

7. Kehren wir jetzt wieder zu der Schrödingergleichung $\mathbf{H}\psi = E\psi$ zurück! Wir haben im vorigen Kapitel gesehen, daß zu jedem Eigenwert von \mathbf{H} eine bis auf eine Ähnlichkeitstransformation eindeutig bestimmte Darstellung der Gruppe der \mathbf{P}_R gehört. Andererseits wissen wir auch, daß wir über diese Ähnlichkeitstransformation ganz frei verfügen dürfen: dies bedeutet nur eine Festlegung der Linearkombinationen der Eigenfunktionen, die wir benutzen wollen.

Es bietet vielfach Vorteile, die Darstellungen der einzelnen Eigenwerte, sofern sie nicht schon irreduzibel sind, in ausreduzierter Form anzunehmen

$$\Delta(R) = \begin{pmatrix} \mathbf{D}^{(1)}(R) & \mathbf{0} & \dots & \mathbf{0} \\ \mathbf{0} & \mathbf{D}^{(2)}(R) & \dots & \mathbf{0} \\ \vdots & \vdots & & \vdots \\ \mathbf{0} & \mathbf{0} & \dots & \mathbf{D}^{(s)}(R) \end{pmatrix}. \quad (13)$$

Hierin sind die $\mathbf{D}^{(1)}(R), \mathbf{D}^{(2)}(R), \dots, \mathbf{D}^{(s)}(R)$ lauter irreduzible (nicht notwendig voneinander verschiedene) Darstellungen, die s irreduziblen Bestandteile von $\Delta(R)$, ihre Dimensionen seien der Reihe nach l_1, l_2, \dots, l_s . Die dieser Form der Darstellung des betrachteten Eigenwerts entsprechenden Linearkombinationen der Eigenfunktionen bezeichnen wir mit

$$\psi_1^{(1)}, \psi_2^{(1)}, \dots, \psi_{l_1}^{(1)}, \psi_1^{(2)}, \psi_2^{(2)}, \dots, \psi_{l_2}^{(2)}, \dots, \psi_1^{(s)}, \psi_2^{(s)}, \dots, \psi_{l_s}^{(s)}.$$

Schreiben wir jetzt für diesen Eigenwert die Gleichung (23) des vorangehenden Kapitels auf, so ergibt sich

$$\mathbf{P}_R \psi_x^{(j)} = \sum_v \mathbf{D}^{(j)}(R)_{vx} \psi_v^{(j)}. \quad (14)$$

[Wegen der Nullen in (13) können die $\mathbf{P}_R \psi_x^{(j)}$ schon als Linearkombinationen der $\psi_v^{(j)}$ mit demselben oberen Index ausgedrückt werden.] Die Eigenfunktion $\psi_x^{(j)}$ gehört zur x -Zeile der Darstellung $\mathbf{D}^{(j)}$ und ihre Partner sind $\psi_1^{(j)}, \psi_2^{(j)}, \dots, \psi_{l_j}^{(j)}$.

Die Gestalt der Transformationsformel (14) legt es nahe, den Eigenwert von (13) als zufällig zusammenfallende s -Eigenwerte anzusehen. Zum ersten Eigenwert gehören $\psi_1^{(1)}, \psi_2^{(1)}, \dots, \psi_{l_1}^{(1)}$, zum zweiten $\psi_1^{(2)}, \psi_2^{(2)}, \dots, \psi_{l_2}^{(2)}$ und zum letzten schließlich die Eigenfunktionen $\psi_1^{(s)}, \psi_2^{(s)}, \dots, \psi_{l_s}^{(s)}$. Zu jedem dieser Eigenwerte gehört eine irreduzible Darstellung. Führen wir diese Bezeichnungs-

weise konsequent für das ganze Eigenwertspektrum durch, so erreichen wir erstens, daß zu jedem Eigenwert eine irreduzible Darstellung zugeordnet ist, und zweitens, daß jede Eigenfunktion zu einer Zeile einer irreduziblen Darstellung gehört, ihre Partner sind die anderen zu diesem Eigenwert gehörigen Eigenfunktionen.

Im allgemeinen wird dieselbe Darstellung sehr vielen Eigenwerten gemeinsam sein. Wir können daher eine weitere Vereinheitlichung der Transformationsformeln erzielen, indem wir die Darstellungen für alle Terme, denen sie gemeinsam sind, in derselben Form annehmen.

Daß zu einem Eigenwert mehrere Eigenfunktionen, nämlich alle Partner voneinander gehören, wird als „normale Entartung“ bezeichnet. Wenn außerdem noch mehrere Eigenwerte zusammenfallen, wie das im Eigenwert von (13) der Fall war, so spricht man von einer zufälligen Entartung. Es wird immer angenommen, daß diese eine sehr seltene Erscheinung ist und für die wirkliche Schrödinger-Gleichung, von einzelnen Ausnahmefällen abgesehen, nicht vorkommt.

8. Wir wollen jetzt, um mit den soeben gewonnenen Begriffen vertrauter zu werden, mit ihrer Hilfe einige Betrachtungen über die Rayleigh-Schrödingersche Störungstheorie anstellen. Wir gehen zuerst von einem Eigenwert E eines „ungestörten“ Problems aus, der keine zufällige Entartung zeigt. Die zugehörige Darstellung der Gruppe der Schrödinger-Gleichung ist dann irreduzibel und die Eigenfunktionen $\psi_{E1}, \psi_{E2}, \dots, \psi_{El}$ gehören zu den verschiedenen Zeilen dieser irreduziblen Darstellung. Wir fügen zu dem ursprünglichen Hamiltonschen Operator H eine „symmetrische Störung“ λV zu, die so beschaffen ist, daß sie die Symmetriegruppe von H nicht stört, d. h. selber ein symmetrischer Operator im Sinne dieses Kapitels ist. Um die Säkulargleichung für die erste Näherung ΔE der Zusatzenergie aufzustellen, müssen wir die Matrixelemente $(\psi_{Ex}, V \psi_{Ex'})$ berechnen. Nach (8a) sind diese aber für $x \neq x'$ alle Null, für $x = x'$ alle einander gleich, etwa v_E , so daß die Säkulargleichung die Form

$$\begin{vmatrix} \lambda v_E - \Delta E & 0 & \dots & 0 \\ 0 & \lambda v_E - \Delta E & \dots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \dots & \lambda v_E - \Delta E \end{vmatrix} = 0$$

und die l -fache Wurzel λv_E hat. Der Eigenwert spaltet also in erster Näherung nicht auf. Aber auch in beliebig hoher Näherung kann er nicht aufspalten, da bei einer Aufspaltung in etwa zwei Eigenwerte E_1 und E_2 mit l_1 und l_2 Eigenfunktionen ($l_1 + l_2 = l$) die l_1 Eigenfunktionen von E_1 sich bei der Anwendung der P unter sich transformieren müßten und ihnen eine Darstellung der Dimension l_1 zugeordnet sein müßte. Diese Darstellung kann aber wegen $l_1 < l$ die ursprüngliche irreduzible Darstellung des ungestörten Eigenwerts nicht enthalten und so müßten die l_1 Eigenfunktionen von E_1 auf sämtliche Eigenfunktionen von E orthogonal sein und könnten aus ihnen oder ihren Linearkombinationen nicht stetig hervorgehen.

Ein Eigenwert mit einer irreduziblen Darstellung kann bei einer „symmetrischen Störung“ nicht aufspalten und behält seine Darstellung bei.

9. Betrachten wir jetzt einen Eigenwert, zu dem eine Darstellung $\Delta(R)$, die $D^{(1)}(R)$, $D^{(2)}(R)$, ... der Reihe nach a_1 , a_2 , ...-mal enthalten soll, gehört. Im Sinne der Ausführungen in 7. können wir auch sagen, daß a_1 Eigenwerte mit der Darstellung $D^{(1)}(R)$ mit a_2 Eigenwerten mit der Darstellung $D^{(2)}(R)$ usw. zufällig zusammenfallen. Lassen wir jetzt die symmetrische Störung λV einwirken, so wird sich hieran höchstens das ändern, daß diese Eigenwerte nicht mehr zusammenfallen werden. Es werden aber auch nach der Störung a_1 Eigenwerte mit der Darstellung $D^{(1)}(R)$, a_2 mit der Darstellung $D^{(2)}(R)$ usw. vorhanden sein, diese $a_1 + a_2 + \dots$ Eigenwerte werden nur im allgemeinen alle verschieden groß sein. Daß auch nach der Störung genau a_1 Eigenwerte mit der Darstellung $D^{(1)}(R)$ vorhanden sein müssen, ergibt sich daraus, daß sich die Anzahl $a_1 l_1$ der Eigenfunktionen, die zu der Darstellung $D^{(1)}(R)$ gehören, nicht ändern darf. Eine Änderung dieser Zahl würde ja bedeuten, daß Eigenfunktionen ihre Zugehörigkeit zu einer irreduziblen Darstellung geändert haben, was aber — wie wir soeben gesehen haben — nicht stetig erfolgen könnte.

In 8. haben wir einen Eigenwert mit einer irreduziblen Darstellung betrachtet. Obwohl zu ihm l Eigenfunktionen gehörten, konnte er bei einer symmetrischen Störung nicht aufspalten. Dies rechtfertigt den Namen „natürliche Entartung“ für das Zugehören der l linear unabhängigen Eigenfunktionen zu einem Eigenwert.

Vom soeben betrachteten Eigenwert mit einer nicht irreduziblen Darstellung sagten wir, daß er eigentlich aus a_1 Eigenwerten der Darstellung $\mathbf{D}^{(1)}(R)$, a_2 Eigenwerten der Darstellung $\mathbf{D}^{(2)}(R)$ usw. besteht. Das Zusammenfallen dieser $a_1 + a_2 + \dots$ Eigenwerte nannten wir eine zufällige Entartung, weil ihr Vorhandensein bei Abwesenheit der Störung quasi einem Zufall zuschreiben ist.

10. Daß von den Eigenfunktionen von $\mathbf{H} + \lambda \mathbf{V}$ angenommen werden kann, daß sie zu einer Zeile einer irreduziblen Darstellung gehören, gilt nicht nur für die exakten Eigenfunktionen, sondern auch von allen sukzessiven Näherungen des Störungsverfahrens. Zunächst gilt es zwar nur für die exakten Eigenfunktionen, für die gesamte Potenzreihe nach λ für diese. Da es aber für die gesamte Potenzreihe bei jedem beliebigen Wert von λ gilt, gilt es auch für alle ihre Glieder separat.

Insbesondere können die „richtigen Linearkombinationen“, die erste Näherung für die Eigenfunktionen eines bestimmten Eigenwerts E so angenommen werden, daß sie aus lauter solchen Eigenfunktionen von E zusammengesetzt sind, die alle zu derselben Zeile derselben irreduziblen Darstellung gehören. Enthält die Darstellung von E eine bestimmte irreduzible Darstellung $\mathbf{D}^{(j)}(R)$ nur einmal, so hat E nur eine einzige Eigenfunktion $\psi_x^{(j)}$, die zu einer, etwa der x -Zeile von $\mathbf{D}^{(j)}(R)$ gehört, und $\psi_x^{(j)}$ ist dann schon eine der „richtigen Linearkombinationen“. Der zugehörige Eigenwert ist

$$(\psi_x^{(j)}, (\mathbf{H} + \lambda \mathbf{V}) \psi_x^{(j)}).$$

Enthält die Darstellung von E die irreduzible Darstellung $\mathbf{D}^{(j)}(R)$ mehrmals, etwa a_j -mal, so hat E auch a_j Eigenfunktionen, $\psi_{x1}^{(j)}, \psi_{x2}^{(j)}, \dots, \psi_{xa_j}^{(j)}$, die zu derselben x -Zeile von $\mathbf{D}^{(j)}(R)$ gehören. Die richtigen Linearkombinationen sind dann Linearkombinationen dieser a_j Eigenfunktionen, die man allerdings ohne Rechnung nicht mehr bestimmen kann.

Auf alle Fälle bleibt es aber vorteilhaft, von vornherein solche Linearkombinationen $\psi_{x\varrho}^{(j)}$ der Eigenfunktionen von E zu benutzen, die zu einer Zeile einer irreduziblen Darstellung gehören. Es verschwindet nämlich wegen (8a)

$$(\psi_{x\varrho}^{(j)}, \mathbf{V} \psi_{x'\varrho'}^{(j)}) = V_{jx\varrho; j'x'\varrho'} = \delta_{jj'} \delta_{xx'} v_{\varrho\varrho'}^j$$

9*

für $j \neq j'$ oder $\kappa \neq \kappa'$, wodurch sich die Säkulargleichung von E

$$|\mathbf{V}_{j\kappa\varrho; j'\kappa'\varrho'} - \Delta E \mathbf{1}| = 0$$

wesentlich vereinfacht: sie zerfällt, wie es sich bei näherem Zusehen zeigt, in lauter kleine sogenannte irreduzible Säkulargleichungen, deren Dimensionen die Zahlen a_j sind, die angeben, wie oft dieselbe irreduzible Darstellung in der Darstellung des Eigenwerts E enthalten ist.

Auch in höheren Näherungen ist es bei einer symmetrischen Störung immer möglich, die Änderung der Eigenwerte und Eigenfunktionen mit der Darstellung $\mathbf{D}^{(j)}(R)$ mit Hilfe der Eigenwerte und Eigenfunktionen dieser selben Darstellung allein zu berechnen. Es genügt sogar, die Eigenfunktionen, die zu einer bestimmten, etwa der ersten Zeile dieser Darstellung gehören, allein zu betrachten. Nach (22), Kap. V, ist z. B. die zweite Näherung

$$F_{k\nu} = E_k + \lambda(\psi_{k\nu}, \mathbf{V}\psi_{k\nu}) + \lambda^2 \sum_{E_l \neq E_k} \frac{|(\psi_l, \mathbf{V}\psi_{k\nu})|^2}{E_k - E_l}.$$

Wenn nun ψ_l zu einer von $\mathbf{D}^{(j)}$ verschiedenen Darstellung oder zu einer anderen Zeile von $\mathbf{D}^{(j)}$ als $\psi_{k\nu}$ gehört, verschwindet $(\psi_l, \mathbf{V}\psi_{k\nu})$, so daß diese Glieder einfach weggelassen werden können.

11. Ist die Störung $\lambda \mathbf{V}$ von \mathbf{H} nicht invariant gegenüber der ganzen Gruppe der \mathbf{P} , sondern nur gegenüber einer Untergruppe, so muß man solche Eigenfunktionen einführen, die zu irreduziblen Darstellungen dieser Untergruppe gehören. Die Eigenfunktionen und Eigenwerte von \mathbf{H} seien schon so geordnet angenommen, daß ihnen irreduzible Darstellungen der ganzen Gruppe der \mathbf{P} zugeordnet sind. Man kann dann die Matrizen, die den Elementen der Untergruppe zugeordnet sind, als Darstellung dieser Untergruppe auffassen. Es gilt — wie für alle \mathbf{P} , so auch für die \mathbf{P}_R der Untergruppe —

$$\mathbf{P}_R \psi_x^{(j)} = \sum_{\lambda} \mathbf{D}^{(j)}(R)_{\lambda x} \psi_{\lambda}^{(j)}.$$

Die $\mathbf{D}^{(j)}(R)$ für die R der Untergruppe sind aber nicht irreduzibel, und um zu solchen Funktionen zu gelangen, die zu irreduziblen Darstellungen der Untergruppe gehören, muß man sie ausreduzieren. Die Anzahl und Art der irreduziblen Bestandteile von $\mathbf{D}^{(j)}(R)$ als Darstellung der Untergruppe gibt uns Anzahl

und Art der Eigenwerte an, in die der betrachtete Eigenwert aufspalten kann.

Wir sehen, daß für die Charakterisierung der Eigenwerte der Schrödinger-Gleichung die Kenntnis der irreduziblen Darstellungen der symmetrischen Gruppe von n Elementen und der dreidimensionalen Drehgruppe notwendig ist. Im folgenden werden wir uns also dieser Aufgabe zuwenden.

12. In diesem ganzen Kapitel wurde von den Operatoren P_R nur die Tatsache, daß sie eine Gruppe bilden, sowie ihr linear-unitärer Charakter [nicht dagegen etwa (22) des vorangehenden Kapitels] vorausgesetzt. Außerdem wurde nur noch angenommen, daß die P_R Eigenfunktionen eines bestimmten Eigenwerts von H in Eigenfunktionen desselben Eigenwerts überführen. Aus diesen Voraussetzungen folgt schon die hier allein zugrunde gelegte Transformationsgleichung (23) Kap. XI und daß die darin auftretenden Koeffizienten eine Darstellung der Gruppe der Operatoren P_R bilden.

Dies wird hier deshalb ausdrücklich bemerkt, weil es später (bei der Theorie des Drehelektrons) vorkommen wird, daß für die Operatoren, die Eigenfunktionen in Eigenfunktionen überführen (sie werden dort O_R genannt), erstens (22) nicht mehr gilt, und daß zweitens die Symmetriegruppe des Konfigurationsraumes zu ihrer Gruppe nicht mehr holomorph, sondern nur isomorph ist. Die Koeffizienten in (23) werden dann eine Darstellung der Gruppe der Operatoren O_R und nicht der Symmetriegruppe des Konfigurationsraumes bilden; alle anderen Sätze dieses Kapitels (z. B. die Orthogonalität der zu verschiedenen irreduziblen Darstellungen gehörigen Eigenfunktionen) bleiben aber unverändert.

XIII. Die symmetrische Gruppe

1. Die Elemente der symmetrischen Gruppe n -ten Grades sind die Permutationen, die Vertauschungen von n Dingen. Ihre Ordnung ist $n!$. Man bezeichnet mit $\begin{pmatrix} 1 & 2 & \dots & n \\ \alpha_1 & \alpha_2 & \dots & \alpha_n \end{pmatrix}$ diejenige Umordnung, bei der 1 durch α_1 , 2 durch α_2 , ..., schließlich n durch α_n ersetzt wird. Mit $\begin{pmatrix} 1 & 2 & \dots & n \\ \alpha_1 & \alpha_2 & \dots & \alpha_n \end{pmatrix}$ ist $\begin{pmatrix} k_1 & k_2 & \dots & k_n \\ \alpha_{k_1} & \alpha_{k_2} & \dots & \alpha_{k_n} \end{pmatrix}$ wesensgleich, da es ja auch jedes k in α_k überführt. Dabei kann k_1, k_2, \dots, k_n eine beliebige Reihenfolge der Zahlen 1, 2, ..., n sein. Unter dem Produkt zweier Permutationen $A = \begin{pmatrix} 1 & 2 & \dots & n \\ \alpha_1 & \alpha_2 & \dots & \alpha_n \end{pmatrix}$ und

$B = \begin{pmatrix} 1 & 2 & \dots & n \\ \beta_1 & \beta_2 & \dots & \beta_n \end{pmatrix}$ versteht man das Nacheinanderausführen der beiden: A führt k in α_k und B dieses in β_{α_k} über, so daß AB das k in β_{α_k} überführt. Die Transformationen (A) des elften Kapitels bilden eine zur symmetrischen Gruppe n -ten Grades holomorphe Gruppe, die Transformation (A) führt den Punkt x_1, x_2, \dots, x_n in den Punkt $x_{\alpha_1}, x_{\alpha_2}, \dots, x_{\alpha_n}$ über und entspricht dabei der Permutation A .

Man kann bekanntlich für die Permutationen eine andere Bezeichnung einführen, sie in „Cyklen auflösen“. Ein Cyklus $(r_1 r_2 \dots r_\lambda)$ ist eine Permutation, die jedes ihrer Elemente r_k durch das darauf folgende r_{k+1} ersetzt, nur das allerletzte Element des Cyklus, r_λ , wird in das allererste, r_1 , übergeführt. Der Cyklus $(r_1 r_2 \dots r_\lambda)$ ist mit der Permutation $\begin{pmatrix} r_1 & r_2 & \dots & r_\lambda \\ r_2 & r_3 & \dots & r_1 \end{pmatrix}$ identisch. Er ist auch mit dem Cyklus $(r_2 r_3 \dots r_\lambda r_1)$ oder auch mit $(r_3 r_4 \dots r_\lambda r_1 r_2)$ gleichbedeutend.

Cyklen, die keine gemeinsamen Elemente enthalten, sind vertauschbar. Zum Beispiel ist $(1 \ 3 \ 5) \ (2 \ 4 \ 6 \ 7) = (2 \ 4 \ 6 \ 7) \ (1 \ 3 \ 5) = (1 \ 3 \ 5 \ 2 \ 4 \ 6 \ 7) \ (3 \ 5 \ 1 \ 4 \ 6 \ 7 \ 2)$.

Eine Permutation in Cyklen auflösen bedeutet: sie als Produkt lauter solcher elementenfremder vertauschbarer Cyklen schreiben. Dabei ist die Reihenfolge der einzelnen Faktoren, der einzelnen Cyklen, sowie das Anfangselement eines jeden Cyklus noch willkürlich wählbar. Man kann diese Auflösung in Cyklen ausführen, indem man etwa mit dem Element 1 beginnt, danach dasjenige Element setzt, in was 1 übergeführt wird, darauf folgt dasjenige, in welches dieses übergeführt wird usw. So kommt man schließlich zu einem Element, das in die 1 übergeführt wird, dieses ist das letzte Glied des ersten Cyklus. Hierauf geht man zu einem beliebigen anderen Element über, das vom Cyklus noch nicht erfaßt wurde und wiederholt damit dasselbe Verfahren. So fährt man fort, bis die ganze Permutation erschöpft ist.

Z. B. ist die Permutation $\begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 3 & 4 & 6 & 2 & 5 & 1 \end{pmatrix}$ in Cyklen aufgelöst $(1 \ 3 \ 6) \ (2 \ 4) \ (5)$ gleich $(3 \ 6 \ 1) \ (2 \ 4) \ (5)$ oder auch $(2 \ 4) \ (5) \ (1 \ 3 \ 6)$, da es auf die Reihenfolge der Cyklen nicht ankommt.

Permutationen, die aus gleich vielen gleich langen Cyklen bestehen, sind in derselben Klasse enthalten. In der Tat lassen sich z. B. ($n = \mu_\varrho$)

$$R = (r_1 r_2 \dots r_{\mu_1}) (r_{\mu_1+1} r_{\mu_1+2} \dots r_{\mu_2}) \dots (r_{\mu_{\varrho-1}+1} \dots r_{\mu_\varrho})$$

und

$$S = (s_1 s_2 \dots s_{\mu_1}) (s_{\mu_1+1} s_{\mu_1+2} \dots s_{\mu_2}) \dots (s_{\mu_{\varrho-1}+1} \dots s_{\mu_\varrho})$$

durch

$$T = \begin{pmatrix} s_1 s_2 \dots s_{\mu_1} & s_{\mu_1+1} s_{\mu_1+2} \dots s_{\mu_2} & \dots & s_{\mu_{\varrho-1}+1} \dots s_{\mu_\varrho} \\ r_1 r_2 \dots r_{\mu_1} & r_{\mu_1+1} r_{\mu_1+2} \dots r_{\mu_2} & \dots & r_{\mu_{\varrho-1}+1} \dots r_{\mu_\varrho} \end{pmatrix}$$

$$T^{-1} = \begin{pmatrix} r_1 r_2 \dots r_{\mu_1} r_{\mu_1+1} r_{\mu_1+2} \dots r_{\mu_2} \dots r_{\mu_{\varrho-1}+1} \dots r_{\mu_\varrho} \\ s_1 s_2 \dots s_{\mu_1} s_{\mu_1+1} s_{\mu_1+2} \dots s_{\mu_2} \dots s_{\mu_{\varrho-1}+1} \dots s_{\mu_\varrho} \end{pmatrix}$$

ineinander transformieren: $S = TRT^{-1}$, und umgekehrt sind die Cyklenlängen einer Permutation, die aus R durch eine Transformation T hervorgeht, wiederum $\mu_1, \mu_2 - \mu_1, \mu_3 - \mu_2, \dots, \mu_\varrho - \mu_{\varrho-1}$.

Wollen wir daher von zwei Permutationen entscheiden, ob sie in derselben Klasse enthalten sind, so können wir in beiden den (oder die) längsten Cyklen voranstellen, den zweitlängsten darauf folgen lassen usw. und den kürzesten an die letzte Stelle setzen. Sind dann die Längen der Cyklen $\lambda_1 = \mu_1, \lambda_2 = \mu_2 - \mu_1, \dots, \lambda_\varrho = \mu_\varrho - \mu_{\varrho-1}$ (mit $\lambda_1 \geqq \lambda_2 \geqq \dots \geqq \lambda_\varrho$ und $\lambda_1 + \lambda_2 + \dots + \lambda_\varrho = \mu_\varrho = n$) in beiden Permutationen paarweise gleich, so gehören sie in dieselbe Klasse, sonst nicht. Die Anzahl der Klassen ist also gleich der voneinander verschiedenen Cykleneinteilungen, der Anzahl der Zahlensysteme $\lambda_1, \lambda_2, \dots, \lambda_\varrho$, die den Bedingungen $\lambda_1 \geqq \lambda_2 \geqq \dots \geqq \lambda_\varrho$ und $\lambda_1 + \lambda_2 + \dots + \lambda_\varrho = n$ genügen. Diese Zahl, die Anzahl der möglichen Zerlegungen von n in positive ganzzählige Summanden ohne Rücksicht auf die Reihenfolge¹⁾, bezeichnet man als „partitio numerorum“ von n . Nach Kap. IX ist die Anzahl der voneinander verschiedenen irreduziblen Darstellungen gleich der Zahl der Klassen, also der partitio numerorum von n .

Z. B. hat die symmetrische Gruppe vierten Grades (der Ordnung 24) fünf verschiedene Klassen. Repräsentanten je einer Klasse sind z. B. $E = (1)(2)(3)(4); (1\ 2)(3)(4); (1\ 2)(3\ 4); (1\ 2\ 3)(4); (1\ 2\ 3\ 4)$. Sie muß also auch fünf irreduzible Darstellungen haben. Die symmetrische Gruppe dritten Grades hat den drei Klassen $E = (1)(2)(3); (1\ 2)(3)$;

¹⁾ Es ist für die Anzahl der Zahlensysteme λ offenbar gleichgültig, ob man die Reihenfolge der λ unberücksichtigt läßt, oder nur eine Reihenfolge $\lambda_1 \geqq \dots \geqq \lambda_\varrho$ in Betracht zieht.

(1 2 3) entsprechend die schon im neunten Kapitel besprochenen drei irreduziblen Darstellungen.

Die Einsercyklen lässt man häufig weg und schreibt z. B. an Stelle von (1 2)(3)(4) einfach (1 2).

2. Die einfachsten Permutationen sind — abgesehen von der Einheit — diejenigen, die nur je zwei Elemente vertauschen. Eine solche Permutation — in Cyklen aufgelöst kann man sie (kl) schreiben — nennt man eine Transposition. Jede Permutation lässt sich als Produkt von genügend vielen Transpositionen schreiben, z. B. eine Permutation, die nur aus einem Cyklus besteht:

$$(\alpha_1 \alpha_2 \dots \alpha_k) = (\alpha_1 \alpha_2) (\alpha_1 \alpha_3) \dots (\alpha_1 \alpha_k),$$

und so auch das Produkt von mehreren Cyklen, also jede Permutation.

Der Begriff der geraden und ungeraden Permutation spielt in der Determinantentheorie eine wichtige Rolle. Der Wert einer Determinante

$$\begin{vmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{vmatrix}$$

ist ja gleich der Summe der $n!$ Produkte

$$|a_{ik}| = \sum \varepsilon_{(\alpha_1 \alpha_2 \dots \alpha_n)} a_{1\alpha_1} a_{2\alpha_2} \dots a_{n\alpha_n},$$

wo $\alpha_1 \alpha_2 \dots \alpha_n$ alle $n!$ Permutationen der n Zahlen 1, 2, ..., n durchläuft und $\varepsilon_{(\alpha_1 \alpha_2 \dots \alpha_n)}$ gleich +1 oder -1 ist, je nachdem $\binom{1 2 \dots n}{\alpha_1 \alpha_2 \dots \alpha_n}$ eine gerade oder ungerade Permutation ist; d. h. je nachdem es als Produkt einer geraden oder ungeraden Zahl von Transpositionen geschrieben werden kann. (Man kann zwar eine Permutation in mannigfach verschiedener Weise in Transpositionen zerlegen, aber die Zerlegungen einer Permutation bestehen entweder alle aus einer geraden oder alle aus einer ungeraden Anzahl von Transpositionen.)

Das Produkt zweier gerader Permutationen ist wieder eine gerade Permutation, da man sie als Produkt von so vielen Transpositionen schreiben kann, wie die beiden Permutationen zusammen enthalten. Die geraden Permutationen bilden also eine Untergruppe, die alternierende Gruppe. Der Index der alternierenden Untergruppe ist 2, da eine ein-eindeutige Zuordnung zwischen den ungeraden und geraden Permutationen — etwa durch Multiplikation mit (12) — hergestellt werden kann. Die alternierende Gruppe

ist ein Normalteiler der symmetrischen Gruppe, weil das konjugierte Element einer geraden Permutation P wieder eine gerade Permutation $S^{-1}PS$ ist, da es als Produkt von doppelt so viel Transpositionen, wie S plus so viel Transpositionen wie P geschrieben werden kann.

Den Cyklus $(\alpha_1 \alpha_2 \alpha_3 \dots \alpha_\lambda) = (\alpha_1 \alpha_2)(\alpha_1 \alpha_3) \dots (\alpha_1 \alpha_\lambda)$ kann man als Produkt von $\lambda - 1$ Transpositionen schreiben. Eine Permutation mit den Cyklenlängen

$$\lambda_1, \lambda_2, \dots, \lambda_q \text{ (mit } \lambda_1 + \lambda_2 + \dots + \lambda_q = n\text{)}$$

kann man also als Produkt von $\lambda_1 - 1 + \lambda_2 - 1 + \dots + \lambda_q - 1$ Transpositionen schreiben: unter diesen Zahlen muß für alle Elemente der alternierenden Gruppe eine gerade Anzahl von ungeraden Zahlen vorhanden sein, unter den $\lambda_1, \lambda_2, \dots, \lambda_q$ also eine gerade Anzahl von geraden Zahlen. Die Permutationen der alternierenden Gruppe enthalten eine gerade Anzahl von geradzahligen Cyklen (Zweiercyklen, Vierercyklen usw.).

Die Faktorgruppe der alternierenden Gruppe hat die Ordnung zwei. Aus ihren beiden irreduziblen Darstellungen kann man zwei Darstellungen der ganzen symmetrischen Gruppe gewinnen, indem man entweder ihrer identischen Darstellung entsprechend sowohl den Elementen der alternierenden Gruppe, wie auch den Elementen der Nebengruppe die Matrix (1) zuordnet, oder indem man ihrer Darstellung $D(E) = (1); D(S) = (-1)$ entsprechend den Elementen der alternierenden Gruppe die Matrix (1), den Elementen der Nebengruppe, den ungeraden Permutationen die Matrix (-1) zuordnet. Im ersten Fall erhält man die identische Darstellung $D^{(0)}(R) = (1)$, die zweite Darstellung nennt man die antisymmetrische $\bar{D}^{(0)}(R) = (\varepsilon_R)$. Diese beiden Darstellungen sind eindimensional.

3. Alle anderen Darstellungen der symmetrischen Gruppe sind mehrdimensional. In einer eindimensionalen Darstellung kann nämlich etwa der Transposition (12) nur entweder die Matrix (1) oder die Matrix (-1) entsprechen, da das Quadrat dieser Matrix die Einheitsmatrix (1) sein muß. Im ersten Fall muß aber jeder Transposition (1), im zweiten Fall jeder Transposition (-1) durch die Darstellung zugeordnet sein, weil alle Transpositionen in derselben Klasse sind und in jeder Darstellung alle denselben Charakter haben müssen. Die Matrizen, die den Transpositionen entsprechen,

bestimmen aber ihrerseits die ganze Darstellung, weil sich alle Gruppenelemente als Produkte von Transpositionen schreiben lassen. Man erhält daher im ersten Fall die identische, im zweiten die antisymmetrische Darstellung.

Die Abelsche Faktorgruppe der alternierenden Gruppe stellt einen sehr wichtigen Zusammenhang zwischen je zwei irreduziblen Darstellungen her. Betrachten wir eine irreduzible Darstellung $\mathbf{D}^{(k)}(R)$, so kann man aus ihr eine weitere, die assoziierte Darstellung $\bar{\mathbf{D}}^{(k)}(R)$ bilden, indem man alle Matrizen, die zu den Elementen der alternierenden Gruppe zugeordnet sind, ungeändert läßt, alle anderen aber mit -1 multipliziert. Die so erhaltenen Matrizen bilden auch eine Darstellung der Gruppe, wie man sich leicht überzeugt, da $\bar{\mathbf{D}}^{(k)}(R)$ eigentlich das direkte Produkt von $\mathbf{D}^{(k)}(R)$ mit $\bar{\mathbf{D}}^{(0)}(R)$, der antisymmetrischen Darstellung ist:

$$\bar{\mathbf{D}}^{(k)}(R) = \mathbf{D}^{(k)}(R) \times \bar{\mathbf{D}}^{(0)}(R) = \varepsilon_R \mathbf{D}^{(k)}(R),$$

da $\bar{\mathbf{D}}^{(0)}(R)$ eine Zahl ± 1 ist.

Diese assoziierten Darstellungen spielen sowohl in der Quantenmechanik wie in der mathematischen Theorie der irreduziblen Darstellungen eine sehr wichtige Rolle und wir wollen die uns interessierenden Darstellungen auch zum Teil mit ihrer Hilfe ableiten.

Die Anzahl der verschiedenen irreduziblen Darstellungen der symmetrischen Gruppe ist gleich der Partitio numerorum von n . Dies sollte auch die Zahl der qualitativ verschiedenen Arten von Eigenwerten sein. Es zeigt sich jedoch, daß nur Eigenwerten mit einigen bestimmten Darstellungen wirkliche Energieniveaus des Atoms entsprechen, den Eigenwerten mit anderen Darstellungen entsprechen keine wirklich existierenden stationären Zustände, sie sind durch ein von der Eigenwertgleichung unabhängiges Prinzip, das Pauliprinzip verboten. Mit Hilfe der Methode, die hier verwendet werden soll, kann man zwar sämtliche irreduziblen Darstellungen der symmetrischen Gruppe bestimmen, wir wollen aber nur die Bestimmung jener Darstellungen ausführen, deren Eigenwerte durch das Pauliprinzip nicht verboten sind. Die genaue Formulierung des Pauliprinzips wird zwar hier noch nicht gegeben werden, doch wird die Methode, mit Hilfe der wir die in Betracht kommenden Darstellungen bestimmen, genau die Überlegungen enthalten, die später bei der Anwendung des Pauliprinzips notwendig sein werden.

4. Haben wir ein System von Variablen, die nur je einen Wert — etwa 1 — annehmen können, so besteht der Variabilitätsbereich aus einem einzigen Punkt und jede Funktion ist vollkommen bestimmt, wenn ihr Wert in diesem Punkt gegeben ist. In diesem Raum gibt es keine zwei linear unabhängigen Funktionen, alle sind im „ganzen Variabilitätsbereich“ konstant und so Vielfache voneinander. Jede Funktion in diesem Raum bleibt ungeändert, wenn man die Werte der Koordinaten vertauscht — da man ja dadurch nur 1 mit 1 ersetzt. Alle Funktionen in diesem Raum gehören zur identischen Darstellung.

Betrachten wir n Variable s_1, s_2, \dots, s_n , die je zweier Werte (etwa +1 und -1) fähig sind, so besteht der gesamte Raum aus 2^n Punkten und wir haben 2^n linear unabhängige Funktionen, etwa jene, die nur in je einem dieser 2^n Punkte 1, in allen anderen Punkten dagegen Null sind. Das skalare Produkt zweier Funktionen φ und g ist in diesem Raum

$$\sum_{s_1=\pm 1} \sum_{s_2=\pm 1} \dots \sum_{s_n=\pm 1} \varphi(s_1 \dots s_n) * g(s_1 \dots s_n) = (\varphi, g).$$

Im Raum eines s_k (er besteht nur aus den beiden Punkten $s_k = -1$ und $s_k = +1$) bilden die beiden Funktionen $\delta_{s_k=-1}$ und $\delta_{s_k=+1}$ ein „vollständiges Orthogonalsystem“; die 2^n Produkte dieser Funktionen $\delta_{s_1 \sigma_1} \delta_{s_2 \sigma_2} \dots \delta_{s_n \sigma_n}$ (mit $\sigma_1 = \pm 1, \sigma_2 = \pm 1, \dots, \sigma_n = \pm 1$)

bilden ein vollständiges Orthogonalsystem im n -dimensionalen Raum der s_1, s_2, \dots, s_n . Die folgenden Formeln lassen sich wesentlich bequemer schreiben, wenn wir nicht die Funktionen $\delta_{s_k=1}, \delta_{s_k=-1}$, sondern die beiden Funktionen 1 und s_k gebrauchen, die ebenfalls orthogonal sind:

$$(1, s_k) = \sum_{s_k=\pm 1} 1 \cdot s_k = 1 \cdot -1 + 1 \cdot 1 = 0.$$

Das vollständige Funktionensystem im Raum der s_1, s_2, \dots, s_n besteht dann aus den 2^n Funktionen $s_1^{\gamma_1} s_2^{\gamma_2} \dots s_n^{\gamma_n}$ (mit γ_k gleich 0 oder 1), die man auch folgendermaßen ordnen kann:

$$\left. \begin{array}{c} 1 \\ s_1, s_2, \dots, s_n \\ s_1 s_2, s_1 s_3, \dots, s_1 s_n, s_2 s_3, s_2 s_4, \dots, s_2 s_n, \dots, s_{n-1} s_n \\ s_1 s_2 s_3, \dots, s_{n-2} s_{n-1} s_n \\ \dots \\ s_1 s_2 s_3 \dots s_n \end{array} \right\} \quad (1)$$

Das sind $1 + \binom{n}{1} + \binom{n}{2} + \cdots + \binom{n}{n} = 2^n$ Funktionen. Wendet man ein \mathbf{P}_R , wo R eine Permutation ist, auf eine dieser Funktionen an, so entsteht daraus eine neue Funktion der s_1, s_2, \dots, s_n , die — wie jede Funktion dieser Variablen — durch diese 2^n linear ausgedrückt werden kann. Die Koeffizienten würden eine 2^n -dimensionale Darstellung der symmetrischen Gruppe geben. Diese Darstellung $\Delta(R)$ ist nicht irreduzibel, sondern enthält noch mehrere irreduzible Bestandteile. Wegen des stark eingeengten Definitionsbereiches der in Betracht gezogenen Funktionen ist immerhin zu erwarten, daß $\Delta(R)$ nicht alle irreduziblen Darstellungen der symmetrischen Gruppe enthalten wird, so daß wir hoffen können, daß man es leichter ausreduzieren kann, als eine ganz beliebige Darstellung. Seine irreduziblen Bestandteile bzw. die assoziierten dieser sind aber die einzigen Darstellungen, die für Elektronen in Betracht kommen.

Wendet man auf eine der Funktionen (1), etwa auf $s_a s_b s_c$, den Operator \mathbf{P}_R an, wo R eine beliebige Permutation ist, d. h. führt man eine „Vertauschung der Variablen“ aus, so erhält man wieder ein Produkt aus drei s , etwa $s_{a'} s_{b'} s_{c'}$, also eine Funktion, die in derselben [dritten¹⁾] Zeile von (1) steht, in der auch $s_a s_b s_c$ war. Will man die $\binom{n}{k}$ Funktionen, die entstehen, wenn man auf alle Funktionen der k -ten Zeile einen Operator \mathbf{P}_R anwendet, durch die Funktionen (1) ausdrücken, so braucht man dazu nur die Funktionen derselben k -ten Zeile, aus der diese Funktionen stammen. Diese Funktionen ergeben daher schon für sich eine Darstellung $\Delta^{(k)}(R)$ der symmetrischen Gruppe mit der Dimension $\binom{n}{k}$, und $\Delta(R)$ zerfällt in die Darstellungen $\Delta^{(0)}(R), \Delta^{(1)}(R), \dots, \Delta^{(n)}(R)$, deren Matrizen, wie gleichzeitig bemerkt werde, so aussehen, daß sie in jeder Zeile eine 1, sonst lauter Nullen enthalten.

Berechnen wir noch den Charakter von $\Delta^{(k)}(R)$! Der Charakter, der dem Element R entspricht, ist offenbar gleich der Anzahl der Funktionen in der k -ten Zeile von (1), die bei der Anwendung von \mathbf{P}_R unverändert bleiben. In den diesen Funktionen ent-

¹⁾ Wir beginnen die Numerierung der Zeilen von (1) bei Null, die letzte Zeile ist die n -te.

sprechenden Spalten steht nämlich in $\Delta^{(k)}(R)$ die 1 in der Hauptdiagonale, in allen anderen Zeilen anderswo, so daß in der Hauptdiagonale eine 0 steht.

Sei jetzt R eine Permutation mit den Cyklenlängen $\lambda_1 = \mu_1$; $\lambda_2 = \mu_2 - \mu_1, \dots, \lambda_q = \mu_q - \mu_{q-1}$ ($\mu_q = n$) etwa die Permutation $(1\ 2\ \dots\ \mu_1)(\mu_1 + 1\ \dots\ \mu_2)\dots(\mu_{q-1} + 1,\ \mu_{q-1} + 2, \dots, \mu_q)$. Soll diese die Funktion $s_1^{\alpha_1} s_2^{\alpha_2} \dots s_n^{\alpha_n}$ ungeändert lassen, so müssen die Exponenten von $s_1, s_2, \dots, s_{\mu_1}$ untereinander gleich sein, ebenso die Exponenten von $s_{\mu_1+1}, s_{\mu_1+2}, \dots, s_{\mu_2}$ usw., schließlich auch die Exponenten von $s_{\mu_{q-1}+1}, s_{\mu_{q-1}+2}, \dots, s_{\mu_q} = s_n$.

Es werden daher von allen Funktionen (1) diejenigen von P_R ungeändert gelassen, die sich in der Gestalt

$$(s_1 s_2 \dots s_{\mu_1})^{\gamma_1} (s_{\mu_1+1} s_{\mu_1+2} \dots s_{\mu_2})^{\gamma_2} \dots (s_{\mu_{q-1}+1} \dots s_{\mu_q})^{\gamma_q} \quad (2)$$

schreiben lassen. (Es sind alle γ Null oder 1.) Wir interessieren uns für die Anzahl der Funktionen der k -ten Zeile von (1), die die Form von (2) haben. Für diese Funktionen ist noch

$$\mu_1 \gamma_1 + (\mu_2 - \mu_1) \gamma_2 + \dots + (\mu_q - \mu_{q-1}) \gamma_q = \lambda_1 \gamma_1 + \lambda_2 \gamma_2 + \dots + \lambda_q \gamma_q = k, \quad (3)$$

so daß ihre Anzahl gleich der Anzahl der Lösungen von (3) ist, wobei jedoch für die Unbekannten $\gamma_1, \gamma_2, \dots, \gamma_q$ nur die Zahlen 0 und 1 zugelassen sind. Dies ist der Charakter von R in $\Delta^{(k)}(R)$, oder auch jeder anderen Permutation mit den Cyklenlängen $\lambda_1, \lambda_2, \dots, \lambda_q$, da diese ja alle in derselben Klasse sind und daher alle dieselbe Spur haben müssen. Man pflegt dies so auszudrücken, daß der Koeffizient der k -ten Potenz von x im Polynom

$$(1 + x^{\lambda_1})(1 + x^{\lambda_2}) \dots (1 + x^{\lambda_q}) \quad (\lambda_1 + \lambda_2 + \dots + \lambda_q = n) \quad (4)$$

die Spur von $\Delta^{(k)}(R)$ ist.

Die Spur von $\Delta^{(k)}(E)$ muß gleich der Dimension $\binom{n}{k}$ der Darstellung sein. Da für E alle Cyklenlängen $\lambda_1 = \lambda_2 = \dots = \lambda_q = 1$ sind, ist (4) gleich $(1+x)^n$, der Koeffizient von x^k darin ist tatsächlich $\binom{n}{k}$. Die Spur, die zu einer Transposition $(1\ 2)(3\ 4)\dots(n)$ zugeordnet ist, ist der Koeffizient von x^k in

$$(1 + x^2)(1 + x)(1 + x) \dots (1 + x) = (1 + x^2)(1 + x)^{n-2}.$$

Er berechnet sich zu

$$\sum_x \Delta^{(k)}(R)_{xx} = \binom{n-2}{k} + \binom{n-2}{k-2}.$$

Nun ist aber $\Delta^{(k)}(R)$ noch keine irreduzible Darstellung. Man kann nämlich aus den $\binom{n}{k}$ Funktionen der k -ten Zeile von (1) solche Linearkombinationen bilden, die bei der Anwendung der P_R unter sich, und zwar nach $\Delta^{(k-1)}(R)$ transformieren. Dies wäre bei einer irreduziblen Darstellung nicht möglich.

Die Linearkombinationen der Funktionen der k -ten Zeile von (1), die sich ebenso wie die Funktionen $s_{a_1} s_{a_2} \dots s_{a_{k-1}}$ der $k-1$ -ten Zeile transformieren, sind

$$F_{a_1 a_2 \dots a_{k-1}} = s_{a_1 a_2 \dots a_{k-1}} s_{a_1} s_{a_2} \dots s_{a_{k-1}}, \quad (5)$$

wo $s_{a_1 a_2 \dots a_{k-1}}$ die Summe aller derjenigen $n - k + 1$ Variablen ist, die unter den $s_{a_1}, s_{a_2}, \dots, s_{a_{k-1}}$ nicht vorkommen. Geht bei der Anwendung eines P_R das s_{a_i} in s_{b_i} über, so geht $s_{a_1} s_{a_2} \dots s_{a_{k-1}}$ in $s_{b_1} s_{b_2} \dots s_{b_{k-1}}$ und auch $s_{a_1 a_2 \dots a_{k-1}}$ in $s_{b_1 b_2 \dots b_{k-1}}$ über, so daß sich die $F_{a_1 a_2 \dots a_{k-1}}$ tatsächlich genau wie die $s_{a_1} s_{a_2} \dots s_{a_{k-1}}$ transformieren.

Im Anhang wird gezeigt, daß für $k \leq \frac{1}{2}n$ die $\binom{n}{k-1}$ Funktionen $F_{a_1 \dots a_{k-1}}$ voneinander linear unabhängig sind. Wir können daher für $k \leq \frac{1}{2}n$ aus den $\binom{n}{k}$ Funktionen $s_{a_1} s_{a_2} \dots s_{a_k}$ solche

$$l_k = \binom{n}{k} - \binom{n}{k-1} \quad (6)$$

Linearkombinationen g_1, g_2, \dots, g_{l_k} bilden¹⁾, die auf alle $F_{a_1 \dots a_{k-1}}$ und aufeinander paarweise orthogonal und normiert sind. Umgekehrt lassen sich die $s_{a_1} s_{a_2} \dots s_{a_k}$ durch die $F_{a_1 \dots a_{k-1}}$ und die g_x linear ausdrücken. Wir wollen noch die $F_{a_1 \dots a_{k-1}}$, die zwar linear unabhängig, aber nicht orthogonal sind, durch orthogonale Funktionen $F_1, F_2, \dots, F_{\binom{n}{k-1}}$ ersetzen und die Darstellung $\Delta^{(k)}(R)$ in der

Form $\bar{\Delta}^{(k)}(R)$ betrachten, die sie annimmt, wenn man an Stelle der $s_{a_1} s_{a_2} \dots s_{a_k}$ die $F_1, F_2, \dots, F_{\binom{n}{k-1}}, g_1, g_2, \dots, g_{l_k}$ als linear

¹⁾ Eine solche Funktion ist z. B. $(s_1 - s_2)(s_3 - s_4) \dots (s_{2k-1} - s_{2k})$.

unabhängige Funktionen einführt: $\Delta^{(k)}(R)$ erleidet ja hierdurch nur eine Ähnlichkeitstransformation.

Da F_x eine Linearkombination der $F_{a_1 \dots a_{k-1}}$ ist, ist $\mathbf{P}_R F_x$ eine Linearkombination der $\mathbf{P}_R F_{a_1 \dots a_{k-1}}$ und läßt sich ebenso wie diese durch die $F_{a_1 \dots a_{k-1}}$, also auch durch die F_x linear ausdrücken. Die Koeffizienten können wir mit $\bar{\Delta}^{(k-1)}(R)_{\lambda x}$ bezeichnen

$$\mathbf{P}_R F_x = \sum_{\lambda=1}^{\binom{n}{k-1}} \bar{\Delta}^{(k-1)}(R)_{\lambda x} F_\lambda, \quad (7)$$

da sie eine zu $\Delta^{(k-1)}(R)$ äquivalente Darstellung bilden; $\bar{\Delta}^{(k)}(R)$ hat dann die Form

$$\bar{\Delta}^{(k)}(R) = \begin{pmatrix} \bar{\Delta}^{(k-1)}(R) & \mathbf{A}(R) \\ \mathbf{0} & \mathbf{D}^{(k)}(R) \end{pmatrix},$$

weil die g in (7) alle mit dem Koeffizienten Null vorkommen. Da weiter $\bar{\Delta}^{(k)}(R) = \bar{\Delta}^{(k)}(R^{-1})^\dagger$ in (7) unitär ist (die $F_1, \dots, F_{\binom{n}{k-1}}$, g_1, \dots, g_{l_k} sind paarweise orthogonal), folgt

$$\begin{aligned} \begin{pmatrix} \bar{\Delta}^{(k-1)}(R) & \mathbf{A}(R) \\ \mathbf{0} & \mathbf{D}^{(k)}(R) \end{pmatrix} &= \begin{pmatrix} \bar{\Delta}^{(k-1)}(R^{-1}) & \mathbf{A}(R^{-1}) \\ \mathbf{0} & \mathbf{D}^{(k)}(R^{-1}) \end{pmatrix}^\dagger \\ &= \begin{pmatrix} \bar{\Delta}^{(k-1)}(R^{-1})^\dagger & \mathbf{0} \\ \mathbf{A}(R^{-1})^\dagger & \mathbf{D}^{(k)}(R^{-1})^\dagger \end{pmatrix}, \end{aligned}$$

also ist $\mathbf{A}(R) = \mathbf{0}$ und

$$\bar{\Delta}^{(k)}(R) = \begin{pmatrix} \bar{\Delta}^{(k-1)}(R) & \mathbf{0} \\ \mathbf{0} & \mathbf{D}^{(k)}(R) \end{pmatrix}. \quad (8)$$

Ist $k \leq \frac{1}{2}n$, so zerfällt die Darstellung $\Delta^{(k)}(R)$ in zwei Darstellungen $\bar{\Delta}^{(k-1)}(R)$ und $\mathbf{D}^{(k)}(R)$ mit den Dimensionen $\binom{n}{k-1}$ und $l_k = \binom{n}{k} - \binom{n}{k-1}$, deren erste zu $\Delta^{(k-1)}(R)$ äquivalent ist. Daher kann man $\bar{\Delta}^{(k-1)}(R)$ seinerseits wieder in zwei Darstellungen $\bar{\Delta}^{(k-2)}$ und $\mathbf{D}^{(k-1)}$ zerlegen, dann $\bar{\Delta}^{(k-2)}$ weiter usw., so daß schließlich $\Delta^{(k)}$ in $\mathbf{D}^{(0)} + \mathbf{D}^{(1)} + \dots + \mathbf{D}^{(k)}$ zerfällt.

Dies gilt für $k \leq \frac{1}{2}n$. Für $k > \frac{1}{2}n$ transformieren sich die $\binom{n}{k} = \binom{n}{n-k}$ Funktionen, die $n-k$ Variable in erster Potenz, die übrigen in nullter enthalten, genau so, wie diejenigen, die k in

erster und $n-k$ in nullter Potenz enthalten. Es ist ja ganz gleichgültig gewesen, was für ein Orthogonalsystem wir für die s gewählt haben, an Stelle von $1, s$ hätten wir auch $s, 1$ benutzen können. Daher ist $\Delta^{(k)}$ äquivalent mit $\Delta^{(n-k)}$ und kann in dieselben Bestandteile zerlegt werden. Für die Zerlegung von $\Delta(R)$ ergibt sich also etwa folgendes Bild ($n = 4$):

$$\Delta(R) = \begin{cases} \Delta^{(0)}(R) = \mathbf{D}^{(0)} & = \mathbf{D}^{(0)} \\ \Delta^{(1)}(R) = \Delta^{(0)} + \mathbf{D}^{(1)} & = \mathbf{D}^{(0)} + \mathbf{D}^{(1)} \\ \Delta^{(2)}(R) = \Delta^{(1)} + \mathbf{D}^{(2)} & = \mathbf{D}^{(0)} + \mathbf{D}^{(1)} + \mathbf{D}^{(2)} \\ \Delta^{(3)}(R) \sim \Delta^{(1)} & = \mathbf{D}^{(0)} + \mathbf{D}^{(1)} \\ \Delta^{(4)}(R) \sim \Delta^{(0)} & = \mathbf{D}^{(0)} \end{cases}$$

oder für ungerades n ($n = 5$) etwa folgendes:

$$\Delta(R) = \begin{cases} \Delta^{(0)}(R) = \mathbf{D}^{(0)} & = \mathbf{D}^{(0)} \\ \Delta^{(1)}(R) = \Delta^{(0)} + \mathbf{D}^{(1)} & = \mathbf{D}^{(0)} + \mathbf{D}^{(1)} \\ \Delta^{(2)}(R) = \Delta^{(1)} + \mathbf{D}^{(2)} & = \mathbf{D}^{(0)} + \mathbf{D}^{(1)} + \mathbf{D}^{(2)} \\ \Delta^{(3)}(R) \sim \Delta^{(2)} & = \mathbf{D}^{(0)} + \mathbf{D}^{(1)} + \mathbf{D}^{(2)} \\ \Delta^{(4)}(R) \sim \Delta^{(1)} & = \mathbf{D}^{(0)} + \mathbf{D}^{(1)} \\ \Delta^{(5)}(R) \sim \Delta^{(0)} & = \mathbf{D}^{(0)}. \end{cases}$$

Wir werden jetzt zeigen, daß die soeben gewonnenen Darstellungen $\mathbf{D}^{(0)}, \mathbf{D}^{(1)}, \dots, \mathbf{D}^{(1/2n)}$ bzw. $\mathbf{D}^{(1/2n-1/2)}$ irreduzibel und voneinander verschieden sind. Zu diesem Zweck wollen wir die Irreduzibilität und Verschiedenheit der auf dieselbe Weise für die symmetrische Gruppe $n-1$ -ten Grades gewonnenen Darstellungen ' $\mathbf{D}^{(0)}(R')$, ' $\mathbf{D}^{(1)}(R')$, ..., ' $\mathbf{D}^{(k)}(R')$, ... für $k \leq \frac{1}{2}(n-1)$ voraussetzen und auch ihre Dimensionen zu $l'_k = \binom{n-1}{k} - \binom{n-1}{k-1}$ annehmen¹⁾.

5. Die Funktionen g_1, g_2, \dots, g_{l_k} können wir, da sie in s_n linear sind, so zerlegen

$$g_x = g'_x s_n + h'_x, \quad (9)$$

daß sowohl in g'_x wie auch in h'_x die Variable s_n in nullter Potenz vorkomme, g'_x und h'_x können dann auch als Funktionen der Variablen s_1, s_2, \dots, s_{n-1} allein betrachtet werden, g'_x ist $k-1$ -ten Grades,

¹⁾ Die gestrichenen Größen beziehen sich immer auf die symmetrische Gruppe $n-1$ -ten Grades bzw. auf Funktionen der $n-1$ Variablen s_1, s_2, \dots, s_{n-1} .

h'_x ist k -ten Grades. Nun ist g_x orthogonal auf alle Funktionen $F_{a_1 a_2 \dots a_{k-1}}$, also auch auf

$$F_{a_1 a_2 \dots a_{k-2} n} = s_{a_1} s_{a_2} \dots s_{a_{k-2}} s_n s_{a_1 a_2 \dots a_{k-2} n},$$

und da h'_x das s_n in nullter Potenz enthält, gilt dies auch für h'_x und muß folglich auch für $g'_x s_n$ gelten. Daher ist g'_x orthogonal zu den $s_{a_1} s_{a_2} \dots s_{a_{k-2}} s_{a_1 a_2 \dots a_{k-2} n}$. Dies ist aber die Definition derjenigen l'_{k-1} Funktionen der s_1, s_2, \dots, s_{n-1} vom Grade $k-1$, die sich voraussetzungsgemäß nach der irreduziblen Darstellung ' $D^{(k-1)}$ ' transformieren. Unter den g'_x sind also nur $l'' \leq l'_{k-1}$ linear unabhängig, und man kann, wenn l_k größer als l'' ist, solche Linearkombinationen \bar{g}_x der g_x bilden, daß

$$\bar{g}_x = \bar{g}'_x s_n + \bar{h}'_x \quad (x = 1, 2, \dots, l'') \quad (9a)$$

und die \bar{g}'_x linear unabhängig sind und

$$\bar{g}_x = \bar{h}'_x \quad (x = l'' + 1, l'' + 2, \dots, l_k) \quad (9b)$$

gilt. Auch die \bar{g}_x sind orthogonal zu den $F_{a_1 a_2 \dots a_{k-1}}$, und dies muß für $x > l''$ auch von den \bar{h}'_x gelten. Wir betrachten insbesondere solche $F_{a_1 a_2 \dots a_{k-1}}$, bei denen n unter den a_1, a_2, \dots, a_{k-1} nicht vorkommt. Dann ist

$$F_{a_1 a_2 \dots a_{k-1}} = s_{a_1} s_{a_2} \dots s_{a_{k-1}} (s_{a_1 a_2 \dots a_{k-1} n} + s_n).$$

Da in \bar{h}'_x das s_n in nullter Potenz ist, ist es orthogonal auf $s_{a_1} s_{a_2} \dots s_{a_{k-1}} s_n$, also für $x > l''$ auch auf $s_{a_1} s_{a_2} \dots s_{a_{k-1}} s_{a_1 a_2 \dots a_{k-1} n}$. Das ist aber die Definition derjenigen Funktionen der s_1, s_2, \dots, s_{n-1} , die voraussetzungsgemäß zu der irreduziblen Darstellung ' $D^{(k)}$ ' gehören, und die \bar{h}'_x müssen für $x > l''$ Linearkombinationen dieser sein.

Nun sei R' eine Permutation, die nur die ersten $n-1$ Variablen vertauscht, s_n dagegen unverändert läßt.

Dann ist

$$\mathbf{P}_{R'} \bar{g}_x = s_n \mathbf{P}_{R'} \bar{g}'_x + \mathbf{P}_{R'} \bar{h}'_x \quad (x = 1, 2, \dots, l''), \quad (10a)$$

$$\mathbf{P}_{R'} \bar{g}_x = \mathbf{P}_{R'} \bar{h}'_x \quad (x = l'' + 1, l'' + 2, \dots, l_k). \quad (10b)$$

Und es ist für alle Permutationen R , also auch für R'

$$\begin{aligned} \mathbf{P}_{R'} \bar{g}_x &= \sum_{\lambda=1}^{l_k} \mathbf{D}^{(k)}(R')_{\lambda x} \bar{g}_\lambda \\ &= \sum_{\lambda=1}^{l''} \mathbf{D}^{(k)}(R')_{\lambda x} \bar{g}'_\lambda s_n + \sum_{\lambda=1}^{l_k} \mathbf{D}^{(k)}(R')_{\lambda x} \bar{h}'_\lambda. \end{aligned} \quad (11)$$

Der Vergleich der rechten Seiten von (10a), (10b) und (11) ergibt durch Gleichsetzen der Koeffizienten von s_n (die $\mathbf{P}_{R'} \bar{h}'_x$ sind Funktionen der s_1, \dots, s_{n-1} allein, enthalten also s_n in nullter Potenz)

$$\mathbf{P}_{R'} \bar{g}'_x = \sum_{\lambda=1}^{l''} \mathbf{D}^{(k)}(R')_{\lambda x} \bar{g}'_\lambda \quad (x \leqq l''), \quad (12a)$$

$$0 = \mathbf{D}^{(k)}(R')_{\lambda x} \quad (x > l''; \lambda \leqq l''). \quad (12b)$$

Der Vergleich der von s_n freien Glieder ergibt für $x > l''$ mit Hilfe von (12b)

$$\mathbf{P}_{R'} \bar{h}'_x = \sum_{\lambda=1}^{l_k} \mathbf{D}^{(k)}(R')_{\lambda x} \bar{h}'_\lambda = \sum_{\lambda=l''+1}^{l_k} \mathbf{D}^{(k)}(R')_{\lambda x} \bar{h}'_\lambda \quad (x > l''). \quad (12c)$$

Aus¹⁾ (12b) ersehen wir, daß $\mathbf{D}^{(k)}(R')$ die Gestalt

$$\mathbf{D}^{(k)}(R') = \begin{pmatrix} \mathbf{A}(R') & \mathbf{0} \\ \mathbf{B}(R') & \mathbf{C}(R') \end{pmatrix}$$

hat. Nehmen wir nun an, daß die $\bar{g}_1, \bar{g}_2, \dots, \bar{g}_{l_k}$ paarweise orthogonal sind (was wir erreichen können, wenn wir auf die zunächst ohne diese Vorschrift gebildeten \bar{g}_x das Schmidtsche Orthogonalisierungsverfahren in der Reihenfolge $\bar{g}_{l_k}, \dots, \bar{g}_2, \bar{g}_1$ anwenden), so muß $\mathbf{D}^{(k)}$ unitär sein und $\mathbf{B}(R')$ muß, wie wir das schon bei (8) gesehen haben, verschwinden.

Nun folgt weiter aus (12a) und (12c), da \bar{g}'_x für $x \leqq l''$ und \bar{h}'_x für $x > l''$ zu ' $\mathbf{D}^{(k-1)}$ ' bzw. zu ' $\mathbf{D}^{(k)}$ ' gehören, daß $\mathbf{A}(R')$ und $\mathbf{C}(R')$ äquivalent zu ' $\mathbf{D}^{(k-1)}(R')$ ' bzw. ' $\mathbf{D}^{(k)}(R')$ ' sein muß. Die Darstellung $\mathbf{D}^{(k)}(R')$, als Darstellung der symmetrischen Gruppe $n-1$ -ten Grades betrachtet, die nur die ersten $n-1$ Variablen vertauscht, zerfällt in zwei verschiedene irreduzible Bestandteile ' $\mathbf{D}^{(k-1)}(R')$ ' und ' $\mathbf{D}^{(k)}(R')$ ', d.h. die Matrizen, die diesen Permutationen entsprechen, haben die Gestalt

$$\mathbf{D}^{(k)}(R') = \begin{pmatrix} ' \mathbf{D}^{(k-1)}(R') & \mathbf{0} \\ \mathbf{0} & ' \mathbf{D}^{(k)}(R') \end{pmatrix}. \quad (13)$$

¹⁾ Aus (12a) sehen wir, daß $l'' = l'_{k-1}$ sein muß. Wäre nämlich $l'' < l'_{k-1}$, so würden die Funktionen \bar{g}'_x nach (12a) zu einer Darstellung gehören, deren Dimension kleiner als die von ' $\mathbf{D}^{(k-1)}$ ' ist. Da sie aber zu ' $\mathbf{D}^{(k-1)}$ ' gehören und diese irreduzibel ist, ist dies nicht möglich.

Die Dimension von $\mathbf{D}^{(k)}$ muß also gleich der Summe der Dimensionen von ' $\mathbf{D}^{(k-1)}$ ' und ' $\mathbf{D}^{(k)}$ ' sein. In der Tat ist

$$\begin{aligned} l_k &= \binom{n}{k} - \binom{n}{k-1} \\ &= \binom{n-1}{k-1} + \binom{n-1}{k} - \binom{n-1}{k-2} - \binom{n-1}{k-1} = l_{k-1} + l'_k. \end{aligned}$$

Betrachten wir jetzt eine Matrix

$$\begin{pmatrix} \mathbf{M}_1 & \mathbf{M}_2 \\ \mathbf{M}_3 & \mathbf{M}_4 \end{pmatrix}, \quad (14)$$

die mit allen $\mathbf{D}^{(k)}(R)$ vertauschbar ist. Die Zeilen-Spalteneinteilung sei in (14) dieselbe wie in (13). Es muß (14) insbesondere auch mit den $\mathbf{D}^{(k)}(R')$ von (13) vertauschbar sein:

$$\begin{aligned} &\begin{pmatrix} \mathbf{D}^{(k-1)}(R') & \mathbf{0} \\ \mathbf{0} & \mathbf{D}^{(k)}(R') \end{pmatrix} \begin{pmatrix} \mathbf{M}_1 & \mathbf{M}_2 \\ \mathbf{M}_3 & \mathbf{M}_4 \end{pmatrix} \\ &= \begin{pmatrix} \mathbf{M}_1 & \mathbf{M}_2 \\ \mathbf{M}_3 & \mathbf{M}_4 \end{pmatrix} \begin{pmatrix} \mathbf{D}^{(k-1)}(R') & \mathbf{0} \\ \mathbf{0} & \mathbf{D}^{(k)}(R') \end{pmatrix}. \end{aligned}$$

Daher gilt für alle Matrizen der irreduziblen Darstellungen ' $\mathbf{D}^{(k-1)}(R')$ ' bzw. ' $\mathbf{D}^{(k)}(R')$ ' der symmetrischen Gruppe $n-1$ -ten Grades, die den Permutationen der s_1, s_2, \dots, s_{n-1} entspricht:

$$\begin{aligned} \mathbf{D}^{(k-1)}(R') \mathbf{M}_1 &= \mathbf{M}_1 \mathbf{D}^{(k-1)}(R'), \\ \mathbf{D}^{(k-1)}(R') \mathbf{M}_2 &= \mathbf{M}_2 \mathbf{D}^{(k)}(R'), \\ \mathbf{D}^{(k)}(R') \mathbf{M}_3 &= \mathbf{M}_3 \mathbf{D}^{(k-1)}(R'), \\ \mathbf{D}^{(k)}(R') \mathbf{M}_4 &= \mathbf{M}_4 \mathbf{D}^{(k)}(R'). \end{aligned}$$

Hieraus folgt aber nach den Sätzen 2 und 3, Kap. IX, daß \mathbf{M}_2 und \mathbf{M}_3 Nullmatrizen, \mathbf{M}_1 und \mathbf{M}_4 Vielfache der Einheitsmatrix sein müssen; (14) muß schon wegen der Vertauschbarkeit mit (13) die Gestalt

$$\begin{pmatrix} m_1 \mathbf{1} & \mathbf{0} \\ \mathbf{0} & m_4 \mathbf{1} \end{pmatrix} \quad (14a)$$

haben. Betrachten wir nun eine Permutation R , die s_n nicht mehr ungeändert läßt, sondern in ein anderes, in \bar{h}'_{l_k} auch in erster Potenz vorkommendes s überführt. Zur linearen Darstellung von $\mathbf{P}_R \bar{h}'_{l_k}$

braucht man dann sicher wenigstens eines der $\bar{g}_1, \bar{g}_2, \dots, \bar{g}_{r_{k-1}}$. Schreibt man daher $D^{(k)}(R)$ wieder in der Form

$$D^{(k)}(R) = \begin{pmatrix} A & B \\ C & D \end{pmatrix}, \quad (13a)$$

so ist B sicher keine Nullmatrix, und (14a) kann nur mit (13a) vertauschbar sein, wenn $m_1 = m_4$, (14) eine konstante Matrix ist. Dies ist aber die Bedingung für die Irreduzibilität von $D^{(k)}(R)$, die wir mithin bewiesen haben.

Es wurde bisher immer angenommen, daß, sobald $k \leq \frac{1}{2}n$ ist, außer ' $D^{(k-1)}(R')$ noch etwas in ' $D^{(k)}(R')$ enthalten ist, und dann gezeigt, daß dies nur $D^{(k)}(R')$ sein kann. Nun ist das erstere nur für $k \leq \frac{1}{2}(n-1)$ sicher, weil dann $l'_{k-1} < l'_k$ ist. Der Fall $k = \frac{1}{2}n$ muß noch gesondert betrachtet werden. In diesem Falle enthält $D^{(k)}(R)$ als Darstellung der Untergruppe, die s_n unverändert läßt, zwar auch noch ' $D^{(k-1)}(R')$, aber es enthält auch nichts weiter. Die Dimension von $D^{(k)}(R)$ ist nämlich gleich der Dimension von ' $D^{(k-1)}(R')$:

$$l_k = \binom{n}{\frac{1}{2}n} - \binom{n}{\frac{1}{2}n-1} = \binom{n-1}{\frac{1}{2}n-1} - \binom{n-1}{\frac{1}{2}n-2} = l'_{k-1},$$

weil

$$\begin{aligned} \binom{n}{\frac{1}{2}n} - \binom{n-1}{\frac{1}{2}n-1} &= \binom{n-1}{\frac{1}{2}n} \\ &= \binom{n-1}{\frac{1}{2}n-1} = \binom{n}{\frac{1}{2}n-1} - \binom{n-1}{\frac{1}{2}n-2} \end{aligned}$$

ist. $D^{(1/2n)}(R)$ ist irreduzibel, weil schon die Matrizen, die zur Untergruppe, die s_n unverändert läßt, gehören, irreduzibel sind.

Daß die Darstellungen $D^{(0)}, D^{(1)}, \dots, D^{(1/2n)}$ bzw. $D^{(1/2n-1/2)}$ alle verschieden sind, sieht man aus (13): schon die Matrizen, die Permutationen R' der s_1, s_2, \dots, s_{n-1} entsprechen, sind in allen diesen Darstellungen inäquivalent.

6. Wir wollen noch den Charakter $\chi^{(k)}(R)$ der irreduziblen Darstellung $D^{(k)}(R)$ berechnen. Nachdem $\Delta^{(k)}(R)$ so transformiert werden kann, daß

$$\bar{\Delta}^{(k)}(R) = \begin{pmatrix} \Delta^{(k-1)}(R) & \mathbf{0} \\ \mathbf{0} & D^{(k)}(R) \end{pmatrix} \quad (8)$$

wird, ist $\chi^{(k)}(R)$ gleich der Differenz der Charaktere von $A^{(k)}(R)$ und $A^{(k-1)}(R)$. Der Charakter von $A^{(k)}(R)$ ist nach (4) gleich dem Koeffizienten von x^k im Polynom ($k \leqq \frac{1}{2}n$):

$$(1 + x^{\lambda_1})(1 + x^{\lambda_2}) \dots (1 + x^{\lambda_q}) \quad (\lambda_1 + \lambda_2 + \dots + \lambda_q = n), \quad (4)$$

und der von $A^{(k-1)}(R)$ ist gleich dem Koeffizienten von x^{k-1} in diesem Ausdruck, oder dem von x^k im x -fachen von (4); $\chi^{(k)}(R)$ ist die Differenz dieser beiden Koeffizienten oder gleich dem Koeffizienten von x^k in der Differenz der beiden Ausdrücke, in

$$(1 - x)(1 + x^{\lambda_1})(1 + x^{\lambda_2}) \dots (1 + x^{\lambda_q}) = \sum_k x^k \chi^{(k)}(R), \quad (15)$$

wo $\lambda_1, \lambda_2, \dots, \lambda_q$ die Zyklenlängen von R sind.

Für die assozierte Darstellung $\bar{D}^{(k)}(R)$ gelten dieselben Ausdrücke mit demselben bzw. entgegengesetzten Vorzeichen, je nachdem R eine gerade oder ungerade Permutation, d. h. je nachdem $\lambda_1 - 1 + \lambda_2 - 1 + \dots + \lambda_q - 1 = n - q$ gerade oder ungerade ist: $\bar{\chi}^{(k)}(R)$ ist der Koeffizient von x^k in

$$(-1)^{n-q}(1 - x)(1 + x^{\lambda_1})(1 + x^{\lambda_2}) \dots (1 + x^{\lambda_q}) = \sum_k x^k \bar{\chi}^{(k)}(R). \quad (15a)$$

$D^{(0)}(R)$ ist die identische, $\bar{D}^{(0)}(R)$ die atisymmetrische Darstellung.

Wir haben auf diese Weise, wie schon am Anfang der Ableitung betont wurde, nicht alle irreduziblen Darstellungen der symmetrischen Gruppe bestimmt, sondern nur jene, die in der Spektroskopie eine Rolle spielen. In der mathematischen Theorie der irreduziblen Darstellungen — wir verdanken sie in erster Linie A. Young und G. Frobenius — ordnet man die einzelnen Darstellungen nicht einfachen Indizes k , sondern den verschiedenen Zerlegungen der Zahl n in positive ganzzählige Summanden zu, deren Anzahl ja gleich der aller irreduziblen Darstellungen ist. Der Darstellung $D^{(k)}(R)$ entspricht dabei die Zerlegung $(n - k) + k$ (wo wegen der Beschränkung $n - k \geqq k$ wieder $k \leqq \frac{1}{2}n$ ist), der Darstellung $\bar{D}^{(k)}(R)$ die partitio $2 + 2 + \dots + 2 + 1 + 1 + \dots + 1$ aus k Zweieren und $n - 2k$ Einsern.

Daß sich bei der Vertauschung der Koordinaten der Elektronen alle in der Natur vorkommenden Eigenfunktionen nach diesen Darstellungen transformieren, hängt damit zusammen, daß sich die Elektronen in einem äußeren Magnetfeld nur in zwei verschiedenen Richtungen einquarteln können. Wenn drei Richtungen möglich sind (wie das z. B. bei dem Stickstoffkern der Fall ist), kommen auch die Darstellungen, die Zerlegungen von n in drei Summanden entsprechen, bzw. ihre assoziierten, die Zerlegungen von n in lauter 1, 2, 3 entsprechen, vor. Ist umgekehrt (wie z. B. bei dem He-Kern) nur eine Einstellungsmöglichkeit da, so kommt nur

die symmetrische Darstellung ($n = n$) und die antisymmetrische ($n = 1 + 1 + \dots + 1$) in Frage.

Man vergleiche die Resultate dieses Kapitels mit den irreduziblen Darstellungen (\dagger), (\ddagger), ($\ddot{\dagger}$) S. 89 bis 90 der symmetrischen Gruppe dritten Grades! Die Darstellung durch lauter (1), die identische ist $D^{(0)}(R)$, ihre assozierte, die antisymmetrische ist $\bar{D}^{(0)}(R)$, die zweiten Grades ist $D^{(1)}(R)$, und in diesem Falle ist $\bar{D}^{(1)}(R)$ dieser äquivalent.

Für die symmetrische Gruppe vierten Grades haben wir

$$D^{(0)}(R) \text{ und } \bar{D}^{(0)}(R) \quad \text{von den Dimensionen } \binom{4}{0} - \binom{4}{-1} = 1,$$

$$D^{(1)}(R) \quad " \quad \bar{D}^{(1)}(R) \quad " \quad " \quad " \quad \binom{4}{1} - \binom{4}{0} = 3,$$

$$D^{(2)}(R) \text{ äquivalent } \bar{D}^{(2)}(R) \quad " \quad " \quad " \quad \binom{4}{2} - \binom{4}{1} = 2,$$

das sind im ganzen fünf inäquivalente irreduzible Darstellungen, den fünf Klassen entsprechend (es ist auch $2 \cdot 1^2 + 2 \cdot 3^2 + 2^2 = 24 = 4!$ also wieder alle. Für $n = 5$ erhalten wir auf diese Weise sechs irreduzible Darstellungen [$D^{(2)}(R)$ ist $\bar{D}^{(2)}(R)$ nicht äquivalent], da aber bereits sieben Klassen da sind, nicht mehr alle. Für noch größere n wird ein immer kleinerer Bruchteil der Darstellungen erfaßt. Trotzdem werden wir mit diesen auskommen, da die anderen in der Spektraltheorie der Atome wegen des Pauliprinzips keine Rolle spielen. Man gewinnt sie ebenso, wie wir diese gewonnen haben, wenn man die Transformationseigenschaften von Funktionen von n Variablen untersucht, die mehr als zwei Werte annehmen dürfen, während unsere s auf die zwei Werte -1 und $+1$ beschränkt waren.

Anhang. Es soll noch gezeigt werden, daß für $k \leq \frac{1}{2}n$ die $\binom{n}{k-1}$ Funktionen

$$F_{a_1 a_2 \dots a_{k-1}} = s_{a_1} s_{a_2} \dots s_{a_{k-1}} s_{a_1 a_2 \dots a_{k-1}} \quad (5)$$

($s_{a_1 a_2 \dots a_{k-1}}$ ist die Summe aller jener s , deren Indizes unter den Zahlen a_1, \dots, a_{k-1} nicht vorkommen) linear unabhängig sind. Nur wenn dies feststeht, können wir nämlich schließen, daß nur $\binom{n}{k} - \binom{n}{k-1}$ Linearkombinationen der $s_{a_1} s_{a_2} \dots s_{a_k}$ auf allen $F_{a_1 \dots a_{k-1}}$ senkrecht stehen. Die $F_{a_1 \dots a_{k-1}}$ sind Linearkombinationen der $s_{a_1} s_{a_2} \dots s_{a_k}$

$$F_{a_1 a_2 \dots a_{k-1}} = \sum_b m_{a_1 \dots a_{k-1}; b_1 \dots b_k} s_{b_1} s_{b_2} \dots s_{b_k}, \quad (a)$$

wo

$$m_{a_1 \dots a_{k-1}; b_1 \dots b_k} = \begin{cases} 1, & \text{wenn die } a_1, a_2, \dots, a_{k-1} \\ & \text{unter den } b_1, b_2, \dots, b_k \\ & \text{vorkommen,} \\ 0 & \text{sonst.} \end{cases} \quad (b)$$

In der Summe in (a) durchlaufen die b_1, b_2, \dots, b_k alle $\binom{n}{k}$ Kombinationen der Zahlen 1, 2, ..., n . Wenn unter den $F_{a_1 \dots a_{k-1}}$ eine lineare Beziehung

$$\sum_a c_{a_1 \dots a_{k-1}} F_{a_1 \dots a_{k-1}} = \sum_{a, b} c_{a_1 \dots a_{k-1}} m_{a_1 \dots a_{k-1}; b_1 \dots b_k} s_{b_1 \dots b_k} = 0 \quad (\text{c})$$

bestehen würde (die Summation ist wieder über die $\binom{n}{k-1}$ Kombinationen der a bzw. über die $\binom{n}{k}$ Kombinationen der b zu erstrecken), so würde für alle Zahlen $x_{b_1 \dots b_k}$

$$\sum_{a, b} c_{a_1 \dots a_{k-1}} m_{a_1 \dots a_{k-1}; b_1 \dots b_k} x_{b_1 \dots b_k} = 0 \quad (\text{d})$$

folgen, wie man erkennt, wenn das skalare Produkt von (c) mit $s_{b_1} s_{b_2} \dots s_{b_k}$ mit $x_{b_1 \dots b_k}$ multipliziert und die so entstandenen Gleichungen für alle Kombinationen der b addiert.

Wir wollen jetzt die $x_{b_1 \dots b_k}$ so wählen, daß

$$\sum_b m_{a_1 \dots a_{k-1}; b_1 \dots b_k} x_{b_1 \dots b_k} = \begin{cases} 1 & \text{für } a_1 = 1, \dots, a_{k-1} = k-1, \\ 0 & \text{sonst,} \end{cases} \quad (\text{e})$$

dann ist (d) mit

$$c_{12 \dots k-1} = 0 \quad (\text{f})$$

gleichbedeutend, und dies muß auch für alle anderen $c_{a_1 \dots a_{k-1}}$ gelten, da sie ja alle gleichberechtigt sind. Damit wäre die lineare Unabhängigkeit der $F_{a_1 \dots a_{k-1}}$ bewiesen.

Über die $x_{b_1 \dots b_k}$ können wir frei verfügen. Die $x_{b_1 \dots b_k}$, von deren Indizes b_1, \dots, b_k genau τ unter den Zahlen 1, 2, ..., $k-1$ vorkommen ($k-\tau$ dagegen größer als $k-1$ sind), wählen wir gleich groß und bezeichnen sie mit x_τ . Nun betrachten wir diejenigen Gleichungen (e), bei denen unter den Zahlen a_1, a_2, \dots, a_{k-1} genau σ unter den Zahlen 1, 2, ..., $k-1$ vorkommen. Da $m_{a_1 \dots a_{k-1}; b_1 \dots b_k}$ nur dann von Null verschieden ist, wenn die a unter den b vorkommen, kommen für (e) nur jene Glieder in Betracht, bei denen auch unter den letzteren σ unter den 1, 2, ..., $k-1$ und $k-1-\sigma$ unter den $k, k+1, \dots, n$ sind. Der einzige Index, dessen Zugehörigkeit noch nicht bestimmt ist, kann entweder unter den ersten oder unter den letzten Zahlen sein. Im ersteren

Falle kann er $k - 1 - \sigma$, im letzteren $n - k + 1 - (k - 1 - \sigma) = n - 2k + 2 + \sigma$ Werte annehmen, da er keinem der a_1, a_2, \dots, a_{k-1} gleich sein kann. So geht (e) in

$$\left. \begin{aligned} (k-1-\sigma)x_{\sigma+1} + (n-2k+2+\sigma)x_\sigma &= 1 \quad (\text{für } \sigma = k-1), \\ (k-1-\sigma)x_{\sigma+1} + (n-2k+2+\sigma)x_\sigma &= 0 \quad (\text{für } \sigma = 0, 1, \dots, k-2) \end{aligned} \right\} \quad (g)$$

über. Dies ergibt $x_{k-1} = \frac{1}{(n-k+1)}$ und

$$-\frac{x_\sigma}{x_{\sigma+1}} = \frac{k-1-\sigma}{n-2k+2+\sigma} \quad (\text{für } \sigma = 0, 1, \dots, k-2);$$

diese Gleichungen lassen sich aber für $n-2k+2 > 0$ oder $k < \frac{1}{2}n+1$ und a fortiori für $k \leq \frac{1}{2}n$ alle erfüllen.

XIV. Die Drehgruppen

1. Man nennt die kontinuierliche Gruppe, die aus der Gesamtheit aller reellen, orthogonalen n -dimensionalen Matrizen gebildet wird, die n -dimensionale Drehgruppe. Die reine Drehgruppe umfaßt nur die orthogonalen Matrizen mit der Determinante 1, während die Drehspiegelungsgruppe auch die mit der Determinante -1, also alle reellen orthogonalen Matrizen enthält. Die Gruppenmultiplikation ist wieder die Matrixmultiplikation, und die Einheit ist die Einheitsmatrix.

Im dritten Kapitel haben wir gesehen, daß jede reelle orthogonale Matrix durch eine unitäre Matrix auf die Diagonalform gebracht werden kann. Die Diagonalelemente haben dann alle den Absolutwert 1, einige sind +1, andere -1, der Rest besteht aus paarweise konjugiert komplexen Zahlen $e^{i\varphi}$ und $e^{-i\varphi}$. Die Eigenvektoren, die zu +1 oder -1 gehören, können reell geschrieben werden, diejenigen, die zu zwei konjugiert komplexen Eigenwerten gehören, konjugiert komplex. Da diese Eigenvektoren, wie alle, aufeinander im Hermiteschen Sinne senkrecht stehen, sind sie im komplex orthogonalen Sinne auf sich selber orthogonal: die Summe der Quadrate ihrer Komponenten ist Null.

Die n -dimensionale orthogonale Matrix bedeutet einen Übergang von einem rechtwinkligen Achsenkreuz zu einem anderen, eine Verdrehung des Achsenkreuzes. Die Orthogonalität der Matrix be-

deutet, daß je zwei Achsen des neuen Koordinatensystems aufeinander senkrecht stehen, und daß die Längenmessung längs der neuen Achsen mit dem ungeänderten alten Maßstab geschieht. Die reine Drehgruppe enthält nur Übergänge von einem „rechtshändigen Achsenkreuz“ zu einem anderen „rechtshändigen“, die Drehspiegelungsgruppe auch von einem rechtshändigen zu einem linkshändigen und umgekehrt.

Wir müssen, um unsere allgemeinen Resultate über kontinuierliche Gruppen anwenden zu können, zunächst Parameter einführen. Dies kann nur in unsymmetrischer Weise geschehen, indem man nicht nur Raumrichtungen (die Koordinatenachsen), sondern sogar unter diesen Koordinatenachsen einige auszeichnen muß. Zunächst bestimmen wir ihre Anzahl, d. h. die Dimension des Raumes, in dem der Variabilitätsbereich der Parameter abgegrenzt werden muß. Betrachten wie eine n -dimensionale reell-orthogonale Matrix! Die erste Zeile ist ein n -dimensionaler Vektor mit der Länge 1 (die X-Achse des neuen Systems), das sind — wegen der letzten Beschränkung $a_{11}^2 + a_{12}^2 + \dots + a_{1n}^2 = 1$ — genau $n - 1$ Parameter. Die zweite Zeile (die Y-Achse) muß auf der ersten senkrecht stehen — das ist eine homogene lineare Gleichung $a_{11}a_{21} + a_{12}a_{22} + \dots + a_{1n}a_{2n} = 0$ für die $a_{21}, a_{22}, \dots, a_{2n}$ — und die Länge 1 haben $a_{21}^2 + a_{22}^2 + \dots + a_{2n}^2 = 1$ — das sind noch $n - 2$ Parameter. Die k -Zeile muß auf $k - 1$ vorangehende Zeilen senkrecht stehen — das sind $k - 1$ homogene lineare Gleichungen — und die Länge 1 haben. Es bleiben für die $a_{k1}, a_{k2}, \dots, a_{kn}$ insgesamt $n - k$ Parameter. Im ganzen haben wir

$$(n - 1) + (n - 2) + (n - 3) + \dots + n - (n - 1) + 0 = \frac{1}{2}n(n - 1)$$

freie Parameter.

2. Im folgenden werden wir uns auf die zwei- und dreidimensionalen Drehgruppen beschränken.

Das allgemeine Element der zweidimensionalen reinen Drehgruppe erhält man durch einen Übergang zu einem neuen Koordinatensystem in der Ebene. Dabei ist¹⁾

$$\left. \begin{aligned} x' &= x \cos \varphi - y \sin \varphi, \\ y' &= x \sin \varphi + y \cos \varphi, \end{aligned} \right\} \quad (1)$$

¹⁾ Der Drehungssinn ist so gewählt, daß die Y-Achse zur X-Achse gedreht wird, weil dies bei der dreidimensionalen Drehgruppe so üblich ist.

wo φ , der Drehwinkel von $-\pi$ bis π , variiert. Das allgemeine Element ist also

$$\begin{pmatrix} \cos \varphi & -\sin \varphi \\ \sin \varphi & \cos \varphi \end{pmatrix}. \quad (2)$$

Geht man mit Hilfe einer weiteren Drehung des Achsenkreuzes um φ' von x', y' zu x'', y'' über, so erhält man das Produkt

$$\left. \begin{aligned} & \begin{pmatrix} \cos \varphi' & -\sin \varphi' \\ \sin \varphi' & \cos \varphi' \end{pmatrix} \begin{pmatrix} \cos \varphi & -\sin \varphi \\ \sin \varphi & \cos \varphi \end{pmatrix} \\ & = \begin{pmatrix} \cos(\varphi + \varphi') & -\sin(\varphi + \varphi') \\ \sin(\varphi + \varphi') & \cos(\varphi + \varphi') \end{pmatrix}, \end{aligned} \right\} \quad (3)$$

eine Drehung mit dem Winkel $\varphi + \varphi'$, wie man aus (3) auch unmittelbar verifiziert.

Die zweidimensionale reine Drehgruppe ist, da sie nur einen Parameter hat, abelsch. Führen wir die Bezeichnung $\{\varphi\}$ des Kapitels X für das Gruppenelement mit dem Parameter φ ein, so lautet (3)

$$\{\varphi'\} \{\varphi\} = \{\varphi' + \varphi\} = \{\varphi\} \{\varphi'\}, \quad (4)$$

wo aber, wenn $\varphi' + \varphi$ nicht zwischen $-\pi$ und π liegt, 2π zu addieren bzw. zu subtrahieren ist, so daß es in sein richtiges Variabilitätsbereich kommt.

Für die Matrix (2) ist φ die komplexe Phase der Eigenwerte $e^{\pm i\varphi}$. Die Eigenvektoren, die Spalten der unitären Matrix u , die (2) auf Diagonalform bringt, bestimmen sich aus

$$|u_{1\alpha}|^2 + |u_{2\alpha}|^2 = 1; u_{1\alpha}^2 + u_{2\alpha}^2 = 0; u_{1\alpha} = \pm i u_{2\alpha}$$

bis auf einen Faktor vom Betrag 1, der aber willkürlich gewählt werden kann, zu $u_{11} = \frac{-i}{\sqrt{2}}$; $u_{21} = \frac{1}{\sqrt{2}}$; $u_{12} = \frac{i}{\sqrt{2}}$; $u_{22} = \frac{1}{\sqrt{2}}$,

und es ist

$$\begin{pmatrix} i/\sqrt{2} & 1/\sqrt{2} \\ -i/\sqrt{2} & 1/\sqrt{2} \end{pmatrix} \begin{pmatrix} \cos \varphi & -\sin \varphi \\ \sin \varphi & \cos \varphi \end{pmatrix} \begin{pmatrix} -i/\sqrt{2} & i/\sqrt{2} \\ 1/\sqrt{2} & 1/\sqrt{2} \end{pmatrix} = \begin{pmatrix} e^{-i\varphi} & 0 \\ 0 & e^{i\varphi} \end{pmatrix}. \quad (5)$$

Die Eigenvektoren sind also für alle Matrizen (2) dieselben. Da die reine zweidimensionale Drehgruppe abelsch ist, bildet jedes Element von (2) eine Klasse für sich.

3. Aus jeder zweidimensionalen Matrix mit der Determinante -1 erhält man eine Matrix mit der Determinante 1, wenn man die erste

Zeile mit -1 multipliziert. Umgekehrt erhält man also die allgemeine orthogonale Matrix mit der Determinante -1 , indem man in der ersten Zeile von (2) die Vorzeichen umkehrt:

$$\begin{pmatrix} -\cos \varphi & \sin \varphi \\ \sin \varphi & \cos \varphi \end{pmatrix}. \quad (2a)$$

Die Matrizen (2) und (2a) für $\varphi = -\pi$ bis π bilden die zweidimensionale Drehspiegelungsgruppe. Die Matrizen (2a) haben alle die Eigenwerte 1 und -1 , sie unterscheiden sich durch ihre Eigenvektoren $\mathbf{u}_{.1} = (\cos \frac{\varphi}{2}, -\sin \frac{\varphi}{2})$ und $\mathbf{u}_{.2} = (\sin \frac{\varphi}{2}, \cos \frac{\varphi}{2})$ von einander, während (2) alle dieselben Eigenvektoren und verschiedene Eigenwerte hatten. Es ist nämlich

$$\begin{pmatrix} -\cos \varphi & \sin \varphi \\ \sin \varphi & \cos \varphi \end{pmatrix} = \begin{pmatrix} \cos \frac{\varphi}{2} & \sin \frac{\varphi}{2} \\ -\sin \frac{\varphi}{2} & \cos \frac{\varphi}{2} \end{pmatrix} \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \cos \frac{\varphi}{2} & -\sin \frac{\varphi}{2} \\ \sin \frac{\varphi}{2} & \cos \frac{\varphi}{2} \end{pmatrix}, \quad (5a)$$

was der Tatsache entspricht, daß jede Drehspiegelung (2a) als eine reine Spiegelung an einer Geraden aufgefaßt werden kann: (5a)

bedeutet, daß man (2a) erhält, wenn man zuerst um $\frac{\varphi}{2}$ dreht,

dann an der Y -Achse spiegelt, dann mit $\frac{\varphi}{2}$ zurückdreht. Statt dessen hätte man an einer Geraden, die mit der Y -Achse den Winkel $-\frac{\varphi}{2}$ einschließt, spiegeln können.

Die zweidimensionale Drehspiegelungsgruppe ist eine gemischtkontinuierliche Gruppe. Man kann für sie am besten in der Weise Parameter einführen, daß man einen kontinuierlichen Parameter φ und einen diskreten d nimmt, der letzte ist gleich der Determinante, also ± 1 . Es ist dann

$$\{\varphi, d\} = \begin{pmatrix} d \cos \varphi & -d \sin \varphi \\ \sin \varphi & \cos \varphi \end{pmatrix}, \quad (6)$$

$$\{\varphi, d\} \{\varphi', d'\} = \{d' \varphi + \varphi', dd'\}.$$

Die Gruppe ist also nicht mehr abelsch, die Matrizen (2a) sind nicht mehr vertauschbar, sie haben ja auch nicht dieselben Eigenvektoren.

Auch die Klasseneinteilung ändert sich: (5 a) zeigt, daß alle Elemente (2 a) in einer Klasse sind, sie lassen sich nämlich alle in $\{0, -1\}$ transformieren. Aber auch die Elemente von (2) bilden nicht mehr je eine Klasse für sich:

$$\begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \cos \varphi & \sin \varphi \\ -\sin \varphi & \cos \varphi \end{pmatrix} \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} \cos \varphi & -\sin \varphi \\ \sin \varphi & \cos \varphi \end{pmatrix}. \quad (7)$$

$\{\varphi, 1\}$ und $\{-\varphi, 1\}$ sind zusammen in einer Klasse. Andere Elemente können dagegen nicht mehr in dieser Klasse sein, da diese andere Eigenwerte haben und so nicht mehr in diese transformiert werden können.

4. Nach den allgemeinen Entwicklungen über das Hurwitzsche Integralkalkül existiert im Bereich der zweidimensionalen reinen Drehgruppe ein invariantes Integral, so daß

$$\int_{-\pi}^{\pi} J(\{\varphi\}) g(\{\varphi\}) d\varphi = \int_{-\pi}^{\pi} J(R\{\varphi\}) g(\{\varphi\}) d\varphi \quad (8).$$

für alle Gruppenelemente R gilt, wenn $g(T)$ durch

$$g(T) = \frac{g(E)}{\frac{\partial p(T \cdot \{\alpha\})}{\partial \alpha}}, \quad \text{für } \alpha = 0, \quad (9)$$

definiert ist und $p(T)$ der Parameter des Elementes T ist.

Im Falle der zweidimensionalen Drehgruppe ergibt die unmittelbare Anschauung, daß gleiche Bereiche von φ gleiches Gewicht haben müssen. In der Tat sei t der Parameter von T , dann ist nach (4) der Parameter $p(T \cdot \{\alpha\}) = t + \alpha$, und dies nach α differenziert ergibt 1, so daß

$$g(T) = g(E) \quad (10)$$

ist. Das invariante Integral lautet also

$$\int_{-\pi}^{\pi} J(\{\varphi\}) d\varphi = \int_{-\pi}^{\pi} J(R \cdot \{\varphi\}) d\varphi. \quad (11)$$

Die irreduziblen Darstellungen der zweidimensionalen reinen Drehgruppe sind alle eindimensional. Dies gilt nämlich für alle, auch für kontinuierliche abelsche Gruppen. Betrachten wir eine mehrdimensionale — etwa zweidimensionale — Darstellung. Wir

können irgendeine Matrix der Darstellung auf Diagonalform bringen. Würde sie zwei ungleiche Diagonalelemente, also etwa die Gestalt

$$\begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix} \quad (*)$$

haben, so müßten alle mit ihr vertauschbaren Matrizen — also alle Matrizen der Darstellung — an den Kreuzungspunkten der verschiedenen Diagonalelemente von (*) lauter Nullen haben, und die Darstellung wäre reduzibel. Soll dies nicht der Fall sein, so müssen die Eigenwerte der ersten Matrix alle gleich und diese eine konstante Matrix sein. Sie hätte, dann schon vor der Transformation die Diagonalform gehabt. Da dies aber für jede Matrix der Darstellung gilt, wären sie alle Vielfache der Einheitsmatrix und die Darstellung a fortiori reduzibel.

Aus (4) folgt, daß wenn in einer Darstellung dem Element $\{\varphi\}$ die Matrix $(f(\varphi))$ entspricht, so muß

$$(f(\varphi)) (f(\varphi')) = (f(\varphi) \cdot f(\varphi')) = (f(\varphi + \varphi')),$$

also

$$f(\varphi) = e^{ik\varphi}$$

sein. Da aber die Matrix für $\varphi = -\pi$ der Matrix für $\varphi = \pi$ gleich sein muß, muß $e^{ik\pi} = e^{-ik\pi}$; $e^{2\pi ik} = 1$ sein. Daraus folgt, daß k eine reelle ganze Zahl ist. Die zweidimensionale reine Drehgruppe hat unendlich viele irreduzible Darstellungen, alle sind von der Dimension 1, die Matrix, die in der m -ten Darstellung dem Element (2) mit dem Drehwinkel φ zugeordnet ist, lautet

$$(e^{im\varphi}).$$

Für alle positiven und negativen ganzzahligen $m = \dots - 4, -3, -2, -1, 0, 1, 2, 3, 4, \dots$ erhält man je eine irreduzible Darstellung.

Die Orthogonalitätsrelationen

$$\int_{-\pi}^{\pi} (e^{im'\varphi})^* e^{im\varphi} d\varphi = 0 \quad \text{für } m \neq m'$$

$$= \int_{-\pi}^{\pi} d\varphi = 2\pi \quad \text{für } m = m'$$

sind die Orthogonalitätsrelationen der Fourierreihe. Die Vollständigkeit der Darstellungskoeffizienten ist auch die Vollständigkeitsrelation der Fourierentwicklung.

5. Die irreduziblen Darstellungen der zweidimensionalen Drehspiegelungsgruppe bestimmen wir hier nach einer Methode, die hierfür zwar etwas kompliziert erscheinen muß, die wir aber später bei der dreidimensionalen Gruppe noch verwenden wollen und die wieder ein Beispiel für den Zusammenhang der Darstellungen mit den zugehörigen Funktionen liefert. Wir betrachten die Gleichung der harmonischen Polynome zweier Variablen

$$\frac{\partial^2 f(x, y)}{\partial x^2} + \frac{\partial^2 f(x, y)}{\partial y^2} = 0, \quad (12)$$

die offenbar invariant allen Transformationen (6) gegenüber ist. Weiter bemerken wir, daß eine Lösung von (12), die etwa homogen vom Grade m in x und y ist, bei Anwendung eines Operators \mathbf{P}_R [mit einem R aus (6)] wieder in ein solches Polynom übergeht, da ja \mathbf{P}_R eine lineare Transformation der Variablen x und y bedeutet:

$\mathbf{P}_{\{\varphi, d\}} f(x d \cos \varphi - y d \sin \varphi, x \sin \varphi + y \cos \varphi) = f(x, y)$, (13)
oder

$\mathbf{P}_{\{\varphi, d\}} f(x, y) = f(x d \cos \varphi + y \sin \varphi, -x d \sin \varphi + y \cos \varphi)$, (14)
so daß auch $\mathbf{P}_R f$ homogen vom Grade m ist, wenn dies für f gilt.

Nun ist (12) nichts anderes als die eindimensionale Wellengleichung mit der imaginären Geschwindigkeit i . Ihre allgemeine Lösung ist

$$f(x, y) = f_-(y - ix) + f_+(y + ix). \quad (15)$$

Soll $f(x, y)$ homogen vom Grade m in x und y sein, so muß bis auf eine Konstante

$$f_-(y - ix) = (y - ix)^m; \quad f_+(y + ix) = (y + ix)^m \quad (16)$$

sein. Die zu diesen Funktionen gehörige Darstellung $\mathfrak{Z}^{(m)}(\{\varphi, d\})$ ist zweidimensional, ihre erste (—)-Spalte bestimmt sich aus (23) Kapitel XI und (14)

$$\begin{aligned} \mathbf{P}_{\{\varphi, d\}} f_-(x, y) &= f_-(x d \cos \varphi + y \sin \varphi, -x d \sin \varphi + y \cos \varphi) \\ &= [-x d \sin \varphi + y \cos \varphi - i(x d \cos \varphi + y \sin \varphi)]^m \\ &= [y(\cos \varphi - i \sin \varphi) - i x d (\cos \varphi - i \sin \varphi)]^m \\ &= (y - i dx)^m e^{-im\varphi} \\ &= \mathfrak{Z}^{(m)}(\{\varphi, d\})_{--} f_- + \mathfrak{Z}^{(m)}(\{\varphi, d\})_{+-} f_+ \end{aligned}$$

zu

$$\left. \begin{aligned} \mathfrak{Z}^{(m)}(\{\varphi, 1\})_{--} &= e^{-im\varphi}; & \mathfrak{Z}^{(m)}(\{\varphi, 1\})_{+-} &= 0, \\ \mathfrak{Z}^{(m)}(\{\varphi, -1\})_{--} &= 0; & \mathfrak{Z}^{(m)}(\{\varphi, -1\})_{+-} &= e^{-im\varphi}. \end{aligned} \right\} (17)$$

Entsprechend kann man auch die andere (+)-Spalte bestimmen. Schließlich lautet die Matrix $\mathfrak{Z}^{(m)}(\{\varphi, 1\})$, die in dieser Darstellung einer reinen Drehung mit φ zugeordnet ist, wenn man die -- Zeile bzw. Spalte zur ersten und die + -Zeile bzw. Spalte zur zweiten macht:

$$\mathfrak{Z}^{(m)}(\{\varphi, 1\}) = \begin{pmatrix} e^{-im\varphi} & 0 \\ 0 & e^{im\varphi} \end{pmatrix}, \quad (18)$$

und die Matrix, die dem Gruppenelement, das in (2a) steht, zugeordnet ist, lautet

$$\mathfrak{Z}^{(m)}(\{\varphi, -1\}) = \begin{pmatrix} 0 & e^{im\varphi} \\ e^{-im\varphi} & 0 \end{pmatrix}. \quad (18a)$$

Die Funktion f_- gehört zur -- (ersten), die Funktion f_+ zur + (zweiten) Zeile von $\mathfrak{Z}^{(m)}$.

Diese Darstellungen sind für $m = 1, 2, 3, \dots$ voneinander verschiedene irreduzible Darstellungen, da mit (18) nur eine Diagonalmatrix, mit (18a) aber -- abgesehen von einer konstanten Matrix -- keine Diagonalmatrix vertauschbar ist. Die Matrizen (6) sind natürlich auch eine „Darstellung“ ihrer eigenen Gruppe, diese ist mit der Darstellung (18), (18a) für $m = 1$ äquivalent und lässt sich mit Hilfe der Matrix in (5) in diese transformieren.

Für $m = 0$ dagegen ist (18), (18a) zwar auch eine Darstellung, jedoch keine irreduzible, da mit (18) dann jede Matrix vertauschbar ist. Bringt man die Matrix (18a), das ist jetzt $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$, auf Diagonalform, so zerfällt die Darstellung in zwei irreduzible Bestandteile

$$\mathfrak{Z}^{(0)}(\{\varphi, 1\}) = (1); \quad \mathfrak{Z}^{(0)}(\{\varphi, -1\}) = (1) \quad (19)$$

und

$$\mathfrak{Z}^{(0')}(\{\varphi, 1\}) = (1); \quad \mathfrak{Z}^{(0')}(\{\varphi, -1\}) = (-1). \quad (20)$$

6. Hierdurch haben wir alle Darstellungen der zweidimensionalen Drehspiegelungsgruppe gewonnen. Diese sind für $m = 1, 2, 3, 4, \dots$ in (18), (18a) gegeben und von der Dimension 2, für $m = 0$ und $m = 0'$ in (19) und (20) von der Dimension 1.

Die Koeffizienten $\mathfrak{Z}^{(m)}(\{\varphi, d\})_{\pm\pm}$ bilden im Raum von φ und d ein vollständiges Funktionssystem, d. h. man kann jede Funktion $g(\varphi, d)$ (φ variiert von $-\pi$ bis π , d ist $+1$ oder -1) durch sie linear ausdrücken. Die Funktionen $\frac{1}{2}(\mathfrak{Z}^{(0)} + \mathfrak{Z}^{(0')})$, $\mathfrak{Z}_{-+}^{(1)}$, $\mathfrak{Z}_{++}^{(1)}$, $\mathfrak{Z}_{--}^{(2)}$, $\mathfrak{Z}_{+-}^{(2)}, \dots$ sind nämlich der Reihe nach $1, e^{-i\varphi}, e^{i\varphi}, e^{-2i\varphi}, e^{2i\varphi}, \dots$

für $d = 1$ und verschwinden für $d = -1$, dagegen sind $\frac{1}{2}(\mathbf{3}^{(0)} - \mathbf{3}^{(0')})$, $\mathbf{3}_{+}^{(1)}$, $\mathbf{3}_{-}^{(1)}$, $\mathbf{3}_{+}^{(2)}$, $\mathbf{3}_{-}^{(2)}$, ... gleich Null für $d = 1$ und der Reihe nach $1, e^{-i\varphi}, e^{i\varphi}, e^{-2i\varphi}, e^{2i\varphi}, \dots$ für $d = -1$. Durch die erste Reihe läßt sich $g(\varphi, 1)$ durch die zweite $g(\varphi, -1)$ linear ausdrücken.

Es folgt hieraus, daß außer (18), (18a), (19) und (20) keine weiteren irreduziblen Darstellungen der zweidimensionalen Drehspiegelungsgruppe existieren.

7. Wir gehen jetzt zur Untersuchung der dreidimensionalen reinen Drehgruppe über. Die Eigenwerte einer reellen orthogonalen dreidimensionalen Matrix α mit der Determinante 1 müssen die Gestalt $1, e^{i\varphi}, e^{-i\varphi}$ haben, da sie alle vom Absolutwert 1 sind und die komplexen paarweise konjugiert komplex sind. Die Phase φ der komplexen Eigenwerte heißt Drehwinkel, der Eigenvektor $\mathbf{v}_{.1}$ des Eigenwerts 1 heißt Drehachse. Seine Komponenten v_{11}, v_{21}, v_{31} bestimmt man wohl am einfachsten, indem man von $\alpha \mathbf{v}_{.1} = \mathbf{v}_{.1}$ ausgeht und mit $\alpha^{-1} = \alpha'$ multipliziert: $\mathbf{v}_{.1} = \alpha' \mathbf{v}_{.1}$. So ergibt sich $(\alpha - \alpha') \mathbf{v}_{.1} = 0$ oder ausgeschrieben

$$\left. \begin{aligned} (\alpha_{12} - \alpha_{21}) v_{21} + (\alpha_{13} - \alpha_{31}) v_{31} &= 0, \\ (\alpha_{21} - \alpha_{12}) v_{11} + (\alpha_{23} - \alpha_{32}) v_{31} &= 0, \\ (\alpha_{31} - \alpha_{13}) v_{11} + (\alpha_{32} - \alpha_{23}) v_{21} &= 0, \end{aligned} \right\} \quad (21)$$

und hieraus

$$v_{11} : v_{21} : v_{31} = \alpha_{23} - \alpha_{32} : \alpha_{31} - \alpha_{13} : \alpha_{12} - \alpha_{21}. \quad (22)$$

Den Drehwinkel φ bestimmt man am besten, indem man die Summe der Eigenwerte gleich der Spur der Matrix setzt:

$$1 + e^{i\varphi} + e^{-i\varphi} = 1 + 2 \cos \varphi = \alpha_{11} + \alpha_{22} + \alpha_{33}, \quad (23)$$

φ soll zwischen 0 und π liegen.

Die zu $e^{i\varphi}$ und $e^{-i\varphi}$ gehörigen Eigenvektoren $\mathbf{v}_{.2}$ und $\mathbf{v}_{.3}$ sind konjugiert komplex: $\mathbf{v}_{.2}^* = \mathbf{v}_{.3}$, dagegen soll $\mathbf{v}_{.1}$ reell angenommen werden, $(\mathbf{v}_{.1}, \mathbf{v}_{.1}) = ((\mathbf{v}_{.1}, \mathbf{v}_{.1})) = 1$.

Die Matrix \mathbf{v} , deren Spalten die Eigenvektoren $\mathbf{v}_{.1}, \mathbf{v}_{.2}, \mathbf{v}_{.3}$ von α sind, bringt α auf die Diagonalform. Es ist daher $\mathbf{v}^\dagger \alpha \mathbf{v} = \mathbf{d}$ eine Diagonalmatrix mit den Eigenwerten $1, e^{i\varphi}, e^{-i\varphi}$ als Diagonalelementen. Setzen wir noch $\mathbf{B} = \mathbf{v} \mathbf{v}_0$ mit

$$\mathbf{v}_0 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1/\sqrt{2} & i/\sqrt{2} \\ 0 & 1/\sqrt{2} & -i/\sqrt{2} \end{pmatrix}, \quad (24)$$

so sind die Spalten von \mathfrak{B} der Reihe nach $\mathbf{v}_{\cdot 1}, (\mathbf{v}_{\cdot 2} + \mathbf{v}_{\cdot 2}^*)/\sqrt{2}$ und $i(\mathbf{v}_{\cdot 2} - \mathbf{v}_{\cdot 2}^*)/\sqrt{2}$, so daß \mathfrak{B} rein reell ist. Außerdem ist \mathfrak{B} als das Produkt der unitären Matrizen \mathbf{v} und \mathbf{v}_0 auch unitär, so daß es als eine reelle unitäre Matrix ein Element der Drehgruppe ist. Transformieren wir nunmehr die Gleichung $\mathbf{v}^\dagger \alpha \mathbf{v} = \mathbf{d}$ mit \mathbf{v}_0 , so erhalten wir

$$\mathfrak{B}^\dagger \alpha \mathfrak{B} = \mathbf{v}_0^\dagger \mathbf{v}^\dagger \alpha \mathbf{v} \mathbf{v}_0 = \mathbf{v}_0^\dagger \mathbf{d} \mathbf{v}_0 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos \varphi & -\sin \varphi \\ 0 & \sin \varphi & \cos \varphi \end{pmatrix} = \epsilon_\varphi. \quad (25)$$

Hierin können wir sogar \mathfrak{B} als eine reine Drehung annehmen, weil wir es, wenn seine Determinante -1 wäre, mit -1 multiplizieren könnten: (25) würde sich dadurch nicht ändern. Wir sehen aus (25), daß alle Drehungen mit demselben Drehwinkel φ in derselben Klasse sind, weil sie alle in ϵ_φ transformiert werden können. Matrizen mit von φ verschiedenem Drehwinkel können dagegen nicht mehr in dieser Klasse sein, weil sie andere Eigenwerte haben und daher nicht in ϵ_φ transformiert werden können.

Der geometrische Sinn dieser Überlegung ist der bekannte Satz, daß im dreidimensionalen Raum jede orthogonale Transformation durch eine Drehung um eine zweckmäßig gewählte Drehachse $\mathbf{v}_{\cdot 1}$ ersetzt werden kann. (Die Drehachse bleibt wegen $\alpha \mathbf{v}_{\cdot 1} = \mathbf{v}_{\cdot 1}$ bei der Drehung ungeändert.) Bringt eine Transformation den Bogen YZ (Abb. 6) nach $Y'Z'$, so muß die Drehachse in der Mittelsenkrechten von ZZ' und YY' , also in ihrem Schnittpunkt C liegen. In der Tat bringt nun die Drehung um C , die Z in Z' überführt, auch Y nach Y' : aus der Kongruenz der beiden Dreiecke ZCY und $Z'CY'$ (die drei Seiten sind gleich) folgt, daß die Winkel ZCY und $Z'CY'$ gleich sind, und daher sind auch die Winkel ZCZ' und YCY' gleich. Eine Drehung mit dem Drehwinkel φ läßt sich in eine andere Drehung mit demselben Drehwinkel transformieren, indem man durch eine Drehung \mathfrak{B} die eine Drehachse zur Deckung mit der anderen bringt, hierauf die Drehung mit φ vornimmt und dann durch \mathfrak{B}^{-1} die Drehachse auf die alte Stelle bringt.

Zur eindeutigen Charakterisierung der Drehung muß man noch der Drehachse einen Richtungssinn geben, wodurch auch das Vor-

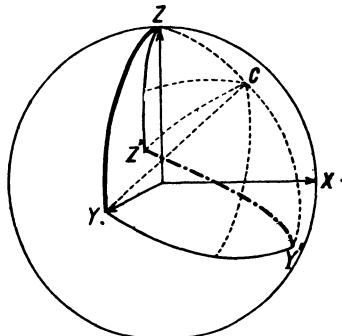


Abb. 6

zeichen des Vektors $v_{.1}$ einen Sinn bekommt. Die Drehung soll, in Richtung der Drehachse gesehen, im Uhrzeigersinn erfolgen.

Auf diesen Tatsachen beruht die im X. Kapitel besprochene Parameterdarstellung (Fig. 1) der dreidimensionalen reinen Drehgruppe. Der Drehung mit dem Drehwinkel φ um die Drehachse $v_{.1}$ entspricht ein Punkt im Abstande¹⁾ φ vom Nullpunkt in der Richtung von $v_{.1}$ von diesem. Der Drehwinkel φ ist durch die Drehung immer eindeutig gegeben. Für die Drehung mit $\varphi = 0$ (die eigentlich keine Drehung ist) ist die Richtung der Drehachse unbestimmt, der entsprechende Punkt des Parameterraumes ist trotzdem eindeutig gegeben: er ist der Mittelpunkt der Kugel.

Auf der Oberfläche der Kugel ist $\varphi = \pi$, und der Sinn der Drehachse ist nicht mehr eindeutig: Drehungen um entgegengesetzte gerichtete Achsen mit dem Drehwinkel π sind äquivalent. Den Antipoden der Kugeloberfläche entspricht daher dieselbe Drehung. Sonst ist die Zuordnung der Drehungen zu Punkten des Parameterraumes überall ein-eindeutig. Elemente gleicher Klassen liegen auf konzentrischen Kugeln.

In diesem Parametersystem können wir auch das Hurwitzsche invariante Integral verhältnismäßig leicht aufstellen. Da die Punkte einer Kugeloberfläche vom Radius φ um den Nullpunkt Drehungen mit gleichem Drehwinkel φ nur um verschiedene gerichtete Drehachsen entsprechen, die Drehachsen im Raum aber gleichberechtigt sind, kann $g(\{\varphi v_{11}, \varphi v_{21}, \varphi v_{31}\})$ nur vom Drehwinkel φ , nicht aber von der Richtung v_{11}, v_{21}, v_{31} abhängen. Es genügt also, $g(\{\varphi, 0, 0\})$ zu bestimmen. Zu diesem Zweck [vgl. (3a) Kap. X.] berechnen wir zuerst $p_x(\{\varphi, 0, 0\} \{e_1, e_2, e_3\})$, und zwar, da wir am Ende e_1, e_2, e_3 doch gegen Null gehen lassen werden, nur für sehr kleine e . In der ersten Potenz genau in e ist [vgl. (22)]

$$\{e_1, e_2, e_3\} = \begin{pmatrix} 1 & -e_3 & e_2 \\ e_3 & 1 & -e_1 \\ -e_2 & e_1 & 1 \end{pmatrix}$$

Wir erhalten für $\{\varphi, 0, 0\} \{e_1, e_2, e_3\} = \epsilon_\varphi \{e_1, e_2, e_3\}$

$$\begin{pmatrix} 1 & -e_3 & e_2 \\ e_3 \cos \varphi + e_2 \sin \varphi & \cos \varphi - e_1 \sin \varphi & -e_1 \cos \varphi - \sin \varphi \\ e_3 \sin \varphi - e_2 \cos \varphi & \sin \varphi + e_1 \cos \varphi & -e_1 \sin \varphi + \cos \varphi \end{pmatrix}.$$

Der Drehwinkel φ' hiervon berechnet sich nach (23)

$$1 + 2 \cos \varphi' = 1 + 2 \cos \varphi - 2 e_1 \sin \varphi \quad \varphi' = \varphi + e_1. \quad (23a)$$

¹⁾ Um die Abb. 1 übersichtlicher zu machen, ist der Abstand dort φ/π , nicht φ . Die linke Hälfte der Figur ist also im Verhältnis $\pi : 1$ vergrößert zu denken.

Aus (22) erhalten wir für die Richtung der Drehachse

$$\begin{aligned} \mathbf{v}'_{11} : \mathbf{v}'_{21} : \mathbf{v}'_{31} &= -2e_1 \cos \varphi - 2 \sin \varphi : e_3 \sin \varphi - e_2 (1 + \cos \varphi) \\ &\quad : -e_3 (1 + \cos \varphi) - e_2 \sin \varphi. \end{aligned} \quad \left. \right\} \quad (22a)$$

Mit Hilfe der Normierung $\mathbf{v}'_{11}^2 + \mathbf{v}'_{21}^2 + \mathbf{v}'_{31}^2 = 1$ ergibt dies, wieder bis zu den ersten Potenzen der e genau

$$\mathbf{v}'_{11} = 1; \quad \mathbf{v}'_{21} = \frac{e_2 (1 + \cos \varphi)}{2 \sin \varphi} - \frac{e_3}{2}; \quad \mathbf{v}'_{31} = \frac{e_2}{2} + \frac{e_3 (1 + \cos \varphi)}{2 \sin \varphi},$$

so daß sich die Parameter von $\{\varphi, 0, 0\} \{e_1 e_2 e_3\}$ zu

$$\varphi + e_1; \quad \varphi \frac{e_2 (1 + \cos \varphi)}{2 \sin \varphi} - \frac{e_3}{2}; \quad \frac{e_2}{2} + \varphi \frac{e_3 (1 + \cos \varphi)}{2 \sin \varphi}$$

berechnen. Für $e_1 = e_2 = e_3 = 0$ erhält man natürlich die Parameter $\varphi, 0, 0$ von \mathbf{e}_φ zurück.

Die fragliche Funktionaldeterminante ist daher für $e_1 = e_2 = e_3 = 0$

$$\begin{aligned} &\frac{\partial(p_1(\{\varphi, 0, 0\} \{e_1 e_2 e_3\}), \dots, p_3(\{\varphi, 0, 0\} \{e_1 e_2 e_3\}))}{\partial(e_1 e_2 e_3)} \\ &= \left| \begin{array}{ccc} 1 & 0 & 0 \\ 0 & \varphi \frac{1 + \cos \varphi}{2 \sin \varphi} & \frac{\varphi}{2} \\ 0 & -\frac{\varphi}{2} & \varphi \frac{1 + \cos \varphi}{2 \sin \varphi} \end{array} \right| = \frac{\varphi^3}{4} \frac{(1 + \cos \varphi)^3 + \sin^2 \varphi}{\sin^2 \varphi} \\ &= \frac{\varphi^3}{2(1 - \cos \varphi)}. \end{aligned}$$

Und so erhalten wir für das Gewicht nach (3a) Kapitel X

$$g(\{\varphi, 0, 0\}) = g(\{\mathbf{v}_{11}\varphi, \mathbf{v}_{21}\varphi, \mathbf{v}_{31}\varphi\}) = \frac{2g(E)(1 - \cos \varphi)}{\varphi^3}. \quad (26)$$

Besonders einfach berechnet sich das invariante Integral über eine Funktion $J(R) = J(\varphi)$, die für alle Elemente einer Klasse denselben Wert hat (wie z. B. der Charakter einer Darstellung). Man kann dann im Raume der Parameter zuerst über eine Kugeloberfläche ($\varphi = \text{Konst.}$), über alle Elemente einer Klasse integrieren, das ergibt $4\pi\varphi^3$, und dann über φ (über die verschiedenen Klassen). Man erhält so

$$g(E) \int_0^\pi J(\varphi) 8\pi(1 - \cos \varphi) d\varphi \quad (27)$$

für das Hurwitzsche Integral.

Eine andere, sehr gebräuchliche Parameterdarstellung ist die durch die Eulerschen Winkel, die in Abb. 2 veranschaulicht ist. Die Drehung mit den Eulerschen Winkeln α, β, γ ist das Produkt der drei Drehungen: mit α um Z , mit β um X und mit γ um Z . Im folgenden soll $\{\alpha \beta \gamma\}$ immer die Drehung mit den Eulerschen

Winkeln $\alpha \beta \gamma$ bedeuten. Dabei variieren im allgemeinen α und γ von $-\pi$ bis π und β von 0 bis π . Ist aber $\beta = 0$, so sind α und γ einzeln nicht bestimmt: die Drehungen $\{\alpha, 0, \gamma\}$ sind alle Drehungen mit $\alpha + \gamma$ um die Z-Achse.

XV. Die Darstellungen der dreidimensionalen reinen Drehgruppe Kugelfunktionen

1. Die irreduziblen Darstellungen der dreidimensionalen Drehgruppe kann man ähnlich wie die der zweidimensionalen mit Hilfe der Laplaceschen Differentialgleichung

$$\frac{\partial^2 f(xyz)}{\partial x^2} + \frac{\partial^2 f(xyz)}{\partial y^2} + \frac{\partial^2 f(xyz)}{\partial z^2} = 0 \quad (1)$$

ableiten, indem man die homogenen Polynome l -ten Grades bestimmt, die (1) befriedigen. Unterwirft man in einem solchen Polynom x, y, z einer orthogonalen Transformation R , so entsteht wieder ein Polynom l -ten Grades, das (1) ebenfalls befriedigt, so daß es linear durch die unveränderten Polynome ausgedrückt werden kann; die Koeffizienten bilden eine Darstellung, die wir mit $D^l(R)$ bezeichnen werden.

Da wir die irreduziblen Darstellungen der dreidimensionalen Drehgruppe noch mit Hilfe einer anderen, von H. Weyl angegebenen Methode bestimmen wollen, sei der Weg über die Laplacesche Differentialgleichung nur skizzenhaft angegeben. Zur Auflösung von (1) führt man gewöhnlich Polarkoordinaten r, ϑ, φ ein, ein Polynom l -ten Grades hat dann die Gestalt $r^l P^l(\vartheta, \varphi)$. Geht man mit diesem Ansatz in die auf Polarkoordinaten transformierte Gleichung (1) ein, so fällt r heraus, und man erhält eine Differentialgleichung für ϑ und φ allein (in der auch l vorkommt). Die $2l + 1$ linear unabhängigen Lösungen¹⁾ dieser Differentialgleichung

$$P_{-l}^l(\vartheta, \varphi), P_{-l+1}^l(\vartheta, \varphi), \dots, P_{l-1}^l(\vartheta, \varphi), P_l^l(\vartheta, \varphi) \quad (2)$$

¹⁾ Siehe z. B. R. Courant und D. Hilbert, Methoden der mathematischen Physik, Berlin 1924, S. 265, 420, 66.

bezeichnet man als Kugelfunktionen l -ten Grades. Sie haben die Gestalt

$$P_m^l(\vartheta, \varphi) = e^{-im\varphi} P_m^l(\vartheta), \quad (3)$$

wo $P_m^l(\vartheta) = P_{-m}^l(\vartheta)$ und für $m \geq 0$

$$P_m^l(\vartheta) = P_{-m}^l(\vartheta) = \frac{\sin^m \vartheta}{2^l l!} \frac{d^{l+m} \sin^2 \vartheta}{(d \cos \vartheta)^{l+m}} \quad (3a)$$

ist. Für $\vartheta = 0$ verschwinden, abgesehen von P_0^l , alle P_m^l , was schon darum der Fall sein muß, weil für $\vartheta = 0$ das Azimut φ unbestimmt ist und der Wert von $P_m^l(\vartheta, \varphi) = e^{-im\varphi} P_m^l(\vartheta)$ nicht von ihm abhängen darf.

Das Wichtige in (3) ist die Abhängigkeit von φ . Wendet man nämlich auf $r^l P_m^l(\vartheta, \varphi)$ den Operator P_R an, wo R eine Drehung um die Z -Achse mit α ist, so bleiben Radius und Polabstand ungeändert, φ geht in $\varphi - \alpha$ über. Daher ist

$$P_{\{\alpha 00\}} r^l P_m^l(\vartheta, \varphi) = r^l e^{-im(\varphi-\alpha)} P_m^l(\vartheta) = e^{im\alpha} r^l P_m^l(\vartheta, \varphi), \quad (4)$$

wenn wir die Drehung mit den Eulerschen Winkeln α, β, γ mit $\{\alpha, \beta, \gamma\}$ bezeichnen.

Die Zeilen und Spalten der zu den Kugelfunktionen l -ten Grades gehörigen $2l+1$ -dimensionalen Darstellung numerieren wir den unteren Indizes der Kugelfunktionen entsprechend von $-l$ bis l . Es ist dann

$$P_{\{\alpha\beta\gamma\}} r^l P_m^l(\vartheta, \varphi) = \sum_{m'=-l}^l \mathfrak{D}^{(l)}(\{\alpha\beta\gamma\})_{m'm} r^l P_m^l(\vartheta, \varphi), \quad (5)$$

was für $\beta = \gamma = 0$ durch Koeffizientenvergleich mit (4) in bekannter Weise

$$\mathfrak{D}^{(l)}(\{\alpha 00\})_{m'm} = e^{im\alpha} \delta_{m'm}$$

ergibt. In der Darstellung $\mathfrak{D}^{(l)}$ sind also die Matrizen, die Drehungen um Z entsprechen auf Diagonalform. Der Drehung mit α entspricht

$$\mathfrak{D}^{(l)}(\{\alpha 00\}) = \begin{pmatrix} e^{-il\alpha} & 0 & \dots & 0 & 0 \\ 0 & e^{-i(l-1)\alpha} & \dots & 0 & 0 \\ \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & \dots & e^{i(l-1)\alpha} & 0 \\ 0 & 0 & \dots & 0 & e^{il\alpha} \end{pmatrix} \quad (6)$$

Wir wollen jetzt zeigen, daß die Darstellungen $\mathfrak{D}^{(l)}$ irreduzibel sind, indem wir zeigen, daß eine Matrix, die mit den $\mathfrak{D}^{(l)}(\{\alpha\beta\gamma\})$

für alle Werte der α, β, γ vertauschbar ist, notwendig eine konstante Matrix sein muß. Zunächst ist mit den Matrizen (6) nur eine Diagonalmatrix vertauschbar, eine mit allen $\mathfrak{D}^{(l)}(\{\alpha \beta \gamma\})$ vertauschbare Matrix ist also sicher eine Diagonalmatrix. Außerdem werden wir gleich sehen, daß in $\mathfrak{D}^{(l)}(\{0 \beta 0\})$ in der 0-Zeile im allgemeinen (d. h. abgesehen von gewissen diskreten Werten von β) keine 0 vorkommt. Mit diesen Matrizen ist nur eine Diagonalmatrix mit lauter gleichen Diagonalelementen, d. h. nur eine konstante Matrix vertauschbar. In der Tat, wäre die Diagonalmatrix mit den Diagonalelementen d_k mit $\mathfrak{D}^{(l)}(\{0 \beta 0\})$ vertauschbar, so würde für die 0-Zeile des Produktes

$$d_0 \mathfrak{D}^{(l)}(\{0 \beta 0\})_{0k} = \mathfrak{D}^{(l)}(\{0 \beta 0\})_{0k} d_k, \quad (\oplus)$$

d. h. $d_0 = d_k$ folgen.

Daß $\mathfrak{D}^{(l)}(\{0 \beta 0\})$ im allgemeinen in der 0-Zeile keine 0 enthält, kann man folgendermaßen einsehen: Ist R eine Drehung um X mit β , so ersetzt P_R den Punkt $r, \vartheta, 0$ durch $r, \vartheta - \beta, 0$. Es sind daher die $r^l P_m^l(\vartheta - \beta)$; Linearkombinationen der $r^l P_{m'}^l(\vartheta)$, die Koeffizienten sind die $\mathfrak{D}^{(l)}(\{0 \beta 0\})_{m' m}$. Setzen wir nunmehr auch $\vartheta = 0$, so ist im allgemeinen $P_m^l(-\beta)$ nicht Null, dagegen sind die $P_{m'}^l(0)$ alle Null, nur $P_0^l(0)$ nicht. Der Koeffizient hiervon, $\mathfrak{D}^{(l)}(\{0 \beta 0\})_{0 m}$, darf also nicht verschwinden, sonst wären auf der rechten Seite der Gleichung alle Glieder Null, während die linke Seite nicht verschwindet.

2. Die Darstellungen $\mathfrak{D}^{(l)}(\{\alpha \beta \gamma\})$ sind also für alle $l = 0, 1, 2, 3, \dots$ irreduzibel. Um ihre Charaktere zu bestimmen, erinnern wir uns daran, daß die Spuren der Matrizen, die Elementen derselben Klasse zugeordnet sind, gleich sind. Da die Klasse in unserem Falle durch den Drehwinkel, den wir mit φ bezeichnen wollen, charakterisiert ist, ist der Charakter $\chi^{(l)}(\varphi)$ eine Funktion des Drehwinkels allein, die wir bestimmen können, wenn wir die Diagonalsumme einer Matrix, die einem Element mit dem Drehwinkel φ zugeordnet ist, berechnen. Eine solche haben wir aber in (6), wenn wir $\alpha = \varphi$ setzen. Wir erhalten so

$$\chi^{(l)}(\varphi) = \sum_{m=-l}^l e^{im\varphi} = 1 + 2 \cos \varphi + 2 \cos 2 \varphi + \cdots + 2 \cos l \varphi. \quad (7)$$

Die Orthogonalitätsrelationen lauten mit Hilfe der im vorhergehenden Kapitel (27) bestimmten Gewichtsfunktion

$$8\pi g(E) \int_0^\pi \chi^{(l)}(\varphi)^* \chi^{(l)}(\varphi) (1 - \cos \varphi) d\varphi = 8\pi^2 g(E) \delta_{ll}.$$

Man kann sie durch einfache Integration leicht verifizieren sowie auch einsehen, daß es außer den $\mathfrak{D}^{(l)}$ keine weiteren irreduziblen Darstellungen geben kann. Der Charakter einer solchen müßte nämlich, als Funktion von φ mit $(1 - \cos \varphi)$ multipliziert, auf alle $\chi^{(l)}$, also auch auf alle $\chi^{(l+1)} - \chi^{(l)}$, d. h. auf die Funktionen $1, 2 \cos \varphi, 2 \cos 2\varphi, 2 \cos 3\varphi, \dots$ im Gebiet von 0 bis π senkrecht stehen und muß daher nach dem Fouriertheorem verschwinden.

Es folgt, daß die $\mathfrak{D}^{(0)}, \mathfrak{D}^{(1)}, \mathfrak{D}^{(2)}, \dots$ die sämtlichen miteinander nichtäquivalenten irreduziblen Darstellungen der dreidimensionalen reinen Drehgruppe sind.

Die identische Darstellung ist $\mathfrak{D}^{(0)}$, dagegen sind die dreidimensionalen orthogonalen Matrizen als die Darstellung ihrer eigenen Gruppe zu $\mathfrak{D}^{(1)}$ äquivalent, was schon aus der Dimension oder dem Vergleich der Charaktere folgt.

Jede Darstellung der dreidimensionalen Drehgruppe setzt sich aus den $\mathfrak{D}^{(0)}, \mathfrak{D}^{(1)}, \mathfrak{D}^{(2)}, \dots$ zusammen und ist bis auf eine Ähnlichkeitstransformation gegeben, wenn man weiß, wie oft darin die einzelnen $\mathfrak{D}^{(0)}, \mathfrak{D}^{(1)}, \mathfrak{D}^{(2)}, \dots$ vorkommen. Diese Zahlen A_0, A_1, A_2, \dots kann man aber schon bestimmen, wenn man die Matrizen kennt, die einer Untergruppe entsprechen, die eine zweidimensionale Drehgruppe ist, wie etwa die Drehungen um die Z -Achse. Kommt darin die Darstellung $(e^{im\varphi})$ der zweidimensionalen Drehgruppe a_m -mal vor, so ist (für $m \geqq 0$) $a_m = A_m + A_{m+1} + \dots$, und $\mathfrak{D}^{(l)}$ ist in der ganzen Darstellung $A_l = a_l - a_{l+1}$ -mal enthalten. Wohlgemerkt kann man so nur schließen, wenn man aus irgendwelchen Gründen schon weiß, daß man es mit einer Darstellung zu tun hat, und man kann dieses Kriterium nicht auf irgendein Matrzensystem anwenden.

Aus (6) kennen wir die Matrizen $\mathfrak{D}^{(l)}(\{\alpha, 0, 0\}) = \mathfrak{D}^{(l)}(\{0, 0, \alpha\})$, und wir würden alle Matrizen $\mathfrak{D}^{(l)}(\{\alpha, \beta, \gamma\})$ kennen, wenn wir noch diejenigen Matrizen $\mathfrak{D}^{(l)}(\{0, \beta, 0\})$ kennen würden, die den Drehungen um X entsprechen. Bezeichnen wir $\mathfrak{D}^{(l)}(\{0, \beta, 0\})_{x1}$ mit $d^{(l)}(\beta)_{x1}$. Die Drehung $\{\alpha, \beta, \gamma\}$ ist das Produkt der drei Drehungen $\{\alpha, 0, 0\}$ $\{0, \beta, 0\}$ $\{0, 0, \gamma\}$, und so entspricht ihr

$$\mathfrak{D}^{(l)}(\{\alpha, \beta, \gamma\}) = \mathfrak{D}^{(l)}(\{\alpha, 0, 0\}) \mathfrak{D}^{(l)}(\{0, \beta, 0\}) \mathfrak{D}^{(l)}(\{0, 0, \gamma\}).$$

Daher ist

$$\mathfrak{D}^{(l)}(\{\alpha \beta \gamma\})_{m' m} = e^{i m' \alpha} \mathcal{U}^{(l)}(\beta)_{m' m} e^{i m' \gamma}. \quad (8)$$

Die Anzahl der Darstellungen der Drehgruppe ist unendlich, weil sie auch unendlich viele Klassen hat.

Isomorphie der zweidimensionalen unitären Gruppe zur Drehgruppe

3. Wir wollen die irreduziblen Darstellungen der dreidimensionalen reinen Drehgruppe noch auf einem anderen, von H. Weyl angegebenen Wege ableiten. Wenn wir nicht die sich natürlich darbietende Methode durch die Laplacesche Differentialgleichung weiter ausbauen, so hat dies seine Ursache darin, daß die Weylsche Methode mit den eigentlichen Darstellungen gleichzeitig die sogenannten „zweideutigen Darstellungen“ abzuleiten gestattet, die im Laufe der späteren Entwicklungen (bei der Theorie des Spins) eine mit den eigentlichen Darstellungen gleich wichtige Rolle spielen werden.

Während man sich bei der symmetrischen Gruppe mit der Bestimmung der Dimensionen und Charaktere einiger Darstellungen begnügen konnte, sind bei der Drehgruppe nicht nur die Charaktere, sondern auch die Koeffizienten aller Darstellungen von Bedeutung. Dies beruht, wie wir später sehen werden, darauf, daß in alle physikalisch sinnvolle Größen Teilchen gleicher Art in gleicher Weise eingehen, verschiedene Raumrichtungen aber nur dann, wenn nicht nur im mechanischen Problem, sondern auch in der Fragestellung keine Richtung ausgezeichnet ist. Dies letztere ist z. B. schon der Fall, wenn man nach der Z-Komponente des Dipolmoments fragt.

Wir beginnen mit drei einfachen Hilfssätzen, die eigentlich der elementaren Theorie der Matrizen angehören.

a) Eine Matrix, die jeden reellen Vektor in einen reellen Vektor überführt, ist selber reell, d. h. alle ihre Koeffizienten sind reell. Wendet man die Matrix auf den k -ten Einheitsvektor, dessen k -te Komponente 1, alle anderen 0 sind, an, so erhält man den Vektor, der die k -te Spalte der Matrix bildet. Diese muß also reell sein. Das gilt aber für alle k .

b) Wir haben in Kapitel III (S. 28 bis 29) gesehen, daß eine Matrix \mathbf{O} komplexorthogonal ist, wenn sie das einfache skalare

Produkt zweier beliebiger Vektoren a und b ungeändert läßt, wenn $((a, b)) = ((Oa, Ob))$ gilt. Es genügt aber schon, wenn sie die Länge eines jeden Vektors v invariant läßt, wenn $((v, v)) = ((Ov, Ov))$ gilt. Sind nämlich a und b zwei beliebige Vektoren, so können wir $v = a + b$ setzen, und es ist dann wegen $((a, b)) = ((b, a))$

$$\begin{aligned} ((v, v)) &= ((a + b, a + b)) = ((a, a)) + 2((a, b)) + ((b, b)) \\ &= ((Ov, Ov)) = ((Oa, Oa)) + 2((Oa, Ob)) + ((Ob, Ob)). \end{aligned}$$

Nachdem aber auch $((a, a)) = ((Oa, Oa))$, $((b, b)) = ((Ob, Ob))$ gelten soll, ist für jedes Vektorenpaar a, b

$$((a, b)) = ((Oa, Ob)),$$

und hieraus folgt, daß O komplexorthogonal ist. [Entsprechend gilt auch, daß U unitär ist, sobald für jeden Vektor $(v, v) = (Uv, Uv)$ gilt.]

Eine Matrix, die jeden reellen Vektor reell läßt und die Länge eines jeden Vektors unverändert läßt, ist eine Drehung. Der geometrische Sinn dieses Satzes ist die einfache Tatsache, daß, sobald in den ursprünglichen und transformierten Figuren alle Längen gleich sind, auch die Winkel alle gleich sein müssen und die Transformation lediglich eine Drehung ist.

c) Es werde noch die allgemeine Gestalt einer zweidimensionalen unitären Matrix

$$u = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

mit der Determinante 1 bestimmt. Aus $a^*c + b^*d = 0$ folgt $c = -b^*d/a^*$, und dies in $ad - bc = 1$ eingesetzt, ergibt $(aa^* + bb^*)d/a^* = 1$. Da weiter $aa^* + bb^* = 1$ ist, folgt $d = a^*$ und dann $c = -b^*$. Die allgemeine zweidimensionale unitäre Matrix mit der Determinante 1 ist also

$$u = \begin{pmatrix} a & b \\ -b^* & a^* \end{pmatrix}, \quad (9)$$

wo noch $|a^2| + |b^2| = 1$ gelten muß.

4. Wir betrachten nunmehr die drei sogenannten Paulischen Matrizen

$$s_x = \begin{pmatrix} 0 & i \\ -i & 0 \end{pmatrix}, \quad s_y = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad s_z = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}. \quad (10)$$

Man kann jede zweidimensionale Matrix mit der Spur 0 als eine Linearkombination $\mathbf{h} = x\mathbf{s}_x + y\mathbf{s}_y + z\mathbf{s}_z = [\mathfrak{r}, \mathfrak{f}]$

$$\mathbf{h} = [\mathfrak{r}, \mathfrak{f}] = \begin{pmatrix} -z & y + ix \\ y - ix & z \end{pmatrix} \quad (10a)$$

dieser Matrizen auffassen, indem man $-\mathbf{h}_{11} = \mathbf{h}_{22} = z$; $\mathbf{h}_{12} + \mathbf{h}_{21} = 2y$ und $\mathbf{h}_{12} - \mathbf{h}_{21} = 2ix$ setzt. Sind insbesondere x, y, z reell, so ist \mathbf{h} hermiteisch

Transformieren wir \mathbf{h} durch eine beliebige unitäre Matrix \mathbf{u} mit der Determinante 1, so erhalten wir wieder eine Matrix $\bar{\mathbf{h}} = \mathbf{u}\mathbf{h}\mathbf{u}^\dagger$ mit der Diagonalsumme 0, die daher auch als Linearkombination von $\mathbf{s}_x, \mathbf{s}_y, \mathbf{s}_z$ geschrieben werden kann:

$$\bar{\mathbf{h}} = \mathbf{u}\mathbf{h}\mathbf{u}^\dagger = \mathbf{u}[\mathfrak{r}, \mathfrak{f}]\mathbf{u}^\dagger = x'\mathbf{s}_x + y'\mathbf{s}_y + z'\mathbf{s}_z = [\mathfrak{r}', \mathfrak{f}], \quad (11)$$

$$\begin{pmatrix} a & b \\ -b^* & a^* \end{pmatrix} \begin{pmatrix} -z & y + ix \\ y - ix & z \end{pmatrix} \begin{pmatrix} a^* & -b \\ b^* & a \end{pmatrix} = \begin{pmatrix} -z' & y' + ix' \\ y' - ix' & z' \end{pmatrix}. \quad (11a)$$

Dabei sind die x', y', z' lineare Funktionen der x, y, z . Die Transformation \mathbf{R}_u , die das $\mathfrak{r}(xyz)$ in das $\mathbf{R}_u\mathfrak{r} = \mathfrak{r}'(x'y'z')$ überführt, kann man aus (11a) berechnen, es ist

$$\left. \begin{aligned} x' &= \frac{1}{2} (a^2 + a^{*2} + b^2 + b^{*2})x + \frac{1}{2}i(a^{*2} - a^2 + b^2 - b^{*2})y + i(a^*b^* - ab)z, \\ y' &= \frac{1}{2}i(a^2 - a^{*2} + b^2 - b^{*2})x + \frac{1}{2}(a^2 + a^{*2} - b^2 - b^{*2})y + (a^*b^* + ab)z, \\ z' &= -i(a^*b - ab^*)x - (a^*b + ab^*)y + (aa^* - bb^*)z. \end{aligned} \right\} \quad (12)$$

Auf die spezielle Form der Matrix \mathbf{R}_u kommt es dabei weniger an¹⁾, wichtig ist nur, daß wegen der Gleichheit der Determinanten von $\bar{\mathbf{h}}$ und \mathbf{h} immer

$$-x'^2 - y'^2 - z'^2 = -x^2 - y^2 - z^2 \quad (13)$$

ist, und hieraus folgt nach b), daß die Transformation \mathbf{R}_u komplexorthogonal sein muß. Man kann sich hiervon an Hand der expliziten Formel (12) auch leicht überzeugen. Weiter ist $\bar{\mathbf{h}}$ hermiteisch, also $\mathfrak{r}'(x'y'z')$ reell, wenn dies für \mathbf{h} gilt, also wenn $\mathfrak{r}(xyz)$ reell ist. Hieraus folgt wieder nach a), daß \mathbf{R}_u rein reell ist, wie man aus (12) ebenfalls ersehen kann. Es ergibt sich nunmehr, daß \mathbf{R}_u eine Drehung ist: durch (11) wird jeder unitären Matrix \mathbf{u} eine dreidimensionale Drehung zugeordnet.

¹⁾ Die komplexen Zahlen a und b , durch die die Drehung in (12) charakterisiert ist, heißen die Cayley-Kleinschen Parameter. Es ist $|a^2| + |b^2| = 1$.

Die Determinante von R_u ist +1, weil sich R_u dadurch, daß man u stetig in die Einheitsmatrix überführt, ebenfalls stetig in die dreidimensionale Einheitsmatrix überführen läßt. Wäre seine Determinante am Anfang dieses Prozesses -1 gewesen, so hätte sie einen Sprung nach +1 machen müssen, was nicht möglich ist: R_u ist für alle u eine reine Drehung.

Die Zuordnung ist dabei so beschaffen, daß dem Produkt $q u$ zweier unitärer Matrizen q und u das Produkt $R_{qu} = R_q R_u$ der entsprechenden Drehungen zugeordnet wird. Nach (11) angewendet auf q an Stelle von u ist

$$q [r, \mathfrak{f}] q^\dagger = [R_q r, \mathfrak{f}],$$

und dies mit u transformiert, ergibt

$$u q [r, \mathfrak{f}] q^\dagger u^\dagger = u [R_q r, \mathfrak{f}] u^\dagger = [R_u R_q r, \mathfrak{f}] = [R_{uq} r, \mathfrak{f}],$$

wo noch (11), angewendet auf $R_q r$ an Stelle von r und $u q$ an Stelle von u , benutzt ist. Es besteht also ein Isomorphismus zwischen der Gruppe¹⁾ der zweidimensionalen unitären Matrizen der Determinante 1 und der ihnen mit Hilfe von (11) oder (12) zugeordneten dreidimensionalen Drehungen. Es ist aber zu bemerken, daß der Isomorphismus nicht zwischen der zweidimensionalen unitären und der ganzen dreidimensionalen Drehgruppe zu bestehen brauchte, da dies bedeuten würde, daß die R_u alle Drehungen durchlaufen, wenn u die ganze unitäre Gruppe bestreicht, was noch keineswegs feststeht, was aber sogleich gezeigt werden soll. Zweitens ist zu bemerken, daß der Isomorphismus natürlich kein Holomorphismus sein muß, weil mehreren unitären Matrizen dieselbe Drehung entsprechen kann, was, wie wir ebenfalls sogleich sehen werden, auch der Fall ist.

Nehmen wir zuerst u als eine Diagonalmatrix $u_1(\alpha)$ an, d. h. setzen wir $b = 0$ und (aus später ersichtlichen Gründen) $a = e^{-\frac{1}{2}i\alpha}$, so ist α wegen $|a|^2 = 1$ reell

$$u_1(\alpha) = \begin{pmatrix} e^{-\frac{1}{2}i\alpha} & 0 \\ 0 & e^{\frac{1}{2}i\alpha} \end{pmatrix}. \quad (14a)$$

Aus (12) ersehen wir, daß die zugeordnete Drehung

$$\begin{pmatrix} \cos \alpha & -\sin \alpha & 0 \\ \sin \alpha & \cos \alpha & 0 \\ 0 & 0 & 1 \end{pmatrix} = \{\alpha, 0, 0\} \quad (14a')$$

¹⁾ Wir nennen sie im folgenden kurz „unitäre Gruppe“.

eine Drehung mit α um Z ist. Wenn wir zweitens u rein reell annehmen, so können wir

$$u_2(\beta) = \begin{pmatrix} \cos \frac{1}{2}\beta & -\sin \frac{1}{2}\beta \\ \sin \frac{1}{2}\beta & \cos \frac{1}{2}\beta \end{pmatrix} \quad (14b)$$

setzen, die zugeordnete Drehung ergibt sich dann aus (12)

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos \beta & -\sin \beta \\ 0 & \sin \beta & \cos \beta \end{pmatrix} = \{0, \beta, 0\} \quad (14b')$$

als eine Drehung mit β um X . Dem Produkt der drei unitären Matrizen $u_1(\alpha) u_2(\beta) u_1(\gamma)$ entspricht das Produkt einer Drehung mit α um Z , mit β um X und mit γ um Z , also die reine Drehung mit den Eulerschen Winkeln α, β, γ . Es folgt hieraus, daß in der in (11) definierten Zuordnung nicht nur jeder unitären zweidimensionalen Matrix eine dreidimensionale Drehung, sondern daß auch umgekehrt jeder reinen Drehung wenigstens eine unitäre Matrix, nämlich der Drehung $\{\alpha, \beta, \gamma\}$ die Matrix

$$\left. \begin{aligned} & \left(e^{-\frac{1}{2}i\alpha} \begin{pmatrix} 0 \\ e^{\frac{1}{2}i\alpha} \end{pmatrix} \right) \left(\begin{pmatrix} \cos \frac{1}{2}\beta & -\sin \frac{1}{2}\beta \\ \sin \frac{1}{2}\beta & \cos \frac{1}{2}\beta \end{pmatrix} \begin{pmatrix} e^{-\frac{1}{2}i\gamma} & 0 \\ 0 & e^{\frac{1}{2}i\gamma} \end{pmatrix} \right) \\ & = \left(\begin{pmatrix} e^{-\frac{1}{2}i\alpha} \cos \frac{1}{2}\beta e^{-\frac{1}{2}i\gamma} & -e^{-\frac{1}{2}i\alpha} \sin \frac{1}{2}\beta e^{\frac{1}{2}i\gamma} \\ e^{\frac{1}{2}i\alpha} \sin \frac{1}{2}\beta e^{-\frac{1}{2}i\gamma} & e^{\frac{1}{2}i\alpha} \cos \frac{1}{2}\beta e^{\frac{1}{2}i\gamma} \end{pmatrix} \right) \end{aligned} \right\} \quad (15)$$

entspricht. Die Isomorphie besteht also tatsächlich zwischen der unitären Gruppe und der ganzen dreidimensionalen Drehgruppe.

Es fragt sich noch, wie vielstufig diese Isomorphie ist, d. h. wie vielen unitären Matrizen dieselbe Drehung zugeordnet ist? Es genügt festzustellen, wie vielen unitären Matrizen u_0 die Einheit der Drehgruppe, die Transformation $x' = x, y' = y, z' = z$ entspricht. Für diese u_0 muß $u_0 h u_0^\dagger = h$ für alle h gelten, was nur sein kann, wenn u_0 eine konstante Matrix, $b = 0$ und $a = a^*$ reell, wegen $|a^2| + |b^2| = 1$ gleich ± 1 ist. Den beiden unitären Matrizen $+1$ und -1 — und nur diesen — entspricht also die Einheit der Drehgruppe. Diese beiden Elemente bilden einen Normalteiler der unitären Gruppe, und den Elementen (und nur diesen), die in derselben Nebengruppe dieses Normalteilers sind, also u und $-u$, entspricht dieselbe Drehung. Daß u und $-u$ dieselbe Drehung entspricht, sieht man auch aus (12) unmittelbar,

sowie auch daraus, daß in (15) die trigonometrischen Funktionen der halben Eulerschen Winkel vorkommen. Durch die Drehung sind die Eulerschen Winkel gewissermaßen nur bis auf Vielfache von 2π , die halben Winkel nur bis auf Vielfache von π , die trigonometrischen Funktionen dieser nur bis auf das Vorzeichen festgelegt.

Unser wichtiges Resultat lautet also folgendermaßen: Es besteht eine zweistufige Isomorphie zwischen der Gruppe der zweidimensionalen unitären Matrizen mit der Determinante 1 und der dreidimensionalen reinen Drehgruppe, bei der jedem unitären Matrizenpaar $u, -u$ eindeutig umkehrbar eine Drehung R_u in der Weise zugeordnet ist, daß aus $u \cdot q = t$ auch $R_u \cdot R_q = R_t$ und aus $R_u \cdot R_q = R_t$ umgekehrt $u \cdot q = \pm t$ folgt. Kennt man die unitäre Matrix u , so erhält man die zugehörige Drehung R_u am besten aus (12), umgekehrt findet man die unitären Matrizen zur Drehung $\{\alpha \beta \gamma\}$ aus (15).

Die Darstellungen der unitären Gruppe

5. Durch diese Isomorphie besteht ein enger Zusammenhang zwischen den Darstellungen der beiden Gruppen. Man kann nämlich — wie dies schon im IX. Kapitel auseinandergesetzt wurde — aus jeder Darstellung $D(R)$ der kleineren Gruppe — das ist in diesem Fall die Drehgruppe — eine Darstellung $U(u)$ der anderen Gruppe gewinnen, indem man allen Elementen (u und $-u$) der zweiten Gruppe, denen im Isomorphismus dasselbe Element R_u der ersten Gruppe entspricht, die Matrix $U(u) = D(R_u)$ zuordnet. Insbesondere wird so den beiden unitären Matrizen 1 und -1 die Einheitsmatrix $D(E)$ zugeordnet. Kennt man umgekehrt alle Darstellungen der unitären Gruppe, so kann man aus diesen jene heraussuchen, in denen den beiden Matrizen u und $-u$ dieselbe Matrix $U(u) = U(-u)$ entspricht. Jede dieser Darstellungen erlaubt eine Darstellung der Drehgruppe zu bilden, die man erhält, wenn man der Drehung R_u die Matrix $D(R_u) = U(u) = U(-u)$ zuordnet. Man kann auf diese Weise sämtliche Darstellungen der Drehgruppe erhalten.

Sei insbesondere die Darstellung $U(u)$ der unitären Gruppe irreduzibel. Das Element $u = -1$ ist mit allen Elementen der Gruppe vertauschbar, und so muß auch $U(-1)$ mit allen $U(u)$

vertauschbar sein. Es ist daher nach den allgemeinen Sätzen über irreduzible Darstellungen eine konstante Matrix. Ihr Quadrat muß ebenso wie das von -1 die Einheitsmatrix¹⁾ $\mathbf{U}(1)$ sein. Es ist daher entweder

$$\mathbf{U}(-1) = \mathbf{U}(1) \quad \text{oder} \quad \mathbf{U}(-1) = -\mathbf{U}(1).$$

Die Darstellungen, für die $\mathbf{U}(-1) = \mathbf{U}(1)$ gilt, nennen wir gerade Darstellungen. Für sie gilt auch $\mathbf{U}(-u) = \mathbf{U}(-1) \cdot \mathbf{U}(u) = \mathbf{U}(1) \cdot \mathbf{U}(u) = \mathbf{U}(u)$, den beiden Elementen u und $-u$ ist bei ihnen immer dieselbe Matrix zugeordnet. Sie geben daher richtiggehende Darstellungen der Drehgruppe, und wir kennen sie implizite alle aus 1.

In den Darstellungen, für die $\mathbf{U}(-1) = -\mathbf{U}(1)$ gilt, ist $\mathbf{U}(-u) = \mathbf{U}(-1) \cdot \mathbf{U}(u) = -\mathbf{U}(u)$. Elementen, die sich im Vorzeichen unterscheiden, entsprechen Matrizen, die ebenfalls entgegengesetztes Vorzeichen haben. Aus diesen Darstellungen der unitären Gruppe — wir nennen sie ungerade — kann man keine richtiggehenden Darstellungen der Drehgruppe bilden, sondern nur sogenannte „zweideutige“ oder „halbzählig“, in denen jeder Drehung $R_u = R_{-u}$ nicht eine Matrix, sondern zwei Matrizen $\mathbf{U}(u)$ und $\mathbf{U}(-u) = -\mathbf{U}(u)$ entsprechen, die sich im Vorzeichen aller Koeffizienten unterscheiden.

Eine ungerade Darstellung der unitären Gruppe bildet sie als ihre eigene Darstellung: $\mathbf{U}(u) = u$. In der entsprechenden zweideutigen Darstellung $\mathfrak{D}^{(1|2)}$ der Drehgruppe ist der Drehung $\{\alpha, \beta, \gamma\}$ diejenige Matrix $u = \mathbf{U}(u)$ selber zugeordnet, der im Isomorphismus R entspricht. Es ist also nach (15)

$$\mathfrak{D}^{(1|2)}(\{\alpha \beta \gamma\}) = \pm \begin{pmatrix} e^{-\frac{1}{2}i\alpha} \cos \frac{1}{2}\beta e^{-\frac{1}{2}i\gamma} & -e^{-\frac{1}{2}i\alpha} \sin \frac{1}{2}\beta e^{\frac{1}{2}i\gamma} \\ e^{\frac{1}{2}i\alpha} \sin \frac{1}{2}\beta e^{-\frac{1}{2}i\gamma} & e^{\frac{1}{2}i\alpha} \cos \frac{1}{2}\beta e^{\frac{1}{2}i\gamma} \end{pmatrix}. \quad (16)$$

Die erste Zeile bzw. Spalte nennt man dabei gewöhnlich die $-\frac{1}{2}$ -Zeile bzw. -Spalte, die zweiten die $+\frac{1}{2}$ -Zeile bzw. -Spalte. In (16) haben wir die erste zweideutige Darstellung der Drehgruppe.

¹⁾ Die Matrix $\mathbf{U}(1)$, die der Einheit der Gruppe zugeordnet ist, ist eine Einheitsmatrix, die die Dimension der Darstellung hat. Wir gebrauchen hier das Zeichen $\mathbf{U}(1)$ an Stelle des einfacheren Zeichens 1 , um eine Verwechslung mit der Einheit 1 der unitären Gruppe (die zweidimensional ist) zu vermeiden.

Für die zweideutigen Darstellungen gilt nicht $\mathfrak{D}(R) \cdot \mathfrak{D}(S) = \mathfrak{D}(RS)$, sondern nur $\mathfrak{D}(R)\mathfrak{D}(S) = \pm \mathfrak{D}(RS)$, weil die Darstellungsmatrizen nur bis auf das Vorzeichen bestimmt sind. Es ist sogar so, daß es nicht möglich ist, das Vorzeichen für alle Matrizen gleichzeitig so festzulegen, daß für das resultierende Matrzensystem das strengere Multiplikationsgesetz der eindeutigen Darstellungen gelte. Die zweideutigen Darstellungen sind also nicht so entstanden, daß man in einer eindeutigen Darstellung die Vorzeichen aller Matrizen unbestimmt gelassen hat. Dies sieht man z. B. aus (16): einer Drehung mit π um Z entsprechen die Matrizen $\pm i\mathbf{s}_z$, dem Quadrate hiervon, einer Drehung mit 2π um Z entspricht also sicher auch die Matrix -1 . Da diese Drehung aber eigentlich keine Drehung ist, sondern alles unverändert läßt und die Einheit der Gruppe bildet, muß ihr auch noch die Einheitsmatrix entsprechen, das Eindeutigmachen der Darstellung ist nicht möglich.

6. Wir bestimmen jetzt die irreduziblen Darstellungen der zweidimensionalen unitären Gruppe.

Haben wir ein homogenes Polynom n -ten Grades von ε und ξ und führen eine unitäre Transformation der Variablen aus

$$\begin{aligned}\varepsilon' &= a\varepsilon + b\xi, \\ \xi' &= -b^*\varepsilon + a^*\xi,\end{aligned}\} \quad (17)$$

so erhalten wir wieder ein homogenes Polynom n -ten Grades. Dies ist sogar für beliebige lineare Transformationen richtig, doch wollen wir uns auf unitäre beschränken. Zu den $n+1$ Polynomen $\varepsilon^n, \varepsilon^{n-1}\xi, \varepsilon^{n-2}\xi^2, \dots, \varepsilon^2\xi^{n-2}, \varepsilon\xi^{n-1}, \xi^n$ gehört daher eine $n+1$ -dimensionale Darstellung der unitären Gruppe. Um sofort die bei der Drehgruppe üblichen Bezeichnungen zu erhalten, nennen wir $n = 2j$, die Dimension der Darstellung ist dann $2j+1$ und j ist halbzahlig¹⁾ oder ganzzahlig. Die Polynome seien

$$f_\mu = \frac{\varepsilon^{j+\mu} \xi^{j-\mu}}{\sqrt{(j+\mu)!(j-\mu)!}}, \quad (18)$$

wo μ die $2j+1$ -Werte $-j, -j+1, -j+2, \dots, j-2, j-1, j$ annehmen kann, es ist bei ganzzahligem j selber ganzzahlig, bei

¹⁾ Darunter versteht man, daß es sich von einer ganzen Zahl um $\pm \frac{1}{2}$ unterscheidet.

halbzahligem j halbzahlig. Der konstante Faktor $[(j + \mu)! (j - \mu)!]^{-\frac{1}{2}}$ wurde zu $\varepsilon^{j+\mu} \xi^{j-\mu}$ deshalb hinzugefügt, weil — wie sich zeigen wird — die Darstellung $\mathfrak{U}^{(j)}$ für die $2j + 1$ -Funktionen (18) hierdurch die unitäre Form annimmt.

Bilden wir nunmehr¹⁾ $\mathbf{P}_{\mathbf{u}} f_{\mu}$ nach (19), Kapitel XI.

$$\begin{aligned}\mathbf{P}_{\mathbf{u}} f_{\mu}(\varepsilon, \xi) &= f_{\mu}(a^* \varepsilon - b \xi, b^* \varepsilon + a \xi) \\ &= \frac{(a^* \varepsilon - b \xi)^{j+\mu} (b^* \varepsilon + a \xi)^{j-\mu}}{\sqrt{(j + \mu)! (j - \mu)!}}.\end{aligned}\quad (19)$$

Um die rechte Seite als Linearkombination der $f_{\mu'}$ auszudrücken, entwickeln wir sie nach dem Binomialtheorem:

$$\sum_{x=0}^{j+\mu} \sum_{x'=0}^{j-\mu} (-1)^x \frac{\sqrt{(j + \mu)! (j - \mu)!}}{x! x'! (j + \mu - x)! (j - \mu - x')!} \cdot a^x a^{*j+\mu-x} b^x b^{*j-\mu-x'} \varepsilon^{2j-x-x'} \xi^x + x'. \quad (19a)$$

Man kann hierin die Summationsgrenzen weglassen und die Summation über alle ganzen Zahlen erstrecken, weil die Binomialkoeffizienten für außerhalb des Summationsbereiches liegende x, x' doch verschwinden. Wenn man noch $j - x - x' = \mu'$ setzt, muß μ' bei ganzzahligem j alle ganzen Zahlen, bei halbzahligem j alle halben Zahlen durchlaufen. Drückt man in (19a) alle Funktionen von ε und ξ mit Hilfe von (18) durch die $f_{\mu'}$ aus, so ergibt sich

$$\mathbf{P}_{\mathbf{u}} f_{\mu}(\varepsilon, \xi) = \sum_{\mu'} \sum_x (-1)^x \frac{\sqrt{(j + \mu)! (j - \mu)! (j + \mu')! (j - \mu')!}}{x! (j - \mu' - x)! (j + \mu - x)! (x + \mu' - \mu)!} \cdot a^{j-\mu'-x} a^{*j+\mu-x} b^x b^{*x+\mu'-\mu} f_{\mu'}(\varepsilon, \xi). \quad (20)$$

Der Koeffizient von $f_{\mu'}$ rechts ist $\mathfrak{U}^{(j)}(\mathbf{u})_{\mu' \mu}$:

$$\mathfrak{U}^{(j)}(\mathbf{u})_{\mu' \mu} = \sum_x (-1)^x \frac{\sqrt{(j + \mu)! (j - \mu)! (j + \mu')! (j - \mu')!}}{(j - \mu' - x)! (j + \mu - x)! x! (x + \mu' - \mu)!} \cdot a^{j-\mu'-x} a^{*j+\mu-x} b^x b^{*x+\mu'-\mu}. \quad (21)$$

¹⁾ Dabei ist \mathbf{u} die unitäre Transformation in (17). In Kapitel XI wurde \mathbf{P}_R eigentlich nur für reelle orthogonale R definiert. In unserem Falle, wo \mathbf{u} unitär ist, folgt aus (18a) Kapitel XI

$$x_i = \sum_j \mathbf{R}_{ji}^* x'_j$$

an Stelle von (18b), wo also \mathbf{R}_{ji}^* an Stelle von \mathbf{R}_{ji} steht.

Etwas einfacher lauten die Ausdrücke für $\mu' = j$, d. h. für die letzte Zeile der Darstellungsmatrizen, weil wegen der Fakultäten im Nenner nur das Glied mit $\kappa = 0$ übrigbleibt:

$$\mathfrak{U}^{(j)}(\mathbf{u})_{j\mu} = \sqrt{\frac{(2j)!}{(j+\mu)!(j-\mu)!}} a^{*j+\mu} b^{*j-\mu}. \quad (21a)$$

Hierdurch haben wir die Koeffizienten der Darstellungen $\mathfrak{U}^{(j)}$ für alle möglichen $j = 0, \frac{1}{2}, 1, \frac{3}{2}, 2, \dots$ bestimmt, und wir müssen nur noch nachweisen, daß sie in (21) in unitärer Form vorliegen, daß sie irreduzibel sind und daß die zweidimensionale unitäre Gruppe außer diesen keine weiteren irreduziblen Darstellungen hat.

7. Zuerst beweisen wir die Unitarität der Darstellungen (21). Der Beweis beruht auf der Tatsache, daß die Polynome f_μ in (18) so angenommen wurden, daß

$$\sum_{\mu=-j}^j f_\mu f_\mu^* = \sum_\mu \frac{1}{(j+\mu)!(j-\mu)!} |\varepsilon^2|^{j+\mu} |\xi^2|^{j-\mu} = \frac{(|\varepsilon^2| + |\xi^2|)^{2j}}{(2j)!}. \quad (22)$$

Da eine unitäre Transformation (17) das skalare Produkt $|\varepsilon^2| + |\xi^2|$ unverändert läßt, ist $\mathbf{P}_u (|\varepsilon^2| + |\xi^2|) = |\varepsilon^2| + |\xi^2|$ und wegen (22) Kapitel XI auch

$$\left. \begin{aligned} (2j)! \mathbf{P}_u \sum_\mu f_\mu f_\mu^* &= \mathbf{P}_u (|\varepsilon^2| + |\xi^2|)^{2j} = [\mathbf{P}_u (|\varepsilon^2| + |\xi^2|)]^{2j} \\ &= (|\varepsilon^2| + |\xi^2|)^{2j} = (2j)! \sum_\mu f_\mu f_\mu^*. \end{aligned} \right\} \quad (23)$$

Andererseits ist wegen derselben Gleichung

$$\left. \begin{aligned} \mathbf{P}_u \sum_\mu f_\mu f_\mu^* &= \sum_\mu \mathbf{P}_u f_\mu (\mathbf{P}_u f_\mu)^* \\ &= \sum_\mu \sum_{\mu' \mu''} \mathfrak{U}_{\mu' \mu}^{(j)} f_{\mu'} \sum_{\mu''} \mathfrak{U}_{\mu'' \mu}^{(j)*} f_{\mu''}^*, \end{aligned} \right\} \quad (23a)$$

da man, anstatt in f_μ^* die Substitution \mathbf{P}_u auszuführen, sie in f_μ ausführen und dann zum konjugiert komplexen übergehen kann. Wenn man die $(2j+1)^3$ Funktionen $f_{\mu'} f_{\mu''}^*$ als linear unabhängig voraussetzt, ergibt der Vergleich von (23) und (23a) direkt

$$\sum_\mu \mathfrak{U}^{(j)}(\mathbf{u})_{\mu' \mu} \mathfrak{U}^{(j)}(\mathbf{u})_{\mu'' \mu}^* = \delta_{\mu' \mu''}, \quad (24)$$

also die Bedingung für die Unitarität von $\mathfrak{U}^{(j)}$. Die Unitarität von $\mathfrak{U}^{(j)}$ ist also bewiesen, sobald gezeigt ist, daß unter den $f_{\mu'} f_{\mu''}^*$ keine linearen Beziehungen bestehen, d. h. daß aus

$$\sum_{\mu' \mu''} c_{\mu' \mu''} \varepsilon^{j+\mu'} \xi^{j-\mu'} \varepsilon^{*\mu''} \xi^{*\mu'-\mu''} = 0 \quad (*)$$

$c_{\mu' \mu''} = 0$ folgt. Dabei müßte (*) für alle komplexen Werte der Variablen ε, ξ gelten, weil auch (23) und (23a) für alle komplexen ε und ξ gelten. Nehmen wir insbesondere ε reell an, so ergibt der Vergleich der $\lambda = 2j + \mu' + \mu''$ Potenzen von ε in (*), wenn man noch durch $\varepsilon^{\lambda} \xi^j \xi^{*\lambda - \lambda}$ dividiert:

$$\sum_{\mu'} c_{\mu', \lambda - 2j - \mu'} (\xi^*/\xi)^{\mu'} = 0.$$

Hieraus folgt aber auch $c_{\mu', \lambda - 2j - \mu'} = 0$ und die lineare Unabhängigkeit der $f_{\mu'} f_{\mu''}^*$, da ξ^*/ξ eine auf dem komplexen Einheitskreis frei veränderliche Variable ist. Man kann für sie auch $e^{i\tau}$ einsetzen, und τ kann dann jeden reellen Wert annehmen. Aus

$$\sum_{\mu'} c_{\mu', \lambda - 2j - \mu'} e^{i\mu' \tau} = 0$$

für alle reellen τ folgt aber das Verschwinden aller c .

8. Die Irreduzibilität des Matrzensystems $\mathfrak{U}^{(j)}$ kann man ebenso beweisen, wie wir in 1. die Irreduzibilität der Darstellungen $\mathfrak{D}^{(l)}$ der Drehgruppe bewiesen haben: indem man nachweist, daß eine Matrix \mathbf{M} , die mit $\mathfrak{U}^{(j)}(\mathbf{u})$ für alle \mathbf{u} , d. h. für alle Werte von a und b , die der Bedingung $|a|^2 + |b|^2 = 1$ genügen, vertauschbar ist, notwendigerweise eine konstante Matrix sein muß. Wir betrachten zuerst ein \mathbf{u} der Gestalt $\mathbf{u}_1(\alpha)$ (14a), setzen also $b = 0$, $a = e^{-1/2} e^{i\alpha}$. Dann bleibt in der Summe in (21) nur das Glied mit $\kappa = 0$ und auch dies nur dann übrig, wenn $\mu' = \mu$ ist, und wir erhalten

$$\mathfrak{U}^{(j)}(\mathbf{u}_1(\alpha))_{\mu' \mu} = \delta_{\mu' \mu} e^{i\mu \alpha}. \quad (25)$$

Die Matrizen, die den unitären Transformationen der Gestalt $\mathbf{u}_1(\alpha)$ in $\mathfrak{U}^{(j)}$ entsprechen, haben also genau die Gestalt (6) mit dem einzigen Unterschied, daß j nicht wie das l in (6) notwendigerweise ganzzahlig sein muß, sondern auch halbzahlig sein kann. Mit diesen Matrizen ist nur eine Diagonalmatrix vertauschbar, so daß \mathbf{M} eine Diagonalmatrix sein muß. Nun sehen wir aus (21a), daß in der letzten Zeile von $\mathfrak{U}^{(j)}$ im allgemeinen kein Koeffizient verschwindet. Hieraus kann man ähnlich, wie dies in \oplus geschah, durch Gleichsetzen der Koeffizienten der j -Zeile von $\mathfrak{U}^{(j)} \mathbf{M}$ und $\mathbf{M} \mathfrak{U}^{(j)}$ schließen, daß

$$\mathfrak{U}_{jk}^{(j)} \mathbf{M}_{kk} = \mathbf{M}_{jj} \mathfrak{U}_{kj}^{(j)}, \quad \mathbf{M}_{kk} = \mathbf{M}_{jj}$$

ist, und \mathbf{M} eine konstante Matrix sein muß. Daher sind die Darstellungen $\mathfrak{U}^{(j)}$ irreduzibel.

9. Auch daß es außer den $\mathfrak{U}^{(j)}$ keine weiteren irreduziblen Darstellungen der unitären Gruppe mehr gibt, können wir genau so zeigen, wie dies in 2. für die Darstellungen $\mathfrak{D}^{(l)}$ der Drehgruppe geschah. Wir bestimmen zuerst die Klasseneinteilung der „unitären Gruppe“, d. h. der Gruppe der zweidimensionalen unitären Matrizen mit der Determinante 1. Jede unitäre Matrix kann durch Transformation mit einer unitären Matrix auf die Diagonalform gebracht werden, unsere Matrizen haben nach dieser Transformation alle die Gestalt $u_1(\alpha)$, wobei, da auch $u_1(-\alpha)$ mit $u_1(\alpha)$ äquivalent ist, α von 0 bis 2π gezählt werden kann. Alle u , die in dasselbe $u_1(\alpha)$ transformiert werden können, sind in derselben Klasse. (Daß für die Transformation nur Gruppenelemente, d. h. nur unitäre Matrizen mit der Determinante 1 in Frage kommen, braucht uns nicht zu stören, da jede unitäre Matrix als das Produkt einer unitären Matrix mit der Determinante 1 und einer konstanten Matrix geschrieben werden kann und die Transformation mit der konstanten Matrix doch weggelassen werden kann.)

Um den Charakter von $\mathfrak{U}^{(j)}$ zu bestimmen, genügt es, die Spur je eines Elementes jeder Klasse zu berechnen. Als das Element aus der Klasse von $u_1(\alpha)$ nehmen wir $u_1(\alpha)$ selber, die zugeordnete Matrix ist in (25) berechnet. Ihre Diagonalsumme ist

$$\xi_j(\alpha) = \sum_{\mu=-j}^j e^{i\mu\alpha}, \quad (26)$$

wo die Summation von der unteren Grenze bis zur oberen in ganzzähligen Schritten auszuführen ist.

Wir sehen hieraus, daß die unitäre Gruppe außer den $\mathfrak{U}^{(j)}$ mit $j = 0, \frac{1}{2}, 1, \frac{3}{2}, \dots$ keine irreduziblen Darstellungen haben kann. Der Charakter einer solchen müßte mit einer Gewichtsfunktion multipliziert auf alle $\xi_j(\alpha)$ und daher auch auf $\xi_0(\alpha), \xi_{1/2}(\alpha), \xi_1(\alpha) - \xi_0(\alpha), \xi_{3/2}(\alpha) - \xi_{1/2}(\alpha), \dots$ senkrecht stehen. Eine Funktion aber, die auf $1, 2 \cos \frac{1}{2}\alpha, 2 \cos \alpha, 2 \cos \frac{3}{2}\alpha, \dots$ im Gebiet von 0 bis 2π senkrecht steht, muß nach dem Fouriertheorem verschwinden.

Die Darstellungen der dreidimensionalen reinen Drehgruppe

10. Jede Darstellung $\mathfrak{U}^{(j)}$ der unitären Gruppe ist gleichzeitig eine — ein- oder zweideutige — Darstellung der Drehgruppe, die der Drehung $\{\alpha \beta \gamma\}$ die Matrix $\mathfrak{U}^{(j)}(u)$ zuordnet, wo u die $\{\alpha \beta \gamma\}$

im Sinne des Isomorphismus entsprechende unitäre Transformation ist. Die Koeffizienten a und b von \mathfrak{u} sind nach (15)

$$a = e^{-1/2 i \alpha} \cos \frac{1}{2} \beta \cdot e^{-1/2 i \gamma}, \quad b = -e^{-1/2 i \alpha} \sin \frac{1}{2} \beta \cdot e^{1/2 i \gamma}, \quad (15 \text{a})$$

und dies muß man in (21) einsetzen, um die Koeffizienten der zu $\{\alpha \beta \gamma\}$ zugeordneten Darstellungsmatrix zu erhalten. Um mit (16) in Übereinstimmung zu bleiben, transformieren wir noch die so entstandene Matrix mit der Diagonalmatrix $\mathbf{M}_{x\lambda} = \delta_{x\lambda} i^{-2x}$, d. h. multiplizieren die μ' -Zeile mit $i^{-2\mu'}$, die μ -Spalte mit $i^{2\mu}$, also den μ', μ -Koeffizienten mit $i^{2(\mu-\mu')} = (-1)^{\mu-\mu'}$. Die so aus $\mathfrak{U}^{(j)}$ entstandene Darstellung nennen wir $\mathfrak{D}^{(j)}(\{\alpha \beta \gamma\})$, ihre Koeffizienten sind

$$\left. \begin{aligned} \mathfrak{D}^{(j)}(\{\alpha \beta \gamma\})_{\mu' \mu} &= \sum_x (-1)^x \frac{\sqrt{(j+\mu)! (j-\mu)! (j+\mu')! (j-\mu')!}}{(j-\mu'-x)! (j+\mu-x)! x! (x+\mu'-\mu)!} \\ &\quad \cdot e^{i\mu' \alpha} \cos^{2j+\mu-\mu'-2x} \frac{1}{2} \beta \cdot \sin^{2x+\mu'-\mu} \frac{1}{2} \beta \cdot e^{i\mu \gamma}. \end{aligned} \right\} \quad (27)$$

Die Darstellung $\mathfrak{D}^{(j)}$ ist $2j+1$ -dimensional, wobei j ganz- oder halbzahlig sein kann. Die Zeilen und Spalten von $\mathfrak{D}^{(j)}$ sind nach den ganzen bzw. halben Zahlen $-j, -j+1, \dots, j-1, j$ benannt. Die Summation über x in (27) ist eigentlich über alle ganzen Zahlen zu erstrecken, wegen des Unendlichwerdens der Fakultäten im Nenner genügt es, sie von der größeren der beiden Zahlen 0 und $\mu-\mu'$ bis zur kleineren von $j-\mu'$ und $j+\mu$ zu führen. Besonders einfach werden die Formeln für $\mu'=j$ und $\mu'=-j$, im ersten Falle kommt nur $x=0$, im zweiten nur $x=j+\mu$ in Frage:

$$\mathfrak{D}^{(j)}(\{\alpha \beta \gamma\})_{j\mu} = \sqrt{\binom{2j}{j-\mu}} e^{ij\alpha} \cos^{j+\mu} \frac{1}{2} \beta \sin^{j-\mu} \frac{1}{2} \beta e^{i\mu \gamma}, \quad (27 \text{a})$$

$$\left. \begin{aligned} \mathfrak{D}^{(j)}(\{\alpha \beta \gamma\})_{-j\mu} &= (-1)^{j+\mu} \sqrt{\binom{2j}{j-\mu}} e^{-ij\alpha} \cos^{j-\mu} \frac{1}{2} \beta \sin^{j+\mu} \frac{1}{2} \beta e^{i\mu \gamma}. \end{aligned} \right\} \quad (27 \text{b})$$

Eine einfache Gestalt haben noch alle Darstellungskoeffizienten, die Drehungen um Z entsprechen. Der Drehung um Z mit α entspricht im Isomorphismus die unitäre Transformation $\mathfrak{u}_1(\alpha)$ und

die Koeffizienten der zu dieser zugeordneten Matrix sind in (25) angegeben. Die Matrix, die in $\mathfrak{D}^{(j)}$ der Drehung $\{\alpha, 0, 0\}$ zugeordnet ist, ist also eine Diagonalmatrix mit den Diagonalelementen $e^{-ij\alpha}, e^{-i(j-1)\alpha}, \dots, e^{i(j-1)\alpha}, e^{ij\alpha}$. Dasselbe erhält man direkt aus (27), indem man $\beta = \gamma = 0$ setzt. Die Matrix $\mathfrak{D}^{(j)}(\{\alpha 0 0\})$ steht schon in (6) explizit hingeschrieben, diese Formel gilt also nicht nur für ganzzahlige j , sondern auch für halbzahliges j . Daselbe gilt auch von (8).

Der Charakter $\chi^{(j)}(\varphi)$ von $\mathfrak{D}^{(j)}$ ergibt sich wieder als die Spur einer Drehung mit dem Drehwinkel φ zu

$$\left. \begin{aligned} \chi^{(j)}(\varphi) &= \sum_{\mu=-j}^j e^{i\mu\varphi} \\ &= \begin{cases} 1 + 2 \cos \varphi + \dots + 2 \cos j\varphi & (j \text{ ganzzahlig}), \\ 2 \cos \frac{1}{2}\varphi + 2 \cos \frac{3}{2}\varphi + \dots + 2 \cos j\varphi & (j \text{ halbzahlig}). \end{cases} \end{aligned} \right\} \quad (28)$$

Richtiggehende Darstellungen liefern nur jene j , für die $\mathfrak{U}^{(j)}(-1) = \mathfrak{U}^{(j)}(\mathbf{u}_1(2\pi))$ die positive Einheitsmatrix ist. Aus (25) zeigt sich, daß dies dann der Fall ist, wenn μ ganzzahlig ist, also bei ganzzahligem j . In diesem Falle ist $\mathfrak{D}^{(j)}$ mit den in 1. abgeleiteten $\mathfrak{D}^{(l)}$ identisch, wie auch aus dem Vergleich der Charaktere hervorgeht.

Für halbzahliges j ist $\mathfrak{D}^{(j)}$ zweideutig, der Drehung $\{\alpha \beta \gamma\}$ ist $\pm \mathfrak{D}^{(j)}(\{\alpha \beta \gamma\})$ zugeordnet. Dies ist natürlich nicht etwa so zu verstehen, daß man die Vorzeichen der Koeffizienten von $\mathfrak{D}^{(j)}$ auch einzeln umkehren kann, man kann nur das Vorzeichen der ganzen Matrix, also aller Koeffizienten gleichzeitig umkehren. Einer Drehung R_u entsprechen ja nur zwei unitäre Matrizen u und $-u$ und diesen ist nur je eine Matrix $\mathfrak{U}^{(j)}(u)$ und $\mathfrak{U}^{(j)}(-u)$ zugeordnet, deren zweite in diesem Falle gleich $-\mathfrak{U}^{(j)}(u)$ ist. Diese beiden Matrizen – aber keine weiteren – entsprechen in $\mathfrak{D}^{(j)}$ der Drehung R_u . Immerhin muß man sich vor Augen halten, daß die „zweideutigen Darstellungen“ eigentlich keine Darstellungen sind. Sie kommen zum ersten Male in der Paulischen Spintheorie vor, wir haben sie hier nur aus dem Grunde gemeinsam mit den eindeutigen Darstellungen abgeleitet, damit wir nicht wieder auf diesen Gegenstand zurückzukommen brauchen.

Die Theorie der Darstellungen der Drehgruppe röhrt von J. Schur her. Die „zweideutigen Darstellungen“ hat zuerst H. Weyl angegeben.

11. Die ersten paar Darstellungen seien hier noch explizite angegeben. Es ist $\mathbf{D}^{(0)}(R) = (1)$; $\mathbf{D}^{(1,2)}(R)$ steht in (16). Weiter ist

$$\mathbf{D}^{(1)}(\{\alpha \beta \gamma\}) = \begin{pmatrix} e^{-i\alpha} \frac{1+\cos \beta}{2} e^{-i\gamma} & -e^{-i\alpha} \frac{\sin \beta}{\sqrt{2}} & e^{-i\alpha} \frac{1-\cos \beta}{2} e^{i\gamma} \\ \frac{1}{\sqrt{2}} \sin \beta e^{-i\gamma} & \cos \beta & -\frac{1}{\sqrt{2}} \sin \beta e^{i\gamma} \\ e^{i\alpha} \frac{1-\cos \beta}{2} e^{-i\gamma} & e^{i\alpha} \frac{\sin \beta}{\sqrt{2}} & e^{i\alpha} \frac{1+\cos \beta}{2} e^{i\gamma} \end{pmatrix}, \quad (29)$$

wo die trigonometrischen Funktionen der halben Winkel schon in solche der ganzen Winkel umgewandelt wurden.

Die Darstellungen der Drehgruppe — wenigstens die eindeutigen — sind eigentlich etwas dem Physiker sehr Geläufiges, weil sie gleichzeitig die Transformationsformeln für Vektoren, Tensoren usw. sind. Bei einem Übergang zu einem neuen Koordinatensystem bleiben ja die Vektor- oder Tensorkomponenten nicht unverändert, sondern setzen sich linear aus den Komponenten des Vektors oder Tensors im alten Koordinatensystem zusammen. Bezeichnen wir die Komponenten im alten System mit T_σ , wo σ auch eine Gesamtheit von mehreren Indizes andeuten kann, so seien die Komponenten T'_ϱ im neuen Koordinatensystem

$$T'_\varrho = \sum_\sigma \mathbf{D}(R)_{\varrho\sigma} T_\sigma, \quad (30)$$

wo die Abhängigkeit der Transformationskoeffizienten von der Drehung R des neuen Achsenkreuzes zum alten explizite zum Ausdruck gebracht ist. Gehen wir wieder — etwa durch die Drehung S — zu einem neuen Achsenkreuz über, so ist

$$T''_\tau = \sum_\varrho \mathbf{D}(S)_{\tau\varrho} T'_\varrho = \sum_{\varrho\sigma} \mathbf{D}(S)_{\tau\varrho} \mathbf{D}(R)_{\varrho\sigma} T_\sigma. \quad (31)$$

Nun sind die T''_τ die Komponenten des Tensors im um SR verdrehten Achsenkreuz, so daß auch

$$T''_\tau = \sum_\sigma \mathbf{D}(SR)_{\tau\sigma} T_\sigma \quad (32)$$

gelten muß. Da (31) und (32) für beliebige Werte der Tensorkomponenten T_σ gelten muß, ist

$$\mathbf{D}(SR)_{\tau\sigma} = \sum_\varrho \mathbf{D}(S)_{\tau\varrho} \mathbf{D}(R)_{\varrho\sigma}; \quad \mathbf{D}(SR) = \mathbf{D}(S) \mathbf{D}(R). \quad (33)$$

Die Transformationsmatrizen der Vektor- oder Tensorkomponenten bilden eine Darstellung der Drehgruppe.

So ist z. B. die Transformationsmatrix für Vektoren die Drehungsmatrix R selber, und diese bilden ja eine Darstellung ihrer eigenen Gruppe,

die zu $\mathfrak{D}^{(1)}$ äquivalent ist. Die „Transformationsmatrix“ für Invarianten ist $\mathfrak{D}^{(0)}$.

Zu den meisten geläufigen Arten von Tensoren gehören aber keine irreduziblen, sondern noch reduzible Darstellungen, weil man aus den Tensorkomponenten noch solche Linearkombinationen bilden kann, die sich unter sich transformieren. Durch die Matrix, durch die diese Linearkombinationen aus den ursprünglichen Komponenten hervorgehen, wird auch die Darstellung ausreduziert.

Betrachten wir z. B. einen Tensor zweiten Grades mit den Komponenten $T_{xx}, T_{xy}, T_{xz}, T_{yx}, T_{yy}, T_{yz}, T_{zx}, T_{zy}, T_{zz}$. Man kann aus ihm einen symmetrischen und einen schiefen Tensor bilden, die sechs Komponenten des ersten sind: $T_{xx}, T_{yy}, T_{zz}, T_{zy} + T_{yz}, T_{yz} + T_{zy}, T_{zx} + T_{xz}$; die drei Komponenten des letzteren sind: $T_{xy} - T_{yx}, T_{yz} - T_{zy}, T_{zx} - T_{xz}$. Die Darstellung dieses ist zu $\mathfrak{D}^{(1)}$ äquivalent und irreduzibel, nicht aber die des symmetrischen Tensors. In der Tat existiert eine Linearkombination $T_{xx} + T_{yy} + T_{zz} = T$ seiner Komponenten, die invariant ist. Die übriggebliebenen fünf Linearkombinationen $T_{xx} - \frac{1}{3}T, T_{yy} - \frac{1}{3}T, T_{xy} + T_{yz}, T_{yz} + T_{xy}, T_{zx} + T_{xz}$ sind die voneinander unabhängigen Komponenten eines symmetrischen Tensors mit der Spur Null, zu ihnen gehört eine mit $\mathfrak{D}^{(2)}$ äquivalente, also irreduzible Darstellung.

Man sieht hieraus auch, warum es nicht zweckmäßig ist, die Zeilen und Spalten der irreduziblen Darstellungen nach den gewöhnlichen Zeichen der Tensorkomponenten zu benennen, auf die sie sich beziehen: die Freiheit ist dabei noch zu groß. Z. B. kann man bei dem vorher angegebenen symmetrischen Tensor mit der Spur 0 die Komponente $T_{xx} - \frac{1}{3}T$ weglassen und an ihrer Stelle $T_{zz} - \frac{1}{3}T$ nehmen.

Auch bei $\mathfrak{D}^{(1)}$ beziehen sich die drei Zeilen nicht auf die X, Y, Z -Komponenten eines Vektors — dann wäre ja $\mathfrak{D}^{(1)}$ rein reell. $\mathfrak{D}^{(1)}$ gibt vielmehr an, wie sich die

$$\left. \begin{array}{l} T_{-1} = \frac{i}{\sqrt{2}} X + \frac{1}{\sqrt{2}} Y, \\ T_0 = \qquad \qquad \qquad Z, \\ T_1 = \frac{i}{\sqrt{2}} X - \frac{1}{\sqrt{2}} Y \end{array} \right\} \quad (34)$$

Komponenten eines Vektors transformieren. Durch die in (34) vorkommende Matrix kann man $\mathfrak{D}^{(1)}$ in die für die X, Y, Z -Komponenten eines Vektors gültige Darstellung, d. h. in R selber transformieren. Man überzeugt sich hiervon, indem man aus (29) $\mathfrak{D}^{(1)}(\{\alpha, 0, 0\})$ und $\mathfrak{D}^{(1)}(\{0, \beta, 0\})$ entnimmt und diese mit der Transformation in (34) von rechts und ihrer Adjungierten von links multipliziert: man erhält im ersten Falle die Matrix (14a'), im zweiten Falle (14b').

XVI. Die Darstellungen des direkten Produktes

1. In den meisten physikalischen Problemen ist nicht eine einzige Symmetrieart, sondern mehrere Arten von Symmetrien vorhanden, die nebeneinander einherlaufen, ohne sich wesentlich zu beeinflussen. Zum Beispiel haben wir im Falle eines Wasserstoffmoleküls eine Differentialgleichung

$$-\frac{\hbar^2}{8\pi^2} \left(\frac{1}{M} \sum_{k=1}^6 \frac{\partial^2}{\partial X_k^2} + \frac{1}{m} \sum_{k=1}^{30} \frac{\partial^2}{\partial x_k^2} \right) \psi + V \psi = E \psi. \quad (*)$$

M bedeutet die Masse der Wasserstoffkerne, X_1, \dots, X_6 ihre Cartesischen Koordinaten, m die Masse der Elektronen, x_1, \dots, x_{30} ihre Cartesischen Koordinaten, der Sauerstoffkern ist wegen seiner großen Masse als ruhendes Anziehungszentrum betrachtet, die von ihm herrührende potentielle Energie in V miteinbezogen. Das Problem (*) hat mehrere Arten von Symmetrien: erstens kann man die Koordinaten der Wasserstoffkerne vertauschen, zweitens die der Elektronen, drittens das ganze System einer Drehung unterwerfen. Hierbei kommt sogar nicht nur die reine, sondern die ganze Drehspiegelungsgruppe in Frage. Es fragt sich also, wie diese Symmetrieeigenschaften nebeneinander zu berücksichtigen sind?

2. Die drei soeben aufgezählten Arten von Operationen haben die Eigenschaft, daß die Operatoren der einen Art mit den Operatoren der anderen Arten vertauschbar sind. Es ist ja offenbar gleichgültig, ob man die Koordinaten der Teilchen zuerst vertauscht und dann eine Drehung ausführt, oder ob man zuerst die Drehung und dann die Vertauschung vornimmt. Es sei daher angenommen, daß die Elemente jeder Operatorengruppe mit allen Elementen der mit ihr zu verknüpfenden Operatorengruppen vertauschbar sind.

Betrachten wir zuerst den Fall, daß (*) nur zwei Gruppen gegenüber invariant ist. Die Elemente der beiden Gruppen seien E', A_2, A_3, \dots, A_n bzw. $E'', B_2, B_3, \dots, B_m$. Dann ist (*) nicht nur den Operatoren $P_{E'} = 1, P_{A_2}, P_{A_3}, \dots, P_{A_n}$ und $P_{E''} = 1, P_{B_2}, P_{B_3}, \dots, P_{B_m}$, sondern allen nm Produkten $P_{A_x} P_{B_\lambda}$ dieser Operatoren gegenüber invariant, wo wegen der oben erwähnten Vertauschbarkeit $P_{A_x} P_{B_\lambda} = P_{B_\lambda} P_{A_x}$ gilt. Die $P_{A_x} P_{B_\lambda}$ bilden nach dem Operatorenmultiplikationsgesetz eine Gruppe, weil das Produkt von zweien wieder unter ihnen vorkommt:

$$P_{A_x} P_{B_\lambda} \cdot P_{A_{x'}} P_{B_{\lambda'}} = P_{A_x} P_{A_{x'}} P_{B_\lambda} P_{B_{\lambda'}} = P_{A_x A_{x'}} P_{B_\lambda B_{\lambda'}}. \quad (1)$$

Die Einheit dieser Gruppe ist der Einheitsoperator $\mathbf{P}_{E'} \mathbf{P}_{E''} = 1$. Man nennt sie das direkte Produkt der Gruppe der \mathbf{P}_{A_x} und der Gruppe der \mathbf{P}_{B_λ} , sie bildet die gesamte Symmetriegruppe von (*).

Allgemein enthält das direkte Produkt zweier Gruppen E' , A_2 , A_3, \dots, A_n und E'' , B_2 , B_3, \dots, B_m als Elemente die Paare $A_x B_\lambda$ der beiden „Faktoren“, d. h. der Gruppen, aus denen es aufgebaut ist. Das Multiplikationsgesetz ist dabei

$$A_x B_\lambda \cdot A_{x'} B_{\lambda'} = A_x A_{x'} \cdot B_\lambda B_{\lambda'} = A_{x''} B_{\lambda''}, \quad (1a)$$

wo $A_{x''} = A_x A_{x'}$ und $B_{\lambda''} = B_\lambda B_{\lambda'}$ ist. Man schreibt dabei noch A_x für $A_x E''$ und B_λ für $E' B_\lambda$. Der Vergleich von (1) und (1a) zeigt, daß die Gruppe der $A_x B_\lambda$ zur Gruppe der $\mathbf{P}_{A_x} \mathbf{P}_{B_\lambda}$ holomorph ist, und wir schreiben daher, wie gewöhnlich, $\mathbf{P}_{A_x} \mathbf{P}_{B_\lambda} = \mathbf{P}_{A_x B_\lambda}$. An Stelle der Darstellungen der Gruppe der $\mathbf{P}_{A_x} \mathbf{P}_{B_\lambda}$ können wir auch die Darstellungen der Gruppe der $A_x B_\lambda$ untersuchen.

3. Suchen wir eine Darstellung dieser Gruppe, so liegt es nahe, dem Element $A_x B_\lambda$ das direkte Produkt der Matrizen $\mathbf{a}(A_x)$ und $\mathbf{b}(B_\lambda)$ zuzuordnen, die in irgendwelchen Darstellungen der einzelnen „Faktoren“ dem Element A_x bzw. B_λ zugeordnet sind. In der Tat bilden die Matrizen $\mathbf{d}(A_x B_\lambda) = \mathbf{a}(A_x) \times \mathbf{b}(B_\lambda)$ eine Darstellung des direkten Produkts, es ist [vgl. (7) Kap. II]:

$$\begin{aligned} \mathbf{a}(A_x) \times \mathbf{b}(B_\lambda) \cdot \mathbf{a}(A_{x'}) \times \mathbf{b}(B_{\lambda'}) &= \mathbf{a}(A_x) \cdot \mathbf{a}(A_{x'}) \times \mathbf{b}(B_\lambda) \cdot \mathbf{b}(B_{\lambda'}) \\ &= \mathbf{a}(A_x A_{x'}) \times \mathbf{b}(B_\lambda B_{\lambda'}). \end{aligned} \quad (2)$$

Das Produkt der zu den Elementen $A_x B_\lambda$ und $A_{x'} B_{\lambda'}$ zugeordneten Matrizen $\mathbf{a}(A_x) \times \mathbf{b}(B_\lambda)$ und $\mathbf{a}(A_{x'}) \times \mathbf{b}(B_{\lambda'})$ ergibt die zum Element $A_x A_{x'} B_\lambda B_{\lambda'} = A_x B_\lambda \cdot A_{x'} B_{\lambda'}$ zugeordnete Matrix.

Die Koeffizienten der Matrix $\mathbf{d}(A_x B_\lambda) = \mathbf{a}(A_x) \times \mathbf{b}(B_\lambda)$ sind

$$\mathbf{d}(A_x B_\lambda)_{\varrho' \sigma'; \varrho \sigma} = \mathbf{a}(A_x)_{\varrho' \varrho} \mathbf{b}(B_\lambda)_{\sigma' \sigma}. \quad (2a)$$

Sind $\mathbf{a}(A_x)$ und $\mathbf{b}(B_\lambda)$ irreduzibel, so ist es auch $\mathbf{d}(A_x B_\lambda)$. Ist nämlich eine Matrix $(\mathbf{M}_{\varrho' \sigma'; \varrho \sigma})$ vertauschbar mit den $\mathbf{d}(A_x B_\lambda)$, so gilt für alle x und λ

$$\sum_{\varrho \sigma} \mathbf{M}_{\varrho' \sigma'; \varrho \sigma} \mathbf{a}(A_x)_{\varrho \varrho''} \mathbf{b}(B_\lambda)_{\sigma \sigma''} = \sum_{\varrho \sigma} \mathbf{a}(A_x)_{\varrho' \varrho} \mathbf{b}(B_\lambda)_{\sigma' \sigma} \mathbf{M}_{\varrho \sigma; \varrho'' \sigma''}. \quad (3)$$

Setzen wir insbesondere zuerst $A_x = E'$, dann $B_\lambda = E''$, so sind $\mathbf{a}(E')$ bzw. $\mathbf{b}(E'')$ Einheitsmatrizen und (3) schreibt sich

$$\sum_{\sigma} \mathbf{M}_{\varrho' \sigma'; \varrho \sigma} \mathbf{b}(B_\lambda)_{\sigma \sigma''} = \sum_{\sigma} \mathbf{b}(B_\lambda)_{\sigma' \sigma} \mathbf{M}_{\varrho' \sigma; \varrho'' \sigma''}, \quad (3a)$$

$$\sum_{\varrho} \mathbf{M}_{\varrho' \sigma'; \varrho \sigma''} \mathbf{a}(A_x)_{\varrho \varrho''} = \sum_{\varrho} \mathbf{a}(A_x)_{\varrho' \varrho} \mathbf{M}_{\varrho \sigma'; \varrho'' \sigma''}. \quad (3b)$$

Die Untermatrizen

$$\begin{pmatrix} \mathbf{M}_{\varrho'1; \varrho''1} & \mathbf{M}_{\varrho'1; \varrho''2} & \cdots \\ \mathbf{M}_{\varrho'2; \varrho''1} & \mathbf{M}_{\varrho'2; \varrho''2} & \cdots \\ \vdots & \vdots & \ddots \end{pmatrix} \quad (\dagger)$$

sind alle (für jedes ϱ' und ϱ'') mit allen $\mathbf{b}(B_\lambda)$ vertauschbar. Ebenso sind nach (3 b) die Matrizen

$$\begin{pmatrix} \mathbf{M}_{1\sigma'; 1\sigma''} & \mathbf{M}_{1\sigma'; 2\sigma''} & \cdots \\ \mathbf{M}_{2\sigma'; 1\sigma''} & \mathbf{M}_{2\sigma'; 2\sigma''} & \cdots \\ \vdots & \vdots & \ddots \end{pmatrix} \quad (\dagger\dagger)$$

für alle σ', σ'' mit allen $\mathbf{a}(A_\alpha)$ vertauschbar, sowohl (\dagger) wie auch ($\dagger\dagger$) sind konstante Matrizen. Es folgt

$$\mathbf{M}_{\varrho'\sigma'; \varrho''\sigma''} = \delta_{\sigma'\sigma''} \mathbf{M}_{\varrho'1; \varrho''1}, \quad (4a)$$

$$\mathbf{M}_{\varrho'\sigma'; \varrho''\sigma''} = \delta_{\varrho'\varrho''} \mathbf{M}_{1\sigma'; 1\sigma''}, \quad (4b)$$

woraus sich

$$\mathbf{M}_{\varrho'\sigma'; \varrho''\sigma''} = \delta_{\sigma'\sigma''} \mathbf{M}_{\varrho'1; \varrho''1} = \delta_{\sigma'\sigma''} \delta_{\varrho'\varrho''} \mathbf{M}_{11; 11} \quad (4)$$

ergibt. Die Matrix \mathbf{M} muß selber eine konstante Matrix sein, $\mathbf{d}(A_\alpha B_\lambda)$ ist irreduzibel.

4. Wir haben jetzt eine Methode, mit welcher man irreduzible Darstellungen einer Gruppe gewinnen kann, die das direkte Produkt zweier Gruppen ist, vorausgesetzt, daß die irreduziblen Darstellungen der „Faktoren“ bekannt sind. Es fragt sich noch, ob man auf diese Weise auch alle irreduziblen Darstellungen des direkten Produktes erhalten kann?

Die Dimensionen der irreduziblen Darstellungen der Gruppe der A seien mit g_1, g_2, \dots , die der Darstellungen der B mit h_1, h_2, \dots bezeichnet. Wenn wir jede Darstellung der ersten Gruppe mit jeder Darstellung der zweiten kombinieren, erhalten wir irreduzible Darstellungen des direkten Produkts mit den Dimensionen $g_1 h_1, g_1 h_2, \dots, g_2 h_1, g_2 h_2, \dots$. Nehmen wir im Sinne eines im IX. Kapitel (S. 91) erwähnten Satzes an, daß die Summe der Quadrate der Dimensionen aller irreduziblen Darstellungen einer Gruppe gleich ihrer Ordnung ist, so ist

$$g_1^2 + g_2^2 + \cdots = n \text{ und } h_1^2 + h_2^2 + \cdots = m,$$

wo n und m die Ordnungen der Gruppen der A bzw. B sind. Die Summe der Quadrate der Dimensionen der durch direkte Produktbildung erhaltenen Darstellungen der Gruppe der $A_x B_1$ ist

$$(g_1 h_1)^2 + (g_1 h_2)^2 + \cdots + (g_2 h_1)^2 + (g_2 h_2)^2 + \cdots + \cdots = g_1^2 m + g_2^2 m + \cdots = nm$$

gleich der Ordnung des direkten Produktes. Hieraus folgt, daß die angegebene Methode tatsächlich alle irreduziblen Darstellungen liefert¹⁾.

Man kann dieser Überlegung auch folgende Form geben, in der sie auch für kontinuierliche Gruppen gilt: Die $g_1^2 + g_2^2 + \cdots$ Koeffizienten der erstenen Darstellungen bilden, als Funktionen der A betrachtet²⁾, ein vollständiges Funktionensystem für die Funktionen der A . Ähnlich bilden die $h_1^2 + h_2^2 + \cdots$ Darstellungskoeffizienten, als Funktionen der B betrachtet, ein vollständiges Funktionensystem der B . Daher bilden alle Produkte beider Funktionensysteme ein vollständiges Funktionensystem beider Variablen.

5. Die Eigenwerte der Differentialgleichung (*) lassen sich in qualitativ verschiedene Klassen einteilen: zu jedem Eigenwert gehört eine Darstellung der Gruppe der Operatoren, die (*) invariant lassen. Zur Charakterisierung der irreduziblen Darstellungen dieser Gruppe (des direkten Produkts der erwähnten drei Gruppen) bedient man sich am besten der drei Symbole, die die drei irreduziblen Darstellungen charakterisieren, aus denen diese Darstellung zusammengesetzt ist. Man wird so von einem Eigenwert von (*) sagen können: er gehört zur symmetrischen Darstellung der Vertauschung der H -Kerne, zur antisymmetrischen der Vertauschung der zehn Elektronen und zur siebendimensionalen der Drehgruppe, und versteht darunter, daß er zu derjenigen Darstellung des direkten Produkts dieser drei Gruppen gehört, die sich aus den aufgezählten Darstellungen der „Faktoren“ zusammensetzt.

Was die Eigenfunktionen eines solchen Eigenwertes anbelangt, so tragen sie — den Zeilen des direkten Produkts dreier Matrizen entsprechend — drei Indizes, die jeweils angeben, zu welcher Zeile

¹⁾ Es gehen hier zwei „direkte Produkte“ durcheinander, die wesentlich verschieden sind: das direkte Produkt zweier Gruppen und das direkte Produkt zweier Matrizen. Die Elemente des direkten Produkts der Gruppen sind die $A_x B_1$. Dem $A_x B_1$ ordnet die Darstellung $a(A_x) \times b(B_1)$ das direkte Produkt von $a(A_x)$ und $b(B_1)$ zu.

²⁾ Eine Funktion der A bedeutet die Zuordnung je einer Zahl J_{A_x} zu jedem Gruppenelement A_x .

der in Frage kommenden Darstellungen der ersten, zweiten und dritten Gruppe sie gehören. Daß zwei Eigenfunktionen, deren Indizes nicht alle drei übereinstimmen, aufeinander orthogonal sind, auch wenn man auf sie beliebige symmetrische Operatoren anwendet, sieht man erstens daraus, daß sie zu verschiedenen Zeilen der Darstellung des direkten Produkts gehören, oder auch daraus — wenn etwa ihre zweiten Indizes verschieden sind —, daß sie zu verschiedenen Zeilen der Darstellung der zweiten Gruppe gehören.

Wendet man auf eine Funktion, die zur $\varrho\sigma$ -Zeile der Darstellung $\alpha(A) \times b(B)$ des direkten Produkts zweier Gruppen gehört, einen Operator $P_A = P_A P_{E''}$ der ersten Gruppe an, so erhält man eine Funktion, welche durch die den 1 σ , 2 σ , 3 σ , ... Zeilen der Darstellung $\alpha(A) \times b(B)$ entsprechenden Funktionen allein ausgedrückt werden kann. Und zwar sind die Koeffizienten dieselben, wie wenn die zweite Gruppe gar nicht da wäre:

$$\begin{aligned} P_A P_{E''} \psi_{\varrho\sigma} &= \sum_{\varrho' \sigma'} \alpha(A)_{\varrho' \varrho} b(E'')_{\sigma' \sigma} \psi_{\varrho' \sigma'} \\ &= \sum_{\varrho' \sigma'} \alpha(A)_{\varrho' \varrho} \delta_{\sigma' \sigma} \psi_{\varrho' \sigma'} = \sum_{\varrho'} \alpha(A)_{\varrho' \varrho} \psi_{\varrho' \sigma}. \end{aligned}$$

Eine Funktion, die zur $\varrho\sigma$ -Zeile der Darstellung $\alpha(A) \times b(B)$ gehört, gehört also auch zur ϱ -Zeile von $\alpha(A)$ [und ebenso zur σ -Zeile von $b(B)$] und hat demgemäß alle Eigenschaften dieser beiden Funktionenklassen.

6. Man muß in der *Störungstheorie* bei der Bildung der „richtigen Linearkombinationen“ aus einer Schar von Funktionen solche Linearkombinationen bilden, die je zu einer Zeile $\varrho\sigma$ einer Darstellung $\alpha(A) \times b(B)$ des direkten Produktes der vorhandenen Symmetriegruppen gehören. Man kann hierbei so verfahren, daß man zunächst solche Linearkombinationen f_1, f_2, f_3, \dots bildet, die zur ϱ -Zeile von $\alpha(A)$ gehören: jede Funktion $\psi_{\varrho\sigma}$ die zur $\varrho\sigma$ -Zeile von $\alpha(A) \times b(B)$ gehört, muß eine Linearkombination der f_1, f_2, f_3, \dots sein. Würde nämlich $\psi_{\varrho\sigma}$ auch solche Funktionen f'_1, f'_2, f'_3, \dots enthalten, die nicht zur Darstellung $\alpha(A)$ oder nicht zu ihrer ϱ -Zeile gehören:

$$\psi_{\varrho\sigma} = c_1 f_1 + c_2 f_2 + c_3 f_3 + \dots + c'_1 f'_1 + c'_2 f'_2 + c'_3 f'_3 \dots, \quad (5)$$

so müßte doch $c'_1 f'_1 + c'_2 f'_2 + c'_3 f'_3 + \dots = 0$ sein. Wenn man nämlich in (5) $c_1 f_1 + c_2 f_2 + c_3 f_3 + \dots$ auf die linke Seite bringt, so gehört die ganze linke Seite zur ϱ -Zeile von $\alpha(A)$ und ist daher

zu allen Gliedern der rechten Seite orthogonal, so daß beide Seiten verschwinden müssen.

7. Wir wollen die Sätze über die Darstellungen des direkten Produkts zur Bestimmung der irreduziblen Darstellungen der dreidimensionalen Drehspiegelungsgruppe benutzen. Die Drehspiegelungsgruppe ist die Gruppe der reellen orthogonalen dreidimensionalen Matrizen mit der Determinante ± 1 . Sie ist das direkte Produkt der reinen Drehgruppe und der mit der Spiegelungsgruppe holomorphen Gruppe, die aus der Einheit E und der Inversion I besteht

$$E = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}; \quad I = \begin{pmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix}.$$

In der Tat entsteht jede orthogonale reelle Matrix aus einer reinen Drehung durch Multiplikation mit E oder I : entweder ist ihre Determinante schon $+1$, dann ist sie schon eine reine Drehung, und wenn ihre Determinante -1 ist, so entsteht sie aus einer reinen Drehung durch Multiplikation mit I . Auch ist es klar, daß E und I mit allen Matrizen der reinen Drehgruppe (sogar mit allen Matrizen überhaupt) vertauschbar sind.

Die Spiegelungsgruppe hat zwei irreduzible Darstellungen: die identische (auch positive genannt) und diejenige, bei der der Einheit die Matrix (1) und I die Matrix (-1) zugeordnet ist. Daher entstehen aus jeder Darstellung $\mathfrak{D}^{(l)}(R)$ der reinen Drehgruppe zwei Darstellungen der Drehspiegelungsgruppe: die Kombinationen von $\mathfrak{D}^{(l)}(R)$ mit der positiven und mit der negativen Darstellung der Spiegelungsgruppe.

Die dreidimensionale Drehspiegelungsgruppe hat je zwei (ein-deutige) irreduzible Darstellungen der Dimensionen $1, 3, 5, \dots$, die man der Reihe nach mit $l = 0_+, 0_-, 1_+, 1_-, 2_+, 2_-, \dots$ bezeichnet. Beide Darstellungen l_+ und l_- sind $2l+1$ -dimensional, in beiden entsprechen reinen Drehungen dieselben Matrizen $\mathfrak{D}^{(l)}(R)$, die ihnen in der $2l+1$ -dimensionalen Darstellung der reinen Drehgruppe entsprechen. In l_+ entspricht auch der Drehspiegelung IR die Matrix $\mathfrak{D}^{(l)}(R)$, die R entspricht, in l_- dagegen entspricht IR die Matrix $-\mathfrak{D}^{(l)}(R)$.

XVII. Die Grundzüge der Atomspektren

1. Wir wollen jetzt unsere Ergebnisse vor allem zur Erklärung der wichtigsten Züge der Atomspektren benutzen. Bevor die Einzelheiten besprochen werden, sei das Ganze einmal in diesem Kapitel programmatisch zusammengefaßt, wobei zunächst auf die Beweise verzichtet werde. Ich hoffe, daß hierdurch auch eine bessere Übersicht über die Gesetzmäßigkeiten selber, wie sie uns von der experimentellen Forschung¹⁾ geliefert wurden, gewonnen werden kann.

Bevor man an die eigentliche Auflösung der Schrödinger-Gleichung herangeht, muß man zuerst die „Schwerpunktskoordinaten abseparieren“.

Die Schrödinger-Gleichung hat in ihrer unveränderten Gestalt [(5a) Kap. IV] nur ein kontinuierliches Spektrum, dem Umstand entsprechend, daß das Atom neben seiner Anregungsenergie als ganzes eine beliebig stetig veränderliche kinetische Energie aufnehmen kann. Will man sich — wie das praktisch immer der Fall ist — auf die Anregungsenergie beschränken, so muß man annehmen, daß das Atom in Ruhe ist, d. h. die Wellenfunktion von den Schwerpunktskoordinaten unabhängig ist und nur von den Koordinatendifferenzen der einzelnen Teilchen abhängt²⁾. Da die Elektronenmasse der Kernmasse gegenüber vernachlässigt werden kann, identifiziert man gewöhnlich die Koordinaten des Kerns mit denen des Schwerpunkts und nimmt an, daß die Wellenfunktion von den Kernkoordinaten unabhängig ist. Man führt diese gewöhnlich gar nicht erst in die Schrödinger-Gleichung ein, sondern betrachtet den Kern als festes Anziehungs-Zentrum, in dessen Feld sich die Elektronen bewegen. Dies ist natürlich nur bei Atomen, bei Gebilden mit nur einem Kern möglich. Die späteren, allgemeinen Entwicklungen sind bis auf die Aufspaltungsbilder in äußeren Feldern von dieser Annahme der Vernachlässigung der „Mitbewegung des Kerns“ unabhängig. Die Auffassung ist dabei die, daß die Wellenfunktion als Variable alle Koordinaten enthält, nur von denen des Schwerpunkts unabhängig ist, also längs Linien, deren Punkte denselben, nur im Raum verlagerten Konfigurationen entsprechen, konstant ist. Dies ist als eine Nebenbedingung anzusehen. Damit das skalare Produkt zweier Funktionen nicht divergiert, muß dann allerdings der Konfigurationsraum auf ein endliches Gebiet abgegrenzt werden, das aber beliebig groß sein kann. Doch spricht man der einfacheren Ausdrucksweise halber öfters so, wie wenn die Wellenfunktion die Koordinaten des Kerns als Variable gar nicht enthielte.

¹⁾ Eine ausführlichere vortreffliche Darstellung findet man in F. Hund's Büchlein: Linienspektren und periodisches System, Berlin 1927, und auch bei L. Pauling und S. Goudsmit: The Structure of Line Spectra. New York 1930.

²⁾ Das ist auch der Grund, warum man bei diesen Problemen nicht auch die Translationsgruppe als eine Symmetriegruppe des Systems einführt: alle Wellenfunktionen sollen Translationen gegenüber invariant sein, also zur identischen Darstellung der Translationsgruppe gehören.

Das einfachste Spektrum hat das Wasserstoffatom, da es aus einem einzigen Elektron besteht, das sich — bei Vernachlässigung der Mitbewegung des Kerns — in einem konstanten Potentialfeld bewegt.

Die Schrödinger-Gleichung lautet

$$\left[-\frac{\hbar^2}{8\pi^2 m} \left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2} \right) - \frac{e^2}{\sqrt{x^2 + y^2 + z^2}} \right] \psi(x, y, z) = E \psi(x, y, z) \quad (1)$$

und läßt sich noch streng lösen. Man erhält so das Spektrum, die möglichen Energiewerte (die „Terme“, wie man sie in der Spektroskopie nennt) und die Eigenfunktionen, d. h. die stationären Zustände des Wasserstoffatoms. Das Spektrum hat einen diskreten Teil mit den Termwerten $E = -Rhc/1^2, -Rhc/2^2, -Rhc/3^2, \dots$, wo R die sogenannte Rydbergkonstante ist:

$$E_N = -\frac{2\pi^2 m e^4}{\hbar^2 N^2} = -\frac{Rhc}{N^2} = -2,15 \cdot 10^{-11} \frac{1}{N^2} \text{ erg.} \quad (2)$$

Die Energien sind negativ, dem entsprechend, daß das Elektron in der Nähe des Kerns ein beträchtliches negatives Potential hat, wogegen man Arbeit aufwenden muß, um es ins Unendliche, an die Stelle mit dem Potential Null zu bringen. Die Abstände zwischen den einzelnen Termen werden mit wachsender Hauptquantenzahl N immer geringer, schließlich konvergiert die Energie für unendlich hohe Laufzahlen zu Null. Physikalisch entspricht dem ein immer stärkeres und stärkeres Enttreiben des Elektrons der Anziehungsphäre des Kerns; wenn das Elektron ganz befreit ist, hat es die Energie Null.

An das diskrete Spektrum (2) schließt sich ein kontinuierliches Spektrum an, das die ganze positive Zahlengerade bedeckt. In den entsprechenden Zuständen ist das Wasserstoffatom ionisiert, die zugehörige positive Energie bedeutet die kinetische Energie des Elektrons, nachdem es sich ins Unendliche entfernt hat. Im kontinuierlichen Spektrum hat man keine im eigentlichen Sinne stationären Zustände: das Elektron entfernt sich nach hinreichend langer Zeit beliebig weit vom Kern. Einem Zustand entspricht ja auch mathematisch eine normierte Wellenfunktion, die Eigenfunktionen des kontinuierlichen Spektrums lassen sich aber nicht normieren.

Das Auftreten von Serien der ungefähren Gestalt (2) mit einer Konvergenzstelle im endlichen und einem daranschließenden, dem

ionisierten Zustand entsprechenden kontinuierlichen Spektrum ist für alle Atomspektren charakteristisch.

Die Eigenwerte (2) sind entartet, d. h. es gehören nicht nur je eine, sondern mehrere linear unabhängige Eigenfunktionen zu jedem Eigenwert. Der Eigenwert mit der Laufzahl („Hauptquantenzahl“) N ist N^2 -fach entartet.

Die zugehörigen normierten Eigenfunktionen seien zur Bequemlichkeit des Lesers hier angegeben. Man schreibt sie am besten in Polarkoordinaten r , ϑ , φ auf, sie lauten mit $\eta = 2r/Nr_0$, wo $r_0 = \hbar^2/4\pi^2 m e^2$ der „Radius der ersten Bohrschen Bahn“ ist:

$$\left. \begin{aligned} \psi_{\mu}^{Nl} &= \frac{e^{-i\mu\varphi}}{\sqrt{2\pi}} C_{l\mu} P_{\mu}^l(\vartheta) C'_{Nl} \eta^l e^{-1/2\eta} L_{N+l}^{2l+1}(\eta), \\ C_{l\mu} &= C_{l,-\mu} = \sqrt{\frac{2l+1}{2}} \sqrt{\frac{(l-\mu)!}{(l+\mu)!}} \quad \text{für } \mu \geq 0, \\ C'_{Nl} &= \sqrt{\frac{(N-l-1)!}{2N}} \left(\frac{2}{Nr_0(N+l)!} \right)^{3/2}, \end{aligned} \right\} \quad (3)$$

wenn wir zur Unterscheidung der N^2 zum Eigenwert E_N gehörigen Eigenfunktionen die Indizes l („azimutale Quantenzahl“) und μ („magnetische Quantenzahl“) einführen. Bei einem festen N kann l die Werte $0, 1, 2, \dots, N-1$ annehmen, μ läuft (unabhängig von N) von $-l$ bis $+l$. Die An-

zahl sämtlicher zu E gehöriger Eigenfunktionen ist $\sum_{l=0}^{N-1} (2l+1) = N^2$.

Die $P_{\mu}^l(\vartheta)$ sind die in Kap. XV (3a) definierten Kugelfunktionen, $C_{l\mu}$ ihre Normierungsfaktoren. Mit $L_{N+l}^{2l+1}(\eta)$ ist die $2l+1$ -te Ableitung des $N+l$ -ten Laguerreschen Polynoms L_{N+l}

$$L_{\nu}(\eta) = (-1)^{\nu} \left(\eta^{\nu} - \frac{\nu^2}{1!} \eta^{\nu-1} + \frac{\nu^2(\nu-1)^2}{2!} \eta^{\nu-2} - \dots + (-1)^{\nu} \nu! \right)$$

bezeichnet, C'_{Nl} ist ihr Normierungsfaktor.

Man erkennt schon hier, daß eine Beziehung zwischen der azimutalen Quantenzahl l und der $2l+1$ -dimensionalen Darstellung der Drehgruppe bestehen muß.

Dasselbe, was vom Wasserstoffatom gilt, gilt auch vom Heliumion, vom zweimal ionisierten Li und allen anderen Systemen, in denen nur ein Elektron und ein Kern vorhanden ist. Man muß nur die potentielle Energie in der Schrödinger-Gleichung (1) durch $-\frac{Ze^2}{r}$, wobei Z die Kernladungszahl ist, die Energieniveaus durch

$$E_N^{(Z)} = -\frac{2\pi^2 m Z^2 e^4}{\hbar^2} \frac{1}{N^2}, \quad (2a)$$

in (3) η durch

$$\eta^{(Z)} = \frac{8\pi^2 m e^3 Z}{h^2 N} r = \frac{2Z}{Nr_0} r \quad (3a)$$

ersetzen und ψ wegen der Normierung mit $Z^{3/2}$ multiplizieren.

2. Die Spektren der Atome mit mehreren — sagen wir n Elektronen lassen sich nicht mehr streng berechnen. Dies röhrt von der verhältnismäßig komplizierten Gestalt der potentiellen Energie

$$V = \sum_{i=1}^n -\frac{e^2 Z}{\sqrt{x_i^2 + y_i^2 + z_i^2}} + \frac{1}{2} \sum_{i \neq j} \frac{e^2}{\sqrt{(x_i - x_j)^2 + (y_i - y_j)^2 + (z_i - z_j)^2}} \quad (4)$$

her. Wäre das zweite Glied in (4), das der gegenseitigen Abstoßung der Elektronen Rechnung trägt, nicht vorhanden, so würden sich die Elektronen nur unter der Wirkung des konstanten Feldes des Keras bewegen. Dann ließe sich die Schrödinger-Gleichung

$$(\mathbf{H}_1 + \mathbf{H}_2 + \cdots + \mathbf{H}_n) \psi(x_1 y_1 z_1 \cdots x_n y_n z_n) = E \psi(x_1 y_1 z_1 \cdots x_n y_n z_n), \quad (5)$$

$$\mathbf{H}_k = -\frac{\hbar^2}{8\pi^2 m} \left(\frac{\partial^2}{\partial x_k^2} + \frac{\partial^2}{\partial y_k^2} + \frac{\partial^2}{\partial z_k^2} \right) - \frac{Z e^3}{\sqrt{x_k^2 + y_k^2 + z_k^2}} \quad (5a)$$

auflösen. Die Eigenwerte wären die Summen, die Eigenfunktionen die Produkte der Eigenwerte bzw. Eigenfunktionen von (5a) und ließen sich mit Hilfe von (2a), (3), (3a) ausdrücken:

$$\psi(x_1 y_1 z_1 \cdots x_n y_n z_n) = \psi_{l_1 \mu_1}^{N_1}(x_1 y_1 z_1) \cdots \psi_{l_n \mu_n}^{N_n}(x_n y_n z_n), \quad (6)$$

$$E = E_{N_1} + E_{N_2} + \cdots + E_{N_n}. \quad (6a)$$

Wenn man nämlich (6) in (5) einsetzt und $\mathbf{H}_k \psi(x_1 \dots z_n)$ bildet, so ergibt sich $E_{N_k} \psi(x_1 \dots z_n)$, weil

$$\mathbf{H}_k \psi_{l_k \mu_k}^{N_k}(x_k y_k z_k) = E_{N_k} \psi_{l_k \mu_k}^{N_k}(x_k y_k z_k)$$

ist, und die anderen Faktoren von $\psi(x_1 \dots z_n)$ bei der Anwendung von \mathbf{H}_k als Konstante zu betrachten sind.

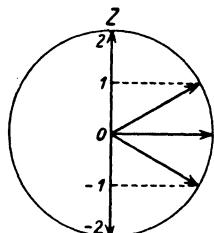
Natürlich stellt (5) eine sehr schlechte Näherung der wirklichen Schrödinger-Gleichung dar. Trotzdem pflegt man — wenigstens im Gedanken — von dieser Näherung auszugehen und die Wechselwirkung der Elektronen als eine „Störung“ zu betrachten.

Zu den meisten Eigenwerten (6 a) gehören sehr viele Eigenfunktionen, weil im Energieausdruck (6 a) die Quantenzahlen l_k, μ_k nicht vorkommen und diesen noch in (6) mehrere Werte gegeben werden können. Außerdem kann man die Hauptquantenzahlen N_k der einzelnen Elektronen beliebig vertauschen, der zugehörige Energiewert bleibt immer noch derselbe. Führt man aber die Wechselwirkung der Elektronen als Störung ein, so wird ein Teil der Entartung aufgehoben, der Term spaltet auf. Von den entstehenden Termen, die größtenteils auch noch entartet sind, kennt man rein theoretisch — abgesehen von einer ganz rohen Abschätzung ihrer Lage — nichts als ihre Symmetrieeigenschaften, die sich in den Transformationseigenschaften der zugehörigen Eigenfunktionen äußern. Diese Transformationen sind: Vertauschung der Elektronen, Ausführung einer reinen Drehung, Inversion¹⁾ (Spiegelung). Demgemäß hat jeder Term drei Darstellungen: eine der symmetrischen Gruppe, eine der reinen Drehgruppe und eine der Spiegelungsgruppe. (Die beiden letzten pflegt man auch zu einer Darstellung der Drehspiegelungsgruppe zusammenzufassen.) Die entsprechenden Quantenzahlen (Charakteristika der Darstellungen) sind²⁾

Multiplettsystem S ,
azimutale Quantenzahl L ,
Spiegelungscharakter w .

3. Die azimutale Quantenzahl hat für die verschiedenen Terme die Werte $L = 0, 1, 2, 3, \dots$. Die entsprechenden Eigenwerte gehören zu den Darstellungen $\mathfrak{D}^{(0)}(R), \mathfrak{D}^{(1)}(R), \mathfrak{D}^{(2)}(R), \dots$ der Drehgruppe³⁾. Man nennt sie der Reihe nach S, P, D, F, \dots Terme. Zu einem S -Term gehört nur eine Eigenfunktion, zu einem P -Term gehören drei, zu einem D -Term fünf usw. Eigenfunktionen. Die $2L + 1$ Eigenfunktionen, die zu einem Term mit der Azimutalquantenzahl L

Abb. 7. Ist der Gesamt-drehimpuls 2, so kann seine Z-Komponente die Werte 2, 1, 0, -1 oder -2 haben (alles in Einheiten von $\hbar/2\pi$)



¹⁾ Das heißt das Umkehren des Vorzeichens aller Koordinaten x_1, \dots, z_n .

²⁾ Es ist üblich, die Quantenzahlen von ganzen Atomen mit großen Buchstaben zu bezeichnen, während für die einzelner Elektronen kleine Buchstaben gebräuchlich sind.

³⁾ Daß dies auch für das $\psi_{l\mu}^N$ gilt, wird (9b), Kap. XIX, zeigen: man sieht dort, daß es zur μ -Zeile von $\mathfrak{D}^{(l)}$ gehört.

gehören, unterscheidet man durch ihre magnetischen Quantenzahlen m , die ebenfalls ganzzahlige Werte annehmen können, und zwar läuft m von $-L$ bis L . Die zugehörige Eigenfunktion gehört zur m -Zeile der irreduziblen Darstellung $\mathbf{D}^{(L)}$.

Der physikalische Sinn der azimutalen Quantenzahl ist der gesamte Drehimpuls. Der magnetischen Quantenzahl dagegen entspricht die Komponente des Drehimpulses in der Z -Achse. Damit scheint eine Raumrichtung überflüssigerweise ausgezeichnet zu sein. Dies röhrt daher, daß wir, um Funktionen zu definieren, die zu einer Zeile einer Darstellung gehören, die Darstellung ganz (nicht nur bis auf eine Ähnlichkeitstransformation) festlegen mußten. Diese Festlegung geschah so, daß angenommen wurde, daß die Matrizen, die Drehungen um die Z -Achse entsprechen, auf Diagonalform gebracht sind [(6) Kap. XV)]. In der Aussage, daß alle Eigenfunktionen eines D -Terms zur Darstellung $\mathbf{D}^{(2)}(R)$ gehören, ist dagegen noch keine Richtungsauszeichnung vorhanden.

Terme, die zur identischen („positiven“) Darstellung der Spiegelungsgruppe gehören, nennt man positive Terme, die anderen negative. Dieser für das Verstehen der Spektren sehr wichtige Begriff hat kein Analogon in der klassischen Theorie, wie etwa der Drehimpuls das Analogon der Azimutalquantenzahl ist. Den Spiegelungscharakter eines Terms fügt man als Index an das Zeichen des Terms an: S_+ , S_- , P_+ , P_- , ... Die zugehörigen Darstellungen der dreidimensionalen Drehspiegelungsgruppe sind 0_+ , 0_- , 1_+ , 1_- , ... Die gewöhnlichsten Terme sind S_+ , P_- , D_+ , F_- , ... usw. Terme¹⁾.

Auch S , das Multiplettsystem, ist ein Begriff, der der klassischen Theorie eigentlich fremd ist. Jeder Term eines n Elektronenproblems hat eine Darstellung der symmetrischen Gruppe n -ten Grades gegenüber. Die Darstellungen kommen zwar nicht alle vor, vielmehr sind in der Natur — aus Gründen, die erst später bei der Besprechung des rotierenden Elektrons und Pauli-prinzips auseinandergesetzt werden können — nur die assoziierten Darstellungen der in XIII. Kapitel mit $\mathbf{D}^{(0)}$, $\mathbf{D}^{(1)}$, $\mathbf{D}^{(2)}$, ..., $\mathbf{D}^{(1/2 \cdot n)}$ (bei gerader Elektronenzahl) oder $\mathbf{D}^{[1/2 \cdot (n-1)]}$ (bei ungerader Elektronenzahl) bezeichneten Darstellungen vorhanden. Bei gerader Elektronenzahl hat ein Term mit $S = 0$ die Dar-

¹⁾ Man nennt sie auch „ungestrichene“ Terme, während S_- , P_+ , D_- , F , ... „gestrichene“ Terme sind.

stellung $\bar{\mathbf{D}}^{(1/2 \cdot n)}$, mit $S = 1$ die Darstellung $\bar{\mathbf{D}}^{(1/2 \cdot n - 1)}$, die zu $S = \frac{1}{2}n$ gehörige Darstellung ist $\bar{\mathbf{D}}^{(0)}$. Um die S -Werte aus der Darstellung direkter ablesen zu können und auch eine Verwechslung mit den Darstellungen der Drehgruppe zu vermeiden, schreiben wir von nun an

$$\bar{\mathbf{D}}^{(k)} = \bar{\mathbf{A}}^{(S)}, \text{ wo } S = \frac{1}{2}n - k. \quad (7)$$

Die Größe S kann bei Atomen mit gerader Elektronenzahl n die Werte $0, 1, 2, \dots, \frac{1}{2}n$ annehmen. Bei ungerader Elektronenzahl werde (7) ebenfalls beibehalten, die möglichen S -Werte sind dann $\frac{1}{2}, \frac{3}{2}, \frac{5}{2}, \dots, \frac{1}{2}n$. Bei dem Wasserstoffatom ist $S = \frac{1}{2}$, die symmetrische Gruppe ersten Grades hat ja auch nur eine Darstellung. Die S -Werte des Terms bestimmen seine „Multiplizität“ $2S + 1$. Bei gerader Elektronenzahl haben wir Singulett-, Triplett-, Quintett- usw. Terme, weil $2S + 1$ die Werte $1, 3, 5, \dots$ annehmen kann, bei ungerader Elektronenzahl gibt es Dublett-, Quartett-, Sextett- usw. Terme. Die Terme des Einelektronenproblems sind alle als Dubletterme anzusprechen. Den Betrag der Multiplizität, also den Wert von $2S + 1$, fügt man dem Zeichen des Terms links oben an: ${}^1S_+$ ist ein positiver Singulett- S -Term, ${}^2P_-$ ein negativer Dublett- P -Term usw. Die zur antisymmetrischen Darstellung $\bar{\mathbf{D}}^{(0)} = \bar{\mathbf{A}}^{(1/2 \cdot n)}$ gehörigen Terme haben die höchste Multiplizität $n + 1$, während bei Singulettermen $S = 0$ und die Darstellung $\bar{\mathbf{D}}^{(1/2 \cdot n)} = \bar{\mathbf{A}}^{(0)}$ ist.

Die Terme haben drei qualitative Charakteristika: S, L, w , weil sie zu verschiedenen Darstellungen $\bar{\mathbf{A}}^{(S)} \times \mathfrak{D}^{(L w)}$ des direkten Produkts der symmetrischen Gruppe und der Drehspiegelungsgruppe gehören. Da aber natürlich mehrere Terme desselben Spektrums zu derselben Darstellung gehören, muß man zu ihrer Unterscheidung noch eine Laufzahl N einführen. Ein Term E_{SLw}^N trägt dann vier Indizes: N, S, L, w . Zu ihm gehören $(2L + 1)g_S$ Eigenfunktionen, wenn g_S die Dimension der Darstellung $\bar{\mathbf{A}}^{(S)}$ ist, zu ihrer Unterscheidung muß man außer der magnetischen Quantenzahl m noch angeben, zu welcher Zeile χ der Darstellung $\bar{\mathbf{A}}^{(S)}$ sie gehören. Eine Eigenfunktion $\psi_{\chi m}^{NSLw}$ sollte also im ganzen sechs Indizes tragen, wobei man aber zumeist einige Indizes unterdrückt (namentlich das χ , das keine physikalische Bedeutung hat). Die experimentellen Merkmale der Terme mit verschiedenen S, L und w

sind wohl hinreichend bekannt; das wichtigste ist, daß optische Übergänge mit beträchtlicher Intensität nur zwischen solchen Termen auftreten, deren azimutale Quantenzahlen entweder gleich oder nur um 1 verschieden sind, die beiden Terme müssen außerdem verschiedenen Spiegelungscharakter und dasselbe Multiplettsystem haben.

Diese Interkombinationsregeln müssen sich auch aus der Quantenmechanik ergeben, und die Aufgabe des nächsten Kapitels wird ihre Ableitung sein.

4. Die Einführung der magnetischen Momente der Elektronen, des Spins (Kap. XX), wird eine tiefgreifende Modifizierung der Schrödingergleichung mit sich bringen.

Am auffallendsten äußert sich die Wirkung der Spins in der sogenannten Feinstruktur der Spektrallinien. An den Stellen, wo nach der einfachen Schrödingerschen Theorie ein Term mit der Azimutalquantenzahl L und dem Multiplettsystem S liegen sollte, findet man in Wirklichkeit ein „Multiplett“, d. h. mehrere nahe benachbarte Terme. Ihre Anzahl ist $2L + 1$ oder $2S + 1$, je nachdem, welche dieser beiden Zahlen kleiner ist: S -Terme ($L = 0$) sind immer einfache, P -Terme ($L = 1$) sind nur im Singulettsystem ($S = 0$) einfache, im Dublettsystem doppelt, im Triplett und allen höheren Systemen dreifach usw. Bei genügend hoher Azimutalquantenzahl $L \geq S$ ist die Multiplizität $2S + 1$.

Den verschiedenen Feinstrukturkomponenten eines Multipletts schreibt man verschiedene Gesamtquantenzahlen J zu

$$J = |L - S|, |L - S| + 1, \dots, L + S - 1, L + S.$$

Das sind $2L + 1$ bzw. $2S + 1$ Zahlen, je nachdem L kleiner oder größer als S ist. Die Gesamtquantenzahl spielt die Rolle des gesamten Drehimpulses, in dem auch der vom Elektronenspin herrührende Drehimpuls enthalten ist.

Die Auswahlregeln für L , S und w sollen für alle $2L + 1$ bzw. $2S + 1$ Terme des Multipletts gelten¹⁾. Dazu kommt noch die Auswahlregel für J , die ähnlich der Regel für L ist: J ändert sich bei einem optischen Übergang um ± 1 oder 0, der Übergang zwischen zwei Termen mit $J = 0$ ist auch verboten.

¹⁾ Die beiden ersten Regeln gelten allerdings nur, solange die Spinkräfte klein sind.

5. Kehren wir jetzt zu den Entwicklungen zurück, die wir am Ende von 2. unterbrochen haben! Wir haben dort eine vereinfachte Schrödinger-Gleichung (5) aufgestellt, deren Lösung (6), (6a) unmittelbar hingeschrieben werden konnte. Die Eigenwerte waren im allgemeinen sehr vielfach entartet, und es wurde schon darauf hingewiesen, daß bei der Berücksichtigung der Wechselwirkung der Elektronen, bei dem Übergang zum richtigen Potential (4) etwa mit Hilfe des Rayleigh-Schrödingerschen Verfahrens, die Eigenwerte (6a) aufgespalten werden und solche Terme entstehen, deren charakteristische Merkmale (S, L, w) soeben besprochen wurden. Die Bestimmung der Anzahl und Art der Terme, die aus einem Term (6a) entstehen, nennt man das Aufbauprinzip.

Bei der Ableitung des Aufbauprinzips darf es nicht vergessen werden, daß die konsequente Anwendung der Schrödinger-Gleichung auch die Energiewerte solcher Zustände liefert, die in der Natur wegen des Pauliprinzips nicht vorkommen. Wir werden aber nur die Anzahl derjenigen Terme bestimmen, die wirklich existieren. (Das sind, wenn man den Spin nicht berücksichtigt, Terme mit den Darstellungen $\bar{D}^{(k)} = \bar{A}^{(1/2, n-k)}$, und wenn man auch den Spin berücksichtigt, nur Eigenwerte mit antisymmetrischen Eigenfunktionen. Vgl. Kap. XXII.)

Das Aufbauprinzip werden wir mit Hilfe der sehr eleganten Slatterschen Methode ableiten.

Das Vektoradditionsmodell

6. Es sei hier noch ein einfacher, weitgehend schematisierter Fall des Aufbauprinzips behandelt, bei dem auf die Gleichheit der Elektronen keine Rücksicht genommen und die Drehgruppe als einzige vorhandene Symmetrie behandelt wird¹⁾.

Wir betrachten zwei Systeme, im einfachsten Falle besteht jedes System aus einem einzigen Elektron, die beide denselben Kern umkreisen. Das erste System habe die Energie E und befindet sich in einem Zustand mit der Azimutalquantenzahl l , die zugehörigen $2l+1$ Eigenfunktionen seien ψ_{-l}, \dots, ψ_l . Es gilt dann

$$\mathbf{P}_R \psi_\mu = \sum_{\mu'} \mathbf{D}^{(0)}(R)_{\mu' \mu} \psi_{\mu'}, \quad (8)$$

¹⁾ Vgl. E. Fues, Zeitschr. f. Phys. **51**, 817, 1928.

wo \mathbf{P}_R eine Drehung der Koordinaten des ersten Systems ist. Die Energie des zweiten Systems sei \bar{E} , die Azimutalquantenzahl \bar{l} , die Eigenfunktionen $\bar{\psi}_{-\bar{l}}, \dots, \bar{\psi}_{\bar{l}}$, und es gilt

$$\bar{\mathbf{P}}_R \bar{\psi}_v = \sum_{v'} \mathfrak{D}^{(\bar{l})}(R)_{vv'}, \bar{\psi}_{v'}. \quad (8a)$$

Die beiden Operatoren $\bar{\mathbf{P}}_R$ und \mathbf{P}_R sind dabei natürlich verschieden, weil \mathbf{P}_R die Variablen der ψ_u , dagegen $\bar{\mathbf{P}}_R$ die der $\bar{\psi}_v$ einer Drehung unterwirft und die beiden Variablensysteme ja gänzlich verschieden sein sollen. Aus diesem Grunde sind auch alle \mathbf{P}_R mit allen $\bar{\mathbf{P}}_R$ vertauschbar, es ist auch $\mathbf{P}_R \bar{\psi}_v = \bar{\psi}_v$ und $\bar{\mathbf{P}}_R \psi_u = \psi_u$, weil eben \mathbf{P}_R auf die Variable der $\bar{\psi}$ und $\bar{\mathbf{P}}_R$ auf die Variable der ψ gar nicht einwirkt.

Betrachtet man die beiden Systeme als ein einziges, so sind die Eigenwerte nach (6a) die Summen, die Eigenfunktionen nach (6) die Produkte der Eigenwerte bzw. der Eigenfunktionen der Einzelsysteme. Zum Eigenwert $E + \bar{E}$ gehören die $(2l+1)(2\bar{l}+1)$ Eigenfunktionen

$$\left. \begin{aligned} & \psi_{-l} \bar{\psi}_{-\bar{l}}, \psi_{-l} \bar{\psi}_{-\bar{l}+1}, \dots, \psi_{-l} \bar{\psi}_{\bar{l}-1}, \psi_{-l} \bar{\psi}_{\bar{l}}, \\ & \dots \end{aligned} \right\} \quad (9)$$

$$\begin{aligned} & \psi_l \bar{\psi}_{-\bar{l}}, \psi_l \bar{\psi}_{-\bar{l}+1}, \dots, \psi_l \bar{\psi}_{\bar{l}-1}, \psi_l \bar{\psi}_{\bar{l}}. \end{aligned}$$

Es fragt sich nun, aus welchen Operatoren die Gruppe des Gesamtsystems bei Einführung einer Wechselwirkung der beiden Systeme noch bestehen wird? Offenbar nicht aus dem ganzen direkten Produkt der beiden Operatorengruppen \mathbf{P}_R und $\bar{\mathbf{P}}_R$, die Elemente $\mathbf{P}_R \bar{\mathbf{P}}_R$ dieses würden ja gleichzeitigen, aber verschiedenen Drehungen der Achsenkreuze der Variablen von ψ und $\bar{\psi}$ entsprechen. Die Gruppe, die wir betrachten müssen, ist vielmehr die, bei der beide Achsenkreuze derselben Drehung unterworfen werden, sie besteht nicht aus allen Operatoren $\mathbf{P}_R \bar{\mathbf{P}}_R$, sondern nur aus den $\mathbf{P}_R \bar{\mathbf{P}}_R$. Die Gruppe der $\mathbf{P}_R \bar{\mathbf{P}}_R$ ist der einfachen Drehgruppe holomorph: aus $RQ = T$ folgt

$$\mathbf{P}_R \bar{\mathbf{P}}_R \cdot \mathbf{P}_Q \bar{\mathbf{P}}_Q = \mathbf{P}_R \mathbf{P}_Q \bar{\mathbf{P}}_R \bar{\mathbf{P}}_Q = \mathbf{P}_T \bar{\mathbf{P}}_T.$$

Wenden wir die Operatoren $\mathbf{P}_R \bar{\mathbf{P}}_R$ auf die Funktionen (9) an, so kann man die entstandenen Funktionen mit Hilfe der unveränderten ausdrücken.

Nach (8), (8a) ist

$$\begin{aligned} & \mathbf{P}_R \bar{\mathbf{P}}_R \psi_\mu \psi_\nu = \mathbf{P}_R \psi_\mu \cdot \bar{\mathbf{P}}_R \bar{\psi}_\nu \\ &= \sum_{\mu'} \mathfrak{D}^{(l)}(R)_{\mu' \mu} \psi_{\mu'} \sum_{\nu'} \mathfrak{D}^{(\bar{l})}(R)_{\nu' \nu} \bar{\psi}_{\nu'} = \sum_{\mu' \nu'} \Delta(R)_{\mu' \nu'; \mu \nu} \psi_{\mu'} \bar{\psi}_{\nu'}. \quad (10) \end{aligned}$$

Die zu den $(2l+1)(2\bar{l}+1)$ Funktionen (9) des Gesamtsystems gehörige Darstellung $\Delta(R)$ ist das direkte Produkt¹⁾ der beiden Darstellungen $\mathfrak{D}^{(l)}$ und $\mathfrak{D}^{(\bar{l})}$ der Einzelsysteme:

$$\Delta(R)_{\mu' \nu'; \mu \nu} = \mathfrak{D}^{(l)}(R)_{\mu' \mu} \mathfrak{D}^{(\bar{l})}(R)_{\nu' \nu}; \quad \Delta(R) = \mathfrak{D}^{(l)}(R) \times \mathfrak{D}^{(\bar{l})}(R). \quad (11)$$

Wir müssen jetzt die irreduziblen Bestandteile von $\Delta(R)$ bestimmen. Das geschieht am einfachsten durch Zerlegung seines Charakters in Charaktere irreduzibler Darstellungen. Der Charakter von $\Delta(R)$, wo R einer Drehung mit dem Drehwinkel φ entspricht, ist gleich

$$\begin{aligned} \sum_{\mu, \nu} \Delta(R)_{\mu \nu; \mu \nu} &= \sum_{\mu} \mathfrak{D}^{(l)}(R)_{\mu \mu} \sum_{\nu} \mathfrak{D}^{(\bar{l})}(R)_{\nu \nu}, \\ &= \chi^{(l)}(\varphi) \chi^{(\bar{l})}(\varphi) = \sum_{\mu=-l}^l e^{i \mu \varphi} \sum_{\nu=-\bar{l}}^{\bar{l}} e^{i \nu \varphi}. \quad (12) \end{aligned}$$

Um diesen Ausdruck umzuformen, pflegt man für jede Exponentialfunktion $e^{ix\varphi}$ (für $x = \dots, -2, -1, 0, 1, 2, \dots$) eine Spalte zu bilden und in diese Spalte so viel Kreuzchen einzutragen, wie vielmehr $e^{ix\varphi}$ in (12) vorkommt. Das kleinste vorkommende x ist $-l - \bar{l}$, das größte $l + \bar{l}$, man braucht also im ganzen $2l + 2\bar{l} + 1$ Spalten. Die Kreuze, die von den $2l + 1$ Gliedern $e^{i(\nu-l)\varphi}, e^{i(\nu-l+1)\varphi}, \dots, e^{i(\nu+\bar{l})\varphi}$ mit demselben ν herühren, schreiben wir in dieselbe Zeile. Wenn wir noch $l > \bar{l}$ annehmen, erhalten wir die

Tabelle 1

$x =$	$-l - \bar{l}$	$-l + \bar{l}$	$l - \bar{l}$	$l + \bar{l}$
$\nu = -\bar{l} \dots$	+	+	+	+
$\nu = 0 \dots$		+	+	+
$\nu = \bar{l} \dots$			+	+

¹⁾ Wir haben es hier mit einer anderen Art des direkten Produkts zu tun als im vorangehenden Kapitel. Dort hatten wir zwei Symmetrien (Drehung R und Spiegelung I) vereinigt, die Gruppe also vergrößert. Hier vereinigen wir zwei Systeme, die die gleiche Symmetrie haben, das Gesamtsystem hat dann auch dieselbe Symmetrie.

Man kann hierbei jedes Kreuzchen in seiner Spalte beliebig verschieben. Klappt man den Teil der Tabelle, der links von der punktierten Linie liegt, um die durch \rightarrow angezeichnete Zeile, die von den Gliedern mit $\nu = 0$ herröhrt, so erhält man ein Schema der Gestalt:

Tabelle 2

$\nu =$	$-l - \bar{l}$	$-l + \bar{l}$	$l - \bar{l}$	$l + \bar{l}$
$\nu = -\bar{l} \dots$		+ + +		
$\nu = 0 \dots$	$\rightarrow + + +$	+ + + + +		
$\nu = \bar{l} \dots$	+ + + +	+ + + + + +	+ + + + + +	+ + + + +

Das erste Kreuzchen in der ν -Zeile ist jetzt in der Spalte $-\nu - l$, die Glieder, die der ν -Zeile der Tabelle 2 entsprechen, sind

$$e^{-i(\nu+l)\varphi} + e^{-i(\nu+l-1)\varphi} + \dots + e^{i(\nu+l-1)\varphi} + e^{i(\nu+l)\varphi} = \chi^{(l+\nu)}(\varphi), \quad (13)$$

sie geben zusammen eben den Charakter einer irreduziblen Darstellung mit $L = l + \nu$. In der ganzen Tabelle sind die irreduziblen Darstellungen mit

$$L = l - \bar{l}, l - \bar{l} + 1, \dots, l + \bar{l} - 1, l + \bar{l} \quad (*)$$

enthalten. Für $\bar{l} \leqq l$ entstehen aus dem Term $E + \bar{E}$ bei Einführung einer Wechselwirkung $2\bar{l} + 1$ Terme mit den Azimutalquantenzahlen (*). Die irreduziblen Bestandteile von $\mathfrak{D}^{(l)}(R) \times \mathfrak{D}^{(\bar{l})}(R)$ sind in diesem Falle: je ein $\mathfrak{D}^{(L)}(R)$ mit den L -Werten (*). Ist $l \leqq \bar{l}$, so vertauscht sich die Rolle von l und \bar{l} , so daß die L -Werte allgemein

$$L = |l - \bar{l}|, |l - \bar{l}| + 1, \dots, l + \bar{l} - 1, l + \bar{l} \quad (14)$$

sind.

Dieses „Vektoradditionsmodell“ (siehe Abb. 8) ist von sehr allgemeiner Gültigkeit und von grundlegender Bedeutung für die gesamte Spektroskopie. Die beiden Systeme, um deren Vereinigung es sich handelt, müssen nicht je aus einem einzigen Elektron bestehen — in diesem Falle gibt das Prinzip in der vorliegenden Form wegen der Nichtbeachtung der Gleichheit der Teilchen gar nicht alle Einzelheiten wieder —, sie können schon selber zusammengesetzte Systeme sein; ja das Prinzip gilt sogar — wie wir sehen

werden — für die Wechselwirkung von Spinquantenzahl und azimuthaler Quantenzahl (wobei die entstehenden L „Gesamtquantenzahl“ genannt werden), oder auch für die Wechselwirkung zwischen Gesamtquantenzahl und Kernspin usw.

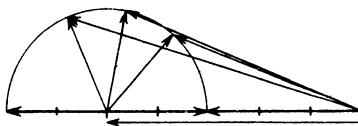


Abb. 8. Die Zusammensetzung von $l = 5$ und $\bar{l} = 2$ ergibt als mögliche L -Werte 3, 4, 5, 6, 7

7. Wir wissen jetzt, daß die Darstellung $\mathfrak{D}^{(l)} \times \mathfrak{D}^{(\bar{l})}$ mit der Darstellung

$$\begin{pmatrix} \mathfrak{D}(|l-\bar{l}|) & 0 & \dots & 0 & 0 \\ 0 & \mathfrak{D}(l-\bar{l}+1) & \dots & 0 & 0 \\ \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & \dots & \mathfrak{D}^{(l+\bar{l}-1)} & 0 \\ 0 & 0 & \dots & 0 & \mathfrak{D}^{(l+\bar{l})} \end{pmatrix} = \mathbf{M}(R), \quad (15)$$

die wir der Kürze halber mit $\mathbf{M}(R)$ bezeichnen, äquivalent ist. Es muß daher eine Matrix S existieren, die sie ineinander transformiert:

$$\mathfrak{D}^{(l)}(R) \times \mathfrak{D}^{(\bar{l})}(R) = S^{-1} \mathbf{M}(R) S. \quad (16)$$

Da $\mathbf{M}(R)$ und auch $\mathfrak{D}^{(l)} \times \mathfrak{D}^{(\bar{l})}$ unitär sind, kann man sogar nach Satz 1 a, Kap. IX, annehmen, daß S unitär, $S^{-1} = S^\dagger$ ist.

Die Matrix S ist eine im weiteren Sinne quadratische Matrix, wie wir sie im zweiten Kapitel besprochen haben. Die Zeilen-Spaltenbenennung von $\mathfrak{D}^{(l)} \times \mathfrak{D}^{(\bar{l})}$ geschieht nämlich durch zwei Indizes μ und ν , und dies muß auch von der Spaltenbenennung von S gelten. Die Zeilen und Spalten von $\mathbf{M}(R)$ tragen auch zwei Indizes, doch sind diese anderer Art: der erste Index L gibt an, welche Darstellung $\mathfrak{D}^{(L)}$ in dieser Zeile steht, der zweite m , um welche Zeile dieser Darstellung es sich handelt. Die Koeffizienten von $\mathbf{M}(R)$ lauten

$$\mathbf{M}(R)_{L'm'; Lm} = \delta_{L'L} \mathfrak{D}^{(L)}(R)_{m'm}. \quad (17)$$

Die Zeilen von \mathbf{S} tragen dementsprechend auch die Indizes L, m , wobei L von $|l - \bar{l}|$ bis $l + \bar{l}$ und m von $-L$ bis L läuft: (16) lautet in den Koeffizienten ausgeschrieben

$$\mathfrak{D}^{(l)}(R)_{\mu' \mu} \mathfrak{D}^{(\bar{l})}(R)_{\nu' \nu} = \sum_{m' m} \sum_L S_{L m'; \mu' \nu'}^* \mathfrak{D}^{(L)}(R)_{m' m} S_{L m; \mu \nu}. \quad (16a)$$

Die Bedeutung der Matrix \mathbf{S} liegt darin, daß man mit ihrer Hilfe aus den $\psi_\mu \bar{\psi}_\nu$ solche Linearkombinationen

$$\Psi_m^L = \sum_{\mu \nu} S_{L m; \mu \nu}^* \psi_\mu \bar{\psi}_\nu. \quad (18)$$

bilden kann, die sich bei der Anwendung der Operatoren $\mathbf{P}_R \bar{\mathbf{P}}_R$, denen gegenüber das System auch nach Einführung der Wechselwirkung invariant ist, nach irreduziblen Darstellungen transformieren:

$$\begin{aligned} \mathbf{P}_R \bar{\mathbf{P}}_R \Psi_m^L &= \sum_{\mu \nu} S_{L m; \mu \nu}^* \mathbf{P}_R \psi_\mu \cdot \bar{\mathbf{P}}_R \bar{\psi}_\nu, \\ &= \sum_{\mu \nu} \sum_{\mu' \nu'} S_{L m; \mu \nu}^* \mathfrak{D}^{(l)}(R)_{\mu' \mu} \mathfrak{D}^{(\bar{l})}(R)_{\nu' \nu} \psi_{\mu'} \psi_{\nu'} \\ &= \sum_{\mu \mu'} \sum_{\nu \nu'} \sum_{L' m'} S_{L m; \mu \nu}^* \mathfrak{D}^{(l)}(R)_{\mu' \mu} \mathfrak{D}^{(\bar{l})}(R)_{\nu' \nu} S_{L' m'; \mu' \nu'} \Psi_{m'}^{L'} \\ &= \sum_{L' m'} [\mathbf{S} \cdot \mathfrak{D}^{(l)}(R) \times \mathfrak{D}^{(\bar{l})}(R) \cdot \mathbf{S}^{-1}]_{L' m'; L m} \Psi_{m'}^{L'}, \\ &= \sum_{L' m'} \mathbf{M}(R)_{L' m'; L m} \Psi_{m'}^{L'} = \sum_{m'} \mathfrak{D}^{(L)}(R)_{m' m} \Psi_m^L. \end{aligned} \quad \left. \right\} \quad (19)$$

Sie bilden daher die Eigenfunktionen erster Näherung (die „richtigen“ Linearkombinationen des Kap. V) des gestörten Gesamtsystems.

Um die Koeffizienten $S_{L m; \mu \nu}^*$ zu bestimmen, wenden wir zuerst auf (18) einen Operator $\mathbf{P}_R \bar{\mathbf{P}}_R$ an, wo R eine Drehung mit α um Z ist. Die linke Seite multipliziert sich dadurch mit $e^{im\alpha}$, und dies muß auch für die rechte Seite gelten:

$$\begin{aligned} \sum_{\mu \nu} S_{L m; \mu \nu}^* e^{im\alpha} \psi_\mu \bar{\psi}_\nu &= \sum_{\mu \nu} S_{L m; \mu \nu}^* \mathbf{P}_R \psi_\mu \bar{\mathbf{P}}_R \bar{\psi}_\nu \\ &= \sum_{\mu \nu} S_{L m; \mu \nu}^* e^{i\mu\alpha} \psi_\mu e^{i\nu\alpha} \bar{\psi}_\nu. \end{aligned} \quad (20)$$

Es ist daher wegen der linearen Unabhängigkeit der $\psi_\mu \bar{\psi}_\nu$,

$$S_{L m; \mu \nu} = 0 \quad \text{für } m \neq \mu + \nu. \quad (20a)$$

Dasselbe erhält man aus (16a), wenn man die Abhängigkeit der Darstellungskoeffizienten von α und γ darin nach (8), Kap. XV,

einträgt und die Glieder mit gleicher Abhängigkeit von α und γ einander gleichsetzt. Wir setzen noch

$$S_{L, \mu + \nu; \mu \nu} = s_{L \mu \nu}, \quad (20\text{b})$$

dann ist (16a)

$$\mathfrak{D}^{(l)}(R)_{\mu' \mu} \mathfrak{D}^{(\bar{l})}(R)_{\nu' \nu} = \sum_{L=|l-\bar{l}|}^{\bar{l}+l} s_{L \mu' \nu'}^* \mathfrak{D}^{(L)}(R)_{\mu' + \nu'; \mu + \nu} s_{L \mu \nu}. \quad (16\text{b})$$

Die Matrix S ist durch (16) noch nicht eindeutig festgelegt. Da $M(R)$ mit einer Diagonalmatrix

$$u = \begin{pmatrix} \omega_{|l-\bar{l}|} 1 & 0 & \dots & 0 & 0 \\ 0 & \omega_{|l-\bar{l}|+1} 1 & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & \omega_{l+\bar{l}-1} 1 & 0 \\ 0 & 0 & \dots & 0 & \omega_{l+\bar{l}} 1 \end{pmatrix};$$

$$u_{L' m'; L m} = \omega_L \delta_{L' L} \delta_{m' m}$$

vertauschbar ist, ändert sich die rechte Seite von (16) nicht, wenn man S durch uS ersetzt. Damit uS unitär bleibe, muß u unitär sein, was dann der Fall ist, wenn die Absolutwerte der ω alle 1 sind. Die Koeffizienten von uS , womit wir S ersetzen wollen, sind

$$(uS)_{L m; \mu \nu} = \omega_L S_{L m; \mu \nu}.$$

Durch zweckmäßige Wahl der ω kann man jedenfalls erreichen, daß

$$S_{L, l-\bar{l}; l, -\bar{l}} = s_{L, l, -\bar{l}} = |s_{L, l, -\bar{l}}| \quad (21)$$

reell positiv sein soll. Dies wollen wir im folgenden annehmen. Wir multiplizieren nun (16b) mit $\mathfrak{D}^{(L')}(R)_{\mu' + \nu'; \mu + \nu}^*$ und integrieren über die ganze Drehgruppe. Es bleibt dann rechts wegen der Orthogonalitätsrelationen der Darstellungskoeffizienten nur ein Glied stehen, und wir erhalten, wenn wir für L' wieder L einsetzen (und für $\int dR = g$ schreiben),

$$\int \mathfrak{D}^{(l)}(R)_{\mu' \mu} \mathfrak{D}^{(\bar{l})}(R)_{\nu' \nu} \mathfrak{D}^{(L)}(R)_{\mu' + \nu'; \mu + \nu}^* dR = g \frac{s_{L \mu' \nu'}^* s_{L \mu \nu}}{2L + 1}. \quad (22)$$

Zur Bestimmung der $s_{L \mu \nu}$ ist es nicht notwendig, das Integral in (22) für alle möglichen Werte der L , μ' , ν' , μ , ν auszuwerten, es genügt, wenn es für ein einziges μ', ν' -Wertepaar und alle L , μ , ν (und l , \bar{l}) bekannt ist. Um die Formeln möglichst einfach zu

machen, setzen wir $\mu' = l$, $\nu' = -\bar{l}$ und erhalten nach (27) bzw. (27a) und (27b), Kap. XV,

$$\sqrt{\binom{2l}{l-\mu}\binom{2\bar{l}}{\bar{l}-\nu}} \sum_{\kappa} (-1)^{\kappa+\bar{l}+\nu} \frac{\sqrt{(L+\mu+\nu)!(L-\mu-\nu)!(L+l-\bar{l})!(L-l+\bar{l})!}}{(L-l+\bar{l}-\kappa)!(L+\mu+\nu-\kappa)!\kappa!(\kappa+l-\bar{l}-\mu-\nu)!} \cdot \int \cos^{2l+2\bar{l}+2\mu-2\kappa} \beta \cdot \sin^{2l-2\mu+2\kappa} \beta dR = g \frac{s_{L,\bar{l},-\bar{l}}^* s_{L,\mu,\nu}}{2L+1}. \quad (23)$$

(Wie zu erwarten war, fielen dabei α und γ heraus.) Was wir noch brauchen, sind Integrale der Gestalt

$$\int \cos^{2a} \beta \sin^{2b} \beta dR.$$

Auch diese liefert uns die Orthogonalitätsrelation der Darstellungs-koeffizienten. Es ist nämlich

$$\frac{g}{2j+1} = \int |\mathfrak{D}^{(j)}(R)_{j\mu}|^2 dR = \binom{2j}{j-\mu} \int \cos^{2j+2\mu} \beta \sin^{2j-2\mu} \beta dR,$$

oder wenn wir für $j + \mu = a$, $j - \mu = b$ setzen,

$$\int \cos^{2a} \beta \sin^{2b} \beta dR = g \frac{b! a!}{(a+b+1)!}. \quad (24)$$

Dies in (23) eingesetzt, ergibt zunächst

$$\sum_{\kappa} (-1)^{\kappa+\bar{l}+\nu} \frac{\sqrt{(2l)!(2\bar{l})!(L+\mu+\nu)!(L-\mu-\nu)!(L+l-\bar{l})!(L-l+\bar{l})!}}{(L+l+\bar{l}+1)!\sqrt{(l-\mu)!(l+\mu)!(l-\nu)!(l+\nu)!}} \cdot \frac{(L+\bar{l}+\mu-\kappa)!(l-\mu+\kappa)!(2L+1)}{(L-l+\bar{l}-\kappa)!(L+\mu+\nu-\kappa)!\kappa!(\kappa+l-\bar{l}-\mu-\nu)!} = s_{L,\bar{l},-\bar{l}}^* s_{L,\mu,\nu}. \quad (25)$$

Um $s_{L,\bar{l},-\bar{l}}$ zu bestimmen, setzen wir hierin noch $\mu = l$, $\nu = -\bar{l}$

$$\frac{2L+1}{(L+l+\bar{l}+1)!} \sum_{\kappa} \frac{(-1)^{\kappa}(L+l-\bar{l})!(L-l+\bar{l})!(L+\bar{l}+l-\kappa)!}{(L-l+\bar{l}-\kappa)!(L+l-\bar{l}-\kappa)!\kappa!} = |s_{L,\bar{l},-\bar{l}}|^2 = (s_{L,\bar{l},-\bar{l}})^2, \quad (25a)$$

das letzte wegen (21). Da weiter, wie im Anhang gezeigt werden soll,

$$\sum_{\kappa} (-1)^{\kappa} \binom{L-l+\bar{l}}{\kappa} \frac{(L+\bar{l}+l-\kappa)!}{(L+l-\bar{l}-\kappa)!} = (2\bar{l})! \binom{2\bar{l}}{L+l-\bar{l}} \quad (26)$$

ist, ergibt sich schließlich

$$s_{L,\bar{l},-\bar{l}} = \sqrt{\frac{(2L+1)(2\bar{l})!(2\bar{l})!}{(L+l+\bar{l}+1)!(l+\bar{l}-L)!}}, \quad (27a)$$

und mit Hilfe von (25)

$$s_{L\mu}^{(l\bar{l})} = \frac{\sqrt{(L+l-\bar{l})! (L-l+\bar{l})! (l+\bar{l}-L)! (L+\mu+\nu)! (L-\mu-\nu)!}}{\sqrt{(L+l+\bar{l}+1)! (l-\mu)! (l+\mu)! (\bar{l}-\nu)! (\bar{l}+\nu)!}} \sum_x \frac{(-1)^x + \bar{l} + \nu}{(L-l+\bar{l}-x)! (L+\mu+\nu-x)! x! (x+l-\bar{l}-\mu-\nu)!}. \quad (27)$$

Die Summation über x ist hierin ebenso wie in (27), Kap. XV, über alle ganzen Zahlen zu erstrecken, wegen des Unendlichwerdens der Fakultäten im Nenner genügt es, sie — ebenso wie dort — von der größeren der beiden Zahlen 0, $\bar{l} - l + \mu + \nu$ bis zur kleineren von $L + \mu + \nu$ und $L - l + \bar{l}$ zu erstrecken. Die Größen s hängen außer von ihren Indizes L , μ , ν noch von den beiden Zahlen l und \bar{l} ab, die angeben, welches direkte Produkt $\mathfrak{D}^{(l)} \times \mathfrak{D}^{(\bar{l})}$ man durch sie ausreduzieren kann. Außerdem sollte s im wesentlichen ungeändert bleiben¹⁾, wenn man l mit \bar{l} und gleichzeitig μ mit ν vertauscht, aus (27) ist dies aber nicht ohne weiteres zu erkennen, weil die Summation über x nicht in geschlossener Form ausgeführt werden kann. Im Falle $\mu + \nu = L$ bleibt jedoch von der ganzen Summe nur ein Glied ($x = L - l + \bar{l}$) stehen, und man erhält

$$s_{L\mu L-\mu}^{(l\bar{l})} = (-1)^{l-\mu} \sqrt{\frac{(2L+1)! (l+\bar{l}-L)! (l+\mu)! (L+\bar{l}-\mu)!}{(L+l+\bar{l}+1)! (L+l-\bar{l})! (L-l+\bar{l})! (l-\mu)! (\bar{l}-L+\mu)!}}. \quad (27b)$$

Es seien noch die Gleichungen, die aus der Unitarität von S für die s folgen, explizite hingeschrieben [(27) zeigt, daß S reell ist]:

$$\sum_\mu s_{L\mu m-\mu}^{(l\bar{l})} s_{L'\mu' m-\mu}^{(l\bar{l})} = \delta_{LL'}; \quad \sum_L s_{L\mu m-\mu}^{(l\bar{l})} s_{L\mu' m-\mu'}^{(l\bar{l})} = \delta_{\mu\mu'}. \quad (28)$$

8. Hiermit haben wir alle in (16 b) und in (18)

$$\Psi_m^L = \sum_\mu s_{L\mu m-\mu}^{(l\bar{l})} \psi_\mu \bar{\psi}_{m-\mu} \quad (18a)$$

auftrtenden Koeffizienten bestimmt. Bezuglich (18 a) ist zu bemerken, daß wir hier einen Fall — und zwar einen der wich-

¹⁾ In Wirklichkeit ist $s_{L\mu\nu}^{(l\bar{l})} = (-1)^{l+\bar{l}-L} s_{L\nu\mu}^{(\bar{l}l)}$: in (21) gehen l und \bar{l} nicht in gleicher Weise ein.

tigsten — vor uns haben, in dem die „richtigen Linearkombinationen“ erster Näherung des Störungsverfahrens aus allgemeinen Überlegungen bestimmt werden können: (18a) gilt für alle Störungen, die keine Raumrichtung auszeichnen, ganz allgemein. Dies beruht darauf, daß wir von vornherein wissen, daß die richtigen Linearkombinationen alle „zu einer Zeile einer irreduziblen Darstellung gehören“ und daß man aus den Funktionen (9) nur eine einzige Linearkombination bilden kann, die zur m -Zeile von $\mathfrak{D}^{(L)}$ gehört — wenn man überhaupt eine bilden kann (wenn L zwischen $|l - \bar{l}|$ und $l + \bar{l}$ liegt). Wenn allerdings im ungestörten Problem außer (9) noch weitere Eigenfunktionen zu demselben Eigenwert gehören, so ist es möglich, daß mehrere Linearkombinationen der verlangten Eigenschaft existieren, und die „richtige“ kann noch eine Linearkombination dieser — aber auch nur dieser — sein.

Die Formel (16b) ist vieler Anwendungen fähig. Zunächst gilt sie nicht nur für eindeutige Darstellungen (für ganzzahlige l), sondern auch für die zweideutigen Darstellungen des Kap. XV. Sie enthält unter anderem die Intensitätsformeln der Multiplettlinien und der Zeemankomponenten (Kap. XXIII).

Daß man $\mathfrak{D}^{(l)}(R)_{\mu',\mu} \mathfrak{D}^{(\bar{l})}(R)_{\nu'}$, durch die Darstellungskoeffizienten ausdrücken kann, ist selbstverständlich: diese bilden ja ein vollständiges Funktionensystem. Auch daß nur die Koeffizienten, die in irgendeiner Darstellung in der $\mu' + \nu'$ -Zeile und in der $\mu + \nu$ -Spalte stehen, in (16b) vorkommen können, ist klar: nur diese haben die richtige Abhängigkeit von α und γ . Außerdem zeigt (16b) noch, daß auch L nur zwischen $|l - \bar{l}|$ und $l + \bar{l}$ variiert werden muß. Sind l und \bar{l} beide ganzzahlig oder beide halbzahlig, so sind die L in (16b) alle ganzzahlig, ist dagegen nur eines ganzzahlig, das andere halbzahlig, so sind die L alle halbzahlig: die Summation ist von der unteren bis zur oberen Grenze immer in ganzzahligen Schritten zu erstrecken.

Für $\bar{l} = 0$ ist (16b) trivial, für $\bar{l} = 1$ seien noch die $s_{L,\mu}^{(l,1)}$, in einer Tabelle zusammengestellt¹⁾.

1) Man kann sich die $s_{L,\mu,\nu}^{(l,1)}$ leicht merken, wenn man sich vor Augen hält, daß sie verschwinden, wenn $|\mu| > l$ und wenn $|\mu + \nu| > L$ ist (nicht wenn beides zutrifft); also immer dann, wenn einer der Darstellungskoeffizienten im Integral (22) sinnlos wird.

Tabelle für die $s_{L\mu\nu}^{(l_1)}$

L	$\nu = -1$	0	+1
$l-1$	$\frac{\sqrt{l+\mu}\sqrt{l+\mu-1}}{\sqrt{2l}\sqrt{2l+1}}$	$-\frac{\sqrt{l-\mu}\sqrt{l+\mu}}{\sqrt{l}\sqrt{2l+1}}$	$\frac{\sqrt{l-\mu-1}\sqrt{l-\mu}}{\sqrt{2l}\sqrt{2l+1}}$
l	$\frac{\sqrt{l-\mu+1}\sqrt{l+\mu}}{\sqrt{2l}\sqrt{l+1}}$	$\frac{\mu}{\sqrt{l}\sqrt{l+1}}$	$\frac{\sqrt{l+\mu+1}\sqrt{l-\mu}}{\sqrt{2l}\sqrt{l+1}}$
$l+1$	$\frac{\sqrt{l-\mu+1}\sqrt{l-\mu+2}}{\sqrt{2l+1}\sqrt{2l+2}}$	$\frac{\sqrt{l-\mu+1}\sqrt{l+\mu+1}}{\sqrt{2l+1}\sqrt{l+1}}$	$\frac{\sqrt{l+\mu+1}\sqrt{l+\mu+2}}{\sqrt{2l+1}\sqrt{2l+2}}$

Anhang. Um noch (26) zu beweisen, gehen wir von der Identität

$$\sum_x \binom{a}{x} \binom{b}{c-x} = \binom{a+b}{c}$$

aus. Links steht der Koeffizient von x^u in $(1+x)^a$, multipliziert mit dem Koeffizienten von x^{c-u} in $(1+x)^b$ und summiert über alle x , d. h. der Koeffizient von x^c in $(1+x)^a \cdot (1+x)^b = (1+x)^{a+b}$, und dies steht auch rechts; a sei eine positive ganze Zahl, b kann auch negativ sein. Weiter ist ($u < 0$)

$$\begin{aligned} \binom{u}{v} &= \frac{u(u-1)\dots(u-v+2)(u-v+1)}{1 \cdot 2 \dots (v-1) \cdot v} \\ &= (-1)^v \frac{(v-u-1)(v-u-2)\dots(1-u)(-u)}{1 \cdot 2 \dots (v-1) \cdot v} \\ &= (-1)^v \binom{v-u-1}{v}. \end{aligned}$$

Es ist daher

$$\begin{aligned} \sum_x (-1)^x \binom{L-l+\bar{l}}{x} \binom{L+\bar{l}+l-x}{L+l-\bar{l}-x} (2\bar{l})! \\ = \sum_x (-1)^{L+l-\bar{l}} (2\bar{l})! \binom{L-l+\bar{l}}{x} \binom{-2\bar{l}-1}{L+l-\bar{l}-x} \\ = (-1)^{L+l-\bar{l}} (2\bar{l})! \binom{L-l-\bar{l}-1}{L+l-\bar{l}} = (2\bar{l})! \binom{2l}{L+l-\bar{l}}, \end{aligned}$$

womit (26) bewiesen ist.

XVIII. Auswahlregeln und die Aufspaltung der Spektrallinien

1. Im VI. Kapitel haben wir mit Hilfe der Schrödingerschen zeitabhängigen Differentialgleichung die Zunahme der Anregungswahrscheinlichkeit $|\alpha_F(t)|^2 = |\psi(t)|^2$ des stationären Zustandes ψ_F unter dem Einfluß eines in der X-Richtung polarisierten Lichtstrahles der Intensität (Energiedichte pro Frequenzeinheit) J berechnet. Es ergab sich dafür [(17) und (6) Kapitel VI], wenn das Atom anfangs ganz im stationären Zustand ψ_E war

$$|\alpha_F(t)|^2 = B_{FE} J t = \frac{8 \pi^3 e^2}{h^3} |\mathbf{X}_{FE}|^2 J t, \quad (1)$$

wo \mathbf{X}_{FE} das sogenannte Matrixelement

$$\mathbf{X}_{FE} = (\psi_F, (x_1 + x_2 + \dots + x_n) \psi_E) \quad (2a)$$

der „X-Komponente des Dipolmoments“ des Überganges $E \rightarrow F$ ist. Ist das Licht in der Y- bzw. Z-Richtung polarisiert, so tritt in (1) an Stelle von \mathbf{X}_{FE}

$$\mathbf{Y}_{FE} = (\psi_F, (y_1 + y_2 + \dots + y_n) \psi_E) \quad (2b)$$

$$\mathbf{Z}_{FE} = (\psi_F, (z_1 + z_2 + \dots + z_n) \psi_E), \quad (2c)$$

ist es in der Richtung mit den Richtungscosinus α_1 , α_2 , α_3 polarisiert, so tritt entsprechend

$$\alpha_1 \mathbf{X}_{FE} + \alpha_2 \mathbf{Y}_{FE} + \alpha_3 \mathbf{Z}_{FE} \quad (2)$$

auf.

Nach der bekannten Einsteinschen Überlegung¹⁾ kann man hieraus die Wahrscheinlichkeit $A_{FE} dt$ dafür berechnen, daß ein Atom, das im angeregten Zustand ψ_F ist, im Laufe der sehr kurzen Zeit dt durch spontane Ausstrahlung in den Zustand ψ_E übergehe. Eigentlich nennt man diese Größe die „Übergangswahrscheinlichkeit“, es ist

$$A_{FE} = \frac{64 \pi^4 c^2 \nu^8}{3 h c^3} (|\mathbf{X}_{FE}|^2 + |\mathbf{Y}_{FE}|^2 + |\mathbf{Z}_{FE}|^2). \quad (1a)$$

Wenn eine Spektrallinie mit der Frequenz $(F - E)/h$ in einem Spektrum nicht auftritt, obwohl die Existenz von Atomen im Zustand ψ_F durch das Auftreten anderer Linien erwiesen ist, wird man

¹⁾ A. Einstein, Verh. d. Deutsch. Phys. Ges., S. 318, 1916; Phys. Zeitschr. 18, 121, 1916.

schließen dürfen, daß die Ausdrücke (2 a), (2 b), (2 c) verschwinden. In weitaus den meisten Fällen folgen diese „Auswahlregeln“ aus den Transformationseigenschaften der beteiligten Eigenfunktionen. Den Transformationseigenschaften der Eigenfunktionen gegenüber der symmetrischen Gruppe, der dreidimensionalen Drehgruppe und der Spiegelungsgruppe entsprechen drei Arten von Auswahlregeln.

Es ist aber zu bemerken, daß das Verschwinden von (2) nicht das absolute Ausfallen der Linie $F \rightarrow E$ zur Folge hat. Bei der Ableitung von (1) ist nämlich eine wesentliche, nicht streng richtige Voraussetzung gemacht worden: die Atomdimensionen wurden gegenüber der Lichtwellenlänge als klein angenommen, und es wurde so gerechnet, als ob das vom Licht herrührende Zusatzpotential in der Richtung des Lichtstrahls konstant wäre, weil es sich nur in Abständen, die in der Größenordnung der Wellenlänge sind, wesentlich ändert. Würde man berücksichtigen, daß dieses Potential in Wirklichkeit in Richtung des Strahles sinusförmig schwankt, so würde sich ein etwas anderer Ausdruck für die Übergangswahrscheinlichkeit (und daher auch für die Lebensdauer) ergeben, zum B_{EF} in (1) würde noch ein Korrektionsglied B' hinzukommen.

Von der Übergangswahrscheinlichkeit, die man nach (1) oder (1 a) berechnet, sagt man, daß sie von der Dipolstrahlung herrührt, B' ist durch Quadrupol und höhere Momente bedingt. Es ist $(\text{Atomdimension}/\text{Wellenlänge})^2$, also etwa 10^7 -mal kleiner als das durch Dipolstrahlung bedingte B_{EF} , und man kann es neben B_{EF} vernachlässigen, wenn (2) nicht verschwindet. Die Übergänge, für die (2) Null ist, sind aber nicht absolut verboten, sondern nur sehr viel schwächer als die gewöhnlichen, durch Dipolstrahlung bedingten Übergänge. Für die Intensität der Quadrupolstrahlung selber ist das Absolutwertquadrat von

$$\frac{2\pi\nu}{c} (\psi_E, (x_1 y_1 + x_2 y_2 + \dots + x_n y_n) \psi_F), \quad (3)$$

das an Stelle von X_{FE} in (1) eingesetzt werden muß, maßgebend¹⁾.

A. Terme verschiedener Multiplizität kombinieren nicht miteinander. Die Terme, die verschiedene Multiplizität $2S + 1$ haben, gehören ja zu verschiedenen Darstellungen der symmetrischen Gruppe, und das Multiplizieren mit $x_1 + x_2 + \dots + x_n$ ist ein gegenüber der Vertauschung von Elektronen symmetrischer Operator, so daß das skalare Produkt (2) nach den Resultaten des XII. Kapitels verschwinden muß. Selbst die Strahlung, die durch Quadrupol und höhere Momente bedingt werden könnte, verschwindet aus diesem Grunde.

¹⁾ Die Quadrupolstrahlung wurde in der Quantenmechanik hauptsächlich von A. Rubinowicz eingehend untersucht. Vgl. z. B. Zeitschr. f. Phys. 61, 338; 65, 662, 1930.

Empirisch ist bekannt, daß dieses sogenannte Interkombinationsverbot nur bei Elementen mit niedriger Ordnungszahl gut erfüllt ist. Bei schwereren Elementen kommen schon verhältnismäßig starke Linien vor, die Terme verschiedener Multiplizität verbinden. Dies röhrt von den Zusatzgliedern in der Schrödinger-Gleichung her, die dem magnetischen Moment des Elektrons Rechnung tragen und die sich bei wachsender Elektronenzahl immer stärker und stärker bemerkbar machen.

B. Gegenüber Drehungen ist das Multiplizieren mit $x_1 + x_2 + \dots + x_n$ kein symmetrischer Operator mehr, so daß die Auswahlregel für die Azimutalquantenzahl L anders als die für das S lauten wird. Die Azimutalquantenzahl von ψ_E sei L , dann gehört im Produkt $(x_1 + x_2 + \dots + x_n)\psi_E$ der zweite Faktor zur Darstellung $\mathfrak{D}^{(L)}$, der erste Faktor ist eine Vektorkomponente und gehört zu $\mathfrak{D}^{(1)}$.

Die $(2\bar{L} + 1)(2L + 1)$ Produkte von je zwei Funktionen, von denen jeweils die erste $f_{\bar{x}}^{(\bar{L})}$ zur \bar{x} -Zeile von $\mathfrak{D}^{(\bar{L})}$ und die zweite $\psi_x^{(L)}$ zur x -Zeile von $\mathfrak{D}^{(L)}$ gehört, transformieren sich nach $\mathfrak{D}^{(\bar{L})} \times \mathfrak{D}^{(L)}$ (vgl. auch die analogen Entwicklungen des vorangehenden Kapitels).

$$\mathbf{P}_R f_{\bar{x}}^{(\bar{L})} \psi_x^{(L)} = \mathbf{P}_R f_{\bar{x}}^{(\bar{L})} \cdot \mathbf{P}_R \psi_x^{(L)} = \sum_{\bar{\lambda}, \lambda} \mathfrak{D}^{(\bar{L})}(R)_{\bar{\lambda} \bar{x}} \mathfrak{D}^{(L)}(R)_{\lambda x} f_{\bar{\lambda}}^{(\bar{L})} \psi_{\lambda}^{(L)}.$$

Mit Hilfe der Matrix \mathbf{S} , die $\mathfrak{D}^{(\bar{L})} \times \mathfrak{D}^{(L)}$ ausreduziert, kann man aus $f_{\bar{x}}^{(\bar{L})} \psi_x^{(L)}$ solche Linearkombinationen $F_{\mu}^{(K)}$ bilden, die zu den irreduziblen Bestandteilen $\mathfrak{D}^{(K)}$ von $\mathfrak{D}^{(\bar{L})} \times \mathfrak{D}^{(L)}$ gehören. Umgekehrt kann man die Funktionen $f_{\bar{x}}^{(\bar{L})} \psi_x^{(L)}$ mit Hilfe der reziproken Matrix \mathbf{S}^{-1} durch die $F_{\mu}^{(K)}$ ausdrücken.

In unserem Falle ist $\bar{L} = 1$ und die irreduziblen Bestandteile von $\mathfrak{D}^{(1)} \times \mathfrak{D}^{(L)}$ sind für $L \neq 0$

$$\mathfrak{D}^{(L-1)}, \mathfrak{D}^{(L)}, \mathfrak{D}^{(L+1)}. \quad (*)$$

Man kann daher $(x_1 + x_2 + \dots + x_n)\psi_E$ als eine Summe von drei Funktionen schreiben, die zu je einer der Darstellungen (*) gehören. Ist nun die Azimutalquantenzahl L' von ψ_F weder $L - 1$ noch L noch $L + 1$, so verschwinden alle drei Teile des skalaren Produkts (2a). Die Azimutalquantenzahl L kann sich bei einem durch Dipolstrahlung bedingten spontanen Übergang nur um ± 1 oder 0 ändern.

Ist $L = 0$ so gehört $(x_1 + x_2 + \dots + x_n) \psi_E$ zur Darstellung $\mathfrak{D}^{(1)}$, da in diesem Falle $\mathfrak{D}^{(1)} \times \mathfrak{D}^{(0)}$ gleich $\mathfrak{D}^{(1)}$ ist. Ist $L' \neq 1$, so verschwindet (2): S-Terme ($L = 0$) kombinieren nur mit P-Termen ($L' = 1$), auch der Übergang $S \rightarrow S$ ist verboten.

Auch diese Regeln gelten nur für die leichten Elemente genau.

Die Ursache für ihr Versagen bei höheren Ordnungszahlen ist auch in den Störungen zu suchen, die durch das magnetische Moment des Elektrons bedingt sind. Die Linien, die trotz dieser Regel auftreten, sind nicht so augenfällig wie die Linien, die das Interkombinationsverbot verletzen, weil nämlich noch andere, auch bei Mitberücksichtigung dieser Störungen gültige Regeln existieren, die von sich aus das Ausfallen der meisten durch dieses Verbot erfaßten Übergänge bewirken.

Die Quadrupol- und höheren Momente verschwinden bei den besprochenen Übergangsverboten nicht. Wir müßten ja explizite die Tatsache benutzen, daß $(x_1 + x_2 + \dots + x_n)$ zur Darstellung $\mathfrak{D}^{(1)}$ gehört. Die entsprechenden, für die Quadrupolstrahlung maßgebenden Ausdrücke $(x_1 y_1 + \dots + x_n y_n)$ usw. gehören aber nicht zu $\mathfrak{D}^{(1)}$ sondern zu $\mathfrak{D}^{(2)}$, die Azimutalquantenzahl L kann sich dementsprechend bei einem durch Quadrupolmomente bedingten Übergang um ± 2 , ± 1 oder 0 ändern. Außerdem ist noch $S \rightarrow S$ und $S \rightarrow P$ durch Quadrupolstrahlung verboten.

C. Bei Dipolstrahlung ändert sich die Spiegelungssymmetrie immer, positive Terme kombinieren nur mit negativen, negative nur mit positiven. Bleibt nämlich ψ_E bei dem Ersetzen der x_k , y_k , z_k durch $-x_k$, $-y_k$, $-z_k$ ungeändert, so wechselt $(x_1 + x_2 + \dots + x_n) \psi_E$ sein Vorzeichen und umgekehrt bleibt bei dieser Operation der Ausdruck $(x_1 + \dots + x_n) \psi_E$ ungeändert, wenn ψ_E sein Vorzeichen umkehrt; er hat zu ψ_E entgegengesetzten Spiegelungscharakter. Wenn das skalare Produkt (2) nicht verschwinden soll, muß auch ψ_F zu ψ_E entgegengesetzten Spiegelungscharakter haben.

Empirisch ist diese Regel als Laporte-Russellsches Auswahlverbot bekannt. Gemäß ihrer Ableitung bezieht sie sich nur auf die Dipolstrahlung¹⁾, dagegen gilt sie auch unter Mitberücksichtigung der magnetischen Momente der Elektronen, also auch bei den schweren Elementen. Ihr widersprechende optische

¹⁾ Für die Übergänge, die durch Quadrupolstrahlung bedingt sind, gilt sogar die entgegengesetzte Regel: der Spiegelungscharakter ändert sich bei diesen Übergängen nicht.

Übergänge sind — trotz des so sehr reichen Materials — kaum bekannt. Die bekanntesten treten im sogenannten Nebuliumspektrum auf, wo sie von metastabilen Zuständen ausgehen, was die Möglichkeit — namentlich bei den Verhältnissen, die in den verdünnten Sternnebeln herrschen — einer extrem langen Abklingungszeit, also kleiner Übergangswahrscheinlichkeit, offenläßt.

Überblicken wir die drei Arten von Auswahlregeln noch einmal, so sehen wir, daß durch sie eigentlich die meisten Linien verboten sind: das Multiplettsystem darf sich nicht ändern, L darf sich nur um ± 1 oder 0 ändern (0 zu 0 ist auch verboten), der Spiegelungscharakter muß sich ändern. So kann z. B. ein $^3S_+$ -Term nur mit $^3P_-$ -Termen kombinieren¹⁾, ein $^4D_-$ -Term nur mit $^4P_+, ^4D_+$ und $^4F_+$ -Termen²⁾ usw. Es werde nochmals betont, daß die magnetischen Momente der Elektronen bisher nicht berücksichtigt und so die sogenannte Feinstruktur der Spektrallinien nicht erfaßt wurde. Die Regeln sollen für alle Feinstrukturkomponenten einer Linie gelten. Die ersten beiden Regeln gelten nur, wenn der Einfluß der erwähnten magnetischen Momente klein ist (bei kleiner Multiplettaufspaltung, d. h. bei den leichten Elementen), die letzte dagegen soll — was gegenwärtig noch nicht eingesehen werden kann — genau gelten.

2. Es sollen noch die Verhältnisse bei dem Einsetzen eines elektrischen oder magnetischen Feldes, also bei der Aufhebung der dreidimensionalen Drehsymmetrie besprochen werden.

Bekanntlich zeigt sich dabei ein Aufspalten der Linien in mehrere Komponenten. Dieses ist im Falle des magnetischen Feldes als Zeeman-Effekt sehr genau bekannt, während die analoge Erscheinung im elektrischen Felde, der Stark-Effekt, in den allermeisten Fällen der Beobachtung viel schlechter zugänglich ist. Von unserem vorläufigen Gesichtspunkt aus können die Verhältnisse natürlich nur sehr schlecht wiedergegeben werden und wir erhalten im wesentlichen nur Aufschluß darüber, wie der Zeeman- und Stark-Effekt wäre, wenn die Elektronen kein magnetisches Moment hätten.

Ein magnetisches Feld in der Z -Achse verringert die Symmetriegruppe des Konfigurationsraumes. Von den Drehungen bleiben nur diejenigen um die Z -Achse übrig. Außerdem bleiben noch — wegen des axialen Charakters des magnetischen Feldvektors —

¹⁾ Durch Quadrupolstrahlung noch mit $^3D_+$ -Termen.

²⁾ Durch Quadrupolstrahlung noch mit $^4S_-, ^4P_-, ^4D_-, ^4F_-, ^4G_-$ -Termen.

die beiden Richtungen Z und $-Z$ gleichberechtigt, so daß die XY -Ebene Symmetrieebene bleibt. Aus demselben Grunde aber ist etwa die YZ -Ebene keine Symmetrieebene mehr, da ein Drehungssinn ausgezeichnet ist. Man sieht dies am klarsten, wenn man die klassische Bahn eines Elektrons im Felde eines Kerns und eines Magnetfeldes betrachtet: durch Spiegelung der Bahn an der zum Felde senkrechten Ebene durch den Kern erhält man eine klassisch mögliche Bahn, nicht dagegen durch Spiegelung an einer zur Feldrichtung parallelen Ebene YZ .

Es folgt hieraus, daß die Inversionssymmetrie des Problems durch das Magnetfeld nicht gestört wird: die Inversion ($x'_k = -x_k$, $y'_k = -y_k$, $z'_k = -z_k$) ist ja gleich dem Produkt einer Drehung mit π um Z ($x'_k = -x_k$, $y'_k = -y_k$, $z'_k = z_k$) und einer Spiegelung an der XY -Ebene ($x'_k = x_k$, $y'_k = y_k$, $z'_k = -z_k$) und

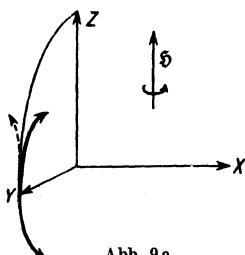


Abb. 9a

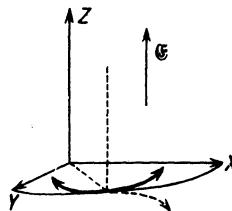


Abb. 9b

Magnetisches Feld in der Z -Richtung:
Spiegelt man die Bahn eines Teilchens
an der XY -Ebene, so erhält man wieder
eine mögliche Bahn, nicht aber durch
Spiegelung an der YZ -Ebene

Elektrisches Feld in der Z -Richtung:
Spiegelt man die Bahn eines Teilchens
an einer Ebene durch die Z -Achse, so
erhält man wieder eine mögliche Bahn,
nicht aber durch Spiegelung an der
 XY -Ebene

daher in der Symmetriegruppe des Systems enthalten. Die gesamte Symmetrie ist das direkte Produkt der Gruppe der reinen Drehungen um Z und der Spiegelungsgruppe (die außer der Einheit noch die Inversion enthält), wozu als dritter Faktor noch die symmetrische Gruppe hinzukommt. Die beiden ersten Gruppen, also auch ihr direktes Produkt ist abelsch.

Wenn aber auch die volle Drehsymmetrie des Problems durch das äußere Feld zerstört wird, haben, so lange das Magnetfeld schwach ist — und die experimentell erzeugbaren Felder sind immer schwach in diesem Sinne —, die Eigenwerte und Eigenfunktionen noch näherungsweise die Lage bzw. die Eigenschaften, die sie ohne

Magnetfeld hatten. Insbesondere kann man auch von einer azimutalen Quantenzahl L sprechen und es gelten auch noch die gewöhnlichen Auswahlregeln für L . Außerdem gehört natürlich jeder Term, auch wenn das äußere Feld beliebig stark ist, zu einer irreduziblen Darstellung der vorher erwähnten drei Gruppen, hat also ebenso wie die Terme des feldlosen Systems ein Multiplettsystem S und einen Spiegelungscharakter. Auch die Auswahlregeln A und C , die aus dem Zugehören der Eigenfunktionen zu Darstellungen der symmetrischen und der Spiegelungsgruppe folgen, bleiben streng erhalten. Hinzu kommt noch die sogenannte magnetische Quantenzahl μ , die angibt, zu welcher Darstellung ($e^{i\mu\varphi}$) der zweidimensionalen reinen Drehgruppe der betrachtete Term gehört. Für μ tritt auch eine neue Auswahlregel auf, sie lautet für in verschiedenen Richtungen polarisiertes Licht verschieden, so daß manche Übergänge (π -Komponenten) nur durch in der Feldrichtung, manche (σ -Komponenten) nur durch senkrecht dazu polarisiertes Licht bewirkt werden. Da die verschiedenen Raumrichtungen nicht gleichberechtigt sind, ist dies nicht weiter verwunderlich.

Für Übergänge mit in der Z -Richtung polarisiertem Licht ist

$$(\psi_F, (z_1 + z_2 + \dots + z_n) \psi_E) \quad (2c)$$

maßgebend. Da das Multiplizieren mit $z_1 + \dots + z_n$ ein in Bezug auf Drehungen um Z symmetrischer Operator ist, müssen ψ_F und ψ_E , wenn (2c) nicht verschwinden soll, zu derselben Darstellung ($e^{i\mu\varphi}$) gehören, dieselbe magnetische Quantenzahl haben. Bei einem Übergang, bei dem das Licht parallel zur Feldrichtung polarisiert ist, ändert sich die magnetische Quantenzahl nicht.

Für die Übergänge, bei denen das Licht senkrecht zur Feldrichtung polarisiert ist (σ -Komponenten), sind \mathbf{X}_{FE} und \mathbf{Y}_{FE} maßgebend. Nun gehört $y_1 + y_2 + \dots + y_n + i(x_1 + x_2 + \dots + x_n)$ zur Darstellung ($e^{i\varphi}$), so daß $[y_1 + \dots + y_n + i(x_1 + \dots + x_n)] \psi_E$ zur Darstellung ($e^{i(u+1)\varphi}$) gehört. Soll

$$\mathbf{Y}_{FE} + i\mathbf{X}_{FE}$$

$$= (\psi_F, [y_1 + y_2 + \dots + y_n + i(x_1 + x_2 + \dots + x_n)] \psi_E)$$

nicht verschwinden, so muß auch ψ_F zur Darstellung ($e^{i(u+1)\varphi}$) gehören. Ebenso schließt man, daß ψ_F zu ($e^{i(u-1)\varphi}$) gehören muß, damit

$$\mathbf{Y}_{FE} - i\mathbf{X}_{FE}$$

$$= (\psi_F, [y_1 + y_2 + \dots + y_n - i(x_1 + x_2 + \dots + x_n)] \psi_E)$$

von Null verschieden sei. Es folgt, daß X_{FE} und Y_{FE} nur endlich sein können, wenn sich die magnetischen Quantenzahlen von ψ_F und ψ_E um 1 unterscheiden. Durch senkrecht zur Feldrichtung polarisiertes Licht werden nur Übergänge mit $\Delta\mu = \pm 1$ angeregt.

Für den umgekehrten Prozeß der Emission folgt hieraus, daß das Licht, das senkrecht zur Feldrichtung ausgestrahlt wird (Transversaleffekt) bei Übergängen $\Delta\mu = 0$ parallel, bei Übergängen $\Delta\mu = 1$ senkrecht zur Feldrichtung polarisiert ist, womit in diesem Fall die Polarisationsrichtung, die ja auch noch senkrecht zur Strahlrichtung sein muß, eindeutig festgelegt ist.

Betrachten wir nun das Licht, das in Richtung des Magnetfeldes ausgestrahlt wird (Longitudinaleffekt)! Dieses muß senkrecht zur Feldrichtung polarisiert sein, es kann also keine π -Komponenten, nur σ -Komponenten enthalten. Der Polarisationszustand der σ -Komponenten ist aber durch die Angabe „senkrecht zur Feldrichtung“ nicht bestimmt. Die Erfahrung zeigt, daß sie teils rechts-, teils linkszirkular polarisiertes Licht enthalten. Dies bedeutet wiederum, daß erstere Übergänge durch linkszirkular polarisiertes, letztere durch rechtszirkular polarisiertes Licht nicht angeregt werden können¹⁾). Nun zeigt eine zu der im VI. Kapitel durchaus analoge Rechnung, daß für Übergänge, die durch in der XY-Ebene zirkular polarisiertes Licht angeregt werden, je nachdem der Drehungssinn von Y zu X oder von X zu Y weist, die Matrixelemente $(Y_{FE} + iX_{FE})/\sqrt{2}$ oder $(Y_{FE} - iX_{FE})/\sqrt{2}$ in (1) an Stelle von X_{FE} treten. Ist also das Licht in Richtung des Feldes gesehen (von unten nach oben, wenn die Z-Achse nach oben zeigt)

¹⁾ Zur Bestimmung des Polarisationszustandes des bei einem Übergang emittierten Lichtes ist es immer nur wesentlich, was für Licht bei dem umgekehrten Prozeß nicht absorbiert wird. Ein Übergang z. B., der zur Z-Achse parallel polarisiertes Licht aussendet, wird — wenn auch schwächer — auch durch Licht angeregt, dessen Polarisationsrichtung zur Z-Achse geneigt ist. Wesentlich ist, daß so ein Übergang nicht durch senkrecht zu Z polarisiertes Licht angeregt werden kann, ebenso wie ein Übergang, der rechtszirkular polarisiertes Licht aussendet, nicht durch linkszirkular polarisiertes Licht angeregt werden kann.

Daß man den Polarisationszustand des emittierten Lichtes auf dem Umweg über den umgekehrten Prozeß bestimmen muß, beruht darauf, daß die Schrödingergleichung in der zugrunde gelegten Gestalt die Emission überhaupt nicht zu erklären imstande ist.

rechtszirkular polarisiert, so bewirkt es einen Sprung mit einer Erhöhung von μ um 1, ist es linkszirkular polarisiert, so bewirkt es ein Fallen von μ um 1. Fällt umgekehrt bei der spontanen Emission μ um 1, so ist das emittierte Licht (in derselben Richtung gesehen) rechts-, steigt es, so ist es linkszirkular polarisiert.

3. Betrachten wir jetzt einen Term des feldlosen Systems und untersuchen wir, wie er sich bei dem Einsetzen des magnetischen Feldes verhalten wird. Der Term E_{SLw} des feldlosen Systems wird durch das Magnetfeld im allgemeinen aufgespalten, und es entstehen mehrere neue Terme aus einem alten Term. Das Multiplettsystem S und der Spiegelungscharakter w werden aber hierdurch nicht berührt, das erste ist für alle neuen Terme, die aus demselben feldlosen Term entstehen, S , der zweite w . Dies folgt daraus, daß die Eigenfunktionen während des ganzen Anwachsens des magnetischen Feldes zu einer Darstellung der symmetrischen Gruppe und der Spiegelungsgruppe gehören und sich außerdem stetig ändern. Eine Änderung der Darstellungseigenschaft würde aber unstetig vor sich gehen müssen.

Es fragt sich noch, welche μ -Werte die aus E_{SLw} entstehenden Terme haben werden? Es sei R eine Drehung um Z mit φ , dann ist nach (6) Kap. XV $\mathfrak{D}^{(L)}(R)$

$$\begin{pmatrix} e^{-iL\varphi} & 0 & \dots & 0 & 0 \\ 0 & e^{-i(L-1)\varphi} & \dots & 0 & 0 \\ \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & \dots & e^{i(L-1)\varphi} & 0 \\ 0 & 0 & \dots & 0 & e^{iL\varphi} \end{pmatrix} \quad (\dagger)$$

und wenn $\psi_{x\mu}$ eine Eigenfunktion von E_{SLw} ist, die zur x -Zeile von $\bar{\mathbf{A}}^{(S)}$ und zur μ -Zeile von $\mathfrak{D}^{(L)}$ gehört, ist

$$\mathbf{P}_R \psi_{x\mu} = \sum_{\mu'} \mathfrak{D}^{(L)}(\{\varphi, 0, 0\})_{\mu' \mu} \psi_{x\mu'} = e^{i\mu\varphi} \psi_{x\mu}, \quad (4)$$

d. h. $\psi_{x\mu}$ gehört zur Darstellung $(e^{i\mu\varphi})$ der Gruppe der Drehungen um Z . Auch diese Tatsache wird sich, während das Feld anwächst, nicht ändern, und da μ bei der Darstellung $\mathfrak{D}^{(L)}$ von $-L$ bis L läuft, wird ein Term mit der Azimutalquantenzahl L in $2L + 1$ Terme $E_{SLw, \mu}$ mit den magnetischen Quantenzahlen $\mu = -L, -L+1, \dots, L-1, L$ aufspalten. Die zu $E_{SLw, \mu}$ ge-

hörigen Eigenfunktionen erster Näherung sind die $\psi_{x\mu}$ selber, da sie zu einer Zeile von $\bar{A}^{(S)}$ und zu einer Darstellung ($e^{i\mu\varphi}$) gehören müssen und man aus den $\psi_{x\mu}$ keine anderen Linearkombinationen mit dieser Eigenschaft bilden kann. Wir haben also wieder einen Fall vor uns, wo die „richtigen Linearkombinationen“ erster Näherung schon durch gruppentheoretische Überlegungen bestimmt werden können.

Dass die Bestimmung der „richtigen Linearkombinationen“ in diesem Falle so einfach war, beruht darauf, dass in $\mathfrak{D}^{(L)}$ die Matrizen, die Drehungen um Z entsprechen — wie dies (†) zeigt —, als Darstellung der Gruppe der Drehungen um Z schon ausreduziert sind. Hätten wir das Magnetfeld etwa in die X -Richtung gelegt, so hätten wir die Matrizen $\mathfrak{d}^{(L)}(\varphi)$, die Drehungen um X entsprechen, ausreduzieren, in die Form (†) bringen müssen. Die Matrix $(T'_{\mu'\mu})$, die zur Ausreduktion dient, erzeugt dann auch die richtigen Linearkombinationen

$$\psi'_{x\mu'} = \sum_{\mu} T'_{\mu'\mu} \psi_{x\mu}$$

für diesen Fall.

Aus den Eigenfunktionen erster Näherung kann man auch die erste Näherung für den Termwert $E_{SLw,\mu}$ berechnen, wenn man die durch das magnetische Feld \mathfrak{H}_z bedingte Modifizierung des Hamiltonschen Operators des Systems kennt. In der klassischen Theorie tritt im Magnetfeld zur feldlosen Hamiltonschen Funktion, wenn man höhere Potenzen der Feldstärke vernachlässigt, das Glied $e/c \cdot (\mathfrak{A}, \mathbf{v}) = e/mc \cdot (\mathfrak{A}_x p_x + \mathfrak{A}_y p_y + \mathfrak{A}_z p_z)$ hinzu, wo \mathfrak{A} das Vektorpotential ist, als dessen Rotation sich die Feldstärke schreiben lässt. In der Quantenmechanik schreibt man hierfür, wenn man in derselben Näherung auch bei Vorhandensein eines äußeren Feldes — $h/2\pi i \cdot \partial/\partial x$ für p_x usw. setzt,

$$\mathbf{V} = \frac{-e}{mc} \left(\mathfrak{A}_x \frac{\hbar}{2\pi i} \frac{\partial}{\partial x} + \mathfrak{A}_y \frac{\hbar}{2\pi i} \frac{\partial}{\partial y} + \mathfrak{A}_z \frac{\hbar}{2\pi i} \frac{\partial}{\partial z} \right), \quad (5)$$

bzw. die entsprechende Summe für mehrere Elektronen¹⁾. Für ein homogenes, in die Z -Achse gerichtetes Magnetfeld von der Intensität \mathfrak{H}_z ist

$$\mathfrak{A}_x = -\frac{1}{2} \mathfrak{H}_z y; \quad \mathfrak{A}_y = \frac{1}{2} \mathfrak{H}_z x, \quad \mathfrak{A}_z = 0.$$

¹⁾ In Wirklichkeit tritt zu (5) noch das Glied $(\mathfrak{A}_x^2 + \mathfrak{A}_y^2 + \mathfrak{A}_z^2)e^2/2mc^2$ hinzu, das unter anderem für den Diamagnetismus verantwortlich ist.

Die erste Näherung für die magnetische Zusatzenergie $E_{SLw,\mu} - E_{SLw}$ berechnet sich dann nach Kap. V, Gleichung (22)

$$E_{SLw,\mu} - E_{SLw} = (\psi_{x\mu}, \mathbf{V} \psi_{x\mu}) = -\frac{-e\hbar\mathfrak{H}_z}{4\pi mc} (\psi_{x\mu}, L_z \psi_{x\mu}), \quad (6)$$

wo

$$L_z = \frac{1}{i} \left(y_1 \frac{\partial}{\partial x_1} + \dots + y_n \frac{\partial}{\partial x_n} - x_1 \frac{\partial}{\partial y_1} - \dots - x_n \frac{\partial}{\partial y_n} \right) \quad (6a)$$

ist. Das in (6) auftretende skalare Produkt läßt sich vollkommen auswerten. Es wird sich nämlich zeigen, daß für jede Funktion f

$$L_z f = \frac{1}{i} \frac{\partial \mathbf{P}_{\{\varphi, 0, 0\}} f}{\partial \varphi} \quad \text{für } \varphi = 0, \quad (7)$$

also gleich der Differenz der Werte von f in einem etwas „verdrehten Zustande“ und dem ursprünglichen Zustand dividiert durch den Drehwinkel. Da

$$\{\varphi, 0, 0\} = \begin{pmatrix} \cos \varphi & -\sin \varphi & 0 \\ \sin \varphi & \cos \varphi & 0 \\ 0 & 0 & 1 \end{pmatrix},$$

ist

$$\mathbf{P}_{\{\varphi, 0, 0\}} f(\dots, x_k, y_k, z_k, \dots)$$

$$= f(\dots, x_k \cos \varphi + y_k \sin \varphi, -x_k \sin \varphi + y_k \cos \varphi, z_k, \dots),$$

und dies nach φ differentiiert ergibt für $\varphi = 0$

$$\left(\frac{\partial \mathbf{P}_{\{\varphi, 0, 0\}} f}{\partial \varphi} \right)_{\varphi=0} = \sum_k y_k \frac{\partial f}{\partial x_k} - x_k \frac{\partial f}{\partial y_k}, \quad (7a)$$

was mit (7) äquivalent ist. Nun ist noch wegen (4)

$$\frac{\partial}{\partial \varphi} \mathbf{P}_{\{\varphi, 0, 0\}} \psi_{x\mu} = \frac{\partial}{\partial \varphi} e^{i\mu\varphi} \psi_{x\mu} = i\mu e^{i\mu\varphi} \psi_{x\mu} \quad (7b)$$

und daher wegen der Normierung $(\psi_{x\mu}, \psi_{x\mu}) = 1$ und (6)

$$E_{SLw,\mu} - E_{SLw} = \frac{e\hbar\mathfrak{H}_z\mu}{4\pi mc} (\mu = -L, -L+1, \dots, L-1, L). \quad (8)$$

Nach (8) soll der Term mit der azimutalen Quantenzahl L in erster Näherung — d. h. wenn man sich auf Glieder beschränkt, die mit der ersten Potenz von \mathfrak{H}_z proportional sind — in $2L+1$ äquidistante Terme aufspalten, deren mittlerer ($\mu = 0$) die Lage des ursprünglichen Terms hat und deren Abstand bei derselben Feldstärke für alle Terme derselbe ist: in (8) kommen ja nur universelle Konstanten vor.

Betrachten wir jetzt die Zeemankomponenten der Linie $F \rightarrow E$, so sehen wir — da der Term F ebenso stark wie E aufspaltet —, daß sich die Linien mit gleicher Änderung von μ alle überdecken. Da sich aber μ bei einem optischen Übergang nur um ± 1 oder 0 ändern kann, hat man im ganzen nur drei getrennte Linien und die beiden seitlichen Komponenten haben für alle Linien denselben Abstand von der mittleren. Das ist das Aufspaltungsbild des sogenannten normalen Zeeman-Effektes.

Dieses Aufspaltungsbild deckt sich mit dem beobachteten nur in verhältnismäßig seltenen Fällen, nämlich nur bei Linien die Singulettterme verbinden. Bei diesen Termen kompensieren sich nämlich die magnetischen Momente der Elektronen, die sonst die Abweichungen verursachen, so daß ihr Einfluß verschwindet. Dies ist auch der Grund, warum diese Terme keine Feinstruktur haben. Bei allen anderen Termen ist die Aufspaltung teils größer, teils kleiner und für verschiedene Terme zumeist verschieden. Daher überdecken sich auch die Linien, die mit derselben Änderung von μ verbunden sind, nicht, und das Aufspaltungsbild des anomalen Zeeman-Effekts ist wesentlich komplizierter. Das Intensitätsverhältnis der einzelnen Zeemankomponenten wollen wir an dieser Stelle nicht berechnen¹⁾, weil sich die meisten Komponenten doch überdecken.

Es werde noch bemerkt, daß durch ein magnetisches Feld die Terme soweit aufgespalten werden, wie das überhaupt durch ein äußeres Feld geschehen kann: die übriggebliebenen Entartungen röhren alle von der symmetrischen Gruppe her und die Gleichheit der Elektronen kann durch äußere Felder nicht zerstört werden.

4. Bei einem homogenen, in der Richtung der Z-Achse zeigenden elektrischen Felde ist die Symmetrie nicht genau dieselbe, wie bei einem magnetischen Felde, da der elektrische Feldvektor polaren Charakter hat. Dadurch wird die Existenz des Inversionszentrums aufgehoben, dagegen bleiben die zum Felde parallelen Ebenen durch den Kern Symmetrieebenen. Die Verhältnisse liegen also umgekehrt wie bei magnetischen Feldern (siehe Abb. 9). Die Symmetriegruppe ist die zweidimensionale Drehspiegelungsgruppe (also nicht abelsch!), während sie im Magnetfeld das direkte Produkt der

¹⁾ Dies soll im Kap. XXIII geschehen. Die Intensitäten der drei beobachtbaren Linien sind wie in der klassischen Theorie des Zeeman-Effekts, die den normalen Zeeman-Effekt genau ergibt.

zweidimensionalen reinen Drehgruppe und der Spiegelungsgruppe war. Jeder Term hat außer dem Multiplettsystem S noch eine elektrische Quantenzahl $m = 0, 0', 1, 2, \dots$, die angibt, zu welcher Darstellung $\mathfrak{Z}^{(m)}$ der zweidimensionalen Drehspiegelungsgruppe der betrachtete Term gehört.

Die irreduziblen Darstellungen der zweidimensionalen Drehspiegelungsgruppe wurden im Kap. XIV bestimmt. Reinen Drehungen mit φ entsprechen in $\mathfrak{Z}^{(0)}, \mathfrak{Z}^{(0')}, \mathfrak{Z}^{(1)}, \mathfrak{Z}^{(2)}, \dots$ der Reihe nach die Matrizen

$$(1), \quad (1), \quad \begin{pmatrix} e^{-i\varphi} & 0 \\ 0 & e^{i\varphi} \end{pmatrix}, \quad \begin{pmatrix} e^{-2i\varphi} & 0 \\ 0 & e^{2i\varphi} \end{pmatrix}, \dots$$

während einer Spiegelung an Y die Matrizen

$$(1), \quad (-1), \quad \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \dots$$

entsprechen.

Um die m -Werte der Terme festzustellen, in die ein Term mit der azimutalen Quantenzahl L aufspaltet, müssen wir feststellen, welche $\mathfrak{Z}^{(m)}$ und wie oft sie in den $\mathfrak{D}^{(L)}(R)$ vorkommen, die Drehspiegelungen um Z entsprechen? Aus der Form (\dagger) von $\mathfrak{D}^{(L)}(R)$, wenn R eine reine Drehung um Z ist, ersehen wir unmittelbar, daß in (\dagger) $\mathfrak{Z}^{(1)}, \mathfrak{Z}^{(2)}, \dots, \mathfrak{Z}^{(L)}$ je einmal enthalten ist, die zur ersten bzw. zweiten Zeile von $\mathfrak{Z}^{(m)}$ gehörigen Eigenfunktionen gehören zur $-m$ - bzw. m -Zeile von $\mathfrak{D}^{(L)}$. Dagegen können wir von der Eigenfunktion, die zur 0-Zeile von $\mathfrak{D}^{(L)}$ gehört, nur sagen, daß sie entweder zu $\mathfrak{Z}^{(0)}$ oder zu $\mathfrak{Z}^{(0')}$ gehört. Zur Entscheidung der Frage, ob das erste, oder das letzte der Fall ist, muß man noch eine Drehspiegelung, etwa $x' = -x$ hinzuziehen, da sich $\mathfrak{Z}^{(0)}$ und $\mathfrak{Z}^{(0')}$ für reine Drehungen überhaupt nicht unterscheiden.

Um die Spur der zu dieser Transformation zugeordneten Matrix in $\mathfrak{D}^{(L)}$ zu finden, bemerken wir, daß sie das Produkt der Inversion und einer Drehung um X mit π ist und ihre Spur daher

$$\begin{aligned} w(1 + 2 \cos \pi + 2 \cos 2\pi + \dots + 2 \cos L\pi) \\ = w(1 - 2 + 2 - \dots + 2 \cdot (-1)^L) = w \cdot (-1)^L \end{aligned} \quad \left. \right\} \quad (9)$$

beträgt, wo w für positive Terme $+1$, für negative -1 ist. Da die Bestandteile $\mathfrak{Z}^{(1)}, \dots, \mathfrak{Z}^{(L)}$ nichts zur Spur der Spiegelungen beitragen, ist der übriggebliebene Term ein 0-Term, wenn $w(-1)^L = +1$, und ein 0'-Term, wenn $w(-1)^L = -1$ ist.

Wir sehen, daß die Aufspaltung der Terme im elektrischen Felde keine so vollkommene wie im magnetischen Felde ist: aus einem Term mit der Azimutalquantenzahl L entstehen nur $L + 1$ Terme.

Die Auswahlregeln, die bei starken elektrischen Feldern gelten, sind den entsprechenden Regeln für Magnetfelder ähnlich: die elektrische Quantenzahl m ändert sich bei einem Übergang mit in der Z -Achse polarisiertem Licht nicht, da das Multiplizieren mit $z_1 + z_2 + \dots + z_n$ ein gegenüber der zweidimensionalen Drehspiegelungsgruppe der Drehungen um Z symmetrischer Operator ist. Daher ist auch ein solcher Übergang von einem 0-Term zu einem 0'-Term verboten. Dagegen ändert sich m bei einem Übergang, bei dem das ausgestrahlte Licht senkrecht zur Feldrichtung polarisiert ist, um ± 1 .

Die Auswahlregel für die Azimutalquantenzahl wird in starken elektrischen Feldern durchbrochen, weil die volle Drehsymmetrie nicht mehr da ist und die Eigenfunktionen zu keinen Darstellungen der dreidimensionalen Drehgruppe mehr gehören. Auch die Laportesche Regel verliert ihre Gültigkeit (während sie durch magnetische Felder nicht gestört wurde); es bleibt aus ihr nur das Verbot des Überganges von einem 0-Term zu einem 0'-Term übrig.

Man kann auch die Eigenwertstörung durch das elektrische Feld mit Hilfe des Rayleigh-Schrödingerschen Verfahrens formal berechnen. Dem Resultat ist nur ein bedingter Sinn beizulegen, da das Verfahren wegen der Gestalt des Störungsgliedes

$$\mathbf{V} = e \mathfrak{E}_z (z_1 + z_2 + \dots + z_n) \quad (10)$$

divergieren muß¹⁾. Stellt man nämlich, etwa bei dem Wasserstoffatom, das Potential als Funktion des Abstandes vom Kern graphisch dar, so sieht man, daß zwar in der Nähe des Kerns ein tiefes Potentialminimum liegt, daß aber das Elektron immer genügend Energie hat, um sich in der Feldrichtung ins Unendliche zu entfernen. Dies macht es schon sehr wahrscheinlich, daß im elektrischen Feld streng genommen überhaupt kein diskretes Spektrum, keine streng stationären Zustände existieren. Die erste bzw. zweite Näherung, die wir nach dem Schrödingerschen Verfahren berechnen können, ist trotzdem nicht ganz sinnlos: sie ergibt Zu-

¹⁾ Vgl. J. R. Oppenheimer, Phys. Rev. **31**, 66, 1928.

stände, die, wenn sie auch keineswegs stationär sind, sich während sehr langer Zeiten stationär verhalten, dementsprechend, daß sich das Elektron im Mittel nur nach verhältnismäßig sehr langer Zeit entschließen wird, den Sprung über die Potentialschwelle zu machen und sich vom Kern zu entfernen.

Berechnet man die erste Näherung für die Energiestörung im elektrischen Feld, so sieht man, daß die Eigenwerte in erster Näherung überhaupt nicht aufspalten. Die Koeffizienten $v_{x' \mu'; x \mu}$,

$$v_{x' \mu'; x \mu} = e \mathfrak{E}_s (\psi_{x' \mu'}, (z_1 + z_2 + \dots + z_n) \psi_{x \mu}) = 0, \quad (11)$$

der Säkulargleichung von (18) Kap. V, sind alle Null, weil $\psi_{x' \mu'}$ und $(z_1 + z_2 + \dots + z_n) \psi_{x \mu}$ verschiedenen Spiegelungscharakter haben. Liegt nämlich keine zufällige Entartung vor, so haben alle Eigenfunktionen desselben Eigenwertes E — etwa $\psi_{x' \mu'}$ und $\psi_{x \mu}$ — denselben Spiegelungscharakter, $\psi_{x' \mu'}$ und $(z_1 + z_2 + \dots + z_n) \psi_{x \mu}$ also verschiedenen, weil $z_1 + \dots + z_n$ den Spiegelungscharakter umkehrt. Wir haben dies bei dem Operator, der die Übergangswahrscheinlichkeiten regelte, gesehen; mit diesem ist (10) bis auf einen konstanten Faktor identisch. Die Eigenwerte von $(v_{x' \mu'; x \mu}) = 0$ sind alle 0. In erster Näherung fallen alle Terme mit dem ungestörten Term zusammen, in der Energiestörung, nach Potenzen der Feldstärke entwickelt, ist der Koeffizient der ersten Potenz Null, die Verschiebung der Terme geht bei kleiner Feldstärke quadratisch gegen Null. Nur bei dem Wasserstoffatom, bei dem durch eine zufällige Entartung Terme verschiedenen Spiegelungscharakters zusammenfallen, ist ein linearer Effekt vorhanden.

Der experimentellen Prüfung der soeben für den Starkeffekt abgeleiteten Gesetzmäßigkeiten stehen — ebenso wie bei dem Zeeman-Effekt — in erster Linie die Komplikationen, die durch das magnetische Moment der Elektronen bedingt sind, im Wege. Das einzige Resultat, das allgemeine Gültigkeit beansprucht, ist das Nichtvorhandensein einer mit der ersten Potenz der Feldstärke proportionalen Termverschiebung, da dies aus der Betrachtung der Spiegelungssymmetrie allein folgte.

5. Für freie Atome bilden konstante magnetische bzw. elektrische Felder wohl die wichtigsten Fälle von äußeren Störungen. Anders bei einem Atom im Kristallverband. Bei diesem ist die

Symmetrie des „äußeren“ Feldes¹⁾), das in diesem Fall von den umgebenden Atomen herröhrt, durch die Kristallsymmetrie gegeben, was zu interessanten Aufspaltungsbildern Anlaß geben kann. Diese wurden von H. Bethe²⁾ für die meisten Symmetrieklassen eingehend untersucht, wir greifen aus seinen Beispielen nur den verhältnismäßig einfachen Fall der rhombisch-hemimorphen Symmetrie, der Symmetrie einer rhombischen Pyramide, heraus.

Die rhombische Pyramide hat drei Symmetrieelemente: die Drehung um Z mit π , die Spiegelung an der ZX - und an der ZY -Ebene. Ihre Symmetriegruppe V_d besteht aus der Einheit und diesen drei Elementen, sie ist der Vierergruppe holomorph, da alle ihre Elemente von der Ordnung zwei sind. Sie hat — da sie ja abelsch ist — vier irreduzible, eindimensionale Darstellungen, die der Reihe nach durch die Matrizen:

	E	Drehung um Z mit π	Spiegelung an der ZX -Ebene	Spiegelung an der ZY -Ebene
I	(1)	(1)	(1)	(1)
II	(1)	(-1)	(-1)	(1)
III	(1)	(-1)	(1)	(-1)
IV	(1)	(1)	(-1)	(-1)

gegeben sind. Die erste ist die identische Darstellung, die zweite und dritte sind gleichberechtigt, da in ihnen nur die Rolle der X - und Y -Achse vertauscht ist, während die vierte eine ausgezeichnete Rolle spielt.

Bringen wir ein Atom an seine Stelle im Kristall, so wirken darauf Kräfte, die die volle Raumsymmetrie aufheben, so daß nur die rhombisch-hemimorphe Symmetrie übrigbleibt. Da die irreduziblen Darstellungen dieser Gruppe alle eindimensional sind, spaltet ein Term mit der Azimutalquantenzahl L in $2L+1$ Terme auf.

Die Frage, die hier beantwortet werden soll, ist die: wie viele Terme der Darstellungseigenschaften I, II, III, IV entstehen im Kristall aus einem Term mit der Azimutalquantenzahl L und dem Spiegelungscharakter w ? Man löst diese Frage nach der all-

¹⁾ Natürlich liegt darin, daß man dieses Feld als „äußeres“ betrachtet und die umgebenden Atome nicht zum System dazurechnet, eine gewisse Vernachlässigung. Es sind im wesentlichen die „Austauschkräfte“, die man so außer acht läßt.

²⁾ H. Bethe, Ann. d. Phys. 3, 133, 1929.

gemeinen Theorie: man bestimmt, wie oft die Darstellungen I, II, III, IV der rhombisch-hemimorphen Gruppe in der Darstellung $\mathfrak{D}^{(L, w)}$ der Drehgruppe enthalten sind, wenn man diese als Darstellung ihrer rhombisch-hemimorphen Untergruppe ansieht. Am einfachsten bestimmt man diese Zahlen $\alpha_I, \alpha_{II}, \alpha_{III}, \alpha_{IV}$ durch die Bestimmung des Charakters von $\mathfrak{D}^{(L, w)}$ für die Operationen von V_d . Für die Einheit ist

$$2L + 1 = \alpha_I + \alpha_{II} + \alpha_{III} + \alpha_{IV}, \quad (12a)$$

für die Drehung um Z mit π bzw. die Spiegelung an der ZX - und ZY -Ebene dagegen nach (9)

$$(-1)^L = \alpha_I - \alpha_{II} - \alpha_{III} + \alpha_{IV}, \quad (12b)$$

$$w(-1)^L = \alpha_I - \alpha_{II} + \alpha_{III} - \alpha_{IV} = \alpha_I + \alpha_{II} - \alpha_{III} - \alpha_{IV}. \quad (12c)$$

Aus (12c) folgt $\alpha_{II} = \alpha_{III}$ und $w(-1)^L = \alpha_I - \alpha_{IV}$, aus (12a) und (12b) $2L + 1 + (-1)^L = 2\alpha_I + 2\alpha_{IV}$. Es ergibt sich so für

S_+ -Terme	$\alpha_I = 1$	$\alpha_{II} = \alpha_{III} = 0$	$\alpha_{IV} = 0$
S_- „	0	0	1
P_+ „	0	1	1
P_- „	1	1	0
D_+ „	2	1	1 usw.

Eine Kontrolle für die Rechnung besteht immer darin, daß die α positive ganze Zahlen sein müssen.

In seiner obengenannten Arbeit hat H. Bethe für fast alle 32 im Kristall vorkommenden Symmetrieverhältnisse die Aufspaltungen bestimmt und aus ihnen weitere Schlüsse gezogen. So gewinnt man z. B. die Auswahlregeln für die Terme der Art I, II, III, IV sehr einfach, indem man etwa für die Strahlung, die in der Z -Achse polarisiert ist, bemerkt, daß das Multiplizieren mit $(z_1 + z_2 + \dots + z_n)$ ein der rhombisch-hemimorphen Gruppe gegenüber symmetrischer Operator ist, so daß nur Terme mit der gleichen Darstellung miteinander kombinieren.

XIX. Teilweise Bestimmung der Eigenfunktionen aus ihren Transformations-eigenschaften

1. Die Transformationseigenschaften der Eigenfunktionen, die im vorangehenden Kapitel eine so wesentliche Rolle gespielt haben, können nur so zustande kommen, daß die Werte der Eigenfunktionen für solche Werte der Argumente, die durch die Transformationen der Gruppe ineinander übergeführt werden, in irgendeiner Weise zusammenhängen. Besteht die Gruppe z. B. aus der Einheit und aus der Transformation $x' = -x$, so gilt für Funktionen, die zur identischen Darstellung gehören, d. h. für gerade Funktionen,

$$g(-x) = g(x), \quad (1)$$

während für Funktionen, die zur negativen Darstellung gehören, d. h. für ungerade Funktionen,

$$f(-x) = -f(x) \quad (1a)$$

gilt. Im allgemeinen folgt aus

$$\mathbf{P}_R \psi_x(x_1, x_2, \dots, x_n) = \sum_{\lambda} \mathbf{D}(R)_{\lambda x} \psi_{\lambda}(x_1, x_2, \dots, x_n) \quad (2)$$

nach Kap. XI (26 a)

$$\psi_x(x'_1, x'_2, \dots, x'_n) = \sum_{\lambda} \mathbf{D}(R)_{\lambda x}^* \psi_{\lambda}(x_1, x_2, \dots, x_n), \quad (3)$$

wo die x'_1, x'_2, \dots, x'_n durch die Transformation R aus den x_1, x_2, \dots, x_n hervorgehen. Zerlegt man also den ganzen Variabilitätsbereich der Argumente der Wellenfunktion, den ganzen Konfigurationsraum in Teile, die aus einem Teil, dem Grundgebiet, durch die Transformationen der Gruppe hervorgehen, so kann man nach (3) die ψ_x überall berechnen, wenn man sie im Grundgebiet kennt. Es bedeutet (3) eine je nach der Größe der Gruppe, der gegenüber das Eigenwertproblem invariant ist, mehr oder weniger wesentliche Reduktion des Variabilitätsbereiches der Argumente x_1, \dots, x_n und bringt gewissermaßen die Transformationseigenschaften der ψ_x explizite zum Ausdruck. Man muß daher aus (3) alle Resultate ableiten können, die aus den Invarianzeigenschaften der ψ folgen. Z. B. erhält man für das skalare Produkt einer ungeraden und einer geraden Funktion

$$\int_{-\infty}^{\infty} g(x)^* f(x) dx, \quad (4)$$

wenn man das Integrationsgebiet in zwei Teile (von $-\infty$ bis 0 und von 0 bis ∞) zerlegt, die auseinander durch die Transformation $x' = -x$ hervorgehen,

$$\int_{-\infty}^{\infty} g(x)*f(x) dx = \int_{-\infty}^0 g(x)*f(x) dx + \int_0^{\infty} g(x)*f(x) dx.$$

Wenn man nun noch im ersten Integral für $-x$ die Variable y einführt und $g(-y)$ und $f(-y)$ mit Hilfe von (1) bzw. (1a) ausdrückt, so geht dies in

$$-\int_0^{\infty} g(y)*f(y) dy + \int_0^{\infty} g(x)*f(x) dx = 0 \quad (5)$$

über: die beiden Teile des Integrals (4) heben sich gegenseitig weg.

Formal ist es natürlich viel einfacher, zu sagen, f und g gehören zu verschiedenen irreduziblen Darstellungen, ihr skalares Produkt muß daher verschwinden. Von (3) auszugehen, hat dagegen — neben der größeren Anschaulichkeit — den Vorteil, daß man zu einer teilweisen Berechnung der Eigenfunktionen kommt, die bei der Verwendung der Drehgruppe bei einfachen Problemen ziemlich weitreichend ist.

2. Mit Hilfe von (3) kann man die Wellenfunktion an allen jenen Stellen berechnen, die aus der Stelle $P = (x_1, y_1, z_1, x_2, y_2, z_2, \dots, x_n, y_n, z_n)$ durch eine Transformation der Gruppe hervorgehen, wenn man die Werte aller ihrer Partner an der Stelle P kennt. Bei der Drehgruppe sind das alle jene Stellen des Konfigurationsraumes, in denen die gegenseitige Lage der Atome eine vorgegebene ist. Die Punkte des Konfigurationsraumes können anschaulich durch ein im Zentrum stehendes n -Bein im dreidimensionalen Raum charakterisiert werden. Jedes Bein zeigt nach derjenigen Stelle des dreidimensionalen Raumes, in dem sich das zugehörige Elektron bei der betreffenden Konfiguration befindet. Die Wellenfunktion in allen Punkten des Konfigurationsraumes kennen, heißt, sie für alle erdenklichen n -Beine zu kennen.

Im Sinne der Entwicklungen auf S. 190 über die „Abseparierung des Schwerpunktes“ soll die Wellenfunktion eigentlich auch die Koordinaten des Kernes als Variable enthalten. Sie soll also nicht nur für diejenigen Lagen des n -Beines definiert sein, bei denen das n -Bein im Zentrum steht, sein Mittelpunkt soll vielmehr die Lage des Kernes bedeuten und jede Stelle im Raum einnehmen können. Da aber die Wellenfunktion für alle jene Lagen des n -Beines gleich sein soll, die auseinander durch Parallel-

verschiebung hervorgehen, genügt es, sie für alle im Zentrum stehenden n -Beine anzugeben. Die Wellenfunktion soll sich ja nicht ändern, wenn man alle x -Koordinaten (auch die des Kernes) oder alle y -Koordinaten oder alle z -Koordinaten um denselben Betrag vergrößert oder verkleinert.

Stellen, die auseinander durch eine Drehung hervorgehen, entspricht dieselbe Gestalt des n -Beines, das nur im Raum verschieden gelagert (verdreht, nicht parallelverschoben!) ist. Als Grundgebiet wählen wir diejenigen Lagen, bei denen das erste Bein (dem ersten Elektron zugeordnet) in der Z -Achse liegt, das zweite Bein in der ZY -Ebene. (Dem entsprechen Punkte des Konfigurationsraumes, für die $x_1 = y_1 = z_2 = 0$ ist.) Sind die Werte der $2L + 1$ zu der Darstellung $\mathfrak{D}^{(L)}(\{\alpha \beta \gamma\})$ gehörigen Wellenfunktionen $\psi_{-L}, \psi_{-L+1}, \dots, \psi_{L-1}, \psi_L$ im Grundgebiet der Reihe nach $G_{-L}, G_{-L+1}, \dots, G_{L-1}, G_L$ [wo also $G_\lambda = \psi_\lambda(0, 0, z_1, 0, y_2, z_2, \dots, x_n, y_n, z_n)$ ist und nur von der gegenseitigen Konfiguration der Teilchen abhängt], so ist der Wert der Wellenfunktion an jener Stelle $x'_1, y'_1, z'_1, \dots, x'_n, y'_n, z'_n$, die aus $0, 0, z_1, 0, y_2, z_2, \dots, x_n, y_n, z_n$ durch die Drehung $\{\alpha \beta \gamma\}$ hervorgeht, nach (3)

$$\psi_\mu(x'_1, y'_1, z'_1, \dots, x'_n, y'_n, z'_n) = \sum_{\lambda=-L}^L \mathfrak{D}^{(L)}(\{\alpha \beta \gamma\})_{\mu \lambda}^* G_\lambda(g). \quad (6)$$

Dabei sind α und β definitionsgemäß Azimut und Polabstand des ersten Elektrons und γ jener Winkel, den die Ebene durch die Z -Achse und das erste Elektron mit der Ebene durch den Ursprung und die ersten zwei Elektronen einschließt. Die G hängen nur von der Gestalt g des n -Beines ab.

Für $L = 0$, für S-Terme lautet (6)

$$\psi(x'_1, y'_1, z'_1, \dots, x'_n, y'_n, z'_n) = G_0(g), \quad (7)$$

die Wellenfunktion hängt in diesem Falle auch selbst nur von der Gestalt des n -Beines, nicht von seiner Lage im Raum ab: S-Zustände sind kugelsymmetrisch¹⁾. Dies ist auch ganz natürlich: zu einem S-Terme gehört nur eine Eigenfunktion und diese kann daher keine Richtung auszeichnen. Bei höheren Azimutalquantenzahlen ist zwar durch die Gesamtheit der Eigenfunktionen auch keine Richtung ausgezeichnet, doch kann man keine Eigenfunktion herauswählen, ohne eine Richtung auszuzeichnen, so daß die einzelnen Eigenfunktionen nicht mehr kugelsymmetrisch sind.

¹⁾ Vgl. A. Unsöld, Ann. d. Phys. 82, 355, 1927.

Aus Gleichung (6) kann man auch die Auswahlregeln für L ableiten. Wir wollen hier sehen, wie weit man mit ihrer Hilfe die Eigenfunktionen explizite bestimmen kann.

3. Wenn die Eigenfunktionen nur von α , β und γ abhängen, sind sie durch (6) vollkommen bestimmt, weil die G_λ reine Zahlen werden.

Die Gleichung des räumlichen Rotators lautet:

$$-\frac{\hbar^2}{8\pi^2 J} \left[\frac{1}{\sin \vartheta} \frac{\partial}{\partial \vartheta} \sin \vartheta \frac{\partial \psi_\mu^{NL}(\vartheta, \varphi)}{\partial \vartheta} + \frac{1}{\sin^2 \vartheta} \frac{\partial^2 \psi_\mu^{NL}(\vartheta, \varphi)}{\partial \varphi^2} \right] = E_L^N \psi_\mu^{NL}(\vartheta, \varphi), \quad (8)$$

wo J das Trägheitsmoment, ϑ und φ Polabstand und Azimut des Rotators sind. Das Grundgebiet ist hier ein einziger Punkt $\vartheta = 0$, die „Normallage“ des Rotators. Bezeichnen wir die Werte der Eigenfunktionen an dieser Stelle mit G_λ^{NL} , so gilt nach (8) Kap. XV und (6)

$$\begin{aligned} \psi_\mu^{NL}(\vartheta, \varphi) &= \sum_\lambda \mathfrak{D}^{(L)}(\varphi, \vartheta, \gamma)_n^\star G_\lambda^{NL} \\ &= \sum_\lambda e^{-i\mu\varphi} d^{(L)}(\vartheta)_{\mu\lambda} e^{-i\lambda\gamma} G_\lambda^{NL}, \end{aligned} \quad (8a)$$

da aber $\psi_\mu^{NL}(\vartheta, \varphi)$ von γ nicht abhängen darf, müssen für $\lambda \neq 0$ die $G_\lambda^{NL} = 0$ sein, so daß (8a) in

$$\psi_\mu^{NL}(\vartheta, \varphi) = e^{-i\mu\varphi} d^{(L)}(\vartheta)_{\mu 0} G_0^{NL} \quad (8b)$$

übergeht, womit die Eigenfunktionen vollkommen mit Hilfe der Darstellungskoeffizienten ausgedrückt sind. Die Gleichung (8b) zeigt auch, daß sich die Eigenfunktionen ψ_μ^{NL} für gleiches L und μ und verschiedene N höchstens durch einen konstanten Faktor unterscheiden, was für Eigenfunktionen verschiedener Eigenwerte nicht möglich wäre. Zu jedem L gehört in diesem Falle nur ein einziger Eigenwert, so daß man den Index N in (8), (8a), (8b) weglassen kann.

Die Lösungen von (8) sind andererseits als die Kugelflächenfunktionen, die Kugelfunktionen vom Grade L bekannt, (8b) zeigt, daß die $\mathfrak{D}^{(l)}(\{\varphi \vartheta \gamma\}_m^*)$, bis auf den Normierungsfaktor mit den Kugelfunktionen $P_m^l(\vartheta, \varphi)$ identisch sind.

Bei näherem Zusehen ist es keineswegs verwunderlich, daß wir (8) ganz ohne Rechnung lösen konnten. Eine Methode zur Bestimmung der Darstellungen $\mathfrak{D}^{(l)}$ war ja die (Kap. XV, 1), daß wir die (8) äquivalente Laplacesche Differentialgleichung gelöst haben. Jetzt haben wir diese Lösung gewissermaßen nur in (8) wieder eingesetzt.

Nicht a priori selbstverständlich dagegen ist es, daß die Lösung von (8) für alle kugelsymmetrischen Probleme maßgebend ist. So besagt z. B. im Falle des Wasserstoffatoms

$$-\frac{\hbar^2}{8\pi^2 m} \left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2} \right) \psi_{\mu}^{Nl} - \frac{e^2}{r} \psi_{\mu}^{Nl} = E_{Nl} \psi_{\mu}^{Nl} \quad (9)$$

die Gleichung (6), daß

$$\psi_{\mu}^{Nl} = \sum_{\lambda} \mathfrak{D}^{(l)}(\{\alpha \beta \gamma\})_{\mu \lambda}^* G_{\lambda}^{Nl}(r) \quad (9a)$$

gilt, da das n -Bein hier in ein ein-Bein ausartet, dessen Gestalt schon durch die Länge r des Beines, den Abstand des Elektrons vom Zentrum gegeben ist. Dabei sind α und β Azimut und Polabstand des Elektrons, während γ wieder bedeutungslos ist und in (9a) daher nicht wirklich vorkommen darf. Daraus folgt, ebenso wie bei (8b), daß $G_{\lambda}^{Nl}(r)$ für $\lambda \neq 0$ verschwinden muß:

$$\psi_{\mu}^{Nl}(r, \vartheta, \varphi) = \mathfrak{D}^{(l)}(\{\varphi, \vartheta, 0\})_{\mu 0}^* G^{Nl}(r) \approx P_m^l(\vartheta, \varphi) G^{Nl}(r). \quad (9b)$$

Nach (3), Kap. XVII, haben die Wasserstoffeigenfunktionen tatsächlich diese Form. Wir sehen gleichzeitig, daß das ψ_{μ}^{Nl} tatsächlich zur μ -Zeile von $\mathfrak{D}^{(l)}$ gehört, wie dies für eine Eigenfunktion mit der magnetischen Quantenzahl μ und der azimutalen l auch sein muß.

Seine volle Wirksamkeit erlangt diese Methode erst bei der Quantenmechanik des Kreisels. Betrachten wir zuerst den asymmetrischen Kreisel! Die Lage des Kreisels kann durch die drei Eulerschen Winkel $\alpha \beta \gamma$ der Drehung gekennzeichnet werden, die den Kreisel aus seiner Normallage (bei der die größte Trägheitsachse mit der Z-Achse, die mittlere mit der Y-Achse und die kleinste mit der X-Achse zusammenfällt) in die fragliche Lage bringt. Die Wellenfunktion wird von diesen drei Winkeln abhängen, und zwar ist nach (6)

$$\begin{aligned} \psi_{\mu}^{Nl}(\alpha \beta \gamma) &= \sum_{\lambda} \mathfrak{D}^{(l)}(\{\alpha \beta \gamma\})_{\mu \lambda}^* G_{\lambda}^{Nl} \\ &= \sum_{\lambda} e^{-i\mu\alpha} \mathfrak{d}^{(l)}(\beta)_{\mu \lambda} e^{-i\lambda\gamma} G_{\lambda}^{Nl}, \end{aligned} \quad (10)$$

wo die G_{λ}^{Nl} Konstanten sind. Man kann sie sowie auch den Eigenwert E_{Nl} bestimmen, indem man (10) in die Schrödingergleichung einsetzt. Man erhält dann $2l+1$ lineare homogene Gleichungen für die $G_{-l}^{Nl}, \dots, G_l^{Nl}$. Die Bedingung für das Verschwinden der Determinante dieses Gleichungssystems ist eine algebraische

Gleichung $(2l+1)$ -ten Grades für die Energie E^{Nl} , so daß $2l+1$ Eigenwerte die Azimutalquantenzahl l haben.

Betrachten wir jetzt einen Kreisel, dessen zwei kleinere Trägheitsmomente gleich sind! Dann ist die „Normallage“ des Kreisels nicht eindeutig definiert, es bleibt noch eine Drehung um Z frei. Dies hat zur Folge, daß eine Eigenfunktion Eigenfunktion bleibt, wenn man in ihr γ durch $\gamma + \gamma_0$ ersetzt.

Auch eine Linearkombination

$$\int_0^{2\pi} \psi_{\mu}^{Nl}(\alpha, \beta, \gamma + \gamma_0) e^{i\nu\gamma_0} d\gamma_0 \\ = \sum_{\lambda} G_{\lambda}^{Nl} e^{-i\mu\alpha} \mathbf{d}^{(l)}(\beta)_{\mu\lambda} e^{-i\lambda\gamma} \int_0^{2\pi} e^{i\gamma_0(\nu-\lambda)} d\gamma_0 \quad (11)$$

solcher Funktionen ist Eigenfunktion. Nun ist aber, wenn G_{ν}^{Nl} nicht Null ist, (11) bis auf eine Proportionalitätskonstante nichts anderes als $\mathfrak{D}^{(l)}(\{\alpha\beta\gamma\})_{\mu\nu}$, selber, so daß die Eigenfunktionen, wenn man $\nu = N$ zur Laufzahl macht,

$\psi_{\mu}^{\nu l}(\alpha\beta\gamma) = \mathfrak{D}^{(l)}(\{\alpha\beta\gamma\})_{\mu\nu}^*, (\nu = -l, -l+1, \dots, l-1, l)$ (11a)
geschrieben werden können. Es wird sich noch später, bei der Betrachtung der Spiegelungssymmetrie, zeigen, daß der Eigenwert $E_{\nu,l} = E_{-v,l}$ ist, so daß im ganzen $l+1$ voneinander verschiedene Eigenwerte dieselbe Azimutalquantenzahl haben.

Sind alle drei Trägheitsmomente gleich, so ist die Normallage überhaupt unbestimmt, und die Eigenfunktionen (11a) bleiben Eigenfunktionen, wenn man $\{\alpha\beta\gamma\}$ durch $\{\alpha\beta\gamma\} R$ ersetzt, wo R eine beliebige Drehung sein kann. Zum Eigenwert von $\mathfrak{D}^{(l)}(\{\alpha\beta\gamma\})_{\mu\nu}^*$ gehört also auch

$$\mathfrak{D}^{(l)}(\{\alpha\beta\gamma\} R)_{\mu\nu}^* = \sum_x \mathfrak{D}^{(l)}(\{\alpha\beta\gamma\})_{\mu x}^* \mathfrak{D}^{(l)}(R)_{x\nu}^*$$

und auch

$$\int \sum_x \mathfrak{D}^{(l)}(\{\alpha\beta\gamma\})_{\mu x}^* \mathfrak{D}^{(l)}(R)_x^* \mathfrak{D}^{(l)}(R)_{x\nu} dR = \text{Konst.} \mathfrak{D}^{(l)}(\{\alpha\beta\gamma\})_{\mu\nu}^*. \quad (12)$$

Daher fallen in diesem Falle alle Eigenwerte $E_{-l,l}, E_{-l+1,l}, \dots, E_{l-1,l}, E_{l,l}$ zusammen, und zu jeder Azimutalquantenzahl gehört nur ein Eigenwert, der aber $(2l+1)^3$ -fach entartet ist.

Sind also wenigstens zwei Trägheitsmomente des Kreisels gleich, so kann man die Eigenfunktionen ohne weiteres explizite angeben (11a). Die zugehörigen Eigenwerte kann man berechnen,

indem man je eine Eigenfunktion jedes Eigenwerts (etwa $\mathbf{D}^{(l)}((\alpha\beta\gamma))_{\mu\nu}^*$) in die Schrödinger-Gleichung einsetzt, für $\alpha\beta\gamma$ Werte (etwa $\alpha = \beta = \gamma = 0$) einführt, für die $\psi_u^{\nu l}(\alpha\beta\gamma)$ nicht verschwindet und durch $\psi_{\mu}^{\nu l}(\alpha\beta\gamma)$ dividiert.

Die Schrödinger-Gleichung des symmetrischen Kreisels kann man mit Hilfe der Jacobischen Polynome auch direkt auflösen¹⁾. Die Beziehung zwischen den Darstellungskoeffizienten und den Jacobischen Polynomen lautet für $\mu \geqq \nu$

$$\mathbf{d}^{(l)}(\beta)_{\mu\nu} = \sqrt{\frac{(l-\nu)!}{(l+\mu)!}} \frac{\cos^{2l+\nu-\mu}\frac{1}{2}\beta \sin^{\mu-\nu}\frac{1}{2}\beta}{(\mu-\nu)!} \cdot F(\mu-l, -\nu-l, \mu-\nu+1, -\operatorname{tg}^2\frac{1}{2}\beta). \quad (13)$$

4. Bevor wir an die Betrachtung der Spiegelungssymmetrie herangehen, sei hier noch eine Formel

$$\mathbf{d}^{(l)}(\pi - \beta)_{\mu\nu} = (-1)^{l-\mu} \mathbf{d}^{(l)}(\beta)_{\mu,-\nu} \quad (14)$$

abgeleitet. In Kap. XV haben wir die Darstellungskoeffizienten vollkommen bestimmt, aus (27) können wir

$$\mathbf{d}^{(l)}(\beta)_{\mu\nu} = \sum_x (-1)^x \frac{\sqrt{(l+\mu)! (l-\mu)! (l+\nu)! (l-\nu)!}}{(l-\mu-x)! (l+\nu-x)! x! (x+\mu-\nu)!} \cdot \cos^{2l-\mu+\nu-2x}\frac{1}{2}\beta \sin^{2x+\mu-\nu}\frac{1}{2}\beta \quad (15)$$

entnehmen. Setzen wir hierin $\pi - \beta$ für β , so geht wegen $\cos(\frac{1}{2}\pi - x) = \sin x$ der cos in (15) in sin und der sin in cos über. Führen wir gleichzeitig an Stelle von x als Summationsindex $x' = l - \mu - x$ ein, so geht (15) in

$$\mathbf{d}^{(l)}(\pi - \beta)_{\mu\nu} = \sum_{x'} (-1)^{l-\mu-x'} \frac{\sqrt{(l+\mu)! (l-\mu)! (l+\nu)! (l-\nu)!}}{x'! (\mu+\nu+x')! (l-\mu-x')! (l-\nu-x')!} \cdot \sin^{2x'+\mu+\nu}\frac{1}{2}\beta \cos^{2l-\mu-\nu-2x'}\frac{1}{2}\beta \quad (16)$$

¹⁾ Die Quantenmechanik des symmetrischen Kreisels wurde von H. Rademacher und F. Reiche, Zeitschr. f. Phys. **39**, 444, 1926; **41**, 453, 1927; von R. de L. Kronig und J. J. Rabi, Phys. Rev. **29**, 262, 1927; C. Maneback, Zeitschr. f. Phys. **28**, 76, 1927 und J. H. van Vleck, Phys. Rev. **33**, 476, 1929 behandelt. Die Quantenmechanik des asymmetrischen Kreisels behandeln E. E. Witmer, Proc. Nat. Acad. **18**, 60, 1932; S. C. Wang, Phys. Rev. **84**, 243, 1929; H. A. Kramers und G. P. Ittmann, Zeitschr. f. Phys. **58**, 553, 1929; **58**, 217, 1929; **60**, 663, 1930; O. Klein, ebenda **58**, 730, 1929; H. Casimir, ebenda **59**, 623, 1930.

über. Was aber jetzt auf der rechten Seite von (16) steht, ist nichts anderes als $(-1)^{l-\mu} \mathbf{d}^{(l)}(\beta)_\mu, -\nu$, womit (14) bewiesen ist.

Bei einem Einkörperproblem ist der Spiegelungscharakter der Wellenfunktion durch ihre Winkelabhängigkeit gegeben. Die Inversion \mathbf{P}_I im Ursprung bedeutet ja bei einem ein-Bein nur das Ersetzen von φ durch $\varphi \pm \pi$ und θ durch $\pi - \theta$, während r , die Länge des Beins, ungeändert bleibt (vgl. Abb. 10). Hierdurch geht (9 b) in

$$\begin{aligned} \mathbf{P}_I \psi_\mu^{Nl}(r \vartheta \varphi) &= e^{-i\mu(\varphi \pm \pi)} \mathbf{d}^{(l)}(\pi - \beta)_{\mu 0} G^{Nl}(r) \\ &= (-1)^\mu e^{-i\mu\varphi} (-1)^{l-\mu} \mathbf{d}^{(l)}(\beta)_{\mu 0} G^{Nl}(r) = (-1)^l \psi_\mu^{Nl}(r \vartheta \varphi) \quad (17) \end{aligned}$$

über: die Terme mit geradem l haben positiven, die mit ungeradem l negativen Spiegelungscharakter, es existieren bei dem Einelektronenproblem nur ungestrichene Terme¹⁾.

Bei zwei unabhängigen Teilchen, bei dem He-Atom gilt nach (6)

$$\psi_\mu = \sum_l \mathfrak{D}^{(L)}(\{\alpha \beta \gamma\})_{\mu l}^* G_l(r_1, r_2, \varepsilon), \quad (18)$$

da die gegenseitige Konfiguration durch die Gestalt eines zwei-Beines gegeben ist, die durch die Längen r_1, r_2 der beiden Beine und den eingeschlossenen Winkel ε charakterisiert werden kann. Nun ändert sich durch eine Spiegelung die Gestalt eines zwei-Beines nicht, nur seine Lage wird verändert. Wenn man im Nullpunkt des Koordinatensystems spiegelt, geht α, β und γ in $\alpha \pm \pi, \pi - \beta$ und $\pi - \gamma$ über [siehe Abb. 10²⁾].

Es ist daher

$$\begin{aligned} \mathbf{P}_I \psi_\mu &= \sum_l \mathfrak{D}^{(L)}(\{\alpha \pm \pi, \pi - \beta, \pi - \gamma\})_{\mu l}^* G_l \\ &= \sum_l (-1)^{L+l} \mathfrak{D}^{(L)}(\{\alpha \beta \gamma\})_{\mu, -l}^* G_l, \quad (18a) \end{aligned}$$

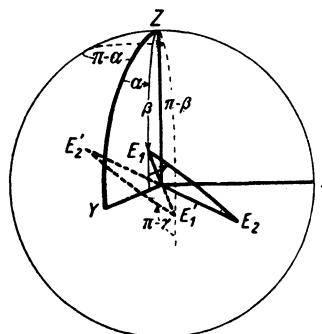


Abb. 10

¹⁾ Daher ist auch ein optischer Übergang mit $\Delta l = 0$ bei Ein-Elektronensystemen verboten: er würde ohne Änderung des Spiegelungscharakters vor sich gehen müssen.

²⁾ In Abb. 10 wurde der Einfachheit halber $r_1 = r_2 = 1$ angenommen, E_1 und E_2 sind die Stellen der beiden Elektronen vor, E'_1 und E'_2 nach der Inversion.

weil nach (14)

$$\begin{aligned} & \mathfrak{D}^{(L)}(\{\alpha \pm \pi, \pi - \beta, \pi - \gamma\})_{\mu \lambda} \\ &= e^{i\mu(\alpha \pm \pi)} (-1)^{L-\mu} \mathfrak{d}^{(L)}(\beta)_{\mu, -\lambda} e^{i\lambda(\pi - \gamma)} \\ &= (-1)^{L+\lambda} \mathfrak{D}^{(L)}(\{\alpha \beta \gamma\})_{\mu, -\lambda} \end{aligned} \quad (14a)$$

ist. Bei positiven Termen ist also wegen $\mathbf{P}_I \psi_\mu = \psi_\mu$

$$G_{-\lambda}(r_1, r_2, \varepsilon) = (-1)^{L+\lambda} G_\lambda(r_1, r_2, \varepsilon), \quad (19)$$

während bei negativen Termen $\mathbf{P}_I \psi_\mu = -\psi_\mu$

$$G_{-\lambda}(r_1, r_2, \varepsilon) = -(-1)^{L+\lambda} G_\lambda(r_1, r_2, \varepsilon) \quad (19a)$$

gilt¹⁾.

Die mittlere der Funktionen G , das G_0 , ist nur bei ungestrichenen Termen (S_+ , P_- , D_+ usw.) von Null verschieden. Hieraus folgt auch, daß He noch keine S_- -Terme hat. In der Tat hat bei S -Termen die Wellenfunktion für alle Lagen des zwei-Beines denselben Wert, die S -Terme haben notwendigerweise positiven Spiegelungscharakter.

Setzt man (6) in die Schrödingersche Differentialgleichung ein, so erhält man im allgemeinen — durch Vergleich der Koeffizienten gleicher Funktionen von $\alpha \beta \gamma$ — genau $2L + 1$ Gleichungen für die $2L + 1$ Funktionen $G_{-L}, G_{-L+1}, \dots, G_{L-1}, G_L$ der Variablen, die die Gestalt des n -Beines beschreiben. Bei dem He kann man mit Hilfe von (19), (19a) die Anzahl der unabhängigen Funktionen wesentlich herabsetzen und hat dann nicht nur bei S_+ , sondern auch bei P_+ -Termen eine, bei P_- , D_- -Termen zwei, bei D_+ , F_+ -Termen drei usw. unbekannte Funktionen²⁾.

Bei mehreren Elektronen ist es nicht möglich, die Inversion des n -Beines durch eine reine Drehung zu ersetzen, da das n -Bein bei der Inversion seine Gestalt ändert, in das „optisch isomere“ n -Bein übergeht, das sich etwa so zum ursprünglichen verhält, wie die rechte Hand zur linken. Bei dem zwei-Bein tritt die Erscheinung der optischen Isomerie noch nicht auf.

¹⁾ Auch bei dem asymmetrischen Kreisel gilt für $l+1$ Eigenwerte mit der Azimutalquantenzahl l , daß $G_{-\lambda} = G_\lambda$, und für l Eigenwerte, daß $G_{-\lambda} = -G_\lambda$ ist. Hierdurch wird die Säkulargleichung $2l+1$ -ten Grades in zwei Gleichungen vom Grade $l+1$ und l zerlegt. Für den symmetrischen Kreisel folgt, daß $\psi_\mu^{-\nu l}$ und $\psi_\mu^{\nu l}$ in (11a) nach (14a) durch Inversion ineinander übergehen, daß sie zum selben Eigenwert gehören.

²⁾ Vgl. G. Breit, Phys. Rev. 35, 569, 1930.

Bezeichnen wir die Koordinaten, die die zu g isomere Gestalt des n -Beines beschreiben, mit \bar{g} , so ist nach (6) und Abb. 10

$$\begin{aligned}\mathbf{P}_I \psi_\mu &= \sum_{\lambda} \mathfrak{D}^{(L)}(\{\alpha \pm \pi, \pi - \beta, \pi - \gamma\})_{\mu, \lambda}^* G_\lambda(\bar{g}) \\ &= \sum_{\lambda} (-1)^{L+\lambda} \mathfrak{D}^{(L)}(\{\alpha \beta \gamma\})_{\mu, -\lambda}^* G_\lambda(\bar{g}).\end{aligned}$$

Andererseits ist

$$\mathbf{P}_I \psi_\mu = w \psi_\mu = \sum_{\lambda} w \mathfrak{D}^{(L)}(\{\alpha \beta \gamma\})_{\mu, \lambda}^* G_\lambda(g).$$

wo für positive Terme $w = +1$, für negative $w = -1$ ist.
Es folgt

$$(-1)^{L+\lambda} G_\lambda(\bar{g}) = w G_{-\lambda}(g). \quad (20)$$

Die Gleichung (20) kann zur expliziten Berechnung der Eigenfunktionen nicht mehr benutzt werden, sie zeigt aber, wie man mit Hilfe der Symmetrieeigenschaften der Eigenfunktionen der Spiegelungsgruppe gegenüber noch über (6) hinausgehende Folgerungen für die Gestalt der Eigenfunktionen ziehen kann. Es war von vornherein klar, daß wir hierdurch nicht mehr allzuviel gewinnen konnten: durch die Spiegelungsgruppe kann man die Wellenfunktion nur an zwei Stellen vergleichen, während die Drehgruppe einen Vergleich einer dreiparametrisch unendlichen Schar von Stellen ermöglicht hat. Demgemäß konnten durch die Drehgruppe drei Variable $(\alpha \beta \gamma)$ eliminiert werden, und es mußte dagegen nur eine Erhöhung der Anzahl der unbekannten Funktionen (G_{-L}, \dots, G_L) in Kauf genommen werden. Durch die Spiegelungsgruppe konnte nur diese Zahl $2L+1$ wieder etwas reduziert werden, was eine verhältnismäßig unwesentliche Vereinfachung ist.

Es liegt nahe, zur weiteren Reduzierung der Anzahl der unbekannten Funktionen die Symmetrie in bezug auf Vertauschung der Elektronen auszunutzen. Dies ist in der Tat bis zu einem gewissen Grade möglich. Die hierzu erforderlichen Überlegungen sind aber, da in (6) das erste und auch das zweite Elektron vor den übrigen ausgezeichnet sind (α und β sind Azimut und Polabstand des ersten Elektrons), nicht mehr ganz so einfach wie die bisherigen, und es werde daher auf ihre Ausführung verzichtet.

XX. Das Drehelektron

Die physikalischen Grundlagen der Paulischen Theorie

1. In den vorangehenden Kapiteln wurden die wichtigsten Eigenschaften der Atomspektren, die ohne Einführung des Drehelektrons behandelt werden können, besprochen. Viele feinere Züge — in erster Reihe die Feinstruktur — konnten jedoch noch nicht erfaßt werden, weil sie mit einer neuen Eigenschaft der Elektronen, mit ihrem magnetischen Moment aufs engste verknüpft sind.

Die Hypothese, daß die Elektronen ein magnetisches Moment und ein Drehmoment — kurz einen „Spin“ — haben, stammt von Goudsmit und Uhlenbeck. Sie haben — noch vor der Entdeckung der Quantenmechanik — bemerkt, daß zur einfachen und möglichst vollständigen Beschreibung der Spektren die Elektronen nicht als bloße negative Punktladungen angesehen werden dürfen, sondern daß ihnen noch ein magnetisches und mechanisches Moment zugeschrieben werden muß. Man kann sich bekanntlich in der Elektrodynamik eine Magnetnadel durch eine um die Achse des magnetischen Moments rotierende Punktladung ersetzt denken. Der Vektor des magnetischen Moments m berechnet sich dann aus dem Drehimpulsvektor \mathbf{d} nach

$$m = \frac{e \mathbf{d}}{2mc}, \quad (\dagger)$$

wo e die Ladung des rotierenden Teilchens und m seine Masse ist. Für das durch den Spin bedingte magnetische Moment der Elektronen soll aber nach Goudsmit und Uhlenbeck (\dagger) nicht mit der normalen Ladung und Masse des Elektrons gelten, es soll vielmehr angenommen werden, daß der Drehimpuls nur halb so groß ist, wie er sich aus (\dagger) ergeben würde: das magnetische Moment soll ein ganzes Bohrsches Magneton

$$|m| = \frac{e \hbar}{4\pi mc}, \quad (1)$$

der Drehimpuls, das mechanische Moment aber nur vom Betrage

$$|\mathbf{d}| = \frac{1}{2} \frac{\hbar}{2\pi} \quad (1a)$$

sein.

Die Quantenmechanik des rotierenden Elektrons zeigt, daß diese Begriffe nicht ganz wörtlich genommen werden dürfen. Schon

die ältere Paulische Theorie verlangt, daß es keinen Versuch geben soll, der die Richtung (also etwa die Richtungscosinus) des mechanischen oder magnetischen Moments zu bestimmen gestatten würde. Möglich ist nur, zwischen einer Richtung und der ihr entgegengesetzten zu entscheiden. Die Frage nach den Wahrscheinlichkeiten für die verschiedenen räumlichen Richtungen des Spins soll demnach nicht sinnvoll, d. h. durch Versuche beantwortbar sein, und es soll nur die Komponente des Spins in irgendeiner Richtung gemessen werden können. Bei dieser Messung — man denke an einen Stern-Gerlachschen Versuch — stellt sich der Spin, wenn man etwa die Z -Komponente mißt, entweder in die $+Z$ - oder in die $-Z$ -Richtung ein und die möglichen Versuchsresultate für die Richtungen sind $+Z$ und $-Z$, für die Komponente des Drehimpulses $+h/4\pi$ und $-h/4\pi$. Hat man das erste Resultat erhalten, so gibt eine zweite, unmittelbar darauf ausgeführte Messung der Z -Komponente wieder sicher $+Z$ und sicher nicht $-Z$. Eine Messung der Y -Komponente gibt dagegen mit gleichen Wahrscheinlichkeiten die beiden möglichen Ergebnisse $+Y$ und $-Y$. In diesem Sinne ist es auch zu verstehen, daß man nicht nach der Wahrscheinlichkeit aller Richtungen gleichzeitig fragen kann: auch wenn der Spin sicher die Z -Richtung hat (die Z -Komponente des Drehimpulses sicher $+h/4\pi$ ist), ist die Wahrscheinlichkeit für die $+Y$ -Richtung noch $1/2$ und auch für alle anderen Richtungen von Null verschieden.

Einen noch mehr symbolischen Charakter erhält der Spin in der Diracschen relativistischen Theorie des Elektrons und durch die an diese Theorie geknüpften Bemerkungen von N. Bohr. Nach dieser Theorie (auf die wir nicht näher eingehen werden) ist die Existenz des magnetischen Moments lediglich ein relativistischer Effekt, der in den Gleichungen bei der gleichmäßigen Behandlung von Raum und Zeit von selber auftritt.

2. In der Paulischen Theorie berücksichtigt man das magnetische Moment durch eine neue Koordinate s , die man außer den Cartesischen Koordinaten x, y, z in die Wellenfunktion $\Phi(x, y, z, s)$ einführt. Während sich der Variabilitätsbereich von x, y, z von $-\infty$ bis ∞ erstreckt, kann s nur die beiden Werte -1 und $+1$ annehmen. Im Falle eines Elektrons besteht daher die Wellenfunktion $\Phi(x, y, z, s)$ eigentlich aus zwei Funktionen von x, y, z : aus $\Phi(x, y, z, -1)$ und $\Phi(x, y, z, 1)$. Daß im Gegensatz zu x, y, z

die Variable s nur zwei Werte annehmen kann, beruht darauf, daß die Komponente des Spins etwa in der Z -Richtung nur zwei Werte ($-1 \cdot h/4\pi$ und $+1 \cdot h/4\pi$) haben kann, während etwa die X -Koordinate des Elektrons alle Werte von $-\infty$ bis ∞ annehmen kann.

Wir müssen vor allem das skalare Produkt zweier Funktionen von xyz und s definieren. Das skalare Produkt zweier Funktionen von $\varphi(xyz)$ und $g(xyz)$ entstand (S. 38) durch Grenzübergang aus der Summe

$$\sum_{xyz} \varphi(xyz)^* g(xyz),$$

wo die Summation über den ganzen Variabilitätsbereich, von $-\infty$ bis ∞ erstreckt werden mußte. Ähnlich ist das skalare Produkt von $\Phi(xyzs)$ und $G(xyzs)$:

$$\sum_{s=\pm 1} \sum_{xyz} \Phi(xyzs)^* G(xyzs), \quad (2)$$

wo die Summation wieder über den ganzen Variabilitätsbereich aller Variablen zu erstrecken ist. Durch Grenzübergang erhält man

$$\begin{aligned} (\Phi, G) &= \sum_{s=\pm 1} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \Phi(x, y, z, s)^* G(x, y, z, s) dx dy dz \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} [\Phi(x, y, z, -1)^* G(x, y, z, -1) \\ &\quad + \Phi(x, y, z, 1)^* G(x, y, z, 1)] dx dy dz. \end{aligned} \quad (3)$$

3. Auch in der Paulischen Theorie bilden die Größen, die nur auf die Cartesischen Koordinaten der Teilchen Bezug haben, eine wichtige spezielle Klasse von physikalischen Größen. Diese Größen — wie etwa die X -Koordinate oder die Geschwindigkeit — haben auch in einer Theorie, die den Spin gar nicht kennt, einen Sinn. Versuche zur Messung dieser Größen nennen wir „spinfreie Versuche“. Diesen Größen entsprechen in der Paulischen Theorie Operatoren, die nur auf die Cartesischen Koordinaten xyz einwirken, so daß man die Spinkoordinate s als einen Parameter behandeln darf.

Es sei hier der Begriff von Operatoren, die nur auf einen Teil der Koordinaten einwirken, etwas genauer erläutert. Jeder Operator X , der auf eine Funktion $f(\xi)$ nur einer Variablen ξ angewendet werden kann (wie etwa Differentiieren nach ξ), kann auch

auf eine Funktion zweier Variablen angewendet werden. Eine Funktion $F(\xi, \sigma)$ zweier Variablen kann nämlich als eine Schar von Funktionen der Variable ξ allein aufgefaßt werden, für jeden speziellen Wert von σ ist $F(\xi, \sigma)$ eine Funktion von ξ allein¹⁾. Wendet man auf alle diese Funktionen von ξ den Operator X an, so erhält man wieder eine Schar von Funktionen von ξ , für jeden σ -Wert eine Funktion. Diese Schar bildet dann die Funktion $XF(\xi, \sigma)$. Daß X nur auf ξ einwirkt, soll bedeuten, daß man bei seiner Anwendung σ in dieser Weise als einen Parameter behandeln darf.

Es sei nun $\Psi_k(xyzs)$ eine Eigenfunktion eines Operators H , der nur auf x, y und z einwirkt. Ist λ_k der zugehörige Eigenwert, so muß die Funktion von x, y, z und s

$$H\Psi_k(x, y, z, s) - \lambda_k \Psi_k(x, y, z, s) = 0 \quad (4)$$

verschwinden, d. h. beide Funktionen der vorher erwähnten Schar (für $s = -1$ und $s = +1$)

$$H\Psi_k(x, y, z, -1) - \lambda_k \Psi_k(x, y, z, -1) = 0,$$

$$H\Psi_k(x, y, z, 1) - \lambda_k \Psi_k(x, y, z, 1) = 0$$

müssen verschwinden. Gehört zu λ_k nur eine Lösung $\psi_k(xyz)$ der Gleichung

$$H\psi_k(xyz) = \lambda_k \psi_k(xyz),$$

so muß sowohl $\Psi_k(x, y, z, -1)$ wie auch $\Psi_k(x, y, z, 1)$ ein konstantes Vielfaches von $\psi_k(xyz)$ sein²⁾:

$$\Psi_k(x, y, z, -1) = u_{-1} \psi_k(xyz); \quad \Psi_k(x, y, z, 1) = u_1 \psi_k(xyz);$$

$$\Psi_k(x, y, z, s) = u_s \psi_k(xyz). \quad (5)$$

Man kann u_{-1} und u_1 in (5) noch ganz beliebig wählen, $u_s \psi_k(xyz)$ bleibt immer noch eine Eigenfunktion von H zum Eigenwert λ_k . Dies zeigt, daß die Einführung der Spinkoordinate s den Eigenwert λ_k zu einem doppelten gemacht hat, es gehören die beiden linear unabhängigen, zueinander orthogonalen Eigenfunktionen

$$\Psi_{k-} = \delta_{s, -1} \psi_k(xyz), \quad (5a)$$

$$\Psi_{k+} = \delta_{s, 1} \psi_k(xyz) \quad (5b)$$

¹⁾ Man denke bei ξ an das Koordinatentripel xyz , bei σ an die Spinkoordinate s .

²⁾ Es ist zunächst etwas störend, daß in (5) s links als Variable, rechts als Index auftritt, es zeigt aber nur erneut, daß man eine Funktion von xyz und s als eine Zuordnung je einer Funktion von xyz zu jedem s -Wert auffassen kann.

dazu. Das skalare Produkt von Ψ_{k-} und Ψ_{k+} verschwindet ja, weil in beiden Gliedern des Integranden von (3) der eine Faktor Null ist.

Die Eigenfunktionen

$$\delta_{s,-1} \psi_1; \delta_{s,+1} \psi_1; \delta_{s,-1} \psi_2; \delta_{s,+1} \psi_2; \delta_{s,-1} \psi_3; \dots$$

entsprechen den möglichen Resultaten der beiden gleichzeitig ausgeführten Messungen: a) der Messung der Größe, der der Operator H zugeordnet ist, b) der Messung der Z -Komponente des Spins. Für Ψ_{k-} ist der Wert der ersten Größe sicher λ_k , der Spin hat sicher die $-Z$ -Richtung, für Ψ_{k+} ist der Wert der ersten Größe auch noch λ_k , der Spin hat aber die $+Z$ -Richtung, d. h. die Wahrscheinlichkeit für die Einstellung in die $-Z$ -Richtung ist Null. Ist die Wellenfunktion schließlich allgemein

$$\begin{aligned}\Phi = & a_1 \Psi_{1-} + a_2 \Psi_{2-} + a_3 \Psi_{3-} + \dots \\ & + b_1 \Psi_{1+} + b_2 \Psi_{2+} + b_3 \Psi_{3+} + \dots\end{aligned}\quad (6)$$

so ist die Wahrscheinlichkeit dafür, daß H den Wert λ_k , der Spin gleichzeitig die $-Z$ -Richtung habe, gleich $|a_k|^2$; die Wahrscheinlichkeit des Wertes λ_k für H und einer Spinrichtung $+Z$ ist $|b_k|^2$.

Die Invarianz der Beschreibung räumlichen Drehungen gegenüber

4. Die soeben gegebene Beschreibung des Elektrons mit einer von xyz und s abhängigen Wellenfunktion zeichnet nicht nur die Koordinatenachsen $X Y Z$ anderen, gleichberechtigten Richtungen gegenüber aus, sondern bevorzugt auch noch die Z -Achse den beiden anderen Achsen gegenüber. Es ist hier also sehr wohl am Platze, zu untersuchen, wie sich die Isotropie des Raumes in dieser Beschreibungsweise äußert, d. h. wie die Wellenfunktion $O_R \Phi$ des Zustandes Φ für einen Beobachter lautet, der die Systeme und Größen genau so wie der erste Beobachter beschreibt, nur daß er seiner Beschreibung ein gegenüber dem Achsenkreuz des ersten Beobachters verdrehtes Achsenkreuz zugrunde legt. Die Lage der beiden Achsenkreuze sei dabei so, daß die Koordinaten des Punktes x, y, z im zweiten Achsenkreuz

$$R_{xx}x + R_{xy}y + R_{xz}z = x',$$

$$R_{yx}x + R_{yy}y + R_{yz}z = y',$$

$$R_{zx}x + R_{zy}y + R_{zz}z = z'$$

sind. (R ist eine reelle orthogonale dreidimensionale Matrix mit der Determinante 1.) Man kann $\mathbf{O}_R \Phi$ entweder als die Wellenfunktion des Zustandes Φ für den zweiten Beobachter oder als die Wellenfunktion des um R verdrehten Zustandes für den ursprünglichen Beobachter definieren.

Solange die Wellenfunktionen nur von den Cartesischen Koordinaten der Teilchen abhingen, war die Operation \mathbf{O}_R lediglich eine Punkttransformation \mathbf{P}_R (vgl. Kap. XI):

$$\mathbf{P}_R \varphi(x' y' z') = \varphi(x y z). \quad (7)$$

In (7) steht, daß die Wellenfunktion des zweiten Beobachters $\mathbf{P}_R \varphi$ im Punkte x', y', z' den gleichen Wert annimmt, den die Wellenfunktion φ des ersten Beobachters an der Stelle x, y, z hat. Dies muß auch so sein, weil der Punkt x, y, z für den zweiten Beobachter x', y', z' heißt.

Jetzt, wo wir außer den Cartesischen auch eine Spinkoordinate haben, können wir mit einer einfachen Punkttransformation nicht auskommen, es hat ja gar keinen Sinn, φ transformieren zu wollen. Deshalb wird \mathbf{O}_R ein allgemeinerer Operator sein, als es \mathbf{P}_R war. Wir werden die Existenz des Operatorensystems \mathbf{O}_R (zu jeder Drehung R gehört ein Operator \mathbf{O}_R) annehmen und es dann mit Hilfe der besprochenen Grundannahmen der Paulischen Theorie und der Forderung, daß die mit verschiedenen Achsenkreuzen arbeitenden Beobachter prinzipiell gleichberechtigt sind, zu bestimmen versuchen. Es wird sich zeigen, daß im wesentlichen nur ein Operatorensystem existiert, das diesen Bedingungen genügt. Seine Bestimmung wird uns über wichtige Eigenschaften des Paulischen Drehelektrons Aufschluß geben.

5. Die Beschreibungsweise des zweiten Beobachters, die die Wellenfunktion $\mathbf{O}_R \Phi$ dem Zustand Φ zuschreibt, muß der ursprünglichen Beschreibungsweise völlig äquivalent sein. Insbesondere muß sie dieselben Übergangswahrscheinlichkeiten zwischen zwei beliebigen Zuständen Ψ, Φ ergeben, wie die erste:

$$|(\Psi, \Phi)|^2 = |(\mathbf{O}_R \Psi, \mathbf{O}_R \Phi)|^2. \quad (8)$$

Es ist dabei wichtig, zu bemerken, daß, obwohl der Zustand Φ , der für den Beobachter mit dem verdrehten Achsenkreuz als der Zustand $\mathbf{O}_R \Phi$ erscheint, durch Angabe seiner Wellenfunktion Φ vollkommen gegeben ist, in $\mathbf{O}_R \Phi$ sicher noch ein konstanter

Faktor c_ϕ vom Absolutwert 1 willkürlich wählbar ist, weil die Wellenfunktionen $\mathbf{O}_R \Phi$ und $c_\phi \mathbf{O}_R \Phi$ denselben physikalischen Zustand charakterisieren. Das bedeutet, daß der Operator \mathbf{O}_R sehr vieldeutig ist, für jede Funktion Φ ist in \mathbf{O}_R noch ein Faktor frei. Es zeigt sich nun (siehe den Anhang am Ende des Kapitels), daß man über diese Freiheit in \mathbf{O}_R so verfügen kann, daß für alle Ψ und alle Φ

$$\begin{aligned} (\Psi, \Phi) &= (\mathbf{O}_R \Psi, \mathbf{O}_R \Phi), \\ \mathbf{O}_R (a\Psi + b\Phi) &= a\mathbf{O}_R \Psi + b\mathbf{O}_R \Phi \end{aligned} \quad | \quad (8a)$$

gilt (a und b sind Konstanten), d. h. daß \mathbf{O}_R ein linearer unitärer Operator wird. Dann unterscheiden sich die beiden Beschreibungsweisen: die für den Beobachter mit dem ursprünglichen, und die andere für den Beobachter mit dem verdrehten Achsenkreuz nur durch eine kanonische Transformation, sind also physikalisch völlig äquivalent. Den Zustand mit der Wellenfunktion Φ nennt der zweite Beobachter $\mathbf{O}_R \Phi$, und der Größe, der der erste Beobachter den Operator \mathbf{H} zuordnet, ordnet er $\mathbf{O}_R \mathbf{H} \mathbf{O}_R^{-1}$ zu.

Umgekehrt sind durch die Forderung (8a), daß \mathbf{O}_R ein linear-unitärer Operator sei, die Konstanten c_ϕ für alle Wellenfunktionen — bis auf eine einzige — eindeutig festgelegt. Ersetzt man eine Wellenfunktion $\mathbf{O}_R \Phi$ durch $c \mathbf{O}_R \Phi$, so muß man gleichzeitig, wenn man an (8a) festhalten will, alle Wellenfunktionen durch ihre c -fachen ersetzen. Ersetzen wir nämlich in (8a) $\mathbf{O}_R \Phi$ durch $c \mathbf{O}_R \Phi$, während wir etwa $\mathbf{O}_R \Psi$ ungeändert lassen, so folgt, wenn (8a) auch für dieses neue System gelten soll:

$$(\Psi, \Phi) = (\mathbf{O}_R \Psi, c \mathbf{O}_R \Phi) = c (\mathbf{O}_R \Psi, \mathbf{O}_R \Phi),$$

was mit (8a) zusammen $c = 1$ ergibt. Im folgenden wollen wir die $\mathbf{O}_R \Phi$ immer so annehmen, daß für sie (8a) gelte, dann ist in allen $\mathbf{O}_R \Phi$ für eine gegebene Drehung R nur eine einzige Konstante frei.

6. Betrachten wir nun die beiden Zustände $\Psi_- = \psi(xyz) \delta_{s,-1}$ und $\Psi_+ = \psi(xyz) \delta_{s,+1}$. Für spinfreie Versuche benehmen sich diese beiden Zustände so, als ob ihre Wellenfunktion $\psi(xyz)$ wäre. Für den Beobachter mit dem verdrehten Achsenkreuz erscheinen sie daher spinfreien Versuchen gegenüber wie ein Zustand mit der Wellenfunktion $\mathbf{P}_R \psi(xyz)$. Die Wellenfunktionen $\mathbf{O}_R \Psi_-$ und $\mathbf{O}_R \Psi_+$ müssen daher nach (5) die Form

$$\begin{aligned} \mathbf{O}_R \delta_{s,-1} \psi(xyz) &= u_{s,-1} \mathbf{P}_R \psi(xyz), \\ \mathbf{O}_R \delta_{s,+1} \psi(xyz) &= u_{s,+1} \mathbf{P}_R \psi(xyz) \end{aligned} \quad | \quad (9)$$

haben, wo $u_{s,-1}$ und $u_{s,1}$ zwar von x, y, z unabhängig sind, für verschiedene ψ aber zunächst noch verschieden sein könnten. Aus der Linearität des \mathbf{O}_R folgt aber, wenn φ ein von ψ verschiedener Zustand ist, für den

$$\mathbf{O}_R \delta_{s,-1} \varphi (xyz) = u_{s,-1} \mathbf{P}_R \varphi (xyz)$$

gilt, daß

$$\begin{aligned} \mathbf{O}_R \delta_{s,-1} (\varphi + \psi) &= \bar{u}_{s,-1} \mathbf{P}_R (\varphi + \psi) = \bar{u}_{s,-1} \mathbf{P}_R \varphi + \bar{u}_{s,-1} \mathbf{P}_R \psi \\ &= \mathbf{O}_R \delta_{s,-1} \varphi + \mathbf{O}_R \delta_{s,-1} \psi = u_{s,-1} \mathbf{P}_R \varphi + u_{s,-1} \mathbf{P}_R \psi \end{aligned}$$

ist. Hieraus folgt aber wegen der linearen Unabhängigkeit von $\mathbf{P}_R \varphi$ und $\mathbf{P}_R \psi$

$$u_{s,-1} = \bar{u}_{s,-1} = u_{s,-1} \text{ und ähnlich } u_{s,1} = \bar{u}_{s,1},$$

daß die u_{st} für alle Wellenfunktionen dieselben sind; die Matrix $\mathbf{u} = \mathbf{u}(R)$ kann nur von der Drehung R abhängen. Ist nun $\Phi(xyzs)$ eine beliebige Wellenfunktion

$\Phi(x, y, z, s) = \delta_{s,-1} \Phi(x, y, z, -1) + \delta_{s,1} \Phi(x, y, z, 1), \quad (10)$
so folgt wiederum aus der Linearität von \mathbf{O}_R und (9):

$$\begin{aligned} \mathbf{O}_R \Phi(x, y, z, s) &= \mathbf{O}_R \delta_{s,-1} \Phi(x, y, z, -1) + \mathbf{O}_R \delta_{s,1} \Phi(x, y, z, 1) \\ &= u_{s,-1} \mathbf{P}_R \Phi(x, y, z, -1) + u_{s,1} \mathbf{P}_R \Phi(x, y, z, 1) \\ \mathbf{O}_R \Phi(x, y, z, s) &= \sum_{t=\pm 1} u_{st} \mathbf{P}_R \Phi(x, y, z, t). \quad (11) \end{aligned}$$

Der Operator \mathbf{O}_R läßt sich also in zwei Faktoren

$$\mathbf{O}_R = \mathbf{Q}_R \mathbf{P}_R \quad (12)$$

spalten, wo \mathbf{P}_R die bekannte, auch in (7) definierte Transformation ist, die nur auf die Lagenkoordinaten der Wellenfunktion einwirkt, und \mathbf{Q}_R durch

$$\mathbf{Q}_R \Phi(xyzs) = \sum_{t=\pm 1} u(R)_{st} \Phi(xyzt) \quad (12a)$$

definiert ist, und nur auf die Spinkoordinate s einwirkt. Da der Variabilitätsbereich von s nur aus zwei Zahlen -1 und $+1$ besteht, ist \mathbf{Q}_R nach (12a) einer zweidimensionalen Matrix

$$\mathbf{u}(R) = \begin{pmatrix} \mathbf{u}(R)_{-1,-1} & \mathbf{u}(R)_{-1,1} \\ \mathbf{u}(R)_{1,-1} & \mathbf{u}(R)_{1,1} \end{pmatrix} \quad (13)$$

äquivalent. Es gilt für zwei beliebige Drehungen R und S

$$\mathbf{P}_S \mathbf{Q}_R = \mathbf{Q}_R \mathbf{P}_S \text{ insbesondere } \mathbf{P}_R \mathbf{Q}_R = \mathbf{Q}_R \mathbf{P}_R. \quad (14)$$

Die Möglichkeit der Zerlegung von \mathbf{O}_R in die beiden Teile \mathbf{P}_R und \mathbf{Q}_R beruht wesentlich auf der Annahme, daß es „spinfreie Versuche“ gibt, die man schon mit einer nur von xyz abhängigen Wellenfunktion beschreiben kann.

Zusammenhang mit der Darstellungstheorie

7. Aus der Unitarität von \mathbf{O}_R folgt, da auch \mathbf{P}_R , also auch \mathbf{P}_R^{-1} unitär ist, daß auch $\mathbf{Q}_R = \mathbf{O}_R \mathbf{P}_R^{-1}$ unitär sein muß. Es ist daher für alle Funktionen Φ, Ψ

$$(\mathbf{Q}_R \Phi, \mathbf{Q}_R \Psi) = (\Phi, \Psi). \quad (15)$$

Es folgt hieraus, daß auch die Matrix $\mathbf{u}(R)$ unitär sein muß. Setzen wir nämlich etwa $\Phi = \delta_{s\sigma} \psi$ und $\Psi = \delta_{s\tau} \psi$, so ist nach (3), wenn ψ normiert ist, $(\Phi, \Psi) = \delta_{\sigma\tau}$ und daher nach (15) und (12a) auch

$$\begin{aligned} \delta_{\sigma\tau} &= (\mathbf{Q}_R \delta_{s\sigma} \psi, \mathbf{Q}_R \delta_{s\tau} \psi) = (\mathbf{u}_{s\sigma} \psi, \mathbf{u}_{s\tau} \psi) \\ &= \sum_{s=\pm 1} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \mathbf{u}_{s\sigma}^* \psi^* \mathbf{u}_{s\tau} \psi dx dy dz = \sum_{s=\pm 1} \mathbf{u}_{s\sigma}^* \mathbf{u}_{s\tau}. \end{aligned}$$

Dies ist aber die Bedingung für die Unitarität von \mathbf{u} .

Wir können uns weiter daran erinnern, daß \mathbf{O}_R durch physikalische Tatsachen und die Gleichungen (8a) nur bis auf eine von R abhängige Konstante vom Betrag 1 festgelegt ist, und wir können daher, ohne am physikalischen Inhalt der Theorie etwas zu ändern, oder (8a) zu beeinträchtigen, jedes \mathbf{O}_R durch $c_R \mathbf{O}_R$ ersetzen, wo $|c_R| = 1$ ist. Den Faktor c_R können wir an \mathbf{Q}_R , d. h. an $\mathbf{u}(R)$ anfügen und mit seiner Hilfe erreichen, daß die Determinante von $\mathbf{u}(R)$ gleich 1 wird.

Um schließlich die Matrix $\mathbf{u}(R)$ gänzlich zu bestimmen, beachten wir noch, daß $\mathbf{O}_R \Phi$ die Wellenfunktion des um R verdrehten Zustandes Φ ist und $\mathbf{O}_S \cdot \mathbf{O}_R \Phi$ die Wellenfunktion des zuerst um R und dann um S , im Ganzen um SR verdrehten Zustandes Φ ist. Der Operator $\mathbf{O}_S \mathbf{O}_R$ ist also mit dem Operator \mathbf{O}_{SR} physikalisch jedenfalls völlig äquivalent. Da er auch die Gleichungen (8a) befriedigt — das Produkt zweier linear-unitärer Operatoren \mathbf{O}_S und \mathbf{O}_R ist auch selber linear-unitär —, kann er sich von \mathbf{O}_{SR} nur um einen konstanten Faktor unterscheiden:

$$\mathbf{O}_{SR} = c_{S,R} \mathbf{O}_S \mathbf{O}_R. \quad (16)$$

Mit Hilfe von (12) folgt nun wegen $\mathbf{P}_{SR} = \mathbf{P}_S \mathbf{P}_R$ und (14)

$$\mathbf{Q}_{SR} \mathbf{P}_{SR} = c_{S,R} \mathbf{Q}_S \mathbf{P}_S \mathbf{Q}_R \mathbf{P}_R; \quad \mathbf{Q}_{SR} = c_{S,R} \mathbf{Q}_S \mathbf{Q}_R$$

oder mit Hilfe von (12a)

$$\sum_{t=\pm 1} \mathbf{u}(SR)_{st} \Phi(xyzt) = c_{S,R} \sum_{r=\pm 1} \sum_{t=\pm 1} \mathbf{u}(S)_{sr} \mathbf{u}(R)_{rt} \Phi(xyzt),$$

$$\mathbf{u}(SR) = c_{S,R} \mathbf{1} \cdot \mathbf{u}(S) \cdot \mathbf{u}(R). \quad (17)$$

Da wir die Determinante aller \mathbf{u} auf 1 normiert haben, folgt aus (17), daß auch die Determinante $|c_{S,R} \mathbf{1}| = 1$, $c_{S,R} = \pm 1$ ist. Es folgt nunmehr

$$\mathbf{u}(SR) = \pm \mathbf{u}(S) \mathbf{u}(R), \quad (17a)$$

dass die Matrizen $\mathbf{u}(R)$ bis auf das Vorzeichen eine Darstellung der dreidimensionalen Drehgruppe bilden.

Dies legt schon den Gedanken nahe, daß $\mathbf{u}(R)$ mit der im Kapitel XV besprochenen Matrix

$$\left. \begin{aligned} \mathbf{u}(\{\alpha \beta \gamma\}) &= \mathbf{D}^{(1/2)}(\{\alpha \beta \gamma\}) \\ &= \begin{pmatrix} e^{-\frac{1}{2}i\alpha} \cos \frac{1}{2}\beta e^{-\frac{1}{2}i\gamma} & -e^{-\frac{1}{2}i\alpha} \sin \frac{1}{2}\beta e^{\frac{1}{2}i\gamma} \\ e^{\frac{1}{2}i\alpha} \sin \frac{1}{2}\beta e^{-\frac{1}{2}i\gamma} & e^{\frac{1}{2}i\alpha} \cos \frac{1}{2}\beta e^{\frac{1}{2}i\gamma} \end{pmatrix} \end{aligned} \right\} \quad (18)$$

identisch ist oder aus ihr doch mit Hilfe einer Ähnlichkeitstransformation hervorgeht. Dies ist in der Tat richtig, wir werden im nächsten Kapitel zeigen, daß jedes zweidimensionale Matrzensystem, das (17a) befriedigt, entweder aus lauter Einheitsmatrizen besteht oder aus $\mathbf{D}^{(1/2)}$ durch eine Ähnlichkeitstransformation hervorgeht. Das erstere kommt aber für uns nicht in Frage, da es z. B. bedeuten würde, daß ein Zustand, bei dem der Spin sicher in der Z -Richtung liegt, auch nach einer beliebigen Drehung noch diese Eigenschaft hat.

Nur für Drehungen mit $\beta = 0$, die die Z -Achse unverändert lassen, darf und muß \mathbf{u} eine Diagonalmatrix sein. Hat der Spin nämlich im ersten Koordinatensystem die $-Z$ -Richtung, ist also $\Phi(x, y, z, 1) = 0$, so muß das auch im zweiten Koordinatensystem der Fall sein. Hieraus folgt aber nach (11) $\mathbf{u}_{1,-1} = 0$, und ähnlich muß auch $\mathbf{u}_{-1,1} = 0$ sein: $\mathbf{u}(\{\alpha 0 0\})$ ist eine Diagonalmatrix, und da es $\mathbf{D}^{(1/2)}(\{\alpha 0 0\})$ äquivalent ist, kann es entweder

$$\mathbf{u}(\{\alpha 0 0\}) = \begin{pmatrix} e^{-\frac{1}{2}i\alpha} & 0 \\ 0 & e^{\frac{1}{2}i\alpha} \end{pmatrix} \text{ oder } \begin{pmatrix} e^{\frac{1}{2}i\alpha} & 0 \\ 0 & e^{-\frac{1}{2}i\alpha} \end{pmatrix} \quad (*)$$

sein. Der zweite Fall würde aber bedeuten, daß der Drehimpuls des Elektrons im Zustand $\delta_{s,-1}\psi$ die $+Z$ und im Zustand $\delta_{s,+1}\psi$ die $-Z$ -Richtung hat, so daß wir diesen Fall ausschließen können. Er würde zur Folge haben, daß dem physikalischen Zustand, dem wir die Wellenfunktion $\Phi(x, y, z, s)$ zuordnen wollen, die Wellenfunktion $\Phi(x, y, z, -s)$ zugeordnet wäre.

Es ist also $\mathbf{u}(\{\alpha 0 0\}) = \mathbf{D}^{(1/2)}(\{\alpha 0 0\})$, und die unitäre Matrix S , die $\mathbf{D}^{(1/2)}$ in \mathbf{u} transformiert, muß mit $\mathbf{D}^{(1/2)}(\{\alpha 0 0\})$ vertauschbar, also eine Diagonalmatrix sein. Ihre beiden Diagonalelemente seien a und a' ($|a| = |a'| = 1$). Dann können wir festlegen, daß wir den Zustand, dessen Wellenfunktion bisher $\Phi(xyzs)$ war, mit $S\Phi(xyzs)$ beschreiben werden, wo

$$S\Phi(x, y, z, -1) = a\Phi(x, y, z, -1),$$

$$S\Phi(x, y, z, 1) = a'\Phi(x, y, z, 1)$$

ist. Dies dürfen wir tun, da wir der komplexen Phase des Verhältnisses $\Phi(x, y, z, -1)/\Phi(x, y, z, 1)$ bisher keine Bedeutung zugeschrieben haben. In der Beschreibung, die so aus der ursprünglichen entstanden und mit ihr völlig äquivalent ist, ist jedenfalls

$$\mathbf{u}(\{\alpha\beta\gamma\}) = \mathbf{D}^{(1/2)}(\{\alpha\beta\gamma\}) \quad (18a)$$

und wir erhalten als Endresultat, daß jede Beschreibung des Spins, die sich auf die in 1., 2. und 3. besprochenen Ideen gründet, mit einer Beschreibung physikalisch völlig äquivalent sein muß, bei der die Wellenfunktion eines um¹⁾ R verdrehten Zustandes Φ durch $\mathbf{O}_R\Phi$ gegeben ist. Dabei ist $\mathbf{O}_R = \mathbf{P}_R\mathbf{Q}_R$ der durch

$$\begin{aligned} \mathbf{O}_R\Phi(x, y, z, s) &= \sum_{t=\pm 1} \mathbf{D}^{(1/2)}(R)_{\frac{1}{2}s, \frac{1}{2}t} \mathbf{P}_R\Phi(x, y, z, t) \\ &= \sum_{t=\pm 1} \mathbf{D}^{(1/2)}(R)_{\frac{1}{2}s, \frac{1}{2}t} \Phi(x'', y'', z'', t) \end{aligned} \quad (19)$$

definierte Operator, wo x'', y'', z'' aus x, y, z durch die Drehung R^{-1} hervorgehen. In $\mathbf{D}^{(1/2)}$ müssen die Indizes $\frac{1}{2}s, \frac{1}{2}t$ auftreten, da die Zeilen und Spalten von $\mathbf{D}^{(1/2)}$ — statt mit $-1, +1$ wie in \mathbf{u} — mit $-\frac{1}{2}$ und $+\frac{1}{2}$ bezeichnet sind.

Sei z. B.

$$\Phi(xyzs) = (x - iy)e^{-r/2r_0} = \begin{cases} (x - iy)e^{-r/2r_0} & \text{für } s = -1, \\ (x - iy)e^{-r/2r_0} & \text{für } s = +1. \end{cases} \quad (\dagger)$$

¹⁾ R ist hier immer eine reine Drehung.

$[(x - iy)e^{-r/2} r_0]$ ist die Eigenfunktion des Wasserstoffatoms mit $N = 2$, $l = 1$, $\mu = 1$, vgl. (3), Kap. XVII] und betrachten wir den Zustand (\dagger) von einem Koordinatensystem, dessen X -Achse die alte X -Achse und dessen Z -Achse die alte Y -Achse ist. Die Drehung R ist dann $\{0, \frac{1}{2}\pi, 0\}$

$$x' = x, \quad y' = -z, \quad z' = y$$

und die reziproke Drehung ist

$$x'' = x, \quad y'' = z, \quad z'' = -y.$$

Die Matrix $D^{(1/2)}(\{0, \frac{1}{2}\pi, 0\})$ ist

$$\begin{pmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix}.$$

Die Wellenfunktion des Zustandes (\dagger) vom neuen Koordinatensystem ist nach (19)

$$\begin{aligned} O_R \Phi(xyzs) &= \begin{cases} \frac{1}{\sqrt{2}}(x - iz)e^{-r/2}r_0 - \frac{1}{\sqrt{2}}(x - iz)e^{-r/2}r_0 & \text{für } s = -1, \\ \frac{1}{\sqrt{2}}(x - iz)e^{-r/2}r_0 + \frac{1}{\sqrt{2}}(x - iz)e^{-r/2}r_0 & \text{für } s = +1 \end{cases} \\ &= \delta_{s1} \sqrt{2}(x - iz)e^{-r/2}r_0. \end{aligned}$$

Im neuen Koordinatensystem hat also der Spin sicher die Richtung Z , im alten hatte er folglich sicher die Richtung Y .

8. Wir wollen nun versuchen, aus der Transformationsformel (19) einige physikalische Folgerungen zu ziehen. Eine sehr wichtige Frage, die mit ihrer Hilfe entschieden werden kann, ist die folgende: Mit welchen Wahrscheinlichkeiten ergibt die Messung der Z' -Komponente des Spins die beiden Werte $+h/4\pi$ und $-h/4\pi$, wenn die Z -Komponente sicher den Wert $+h/4\pi$ hat? Mit anderen Worten: was ist der Wahrscheinlichkeitszusammenhang zwischen den Spinkomponenten für zwei verschiedene Richtungen Z' und Z , die miteinander einen bestimmten Winkel β einschließen? Wenn der Spin die Z -Richtung hat, hat die Wellenfunktion die Form $\Phi(xyzs) = \delta_{s1}\varphi(xyz)$, und wenn wir diesen Zustand von einem um $\{0 \beta 0\}$ verdrehten Koordinatensystem betrachten, so wird wegen $P_{\{0 \beta 0\}} \Phi(xyzs) = \delta_{s1} P_{\{0 \beta 0\}} \varphi(xyz)$

$$\left. \begin{aligned} O_{\{0 \beta 0\}} \Phi(x, y, z, -1) &= \cos \frac{1}{2}\beta P_{\{0 \beta 0\}} \Phi(x, y, z, -1) - \sin \frac{1}{2}\beta P_{\{0 \beta 0\}} \Phi(x, y, z, 1) \\ &= -\sin \frac{1}{2}\beta \cdot P_{\{0 \beta 0\}} \varphi(xyz), \\ O_{\{0 \beta 0\}} \Phi(x, y, z, 1) &= \sin \frac{1}{2}\beta P_{\{0 \beta 0\}} \Phi(x, y, z, -1) + \cos \frac{1}{2}\beta P_{\{0 \beta 0\}} \Phi(x, y, z, 1) \\ &= \cos \frac{1}{2}\beta \cdot P_{\{0 \beta 0\}} \varphi(xyz). \end{aligned} \right\} (20)$$

Nun kann der zweite Beobachter die Wahrscheinlichkeit für die Einstellung des Spins in seine Z -Achse — das ist eine Richtung, die mit der ursprünglichen Z -Richtung den Winkel β einschließt — aus der Wellenfunktion $O_{\{\alpha\beta\gamma\}}\Phi$ direkt berechnen. Er erhält aus (20) $|\cos \frac{1}{2}\beta|^2$ für die Wahrscheinlichkeit der $+Z'$ -Richtung und $|- \sin \frac{1}{2}\beta|^2$ für die Wahrscheinlichkeit der $-Z'$ -Richtung. Ist die Wahrscheinlichkeit für eine bestimmte Richtung des Spins gleich 1, so ist sie für eine gegen diese um β geneigte Richtung gleich $\cos^2 \frac{1}{2}\beta$. Für $\beta = 0$ ist dies 1, dann fallen ja die beiden Richtungen zusammen, für $\beta = \frac{1}{2}\pi$, wenn die beiden Richtungen senkrecht aufeinander stehen, gleich $\frac{1}{2}$, für $\beta = \pi$, wenn die beiden Richtungen entgegengesetzt sind, ist diese Wahrscheinlichkeit Null.

Es ist noch die Frage von Interesse, wann eine Richtung existiert, in die sich der Spin sicher nicht einstellt? Sei diese Richtung etwa die Z' -Richtung, so hat die Wellenfunktion $O_{\{\alpha\beta\gamma\}}\Phi$ für ein Koordinatensystem, dessen Z -Achse Z' ist, die Gestalt

$$O_{\{\alpha\beta\gamma\}}\Phi(xyzs) = \delta_{s,-1}\varphi(xyz).$$

Die Wellenfunktion Φ selber ist, wenn wir zur Abkürzung $\{\alpha\beta\gamma\} = R$ schreiben,

$$\begin{aligned}\Phi(x,y,z,s) &= O_{R^{-1}}O_R\Phi = O_{R^{-1}}\delta_{s,-1}\varphi(xyz) \\ &= D^{(1/2)}(R^{-1})_{\frac{1}{2}s,-\frac{1}{2}}P_{R^{-1}}\varphi(xyz).\end{aligned}$$

$\Phi(x,y,z,-1)$ und $\Phi(x,y,z,+1)$ unterscheiden sich nur um einen konstanten von xyz unabhängigen Faktor

$$\begin{aligned}&\Phi(x,y,z,-1)/\Phi(x,y,z,1) \\ &= D^{(1/2)}(R)_{-\frac{1}{2}-\frac{1}{2}}/D^{(1/2)}(R)_{\frac{1}{2}-\frac{1}{2}} = e^{-i\alpha} \operatorname{ctg} \frac{1}{2}\beta,\end{aligned}\quad (\dagger\dagger)$$

der, wie ($\dagger\dagger$) zeigt, sowohl dem Absolutwert wie der komplexen Phase nach ganz beliebig sein kann. Daß sich $\Phi(x,y,z,-1)$ und $\Phi(x,y,z,+1)$ nur durch einen Faktor unterscheiden, bedeutet, daß bei gleichzeitiger Messung der Z -Komponente des Spins und einer beliebigen spinfreien Größe, die Wahrscheinlichkeiten für die letztere nach 3. von der Spinrichtung statistisch unabhängig sind. In diesem Falle existiert also immer eine Richtung — Azimut α und Polabstand β sind in ($\dagger\dagger$) gegeben —, in die sich der Spin sicher nicht einstellt, sonst aber existiert keine solche Richtung.

9. Es ist vielleicht nicht überflüssig, hervorzuheben, wie weitgehend konkrete Aussagen man für das Benehmen des Drehelektrons lediglich auf Grund der allgemeinen Prinzipien der Quantenmechanik und einiger qualitativer Vorstellungen mit Hilfe von Invarianzforderungen erhalten kann. Die beiden eben abgeleiteten Resultate — namentlich das erstere über den Wahrscheinlichkeitszusammenhang verschiedener Spinrichtungen — ließen sich, wenigstens im Prinzip, auch experimentell prüfen.

Die Bestimmung der Operatoren \mathbf{O}_R gelang uns unter der Voraussetzung, daß sich die verschiedenen Achsenkreuze physikalisch nicht unterscheiden. Äußere Felder, die die Isotropie des Raumes stören, könnten auch eine Modifikation der Operatoren \mathbf{O}_R nach sich ziehen. Natürlich wird, solange das äußere Feld schwach ist, der Operator, der den Übergang zum verdrehten Achsenkreuz bewirkt, noch näherungsweise durch (19) bestimmt sein. Die Gültigkeit von (19) wird aber auch in starken Feldern angenommen.

Es werde noch ein Gedanke, der bei der Ableitung von (19) von entscheidender Bedeutung ist, aber im mathematischen Formalismus etwas untergetaucht war, explizite hervorgehoben. Das ist die Tatsache, daß die Gleichheit zweier Koordinatensysteme auch die Gleichheit der Operatoren \mathbf{O}_R für beide nach sich zieht, die den Übergang zu einem in der gleichen Weise verdrehten Koordinaten- system vermitteln.

Die Operatoren \mathbf{O}_R sind zwar linear-unitär, aber keine Punkttransformationen, wie sie die \mathbf{P}_R waren. Deshalb gilt für sie (22) Kap. XI,

$$\mathbf{O}_R \Phi \Psi = \mathbf{O}_R \Phi \cdot \mathbf{O}_R \Psi$$

im allgemeinen 'keineswegs'. Außerdem ist zu bemerken, daß einer Drehung R nicht ein, sondern eigentlich zwei Operatoren \mathbf{O}_R und $-\mathbf{O}_R$ zugeordnet sind, weil das $\mathfrak{D}^{1/2}(\{\alpha \beta \gamma\})$, das in (19) vorkommt, durch die Drehung nur bis auf das Vorzeichen bestimmt ist. Es gilt auch nicht $\mathbf{O}_{SR} = \mathbf{O}_S \mathbf{O}_R$, sondern nur

$$\mathbf{O}_{SR} = \pm \mathbf{O}_S \mathbf{O}_R, \quad (16a)$$

und man kann von den Operatoren $+\mathbf{O}_R$ und $-\mathbf{O}_R$ auch nicht willkürlich je einen in der Weise weglassen, daß für den Rest (16a) durchweg mit dem oberen Vorzeichen gelte.

10. Die Z -Komponente des Spins ist eine „physikalische Größe“, ebenso wie der Impuls oder die Koordinate. Nach der statistischen Deutung der Quantenmechanik muß ihr also ein linearer hermiteischer Operator ent-

sprechen, den wir mit $\frac{h}{4\pi} \cdot \mathbf{s}_z$ bezeichnen wollen. Die Eigenwerte von \mathbf{s}_z sind den möglichen Werten $-h/4\pi$ und $+h/4\pi$ der Z -Komponente des Spins entsprechend -1 und $+1$. Die Eigenfunktionen zum ersten Eigenwert sind alle Funktionen $\Psi_-(xyzs) = \delta_{s,-1}\psi(xyz)$, die nur für $s = -1$, die Eigenfunktionen zum zweiten Eigenwert alle Funktionen $\Psi_+(xyzs) = \delta_{s,1}\psi'(xyz)$, die nur für $s = 1$ von Null verschieden sind. Es ist also

$$\begin{aligned}\mathbf{s}_z \delta_{s,-1}\psi(xyz) &= -\delta_{s,-1}\psi(xyz) \\ \mathbf{s}_z \delta_{s,1}\psi'(xyz) &= \delta_{s,1}\psi'(xyz),\end{aligned}$$

und für ein beliebiges

$$\Phi(x, y, z, s) = \delta_{s,-1}\Phi(x, y, z, -1) + \delta_{s,1}\Phi(x, y, z, 1)$$

gilt wegen der Linearität

$$\begin{aligned}\mathbf{s}_z \Phi(xyzs) &= \mathbf{s}_z(\delta_{s,-1}\Phi(x, y, z, -1) + \delta_{s,1}\Phi(x, y, z, 1)) \\ &= -\delta_{s,-1}\Phi(x, y, z, -1) + \delta_{s,1}\Phi(x, y, z, 1) \\ \mathbf{s}_s \Phi(xyzs) &= \sum_{t=\pm 1} t \delta_{st} \Phi(xyzt) = s \Phi(xyzs).\end{aligned}\quad (21)$$

Da \mathbf{s}_z nur auf die Spinkoordinaten einwirkt, hat es ebenso wie \mathbf{Q}_R eine Matrixform

$$\mathbf{s}_z = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}. \quad (21a)$$

Bestimmen wir nun den Operator \mathbf{h} , der der Z' -Komponente des Spins entspricht. Für den Beobachter mit dem Achsenkreuz, dessen Z -Achse Z' ist, ist dieser Operator einfach \mathbf{s}_z , da für diesen Beobachter definitionsgemäß alle Operatoren genau so wie für den ersten Beobachter lauten, nur daß sie auf sein eigenes Achsenkreuz bezogen sind. Andererseits entsteht aber dieser Operator durch Transformation mit \mathbf{O}_R aus \mathbf{h} , so daß

$$\mathbf{s}_z = \mathbf{O}_R \mathbf{h} \mathbf{O}_R^{-1}; \quad \mathbf{h} = \mathbf{O}_{R^{-1}} \mathbf{s}_z \mathbf{O}_R$$

und hieraus wegen (12) und der Vertauschbarkeit von \mathbf{P}_R mit \mathbf{s}_z

$$\mathbf{h} = \mathbf{Q}_{R^{-1}} \mathbf{P}_{R^{-1}} \mathbf{s}_z \mathbf{P}_R \mathbf{Q}_R = \mathbf{Q}_{R^{-1}} \mathbf{s}_z \mathbf{Q}_R \quad (22)$$

folgt. Wenn man für alle in (22) vorkommenden Operatoren — sie wirken alle nur auf s ein — die Matrixform verwendet, erhält man:

$$\mathbf{h} = \mathbf{u}(R)^\dagger \mathbf{s}_z \mathbf{u}(R).$$

Es folgt nun aus (11), Kapitel XV, daß unser \mathbf{h} mit dem dort verwendeten identisch wird, wenn man $\bar{\mathbf{h}} = \mathbf{s}_z$, also $x' = y' = 0$; $z' = 1$ setzt. Das $\mathbf{r} = (x, y, z)$ in (11) ist der Vektor, der aus \mathbf{r}' (mit den Komponenten $x' = y' = 0$; $z' = 1$) durch die Transformation R^{-1} hervorgeht, also der Einheitsvektor in der Z' -Richtung. Daher lautet die Definitionsgleichung (10a) von \mathbf{h} in Kapitel XV

$$\mathbf{h} = \alpha_1 \mathbf{s}_x + \alpha_2 \mathbf{s}_y + \alpha_3 \mathbf{s}_z, \quad (22a)$$

wo $\alpha_1, \alpha_2, \alpha_3$ die Richtungscosinus von Z' sind. Aus (22a) sehen wir, daß sich der Operator für die Z' -Komponente des Spins ebenso aus den Operatoren für die X, Y, Z -Komponenten [vgl. (10), Kap. XV]

$$s_x = \begin{pmatrix} 0 & i \\ -i & 0 \end{pmatrix}, \quad s_y = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad s_z = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}$$

zusammensetzt, wie etwa der Operator für die Z' -Komponente der Koordinate (Multiplizieren mit $\alpha_1 x + \alpha_2 y + \alpha_3 z$) aus den Operatoren für die X, Y, Z -Koordinate (Multiplizieren mit x, y, z) gebildet wird. Operatoren dieser Art nennt man „Vektoroperatoren“.

Die unmittelbare Bedeutung von (11), Kapitel XV, ist die, daß der Operator $[R r, f] = [r', f]$, der der $R r$ -Komponente des Spins entspricht, aus dem Operator $[r, f]$, der der r -Komponente entspricht, durch Transformation mit $u(R)^{-1}$, oder wie wir es jetzt sagen würden, mit O_R^{-1} hervorgeht.

In der Theorie des Spins pflegt man öfters direkt von (22a) auszugehen und die ganze Theorie auf diese Gleichung zu gründen.

Anhang. Es sei zunächst $\bar{\Phi}$ die Wellenfunktion des Zustandes Φ für den zweiten Beobachter. Dann muß nach (8) für alle Funktionen Ψ, Φ

$$|(\Psi, \Phi)| = |(\bar{\Psi}, \bar{\Phi})| \quad (8')$$

sein. Eigentlich muß (8') nur gelten, wenn Ψ und Φ physikalischen Zuständen entsprechen, also normiert sind. Sonst kann man von der „zweiten Beschreibungsweise“ des Zustandes Φ gar nicht sprechen, weil eben nur normierte Φ Zustände darstellen. Es ist indessen zweckmäßig, auch für nicht normierte Φ' ein $\bar{\Phi}'$ zu definieren, und zwar sei $\bar{\Phi}'$, wenn $\Phi' = a\Phi$, und Φ normiert ist, gleich $a\bar{\Phi}$. Dann gilt (8') für alle Funktionen.

Weiter ändert sich an (8') nichts, wenn man die Ψ und Φ mit Konstanten vom Betrag 1 multipliziert. Es soll nun gezeigt werden, daß man diese Konstanten c_Ψ, c_Φ so wählen kann, daß für alle $c_\Psi \bar{\Psi} = O_R \Psi$ und $c_\Phi \bar{\Phi} = O_R \Phi$ nicht nur

$$|(O_R \Psi, O_R \Phi)| = |(\Psi, \Phi)|, \quad (8)$$

sondern auch

$$(O_R \Psi, O_R \Phi) = (\Psi, \Phi),$$

$$O_R(a\Psi + b\Phi) = aO_R \Psi + bO_R \Phi \quad (8a)$$

gilt (a und b sind Konstanten). Die Schwierigkeit bei dem Übergang von (8) nach (8a) besteht darin, daß (8) nur die Gleichheit der Absolutwerte von $(O_R \Psi, O_R \Phi)$ und (Ψ, Φ) besagt, während nach (8a) auch die komplexen Phasen gleich sein sollen, und dies für alle Funktionenpaare gleichzeitig erreicht werden muß.

Bilden die Funktionen $\Psi_1, \Psi_2, \Psi_3, \dots$ ein vollständiges Orthogonalsystem, so gilt dies auch von den Funktionen $\bar{\Psi}_1, \bar{\Psi}_2, \bar{\Psi}_3, \dots$. Aus $(\Psi_i, \Psi_k) = \delta_{ik}$ folgt nach (8') auch $(\bar{\Psi}_i, \bar{\Psi}_k) = \delta_{ik}$, und wenn keine Funktion existiert, die zu allen Ψ_i orthogonal ist, so kann es auch keine Funktion geben, die auf alle $\bar{\Psi}_i$ senkrecht steht.

Betrachten wir nun die Funktionen \bar{F}_x , die den $F_x = \Psi_1 + \Psi_x$ für $x = 2, 3, 4, \dots$ zugeordnet sind. Wenn wir \bar{F}_x nach dem vollständigen Orthogonalsystem $\bar{\Psi}_1, \bar{\Psi}_2, \dots$ entwickeln, so sind alle Entwicklungskoeffizienten $(\bar{\Psi}_k, \bar{F}_x)$ bis auf die von $\bar{\Psi}_1$ und $\bar{\Psi}_x$ Null, und diese sind vom Absolutwert 1, weil auch (Ψ_k, F_x) nur für $\lambda = 1$ und $\lambda = x$ nicht verschwindet und für diese Werte von λ den Wert 1 hat. Es ist also

$$\bar{F}_x = y_x(\bar{\Psi}_1 + x_x \bar{\Psi}_x); |y_x| = |x_x| = 1 \text{ (für } x = 2, 3, \dots\text{). (A)}$$

Wir verfügen jetzt über einige c , und zwar sei $c_{\Psi_1} = 1; c_{\Psi_x} = x_x; c_{F_x} = 1/y_x$

$$\left. \begin{aligned} O_R \Psi_1 &= \bar{\Psi}_1; & O_R \Psi_x &= c_{\Psi_x} \bar{\Psi}_x = x_x \bar{\Psi}_x, \\ O_R (\Psi_1 + \Psi_x) &= O_R F_x = c_{F_x} \bar{F}_x = \bar{F}_x / y_x = O_R \Psi_1 + O_R \Psi_x. \end{aligned} \right\} \text{ (B)}$$

Nun sei Φ eine beliebige Funktion, die wir uns nach den Ψ_x entwickelt denken

$$\Phi = a_1 \Psi_1 + a_2 \Psi_2 + a_3 \Psi_3 + \dots \quad (\text{C})$$

Entwickeln wir auch $\bar{\Phi}$, und zwar nach dem vollständigen Orthogonalsystem der $O_R \Psi_1, O_R \Psi_2, \dots$

$$\bar{\Phi} = \bar{a}_1 O_R \Psi_1 + \bar{a}_2 O_R \Psi_2 + \bar{a}_3 O_R \Psi_3 + \dots \quad (\text{C}')$$

so ist

$$|\bar{a}_x| = |(O_R \Psi_x, \bar{\Phi})| = |(x_x \bar{\Psi}_x, \bar{\Phi})| = |(\Psi_x, \Phi)| = |a_x| \quad (\text{D})$$

und insbesondere $|\bar{a}_1| = |a_1|$. Wählen wir daher $c_\Phi = a_1/\bar{a}_1$, so wird

$$O_R \Phi = c_\Phi \bar{\Phi} = a_1 O_R \Psi_1 + a'_2 O_R \Psi_2 + a'_3 O_R \Psi_3 + \dots \quad (\text{E})$$

der Koeffizient von $O_R \Psi_1$ gleich a_1 , und es gilt auch weiter $|a'_x| = |a_x|$. Außerdem muß nach (8) und (B)

$$\begin{aligned} |a_1 + a'_x|^2 &= |(O_R \Psi_1 + O_R \Psi_x, O_R \Phi)|^2 = |(O_R F_x, O_R \Phi)|^2 \\ &= |(F_x, \Phi)|^2 = |a_1 + a_x|^2, \end{aligned}$$

oder

$$|a_1|^2 + a_1^* a'_x + a_1 a'_x + |a'_x|^2 = |a_1|^2 + a_1^* a_x + a_1 a_x^* + |a_x|^2$$

sein. Man kann hieraus wegen $a_x' a_x'^* = a_x a_x^*$ das $a_x'^*$ eliminieren und erhält eine Gleichung zweiten Grades

$$a_1^* a_x'^2 - (a_1^* a_x + a_1 a_x^*) a_x' + a_1 |a_x|^2 = 0 \quad (\text{F})$$

für a_x' . Aus (F) folgt entweder

$$a_x' = a_x \quad \text{oder} \quad a_x' = a_x^* a_1 / a_1^*. \quad (\text{G})$$

Im ersten Falle ist für jedes $\Phi = \sum_x a_x \Psi_x$ und $\Psi = \sum_x b_x \Psi_x$

$$\mathbf{O}_R \Phi = \sum_x a_x \mathbf{O}_R \Psi_x; \quad \mathbf{O}_R \Psi = \sum_x b_x \mathbf{O}_R \Psi_x. \quad (\text{E}')$$

Ebenso ist auch

$$\begin{aligned} \mathbf{O}_R (a\Phi + b\Psi) &= \mathbf{O}_R \left(\sum_x (a a_x + b b_x) \Psi_x \right) = \sum_x (a a_x + b b_x) \mathbf{O}_R \Psi_x \\ &= a \mathbf{O}_R \Phi + b \mathbf{O}_R \Psi, \end{aligned}$$

so daß \mathbf{O}_R tatsächlich linear ist. Weiter ist

$$\begin{aligned} (\mathbf{O}_R \Psi, \mathbf{O}_R \Phi) &= \left(\sum_x b_x \mathbf{O}_R \Psi_x, \sum_\lambda a_\lambda \mathbf{O}_R \Psi_\lambda \right) \\ &= \sum_{x\lambda} b_x^* a_\lambda \delta_{x\lambda} = \sum_x b_x^* a_x, \end{aligned}$$

und es ist auch

$$(\Psi, \Phi) = \left(\sum_x b_x \Psi_x, \sum_\lambda a_\lambda \Psi_\lambda \right) = \sum_{x\lambda} b_x^* a_\lambda \delta_{x\lambda} = \sum_x b_x^* a_x,$$

\mathbf{O}_R ist auch unitär, womit (8a) bewiesen ist.

Wir müssen nur noch zeigen, daß die zweite Alternative in (G) nicht auftreten kann. Zu diesem Zweck ersetzen wir für diesen Fall

$$\mathbf{O}_R \Phi = \mathbf{O}_R \sum_x a_x \Psi_x = \frac{a_1}{a_1^*} \sum_x a_x^* \mathbf{O}_R \Psi_x$$

durch

$$\mathbf{O}_R \Phi = \mathbf{O}_R \sum_x a_x \Psi_x = \sum_x a_x^* \mathbf{O}_R \Psi_x. \quad (\text{I})$$

d. h. multiplizieren $\mathbf{O}_R \Phi$ mit a_1^*/a_1 . Dadurch kann sich ja der Inhalt der Beschreibung nicht ändern.

Nun betrachten wir zwei Eigenfunktionen des Hamiltonschen Operators, zwei stationäre Zustände $\chi = \sum_x u_x \Psi_x$ und $\chi' = \sum_x u'_x \Psi_x$ mit verschiedenen Energien E und E' . Es ist

$$e^{2\pi i E t/\hbar} \chi + e^{2\pi i E' t/\hbar} \chi' = \sum_x (u_x e^{2\pi i E t/\hbar} + u'_x e^{2\pi i E' t/\hbar}) \Psi_x \quad (\text{II})$$

254 XX. Äquiv. Beschreib. untersch. sich nur durch eine kanon. Transform.

eine Lösung der zeitabhängigen Schrödingerschen Differentialgleichung. In der zweiten Beschreibungsweise würde nach (I)

$$\mathbf{O}_R \chi = \sum_x u_x^* \mathbf{O}_R \Psi_x$$

dem Zustand χ und

$$\mathbf{O}_R \chi' = \sum_x u'_x^* \mathbf{O}_R \Psi_x$$

dem Zustand χ' entsprechen. Die Energien sind auch in der zweiten Beschreibungsweise E bzw. E' . Daher muß auch

$$\begin{aligned} & e^{2\pi i Et/\hbar} \mathbf{O}_R \chi + e^{2\pi i E't/\hbar} \mathbf{O}_R \chi' \\ &= \sum_x (u_x^* e^{2\pi i Et/\hbar} + u'_x^* e^{2\pi i E't/\hbar}) \mathbf{O}_R \Psi_x \end{aligned} \quad (\text{III})$$

eine Lösung der Schrödinger-Gleichung sein, die für $t = 0$ denselben Zustand wie (II) repräsentiert und demnach für alle späteren Zeiten denselben Zustand darstellen sollte. Dies ist aber nicht richtig, weil dem Zustand (II) nach (I)

$$\sum_x (u_x e^{2\pi i Et/\hbar} + u'_x e^{2\pi i E't/\hbar})^* \mathbf{O}_R \Psi_x$$

entspricht, und dies ist für $t \neq 0$ nur dann mit (III) identisch, wenn $E = E'$ ist. Die zweite Alternative von (G) führt auf einen Widerspruch, bei der in (B) und (E) getroffenen Wahl der c muß in (G) die erste Alternative zutreffen, aus der der linear-unitäre Charakter des \mathbf{O}_R folgt.

Wir erhalten somit das wichtige Resultat, daß zwei physikalisch äquivalente Beschreibungen — nach entsprechender Veränderung der freien Konstanten der Wellenfunktionen — durch eine kanonische Transformation ineinander überführt werden können.

XXI. Die Gesamtquantenzahl

1. In der Transformationsformel (19) des vorangehenden Kapitels

$$\begin{aligned} \mathbf{O}_R \Phi(x, y, z, s) &= \sum_{t=\pm 1} \mathfrak{D}^{(1/2)}(R)_{\frac{1}{2}s, \frac{1}{2}t} \mathbf{P}_R \Phi(x, y, z, t) \\ &= \sum_{t=\pm 1} \mathfrak{D}^{(1/2)}(R)_{\frac{1}{2}s, \frac{1}{2}t} \Phi(x'', y'', z'', t) \end{aligned} \quad (1)$$

ist R eine reine Drehung. Wollen wir die Wellenfunktion in einem Koordinatensystem haben, das durch eine Drehspiegelung

aus dem ursprünglichen entsteht, so können wir zuerst zu dem Koordinatensystem übergehen, das durch Inversion

$$x' = -x, \quad y' = -y, \quad z' = -z$$

aus dem ursprünglichen hervorgeht und dieses dann einer reinen Drehung unterwerfen. Die Frage ist also eigentlich nur die: wie lautet die Wellenfunktion $\mathbf{O}_I \Phi$ des Zustandes Φ für einen Beobachter, dessen Achsenkreuz dem ursprünglichen entgegengerichtete Achsen hat.

Betrachten wir zuerst den Zustand $u_s \psi(xyz)$. Spinfreien Versuchen gegenüber benimmt er sich für den ersten Beobachter so, als ob seine Wellenfunktion ψ wäre, und daher für den zweiten Beobachter so, als ob seine Wellenfunktion $\mathbf{P}_I \psi$ wäre, wo

$$\mathbf{P}_I \psi(xyz) = \psi(-x, -y, -z) \quad (2)$$

ist. Daher muß $\mathbf{O}_I u_s \psi(xyz) = u'_s \mathbf{P}_I \psi(xyz)$ sein. Das magnetische Moment hat für $u_s \psi(xyz)$ eine bestimmte Richtung, bei der Inversion des Achsenkreuzes geht diese Richtung in die entgegengesetzte Richtung über, weil das magnetische Moment ein axialer Vektor ist. Die entgegengesetzte Richtung heißt aber im neuen Achsenkreuz genau so, wie im alten die ursprüngliche Richtung hieß: der zweite Beobachter nennt die Richtung des Spins ebenso, wie sie der erste nannte, der Faktor u'_s vor $\mathbf{P}_I \psi$ in $\mathbf{O}_I u_s \psi(xyz)$ ist gleich u_s .

Man kann einen magnetischen Dipol bekanntlich immer durch einen Stromkreis ersetzen. Liegt dieser Stromkreis etwa in der XY-Ebene und läuft er von X nach Y, so liegt er auch in der X'Y'-Ebene und läuft auch von X' nach Y'.

Es gilt demnach für alle u_s und alle $\psi(xyz)$

$$\mathbf{O}_I u_s \psi(xyz) = u_s \mathbf{P}_I \psi(xyz) = \mathbf{P}_I u_s \psi(xyz), \quad (3)$$

bis auf eine Konstante, die noch von u und ψ abhängen könnte. Man kann aber genau so, wie dies in 6. des vorangehenden Kapitels geschah, einsehen, daß diese Konstante sowohl für alle u wie auch für alle ψ gleich groß sein muß, wenn man an der Linearität von \mathbf{O}_I festhält. Da aber ein Faktor in \mathbf{O}_I doch gänzlich willkürlich ist, können wir diese Konstante ganz weglassen. Da man weiter jede Funktion $\Phi(xyzs)$ als eine Linearkombination von Funktionen der Gestalt $u_s \psi(xyz)$ schreiben kann, folgt aus (3) und der Linearität von \mathbf{O}_I und \mathbf{P}_I , daß $\mathbf{O}_I = \mathbf{P}_I$

$$\mathbf{O}_I \Phi(x, y, z, s) = \mathbf{P}_I \Phi(x, y, z, s) = \Phi(-x, -y, -z, s) \quad (4)$$

ist. Der Operator \mathbf{O}_I , der den Übergang zum Achsenkreuz bewirkt, das durch Inversion aus dem ursprünglichen hervorgeht, wirkt auf die Spinkoordinate s gar nicht ein und ist durch (4) gegeben. Es ist $\mathbf{O}_I^2 = 1$; $\mathbf{O}_I \mathbf{O}_I \Phi = \Phi$; der Einheitsoperator bildet mit \mathbf{O}_I zusammen eine Gruppe, die zur Spiegelungsgruppe holomorph ist.

In (1) und (4) haben wir die Transformationsformeln für die Wellenfunktion bei einer beliebigen Änderung des Achsenkreuzes. Die Formeln (1) und (4) gelten übrigens nicht nur für Elektronen, sondern auch für Protonen. Das magnetische Moment der Protonen ist aber wegen ihrer etwa 1840 mal größeren Masse 1840 mal kleiner und ist daher keiner so unmittelbaren Beobachtung zugänglich, wie der Elektronenspin. Im folgenden wird dieser „Kernspin“ immer vernachlässigt.

Auch in der Diracschen relativistischen Theorie des Elektrons bleiben (1) und (4) im wesentlichen ungeändert¹⁾. Bei Dirac besteht eine Wellenfunktion nicht aus zwei Funktionen $\Phi(x, y, z, -1)$ und $\Phi(x, y, z, 1)$ von x, y, z , sondern aus vier. Man kann das so berücksichtigen, daß man außer s noch eine fünfte Koordinate s' einführt, die ebenfalls zweier Werte fähig ist. Für reine Drehungen gilt dann (1) ungeändert, s' nimmt an der Transformation gar nicht teil. Bei der Inversion vertauschen sich dagegen die beiden s' -Werte noch.

2. Die Formeln (1) und (4) beziehen sich auf ein System, das nur ein einziges Elektron enthält. Im Falle mehrerer Elektronen enthält die Wellenfunktion $\Phi(x_1, y_1, z_1, s_1, x_2, y_2, z_2, s_2, \dots, x_n, y_n, z_n, s_n)$ außer den Cartesischen Koordinaten auch die Spinkoordinaten aller Teilchen. Das skalare Produkt zweier Funktionen Φ und G wird

$$(\Phi, G) = \sum_{s_1=\pm 1} \sum_{s_2=\pm 1} \dots \sum_{s_n=\pm 1} \iint \dots \int \Phi(x_1 \dots s_n)^* G(x_1 \dots s_n) dx_1 \dots dz_n. \quad (5)$$

Ebenso wie in der einfachen Theorie, die den Spin nicht berücksichtigt, der Operator \mathbf{P}_R auf alle Koordinatentripel und zwar in gleicher Weise eingewirkt hat, wird jetzt der Operator \mathbf{O}_R , der den Übergang in der Paulischen Theorie zu einem anderen Achsenkreuz vermittelt, auf alle Koordinaten x_k, y_k, z_k und s_k so ein-

¹⁾ J. A. Gaunt: Proc. Roy. Soc. **124**, 163, 1929.

wirken, wie das \mathbf{O}_R in (1) oder (4) auf xyz und s eingewirkt hat. Es wird also

$$\begin{aligned} \mathbf{O}_R \Phi(x_1 y_1 z_1 s_1 \dots x_n y_n z_n s_n) = \\ \sum_{\dots t_n} \mathfrak{D}^{(1/2)}(R)_{\frac{1}{2}s_1, \frac{1}{2}t_1} \dots \mathfrak{D}^{(1/2)}(R)_{\frac{1}{2}s_n, \frac{1}{2}t_n} \mathbf{P}_R \Phi(x_1 y_1 z_1 t_1 \dots x_n y_n z_n t_n), \quad (6) \end{aligned}$$

bzw.

$$\begin{aligned} \mathbf{O}_I \Phi(x_1 y_1 z_1 s_1 \dots x_n y_n z_n s_n) = \mathbf{P}_I \Phi(x_1 y_1 z_1 s_1 \dots x_n y_n z_n s_n) \\ = \Phi(-x_1, -y_1, -z_1, s_1, \dots, -x_n, -y_n, -z_n, s_n) \quad (7) \end{aligned}$$

gelten. Der Operator \mathbf{O}_R ist das Produkt zweier Operatoren \mathbf{P}_R und \mathbf{Q}_R , deren erster nur auf die Cartesischen Koordinaten einwirkt:

$$\begin{aligned} \mathbf{P}_R \Phi(x'_1 y'_1 z'_1 s_1 \dots x'_n y'_n z'_n s_n) = \Phi(x_1 y_1 z_1 s_1 \dots x_n y_n z_n s_n), \quad (6a) \\ \text{wo die } x'_k y'_k z'_k \text{ aus den } x_k y_k z_k \text{ durch die Drehung } R \text{ hervorgehen,} \\ \text{und deren zweiter nur die Spinkoordinaten berührt:} \end{aligned}$$

$$\begin{aligned} \mathbf{Q}_R \Phi(x_1 y_1 z_1 s_1 \dots x_n y_n z_n s_n) \\ = \sum_{t_1=\pm 1} \dots \sum_{t_n=\pm 1} \mathfrak{D}^{(1/2)}(R)_{\frac{1}{2}s_1, \frac{1}{2}t_1} \dots \mathfrak{D}^{(1/2)}(R)_{\frac{1}{2}s_n, \frac{1}{2}t_n} \Phi(x_1 y_1 z_1 t_1 \dots x_n y_n z_n t_n). \quad (6b) \end{aligned}$$

Da das System der Spinkoordinaten 2^n verschiedene Wertesysteme annehmen kann, ist \mathbf{Q}_R einer 2^n -dimensionalen Matrix äquivalent, deren Zeilen und Spalten den Wertesystemen der Spinkoordinaten entsprechend mit n Indizes, die alle die Werte ± 1 haben können, numeriert sind. Die Matrixform von \mathbf{Q}_R ist

$$\begin{aligned} \mathbf{Q}_R = \mathfrak{D}^{(1/2)}(R) \times \mathfrak{D}^{(1/2)}(R) \times \dots \times \mathfrak{D}^{(1/2)}(R) \\ (\mathbf{Q}_R)_{s_1 s_2 \dots s_n; t_1 t_2 \dots t_n} = \mathfrak{D}^{(1/2)}(R)_{\frac{1}{2}s_1, \frac{1}{2}t_1} \dots \mathfrak{D}^{(1/2)}(R)_{\frac{1}{2}s_n, \frac{1}{2}t_n}. \quad (6c) \end{aligned}$$

Die Operatoren \mathbf{P} sind wieder mit den Operatoren \mathbf{Q} alle vertauschbar

$$\mathbf{P}_S \mathbf{Q}_R = \mathbf{Q}_R \mathbf{P}_S \text{ insbesondere } \mathbf{O}_R = \mathbf{P}_R \mathbf{Q}_R = \mathbf{Q}_R \mathbf{P}_R, \quad (8)$$

und der Operator $\mathbf{O}_I = \mathbf{P}_I$ ist mit allen \mathbf{P}_R und daher auch wegen (8) mit allen \mathbf{O}_R vertauschbar, wenn R eine reine Drehung ist.

Die \mathbf{Q}_R sind durch die Drehung R nur bis auf das Vorzeichen bestimmt, weil in $\mathfrak{D}^{(1/2)}(R)$ das Vorzeichen frei ist. Bei gerader Elektronenzahl kann man diese Zweideutigkeit aufheben, indem man festsetzt, daß alle $\mathfrak{D}^{(1/2)}(R)$ in (6) und (6c) mit demselben Vorzeichen zu nehmen sind. Bei ungerader Elektronenzahl ist es nicht möglich, die \mathbf{Q}_R in dieser Weise eindeutig zu machen.

3. Geht man zuerst zu einem um R und dann zu einem in bezug auf dieses um S verdrehten Koordinatensystem über, so geht die Wellenfunktion Φ zuerst in $\mathbf{O}_R \Phi$ und dann in $\mathbf{O}_S \mathbf{O}_R \Phi$ über. Dasselbe Koordinatensystem erhält man aber auch, wenn man sofort um SR dreht. In diesem Fall erhält man für die Wellenfunktion $\mathbf{O}_{SR} \Phi$, und dies kann sich von $\mathbf{O}_S \mathbf{O}_R \Phi$ nur durch eine Konstante unterscheiden. Da sowohl $\mathbf{O}_S \mathbf{O}_R$ als auch \mathbf{O}_{SR} linear-unitär sind, ist diese Konstante für alle Wellenfunktionen dieselbe, sie kann nur noch von den beiden Drehungen S und R abhängen

$$\mathbf{O}_{SR} = c_{S,R} \mathbf{O}_S \mathbf{O}_R. \quad (9)$$

Da der Übergang zur Beschreibung mit einem anderen Koordinatensystem immer durch einen linear-unitären Operator bewirkt werden kann, enthält (9) noch nichts von den speziellen Annahmen der Paulischen Theorie und ist eine notwendige Folge der Invarianz des Gleichungssystems räumlichen Drehungen gegenüber. Wir werden diese Gleichung noch am Schlusse dieses Kapitels genauer untersuchen und zeigen, wie man aus ihr Folgerungen ziehen kann, die in jeder quantenmechanischen Theorie gelten müssen.

Man kann (9) natürlich auch rechnerisch verifizieren, es ist zunächst nach (8)

$$\mathbf{O}_S \mathbf{O}_R = \mathbf{P}_S \mathbf{Q}_S \mathbf{P}_R \mathbf{Q}_R = \mathbf{P}_S \mathbf{P}_R \mathbf{Q}_S \mathbf{Q}_R = \mathbf{P}_{SR} \mathbf{Q}_S \mathbf{Q}_R.$$

Bei gerader Elektronenzahl bilden die Matrizen (6c), die die Matrixform von \mathbf{Q} sind, eine eindeutige Darstellung der Drehgruppe, so daß $\mathbf{Q}_S \mathbf{Q}_R = \mathbf{Q}_{SR}$ ist, und

$$\mathbf{O}_S \mathbf{O}_R = \mathbf{P}_{SR} \mathbf{Q}_S \mathbf{Q}_R = \mathbf{P}_{SR} \mathbf{Q}_{SR} = \mathbf{O}_{SR} \quad (10a)$$

gilt, in diesem Fall ist in (9) die Konstante $c_{S,R} = 1$, und die Operatoren \mathbf{O}_R bilden eine zur reinen Drehgruppe holomorphe Gruppe. Man kann daher in diesem Fall Funktionen definieren, die in bezug auf die Operatoren \mathbf{O}_R zu einer Zeile einer irreduziblen Darstellung oder auch zu einer irreduziblen Darstellung der Drehgruppe gehören.

Bei ungerader Elektronenzahl bilden die Matrizen (6c) nur eine zweideutige Darstellung der reinen Drehgruppe, es ist $\mathbf{Q}_S \mathbf{Q}_R = \pm \mathbf{Q}_{SR}$ und daher ist

$$\mathbf{O}_S \mathbf{O}_R = \mathbf{P}_{SR} \mathbf{Q}_S \mathbf{Q}_R = \pm \mathbf{P}_{SR} \mathbf{Q}_{SR} = \pm \mathbf{O}_{SR}, \quad (10b)$$

die Konstante in (9) ist $c_{S,R} = \pm 1$, die Operatoren \mathbf{O}_R sind zur Drehgruppe eigentlich nicht mehr holomorph. Es entsprechen wegen der Zweideutigkeit der \mathbf{Q}_R auch je zwei Operatoren $+\mathbf{O}_R$ und $-\mathbf{O}_R$ jeder Drehung R . Da auch in der Isomorphie der unitären Gruppe¹⁾ zur Drehgruppe jeder Drehung R zwei unitäre Matrizen $\mathbf{u} = \mathfrak{D}^{(1/2)}(R)$ und $\mathbf{v} = -\mathfrak{D}^{(1/2)}(R)$ entsprechen, kann man versuchen, die \mathbf{O} in eindeutiger Weise den \mathbf{u} zuzuordnen. In der Tat gelingt dies: man muß dabei dem \mathbf{u} das $\mathbf{O}_u = \mathbf{Q}_u \mathbf{P}_{R_u}$ zuordnen und als Matrixform von \mathbf{Q}_u nach (6c) $\mathbf{u} \times \mathbf{u} \times \dots \times \mathbf{u}$ annehmen, während R_u die \mathbf{u} im Sinne des Isomorphismus zugeordnete Drehung ist. Dann ist \mathbf{Q}_u dem \mathbf{u} jedenfalls eindeutig zugeordnet. Da auch R_u dem \mathbf{u} eindeutig zugeordnet ist, gilt dies auch für die Operatoren \mathbf{P}_{R_u} . Es ist auch

$$(\mathbf{u} \times \mathbf{u} \times \dots \times \mathbf{u}) \cdot (\mathbf{v} \times \mathbf{v} \times \dots \times \mathbf{v}) = \mathbf{u}\mathbf{v} \times \mathbf{u}\mathbf{v} \times \dots \times \mathbf{u}\mathbf{v},$$

und aus $R_u R_v = R_{uv}$ folgt auch $\mathbf{P}_{R_u} \mathbf{P}_{R_v} = \mathbf{P}_{R_{uv}}$ und daher auch

$$\mathbf{O}_u \mathbf{O}_v = \mathbf{O}_{uv}.$$

Bei ungerader Elektronenzahl gehören also die Funktionen $f_{-j}, f_{-j+1}, \dots, f_{j-1}, f_j$, für die

$$\mathbf{O}_u f_\mu^{(j)} = \sum_{\mu'=-j}^j \mathfrak{U}^{(j)}(\mathbf{u})_{\mu'\mu} f_{\mu'}^{(j)} \quad (11)$$

gilt, zu den verschiedenen Zeilen der Darstellung $\mathfrak{U}^{(j)}$ der unitären Gruppe. Sie erfüllen daher ebenfalls die Relationen, die im Kapitel XII für Funktionen abgeleitet wurden, welche zu einer irreduziblen Darstellung irgendeiner Gruppe gehören.

An Stelle der Gleichung (11) wird in der Folge immer die Gleichung

$$\mathbf{O}_R f_\mu^{(j)} = \pm \sum_{\mu'=-j}^j \mathfrak{D}^{(j)}(R)_{\mu'\mu} f_{\mu'}^{(j)} \quad (11a)$$

aufreten, aus der (11) nur mit dem \pm -Zeichen

$$\mathbf{O}_u f_u^{(j)} = \pm \sum_{\mu'} \mathfrak{U}^{(j)}(\mathbf{u})_{\mu'\mu} f_{\mu'}^{(j)} \quad (11b)$$

zu folgen scheint. Es ist aber zu bemerken, daß eigentlich immer (11) abgeleitet wird, und zweitens folgt aus (11a) nicht nur (11b), sondern auch (11). Würde nämlich in (11b) für

¹⁾ Genauer: der Gruppe der zweidimensionalen unitären Matrizen mit der Determinante 1.

irgendein u das untere Zeichen gelten, so könnte man u stetig verändern, bis es in die Einheit übergeht. Dabei ändern sich beide Seiten von (11b) stetig, so daß man immerfort bei dem unteren Vorzeichen bleiben müßte. Für $u = 1$ bedeutet aber (11b) mit dem unteren Vorzeichen

$$\mathbf{O}_1 f_u^{(j)} = - \sum_{\mu'} \delta_{\mu' \mu} f_{\mu'}^{(j)} = - f_u^{(j)},$$

was sicher nicht richtig ist, da \mathbf{O}_1 der Einheitsoperator ist, der jede Funktion ungeändert läßt. Es muß daher in (11b) überall das obere Vorzeichen gelten; (11a) ist mit (11) identisch und wird nur deshalb an seiner Stelle benutzt, weil es an die Bedeutung der Operationen \mathbf{O} als räumliche Drehungen erinnert.

Setzen wir noch in (11) $u = -1$. Es ist dann \mathbf{O}_{-1} der negative Einheitsoperator, weil $\mathbf{P}_{R_{-1}} = \mathbf{P}_E$ der positive und $-1 \times -1 \times \cdots \times -1$ in (6c) der negative Einheitsoperator ist. Es folgt dann aus (11), daß $\mathbf{U}^{(j)}(-1) = -1$ ist und hieraus nach dem XV. Kapitel, daß j halbzahlig sein muß. Bei ungerader Elektronenzahl kann eine Wellenfunktion in bezug auf die \mathbf{O} nur zu einer ungeraden Darstellung der unitären Gruppe, also zu einer zweideutigen Darstellung der Drehgruppe gehören. Bei gerader Elektronenzahl kommen natürlich nur regelmäßige eindeutige Darstellungen der Drehgruppe in Frage (oder, wenn man will, gerade Darstellungen der unitären Gruppe).

Die Komplikation mit den zweideutigen Darstellungen röhrt daher, daß die $c_{S,R}$ in (9) nicht nur +1, sondern auch -1 sein können, die Operatoren \mathbf{O} , die die Invarianz der Beschreibung räumlichen Drehungen gegenüber zum Ausdruck bringen, bilden keine zur Drehgruppe, sondern eine zur unitären Gruppe holomorphe Gruppe.

4. Der Hamiltonsche Operator \mathbf{H} der Schrödingerschen Eigenwertgleichung $\mathbf{H}\Psi = E\Psi$ für die Energie E ist bei Mitberücksichtigung des Spins nicht mehr der einfache, nur auf die Cartesischen Koordinaten einwirkende Operator, der den bisherigen Betrachtungen zugrunde lag. Die Kräfte, die von den magnetischen Momenten der Elektronen ausgehen, machen noch Zusatzglieder notwendig, auf deren Bedeutung wir noch einmal zu sprechen kommen werden. Obwohl die genaue Gestalt dieser Glieder noch nicht feststeht, ist immerhin so viel sicher wahr, daß, solange keine äußeren

elektrischen oder magnetischen Felder vorhanden sind, keine Raumrichtung ausgezeichnet sein kann, und wenn Ψ_μ ein stationärer Zustand ist, so ist auch der verdrehte Zustand $\mathbf{O}_R \Psi_\mu$ oder $\mathbf{O}_u \Psi_\mu$ stationär, und beide haben dieselbe Energie. Daraus folgt, daß $\mathbf{O}_R \Psi_\mu$ bzw. $\mathbf{O}_u \Psi_\mu$ durch die anderen Eigenfunktionen desselben Eigenwertes linear ausgedrückt werden kann:

$$\mathbf{O}_R \Psi_\mu = \sum_v \mathbf{D}(R)_{v\mu} \Psi_v, \text{ bzw. } \mathbf{O}_u \Psi_\mu = \sum_v \mathbf{D}(u)_{v\mu} \Psi_v. \quad (12)$$

Aus $\mathbf{O}_S \mathbf{O}_R = \mathbf{O}_{SR}$ bzw. bei ungerader Elektronenzahl aus $\mathbf{O}_S \mathbf{O}_R = \pm \mathbf{O}_{SR}$ oder $\mathbf{O}_u \mathbf{O}_v = \mathbf{O}_{uv}$ kann man in bekannter Weise schließen, daß

$$\begin{aligned} \mathbf{D}(S) \mathbf{D}(R) &= \mathbf{D}(SR) \\ \text{bzw.} \quad \mathbf{D}(S) \mathbf{D}(R) &= \pm \mathbf{D}(SR) \quad \text{oder} \quad \mathbf{D}(u) \mathbf{D}(v) = \mathbf{D}(uv). \end{aligned} \quad \left. \right\} (13)$$

Die Matrizen $\mathbf{D}(R)$ bilden bei gerader Elektronenzahl eine Darstellung der Drehgruppe, bei ungerader Elektronenzahl bilden sie eine zweideutige Darstellung der Drehgruppe und eine eindeutige Darstellung der unitären Gruppe.

Man folgert weiter, ebenso wie im Kapitel XII, daß diese Darstellungen als irreduzibel angenommen werden können¹⁾: $\mathbf{D}(R)$ kann bei gerader Elektronenzahl $\mathfrak{D}^{(0)}, \mathfrak{D}^{(1)}, \mathfrak{D}^{(2)}, \dots$, sein, bei ungerader Elektronenzahl kann $\mathbf{D}(R)$ gleich $\mathfrak{D}^{(1/2)}, \mathfrak{D}^{(3/2)}, \mathfrak{D}^{(5/2)}, \dots$ [und $\mathbf{D}(u)$ gleich $\mathfrak{U}^{(1/2)}, \mathfrak{U}^{(3/2)}, \mathfrak{U}^{(5/2)}, \dots$] sein:

$$\mathbf{O}_R \Psi_\mu^j = \sum_{\mu'} \mathfrak{D}^{(j)}(R)_{\mu'\mu} \Psi_{\mu'}^j. \quad (12a)$$

Den oberen Index dieser Darstellungen nennt man die Gesamtquantenzahl und bezeichnet ihn mit dem Buchstaben j oder J , sie ist bei gerader Elektronenzahl ganzzahlig, bei ungerader Elektronenzahl halbzahlig. Die Zeile μ , zu der die Eigenfunktion Ψ_μ^j gehört, nennt man auch hier die magnetische Quantenzahl. Auch μ ist bei gerader Elektronenzahl ganzzahlig, bei ungerader halbzahlig.

5. Es sei \mathbf{S} ein den \mathbf{O}_R gegenüber symmetrischer Operator, also ein Skalar, der durch die Veränderung des Achsenkreuzes

¹⁾ Indem man einen Eigenwert als mehrere zufällig zusammenfallende Eigenwerte ansieht.

nicht berührt wird. Wir wissen dann, daß die sogenannten Matrixelemente

$$S_{Nj\mu; N'j'\mu'} = (\Psi_\mu^{Nj}, \mathbf{S} \Psi_{\mu'}^{N'j'}) = \delta_{jj'} \delta_{\mu\mu'} S_{Nj; N'j} \quad (14)$$

für zwei Eigenfunktionen, die zu verschiedenen Darstellungen $\mathfrak{D}^{(j)}$ und $\mathfrak{D}^{(j')}$, oder zu verschiedenen Zeilen derselben Darstellung gehören, verschwindet. Ist dagegen in (14) $j = j'$ und $\mu = \mu'$, so ist (14) von der magnetischen Quantenzahl unabhängig für alle μ gleich.

Es liegt nahe, nach der analogen Formel für Vektor- und Tensoroperatoren zu fragen.

Der skalare Operator war dadurch definiert, daß er unabhängig vom Achsenkreuz war. Z. B. ist die Energie eine vom Achsenkreuz unabhängige Größe. Dagegen ist es z. B. die X-Komponente des Dipolmoments nicht. Der ersten Größe entsprach für alle Beobachter derselbe Operator. Da andererseits der erste Beobachter der Größe, welcher der zweite Beobachter \mathbf{S} zuordnet, $\mathbf{O}_R^{-1} \mathbf{S} \mathbf{O}_R$ zuordnet, muß

$$\mathbf{O}_R^{-1} \mathbf{S} \mathbf{O}_R = \mathbf{S}; \quad \mathbf{S} \mathbf{O}_R = \mathbf{O}_R \mathbf{S} \quad (15)$$

sein: ein symmetrischer Operator ist mit allen Transformationsoperatoren vertauschbar.

Sind dagegen \mathbf{V}_x , \mathbf{V}_y , \mathbf{V}_z die X, Y, Z-Komponenten eines Vektoroperators, so sind die X' , Y' , Z' -Komponenten dieses Operators

$$\left. \begin{aligned} \mathbf{O}_R^{-1} \mathbf{V}_x \mathbf{O}_R &= R_{xx} \mathbf{V}_x + R_{xy} \mathbf{V}_y + R_{xz} \mathbf{V}_z, \\ \mathbf{O}_R^{-1} \mathbf{V}_y \mathbf{O}_R &= R_{yx} \mathbf{V}_x + R_{yy} \mathbf{V}_y + R_{yz} \mathbf{V}_z, \\ \mathbf{O}_R^{-1} \mathbf{V}_z \mathbf{O}_R &= R_{zx} \mathbf{V}_x + R_{zy} \mathbf{V}_y + R_{zz} \mathbf{V}_z. \end{aligned} \right\} \quad (16)$$

Die \mathbf{V}_x , \mathbf{V}_y , \mathbf{V}_z transformieren sich nicht wie \mathbf{S} nach $\mathfrak{D}^{(0)}$, d. h. sie bleiben bei dem Übergang zu einem neuen Achsenkreuz nicht ungeändert, sondern erleiden eine Transformation mit der Drehungsmatrix \mathbf{R} . Nun ist $\mathfrak{D}^{(1)}$ als Darstellung der Drehgruppe der Darstellung durch die Matrizen \mathbf{R} äquivalent und es empfiehlt sich, für die spätere Rechnung nicht die X, Y, Z-Komponenten, sondern die

$$\left. \begin{aligned} \mathbf{V}^{(-1)} &= \frac{i}{\sqrt{2}} \mathbf{V}_x + \frac{1}{\sqrt{2}} \mathbf{V}_y, \\ \mathbf{V}^{(0)} &= \mathbf{V}_z, \\ \mathbf{V}^{(1)} &= \frac{i}{\sqrt{2}} \mathbf{V}_x - \frac{1}{\sqrt{2}} \mathbf{V}_y \end{aligned} \right\} \quad (17)$$

Komponenten zu benutzen, für die nach (34), Kapitel XV, an Stelle von (16)

$$\mathbf{O}_R^{-1} \mathbf{V}^{(\varrho)} \mathbf{O}_R = \sum_{\sigma=-1}^1 \mathfrak{D}^{(1)}(R)_{\varrho \sigma} \mathbf{V}^{(\sigma)} \quad (16 \text{ a})$$

gilt. Man kann auch noch etwas allgemeiner einen irreduziblen Tensoroperator ω -ten Grades betrachten, der dadurch definiert ist, daß sich seine $2\omega + 1$ Komponenten $\mathbf{T}^{(\varrho)}$ bei einer Drehung des Achsenkreuzes nach

$$\mathbf{O}_R^{-1} \mathbf{T}^{(\varrho)} \mathbf{O}_R = \sum_{\sigma=-\omega}^{\omega} \mathfrak{D}^{(\omega)}(R)_{\varrho \sigma} \mathbf{T}^{(\sigma)} \quad (16 \text{ b})$$

transformieren. Setzt man in (16 b) R^{-1} an Stelle von R , so erhält man wegen $\mathbf{O}_{R^{-1}} = \mathbf{O}_R^{-1}$ und $\mathfrak{D}^{(\omega)}(R^{-1})_{\varrho \sigma} = \mathfrak{D}^{(\omega)}(R)_{\sigma \varrho}^*$

$$\mathbf{O}_R \mathbf{T}^{(\varrho)} \mathbf{O}_R^{-1} = \sum_{\sigma=-\omega}^{\omega} \mathfrak{D}^{(\omega)}(R)_{\sigma \varrho}^* \mathbf{T}^{(\sigma)}. \quad (16 \text{ c})$$

Aus diesen Gleichungen müssen wir die zu (14) analoge Gleichung für Vektor- und Tensoroperatoren ableiten. Um (16 c) einführen zu können, wenden wir auf beide Teile des skalaren Produkts

$$\mathbf{T}_{Nj\mu; N'j'\mu'}^{(\varrho)} = (\Psi_{\mu}^{Nj}, \mathbf{T}^{(\varrho)} \Psi_{\mu'}^{N'j'}) \quad (18)$$

den unitären Operator \mathbf{O}_R an und erhalten

$$\begin{aligned} \mathbf{T}_{Nj\mu; N'j'\mu'}^{(\varrho)} &= (\mathbf{O}_R \Psi_{\mu}^{Nj}, \mathbf{O}_R \mathbf{T}^{(\varrho)} \mathbf{O}_R^{-1} \mathbf{O}_R \Psi_{\mu'}^{N'j'}) \\ &= \sum_{\nu} \sum_{\sigma} \sum_{\nu'} \mathfrak{D}^{(j)}(R)_{\nu \mu}^* \mathfrak{D}^{(\omega)}(R)_{\sigma \varrho}^* \mathfrak{D}^{(j')}(R)_{\nu' \mu'} \mathbf{T}_{Nj\nu; N'j'\nu'}^{(\varrho)}. \end{aligned} \quad (18 \text{ a})$$

Die analoge Formel für skalare Operatoren haben wir über alle Drehungen integriert, die Orthogonalitätsrelationen ergaben dann direkt (14). Um die für (18 a) notwendigen Integrale über das Produkt dreier Darstellungskoeffizienten auszuwerten, können wir zuerst nach (16 b), Kap. XVII das Produkt der ersten beiden als eine Linearkombination

$$\mathfrak{D}^{(j)}(R)_{\nu \mu}^* \mathfrak{D}^{(\omega)}(R)_{\sigma \varrho}^* = \sum_{L=|j-\omega|}^{j+\omega} s_{L\nu\sigma}^{(j\omega)} \mathfrak{D}^{(L)}(R)_{\nu+\sigma, \mu+\varrho}^* s_{L\mu\varrho}^{(j\omega)}$$

schreiben. Setzt man dies in (18 a) ein, so erhält man wegen der Orthogonalitätsrelationen bei der Integration über alle Drehungen

$$\mathbf{T}_{Nj\mu; N'j'\mu'}^{(\varrho)} = \sum_{L=|j-\omega|}^{j+\omega} s_{L\mu\varrho}^{(j\omega)} \sum_{\nu \sigma \nu'} s_{L\nu\sigma}^{(j\omega)} \frac{\delta_{Lj'} \delta_{\nu+\sigma, \nu'} \delta_{\mu+\varrho, \mu'}}{2j'+1} \cdot \mathbf{T}_{Nj\nu; N'j'\nu'}^{(\varrho)},$$

wo noch mit $\int dR$ gekürzt ist. Dieser Ausdruck verschwindet, wenn j' nicht zwischen den Grenzen $|j - \omega|$ und $j + \omega$ liegt, und ist für $|j - \omega| \leq j' \leq j + \omega$ gleich

$$\mathbf{T}_{Nj\mu; N'j'\mu'}^{(0)} = s_{j'\mu\varrho}^{(j\omega)} \delta_{\mu+\varrho, \mu'} T_{Nj; N'j'}, \quad (19)$$

wo $T_{Nj; N'j'}$ von μ, μ' und ϱ nicht mehr abhängt.

Diese Formel ist von großer Allgemeinheit¹⁾. Sie gestattet das Verhältnis $\mathbf{T}_{Nj\mu; N'j'\mu'}^{(0)} / \mathbf{T}_{Nj\nu; N'j'\nu'}^{(0)}$ solcher „Matrixelemente“, d. h. skalarer Produkte (18) zahlenmäßig anzugeben, deren erste Faktoren (Ψ_μ^{Nj}) verschiedene Eigenfunktionen desselben Eigenwertes, deren Operatoren verschiedene Komponenten desselben irreduziblen Tensors und deren zweite Faktoren ($\Psi_\mu^{N'j'}$) ebenfalls Eigenfunktionen ein und desselben Eigenwertes sind.

Kehren wir zu den Vektoroperatoren zurück, indem wir in (19) $\omega = 1$ setzen, so erhalten wir mit Hilfe der Tabelle auf S. 208 die zu (14) analogen Formeln für Vektoroperatoren:

$$\left. \begin{aligned} \mathbf{V}_{Nj\mu; N'j-1\mu-1}^{(-1)} &= \sqrt{j+\mu} \sqrt{j+\mu-1} V'_{Nj; N'j-1}, \\ \mathbf{V}_{Nj\mu; N'j-1\mu}^{(0)} &= -\sqrt{j+\mu} \sqrt{j-\mu} \sqrt{2} V'_{Nj; N'j-1}, \\ \mathbf{V}_{Nj\mu; N'j-1\mu+1}^{(1)} &= \sqrt{j-\mu-1} \sqrt{j-\mu} V'_{Nj; N'j-1}. \end{aligned} \right\} (19a)$$

$$\left. \begin{aligned} \mathbf{V}_{Nj\mu; N'j\mu-1}^{(-1)} &= \sqrt{j-\mu+1} \sqrt{j+\mu} V'_{Nj; N'j}, \\ \mathbf{V}_{Nj\mu; N'j\mu}^{(0)} &= \mu \sqrt{2} V'_{Nj; N'j}, \\ \mathbf{V}_{Nj\mu; N'j\mu+1}^{(1)} &= -\sqrt{j+\mu+1} \sqrt{j-\mu} V'_{Nj; N'j}. \end{aligned} \right\} (19b)$$

$$\left. \begin{aligned} \mathbf{V}_{Nj\mu; N'j+1\mu-1}^{(-1)} &= \sqrt{j-\mu+1} \sqrt{j-\mu+2} V'_{Nj; N'j+1}, \\ \mathbf{V}_{Nj\mu; N'j+1\mu}^{(0)} &= \sqrt{j-\mu+1} \sqrt{j+\mu+1} \sqrt{2} V'_{Nj; N'j+1}, \\ \mathbf{V}_{Nj\mu; N'j+1\mu+1}^{(1)} &= \sqrt{j+\mu+1} \sqrt{j+\mu+2} V'_{Nj; N'j+1}. \end{aligned} \right\} (19c)$$

Alle hier nicht angeführten Matrixelemente eines Vektoroperators verschwinden. Auch der Koeffizient mit $j = j' = 0$ ist Null. Die $V_{Nj; N'j'}$ kann man natürlich aus allgemeinen Überlegungen nicht mehr bestimmen. Während die Matrixelemente eines skalaren Operators Null waren, wenn die Gesamtquantenzahl

¹⁾ In dieser Allgemeinheit stammen die Formeln von C. Eckart, Rev. Mod. Phys. 2, 305, 1930.

oder wenn die magnetische Quantenzahl von Zeile und Spalte (j und j' bzw. μ und μ') überhaupt verschieden war, können sich diese Zahlen bei einem Vektoroperator schon um 1 unterscheiden.

Bei der Ableitung von (19) wurde über die spezielle Form der Operatoren \mathbf{O}_R nichts vorausgesetzt und (19) muß daher auch in der Theorie, die den Spin nicht berücksichtigt, gelten, wenn man die \mathbf{O}_R durch die \mathbf{P}_R und die Gesamtquantenzahl j durch die azimutale l ersetzt. Wir haben auch schon öfters über das Verschwinden von Matrixelementen eines Vektoroperators Aussagen gemacht. Z. B. sind das Multiplizieren mit

$$x_1 + x_2 + \cdots + x_n, \quad y_1 + y_2 + \cdots + y_n, \quad z_1 + z_2 + \cdots + z_n,$$

die drei Komponenten eines Vektoroperators. In der Tat fanden wir, daß die Übergangswahrscheinlichkeit durch Strahlung vom Zustand ψ_F in den Zustand ψ_E , für die

$$(\psi_F, (x_1 + x_2 + \cdots + x_n) \psi_E) \text{ usw.}$$

maßgebend ist, verschwindet, wenn die Differenz der azimutalen Quantenzahlen von ψ_F und ψ_E nicht 0 oder ± 1 ist. Später sahen wir, daß die magnetische Quantenzahl unverändert bleibt, wenn das Licht in der Z -Richtung polarisiert ist ($\varrho = 0$) und sich mit ± 1 ändert, wenn das Licht in der X - oder Y -Richtung polarisiert ist.

Der durch das Magnetfeld \mathfrak{H}_z in der Z -Richtung bedingte Zusatzoperator in der Schrödinger-Gleichung

$$\begin{aligned} \mathbf{V}_z = \mathbf{V}^{(0)} = -\frac{e\hbar\mathfrak{H}_z}{4\pi m c i} & \left[-y_1 \frac{\partial}{\partial x_1} - y_2 \frac{\partial}{\partial x_2} - \cdots - y_n \frac{\partial}{\partial x_n} \right. \\ & \left. + x_1 \frac{\partial}{\partial y_1} + x_2 \frac{\partial}{\partial y_2} + \cdots + x_n \frac{\partial}{\partial y_n} \right] \end{aligned}$$

ist die Z -Komponente eines Vektoroperators. Hier haben wir die Matrixelemente $V_{Nl\mu; Nl\mu}$ auch wirklich berechnet. Aus der mittleren Gleichung von (19b) sehen wir, daß sie proportional zur magnetischen Quantenzahl μ sein müssen. Dies haben wir auch gefunden, die Proportionalitätskonstante war in diesem Fall unabhängig von N und l gleich $e\hbar\mathfrak{H}_z/4\pi m c$.

Für die Theorie ohne Spin ist (19) gleichsam der Ausdruck der Formel (6) des XIX. Kapitels, worin die Abhängigkeit der Eigenfunktionen von der Lage des Konfigurations- n -Beines explizite bestimmt ist.

6. Es ist interessant zu bemerken¹⁾, daß schon aus der ganz allgemeinen Gleichung (9) die Existenz einer Gesamtquantenzahl folgt. Natürlich kann man aus (9) nichts über die Ganzzahligkeit oder Halbzahligkeit von j erfahren, in (9) geht ja die Elektronenzahl gar nicht ein.

¹⁾ Der Rest dieser Kapitels ist für die folgenden nicht notwendig.

Benutzt man an Stelle von (10a) oder (10b) die Gleichung (9) $\mathbf{O}_{SR} = c_{S,R} \mathbf{O}_S \mathbf{O}_R$, so erhält man an Stelle von (13), daß die in (12) definierten Matrizen $\mathbf{D}(R)$

$$\mathbf{D}(SR) = c_{S,R} \mathbf{D}(S) \mathbf{D}(R) \quad (13a)$$

bloß eine Darstellung bis auf einen Faktor der Drehgruppe bilden. Der Einheit E entspricht dabei noch die Einheitsmatrix. Es soll nun gezeigt werden, daß man aus den Matrizen, die (13a) erfüllen, durch Multiplikation einer jeden Matrix $\mathbf{D}(R)$ mit einer zweckmäßig gewählten Zahl c_R ein Matrzensystem $\bar{\mathbf{D}}(R) = c_R \mathbf{D}(R)$ erhalten kann, das eine Darstellung der unitären Gruppe bildet, und für das

$$\bar{\mathbf{D}}(S) \bar{\mathbf{D}}(R) = \pm \bar{\mathbf{D}}(SR) \quad (20)$$

gilt. Man kann daher nach Kapitel XV dieses Matrzensystem durch eine Ähnlichkeitstransformation in die Darstellungen $\mathfrak{D}^{(0)}$, $\mathfrak{D}^{(1/2)}$, $\mathfrak{D}^{(1)}$, $\mathfrak{D}^{(3/2)}$, ... zerlegen. Dies bedeutet, daß die Darstellungen bis auf einen Faktor, auf die (9) zunächst führt, bei der Drehgruppe im wesentlichen die zweideutigen Darstellungen des XV. Kapitels sind.

Eine zweidimensionale Darstellung muß demnach entweder lauter konstante Matrizen enthalten (wenn sie $\mathfrak{D}^{(0)}$ zweimal enthält) oder $\mathfrak{D}^{(1/2)}$ äquivalent sein, wie wir es im vorangegangenen Kapitel vorausgesetzt haben.

Wir bilden aus den $\mathbf{D}(R)$ zuerst $\mathbf{D}(R) = c_R \mathbf{D}(R)$ und wählen c_R gleich der $-1/\lambda$ -ten Potenz der Determinante von $\mathbf{D}(R)$, wo λ die Dimension der $\mathbf{D}(R)$ ist. so daß die Determinante $|\bar{\mathbf{D}}(R)| = 1$ wird:

$$\begin{aligned} |\bar{\mathbf{D}}(R)| &= |c_R \mathbf{1} \cdot \mathbf{D}(R)| = |c_R \cdot \mathbf{1}| \cdot |\mathbf{D}(R)| \\ &= c_R^\lambda \cdot |\mathbf{D}(R)| = 1. \end{aligned} \quad (21)$$

Hierdurch sind c_R und $\bar{\mathbf{D}}(R)$ noch nicht eindeutig, sondern wegen der Mehrdeutigkeit der λ -ten Wurzel nur bis auf eine λ -te Einheitswurzel ω bestimmt, so daß einem Gruppenelement R nicht mehr eine, sondern zunächst λ Matrizen zugeordnet sind, nämlich alle Vielfache von $\mathbf{D}(R)$, deren Determinante 1 ist.

Multipliziert man ein $\bar{\mathbf{D}}(S)$ mit einem $\bar{\mathbf{D}}(R)$, so erhält man ein $\bar{\mathbf{D}}(SR)$: das Produkt ist wegen (13a) jedenfalls ein Vielfaches von $\bar{\mathbf{D}}(SR)$, und seine Determinante ist das Produkt der Determinanten von $\bar{\mathbf{D}}(S)$ und $\bar{\mathbf{D}}(R)$, also 1.

Wir haben aus der Darstellung bis auf einen Faktor $\mathbf{D}(R)$ eine mehrdeutige und zwar λ -deutige Darstellung gewonnen, das Produkt eines jeden $\bar{\mathbf{D}}(S)$ mit einem jeden $\bar{\mathbf{D}}(R)$ gibt ein $\bar{\mathbf{D}}(SR)$.

Man kann diese Vieldeutigkeit der Darstellung dadurch zu verringern versuchen, daß man von den λ Matrizen $\bar{\mathbf{D}}(R)$ einige herauswählt und beibehält, die anderen einfach wegläßt. Dies kann natürlich nicht ganz willkürlich geschehen, sondern nur so, daß eine beliebige der behaltenen Matrizen $\bar{\mathbf{D}}(S)$ mit einer beliebigen, ebenfalls behaltenen Matrix $\bar{\mathbf{D}}(R)$ multipliziert eine ebenfalls behaltene Matrix $\bar{\mathbf{D}}(SR)$ gibt.

Bei kontinuierlichen Gruppen kann man nach H. Weyl¹⁾ auf Grund der Stetigkeitseigenschaft der Darstellungen diese Auswahl folgendermaßen treffen.

Sind S und S' zwei benachbarte Gruppenelemente, $S \sim S'$, so waren auch

$$\mathbf{D}(S) \sim \mathbf{D}(S') \quad \text{und} \quad |\mathbf{D}(S)| \sim |\mathbf{D}(S')|.$$

Aus der letzteren Gleichung folgt, daß die λ Werte von c_S mit den λ Werten von $c_{S'}$ paarweise benachbart sind. Auch die λ Matrizen $\bar{\mathbf{D}}(S)$ sind mit den λ Matrizen $\bar{\mathbf{D}}(S')$ paarweise benachbart, in der Weise, daß einem $\bar{\mathbf{D}}(S)$ ein und nur ein $\bar{\mathbf{D}}(S')$ benachbart ist, die anderen $\lambda - 1$ sich dagegen davon wesentlich unterscheiden, weil sie durch Multiplikation mit einer von 1 wesentlich verschiedenen Zahl (einer λ -ten Einheitswurzel) daraus hervorgehen.

Verbinden wir nun im Parameterraume die Einheit $E = S(0)$ durch eine stetige Linie $S(t)$ mit dem Element $S = S(1)$, so können wir von den Matrizen $\bar{\mathbf{D}}(S(t))$ fordern, daß auch sie eine stetige Folge durchlaufen sollen. Dann entsteht längs eines bestimmten Weges $S(t)$ aus $\bar{\mathbf{D}}(S(0)) = \bar{\mathbf{D}}(E) = 1$ nur eine bestimmte der λ Matrizen $\bar{\mathbf{D}}(S)$, die wir mit $\bar{\mathbf{D}}(S)_{S(t)}$ bezeichnen wollen. Deformiert man den Weg $S(t)$ stetig bei festgehaltenem Anfang und Ende, so ändert sich $\bar{\mathbf{D}}(S)_{S(t)}$ gar nicht, da es sich bei einer stetigen Deformation des Weges nur stetig ändern könnte, der Übergang zu einem anderen $\bar{\mathbf{D}}(S)$ aber notwendig einen Sprung bedeuten würde.

¹⁾ H. Weyl, Mathem. Zeitschr. **28**, 271; **24**, 328, 377, 789, 1925;
V. Schreier, Abhandl. mathem. Sem. Hamburg **4**, 15, 1926; **5**, 233, 1927.

Das Produkt $\bar{\mathbf{D}}(S)_{S(t)} \bar{\mathbf{D}}(R)_{R(t)}$ ist eine der Matrizen $\bar{\mathbf{D}}(SR)$, die ebenfalls stetig aus $\bar{\mathbf{D}}(E) = 1$ entsteht. Der zugehörige Weg führt zuerst von E über $S(t)$ nach S — dabei geht $\bar{\mathbf{D}}(E) = 1$ stetig in $\bar{\mathbf{D}}(S)_{S(t)}$ über —, dann geht der Weg über die Punkte $S.R(t)$ nach SR — dabei geht $\bar{\mathbf{D}}(S)_{S(t)} = \bar{\mathbf{D}}(S)_{S(t)} \cdot 1$ über die Matrizen $\bar{\mathbf{D}}(S)_{S(t)} \bar{\mathbf{D}}(R(t))$ stetig in $\bar{\mathbf{D}}(S)_{S(t)} \bar{\mathbf{D}}(R)_{R(t)}$ über:

$$\bar{\mathbf{D}}(S)_{S(t)} \bar{\mathbf{D}}(R)_{R(t)} = \bar{\mathbf{D}}(SR)_{S(t), S.R(t)} \quad (22)$$

Lassen sich alle Wege von E nach S stetig ineinander deformieren, ist der Parameterraum einfach zusammenhängend, so existiert nur ein einziges $\bar{\mathbf{D}}(S) = \bar{\mathbf{D}}(S)_{S(t)}$, das stetig von $\bar{\mathbf{D}}(E) = 1 = \bar{\mathbf{D}}(E)$ aus erreicht werden kann. Diese $\bar{\mathbf{D}}(S)$ bilden dann eine eindeutige Darstellung der Gruppe.

Ist der Parameterraum mehrfach zusammenhängend, existieren zwei oder mehrere Wege $S_1(t), S_2(t), \dots$, die man nicht stetig ineinander deformieren kann, so können die entsprechenden Matrizen $\bar{\mathbf{D}}(S)_{S_1(t)}, \bar{\mathbf{D}}(S)_{S_2(t)}, \dots$ auch voneinander verschieden sein. Die Darstellung ist so vieldeutig, wie nicht ineinander deformierbare Wege von E nach S führen.

7. Die Gleichung (22) legt in diesem Fall den Gedanken nahe, an Stelle der ursprünglichen Gruppe eine „Überlagerungsgruppe“ zu betrachten, die für jedes Element S der Gruppe so viele Elemente $S_{S_1(t)}, S_{S_2(t)}, \dots$ hat, wie nicht ineinander deformierbare Wege $S_1(t), S_2(t), \dots$ von E nach S führen. Die Multiplikationsregel der Überlagerungsgruppe ist

$$S_{S_t(t)} R_{R_k(t)} = S R_{S_t(t), S.R_k(t)}. \quad (22a)$$

Ordnet man dem Element $S_{S_t(t)}$ der Überlagerungsgruppe die Matrix $\bar{\mathbf{D}}(S)_{S_t(t)}$ zu, so bilden diese nach (22) eine regelrechte, eindeutige Darstellung der Überlagerungsgruppe. Die Darstellungen bis auf einen Faktor einer kontinuierlichen Gruppe lassen sich durch Multiplikation mit entsprechend gewählten Zahlen in regelrechte Darstellungen der Überlagerungsgruppe verwandeln. Kennt man alle Darstellungen der Überlagerungsgruppe, so kennt man auch alle Darstellungen bis auf einen Faktor der ursprünglichen Gruppe.

Der Parameterraum (Abb. 1, S. 99) der dreidimensionalen reinen Drehgruppe ist zweifach zusammenhängend. Man kann von E zu einem beliebigen Punkt entweder (vgl. Abb. 11) direkt (I), oder über

einen Sprung zum Antipoden (II) gelangen, und diese beiden Wege lassen sich nicht ineinander deformieren. (Ein Sprung zum Antipoden ist nicht als eine Unterbrechung einer stetigen Linie im Parameterraume der dreidimensionalen Drehgruppe anzusehen, da die Antipoden derselben Drehung entsprechen.) Ein Weg mit zwei Sprüngen zum Antipoden lässt sich dagegen schon stetig in einen Weg ohne Sprung verwandeln (vgl. Abb. 12).

Die Überlagerungsgruppe hat demnach doppelt so viele Elemente wie die Drehgruppe und ist folglich der Gruppe der $\mathfrak{D}^{(1/2)}(R)$ holomorph. Diese ist ja als zweideutige Darstellung der Drehgruppe sicher eine regelrechte Darstellung der Überlagerungsgruppe und zwar eine treue Darstellung, da sie jeder Drehung R zwei voneinander und von allen anderen $\mathfrak{D}^{(1/2)}(S)$ verschiedene Matrizen $\pm \mathfrak{D}^{(1/2)}(R)$ zuordnet.

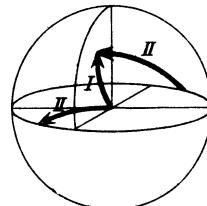


Abb. 11

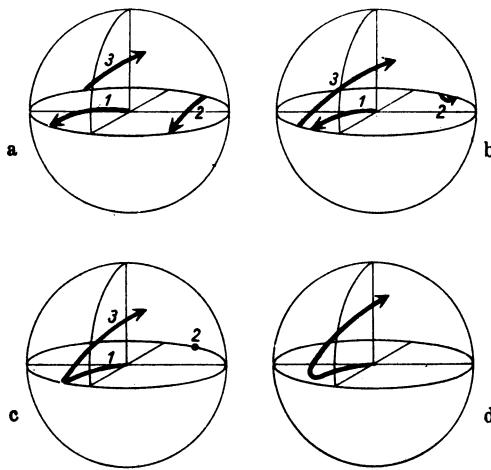


Abb. 12

Die $\mathfrak{D}^{(1/2)}(R)$ bilden die unitäre Gruppe, diese ist also die Überlagerungsgruppe der dreidimensionalen Drehgruppe. Ihre Darstellungen lassen sich in $\mathfrak{U}^{(0)}$, $\mathfrak{U}^{(1/2)}$, $\mathfrak{U}^{(1)}$, ... zerlegen und dementsprechend lässt sich $\bar{\mathcal{D}}(R) = c_R \mathcal{D}(R)$ in $\pm \mathfrak{D}^{(0)}$, $\pm \mathfrak{D}^{(1/2)}$, $\pm \mathfrak{D}^{(1)}$, ... zerlegen. Wenn man $\bar{\mathcal{D}}(R)$ in der ausreduzierten Form

annimmt — dies bedeutet ja nur einen Übergang zu einem neuen System linear unabhängiger Eigenfunktionen —, erhält man für diese

$$\mathbf{O}_R \Psi_{\mu}^{(j)} = \frac{1}{c_R} \sum_{\mu'} \mathfrak{D}^{(j)}(R)_{\mu' \mu} \Psi_{\mu'}^{(j)}. \quad (12b)$$

und man kann j die Gesamtquantenzahl der Eigenfunktionen $\Psi_{-j}^{(j)}$, $\Psi_{-j+1}^{(j)}, \dots, \Psi_{j-1}^{(j)}, \Psi_j^{(j)}$ nennen

Obwohl (12b) mit (12a) nicht ganz äquivalent ist, genügt es schon zur Ableitung der meisten Regeln für die Gesamtquantenzahl.

XXII. Die Feinstruktur der Spektrallinien

1. Im Kapitel XVIII haben wir die Auswahlregeln für Azimutalquantenzahl, Spiegelungscharakter und Multiplettsystem abgeleitet, wie sie in einer Theorie, die den Spin unberücksichtigt läßt, gelten. Wenn man die Kräfte, die von den magnetischen Momenten der Elektronen herrühren, berücksichtigt, können diese Regeln nicht mehr als streng richtig angesehen werden, weil bei ihrer Ableitung die Voraussetzung gemacht worden ist, daß mit Ψ auch $\mathbf{P}_R \Psi$ eine Eigenfunktion des Energieoperators zum Eigenwert von Ψ ist, da $\mathbf{P}_R \Psi$ mit dem Zustand Ψ — bis auf eine Drehung — identisch ist.

Nun wissen wir, daß, wenn man auch den Spin berücksichtigt, nicht \mathbf{P}_R , sondern \mathbf{O}_R die Drehung des Zustandes bewirkt, \mathbf{P}_R verdreht gewissermaßen nur die Cartesischen Koordinaten des Systems. Demgemäß wäre $\mathbf{P}_R \Psi$ nur dann eine Eigenfunktion von \mathbf{H} , wenn \mathbf{H} eine „spinfreie Größe“ wäre. In Wirklichkeit treten in \mathbf{H} auch Glieder auf, die vom Spin herrühren, und $\mathbf{P}_R \Psi$ ist keine Eigenfunktion von \mathbf{H} zum Eigenwert von Ψ und läßt sich daher auch nicht als eine Linearkombination der zu diesem Eigenwert gehörigen Eigenfunktionen schreiben. Die Eigenfunktionen gehören daher in bezug auf die \mathbf{P}_R auch zu keiner Darstellung der Drehgruppe, und der Begriff der azimutalen Quantenzahl ist, wenn man auch den Spin berücksichtigt, strenggenommen sinnlos. Nur solange die vom Spin herrührenden Glieder klein sind und die Schrödingergleichung näherungsweise auch bei ihrer Vernachlässigung aufgelöst werden kann — dies ist allerdings die Regel —, hat der Begriff der azimutalen Quantenzahl (und des Multiplettsystems) einen Sinn, nur

dann gelten auch die Auswahlregeln des XVIII. Kapitels. Dies soll im folgenden noch genauer ausgeführt werden.

Man kann natürlich mit Hilfe der Operatoren \mathbf{O}_R , die einer wirklichen Drehung entsprechen, des Operators \mathbf{O}_I der Inversion und der Operatoren \mathbf{O}_P , die Vertauschungen aller vier Koordinaten zweier oder mehrerer Elektronen entsprechen:

$$\begin{aligned} \mathbf{O}_P \Psi(x_{\alpha_1} y_{\alpha_1} z_{\alpha_1} s_{\alpha_1} \dots x_{\alpha_n} y_{\alpha_n} z_{\alpha_n} s_{\alpha_n}) \\ = \Psi(x_1 y_1 z_1 s_1 \dots x_n y_n z_n s_n) \end{aligned} \quad (1)$$

[P ist die Permutation $\begin{pmatrix} 1 & 2 & \dots & n \\ \alpha_1 & \alpha_2 & \dots & \alpha_n \end{pmatrix}$], genau dieselben Überlegungen ausführen, die wir im Kapitel XVIII mit Hilfe der Operatoren \mathbf{P} ausgeführt haben. Die so erhaltenen Sätze werden natürlich auch bei Miterücksichtigung der Spinglieder gelten. Die Operatoren \mathbf{O}_R , die Drehungen entsprechen, sind mit den Operatoren \mathbf{O}_P , die Vertauschungen der Elektronen entsprechen, und beide mit \mathbf{O}_I vertauschbar, so daß die gesamte Symmetrie, genau wie im Kapitel XVIII, das direkte Produkt der Gruppe der \mathbf{O}_R , der symmetrischen und der Spiegelungsgruppe ist.

Will man eine Vertauschung sowohl der Cartesischen wie auch der Spinkoordinaten der Elektronen ausführen, so kann man die Vertauschung der Cartesischen Koordinaten und der Spinkoordinaten gesondert vornehmen, \mathbf{O}_P läßt sich (ähnlich wie \mathbf{O}_R) in zwei Faktoren

$$\mathbf{O}_P = \mathbf{P}_P \mathbf{Q}_P = \mathbf{Q}_P \mathbf{P}_P \quad (2)$$

zerlegen, wobei \mathbf{Q}_P nur auf die Spinkoordinaten

$$\mathbf{Q}_P \Psi(x_1 y_1 z_1 s_{\alpha_1} \dots x_n y_n z_n s_{\alpha_n}) = \Psi(x_1 y_1 z_1 s_1 \dots x_n y_n z_n s_n) \quad (2a)$$

und \mathbf{P}_P nur auf die Cartesischen Koordinaten einwirkt:

$$\begin{aligned} \mathbf{P}_P \Phi(x_{\alpha_1} y_{\alpha_1} z_{\alpha_1} \sigma_1 \dots x_{\alpha_n} y_{\alpha_n} z_{\alpha_n} \sigma_n) \\ = \Phi(x_1 y_1 z_1 \sigma_1 \dots x_n y_n z_n \sigma_n). \end{aligned} \quad (2b)$$

Setzt man nämlich in (2b) $\mathbf{Q}_P \Psi$ für Φ ein, so erhält man

$$\mathbf{P}_P \mathbf{Q}_P \Psi(\dots x_{\alpha_k} y_{\alpha_k} z_{\alpha_k} \sigma_k \dots) = \mathbf{Q}_P \Psi(\dots x_k y_k z_k \sigma_k \dots),$$

und wenn man hierin noch $\sigma_k = s_{\alpha_k}$ einführt, so ergibt es mit Hilfe von (2a) und (1) direkt (2).

2. Es ist eine wesentliche Vereinfachung für das Folgende, daß die Eigenfunktionen aller wirklich existierenden Zustände den \mathbf{O}_P gegenüber zur antisymmetrischen Darstellung gehören:

$$\mathbf{O}_P \Phi = \varepsilon_P \Phi; \quad (\mathbf{O}_P - \varepsilon_P) \Phi = 0, \quad (3)$$

wo ε_P gleich +1 oder -1 ist, je nachdem P eine gerade oder ungerade Permutation ist. Man nennt Funktionen, die (3) befriedigen, kurz antisymmetrische Funktionen; die Aussage, daß alle Wellenfunktionen antisymmetrisch sind, ist der Inhalt des Pauliprinzips¹⁾.

Das Pauliprinzip ist natürlich keineswegs eine Folge der bisher eingeführten Prinzipien der Quantenmechanik; gegenüber der zeitabhängigen Schrödinger-Gleichung, die einer Bewegungsgleichung entspricht, spielt es etwa die Rolle einer Anfangsbedingung, die in jedem System erfüllt ist. Wenn aber zu irgendeiner Zeit (3) erfüllt ist, so bleibt es immer erfüllt; aus

$$\frac{h}{2\pi i} \frac{\partial \Phi}{\partial t} = \mathbf{H}\Phi$$

folgt, da \mathbf{H} ein den \mathbf{O}_P gegenüber symmetrischer Operator und mit ihnen vertauschbar ist:

$$\begin{aligned} \frac{h}{2\pi i} \frac{\partial}{\partial t} (\mathbf{O}_P - \varepsilon_P) \Phi &= (\mathbf{O}_P - \varepsilon_P) \frac{h}{2\pi i} \frac{\partial \Phi}{\partial t} \\ &= (\mathbf{O}_P - \varepsilon_P) \mathbf{H}\Phi = \mathbf{H}(\mathbf{O}_P - \varepsilon_P) \Phi. \end{aligned} \quad (3a)$$

Hieraus folgt aber, daß $(\mathbf{O}_P - \varepsilon_P) \Phi$ immer verschwindet, wenn es zu irgendeiner Zeit Null war. Aus (3a) schließt man, daß das skalare Produkt

$$((\mathbf{O}_P - \varepsilon_P) \Phi, (\mathbf{O}_P - \varepsilon_P) \Phi) \quad (\dagger)$$

zeitlich konstant ist und immer Null bleibt, wenn es einmal Null war. Aus dem Verschwinden von (†) folgt aber auch das Verschwinden von $(\mathbf{O}_P - \varepsilon_P) \Phi$. Man sieht, daß das Pauliprinzip mit der quantenmechanischen Bewegungsgleichung jedenfalls verträglich ist.

Eine wichtige Folge davon, daß alle Wellenfunktionen zur antisymmetrischen Darstellung gehören, tritt bei einer Zerteilung eines Systems in mehrere Teilsysteme zutage. Man denke etwa an ein System aus zwei He-Atomen, die zunächst in Wechselwirkung waren, und die man nachher voneinander entfernt hat. Die Wellenfunktion gehöre vor der Trennung zur irreduziblen Darstellung $\mathbf{D}(P)$ der symmetrischen Gruppe vierten Grades. Man kann dann die Frage aufwerfen, zu welchen Darstellungen der symmetrischen Gruppe zweiten Grades der Zustand des einen He-Atoms nach der Trennung gehören kann. Ist $\mathbf{D}(P)$ die antisymmetrische Darstellung, so ist der Zustand beider He-Atome nach der Trennung auch sicher antisymmetrisch. Die Zugehörigkeit der Teilsysteme zu

¹⁾ Vgl. W. Heisenberg, ZS. f. Phys. 38, 411, 1926 und P. A. M. Dirac, Proc. Roy. Soc. 112, 661, 1926.

einer bestimmten Darstellung ist durch die Zugehörigkeit des Gesamtsystems zur antisymmetrischen Darstellung eindeutig vorgegeben. Ähnliches gilt, wenn $D(P)$ die symmetrische (identische) Darstellung ist, aber bei keiner weiteren Darstellung.

Warum alle Wellenfunktionen antisymmetrisch sind und nicht etwa alle symmetrisch, läßt sich natürlich auf Grund allgemeiner Überlegungen nicht sagen und muß als eine Erfahrungstatsache angesehen werden.

3. Im folgenden denken wir uns den Hamiltonschen Operator in zwei Teile zerlegt:

$$\mathbf{H} = \mathbf{H}_0 + \mathbf{H}_1. \quad (4)$$

Der erste Teil ist der gewöhnliche Schrödingersche Operator, er trägt nur der Wechselwirkung der Ladungen und der kinetischen Energie der Elektronen Rechnung

$$\mathbf{H}_0 = -\frac{\hbar^2}{4\pi^2} \sum_k \frac{1}{2m_k} \left(\frac{\partial^2}{\partial x_k^2} + \frac{\partial^2}{\partial y_k^2} + \frac{\partial^2}{\partial z_k^2} \right) + V(x_1 \dots z_n) \quad (4a)$$

Er ist ein spinfreier Operator. Der zweite Teil \mathbf{H}_1 enthält die Energie der magnetischen Momente der Elektronen, wird im folgenden im Verhältnis zu \mathbf{H}_0 als klein angenommen und als eine „Störung“ behandelt. Diese Störung ist die Ursache der Feinstruktur, sie bewirkt das Aufspalten der Eigenwerte der einfachen Schrödinger-Gleichung (4a), der „Grobstrukturterme“ in mehrere Feinstrukturkomponenten.

Vor der Anwendung der Störungstheorie muß man immer, wenn mehrere Eigenfunktionen zu einem Eigenwert des ungestörten Problems gehören (was bei \mathbf{H}_0 immer der Fall ist), die sogenannten richtigen Linearkombinationen bestimmen. Dies wird die Hauptaufgabe dieses Kapitels sein. Von den richtigen Linearkombinationen kann man annehmen, daß sie in bezug auf die O_P zu einer irreduziblen Darstellung der symmetrischen Gruppe gehören — wir brauchen wegen des Pauliprinzipes nur jene, die antisymmetrisch sind. Wenn die Isotropie des Raumes nicht gestört ist, kann man außerdem annehmen, daß sie in bezug auf die O_R zu einer Zeile einer irreduziblen Darstellung $\mathfrak{D}^{(J)}$ der Drehgruppe gehören. In diesem Fall zeigt es sich, daß man aus den Eigenfunktionen eines Eigenwerts von \mathbf{H}_0 nur eine antisymmetrische Linearkombination bilden kann, die zu einer bestimmten Zeile einer Darstellung $\mathfrak{D}^{(J)}$ gehört, so daß die richtigen Linearkombinationen für das Störungsverfahren allgemein bestimmt werden können.

Sei E ein Eigenwert von \mathbf{H}_0 und $\psi(x_1 y_1 z_1 \dots x_n y_n z_n)$ eine zugehörige Eigenfunktion, eine Funktion der Cartesischen Koordinaten allein, wie wir sie immer betrachtet haben. Wir erhalten eine Eigenfunktion von \mathbf{H}_0 zum Eigenwert E , die Funktion aller Koordinaten $x_1 y_1 z_1 s_1 \dots x_n y_n z_n s_n$ ist, indem wir ψ mit einer beliebigen Funktion $f(s_1 \dots s_n)$ der Spinkoordinaten multiplizieren. Da \mathbf{H}_0 ein spinfreier Operator ist, kann man $f(s_1 \dots s_n)$ aus \mathbf{H}_0 wie einen konstanten Faktor herausziehen:

$$\mathbf{H}_0 \psi f = f \mathbf{H}_0 \psi = f E \psi = E \psi f. \quad (5)$$

Nun existieren im ganzen 2^n linear unabhängige Funktionen der $s_1 \dots s_n$, etwa

$$f_{\sigma_1 \sigma_2 \dots \sigma_n} = \delta_{s_1 \sigma_1} \delta_{s_2 \sigma_2} \dots \delta_{s_n \sigma_n} (\sigma_1 = \pm 1, \sigma_2 = \pm 1, \dots, \sigma_n = \pm 1), \quad (6)$$

durch die man alle Funktionen von $s_1 \dots s_n$ linear ausdrücken kann, wie wir dies schon im XIII. Kapitel bei der Bestimmung irreduzibler Darstellungen der symmetrischen Gruppe gesehen haben. Ist daher

$$f_1, f_2, \dots, f_{2^n} \quad (6a)$$

ein vollständiges Orthogonalsystem der Funktionen der s [etwa die Funktionen (6) selber], so kann man aus ψ folgende 2^n Eigenfunktionen von \mathbf{H}_0 bilden:

$$\psi f_1, \psi f_2, \dots, \psi f_{2^n}. \quad (7)$$

Sind mehrere Funktionen von $x_1 y_1 z_1 \dots x_n y_n z_n$ Eigenfunktionen von \mathbf{H}_0 zum Eigenwert E , so kann man nach (7) aus jeder 2^n linear unabhängige Eigenfunktionen bilden, die auch die Spinkoordinaten als Variable enthalten. Durch die Einführung der Spinkoordinate wird die Zähligkeit der Eigenwerte der spinfreien Operatoren 2^n -mal vergrößert. Dies entspricht dem Umstand, daß $\mathbf{H}_0 \Psi = E \Psi$ nur die Bewegung der Cartesischen Koordinaten vorschreibt, die n Spinkoordinaten können noch alle etwa zwischen der $+Z$ und der $-Z$ -Richtung wählen.

4. Betrachten wir zuerst ein System, das außer der Gleichheit der Elektronen keine weitere Symmetrie aufweist, die räumliche Symmetrie sei durch äußere Felder aufgehoben. Es kann angenommen werden, daß die Funktionen ψ_1, ψ_2, \dots von $x_1 y_1 z_1 \dots x_n y_n z_n$, die Eigenfunktionen eines bestimmten Eigenwerts von \mathbf{H}_0 sind, zu einer irreduziblen Darstellung der symmetrischen Gruppe n ten Grades gehören

$$\mathbf{P}_P \psi_x = \sum_{x'} \mathbf{D}(P)_{x' x} \psi_{x'} \quad (8)$$

Diese Gleichungen gelten auch, wenn man die ψ_x durch $\psi_x f_\lambda$ ersetzt, weil auch bei der Anwendung der \mathbf{P} eine Funktion der Spinkoordinaten wie ein konstanter Faktor herausgezogen werden kann.

Die Eigenfunktionen $\psi_x f_\lambda$ von \mathbf{H}_0 müssen auch in bezug auf die \mathbf{O}_P zu einer Darstellung gehören, da die Elektronen auch bei Mitberücksichtigung des Spins gleichwertig sind. Der Zustand $\mathbf{O}_P \psi_x f_\lambda$, in dem im Verhältnis zu $\psi_x f_\lambda$ die Elektronen lediglich ihre Rollen vertauscht haben, muß auch eine Eigenfunktion von \mathbf{H}_0 zum Eigenwert der $\psi_x f_\lambda$ sein und muß daher durch die $\psi_x' f_{\lambda'}$ linear ausgedrückt werden können. In der Tat kann man in

$$\mathbf{O}_P \psi_x f_\lambda = \mathbf{P}_P \mathbf{Q}_P \psi_x f_\lambda = \mathbf{P}_P \psi_x \cdot \mathbf{Q}_P f_\lambda \quad (9)$$

für die $\mathbf{P}_P \psi_x$ die Ausdrücke (8) einführen und auch die $\mathbf{Q}_P f_\lambda$ kann man durch die $f_{\lambda'}$ ausdrücken, wie man ja jede Funktion der s durch die $f_{\lambda'}$ ausdrücken kann. Um aber hierbei ein möglichst einfaches Koeffizientensystem zu haben, wird man von einem solchen Orthogonalsystem (6a) der s ausgehen, dessen Funktionen in bezug auf die Operatoren \mathbf{Q}_P zu irreduziblen Darstellungen der symmetrischen Gruppe gehören.

5. Ein solches Orthogonalsystem von Funktionen der s haben wir im Kapitel XIII bestimmt. Wir gingen dabei an Stelle von (6) von dem Orthogonalsystem

$$s_1^{\gamma_1} s_2^{\gamma_2} \dots s_n^{\gamma_n} \quad (\gamma_1, \gamma_2, \dots, \gamma_n = 0 \text{ oder } 1) \quad (6b)$$

aus und ordneten diese Funktionen zunächst so, daß in eine Zeile lauter Funktionen gleichen Grades kamen. In die k -te Zeile kamen lauter Funktionen k -ten Grades ($\gamma_1 + \gamma_2 + \dots + \gamma_n = k$), im ganzen $\binom{n}{k}$ Funktionen. Es wurde dann gezeigt, daß man aus den Funktionen k -ten Grades solche Linearkombinationen bilden kann, die für $k \leq \frac{1}{2} n$ zu je einer Zeile der Darstellungen

$$\mathbf{D}^{(0)}, \mathbf{D}^{(1)}, \mathbf{D}^{(2)}, \dots, \mathbf{D}^{(k)} \quad (\dagger)$$

gehören. Da die Dimension von $\mathbf{D}^{(i)}$

$$l_i = \binom{n}{i} - \binom{n}{i-1} \quad (10)$$

ist, sind dies im ganzen eben $l_0 + l_1 + l_2 + \dots + l_k = \binom{n}{k}$ Funktionen. Für $k \geq \frac{1}{2} n$ traten (vgl. die Tabellen auf S. 144) an Stelle der Darstellungen (\dagger) die Darstellungen

$$\mathbf{D}^{(0)}, \mathbf{D}^{(1)}, \mathbf{D}^{(2)}, \dots, \mathbf{D}^{(n-k)}. \quad (\ddagger\ddagger)$$

Wir bezeichnen die Funktion k -ten Grades, die zur λ -Zeile von $\mathbf{D}^{(i)}$ gehört, mit $g_{\lambda k}^{(i)}$, dann ist

$$\mathbf{Q}_P g_{\lambda k}^{(i)} = \sum_{\lambda' = 1}^k \mathbf{D}^{(i)}(P)_{\lambda' \lambda} g_{\lambda' k}^{(i)} \quad (i = 0, 1, 2, \dots, k \text{ bzw. } n - k). \quad (11)$$

Das Funktionensystem (6 b) wurde lediglich darum an Stelle der Funktionen (6) benutzt, weil man in den Funktionen (6 b) die Faktoren s_ϱ^γ , deren $\gamma_\varrho = 0$ ist, einfach weglassen kann, wodurch sich die Formeln etwas vereinfacht haben. Wir wollen aber jetzt wieder zu den Funktionen (6) zurückkehren, überall $s_\varrho^0 = 1$ durch $\delta_{s_\varrho, -1}$ und $s_\varrho^1 = s_\varrho$ durch $\delta_{s_\varrho, 1}$, d. h. s_ϱ^γ durch $\delta_{s_\varrho, 2\gamma - 1}$ ersetzen, also überall an Stelle von

$$F(s_1 s_2 \dots s_n) = \sum_{\gamma_\varrho = 0, 1} c_{\gamma_1 \gamma_2 \dots \gamma_n} s_1^{\gamma_1} s_2^{\gamma_2} \dots s_n^{\gamma_n} \quad (12)$$

die Funktion

$$\mathbf{U}F(s_1 s_2 \dots s_n) = \sum_{\gamma_\varrho = 0, 1} c_{\gamma_1 \gamma_2 \dots \gamma_n} \delta_{s_1, 2\gamma_1 - 1} \delta_{s_2, 2\gamma_2 - 1} \dots \delta_{s_n, 2\gamma_n - 1} \quad (12a)$$

schreiben. Dadurch ändern sich die Transformationseigenschaften der Funktionen nicht, da das Ersetzen von (12) durch (12a) mit einer Permutation der Variablen offenbar vertauschbar ist. Es gilt daher auch, wenn wir für

$$\mathbf{U}g_{\lambda k}^{(i)} = f_{\lambda, k - \frac{1}{2}n}^{(\frac{1}{2}n - i)}, \quad \mathbf{U}g_{\lambda, \frac{1}{2}n + m}^{(\frac{1}{2}n - S)} = f_{\lambda, m}^{(S)} \quad (13)$$

und für

$$\mathbf{D}^{(i)}(P) = \mathbf{A}^{(\frac{1}{2}n - i)}(P)^*, \quad \mathbf{D}^{(\frac{1}{2}n - S)}(P) = \mathbf{A}^{(S)}(P)^* \quad (13a)$$

schreiben¹⁾, nach (11)

$$\mathbf{Q}_P f_{\lambda m}^{(S)} = \sum_{\lambda'} \mathbf{A}^{(S)}(P)_{\lambda' \lambda}^* f_{\lambda' m}^{(S)}. \quad (11a)$$

Bei geradem n sind S und m beide ganzzahlig, bei ungeradem n beide halbzahlig.

Die Funktion $g_{\lambda, \frac{1}{2}n + m}^{(\frac{1}{2}n - S)}$ ist vom Grade $\frac{1}{2}n + m$, d. h. wenn man sie in der Form (12) schreibt, kommen in ihr nur jene Glieder vor, die $\frac{1}{2}n + m$ Faktoren s_ϱ^0 (und $\frac{1}{2}n - m$ Faktoren s_ϱ^0) enthalten. Daher kommen in $\mathbf{U}g_{\lambda, \frac{1}{2}n + m}^{(\frac{1}{2}n - S)} = f_{\lambda m}^{(S)}$ nur solche Glieder vor, die

¹⁾ Die $\mathbf{D}^{(i)}(P)$ sind eigentlich rein reell; der Übergang zum konjugiert Komplexen erfolgt in (13a) nur, damit der reelle Charakter später nicht benutzt werden muß.

$\frac{1}{2}n + m$ Faktoren $\delta_{s_0, 1}$ (und $\frac{1}{2}n - m$ Faktoren $\delta_{s_0, -1}$) enthalten: $f_{\lambda m}^{(S)}$ kann nur für solche Wertesysteme der s_0 von Null verschieden sein, bei denen genau $\frac{1}{2}n + m$ unter den s gleich $+1$ (und $\frac{1}{2}n - m$ gleich -1) sind, so daß ihre Summe $\frac{1}{2}n + m - (\frac{1}{2}n - m) = 2m$ wird:

$$f_{\lambda m}^{(S)}(s_1 s_2 \dots s_n) = 0 \text{ für } s_1 + s_2 + \dots + s_n \neq 2m. \quad (14)$$

Nimmt man für das vollständige Orthogonalsystem der Funktionen der s die Funktionen

$$f_{\lambda m}^{(S)} \left\{ \begin{array}{l} \lambda = 1, 2, \dots, \left(\frac{1}{2}n - S\right) - \left(\frac{1}{2}n - S - 1\right) \\ m = -S, -S + 1, \dots, S - 1, S \\ S = \frac{1}{2}n, \frac{1}{2}n - 1, \frac{1}{2}n - 2, \dots, \frac{1}{2} \text{ oder } 0, \end{array} \right\} \quad (*)$$

so erhält man für (9) mit Hilfe von (8) und (11a)

$$\mathbf{O}_P \psi_x f_{\lambda m}^{(S)} = \sum_{x'} \sum_{\lambda'} \mathbf{D}(P)_{x' x} \mathbf{A}^{(S)}(P)_{\lambda' \lambda}^* \psi_{x'} f_{\lambda' m}^{(S)}. \quad (9a)$$

Die $\psi_x f_{\lambda m}^{(S)}$ transformieren sich nach dem direkten Produkt

$$\mathbf{D}(P) \times \mathbf{A}^{(S)}(P)^*.$$

6. Von den Eigenfunktionen erster Näherung, von den richtigen Linearkombinationen

$$\sum_{\substack{x' S' \\ \lambda' m}} a_{x' S' \lambda' m} \psi_{x'} f_{\lambda' m}^{(S')} \quad (15)$$

des Störungsverfahrens (das zur Einführung der Spinkräfte dient) kann angenommen werden, daß sie zu irreduziblen Darstellungen der Gruppe der \mathbf{O}_P gehören. Da uns wegen des Pauliprinzips nur solche Eigenwerte beschäftigen werden, deren Darstellung die antisymmetrische ist, genügt es, die antisymmetrischen Linearkombinationen (15) zu bestimmen: die erste Näherung der dem Pauliprinzip genügenden Eigenfunktionen wird eine Linearkombination dieser sein.

Nehmen wir daher an, daß (15) antisymmetrisch ist, so folgt aus (9a) wegen der linearen Unabhängigkeit der $\psi_{x'} f_{\lambda' m}^{(S')}$

$$\sum_{x \lambda} a_{x S' \lambda m} \mathbf{D}(P)_{x' x} \mathbf{A}^{(S')}(P)_{\lambda' \lambda}^* = \varepsilon_P a_{x' S' \lambda' m}. \quad (16)$$

Wenn man die zu $\mathbf{A}^{(S')}(P)$ assoziierte Darstellung mit

$$\overline{\mathbf{A}}^{(S')}(P) = \varepsilon_P \mathbf{A}^{(S')}(P) \quad (17)$$

bezeichnet und (16) mit ε_P multipliziert, folgt wegen $\varepsilon_P^2 = 1$

$$\sum_{\pi \lambda} a_{\pi' S' \lambda' m} \mathbf{D}(P)_{\pi' \pi} \bar{\mathbf{A}}^{(S')}(P)_{\lambda' \lambda}^* = a_{\pi' S' \lambda' m}. \quad (18)$$

Summiert man dies über alle Permutationen P , so folgt aus den Orthogonalitätsrelationen, daß die linke Seite verschwindet, wenn $\mathbf{D}(P)$ und $\bar{\mathbf{A}}^{(S')}(P)$ nicht äquivalent sind. Ist $\mathbf{D}(P)$ keiner der Darstellungen $\bar{\mathbf{A}}^{(S')}(P)$ äquivalent, so verschwinden alle $a_{\pi' S' \lambda' m}$ und man kann dann aus den $\psi_x f_{\lambda m}^{(S')}$ überhaupt keine antisymmetrische Linearkombination bilden. Aus den Eigenfunktionen von \mathbf{H}_0 , die in bezug auf die \mathbf{P}_P zu einer irreduziblen Darstellung gehören, die keiner der Darstellungen $\bar{\mathbf{A}}^{(\frac{1}{2}n)}, \bar{\mathbf{A}}^{(\frac{1}{2}n-1)}, \bar{\mathbf{A}}^{(\frac{1}{2}n-2)}, \dots$ äquivalent ist, kann man mit Hilfe der Funktionen der s keine antisymmetrischen Eigenfunktionen bilden. Diese Eigenfunktionen und ihre Eigenwerte scheiden daher wegen des Pauliprinzips für das Folgende aus.

Wir wollen daher annehmen, daß $\mathbf{D}(P)$ einer der $\bar{\mathbf{A}}^{(S')}(P)$, etwa $\bar{\mathbf{A}}^{(S)}(P)$ äquivalent ist und sogar, daß es damit identisch ist, da eine Ähnlichkeitstransformation von $\mathbf{D}(P)$ nur eine bestimmte Wahl der linear unabhängigen Eigenfunktionen ψ_x bedeutet. Man nennt S das Multiplettsystem der Eigenfunktionen ψ_x . Setzt man

$$\mathbf{D}(P) = \bar{\mathbf{A}}^{(S)}(P) \quad (19)$$

in (18) ein und summiert jetzt über alle Permutationen, so ergibt sich, wenn man die Dimension von $\bar{\mathbf{A}}^{(S)}(P)$ für den Außenblick mit g_S bezeichnet:

$$\begin{aligned} \sum_{\pi \lambda} a_{\pi' S' \lambda' m} \frac{n!}{g_S} \delta_{SS'} \delta_{\pi' \pi} \delta_{\lambda' \lambda} &= n! a_{\pi' S' \lambda' m}, \\ a_{\pi' S' \lambda' m} &= \delta_{SS'} \delta_{\pi' \pi} \sum_{\lambda} \frac{a_{\pi S \lambda m}}{g_S} = \delta_{SS'} \delta_{\pi' \pi} b_m, \end{aligned} \quad (20)$$

wo b_m von S', π', λ' unabhängig ist. Die antisymmetrischen Linearkombinationen (15) der $\psi_x f_{\lambda m}^{(S')}$ lauten mithin:

$$\sum_{\substack{\pi' S' \\ \lambda' m}} \delta_{SS'} \delta_{\pi' \pi} b_m \psi_x f_{\lambda m}^{(S')} = \sum_m b_m \sum_{\pi} \psi_x f_{\pi m}^{(S)}. \quad (20a)$$

Das sind den $2S + 1$ Konstanten b_m entsprechend $2S + 1$ linear unabhängige antisymmetrische Funktionen

$$\Xi_m^S = \sum_{\pi} \psi_x f_{\pi m}^{(S)}. \quad (20b)$$

Will man aus den Eigenfunktionen $\psi_1, \psi_2, \psi_3, \dots$ eines Eigenwerts von \mathbf{H}_0 antisymmetrische Funktionen bilden, so muß man ψ_x , das zur x -Zeile von $\bar{\mathbf{A}}^{(S)}$ gehört, mit einer Funktion $f_{x'm}^{(S)}$ der s multiplizieren, die zur x -Zeile der assoziierten Darstellung $\mathbf{A}^{(S)*}$ gehört, und diese Produkte für alle x (für alle Partner) addieren. Die anderen, zu diesen orthogonalen $g_S \cdot 2^n - 2S - 1$ Linearkombinationen der $\psi_x f_{x'm}^{(S)}$ gehören in bezug auf die \mathbf{O}_P zu Darstellungen, die von der antisymmetrischen Darstellung verschieden sind.

7. Wenn man zu \mathbf{H}_0 die Spinglieder \mathbf{H}_1 als eine Störung hinzufügt, so bleibt er kein spinfreier Operator mehr und der Eigenwert E wird in mehrere Eigenwerte aufzuteilen, im allgemeinen in solche, zu denen irreduzible Darstellungen der Gruppe der Symmetrioperatoren \mathbf{O}_P gehören. Da in Wirklichkeit nur antisymmetrische Wellenfunktionen existieren, werden nur jene Terme eine Realität haben, zu denen die antisymmetrische Darstellung gehört. Gehörten die Funktionen von $x_1 y_1 z_1 \dots x_n y_n z_n$, die Eigenfunktionen von \mathbf{H}_0 zu E waren, zur Darstellung $\bar{\mathbf{A}}^{(S)}(P)$, so werden an der Stelle, wo E liegen sollte, $2S + 1$ nahe benachbarte Terme liegen. Zu jedem dieser Terme gehört in erster Näherung eine Linearkombination der Ξ_m von (20b). Die richtige Linearkombination kann man allerdings — da keine Symmetrie, die uns dabei helfen würde, angenommen wurde, ohne Auflösung der $2S + 1$ -dimensionalen Säkulargleichung von

$$(\mathbf{H}_1)_{m'm} = (\Xi_{m'}, \mathbf{H}_1 \Xi_m)$$

nicht bestimmen.

8. Wir gehen jetzt zu einem System über, bei dem außer der Gleichheit der Elektronen die volle Drehsymmetrie vorhanden ist. Die Funktionen von $x_1 y_1 z_1 \dots x_n y_n z_n$, die Eigenfunktionen von \mathbf{H}_0 sind, haben dann außer dem Multiplettsystem S auch eine azimutale Quantenzahl L und können so gewählt werden, daß sie den Gleichungen

$$\mathbf{P}_P \psi_{x\mu} = \sum_{x'} \bar{\mathbf{A}}^{(S)}(P)_{x'x} \psi_{x'\mu}; \quad \mathbf{P}_R \psi_{x\mu} = \sum_{\mu'} \mathbf{D}^{(L)}(R)_{\mu'\mu} \psi_{x\mu'} \quad (21)$$

genügen (P ist eine Permutation, R eine Drehung). Die Funktionen $\psi_{x\mu}$ kann man ebenso wie vorher durch Multiplikation mit einer beliebigen Funktion der s — wir wollen wieder die Funktionen $f_{x'm}^{(S)}$ verwenden — zu Funktionen $\psi_{x\mu} f_{x'm}^{(S)}$ aller Koordinaten ergänzen, die ebenfalls Eigenfunktionen von \mathbf{H}_0 sind. Die $\psi_{x\mu} f_{x'm}^{(S)}$

können auch an Stelle der $\psi_{\lambda\mu}$ in (21) eingesetzt werden, da \mathbf{P} auf die Spinkoordinaten nicht einwirkt und $f_{\lambda m}^{(S)}$ daher in (21) wie ein konstanter Faktor behandelt werden kann.

Die $\psi_{\lambda\mu} f_{\lambda m}^{(S)}$ müssen auch in bezug auf die \mathbf{O}_R zu einer Darstellung der Drehgruppe gehören, da die bloße Einführung der Spinkoordinaten die Gleichheit der Raumrichtungen nicht stört. In der Tat ist [vgl. die analoge Gleichung (9)]

$$\mathbf{O}_R \psi_{\lambda\mu} f_{\lambda m}^{(S)} = \mathbf{P}_R \psi_{\lambda\mu} \cdot \mathbf{Q}_R f_{\lambda m}^{(S)}, \quad (22)$$

und man kann hierin die $\mathbf{P}_R \psi_{\lambda\mu}$ mit Hilfe von (21) durch die unveränderten $\psi_{\lambda\mu}$ und auch $\mathbf{Q}_R f_{\lambda m}^{(S)}$ — wie jede Funktion der s — mit Hilfe der $f_{\lambda m}^{(S)}$ ausdrücken. Wir wollen jetzt die Koeffizienten bestimmen.

Wenn man die $\mathbf{Q}_R f_{\lambda m}^{(S)}$ durch die $f_{\lambda m'}^{(S')}$ ausdrücken will, braucht man dazu nur die $f_{\lambda m'}^{(S')}$, die ebenfalls zur λ -Zeile von $\mathbf{A}^{(S)}$ gehören, weil ja \mathbf{Q}_R ein den Vertauschungen der s gegenüber symmetrischer Operator ist und die Transformationseigenschaften den \mathbf{Q}_P gegenüber nicht ändert. Es muß daher

$$\mathbf{Q}_R f_{\lambda m}^{(S)} = \sum_{m'=-S}^S \mathfrak{D}^{(S)}(R)_{m'm} f_{\lambda m'}^{(S')} \quad (m = -S, -S+1, \dots, S-1, S) \quad (23)$$

sein, und die Matrizen $\mathfrak{D}^{(S)}(R)$ bilden wegen $\mathbf{Q}_R \mathbf{Q}_{R'} = \pm \mathbf{Q}_{R'R}$ eine $2S+1$ -dimensionale ein- oder zweideutige Darstellung der Drehgruppe, die — wie sogleich gezeigt werden soll — die irreduzible Darstellung $\mathfrak{D}^{(S)}(R)$ ist.

Es sei nämlich R eine Drehung mit α um Z , dann ist

$$\begin{aligned} \mathbf{Q}_R f_{\lambda m}^{(S)}(s_1 \dots s_n) &= \sum_{t_\rho=\pm 1} \dots \mathfrak{D}^{(1/2)}(R)_{\frac{1}{2}s_\rho, \frac{1}{2}t_\rho} \dots f_{\lambda m}^{(S)}(t_1 \dots t_n) \\ &= \sum_{t_\rho=\pm 1} \delta_{s_1 t_1} e^{i \frac{1}{2} s_1 \alpha} \dots \delta_{s_n t_n} e^{i \frac{1}{2} s_n \alpha} f_{\lambda m}^{(S)}(t_1 \dots t_n) \\ &= e^{i \frac{1}{2} (s_1 + \dots + s_n) \alpha} f_{\lambda m}^{(S)}(s_1 \dots s_n) \end{aligned}$$

$$\mathbf{Q}_R f_{\lambda m}^{(S)}(s_1 \dots s_n) = e^{i m \alpha} f_{\lambda m}^{(S)}(s_1 \dots s_n), \quad (24)$$

wo noch für $\frac{1}{2}(s_1 + s_2 + \dots + s_n) = m$ gesetzt werden konnte, weil nach (14) $f_{\lambda m}^{(S)}$ für andere Wertesysteme der s doch verschwindet.

Die Darstellung in (23) ist für $R = \{\alpha 0\}$ eine Diagonalmatrix mit den Diagonalelementen $e^{-i s_1 \alpha}, e^{-i(s-1)\alpha}, \dots, e^{i(s-1)\alpha}, e^{i s \alpha}$, woraus tatsächlich ihre Äquivalenz mit der irreduziblen Darstellung $\mathfrak{D}^{(S)}(R)$ folgt. Zudem zeigt (24), daß $f_{\lambda m}^{(S)}$ zur m -Zeile von

$\mathfrak{D}^{(S)}(R)$ gehört, so daß in (23) $\mathfrak{D}^{(S)}(R)$ zu Recht besteht. Die Funktionen von $s_1 s_2 \dots s_n$, die in bezug auf Vertauschung der Variablen zur Darstellung $A^{(S)*} = D^{(1/2n-S)}$ gehören, gehören in bezug auf die Drehungen Q_R zur Darstellung $\mathfrak{D}^{(S)}$ der Drehgruppe. Sie gehören eigentlich zur Darstellung $A^{(S)*} \times \mathfrak{D}^{(S)}$ des direkten Produktes beider Gruppen: aus (23) folgt, wenn man darauf Q_P anwendet, mit Hilfe von (11a)

$$\begin{aligned} Q_P Q_R f_{\lambda m}^{(S)} &= \sum_{m'} \mathfrak{D}^{(S)}(R)_{m' m} Q_P f_{\lambda m'}^{(S)} \\ &= \sum_{m'} \sum_{\lambda'} \mathfrak{D}^{(S)}(R)_{m' m} A^{(S)}(P)_{\lambda' \lambda}^* f_{\lambda' m'}^{(S)}, \end{aligned} \quad (24a)$$

däß $f_{\lambda m}^{(S)}$ zur λ, m -Zeile von $A^{(S)}(P)^* \times \mathfrak{D}^{(S)}(R)$ gehört.

Man kann diesen Tatbestand an Hand der Schemata auf S. 144 sehr schön erläutern. Man kann sich nämlich rechts an die Stellen, wo die Symbole $D^{(i)}$ stehen, diejenigen Funktionen eingetragen denken, die zu diesem $D^{(i)}$ gehören. Dann stehen etwa in der i -ten Spalte lauter Funktionen, die in bezug auf Vertauschungen der Variablen zu $D^{(i)} = A^{(1/2n-i)}$ gehören, an der Stelle des $D^{(i)}$ der k -ten Zeile steht $g_{1k}^{(i)}, g_{2k}^{(i)}, g_{3k}^{(i)}, \dots$. Ersetzt man noch die $g_{\lambda k}^{(i)}$ durch die $f_{\lambda, k-1/2n}^{(1/2n-i)} = U g_{\lambda k}^{(i)}$ der S. 276, so stehen in der $k = (1/2n + m)$ -ten Zeile lauter Funktionen, die in bezug auf Drehungen um Z zu $(e^{im\varphi})$ gehören. Daraus, daß jedes $D^{(i)} = A^{(1/2n-i)}$ in jeder Zeile höchstens einmal vorkommt, sieht man, daß höchstens eine Funktion der s existiert, die zu einer bestimmten Zeile von $A^{(1/2n-i)}$ und in bezug auf Drehungen zu $(e^{im\varphi})$ gehört. Da $A^{(1/2n-i)}$ in den Zeilen $i, i+1, \dots, n-i-1, n-i$ vorkommt, kann $m = k - 1/2n$ die Werte $-1/2n+i, -1/2n+i+1, \dots, 1/2n-i-1, 1/2n-i$ annehmen. In bezug auf dreidimensionale Drehungen gehören die Funktionen, die in der i -ten Spalte an den entsprechenden Stellen verschiedener Zeilen stehen, zu den verschiedenen Zeilen von $\mathfrak{D}^{(1/2n-i)}$ und sind Partner voneinander.

Hätten wir an Stelle der Funktionen $f = U g$ direkt die Funktionen g des Kap. XIII benutzt, so müßten wir anstatt „Drehungen um Z “ immer „Drehungen um Y “ sagen, sonst hätte sich nichts geändert.

Eine antisymmetrische Funktion der $x_1 y_1 z_1 s_1 \dots x_n y_n z_n s_n$, die in bezug auf die Vertauschung P_P der Cartesischen Koordinaten allein zur Darstellung $\overline{A}^{(S)}$ gehört, das Multiplettsystem S hat, gehört nach (20b) und (11a) in bezug auf die Vertauschung Q_P der Spinkoordinaten zu der Darstellung $A^{(S)*}$ und in bezug auf Drehungen der Spinkoordinaten nach (24a) zu $\mathfrak{D}^{(S)}$. Das Multiplettsystem ist daher nicht bloß eine Symmetrie in bezug auf die Permutationen der Variablen (entweder der Cartesischen oder der Spinkoordinaten), sondern wegen der besonderen Struktur der Funktionen der s auch eine Symmetrie in bezug auf

die Drehung der Spinkoordinaten; ebenso wie die azimutale Quantenzahl der Ausdruck für eine Symmetrie in bezug auf Drehung der Cartesischen Koordinaten ist. Ein wichtiger Unterschied gegenüber der azimutalen Quantenzahl besteht darin, daß die wichtigsten Größen, wie auch H_0 , spinfreie Größen sind, die allen \mathbf{Q} gegenüber, auch wenn die Isotropie des Raumes durch äußere Felder aufgehoben ist, invariant sind.

Daß aus der Zugehörigkeit der Funktionen $f_{\lambda m}^{(S)}$ zu einer irreduziblen Darstellung $A^{(S)}$ der symmetrischen Gruppe ihre Zugehörigkeit zu der irreduziblen Darstellung $\mathfrak{D}^{(S)}$ der Drehgruppe und umgekehrt folgt, ist eine Eigenschaft der Funktionen der Variablen s , die nur zweier Werte fähig sind. Würden die s mehrere Werte annehmen können, so würde eine Funktion, die zu einer bestimmten Darstellung der symmetrischen Gruppe gehört, noch zu verschiedenen Darstellungen der Drehgruppe gehören können, und auch umgekehrt, eine Funktion, die zu einer bestimmten Darstellung der Drehgruppe gehört, könnte noch zu verschiedenen Darstellungen der symmetrischen Gruppe gehören. Die Transformationseigenschaften gegenüber Drehungen und Permutationen würden sich nicht gegenseitig bestimmen.

9. Bilden wir jetzt nach (20b) antisymmetrische Linear-kombinationen der $\psi_{k\mu} f_{\lambda m}^{(S)}$, so gibt es deren $(2S+1)(2L+1)$, weil man für jedes μ die Ausdrücke (20b) bilden kann:

$$\Xi_{m\mu}^{SL} = \sum_s \psi_{s\mu}^{SL} f_{\lambda m}^{(S)} (\mu = -L, \dots, L; m = -S, \dots, S). \quad (25)$$

Unterwirft man die Zustände $\Xi_{m\mu}^{SL}$ einer Drehung

$$\begin{aligned} \mathbf{O}_R \Xi_{m\mu}^{SL} &= \sum_s \mathbf{P}_R \psi_{s\mu}^{SL} \cdot \mathbf{Q}_R f_{\lambda m}^{(S)} \\ &= \sum_s \sum_{\mu' m'} \mathfrak{D}^{(L)}(R)_{\mu' \mu} \psi_{s\mu'}^{SL} \mathfrak{D}^{(S)}(R)_{m' m} f_{\lambda m'}^{(S)} \\ &= \sum_{\mu' m'} \mathfrak{D}^{(L)}(R)_{\mu' \mu} \mathfrak{D}^{(S)}(R)_{m' m} \Xi_{m'\mu'}^{SL}, \end{aligned}$$

so transformieren sie sich nach dem direkten Produkt $\mathfrak{D}^{(L)} \times \mathfrak{D}^{(S)}$. Die Gesamtquantenzahlen J der entstehenden antisymmetrischen Eigenwerte erhalten wir daher durch Ausreduktion von $\mathfrak{D}^{(L)} \times \mathfrak{D}^{(S)}$. Diese Ausreduktion haben wir aber schon in Kap. XVII ausgeführt; die irreduziblen Bestandteile haben die oberen Indizes

$$J = |L-S|, |L-S|+1, \dots, L+S-1, L+S, \quad (26)$$

während die zugehörigen Linearkombinationen der $\Xi_{m\mu}^{SL}$ nach (18a) S. 206

$$\Psi_m^J = \sum_{\mu} s_{J\mu m-\mu}^{(L S)} \Xi_{m-\mu \mu}^{SL} \quad (27)$$

sind. Die Koeffizienten $s_{J\mu m-\mu}^{(L S)}$ wurden ebenfalls im XVII. Kapitel berechnet [Gl. (27), (27b)].

Aus einem Term mit der Azimutalquantenzahl L und dem Multiplettsystem S entstehen durch die Einführung des Spins $2L + 1$ bzw. $2S + 1$ (je nachdem, welche Zahl kleiner ist) sog. Feinstrukturkomponenten mit den Gesamtquantenzahlen (26). Die zugehörigen Eigenfunktionen erster Näherung sind durch (27) gegeben.

Obwohl also bei voller räumlicher Symmetrie mehr Eigenfunktionen zu einem Eigenwert gehören, als ohne diese, kann man die richtigen Linearkombinationen für das Störungsverfahren doch viel leichter, eben schon auf Grund der räumlichen Symmetrie bestimmen.

Aus (26) ersieht man, daß die Gesamtquantenzahlen J der Terme, die aus einem Term mit der Azimutalquantenzahl L und dem Multiplettsystem S durch Einführung des Spins entstehen, mit Hilfe des Vektoradditionsmodells erhalten werden können. Man muß dabei nach der Abb. 8 die beiden Vektoren L und S addieren¹⁾ und erhält als Resultante die Vektoren J . Es wird L als der Drehimpuls, der von der Bahnbewegung herrührt, und S als der Drehimpuls, der vom Spin der Elektronen stammt, aufgefaßt; J ist der gesamte Drehimpuls.

10. Bestimmen wir noch den Spiegelungscharakter der Funktionen (27)! Ist der Spiegelungscharakter der Funktionen $\psi_{x\mu}$ gleich w (er ist für alle $\psi_{x\mu}$ derselbe)

$$\mathbf{P}_I \psi_{x\mu} = w \psi_{x\mu},$$

so gilt dies auch für die Funktionen (27), weil sie ja, als Funktionen der Cartesischen Koordinaten betrachtet, Linearkombinationen der $\psi_{x\mu}$ sind, und $\mathbf{O}_I = \mathbf{P}_I$ ist. Der Spiegelungscharakter ändert sich also durch Einführung der Spinkoordinate nicht, er ist für alle Feinstrukturkomponenten eines Terms derselbe, wie vor Einführung des Spins für den entsprechenden Grobstrukturterm.

11. Es sei hier noch die Größenordnung des Abstandes der Feinstrukturkomponenten voneinander, die sogenannte Multiplett-aufspaltung, roh abgeschätzt.

Es handelt sich dabei, klassisch gesprochen, um die Wechselwirkungsenergie von Magnettadeln (den Spins) und von Stromkreisen, die von den Elektronen, die den Kern umkreisen, gebildet werden. Die Feldstärke, die durch die Stromkreise erzeugt wird, ist $\sim ev/r^2 c$, wo e Ladung, v Geschwindigkeit des Elektrons und r

¹⁾ Auf Abb. 8 (S. 202) steht l für L ; \bar{l} für S und L für J .

sein Abstand vom Aufpunkt ist. Man kann für r den Radius der ersten Bohrschen Bahn einsetzen, der auch in der Quantenmechanik eine mittlere Entfernung der Atombestandteile voneinander ist, v kann man aus $mv r \sim h/2\pi$ abschätzen. Man erhält so für die Feldstärke

$$\sim \frac{32\pi^5 m^9 e^7}{h^5 c}$$

und die Energie eines magnetischen Dipols von der Stärke $eh/4\pi mc$ ist $8\pi^4 m e^8/h^4 c^2$. Die genaue Rechnung mit Hilfe der Diracschen relativistischen Theorie gibt für die Energiedifferenz der beiden Terme $N = 2$, $l = 1$, $j = 1/2$ und $j = 3/2$ bei dem Wasserstoffatom den Wert $\pi^4 m e^8/2 h^4 c^2$. Die Feinstrukturaufspaltung ist also nur der etwa

$$\sim \left(\frac{2\pi e^2}{hc} \right)^2 = \alpha^2 = \left(\frac{1}{137,3} \right)^2$$

Teil der Grobstruktur, der gesamten Energie des Wasserstoffatoms. Man nennt α die Sommerfeldsche Feinstrukturkonstante.

Man gelangt zu einer Unterscheidung verschiedener physikalischer Effekte nach ihrer Größenordnung, indem man die Energie nach Potenzen der Feinstrukturkonstante entwickelt. Fast jede Potenz trägt einem neuen physikalischen Effekt Rechnung. Die nullte Potenz enthält nur die Ruhenergie mc^2 des Elektrons. Die erste Potenz hat den Koeffizienten Null, die zweite Potenz ist das, was die gewöhnliche Schrödingersche Theorie liefert, sie ist proportional $mc^2 \alpha^2 = 4\pi^2 m e^4/h^2$ und ist das einzige Glied, in dem die Lichtgeschwindigkeit nicht vorkommt. Der Koeffizient des dritten Gliedes ist wieder Null; das vierte Glied trägt, wie wir eben gesehen haben, der Energie der magnetischen Momente der Elektronen Rechnung, und zwar gibt es die erste Näherung des Störungsverfahrens. Die fünfte Potenz enthält die durch die Dipolstrahlung bedingte Verbreiterung der Terme¹⁾. Das sechste Glied ist die zweite Näherung

¹⁾ Die Summe der Verbreiterungen zweier Terme ergibt die natürliche Breite der sie verbindenden Linie. Die Breite eines Terms ist die h -fache Summe der Übergangswahrscheinlichkeiten der von diesem Term ausgehenden Linien.

Wenn man in (1a) des Kap. XVIII für $h\nu \sim mc^2 \alpha^2$ und für die Matrixelemente von X den Radius der ersten Bohrschen Bahn einsetzt, sieht man, daß die Übergangswahrscheinlichkeiten

$$\sim \frac{64\pi^4 e^2 m^3 c^6 \alpha^6}{3 h c^3 h^3} \frac{h^4}{16\pi^4 m^2 e^4} = \frac{1}{h} \frac{4m c^3 h \alpha^6}{3 e^2} \sim \frac{m c^3}{h} \alpha^5$$

mit der fünften Potenz der Feinstrukturkonstante gehen.

der Spinwechselwirkung, das siebente die Verbreiterung der Terme durch Quadrupolstrahlung. Das achte Glied ist die dritte Näherung bei der Berechnung der Energie des Spins, die neunte Potenz die Verbreiterung der Terme durch Linien, die das Interkombinationsverbot verletzen usw.

Natürlich kann ein Glied, das eine höhere Potenz der Feinstrukturkonstante enthält, auch größer sein als ein Glied mit einer niedrigeren Potenz, nämlich, wenn sein Koeffizient groß genug ist. Die Regel ist aber doch die, daß das Glied mit der höheren Potenz von α das kleinere ist. Immerhin steigt der Koeffizient mancher Glieder (z. B. der Feinstrukturaufspaltung) hauptsächlich mit steigender Ordnungszahl regelmäßig an, bei anderen Gliedern (Verbreiterung der Terme durch Strahlung) ist das nicht, oder nicht in dem Maße der Fall. Daher ist z. B. die zweite Näherung in der Feinstrukturaufspaltung — abgesehen von den allerersten Elementen — wesentlich größer als die Verbreiterung durch Strahlung; die Linien, die das Interkombinationsverbot verletzen, sind fast immer stärker als die Quadrupollinien, so daß die Entwicklung der Effekte nach Potenzen von α oft nicht mehr als ein Einteilungsprinzip ist. Wenn das Glied, das früher in der Reihe vorkommt, größer als das später vorkommende ist, sagt man, daß die Koppelung normal ist.

12. Wenn die Eigenfunktionen von \mathbf{H} genau durch (27) gegeben wären, würden sie in bezug auf die \mathbf{P}_P bzw. \mathbf{P}_R zu den irreduziblen Darstellungen $\overline{\mathbf{A}}^{(S)}(P)$ bzw. $\mathbf{D}^{(L)}(R)$ gehören, und die Auswahlregeln für das Multiplettsystem und die azimutale Quantenzahl wären genau richtig. In Wirklichkeit ist (27) nur die erste Näherung für die Eigenfunktion. Wenn sich die zweite Näherung

$$\Phi_m^{NJ} = \Psi_m^{NJ} + \sum_{N' \neq N} \sum_{J' m'} \frac{(\Psi_{m'}^{N' J'}, \mathbf{H}_1 \Psi_m^{NJ})}{E_N - E_{N'}} \quad (28)$$

nicht wesentlich von der ersten unterscheidet, so kann man annehmen, daß die erste Näherung schon nahezu richtig ist¹⁾. In diesem Fall dürfen auch Linien, die gegen das Auswahlverbot für S oder L verstößen, nicht mit merklicher Intensität auftreten.

¹⁾ $E_{N'}$ durchläuft in (28) alle Eigenwerte der einfachen Schrödinger-Gleichung; die Indizes S und L , die das Multiplettsystem und die Azimutalquantenzahl angeben, sind in das N' mit einbezogen.

Nun genügt es, in (28) über N' zu summieren; da \mathbf{H}_1 den \mathbf{O}_R gegenüber symmetrisch ist, verschwindet $(\Psi_m^{N'J'}, \mathbf{H}_1 \Psi_m^{NJ})$ doch für $J' \neq J$ oder $m' \neq m$. Weiter ist $(\Psi_m^{NJ}, \mathbf{H}_1 \Psi_m^{NJ})$ die erste Näherung für die Energietörung durch den Spin, also die Multiplettaufspaltung, und $(\Psi_m^{N'J}, \mathbf{H}_1 \Psi_m^{N'J})$ hat dieselbe Größenordnung. Dagegen ist $E_N - E_{N'}$ der Abstand des nächsten Grobstrukturterms, der einen Eigenwert mit der Gesamtquantenzahl J liefert. Wenn die erste Größe wesentlich kleiner als die zweite ist, ist die Näherung (27) eine gute, sonst aber nicht. Es kommt also dabei wesentlich auf die zufällige Nähe eines Terms mit demselben J , aber anderem S oder L an. [Wenn S und L von Ψ_m^{NJ} gleich dem S bzw. L von $\Psi_m^{N'J}$ sind, stört das entsprechende Zusatzglied in (28) die Transformationseigenschaft von $\Psi_m^{N'J}$ den \mathbf{P} gegenüber nicht.]

XXIII. Auswahl- und Intensitätsregeln bei Mitberücksichtigung des Spins

Die Auswahl-, Intensitäts- und Aufspaltungsregeln können bei Mitberücksichtigung des Spins in zwei Klassen eingeteilt werden. Die Regeln der ersten Klasse (1 bis 4) folgen aus reinen Symmetrieverlegungen, ohne irgendwelche Annahmen über die Größe der Spinkräfte. Diese Regeln haben sowohl ihrem Inhalt wie ihrer Begründung nach große Ähnlichkeit mit den Regeln der einfachen Theorie (Kap. XVIII), welche auch auf der Invarianz des Energieoperators räumlichen Drehungen und Spiegelungen gegenüber beruhen. Die Regeln der zweiten Klasse (5 bis 7) beruhen nicht auf der Isotropie des Raumes allein, sondern enthalten noch wesentlich die Voraussetzung, daß die Spinkräfte klein sind im Verhältnis zu den elektrostatischen Kräften der einfachen Theorie, und daß sie daher die Eigenfunktionen und Eigenwerte der einfachen Theorie nicht wesentlich verändern können.

1. Das Auswahlverbot für die Gesamtquantenzahl lautet genau so, wie das Auswahlverbot für die Azimutalquantenzahl in der einfachen Theorie. Bei einem durch Dipolstrahlung bedingten Übergang ändert sich j um ± 1 oder 0, mit dem Zusatz, daß auch der Übergang zwischen zwei Termen mit $j = 0$ verboten ist. Für die Dipolstrahlung sind die Matrixelemente eines Vektoroperators (des

Multiplizierens mit $x_1 + x_2 + \dots + x_n$ usw.) charakteristisch, und diese verschwinden, wenn die obengenannten Bedingungen nicht erfüllt sind.

Auch das Auswahlverbot für den Spiegelungscharakter (Laportesche Regel) bleibt erhalten: der Operator \mathbf{O}_I ist ja identisch mit dem Operator \mathbf{P}_I der einfachen Schrödingerschen Theorie. Es verschwinden sogar alle Matrixelemente

$$(\Psi_F, \mathbf{V}_x \Psi_E), (\Psi_F, \mathbf{V}_y \Psi_E), (\Psi_F, \mathbf{V}_z \Psi_E) \quad (1)$$

eines polaren Vektoroperators, wenn der Spiegelungscharakter von Ψ_F und Ψ_E nicht entgegengesetzt ist. Führt man nämlich eine Inversion des Achsenkreuzes aus, so behält ein polarer Vektor seine Richtung bei, seine Komponenten wechseln also das Vorzeichen, es ist

$$\mathbf{O}_I \mathbf{V}_x \mathbf{O}_I^{-1} = -\mathbf{V}_x.$$

Ist daher $\mathbf{O}_I \Psi_F = w_F \Psi_F$ und $\mathbf{O}_I \Psi_E = w_E \Psi_E$, so folgt aus dem unitären Charakter von \mathbf{O}_I :

$$\begin{aligned} (\Psi_F, \mathbf{V}_x \Psi_E) &= (\mathbf{O}_I \Psi_F, \mathbf{O}_I \mathbf{V}_x \mathbf{O}_I^{-1} \cdot \mathbf{O}_I \Psi_E) \\ &= -w_F w_E (\Psi_F, \mathbf{V}_x \Psi_E), \end{aligned}$$

so daß (1) verschwinden muß, wenn Ψ_F und Ψ_E gleichen Spiegelungscharakter haben. (Genau entsprechend zeigt man, daß die Matrixelemente eines axialen Vektoroperators Null sind, wenn Ψ_F und Ψ_E verschiedenen Spiegelungscharakter haben.)

Da sich der Spiegelungscharakter eines Grobstrukturterms auf alle seine Feinstrukturkomponenten überträgt, gilt die Laportesche Regel auch bei Mitberücksichtigung des Spins für alle Feinstrukturkomponenten eines Grobstrukturterms in gleicher Weise.

In einem Termschema, das aus lauter ungestrichenen $w = (-1)^L$ [oder aus lauter gestrichenen $w = -(-1)^L$] Dublettermen besteht, wie z. B. in jedem wasserstoffähnlichen Spektrum, haben die Auswahlregeln für j und w auch die strenge Gültigkeit der Auswahlregel für L zur Folge. Wegen der Laporteschen Regel kann sich L nur um eine ungerade Zahl ändern: eine Änderung um 1 ist erlaubt, bei einer Änderung um 3 oder mehr müßte sich aber wegen $j = L \pm \frac{1}{2}$ auch j um 2 oder mehr ändern, was nicht möglich ist.

Die Regel, daß sich die Transformationseigenschaft gegenüber der Vertauschung der Elektronen nicht ändert, ist schon darin enthalten, daß die Wellenfunktionen, die alle antisymmetrisch sind, auch bei einer beliebigen Störung (nicht nur bei der Einwirkung von Licht) antisymmetrisch bleiben.

2. In einem zu Z parallelen magnetischen Feld spaltet der Term mit der Gesamtquantenzahl j in $2j+1$ Zeemankomponenten auf. Die „richtigen Linearkombinationen“ bei der Einführung der magnetischen Zusatzenergie als Störung sind, ebenso wie in der einfachen Theorie, die Ψ_μ^j selber. Wenn man auf ein Ψ_μ^j den Operator \mathbf{O}_R anwendet, wo R eine Drehung um Z mit α ist, multipliziert sich Ψ_μ^j lediglich mit $e^{i\alpha}$. Das Magnetfeld hebt die Entartung vollkommen auf, im Magnetfeld gehört nur je eine Eigenfunktion zu jedem Eigenwert.

Wenn man den durch das Magnetfeld bedingten Zusatzoperator \mathbf{H}_2 nach den Komponenten \mathfrak{H}_x , \mathfrak{H}_y , \mathfrak{H}_z der Feldstärke in eine Reihe entwickelt,

$$\mathbf{H}_2 = \mathfrak{H}_x \mathbf{V}_x + \mathfrak{H}_y \mathbf{V}_y + \mathfrak{H}_z \mathbf{V}_z + \mathfrak{H}_x^3 \mathbf{V}_{xx} + \dots, \quad (2)$$

so müssen die Koeffizienten \mathbf{V}_x , \mathbf{V}_y , \mathbf{V}_z der ersten Potenz einen axialen Vektoroperator bilden, da ja \mathfrak{H} auch ein axialer Vektor ist und das ganze \mathbf{H}_2 ein Skalar sein muß. Die Energiestörung erster Näherung in einem Magnetfeld, das die Z -Richtung hat,

$$\mathfrak{H}_z (\Psi_\mu^j, \mathbf{V}_z \Psi_\mu^j),$$

ist nach der Formel für die Matrixelemente von Vektoroperatoren [vgl. die mittlere Formel (19b), Kap. XXI] proportional zu μ . Die Aufspaltung im Magnetfeld ist daher bei kleiner Feldstärke proportional zur ersten Potenz der Feldstärke und erfolgt in $2j+1$ äquidistante Komponenten. Sie ist jedoch im Gegensatz zum Aufspaltungsbild der einfachen Theorie nicht mehr für alle Terme gleich groß und lässt sich nicht mehr allgemein berechnen. Nur im Falle „normaler Kopplung“, wenn (27) des vorangehenden Kapitels eine gute Näherung für die Eigenfunktion ist, kann man sie zahlenmäßig angeben.

Anders die Intensitätsverhältnisse der Zeemankomponenten! Die Stärke der Linie, die von der μ -Komponente des oberen Terms zur μ' -Komponente des unteren Terms führt, ist, abgesehen von einem konstanten universellen Faktor, je nachdem das Licht in der Z -Richtung oder rechts oder links um Z zirkular polarisiert ist, das Quadrat des Absolutwertes von

$$(\Psi_\mu^{Nj}, \sum_k z_k \Psi_{\mu'}^{N'j'}), \frac{1}{\sqrt{2}} (\Psi_\mu^{Nj}, \sum_k (ix_k + y_k) \Psi_{\mu'}^{N'j'}), \\ \frac{1}{\sqrt{2}} (\Psi_\mu^{Nj}, \sum_k (ix_k - y_k) \Psi_{\mu'}^{N'j'}).$$

Das sind aber die Matrixelemente der drei verschiedenen $(0, -1, +1)$ Komponenten eines Vektoroperators. Ihre Verhältnisse für verschiedene μ, μ' und Polarisierung kann man direkt aus (19), Kap. XXI, entnehmen. Aus diesen Formeln ersieht man auch, daß sich die magnetische Quantenzahl μ nur um 0 (π -Komponente) oder um ± 1 (σ -Komponente) ändern kann. Die Intensitätsverhältnisse sind bei einem $j \rightarrow j - 1$ -Übergang z. B.

$$\left. \begin{aligned} A_{\mu \rightarrow \mu-1} &= (j+\mu)(j+\mu-1); \\ A_{\mu \rightarrow \mu} &= 2(j+\mu)(j-\mu); \\ A_{\mu \rightarrow \mu+1} &= (j-\mu-1)(j-\mu). \end{aligned} \right\} \quad (3)$$

Die Summe der drei Ausdrücke in (3), der Übergangswahrscheinlichkeiten vom oberen Zustand mit der magnetischen Quantenzahl μ in alle Zeemanterme des unteren Niveaus, ist unabhängig von μ für alle Zeemanterme des oberen Niveaus gleich. Dasselbe gilt auch für die Summe der Übergangswahrscheinlichkeiten aller Linien, die den unteren Zeemanterm gemeinsam haben.

Bei näherem Zusehen ist dieser Summensatz der Übergangswahrscheinlichkeiten auch nicht so überraschend. Die Summe der drei Ausdrücke in (3) ist ja die gesamte Wahrscheinlichkeit für den Übergang des Zustandes Ψ_{μ}^{Nj} in einen Zustand mit der Energie des unteren Niveaus. Da aber die Zustände Ψ_{μ}^{Nj} mit verschiedenen μ bei einer Drehung des Achsenkreuzes in Linearkombinationen voneinander übergehen und sich so im wesentlichen nur durch eine Drehung unterscheiden, die gesamte Übergangswahrscheinlichkeit andererseits von einer Drehung unabhängig sein muß, darf sie auch von μ nicht abhängen.

Man verifiziert diese Regeln am einfachsten, indem man von (18a), Kap. XXI, ausgeht und

$$|\mathbf{T}_{Nj\mu; N'j'\mu'}^{(o)}|^2 = \sum_{r\sigma r'} \sum_{\lambda\tau\lambda'} \mathfrak{D}^{(j)}(R)_{r\mu}^* \mathfrak{D}^{(j)}(R)_{\lambda\mu} \mathfrak{D}^{(\omega)}(R)_{\sigma\varrho}^* \mathfrak{D}^{(\omega)}(R)_{\tau\varrho} \mathfrak{D}^{(j')}(\mathbf{R})_{r'\mu'}^* \mathfrak{D}^{(j')}(\mathbf{R})_{\lambda'\mu'}^* \mathbf{T}_{Nj\nu; N'j'\nu'}^{(o)} \mathbf{T}_{Nj\lambda; N'j'\lambda'}^{(t)} ^*$$

bildet. Wenn man jetzt über μ und ϱ summiert, so erhält man rechts wegen der Unitarität von $\mathfrak{D}^{(j)}$ und $\mathfrak{D}^{(\omega)}$ an Stelle der ersten vier Faktoren $\delta_{rr'} \delta_{\sigma\tau}$. Wenn man dann noch über alle Drehungen integriert, ergibt sich wegen der Orthogonalitätsrelationen nach Kürzung mit $\int dR$:

$$\sum_{\mu\varrho} |\mathbf{T}_{Nj\mu; N'j'\mu'}^{(o)}|^2 = \frac{1}{2j'+1} \sum_{r\sigma r'} |\mathbf{T}_{Nj\nu; N'j'\nu'}^{(o)}|^2, \quad (*)$$

also direkt die Unabhängigkeit der Summe (*) von μ' .

3. Ein zu Z paralleles elektrisches Feld spaltet bei gerader Elektronenzahl einen Term mit der Gesamtquantenzahl j — ähnlich wie in der im Kap. XVIII behandelten einfachen Theorie — in $j + 1$ Komponenten auf, die zu den Darstellungen

$$\mathbf{3}^{(j)}, \mathbf{3}^{(j-1)}, \dots, \mathbf{3}^{(2)}, \mathbf{3}^{(1)}, \mathbf{3}^{(0)} \text{ oder } \mathbf{3}^{(0')}$$

der zweidimensionalen Drehspiegelungsgruppe gehören. Der letzte Term gehört zu $\mathbf{3}^{(0)}$ oder zu $\mathbf{3}^{(0')}$, je nachdem $w(-1)^j$ gleich $+1$ oder -1 ist. Auch die Auswahlregeln des Kap. XVIII lassen sich genau so, wie es dort geschah, ableiten, mit dem einzigen Unterschied, daß überall j an Stelle von L zu setzen ist.

Bei ungerader Elektronenzahl sollen die Verhältnisse etwas genauer untersucht werden. Nicht wegen des Resultates, das man fast sofort hinschreiben könnte, sondern weil in ähnlichen Fällen Zweifel über die richtige Anwendung der Theorie entstehen könnten.

Die Schwierigkeit besteht darin, daß den Drehungen R bei ungerader Elektronenzahl eine Matrix $\pm \mathbf{D}^{(j)}(R)$ mit unbestimmtem Vorzeichen zugeordnet ist. Dasselbe gilt im allgemeinen auch für Drehspiegelungen $\mathbf{D}^{(j,w)}(RI) = \pm w \mathbf{D}^{(j)}(R)$, nur der Inversion kann eindeutig $+w\mathbf{1}$ zugeordnet werden. Nun hebt das elektrische Feld die Inversionssymmetrie auf, und obwohl noch viele Drehspiegelungen übrigbleiben, ist es für die zugeordneten Matrizen eben wegen ihrer Zweideutigkeit gleichgültig, ob $w = +1$ oder $w = -1$ ist. Man hat daher das Empfinden, daß durch die Zweideutigkeit der Darstellungen etwas Wesentliches verlorengegangen ist. Bei gewissen Symmetrien ist dies auch der Fall, nicht aber, wie wir sehen werden, bei der hier vorliegenden Symmetrie.

Um zu eindeutigen Darstellungen zu gelangen, kann man von der Tatsache ausgehen, daß die Drehsymmetrie bei ungerader Elektronenzahl durch die Operatoren \mathbf{O}_u , die eine zu der zweidimensionalen unitären Gruppe holomorphe Gruppe bilden, in Evidenz gesetzt wird. Um auch die Invarianz Drehspiegelungen gegenüber zum Ausdruck zu bringen, braucht man noch die Operatoren $\mathbf{O}_I \mathbf{O}_u$. Die Gesamtheit der Operatoren $\mathbf{O}_u = 1 \mathbf{O}_u$; $\mathbf{O}_I \mathbf{O}_u$ ist das direkte Produkt der Spiegelungsgruppe $\mathbf{O}_E = 1, \mathbf{O}_I$ und der Gruppe der \mathbf{O}_u . Bezeichnet man das allgemeine Element der Gruppe des direkten Produktes der Spiegelungsgruppe und der unitären Gruppe mit¹⁾ \mathfrak{z} , so kann die ganze

¹⁾ Die \mathfrak{z} sind also Elemente einer abstrakten Gruppe, keine Matrizen, sondern Paare $J \mathfrak{u}$, wo J entweder E oder I und \mathfrak{u} ein Element der unitären Gruppe ist. Das Multiplikationsgesetz ist (vgl. Kap. XVI) $J \mathfrak{u} \cdot J_1 \mathfrak{u}_1 = J J_1 \mathfrak{u} \mathfrak{u}_1$.

Drehspiegelungssymmetrie durch die \mathbf{O}_δ , die eine zu der Gruppe der \mathfrak{z} holomorphe Gruppe bilden, in Evidenz gesetzt werden. Die \mathfrak{z} und \mathbf{O}_δ entsprechen entweder einer reinen Drehung — in diesem Fall hat \mathfrak{z} die Gestalt $E \mathbf{u}$ — oder einer Drehspiegelung, dann hat \mathfrak{z} die Gestalt $I \mathbf{u}$. Umgekehrt entsprechen aber zwei \mathfrak{z} bzw. \mathbf{O}_δ jeder Drehung und Drehspiegelung.

Wenn man nun ein äußeres Feld einführt, bleiben nur jene \mathfrak{z} bestehen¹⁾, die solchen Drehungen bzw. Drehspiegelungen entsprechen, die auch bei Vorhandensein des äußeren Feldes zur Symmetriegruppe des Systems gehören. Die diesen entsprechenden Matrizen $\mathbf{D}(\mathfrak{z})$

$$\mathbf{O}_\delta \Psi_\mu = \sum_{\mu'} \mathbf{D}(\mathfrak{z})_{\mu' \mu} \Psi_\mu \quad (4)$$

bilden eine (eindeutige!) Darstellung der Gruppe der entsprechenden \mathfrak{z} , und die verschiedenen Terme des Systems mit Feld gehören zu den verschiedenen Darstellungen dieser Gruppe. Zu dieser Gruppe ist die Symmetriegruppe des Systems nicht mehr holomorph, sondern (zweistufig) isomorph, weil je zwei \mathfrak{z} derselben Drehung oder Drehspiegelung entsprechen.

Im Falle eines homogenen, zu Z parallelen elektrischen Feldes gehören die Drehungen um Z und die Spiegelungen an den Ebenen durch Z zur Symmetriegruppe. Der Drehung um Z mit α entsprechen die \mathfrak{z} [die Gestalt der Matrizen $\mathbf{u} = \mathfrak{D}^{(1/2)}$ entnehme man (16), Kap. XV]

$$\mathfrak{z}_\alpha = E \begin{pmatrix} e^{-\frac{1}{2}i\alpha} & 0 \\ 0 & e^{\frac{1}{2}i\alpha} \end{pmatrix}; \quad \mathfrak{z}'_\alpha = E \begin{pmatrix} -e^{-\frac{1}{2}i\alpha} & 0 \\ 0 & -e^{\frac{1}{2}i\alpha} \end{pmatrix} (-\pi < \alpha \leq \pi). \quad (5)$$

Die Spiegelung an der ZY -Ebene ist das Produkt einer Inversion und einer Drehung um X mit π . Die entsprechenden \mathfrak{z} sind daher

$$\mathfrak{z}_x = I \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}; \quad \mathfrak{z}'_x = I \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}. \quad (5a)$$

Den Spiegelungen an anderen Ebenen entsprechen Produkte von (5) und (5a). Den Gruppenelementen (5) bzw. (5a) entsprechen in der Darstellung des direkten Produkts der Spiegelungsgruppe und der

¹⁾ Das heißt nur diese führen Eigenfunktionen in Eigenfunktionen über.

unitären Gruppe, die zu einem Term mit dem Spiegelungscharakter w und der Gesamtquantenzahl j gehört:

$$\mathbf{D}(\mathfrak{z}_a) = \begin{pmatrix} e^{-ij\alpha} & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & e^{ij\alpha} \end{pmatrix}; \quad \mathbf{D}(\mathfrak{z}'_a) = - \begin{pmatrix} e^{-ij\alpha} & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & e^{ij\alpha} \end{pmatrix} \quad (6)$$

bzw. [man setze in (21), Kap. XV, $a = 0$, $b = -1$ für \mathfrak{z}_x und $a = 0$, $b = 1$ für \mathfrak{z}'_x und multipliziere mit w]

$$\mathbf{D}(\mathfrak{z}_x) = \begin{pmatrix} 0 & 0 & \dots & 0 & -w \\ 0 & 0 & \dots & w & 0 \\ \vdots & \vdots & & \vdots & \vdots \\ 0 & -w & \dots & 0 & 0 \\ w & 0 & \dots & 0 & 0 \end{pmatrix} = -\mathbf{D}(\mathfrak{z}'_x). \quad (6a)$$

Die Matrizen (6), (6a) und ihre Produkte bilden eine Darstellung jener Untergruppe des direkten Produkts der Spiegelungsgruppe und der unitären Gruppe, deren Elemente den durch das elektrische Feld nicht aufgehobenen Symmetrieelementen des Systems entsprechen. Wir können diese Darstellung ausreduzieren, indem wir die Zeilen und Spalten vertauschen, so daß ihre Reihenfolge $-j, j, -j+1, j-1, \dots, -\frac{1}{2}, \frac{1}{2}$ anstatt $-j, -j+1, \dots, j-1, j$ wird. Dann zerfällt sie in lauter zweireihige irreduzible Darstellungen:

$$\mathbf{Z}^{(m)}(\mathfrak{z}_a) = \begin{pmatrix} e^{-im\alpha} & 0 \\ 0 & e^{im\alpha} \end{pmatrix}; \quad \mathbf{Z}^{(m)}(\mathfrak{z}'_a) = \begin{pmatrix} -e^{-im\alpha} & 0 \\ 0 & -e^{im\alpha} \end{pmatrix} \quad (7)$$

und

$$\mathbf{Z}^{(m)}(\mathfrak{z}_x) = (-1)^{j-m} \begin{pmatrix} 0 & -w \\ w & 0 \end{pmatrix}; \quad \mathbf{Z}^{(m)}(\mathfrak{z}'_x) = (-1)^{j-m} \begin{pmatrix} 0 & w \\ -w & 0 \end{pmatrix}, \quad (7a)$$

wobei m die Werte

$$m = j, j-1, j-2, \dots, \frac{3}{2}, \frac{1}{2} \quad (8)$$

durchläuft, so daß ein Term mit der Gesamtquantenzahl j in $j + \frac{1}{2}$ Starkoeffektkomponenten aufspaltet, deren elektrische Quantenzahlen in (8) stehen.

Nun ist es in diesem Falle so, daß die Darstellungen $\mathbf{Z}^{(m)}$ für $w = +1$ und $w = -1$ äquivalent sind, weil sie durch Transformation mit der Matrix

$$\begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}$$

ineinander übergehen. Die Terme mit denselben elektrischen Quantenzahlen, die aus positiven bzw. negativen Termen entstehen, sind in starken elektrischen Feldern bezüglich ihrer Transformations-eigenschaften und daher auch bezüglich ihrer Auswahlregeln nicht mehr zu unterscheiden. Dieses Resultat war auch zu erwarten, da auch bei gerader Elektronenzahl die Gültigkeit der Laporteschen Regel im allgemeinen aufgehoben wurde, es blieb nur ein Unterschied zwischen 0 und 0'-Termen erhalten. Bei ungerader Elektronenzahl fällt auch dieser Unterschied weg, weil überhaupt keine 0 oder 0'-Terme vorkommen.

4. Wenn man den durch das elektrische Feld erzeugten Störungsoperator in eine Reihe, wie (2), entwickelt, so folgt aus dem polaren Charakter des elektrischen Feldvektors, daß die Koeffizienten V_x , V_y , V_z polaren Charakter haben müssen. Daher verschwinden die Matrixelemente

$$(\Psi_{\mu}^{Nj}, V_z \Psi_{\mu'}^{Nj}),$$

die einen mit der ersten Potenz der Feldstärke proportionalen Effekt geben könnten, weil Ψ_{μ}^{Nj} und $\Psi_{\mu'}^{Nj}$ gleichen Spiegelungscharakter haben.

Die Aufspaltung erfolgt erst in der zweiten Näherung, die zu E^2 proportional ist, und es läßt sich zeigen, daß sie sich in dieser Näherung in einer Verschiebung und in einer zu μ^2 proportionalen Aufspaltung äußert.

5. Den größten Teil der bisher abgeleiteten Regelmäßigkeiten, soweit sie sich auf den isotropen Fall beziehen, kann man vom mathematischen Standpunkt aus in der Formel (19) des Kap. XXI

$$T_{Nj\mu; N'j'\mu'}^{(\varrho)} = s_{j'\mu'\varrho}^{(j\omega)} \delta_{\mu+\varrho, \mu'} T_{Nj; N'j'} \quad (9)$$

zusammenfassen, die das Verhältnis von Matrixelementen

$$\frac{T_{Nj\mu; N'j'\mu'}^{(\varrho)}}{T_{Nj\nu; N'j'\nu'}^{(\sigma)}} = \frac{(\Psi_{\mu}^{Nj}, T^{(\varrho)} \Psi_{\mu'}^{Nj'})}{(\Psi_{\nu}^{Nj}, T^{(\sigma)} \Psi_{\nu'}^{Nj'})} \quad (9a)$$

zu bestimmen gestattet, wenn die entsprechenden Eigenfunktionen Ψ_{μ}^{Nj} und Ψ_{ν}^{Nj} bzw. $\Psi_{\mu'}^{Nj'}$ und $\Psi_{\nu'}^{Nj'}$ zu demselben Eigenwert E_j^N bzw. $E_{j'}^{N'}$ gehören, also Partner voneinander sind, und wenn die Operatoren $T^{(\varrho)}$ und $T^{(\sigma)}$ Komponenten desselben irreduziblen Tensoroperators

$$O_R^{-1} T^{(\varrho)} O_R = \sum_{\sigma} \mathfrak{D}^{(\omega)}(R)_{\varrho\sigma} T^{(\sigma)} \quad (9b)$$

sind. Ein irreduzibler Tensor ist unter anderem ein Skalar (die zugehörige Darstellung ist $\mathfrak{D}^{(0)}$) oder auch ein Vektor, dessen Darstellung $\mathfrak{D}^{(1)}$ äquivalent ist, usw. Die $T_{Nj; N'j'}$ sind dabei Zahlen, die man mit allgemeinen Methoden nicht mehr bestimmen kann, und die auch von der speziellen Gestalt des Operatorensystems \mathbf{T} und des zugrunde gelegten Hamiltonschen Operators abhängen.

Wir hatten bisher nur eine einzige Formel, die sich nicht auf die Form (9) bringen läßt, das ist die Formel (8) für den normalen Zeeman-Effekt im Kap. XVIII. Bei ihrer Ableitung mußte bemerkt werden, daß der fragliche Vektoroperator L_z durch

$$L_z = \frac{1}{i} \frac{\partial}{\partial \varphi} P_{\{q_00\}} \text{ für } \varphi = 0 \quad (10)$$

gegeben ist. Sonst ist aber die Ableitung von Regelmäßigkeiten, die über (9) hinausgehen, nur mit Hilfe weiterer Annahmen möglich.

Ein Fall, in welchem eine solche Annahme zulässig ist, ist der sogenannte normale Kopplungsfall von Spin und Bahndrehimpuls, das ist der schon im vorangehenden Kapitel behandelte Fall, daß die Feinstrukturaufspaltung der Terme klein gegen den Abstand des jeweils nächstgelegenen Grobstrukturterms ist. Dann läßt sich nicht nur eine Gesamtquantenzahl definieren, auch die Begriffe des Multiplettsystems und der Azimutalquantenzahl werden sinnvoll. Es empfiehlt sich, dies auch äußerlich zum Ausdruck zu bringen, indem man die einfache Laufzahl N der Terme mit derselben Gesamtquantenzahl durch ein dreifaches Symbol NSL ersetzt, wo wie gewöhnlich S das Multiplettsystem, L die azimutale Quantenzahl ist, und N nur noch zur Unterscheidung der Terme mit gleichen S , L und J dient¹⁾. Im Rest dieses Kapitels wollen wir uns mit dem „normalen Kopplungsfall“ beschäftigen.

Die Eigenfunktionen haben nach (27) des vorangehenden Kapitels die Gestalt

$$\psi_m^{NSLJ} = \sum_{\mu} s_{J, \mu, m-\mu}^{(L, S)} E_{m-\mu, \mu}. \quad (11)$$

Die $E_{-S, \mu}^{NSL}$, $E_{-S+1, \mu}^{NSL}$, ..., $E_{S, \mu}^{NSL}$ sind Partner in bezug auf die Q_R und gehören zu den verschiedenen Zeilen von $\mathfrak{D}^{(S)}$, für die $E_r^{-L}, E_r^{-L+1}, \dots, E_r^L$ gilt dasselbe in bezug auf die Opera-

¹⁾ Wir benutzen hier wieder — wie immer im normalen Kopplungsfall — das Zeichen J für die Gesamtquantenzahl.

toren \mathbf{P}_R und die Darstellung $\mathfrak{D}^{(L)}$. Wenn (11) gilt, kann man noch das Verhältnis von Matrixelementen

$$(\Psi_m^{NSLJ}, \mathbf{T}^{(\sigma\varrho)} \Psi_{m'}^{N'S'L'J'}) = \mathbf{T}_{NSLJm; N'S'L'J'm'}^{(\sigma\varrho)}, \quad (12)$$

mit verschiedenen $J, J', m, m', \sigma, \varrho$, aber gleichen NSL und $N'S'L'$ berechnen, wenn die $\mathbf{T}^{(\sigma\varrho)}$ Komponenten eines Tensors sind, der in bezug auf die \mathbf{Q}_R vom Grade q , in bezug auf die \mathbf{P}_R vom Grade p und in bezug auf beide irreduzibel ist:

$$\mathbf{Q}_R^{-1} \mathbf{T}^{(\sigma\varrho)} \mathbf{Q}_R = \sum_{\sigma'} \mathfrak{D}^{(q)}(R)_{\sigma\sigma'} \mathbf{T}^{(\sigma'\varrho)}, \quad (13a)$$

$$\mathbf{P}_R^{-1} \mathbf{T}^{(\sigma\varrho)} \mathbf{P}_R = \sum_{\varrho'} \mathfrak{D}^{(p)}(R)_{\varrho\varrho'} \mathbf{T}^{(\sigma\varrho')}. \quad (13b)$$

In Wirklichkeit, d. h. in bezug auf die \mathbf{O}_R , muß dagegen \mathbf{T} nicht irreduzibel sein, es gehört

$$\begin{aligned} \mathbf{O}_R^{-1} \mathbf{T}^{(\sigma\varrho)} \mathbf{O}_R &= \mathbf{Q}_R^{-1} \mathbf{P}_R^{-1} \mathbf{T}^{(\sigma\varrho)} \mathbf{P}_R \mathbf{Q}_R \\ &= \sum_{\varrho'} \mathbf{Q}_R^{-1} \mathfrak{D}^{(p)}(R)_{\varrho\varrho'} \mathbf{T}^{(\sigma\varrho')} \mathbf{Q}_R = \sum_{\sigma'\varrho'} \mathfrak{D}^{(q)}(R)_{\sigma\sigma'} \mathfrak{D}^{(p)}(R)_{\varrho\varrho'} \mathbf{T}^{(\sigma'\varrho')} \end{aligned}$$

zu einem direkten Produkt zweier irreduzibler Darstellungen.

Da die $\Xi_{v\mu}^{NSL}$ für $v = -S, \dots, S$ in bezug auf die \mathbf{Q}_R zu $\mathfrak{D}^{(S)}$ gehören und Partner voneinander sind, gilt für sie wegen (13a) die (9) analoge Gleichung

$$(\Xi_{v\mu}^{NSL}, \mathbf{T}^{(\sigma\varrho)} \Xi_{v'\mu'}^{N'S'L'}) = \delta_{v+\sigma, v'} s_{S'v\sigma}^{(S\varrho)} t_{NSL\mu; N'S'L'\mu'}^{(\varrho)}. \quad (14a)$$

Ähnlich gilt wegen (13b) und weil die $\Xi_{v\mu}^{NSL}$ für $\mu = -L, \dots, L$ in bezug auf die \mathbf{P}_R zu $\mathfrak{D}^{(L)}$ gehören

$$(\Xi_{v\mu}^{NSL}, \mathbf{T}^{(\sigma\varrho)} \Xi_{v'\mu'}^{N'S'L'}) = \delta_{\mu+\varrho, \mu'} s_{L'\mu\varrho}^{(Lp)} \bar{t}_{NSL\tau; N'S'L'\tau'}^{(\sigma)}. \quad (14b)$$

Durch Vereinigung von (14a) und (14b) erhält man

$$(\Xi_{v\mu}^{NSL}, \mathbf{T}^{(\sigma\varrho)} \Xi_{v'\mu'}^{N'S'L'}) = \delta_{v+\sigma, v'} \delta_{\mu+\varrho, \mu'} s_{S'v\sigma}^{(S\varrho)} s_{L'\mu\varrho}^{(Lp)} t_{NSL; N'S'L'}. \quad (14)$$

Es ist daher wegen (11)

$$\begin{aligned} &(\Psi_m^{NSLJ}, \mathbf{T}^{(\sigma\varrho)} \Psi_{m'}^{N'S'L'J'}) \\ &= \sum_{\mu\mu'} s_{J\mu m-\mu}^{(LS)} s_{J'\mu' m'-\mu'}^{(L'S')} \delta_{m-\mu+\sigma, m'-\mu'} \delta_{\mu+\varrho, \mu'} s_{S'm-\mu\sigma}^{(S\varrho)} s_{L'\mu\varrho}^{(Lp)} t_{NSL; N'S'L'} \\ &= \sum_{\mu} s_{J\mu m-\mu}^{(LS)} s_{J'\mu+q m-\mu+\sigma}^{(L'S')} \delta_{m+\sigma+\varrho, m'} s_{S'm-\mu\sigma}^{(S\varrho)} s_{L'\mu\varrho}^{(Lp)} t_{NSL; N'S'L'}. \quad (15) \end{aligned}$$

Durch diese Formel kann man im Prinzip das Verhältnis von Matrixelementen (12) mit gleichen NSL und $N'S'L'$ immer bestimmen.

Dabei sind, wie auch in (9), alle $s_{J\mu r}^{(LS)}$, deren erster unterer Index größer als die Summe der beiden oberen Indizes oder kleiner

als der Absolutwert ihrer Differenz ist ($J > L + S$ oder $J < |L - S|$), gleich Null zu setzen. Dasselbe gilt für $|\mu| > L$, für $|\nu| > S$ oder $|\mu + \nu| > J$.

6. Zunächst mag es einem so scheinen, als ob die durch (13) definierte Operatorenklasse eine sehr gekünstelte wäre. Bei näherem Zusehen zeigt es sich aber, daß fast alle wichtigen Operatoren Tensorkomponenten oder Summen von Tensorkomponenten dergestalter Operatoren sind. Vor allem sind die spinfreien Operatoren den \mathbf{Q}_R gegenüber alle symmetrische Operatoren, also Skalare, so daß für diese (13a) mit $q = 0$ gilt. Sie transformieren sich daher unter der Wirkung der \mathbf{P}_R ebenso wie unter der Wirkung der \mathbf{O}_R , und sind auch in bezug auf die ersteren Skalare, Vektoren usw., wenn sie es in Wirklichkeit sind.

Bei spinfreien Operatoren, wie immer, wenn $q = 0$ gilt, zeigt (15) unmittelbar — was wir auch schon von früher her wissen —, daß die skalaren Produkte (15) verschwinden, wenn $S' \neq S$ ist. Ist auch $p = 0$, d. h. der Operator auch in bezug auf die \mathbf{P}_R , also auch in bezug auf die \mathbf{O}_R ein Skalar, so muß auch $L' = L$ sein, und da auch $\varrho = \sigma = 0$ gilt (ein Skalar hat nur eine 0-Komponente), kann man die Summation in (15) wegen $s_{L\mu_0}^{(L_0)} = s_{S'm-\mu_0}^{(S_0)} = 1$ mit Hilfe der Orthogonalitätsrelation der s (28), Kap. XVII,

$$\sum_{\mu} s_{J\mu m-\mu}^{(L_S)} s_{J'\mu m-\mu}^{(L_S)} = \delta_{J,J'} \quad (16)$$

ausführen. Man erhält für die Matrixelemente eines Operators, der beiderseits ein Skalar ($p = q = 0$) ist, daß sie

$$(\Psi_m^{NSLJ}, \mathbf{T} \Psi_{m'}^{N'S'L'J'}) = \delta_{SS'} \delta_{LL'} \delta_{JJ'} \delta_{mm'} t_{NSL; N'S'L'} \quad (17)$$

erstens für $J' \neq J$ oder $m' \neq m$ verschwinden und für $J' = J$ und $m' = m$ von m unabhängig sind — das ist nur die Regel für Operatoren, die den \mathbf{O}_R gegenüber symmetrisch sind —, und zweitens, daß sie auch von J unabhängig für alle Feinstrukturkomponenten eines Grobstrukturterms gleich sind. Hierzu genügt es nicht mehr, daß \mathbf{T} in bezug auf die $\mathbf{O}_R = \mathbf{P}_R \mathbf{Q}_R$ ein Skalar sei, es muß sowohl in bezug auf die \mathbf{P}_R wie auch in bezug auf die \mathbf{Q}_R ein Skalar sein, und auch der „normale Kopplungsfall“ vorliegen.

Ein Operator, der beiderseits ein Skalar ist, ist z. B. der Hamiltonsche Operator H_0 der einfachen Schrödingerschen Theorie. Es ist in diesem Falle

$$(\Psi_m^{NSLJ}, H_0 \Psi_{m'}^{N'S'L'J'}) = E^{NSL} \delta_{NN'} \delta_{SS'} \delta_{LL'} \delta_{JJ'} \delta_{mm'}.$$

wo E^{NSL} ein Eigenwert der einfachen Schrödinger-Gleichung ist, der — da diese ja keine Feinstrukturauflösung gibt — für alle Feinstrukturkomponenten gleich ist.

Die Hönls-Kronigschen Intensitätsformeln

Ist $T^{(\sigma\varrho)} = V^{(\varrho)}$ in bezug auf die Q_R ein Skalar, in bezug auf die P_R ein Vektor, wie z. B. der Operator des Multiplizierens mit

$$\frac{1}{\sqrt{2}} \sum_k (y_k + i x_k), \quad \sum_k x_k, \quad \frac{1}{\sqrt{2}} \sum_k (-y_k + i x_k),$$

der für die Übergangswahrscheinlichkeiten maßgebend ist, so ist $p = 1$, $q = 0$, und man kann für (15)

$$V_{N' S L J m; N' S' L' J' m'}^{(0)} = \delta_{S S'} \delta_{m+\varrho, m'} \sum_{\mu} s_{J \mu m-\mu}^{(L S)} s_{J' \mu+\varrho m-\mu}^{(L' S)} s_{L' \mu \varrho}^{(L 1)} v_{N S L; N' S L'} \quad (15a)$$

schreiben. Es muß diesmal, damit die Matrixelemente (15a) nicht verschwinden sollen, $S' = S$ und $L' = L$ oder $L' = L \pm 1$ (und $J' = J$ oder $J' = J \pm 1$) sein. Da wir die Verhältnisse der Matrixelemente

$$V_{N' S L J m; N' S L' J' m'}^{(0)} = s_{J' m \varrho}^{(J 1)} \delta_{m+\varrho, m'} V_{N S L J; N' S L' J'} \quad (18)$$

für verschiedene m , m' , ϱ schon kennen, können wir für diese spezielle Werte einsetzen und nur für diese speziellen Werte die Verhältnisse für verschiedene J , J' berechnen. Die Formeln für die s werden am einfachsten, wenn wir $m = J$, $m' = J'$ und $\varrho = m' - m = J' - J$ einsetzen. Es ist dann z. B. für $L' = L - 1$, $J' = J + 1$ wegen (27b), Kap. XVII, und der Tabelle auf S. 208

$$\begin{aligned} & s_{J \mu J-\mu}^{(L S)} s_{L-1 \mu 1}^{(L 1)} \\ &= \frac{(-1)^{L-\mu} \sqrt{(L+S-J)! (2J+1)!}}{\sqrt{(J+L+S+1)! (J+S-L)! (J-S+L)!}} \sqrt{\frac{(L+\mu)! (S+J-\mu)!}{(L-\mu)! (S-J+\mu)!}} \frac{\sqrt{(L-\mu-1)(L-\mu)}}{\sqrt{2} L (2L+1)} \\ &= \frac{(-1)^{L-1-(\mu+1)} \sqrt{(L-1+S-J-1)! (2J+3)!}}{\sqrt{(J+1+L-1+S+1)! (J+1+S-L+1)! (J+1-S+L-1)!}} \sqrt{\frac{(L+\mu)! (S+J-\mu)!}{(L-\mu-2)! (S-J+\mu)!}} \\ & \quad \sqrt{\frac{(L+S-J-1)(L+S-J)(J+S-L+1)(J+S-L+2)}{(2J+2)(2J+3)2L(2L+1)}} \\ &= s_{J+1 \mu+1 J-\mu}^{(L-1 S)} \sqrt{\frac{(L+S-J-1)(L+S-J)(J+S-L+1)(J+S-L+2)}{(2J+2)(2J+3)2L(2L+1)}}, \end{aligned}$$

und daher erhält man für die Summe (15a) mit Hilfe von (16)

$$\begin{aligned} \mathcal{V}_{N' S L J J; N' S L - 1 J + 1 J + 1}^{(1)} \\ = \sum_{\mu} s_{J, \mu, J - \mu}^{(L, S)} s_{L - 1, \mu, 1}^{(L, 1)} s_{J + 1, \mu + 1, J - \mu}^{(L - 1, S)} v_{N S L; N' S L - 1} \\ = \sqrt{\frac{(L + S - J - 1)(L + S - J)(J + S - L + 1)(J + S - L + 2)}{(2J + 2)(2J + 3)2L(2L + 1)}} v_{N S L; N' S L - 1}, \end{aligned}$$

und so wegen $s_{J + 1, J, 1}^{(J, 1)} = 1$ für das $\mathcal{V}_{N S L J; N' S L' J'}$ in (18)

$$\begin{aligned} \mathcal{V}_{N S L J; N' S L - 1 J + 1} \\ = \sqrt{\frac{(L + S - J - 1)(L + S - J)(J + S - L + 1)(J + S - L + 2)}{(2J + 2)(2J + 3)2L(2L + 1)}} v_{N S L; N' S L - 1}. \quad (19) \end{aligned}$$

Ganz entsprechend kann man die anderen Formeln

$$\left. \begin{aligned} \mathcal{V}_{N S L J; N' S L - 1 J} \\ = \sqrt{\frac{(L + S - J)(J + S - L + 1)(J - S + L)(J + L + S + 1)}{2J(2J + 2)L(2L + 1)}} v_{N S L; N' S L - 1}, \\ \mathcal{V}_{N S L J; N' S L - 1 J - 1} \\ = \sqrt{\frac{(J - S + L - 1)(J - S + L)(J + L + S)(J + L + S + 1)}{2J(2J - 1)2L(2L + 1)}} v_{N S L; N' S L - 1}, \end{aligned} \right\} \quad (19')$$

ableiten, wodurch für $L' = L - 1$ die miteinander vergleichbaren Matrixelemente auf eine gemeinsame Größe $v_{N S L; N' S L - 1}$ zurückgeführt sind. Ähnlich erhält man für $L' = L$

$$\left. \begin{aligned} \mathcal{V}_{N S L J; N' S L J + 1} \\ = \sqrt{\frac{(L + S - J)(J - S + L + 1)(J + S - L + 1)(J + L + S + 2)}{(2J + 2)(2J + 3)2L(L + 1)}} v_{N S L; N' S L}, \\ \mathcal{V}_{N S L J; N' S L J - 1} \\ = - \sqrt{\frac{(L + S - J + 1)(J - S + L)(J + S - L)(J + L + S + 1)}{2J(2J - 1)2L(L + 1)}} v_{N S L; N' S L}, \end{aligned} \right\} \quad (19a)$$

nur daß $s_{L, \mu, J' - J}^{(L, 1)}$ zum Teil mit dem ersten, zum Teil mit dem zweiten Faktor in (15a) zu vereinigen ist. Nur die Ableitung der entsprechenden Formel für $L' = L$, $J' = J$ muß etwas anders geschehen: man muß $s_{L, \mu, 0}^{(L, 1)}$ in die beiden Teile

$$s_{L, \mu, 0}^{(L, 1)} = \frac{\mu}{\sqrt{L(L + 1)}} = \frac{L}{\sqrt{L(L + 1)}} - \frac{\sqrt{L - \mu}\sqrt{L - \mu}}{\sqrt{L(L + 1)}}$$

zerlegen. Die Summation über das erste Glied läßt sich in (15a) mit Hilfe der Orthogonalitätsrelation (16) direkt ausführen und gibt

$$v_{NSL; N'SL} L/\sqrt{L(L+1)}.$$

Auch über das zweite Glied läßt sich die Summation mit Hilfe von

$$\begin{aligned} s_{J\mu J-\mu}^{(L S)} \sqrt{L-\mu} \\ = \frac{(-1)^{L-\mu} \sqrt{(L+S-J)!(2J+1)!}}{\sqrt{(J+L+S+1)!(J+S-L)!(J-S+L)!}} \sqrt{\frac{(L+\mu)!(J+S-\mu)!}{(L-\mu-1)!(S-J+\mu)!}} \\ = -s_{J+\frac{1}{2}, \mu+\frac{1}{2}, J-\mu}^{(L-\frac{1}{2}, S)} \sqrt{\frac{(L+S-J)(J+S-L+1)}{2J+2}} \end{aligned}$$

durch Anwendung der Orthogonalitätsrelation (16) ausführen. Man erhält so für (15a)

$$\begin{aligned} \sum_{\mu} (s_{J\mu J-\mu}^{(L S)})^2 s_{L\mu 0}^{(L 1)} &= \frac{L}{\sqrt{L(L+1)}} - \frac{(L+S-J)(J+S-L+1)}{(2J+2)\sqrt{L(L+1)}} \\ &= \frac{J(J+1)+L(L+1)-S(S+1)}{2(J+1)\sqrt{L(L+1)}} \end{aligned}$$

und hieraus schließlich

$$V_{NSLJ; N'SLJ} = \frac{J(J+1)+L(L+1)-S(S+1)}{2\sqrt{J(J+1)}\sqrt{L(L+1)}} v_{NSL; N'SL}. \quad (19b)$$

Das Verhältnis der Matrixelemente für $L' = L + 1$ erhält man entweder in derselben Weise durch direkte Rechnung oder durch die Bemerkung, daß wegen des hermiteischen Charakters von $V^{(0)}$

$$V_{N'SL-1J'm'; NSLJm}^{(0)} = V_{NSLJm; N'SL-1J'm'}^{(0)*}$$

diese Verhältnisse aus den Verhältnissen (19) zu berechnen sind.

Diese Formeln (19) enthalten die berühmten Hönl-Kronig-schen Intensitätsformeln für das Verhältnis der Intensitäten der Feinstrukturkomponenten einer Linie. Um die gesamte Intensität einer Feinstrukturkomponente $NSLJ \rightarrow N'S'L'J'$ zu erhalten, muß man die Intensitäten $|V_{NSLJm; N'SL'J'm'}^{(0)}|^2$ der einzelnen Zeemankomponenten über alle m, m' und ϱ summiieren

$$\begin{aligned} \sum_{m'm} \sum_{\varrho} |V_{NSLJm; N'SL'J'm'}^{(0)}|^2 &= \sum_{mm'} |V_{NSLJ; N'SL'J'} s_{J'm'm'-m}^{(J1)}|^2 \\ &= |V_{NSLJ; N'SL'J'}|^2 \sum_{m'} 1 = (2J'+1) |V_{NSLJ; N'SL'J'}|^2. \end{aligned}$$

Die Gesamtintensität der Linie $J \rightarrow J'$ ist daher im wesentlichen durch die $V_{NSLJ; N'SL'J'}$ bestimmt.

Die Landésche g -Formel

7. Eine andere Anwendung der Formel (19b) bezieht sich auf den Zeeman-Effekt. Die Wechselwirkung des Magnetfeldes mit dem Atom gibt zu zwei Zusatzgliedern für den Hamiltonschen Operator Anlaß. Das erste Glied ist $\mathbf{V} = \eta \mathfrak{H} \mathbf{L}_z$, wo $\eta = \frac{e\hbar}{4\pi m_0 c}$ und m_0 die Masse des Elektrons ist. Dieses Glied röhrt von der Wechselwirkung des Magnetfeldes mit den durch die Bewegung der Elektronen erzeugten Strömen her und hatte auch in der einfachen Theorie (Kap. XVIII) dieselbe Gestalt. Es war [vgl. (7) und (7b), Kap. XVIII]

$$\mathbf{L}_z \psi_{x\mu}^{NSL} = \mu \psi_{x\mu}^{NSL} \quad (20)$$

einfach die Multiplikation mit der Z -Komponente des Drehimpulses. Daher ist nach (25) des vorangehenden Kapitels

$$\mathbf{L}_z \mathfrak{E}_{r\mu}^{NSL} = \sum_x \mathbf{L}_z \psi_{x\mu}^{NSL} f_{x\nu}^{(S)} = \sum_x \mu \psi_{x\mu}^{NSL} f_{x\nu}^{(S)} = \mu \mathfrak{E}_{r\mu}^{NSL},$$

so daß $v_{NSL; NSL}$ in diesem Falle, wie der Vergleich mit (14) zeigt,

$$(\mathfrak{E}_{r\mu}^{NSL}, \mathbf{L}_z \mathfrak{E}_{r\mu}^{NSL}) = \mu = v_{NSL; NSL} \frac{\mu}{\sqrt{L(L+1)}} \\ v_{NSL; NSL} = \sqrt{L(L+1)}$$

wird. Es ergibt sich so schließlich nach (19b)

$$V_{NSLJ; NSLJ} = \frac{J(J+1) + L(L+1) - S(S+1)}{2\sqrt{J(J+1)}}. \quad (20a)$$

Die Matrixelemente von \mathbf{L}_z sind daher nach (18)

$$(\Psi_m^{NSLJ}, \mathbf{L}_z \Psi_m^{NSLJ}) = m \frac{J(J+1) + L(L+1) - S(S+1)}{2J(J+1)}. \quad (21)$$

Das zweite Glied $\bar{\mathbf{V}}$ im magnetischen Zusatzoperator für den Hamiltonschen Operator röhrt von der Wechselwirkung des Magnetfeldes mit den magnetischen Momenten der Elektronen her: es ist gleich dem skalaren Produkt des magnetischen Momentes mit der Feldstärke, also ($\eta = e\hbar/4\pi m_0 c$)

$$\eta (\mathbf{s}_x \mathfrak{H}_x + \mathbf{s}_y \mathfrak{H}_y + \mathbf{s}_z \mathfrak{H}_z) \quad (22)$$

bzw. bei mehreren Elektronen gleich der entsprechenden Summe von Gliedern (22). Hat das Magnetfeld die Z -Richtung, so ist $\bar{\mathbf{V}} = \eta \mathfrak{H} \mathbf{S}_z$, wobei \mathbf{S}_z die Z -Komponente des Gesamtspins, das Multiplizieren mit

$$s_1 + s_2 + \cdots + s_n \quad (22a)$$

bedeutet, weil \mathbf{s}_z nach (21), Kap. XX, das Multiplizieren mit s ist. Es war

$$\mathbf{L}_z \Psi = \frac{1}{i} \frac{\partial}{\partial \alpha} \mathbf{P}_{\{\alpha 00\}} \Psi \quad (\text{für } \alpha = 0), \quad (23a)$$

und ähnlich soll gezeigt werden, daß

$$\frac{1}{2} \mathbf{S}_z \Psi = \frac{1}{i} \frac{\partial}{\partial \alpha} \mathbf{Q}_{\{\alpha 00\}} \Psi \quad (\text{für } \alpha = 0) \quad (23b)$$

gilt. Der Faktor $\frac{1}{2}$ in (23b) röhrt daher, daß der Spin nur einen Drehimpuls $h/4\pi$ hat. Es ist wegen $\mathbf{D}^{(1/2)}(\{\alpha 00\})_{\frac{1}{2}s_k, \frac{1}{2}t_k} = \delta_{st} e^{\frac{1}{2}i s_k \alpha}$ nach der Definitionsgleichung von \mathbf{Q}_R (6b), Kap. XXI

$$\begin{aligned} & \mathbf{Q}_{\{\alpha 00\}} \Psi (\dots, x_k, y_k, z_k, s_k, \dots) \\ &= \sum_{t_1 \dots t_n = \pm 1} \dots \mathbf{D}^{(1/2)}(\{\alpha 00\})_{\frac{1}{2}s_k, \frac{1}{2}t_k} \dots \Psi (\dots, x_k, y_k, z_k, t_k, \dots) \\ &= \sum_{t_1 \dots t_n} \dots \delta_{s_k t_k} e^{\frac{1}{2}i s_k \alpha} \dots \Psi (\dots, x_k, y_k, z_k, t_k, \dots) = e^{\frac{1}{2}i(s_1 + \dots + s_n)\alpha} \Psi, \end{aligned}$$

und hieraus folgt unmittelbar (23b).

Es gilt daher auch

$$(\Xi_{r\mu}^{NSL}, \frac{1}{2} \mathbf{S}_z \Xi_{r\mu}^{NSL}) = \nu. \quad (22b)$$

Nun ist \mathbf{S}_z in bezug auf die \mathbf{P}_R ein Skalar und in bezug auf die \mathbf{Q}_R ein Vektor, während \mathbf{L}_z in bezug auf die \mathbf{P}_R ein Vektor und in bezug auf die \mathbf{Q}_R ein Skalar war. Bei der Berechnung der Ausdrücke $(\Psi_m^{NSLJ}, \mathbf{S}_z \Psi_m^{NSLJ})$ aus den Größen (22b) muß daher die Rolle von L und S vertauscht werden, und man erhält an Stelle von (21)

$$(\Psi_m^{NSLJ}, \frac{1}{2} \mathbf{S}_z \Psi_m^{NSLJ}) = m \frac{J(J+1) + S(S+1) - L(L+1)}{2J(J+1)}. \quad (24)$$

Die Matrixelemente des ganzen magnetischen Zusatzoperators $\mathbf{V} + \bar{\mathbf{V}} = \eta \mathfrak{H} \mathbf{L}_z + \eta \mathfrak{H} \mathbf{S}_z$ lassen sich aus (21) und (24) berechnen. Die Verschiebung der Zeemankomponente mit der magnetischen Quantenzahl m ergibt sich zu

$$\Delta E_m = \frac{e \hbar \mathfrak{H}}{4 \pi m_0 c} m \frac{3J(J+1) + S(S+1) - L(L+1)}{2J(J+1)}. \quad (25)$$

Dies ist die bekannte Landésche g -Formel. Sie besagt, daß wegen der doppelten Größe des mit dem Spin verbundenen magne-

tischen Moments die verschiedenen Terme im Magnetfeld verschieden aufspalten. Man erhält die richtige Aufspaltung, wenn man die Aufspaltung $\eta \mathfrak{H} m$ des normalen Zeemaneffekts mit

$$g = 1 + \frac{J(J+1) + S(S+1) - L(L+1)}{2J(J+1)} \quad (25a)$$

multipliziert.

Die Ableitung der Formeln (19) und anderer analoger Formeln kann oft einfacher, als es im Text geschah, von der Bemerkung ausgehend erfolgen, daß die zu berechnenden Verhältnisse von Matrixelementen nach (15) vom speziellen mechanischen Problem wie auch von der Gestalt der Operatoren $T^{(q,o)}$ unabhängig sind und nur von ihren Transformationseigenschaften abhängen. In der Tat sind die zu berechnenden Verhältnisse lediglich aus den $s_{J,\mu,\nu}^{(L,S)}$ usw. zusammengesetzte Ausdrücke. Die s sind aber reine, in (27), Kap. XVII, gegebene Zahlen, so daß diese Verhältnisse für alle Operatoren mit gleichen Transformationseigenschaften dieselben sind. Man kann sie für irgendeinen Tensor der verlangten Transformationseigenschaft (13) berechnen und das Resultat auf alle Tensoren mit gleichem p und q übertragen.

Die Intervallregel

8. Als ein Beispiel hierzu sei noch die Landésche Intervallregel abgeleitet, d.h. das Verhältnis der Termverschiebungen

$$(\Psi_m^{NSLJ}, H_1 \Psi_m^{NSLJ}) = \Delta E_J^{NSL}$$

der verschiedenen Feinstrukturkomponenten desselben Grobstrukturterms berechnet, wo H_1 der durch die magnetischen Momente der Elektronen bedingte Zusatzoperator zum einfachen Schrödingerischen Energieoperator ist.

Der Zusatzoperator H_1 besteht aus zwei Teilen. Der erste Teil trägt der Wechselwirkung der magnetischen Momente der Elektronen mit den durch ihre Bewegung erzeugten Strömen, der zweite Teil der Wechselwirkung der magnetischen Momente untereinander Rechnung.

Der erste und fast immer wesentlichere Teil besteht aus einer Summe von n Ausdrücken $\mathbf{B} = \mathbf{B}_1 + \mathbf{B}_2 + \dots + \mathbf{B}_n$, wo \mathbf{B}_k der Wechselwirkung des magnetischen Moments des k -ten Elektrons mit den Strömen entspricht; \mathbf{B}_k wirkt, abgesehen von den Cartesi-

schen Koordinaten, nur auf die Spinkoordinate des k -ten Elektrons ein¹⁾:

$$\mathbf{B}_k = \mathbf{s}_{kx} \mathbf{V}_{kx} + \mathbf{s}_{ky} \mathbf{V}_{ky} + \mathbf{s}_{kz} \mathbf{V}_{kz}, \quad (26)$$

wo die \mathbf{V}_{kx} , \mathbf{V}_{ky} , \mathbf{V}_{kz} spinfreie Operatoren sind. Da \mathbf{B}_k in bezug auf die \mathbf{O}_R ein Skalar sein muß, und \mathbf{s}_{kx} , \mathbf{s}_{ky} , \mathbf{s}_{kz} die Komponenten eines Vektoroperators sind, muß dies auch für \mathbf{V}_{kx} , \mathbf{V}_{ky} , \mathbf{V}_{kz} gelten.

An Stelle der Verhältnisse der Matrixelemente von \mathbf{B}

$$(\Psi_m^{NSLJ}, \mathbf{B} \Psi_m^{NSLJ}) : (\Psi_m^{NSLJ'}, \mathbf{B} \Psi_m^{NSLJ'}),$$

die wir zu berechnen haben, können wir auch diese Verhältnisse für ein \mathbf{B}_k berechnen, diese Verhältnisse sind ja für alle k gleich. Weiter sind $\mathbf{s}_{kx} \mathbf{V}_{kx}$, $\mathbf{s}_{ky} \mathbf{V}_{ky}$, $\mathbf{s}_{kz} \mathbf{V}_{kz}$ die xx , yy , zz -Komponenten eines Tensors, der (13) erfüllt und ein beiderseitiger Vektor ($p = q = 1$) ist. Die Ausdrücke

$$\begin{aligned} & (\Psi_m^{NSLJ}, \mathbf{s}_{kx} \mathbf{V}_{kx} \Psi_m^{NSLJ}); (\Psi_m^{NSLJ}, \mathbf{s}_{ky} \mathbf{V}_{ky} \Psi_m^{NSLJ}); \\ & (\Psi_m^{NSLJ}, \mathbf{s}_{kz} \mathbf{V}_{kz} \Psi_m^{NSLJ}) \end{aligned} \quad (27)$$

stehen daher sowohl untereinander, wie auch mit den entsprechenden Ausdrücken, in denen J' an Stelle von J steht, in einem konstanten, für alle gleichgearteten Tensoren gemeinsamen Verhältnis. Dasselbe gilt von der Summe der drei Ausdrücke in (27), so daß es genügt, das Verhältnis der Matrixelemente

$$(\Psi_m^{NSLJ}, (\mathbf{T}^{(xx)} + \mathbf{T}^{(yy)} + \mathbf{T}^{(zz)}) \Psi_m^{NSLJ}) \quad (28)$$

für verschiedene J zu berechnen, wenn \mathbf{T} ein beliebiger beiderseitiger Vektoroperator ist. Diesen Operator werden wir natürlich so wählen, daß die Berechnung von (28) möglichst leicht auszuführen ist. Es sei

$$\mathbf{T}^{(xx)} = \mathbf{L}_x \mathbf{S}_x; \quad \mathbf{T}^{(yy)} = \mathbf{L}_y \mathbf{S}_y; \quad \mathbf{T}^{(zz)} = \mathbf{L}_z \mathbf{S}_z; \dots \quad (29)$$

¹⁾ Jeder Operator, der nur auf die k -te Spinkoordinate einwirkt, läßt sich in der Gestalt $\mathbf{S} + \mathbf{s}_{kx} \mathbf{V}_{kx} + \mathbf{s}_{ky} \mathbf{V}_{ky} + \mathbf{s}_{kz} \mathbf{V}_{kz}$ schreiben, wo \mathbf{S} , \mathbf{V}_{kx} , \mathbf{V}_{ky} , \mathbf{V}_{kz} spinfreie Operatoren sind. In (26) fehlt das Glied \mathbf{S} , das aber in \mathbf{B}_k ein beiderseitiger Skalar sein müßte und nach (17) nur eine für alle Feinstrukturkomponenten gleiche Verschiebung und keine Aufspaltung geben würde.

Zunächst ist

$$\frac{\partial}{\partial \alpha} \mathbf{O}_{\{\alpha 00\}} = \mathbf{Q}_{\{\alpha 00\}} \frac{\partial}{\partial \alpha} \mathbf{P}_{\{\alpha 00\}} + \mathbf{P}_{\{\alpha 00\}} \frac{\partial}{\partial \alpha} \mathbf{Q}_{\{\alpha 00\}}, \quad (30)$$

und für $\alpha = 0$ ergibt das nach (23 a), (23 b):

$$\mathbf{L}_z + \frac{1}{2} \mathbf{S}_z = \frac{1}{i} \frac{\partial}{\partial \alpha} \mathbf{O}_{\{\alpha 00\}} \quad (\text{für } \alpha = 0). \quad (30a)$$

Hieraus folgt dann, da $\mathbf{O}_{\{\alpha 00\}} \Psi_m = e^{im\alpha} \Psi_m$ ist,

$$(\mathbf{L}_z + \frac{1}{2} \mathbf{S}_z) \Psi_m = m \Psi_m \quad (31)$$

und

$$(\mathbf{L}_z + \frac{1}{2} \mathbf{S}_z)^2 \Psi_m = m^2 \Psi_m, \quad (31a)$$

woraus man

$$\begin{aligned} & \sum_{m=-J}^J (\Psi_m^{NSLJ}, (\mathbf{L}_z + \frac{1}{2} \mathbf{S}_z)^2 \Psi_m^{NSLJ}) \\ &= \sum_{m=-J}^J m^2 = \frac{J(J+1)(2J+1)}{3} \end{aligned} \quad (32)$$

gewinnt. Man kann hierin auch x und y für z einsetzen, da nach der Summation über m keine Achse mehr ausgezeichnet ist. Ist nämlich \mathbf{O}_R eine Drehung, die Z in X überführt, so ist (32) gleich

$$\begin{aligned} & \sum_m (\mathbf{O}_{R^{-1}} \Psi_m^{NSLJ}, \mathbf{O}_{R^{-1}} (\mathbf{L}_z + \frac{1}{2} \mathbf{S}_z)^2 \mathbf{O}_R \cdot \mathbf{O}_{R^{-1}} \Psi_m^{NSLJ}) \\ &= \sum_{m' m''} \mathfrak{D}^{(J)}(R^{-1})_{m' m}^* \mathfrak{D}^{(J)}(R^{-1})_{m'' m} (\Psi_{m'}^{NSLJ}, (\mathbf{L}_x + \frac{1}{2} \mathbf{S}_x)^2 \Psi_{m''}^{NSLJ}) \\ &= \sum_{m' m''} \delta_{m' m''} (\Psi_{m'}^{NSLJ}, (\mathbf{L}_x + \frac{1}{2} \mathbf{S}_x)^2 \Psi_{m''}^{NSLJ}) \\ &= \sum_m (\Psi_m^{NSLJ}, (\mathbf{L}_x + \frac{1}{2} \mathbf{S}_x)^2 \Psi_m^{NSLJ}). \end{aligned}$$

Daher wird

$$\begin{aligned} & \sum_m (\Psi_m^{NSLJ}, [(\mathbf{L}_x + \frac{1}{2} \mathbf{S}_x)^2 + (\mathbf{L}_y + \frac{1}{2} \mathbf{S}_y)^2 + (\mathbf{L}_z + \frac{1}{2} \mathbf{S}_z)^2] \Psi_m^{NSLJ}) \\ &= J(J+1)(2J+1). \end{aligned} \quad (33)$$

Da nun $(\mathbf{L}_x + \frac{1}{2} \mathbf{S}_x)^2 + (\mathbf{L}_y + \frac{1}{2} \mathbf{S}_y)^2 + (\mathbf{L}_z + \frac{1}{2} \mathbf{S}_z)^2$ ein Skalar, d. h. ein in bezug auf die \mathbf{O}_R symmetrischer Operator ist, sind die $2J+1$ Glieder der Summe links in (33) alle gleich

$$\begin{aligned} & (\Psi_m^{NSLJ}, [(\mathbf{L}_x + \frac{1}{2} \mathbf{S}_x)^2 + (\mathbf{L}_y + \frac{1}{2} \mathbf{S}_y)^2 + (\mathbf{L}_z + \frac{1}{2} \mathbf{S}_z)^2] \Psi_m^{NSLJ}) \\ &= J(J+1). \end{aligned} \quad (33a)$$

Ebenso zeigt man, daß wegen

$$\mathbf{L}_z \Xi_{\nu\mu}^{NSL} = \mu \Xi_{\nu\mu}^{NSL} \quad (34)$$

auch

$$(\Xi_{\nu\mu}^{NSL}, (\mathbf{L}_x^2 + \mathbf{L}_y^2 + \mathbf{L}_z^2) \Xi_{\nu\mu}^{NSL}) = L(L+1) \quad (35)$$

und nach (17) und (14) auch

$$(\Psi_m^{NSLJ}, (\mathbf{L}_x^2 + \mathbf{L}_y^2 + \mathbf{L}_z^2) \Psi_m^{NSLJ}) = L(L+1), \quad (35a)$$

gilt, da $\mathbf{L}_x^2 + \mathbf{L}_y^2 + \mathbf{L}_z^2$ ein beiderseitiger Skalar ist. Entsprechend folgt aus

$$\frac{1}{2} \mathbf{S}_z \Xi_{\nu\mu}^{NSL} = \frac{1}{i} \frac{\partial}{\partial \alpha} \mathbf{Q}_{\{\alpha 00\}} \Xi_{\nu\mu}^{NSL} = \nu \Xi_{\nu\mu}^{NSL} \quad (\alpha = 0), \quad (36)$$

$$(\Psi_m^{NSLJ}, (\frac{1}{4} \mathbf{S}_x^2 + \frac{1}{4} \mathbf{S}_y^2 + \frac{1}{4} \mathbf{S}_z^2) \Psi_m^{NSLJ}) = S(S+1). \quad (36a)$$

Durch Abziehen von (35a) und (36a) aus (33a) ergibt sich schließlich

$$\begin{aligned} & (\Psi_m^{NSLJ}, (\mathbf{L}_x \mathbf{S}_x + \mathbf{L}_y \mathbf{S}_y + \mathbf{L}_z \mathbf{S}_z) \Psi_m^{NSLJ}) \\ & \qquad \qquad \qquad = J(J+1) - L(L+1) - S(S+1). \end{aligned} \quad (37)$$

Nach dem Vorangehenden stehen auch die Verschiebungen ΔE_J^{NSL} der Feinstrukturkomponenten desselben Grobstrukturterms in dem durch (37) angegebenen Verhältnis:

$$\Delta E_J^{NSL} = \epsilon_{NSL} [J(J+1) - L(L+1) - S(S+1)]. \quad (37a)$$

Die Differenzen der Termverschiebungen zweier aufeinanderfolgender Feinstrukturkomponenten

$$\Delta E_{J+1}^{NSL} - \Delta E_J^{NSL} = 2\epsilon_{NSL}(J+1) \quad (37b)$$

verhalten sich demnach so wie die größeren der entsprechenden Gesamtquantenzahlen. Das ist die Landésche Intervallregel.

Die Landésche Intervallregel kann nur im normalen Koppellungsfall Gültigkeit beanspruchen, d. h. wenn die Feinstrukturauflösung klein ist im Verhältnis zu den Abständen der Grobstrukturterme voneinander. Sie enthält zudem noch die Annahme, daß man die Wechselwirkung der magnetischen Momente der Spins untereinander vernachlässigen kann, eine Annahme, die — wie Heisenberg gezeigt hat¹⁾ — bei den ganz leichten Elementen, insbesondere bei He, versagt. Die Intervallregel gilt daher am besten bei Elementen mittlerer Ordnungszahl.

¹⁾ Vgl. Zeitschr. f. Phys. **39**, 499, 1926.

Die Wechselwirkung der magnetischen Momente der Spins besteht ihrerseits aus zwei Teilen. Der erste Teil ist ein beiderseitiger Skalar, er beeinflußt daher die Feinstruktur gar nicht. Der zweite Teil gehört beiderseits zu $\mathfrak{D}^{(2)}$. Man erhält im ganzen eine Termverschiebung, die, abgesehen von einem von J unabhängigen Glied, zu $[J(J+1) - L(L+1) - S(S+1)]^2$ proportional ist¹⁾. Das Verhältnis der Proportionalitätskonstante dieses Gliedes zu ϵ_{NSL} läßt sich aus allgemeinen Überlegungen nicht mehr bestimmen, so daß man, wenn man die Wirkung dieses Gliedes nicht vernachlässigen will, auf eine zweikonstantige Formel angewiesen ist.

XXIV. Das Aufbauprinzip

1. Das Aufbauprinzip soll eine Abschätzung der Lage der Energieniveaus der Atome ermöglichen. Durch ihre Auswahlregeln, Aufspaltungsbilder usw. kann man zwar die Charakteristika der einzelnen Terme, wie z. B. die Azimutalquantenzahl, im Prinzip bestimmen, doch ist auch hierbei eine Abschätzung der Termwerte, d. h. der Stellen, wo man einen Term vorgegebener Art zu suchen hat, von großer Bedeutung.

Die eigentliche Bedeutung des Aufbauprinzips liegt aber nicht in seiner Verwendbarkeit bei der Analyse verwickelter Spektren, sondern in der Tatsache, daß die Lage der Energieniveaus die wichtigsten physikalischen und chemischen Eigenschaften des Atoms bestimmt. So z. B. beruht der stark elektropositive Charakter der Alkalien auf ihrer Fähigkeit, unter verhältnismäßig geringer Energieaufnahme ein Elektron abspalten zu können, also darauf, daß der Grundzustand nicht sehr tief unter dem tiefsten ionisierten Zustand liegt. Umgekehrt erklärt sich die Reaktionsträgheit der Edelgase durch die besonders hohe Lage der angeregten und ionisierten Zustände über dem Normalzustand. Der Ausgangspunkt für diese, die moderne Physik geradezu beherrschenden Gedankengänge — deren Wiedergabe leider außerhalb des Rahmens dieses Buches liegt — war N. Bohrs Erklärung der Grundzüge des periodischen Systems der Elemente. Die wichtigsten Schritte bei der Aufdeckung des Aufbauprinzips waren u. a. die Landé-Sommerfeldsche Entdeckung

¹⁾ Der Beweis soll dem Leser überlassen werden.

des Vektormodells, die Russell-Saundersche Formulierung des normalen Kopplungsfalles, und das Paulische Verbot äquivalenter Bahnen. Die endgültige Fassung des Aufbauprinzips verdanken wir den Untersuchungen von F. Hund.

Bei der Abschätzung der Lage der Energieniveaus, der Eigenwerte der Schrödinger-Gleichung

$$\sum_k \left\{ -\frac{\hbar^2}{8\pi^2 m} \left[\frac{\partial^2}{\partial x_k^2} + \frac{\partial^2}{\partial y_k^2} + \frac{\partial^2}{\partial z_k^2} \right] - \frac{Z e^2}{r_k} \right\} \psi + \frac{1}{2} \sum_{i \neq k} \frac{e^2}{r_{ik}} \psi = E \psi, \quad (1)$$

geht man zunächst von einer vereinfachten Gleichung aus

$$\begin{aligned} \mathbf{H}_0 \psi &= (\mathbf{H}_1 + \mathbf{H}_2 + \cdots + \mathbf{H}_n) \psi = E \psi, \\ \mathbf{H}_k &= -\frac{\hbar^2}{8\pi^2 m} \left[\frac{\partial^2}{\partial x_k^2} + \frac{\partial^2}{\partial y_k^2} + \frac{\partial^2}{\partial z_k^2} \right] - \frac{Z e^2}{r_k} \end{aligned} \quad (1a)$$

(Z ist die Kernladungszahl) in der die Wechselwirkungsenergie der Elektronen

$$\mathbf{W} = \sum_{k=2}^n \sum_{i=1}^{k-1} \frac{e^2}{r_{ik}} \quad (1b)$$

zunächst vernachlässigt ist und sucht diese dann durch ein Störungsverfahren zu berücksichtigen (vgl. Kap. XVII).

Dieses Vorgehen ist nur dann berechtigt, wenn die vom Kern herrührende potentielle Energie der Wechselwirkungsenergie \mathbf{W} gegenüber groß ist. Am besten ist diese Voraussetzung bei hoher Kernladungszahl Z und kleiner Elektronenzahl, also bei stark ionisierten Atomen, erfüllt. Das Näherungsverfahren, das hier verwendet wird, unterscheidet sich also ganz wesentlich von den bisher benutzten Näherungsverfahren: es hat nichts mit der Kleinheit der Feinstrukturkonstante

$$\alpha = \frac{2\pi e^2}{hc} = 7,28 \cdot 10^{-3}$$

zu tun, die für die Berechtigung der bisher gemachten Vernachlässigungen maßgebend war. In die Eigenwerte von (1) gehen die Naturkonstanten e , h , m , schon aus Dimensionsgründen nur in der Kombination

$$4\pi^2 \frac{mc^4}{h^2} \quad (*)$$

ein [c kommt in (1) gar nicht vor], und so lässt sich sowohl aus den Eigenwerten von (1a) wie auch aus jedem weiteren Näherungswert

des Störungsverfahrens der Faktor (*) herausziehen. Es handelt sich also nicht um eine Potenzreihenentwicklung der Energie nach der Feinstrukturkonstante. Die Voraussetzung dafür, daß (1a) eine gute Näherung geben soll, ist sogar gewissermaßen entgegengesetzt der Voraussetzung, die für die Güte der Näherung maßgebend war, die bei der Feinstrukturaufspaltung verwendet wurde. Die Näherung ist hier um so besser, je weniger die Störung \mathbf{W} ausmacht, je näher im Endresultat jene Eigenwerte von (1) liegen, die aus demselben Eigenwert von (1a) durch die Störung entstanden sind. Bei der Feinstrukturaufspaltung ist die Näherung dagegen um so besser, je weiter im dortigen Ausgangsproblem (1) die Eigenwerte voneinander liegen.

Ein Umstand, der der Termberechnung mit Hilfe des Aufbauprinzips sehr zustatten kommt, ist der, daß man einem großen Teil der Elektronenwechselwirkung durch eine Modifizierung des Kernpotentials Rechnung tragen kann. Wenn z. B. in einem Li-Atom zwei sogenannte K -Elektronen ($N = 1$) und ein höher angeregtes Elektron vorhanden sind, wird das letztere fast im Laufe seiner ganzen Bewegung unter der Wirkung eines Kernpotentials $1/e/r$ stehen, weil das vom Kern eigentlich herrührende Potential $3/e/r$ durch die beiden K -Elektronen, die fast sicher in der Nähe des Kerns sind, zu $2/3$ „abgeschirmt“ wird. Man wird also an Stelle des Coulombschen Potentials $-Ze^2/r$ in (1a) — besser ein anderes, modifiziertes Potential setzen. Dabei muß natürlich auch (1b) entsprechend verändert werden, damit es zusammen mit (1a) wieder (1) ergibt. Die Theorie der Abschirmung wurde in erster Linie von Hartree in die Quantenmechanik eingebaut, sie gibt oft überraschend gute Werte für die Terme und andere atomare Eigenschaften¹⁾.

Die nun folgende Ableitung des Aufbauprinzips röhrt von Slater²⁾ her. Obwohl weder das ungestörte Problem (1a), noch die Störung (1b) die Spinkoordinaten irgendwie enthält, führt Slater — im Gegensatz zu allen anderen Entwicklungen des Aufbauprinzips — die Spinkoordinate von vornherein ein. Durch diese scheinbare Komplikation erhält man in Wirklichkeit eine sehr große

¹⁾ D. R. Hartree: Proc. Cambr. Phil. Soc. **24**, 89, 1928; vgl. auch J. C. Slater, Phys. Rev. **35**, 210, 1930; V. Fock, Zeitschr. f. Phys. **61**, 126, 1930.

²⁾ J. C. Slater: Phys. Rev. **34**, 1293, 1929.

Vereinfachung, weil man sich so von vornherein auf die antisymmetrischen Eigenfunktionen beschränken kann. Man erhält so als Eigenfunktionen nicht die spinfreien Eigenfunktionen $\psi_{x,u}^{SL}$, die in bezug auf Vertauschung der Cartesischen Koordinaten der Elektronen zu einer bestimmten Zeile einer Darstellung $\overline{A}^{(S)}$ der symmetrischen Gruppe gehören, sondern direkt die im XXII. Kapitel abgeleiteten $\Xi_{v,u}^{SL}$, die auch die Spinkoordinaten als Variable enthalten und die in bezug auf Vertauschung aller Koordinaten der Elektronen antisymmetrisch sind. Das Multiplettsystem der $\Xi_{v,u}^{SL}$ äußert sich bei den Drehungen Q_R der Spinkoordinaten: $\Xi_{v,u}^{SL}$ gehört in bezug auf die Q_R zur v -Zeile von $\mathfrak{D}^{(S)}$.

2. Die Eigenfunktionen (vgl. Kap. XVII, 2) der Operatoren H_k seien alle durchnumeriert

$$H_k \psi_b(x_k y_k z_k) = E_b \psi_b(x_k y_k z_k), \quad (2)$$

die „Bahn“ b bedeutet dann die Zusammenfassung der Hauptquantenzahl N , der Azimutalquantenzahl l und der magnetischen Quantenzahl μ . Es wird dabei angenommen, daß die zufällige Entartung, das Zusammenfallen der Eigenwerte mit gleichen N , aber verschiedenen l , die bei dem Wasserstoffatom im rein Coulombschen Zentralfeld auftritt, durch die Abschirmungskräfte aufgehoben wird. Dann liegen die Terme mit verschiedenen l auch bei gleichem N getrennt, und zwar lehrt die detailliertere Theorie, wie auch die Erfahrung, daß bei gleichem N die Terme mit kleineren l tiefer, mit größeren l höher liegen. Da sich die H_1, H_2, \dots, H_n nur dadurch unterscheiden, daß sie auf verschiedene Variable einwirken, sind die Eigenwerte aller H_k dieselben Zahlen, und auch ihre Eigenfunktionen sind dieselben Funktionen, nur jeweils anderer Variablen.

Durch Einführung der Spinkoordinate s_k entstehen je zwei Eigenfunktionen

$$\psi_{b\sigma}(x_k y_k z_k s_k) = \psi_b(x_k y_k z_k) \delta_{s_k \sigma} \quad (\sigma = \pm 1) \quad (3)$$

aus jeder Eigenfunktion $\psi_b(x_k y_k z_k)$. Die Eigenfunktionen von H_0 sind dann als Funktionen aller Koordinaten $x_1 y_1 z_1 s_1 \dots x_n y_n z_n s_n$ die Produkte

$$\psi_{b_1 \sigma_1 b_2 \sigma_2 \dots b_n \sigma_n}$$

$$= \psi_{b_1 \sigma_1}(x_1 y_1 z_1 s_1) \psi_{b_2 \sigma_2}(x_2 y_2 z_2 s_2) \dots \psi_{b_n \sigma_n}(x_n y_n z_n s_n) \quad (4)$$

der Eigenfunktionen der H_k . Je zwei Eigenfunktionen $\psi_{b_1 \sigma_1 \dots b_n \sigma_n}$ und $\psi_{b'_1 \sigma'_1 \dots b'_n \sigma'_n}$ sind orthogonal, wenn sie überhaupt verschieden sind: ist $b_i \neq b'_i$, so verschwindet ihr skalares Produkt schon bei der Integration über x_i, y_i, z_i , ist $\sigma_i \neq \sigma'_i$, so verschwindet es bei der Summation über s_i . Die Eigenfunktionen (4) gehören bei allen 2^n Wertesystemen der $\sigma_1, \sigma_2, \dots, \sigma_n$ zum Eigenwert

$$E = E_{b_1} + E_{b_2} + \dots + E_{b_n}. \quad (4a)$$

Außerdem gehören wegen der Invarianz des Hamiltonoperators gegenüber den Vertauschungen O_P der Elektronen noch alle jene Eigenfunktionen zum Eigenwert (4a), die aus den $\psi_{b_1 \sigma_1 \dots b_n \sigma_n}$ durch Anwendung der Operatoren O_P hervorgehen:

$$O_P \psi_{b_1 \sigma_1 \dots b_n \sigma_n}(1', 2', \dots, n') = \psi_{b_1 \sigma_1}(1) \dots \psi_{b_n \sigma_n}(n), \quad (5)$$

wo P die Permutation $\begin{pmatrix} 1 & 2 & \dots & n \\ 1' & 2' & \dots & n' \end{pmatrix}$ ist und k an Stelle der vier Variablen $x_k y_k z_k s_k$ steht. Durch Umordnung der Faktoren geht die rechte Seite von (5) in $\psi_{b_1, \sigma_1}(1') \dots \psi_{b_n, \sigma_n}(n')$ über. Wenn man dann noch für $1', 2', \dots, n'$ wieder $1, 2, \dots, n$ einsetzt, erhält man

$$O_P \psi_{b_1 \sigma_1 \dots b_n \sigma_n}(1, 2, \dots, n) = \psi_{b_1, \sigma_1}, \dots, b_n, \sigma_n(1, 2, \dots, n). \quad (5a)$$

Der Eigenwert von (5a)

$$E_{b_1}, + E_{b_2}, + \dots + E_{b_n}$$

ist ja identisch mit dem Eigenwert (4a).

Man kann jene Eigenfunktionen des Eigenwerts (4a), die durch Vertauschung der Elektronen auseinander hervorgehen,

$$\psi_{b_1 \sigma_1 \dots b_n \sigma_n}, O_{P_1} \psi_{b_1 \sigma_1 \dots b_n \sigma_n}, O_{P_2} \psi_{b_1 \sigma_1 \dots b_n \sigma_n}, \dots \quad (6)$$

zusammenfassen. Wir nennen ihre Gesamtheit eine Konfiguration. Eine Konfiguration ist dann durch n Symbole $(b_k \sigma_k)$

$$(b_1 \sigma_1)(b_2 \sigma_2) \dots (b_n \sigma_n) = (N_1 l_1 \mu_1 \sigma_1) (N_2 l_2 \mu_2 \sigma_2) \dots (N_n l_n \mu_n \sigma_n) (\dagger)$$

ohne Rücksicht auf die Reihenfolge derselben charakterisiert. Es ist ja klar, daß man, von der Eigenfunktion (5a) ausgehend, in der schon eine Vertauschung ausgeführt ist, genau dieselben Eigenfunktionen durch Vertauschung der Variablen erhalten kann, die

man aus $\psi_{b_1 \sigma_1 \dots b_n \sigma_n}$ auf diese Weise erhält. Man kann daher im Symbol (\dagger) für die Konfiguration irgendeine Reihenfolge der b vorschreiben, z. B. die, die durch

$$\left. \begin{array}{l} N_i \leqq N_{i+1}; \\ l_i \leqq l_{i+1} \text{ für } N_i = N_{i+1}; \\ \mu_i \leqq \mu_{i+1} \text{ für } N_i = N_{i+1}, \text{ und } l_i = l_{i+1} \end{array} \right\} \quad (7)$$

und

$$\sigma_i \leqq \sigma_{i+1} \text{ für } N_i = N_{i+1}, l_i = l_{i+1}, \mu_i = \mu_{i+1} \quad (7a)$$

gegeben ist. Jede Eigenfunktion gehört in eine und nur eine Konfiguration, und Eigenfunktionen verschiedener Konfigurationen sind zueinander orthogonal, weil alle voneinander verschiedenen Funktionen der Gestalt (4) orthogonal sind.

Wenn man auf die Funktionen einer Konfiguration einen Operator O_P anwendet, kann man die entstandenen Funktionen durch die unveränderten Funktionen der Konfiguration linear ausdrücken: sie sind ja noch immer Funktionen der Konfiguration. Zu den Funktionen (6) gehört daher eine Darstellung der Gruppe der O_P , der symmetrischen Gruppe n -ten Grades. Mit Hilfe der Matrix, die diese Darstellung ausreduziert, kann man aus den Funktionen (6) solche Linearkombinationen bilden, die zu irreduziblen Darstellungen der Gruppe der O_P gehören. Umgekehrt kann man mit Hilfe dieser „irreduziblen“ Funktionen die Funktionen (6) ausdrücken, sie erscheinen dann zerlegt in lauter Funktionen, die zu irreduziblen Darstellungen gehören. Wir können daher an Stelle der Funktionen (6) ihre irreduziblen Bestandteile, die Linearkombinationen von ihnen sind, benutzen. Wegen des Pauliprinzips brauchen wir von diesen irreduziblen Linearkombinationen nur die antisymmetrischen, d. h. die antisymmetrischen Bestandteile der $O_{P_x} \psi_{b_1 \sigma_1 \dots b_n \sigma_n}$, die sich nach (6), Kapitel XII, zu

$$\sum_P \varepsilon_P O_P O_{P_x} \psi_{b_1 \sigma_1 \dots b_n \sigma_n} = \sum_P \varepsilon_P O_{PP_x} \psi_{b_1 \sigma_1 \dots b_n \sigma_n} \quad (8)$$

berechnen, wo ε_P für gerade Permutationen gleich +1, für ungerade Permutationen gleich -1 ist.

Zunächst ist es nun allgemein wahr, daß die antisymmetrischen Bestandteile (8) aller Funktionen derselben Konfiguration bis auf

das Vorzeichen identisch sind: für $P_x = E$ ist ja (8) nichts anderes als die Determinante¹⁾

$$\sqrt{n!} \cdot \chi_{b_1 \sigma_1 \dots b_n \sigma_n} = \begin{vmatrix} \psi_{b_1 \sigma_1}(1) & \psi_{b_1 \sigma_1}(2) & \dots & \psi_{b_1 \sigma_1}(n) \\ \psi_{b_2 \sigma_2}(1) & \psi_{b_2 \sigma_2}(2) & \dots & \psi_{b_2 \sigma_2}(n) \\ \vdots & \vdots & & \vdots \\ \psi_{b_n \sigma_n}(1) & \psi_{b_n \sigma_n}(2) & \dots & \psi_{b_n \sigma_n}(n) \end{vmatrix} \quad (8a)$$

und $\mathbf{O}_{P_x} \psi_{b_1 \sigma_1 \dots b_n \sigma_n}$ unterscheidet sich von $\psi_{b_1 \sigma_1 \dots b_n \sigma_n}$ nur dadurch, daß die Variablen vertauscht sind, so daß sich auch die entsprechende antisymmetrische Linearkombination nur darin von (8a) unterscheidet, daß die Funktionen der Variablen $x_k y_k z_k s_k$ nicht in der k -ten, sondern in einer anderen Spalte stehen. Das kann aber durch Umordnung der Spalten, was höchstens eine Vorzeichenänderung bedeutet, ausgeglichen werden.

Es folgt nun umgekehrt, daß man aus den Funktionen (6) nur die eine antisymmetrische Linearkombination (8a) bilden kann: ist F antisymmetrisch, so folgt durch Gleichsetzen der antisymmetrischen Bestandteile der beiden Gleichungsseiten von

$$F = \sum_P c_P \mathbf{O}_P \psi_{b_1 \sigma_1 \dots b_n \sigma_n},$$

daß es bis auf eine Konstante gleich $\chi_{b_1 \sigma_1 \dots b_n \sigma_n}$ ist, weil der antisymmetrische Bestandteil aller Glieder auf der rechten Seite $\chi_{b_1 \sigma_1 \dots b_n \sigma_n}$ ist.

Man kann also aus den Funktionen einer Konfiguration höchstens eine antisymmetrische Linearkombination bilden, und man kann nicht einmal eine bilden, wenn in dem zugrunde gelegten System der σ für irgendein Paar i, k , für das $b_i = b_k$ (d. h. $N_i = N_k$; $l_i = l_k$; $\mu_i = \mu_k$) gilt, auch $\sigma_i = \sigma_k$ ist. Dann sind nämlich zwei Zeilen der Determinante (8a) gleich und sie verschwindet.

Es geben also jene Konfigurationen, in denen für kein Paar i, k gleichzeitig $b_i = b_k$; $\sigma_i = \sigma_k$ ist, genau einen im Sinne des Pauliprinzips erlaubten Zustand, die anderen überhaupt keinen. Dies ist die ursprüngliche Formulierung des Paulischen Äquivalenzprinzips, das vor der Quantenmechanik nur für den hier als erste Näherung zugrunde

¹⁾ Der Faktor $\sqrt{n!}$ wird hinzugefügt, damit $\chi_{b_1 \sigma_1 \dots b_n \sigma_n}$ normiert bleibe.

gelegten Grenzfall formuliert werden konnte, daß die Wechselwirkung der Elektronen vernachlässigt und jedem Elektron eine „Bahn“ zugeschrieben werden kann. In der Quantenmechanik ist das Pauliprinzip in seiner ursprünglichen Gestalt ein Spezialfall der allgemeinen, für alle Systeme bei beliebiger Wechselwirkung gültigen Forderung der Antisymmetrie der Wellenfunktionen¹⁾.

Insbesondere folgt hieraus, daß jene Kombinationen $b_1 b_2 \dots b_n$, in denen drei Bahnen $b_i = b_j = b_k$ identisch sind, überhaupt zu keiner „erlaubten“ Konfiguration Anlaß geben. In einer erlaubten Konfiguration müßte dann nämlich $\sigma_i \neq \sigma_j; \sigma_i \neq \sigma_k$ und $\sigma_j \neq \sigma_k$ sein, was — da die σ nur zwei Werte —1 und +1 haben können — nicht möglich ist. Eine Bahn darf höchstens zweimal besetzt sein.

3. Jetzt wären wir in der Lage, die Anzahl der erlaubten Zustände zu bestimmen, deren Energie bei verschwindender Wechselwirkung der Elektronen (4 a) ist. Wenn die Eigenwerte E_k der H_k als Funktionen der $x_k y_k z_k$ allein betrachtet alle einfach wären, so müßte man nur die erlaubten unter den 2^n Konfigurationen abzählen, die durch die verschiedenen Wertesysteme der σ gegeben sind. Gehören zu einigen Eigenwerten E_k mehrere Eigenfunktionen, so muß man mit jeder möglichen Kombination der b , deren Energiesumme $E_{b_1} + E_{b_2} + \dots + E_{b_n} = E$ ist, diese Abzählung ausführen²⁾.

Nun interessieren wir uns aber nicht nur für die Zahl der Zustände, die aus den Zuständen mit der Energie (4 a) bei der Wechselwirkung der Elektronen entstehen, sondern auch für ihre Art, d. h. für ihr Multiplettsystem und, wenn — wie bei den Atomen, die uns in erster Linie interessieren — die Drehsymmetrie nicht gestört ist, auch für die azimutalen Quantenzahlen. Ihre Bestimmung geschieht wieder durch die Betrachtung der Symmetrieverhältnisse.

¹⁾ Diese Bemerkung, die geschichtlich der Ausgangspunkt der gruppentheoretischen Behandlung der Spektraltheorie war, verdankt man W. Heisenberg und P. A. M. Dirac.

²⁾ Gewöhnlich wird angenommen, daß nur solche Kombinationen der b gleiche Energiesumme

$$E_{b_1} + E_{b_2} + \dots + E_{b_n} = E_{b'_1} + E_{b'_2} + \dots + E_{b'_n} \quad (4)$$

geben, in denen schon die einzelnen Energien [auf der linken und rechten Seite von (4)] paarweise gleich sind.

Es wird im folgenden — außer der Symmetrie in bezug auf Vertauschungen der Elektronen — nur noch die Symmetrie in bezug auf Drehungen \mathbf{Q}_R und \mathbf{P}_R der Spinkoordinaten und der Cartesischen Koordinaten ausgenutzt. Die Drehung der Cartesischen Koordinaten ist nur im kugelsymmetrischen Fall möglich, die der Spinkoordinaten dagegen immer, weil weder das Ausgangsproblem (1a) noch die Störung (1b) die Spinkoordinaten enthalten und das ganze Problem demnach ganz unabhängig von der Kugelsymmetrie sogar in bezug auf jeden Operator invariant ist, der nur auf die Spinkoordinaten einwirkt. Doch kann man sich auf die Drehungen \mathbf{Q}_R beschränken, weil schon diese genügen, um das Multiplettsystem der einzelnen gestörten Terme zu bestimmen. Die ganze Symmetrie, die betrachtet werden soll, ist das direkte Produkt der \mathbf{O}_P mit den \mathbf{Q}_R und — im isotropen Fall — noch mit den \mathbf{P}_R .

4. Betrachten wir zuerst den anisotropen Fall. Wendet man auf die antisymmetrischen Linearkombinationen $\chi_{b_1 \sigma_1 b_2 \sigma_2 \dots b_n \sigma_n}$ eines ungestörten Terms E den Operator \mathbf{Q}_R an, so entsteht wieder eine antisymmetrische Eigenfunktion zu demselben Eigenwert, die man daher durch die unveränderten antisymmetrischen Linearkombinationen ausdrücken kann. Die Koeffizienten bilden eine Darstellung der Drehgruppe und die Zahl A_S , die angibt, wie oft in dieser Darstellung die irreduzible $\mathfrak{D}^{(S)}$ enthalten ist, gibt gleichzeitig die Zahl der antisymmetrischen, in bezug auf die \mathbf{Q}_R zu $\mathfrak{D}^{(S)}$ gehörenden Terme an, die aus $E = E_{b_1} + E_{b_2} + \dots + E_{b_n}$ durch die Störung entstehen. Die irreduziblen Bestandteile einer Darstellung der Drehgruppe kann man aber schon bestimmen, wenn man die Matrizen kennt, die Drehungen um Z entsprechen: kommt in diesen Matrizen die Darstellung $(e^{i m \varphi})$ der zweidimensionalen Drehgruppe a_m -mal vor, so ist in der ganzen Darstellung die irreduzible Darstellung $\mathfrak{D}^{(S)}$ genau $A_S = a_s - a_{s+1}$ -mal enthalten (vgl. Kap. XV, 1).

Ist aber R eine reine Drehung um Z mit φ , so bedeutet \mathbf{Q}_R für jede Funktion der Konfiguration (6) — wie wir schon öfters nachgerechnet haben — lediglich eine Multiplikation mit $e^{\frac{1}{2} i (\sigma_1 + \dots + \sigma_n) \varphi}$, und dies gilt dann auch für die antisymmetrische Linearkombination $\chi_{b_1 \sigma_1 \dots b_n \sigma_n}$, so daß diese zur Darstellung der Drehungen um Z mit $m = \frac{1}{2}(\sigma_1 + \dots + \sigma_n)$ gehört. Hat der Eigenwert E im ganzen a_m erlaubte Konfigurationen mit der Summe $\sigma_1 + \sigma_2 + \dots + \sigma_n = 2m$, so entstehen aus ihm bei der Störung

$a_s - a_{s+1}$ in bezug auf Drehungen der Spinkoordinaten zu $\mathfrak{D}^{(S)}$ gehörende Terme.

Da sich diese Betrachtung zunächst auf den anisotropen Fall bezieht sei hier ein Beispiel betrachtet, bei dem die Eigenwerte E_k der H_k alle einfach sind. Hat man etwa vier Elektronen in einer doppelt besetzten und zwei einfach besetzten Bahnen

$$b_1 = b_2 \neq b_3 \neq b_4; \quad b_1 \neq b_4,$$

so geben folgende Kombinationen der σ je eine antisymmetrische Eigenfunktion¹⁾.

Tabelle 1

σ_1	σ_2	σ_3	σ_4	$\frac{1}{2}(\sigma_1 + \sigma_2 + \sigma_3 + \sigma_4)$
-1	+1	-1	-1	-1
-1	+1	-1	+1	0
-1	+1	+1	-1	0
-1	+1	+1	+1	+1

Es ist also

$$a_0 = 2; \quad a_1 = 1; \quad a_2 = a_3 = \dots = 0,$$

es entsteht $a_1 - a_2 = 1$ Term mit der Darstellung $\mathfrak{D}^{(1)}$ und $a_0 - a_1 = 1$ Term mit der Darstellung $\mathfrak{D}^{(0)}$.

Einen Term, zu dem $2S + 1$ in bezug auf die Q_R zu $\mathfrak{D}^{(S)}$ gehörende antisymmetrische Eigenfunktionen gehören, hat das Multiplettsystem S . Wenn man die Spinwechselwirkung einführt, spaltet er nämlich im allgemeinen (im anisotropen Falle!) in $2S + 1$ Feinstrukturkomponenten auf. Man kann aber auch leicht verifizieren, daß diese Definition des Multiplettsystems mit der früher benutzten übereinstimmt. Nach Kap. XXII gehört jede Funktion der s , die in bezug auf die Q_R zu $\mathfrak{D}^{(S)}$ gehört, in bezug auf die Q_P zu $A^{(S)*}$, und jede antisymmetrische Funktion F , die in bezug auf die Q_P zu $A^{(S)*}$ gehört,

$$\sum_P \sum_x A^{(S)}(P)_{xx} Q_P F = \frac{n!}{g_S} F, \quad (9)$$

¹⁾ Da $b_1 = b_2$ ist, kann nach (7a) bei der Abzählung der Konfigurationen $\sigma_1 \leqq \sigma_2$ angenommen werden, $\sigma_1 = \sigma_2$ ist aber nach dem Pauliprinzip verboten, so daß nur $\sigma_1 < \sigma_2$ übrigbleibt.

gehört in bezug auf die \mathbf{P}_P zu $\overline{\mathbf{A}}^{(S)}$, und dies war die Definition des Multiplettsystems. Das letztere folgt aus (9) mit Hilfe der Antisymmetrie $\mathbf{O}_P F = \varepsilon_P F$ und $\mathbf{Q}_P = \mathbf{P}_{P-1} \mathbf{O}_P$

$$\begin{aligned} \sum_P \sum_x \mathbf{A}^{(S)}(P)_{xx}^* \mathbf{P}_{P-1} \mathbf{O}_P F &= \sum_P \sum_x \mathbf{A}^{(S)}(P^{-1})_{xx} \mathbf{P}_{P-1} \varepsilon_P F \\ &= \sum_P \sum_x \overline{\mathbf{A}}^{(S)}(P^{-1}) \mathbf{P}_{P-1} F, \end{aligned} \quad (9a)$$

weil $\varepsilon_P = \varepsilon_{P-1}$ und $\mathbf{A}^{(S)}$ unitär ist.

Die richtigen Linearkombinationen der $\chi_{b_1 \sigma_1 \dots b_n \sigma_n}$ gehen bei Einführung der Störung \mathbf{W} direkt in die Funktionen Ξ_m^S des Kap. XXII über, die wir dort aus den fertigen spinfreien Eigenfunktionen der Schrödinger-Gleichung (1) gebildet haben.

Die Bedeutung der Slaterschen Methode liegt darin, daß man mit ihrer Hilfe die Betrachtung der Symmetrie der Schrödinger-Gleichung (1) gegenüber Vertauschungen der Cartesischen Koordinaten allein gänzlich vermeiden bzw. sie durch Betrachtung der Invarianz gegenüber Drehungen \mathbf{Q}_R der Spinkoordinaten ersetzen kann. So könnte man z. B. das Interkombinationsverbot darauf zurückführen, daß Eigenfunktionen verschiedener Multiplettsysteme in bezug auf die \mathbf{Q}_R zu verschiedenen Darstellungen gehören, und das Multiplizieren mit $x_1 + x_2 + \dots + x_n$ in bezug auf die \mathbf{Q}_R symmetrisch ist, weil es auf die Spinkoordinaten gar nicht einwirkt. (Wir haben das Interkombinationsverbot daraus gefolgert, daß die Eigenfunktionen verschiedener Multiplettsysteme in bezug auf Vertauschungen \mathbf{P}_P der Cartesischen Koordinaten zu verschiedenen Darstellungen gehören.) Wir haben die Slatersche Methode nicht in der angedeuteten extremen Form verwendet, weil wir uns mit allen Symmetrieeigenschaften beschäftigen und sie nach Möglichkeit alle ausnutzen wollten.

Obwohl es natürlich Fälle geben muß, in denen man bei Mitberücksichtigung der Vertauschbarkeit \mathbf{P}_P der Cartesischen Koordinaten weitergehende Schlüsse als mit Hilfe der \mathbf{Q}_R allein erhalten kann, ist es doch sehr interessant zu sehen, wie weitgehend man die \mathbf{P}_P durch die \mathbf{Q}_R ersetzen kann.

Eine wichtige Voraussetzung für die Verwendbarkeit der Slaterschen Methode ist die, daß die betrachteten Teilchen nur zwei Einstellungsmöglichkeiten haben. Hätte man es mit Teilchen zu tun, die mehr als zwei, etwa drei Einstellungsmöglichkeiten haben (Drehimpuls gleich $\hbar/2\pi$ an Stelle von $\frac{1}{2}\hbar/2\pi$, wie das etwa bei Stickstoffkernen der Fall ist), so würde zwar in der Definitionsgleichung der Operatoren \mathbf{Q}_R [(6b), Kap. XXI] $\mathfrak{D}^{(1)}$ an Stelle von $\mathfrak{D}^{(\frac{1}{2})}$ stehen, die ganze Betrachtung ließe sich aber noch genau so ausführen, wie es hier geschah, mit dem einzigen Unterschied, daß die Äquivalenz der \mathbf{Q}_R mit den \mathbf{P}_P nicht mehr gezeigt werden könnte, weil sie auch nicht vorhanden wäre. In der Tat fallen in diesem Falle mehrere Terme

der Schrödinger-Gleichung (1) zusammen, deren Zusammenfallen man durch Betrachtung der \mathbf{Q}_R allein nicht voraussehen kann. Man müßte dann entweder weitere Operatoren einführen, wodurch die Theorie dann ihre Einfachheit verlieren würde, oder aber man müßte die Symmetrie den Operatoren \mathbf{P}_P gegenüber zugrunde legen, wie das in den ersten Ableitungen des Aufbau-prinzips¹⁾ geschah. Auch aus diesem Grunde wurde in diesem Buch auf die Einführung der Vertauschung der Cartesischen Koordinaten der Teilchen und eine Betrachtung der Darstellungen der symmetrischen Gruppe nicht verzichtet und die Äquivalenz der \mathbf{Q}_R mit den \mathbf{P}_P bei Elektronen explizite bewiesen.

5. Der Unterschied des kugelsymmetrischen Falles vom asymmetrischen besteht darin, daß die Eigenwerte $E_{l_k}^{N_k}$ der \mathbf{H}_k , auch als Funktionen der $x_k y_k z_k$ allein betrachtet, nicht mehr einfach, sondern $2l_k + 1$ -fach sind²⁾, und daß man auch bei den gestörten Termen nicht nur das Multiplettsystem, sondern auch die Azimutalquantenzahl bestimmen muß. Die ungestörten Terme sind demnach durch die Angabe der Haupt- und Azimutalquantenzahlen

$$N_1 l_1, N_2 l_2, \dots, N_n l_n \quad (\dagger\dagger)$$

der Bahnen gegeben. Um alle Konfigurationen des ungestörten Terms ($\dagger\dagger$) zu erhalten, muß man in den Symbolen der Konfigurationen

$(N_1 l_1 \mu_1 \sigma_1) (N_2 l_2 \mu_2 \sigma_2) \cdots (N_n l_n \mu_n \sigma_n) = (b_1 \sigma_1) (b_2 \sigma_2) \cdots (b_n \sigma_n)$ (\dagger) die μ und die σ alle möglichen Werte ($|\mu_k| \leqq l_k; \sigma_k = \pm 1$) durchlaufen lassen. Dabei kann indessen für die N_k, l_k, μ_k (7), und wenn man nur die erlaubten Konfigurationen mitrechnen will, für die σ sogar (7a) mit dem oberen Zeichen

$\sigma_i < \sigma_{i+1}$ für $N_i = N_{i+1}, l_i = l_{i+1}, \mu_i = \mu_{i+1}$ (7b) angenommen werden. Jede erlaubte Konfiguration gibt eine antisymmetrische Eigenfunktion (8a).

Wenn man auf alle diese antisymmetrischen Linearkombinationen der Eigenfunktionen von ($\dagger\dagger$) die Operatoren $\mathbf{Q}_R \mathbf{P}_R'$

¹⁾ Vgl. E. Wigner, Zeitschr. f. Phys. **48**, 624, 1927; §§ 21 bis 25. M. Delbrück, ebenda **51**, 181, 1928.

²⁾ Wenn \mathbf{H}_k wirklich die in (1b) angegebene Form hätte, würden sogar alle Eigenwerte $E_{l_k}^{N_k}$ ($l_k = 0, 1, 2, \dots, N_k - 1$) mit gleichem N_k zusammenfallen. Doch wird angenommen, daß das Coulombsche Feld durch die Abschirmung soweit modifiziert wird, daß die Eigenwerte $E_{l_k}^{N_k}$ alle getrennt sind.

anwendet und die entstehenden Funktionen durch die unveränderten ausdrückt

$$\begin{aligned} & \mathbf{Q}_R \mathbf{P}_{R'} \chi_{N_1 l_1 \mu_1 \sigma_1 \dots N_n l_n \mu_n \sigma_n} \\ &= \sum_{\mu' \sigma'} \Delta(R, R')_{\mu'_1 \sigma' \dots \mu'_n \sigma'_n; \mu_1 \sigma_1 \dots \mu_n \sigma_n} \chi_{N_1 l_1 \mu'_1 \sigma'_1 \dots N_n l_n \mu'_n \sigma'_n}, \quad (10) \end{aligned}$$

so bilden die Matritzen $\Delta(R, R')$ eine Darstellung des direkten Produktes der Gruppen der \mathbf{Q}_R und der $\mathbf{P}_{R'}$. Die Zahl A_{SL} , die angibt, wie oft in $\Delta(R, R')$ die irreduzible Darstellung $\mathfrak{D}^{(S)}(R) \times \mathfrak{D}^{(L)}(R')$ enthalten ist, ist gleichzeitig die Zahl der aus dem Term ($\dagger\dagger$) durch die Störung \mathbf{W} entstehenden Terme mit dem Multiplettsystem S und der Azimutalquantenzahl L . Die entsprechenden Linearkombinationen der χ bilden dann wiederum die erste Näherung für die Berechnung der im Kap. XXII mit $E_{v\mu}^{SL}$ bezeichneten Funktionen.

Zur Bestimmung der Zahlen A_{SL} genügt es wiederum — wie sich zeigen wird —, jene Matritzen $\Delta(R, R')$ zu bestimmen, bei denen R und R' Drehungen um Z sind. Ist R eine Drehung um Z mit α , so wissen wir, daß

$$\begin{aligned} & \mathbf{Q}_{\{\alpha' 00\}} \chi_{N_1 l_1 \mu_1 \sigma_1 \dots N_n l_n \mu_n \sigma_n} \\ &= e^{\frac{1}{2} i(\sigma_1 + \dots + \sigma_n) \alpha} \chi_{N_1 l_1 \mu_1 \sigma_1 \dots N_n l_n \mu_n \sigma_n} \quad (11) \end{aligned}$$

gilt. Bei der Berechnung von $\mathbf{P}_{\{\alpha' 00\}} \psi_{\mu_1 \sigma_1}^{N_1 l_1}(1) \dots \psi_{\mu_n \sigma_n}^{N_n l_n}(n)$ kann man $\mathbf{P}_{\{\alpha' 00\}}$ auf die einzelnen Faktoren getrennt anwenden, der Faktor $\psi_{\mu_k \sigma_k}^{N_k l_k}$ mit der magnetischen Quantenzahl μ_k reproduziert sich dabei bis auf einen Faktor $e^{i \mu_k \alpha'}$. Es gilt daher

$$\begin{aligned} & \mathbf{P}_{\{\alpha' 00\}} \chi_{N_1 l_1 \mu_1 \sigma_1 \dots N_n l_n \mu_n \sigma_n} \\ &= e^{i(\mu_1 + \dots + \mu_n) \alpha'} \chi_{N_1 l_1 \mu_1 \sigma_1 \dots N_n l_n \mu_n \sigma_n}, \quad (11a) \end{aligned}$$

weil dies für alle Eigenfunktionen der Konfiguration

$$(N_1 l_1 \mu_1 \sigma_1) (N_2 l_2 \mu_2 \sigma_2) \dots (N_n l_n \mu_n \sigma_n),$$

also auch für alle ihre Linearkombinationen gilt. Die Gleichung (11a) bringt zum Ausdruck, daß sich die Z -Komponenten der Drehimpulse der einzelnen Elektronen einfach addieren. Ans (11) und (11a) folgt

$$\begin{aligned} & \mathbf{Q}_{\{\alpha 00\}} \mathbf{P}_{\{\alpha' 00\}} \chi_{N_1 l_1 \mu_1 \sigma_1 \dots N_n l_n \mu_n \sigma_n} \\ &= e^{\frac{1}{2} i(\sigma_1 + \sigma_2 + \dots + \sigma_n) \alpha} e^{i(\mu_1 + \mu_2 + \dots + \mu_n) \alpha'} \chi_{N_1 l_1 \mu_1 \sigma_1 \dots N_n l_n \mu_n \sigma_n}. \quad (12) \end{aligned}$$

Die Matrizen $\Delta(R, R')$, die Drehungen der Spinkoordinaten um Z mit α und der Cartesischen Koordinaten mit α' entsprechen, sind Diagonalmatrizen, und das Diagonalelement, das der Eigenfunktion $\chi_{N_1 l_1 \mu_1 \sigma_1 \dots N_n l_n \mu_n \sigma_n}$ entspricht, ist $e^{i(\nu\alpha + \mu\alpha')}$, wo

$$\nu = \frac{1}{2}(\sigma_1 + \dots + \sigma_n) \text{ und } \mu = \mu_1 + \dots + \mu_n$$

ist. Die Diagonalsumme dieser Matrizen erhält man daher durch Addition der $e^{i(\nu\alpha + \mu\alpha')}$ aller erlaubten Konfigurationen. Man kann den so erhaltenen Charakter von $\Delta(R, R')$, wo R und R' Drehungen um Z sind, in einer Tabelle (s. Tabelle 3), ähnlich der auf S. 200, zusammenfassen, indem man für jedes ν eine Zeile und jedes μ eine Spalte bestimmt und in das Kreuzungsquadrat so viele Kreuzchen einträgt, wie erlaubte Konfigurationen mit $\frac{1}{2}(\sigma_1 + \sigma_2 + \dots + \sigma_n) = \nu$ und $\mu_1 + \mu_2 + \dots + \mu_n = \mu$ vorhanden sind.

Nehmen wir als Beispiel zwei Elektronen, deren ungestörte Energie E_1^2 gleich sei und einem p -Zustand ($l = 1$) entspreche. Die erlaubten Konfigurationen sind dann ¹⁾:

Tabelle 2

	μ_1	σ_1	μ_2	σ_2	μ	ν		μ_1	σ_1	μ_2	σ_2	μ	ν
1	(-1 -1)	(-1 1)	-2	0	9	(-1 1)	(1 1)	0	1				
2	(-1 -1)	(0 -1)	-1	-1	10	(0 -1)	(0 1)	0	0				
3	(-1 -1)	(0 1)	-1	0	11	(0 -1)	(1 -1)	1	-1				
4	(-1 -1)	(1 -1)	0	-1	12	(0 -1)	(1 1)	1	0				
5	(-1 -1)	(1 1)	0	0	13	(0 1)	(1 -1)	1	0				
6	(-1 1)	(0 -1)	-1	0	14	(0 1)	(1 1)	1	1				
7	(-1 1)	(0 1)	-1	1	15	(1 -1)	(1 1)	2	0				
8	(-1 1)	(1 -1)	0	0									

Die letzten beiden Spalten enthalten $\mu_1 + \mu_2 = \mu$ bzw. $\frac{1}{2}(\sigma_1 + \sigma_2) = \nu$. In der nun folgenden Tabelle ist für jede Zeile der Tabelle 2 ein Kreuzchen in das entsprechende Quadrat eingetragen.

1) In der Tabelle 2 sind in den Zeichen der Konfigurationen die Haupt- und Azimutalquantenzahlen weggelassen, weil sie für beide Elektronen 2 bzw. 1 sind. Es steht daher $(\mu_k \sigma_k)$ für $(N_k l_k \mu_k \sigma_k) = (2 1 \mu_k \sigma_k)$.

Tabelle 3

$\nu \setminus \mu$	-2	-1	0	1	2
-1		+	+	+	
0	+	++	+++	++	
1		+	+	+	+

In dieser Tabelle sind die Charaktere der Gruppenelemente $Q_{\{\alpha 00\}} P_{\{\alpha' 00\}}$ in der Darstellung $\Delta(R, R')$ zusammengefaßt. Der Charakter dieses Elements in der irreduziblen Darstellung $D^{(S)} \times D^{(L)}$ ist andererseits

$$\sum_{\nu\mu} [D^{(S)}(\{\alpha 00\}) \times D^{(L)}(\{\alpha' 00\})]_{\nu\mu} = \sum_{\mu=-L}^L \sum_{\nu=-S}^S e^{i(\nu\alpha + \mu\alpha')}, \quad (13)$$

und dem würde in der Tab. 3 ein mit Kreuzchen einfach angefülltes Rechteck entsprechen, das sich von $\nu = -S$ bis $\nu = S$ und von $\mu = -L$ bis $\mu = L$ erstreckt. Wenn man den in der Tabelle 3 dargestellten Charakter als Summe lauter irreduzibler Charaktere, d. h. rechteckiger Kreuzchenfelder schreibt, wird die Anzahl $a_{\nu\mu}$ der Kreuzchen im $\nu \mu$ -Quadrat gleich der Summe der Anzahlen A_{SL} der in $\Delta(R, R')$ enthaltenen Darstellungen $D^{(S)} \times D^{(L)}$ mit $S \geqq |\nu|$; $L \geqq |\mu|$ sein:

$$a_{\nu\mu} = A_{SL} + A_{S+1L} + A_{S+2L} + \dots$$

$$+ A_{SL+1} + A_{S+1L+1} + \dots + A_{SL+2} + A_{S+1L+2} + \dots \quad (14)$$

Dabei ist A_{SL} gleichzeitig die Anzahl der bei der Ausführung des Störungsverfahrens aus dem Term $(\dagger\dagger)$ entstehenden Terme mit dem Multiplettsystem S und der Azimutalquantenzahl L . Nach (14) ist

$$A_{SL} = a_{SL} - a_{S+1L} - a_{SL+1} + a_{S+1L+1}. \quad (14a)$$

Es zeigt sich also tatsächlich, daß die irreduziblen Bestandteile der Darstellung $\Delta(R, R')$ schon durch die Charaktere derjenigen Elemente bestimmt sind, bei denen R und R' Drehungen um Z sind; schon diese Charaktere kann man nur in einer Weise in irreduzible Charaktere (13) zerlegen¹⁾.

¹⁾ Dies beruht darauf, daß jede Klasse der Gruppe des direkten Produktes der Q_R und $P_{R'}$ ein Element enthält, bei dem R und R' beide Drehungen um Z sind. Da aber alle Elemente derselben Klasse in jeder Darstellung denselben Charakter haben, bestimmen eigentlich die Charaktere dieser Elemente den ganzen Charakter.

Für den Term der Tabelle 3 ergibt (14a) $A_{11} = 1$; $A_{02} = 1$; $A_{00} = 1$; alle anderen A_{SL} sind Null. Zwei äquivalente p -Elektronen¹⁾ geben daher einen 3P -, einen 1D - und 1S -Term. An Stelle der Kreuzchen der Tabelle 3 ist in der nun folgenden Tabelle das Symbol des Terms eingetragen, für den das entsprechende Kreuzchen verwendet wurde (d. h. zu dessen Darstellungscharakter das durch das Kreuzchen repräsentierte Glied $e^{i(\nu\alpha + \mu\sigma)}$ gebraucht wurde).

Tabelle 4

$\nu \backslash \mu$	-2	-1	0	1	2
-1		3P	3P	3P	
0	1D	1D 3P	1S 1D 3P	1D 3P	1D
1		3P	3P	3P	

6. Außer dem Multiplettsystem und der Azimutalquantenzahl muß noch der Spiegelungscharakter der entstehenden Terme bestimmt werden. Für die Eigenfunktionen des Einelektronenproblems ist nach (17), Kap. XIX, der Spiegelungscharakter durch die Azimutalquantenzahl l gegeben, es ist $w = (-1)^l$. Daher gilt unabhängig von den μ_k und σ_k , wenn $\mathbf{O}_I = \mathbf{P}_I$ die Inversion bedeutet:

$$\begin{aligned} \mathbf{O}_I \psi_{\mu_1 \sigma_1}^{N_1 l_1}(1) \dots \psi_{\mu_n \sigma_n}^{N_n l_n}(n) &= \mathbf{P}_I \psi_{\mu_1 \sigma_1}^{N_1 l_1}(1) \dots \mathbf{P}_I \psi_{\mu_n \sigma_n}^{N_n l_n}(n) \\ &= (-1)^{l_1 + l_2 + \dots + l_n} \psi_{\mu_1 \sigma_1}^{N_1 l_1}(1) \psi_{\mu_2 \sigma_2}^{N_2 l_2}(2) \dots \psi_{\mu_n \sigma_n}^{N_n l_n}(n). \end{aligned} \quad (15)$$

Der Spiegelungscharakter aller Eigenfunktionen des Eigenwerts ($\dagger\dagger$) ist gleich $(-1)^{l_1 + l_2 + \dots + l_n}$, und dies wird auch der Spiegelungscharakter aller daraus entstehenden gestörten Terme sein. Der Spiegelungscharakter der aus ($\dagger\dagger$) entstehenden Terme ist positiv, wenn die Summe der Azimutalquantenzahlen der einzelnen Bahnen $l_1 + l_2 + \dots + l_n$ gerade ist, negativ, wenn sie ungerade ist. Es folgt hieraus u. a., daß optische Übergänge zwischen Termen, die aus demselben ungestörten Term ($\dagger\dagger$) entstehen, durch die Laporte-Regel verboten sind.

7. Zum Schluß sei die Berechnung der ersten Näherung für die Energiestörung skizziert.

Bezeichnen wir die Funktionen $\chi_{N_1 l_1 \mu_1 \sigma_1 \dots N_n l_n \mu_n \sigma_n}$, für die $\mu_1 + \mu_2 + \dots + \mu_n = \mu$ und $\sigma_1 + \sigma_2 + \dots + \sigma_n = 2\nu$ ist, deren

¹⁾ Das heißt Bahnen mit gleicher Energie und mit der Azimutalquantenzahl $l = 1$.

Kreuzchen also im $\nu\mu$ -Quadrat der Tabelle 3 stehen, mit $\chi_{\nu\mu 1}$, $\chi_{\nu\mu 2}$, $\chi_{\nu\mu 3}$, ... Die richtigen Linearkombinationen $f_{\nu\mu 1}$, $f_{\nu\mu 2}$, $f_{\nu\mu 3}$, ..., die zur $\nu\mu$ -Zeile irgendeiner Darstellung $\mathfrak{D}^{(S)} \times \mathfrak{D}^{(L)}$ gehören, können alle als Linearkombinationen der $\chi_{\nu\mu 1}$, $\chi_{\nu\mu 2}$, ... allein geschrieben werden:

$$f_{\nu\mu x} = \sum_{\lambda} u_{x\lambda} \chi_{\nu\mu\lambda}. \quad (16)$$

Dabei ist die Transformationsmatrix u unitär, weil sowohl die $\chi_{\nu\mu\lambda}$ wie auch die $f_{\nu\mu x}$ aufeinander paarweise orthogonal sind:

$$\left. \begin{aligned} \delta_{xx'} &= (f_{\nu\mu x}, f_{\nu\mu x'}) = \sum_{\lambda\lambda'} (u_{x\lambda} \chi_{\nu\mu\lambda}, u_{x'\lambda'} \chi_{\nu\mu\lambda'}) \\ &= \sum_{\lambda\lambda'} u_{x\lambda}^* u_{x'\lambda'} \delta_{\lambda\lambda'} = \sum_{\lambda} u_{x\lambda}^* u_{x'\lambda}. \end{aligned} \right\} \quad (17)$$

Die Störungsenergie erster Näherung des Eigenwerts von $f_{\nu\mu x}$ ist $(f_{\nu\mu x}, \mathbf{W} f_{\nu\mu x})$. Die Summe der Störungsenergien für alle x , d. h. für alle Terme mit $S \geqq |\nu|$; $L \geqq |\mu|$, lässt sich auch ohne Kenntnis von u berechnen

$$\left. \begin{aligned} \sum_x (f_{\nu\mu x}, \mathbf{W} f_{\nu\mu x}) &= \sum_x \sum_{\lambda\lambda'} u_{x\lambda}^* u_{x\lambda'} (\chi_{\nu\mu\lambda}, \mathbf{W} \chi_{\nu\mu\lambda'}) \\ &= \sum_{\lambda\lambda'} \delta_{\lambda\lambda'} (\chi_{\nu\mu\lambda}, \mathbf{W} \chi_{\nu\mu\lambda'}) = \sum_{\lambda} (\chi_{\nu\mu\lambda}, \mathbf{W} \chi_{\nu\mu\lambda}). \end{aligned} \right\} \quad (18)$$

Hieraus folgt für die Summe der Störungsenergien der aus (††) entstehenden Terme mit dem Multiplettsystem S und der Azimutalquantenzahl L [ähnlich wie in (14a)]

$$\begin{aligned} \sum_x \Delta E_{SLx} &= \sum_{\lambda} [(\chi_{SL\lambda}, \mathbf{W} \chi_{SL\lambda}) - (\chi_{S+1L\lambda}, \mathbf{W} \chi_{S+1L\lambda}) \\ &\quad - (\chi_{SL+1\lambda}, \mathbf{W} \chi_{SL+1\lambda}) + (\chi_{S+1L+1\lambda}, \mathbf{W} \chi_{S+1L+1\lambda})]. \end{aligned} \quad (18a)$$

Entsteht aus dem Term (††) des ungestörten Problems nur ein einziger Term mit dem Multiplettsystem S und der Azimutalquantenzahl L , so ist seine Energie direkt durch (18a) gegeben und mithin auf Quadraturen zurückgeführt.

Bei der Berechnung der skalaren Produkte für (18a)

$$(\chi_{N_1 l_1 \mu_1 \sigma_1 \dots N_n l_n \mu_n \sigma_n}, \mathbf{W} \chi_{N_1 l_1 \mu_1 \sigma_1 \dots N_n l_n \mu_n \sigma_n})$$

$$= \frac{1}{n!} \sum_{P P'} (\varepsilon_P \mathbf{O}_P \psi_{\mu_1 \sigma_1}^{N_1 l_1}(1) \dots \psi_{\mu_n \sigma_n}^{N_n l_n}(n), \mathbf{W} \varepsilon_{P'} \mathbf{O}_{P'} \psi_{\mu_1 \sigma_1}^{N_1 l_1}(1) \dots \psi_{\mu_n \sigma_n}^{N_n l_n}(n)) \quad (19)$$

kann man den unitären Operator $\varepsilon_P \mathbf{O}_P$ auf den anderen Faktor als $\varepsilon_P^{-1} \mathbf{O}_P^{-1}$ hinüberziehen, und da er mit \mathbf{W} vertauschbar ist, mit $\varepsilon_{P'} \mathbf{O}_{P'}$ zu $\varepsilon_{P-1 P'} \mathbf{O}_{P-1 P'}$ vereinigen. Nun kann man die Summation anstatt über P' , über $P-1 P' = T$ führen, die Summation über P ergibt dann einfach $n!$. Es wird so aus (19)

$$\left(\psi_{\mu_1 \sigma_1}^{N_1 l_1}(1) \cdots \psi_{\mu_n \sigma_n}^{N_n l_n}(n), \sum_{i \neq k} \mathbf{W}_{ik} \sum_T \varepsilon_T \mathbf{O}_T \psi_{\mu_1 \sigma_1}^{N_1 l_1}(1) \cdots \psi_{\mu_n \sigma_n}^{N_n l_n}(n) \right),$$

wo \mathbf{W} in eine Summe von Gliedern \mathbf{W}_{ik} zerlegt ist, die den Wechselwirkungen der einzelnen Elektronenpaare entsprechen.

Betrachten wir das Glied mit einem bestimmten \mathbf{W}_{ik} . Es lautet, mehr im einzelnen hingeschrieben:

$$\sum_T \sum_{s_1 \cdots s_n} \int \cdots \int \psi_{\mu_1 \sigma_1}^{N_1 l_1}(1)^* \cdots \psi_{\mu_n \sigma_n}^{N_n l_n}(n)^* \cdot \mathbf{W}_{ik} \varepsilon_T \mathbf{O}_T \psi_{\mu_1 \sigma_1}^{N_1 l_1}(1) \cdots \psi_{\mu_n \sigma_n}^{N_n l_n}(n) dx_1 \cdots dz_n. \quad (20)$$

Wenn T hierin nicht nur das i -te und k -te Elektron berührt [also weder die Einheit, noch die Transposition (ik) ist], sondern noch etwa j in j' überführt, so verschwindet (20) bei der Integration über $x_j y_j z_j$ und Summation über s_j wegen der Orthogonalität der Eigenfunktionen $\psi_{\mu_j \sigma_j}^{N_j l_j}$ und $\psi_{\mu_{j'} \sigma_{j'}}^{N_{j'} l_{j'}}$. In einer erlaubten Konfiguration darf ja nicht gleichzeitig $N_j = N_{j'}$; $l_j = l_{j'}$; $\mu_j = \mu_{j'}$ und $\sigma_j = \sigma_{j'}$ sein. Bei dem Glied \mathbf{W}_{ik} genügt es also, T gleich der Einheit und gleich der Transposition (ik) zu setzen. In beiden Fällen gibt die Integration über die Cartesischen Koordinaten und Summation über die Spinkoordinaten aller von i und k verschiedenen Elektronen lediglich den Faktor 1, und (20) geht über in

$$\sum_{s_i, s_k} \int \cdots \int \psi_{\mu_i \sigma_i}^{N_i l_i}(i) \psi_{\mu_k \sigma_k}^{N_k l_k}(k) \mathbf{W}_{ik} \left[\psi_{\mu_i \sigma_i}^{N_i l_i}(i) \psi_{\mu_k \sigma_k}^{N_k l_k}(k) - \psi_{\mu_k \sigma_k}^{N_k l_k}(i) \psi_{\mu_i \sigma_i}^{N_i l_i}(k) \right] dx_i dy_i dz_i dx_k dy_k dz_k. \quad (20a)$$

Wenn man hierin noch $\psi_{\mu_i \sigma_i}^{N_i l_i}(i) = \psi_{\mu_i}^{N_i l_i}(\mathbf{r}_i) \delta_{s_i \sigma_i}$ usw. einsetzt, wo \mathbf{r}_i für die Cartesischen Koordinaten $x_i y_i z_i$ des i -ten Elektrons steht, kann man die Summation über s_i, s_k ausführen und erhält

$$\int \cdots \int \psi_{\mu_i}^{N_i l_i}(\mathbf{r}_i) \psi_{\mu_k}^{N_k l_k}(\mathbf{r}_k) \mathbf{W}_{ik} \left[\psi_{\mu_i}^{N_i l_i}(\mathbf{r}_i) \psi_{\mu_k}^{N_k l_k}(\mathbf{r}_k) - \delta_{\sigma_i \sigma_k} \psi_{\mu_k}^{N_k l_k}(\mathbf{r}_i) \psi_{\mu_i}^{N_i l_i}(\mathbf{r}_k) \right] dx_i dy_i dz_i dx_k dy_k dz_k. \quad (21)$$

Indem man die Integrale (21) für alle $\binom{n}{2}$ Paare ik und alle erlaubten Konfigurationen $(N_1 l_1 \mu_1 \sigma_1)$ $(N_2 l_2 \mu_2 \sigma_2) \cdots (N_n l_n \mu_n \sigma_n)$, für die $\mu_1 + \mu_2 + \cdots + \mu_n = \mu$ und $\sigma_1 + \sigma_2 + \cdots + \sigma_n = 2\nu$ ist, addiert, erhält man die Summe (18) der Störungsenergien aller aus $(\dagger\dagger)$ entstehenden Terme, deren $S \geqq |\nu|$ und $L \geqq |\mu|$ ist. Aus diesen Summen kann man nach (18a) auch die Summe der Energiestörungen erster Näherung der Terme mit einem bestimmten S und L berechnen.

Bezüglich weiterer Einzelheiten der Rechnung, insbesondere des Vergleichs der Integrale in (21), den man in gewissen Fällen ohne sie explizite zu berechnen ausführen kann, muß der Leser auf die Slatersche Originalarbeit verwiesen werden. Er wird dort auch zahlreiche interessante Beispiele durchgerechnet finden.

Formelsammlung

Störungstheorie.

$$\mathbf{V}_{lk} = (\psi_l, \mathbf{V} \psi_k) \quad (\text{V}, 8)$$

$$F_k = E_k + \lambda \mathbf{V}_{kk} + \lambda^2 \sum_{l \neq k} \frac{|\mathbf{V}_{lk}|^2}{E_k - E_l} \quad (\text{V}, 10)$$

$$\varphi_k = \psi_k + \lambda \sum_{l \neq k} \frac{\mathbf{V}_{lk}}{E_k - E_l} \psi_l. \quad (\text{V}, 11)$$

Gruppentheorie. \sum_R bedeutet bei endlichen Gruppen Summation über alle Gruppenelemente, bei kontinuierlichen Gruppen das Hurwitzsche Integral.

$$\sum_R J_R = \sum_R J_{SR}. \quad (\text{VII}, 1; \text{X}, 5)$$

Orthogonalitätsrelationen der unitären, irreduziblen Darstellungen einer Gruppe von der Ordnung h

$$\sum_R \mathbf{D}^{(j)}(R)_{\mu' \nu'}^* \mathbf{D}^{(j)}(R)_{\mu \nu} = \frac{h}{l_j} \delta_{j' j} \delta_{\mu' \mu} \delta_{\nu' \nu}, \quad (\text{IX}, 32)$$

l_j ist die Dimension von $\mathbf{D}^{(j)}$. Für die Charaktere $\chi^{(j)}(R)$ $= \sum_\mu \mathbf{D}^{(j)}(R)_{\mu \mu}$ gilt

$$\sum_R \chi^{(j)}(R)^* \chi^{(j)}(R) = h \delta_{j' j}. \quad (\text{IX}, 33)$$

Bei kontinuierlichen Gruppen ist $h = \sum_R 1$ durch $\int dR$ zu ersetzen
(X, 11, 12).

Darstellungen und Eigenfunktionen. Aus

$$\mathbf{P}_R f(x'_1, x'_2, \dots, x'_n) = f(x_1, x_2, \dots, x_n), \quad (\text{XI}, 19)$$

wo

$$x'_j = \sum_i \mathbf{R}_{ji} x_i \quad x_i = \sum_j \mathbf{R}_{ji}^* x'_j \quad (\text{XI}, 18)$$

ist, folgt

$$\mathbf{P}_{SR} = \mathbf{P}_S \mathbf{P}_R. \quad (\text{XI}, 20)$$

Aus

$$\mathbf{P}_R \psi_v = \sum_x \mathbf{D}(R)_{xv} \psi_x \quad (\text{XI}, 23)$$

und $\mathbf{P}_{SR} = \mathbf{P}_S \mathbf{P}_R$ folgt

$$\mathbf{D}(SR) = \mathbf{D}(S) \mathbf{D}(R). \quad (\text{XI}, 25)$$

Aus

$$\mathbf{P}_R f_x^{(j)} = \sum_k \mathbf{D}^{(j)}(R)_{kx} f_k^{(j)}; \quad \mathbf{P}_R g_{x'}^{(j')} = \sum_{k'} \mathbf{D}^{(j')}(R)_{k'x'} g_{k'}^{(j')}.$$

folgt

$$(f_x^{(j)}, g_{x'}^{(j')}) = \frac{1}{l_j} \delta_{jj'} \delta_{xx'} \sum_k (f_k^{(j)}, g_k^{(j')}). \quad (\text{XII}, 8)$$

Irreduzible Darstellungen der dreidimensionalen Drehgruppe.

$$\mathbf{D}^{(j)}(\{\alpha \beta \gamma\})_{m'm} = e^{i m' \alpha} \mathbf{d}^{(j)}(\beta)_{m'm} e^{i m \gamma}, \quad (\text{XV}, 8)$$

$$\mathbf{D}^{(1/2)}(\{\alpha \beta \gamma\})$$

$$= \begin{pmatrix} e^{-1/2 i \alpha} \cos \frac{1}{2} \beta e^{-1/2 i \gamma} - e^{-1/2 i \alpha} \sin \frac{1}{2} \beta e^{1/2 i \gamma} \\ e^{1/2 i \alpha} \sin \frac{1}{2} \beta e^{-1/2 i \gamma} & e^{1/2 i \alpha} \cos \frac{1}{2} \beta e^{1/2 i \gamma} \end{pmatrix}, \quad (\text{XV}, 16)$$

$$\mathbf{D}^{(j)}(\{\alpha \beta \gamma\})_{j\mu} = \sqrt{\binom{2j}{j-\mu}} e^{ij\alpha} \cos^{j+\mu} \frac{1}{2} \beta \sin^{j-\mu} \frac{1}{2} \beta e^{i\mu\gamma}, \quad (\text{XV}, 27\text{a})$$

$$\chi^{(j)}(\varphi) = \sum_{\mu=-j}^j e^{i\mu\varphi}. \quad (\text{XV}, 28)$$

In der Darstellung $\mathbf{D}^{(l)} \times \mathbf{D}^{(\bar{l})}$ sind die Darstellungen $\mathbf{D}^{(L)}$ mit

$$L = |l - \bar{l}|, \quad |l - \bar{l}| + 1, \quad \dots, \quad l + \bar{l} - 1, \quad l + \bar{l} \quad (\text{XVII}, 14)$$

je einmal enthalten.

$$\mathbf{D}^{(l)}(R)_{\mu'\mu} \mathbf{D}^{(\bar{l})}(R)_{\nu'\nu} = \sum_{L=|l-\bar{l}|}^{l+\bar{l}} s_{L\mu'\nu'}^{(l\bar{l})} \mathbf{D}^{(L)}(R)_{\mu'+\nu';\mu+\nu} s_{L\mu\nu}^{(\bar{l})}, \quad (\text{XVII}, 16\text{b})$$

$$s_{L\mu L-\mu}^{(l\bar{l})} = \frac{(-1)^{l-\mu} \sqrt{(2L+1)! (l+\bar{l}-L)!}}{\sqrt{(L+l+\bar{l}+1)! (L+l-\bar{l})! (L-\bar{l}+l)!}} \sqrt{\frac{(l+\mu)! (l+L-\mu)!}{(l-\mu)! (\bar{l}-L+\mu)!}}, \quad (\text{XVII}, 27\text{b})$$

$$\sum_{\mu} s_{L\mu m-\mu}^{(l\bar{l})} s_{L'\mu m-\mu}^{(\bar{l}\bar{l})} = \delta_{LL'}. \quad (\text{XVII}, 28)$$

Paulische Spintheorie.

$$\begin{aligned} & \mathbf{Q}_R \Phi(x_1 y_1 z_1 s_1 \dots x_n y_n z_n s_n) \\ = & \sum_{t_1 \dots t_n} \mathfrak{D}^{(1/2)}(R)_{1/2 s_1, 1/2 t_1} \dots \mathfrak{D}^{(1/2)}(R)_{1/2 s_n, 1/2 t_n} \Phi(x_1 y_1 z_1 t_1 \dots x_n y_n z_n t_n). \end{aligned} \quad (\text{XXI}, 6\text{ b})$$

$$\mathbf{O}_R = \mathbf{P}_R \mathbf{Q}_R = \mathbf{Q}_R \mathbf{P}_R. \quad (\text{XXI}, 8)$$

Irreduzibler Tensor.

$$\begin{aligned} \mathbf{O}_R^{-1} \mathbf{T}^{(\varrho)} \mathbf{O}_R &= \sum_{\sigma=-\omega}^{\omega} \mathfrak{D}^{(\omega)}(R)_{\varrho \sigma} \mathbf{T}^{(\sigma)}. \quad (\text{XXI}, 16\text{ b}) \\ \mathbf{T}_{Nj;\mu; N'j'\mu'}^{(\varrho)} &= (\Psi_{\mu}^{Nj}, \mathbf{T}^{(\varrho)} \Psi_{\mu'}^{N'j'}) \\ &= s_{j'\mu\varrho}^{(j\omega)} \delta_{\mu+\varrho, \mu'} T_{Nj; N'j'}. \end{aligned} \quad (\text{XXI}, 19)$$

Dabei ist für $s_{j'\mu\varrho}^{(j\omega)}$ Null zu setzen, wenn

$$|j - \omega| < j' \text{ oder } j' > j + \omega$$

ist.

Infinitesimale Drehungen. Infinitesimale Drehung der Cartesischen Koordinaten:

$$\mathbf{L}_z \Psi = \frac{1}{i} \frac{\partial}{\partial \alpha} \mathbf{P}_{\{\alpha \neq 0\}} \Psi \quad (\text{für } \alpha = 0), \quad (\text{XVIII}, 7)$$

der Spinkoordinaten:

$$\frac{1}{2} (s_1 + s_2 + \dots + s_n) \Psi = \frac{1}{2} \mathbf{S}_z \Psi = \frac{1}{i} \frac{\partial}{\partial \alpha} \mathbf{Q}_{\{\alpha \neq 0\}} \Psi \quad (\text{für } \alpha = 0)$$

des ganzen Zustandes: (XXIII, 23 b)

$$\mathbf{L}_z + \frac{1}{2} \mathbf{S}_z = \frac{1}{i} \frac{\partial}{\partial \alpha} \mathbf{O}_{\{\alpha \neq 0\}} \quad (\text{für } \alpha = 0). \quad (\text{XXIII}, 30\text{ a})$$

Sachregister

- | | |
|---|---|
| <p>Abelzsche Gruppe 64, 157.
Abschirmung 308.
Abseparieren des Schwerpunktes 190
 227.
Addition von Vektoren 1.
— — Matrizen 8.
Adjungierte Matrix 26.
Ähnlichkeitstransformation 10.
Alternierende Gruppe 136.
Antisymmetrische Darstellung 137.
— Eigenwerte und Eigenfunktionen
 112, 272.
Äquivalenz von Darstellungen 80, 95.
Äquivalenzprinzip 312.
Assoziativgesetz 6, 64.
Assozierte Darstellung 138, 195, 277.
Aufbauprinzip 198, 306.
Aufspalten der Eigenwerte 49, 130.
— der Spektrallinien 213, 288, 290,
 300.
Ausreduzieren 93.
Auswahlregeln 197, 210, 228, 233, 285,
 286.
— in äußeren Feldern 215, 289, 293.
Axialer Vektor 213, 255, 287.
Azimutale Quantenzahl 192, 194, 211,
 230, 270.</p> <p>Bahn 309, 313, Bohrsche — 192.
Beiderseitiger Skalar 296.
Bein, n-Bein 227.
Bohrsche Frequenzbedingung 60.</p> <p>Cayley-Kleinsche Parameter 170.
Charakter der Darstellungen 91, 126.
— — — der symmetrischen Gruppe
 148.
— — — der dreidimensionalen Dreh-
 gruppe 166, 181.</p> | <p>Charakter der Darstellungen der uni-
 tären Gruppe 179.</p> <p>Darstellung 79, 110, 120.
 bis auf einen Faktor 266.
—, mehrdeutige 267.
— des direkten Produktes 185.
— der dreidimensionalen Drehgruppe
 164, 179.
— — — Drehspiegelungsgruppe 189,
 196.
— — — symmetrischen Gruppe 139, 275.
— — — zweidimensionalen Drehgruppe
 156.
— — — Drehspiegelungsgruppe 158,
 221, 290.
— — — unitären Gruppe 173.
— eines Eigenwerts 128.
— kontinuierlicher Gruppen 103, 109.
Diagonalf orm von Matrizen 24.
Diagonalmatrix 9.
Diagonalsumme 10.
Diedergruppe 70.
Dimension von Matrizen 3.
— — Darstellungen 80.
Dipolstrahlung 209.
Diracsches Elektron 256.
Direktes Produkt von Matrizen 24, 185.
— — — Gruppen 184, 196, 290, 314.
Diskretes Spektrum 41.
Drehachse 160.
Drehelektron 236.
Drehgruppe 98, 152, 160.
Drehimpuls 195, 236, 283.
Drehspiegelung 152, 290.
Drehung, Begriff der 98, 152.
— des Koordinatensystems 115, 194,
 240.
— — Zustand 241.</p> |
|---|---|

- | | |
|--|---|
| <p>Drehung der Cartesischen Koordinaten 244, 270, 314.
 — — Spinkoordinaten 244, 270, 281, 316.</p> <p>Drehwinkel 24.</p> <p>Eigendifferential 41.</p> <p>Eigenfunktionen 37, 41, 54.
 — und Darstellungen 110.</p> <p>Eigenvektor 23.</p> <p>Eigenwerte der Schrödinger-Gleichung 37.
 — einer Matrix 22.
 — eines Operators 40.
 —, verschiedene Arten von 119, 133, 187.</p> <p>Einfache Gruppe 74.</p> <p>Einfachkontinuierliche Gruppe 98.</p> <p>Einheit der Gruppe 64.</p> <p>Einheitsmatrix 6.</p> <p>Einstein'sche Übergangswahrscheinlichkeiten 204.</p> <p>Elektrische Quantenzahl 221.</p> <p>Elementarteiler 24.</p> <p>Elemente der Gruppe 63.</p> <p>Energiewerte 34.</p> <p>Entartung 42, normale und zufällige 129, 309.</p> <p>Erlaubte Konfiguration 312.</p> <p>Eulersche Winkel 99, 163.</p> <p>Faktorgruppe 75, 78.</p> <p>Feinstruktur 197, 270, 279.</p> <p>Feinstrukturkonstante 284, 307.</p> <p>Funktionen einer Matrix 11, 25.</p> <p>Funktionenvektor 38.</p> <p>g-Formel 301.</p> <p>Gemischtkontinuierliche Gruppe 98, 102, 109.</p> <p>Gerade Darstellung der unitären Gruppe 174.
 — Elektronenzahl 258.
 — Permutation 136.</p> <p>Gesamtquantenzahl 197, 261, 286.</p> <p>Gestrichene Terme 195, 233.</p> <p>Grobstruktur 273.</p> <p>Größe, physikalische 52.</p> | <p>Grundgebiet 226.</p> <p>Gruppe 63, endliche — 64, kontinuierliche — 97.</p> <p>Gruppenelement 63.</p> <p>Gruppenparameter 97.</p> <p>Gruppenmultiplikation 64.</p> <p>Gruppentafel 63.</p> <p>Halbzählig Darstellung 174.</p> <p>Hamiltonscher Operator 37, 52.
 — — des Magnetfeldes 218, 300.</p> <p>Hauptachsentransformation 22.
 — von unitären und hermiteischen Matrizen 29.</p> <p>Hauptquantenzahl 191.</p> <p>He-Atom 233, 305.</p> <p>Heisenbergsche Matrizen 34.</p> <p>Hermiteische Matrix 26.</p> <p>Hermiteischer Operator 39.</p> <p>Holomorphie 69, 76.</p> <p>Hurwitzsches Integral 104, 156, 162.</p> <p>Identische Darstellung 137, 167.</p> <p>Index der Untergruppe 67.</p> <p>Infinitesimalgruppe 100.</p> <p>Innere Quantenzahl ist Gesamtquantenzahl 197, 261, 286.</p> <p>Integral über kontinuierliche Gruppen 104, 156, 162.</p> <p>Intensitätsregeln 289, 297.</p> <p>Interkombinationsverbot 211, 316.</p> <p>Intervallregel 302.</p> <p>Invariante Operatoren 116.</p> <p>Invarianten gegenüber Ähnlichkeitstransformationen 22.</p> <p>Invarianzforderungen 249, 258.</p> <p>Inversion 189, 194, 255.</p> <p>Irreduzible Bestandteile 95.
 — Darstellung 80, 120, 128.
 — Säkulargleichung 132.
 — Tensoren 182, 263, 293.</p> <p>Isomere, optische 234.</p> <p>Jacobische Polynome 232.</p> <p>Kanonische Transformation 56, 242, 254.</p> <p>Kernspin 256.</p> |
|--|---|

- | | |
|---|--|
| <p>Klasse 72, 135, 154, 156, 161, 179, 320.</p> <p>Koeffizient der Matrix 3.</p> <p>Kommutativgesetz 5, 64.</p> <p>Komplex von Elementen 75.</p> <p>Komplexorthogonale Matrix 26.</p> <p>Komponente des Drehimpulses 195.</p> <p>— — Vektors 1.</p> <p>Konfiguration 310.</p> <p>Konfigurationsraum 35, 114.</p> <p>Konjugierte Elemente 72.</p> <p>Kontinuierliche Gruppe 97.</p> <p>Kontinuierliches Spektrum 42, 191.</p> <p>Kreisel 230.</p> <p>Kristall 223.</p> <p>Kugelfunktionen 164, 229.</p> <p>Kugelsymmetrie der S-Zustände 228.</p> <p>Landésche g-Formel 301.</p> <p>Laportesche Regel 212, 220, 287, 293, 321.</p> <p>Lineare Abhängigkeit und Unabhängigkeit von Funktionen 38.</p> <p>— — — — Vektoren 11.</p> <p>Linearkombination, richtige 51, 131, 188, 203, 218, 277, 283.</p> <p>Linienbreite 62, 284.</p> <p>Longitudinaleffekt 216.</p> <p>Magnetische Quantenzahl 192, 215, 261.</p> <p>Magnetisches Moment 236.</p> <p>Matrix 1.</p> <p>Matrixelemente 57, 209, 264, 293.</p> <p>Messung einer Größe 53.</p> <p>Mitbewegung des Kerns 190.</p> <p>Multiplett 197.</p> <p>Multiplettaufspaltung 283.</p> <p>Multiplettsystem 194, 270, 278, 315.</p> <p>Multiplizität 196.</p> <p>n-Bein 227, 265.</p> <p>Nebengruppe 66.</p> <p>— des Normalteilers 74.</p> <p>Negative Terme 195, 321.</p> <p>Nichtquadratische Matrix 15.</p> <p>Normale Kopplung 285, 294.</p> | <p>Normaler Zeeman-Effekt 220.</p> <p>Normalteiler 73.</p> <p>Normierung 27.</p> <p>Nullmatrix 8.</p> <p>O, Definition 243, 257, 271.</p> <p>Operator, hermitischer 39, 52.</p> <p>—, linearer 3, 38, 242, 253.</p> <p>—, unitärer 56, 242, 253.</p> <p>Ordnung der Gruppe 64.</p> <p>— eines Gruppenelements 65.</p> <p>Orthogonale Funktionen 38.</p> <p>— Matrix 26.</p> <p>— Vektoren 27, 31.</p> <p>Orthogonalisierungsverfahren, Schmidtsches 31.</p> <p>Orthogonalitätsrelationen der Charaktere 92.</p> <p>— — Darstellungs-koeffizienten 87.</p> <p>— — — bei kontinuierlichen Gruppen 110.</p> <p>— — Eigenfunktionen 42.</p> <p>— für Funktionen 123.</p> <p>Orthogonalsystem 42.</p> <p>— der Funktionen der s 275.</p> <p>P, Definition und Eigenschaften 114.</p> <p>Parameter für Gruppen 97.</p> <p>Parameterraum 97, 153.</p> <p>Partitio numerorum 135, 151.</p> <p>Partner 120.</p> <p>Pauliprinzip 138, 195, 272, 311, 313.</p> <p>Paulische Matrizen 169, 251.</p> <p>Periode eines Gruppenelements 65, 100.</p> <p>Permutationsgruppe 70, 133.</p> <p>Physikalische Größe 52.</p> <p>Polarer Vektor 220, 287, 293.</p> <p>Positive Terme 195.</p> <p>Potenzen von Matrizen 9.</p> <p>Produkt, direktes von Matrizen 19.</p> <p>—, skalares von Funktionen 38.</p> <p>— — — Vektoren 27.</p> <p>— von Komplexen 77.</p> <p>— — Matrizen 5.</p> <p>— — Permutationen 71, 133.</p> <p>— — Zahl und Matrix 9.</p> <p>— — — — Vektor 1.</p> <p>Protonenspin 256.</p> |
|---|--|

- Q**, Definition 243, 257, 271.
Quadratintegrbarkeit 36, 41.
Quadratische Matrix 15.
 — — im weiteren Sinne 17.
Quadrupolstrahlung 210, 285.
Quantenzahlen 192—197, 221.
- Reducible Darstellung** 89.
Reell orthogonale Matrix 26, 33, 152.
Reine Drehung 152.
Relativistische Theorie des Elektrons 256.
Reziproke Matrix 7.
Rhomatische Symmetrie 224.
Richtige Linearkombinationen 51, 131, 188, 203, 218, 277, 283.
Rotator 229.
Rydbergkonstante 191.
- s**, die Variable 139, 237, 274, 281.
Säkulargleichung 22, 50, 129.
Schiefe Matrix 26.
Schmidtsches Verfahren 31.
Schrödingergleichung 37.
Schursches Lemma 83.
Schwerpunktskoordinaten 190.
Skalar 261, 295.
Skalares Produkt von Funktionen 38, 139, 238, 256.
 — — — Vektoren 27.
Spektrum 40.
Spiegelungscharakter 194, 233, 283, 321.
Spiegelungsgruppe 68, 189.
Spin 197, 236.
Spinfreie Versuche und spinfreie Größe 238, 270.
Spinkräfte 273, 283, 302.
Spinoperatoren 250.
Spinvariable 237.
Spur 10.
Starkeffekt 213, 222, 290.
Statistische Deutung 51.
Stetigkeit 97, 103, 267.
Störungstheorie 44, 129, 188.
Summensatz 289.
Symmetriegruppe 69.
 — des Konfigurationsraumes 114.
- Symmetrische Eigenwerte und Eigenfunktionen** 112.
 — Gruppe 71, 133.
 — Matrix 26, 33.
 — Operatoren 117, 124, 261, 304.
Symmetrieverhältnisse 194, 314.
 — in äußeren Feldern 214, 224.
- Tensoren** 182.
Tensoroperator 262, 293.
Term 34, 191.
 —, positiver und negativer bzw. gestrichener und ungestrichener 195.
 —, Singulett-, Doublet- usw. 196.
Transformation, eigentliche 2.
 —, lineare 2.
 — von Vektoren und Tensoren 182.
Transformationstheorie 51.
Transponierte Matrix 25.
Transposition 136.
Transversaleffekt 216.
- Überlagerungsgruppe** 268.
Übergang zu einem neuen Koordinatensystem 55.
Übergangswahrscheinlichkeit durch Strahlung 35, 57, 209.
 — — — Messung 54, 241.
Übermatrizen 20.
Unabhängigkeit, lineare 11, 38.
Unendliche Gruppe 97.
Ungerade Darstellung der unitären Gruppe 174.
 — Elektronenzahl 259.
 — Permutationen 136.
Ungestrichene Terme 195, 233.
Unitäre Gruppe 168, 260.
 — — —, Darstellungen 173.
 — Matrix 26.
Unitarisierbarkeit der Darstellungen 83.
Untergruppe 65.
Untermatrizen 20.
- Vektor** 1, axialer und polarer 213, 220, physikalischer 182.
Vektoradditionsmodell 198, 283, 307.
Vektoroperator 251, 262, 297, 301.
Verbreiterung der Terme und Linien 284.

- | | |
|---|--|
| <p>Vereinigung von Symmetrien 184.
 — — Systemen 199.</p> <p>Vergleich von Matrixelementen 264,
 293.</p> <p>Vertauschbarkeit 5, 124.</p> <p>Vertauschung der Elektronen 115, 194,
 271, 281, 310, 316.</p> <p>Vertauschungsrelationen 35.</p> <p>Vierergruppe 69.</p> <p>Vollständiges Orthogonalsystem 43.
 — Vektorensystem 13.</p> <p>Wahrscheinlichkeiten, quantenmecha-
 nische 53.</p> <p>Wahrscheinlichkeitszusammenhang ver-
 schiedener Spinrichtungen 247.</p> <p>Wasserstoffatom 191, 230.</p> <p>Wechselwirkung der Elektronen 307.
 — von Licht und Atom 58.
 — des Magnetfeldes mit dem Atom
 218, 300.
 — der Spin mit den Bahnen 283, 302.
 — der Spin untereinander 304.</p> | <p>Wege im Parameterraum 98, 102,
 268.</p> <p>Wellenfunktion 37, 51.</p> <p>Zeeman-Effekt 213, 288, 300.</p> <p>Zerlegung von n in Summanden 135,
 149.</p> <p>Zerteilung eines Systems 272.</p> <p>Zufällige Entartung 129, 309.</p> <p>Zugehörigkeit zu einer irreduziblen
 Darstellung 126.
 — — — Zeile einer irreduziblen Dar-
 stellung 120.</p> <p>Zusammenhängend, einfach oder mehr-
 fach 99, 268.</p> <p>Zusammensetzung von Transformationen
 4.</p> <p>Zustand 51.</p> <p>Zweideutige Darstellung 174, 207, 258.</p> <p>Zweidimensionale Drehgruppe 153.
 — Drehspiegelungsgruppe 155, 220.
 — unitäre Gruppe 168.</p> <p>Zyklische Gruppe 67.</p> |
|---|--|