

Struktur der Materie in Einzeldarstellungen

Max Born · Pascual Jordan

Elementare Quantenmechanik

**STRUKTUR DER MATERIE
IN EINZELDARSTELLUNGEN**

HERAUSGEGEBEN VON
M. BORN-GÖTTINGEN UND J. FRANCK-GÖTTINGEN

IX

**ELEMENTARE
QUANTENMECHANIK**

(ZWEITER BAND DER VORLESUNGEN
ÜBER ATOMMECHANIK)

VON

DR. MAX BORN
PROFESSOR AN DER UNIVERSITÄT
GÖTTINGEN

UND **DR. PASCUAL JORDAN**
PROFESSOR AN DER UNIVERSITÄT
ROSTOCK



SPRINGER-VERLAG BERLIN HEIDELBERG GMBH 1930

ISBN 978-3-662-00271-1 ISBN 978-3-662-00291-9 (eBook)
DOI 10.1007/978-3-662-00291-9

ALLE RECHTE, INSBESONDERE DAS DER ÜBERSETZUNG
IN FREMDE SPRACHEN, VORBEHALTEN.

COPYRIGHT 1930 BY SPRINGER-VERLAG BERLIN HEIDELBERG
URSPRÜNGLICH ERSCHIENEN BEI JULIUS SPRINGER IN BERLIN 1930

NIELS BOHR

GEWIDMET

Vorwort.

Dieses Buch ist die Fortsetzung der im Jahre 1925 erschienenen „Vorlesungen über Atommechanik“; es ist der dort im Vorwort angekündigte „zweite Band“, dessen „virtuelle Existenz dazu dienen sollte, Ziel und Sinn dieses Buches deutlich zu machen“. Die Hoffnung, daß der Schleier, der damals noch über der eigentlichen Struktur der Atomgesetze hing, bald fallen müsse, hat sich in überraschend schneller und gründlicher Weise erfüllt. Kurz nach dem Erscheinen des ersten Bandes — und im Zusammenhang mit den im Göttinger Kreise bei der Ausarbeitung des Werkes betriebenen Studien — entstand die Quantenmechanik, an der gleich nach HEISENBERGS grundlegender Abhandlung die beiden Verfasser dieses Buches mitarbeiten konnten. Bald darauf erwuchs ganz unabhängig aus den kühnen Gedanken DE BROGLIES SCHRÖDINGERS Wellenmechanik. Es folgte eine stürmische Entwicklung, die in wenigen Jahren eine ungeheure Fülle von Ergebnissen ans Licht brachte. Dabei erwies sich die SCHRÖDINGERSche Methode als besonders fruchtbar und übte durch ihre Ähnlichkeit mit den bekannten Vorstellungsweisen der älteren theoretischen Physik (die Verwendung von Differentialgleichungen in einem Kontinuum) eine große Anziehungskraft aus. So ist es gekommen, daß es bereits einige zusammenfassende Darstellungen gibt, die fast ausschließlich die Wellenmechanik behandeln (wir nennen die in deutscher Sprache geschriebenen Bücher: A. HAAS, Materiewellen und Quantenmechanik; A. SOMMERFELD, Atombau und Spektrallinien, Wellenmechanischer Ergänzungsband; L. DE BROGLIE, Einführung in die Wellenmechanik; J. FRENKEL, Einführung in die Wellenmechanik).

Um so mehr schien es angebracht, ein Bild der neuen Mechanik zu entwerfen, das von einem ganz anderen Gesichtspunkt aufgenommen ist. Wir sind überzeugt, daß die wahre Verbindung

zwischen klassischer und neuer Mechanik nicht so sehr auf der formalen Ähnlichkeit der wellenmechanischen Differentialgleichungen und Randwertprobleme mit den Methoden der klassischen Kontinuumsphysik beruht, als auf dem Korrespondenzprinzip von NIELS BOHR; ja, wir sehen in der neuen Mechanik die strenge Durchführung des BOHRschen Programms, wie es im 1. Bande dieses Buches entwickelt worden ist. Wir sind der Meinung, daß die neue Mechanik keine Annäherung an die klassischen Vorstellungen und keine Abkehr von den BOHRschen Begriffsbildungen (stationäre Zustände, Quantensprünge, Übergangswahrscheinlichkeiten usw.) bedeutet, sondern gerade ihren systematischen Ausbau zu einem logisch geschlossenen System. Dies haben wir zum Ausdruck gebracht, indem wir den Begründer der modernen Atomtheorie baten, ihm dieses Buch widmen zu dürfen.

Bei der Durchführung unserer Absicht haben wir uns von dem Gedanken leiten lassen, einmal zu versuchen, wie weit man mit elementaren, d. h. in der Hauptsache algebraischen Mitteln kommen kann; wir wollten alle grundlegenden Begriffe und Vorstellungen in engem Anschluß an das Korrespondenzprinzip entwickeln und durch Beispiele erläutern, dann erst die wellentheoretischen Methoden SCHRÖDINGERS heranziehen. Aber die Beispiele wuchsen uns unter den Händen in solchem Maße, daß wir einen ganzen Band vom üblichen Umfang dieser Sammlung beisammen hatten, ehe wir noch zur Wellenmechanik gelangen konnten. Wir waren also gezwungen, für diese einen neuen Band in Aussicht zu nehmen, den wir schreiben wollen, sobald Zeit und Kraft es gestatten. Hier aber sei betont, daß die Vermeidung der Wellenmechanik keinerlei Werturteil bedeuten soll, sondern nur einen Versuch, die Zusammenhänge von einem anderen Standpunkt aus zu beleuchten, der in gewisser Hinsicht als elementarer bezeichnet werden kann. Wer von dieser Seite in die neue Atommechanik eingedrungen ist, wird dann um so leichter die wundervollen Gedanken DE BROGLIES und SCHRÖDINGERS auffassen können.

Wir haben nach einer kurzen Einleitung die mathematischen Hilfsmittel in einem besonderen Kapitel zusammengestellt, weil es uns nicht angebracht schien, es dem Leser zu überlassen, sie aus mathematischen Lehrbüchern zusammenzusuchen; für einen Anhang aber schienen sie uns zu wichtig. (In den Anhängen haben

wir nur solche Gegenstände behandelt, die eigentlich in den ersten Band gehören, aber dort fehlen, sowie einen kleinen mathematischen Hilfssatz.) Im folgenden Kapitel sind möglichst solche Gegenstände behandelt, die in den genannten Büchern über Wellenmechanik wenig oder gar keine Beachtung finden. Natürlich mußten die wichtigsten Eigenschaften des Wasserstoffatoms als Paradigma abgeleitet werden (in engem Anschluß an PAULIS diesbezügliche Arbeit); sie erscheinen aber als Anwendung von allgemeinen Sätzen über den Drehimpuls, die für beliebige Atome gelten — wobei der Spin des Elektrons allerdings noch unberücksichtigt bleibt. Eine strenge Ableitung dieser Sätze auf wellenmechanischer Grundlage, wie sie WEYL in seinem inhaltreichen aber schwierigen Werk „Gruppentheorie und Quantenmechanik“ gibt, erfordert einen beträchtlich größeren Aufwand an abstrakter Mathematik, als unsere Darstellung. Das folgende Kapitel über Störungsrechnung hat sich zu einer vollständigen Übersicht über die Dispersionsoptik, insbesondere die elektro- und magneto-optischen Effekte ausgewachsen und bietet manches, was in der Literatur nicht zu finden ist. Dann kommt ein Kapitel über die statistische Deutung der Quantenmechanik, wobei wir auch den von J. v. NEUMANN durchgeführten weiteren Ausbau der statistischen Transformationstheorie gebührend berücksichtigen. Das letzte Kapitel enthält die Quantenmechanik des Strahlungsfeldes, welche zur Aufklärung des Lichtquantenproblems führt, und im Anschluß daran die strenge DIRACsche Theorie von Emission und Absorption des Lichts; es führt so gewissermaßen zum Ausgangspunkt der HEISENBERGsehen Theorie zurück, deren Grundannahmen von dem schrittweise gewonnenen höheren Standpunkt aus gerechtfertigt werden.

Wir haben mehreren freundlichen und eifrigen Helfern zu danken: zunächst den Herren W. PAULI JR. und J. v. NEUMANN, denen wir wichtige Hinweise und Anregungen verdanken; sodann Herrn Dr. L. ROSENFELD, der von Anfang fast bis zum Abschluß mitgewirkt und manche Abschnitte, besonders über Optik, selbständig entworfen hat; Fräulein TATJANA EHRENFEST, die ihn am Schluß ablöste, den ganzen Text kritisch durchsah und viele wertvolle Ratschläge gab; Herrn Dr. W. HEITLER, der an den magneto-optischen Paragraphen mitwirkte; Fräulein MARIA GÖPPERT, die DIRACs Theorie der Emission, Absorption und

Dispersion darstellen half. Außer diesen haben auch die Herren Dr. W. GORDON und H. A. SENFTLEBEN beim Lesen der Korrekturen geholfen.

Dem Verlage danken wir für großes Entgegenkommen bei der Drucklegung, die sich infolge äußerer Hinderungen sehr verzögerte, und für die schöne Ausstattung des Werkes.

Göttingen, Rostock, den 6. Dezember 1929.

M. BORN. P. JORDAN.

Inhaltsverzeichnis

Erstes Kapitel: **Physikalische Grundlegung.**

	Seite
§ 1. Rückblick auf die Bohrsche Theorie	1
§ 2. Vorläufer der Quantenmechanik	8
§ 3. Heisenbergs korrespondenzmäßige Kinematik	13

Zweites Kapitel: **Mathematische Grundlagen.**

§ 4. Das Rechnen mit quadratischen Matrizen	21
§ 5. Rechteckige Matrizen und Übermatrizen	27
§ 6. Matrizen mit Symmetrien	30
§ 7. Matrizenfunktionen	35
§ 8. Einzelmatrizen	39
§ 9. Vektoren im Hilbert-Raum	41
§ 10. Hilbert-Tensoren und metrische Eigenschaften des Hilbert-Raums	49
§ 11. Hauptachsentransformation der hermiteschen Formen	54
§ 12. Extremaleigenschaften der Eigenwerte	62
§ 13. Gleichzeitige Hauptachsentransformation vertauschbarer hermitischer oder unitärer Matrizen	65
§ 14. Funktionen vertauschbarer hermitescher oder unitärer Matrizen	68
§ 15. Einfache (irreduzible) und zerlegbare (reduzible) Matrizen-systeme	73
§ 16. Simultane Diagonaltransformation zweier Formen. Reelle Formen	79

Drittes Kapitel: **Die Gesetze der Matrizenmechanik**

§ 17. Die kanonischen Vertauschungsregeln	83
§ 18. Folgerungen aus den Vertauschungsregeln	90
§ 19. Die kanonischen Bewegungsgleichungen für abgeschlossene Systeme	96
§ 20. Existenz und Bestimmtheit der Lösung der Bewegungsgleichungen. Zurückführung auf ein Eigenwertproblem	101
§ 21. Integrale der Bewegungsgleichungen und kontinuierliche Gruppen	107
§ 22. Systeme mit zeitabhängiger Energie	114
§ 23. Der lineare harmonische Oszillator. Existenz- und Eindeutigkeitsuntersuchung der kanonischen Matrizen-systeme	122
§ 24. Systeme harmonischer Oszillatoren	129

Viertes Kapitel: Die Sätze über den Drehimpuls.

§ 25. Vertauschungsregeln	132
§ 26. Vertauschbarkeit des Drehimpulses mit drehinvarianten Größen	135
§ 27. Zerlegung der Drehimpulsmatrizen in einfache Bestandteile	139
§ 28. Auswahlregeln für die Drehimpuls-Quantenzahlen	143
§ 29. Schwingungsamplituden beim Zeeman-Effekt	147
§ 30. Dreh- und Spiegelungsinvarianten von Raumvektoren und Raumtensoren	151
§ 31. Intensitäten beim Zeeman-Effekt	158
§ 32. Ganzzahligkeit der inneren Quantenzahl bei Systemen von Massenpunkten	162
§ 33. Separation der Variablen bei der Zentralbewegung eines Massen- punktes	168
§ 34. Die zweiatomige Molekel	171
§ 35. Integrale der Bewegungsgleichungen beim Wasserstoffatom	177
§ 36. Ableitung der Balmer-Formel	184
§ 37. Verhalten des H-Atoms in gekreuzten Feldern	190

Fünftes Kapitel: Störungstheorie.

§ 38. Störungsrechnung für abgeschlossene nicht entartete Systeme	194
§ 39. Der anharmonische Oszillator	200
§ 40. Koppelung von Rotationen und Schwingungen bei der zwei- atomigen Molekel	203
§ 41. Störungstheorie für abgeschlossene Systeme mit entarteter Energie	205
§ 42. Starkeffekt und elektrische Polarisierbarkeit	212
§ 43. Zeeman-Effekt und Magnetisierung	225
§ 44. Verhalten des H-Atoms in gekreuzten Feldern	234
§ 45. Störungstheorie nicht abgeschlossener Systeme	236
§ 46. Dispersion des Lichts	240
§ 47. Optische Aktivität	250
§ 48. Elektrische Doppelbrechung (Kerr-Effekt)	259
§ 49. Magnetische Doppelbrechung (Faraday-Effekt)	267
§ 50. Intensität der Streustrahlung	276
§ 51. Intensität und Polarisation des Fluoreszenzlichts	282

Sechstes Kapitel: Statistische Deutung der Quantenmechanik.

§ 52. Meßbare Größen und Eigenwerte	288
§ 53. Wahrscheinlichkeiten quantenmechanischer Größen in statio- nären Zuständen	292
§ 54. Allgemeine Relativwahrscheinlichkeiten zweier Größen	294
§ 55. Wahrscheinlichkeitsvektoren	300
§ 56. Grenzen der Meßbarkeit	302
§ 57. Statistik quantenmechanischer Gemenge	305
§ 58. Der reine Fall als spezielles Gemenge	312
§ 59. Das allgemeine Gemenge als Mischung reiner Fälle	317

	Seite
§ 60. Quantenmechanik und Determinismus	322
§ 61. Übergangswahrscheinlichkeit bei äußerer Einwirkung	326
§ 62. Der Adiabatensatz	333
§ 63. Wahrscheinlichkeit der Lichtabsorption	340
§ 64. Innere Resonanz entarteter Systeme	344
§ 65. Resonanz zwischen gekoppelten Systemen	356

Siebentes Kapitel: Einleitung in die Quantentheorie des Lichtes.

§ 66. Einleitung	365
§ 67. Das klassische Modell des Strahlungsfeldes	367
§ 68. Asymptotische Verteilung der Eigenschwingungen	372
§ 69. Die Plancksche Formel	374
§ 70. Andere Ableitung des Planckschen Gesetzes	377
§ 71. Quantenmechanik der Eigenschwingungen	383
§ 72. Einsteins thermodynamische Schwankungsgesetze der Strahlung	386
§ 73. Wellentheoretische Ableitung des zweiten Einsteinschen Schwan- kungsgesetzes	392
§ 74. Absorption und Emission von Strahlung durch Atome	400
§ 75. Zerstreuung und Dispersion	404
§ 76. Thermisches Gleichgewicht zwischen Strahlung und Atomen	408

Anhang.

I. Emission und Absorption elektromagnetischer Wellen nach der klassischen Theorie	413
II. Die Jacobi-Poissonschen Klammersymbole in der klassischen Mechanik	420
III. Hilfssatz zum Adiabatensatz	427
Sachverzeichnis	430

Erstes Kapitel. Physikalische Grundlegung.

§ 1. Rückblick auf die BOHRsche Theorie. Der erste Band dieses Werkes schloß mit der Einsicht ab, daß die darin entwickelten Prinzipien der Quantentheorie zu keiner vollständigen Darstellung der physikalischen Tatsachen führen; sie ermöglichen zwar eine allgemeine Übersicht über die qualitativen Erfahrungen im Bereiche der Atomprozesse, aber sie geben außer in seltenen Sonderfällen (Energienstufen beim Ein-Elektron-Problem) keine quantitativ richtigen Ergebnisse.

Die Fortentwicklung der Quantentheorie, die den Gegenstand dieses Bandes bilden soll, brauchte daher eine ganze Reihe neuer Gedanken. Aber ehe wir diese darstellen, wollen wir betonen, daß die grundlegenden Annahmen, welche die Quantentheorie von der klassischen Theorie unterscheiden, nicht nur ihre Geltung bewahrt haben, sondern erst recht fest begründet worden sind: Wir meinen das Auftreten von Diskontinuitäten (Quantenzuständen und Quanten-Sprüngen), die Existenz von Energiequanten, bestimmt durch die PLANCKsche Konstante h .

Es ist also keine Rede von einer „Erklärung“ der fremdartigen Quantengesetze durch eine Zurückführung auf klassische Vorstellungen; im Gegenteil ist der prinzipielle, primäre Charakter der quantentheoretischen Grundannahmen erst durch die neuere Entwicklung klar zum Vorschein gekommen. Der Fortschritt besteht gerade im Abstreifen der Reste klassischer Betrachtungsweise; dadurch ist eine in sich geschlossene Theorie entstanden, die alle Atomvorgänge widerspruchsfrei zu beschreiben gestattet und die klassische Theorie als speziellen Grenzfall enthält.

Die klassische Theorie beruht ja im wesentlichen auf Erfahrungen mit greifbaren und sichtbaren, „makroskopischen“ Körpern. Als nun die experimentelle Forschung der letzten Jahrzehnte die Realität der Atome und Elektronen zur Gewißheit machte, war es selbstverständlich zunächst geboten, diese kleinen Partikel

genau so wie gewöhnliche Körper zu behandeln. Aber die wachsende Erfahrung lehrte bald, daß die Realität der Atome doch von anderer Art sein müßte als die der makroskopischen Körper; die bekannten physikalischen Begriffe wollten nicht mehr passen. Man folgte den neuartigen Erscheinungen der Atomphysik mit den geläufigen Vorstellungsweisen so weit wie möglich; aber die Erscheinungen selbst zwangen allmählich zu Verallgemeinerungen und Umwandlungen. So entstand die im ersten Bande dargestellte BOHRsche Theorie; diese stellt eine Verknüpfung der klassischen Gesetze mit den von PLANCK und EINSTEIN aufgedeckten Gesetzmäßigkeiten der Energie- und Wirkungsquanten dar, also zweier Vorstellungsweisen, die sich im Grunde durchaus widersprechen. Und doch macht dieses Verfahren, wie BOHR zeigen konnte, einen großen Teil der Tatsachen der Atomphysik verständlich. Die Brücke bildet das von BOHR formulierte *Korrespondenzprinzip*, das den diskontinuierlichen Quantenprozessen gewisse Eigenschaften der kontinuierlichen klassischen Prozesse zuordnet.

Will man einsehen, wie die neue Theorie eine natürliche Fortentwicklung der BOHRschen Ideen bedeutet, ist es notwendig, sich die ältere Theorie deutlich vor Augen zu halten. Darum wollen wir hier ihre wesentlichen Züge noch einmal kurz zusammenstellen.

Als Modell eines Atoms dient nach RUTHERFORD ein System, bestehend aus einem schweren Kern und einer Anzahl Elektronen, die um den Kern kreisen. Diese Bewegung ist nach den Gesetzen der klassischen Theorie näherungsweise mehrfach periodisch, aber nicht exakt, weil das System dauernd und stetig Energie durch Aussendung elektromagnetischer Kugelwellen verliert. Eine solche Welle läßt sich als Superposition monochromatischer Kugelwellen auffassen, deren Frequenzen mit denen der Elektronenbewegung im Atom übereinstimmen.

Bei Systemen, die mehrfach periodische Bewegungen ausführen können, gibt es kanonisch konjugierte Winkel- und Wirkungsvariable w_k, J_k , die durch einfache Periodizitätsforderungen im wesentlichen¹ eindeutig festgelegt sind. Die Energie wird eine

¹ „Im wesentlichen“ bedeutet: Die Wirkungsvariablen J_k sind eindeutig bestimmt bis auf eine lineare Transformation, die nebst ihrer Umkehrung ganzzahlige Koeffizienten hat (sogenannte „unimodulare“ Transformation).

Funktion $W(J_1, J_2, \dots) = W(J)$ der Wirkungsvariablen allein. Daher werden diese auf Grund der kanonischen Bewegungsgleichungen zeitlich konstant und die Winkelvariablen lineare Funktionen der Zeit:

$$(1) \quad w_k = \nu_k t + \delta_k,$$

wobei die Koeffizienten

$$(2) \quad \nu_k = \frac{\partial W}{\partial J_k}$$

Funktionen der J_k und die δ_k neue Integrationskonstanten sind. Da die geometrischen Bestimmungsstücke der Bewegung periodisch von den w_k abhängen, so bedeuten die ν_k die *Grundfrequenzen*, die δ_k die *Phasenkonstanten* der Bewegung. Die Gesamtheit der Frequenzen des Systems wird gegeben durch

$$(3) \quad \sum_k \tau_k \nu_k = (\tau \nu),$$

wo τ_1, τ_2, \dots ganze Zahlen sind.

Die Amplituden der monochromatischen Kugelwellen, die von dem Atom emittiert werden, sind in erster Annäherung proportional den Fourierkoeffizienten des *Moments des schwingenden elektrischen Dipols* des Atoms. Dieses läßt sich darstellen durch die Fourierreihe

$$(4) \quad \mathfrak{P} = \sum_{\tau} \mathfrak{P}_{\tau} e^{2\pi i(\tau w)},$$

wo das System der ganzen Zahlen τ_1, τ_2, \dots kurz mit τ bezeichnet, also $\mathfrak{P}_{\tau} = \mathfrak{P}_{\tau_1, \tau_2, \dots}$ und

$$(\tau w) = \sum_k \tau_k w_k$$

gesetzt ist.

Damit \mathfrak{P} *reell* ist, muß

$$(5) \quad \mathfrak{P}_{-\tau} = \mathfrak{P}_{\tau}^*$$

angenommen werden¹.

¹ Wir bezeichnen hier, wie überall in diesem Buche, die zu einer komplexen Zahl $z = x + iy$ konjugierte mit $z^* = x - iy$. In der Mathematik ist statt dessen das Zeichen \bar{z} gebräuchlich; da aber in der Physik das Überstreichen zur Bezeichnung der zeitlichen oder statistischen Mittelbildung unentbehrlich ist und der Stern sonst nicht gebraucht wird, haben wir uns nach reiflicher Überlegung zu dieser Festsetzung entschlossen.

Setzt man in (4) den Ausdruck (1) für w_k ein, so erhält man den zeitlichen Ablauf des elektrischen Moments:

$$(6) \quad \mathfrak{P} = \sum_{\tau} \mathfrak{P}_{\tau} e^{2\pi i[(\tau\nu)t + (\tau\delta)]}.$$

Die Intensität der monochromatischen Kugelwelle ist $|\mathfrak{P}_{\tau}|^2$ proportional, d. h. dem skalaren Produkt des Vektors \mathfrak{P}_{τ} mit dem konjugiert komplexen \mathfrak{P}_{τ}^* :

$$(7) \quad |\mathfrak{P}_{\tau}|^2 = |P_{xz}|^2 + |P_{yz}|^2 + |P_{z\tau}|^2.$$

Der genaue Wert der Ausstrahlung innerhalb einer Zeit Δt , in der viele Schwingungen der Frequenz $(\nu\tau)$ vor sich gehen, aber nur eine relativ geringe Änderung der Atomenergie eintritt, läßt sich auf Grund der klassischen Elektrodynamik berechnen, indem man die bekannte Formel von HERTZ für die Emission eines Dipols anwendet¹; man erhält für die Energie, die von der Kugelwelle der Frequenz $(\tau\nu)$ während der Zeit Δt in den Raum fortgetragen wird:

$$(8) \quad \Delta E_{(\tau\nu)}^{(e)} = \frac{64\pi^4}{3c^3} (\tau\nu)^4 |\mathfrak{P}_{\tau}|^2 \cdot \Delta t.$$

Dieselben Größen \mathfrak{P}_{τ} bestimmen aber auch die *Absorption* des Atoms, und zwar ergibt die klassische Theorie folgendes (s. Anhang I):

In einem Strahlungsfeld der monochromatischen Raumdichte $\varrho(\omega)$ ist die während der Zeit Δt absorbierte Energie:

$$(8a) \quad \Delta E_{(\tau\nu)}^{(a)} = \frac{8\pi^3}{3} \sum_k \tau_k \frac{\partial}{\partial J_k} [(\tau\nu) \cdot |\mathfrak{P}_{\tau}|^2 \cdot \varrho(\tau\nu)] \cdot \Delta t.$$

Diese Formeln (8) und (8a) sind Folgen der MAXWELLSCHEN Gleichungen des Feldes. Aber die vorher genannten Hauptpunkte, nämlich die Formel (3) für das Spektrum des Atoms, die *stetige* Energieabgabe und die *gleichzeitige* Emission *aller* Spektrallinien, sind von viel allgemeinerer Art; sie beruhen unmittelbar auf den Grundvorstellungen der klassischen Physik, verbunden mit der — durch die Erfahrung erzwungenen — Annahme, daß die Atome aufgebaut sind aus Elementarteilchen, deren Ausdehnung klein ist gegen ihre mittleren Abstände.

¹ Da diese Formel und die entsprechende für Absorption im 1. Bande nicht enthalten sind, werden sie hier im Anhang I abgeleitet.

Nun widerlegt aber die Erfahrung gerade diese allgemeinen Aussagen der klassischen Theorie. Und zwar ersetzt sie diese Aussagen durch folgende Feststellungen:

Statt der Formel (3) gilt empirisch das *Kombinationsgesetz*, nach welchem die Frequenzen aller Spektrallinien eines Atoms darzustellen sind als Differenzen von gewissen, dem Atom eigentümlichen *Termen* T_n :

$$(9) \quad \nu_{mn} = T_n - T_m.$$

Statt der gleichzeitigen Emission aller Spektrallinien vermag man experimentell in mannigfachster Weise die Emission nur eines Teils des Spektrums zu erzeugen (in gewissen Fällen bis herab zu einer einzigen Spektrallinie).

Und was endlich die stetige Energieausstrahlung betrifft, so ist auch diese Vorstellung mit elementaren Erfahrungen im Widerspruch, wie zuerst PLANCK durch theoretische Erwägungen bei der Untersuchung der Gesetze der schwarzen Strahlung erschlossen hat. Die von ihm postulierten quantenhaften Energieänderungen stehen in schroffem Gegensatz zu den Grundlagen der klassischen Theorie. Aber ihre empirische Existenz wurde durch EINSTEINS Deutung des lichtelektrischen Effekts und eine große Reihe weiterer experimenteller Entdeckungen über allen Zweifel erhoben.

In diese zunächst unverständlichen Ergebnisse hat BOHR im Anschluß an die PLANCK-EINSTEINSchen Gedanken in folgender Weise eine logische Ordnung bringen können.

Man muß sich vorstellen, daß ein isoliertes Atom im Gegensatz zum klassischen Verhalten nur in gewissen diskreten Zuständen auftreten kann. Jeder solche stationäre Zustand n besitzt eine bestimmte Energie W_n , und die optischen Spektraltermine T_n sind diesen Energiestufen proportional; es gilt

$$(10) \quad W_n = -h T_n,$$

wo h das PLANCKsche Wirkungsquantum ist. Dann ist (9) gleichbedeutend mit

$$(11) \quad \nu_{nm} = \frac{1}{h} (W_n - W_m).$$

Die Emission von Strahlung ist in Verbindung zu bringen mit endlichen, un stetigen Zustandsänderungen des Atoms, wobei die

Frequenz der ausgesandten Strahlung mit Hilfe von (11) aus der Energieabnahme des Atoms zu berechnen ist. Entsprechendes gilt für die Absorption von Strahlung durch Atome. Ebensolche sprungartigen Zustandsänderungen der Atome sind auch anzunehmen bei energieändernden Prozessen anderer Art, bei denen keine Strahlung wirksam ist, wie bei Zusammenstößen von Atomen mit Elektronen oder von Atomen untereinander.

Diese Annahmen genügen bereits, um ein fast unübersehbar großes Gebiet experimenteller Erfahrung über die Anregungsbedingungen spektraler Emissionen weitgehend verständlich zu machen, während die klassische Theorie vollkommen versagte¹.

Was nun die theoretische Berechnung der Zustandsenergien W_n anbetraf, so hat man sich, wie im 1. Bande gezeigt wurde, zunächst in der Weise geholfen, daß man auf die inneratomaren Bewegungen die klassische Mechanik anwandte und dabei die stetige Ausstrahlung fortließ. Besondere Überlegungen erfordert dann die Auswahl der „stationären“ Bewegungen aus der kontinuierlichen Mannigfaltigkeit der klassisch möglichen Bahnen. Sucht man nach mechanischen Größen, die sich ganzen Zahlen proportional setzen lassen, so bieten sich sogleich die Wirkungsvariablen J_k dar, weil diese bis auf lineare, ganzzahlige Transformationen eindeutig definiert sind. Sie haben überdies die Eigenschaft, dem EHRENFESTSchen Adiabatenpostulat zu genügen, d. h. sie bleiben bei hinreichend langsamer Beeinflussung des Systems durch äußere Kräfte mit beliebiger Annäherung konstant. Auf Grund dieser Eigenschaften der J_k ist es möglich, Quantenbedingungen der Form

$$(12) \quad J_k = h n_k$$

vorzuschreiben, wo n_k ganze Zahlen, die Quantenzahlen sind. Dadurch wird aus dem Kontinuum der klassischen Bewegungen eine diskrete Folge von „stationären“ ausgesondert mit den Energiewerten

$$(13) \quad W(J_1, J_2, \dots) = W(h n_1, h n_2, \dots) = W_{n_1, n_2, \dots} = W_n.$$

¹ Eine Übersicht über diese Erfahrungen und die Grundlagen ihrer Deutung findet man in dem Buche von J. FRANCK und P. JORDAN: Anregung von Quantensprüngen durch Stöße (diese Sammlung Nr. III), und in dem Buche von P. PRINGSHEIM: Fluoreszenz und Phosphoreszenz (diese Sammlung Nr. VI).

Nunmehr erkennt man, daß die BOHRsche Frequenzbedingung (11) eine rationelle Verallgemeinerung der klassischen Frequenzbedingung (3) ist. Schreibt man nämlich (11) in der Form

$$\begin{aligned} \nu_{n, n-\tau} &= \frac{1}{h} (W_n - W_{n-\tau}) \\ &= \frac{1}{h} \{ W(J_1, J_2, \dots) - W(J_1 - \tau_1 h_1, J_2 - \tau_2 h_2, \dots) \}, \end{aligned}$$

so erhält man im Limes $h \rightarrow 0$

$$(14) \quad \nu = \sum_k \tau_k \frac{\partial W}{\partial J_k},$$

was auf Grund von (2) mit (3) gleichbedeutend ist.

Diese Zuordnung zwischen den wirklichen Quantenfrequenzen des Atoms und den behelfsmäßig aus dem klassischen Modell berechneten Frequenzen ($\tau\nu$) ermöglichte BOHR auch eine Zuordnung (das sog. *Korrespondenzprinzip*) der klassischen Fourierkoeffizienten \mathfrak{F}_τ zu den wirklichen Quantenfrequenzen (allerdings nur mit einer gewissen Mehrdeutigkeit): Man wird dem Quantenübergang vom Zustande n_1, n_2, \dots nach dem Zustande $n_1 - \tau_1, n_2 - \tau_2, \dots$ den klassischen Fourierkoeffizienten

$$(15) \quad \mathfrak{F}_\tau(J) = \mathfrak{F}_{\tau_1, \tau_2, \dots}(J_1, J_2, \dots)$$

zuordnen, wobei man die J_k entweder dem Anfangszustande

$$(16a) \quad J_k = h n_k$$

oder dem Endzustande

$$(16b) \quad J_k = h (n_k - \tau_k)$$

anpassen oder schließlich einen Zwischenwert wählen kann.

Diese Zuordnung kann man nun nach BOHR benützen, um die Aussagen der klassischen Theorie bezüglich der *Strahlungsintensitäten* auf die wirkliche, in Quantensprüngen erfolgende Ausstrahlung des Atoms anzuwenden, und zwar natürlich in folgendem Sinne: Wenn viele Atome im Zustande n sind und ein Zustand $n - \tau$ kleinerer Energie existiert, so wird man eine Strahlungsenergie im Mittel pro Atom der Frequenz $\nu_{n, n-\tau}$ definieren können für eine Zeit Δt , die groß ist gegen eine Schwingungsperiode, aber klein gegen die mittlere Lebensdauer der Atome im Zustande n . Diese mittlere Strahlungsintensität soll

nach BOHR angenähert der entsprechenden klassischen gleich sein. Freilich ist im allgemeinen die Genauigkeit einer solchen Intensitätsberechnung von vornherein durch die Unbestimmtheit der Zuordnung (16a) oder (16b) beschränkt. Trotzdem hat es sich gezeigt, daß die mit dieser Methode erzielten Resultate, insbesondere bezüglich der sog. *Auswahlregeln* bei Rotations- und Schwingungsfrequenzen (sowie z. B. beim Starkeffekt des Wasserstoffatoms) mit den Beobachtungen bemerkenswert gut übereinstimmen.

§ 2. Vorläufer der Quantenmechanik. Die Erfolge des BOHRschen Korrespondenzprinzips bei der Abschätzung von Intensitäten spektraler Feinstrukturen und bei andern, mehr qualitativen Betrachtungen legten den Gedanken nahe, daß hinter diesem halbklassischen Formalismus eine nicht sehr davon verschiedene, allgemeine Quantenmechanik verborgen sein müsse, in der die Begriffe der Frequenz und Amplitude von Schwingungen ebenfalls sinnvoll sind. So wird man erwarten, daß es in der wahren Quantenmechanik *exakt definierte Dipolamplituden*

$$(1) \quad \mathfrak{P}(n, n - \tau) = \mathfrak{P}(n_1, n_2, \dots; n_1 - \tau_1, n_2 - \tau_2, \dots)$$

geben wird, aus denen sich die Emission und Absorption von Strahlung in ähnlicher Weise berechnen lassen wie in der klassischen Theorie; und zwar insbesondere die emittierte Energie entsprechend der klassischen Formel § 1 (8) gemäß:

$$(2) \quad \Delta E_{n, n-\tau}^{(e)} = \frac{64 \pi^4}{3 c^3} \nu_{n, n-\tau}^4 |\mathfrak{P}(n, n - \tau)|^2 \Delta t.$$

Das quantentheoretische Analogon der Absorptionsformel § 1 (8a) liegt nicht so auf der Hand; wir kommen unten darauf zurück.

Durch die empirisch meßbaren Strahlungsintensitäten (in verdünnten Gasen) sind zwar nach (2) die $\mathfrak{P}(n, n - \tau)$ als komplexe Vektoren noch nicht bestimmt, sondern nur ihre absoluten Beträge. Aber nach HEISENBERG und KRAMERS¹ gewinnt man aus einer korrespondenzmäßigen Untersuchung der Dispersion polarisierten Lichts durch Atome die Überzeugung, daß auch gewisse Phasenbeziehungen zwischen den zu verschiedenen Übergängen gehörigen Lichtemissionen empirisch beobachtbar sind; das bedeutet aber, daß die quantentheoretischen Amplitudenvektoren $\mathfrak{P}(n, n - \tau)$ selbst beobachtbare Größen sind, und zwar

¹ HEISENBERG, W. u. H. A. KRAMERS: Z. Phys. Bd. 31, S. 681. 1925.

in ähnlichem Maße empirisch bestimmbar wie die klassischen Amplitudenvektoren \mathfrak{F}_τ . Wie ein Blick auf die Formel § 1 (6) zeigt, kann man wegen der Willkür der Phasenkonstanten δ die klassischen Amplituden \mathfrak{F}_τ ersetzen durch

$$\mathfrak{F}'_\tau = \mathfrak{F}_\tau e^{2\pi i(\tau\alpha)}, \quad (\tau\alpha) = \sum_k \tau_k \alpha_k,$$

wobei die α_k reelle Phasenkonstanten sind. Ein entsprechendes Maß von Willkür wird bei den Quantenamplituden erlaubt sein: man wird $\mathfrak{F}(n, n - \tau)$ ersetzen dürfen durch

$$(3) \quad \mathfrak{F}'(n, n - \tau) = \mathfrak{F}(n, n - \tau) e^{2\pi i(\alpha_n - \alpha_{n-\tau})}.$$

Dabei ist also jedem Zustande n eine willkürliche, reelle Phasenkonstante α_n zugeordnet.

Die Einführung der Quantenamplituden ermöglicht es, bis zu einem gewissen Grade zu erraten, wie die exakten quantentheoretischen Analoga zu bekannten klassischen Formeln lauten werden. Dies ist zuerst von KRAMERS und HEISENBERG¹ durch Aufstellung der allgemeinen quantentheoretischen Dispersionsformel gezeigt worden, wobei sie einen älteren Ansatz von LADENBURG² in die korrespondenzmäßig richtige Form brachten und wesentlich verallgemeinerten. Diese Dispersionsformel wurde der Ausgangspunkt verschiedener ähnlicher Korrespondenzbetrachtungen³, die wichtige Unterlagen für die Begründung der Quantenmechanik geliefert haben.

Als einfaches Beispiel hierfür betrachten wir die Absorptionsformel § 1 (8a); doch kann man nicht den monochromatischen Anteil direkt analogisieren, sondern muß zunächst die Gesamtabsorption

$$(4) \quad \Delta E^{(a)} = \frac{8\pi^3}{3} \sum_{(\nu\tau) > 0} \left(\tau, \frac{\partial}{\partial J} [(\nu\tau) \cdot |\mathfrak{F}_\tau|^2 \cdot \varrho(\nu\tau)] \right) \cdot \Delta t$$

¹ KRAMERS, H. A.: Nature Bd. 113, S. 673. 1924; Bd. 114, S. 310. 1924. KRAMERS, H. A. u. W. HEISENBERG: Z. Phys. Bd. 31, S. 681. 1925.

² LADENBURG, R.: Z. Phys. Bd. 4, S. 451. 1921. LADENBURG, R. u. F. REICHE: Naturwissensch. Bd. 11, S. 584. 1923.

³ BORN, M.: Z. Phys. Bd. 26, S. 379. 1924. VLECK, J. v.: J. Opt. Soc. Bd. 9, S. 27. 1924; Phys. Rev. Bd. 24, S. 330. 1924. BORN, M. u. P. JORDAN: Z. Phys. Bd. 33, S. 479. 1925. JORDAN, P.: Z. Phys. Bd. 33, S. 506. 1925. PAULI, W. jr.: Det Kgl. Danske Vid. Selskab Bd. 7, S. 3. 1925. THOMAS, W.: Naturwissensch. Bd. 13, S. 627. 1925. KUHN, W.: Z. Phys. Bd. 33, S. 408. 1925. REICHE, F. u. W. THOMAS: Z. Phys. Bd. 34, S. 510. 1925.

bilden. Hier ist das Operatorsymbol

$$(5) \quad \left(\tau, \frac{\partial}{\partial J} \right) = \sum_k \tau_k \frac{\partial}{\partial J_k}$$

verwandt. Zu (4) kann man das Quantenanalogon bilden und dieses dann in monochromatische Bestandteile zerlegen¹.

Die Differentiationsoperation $\left(\tau, \frac{\partial}{\partial J} \right)$ wird durch eine Differenzbildung zu ersetzen sein, ähnlich wie die klassische Frequenz nach § 1 (14) durch die Quantenfrequenz § 1 (11) zu ersetzen ist. So erhält man aus (4) für die Gesamtabsorption

$$\Delta E^{(a)} = \frac{8\pi^3}{3h} \sum_{\tau > 0} \{ \nu_{n+\tau, n} \cdot |\mathfrak{P}(n+\tau, n)|^2 \varrho(\nu_{n+\tau, n}) - \nu_{n, n-\tau} \cdot |\mathfrak{P}(n, n-\tau)|^2 \varrho(\nu_{n, n-\tau}) \} \Delta t.$$

Ersetzt man hier in der zweiten Summe τ durch $-\tau$ und beachtet, daß $\nu_{n+\tau, n} = -\nu_{n, n+\tau}$ ist, so kann man beide Summen in eine zusammenziehen, in der die Summationsindizes τ positive und negative Werte durchlaufen:

$$(6) \quad \Delta E^{(a)} = \frac{8\pi^3}{3h} \sum_{\tau} \nu_{n+\tau, n} \cdot |\mathfrak{P}(n+\tau, n)|^2 \cdot \varrho(\nu_{n+\tau, n}) \cdot \Delta t.$$

Nunmehr kann man wieder in monochromatische Bestandteile zerlegen; wir schreiben

$$(7) \quad \Delta E^{(a)} = \sum_{\tau} \Delta E_{n, n+\tau}^{(a)},$$

wo

$$(8) \quad \Delta E_{n, n+\tau}^{(a)} = \frac{8\pi^3}{3h} \nu_{n+\tau, n} \cdot |\mathfrak{P}(n+\tau, n)|^2 \cdot \varrho(\nu_{n+\tau, n}) \cdot \Delta t$$

die Absorption beim Sprunge $n \rightarrow n + \tau$ darstellt. Sie ist positiv für $\tau > 0$, weil dann $\nu_{n+\tau, n} > 0$ ist; das ist die eigentliche Absorption. Dagegen liefern Sprünge zu tieferen Energiestufen, $\tau < 0$, eine negative Absorption (erzwungene Emission), in Übereinstimmung mit dem Ansatz EINSTEINS zur Ableitung der PLANCKschen Strahlungsformel (Bd. I, § 2, S. 11, 12). Die EINSTEINSchen Übergangskoeffizienten sind durch (2) und (8) vollständig auf

¹ Vgl. die soeben zit. Arbeiten von J. VAN VLECK, BORN und JORDAN.

Atomeigenschaften zurückgeführt. Die in Bd. 1, § 2 mit A_{12} bezeichnete spontane Übergangswahrscheinlichkeit erhält man aus (2) durch Division mit der emittierten Energie $h\nu_{n, n-\tau}$

$$(9) \quad A_{n, n-\tau} = \frac{64 \pi^4}{3 \hbar c^3} \nu_{n, n-\tau}^3 |\mathfrak{F}(n, n-\tau)|^2.$$

Die Wahrscheinlichkeit der erzwungenen Übergänge (für $\varrho = 1$), dort mit $B_{12} = B_{21}$ bezeichnet, erhält man ebenso aus (8) durch Division mit der bei dem Sprunge verbrauchten Energie $h\nu_{n, n-\tau}$:

$$(10) \quad B_{n, n-\tau} = B_{n-\tau, n} = \frac{8 \pi^3}{3 \hbar^2} \cdot |\mathfrak{F}(n, n-\tau)|^2.$$

Durch Division beider Ausdrücke folgt

$$(11) \quad \frac{A_{n, n-\tau}}{B_{n, n-\tau}} = \frac{8 \pi \hbar}{c^3} \nu_{n, n-\tau}^3$$

in Übereinstimmung mit der in Bd. 1, § 2, S. 13 angegebenen Formel, die dort rückwärts aus der PLANCKSchen Strahlungsgleichung erschlossen wurde.

Als zweites Beispiel für das Erraten der richtigen Quantengesetze durch Analogisieren der klassischen Formeln betrachten wir den Fall, wo ein Atom in ein konstantes elektrisches Feld \mathfrak{E} gebracht wird¹; dann ist die Störungsfunktion, ausgedrückt durch die Winkel- und Wirkungsvariablen

$$(12) \quad H^{(1)} = -\mathfrak{F}\mathfrak{E} = \sum_{\tau} H_{\tau}^{(1)} e^{2\pi i(\tau w)},$$

mit

$$(12a) \quad H_{\tau}^{(1)} = -\mathfrak{F}_{\tau}\mathfrak{E}.$$

Die Energiekonstante W läßt sich, wie in Bd. 1, § 41, gezeigt wurde, durch sukzessive Näherungen berechnen:

$$W = W^0 + W^{(1)} + W^{(2)} + \dots,$$

und zwar ergibt sich² nach Bd. 1, § 41, (10) und (24)

$$(13) \quad W^{(1)} = \overline{H^{(1)}} = H_0^{(1)}(J) = H_0^{(1)}(h\nu),$$

$$(13a) \quad W^{(2)} = - \sum_{(\tau\nu) > 0} \sum_k \tau_k \frac{\partial}{\partial J_k} \left(\frac{|H_{\tau}^{(1)}|^2}{(\tau\nu)} \right).$$

¹ BORN, M.: Z. Phys. Bd. 26, S. 379. 1924.

² Für irgend eine Zeitfunktion $f(t)$ soll $\overline{f(t)}$ den zeitlichen Mittelwert bedeuten.

Man errät nun leicht, wie diejenigen Quantenformeln lauten werden, die für $\hbar \rightarrow 0$ in diese Ausdrücke übergehen. Man hat wieder die $H_\tau^{(1)}$ durch Einführung von Quantenamplituden zu analogisieren:

$$(14) \quad H^{(1)}(n, n - \tau) = -\mathfrak{P}(n, n - \tau) \mathfrak{E}.$$

Dann hat man offenbar zunächst an Stelle von (13):

$$(15) \quad W^{(1)} = H^{(1)}(n, n).$$

Ferner wird der Differentiation in (13a) wie oben eine Differenzbildung entsprechen; so erhält man statt (13a)

$$W^{(2)} = -\frac{1}{\hbar} \sum_{\tau_k > 0} \left(\frac{|H^{(1)}(n + \tau, n)|^2}{\nu_{n+\tau, n}} - \frac{|H^{(1)}(n, n - \tau)|^2}{\nu_{n, n-\tau}} \right),$$

oder, ohne Beschränkung der Summation auf positive Werte von τ_k , mit Rücksicht auf $\nu_{n+\tau, n} = -\nu_{n, n+\tau}$:

$$(16) \quad W^{(2)} = \sum_{\tau} \frac{|H^{(1)}(n, n - \tau)|^2}{\hbar \nu_{n, n-\tau}}.$$

Hierhin gehörte auch als weiteres Beispiel der sogenannte „Summensatz“ von THOMAS und KUHN¹, der gewisse Beziehungen zwischen den sämtlichen Übergangswahrscheinlichkeiten eines Atoms ausspricht. Wir wollen jedoch hier darauf nicht eingehen, da wir später in der systematischen Quantenmechanik darauf zurückzukommen haben.

Auch eine große Reihe *speziellerer* Formeln über die Intensitäts- und Polarisationsverhältnisse von Feinstrukturen der Spektrallinien, sowohl für natürliche Multipletts als auch für Aufspaltungen im Magnetfeld (ZEEMAN-Effekt), ferner für Bandenstrukturen von Molekeln, sind durch verfeinerte Korrespondenzbetrachtungen im Anschluß an empirische, gleichfalls korrespondenzmäßig verständliche „Summenregeln“ gefunden worden. Diese Ableitungen beruhen auf der Betrachtung von Modellen, an denen verhältnismäßig einfache, aus überlagerten Rotationen und Präzessionen bestehende Bewegungen auftreten. In der klassischen Theorie bedingt der kinematische Charakter dieser Bewegungsformen einfache mathematische Beziehungen zwischen den

¹ THOMAS, W.: Naturwissensch. Bd. 13, S. 627. 1925. KUHN, W.: Z. Phys. Bd. 33, S. 408. 1925.

Fourierkoeffizienten der rechtwinkligen Koordinaten des bewegten Elektrons. Die Intensitätsformeln ergaben sich durch eine sinn-gemäße, die empirischen Kenntnisse berücksichtigende Über-
setzung dieser klassischen Beziehungen in korrespondierende
quantenmechanische Formeln¹. Als ein allgemeineres Ergebnis
dieser Untersuchungen ist die von HEISENBERG als *spektroskopische Stabilität* bezeichnete Gesetzmäßigkeit hervorzuheben, auf
die wir noch zurückkommen werden².

Alle diese Resultate sind nachher durch die vollständige
Quantenmechanik bestätigt worden. Es sind Vorläufer dieser
Theorie, aus denen sie in natürlicher Weise erwachsen ist.

§ 3. HEISENBERGS korrespondenzmäßige Kinematik. Es handelt
sich nun darum, das durch das Korrespondenzprinzip geleitete
und kontrollierte Raten der richtigen Quantenformeln durch ein
exaktes Verfahren zu ersetzen. Den entscheidenden Schritt hierzu
tat HEISENBERG³, indem er sich von dem zu engen Bilde der
klassischen Bewegungen im Atom befreite und eine neue, natür-
liche Deutung der atomphysikalischen Tatsachen anbahnte.

Unsere Kenntnis vom Bau der Atome beruht in überwiegendem
Maße auf der von ihnen emittierten Strahlung. Würde die klas-
sische Theorie bestehen, so müßten wir erwarten, daß die Fre-
quenzen des ausgestrahlten Lichts mit den Frequenzen ($\tau\nu$) der
atomaren Bewegungen übereinstimmen. Nun lassen sich aber die
empirischen Spektrallinien nicht in der Form ($\tau\nu$) darstellen,
sondern folgen der Kombinationsregel § 1, (9).

Man kann nun entweder die Identität der Lichtfrequenzen
mit den Bewegungsfrequenzen aufgeben — das hat die Theorie
von BOHR getan —; oder man muß auf die klassische Beschrei-

¹ Von den Arbeiten, die zu dieser Richtung gehören, sind folgende die
wichtigsten: HÖNL, H.: Z. Phys. Bd. 31, S. 340. 1925. GOUDSMIT, J. u.
R. DE L. KRONIG: Naturwissensch. Bd. 13, S. 90. 1925. KRONIG, R. DE L.:
Z. Phys. Bd. 31, S. 885. 1925; Bd. 33, S. 261. 1925. SOMMERFELD, A. u.
H. HÖNL: Berl. akad. Ber. 1925, IX, S. 141. RUSSELL, H. N.: Proc. nat.
Acad. Sci. U. S. A. Bd. 11, S. 314 u. 322. 1925. HEISENBERG, W.: Z. Phys.
Bd. 31, S. 617. 1925. FOWLER, R. H.: Phil. Mag. Bd. 49, S. 1272; Bd. 50,
S. 1079. 1925. HÖNL, H. u. F. LONDON: Z. Phys. Bd. 33, S. 803. 1925.
HÖNL, H.: Ann. Phys. Bd. 79, S. 273. 1926.

² HEISENBERG, W.: Z. Phys. Bd. 31, S. 617. 1925.

³ HEISENBERG, W.: Z. Phys. Bd. 33, S. 879. 1925.

bung der inneratomaren Bewegungen durch Fouriersche Reihen verzichten — das ist der von HEISENBERG eingeschlagene Weg. Zur Begründung der Ansicht, daß die Frequenzen des emittierten Lichts tatsächlich den Frequenzen der Prozesse im Atom entsprechen, hat HEISENBERG das folgende heuristische Prinzip herangezogen: Eine physikalische Theorie soll Beziehungen zwischen beobachtbaren Größen darstellen; daher soll sie nach Möglichkeit nur solche Begriffe einführen, die eine unmittelbare Deutung im Experiment zulassen. Eine solche Regel hat natürlich Sinn nur im Hinblick auf eine unfertige, erst zu schaffende Theorie; denn sobald ein theoretisches Begriffssystem abgeschlossen vorliegt, besteht kein Unterschied mehr zwischen beobachtbaren und nicht beobachtbaren Größen. Jede Messung wird ja erst durch die theoretische Interpretation sinnvoll. Wenn man aber das theoretische Begriffssystem, das einem Erfahrungskomplex entspricht, noch nicht kennt, besteht ein Unterschied zwischen den Begriffen: Es gibt solche, die notwendig erscheinen, um überhaupt das Resultat von Experimenten aussprechen zu können, und andere, die nur bei der theoretischen Diskussion der Zusammenhänge eingeführt werden. Der ersten Klasse von Begriffen entsprechen Größen, die man „direkt beobachtbar“ nennen wird; diese sucht man solange wie möglich zu erhalten. Die übrigen Begriffe wird man eher zu opfern bereit sein.

Nun gehören offenbar die von einem Atom ausgesandten Lichtfrequenzen zu den in diesem Sinne direkt beobachtbaren Größen; ohne diesen Begriff läßt sich ja überhaupt die Messung mit einem Spektralapparat nicht beschreiben. Dagegen sind die Umlauffrequenzen der Elektronen in ihren Bahnen keineswegs von dieser Art; sie haben vielmehr hypothetischen Charakter und waren nur eingeführt, um das Verhalten des Atoms in Anlehnung an bekannte Gesetze, nämlich die der klassischen Mechanik, zu berechnen. Aber die Gültigkeit dieser Gesetze ist mehr als zweifelhaft. Man wird daher dazu gedrängt, die Existenz der Umlauffrequenzen ($\tau\nu$) aufzugeben und anzunehmen, daß die Vorgänge im Atom keine andern Frequenzen enthalten als die im emittierten Licht erscheinenden $\nu_{n, n-\tau}$. Hierfür ist es aber nötig, schon den elementaren Begriff der *Bewegung* abzuändern; denn wenn Bewegungen durch Angabe der Koordinaten $q(t)$ als Zeitfunktionen beschrieben werden können, so führt die FOURIER-

zerlegung

$$(1) \quad q(t) = \sum_{\tau} q_{\tau}(J) e^{2\pi i(\tau\nu)t}$$

immer auf die Frequenzen $(\tau\nu)$. Man muß also zunächst eine neue *Kinematik* aufbauen. Hierzu dient der im vorigen Paragraph eingeführte Begriff der quantentheoretischen Amplitude.

Man kann ja auch in der klassischen Theorie eine (mehrfach periodische) Größe q durch die Angabe der Amplituden ihrer harmonischen Bestandteile, d. h. der in (1) auftretenden FOURIER-Koeffizienten

$$(2) \quad q_{\tau}(J), \quad J_k = n_k \hbar$$

vollständig charakterisieren. Bildet man die Summe und das Produkt zweier FOURIER-Reihen der Form (1)

$$(3) \quad \begin{cases} q = q^{(1)} + q^{(2)}, \\ Q = q^{(1)} \cdot q^{(2)}, \end{cases}$$

so sind diese wieder FOURIER-Reihen derselben Form (1) mit denselben Frequenzen $(\tau\nu)$, und ihre Koeffizienten lassen sich aus denen von $q^{(1)}$ und $q^{(2)}$ berechnen:

$$(4) \quad \begin{cases} q_{\tau}(J) = q_{\tau}^{(1)}(J) + q_{\tau}^{(2)}(J), \\ Q_{\tau}(J) = \sum_{\sigma} q_{\sigma}^{(1)}(J) q_{\tau-\sigma}^{(2)}(J). \end{cases}$$

Demnach läßt sich auch jedes Polynom und mit Hilfe von Grenzübergängen (unendlichen Reihen) jede Funktion von Größen, die zu denselben Grundfrequenzen gehören, durch FOURIER-Koeffizienten beschreiben, die sich aus den FOURIER-Koeffizienten der ursprünglichen Größen herstellen lassen.

Dieses Verfahren läßt sich nun nach HEISENBERG auf die quantentheoretischen Größen übertragen. Zwar existiert hier zur Beschreibung einer Größe q keine, einem Zustande zugeordnete Zeitfunktion, wohl aber ein System bestimmter, je zwei Zuständen zugeordneter Amplituden

$$(5) \quad q_{n, n-\tau}.$$

Man kann nun Addition und Multiplikation für solche Systeme nach Regeln definieren, die korrespondenzmäßige Verallgemeinerungen von (4) sind: Man nennt symbolisch eine Größe q die

Summe

$$(6a) \quad q = q^{(1)} + q^{(2)},$$

bzw. Q das Produkt

$$(6b) \quad Q = q^{(1)} q^{(2)}$$

zweier Größen $q^{(1)}$ und $q^{(2)}$, wenn

$$(7) \quad \begin{cases} q_{n, n-\tau} = q_{n, n-\tau}^{(1)} + q_{n, n-\tau}^{(2)}, \\ Q_{n, n-\tau} = \sum_{\sigma} q_{n, n-\sigma}^{(1)} q_{n-\sigma, n-\tau}^{(2)} \end{cases}$$

ist. Dabei ist die Summe σ über alle möglichen Zustände des Atoms zu erstrecken.

Um aus (7) durch Grenzübergang $\hbar \rightarrow 0$ die klassischen Formeln (4) zu erhalten, muß man annehmen, daß im Limes $\hbar \rightarrow 0$ die Größe

$$q_{n, n-\tau} = q(J, J - \tau \hbar)$$

nur wenig von dem Absolutwert der beiden Argumente $J, J - \tau \hbar$, wohl aber merklich von ihrer Differenz $\tau \hbar$ abhängt; man kann dann für kleine \hbar die Schreibweise $q_{\tau}(J)$ wählen und

$$q_{n-\sigma, n-\tau}^{(2)} = q^{(2)}(J - \sigma \hbar, J - \tau \hbar)$$

durch

$$q_{\tau-\sigma}(J)$$

ersetzen. Dann gehen durch diesen Grenzprozeß die Formeln (7) direkt in die klassischen Formeln (4) über.

Die Gleichungen (7) werden übersichtlicher, wenn man $n - \tau = m, n - \sigma = k$ setzt:

$$(8) \quad \begin{cases} q_{nm} = q_{nm}^{(1)} + q_{nm}^{(2)}, \\ Q_{nm} = \sum_k q_{nk}^{(1)} q_{km}^{(2)}. \end{cases}$$

In dieser Form lassen sie erkennen, daß die HEISENBERGSche Multiplikation der quantenmechanischen Amplituden dasselbe ist wie die von den Mathematikern viel untersuchte *Multiplikation von Matrizen*. Man bezeichnet daher auch das Schema der Amplituden $q_{n, n-\tau} = q_{nm}$ als *Matrix* der Größe q . Numeriert man die Zustände des Atoms in irgend einer Reihenfolge mit

$n = 0, 1, 2, \dots$, so bilden die q_{nm} ein quadratisches Schema, die Matrix¹

$$(9) \quad q = (q_{nm}) = \begin{pmatrix} q_{00} & q_{01} & q_{02} & \dots & \dots & \dots \\ q_{10} & q_{11} & q_{12} & \dots & \dots & \dots \\ q_{20} & q_{21} & q_{22} & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots & \dots \end{pmatrix}.$$

Die *Diagonalelemente* q_{nn} entsprechen dabei verschwindenden Frequenzen, $\nu_{nn} = 0$; sie sind nicht den *Übergängen*, sondern den *Zuständen* zugeordnet.

Es ist nun sehr wichtig, daß die Addition und Multiplikation der Matrizen genau wie in der klassischen Theorie die Eigenschaft hat, keine neuen Frequenzen einzuführen. Lag es dort an der Ganzzahligkeit der τ -Werte im Ausdruck $(\tau\nu)$, so hier an der Gültigkeit der Kombinationsregel § 1, (9). Versieht man nämlich das Matrixelement q_{nm} mit dem zugehörigen Zeitfaktor $e^{2\pi i\nu_{nm}t}$, so sieht man aus den Gleichungen (8), daß derselbe Faktor bei den Elementen der Summe und des Produkts auftritt. Bei der Summe ist das trivial. Beim Produkt hat man rechter Hand in jedem Summenglied den Faktor

$$e^{2\pi i\nu_{nk}t} e^{2\pi i\nu_{km}t} = e^{2\pi i(\nu_{nk} + \nu_{km})t}$$

anzubringen; aus dem Kombinationsgesetz der Frequenzen folgt aber

$$(10) \quad \nu_{nk} + \nu_{km} = \nu_{nm}.$$

Diese Eigenschaft der Quantenfrequenzen ist also eine starke Stütze der Multiplikationsregel von HEISENBERG.

Man bemerkt ferner, daß die Definition (8) der Addition und Multiplikation invariant bleibt, wenn man für *alle* vorkommenden Größen die Matrixelemente q_{nm} gemäß § 2, (3) ersetzt durch

$$(11) \quad q'_{nm} = q_{nm} e^{2\pi i(\delta_n - \delta_m)}$$

mit beliebigen reellen Phasenkonstanten δ_n .

Die Matrizenmultiplikation unterscheidet sich von der Multiplikation gewöhnlicher Zahlen dadurch, daß es dabei auf die

¹ Wir benutzen neben q_{nm} auch die Schreibweise $q(n, m)$, je nachdem es für die Deutlichkeit der Formeln angemessen erscheint.

Reihenfolge der Faktoren ankommt: Die Produkte $q^{(1)}q^{(2)}$ und $q^{(2)}q^{(1)}$ mit den Elementen

$$(q^{(1)}q^{(2)})_{nm} = \sum_k q_{nk}^{(1)} q_{km}^{(2)}$$

$$(q^{(2)}q^{(1)})_{nm} = \sum_k q_{nk}^{(2)} q_{km}^{(1)}$$

sind im allgemeinen voneinander *verschieden*. Auf diesem Umstande beruht der Hauptunterschied zwischen den Formeln der neuen Theorie und denen der alten Mechanik.

Unter dem *zeitlichen Mittelwert* \bar{q} einer Matrix q versteht man die Matrix, deren Elemente die Zeitmittel der $q_{nm}e^{2\pi i\nu_{nm}t}$ sind; diese verschwinden, wenn $\nu_{nm} \neq 0$ ist. Nimmt man an, daß alle Terme T_n voneinander verschieden sind, so ist $\nu_{nm} = 0$ nur für $n = m$; dann bleiben bei der Mittelbildung nur die Diagonalelemente übrig, und man erhält für \bar{q} eine sogenannte *Diagonalmatrix*

$$(12) \quad \bar{q} = (q_{nm} \delta_{nm}) = \begin{pmatrix} q_{00} & 0 & 0 & \dots & \dots \\ 0 & q_{11} & 0 & \dots & \dots \\ 0 & 0 & q_{22} & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots \end{pmatrix}.$$

Hier ist das WEIERSTRASSsche Symbol

$$\delta_{nm} = \begin{cases} 1 & \text{für } n = m \\ 0 & \text{für } n \neq m \end{cases}$$

benutzt. Das *Diagonalelement* q_{nn} einer Matrix ist daher physikalisch zu deuten als *Zeitmittelwert der Größe q im Zustande n* . Das gilt für den Fall ungleicher Terme. Es ist aber bekannt, daß auch übereinstimmende Terme vorkommen können. Dieser Fall wurde in Bd. I, S. 101 als „Entartung“ bezeichnet. Genauer sprechen wir künftig in diesem Falle von einem mechanischen System mit „*entarteter Energie*“. Es gibt dann Nullfrequenzen, $\nu_{nm} = 0$ für $n \neq m$, und \bar{q} hat Elemente außerhalb der Diagonale. Die physikalische Bedeutung dieser Verhältnisse kann erst an späterer Stelle geklärt werden. Die *zeitliche Ableitung* \dot{q} einer Matrix q ist naturgemäß als die Matrix mit den Elementen

$$(13) \quad \dot{q}_{nm} = 2\pi i \nu_{nm} q_{nm}$$

zu definieren. Damit eine FOURIER-Reihe wie (1) eine reelle

Größe darstellt, müssen die Koeffizienten offenbar die Bedingung [s. auch § 1, (5)]

$$(14) \quad q_{-\tau}(J) = q_{\tau}(J)^*$$

erfüllen. Da hier τ der Differenz $n - m$ der Indizes von q_{nm} entspricht, so korrespondiert mit (14) die Forderung

$$(15) \quad q_{mn} = q_{nm}^*.$$

Solche Matrizen nennt man *hermitesche Matrizen*; wir werden ihre mathematischen Eigenschaften im nächsten Kapitel besonders untersuchen. Hier bemerken wir nur, daß, wenn q hermitisch ist, auch die in (13) definierte zeitliche Ableitung \dot{q} hermitisch wird.

Durch die vorstehenden Definitionen ist es möglich, mit Matrizen in nahezu derselben Weise zu rechnen wie mit Zahlen und eine Algebra und Analysis der Matrizen aufzubauen. Sodann kann man mit HEISENBERG versuchen, das *Analogon der Bewegungsgleichungen der klassischen Mechanik* zu formulieren. Man könnte ja auch in der klassischen Theorie die mechanischen Gleichungen als Beziehungen zwischen den FOURIER-Koeffizienten der zu bestimmenden Größen darstellen. Eine solche Formulierung erfüllt in vollkommener Weise die Forderung, daß die Theorie nur Beziehungen zwischen beobachtbaren Größen angeben soll: die Atommechanik wird in dieser Gestalt ein System von Beziehungen zwischen Energien, Frequenzen und Amplituden. HEISENBERG hat sich in seiner grundlegenden Arbeit auf den Standpunkt gestellt, daß diese Größen bereits alles darstellen, was man in der Atomphysik beobachten kann. Der heutige Standpunkt ist allerdings ein anderer: man kann — wie von HEISENBERG selber später klargestellt wurde — sehr wohl z. B. auch den Ort eines Elektrons im Atom beobachten.

Wie wir später sehen werden, besteht nur eine prinzipielle Beschränkung für die *gleichzeitige* Beobachtbarkeit verschiedener Größen (wie Zustandsenergie und Ort des Elektrons). Hier soll zunächst der ursprüngliche HEISENBERGSche Standpunkt festgehalten werden, daß zwar Energien, Frequenzen, Amplituden beobachtbar sind, Koordinaten, Geschwindigkeiten, Impulse von Teilchen innerhalb der Atome aber nicht. Diese Beschränkung unseres Stoffes hängt eng damit zusammen, daß der HEISENBERGSche Ansatz nur das quantenmechanische Analogon der *periodi-*

schen bzw. mehrfach periodischen Bewegungen der klassischen Theorie ist. Später werden wir allgemeinere Methoden kennen lernen, die auch die aperiodischen Bewegungen (z. B. die Trägheitsbewegung oder die „Hyperbelbahnen“ beim Wasserstoffatom) quantenmechanisch darzustellen erlauben. In diesem Teil des Buchs wollen wir uns bezüglich dieser Fragen mit einigen Hinweisen begnügen, die an folgende Bemerkung anknüpfen: Bekanntlich kann man eine gewisse Klasse nichtperiodischer Funktionen $f(t)$ durch ein *FOURIER-Integral*

$$f(t) = \int_{-\infty}^{\infty} g(\nu) e^{2\pi i \nu t} d\nu$$

(anstatt durch eine *FOURIERSche* Reihe) darstellen. Man wird dann mit *HEISENBERG* annehmen, daß in der Quantenmechanik analoge Größen dargestellt werden durch „Matrizen mit kontinuierlichem Index“, d. h. durch Funktionen $g(n, m)$ zweier kontinuierlich veränderlichen Variablen n, m , die an die Stelle der diskreten Indizes n, m der eigentlichen Matrizen (9) treten. Der *BOHRsche* Satz, daß eine eigentliche „Quantelung“ mit diskreten Energiewerten nur bei den (mehrfach) *periodischen* Bewegungen eintritt, erweist sich damit als unmittelbare Folgerung aus dem Matrizenansatz. Überhaupt zeigt sich erst von dem allgemeinsten Standpunkte der Theorie, der periodische und aperiodische Vorgänge in gleicher Weise umfaßt, wie natürlich die *HEISENBERG*-schen Begriffsbildungen dem Wesen der Quantenerscheinungen angepaßt sind.

Die in diesem Teil des Buchs entwickelte Theorie gibt also nur ein unvollständiges Bild der physikalischen Zusammenhänge, das später ausgebaut werden muß. Doch enthält dieses Bild so viele wesentliche Züge des ganzen Geschehens, daß es gerechtfertigt erscheint, es vollständig auszuführen. Hierzu sind einige mathematische Vorbereitungen notwendig, die wir hier einschalten, um später den physikalischen Gedankengang nicht unterbrechen zu müssen¹.

¹ Der Leser, der sich scheut, das folgende, ziemlich umfangreiche Kapitel rein mathematischen Inhalts durchzulesen, mag sich zunächst auf die ersten Abschnitte beschränken und die Lektüre des Restes je nach Bedarf nachholen. Jedoch wollen wir gleich hier versichern, daß in den mathematischen Abschnitten dieses Buches nur das zum Verständnis der Quantenmechanik wirklich Notwendige zu finden ist.

Zweites Kapitel.

Mathematische Grundlagen.

§ 4. Das Rechnen mit quadratischen Matrizen. Die folgenden Betrachtungen über das Rechnen mit Matrizen sind insofern formaler Natur, als wir Konvergenzfragen ganz unberücksichtigt lassen. Die Formeln gelten unter allen Umständen für endliche, quadratische Matrizen; bei unendlichen Matrizen muß vorausgesetzt werden, daß alle vorkommenden Summen (z. B. bei der Produktbildung) konvergieren.

Wir gehen davon aus, daß das Verschwinden einer Matrix bedeuten soll, daß alle Elemente Null sind:

$$(1) \quad A = 0, \quad A_{nm} = 0.$$

Sodann schreiben wir die (bereits in § 3 gegebene) Definition der Addition und Multiplikation nochmals an:

$$(2) \quad \left\{ \begin{array}{l} (A + B)_{nm} = A_{nm} + B_{nm}, \\ (A B)_{nm} = \sum_k A_{nk} B_{km}. \end{array} \right.$$

Zwei (oder mehr) quadratische *endliche* Matrizen können also dann und nur dann addiert oder multipliziert werden, wenn sie dieselbe Anzahl von Zeilen besitzen. Wir werden solche Matrizen kurz als *gleichartig* bezeichnen. Unendliche Matrizen sind immer gleichartig¹. Führt man in beiden Matrizen ein und dieselbe Permutation der Zeilen und Spalten aus, so bleiben die Gleichungen (2) ungeändert; bei der ersten ist das trivial, bei der zweiten folgt es daraus, daß man den Summationsindex k auch über die permutierte Reihenfolge der Indizes laufen lassen kann. Daraus geht hervor, daß jede, durch wiederholte Anwendung von Additionen und Multiplikationen entstehende Gleichung zwischen Matrizen gegen eine auf beide Indizes angewandte Permutation invariant ist. Die Reihenfolge der Diagonalglieder ist also ganz unwesentlich.

Man kann daher die Diagonale in beliebiger Weise ordnen, etwa als einfache, abzählbare Reihe, wie 1, 2, 3, ..., oder als zweiseitige Reihe wie ... - 2, - 1, 0, 1, 2, ... oder als mehrere

¹ Wir meinen hier die *eigentlichen* Matrizen mit diskreten Indizes. Später werden wir den Begriff der Matrix in erweitertem Sinne gebrauchen, wobei dann der Begriff der Gleichartigkeit wieder eine Rolle spielen wird.

einfache Reihen, wie 1, 2, 3, . . . 1, 2, 3, . . . usw., kurz mit beliebigem „Ordnungstypus“. In jedem Einzelfalle zeigt sich gewöhnlich eine bestimmte Ordnung der Diagonalelemente als vorteilhaft.

Ferner kann man jeden Index in mehrere aufspalten, also etwa n durch $n_1 n_2 \dots n_f$ ersetzen; man hat dann Matrizen „in f Dimensionen“, mit Elementen $A_{n_1 n_2 \dots n_f; m_1 m_2 \dots m_f}$. Umgekehrt kann man immer eine solche mehrdimensionale Matrix in eine eindimensionale umordnen. Die Multiplikationsregel für mehrdimensionale Matrizen lautet:

$$(3) \quad (A B)_{n_1, n_2, \dots, n_f; m_1, m_2, \dots, m_f} = \sum_{k_1, \dots, k_f} A_{n_1, \dots, n_f; k_1, \dots, k_f} B_{k_1, \dots, k_f; m_1, \dots, m_f}.$$

Ein besonders wichtiges Beispiel für den Fall $f = 2$ wird im folgenden Paragraphen ausführlich behandelt (Matrix-Matrizen, Übermatrizen). Überall, wo es die Deutlichkeit gestattet, werden wir die eindimensionale Schreibweise der Matrizen bevorzugen, also das Indizessystem n_1, \dots, n_f durch einen Index n ersetzen.

Im Gegensatz zu der gewöhnlichen Zahlenrechnung gilt in der Matrizenrechnung *nicht* der Satz, daß aus $AB = 0$ entweder $A = 0$ oder $B = 0$ folgt. So ist z. B. das Produkt der beiden nicht verschwindenden Matrizen

$$A = \begin{pmatrix} 1 & 0 & 0 & \dots \\ 0 & 0 & 0 & \dots \\ 0 & 0 & 0 & \dots \\ \dots & \dots & \dots & \dots \end{pmatrix}, \quad B = \begin{pmatrix} 0 & 0 & 0 & \dots \\ 0 & 1 & 0 & \dots \\ 0 & 0 & 0 & \dots \\ \dots & \dots & \dots & \dots \end{pmatrix}$$

gleich Null. Wenn es zu einer Matrix A eine andere, von 0 verschiedene B gibt, so daß entweder $AB = 0$ oder $BA = 0$, so nennt man A einen *Teiler der Null* oder *Nullteiler*.

Für die beiden Grundoperationen (2) gilt das *assoziative Gesetz*:

$$(4) \quad \begin{cases} (A + B) + C = A + (B + C), \\ (A B) C = A (B C). \end{cases}$$

Für die Addition gilt das *kommutative Gesetz*:

$$(5) \quad A + B = B + A;$$

für die Multiplikation dagegen im allgemeinen *nicht*. Wenn in

besonderen Fällen für zwei Matrizen A und B

$$(6) \quad AB = BA$$

ist, so heißen A und B *vertauschbar*.

Die Differenz $AB - BA$ ist ein Maß der Abweichung von der Vertauschbarkeit und soll durch ein Klammersymbol abgekürzt werden; wir setzen unter Hinzufügung eines konstanten, später geeignet zu wählenden Zahlfaktors¹ $\frac{1}{\varkappa}$:

$$(7) \quad [A, B] = \frac{1}{\varkappa} (AB - BA).$$

Das *distributive Gesetz* zwischen Addition und Multiplikation folgt unmittelbar aus den Definitionen (2), und zwar gilt

$$(8) \quad \begin{cases} A(B + C) = AB + AC, \\ (B + C)A = BA + CA, \end{cases}$$

Diese beiden Formeln sind wegen der Ungültigkeit des kommutativen Gesetzes der Multiplikation voneinander logisch unabhängig.

Nach (4) ist auch eine Summe

$$(9) \quad A = A_1 + A_2 + \cdots + A_s$$

oder ein Produkt

$$(9a) \quad A' = A_1 A_2 \dots A_s$$

eindeutig definiert, ohne daß man angeben muß, welche Additionen oder Multiplikationen *zuerst* auszuführen sind. Ist insbesondere $A_1 = A_2 = \cdots = A_s$, so schreiben wir kurz

$$(10) \quad \begin{cases} A = A_1 + A_1 + \cdots + A_1 = s A_1, \\ A' = A_1 A_1 \dots A_1 = A_1^s. \end{cases}$$

Als *Einheitsmatrix* bezeichnet man die Matrix

$$(11) \quad E = (\delta_{nm}); \quad \delta_{nm} = \begin{cases} 1 & \text{für } n = m, \\ 0 & n \neq m. \end{cases}$$

¹ Die Definition der Multiplikation einer Matrix mit einem Zahlfaktor werden wir sogleich nachholen. [Vgl. (14), (15), (15a).]

Sie hat die Eigenschaft, daß für jede beliebige Matrix A die Gleichung

$$(12) \quad EA = AE = A$$

besteht. Man schreibt deshalb statt E oft auch einfach 1.

Bei endlichen Matrizen A existiert eine *Determinante* $\text{Det } A$. Wegen der bekannten Multiplikationsregel der Determinanten gilt

$$(13) \quad \text{Det}(AB) = \text{Det } A \cdot \text{Det } B.$$

Als *Repräsentant einer gewöhnlichen Zahl* in der Matrizenrechnung definieren wir die Matrix

$$(14) \quad c = (c \delta_{nm}),$$

die aus E durch Multiplikation jedes Elements mit c hervorgeht. Dann ist das *Produkt der Zahl c mit der Matrix A* als das Matrixprodukt

$$(15) \quad cA = \left(\sum_k c \delta_{nk} A_{km} \right) = (c A_{nm})$$

zu bestimmen, d. h. jedes Element von cA entsteht aus dem von A durch Multiplikation mit der Zahl c . Offenbar ist

$$(15a) \quad cA = A c.$$

Für ganzzahlige c ist die Definition (14) im Einklang mit der ersten Gleichung (10).

Diese Matrizen c sind spezielle Fälle von *Diagonalmatrizen*, d. h. solchen, bei denen alle außerhalb der Diagonale stehenden Elemente verschwinden:

$$(16) \quad a = (a_n \delta_{nm}).$$

Sind insbesondere die Diagonalelemente komplexe Zahlen vom Betrage 1, so sprechen wir von einer *Phasenmatrix*:

$$(17) \quad f = (e^{i\varphi_n} \delta_{nm}).$$

Wenn zu einer Matrix A eine zweite Matrix A' existiert, für die

$$(18) \quad AA' = A'A = E$$

ist, so nennt man A' *Reziproke* von A und schreibt

$$(19) \quad A' = A^{-1}.$$

Es gibt nur *eine* Reziproke; denn wäre außer (18) noch $AA'' = E$, so folgte daraus durch vordere Multiplikation mit

A' , daß $A'AA'' = A'E$, folglich wegen (18): $EA'' = A'E$, also $A'' = A'$.

Bei *endlichen* Matrizen ist die notwendige und hinreichende Bedingung für die Existenz der Reziproken das Nichtverschwinden der Determinante:

$$\text{Det } A \neq 0.$$

Bei *unendlichen* Matrizen hat im allgemeinen der Determinantenbegriff keinen Sinn. Bei diesen kann es vorkommen, daß es zu einer Matrix A eine andere A' gibt, für die zwar $AA' = E$, aber zugleich $A'A \neq E$ ist, oder umgekehrt; dann existiert nach unserer Definition *keine* Reziproke. Bei endlichen Matrizen aber folgt aus $AA' = E$ immer auch $A'A = E$.

Aus der Definition folgt unmittelbar für die Reziproke eines Produkts

$$(20) \quad (AB)^{-1} = B^{-1}A^{-1}.$$

Die Summe der Diagonalglieder einer Matrix A nennt man die *Spur* von A ,

$$(21) \quad Sp A = \sum_n A_{nn}.$$

Für diese Größe gelten die Regeln

$$(22) \quad \begin{cases} Sp(A + B) = Sp A + Sp B, \\ Sp(AB) = Sp(BA); \end{cases}$$

für die letzte Gleichung kann man nach (7) auch schreiben

$$(22a) \quad Sp[A, B] = 0.$$

Wir bemerken aber ausdrücklich, daß dies *nur für endliche Matrizen* allgemein gilt; bei unendlichen können die Reihen $Sp(AB)$ und $Sp(BA)$ divergieren, auch wenn die Produkte AB und BA existieren. Offenbar kann schließlich nach (22) auch

$$(23) \quad Sp(ABA^{-1}) = Sp(B)$$

behauptet werden.

Die Spur der endlichen Einheitsmatrix $E^{(N)}$ von N Zeilen und Spalten ist gleich N :

$$(24) \quad Sp E^{(N)} = N.$$

Wenn die Elemente einer Matrix von einem Parameter abhängen, so versteht man unter der *Ableitung der Matrix nach*

dem Parameter die Matrix, deren Elemente die Ableitungen der Elemente der gegebenen Matrix nach dem Parameter sind. Sehr häufig ist der Fall, daß der Parameter die Zeit t ist. Unter der Ableitung der Matrix $x = (x_{nm}(t))$ nach der Zeit verstehen wir also die Matrix

$$(25) \quad \dot{x} = (\dot{x}_{nm}(t)).$$

Hierfür gelten die Rechenregeln

$$(26) \quad \left\{ \begin{array}{l} \frac{d}{dt}(x + y) = \dot{x} + \dot{y}, \\ \frac{d}{dt}xy = x\dot{y} + \dot{x}y, \end{array} \right.$$

die sofort aus der Definition folgen.

Aus der Definition der Reziproken $xx^{-1} = 1$ folgt nach (25) $\dot{x}x^{-1} + x\dot{x}^{-1} = 0$, also

$$(27) \quad \dot{x}^{-1} = -x^{-1}\dot{x}x^{-1}.$$

In § 3 hatten wir insbesondere Matrizen mit rein-periodischen (harmonischen) Elementen betrachtet von der Form

$$(28) \quad x = (x_{nm}) = (a_{nm}e^{2\pi i\nu_{nm}t}), \quad \hbar\nu_{nm} = W_n - W_m.$$

Man kann die zeitliche Ableitung solcher Matrizen durch das Klammersymbol (7) ausdrücken, indem man die Diagonalmatrix

$$(29) \quad W = (W_n\delta_{nm})$$

einführt; dann ist offenbar

$$\begin{aligned} \dot{x} &= (\dot{x}_{nm}) = (2\pi i\nu_{nm}x_{nm}) \\ &= \frac{2\pi i}{\hbar}((W_n - W_m)x_{nm}) \\ &= \frac{2\pi i}{\hbar}(Wx - xW). \end{aligned}$$

Wählt man nun für den in (7) vorläufig unbestimmt gelassenen Faktor \varkappa den Wert:

$$(30) \quad \varkappa = \frac{\hbar}{2\pi i},$$

so wird

$$(31) \quad \dot{x} = [W, x].$$

Wir werden später diese Wahl von \varkappa als sehr zweckmäßig für die Formulierung der Quantenmechanik erkennen.

Im allgemeinen werden wir jedoch nicht, wie in (28), die mit den Zeitfaktoren $e^{2\pi i \nu_{nm} t}$ versehenen Matrizen für die Quantenmechanik benutzen, sondern mit Matrizen $q = (q(nm))$ rechnen, deren Zeitableitung *definitionsgemäß* durch § 3 (13) gegeben ist. Dabei gilt offenbar ebenfalls die Beziehung $\dot{q} = [W, q]$.

§ 5. Rechteckige Matrizen und Übermatrizen. Bisher haben wir von endlichen Matrizen vorausgesetzt, daß sie „quadratisch“ sind, d. h. daß die Anzahl der Zeilen gleich der der Spalten ist. Man kann aber auch *rechteckige Matrizen* bilden. Diese haben zwar keine unmittelbare physikalische Bedeutung, sind aber für manche Rechnungen als Hilfsbegriff nützlich. Wir nennen die Anzahl der Zeilen r und die der Spalten s ; dann schreiben wir die rechteckige Matrix

$$(1) \quad A = \begin{pmatrix} A_{11} & A_{12} & \dots & A_{1s} \\ A_{21} & A_{22} & \dots & A_{2s} \\ \dots & \dots & \dots & \dots \\ A_{r1} & A_{r2} & \dots & A_{rs} \end{pmatrix}.$$

Man kann zwei solche Matrizen, A mit r, s und B mit r', s' offenbar nach der Regel § 4, (2) *addieren*, wenn $r = r', s = s'$ ist und ferner *in der Reihenfolge AB multiplizieren*, wenn $s = r'$ ist; im letzteren Falle definiert man als *Produkt eine neue Matrix* von r Zeilen und s' Spalten:

$$(2) \quad (AB)_{nm} = \sum_{k=1}^s A_{nk} B_{km}, \quad \begin{cases} n = 1, 2, \dots, r \\ m = 1, 2, \dots, s' \end{cases}.$$

Soll es auch möglich sein, das Produkt BA zu bilden, so muß noch weiter $s' = r$ sein. Bildet man nun die Spuren der beiden Produkte

$$Sp(A B) = \sum_{n=1}^r \sum_{k=1}^s A_{nk} B_{kn},$$

$$Sp(B A) = \sum_{n=1}^s \sum_{k=1}^r B_{nk} A_{kn},$$

so sieht man, daß sie einander gleich sind: *Wenn es möglich ist, aus zwei rechteckigen Matrizen A, B die beiden Produkte AB*

und BA zu bilden, so gilt

$$(3) \quad Sp(AB) = Sp(BA).$$

Diese Regeln benutzen wir, um das Produkt zweier endlicher quadratischer oder zweier unendlicher Matrizen in eine Matrix von Produkten rechteckiger Teilmatrizen aufzulösen. Wir spalten den Index n in zwei r, ϱ auf, indem wir die Zahlenreihe

$$n = 1, 2, 3, \dots$$

zunächst in Gruppen mit den Nummern $r = 1, 2, \dots$ zu je $g_1, g_2, \dots, g_r, \dots$ Zahlen teilen und die Zahlen jeder Gruppe mit $\varrho = 1, 2, \dots, g_r$ durchnummerieren:

$$\begin{array}{l|ll} n = 1, 2, \dots, g_1 & g_1 + 1, g_1 + 2, \dots, g_1 + g_2 & g_1 + g_2 + 1, \dots \\ r = & 1 & 2 & 3 \\ \varrho = 1, 2, \dots, g_1 & 1, & 2, \dots & g_2 & 1, \dots \end{array}$$

Ebenso spalten wir den Index m in s, σ mit den Teilungszahlen $\bar{g}_1, \bar{g}_2, \dots, \bar{g}_s, \dots$.

Das Element einer quadratischen oder unendlichen Matrix $A_{n,m} = A_{\varrho,\sigma}^{(r,s)}$ läßt sich dann als Element einer rechteckigen Matrix

$$(4) \quad A^{(r,s)} = \begin{pmatrix} A_{11}^{(r,s)}, A_{12}^{(r,s)}, \dots, A_{1,g_s}^{(r,s)} \\ \cdot & \cdot & \cdot & \cdot \\ A_{g_r,1}^{(r,s)}, A_{g_r,2}^{(r,s)}, & & & A_{g_r,g_s}^{(r,s)} \end{pmatrix} = (A_{\varrho,\sigma}^{(r,s)})$$

von g_r Reihen und \bar{g}_s Spalten auffassen, und die ganze Matrix A schreibt sich als *Übermatrix* (Matrizen-Matrix):

$$(5) \quad A = (A^{(r,s)}) = \begin{pmatrix} A^{(1,1)} & A^{(1,2)} & \dots \\ A^{(2,1)} & A^{(2,2)} & \dots \\ \dots & \dots & \dots \end{pmatrix}.$$

Die $A^{(r,s)}$ nennen wir „*Untermatrizen*“ oder auch „*Teilmatrizen*“.

Hat man zwei allgemeine Übermatrizen A, B mit den Teilungszahlen g_r, \bar{g}_t und \bar{g}_t, \bar{g}_s , so daß also die Spaltenteilung von A mit der Zeilenteilung von B übereinstimmt, so ist das Element des Produkts

$$\begin{aligned} (AB)_{\varrho\sigma}^{(r,s)} &= \sum_t \sum_{\tau=1}^{\bar{g}_t} A_{\varrho\tau}^{(r,t)} B_{\tau\sigma}^{(t,s)} \\ &= \left(\sum_t A^{(r,t)} B^{(t,s)} \right)_{\varrho\sigma}. \end{aligned}$$

In der Klammer steht das Produkt der beiden Matrizen A, B so gebildet, als wenn die Teilmatrizen $A^{(r,t)}, B^{(t,s)}$ Zahlen wären. AB ist also eine Matrizenmatrix mit den Teilungszahlen g_r, \bar{g}_s von der Form

$$(6) \quad AB = (A^{(r,t)}) \cdot (B^{(t,s)}) = \left(\sum_t A^{(r,t)} B^{(t,s)} \right).$$

Wir werden später hauptsächlich mit solchen Fällen zu tun haben, wo bei allen in Betracht kommenden Matrizen Zeilen und Spalten in *derselben* Weise eingeteilt sind, d. h. wo $\bar{g}_r = g_r$ ist.

Sind in einem solchen Falle alle $A^{(r,s)} = 0$ für $r \neq s$, so sprechen wir von einer *Stufenmatrix*:

$$(7) \quad A = (A^{(r)} \delta_{r,s}) = \begin{pmatrix} A^{(1)} & 0 & \dots \\ 0 & A^{(2)} & \dots \\ \dots & \dots & \dots \end{pmatrix}.$$

Sie kann als Verallgemeinerung der Diagonalmatrix aufgefaßt werden.

Sind A, B zwei Stufenmatrizen mit gleichen Stufen, also $g_1 = \bar{g}_1 = \bar{\bar{g}}_1, g_2 = \bar{g}_2 = \bar{\bar{g}}_2, \dots$ und

$$(8) \quad A^{(r,s)} = A^{(r)} \delta_{rs}, \quad B^{(r,s)} = B^{(r)} \delta_{rs},$$

so ist AB wieder eine Stufenmatrix derselben Art:

$$(9) \quad (A^{(r)} \delta_{rs}) \cdot (B^{(r)} \delta_{rs}) = (A^{(r)} B^{(r)} \delta_{rs}).$$

Ist P eine beliebige Matrix und A eine Stufenmatrix mit der Stufengröße g_r , wie (7), so kann man die Zeilen von P ebenfalls nach je g_r abteilen, die Spalten nach beliebigen Zahlen \bar{g}_s ; dann wird nach (6)

$$(10) \quad AP = (A^{(r)} P^{(r,s)}).$$

In analoger Weise erhält man

$$(11) \quad PA = (P^{(r,s)} A^{(s)}),$$

Formeln, die ganz analog zu denen sind, die man bei der Multiplikation einer beliebigen Matrix mit einer Diagonalmatrix erhält; nur sind hier die Elemente selber Matrizen, also im allgemeinen nicht vertauschbar.

Ist aber speziell jede Untermatrix $A^{(r)}$ ein Vielfaches der g_r -dimensionalen Einheitsmatrix, $A^{(r)} = a^{(r)} E^{(g_r)}$, so kann man $A^{(r)}$ mit $P^{(r,s)}$ vertauschen. In diesem Falle ist A eine *geordnete*

Diagonalmatrix, in deren Diagonale die Zahlen

$$a^{(1)} a^{(1)} \dots a^{(1)} a^{(2)} a^{(2)} \dots a^{(2)} a^{(3)} \dots$$

stehen, und zwar g_1 mal $a^{(1)}$, g_2 mal $a^{(2)}$ usw. Man hat also den

Satz: Für eine geordnete Diagonalmatrix A und eine beliebige Matrix P gilt

$$(12) \quad \kappa [A, P] = (A P - P A) = ((a^{(r)} - a^{(s)}) P^{(r, s)}),$$

wo $a^{(r)}$ eine g_r -faches Diagonalelement von A ist ($a^{(r)} \neq a^{(s)}$ für $r \neq s$).

Als Korollar hat man:

$$(13) \quad \text{Aus } [A, P] = 0 \text{ folgt } P^{(r, s)} = 0 \text{ für } r \neq s.$$

Die Gleichung für P bei gegebenem Q

$$(14) \quad \kappa [A, P] = A P - P A = Q$$

ist dann und nur dann lösbar, wenn $Q^{(r, r)} = 0$; die Lösung lautet

$$(15) \quad P^{(r, s)} = \frac{Q^{(r, s)}}{a^{(r)} - a^{(s)}}, \quad r \neq s$$

und $P^{(r, r)}$ bleibt unbestimmt.

Hiervon werden wir später viel Gebrauch machen.

Aus den beiden gleichstufigen Matrizen A, B (8) und einer beliebigen Matrix P erhält man nach (10) und (11)

$$(16) \quad A P B = (A^{(r)} P^{(r, s)} B^{(s)}).$$

§ 6. Matrizen mit Symmetrien. Wir haben in § 3 gesehen, daß in der Atommechanik eine besondere Klasse von Matrizen eine Rolle spielt, die hermiteschen, deren Elemente bei Vertauschung der Indizes in den konjugiert komplexen Wert übergehen. Wir wollen hier diese und ähnliche Arten von Matrizen mit Symmetrieeigenschaften der Elemente untersuchen.

Wir bezeichnen die aus einer Matrix $A = (A_{nm})$ durch Vertauschung von Zeilen und Spalten hervorgehende *transponierte* oder *gestürzte Matrix* mit \tilde{A} :

$$(1) \quad \tilde{A}_{nm} = A_{mn}.$$

Es gelten die Regeln:

$$(2) \quad \left\{ \begin{array}{l} \overline{A + B} = \tilde{A} + \tilde{B}, \\ \widetilde{A B} = \tilde{B} \tilde{A}. \end{array} \right.$$

Für die Matrix, die aus \tilde{A} dadurch entsteht, daß man jedes Element durch die konjugiert komplexe Zahl ersetzt, gebrauchen wir ein besonderes Zeichen:

$$(3) \quad \tilde{A}^* = A^\dagger,$$

und nennen sie die zu A *adjungierte Matrix*. Wegen (2) gilt dann auch:

$$(4) \quad \begin{cases} (A + B)^\dagger = A^\dagger + B^\dagger, \\ (AB)^\dagger = B^\dagger A^\dagger. \end{cases}$$

Eine Matrix, für die

$$(5) \quad A^\dagger = A$$

gilt, heißt „*sich selbst adjungiert*“ oder „*hermitisch*“; man nennt sie auch „*symmetrisch*“ im übertragenen Sinne (nämlich für komplexe Elemente) oder endlich „*reell*“, letzteres, weil die zugehörige hermitische Form, wie wir im § 9 sehen werden, reelle Werte hat.

Aus (4) folgt unmittelbar, daß die Summe zweier hermitescher Matrizen A und B auch hermitisch ist, das Produkt aber dann und nur dann, wenn A, B vertauschbar sind:

$$(6) \quad \begin{cases} (A + B)^\dagger = A + B, & \text{für } A^\dagger = A, B^\dagger = B, \\ (AB)^\dagger = AB, & \text{für } A^\dagger = A, B^\dagger = B, [A, B] = 0. \end{cases}$$

Jede Matrix A läßt sich auf eine und nur eine Weise darstellen in der Form

$$(7) \quad A = A_1 + iA_2$$

mit hermiteschen Matrizen A_1, A_2 ; nämlich:

$$A_1 = \frac{1}{2}(A + A^\dagger), \quad A_2 = \frac{1}{2i}(A - A^\dagger).$$

Eine Matrix, für die

$$(8) \quad A^\dagger = -A$$

ist, heißt „*schiefsymmetrisch*“ oder „*alternierend*“ oder kurz „*schief*“, im übertragenen Sinne (für komplexe Elemente); man sagt auch „*rein imaginär*“, wie man die hermiteschen Matrizen „*reell*“ nennt. Die schiefen Matrizen sind nicht wesentlich von den hermiteschen verschieden; denn wenn A schief ist, so ist offenbar iA hermitisch, und umgekehrt. Wir werden daher von

dem Begriff der schiefen Matrizen so gut wie keinen Gebrauch machen.

Bildet man aus einer hermiteschen Matrix A mit Hilfe einer beliebigen Matrix S die Matrix

$$(9) \quad B = S^\dagger A S,$$

so ist diese wieder hermitisch. Denn man hat

$$(9a) \quad B^\dagger = S^\dagger A^\dagger S = S^\dagger A S = B.$$

Die Spur des Produkts einer von Null verschiedenen Matrix A mit ihrer Adjungierten ist eine positive Zahl:

$$(10) \quad Sp(AA^\dagger) = \sum_{nm} |A_{nm}|^2.$$

Für die *Reziproke der Adjungierten* hat man einen besonderen Namen; man nennt

$$(11) \quad \check{A} = (A^\dagger)^{-1}$$

die zu A *kontragrediente Matrix*. Da nach (4)

$$A^\dagger (A^{-1})^\dagger = (A^{-1}A)^\dagger \Rightarrow E^\dagger = E$$

ist, so hat man

$$(12) \quad \check{A} = (A^\dagger)^{-1} = (A^{-1})^\dagger.$$

Danach ist umgekehrt A kontragredient zu \check{A} . Ferner folgt aus (4) mit Rücksicht auf § 4, (20), daß

$$(13) \quad \text{mit } AB = C \quad \text{auch } \check{A}\check{B} = \check{C}$$

ist. Die zu A, B, \dots kontragredienten Matrizen $\check{A}, \check{B}, \dots$ verhalten sich also hinsichtlich der Multiplikation, wie man in der Gruppentheorie sagt, „isomorph“ zu den A, B, \dots .

Zu jeder hermiteschen Matrix, insbesondere zu jeder reellen Diagonalmatrix, ist ihre Reziproke kontragredient.

Eine Matrix, die mit ihrer kontragredienten Matrix identisch ist, heißt *unitär* (oder hermitisch orthogonal); man hat also für *unitäre Matrizen* folgende äquivalente Definitionen:

$$(14) \quad \check{U} = U, \quad U^\dagger = U^{-1}, \quad UU^\dagger = U^\dagger U = E.$$

Zugleich mit U ist wegen (12) auch U^{-1} eine unitäre Matrix.

Als Beispiel einer unitären Matrix erwähnen wir die *Permutationsmatrix*, die dadurch entsteht, daß man die Zeilen (oder die

Spalten, aber nicht beides) der Einheitsmatrix einer beliebigen Permutation unterwirft.

Eine *unitäre Diagonalmatrix* ist offenbar eine *Phasenmatrix* [s. § 4, (17)].

Wir stellen eine Reihe von

Sätzen über unitäre Matrizen

zusammen.

a) *Das Produkt zweier unitärer Matrizen ist wieder unitär:*

$$(15) \quad (U_1 U_2)^\dagger = (U_1 U_2)^{-1}.$$

Das folgt, wie (13), aus (4) und § 4, (20).

b) *Ist A eine beliebige und U eine unitäre Matrix, so folgt*

$$(16) \quad \text{aus } B = U^{-1} A U \quad \text{auch } B^\dagger = U^{-1} A^\dagger U.$$

Denn es ist nach (12)

$$B^\dagger = U^\dagger A^\dagger (U^{-1})^\dagger = U^{-1} A^\dagger U.$$

Aus diesem Satze folgt unmittelbar:

c) *Wenn A hermitisch ist und U unitär, so ist auch*

$$(17) \quad B = U^{-1} A U$$

hermitisch.

Eine Umkehrung des Satzes lautet:

d) *Sind A und B hermitisch und gilt*

$$(18) \quad B = U^{-1} A U,$$

so ist $U U^\dagger$ mit A und $U^\dagger U$ mit B vertauschbar. Denn man hat dann

$$B^\dagger = U^\dagger A (U^{-1})^\dagger = B = U^{-1} A U$$

und schließt daraus durch vordere Multiplikation mit U und hintere mit U^\dagger :

$$(18a) \quad U U^\dagger A = A U U^\dagger;$$

daraus folgt sofort

$$(18b) \quad U^\dagger U B = B U^\dagger U$$

e) *Ist U eine endliche unitäre Matrix, so ist die Determinante von U eine Zahl vom Betrage 1:*

$$(19) \quad |\text{Det } U| = 1.$$

Denn man hat

$$\text{Det } U^\dagger = \text{Det } \tilde{U}^* = (\text{Det } \tilde{U})^* = (\text{Det } U)^*,$$

also

$$\begin{aligned} |\text{Det } U|^2 &= (\text{Det } U) \cdot (\text{Det } U)^* \\ &= (\text{Det } U) \cdot (\text{Det } U^\dagger) \\ &= \text{Det } (U U^\dagger) = 1. \end{aligned}$$

Man kann den Begriff und das Symbol der adjungierten Matrix auch auf rechteckige Matrizen übertragen. Wir definieren zunächst die Transponierte wieder durch die Formel (1) und dann die Adjungierte nach (3). Die Formeln (4) bleiben erhalten. Zu beachten ist, daß die Adjungierte ebenso viele Zeilen hat wie die ursprüngliche Matrix Spalten, und umgekehrt. Wenn also für eine Matrix $A = A^\dagger$ gilt, ist sie notwendig quadratisch.

Schreibt man nun eine quadratische Matrix A als Übermatrix mit den Diagonal-Teilungszahlen g_r ,

$$(20) \quad A = (A^{(r,s)}), \quad A^{(r,s)} = (A_{\varrho,\sigma}^{(r,s)}), \quad \begin{array}{l} \varrho = 1, 2, \dots, g_r, \\ \sigma = 1, 2, \dots, g_s, \end{array}$$

so entsteht die Frage, wie die transponierte und die adjungierte Übermatrix aussieht.

Offenbar hat man

$$(21) \quad \tilde{A} = (\tilde{A}^{(r,s)}) = (\overline{A^{(s,r)}}), \quad \overline{A_{\varrho,\sigma}^{(s,r)}} = A_{\sigma,\varrho}^{(s,r)},$$

also auch

$$(22) \quad A^\dagger = (A^{\dagger(r,s)}) = (A^{(s,r)\dagger}), \quad A_{\sigma,\varrho}^{(s,r)\dagger} = A_{\varrho,\sigma}^{(s,r)*}.$$

Ist A hermitisch, so folgt für die Untermatrizen:

$$(23) \quad A^{(r,s)} = A^{(s,r)\dagger};$$

insbesondere sind die Diagonal-Untermatrizen selbst hermitisch:

$$(24) \quad A^{(r,r)} = A^{(r,r)\dagger}.$$

Für eine unitäre Übermatrix $U = U^{(r,s)}$ hat man:

$$(25) \quad \sum_t U^{(r,t)} U^{(s,t)\dagger} = \sum_t U^{(t,r)\dagger} U^{(t,s)} = \delta_{rs} E^{(r)},$$

wo $E^{(r)}$ die g_r -dimensionale Einheitsmatrix ist.

Handelt es sich insbesondere um eine unitäre Stufenmatrix $U = (U^{(r)} \delta_{rs})$, so folgt aus (25)

$$(26) \quad U^{(r)} U^{(r)\dagger} = U^{(r)\dagger} U^{(r)} = E^{(r)},$$

d. h. die Diagonal-Untermatrizen sind selbst unitär.

§ 7. Matrizenfunktionen. Aus Matrizen X_1, X_2, \dots, X_f , die man sich als variabel vorstellt, kann man durch endlich viele Additionen und Multiplikationen Polynome mit Zahlenkoeffizienten

$$(1) \quad P(X_1, X_2, \dots, X_f) = \sum_{\nu} \sum_{n_1, n_2, \dots, n_{\nu}} c_{n_1, n_2, \dots, n_{\nu}}^{(\nu)} X_{n_1} X_{n_2} \dots X_{n_{\nu-1}} X_{n_{\nu}}$$

bilden; dabei ist n_1, n_2, \dots, n_{ν} irgend eine Anordnung der Zahlen $1, 2, \dots, f$ mit Wiederholungen und ν kennzeichnet den Grad des betreffenden Gliedes¹.

Wir definieren das *adjungierte Polynom* durch

$$(2) \quad P^{\dagger}(X_1, X_2, \dots, X_f) = \sum_{\nu} \sum_{n_1, n_2, \dots, n_{\nu}} c_{n_1, n_2, \dots, n_{\nu}}^{(\nu)*} X_{n_{\nu}} X_{n_{\nu-1}} \dots X_{n_2} X_{n_1},$$

d. h. durch die Regel: *Man kehre in allen Gliedern die Reihenfolge der Matrixfaktoren um und ersetze die Zahlenkoeffizienten durch ihre konjugiert komplexen Werte.*

Das Polynom (2) hat offenbar nach § 6, (3), (4) die Eigenschaft, daß es die zur Matrix $P(X_1, \dots, X_f)$ adjungierte Matrix darstellt, wenn man in $P^{\dagger}(X_1, \dots, X_f)$ die adjungierten Argumentmatrizen $X_1^{\dagger}, \dots, X_f^{\dagger}$ einsetzt:

$$(3) \quad P^{\dagger}(X_1^{\dagger}, X_2^{\dagger}, \dots, X_f^{\dagger}) = (P(X_1, X_2, \dots, X_f))^{\dagger}.$$

Das Polynom $P(X_1, \dots, X_f)$ heißt *sich selbst adjungiert*, wenn identisch in den X_1, X_2, \dots, X_f

$$(4) \quad P^{\dagger}(X_1, \dots, X_f) = P(X_1, \dots, X_f)$$

ist. Dann ist nämlich für sich selbst adjungierte, d. h. hermitische Argumentmatrizen $X_1^{\dagger} = X_1, \dots, X_f^{\dagger} = X_f$:

$$\begin{aligned} (P(X_1, \dots, X_f))^{\dagger} &= P^{\dagger}(X_1^{\dagger}, \dots, X_f^{\dagger}), \\ &= P^{\dagger}(X_1, \dots, X_f), \\ &= P(X_1, \dots, X_f). \end{aligned}$$

¹ Diese Schreibweise ist nötig, weil die Matrizen X_k im allgemeinen nicht kommutativ sind.

Ausgehend von den Polynomen kann man durch Grenzübergang zu unendlichen Potenzreihen — falls sie konvergieren — allgemeinere *Matrizenfunktionen* definieren:

$$f(X_1, \dots, X_f).$$

Für diese gilt der

Satz 1: Ersetzt man in einer Matrizenfunktion $f(X_1, \dots, X_f)$ die Argumente durch

$$(5) \quad Y_1 = T^{-1} X_1 T, \dots, Y_f = T^{-1} X_f T,$$

wo T eine beliebige Matrix ist, so geht die Funktion über in

$$(5a) \quad f(Y_1, \dots, Y_f) = T^{-1} f(X_1, \dots, X_f) T.$$

Die Behauptung ist offenbar richtig, wenn f eine Summe $X_1 + X_2$ oder ein Produkt $X_1 X_2$ ist:

$$T^{-1} X_1 T + T^{-1} X_2 T = T^{-1} (X_1 + X_2) T,$$

$$T^{-1} X_1 T \cdot T^{-1} X_2 T = T^{-1} X_1 X_2 T.$$

Daher gilt sie auch für Polynome und überträgt sich von da durch Grenzübergang auf beliebige Potenzreihen.

Auch die Begriffe der *adjungierten Funktion* und der *Sichselbstadjungiertheit* übertragen sich sofort auf beliebige Funktionen. Matrixfunktionen *vertauschbarer* Argumente, definiert durch die Potenzreihe

$$(6) \quad f(X_1, \dots, X_f) = \sum_{n_1 \dots n_f} c_{n_1 \dots n_f} X_1^{n_1}, \dots, X_f^{n_f},$$

sind offenbar sich selbst adjungiert, wenn die Koeffizienten c reell sind. Speziell ist jede Funktion eines Arguments mit reellen Koeffizienten

$$(7) \quad f(X) = \sum_n c_n X^n$$

sich selbst adjungiert. Als Beispiel einer solchen betrachten wir die *Exponentialfunktion*

$$(8) \quad e^X = \sum_{n=0}^{\infty} \frac{X^n}{n!}.$$

Wenn X und Y vertauschbare Matrizen sind, so gilt offenbar, wie bei Zahlen

$$(9) \quad e^X e^Y = e^{X+Y}.$$

Da X und $-X$ vertauschbar sind, und nach (8) $e^0 = 1$ ist, so folgt aus (9), daß e^{-X} die Reziproke zu e^X ist.

Ist A eine hermitesche Matrix, so ist es auch e^A wegen der Selbstadjungiertheit der Funktion e^X . Dagegen ist e^{iA} unitär; denn aus $A = A^\dagger$ folgt $(e^{iA})^\dagger = e^{-iA^\dagger} = e^{-iA}$. Wir werden nachher die Umkehrung beweisen, daß jede unitäre Matrix sich durch eine hermitesche A in der Form $U = e^{iA}$ darstellen läßt.

Man kann auch auf andere Weise aus hermiteschen Matrizen unitäre gewinnen und umgekehrt. Nach CAYLEY gilt z. B. folgendes: Die Matrix

$$(10) \quad U = \frac{1 - iA}{1 + iA}$$

ist dann und nur dann unitär, wenn A hermitisch ist. Dabei ist die Schreibweise als Bruch erlaubt, weil Zähler und Nenner als Funktionen derselben Matrix vertauschbar sind. Man hat nach (3)

$$(11) \quad U^\dagger = \frac{1 + iA^\dagger}{1 - iA^\dagger};$$

also folgt für $A = A^\dagger$ in der Tat

$$U U^\dagger = U^\dagger U = 1.$$

Man kann ferner die Gleichung (10) auflösen und hat

$$(12) \quad A = i \frac{U - 1}{U + 1}, \quad A^\dagger = -i \frac{U^\dagger - 1}{U^\dagger + 1};$$

daraus folgt für $U^\dagger = U^{-1}$ auch $A = A^\dagger$. Nach (12) kann jede unitäre Matrix U in der Form (10) geschrieben werden, abgesehen von den singulären Fällen, bei denen $U + 1$ keine Reziproke hat. Die Formel (10) zeigt daher, wieviele *unabhängige Parameter* eine unitäre Matrix enthält: Nämlich ebensoviele, wie eine hermitesche Matrix von gleich vielen Zeilen und Spalten. Ist f deren Anzahl, so stehen

$$\frac{1}{2} f(f - 1)$$

komplexe Zahlen oberhalb der Diagonale und f reelle Zahlen in der Diagonale; im ganzen hat man also $f(f - 1) + f = f^2$ reelle Konstanten.

An die hier entwickelte Matrizenalgebra kann man eine *Matrizenanalysis* anschließen. Differentiationsartige Prozesse kann man natürlich in mannigfaltiger Weise definieren. Hier kommt vor allem der einfachste solche Prozeß in Betracht, bei

dem die Variation der unabhängigen Matrixvariablen nur in der Diagonale vorgenommen wird. Ist $f(x, y, z, \dots)$ eine Funktion der unabhängigen Matrizen x, y, z, \dots , so definieren wir als *partielle Ableitung nach x* :

$$(13) \quad \frac{\partial f}{\partial x} = \lim_{a \rightarrow 0} \frac{f(x + aE, y, z, \dots) - f(x, y, z, \dots)}{a};$$

dabei ist a eine gewöhnliche Zahl¹.

Hiernach ist z. B.

$$(14) \quad \begin{aligned} \frac{dx}{dx} &= E = 1, \\ \frac{dx^2}{dx} &= \left(\lim_{a \rightarrow 0} \frac{1}{a} \sum_k \{ (x_{nk} + a\delta_{nk})(x_{km} + a\delta_{km}) - x_{nk}x_{km} \} \right) \\ &= (2x_{nm}) = 2x. \end{aligned}$$

Man sieht, daß die gewöhnlichen Regeln für die Differentiation von Summen und Produkten gelten, wobei man nur zu beachten hat, daß die Reihenfolge der Faktoren gewahrt bleibt:

$$(15) \quad \left\{ \begin{aligned} \frac{\partial}{\partial x} (f + g) &= \frac{\partial f}{\partial x} + \frac{\partial g}{\partial x}, \\ \frac{\partial}{\partial x} fg &= f \frac{\partial g}{\partial x} + \frac{\partial f}{\partial x} g. \end{aligned} \right.$$

¹ Man kann nach einer Bemerkung von REICHENBACH die Definition (13) auch durch folgende ersetzen: Man bilde aus den Matricelementen f_{nm} von f , welche Funktionen der Matricelemente x_{ki} von x sind, zunächst die Ableitungen $\frac{\partial f_{nm}}{\partial x_{ki}}$. Man erhält dann das Element $\left(\frac{\partial f}{\partial x}\right)_{nm}$ von $\frac{\partial f}{\partial x}$, indem man über $k = l$ summiert:

$$\left(\frac{\partial f}{\partial x}\right)_{nm} = \sum_k \frac{\partial f_{nm}}{\partial x_{kk}}.$$

Wenn man aber statt dessen definitionsgemäß

$$\left(\frac{\partial f}{\partial x}\right)_{nm} = \sum_k \frac{\partial f_{kk}}{\partial x_{mn}}$$

setzt, so bekommt man eine andere Differentiation, deren Gesetze gleichfalls verschiedentlich untersucht worden sind. (Vgl. BORN, M. u. P. JORDAN: Z. Phys. Bd. 34, S. 858. 1925; BORN, M., W. HEISENBERG u. P. JORDAN: Z. Phys. Bd. 35, S. 557. 1926. LANZOS, K.: Z. Phys. Bd. 36, S. 401. 1926. LONDON, F.: Z. Phys. Bd. 37, S. 915. 1926.) Erwähnt sei lediglich, daß für diese „Differentiation zweiter Art“ die Formel (16) des Textes erhalten bleibt, dagegen (15) i. A. nicht.

Die erste von diesen Gleichungen ist trivial, die zweite ergibt sich genau wie in der gewöhnlichen Analysis aus der Umformung

$$\begin{aligned} & f(x + aE, y \dots) g(x + aE, y \dots) - f(x, y \dots) g(x, y \dots) \\ &= f(x + aE, y \dots) \{g(x + aE, y \dots) - g(x, y \dots)\} \\ &\rightarrow \{f(x + aE, y \dots) - f(x, y \dots)\} g(x, y \dots). \end{aligned}$$

Aus (15) sieht man, daß die gewöhnliche Regel für die Differentiation von Potenzen

$$(16) \quad \frac{dx^n}{dx} = nx^{n-1}$$

gilt. Daher kann man auch Polynome und (gleichmäßig konvergente) Potenzreihen von Matrizen in gewöhnlicher Weise differenzieren. Z. B. erhält man für die durch (8) definierte Exponentialfunktion durch gliedweise Differentiation der Reihe

$$(17) \quad \frac{de^x}{dx} = e^x.$$

Wir bemerken noch, daß *keine* einfache Regel für die Differentiation einer Funktion von Funktionen gilt, wie in der Zahlenanalysis. Ist z. B.

$$F(f(x), \varphi(x), \psi(x)) = f(x) \cdot \varphi(x) \cdot \psi(x),$$

so folgt aus (15)

$$\frac{dF}{dx} = \frac{df}{dx} \varphi \psi + f \frac{d\varphi}{dx} \psi + f \varphi \frac{d\psi}{dx},$$

und hier läßt sich im allgemeinen das mittlere Glied nicht als Produkt von $\frac{d\varphi}{dx}$ mit der partiellen Ableitung $\frac{\partial F}{\partial \varphi} = f \psi$ schreiben.

Dieselbe Bemerkung gilt übrigens auch für die früher [§ 4, (25)] definierte Differentiation nach einem Parameter, da diese offenbar ein Spezialfall der allgemeinen Matrizendifferentiation ist; endlich ist sie auch anzuwenden auf die in (13), Kap. I § 3 definierte Differentiation nach der *Zeit*, da diese formal als Differentiation nach einer als Parameter betrachteten Größe t dargestellt werden kann.

§ 8. Einzeimatrizen. Als Beispiel für die im vorstehenden entwickelten Begriffsbildungen betrachten wir eine besondere Klasse

von hermiteschen Matrizen, die *Einzelmatrizen*¹; unter ihnen kommen die Nullmatrix und die Einheitsmatrix als Sonderfälle vor, und sie haben sämtlich mit diesen die Eigenschaft gemein, daß sie mit ihrem Quadrat übereinstimmen. Wir definieren also eine *Einzelmatrix* F durch die Forderungen

$$(1) \quad F^\dagger = F, \quad F^2 - F = F(1 - F) = 0.$$

Allgemeinere Beispiele als $F = 0$ und $F = E$ erhält man, indem man Diagonalmatrizen bildet, deren Diagonalelemente teilweise gleich 1 und teilweise gleich 0 sind. Diese besitzen offenbar keine Reziproke, außer im Falle der Einheitsmatrix, wo keine Null in der Diagonale vorkommt. Das gilt allgemein; denn wenn es zu einer Einzelmatrix F_1 eine andere Matrix A gibt, so daß

$$F_1 A = E$$

ist, so folgt

$$F_1 = F_1 E = F_1^2 A = F_1 A = E.$$

Dies Resultat formulieren wir als ersten einer Reihe einfacher Sätze über Einzelmatrizen:

α) Wenn eine Einzelmatrix eine Reziproke hat, so ist sie die Einheitsmatrix.

β) Mit F ist auch $1 - F$ eine Einzelmatrix und umgekehrt.

Denn es wird ja

$$(1 - F)^2 = 1 - 2F + F^2 = 1 - F$$

und andererseits

$$1 - (1 - F) = F.$$

γ) Das Produkt zweier Einzelmatrizen $F_1 F_2$ ist dann und nur dann wieder eine Einzelmatrix, wenn F_1, F_2 vertauschbar sind.

Denn erstens ist die Vertauschbarkeit nach § 6, (6) notwendig und hinreichend bereits dafür, daß $F_1 F_2$ hermitisch ist. Zweitens hat man dann

$$(F_1 F_2)^2 = F_1 F_2 F_1 F_2 = F_1^2 F_2^2 = F_1 F_2.$$

δ) Die Summe $F_1 + F_2$ ist dann und nur dann wieder eine Einzelmatrix, wenn $F_1 F_2 = 0$ ist (natürlich auch: wenn $F_2 F_1 = 0$ ist).

¹ Von den Mathematikern auch als „idempotente“ (hermitesche) Matrizen bezeichnet.

Aus $F_1 F_2 = 0$ folgt nämlich zunächst, daß auch $(F_1 F_2)^2 = 0$, also $F_1 F_2$ eine Einzelmatrix ist. Nach γ) ist $F_2 F_1 = F_1 F_2 = 0$. Somit wird daher

$$(F_1 + F_2)^2 = F_1^2 + F_2^2 + F_1 F_2 + F_2 F_1 = F_1 + F_2.$$

Umgekehrt folgt daraus, daß F_1, F_2 sowie $F_1 + F_2$ Einzelmatrizen sind, offenbar nach derselben Formel

$$F_1 F_2 + F_2 F_1 = 0.$$

Multipliziert man dies hinten mit F_2 , so erhält man

$$F_1 F_2 + F_2 F_1 F_2 = 0$$

und daraus durch vordere Multiplikation mit F_2

$$F_2 F_1 F_2 + F_2 F_1 F_2 = 2 F_2 F_1 F_2 = 0.$$

Dies in die vorhergehende Gleichung eingesetzt, liefert

$$F_1 F_2 = 0, \quad \text{w. z. b. w.}$$

ϵ) Die Differenz $F_1 - F_2$ ist dann und nur dann wieder eine Einzelmatrix, wenn $F_1 F_2 = F_2$ (oder, was nach γ) dasselbe bedeutet, $F_2 F_1 = F_2$) ist.

Nach β) ist nämlich $F_1 - F_2$ dann und nur dann Einzelmatrix, wenn $1 - (F_1 - F_2) = (1 - F_1) + F_2$ es ist; also ist nach δ) notwendig und hinreichend, daß $(1 - F_1) F_2 = 0$ oder $F_1 F_2 = F_2$.

§ 9. Vektoren im HILBERT-Raum. Wir haben die Matrizen als Repräsentanten physikalischer Größen eingeführt. Nun liegt die Frage nahe, ob diese Beziehung eindeutig ist. In den ursprünglichen HEISENBERGSchen Aufstellungen wurde in der Tat einer bestimmten physikalischen Größe nur eine, bis auf Phasenkonstanten [vgl. § 3, (II)] fest bestimmte Matrix zugeordnet. In späteren Entwicklungen (Kap. VI) wird sich jedoch eine etwas andere Auffassung bewähren: Jeder Größe ist nicht *eine* Matrix zugeordnet, sondern ein System von Matrizen, die auseinander durch „Transformation“ hervorgehen. Man übersieht diese Transformationstheorie, deren mathematischen Teil wir hier zu entwickeln haben¹, am besten im Zusammenhange einer allgemeinen *Vektor- und Tensorrechnung*, ähnlich der, die in der Relativitätstheorie gebraucht wird.

¹ Unter Beschränkung auf Größen mit diskreten Eigenwerten. Vgl. die diesbezüglichen Bemerkungen § 4.

Wir denken uns einen Raum von zunächst endlich viel, etwa N Dimensionen (später werden wir die Betrachtungen auf unendlich viele Dimensionen übertragen). In diesem Raume soll ein Punkt bezüglich eines bestimmten Koordinatensystems Σ durch eine Reihe komplexer Zahlen x_1, x_2, \dots, x_N als Koordinaten gegeben sein. Statt von einem „Punkt“ werden wir meist von einem „Vektor“ x mit den Komponenten x_1, x_2, \dots sprechen.

Wir nennen den betrachteten Raum einen *affinen* Vektorraum, wenn die Einführung eines neuen Koordinatensystems Σ' durch lineare Gleichungen der Form

$$(1) \quad x_n = \sum_k A_{nk} x'_k$$

erfolgt; dabei sind die A_{nk} die Elemente einer Matrix A , die eine *Reziproke* A^{-1} besitzt. Statt (1) schreiben wir auch kurz

$$(2) \quad x = A x'.$$

Der Vektor x im affinen Raum ist also zu definieren als der *Inbegriff* der Komponentenreihen x_1, x_2, \dots , die durch die Transformationen (1) ineinander übergehen.

Hat man ein weiteres Koordinatensystem Σ'' , das mit Σ' durch die Transformation

$$(3) \quad x'_k = \sum_m B_{km} x''_m \quad \text{oder} \quad x' = B x''$$

verknüpft ist, so erhält man durch Elimination von x' die Verknüpfung von Σ mit Σ'' :

$$(4) \quad x_n = \sum_{km} A_{nk} B_{km} x''_m = \sum_m (AB)_{nm} x''_m$$

oder

$$(5) \quad x = A B x''.$$

Die Matrixmultiplikation erscheint hier in natürlichster Weise als Zusammensetzung von Transformationsmatrizen. Doch werden wir sogleich noch andere Auffassungsweisen betrachten. Zunächst kann man die Gleichung (2) anders deuten, nicht als Koordinatentransformation desselben Vektors, sondern als *affine Abbildung des Vektorraums auf sich selbst*. Man wird dann besser schreiben

$$(6) \quad y = A x,$$

wo jetzt x und y *verschiedene* Vektoren sind. Die Matrix A spielt dann die Rolle eines *Operators*, der aus dem Vektor x den Vektor y

erzeugt. Dann erscheint das Produkt AB als Aufeinanderfolge der Operationen: erst B , dann A , angewandt auf x . Diese Auffassung wird in einem späteren Teil des Buchs bevorzugt werden.

Hier werden wir eine dritte Deutung ins Auge fassen, die mit der Tensorrechnung zusammenhängt.

Wir spezialisieren den affinen Vektorraum durch Einführung einer *Metrik*, und zwar ist es eine der euklidischen Metrik nahe verwandte, aber dadurch unterschieden, daß man bei der Definition der Länge eines Vektors der komplexen Natur der Komponenten Rechnung tragen muß.

Als Vorbereitung führen wir einige gebräuchliche Begriffe ein.

Die *Summe* zweier Vektoren x und y ist der Vektor $x + y$ mit den Komponenten $x_n + y_n$. Diese Addition ist kommutativ und assoziativ. Das *skalare Produkt* zweier Vektoren x und y ist

$$(9) \quad (x, y) = \sum_n x_n y_n^*.$$

Außer wenn (x, y) reell ist, hängt der Wert von (x, y) von der *Reihenfolge* ab:

$$(8) \quad (y, x) = \sum_n y_n x_n^* = (x^*, y^*) = (x, y)^*.$$

Zwischen beiden Operationen gilt das distributive Gesetz

$$(x + y, z) = (x, z) + (y, z).$$

Wenn

$$(9) \quad (x, y) = 0$$

ist, so sagen wir, die Vektoren x und y seien *senkrecht*, *normal* oder *orthogonal* zueinander.

Die Gesamtheit aller auf einem festen Vektor a normalen Vektoren x , für die also $(a, x) = 0$ ist, erfüllen einen *linearen Unterraum* (*Hyperebene*) des gegebenen Vektorraumes; hat dieser endlich viele, etwa N Dimensionen, so hat der Unterraum $(a, x) = 0$ $N - 1$ Dimensionen. Durch eine Reihe von Gleichungen

$$(a_1, x) = 0, (a_2, x) = 0, \dots (a_n, x) = 0$$

wird, wenn diese linear unabhängig sind, ein Unterraum von $N - n$ Dimensionen festgelegt.

Das skalare Produkt eines Vektors mit sich selbst ist

$$(10) \quad (x, x) = \sum_n x_n x_n^* = \sum_n |x_n|^2 \geq 0;$$

man nennt die reelle Größe

$$(11) \quad |x| = \sqrt{(x, x)}$$

die *Länge des Vektors* x .

Ersetzt man in (x, y) den Vektor x durch Ax , wo A irgend eine Matrix bedeutet, so wird

$$(12) \quad (Ax, y) = \sum_{mn} A_{mn} x_n y_m^* = (x, A^\dagger y)$$

eine zur Matrix A gehörige *Bilinearform*. Wenn A hermitisch, also $A = A^\dagger$ ist, hat die Form für $y = x$ reelle Werte; denn es ist nach (8) und (12)

$$(13) \quad (Ax, x)^* = (x, Ax) = (x, A^\dagger x) = (Ax, x).$$

Solche *hermitische Formen*

$$(14) \quad (Ax, x) = \sum_{mn} A_{mn} x_m^* x_n$$

verhalten sich in ihren Haupteigenschaften wesentlich wie reelle quadratische Formen. Man kann sie daher, wie diese, durch eine Fläche zweiter Ordnung veranschaulichen, deren Gleichung $(Ax, x) = 1$ lautet. Nur muß man mit Rücksicht auf die komplexe Natur der Koeffizienten und Variablen bei dem Gebrauch der geometrischen Ausdrucksweise eine gewisse Vorsicht walten lassen.

Transformiert man eine hermitische Form (Ax, x) auf neue Variable durch $x = Sx'$, so erhält man nach (13)

$$(15) \quad (Ax, x) = (ASx', Sx') = (S^\dagger ASx', x') = (Bx', x'),$$

wo also

$$(16) \quad B = S^\dagger AS$$

ist; diese Matrix ist nach § 6, (9a) wieder hermitisch:

$$(17) \quad B^\dagger = B.$$

Der Einheitsmatrix entspricht als hermitische Form das skalare Produkt eines Vektors mit sich selbst, oder die *Einheitsform* (10).

Durch die Transformation $x = Sx'$ geht die Länge (11) nach (15) über in

$$\sqrt{(S^\dagger Sx'; x')}.$$

Umkehrbare Transformationen des Vektorraums bzw. Koordinatentransformationen $x = Ux'$, die die *Länge invariant lassen*,

heißen *hermitisch orthogonal* oder *unitär*; für sie gilt also

$$(18) \quad (x, x) = (U^\dagger U x', x') = (x', x'),$$

d. h. es muß

$$(19) \quad U^\dagger U = 1$$

sein. Dies hat weiter

$$(19a) \quad U U^\dagger = 1$$

zur Folge; man multipliziere nämlich (19) links mit U und dividiere rechts durch U , was wegen der vorausgesetzten Umkehrbarkeit statthaft ist. *Die Matrix U ist also in dem durch § 6, (14) definierten Sinne unitär.*

Ist U eine endliche Matrix, so braucht man zur Ableitung von (19a) aus (19) die Umkehrbarkeit nicht besonders voraussetzen [vgl. die Bemerkungen zu § 4, (19)]; bei unendlichen Matrizen aber muß die Umkehrbarkeit besonders vorausgesetzt werden, damit aus (19) auch (19a) folgt.

Als Beispiel einer unitären Transformation $x = U x'$ erwähnen wir die *Permutation der Variablen*; wählt man nämlich für U die in § 6, S. 32 definierte *Permutationsmatrix*, so sind die Komponenten x_n bis auf die Reihenfolge mit den x'_n identisch.

Nunmehr sind wir in der Lage, die Metrik zu definieren. *Wir fordern, daß die Länge eines Vektors $\sqrt{(x, x)}$ invariant ist. Die zugelassenen Koordinatentransformationen sind infolgedessen die unitären.* Der so definierte Raum soll *HILBERT-Raum* genannt werden¹.

Zum Unterschiede von Vektoren und Tensoren im gewöhnlichen 3-dimensionalen Raum, wie sie in der Physik auftreten und die wir *Raumvektoren* und *Raumtensoren* nennen, sprechen wir im vieldimensionalen (bzw. ∞ -dimensionalen) *HILBERT-Raum* von *HILBERT-Vektoren* und *HILBERT-Tensoren*. Letztere werden wir im folgenden Paragraphen untersuchen.

Die *Achsen* eines Koordinatensystems im *HILBERT-Raum* sind die Vektoren

$$(20) \quad e_1 = (1, 0, 0, \dots 0), \quad e_2 = (0, 1, 0, \dots 0), \quad \dots \\ e_N = (0, 0, \dots 1).$$

¹ HILBERT hat nämlich in der Theorie der Integralgleichungen den euklidischen Raum von ∞ vielen Dimensionen mit großem Erfolge benutzt.

Sie bilden ein *vollständiges (normiertes) Orthogonalsystem*, d. h. es gilt

$$(21) \quad (e_k, e_l) = 0, \quad |e_k|^2 = 1$$

und *jeder* Vektor x (x_1, \dots, x_N) ist linear durch die Achsenvektoren darstellbar:

$$(22) \quad x = x_1 e_1 + x_2 e_2 + \dots + x_N e_N.$$

Außer diesem gibt es noch unendlich viele vollständige Orthogonalsysteme im HILBERT-Raum. Es gilt der Satz:

Im Raume von N Dimensionen seien n ($n \leq N$) linear unabhängige Vektoren x_1, x_2, \dots, x_n gegeben. Dann kann man immer n linear unabhängige, lineare Kombinationen y_1, \dots, y_n von x_1, \dots, x_n wählen, die normiert und aufeinander orthogonal sind, d. h. den Gleichungen

$$(23) \quad (y_k, y_l) = \delta_{kl}, \quad k, l = 1, \dots, n$$

genügen. Dazu bilde man aus den x_k der Reihe nach¹

$$(24) \quad \left\{ \begin{array}{ll} z_1 = x_1, & y_1 = \frac{z_1}{\sqrt{(z_1, z_1)}}, \\ z_2 = x_2 - (x_2, y_1) y_1, & y_2 = \frac{z_2}{\sqrt{(z_2, z_2)}}, \\ z_3 = x_3 - (x_3, y_1) y_1 - (x_3, y_2) y_2, & y_3 = \frac{z_3}{\sqrt{(z_3, z_3)}}, \\ \dots & \dots \end{array} \right.$$

Keiner der Vektoren z_k , also auch keiner der y_k ($k = 1, 2, \dots, n$) kann identisch verschwinden, weil sonst eine lineare Beziehung zwischen den x_k bestände, gegen die Voraussetzung. Ferner sieht man, daß jedes z_k auf allen vorausgehenden y_l orthogonal ist; dasselbe gilt also auch für y_k , das überdies normiert ist.

Wenn $n = N$ ist, kann man keinen weiteren, von Null verschiedenen Vektor y finden, der zu allen y_1, y_2, \dots, y_N orthogonal wäre. Denn y müßte (weil N die Dimensionszahl ist) linear von den y_1, \dots, y_N abhängen, etwa $y = \sum_{k=1}^N c_k y_k$, und aus $(y, y_l) = 0$

¹ Das Verfahren stammt von E. SCHMIDT: Math. Ann. Bd. 63, S. 433. 1907.

folgt dann, da $(y_k, y_l) = \delta_{kl}$ ist: $c_k = 0$, $k = 1, 2, \dots, N$. Das System y_1, y_2, \dots, y_N ist daher ein (normiertes) *vollständiges Orthogonalsystem*.

Jedes vollständige Orthogonalsystem y_1, \dots, y_N kann man durch eine unitäre Transformation („Koordinatentransformation“) in das System e_1, e_2, \dots, e_N überführen.

Es sind jedenfalls alle y_k linear von den e_k abhängig und umgekehrt; man hat daher eine umkehrbare lineare Transformation

$$y_k = U e_k,$$

bei der die l -te Komponente des k -ten Vektors durch

$$y_{kl} = \sum_j U_{lj} \delta_{jk} = U_{lk}$$

gegeben ist. Die Komponenten von y_k bilden also die k -te Spalte der Matrix U . Ferner ist

$$(U^\dagger U)_{lk} = \sum_j U_{lj}^\dagger U_{jk} = \sum_j U_{jl}^* U_{jk} = \sum_j y_{kj} y_{lj}^* = (y_k, y_l) = \delta_{kl};$$

da U eine Reziproke hat, ist sie also nach § 6, (14) unitär.

Demnach können zwei vollständige Orthogonalsysteme $y_k = U e_k$, $y'_k = U' e_k$ durch die unitäre Transformation $U U'^{-1}$ ineinander übergeführt werden.

Die vollständigen Orthogonalsysteme sind also mit den Achsen aller erlaubten Koordinatensysteme im HILBERT-Raume identisch.

Wir leiten einen später viel benutzten Satz über einparametrische Scharen von Orthogonalsystemen ab.

Sei $H(t)$ eine hermitesche, von dem Parameter t (den wir Zeit nennen wollen) abhängige Matrix, und \varkappa eine rein-imaginäre Zahl. Ein Vektor $x(t)$ genüge der Differentialgleichung

$$(25) \quad \varkappa \dot{x} + Hx = 0.$$

Wir behaupten zunächst, daß das skalare Produkt irgend zweier Lösungen $x_1(t)$, $x_2(t)$ von t unabhängig ist.

Multipliziert man nämlich

$$\varkappa \dot{x}_1 + Hx_1 = 0, \quad \varkappa \dot{x}_2 + Hx_2 = 0$$

skalar mit x_2 bzw. x_1 und subtrahiert, so kommt

$$(\varkappa \dot{x}_1, x_2) - (x_1, \varkappa \dot{x}_2) + (Hx_1, x_2) - (x_1, Hx_2) = 0.$$

Nun ist nach (7)

$$(\varkappa \dot{x}_1, x_2) = \varkappa (\dot{x}_1, x_2), \quad (x_1, \varkappa \dot{x}_2) = -\varkappa (x_1, \dot{x}_2),$$

und wegen $H = H^\dagger$

$$(x_1, H x_2) = (H x_1, x_2).$$

Also folgt

$$(\dot{x}_1, x_2) + (x_1, \dot{x}_2) = \frac{d}{dt} (x_1, x_2) = 0,$$

$$(26) \quad (x_1, x_2) = \text{konst.}$$

Wenn also x_1, x_2 für einen bestimmten Wert von t orthogonal sind, so bleiben sie immer orthogonal. Daraus folgt der Satz:

Bilden die Lösungsvektoren x_1, x_2, \dots der Gleichung (25) zu einer bestimmten Zeit t_0 ein Orthogonalsystem, so bilden sie zu jeder Zeit t ein Orthogonalsystem. Gibt man zur Zeit t_0 ein vollständiges Orthogonalsystem x_1^0, x_2^0, \dots und bestimmt die Lösung $x_k(t)$, für die $x_k(t_0) = x_k^0$ wird, so bilden die Vektoren $x_1(t), x_2(t), \dots$ für beliebige Zeiten aus Stetigkeitsgründen ebenfalls ein *vollständiges* Orthogonalsystem¹. Nach dem oben Bewiesenen gibt es also eine unitäre Matrix $U(t)$, so daß

$$(27) \quad x_k(t) = U(t) x_k^0$$

ist.

Wir fügen noch eine Bemerkung an, die sich an die Eigenschaft der unitären Transformationen anknüpft, längentreu zu sein. Wir haben in § 6, (11) die *kontragredienten Matrizen* formal definiert. Man kann sie aber auch durch eine geometrische Forderung ähnlicher Art wie die Invarianz der Länge kennzeichnen. Dazu betrachtet man Abbildungen des Vektorraums auf sich selbst und stellt die Forderung, die simultane Transformation zweier Vektoren x und y zu finden, durch welche das *skalare Produkt* nicht geändert wird. Sei also

$$(28) \quad X = A x, \quad Y = \check{A} y$$

und man fordere

$$(29) \quad (X, Y) = (A x, \check{A} y) = (x, y),$$

¹ Vgl. auch Kap. III, § 22.

so schließt man zunächst auf

$$(30) \quad (x, A^\dagger \check{A} y) = (x, y)$$

und daraus auf

$$(31) \quad A^\dagger \check{A} = 1 \quad \text{oder} \quad A = (\check{A}^\dagger)^{-1}$$

in Übereinstimmung mit der in § 6, (11) gegebenen Definition.

§ 10. HILBERT-Tensoren und metrische Eigenschaften des HILBERT-Raums. Die Tensorrechnung im HILBERT-Raum ist der gewöhnlichen im euklidischen Raume sehr ähnlich, nur muß man darauf achten, daß alle Definitionen mit dem Auftreten komplexer Zahlen in der Längendefinition im Einklang sind.

Während sich ein HILBERT-Vektor mit den Komponenten x_n gemäß

$$(1) \quad x_n = \sum_k U_{nk} x'_k$$

transformiert, so transformiert sich das System der komplex konjugierten Werte x_n^* der x_n gemäß

$$(1') \quad x_n^* = \sum_l U_{nl}^* x_l'^* = \sum_l x_l'^* U_{ln}^{-1}.$$

Die Analogie zur gewöhnlichen Tensorrechnung tritt am deutlichsten zutage, wenn wir die x_n fernerhin als *kovariante Komponenten*, die sich gemäß (1') transformierenden komplex konjugierten x_n^* dagegen als *kontravariante* Komponenten des Vektors bezeichnen. Statt x_n^* wollen wir demgemäß auch $x_{\bar{n}}$ schreiben.

Wir definieren allgemein einen *gemischten Tensor* ($r + s$)-ter Stufe: Er ist in einem bestimmten Koordinatensystem des HILBERT-Raumes darzustellen durch das Schema seiner *r-fach kovarianten und s-fach kontravarianten Komponenten*

$$A_{n_1 n_2 \dots n_r \bar{m}_1 \bar{m}_2 \dots \bar{m}_s},$$

die sich bei Änderung des Koordinatensystems transformieren wie die Produkte

$$x_{n_1}^{(1)} x_{n_2}^{(2)} \dots x_{n_r}^{(r)} y_{\bar{m}_1}^{(1)} y_{\bar{m}_2}^{(2)} \dots y_{\bar{m}_s}^{(s)}$$

der kovarianten Komponenten

$$x_{n_1}^{(1)}, x_{n_2}^{(2)}, \dots, x_{n_r}^{(r)}$$

von r Vektoren und der kontravarianten Komponenten

$$y_{\underline{m}_1}^{(1)}, y_{\underline{m}_2}^{(2)}, \dots, y_{\underline{m}_s}^{(s)}$$

von s weiteren Vektoren. D. h. also, es gilt:

$$(2) \quad \left\{ \begin{array}{c} A_{n_1 \dots n_r \underline{m}_1 \dots \underline{m}_s} \\ = \sum_{\substack{k_1 \dots k_r \\ l_1 \dots l_s}} U_{n_1 k_1 \dots} U_{n_r k_r} A'_{k_1 \dots k_r \underline{l}_1 \dots \underline{l}_s} U_{l_1 \underline{m}_1}^{-1} \dots U_{l_s \underline{m}_s}^{-1} \end{array} \right.$$

Derselbe Tensor kann auch durch seine *adjungierten* (s -fach kovarianten und r -fach kontravarianten) Komponenten

$$(3) \quad A_{\underline{m}_1 \underline{m}_2 \dots \underline{m}_s n_1 n_2 \dots n_r}^\dagger = A_{n_1 n_2 \dots n_r \underline{m}_1 \underline{m}_2 \dots \underline{m}_s}^*$$

dargestellt werden.

Eine auf Tensorkomponenten angewandte Operation heißt *kovariant*, wenn die entstehenden Größen wieder die Komponenten eines Tensors sind. Es gibt drei elementare kovariante Operationen:

Die *Addition* gleichartiger Tensoren:

$$(4) \quad C_{n_1 \dots n_r \underline{m}_1 \dots \underline{m}_s} = A_{n_1 \dots n_r \underline{m}_1 \dots \underline{m}_s} + B_{n_1 \dots n_r \underline{m}_1 \dots \underline{m}_s}.$$

Die (*direkte*) *Multiplikation* beliebiger Tensoren:

$$(5) \quad C_{n_1 \dots n_r \underline{v}_1 \dots \underline{v}_\rho \underline{m}_1 \dots \underline{m}_s \underline{\mu}_1 \dots \underline{\mu}_\sigma} = A_{n_1 \dots n_r \underline{m}_1 \dots \underline{m}_s} B_{\underline{v}_1 \dots \underline{v}_\rho \underline{\mu}_1 \dots \underline{\mu}_\sigma}.$$

Die *Verjüngung* gemischter Tensoren; d. h. Summation über ein Paargleich gesetzter Indizes, wo ein Index kovariant (nicht unterstrichen) und ein Index kontravariant (unterstrichen) ist. Also etwa:

$$(6) \quad C_{n_2 \dots n_r \underline{m}_2 \dots \underline{m}_s} = \sum_{n_1 = \underline{m}_1} A_{n_1 n_2 \dots n_r \underline{m}_1 \underline{m}_2 \dots \underline{m}_s}.$$

Der kovariante Charakter dieser Bildungen leuchtet unmittelbar ein.

Die Tensoren erster Stufe sind die kovarianten bzw. kontravarianten Vektoren. Die *gemischten Tensoren zweiter Stufe*, also diejenigen mit einem kovarianten und einem kontravarianten Index, entsprechen den *Matrizen*: Die Komponenten $A_{n \underline{m}}$ des Tensors in einem bestimmten Koordinatensystem bilden eine Matrix $A = (A_{n \underline{m}})$, die bei unitären Transformationen des HILBERT-Raums zufolge (2) dem Transformationsgesetz

$$(7) \quad A = U A' U^{-1}$$

unterworfen ist. Das Schema der adjungierten Komponenten liefert die zu A adjungierte Matrix A^\dagger :

$$A_{\underline{m} \underline{n}}^\dagger = A_{\underline{n} \underline{m}}^*.$$

Durch Verjüngung eines gemischten Tensors zweiter Stufe erhält man die invariante *Spur* der zugehörigen Matrizen A :

$$Sp A = \sum_{\underline{k}} A_{\underline{k} \underline{k}}.$$

Aus zwei gemischten Tensoren zweiter Stufe, die in einem gewissen Koordinatensystem durch die Matrizen A , B dargestellt sein mögen, erhält man durch direkte Multiplikation und nachfolgende Verjüngung einen neuen gemischten Tensor zweiter Stufe. Dabei gibt es zwei Möglichkeiten (Verjüngung nach dem einen oder dem andern Indexpaar); die Matrix des neuen Tensors ist das Produkt AB oder BA der ursprünglichen Matrizen:

$$(8) \quad (A B)_{\underline{n} \underline{m}} = \sum_{\underline{k}} A_{\underline{n} \underline{k}} B_{\underline{k} \underline{m}}; \quad (B A)_{\underline{n} \underline{m}} = \sum_{\underline{k}} A_{\underline{k} \underline{m}} B_{\underline{n} \underline{k}}.$$

Alle Sätze des § 7 über *Matrixfunktionen* bleiben infolge des Satzes 1, § 7 für die entsprechenden Tensorfunktionen bestehen. Insbesondere ist z. B. die Definitionsgleichung der Einzelmatrizen $F^2 = F$ kovariant, und da auch die Bedingung der Hermitizität kovariant ist, besitzen die durch $F^2 = F = F^\dagger$ definierten *Einzeltensoren* in jedem Koordinatensystem eine Einzelmatrix als Komponentenschema. Insbesondere ist das Komponentenschema des *invarianten Einheitstensors* in jedem Koordinatensystem gleich der *Einheitsmatrix*.

Die *Determinante* einer Matrix ist (falls sie existiert) eine Invariante des zugehörigen Tensors, da

$$\text{Det}(U^{-1} A U) = \text{Det}(U^{-1}) \text{Det}(A) \text{Det}(U) = \text{Det}(A).$$

Durch *direkte Multiplikation eines kovarianten und eines kontravarianten Vektors* erhält man einen gemischten Tensor zweiter Stufe

$$(9) \quad C_{\underline{n} \underline{m}} = x_{\underline{n}} y_{\underline{m}} = x_{\underline{n}} y_{\underline{m}}^*.$$

Wir schreiben symbolisch

$$(9a) \quad C = x \times y.$$

Nachfolgende Verjüngung liefert das skalare Produkt (x, y) . Durch Multiplikation eines gemischten Tensors zweiter Stufe mit einem Vektor und nachfolgende Verjüngung ergibt sich ein neuer Vektor Ax mit den Komponenten:

$$(Ax)_n = \sum_k A_{nk} x_k.$$

Die Bilinearform endlich

$$(Ax, y) = \sum_{kl} A_{kl} x_l y_k = \sum_{kl} A_{kl} x_l y_k^*$$

entsteht z. B. durch zweimalige Verjüngung des direkten Produktes von A mit dem Tensor C in (9).

Wir wollen jetzt einige metrische Eigenschaften des HILBERT-Raums ins Auge fassen.

Das skalare Produkt erfüllt folgende *Ungleichungen*:

$$(10) \quad |(x, y)| \leq \sqrt{(x, x) \cdot (y, y)} = |x| \cdot |y|,$$

$$(11) \quad |x + y| \leq |x| + |y|,$$

von denen die erste die *SCHWARZsche Ungleichung* genannt wird. Um sie zu beweisen, bilde man mit Hilfe zweier reeller Zahlen α, β

$$|\alpha x + \beta y|^2 = \alpha^2 (x, x) + \alpha \beta \{ (x, y) + (y, x) \} + \beta^2 (y, y).$$

Die linke Seite ist ≥ 0 , also muß die Diskriminante der rechtsstehenden quadratischen Form

$$(x, x) \cdot (y, y) - \frac{1}{4} \{ (x, y) + (y, x) \}^2 \geq 0$$

sein, oder mit $(x, y) = \xi + i\eta$, $(y, x) = \xi - i\eta$:

$$\xi \leq |x| \cdot |y|.$$

Ersetzen wir hier x durch $e^{i\varphi} x$, wo φ reell ist, so bleibt die rechte Seite ungeändert, während $\xi = \frac{1}{2} \{ (x, y) + (y, x) \}$ in

$$\frac{1}{2} \{ e^{i\varphi} (x, y) + e^{-i\varphi} (y, x) \} = \xi \cos \varphi - \eta \sin \varphi$$

übergeht. Also gilt für alle φ

$$\xi \cos \varphi - \eta \sin \varphi \leq |x| \cdot |y|.$$

Das Maximum der linken Seite ist

$$\sqrt{\xi^2 + \eta^2} = |(x, y)|,$$

also gilt in der Tat die Ungleichung (10). Ferner haben wir

$$\begin{aligned} (x + y, x + y) &= (x, x) + (x, y) + (y, x) + (y, y) \\ &= (x, x) + 2\xi + (y, y) \\ &\leq (x, x) + 2\sqrt{(x, x)(y, y)} + (y, y) \\ &= \left\{ \sqrt{(x, x)} + \sqrt{(y, y)} \right\}^2 \end{aligned}$$

und daraus folgt (11) durch Wurzelziehen.

Eine hermitesche Form und die zugehörige Matrix A heißen *positiv definit* (kurz: *definit*), wenn die Form für alle Vektoren x von endlicher Länge, etwa der Länge 1, nur positive Werte annimmt:

$$(12) \quad (Ax, x) > 0 \quad \text{für} \quad (x, x) = 1.$$

Läßt man auch den Wert Null für die Form zu, so nennt man sie *semidefinit*. *Die Einheitsform ist trivialerweise positiv definit. Die Summe zweier definiten Formen ist wieder definit.* Ferner gilt: *Ist A eine beliebige Matrix, so ist AA^\dagger semidefinit,* denn

$$(AA^\dagger x, x) = (A^\dagger x, A^\dagger x) \geq 0.$$

Daraus folgt weiter, daß $AA^\dagger = 0$ nur für $A = 0$ stattfindet; denn für $AA^\dagger = 0$ gilt identisch in x

$$0 = (AA^\dagger x, x) = (A^\dagger x, A^\dagger x),$$

also ist $A^\dagger x = 0$ identisch in x , was $A^\dagger = 0$, mithin $A = 0$, zur Folge hat.

Ist A eine hermitesche semidefinite Matrix, so gilt für irgend zwei Vektoren x, y

$$(13) \quad |(x, Ay)| \leq \sqrt{(x, Ax)(y, Ay)}.$$

Der Beweis verläuft ganz analog dem der Formel (10). Man bilde die quadratische Form

$$\begin{aligned} (\alpha x + \beta y, A(\alpha x + \beta y)) &= \alpha^2(x, Ax) + \alpha\beta\{(x, Ay) + (y, Ax)\} \\ &\quad + \beta^2(y, Ay) \end{aligned}$$

mit irgend zwei reellen Zahlen α, β ; diese ist ≥ 0 , wenn A semidefinit ist. Folglich ist die Diskriminante

$$(x, Ax) \cdot (y, Ay) - \frac{1}{4} \{(x, Ay) + (y, Ax)\}^2 \geq 0.$$

Setzt man nun $(x, Ay) = \xi + i\eta$, so ist wegen $A = A^\dagger$

$$(y, Ax) = (Ay, x) = (x, Ay)^* = \xi - i\eta;$$

also hat man

$$\xi^2 \leq (x, Ax) \cdot (y, Ay).$$

Ersetzt man nun wieder x durch $e^{i\varphi}x$ mit reellem φ , so erhält man wie oben, S. 52,

$$\xi^2 + \eta^2 = |(x, Ay)|^2 \leq (x, Ax) \cdot (y, Ay)$$

und daraus (13).

Als Korollar folgt hieraus:

Wenn A eine hermitesche, semidefinite Matrix ist und für einen Vektor $Ay \neq 0$ ist, so ist

$$(14) \quad (y, Ay) > 0.$$

Es ist nämlich jedenfalls $(y, Ay) \geq 0$; wäre aber $(y, Ay) = 0$, so würde aus (13) die Beziehung $(x, Ay) = 0$ für jeden Vektor x folgen, also $Ay = 0$ gegen die Voraussetzung.

§ 11. Hauptachsentransformation der hermiteschen Formen.

Die Analogie der hermiteschen Formen mit den reellen quadratischen Formen führt zu der Frage, ob es etwas der Hauptachsentransformation der Flächen 2. Ordnung Entsprechendes auch bei den hermiteschen Formen gibt. Während der Übergang von einem reellen Koordinatensystem zu einem andern, dagegen gedrehten durch eine reelle orthogonale Transformation vermittelt wird, werden wir hier mit unitären Transformationen zu tun haben. Man wird also zu der Frage geführt: Gibt es eine unitäre Transformation $x = Uy$, durch die die hermitesche Form (Ax, x) in die Gestalt (ay, y) gebracht wird, wo a eine Diagonalmatrix bedeutet? Oder geometrisch: Gibt es zu einem hermiteschen Tensor 2. Stufe A ein Koordinatensystem, in dem die darstellende Matrix Diagonalf orm hat? Nach § 10, (7) hat dann U den Gleichungen

$$(1) \quad U^\dagger = U^{-1}, \quad U^{-1}AU = a$$

zu genügen, wo $a = (a_n \delta_{nm})$ ist.

Schreibt man die letztere

$$(2) \quad AU = Ua,$$

so bedeutet sie für die Elemente die Relation

$$(3) \quad \sum_k A_{nk} U_{km} = U_{nm} a_m.$$

Dies kann man auch als ein System von Vektorgleichungen lesen, indem man bemerkt, daß die Elemente U_{km} der m -ten Spalte von U die Komponenten eines Vektors U_m bilden, so daß die Gleichungen (3) sich in der Form

$$(4) \quad A U_m = a_m U_m \quad m = 1, 2, \dots$$

schreiben lassen. Transformiert man nämlich hierin A unitär auf ein andres Koordinatensystem mittels $A' = V^{-1} A V$, so geht (4) über in

$$V A' V^{-1} U_m = a_m U_m \quad \text{oder} \quad A' V^{-1} U_m = a_m V^{-1} U_m$$

und das hat wieder die Form (4), wenn man

$$U'_m = V^{-1} U_m \quad \text{oder} \quad U_m = V U'_m$$

setzt, d. h. U_m wie einen Vektor transformiert [s. § 9, (2)]. Alle diese Vektoren U_m sind offenbar wegen der Unitarität von U aufeinander normal. Die ursprüngliche Aufgabe ist also mit der Forderung äquivalent, eine Zahlenreihe a_1, a_2, \dots und ein System aufeinander normaler Vektoren U_1, U_2, \dots so zu bestimmen, daß die Vektorgleichungen (4) erfüllt sind. Diese haben die Form

$$(5) \quad A x = a x;$$

man hat zu untersuchen, ob es Zahlen a und Vektoren x gibt, die ihr genügen.

Im folgenden soll bis zu ausdrücklichem Widerruf angenommen werden, daß die Zahl der Dimensionen des betrachteten Vektorraumes endlich, gleich N ist. Dann bedeutet (5) für festes a ein System von N linearen, homogenen Gleichungen für die N Unbekannten x_1, x_2, \dots, x_N . Sie sind dann und nur dann lösbar, wenn die Determinante verschwindet:

$$(6) \quad \text{Det} (A - a E) = \begin{vmatrix} A_{11} - a & A_{12} & \dots & A_{1N} \\ A_{21} & A_{22} - a & \dots & A_{2N} \\ \dots & \dots & \dots & \dots \\ A_{N1} & A_{N2} & \dots & A_{NN} - a \end{vmatrix} = 0.$$

Das ist eine algebraische Gleichung N -ten Grades, die N Wurzeln $a = a_1, a_2, \dots, a_N$ hat, von denen auch einige gleich sein können; man nennt sie *Eigenwerte*, ihre Gesamtheit das *Spektrum* der Matrix bzw. Form A .

Die *Eigenwerte sind reell*; denn wenn a ein Eigenwert ist, dann ist (5) durch einen Vektor x erfüllt, und es folgt durch skalare Multiplikation mit x aus (5)

$$(7) \quad (Ax, x) = a(x, x).$$

a ist also als Quotient der beiden hermiteschen Formen (Ax, x) und (x, x) eine reelle Zahl.

Unterwirft man x der unitären Transformation $x = Uy$, so geht nach § 9, (16) A über in $B = U^\dagger AU = U^{-1}AU$; daher wird

$$\begin{aligned} \text{Det}(B - aE) &= \text{Det}(U^{-1}(A - aE)U) \\ &= \text{Det}(U^{-1}) \cdot \text{Det}(A - aE) \cdot \text{Det}U, \end{aligned}$$

also wegen $\text{Det}(U^{-1}) = (\text{Det}U)^{-1}$:

$$(8) \quad \text{Det}(B - aE) = \text{Det}(A - aE).$$

Die *Determinantengleichung (6) ist also unitär invariant, mithin auch ihre Koeffizienten und ihre Wurzeln: die Eigenwerte a_k , einschließlich ihrer Vielfachheit.*

Unter den Koeffizienten, den elementarsymmetrischen Funktionen der Wurzeln, findet sich auch, als Faktor von a^{N-1} , die Spur der Matrix A [s. § 4, (21)]:

$$(9) \quad \sum_k a_k = \sum_n A_{nn} = SpA.$$

Daß diese Größe invariant ist, ist ein Spezialfall des allgemeinen Satzes [§ 4, (23)], daß für beliebige Matrizen $Sp(B^{-1}AB) = SpA$ gilt.

Nun können wir den Hauptsatz formulieren und durch vollständige Induktion beweisen.

Satz 1: *Es gibt eine unitäre Transformation*

$$(10) \quad x = Uy, \quad y = U^{-1}x = U^\dagger x,$$

durch welche die zur Matrix A gehörige hermitesche Form auf eine Summe von Quadraten gebracht wird:

$$(11) \quad (Ax, x) = (ay, y), \quad U^{-1}AU = a,$$

oder ausführlich:

$$(11a) \quad \sum_{mn} A_{mn} x_m^* x_n = \sum_k a_k y_k y_k^*.$$

Das Koordinatensystem, in dem der Tensor A durch die Diagonalmatrix a dargestellt wird, heißt Hauptachsensystem.

Es sei a_1 ein Eigenwert und x_1 ein zugehöriger Lösungsvektor der Gleichung (5),

$$(12) \quad A x_1 = a_1 x_1, \quad x_1 \neq 0$$

den wir durch

$$(x_1, x_1) = 1$$

normieren. Den Einheitsvektor x_1 können wir (s. § 9) zu einem normierten vollständigen Orthogonalsystem x_1, x_2, \dots, x_n ergänzen. Es sei U die unitäre Transformation, welche diese Vektoren x_i in das spezielle Orthogonalsystem e_i [s. § 9, (20)] überführt:

$$(13) \quad x_i = U e_i.$$

Dann ist nach § 6, (c) auch

$$(14) \quad A' = U^{-1} A U$$

hermitisch, und aus (12) folgt durch vordere Multiplikation mit U^{-1} und Beachtung von (13) und (14)

$$A' e_1 = a_1 e_1.$$

Diese Gleichung besagt, daß die erste Zeile (und infolge der Hermitizität auch die erste Spalte) von A' so lautet:

$$a_1, 0, 0, \dots, 0.$$

Somit hat A' die Form (als Übermatrix geschrieben):

$$(15) \quad A' = \begin{pmatrix} a_1 & 0 \\ 0 & \bar{A} \end{pmatrix},$$

wo \bar{A} eine $(N - 1)$ -dimensionale hermitesche Matrix ist.

Nehmen wir nun an, daß der Satz 1 für $N - 1$ -dimensionale hermitesche Matrizen, also insbesondere für \bar{A} gilt. Dann gibt es eine unitäre Matrix \bar{V} derart, daß

$$(16) \quad \bar{V}^{-1} \bar{A} \bar{V} = \begin{pmatrix} a_2 & 0 & \dots & 0 \\ 0 & a_3 & \dots & 0 \\ \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & \dots & a_N \end{pmatrix}$$

ist. Setzen wir (als Übermatrix)

$$(17) \quad V = \begin{pmatrix} 1 & 0 \\ 0 & \bar{V} \end{pmatrix},$$

so ist nach (15) und (16)

$$(18) \quad V^{-1} A' V = (a_k \delta_{k\iota}).$$

Nach (14) ist aber

$$V^{-1} A' V = V^{-1} U^{-1} A U V = (U V)^{-1} A U V,$$

und mit U und \bar{V} ist auch UV unitär. Somit ist unter Voraussetzung der Gültigkeit des Satzes für $N-1$ eine unitäre Transformation gefunden, die die N -dimensionale hermitesche Matrix A auf die Diagonalform bringt. Da der Satz für $N=1$ trivial ist, so ist der Beweis erledigt.

Bezeichnen wir wieder die transformierende Matrix mit U , wie in (10), so kann man die Elemente U_{km} einer einzelnen Spalte m als Komponenten eines Vektors U_m auffassen. Diese Vektoren bilden offenbar ein vollständiges Orthogonalsystem und genügen den Gleichungen (4); sie heißen *Eigenvektoren* der hermiteschen Matrix A . Sind mehrere, etwa g_r Eigenwerte einander gleich, so sprechen wir auch von einem g_r -fachen *Eigenwert*; zu einem solchen gehören offenbar g_r *linear unabhängige Eigenvektoren*.

Die in Satz 1 gegebene Darstellung einer hermiteschen Form als Summe von Quadraten ist aber nicht vollständig eindeutig. Zwar sind die Eigenwerte a_k eindeutig bestimmt, da sie notwendig der Gleichung (6) genügen müssen; wenn es also zwei unitäre Transformationen $x = U_1 y$, $x = U_2 z$ gibt, die die Form (Ax, x) auf die Diagonalgestalt bringen, so muß

$$(19) \quad (ay, y) = (az, z)$$

sein. Doch brauchen dazu U_1 und U_2 nicht identisch zu sein. Wir wollen die noch bleibende Willkür feststellen.

Durch Elimination von x erhält man eine Transformation von y nach z , die wieder unitär ist:

$$(20) \quad y = \hat{U} z, \quad \hat{U} = U_1^{-1} U_2, \quad \hat{U} \hat{U}^\dagger = 1.$$

Setzt man das in (19) ein, so folgt nach § 10, (7)

$$(21) \quad \hat{U}^{-1} a \hat{U} = a$$

oder

$$a \overset{\circ}{U} - \overset{\circ}{U} a = 0.$$

Das gibt für die Elemente die Relation:

$$(a_k - a_l) \overset{\circ}{U}_{kl} = 0.$$

Sobald also zu den Indizes k, l verschiedene Eigenwerte gehören, ist $\overset{\circ}{U}_{kl} = 0$ [s. auch § 5, (13)]. Wir bezeichnen nun die nach der Größe geordneten, verschiedenen Eigenwerte mit einem oberen Index r :

$$(22) \quad a^{(1)} < a^{(2)} < \dots < a^{(M)} \quad M \leq N.$$

Die g_r Eigenwerte a_k , die gleich $a^{(r)}$ sind, unterscheiden wir durch einen zweiten Index $\rho = 1, 2, \dots, g_r$. Dann können wir das über die Matrix $\overset{\circ}{U}$ gewonnene Resultat so aussprechen: $\overset{\circ}{U}$ muß durch Aneinanderreihen von M je g_r -reihigen Matrizen $U^{(r)}$ längs der Diagonalen entstehen, also eine *unitäre Stufenmatrix* [s. § 5, (7)] sein:

$$(23) \quad \overset{\circ}{U} = \begin{pmatrix} U^{(1)} & 0 & \dots & 0 \\ 0 & U^{(2)} & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & U^{(M)} \end{pmatrix}.$$

Da $\overset{\circ}{U}$ unitär in N Dimensionen ist, sind es die $U^{(r)}$ in g_r Dimensionen.

Die unitäre Transformation (20) zerfällt also in M je g_r -dimensionale Transformationen, die man

$$(24) \quad y^{(r)} = U^{(r)} z^{(r)}$$

schreiben kann, wo die g_r -dimensionalen Vektoren $y^{(r)}$ die Komponenten $y_1^{(r)}, y_2^{(r)}, \dots, y_{\rho}^{(r)}, \dots, y_{g_r}^{(r)}$ haben. Daß umgekehrt die Form

$$(25) \quad \sum_{k=1}^N a_k y_k y_k^* = \sum_{r=1}^M a^{(r)} \sum_{\rho=1}^{g_r} y_{\rho}^{(r)} y_{\rho}^{(r)*}$$

gegen die Transformationen (24) invariant ist, leuchtet ein. Demnach gilt der

Satz 2: Die unitäre Stufenmatrix (23) stellt die Willkür dar, die bei der Transformation auf Hauptachsen noch übrig bleibt.

Um eine eindeutige Normalform zu erhalten, führen wir die *Einzelformen*

$$(26) \quad (E^{(r)} y, y) = \sum_{\rho=1}^{g_r} y_{\rho}^{(r)} y_{\rho}^{(r)*}$$

ein, so genannt, weil ihre Matrix $E^{(r)}$ eine (N -dimensionale) Einzelmatrix ist, deren Elemente sämtlich Null sind, außer den zu $a^{(r)}$ gehörigen Diagonalelementen, die gleich 1 sind; daher gilt

$$(27) \quad E^{(r)} E^{(s)} = 0 \quad \text{für} \quad r \neq s.$$

Nun schreibt sich die Form A nach (25)

$$(28) \quad (Ax, x) = \sum_r a^{(r)} (E^{(r)} y, y)$$

oder, wenn man mittels $y = U^{-1}x$ zurücktransformiert, nach § 10, (7):

$$(29) \quad (Ax, x) = \sum_r a^{(r)} (U E^{(r)} U^{-1} x, x).$$

Nun ist mit $E^{(r)}$ auch

$$(30) \quad F^{(r)} = U E^{(r)} U^{-1}$$

eine Einzelmatrix als Darstellung desselben Einzeltensors im ursprünglichen Koordinatensystem (s. § 10). Ferner sind je zwei verschiedene der $F^{(r)}$ aufeinander „orthogonal“, d. h. es gilt

$$(31) \quad F^{(r)} F^{(s)} = 0 \quad \text{für} \quad r \neq s,$$

wie sofort aus (27) folgt.

Man kann also die Formel (29) so schreiben:

$$(32) \quad (Ax, x) = \sum_r a^{(r)} (F^{(r)} x, x)$$

oder für die Matrizen:

$$(32a) \quad A = \sum_r a^{(r)} F^{(r)}.$$

Dieselben Überlegungen, angewandt auf die Einheitsform, liefern

$$(33) \quad (x, x) = \sum_r (F^{(r)} x, x),$$

oder

$$(33a) \quad 1 = \sum_r F^{(r)}.$$

Damit ist eine eindeutige Normalform der hermiteschen Matrix

(Form) A erreicht; wir sprechen das Resultat in kovarianter Form so aus:

Satz 3: *Zu den verschiedenen Eigenwerten $a^{(r)}$ eines hermiteschen Tensors A gibt es ein System aufeinander orthogonaler Einzeltensoren $F^{(r)}$, deren Summe die Einheit ist und die erlauben, den Tensor eindeutig in die Gestalt*

$$A = \sum_r a^{(r)} F^{(r)}$$

zu bringen.

Man nennt das System der Tensoren $F^{(r)}$ die zum hermiteschen Tensor A gehörige Zerlegung der Einheit.

Aus der Hauptachsendarstellung (11a) liest man folgenden Satz ab:

Satz 4: *Eine hermitesche Form ist dann und nur dann positiv definit, wenn alle Eigenwerte a_k positiv sind.*

Wäre nämlich ein Eigenwert $a_l \leq 0$, so könnte man alle Variablen y_k gleich Null wählen, außer y_l und dies gleich 1; dann wäre $(x, x) = (y, y) = 1$ und $(Ax, x) = a_l \leq 0$.

Zusatz: *Eine hermitesche Form ist semidefinit, wenn kein Eigenwert negativ ist.*

Alle diese Sätze gelten zunächst für endliche Matrizen und Formen von endlich vielen Variablen. Die Übertragung auf unendliche Matrizen und Formen mit unendlich vielen Variablen erfordert besondere Untersuchungen; wir werden darauf später von einem allgemeineren Standpunkte (dem der Operatorenrechnung) aus eingehen. Hier sei nur vorweg genommen, daß selbst unter sehr allgemeinen Voraussetzungen die Hauptsätze richtig bleiben, vor allem der Satz über die Zerlegbarkeit einer hermiteschen Form in Einzelformen (die man bei endlichen Matrizen als „Hauptachsentransformation“ bezeichnen kann), mit dem einzigen Unterschied, daß die Eigenwerte $a^{(r)}$ im allgemeinen außer einer diskontinuierlichen Wertereihe auch einen kontinuierlichen Bereich erfüllen können. Dann treten z. B. in den Formeln (32a), (33a), die die Zerlegung der Einheit für eine Form A darstellen, neben der Summe noch Integralbestandteile auf:

$$A = \sum_r a^{(r)} F^{(r)} + \int a dF(a),$$

$$1 = \sum_r F^{(r)} + \int dF(a),$$

wo $dF(a)$ differentielle Einzelformen sind. So wichtig gerade dieser Umstand für das Verständnis der physikalischen Tatsachen (kontinuierliche Spektren) ist, soll doch vorläufig von dieser Komplikation abgesehen werden. In der Tat gibt es weite Klassen unendlicher hermitischer Formen, die nur diskrete Eigenwerte, ein „Punktspektrum“, haben.

Wir werden im folgenden bis auf weiteres voraussetzen, daß wir es mit endlichen oder solchen unendlichen Matrizen zu tun haben, die nur diskrete, abzählbare Eigenwerte besitzen.

§ 12. Extremaleigenschaften der Eigenwerte. Die geometrische Darstellung der reellen quadratischen Formen durch Flächen 2. Ordnung erlaubt es, die Eigenwerte (Hauptachsen) durch Extremaleigenschaften zu definieren. Fragt man nach dem größten Radiusvektor vom Nullpunkt nach der Fläche, d. h.

$$(x, x) = \text{Max.} \quad \text{für} \quad (Ax, x) = 1,$$

so entspricht dieser offenbar dem kleinsten Eigenwert; denn man erhält nach § 11, (7) für $(Ax, x) = 1$ als Länge des Radiusvektors

$$\sqrt{(x, x)} = \frac{1}{\sqrt{a}}.$$

Ist die größte Hauptachse gefunden, so kann man die nächstgrößte so charakterisieren: Man suche in der „Ebene“, die auf der größten Hauptachse senkrecht ist, den größten Radiusvektor der Schnittfigur. In dieser Weise kann man der Reihe nach alle Eigenwerte erhalten.

All das läßt sich ohne weiteres auf die hermitischen Formen übertragen. Da aber hier die geometrische Sprechweise doch nur einen beschränkten Sinn hat, ist es bequemer, das Extremalproblem etwas anders zu formulieren, nämlich so, daß die Eigenwerte selbst als Minimalwerte erscheinen. Wir behaupten:

Der kleinste Eigenwert a_1 ist das Minimum der Form

$$(1) \quad (Ax, x) = (ay, y)$$

unter der Nebenbedingung

$$(2) \quad (x, x) = (y, y) = 1.$$

Es ist nämlich für $a_1 \leq a_k$ ($k = 2, 3, \dots$)

$$(ay, y) = \sum_{k=1}^N a_k |y_k|^2 \geq a_1 \sum_{k=1}^N |y_k|^2,$$

also wegen $\sum_{k=1}^N |y_k|^2 = (y, y) = 1$:

$$(ay, y) \geq a_1,$$

und der Wert a_1 wird für den Vektor y mit den Komponenten $1, 0, 0, \dots$ erreicht. Im x -Koordinatensystem hat dieser auf Grund der Transformationsgleichungen $x = Uy$ die Komponenten $x_n = U_{n1}$, ist also identisch mit dem Vektor U_1 .

Wir fügen nun die Nebenbedingung

$$(3) \quad (x, U_1) = 0$$

hinzu und suchen wieder das Minimum der Form. Diese Nebenbedingung lautet im y -Koordinatensystem, wo U_1 die Komponenten $1, 0, 0, \dots$ hat, einfach $y_1 = 0$. Die Form reduziert sich also auf

$$\sum_{k=2}^N a_k |y_k|^2$$

und hat als Minimum den kleinsten der übrigen Eigenwerte a_2, a_3, \dots , der auch gleich a_1 sein kann. Der zugehörige Lösungsvektor ist $U_2(0, 1, 0, 0, \dots)$. Fügt man nun die weitere Nebenbedingung

$$(4) \quad (x, U_2) = 0 \quad \text{oder} \quad y_2 = 0$$

hinzu, so erhält man als Minimum der Form den nächsten Eigenwert a_3 usw.

Satz 1: Der kleinste Eigenwert der Form (Ax, x) ist ihr Minimum unter der Nebenbedingung $(x, x) = 1$; wird es für den Vektor $x = U_1$ erreicht, so ist der zweitkleinste Eigenwert das Minimum der Form unter den beiden Nebenbedingungen $(x, x) = 1, (U_1, x) = 0$. In analoger Weise erhält man alle Eigenwerte durch Hinzufügen der Nebenbedingungen $(U_2, x) = 0, (U_3, x) = 0, \dots$, wo die Vektoren U_2, U_3, \dots jedesmal die vorangehende Minimalaufgabe lösen.

Anstatt in dieser Weise die Eigenwerte durch eine Reihe von Minimalproblemen zu charakterisieren, von denen jedes die Kenntnis der Lösung der vorhergehenden voraussetzt, kann man auch direkt den k -ten Eigenwert in der nach ansteigender Größe

geordneten Reihe $a_1 \leq a_2 \leq \dots \leq a_N$ als Lösung eines bestimmten Extremalproblems bestimmen. Nach R. COURANT¹ denke man sich außer der Nebenbedingung $(x, x) = 1$ noch $k - 1$ weitere lineare Bedingungen

$$(5) \quad (x, \alpha_p) = 0, \quad p = 1, 2, \dots, k - 1$$

vorgegeben. Man suche zunächst das Minimum der Form (Ax, x) bei allen diesen Nebenbedingungen; dieses wird eine Funktion der Vektoren α_p (mit den Komponenten α_{np}) sein. Sodann suche man dieses Minimum von (Ax, x) durch geeignete Wahl der α_{np} zum Maximum zu machen.

Bezogen auf die Hauptachsen hat man also zunächst

$$(6) \quad (ay, y) = \sum_n a_n |y_n|^2 = \text{Min.},$$

$$(7) \quad (y, y) = \sum_n |y_n|^2 = 1,$$

$$(8) \quad (y, \beta_p) = \sum_n \beta_{np} y_n = 0, \quad p = 1, 2, \dots, k - 1.$$

Wählt man $y_{k+1} = \dots = y_N = 0$, so ergeben sich aus (8) bei beliebigen β_{np} jedesmal $k - 1$ homogene Gleichungen für die k Unbekannten y_1, \dots, y_k ; sie haben sicher eine Lösung, die nur bis auf einen Faktor bestimmt ist, und dieser kann so gewählt werden, daß (7) erfüllt ist. Für diese Werte wird

$$(ay, y) = \sum_{n=1}^k a_n |y_n|^2 \leq a_k \sum_{n=1}^N |y_n|^2 = a_k.$$

Das zunächst gesuchte Minimum ist also für kein Wertsystem der β_{np} größer als a_k . Es wird aber gerade gleich a_k , wenn für (8) die Gleichungen

$$y_1 = \dots = y_{k-1} = 0$$

genommen werden. So ergibt sich der

Satz 2: *Der k -te Eigenwert a_k der hermiteschen Form (Ax, x) ist der größte Wert, den das Minimum von (Ax, x) annehmen kann, wenn dem Vektor x außer der Bedingung $(x, x) = 1$ noch vorgeschrieben wird, auf $k - 1$ beliebigen Vektoren α_p orthogonal zu sein: $(x, \alpha_p) = 0$.*

¹ R. COURANT hat die entsprechenden Betrachtungen für reelle quadratische Formen angestellt (s. etwa COURANT, R. u. D. HILBERT: Methoden der mathematischen Physik I, 1. Kap., § 4).

Auf Grund dieser Maximum-Minimum-Eigenschaft läßt sich besonders leicht übersehen, in welcher Weise die Eigenwerte sich verändern, wenn die Variabilität der x_k durch j unabhängige „Bindungen“

$$(9) \quad (x, \gamma_m) = 0, \quad m = 1, 2, \dots, j$$

eingeschränkt wird, so daß sich (Ax, x) auf eine hermitesche Form von $N - j$ unabhängigen Variablen $(A'x, x)$ reduziert. Der k -te Eigenwert von $(A'x, x)$ entsteht durch dasselbe Maximum-Minimum-Problem, wobei nur der Bereich der zur Konkurrenz zugelassenen Wertsysteme x_1, x_2, \dots, x_N durch (9) verengert ist. Daher unterschreitet das einzelne Minimum und somit auch der Eigenwert (als Maximalwert dieses Minimums) bei $(A'x, x)$ sicher nicht die entsprechende Größe bei (Ax, x) .

Ferner ist a_{j+k} das größte Minimum, das (Ax, x) besitzen kann, wenn außer $(x, x) = 1$ noch $k + j - 1$ beliebige lineare Bedingungen für x vorgeschrieben sind, und daher sicher nicht kleiner als a'_k , für welches j dieser Bedingungen durch die festen Gleichungen (9) dargestellt werden. Also ist

$$a_k \leq a'_k \leq a_{k+j},$$

oder in Worten:

Satz 3: *Geht eine hermitesche Form (Ax, x) durch j Bedingungen $(x, \gamma_m) = 0$ in eine hermitesche Form $(A'x, x)$ von $N - j$ Variablen über, so sind die Eigenwerte $a'_1, a'_2, \dots, a'_{N-j}$ der letzteren nicht kleiner als die entsprechenden Zahlen der Reihe a_1, a_2, \dots, a_{N-j} und nicht größer als die entsprechenden Zahlen der Reihe a_{j+1}, \dots, a_N .*

Anstatt ein Maximum-Minimum-Problem zur Kennzeichnung der Eigenwerte zu benutzen, kann man ebensogut ein Minimum-Maximum-Problem zum Ausgang nehmen. Man erhält dann wiederum die a_k , nur in umgekehrter Reihenfolge.

Diese Sätze gestatten, den Einfluß von „Bindungen“ auf die Lage der Eigenwerte abzuschätzen.

§ 13. Gleichzeitige Hauptachsentransformation vertauschbarer hermitescher oder unitärer Matrizen. Es seien eine Reihe von hermiteschen Matrizen (Tensoren) $A^{(1)}, A^{(2)}, \dots, A^{(l)}$ gegeben, die alle untereinander vertauschbar sind:

$$(1) \quad [A^{(k)}, A^{(l)}] = 0, \quad k, l = 1, 2, \dots, f.$$

Wir zeigen, daß sie sich gleichzeitig auf Hauptachsen transformieren lassen.

Zunächst denke man sich $A^{(1)}$ auf die Diagonalform gebracht mit Hilfe der unitären Matrix U ; es sei also

$$(2) \quad U^{-1} A^{(1)} U = a^{(1)} = (a_n^{(1)} \delta_{nm}).$$

Zugleich gehe $A^{(2)}$ über in

$$(3) \quad U^{-1} A^{(2)} U = B^{(2)}.$$

U ist nur bis auf einen Faktor $\overset{\circ}{U}$ bestimmt; dabei ist $\overset{\circ}{U}$ die durch § 11, (23) dargestellte unitäre Stufenmatrix, deren Stufenlängen die Gewichte (Vielfachheiten) $g_r^{(1)}$ von $A^{(1)}$ sind. Aus (1) folgt nun für $k = 1, l = 2$

$$(4) \quad \begin{aligned} U^{-1} [A^{(1)}, A^{(2)}] U &= 0, \\ [U^{-1} A^{(1)} U, U^{-1} A^{(2)} U] &= 0, \\ [a^{(1)}, B^{(2)}] &= 0. \end{aligned}$$

Das bedeutet aber

$$(4a) \quad (a_n^{(1)} - a_m^{(1)}) B_{nm}^{(2)} = 0.$$

Folglich ist $B_{nm}^{(2)} = 0$, wenn $a_n^{(1)} \neq a_m^{(1)}$; wenn man $a^{(1)}$ als geordnete Diagonalmatrix (mit den Multiplizitäten $g_r^{(1)}$) wählt, ist also $B^{(2)}$ eine Stufenmatrix

$$(5) \quad B^{(2)} = \begin{pmatrix} B^{(2,1)} & 0 & \cdot & 0 & \dots \\ 0 & B^{(2,2)} & 0 & \dots \\ \cdot & \cdot & \cdot & \cdot & \cdot \end{pmatrix},$$

wobei $B^{(2,r)}$ eine $g_r^{(1)}$ -reihige Matrix ist. $B^{(2)}$ ist also mit der durch § 11, (23) definierten unitären Matrix $\overset{\circ}{U}$ stufengleich. Daher kann man in U den noch willkürlichen Faktor $\overset{\circ}{U}$ so bestimmen, daß $B^{(2)}$ eine Diagonalmatrix $a^{(2)}$ wird. Dann hat man

$$(6) \quad U^{-1} A^{(2)} U = a^{(2)} = (a_n^{(2)} \delta_{mn}).$$

Diese Schlußweise kann fortgesetzt werden: Die Transformationsmatrix U ist durch (2) und (6) noch nicht eindeutig bestimmt, sondern nur bis auf einen Faktor, dessen Elemente sämtlich verschwinden, wenn für die Indizes n, m entweder $a_n^{(1)} \neq a_m^{(1)}$ oder $a_n^{(2)} \neq a_m^{(2)}$ ist. Ist nun

$$(7) \quad U^{-1} A^{(3)} U = B^{(3)},$$

so folgt aus der Vertauschbarkeit von $A^{(3)}$ mit $A^{(1)}$ und $A^{(2)}$ wie oben, daß auch die Elemente $B_{nm}^{(3)}$ verschwinden, wenn entweder $a_n^{(1)} \neq a_m^{(1)}$ oder $a_n^{(2)} \neq a_m^{(2)}$ ist. Also reicht die Willkür in U aus, um $B^{(3)}$ in eine Diagonalmatrix $a^{(3)}$ zu verwandeln. Das Verfahren läßt sich fortsetzen, bis alle Matrizen $A^{(1)}, A^{(2)}, \dots, A^{(f)}$ auf die Diagonalform gebracht sind, und wir haben den

Satz 1: *Zu jedem System $A^{(1)}, A^{(2)}, \dots, A^{(f)}$ hermitescher Matrizen, die paarweise miteinander vertauschbar sind, gibt es eine unitäre Matrix U , die sie zugleich auf die Diagonalform bringt:*

$$(8) \quad U^{-1} A^{(k)} U = a^{(k)}, \quad k = 1, 2, \dots, f.$$

Eine wichtige Anwendung dieses Satzes ist der Nachweis, daß man jede unitäre Matrix U durch eine unitäre Transformation auf die Diagonalform bringen kann. Es sind nämlich wegen $U U^\dagger = U^\dagger U = 1$ die Matrizen

$$(9) \quad A = \frac{1}{2}(U + U^\dagger), \quad B = \frac{1}{2i}(U - U^\dagger)$$

hermitisch und miteinander vertauschbar. Also kann man sie nach Satz 1 simultan auf Hauptachsen transformieren; dann sind aber auch

$$(10) \quad U = A + iB, \quad U^\dagger = A - iB$$

in Diagonalmatrizen (Phasenmatrizen) verwandelt.

Satz 2: *Jede unitäre Matrix läßt sich durch eine unitäre Transformation in eine Phasenmatrix transformieren.*

Hieraus folgt die Umkehrung eines in § 7 bewiesenen Satzes, wonach man aus jeder hermiteschen Matrix A eine unitäre durch den Ansatz

$$(11) \quad U = e^{iA}$$

gewinnen kann: Man kann auch umgekehrt zu jeder unitären Matrix U eine hermitesche A so bestimmen, daß die Gleichung (11) gilt. Dazu nehme man U nach Satz 2 in der Diagonalform $U = (e^{i a_n} \delta_{nm})$ an, wo a_n reelle Konstanten sind. Dann hat die hermitesche Diagonalmatrix $A = (a_n \delta_{nm})$ offenbar die gewünschte Eigenschaft, und diese bleibt bei unitärer Transformation erhalten.

Satz 3: Zu jeder unitären Matrix U gibt es eine hermitische Matrix A und umgekehrt, so daß

$$U = e^{iA}$$

ist.

Den Satz 2 kann man ohne weiteres auf den Fall mehrerer vertauschbarer unitärer Matrizen verallgemeinern:

Satz 4: Eine Reihe miteinander vertauschbarer unitärer Matrizen $U^{(1)}, U^{(2)}, \dots, U^{(f)}$ kann durch eine einzige unitäre Transformation U in eine Reihe von Phasenmatrizen transformiert werden:

$$(12) \quad U^{-1} U^{(k)} U = \varepsilon^{(k)}, \quad |\varepsilon_n^{(k)}| = 1.$$

Mit $U^{(1)}, \dots, U^{(f)}$ sind auch $U^{(1)-1} = U^{(1)\dagger}, \dots, U^{(f)-1} = U^{(f)\dagger}$ miteinander und mit $U^{(1)}, \dots, U^{(f)}$ vertauschbar, also auch alle

$$A_1 = \frac{1}{2} (U^{(1)} + U^{(1)\dagger}), \dots, A_f = \frac{1}{2} (U^{(f)} + U^{(f)\dagger}),$$

$$B_1 = \frac{1}{2i} (U^{(1)} - U^{(1)\dagger}), \dots, B_f = \frac{1}{2i} (U^{(f)} - U^{(f)\dagger}).$$

Diese sind auch hermitisch, also können sie gleichzeitig auf die Diagonalform gebracht werden, und mit ihnen

$$U^{(1)} = A_1 + iB_1, \dots, U^{(f)} = A_f + iB_f.$$

§ 14. Funktionen vertauschbarer hermitischer oder unitärer Matrizen. Wir hatten in § 7 Matrizenfunktionen mit Hilfe von Polynomen und Grenzwerten von Polynomen definiert. Man kann nun die Hauptachsentransformation der hermitischen oder der unitären Matrizen benutzen, um Funktionen von solchen Matrizen (oder von Tensoren), sofern sie vertauschbar sind, unabhängig von der Darstellbarkeit als Grenzwert von Polynomen (etwa als Potenzreihe) zu erklären. Betrachten wir zunächst Funktionen eines einzigen Arguments. Es sei eine eindeutige Funktion $f(\xi)$ des reellen Arguments ξ gegeben; wir können dann willkürfrei bestimmen, was unter derselben Funktion der hermitischen Matrix A , $f(A)$ zu verstehen sei.

U sei eine unitäre Matrix, die A auf die Diagonalform $U^{-1}AU = a$ bringt; ist dann $f(a) = f(U^{-1}AU)$ die Diagonalmatrix, die in der Diagonale die Elemente $f(a_n)$ hat, so definieren wir

$$(1) \quad f(A) = U f(a) U^{-1} = U \cdot f(U^{-1}AU) \cdot U^{-1}.$$

Die Definition ist unabhängig von der noch in U steckenden Willkür, dargestellt durch die unitäre Stufenmatrix \hat{U} , § 11, (23); denn $f(a)$ ist eine Diagonalmatrix, deren Diagonalelemente immer innerhalb jeder Stufe von U konstant¹, nämlich gleich $f(a^{(r)})$ sind; für jede Teiltransformation $U^{(r)}$ hat man also $U^{(r)}f(a^{(r)})U^{(r)-1} = f(a^{(r)})$. Eine kovariante Formulierung der Definition (1) gewinnt man mit Hilfe der Zerlegung des Tensors A in Einzeltensoren (§ 11, Satz 3). Ist

$$(2) \quad A = \sum_r a^{(r)} F^{(r)},$$

so hat man offenbar

$$(3) \quad f(A) = \sum_r f(a^{(r)}) F^{(r)}.$$

Hier springt die Eindeutigkeit der Definition sofort ins Auge.

Ist $f(\xi)$ eine Funktion, die aus ihrem Argumente ξ durch endlich viele rationale Operationen (Addition, Multiplikation Division) herzustellen ist, so stimmt offenbar für eine beliebige Matrix A die neue Definition von $f(A)$ mit der früheren überein. Dasselbe gilt für eine solche Potenzreihe $f(\xi)$, die für alle (reellen) ξ konvergiert, sofern sie auch für A konvergent ist. (Nur unter dieser letzten Bedingung ist die frühere Definition $f(A)$ überhaupt anwendbar).

Betrachten wir jetzt eine Funktion $f(A)$, die für eine spezielle Matrix A mit bestimmten Eigenwerten a_n gebildet werden soll, so können wir behaupten: Wenn man die Funktion $f(\xi)$ für die in Frage kommenden Stellen $\xi = a_n$ durch eine konvergente Potenzreihe darstellen kann (bei den endlichen Matrizen kann man, wie wir weiter unten [Formel (7), (8)] zeigen werden, immer mit einem Polynom $(N - 1)$ -ten Grades in ξ auskommen) und wenn diese Potenzreihe auch dann konvergiert, wenn man ξ durch die Matrix A ersetzt, so kommt man auf die ursprüngliche

¹ Wenn $f(a^{(r)})$ eine mehrdeutige Funktion ist, braucht das nicht der Fall zu sein. Das hat zur Folge, daß mehrdeutige Funktionen einer und derselben entarteten Matrix nicht miteinander vertauschbar zu sein brauchen. So sind z. B. die Matrizen $x_1 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$ und $x_2 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ beide Einheitswurzeln, sowohl nach der alten $[x_1^2 = x_2^2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}]$, wie auch nach der neuen Definition (vgl. Satz 2), aber $x_1 x_2 = -x_2 x_1$.

Potenzreihen-Definition von $f(A)$ zurück. Doch ist die neue Definition (1) oder (3) wichtig für die spätere Behandlung unendlicher Matrizen, besonders wenn es sich um unstetige Funktionen $f(\xi)$ handelt, z. B. um die Funktion, die in einem gewissen Intervalle gleich 1 und sonst gleich 0 ist.

Als Anwendung dieser Betrachtungen diene die Bemerkung, daß eine hermitesche Matrix A nur dann eine Reziproke A^{-1} besitzt, wenn alle ihre Eigenwerte von Null verschieden sind. Denn in diesem Falle ist, wie oben bemerkt wurde, $f(\xi) = \xi^{-1}$, d. h. $f(a_n) = a_n^{-1}$, und wenn ein $a_n = 0$ ist, so ist das entsprechende $f(a_n) = \infty$. Dies stimmt für endliche Matrizen mit der gewöhnlichen Formulierung der Bedingung für die Existenz von A^{-1} überein, daß $\text{Det } A \neq 0$ sein muß. Es ist nämlich $\text{Det } A = \text{Det } (U^{-1}AU) = \text{Det } a = a_1 a_2 \dots$.

Die Definition (1) kann nach § 13, Satz 1 auf Funktionen $f(A^{(1)}, A^{(2)}, \dots, A^{(r)})$ von r untereinander vertauschbaren Matrizen $A^{(k)}$ ausgedehnt werden, sofern $f(\xi^{(1)}, \xi^{(2)}, \dots, \xi^{(r)})$ als eindeutige Funktion von r gewöhnlichen Zahlen gegeben ist. Dann gibt es nämlich eine unitäre Matrix U , die alle $A^{(k)}$ gleichzeitig auf Hauptachsen bringt:

$$U^{-1} A^{(k)} U = a^{(k)},$$

und man hat zu definieren:

$$(4) \quad f(A^{(1)}, A^{(2)}, \dots, A^{(r)}) = U f(a^{(1)}, a^{(2)}, \dots, a^{(r)}) U^{-1} \\ = U f(U^{-1}A^{(1)}U, \dots, U^{-1}A^{(r)}U) U^{-1},$$

wo $f(a^{(1)}, a^{(2)}, \dots, a^{(r)})$ die Diagonalmatrix mit den Elementen $f(a_n^{(1)}, a_n^{(2)}, \dots, a_n^{(r)})$ bedeutet.

Im Anschluß hieran beweisen wir den

Satz 1: Mehrere vertauschbare Matrizen $A^{(1)}, A^{(2)}, \dots, A^{(r)}$ sind darstellbar als Funktionen einer einzigen Matrix B :

$$(5) \quad A^{(1)} = f_1(B), \quad A^{(2)} = f_2(B), \quad \dots, \quad A^{(r)} = f_r(B).$$

Zum Beweise bringe man $A^{(1)}, \dots, A^{(r)}$ gleichzeitig auf die Diagonalform:

$$U^{-1} A^{(k)} U = a^{(k)}.$$

Wir wählen nun eine Diagonalmatrix b mit lauter verschiedenen Diagonalelementen b_1, b_2, \dots und definieren die Funktionen f_k

in der Weise, daß

$$f_k(b_1) = a_1^{(k)}, \quad f_k(b_2) = a_2^{(k)}, \dots$$

Diese Funktionen leisten das Gewünschte; denn man hat für $B = UbU^{-1}$

$$A^{(k)} = U a^{(k)} U^{-1} = U f_k(b) U^{-1} = f_k(UbU^{-1}) = f_k(B).$$

Wir wollen eine später nützliche Bezeichnung einführen. Die Systeme von zusammengehörigen Eigenwerten der vertauschbaren Matrizen $A^{(1)}, A^{(2)}, \dots, A^{(r)}$, die sich bei der gleichzeitigen Hauptachsentransformation ergeben, also die Zahlensysteme

$$(6) \quad a_n^{(1)}, \quad a_n^{(2)}, \dots, a_n^{(r)}$$

nennen wir kurz die *Eigenwerte des Matrixsystems* $A^{(1)}, \dots, A^{(r)}$. Die Berechtigung dieser Bezeichnung ergibt sich daraus, daß offenbar die in der Transformationsmatrix U verbleibende Willkür lediglich Permutationen der Zahlenreihen (6) untereinander zuläßt.

Alle diese Definitionen und Sätze übertragen sich ohne weiteres auf Funktionen von unitären Matrizen.

Wir fügen noch einige einfache Bemerkungen über Matrixfunktionen eines Arguments an, von denen man oft Gebrauch macht. Sind $a^{(1)}, a^{(2)}, \dots, a^{(M)}$ die *verschiedenen* Eigenwerte einer (endlichen) hermitischen Matrix A , so betrachten wir die Funktion

$$(7) \quad \begin{aligned} F(\xi) &= (\xi - a^{(1)}) (\xi - a^{(2)}) \dots (\xi - a^{(M)}) \\ &= \xi^M + c_1 \xi^{M-1} + \dots + c_{M-1} \xi + c_M. \end{aligned}$$

Bilden wir jetzt nach der Definition (1) die entsprechende Matrixfunktion $F(A)$, so verschwindet diese, da $F(\xi)$ gerade so konstruiert ist, daß es Null wird, wenn man für ξ irgend einen Eigenwert a_n von A einsetzt. So erhalten wir die sogenannte **HAMILTON-CAYLEYsche Gleichung**

$$(8) \quad F(A) = 0.$$

Jede hermitische Matrix mit M voneinander verschiedenen Eigenwerten genügt also einer Gleichung M -ten Grades, ein Resultat, das wir oben schon erwähnt haben.

Allgemeiner kann man offenbar (auch für unendliche Matrizen mit diskreten Eigenwerten) sagen: Ist $R(\xi)$ eine Funktion der

reellen Variablen ξ , die für alle Eigenwerte $a^{(k)}$ einer hermiteschen Matrix A verschwindet, so gilt $R(A) = 0$. Es gilt aber auch die wichtigere Umkehrung: wenn eine hermitesche Matrix einer Gleichung $R(A) = 0$ genügt, so folgt aus (1) sofort $R(a) = 0$, d. h. alle Eigenwerte sind Wurzeln der Gleichung $R(\xi) = 0$. Wir haben also den

Satz 2: *Eine hermitesche Matrix A genügt dann und nur dann einer Gleichung*

$$R(A) = 0,$$

wenn alle Eigenwerte Wurzeln der Gleichung

$$R(\xi) = 0$$

sind.

Daraus folgt, daß für eine endliche Matrix die HAMILTON-CAYLEYSche Gleichung die Gleichung niedersten Grades ist, der die Matrix genügen kann.

Wir machen eine Anwendung auf die *Einzelmatrizen* F (s. § 8), die definitionsgemäß der Gleichung

$$(9) \quad R(F) = F^2 - F = 0$$

genügen. Aus Satz 2 folgt dann sogleich:

Satz 3: *Jeder Eigenwert einer Einzelmatrix ist entweder gleich 0 oder gleich 1. Man kann daher jede Einzelmatrix durch unitäre Transformation in eine Diagonalmatrix $E^{(m)}$ transformieren, bei der m Stellen der Diagonale mit 1, alle übrigen mit 0 besetzt sind.*

Dieser Satz gibt die anschaulichste Erläuterung des eigentlichen Wesens der Einzelmatrizen.

Man kann die Elemente der Einzelmatrix F durch die Komponenten derjenigen Eigenvektoren ausdrücken, die zu den Eigenwerten 1 gehören. Jeder Vektor y läßt sich nämlich durch das vollständige Orthogonalsystem der Eigenvektoren $x^{(1)}, x^{(2)}, \dots$ ausdrücken in der Form

$$y = \sum_n (y, x^{(n)}) x^{(n)}.$$

Daher wird

$$F y = \sum_n (y, x^{(n)}) F x^{(n)},$$

und hier ist $Fx^{(n)} = 0$, außer, wenn $x^{(n)}$ zum Eigenwert 1 gehört, wo dann $Fx^{(n)} = x^{(n)}$ ist. Folglich wird

$$Fy = \sum_{(n)} (y, x^{(n)}) x^{(n)},$$

wo die Summe jetzt nur über die zum Eigenwert 1 gehörigen Eigenvektoren zu erstrecken ist. Das bedeutet aber für die Elemente:

$$(10) \quad F_{kl} = \sum_{(n)} x_k^{(n)} x_l^{(n)*},$$

oder mit der Bezeichnung des direkten Produktes (s. § 10, (9a)):

$$(11) \quad F = \sum_{(n)} x^{(n)} \times x^{(n)*}.$$

§ 15. Einfache (irreduzible) und zerlegbare (reduzible) Matrixensysteme. Für Systeme beliebiger gleichartiger (s. § 4), im allgemeinen *nicht vertauschbarer* Matrizen T_1, T_2, \dots haben wir eine wichtige kovariante Begriffsbildung zu formulieren und einige Sätze darüber abzuleiten. *Man nennt ein System von gleichartigen Matrizen T_1, T_2, \dots , die sämtlich hermitisch oder sämtlich unitär sind, einfach oder irreduzibel, wenn es nicht möglich ist, sie alle gleichzeitig durch eine unitäre Transformation auf die Form von Stufenmatrizen*

$$(1) \quad T_k = \begin{pmatrix} T_k^{(1)} & 0 \\ 0 & T_k^{(2)} \end{pmatrix}$$

mit übereinstimmender Stufenteilung zu bringen. Ist eine Transformation auf die Form (1) möglich, so wird das Matrixensystem als zerlegbar oder reduzibel bezeichnet.

Die Matrizen (1) sind offenbar mit der gleichstufigen Diagonalmatrix

$$(2) \quad a = \begin{pmatrix} a' E^{(1)} & 0 \\ 0 & a'' E^{(2)} \end{pmatrix}$$

vertauschbar, wo a', a'' beliebige Zahlen und $E^{(1)}, E^{(2)}$ die entsprechenden Einheitsmatrizen zu den Teilmatrizen $T_k^{(1)}$ bzw. $T_k^{(2)}$ sind. Daraus folgt, daß ein Matrixensystem sicher nicht die Form (1) haben kann, also irreduzibel ist, wenn man weiß, daß es keine andere, mit allen Matrizen des Systems vertauschbare Matrix gibt als ein Vielfaches der Einheitsmatrix.

Es gilt aber auch das Umgekehrte: Wenn eine Matrix A mit allen Matrizen T_k des einfachen Systems vertauschbar ist, so ist A bis auf einen Zahlenfaktor gleich der Einheitsmatrix.

Man kann A sogleich als hermitisch voraussetzen. Wäre nämlich A nicht hermitisch, so könnte man statt A die hermitischen Matrizen $A_1 = A + A^\dagger$ und $A_2 = i(A - A^\dagger)$ betrachten; denn mit A muß auch A^\dagger und daher auch A_1 und A_2 mit allen T_1, T_2, \dots vertauschbar sein, weil diese entweder alle als hermitisch oder alle als unitär vorausgesetzt wurden.

Eine hermitische Matrix A aber kann man durch eine unitäre Transformation auf die Diagonalform $(a_n \delta_{nm})$ mit reellen a_n bringen. Transformiert man die T_k durch dieselbe unitäre Transformation, so bleiben die Voraussetzungen $[T_k, A] = 0$ erhalten und nehmen die einfache Gestalt

$$(3) \quad (a_n - a_m) T_k(n, m) = 0$$

an. Wir denken uns nun A als geordnete Diagonalmatrix, so daß die einander gleichen a_n aufeinander folgen. Gäbe es nun solche a_m , die nicht gleich dem zuerst angeschriebenen a_1 sind, so wäre für alle Indizes n , für die $a_n = a_1$ ist, nach (3)

$$T_k(n, m) = 0, \quad T_k(m, n) = 0,$$

d. h. alle T_k hätten die verbotene Form (1). Folglich sind alle a_n einander gleich, etwa gleich a , und man hat $A = aE$. Diese Eigenschaft bleibt bei unitären Transformationen erhalten.

Wir haben also das Resultat gewonnen:

Satz 1: Ein System von gleichartigen Matrizen T_1, T_2, \dots ist dann und nur dann einfach, wenn jede Matrix A , die mit allen T_k vertauschbar ist, bis auf einen Zahlenfaktor die Einheitsmatrix ist.

Aus diesem Satze oder auch unmittelbar aus (1) entnimmt man den

Zusatz: Ein einfaches oder zerlegbares Matrixensystem T_1, T_2, \dots bleibt einfach oder zerlegbar, wenn man ihm Summen oder Produkte oder sonstige Funktionen der T_1, T_2, \dots hinzufügt. Auch kann man dem System T_1, T_2, \dots die zugehörige Einheitsmatrix $T_0 = E$ hinzufügen, ohne seine Einfachheit bzw. Zerlegbarkeit aufzuheben.

In vielen mathematischen und physikalischen Problemen spielt ein Prozeß eine Rolle, der aus f Systemen von jeweils gleichartigen Matrizen

$$(4) \quad \begin{cases} t_0^{(1)}, t_1^{(1)}, t_2^{(1)}, \dots \\ \dots \dots \dots \dots \dots \\ t_0^{(k)}, t_1^{(k)}, t_2^{(k)}, \dots \\ \dots \dots \dots \dots \dots \\ t_0^{(f)}, t_1^{(f)}, t_2^{(f)}, \dots \end{cases}$$

ein neues Matrixensystem abzuleiten gestattet, welches wir als „Verschmelzung“ der Matrixensysteme (4) bezeichnen wollen. Wir betonen ausdrücklich, daß zwei zu verschiedenen k gehörige Matrizen (4) *nicht* gleichartig zu sein brauchen.

Die zum System k gehörigen Indizes seien n_k, m_k , so daß für alle s

$$(5) \quad t_s^{(k)} = (t_s^{(k)}(n_k, m_k)).$$

Wir bezeichnen abkürzend mit Δ_k das, was übrig bleibt, wenn man in dem Produkt

$$\delta_{n_1 m_1} \delta_{n_2 m_2} \dots \delta_{n_f m_f}$$

den Faktor $\delta_{n_k m_k}$ ausläßt.

Dann bilden wir das „Verschmelzung“ genannte System gleichartiger Matrizen gemäß der Vorschrift:

$$(6) \quad \begin{cases} T_s^{(k)} = (T_s^{(k)}(n_1, n_2, \dots, n_f; m_1, m_2, \dots, m_f)); \\ T_s^{(k)}(n_1, n_2, \dots, n_f; m_1, m_2, \dots, m_f) = \Delta_k t_s^{(k)}(n_k, m_k). \end{cases}$$

Bilden wir aus den $t_0^{(k)}, t_1^{(k)}, \dots$ für ein festes k eine Funktion

$$(7) \quad F(t_0^{(k)}, t_1^{(k)}, \dots) = (f(n_k, m_k)),$$

so ist dieselbe Funktion der $T_s^{(k)}$ gegeben durch

$$(8) \quad \begin{cases} F(T_0^{(k)}, T_1^{(k)}, \dots) = (F(n_1, n_2, \dots, n_f; m_1, m_2, \dots, m_f)); \\ F(n_1, \dots, n_f; m_1, \dots, m_f) = \Delta_k f(n_k, m_k). \end{cases}$$

Ferner geht $T_s^{(k)\dagger}$ durch den gleichen Prozeß aus $t_s^{(k)\dagger}$ hervor, wie $T_s^{(k)}$ aus $t_s^{(k)}$. Danach ergibt sich

Satz 2: *Besteht zwischen den Matrizen $t_0^{(k)}, t_1^{(k)}, t_2^{(k)}, \dots$ eine Gleichung*

$$F(t_0^{(k)}, t_1^{(k)}, \dots; t_0^{(k)\dagger}, t_1^{(k)\dagger}, \dots) = 0,$$

so bleibt diese bestehen, wenn man die $t_s^{(k)}$ durch die $T_s^{(k)}$ nach (6) ersetzt.

Als Sonderfälle dieses Satzes hat man:

Zusatz: $T_s^{(k)}$ ist gleichzeitig mit $t_s^{(k)}$ hermitisch bzw. unitär.

Man entnimmt ferner aus (6) mit Rücksicht auf die Definition von Δ_k den

Satz 3: Es ist stets

$$(9) \quad [T_s^{(k)}, T_r^{(l)}] = 0 \quad \text{für } k \neq l.$$

Ferner gilt

Satz 4: Die Verschmelzung von f einfachen Matrixsystemen ist wieder ein einfaches Matrixsystem.

Sei nämlich $A = (A(n_1, n_2, \dots, n_f; m_1, m_2, \dots, m_f))$ eine mit allen $T_s^{(k)}$ vertauschbare Matrix:

$$(10) \quad A T_s^{(k)} - T_s^{(k)} A = 0.$$

Nun wird nach (6)

$$(11) \quad \left\{ \begin{array}{l} A T_s^{(k)}(n_1, \dots, n_f; m_1, \dots, m_f) \\ \quad = \sum_{p_k} A(n_1, \dots, n_f; m_1, \dots, m_{k-1}, p_k, m_{k+1}, \dots, m_f) t_s^{(k)}(p_k, m_k), \\ T_s^{(k)} A(n_1, \dots, n_f; m_1, \dots, m_f) \\ \quad = \sum_{p_k} t_s^{(k)}(n_k, p_k) A(n_1, \dots, n_{k-1}, p_k, n_{k+1}, \dots, n_f; m_1, \dots, m_f). \end{array} \right.$$

Nehmen wir also aus der Matrix A alle Elemente heraus, bei denen sämtliche Indizes außer n_k und m_k bestimmte spezielle Werte haben, so erhalten wir, wie (11) zeigt, eine Teilmatrix, die mit $t_s^{(k)}$ gleichartig und vertauschbar ist, wenn (10) gilt. Also muß diese Teilmatrix wegen der vorausgesetzten Einfachheit des Systems $t_0^{(k)}, t_1^{(k)}, \dots$ nach Satz 1 der Einheitsmatrix $E^{(k)}$ proportional sein oder in Formeln

$$\begin{aligned} & A(n_1, n_2, \dots, n_f; m_1, m_2, \dots, m_f) \\ &= \delta_{n_k m_k} A'(n_1, \dots, n_{k-1}, n_{k+1}, \dots, n_f; m_1, \dots, m_{k-1}, m_{k+1}, \dots, m_f), \end{aligned}$$

wo also A' von n_k, m_k unabhängig ist. Diese Gleichung muß aber für alle $k = 1, 2, \dots, f$ gelten und führt somit zu

$$\begin{aligned} & A(n_1, n_2, \dots, n_f, m_1, m_2, \dots, m_f) \\ &= a \delta_{n_1 m_1} \delta_{n_2 m_2} \dots \delta_{n_f m_f} = a E_{n_1 \dots n_f, m_1 \dots m_f} \end{aligned}$$

als notwendiger Bedingung für die Gültigkeit von (10). Das ist

aber nach Satz 1 gleichbedeutend mit der behaupteten Einfachheit des Systems aller $T_s^{(k)}$.

Eine wichtige, wenn auch fast triviale Eigenschaft des Prozesses der Bildung von Verschmelzungen sei hervorgehoben: Wenn man statt der Verschmelzung *aller* f Matrizen-systeme (4) zunächst die Verschmelzung etwa der ersten f' ($f' < f$) dieser Matrizen-systeme bildet und das so entstehende System mit den übrigen verschmilzt (die ihrerseits gleichfalls zuvor schon teilweise verschmolzen werden dürfen), so entsteht *dasselbe* System wie durch Verschmelzung aller einzelnen Systeme (4). Indem wir die in den Zeilen (4) stehenden Matrizen-systeme $\mathfrak{I}^{(1)}, \mathfrak{I}^{(2)}, \dots, \mathfrak{I}^{(f)}$ nennen und die Verschmelzung von $\mathfrak{I}^{(1)}$ und $\mathfrak{I}^{(2)}$ mit $\mathfrak{I}^{(1)} \vee \mathfrak{I}^{(2)}$ bezeichnen, können wir diesen Sachverhalt ausdrücken als ein *assoziatives Gesetz*

$$(\mathfrak{I}^{(1)} \vee \mathfrak{I}^{(2)}) \vee \mathfrak{I}^{(3)} = \mathfrak{I}^{(1)} \vee (\mathfrak{I}^{(2)} \vee \mathfrak{I}^{(3)}) = \mathfrak{I}^{(1)} \vee \mathfrak{I}^{(2)} \vee \mathfrak{I}^{(3)}.$$

Die Verschmelzung aller Systeme (4) ist entsprechend durch

$$\mathfrak{I} = \mathfrak{I}^{(1)} \vee \mathfrak{I}^{(2)} \vee \dots \vee \mathfrak{I}^{(k)} \vee \dots \vee \mathfrak{I}^{(f)}$$

zu bezeichnen.

Für die durch (6) definierten Matrizen $T_s^{(k)}$ von \mathfrak{I} werden wir auch die Bezeichnung

$$(12) \quad T_s^{(k)} = E^{(1)} \times E^{(2)} \times \dots \times E^{(k-1)} \times t_s^{(k)} \times E^{(k+1)} \times \dots \times E^{(f)}$$

gebrauchen, wo die $E^{(l)}$ wieder die Einheitsmatrizen der Systeme $\mathfrak{I}^{(l)}$ sind.

Wir betrachten nun ein zerlegbares Matrizen-system T_1, T_2, \dots . Durch eine unitäre Transformation kann man alle T_k in die Form von gleichstufigen Stufenmatrizen

$$(13) \quad T_k = (T_k^{(r)} \delta_{rs})$$

bringen, und zwar derart, daß jedes der zu einem bestimmten Werte r gehörigen Matrizen-systeme $T_1^{(r)}, T_2^{(r)}, \dots$ *einfach* ist. Wir nennen dann diese Systeme $T_1^{(r)}, T_2^{(r)}, \dots$ *die einfachen oder irreduzibeln Bestandteile* des Matrizen-systems T_1, T_2, \dots . Das in die Gestalt (13) transformierte Matrizen-system wird auch als die *direkte Summe* der irreduzibeln Bestandteile bezeichnet.

In dem besonderen Falle, daß alle T_k miteinander *vertauschbar* sind, kann die Zerlegung (13) so weit geführt werden, daß die

Stufengrößen gleich 1 und die Matrizen $T_k^{(r)}$ gewöhnliche Zahlen werden: das ist offenbar nur eine andere Formulierung des Inhalts von Satz 1, § 13. In diesem Falle sind die irreduzibeln Bestandteile einfach die *Eigenwerte des Matrixsystems* (s. § 14, S. 71). Die Transformation eines Matrixsystems in die Gestalt (13) ist also als Verallgemeinerung der Hauptachsentransformation einer einzigen Matrix oder eines Systems vertauschbarer Matrizen anzusehen.

Im allgemeinen Falle ist offenbar nicht nur die Reihenfolge der irreduzibeln Bestandteile in (13) willkürlich, sondern es kann überdies jeder irreduzible Bestandteil $T_1^{(r)}, T_2^{(r)}, \dots$ noch mit einer gleichartigen, sonst beliebigen unitären Matrix $U^{(r)}$ transformiert werden. In Verallgemeinerung unserer früheren Sätze über die eindeutige Bestimmtheit der Eigenwerte läßt sich jedoch zeigen, daß *im übrigen* die irreduzibeln Bestandteile *eindeutig* bestimmt sind. Dies ist ein tiefer liegender mathematischer Satz, dessen Beweis wir hier nicht geben wollen, da die weiter ausholenden Überlegungen und Begriffsbildungen, die dazu erforderlich sind, später¹ in andern Zusammenhängen noch eingehend betrachtet werden müssen.

Wir formulieren und beweisen endlich den

Satz 5: *Es sei $T_s^{(k)}$ ($k = 1, 2, \dots, f$; $s = 1, 2, \dots, s_k$) ein einfaches System von hermiteschen Matrizen, bestehend aus f Teilsystemen $T_s^{(k)}$ (mit jeweils bestimmtem k) der folgenden Eigenschaften:*

a) *Zwei Matrizen aus zwei verschiedenen Teilsystemen sind stets vertauschbar:*

$$(14) \quad [T_s^{(k)}, T_r^{(l)}] = 0 \quad \text{für } k \neq l.$$

b) *Jedes einzelne Teilsystem $T_s^{(k)}$ (mit festem k) besitzt einfache Bestandteile $t_s^{(k)}$, die sich alle unitär ineinander transformieren lassen.*

Dann geht das System $T_s^{(k)}$ durch eine unitäre Transformation hervor aus der Verschmelzung der Systeme $t_s^{(k)}$.

Der *Beweis* braucht nur für den Fall $f = 2$ ausgeführt zu werden, da man nach den obigen Bemerkungen eine Verschmelzung

¹ In der geplanten Fortsetzung dieses Buchs.

von f Systemen stets durch wiederholte Verschmelzung von je zwei Systemen erzeugen kann. Wir transformieren also das System

$$T_1^{(1)}, T_2^{(1)}, \dots, T_{s_1}^{(1)}; \\ T_1^{(2)}, T_2^{(2)}, \dots, T_{s_2}^{(2)}$$

unitär derart, daß das Teilsystem $T_s^{(1)}$ ausreduziert erscheint; wegen b) kann dabei mit entsprechender Indizierung der Matrixelemente die Form

$$(15) \quad \begin{cases} U^{-1} T_s^{(1)} U = (T'_s(n_1 n_2; m_1 m_2)) \\ \qquad \qquad \qquad = (\delta_{n_s m_s} t_s^{(1)}(n_1, m_1)) \end{cases}$$

hergestellt werden. Andererseits sei

$$(16) \quad U^{-1} T_r^{(2)} U = (T''_r(n_1 n_2; m_1 m_2)).$$

Aus der Voraussetzung (14) wird dann:

$$\sum_{k_1} \{ T''_r(n_1 n_2; k_1 m_2) t_s^{(1)}(k_1, m_1) - t_s^{(1)}(n_1 k_1) T''_r(k_1 n_2; m_1 m_2) \} = 0;$$

jeder Ausschnitt der Matrix $T''_r(n_1 n_2; m_1 m_2)$ mit festen n_2, m_2 ist also eine mit dem einfachen System der $t_s^{(1)}$ vertauschbare Matrix, und folglich ein Vielfaches der Einheitsmatrix. Wir erhalten:

$$T''_r(n_1 n_2; m_1 m_2) = \delta_{n_1 m_1} t''_r(n_2, m_2).$$

Das bedeutet aber, daß das System $T_s^{(1)}, T_r^{(2)}$ aus $t_s^{(1)}$ und t''_r durch Verschmelzung hervorgeht. Soll nun das Gesamtsystem *einfach* sein, so muß auch das System t''_r ($r = 1, 2, \dots$) einfach sein, also mit $t_r^{(2)}$ (bis auf eine unitäre Transformation) übereinstimmen.

§ 16. Simultane Diagonaltransformation zweier Formen. Reelle Formen. Die unitäre Diagonaltransformation einer hermiteschen Form kann man auffassen als simultane Diagonaltransformation der beiden Formen

$$(x, x) \quad \text{und} \quad (A x, x),$$

von denen die erste definit ist und von vornherein die Diagonalform hat. Man kann nun leicht sehen, daß es nicht darauf ankommt, daß die erste Form gerade die Einheitsform ist; vielmehr lassen sich zwei Formen

$$(A x, x) \quad \text{und} \quad (B x, x),$$

von denen die eine definit ist, stets gleichzeitig (durch eine im allgemeinen *nicht unitäre* Transformation) auf eine Summe von Quadraten transformieren. Sei etwa A definit, dann transformiere man zunächst A unitär auf die Diagonalform durch $x = U_1 x'$, so daß

$$(1) \quad (Ax, x) = (ax', x'), \quad U_1 U_1^\dagger = 1, \quad U_1^{-1} A U_1 = a.$$

Dann transformiere man die neue Form durch

$$x' = a^{-\frac{1}{2}} x'' \quad \text{oder} \quad a^{\frac{1}{2}} x' = x'',$$

ausführlich

$$\sqrt{a_n} x'_n = x''_n,$$

in die Einheitsform,

$$(2) \quad (ax', x') = \sum_n a_n |x'_n|^2 = (x'', x'') = \sum_n |x''_n|^2;$$

wegen der Voraussetzung, daß A definit ist, ist die Diagonalmatrix $a^{\frac{1}{2}} = (\sqrt{a_n} \delta_{nm})$ reell. Durch die Aufeinanderfolge dieser beiden Transformationen $x = U_1 a^{-\frac{1}{2}} x''$ geht B über in

$$(3) \quad B' = a^{\frac{1}{2}} U_1^{-1} B U_1 a^{-\frac{1}{2}}.$$

Diese Matrix bringe man nun unitär durch $x'' = U_2 \xi$ auf die Diagonalform

$$(4) \quad (B' x'', x'') = (b \xi, \xi),$$

wobei

$$(5) \quad (Ax, x) = (x'', x'') = (\xi, \xi)$$

erhalten bleibt.

Das Ergebnis ist also, daß durch die Transformation

$$(6) \quad x = U_1 a^{-\frac{1}{2}} U_2 \xi$$

die definite Form (Ax, x) in die Einheitsform (ξ, ξ) , die Form (Bx, x) in eine Diagonalform $(b \xi, \xi)$ übergeht.

Natürlich kann man auch *beide* Formen A und B in Diagonalmatrizen mit von 1 verschiedenen Elementen verwandeln; man braucht nur nachträglich noch mit $\xi = \sqrt{\alpha} \xi'$ zu transformieren, wo α eine Diagonalmatrix mit positiven Elementen ist, und erhält mit $\beta = \alpha b$

$$(7) \quad (Ax, x) = (\alpha \xi', \xi'), \quad (Bx, x) = (\beta \xi', \xi').$$

Da offenbar die Diagonalmatrix b , d. h. die b_n und ihre Vielfachheit, eindeutig bestimmt ist, so ist $\alpha^{-1}\beta$ unabhängig von der transformierenden Matrix.

Wir haben also bewiesen:

Satz 1: Zwei hermitesche Formen, (Ax, x) und (Bx, x) , von denen die eine positiv definit ist, lassen sich immer simultan in Summen von Quadraten mit reellen Koeffizienten transformieren:

$$(8) \quad (Ax, x) = (\alpha \xi, \xi), \quad (Bx, x) = (\beta \xi, \xi), \quad x = S \xi;$$

dabei ist das Produkt $\alpha^{-1}\beta$ von der Transformationsmatrix S unabhängig.

Ein ähnlicher Satz gilt für den Fall, daß man zwei Formen verschiedener Variabelnreihen, (Ax, x) und (By, y) hat, von denen eine definit ist; man kann zeigen, daß diese sich gleichzeitig auf die Diagonalform bringen lassen, während die bilineare Einheitsform (x, y) invariant ist, also mit Hilfe von kontragredienten Transformationen $x = S \xi$, $y = \check{S} \eta$.

Hierzu transformiere man zunächst, ganz wie oben, die Form A auf die Einheitsform mittels $x = U_1 a^{-\frac{1}{2}} x'$, wo U_1 unitär und a die Eigenwertmatrix von A ist; zugleich transformiere man y kontragredient, d. h. man setze $y = U_1 a^{\frac{1}{2}} y'$, da nach § 6, (11), (14) $U_1 = U_1$, $\check{a}^{-\frac{1}{2}} = a^{\frac{1}{2}}$ ist. Dann hat man

$$(9) \quad \begin{cases} (Ax, x) = (x', x') \\ (By, y) = (B' y', y'), \quad B' = a^{-\frac{1}{2}} U_1^{-1} B U_1 a^{\frac{1}{2}}. \end{cases}$$

Nunmehr transformiere man $(B' y', y')$ unitär durch $y' = U_2 \eta$ auf die Diagonalform $(b \eta, \eta)$ und lasse zugleich x' kontragredient in ξ übergehen; da aber $\check{U}_2 = U_2$, so hat man $x' = U_2 \xi$ und erhält

$$(10) \quad \begin{cases} (Ax, x) = (\xi, \xi), \quad x = U_1 a^{-\frac{1}{2}} U_2 \xi, \\ (By, y) = (b \eta, \eta), \quad y = U_1 a^{\frac{1}{2}} U_2 \eta. \end{cases}$$

Auch hier kann man endlich mit Hilfe der Diagonalmatrix α mit positiven Elementen durch die Transformationen

$$(11) \quad \xi = \alpha^{\frac{1}{2}} \xi', \quad \eta = \check{\alpha}^{\frac{1}{2}} \eta' = \alpha^{-\frac{1}{2}} \eta'$$

die Gestalt

$$(12) \quad \begin{cases} (A x, x) = (\alpha \xi', \xi'), \\ (B y, y) = (\beta \eta', \eta') \end{cases}$$

erreichen, wo

$$(13) \quad \beta = \alpha^{-1} b.$$

Da offenbar b eindeutig bestimmt ist, so gilt dasselbe für das Produkt $\alpha \beta$, d. h. für die Produkte entsprechender Diagonalglieder $\alpha_n \beta_n$ und ihre Vielfachheit.

Satz 2: *Zwei hermitische Formen verschiedener Variablenreihen, $(A x, x)$ und $(B y, y)$, von denen die eine positiv definit ist, lassen sich simultan kontragredient in Quadratsummen mit reellen Koeffizienten transformieren:*

$$(14) \quad \begin{cases} (A x, x) = (\alpha \xi, \xi), & x = S \xi, \\ (B y, y) = (\beta \eta, \eta), & y = \check{S} \eta; \end{cases}$$

dabei ist das Produkt $\alpha \beta = (\alpha_n \beta_n \delta_{nm})$ von der Transformationsmatrix S unabhängig. Diese Matrix S läßt nämlich noch eine gewisse Willkür zu: Man kann zunächst durch $\xi = \alpha^{-\frac{1}{2}} \xi'$ die Form $(\alpha \xi, \xi)$ in (ξ', ξ') verwandeln, sodann ξ' beliebig unitär transformieren in $\xi' = U \xi''$, endlich (ξ'', ξ'') in $(\gamma \xi''', \xi''')$ mit einer beliebigen reellen Diagonalmatrix $\gamma = (\gamma_n \delta_{nm})$ verwandeln durch $\xi'' = \gamma^{\frac{1}{2}} \xi'''$. Das gibt zusammen die Transformation $\xi = T \xi'''$ mit $T = \alpha^{-\frac{1}{2}} U \gamma^{\frac{1}{2}} = \check{T}$. Daher kann man zu S hinten den Faktor $T = \alpha^{-\frac{1}{2}} U \gamma^{\frac{1}{2}}$ hinzufügen, wo U eine beliebige unitäre und γ eine reelle Diagonalmatrix ist.

Wir haben uns in allen diesen Erörterungen auf Matrizen mit komplexen Elementen bezogen. Sämtliche Sätze gelten natürlich insbesondere auch für *reelle Matrizen*, lassen aber bei diesen noch Verschärfungen zu von folgender Art:

Reelle hermitische Matrizen sind *symmetrische Matrizen* im eigentlichen Sinne; man ordnet einer symmetrischen Matrix $A = (A_{nm})$ zweckmäßigerweise nicht eine hermitische Form $\sum_{nm} A_{nm} x_n^* x_m$ mit komplexen x_n zu, sondern eine *quadratische Form* $\sum_{nm} A_{nm} x_n x_m$ mit reellen x_n . Das skalare Produkt definiert man in diesem reellen Bereiche durch $(x, y) = \sum_n x_n y_n$; dann

kann man die quadratische Form wieder (Ax, x) schreiben. Eine unitäre Matrix mit reellen Koeffizienten ist eine *orthogonale Matrix* im gewöhnlichen Sinne. Der Satz von der Hauptachsentransformation (Einzelform-Zerlegung) hermitischer Formen verschärft sich dann in der Weise, daß eine *reelle quadratische Form durch eine orthogonale Transformation auf Hauptachsen gebracht werden kann*. Dagegen läßt der Satz von der Hauptachsentransformation *unitärer* Matrizen keine solche Verschärfung zu. Man kann zwar aus ihm schließen, daß jede reelle *orthogonale* Matrix auf die Diagonalforn gebracht werden kann; doch ist diese Transformation — von trivialen Ausnahmen abgesehen — *nicht* mehr unter Beschränkung auf das reelle Gebiet möglich.

Die in diesem Paragraphen abgeleiteten Sätze über die Simultantransformation zweier hermitischer Formen, deren eine definit ist, gelten ebenso für reelle quadratische Formen mit der Verschärfung, daß die Transformationsmatrix S reell ist, also auch ihre kontragrediente $\check{S} = \tilde{S}^{-1}$.

Drittes Kapitel.

Die Gesetze der Matrizenmechanik¹.

§ 17. Die kanonischen Vertauschungsregeln. Nachdem wir die Grundzüge der Algebra und Analysis der Matrizen kennen gelernt haben, wenden wir uns nun der eigentlichen Aufgabe zu, der Formulierung einer *Matrizenmechanik*. Diese muß eine Verallgemeinerung der durch Quantenbedingungen *ergänzten* klassischen Mechanik sein; wir werden daher die neuen Gesetze durch korrespondenzmäßige Übertragung der mechanischen Bewegungsgleichungen und der Quantenbedingungen gewinnen.

Wir beginnen mit dem einfachen Fall eines Systems von nur *einem* Freiheitsgrad. Dieses beschreiben wir, wie in der klassischen Theorie, durch zwei Größen q und p , die wir, wie dort, *Koordinate* und *Impuls* nennen, die aber nicht Zahlen, sondern *hermitische Matrizen* sein sollen:

$$(1) \quad q^\dagger = q, \quad p^\dagger = p.$$

¹ BORN, M. u. P. JORDAN: Z. Phys. Bd. 34, S. 858. 1925. DIRAC, P. A. M.: Proc. Roy. Soc. Lond. (A) Bd. 109, S. 642. 1925. BORN, M., W. HEISENBERG u. P. JORDAN: Z. Phys. Bd. 35, S. 557. 1926. Vgl. auch den Bericht von W. HEISENBERG: Math. Ann. Bd. 95, S. 683. 1926.

Dann korrespondieren sie nach § 3 mit *reellen* Größen der klassischen Theorie. Alle mechanischen Eigenschaften des Systems werden sich dann als (sich selbst adjungierte) Funktionen von q und p , also wieder als (hermitische) Matrizen darstellen lassen. Da nun in der Matrizen­theorie im allgemeinen das kommutative Gesetz der Multiplikation nicht gilt, wird die erste Frage sein: sind q und p vertauschbar oder nicht? Und im letzteren Falle, welches ist der Wert des Klammersymbols $[p, q]$?

Die Antwort hierauf erhält man durch korrespondenzmäßige Übertragung der Quantenbedingung, die für ein periodisches System in der BOHRschen Theorie lautet:

$$(2) \quad J = \oint p \, dq = h n.$$

Dabei ist also vorausgesetzt (s. Bd. 1, § 9), daß die Energiegleichung $H(p, q) = W$ geschlossene Kurven in der p, q -Ebene definiert, und die Forderung (2) besagt, daß man aus dieser Kurvenschar durch Wahl der Energiekonstanten W solche Kurven auswählen soll, für die das über einen Umlauf erstreckte Integral ein ganzzahliges Vielfaches der PLANCKschen Konstanten h wird.

Nun formulieren wir diese Bedingung um in eine Gleichung zwischen den Fourierkoeffizienten von q und p . Es sei

$$(3) \quad q(t) = \sum_{\tau} q_{\tau} e^{2\pi i \cdot \nu \tau \cdot t}, \quad p(t) = \sum_{\tau} p_{\tau} e^{2\pi i \cdot \nu \tau \cdot t};$$

hier sind p_{τ}, q_{τ}, ν Funktionen der durch (2) definierten Wirkungsvariablen J . Andererseits folgt durch Einsetzen von (3) in (2):

$$\begin{aligned} J &= \oint p \dot{q} \, dt = \int_0^{\frac{1}{\nu}} dt \sum_{\tau, \sigma} p_{\tau} 2\pi i \nu \sigma q_{\sigma} e^{2\pi i (\tau + \sigma) \nu t} \\ &= - \sum_{\tau} 2\pi i \cdot \tau \cdot p_{\tau} q_{-\tau}. \end{aligned}$$

Differenzieren wir dies nach J , so folgt

$$1 = - 2\pi i \sum_{\tau} \tau \frac{d}{dJ} (p_{\tau} q_{-\tau}).$$

Hierauf kann man nun die in § 2 an mehreren Beispielen erläuterten Übertragungsprinzipien anwenden¹: Man ersetze $p_{\tau}(J)$

¹ HEISENBERG, W.: Z. Phys. Bd. 33, S. 879. 1925.

durch $p_{n,n-\tau}, q_{-\tau}$ (J) durch $q_{n-\tau,n}$ und den Differentialoperator $\tau \frac{\partial}{\partial J}$ durch Differenzbildung. So erhält man

$$1 = -\frac{2\pi i}{h} \left\{ \sum_{\tau} p_{n+\tau,n} q_{n,n+\tau} - \sum_{\tau} p_{n,n-\tau} q_{n-\tau,n} \right\}.$$

Der Summationsbuchstabe τ durchläuft positive und negative Werte derart, daß in der ersten Summe die Zahl $n + \tau$ und in der zweiten Summe die Zahl $n - \tau$ stets einem möglichen Werte des Index (Zustandes) entspricht. Damit jetzt τ in beiden Summen denselben Wertebereich durchläuft, schreibe man im ersten Gliede $-\tau$ statt τ ; dann wird

$$(4) \quad \frac{2\pi i}{h} \sum_{\tau} (p_{n,n-\tau} q_{n-\tau,n} - q_{n,n-\tau} p_{n-\tau,n}) = 1.$$

Dies ist im wesentlichen der in § 2 bereits erwähnte THOMAS-KUHNsche *Summensatz*¹, der von diesen Verfassern auf Grund seines Zusammenhangs mit der Dispersionstheorie aufgefunden wurde (s. § 46). Der Summationsbuchstabe τ durchläuft jetzt in beiden Gliedern der Differenz positive und negative Werte derart, daß durch $n - \tau$ *jeder* stationäre Zustand des Systems dargestellt wird. Die linke Seite ist also das n -te Diagonalelement der Matrix

$$(5) \quad \frac{1}{\varkappa} (p q - q p) = [p, q],$$

des in § 4, (7) eingeführten Maßes der Nichtvertauschbarkeit, und zwar ist die Konstante [s. auch § 4, (30)]

$$(6) \quad \varkappa = \frac{h}{2\pi i},$$

ein Wert, den wir im folgenden durchweg beibehalten werden. Da die Diagonalelemente von $[p, q]$ gleich 1 sind, so sind also p und q sicher *nicht* kommutativ.

Eine Aussage über die übrigen Elemente von $[p, q]$ kann man auf diesem korrespondenzmäßigen Wege nicht bekommen. Wir machen die Hypothese (die wir später durch andere Betrachtungen stützen werden), daß alle übrigen Elemente von $[p, q]$

¹ THOMAS, W.: Naturwissensch. Bd. 13, S. 627. 1925. KUHN, W.: Z. Phys. Bd. 33, S. 408. 1925.

verschwinden¹; das heißt, die Vertauschbarkeit soll gewissermaßen „beinahe“ gewahrt sein. Demnach formulieren wir die „verschärfte“ Quantenbedingung oder besser die *kanonische Vertauschungsregel für einen Freiheitsgrad*

$$(7) \quad [p, q] = E = 1.$$

Die Funktion $[p, q]$ ist *sich selbst adjungiert* [s. § 7, (4)]; denn man hat nach § 7, (2)

$$[p, q]^\dagger = \frac{1}{\varkappa^*} (q p - p q) = \frac{1}{\varkappa} (p q - q p) = [p, q],$$

weil \varkappa nach (6) rein imaginär ist.

Daher ist die Vertauschungsregel (7) im Einklang mit der Forderung (1), daß p, q hermitische Matrizen sind.

Bei Systemen von mehreren, etwa f Freiheitsgraden benützen wir zur Beschreibung $2f$ hermitische Matrizen q_k, p_k , die f Paare konjugierter Koordinaten und Impulse bilden. Es liegt nahe, in Analogie zu (7) für diese folgende *kanonische Vertauschungsregeln*² anzusetzen:

$$(8) \quad \begin{aligned} [p_k, q_l] &= \delta_{kl}, \\ [p_k, p_l] &= 0, \quad [q_k, q_l] = 0. \end{aligned}$$

Da die Klammersymbole sämtlich sich selbst adjungierte Funktionen der p_k, q_k sind, sind diese Bedingungen im Einklang mit der Forderung, daß alle p_k, q_k hermitische Matrizen sind.

Um diese Vertauschungsregeln zu rechtfertigen, kann man zunächst den Fall betrachten, daß das mechanische System aus f Teilsystemen von je einem Freiheitsgrad besteht, die *völlig ungekoppelt* sind und einzeln durch die Matrixgrößen $q_k(n_k, m_k), p_k(n_k, m_k)$ mit

$$(9) \quad [p_k, q_k] = 1$$

dargestellt werden. Auf welche Weise man nun Matrizen zur Darstellung des Gesamtsystems zu bilden hat, erkennt man durch eine Korrespondenzüberlegung. Es seien in der klassischen Theorie f ungekoppelte Systeme von je einem Freiheitsgrad vorgelegt;

¹ BORN, M. u. P. JORDAN: Z. Phys. Bd. 34, S. 858. 1925. DIRAC, P. A. M.: Proc. Roy. Soc. Lond. (A) Bd. 109, S. 642. 1925.

² DIRAC, P. A. M.: Proc. Roy. Soc. Lond. (A) Bd. 109, S. 642. 1925. BORN, M., W. HEISENBERG u. P. JORDAN: Z. Phys. Bd. 35, S. 557. 1926.

J_k, w_k seien Wirkungs- und Winkelvariable für das k -te System, so daß q_k, p_k periodische Funktionen von w_k sind, deren FOURIER-Koeffizienten von J_k abhängen, also z. B.

$$q_k(w_k, J_k) = \sum_{\tau_k} q_{\tau_k}^{(k)}(J_k) e^{2\pi i \tau_k w_k}.$$

Dann kann man die f Systeme zu einem einzigen zusammengefaßt denken, indem man jedes der q_k, p_k formal als Funktion aller $2f$ Größen $J_1, \dots, J_f, w_1, \dots, w_f$ auffaßt, also z. B. schreibt:

$$q_k(w_1, \dots, w_f; J_1, \dots, J_f) = \sum_{\tau_1, \dots, \tau_f} q_{\tau_1, \dots, \tau_f}^{(k)}(J_1, \dots, J_f) e^{2\pi i (w_1 \tau_1 + \dots + w_f \tau_f)}.$$

Damit dies nun mit $q_k(w_k, J_k)$ identisch ist, müssen die FOURIER-Koeffizienten offenbar folgende Bedingungen erfüllen:

$$(10) \quad q_{\tau_1, \dots, \tau_f}^{(k)}(J_1, \dots, J_f) = \begin{cases} q_{\tau_k}^{(k)}(J_k), & \text{wenn } \tau_l = 0 \text{ für alle } l \neq k, \\ 0, & \text{wenn } \tau_l \neq 0 \text{ für mindestens ein } l \neq k. \end{cases}$$

Dies hat man sinngemäß in die Quantenmechanik zu übertragen: Die beiden Matrizen q_k, p_k des k -ten Teilsystems, dem der Index n_k zur Unterscheidung von Zeilen und Spalten zugehört, ersetzt man durch Matrizen Q_k, P_k , deren Zeilen und Spalten durch das Indexsystem $n_1, n_2, \dots, n_l, \dots, n_f$ unterschieden werden, entsprechend allen verschiedenen Zuständen des Gesamtsystems. Da ferner J_k mit $h n_k$ korrespondiert und τ_k mit $n_k - m_k$, so übertragen sich die klassischen Formeln (10) für die FOURIER-Koeffizienten folgendermaßen auf die Matrixelemente:

$$(11) \quad \begin{cases} Q_k(n_1, n_2, \dots, n_f; m_1, m_2, \dots, m_f) \\ P_k(n_1, n_2, \dots, n_f; m_1, m_2, \dots, m_f) \end{cases} = \begin{cases} q_k(n_k, m_k), & \text{wenn } n_l = m_l \text{ für alle } l \neq k, \\ 0, & \text{wenn } n_l \neq m_l \text{ für mindestens ein } l \neq k; \\ p_k(n_k, m_k), & \text{wenn } n_l = m_l \text{ für alle } l \neq k, \\ 0, & \text{wenn } n_l \neq m_l \text{ für mindestens ein } l \neq k. \end{cases}$$

Dies ist aber nichts anderes als die in § 15, (6) definierte *Verschmelzung* der Matrixensysteme $q_1, p_1; q_2, p_2; \dots, q_f, p_f$, und man kann mit dem dort gebrauchten Symbol Δ_k schreiben

$$(11a) \quad \begin{cases} Q_k(n_1, \dots, n_f; m_1, \dots, m_f) = \Delta_k q_k(n_k, m_k), \\ P_k(n_1, \dots, n_f; m_1, \dots, m_f) = \Delta_k p_k(n_k, m_k); \end{cases}$$

oder, in der Bezeichnungsweise von § 15, (12):

$$(11 \text{ b}) \begin{cases} Q_k = E^{(1)} \times \dots \times E^{(k-1)} \times q_k \times E^{(k+1)} \times \dots \times E^{(f)}, \\ P_k = E^{(1)} \times \dots \times E^{(k-1)} \times p_k \times E^{(k+1)} \times \dots \times E^{(f)}. \end{cases}$$

Aus § 15, (8) und (9) folgt dann sofort unter Voraussetzung von (9) die Gültigkeit von (8).

Daß überdies nach § 15, Satz 4, die durch (11) definierten Matrizen wieder ein einfaches System bilden, wenn jedes der Paare q_k, p_k ein einfaches System darstellt, wird sogleich von Wichtigkeit werden (§ 18).

Für eigentliche mechanische Systeme, deren Freiheitsgrade sich beeinflussen, kann man den Ansatz (8) durch eine Korrespondenzbetrachtung stützen¹. Es handelt sich darum, festzustellen, was die Gleichungen (8) im Grenzfall der klassischen Mechanik bedeuten. Hierzu muß man die Matrixelemente mit großen Werten von n und m ins Auge fassen.

Das Matrixelement x_{nm} soll in den FOURIER-Koeffizienten $x_\tau(J)$ der korrespondierenden klassischen Größe $x(J, w)$ übergehen, wo J_k, w_k die Wirkungs- und Winkelvariablen sind und $\tau = n - m$ ist. Indem wir nun, wie oben bei dem aus f ungekoppelten Einzelsystemen gebildeten System, den Index n in so viele Indizes n_1, n_2, \dots spalten, als es unabhängige Wirkungsvariable des korrespondierenden klassischen Systems gibt, setzen wir

$$(12) \quad x_{nm} = x_{n, n-\tau} = x_\tau(n\hbar) = x_\tau(J)$$

und führen den Grenzprozeß so, daß

$$(13) \quad \begin{cases} n_k \rightarrow \infty, \text{ zugleich } \hbar \rightarrow 0, \\ \text{aber } n_k \hbar = J_k \text{ endlich} \end{cases}$$

bleibt. Die mit τ, σ bezeichneten Indizes sind endliche Zahlen.

Für irgend zwei Größen x, y bilden wir nun

$$\begin{aligned} (xy - yx)_{nm} &= \sum_k x_{nk} y_{km} - \sum_l y_{nl} x_{lm} \\ &= \sum_{\tau+\sigma=n-m} (x_{n, n-\tau} y_{n-\tau, n-\tau-\sigma} - y_{n, n-\sigma} x_{n-\sigma, n-\sigma-\tau}) \\ &= \sum_{\tau+\sigma=n-m} \{ (x_{n, n-\tau} - x_{n-\sigma, n-\tau-\sigma}) y_{n-\tau, n-\tau-\sigma} \\ &\quad - (y_{n, n-\sigma} - y_{n-\tau, n-\tau-\sigma}) x_{n-\sigma, n-\tau-\sigma} \}. \end{aligned}$$

¹ DIRAC, P. A. M.: Proc. Roy. Soc. Lond. (A) Bd. 111, S. 281. 1926.

Nach (12) ist

$$\frac{1}{h}(x_{n, n-\tau} - x_{n-\sigma, n-\tau-\sigma}) = \frac{1}{h}(x_{\tau}(J) - x_{\tau}(J - \sigma h)) \rightarrow \sum_k \sigma_k \frac{\partial x_{\tau}(J)}{\partial J_k}$$

und

$$y_{n-\tau, n-\tau-\sigma} = y_{\sigma}(J - \tau h) \rightarrow y_{\sigma}(J).$$

Mithin wird

$$(14) \quad \frac{2\pi i}{h}(xy - yx)_{nm} \rightarrow 2\pi i \sum_{\substack{\tau+\sigma \\ =n-m}} \sum_k \left\{ \sigma_k y_{\sigma}(J) \frac{\partial x_{\tau}(J)}{\partial J_k} - \tau_k x_{\tau}(J) \frac{\partial y_{\sigma}(J)}{\partial J_k} \right\}.$$

Denkt man sich nun die x, y korrespondierenden klassischen Größen als mehrfache FOURIER-Reihen nach den Winkelvariablen w_k angesetzt, z. B.

$$x(w, J) = \sum_{\tau} x_{\tau}(J) e^{2\pi i(w\tau)},$$

so sind die Ableitungen $\frac{\partial x}{\partial J_k}, \frac{\partial x}{\partial w_k}$ FOURIER-Reihen derselben Form mit den Koeffizienten

$$(15) \quad \begin{cases} \left(\frac{\partial x}{\partial J_k} \right)_{\tau} = \frac{\partial x_{\tau}(J)}{\partial J_k}, \\ \left(\frac{\partial x}{\partial w_k} \right)_{\tau} = 2\pi i \tau_k x_{\tau}(J). \end{cases}$$

Daher kann man (14) so schreiben:

$$(16) \quad \frac{2\pi i}{h}(xy - yx)_{nm} \rightarrow \sum_k \sum_{\substack{\tau+\sigma \\ =n-m}} \left\{ \left(\frac{\partial y}{\partial w_k} \right)_{\sigma} \left(\frac{\partial x}{\partial J_k} \right)_{\tau} - \left(\frac{\partial x}{\partial w_k} \right)_{\tau} \left(\frac{\partial y}{\partial J_k} \right)_{\sigma} \right\}.$$

Rechts steht aber der zum Index $n - m$ gehörige FOURIER-Koeffizient von

$$\sum_k \left(\frac{\partial x}{\partial J_k} \frac{\partial y}{\partial w_k} - \frac{\partial x}{\partial w_k} \frac{\partial y}{\partial J_k} \right);$$

das ist die in der klassischen Mechanik als JACOBI-POISSON'Sches Klammersymbol $[x, y]$ bezeichnete Größe (s. Anhang 2). Da diese

kanonisch invariant ist, hat man für irgendwelche kanonischen Variablen p_k, q_k :

$$(17) \quad \frac{2\pi i}{h} (xy - yx) \rightarrow \sum_k \left[\begin{array}{cc} \frac{\partial x}{\partial p_k} & \frac{\partial x}{\partial q_k} \\ \frac{\partial y}{\partial p_k} & \frac{\partial y}{\partial q_k} \end{array} \right],$$

wobei der Pfeil bedeuten soll, daß das Element n, m der Matrix für große n, m asymptotisch in den FOURIER-Koeffizienten ($n - m$) der klassischen Größe übergeht. Es ist daher gerechtfertigt, für die korrespondierenden Größen der klassischen Theorie und der Quantentheorie dasselbe Zeichen $[x, y]$ zu verwenden.

Tut man das aber, so sieht man, daß die Vertauschungsregeln (8) mit den Identitäten korrespondieren, die man erhält, wenn man in den JACOBI-POISSONSchen Klammern x und y gleich irgendwelchen der Größen q_k, p_k setzt (s. Anhang 2)¹.

Damit sind die Vertauschungsrelationen als sinngemäße Verallgemeinerungen klassischer Formeln nachgewiesen, und zwar gerade der Bedingungen dafür, daß ein Variabelnsystem p_k, q_k *kanonisch* ist. Man wird daher auch in der Quantentheorie jedes System von Matrixpaaren p_k, q_k , die den Vertauschungsrelationen (8) genügen, als *kanonische Variable* bezeichnen. Wir werden im § 19 sehen, daß für jedes System kanonischer Variablen die Grundgleichungen der Mechanik auf eine Form gebracht werden können, die mit der kanonischen Form der klassischen Bewegungsgleichungen formal übereinstimmt.

§ 18. Folgerungen aus den Vertauschungsregeln. Wir wollen aus den Vertauschungsregeln einige Folgerungen ziehen.

Zunächst sieht man, daß keine der Matrizen q_k, p_k endlich sein darf. Denn wäre es der Fall, so wäre auch die Matrix $[p, q]$ endlich, und ihre Spur nach § 4, (22a) gleich Null; da aber $[p, q] = E$ sein soll und SpE gleich der Ordnung der Matrix ist, so hat man einen Widerspruch. Wir haben also das Ergebnis: *Die Matrizen der Quantenmechanik sind unendliche Matrizen.*

Später² werden wir allerdings sehen, daß es gelingt, auch end-

¹ Natürlich darf man nicht verlangen, daß das klassische Klammer-symbol $[x, y]$ auch für zwei beliebige mechanische Größen x und y mit dem entsprechenden quantenmechanischen $[x, y]$ genau übereinstimmt.

² In der geplanten Fortsetzung des Buches.

liche Matrizen in die Quantenmechanik einzuführen, und zwar dadurch, daß man unsere Definition kanonisch konjugierter Größen durch die Vertauschungsregeln durch eine andere, allgemeinere ersetzt. Vorläufig aber halten wir an der Unendlichkeit der Matrizen fest.

Keine der Matrizen p_k, q_k kann eine Diagonalmatrix sein, solange wir daran festhalten, daß der Index n einen diskreten Wertebereich durchläuft (wovon wir später absehen werden). Denn wäre etwa $p_k = (p_k(n) \delta_{nm})$, so hätte man

$$[p_k, q_k]_{nm} = \frac{1}{\varkappa} (p_k(n) - p_k(m)) q_k(n, m) = \delta_{nm},$$

was bei endlichem $q_k(n, m)$ für $n = m$ einen Widerspruch gibt. Dieser Umstand ist z. B. von Bedeutung für die Frage nach der Definition von Wirkungs- und Winkelvariablen J_k, w_k in der Quantenmechanik.

Wir haben nun einige formale Eigenschaften der quantentheoretischen Klammersymbole

$$(1) \quad [x, y] = \frac{1}{\varkappa} (xy - yx), \quad \varkappa = \frac{\hbar}{2\pi i}$$

anzugeben. Es gelten, wie in der klassischen Theorie [s. Anhang 2, (11)] die trivialen Identitäten

$$(2) \quad \begin{cases} [x, y] = -[y, x], \\ [x + y, z] = [x, z] + [y, z], \\ [xy, z] = [x, z]y + x[y, z]. \end{cases}$$

Aus diesen folgt, daß man auch hier die partielle Differentiation einer Funktion $x(p_1, \dots, p_f, q_1, \dots, q_f)$ nach den Argumenten auf Bildung von Klammersymbolen zurückführen kann [entsprechend den klassischen Gleichungen Anhang 2, (12)]:

$$(3) \quad \frac{\partial x}{\partial q_r} = -[x, p_r], \quad \frac{\partial x}{\partial p_r} = [x, q_r].$$

Zunächst sieht man, daß diese Gleichungen sich für $x = p$ und $x = q$ einfach auf die Vertauschungsregeln § 17, (8) reduzieren. Sodann nehme man an, sie seien für zwei Funktionen φ und ψ schon bewiesen; dann folgt aus (3) ihre Gültigkeit für $\varphi + \psi$

und $\varphi\psi$. Für $\varphi + \psi$ ist das trivial. Für $\varphi\psi$ hat man nach § 7, (15)

$$\frac{\partial \varphi\psi}{\partial q_r} = \varphi \frac{\partial \psi}{\partial q_r} + \frac{\partial \varphi}{\partial q_r} \psi = \varphi [p_r, \psi] + [p_r, \varphi] \psi.$$

Andererseits ist nach (2)

$$[\varphi\psi, p_r] = [\varphi, p_r] \psi + \varphi [\psi, p_r],$$

also

$$\frac{\partial \varphi\psi}{\partial q_r} = -[\varphi\psi, p_r].$$

In derselben Weise behandelt man die Ableitung nach p_r .

Daraus folgt, daß (3) für Polynome, also auch für beliebige, durch Potenzreihen darstellbare Funktionen gilt.

Endlich hat man auch das Analogon zur JACOBISCHEN Identität [s. Anhang 2, (17)]:

$$(4) \quad [x, [y, z]] + [y, [z, x]] + [z, [x, y]] = 0.$$

Der Beweis ist hier viel einfacher als in der klassischen Theorie: man sieht beim Ausschreiben der Klammersymbole, daß sich je zwei Glieder fortheben.

Nach Feststellung dieser formalen Regeln haben wir nun Schlüsse über den Charakter der Matrizen q_k, p_k aus den kanonischen Vertauschungsregeln zu ziehen.

Es gilt folgender wichtige

Satz 1: Wenn die hermiteschen Matrizen q_k^0, p_k^0 ($k = 1, 2, \dots, f$) ein kanonisches System bilden, also den Vertauschungsregeln

$$(5) \quad [p_k^0, q_l^0] = \delta_{kl}, \quad [p_k^0, p_l^0] = [q_k^0, q_l^0] = 0$$

genügen, und wenn U eine unitäre Matrix ist, also

$$(6) \quad U^\dagger = U^{-1},$$

so ist

$$(7) \quad q_k = U^{-1} q_k^0 U, \quad p_k = U^{-1} p_k^0 U$$

wieder ein hermitesches, kanonisches Matrixensystem.

Man sieht nämlich nach § 7, Satz 1 sofort, daß aus (7) auch für die q_k, p_k die Gültigkeit der kanonischen Vertauschungsregeln folgt. Daß die p_k, q_k überdies hermitisch werden, folgt aus § 6, Satz c, (17).

Wenn ein kanonisches Matrizenystem q_k, p_k zerlegbar (reduzibel) ist (vgl. § 15), also mit geeignetem unitärem U in der Form (7) geschrieben werden kann derart, daß

$$(8) \quad q_k^0 = \begin{pmatrix} q_k^{(1)} & 0 \\ 0 & q_k^{(2)} \end{pmatrix}, \quad p_k^0 = \begin{pmatrix} p_k^{(1)} & 0 \\ 0 & p_k^{(2)} \end{pmatrix}$$

ist, so sind die irreduziblen Bestandteile des Systems q_k^0, p_k^0 selbst wieder kanonische Matrizenysteme. Zur vollständigen mathematischen Untersuchung der kanonischen Matrizenysteme genügt also die Betrachtung der *einfachen* kanonischen Matrizenysteme.

Wir werden im folgenden stets voraussetzen, daß die kanonischen Matrizen q_k, p_k ein einfaches oder irreduzibles System bilden.

Nunmehr können wir beweisen, daß Transformationen, die einem Matrizenystem q_k, p_k den kanonischen Charakter und gleichzeitig die Realität wahren sollen, notwendig unitär sein müssen:

Satz 2: *Soll ein (einfaches, hermitisches) kanonisches Matrizenystem q_k^0, p_k^0 durch eine Transformation*

$$(9) \quad q_k = U^{-1} q_k^0 U, \quad p_k = U^{-1} p_k^0 U$$

wieder in hermitische Matrizen q_k, p_k übergehen, so unterscheidet sich U von einer unitären Matrix nur durch einen Zahlenfaktor. Denn aus § 6, Satz d, folgt, daß die Matrix $UU^\dagger = V$ mit allen q_k^0, p_k^0 vertauschbar ist. Ferner ist $V^\dagger = (UU^\dagger)^\dagger = UU^\dagger = V$, d. h. V ist hermitisch. Folglich ist nach § 15, Satz 1, V bis auf einen Zahlenfaktor a gleich der Einheitsmatrix, also $UU^\dagger = aE$. Hier ist a erstens reell wegen des hermitischen Charakters von $V = UU^\dagger$, und zweitens auch positiv, da UU^\dagger definit ist (vgl. § 10, S. 53). Dividiert man also U durch die reelle Zahl \sqrt{a} , so wird es unitär.

Für die späteren Entwicklungen ist die Frage wichtig, ob es in Umkehrung der Sätze 1, 2 stets möglich ist, zwei einfache kanonische Matrizenysteme q_k, p_k und q_k^0, p_k^0 durch eine unitäre Transformation ineinander überzuführen. Mit gewissen Einschränkungen werden wir diese Frage später (§ 23) bejahen können.

Matrizentransformationen der Gestalt (9) treten in der Quantenmechanik in ganz verschiedenen Bedeutungen auf, die scharf auseinander gehalten werden müssen. Erstens können wir für

die unitäre Matrix U eine *Funktion der Matrizen* p_k^0, q_k^0 wählen; dabei werden die p_k^0, q_k^0 als *unbestimmte* (variable) Matrizen aufgefaßt, die lediglich den kanonischen Vertauschungsregeln genügen müssen. Dann werden auch die p_k, q_k gewisse Funktionen der p_k^0, q_k^0 :

$$(10) \quad p_k = p_k(p_k^0, q_k^0), \quad q_k = q_k(p_k^0, q_k^0),$$

welche die Eigenschaft besitzen, ihrerseits wieder die kanonischen Vertauschungsregeln zu befriedigen. Als Umkehrung der Gleichungen (10) werden gewisse Gleichungen

$$(10') \quad p_k^0 = p_k^0(p_k, q_k), \quad q_k^0 = q_k^0(p_k, q_k)$$

die p_k^0, q_k^0 als Funktionen der p_k, q_k darstellen. Jede als Funktion der kanonischen Grundgrößen p_k^0, q_k^0 definierte mechanische Größe kann auch als Funktion der p_k, q_k dargestellt werden (und umgekehrt). Die Einführung der p_k, q_k an Stelle der p_k^0, q_k^0 entspricht dem klassischen Begriff der *kanonischen Transformationen* (z. B. Übergang von rechtwinkligen Koordinaten zu schiefwinkligen oder zu Polarkoordinaten). Ihre Darstellung durch die Formeln (9) ist allerdings äußerlich ganz verschieden von allen in der gewöhnlichen Mechanik bekannten Darstellungen kanonischer Transformationen. Doch besteht trotzdem ein enger Korrespondenzzusammenhang, auf den wir später zurückkommen.

In ganz anderem Sinne werden wir in Kap. VI *Matrizentransformationen* (9) betrachten. Wir werden dort nämlich, wie schon zu Anfang des § 9 angedeutet wurde, nicht mehr eindeutig bestimmte Matrizen, sondern *Tensoren des HILBERT-Raums* als mathematische Darstellungen physikalischer Größen benutzen; dabei betrachten wir q_k und q_k^0 als verschiedene Matrixdarstellungen (Komponentenschemata) ein und desselben Tensors, also ein und derselben physikalischen Größe, in verschiedenen Koordinatensystemen des HILBERT-Raums. Die Gleichung (9) mit einer bestimmten vorgegebenen Matrix U entspricht dann einer bestimmten unitären Transformation des HILBERT-Raumes.

Wir wollen hier noch einen Spezialfall der *kanonischen Transformationen* betrachten, nämlich linear-homogene „Punkttransformationen“, bei der die q_k linear-homogene Funktionen der q_k^0 sind. Damit diese sich selbst adjungiert sind, müssen die Koeffizienten reelle Zahlen sein. Unsere Aufgabe ist die Bestimmung

geeigneter p_k zu den q_k . Diese Aufgabe kann durch Aufsuchung einer geeigneten Matrix U entsprechend (7) gelöst werden, doch schlagen wir hier einen anderen Weg ein. Ist f die Anzahl der Freiheitsgrade und $S = (S_{kl})$ eine beliebige *reelle* Matrix von f Reihen und f Spalten,

$$\check{S} = (\check{S}_{kl})$$

ihre kontragrediente Matrix, so behaupten wir, daß die lineare Transformation

$$(11) \quad q_k = \sum_{l=1}^f S_{kl} q_l^0, \quad p_k = \sum_{l=1}^f \check{S}_{kl} p_l^0$$

oder kurz

$$(11a) \quad q = S q^0, \quad p = \check{S} p^0$$

die Vertauschungsregeln in sich überführt. Zunächst sind nämlich alle q_k unter sich und alle p_k unter sich vertauschbar, wenn es die q_k^0 und p_k^0 sind. Sodann hat man

$$\begin{aligned} p_k q_l - q_l p_k &= \sum_{mn} (\check{S}_{km} p_m^0 S_{ln} q_n^0 - S_{ln} q_n^0 \check{S}_{km} p_m^0) \\ &= \sum_{mn} \check{S}_{km} S_{ln} (p_m^0 q_n^0 - q_n^0 p_m^0) \\ &= \varkappa \sum_n \check{S}_{kn} S_{ln}; \end{aligned}$$

nun ist aber nach Definition [§ 6, (11)] $\check{S} S^+ = 1$ oder, da S reell ist, $\check{S} \check{S} = 1$, d. h.

$$\sum_n \check{S}_{kn} S_{ln} = \delta_{kl}.$$

Daher erhält man in der Tat

$$[p_k, q_l] = \delta_{kl}.$$

Ist insbesondere S eine *reelle, orthogonale Matrix* O , so ist [s. § 6, (14)] $\check{O} = O$ und man hat statt (11a)

$$(12) \quad q = O q^0, \quad p = O p^0.$$

Das Ergebnis formulieren wir als

Satz 3: *Eine lineare reelle Transformation der Koordinatenmatrizen q_k und eine dazu kontragrediente Transformation der Impulsmatrizen p_k läßt die Vertauschungsregeln invariant. Ins-*

besondere gilt das für simultane, reelle, orthogonale Transformationen der q_k und p_k .

Dieser Satz ist offenbar wesentlich dafür, daß man die kanonischen Vertauschungsregeln auf die *rechtwinkligen Raumkoordinaten* x, y, z eines Massenpunktes und die zugehörigen Impulse p_x, p_y, p_z anwenden kann: *Bei einer Drehung des dreidimensionalen räumlichen Koordinatensystems bleiben, wie es verlangt werden muß, nach Satz 3 die Vertauschungsregeln invariant.*

§ 19. Die kanonischen Bewegungsgleichungen für abgeschlossene Systeme. Wir gelangen nunmehr zur Übertragung der Grundgleichungen der klassischen Mechanik in die Matrizenlehre.

Ein bestimmtes System (etwa ein Wasserstoffatom) wird in seinen mechanischen Eigenschaften am einfachsten dadurch beschrieben, daß man die *Energie als Funktion der Koordinaten und Impulse*, die sogenannte HAMILTONSche Funktion, angibt:

$$(1) \quad H(q, p) = H(q_1, \dots, q_f; p_1, \dots, p_f);$$

wir nehmen hier zunächst an, daß diese HAMILTONSche Funktion nicht explizit von der Zeit abhängig sei. H soll also in bestimmter, eindeutig definierter Weise aus den Argumentmatrizen q_k, p_k durch die Rechenprozesse der Matrizenaddition und Multiplikation aufgebaut sein. Auf welche Weise man zu der Kenntnis kommt, daß gerade eine bestimmte Funktion H für ein physikalisch gegebenes System die richtige Energiefunktion ist, diese Frage liegt — genau wie in der klassischen Theorie — außerhalb des Rahmens der Mechanik im engeren Sinne. Man muß annehmen, daß die Funktion H entweder aus einer tiefer gehenden physikalischen Theorie — etwa einer quantenmäßigen Formulierung der Elektronentheorie — deduktiv ableitbar ist; oder aber, daß man, geleitet durch Analogien zur klassischen Mechanik, auf induktivem Wege, durch Probieren und Vergleich der Folgerungen mit der Erfahrung, die richtige Form von H gefunden hat. Wir werden auf diese Fragen noch verschiedentlich zurückkommen. An dieser Stelle wollen wir lediglich untersuchen, wie bei gegebener Energiefunktion H die Quantenmechanik aufzubauen ist.

Durch eine Bedingung haben wir freilich die Willkür in der Wahl von H einzuschränken: H soll eine sich selbst adjungierte Funktion der Argumentmatrizen sein. Denn dann wird H eine

hermitesche Matrix, sobald man hermitesche Matrizen für q_k, p_k einsetzt, und korrespondiert mit einer *reellen* Energiefunktion der klassischen Theorie.

Nunmehr kann man *die kanonischen Bewegungsgleichungen* ohne Änderung aus der klassischen Mechanik übernehmen:

$$(2) \quad \dot{q}_k = \frac{\partial H}{\partial p_k}, \quad \dot{p}_k = -\frac{\partial H}{\partial q_k}.$$

Neben diesen Gleichungen haben wir entsprechend dem HEISENBERGSCHEN Ansatz (13), § 3 die Beziehungen

$$(3) \quad \begin{cases} \dot{q}_k(nm) = 2\pi i \nu(nm) q_k(nm), \\ \dot{p}_k(nm) = 2\pi i \nu(nm) p_k(nm), \\ h\nu(nm) = W_n - W_m \end{cases}$$

mit zunächst unbestimmt gelassenen Termen $\frac{1}{h} W_n$ zu fordern.

Verbinden wir diese Gleichungen (2), (3) mit der Forderung, daß die Matrizen q_k, p_k ein einfaches kanonisches System bilden sollen, so können wir folgende Schlüsse ziehen:

Gemäß § 4, (31) ist auf Grund von (3):

$$(4) \quad \dot{q}_k = [W, q_k], \quad \dot{p}_k = [W, p_k],$$

wo W wie in § 4, (29) die Diagonalmatrix

$$(5) \quad W = (W_n \delta_{nm})$$

bezeichnet.

Gemäß § 18, (3) ist andererseits auf Grund von (2):

$$(6) \quad \dot{q}_k = [H, q_k], \quad \dot{p}_k = [H, p_k].$$

Es wird dann nach (4), (6):

$$[W - H, q_k] = 0, \quad [W - H, p_k] = 0.$$

Da aber die q_k, p_k ein *einfaches* Matrixensystem bilden, muß, nach § 15, Satz 1, $W - H$ ein Vielfaches der Einheitsmatrix sein; also

$$(7) \quad H = W + aE$$

mit einem reellen Zahlfaktor a .

In dieser Gleichung (7) sind in Rücksicht auf (5) zwei wichtige physikalische Sätze enthalten:

I. *Energiesatz*: Die Energie H ist eine Diagonalmatrix, also zeitlich konstant.

II. *BOHRsche Frequenzbedingung*: Die mit h multiplizierten Spektraltermine (die W_n) stimmen bis auf eine additive Konstante überein mit den Energiewerten $H(nn)$ des Systems.

Da die additive Konstante in den Spektraltermen willkürlich ist, können und wollen wir im folgenden stets die W_n so normieren, daß

$$(8) \quad H = W$$

wird.

Es ist nun für die weitere Entwicklung von entscheidender Bedeutung, daß man die eben ausgeführte Überlegung umkehren kann. Wir betrachten¹ nämlich jetzt als ursprüngliche Forderungen die, daß erstens die q_k, p_k ein einfaches kanonisches System bilden sollen, und daß zweitens die Energie $H(q, p)$ eine Diagonalmatrix sei:

$$(9) \quad H = W = \text{Diagonalmatrix.}$$

Ferner definieren wir gemäß (3) die Frequenzen $\nu(nm)$ und die zeitlichen Ableitungen \dot{q}_k, \dot{p}_k . Man sieht dann: Als Folgerungen ergeben sich die kanonischen Bewegungsgleichungen (2). Diese Formulierungsweise werden wir im folgenden als die eigentlich sinngemäße erkennen.

Es folgt jetzt sofort: Die kanonischen Bewegungsgleichungen (2) sind invariant gegen kanonische Transformationen. Es sei nämlich

$$(10) \quad Q_k = Q_k(q, p), \quad P_k = P_k(q, p)$$

eine kanonische Transformation im Sinne von § 18, d. h. es mögen die Funktionen Q_k, P_k der q_l, p_l stets ein kanonisches System bilden, wenn die Argumentmatrizen q_l, p_l ein kanonisches System sind². Ferner sei auf Grund von (10)

$$(11) \quad H'(Q, P) = H(q, p).$$

Dann sind die kanonischen Gleichungen (2) in den q_k, p_k gleich-

¹ Vgl. BORN, M. u. P. JORDAN: Z. Phys. Bd. 34, S. 858. 1925.

² Die oben besprochene Frage, ob bzw. unter welchen Bedingungen (10) in der Form $Q_k = U q_k U^{-1}$, $P_k = U p_k U^{-1}$ geschrieben werden kann, spielt an dieser Stelle keine Rolle.

wertig mit den kanonischen Gleichungen

$$(12) \quad \dot{Q}_k = \frac{\partial H'}{\partial P_k}, \quad \dot{P}_k = -\frac{\partial H'}{\partial Q_k}$$

in den Q_k, P_k . Das folgt unmittelbar daraus, daß wir die Gleichungen (2), also auch entsprechend (12), ersetzen können durch die Forderung (9), deren Inhalt wegen (11) unabhängig ist von der Wahl des kanonischen Systems der Argumentmatrizen.

Es seien $H^{(1)}(q^{(1)}, p^{(1)})$, $H^{(2)}(q^{(2)}, p^{(2)})$, . . . , $H^{(r)}(q^{(r)}, p^{(r)})$ die HAMILTON-Funktionen von r verschiedenen, völlig getrennten mechanischen Systemen; dabei möge das s -te System f Freiheitsgrade haben. Es wird also zum s -ten System, wenn es isoliert betrachtet wird, ein gewisses kanonisches Matrixsystem

$$q_k^{(s)}, p_k^{(s)} \quad (k = 1, 2, \dots, f_s)$$

gehören, dessen Einsetzung als Argumentsystem die Funktion $H^{(s)}(q, p)$ zur Diagonalmatrix macht:

$$(13) \quad H^{(s)}(q^{(s)}, p^{(s)}) = W^{(s)}.$$

Die zum Matrixsystem

$$q_k^{(s)}, p_k^{(s)} \quad (k = 1, 2, \dots, f_s)$$

gehörige Einheitsmatrix sei $E^{(s)}$. Wir wollen jetzt aber alle r Systeme zusammengenommen als ein einziges System auffassen. Wir haben dann entsprechend den Betrachtungen von § 17 statt der r einzelnen Matrixsysteme $q_k^{(s)}, p_k^{(s)}$ ihre *Verschmelzung* einzuführen. Dabei möge jeweils $Q_k^{(s)}$ bzw. $P_k^{(s)}$ an Stelle von $q_k^{(s)}, p_k^{(s)}$ treten. Daß die durch den Prozeß der Verschmelzung gebildeten Matrizen $Q_k^{(s)}, P_k^{(s)}$ ein kanonisches System bilden, nachdem die einzelnen Systeme $q_k^{(s)}, p_k^{(s)}$ kanonisch waren, folgt aus § 15, Satz 2.

Nunmehr aber ist auch zu sehen, daß die neuen $Q_k^{(s)}, P_k^{(s)}$ — wie es physikalisch verlangt werden muß — die zur HAMILTON-Funktion

$$(14) \quad H(Q, P) = \sum_{s=1}^r H^{(s)}(Q^{(s)}, P^{(s)})$$

gehörigen *kanonischen Bewegungsgleichungen* ohne weiteres erfüllen, nachdem für jedes s die Gleichung (13) erfüllt war. Es werden nämlich die Summanden in (14) einzeln Diagonalmatrizen:

$$(15) \quad H^{(s)}(Q^{(s)}, P^{(s)}) = E^{(1)} \times \dots \times E^{(s-1)} \times W^{(s)} \times E^{(s+1)} \times \dots \times E^{(r)}.$$

Wir haben in § 3 hervorgehoben, daß die ganze hier entwickelte Theorie, die nur Matrizen mit diskreten Indizes benützt, nur zur Beschreibung „periodischer“ Bewegungen geeignet ist; wir haben ferner angedeutet, daß man sich durch Benützung von allgemeineren Matrizen mit kontinuierlich veränderlichen Indizes von dieser Beschränkung befreien und aperiodische Vorgänge in formal ähnlicher Weise behandeln kann. Hier jedoch wollen wir noch den quantenmechanischen *Virialsatz* aussprechen und beweisen, der in charakteristischer Weise auf der Voraussetzung „periodischer“ Bewegungen bzw. diskreter Matrizen beruht und sich keineswegs auf aperiodische Bewegungen ausdehnen läßt. Wir behaupten¹:

Bei einer Energiefunktion der Form

$$(16) \quad H = T(p) + \Phi(q) = \frac{1}{2} \sum_k \frac{p_k^2}{m_k} + \Phi(q),$$

bei welcher die potentielle Energie $\Phi(q)$ eine homogene Funktion ϱ -ten Grades der Koordinaten ist, besteht zwischen den Zeitmittelwerten \bar{T} , $\bar{\Phi}$ der kinetischen und der potentiellen Energie die Beziehung

$$(17) \quad \bar{T} = \frac{\varrho}{2} \bar{\Phi}.$$

Denn es gilt nach den kanonischen Bewegungsgleichungen

$$\begin{aligned} \frac{d}{dt} \sum_k p_k q_k &= \sum_k (\dot{p}_k q_k + p_k \dot{q}_k) \\ &= \sum_k \left(-q_k \frac{\partial \Phi}{\partial q_k} + \frac{p_k^2}{m_k} \right); \end{aligned}$$

und nach der EULERSchen Formel

$$\Phi(q) = \frac{1}{\varrho} \sum_k q_k \frac{\partial \Phi}{\partial q_k}$$

wird

$$(18) \quad \frac{d}{dt} \sum_k p_k q_k = 2T - \varrho \Phi.$$

¹ BORN, M., W. HEISENBERG u. P. JORDAN: Z. Phys. Bd. 35, S. 557. 1926.

Nun gilt, wie schon aus § 3, (12), (13) hervorgeht, für jede diskrete Matrix a :

$$(19) \quad \overline{\frac{d}{dt} a} = 0,$$

woraus die Behauptung (17) folgt. Man sieht, daß der wesentliche Punkt des Beweises die Annahme „periodischer“ p_k, q_k ist. Der Satz selbst sowie seine Beschränkung auf periodische (bzw. mehrfach periodische) Bewegungen besteht ebenso in der klassischen Theorie.

Als Beispiel nehmen wir den *harmonischen Oszillator*, der ausführlich in § 23 behandelt werden wird, oder allgemeiner ein *System von harmonischen Oszillatoren* (§ 24). Hier ist $\varrho = 2$; es wird also (wie in der klassischen Theorie):

$$(20) \quad \overline{T} = \overline{\Phi}.$$

Als Verallgemeinerung von (17) für den Fall einer *beliebigen* potentiellen Energie $\Phi(q)$ gilt offenbar¹:

$$(21) \quad \overline{T} = \frac{1}{2} \sum_k \overline{q_k \frac{\partial \Phi}{\partial q_k}}.$$

§ 20. Existenz und Bestimmtheit der Lösung der Bewegungsgleichungen. Zurückführung auf ein Eigenwertproblem². Die kanonischen Bewegungsgleichungen führten uns auf das Problem der Konstruktion eines kanonischen Matrixsystems q_k, p_k , welches für eine bestimmte HAMILTON-Funktion die Beziehung

$$(1) \quad H(q, p) = W = \text{Diagonalmatrix}$$

befriedigt. Wir haben nun zu zeigen, daß dieses mathematische Problem im allgemeinen (abgesehen von etwaigen singulären Fällen usw.) wirklich lösbar ist. Allerdings werden wir uns dabei vorläufig mit rein formalen Überlegungen begnügen müssen, ohne auf Konvergenzfragen usw. einzugehen.

Wir setzen voraus, daß jedenfalls ein kanonisches Matrixsystem konstruierbar sei. Später (§ 23) werden wir in der Tat ein spezielles kanonisches Matrixsystem von einem Freiheits-

¹ FINKELSTEIN, B. N.: Z. Phys. Bd. 50, S. 293. 1928.

² BORN, M., W. HEISENBERG u. P. JORDAN: Z. Phys. Bd. 35, S. 557. 1926.

grade explizit angeben; durch den Prozeß der Verschmelzung können wir dann gemäß § 17 ein kanonisches System von f Freiheitsgraden gewinnen.

Es sei also q_k^0, p_k^0 ein spezielles kanonisches System; wir gebrauchen die Bezeichnung

$$(2) \quad H(q^0, p^0) = H_0.$$

Dann suchen wir eine Lösung der Bewegungsgleichungen bzw. der Gleichung (1) durch den Ansatz

$$(3) \quad q_k = U^{-1} q_k^0 U, \quad p_k = U^{-1} p_k^0 U$$

mit $U^{-1} = U^\dagger$.

Diese q_k, p_k werden dann und nur dann (1) befriedigen, wenn

$$(4) \quad U^{-1} H_0 U = W = \text{Diagonalmatrix}$$

wird. Damit ist aber das Problem der Lösung der Bewegungsgleichungen *zurückgeführt auf das Problem der Hauptachsentransformation der hermiteschen Matrix H_0* .

Wir haben dieses mathematische Problem in § 11, Kap. II, ausführlich behandelt für *endliche* Matrizen; es ist damals auch gesagt worden, daß die mathematische Theorie sich entsprechend auf unendliche Matrizen ausdehnen läßt. Das mathematische Problem, auf das wir durch die physikalisch sinngemäße Fortentwicklung der HEISENBERGSchen Ansätze zur Quantenmechanik geführt wurden, ist also wirklich ein allgemein lösbares Problem. Allerdings muß, wie schon in § 11 bemerkt wurde, mit der Möglichkeit eines *kontinuierlichen Eigenwertspektrums* der Matrix H_0 gerechnet werden. Physikalisch bedeutet das, entsprechend unseren früheren Bemerkungen, das Auftreten *aperiodischer* Bewegungen (z. B. Hyperbelbahnen beim H -Atom). Wir werden später auf diese Verhältnisse ausführlicher eingehen. Hier wollen wir uns zunächst auf Systeme mit diskreten Energiewerten beschränken, weil bei diesen viele physikalische Zusammenhänge in besonders durchsichtiger Form erläutert werden können.

Es wird in § 23 bewiesen, daß zwei verschiedene kanonische Systeme von je f Freiheitsgraden unter gewissen zusätzlichen Bedingungen (die ganz dem physikalischen Sinn der Sache entsprechen) gewiß unitär ineinander transformiert werden können. Infolgedessen erhalten wir durch die soeben beschriebene Konstruktion *alle* zulässigen Lösungen unseres Problems; und wir

können die *Mannigfaltigkeit der Lösungen* unseres Problems übersehen, indem wir eben die Mannigfaltigkeit der durch diese Konstruktion gelieferten Lösungen betrachten.

Willkürlich ist zunächst die *Reihenfolge* (Numerierung) der Eigenwerte der Energie; wir dürfen also voraussetzen, daß W als *geordnete* Diagonalmatrix geschrieben sei, so daß übereinstimmende Eigenwerte stets zusammenliegen. Die Vielfachheit (das Gewicht) des r -ten Eigenwertes sei g_r .

Wir wollen g_r als endlich voraussetzen, oder wenigstens annehmen, daß die Sätze von § 11 anwendbar sind.

Ist p_k, q_k eine Lösung des mechanischen Problems, so erhält man nach § 11, Satz 2 die *allgemeinste Lösung* in der Form

$$(5) \quad P_k = \mathring{U}^{-1} p_k \mathring{U}, \quad Q_k = \mathring{U}^{-1} q_k \mathring{U},$$

dabei ist \mathring{U} eine *unitäre Stufenmatrix* mit den Stufengrößen g_r :

$$(6) \quad \mathring{U} = \begin{pmatrix} U^{(1)} & 0 & \dots \\ 0 & U^{(2)} & \dots \\ \dots & \dots & \dots \end{pmatrix},$$

wo $U^{(r)}$ eine g_r -reihige unitäre Matrix bedeutet. Ist insbesondere jeder Eigenwert *einfach*, so bleibt als Willkür nur eine *Phasenmatrix*.

Spaltet man nun die p_k, q_k ; P_k, Q_k in rechteckige Matrizen mit g_s Zeilen und g_r Spalten entsprechend der Vielfachheit der Eigenwerte, also z. B.

$$(7) \quad p_k = (p_k^{(r,s)}), \quad p_k^{(r,s)} = (p_k^{(r,s)}(\varrho, \sigma)), \quad \begin{matrix} \varrho = 1, 2 \dots g_r, \\ \sigma = 1, 2 \dots g_s, \end{matrix}$$

so erhält man nach § 5, (16)

$$(8) \quad \begin{cases} P_k^{(r,s)} = (U^{(r)})^{-1} p_k^{(r,s)} U^{(s)}, \\ Q_k^{(r,s)} = (U^{(r)})^{-1} q_k^{(r,s)} U^{(s)}. \end{cases}$$

Wir suchen nun die *einfachsten, von jeder Unbestimmtheit freien Funktionen der Matrixelemente* der p_k, q_k . Wir nennen diese *relativ-invariant* bezüglich des mechanischen Systems. Die allereinfachsten sind

$$(9) \quad Sp p_k^{(r,r)}, \quad Sp q_k^{(r,r)}.$$

Denn $p_k^{(r,r)}$, $q_k^{(r,r)}$ sind (quadratische, hermitesche) Matrizen, die sich nach (8) unitär transformieren, wobei nach § 4, (23) ihre Spur invariant bleibt.

Die nächst einfachen Größen dieser Art sind

$$(10) \quad Sp(q_k^{(r,s)} q_k^{(s,r)}), \quad Sp(p_k^{(r,s)} p_k^{(s,r)}).$$

Denn nach (8) gelten die Transformationsgleichungen

$$Q_k^{(r,s)} Q_k^{(s,r)} = (U^{(r)})^{-1} q_k^{(r,s)} q_k^{(s,r)} U^{(r)},$$

$$P_k^{(r,s)} P_k^{(s,r)} = (U^{(r)})^{-1} p_k^{(r,s)} p_k^{(s,r)} U^{(r)};$$

die Spuren (10) sind also nach § 4, (23) invariant.

Die Spuren

$$(10a) \quad Sp(q_k^{(r,s)} q_k^{(s,r)}) = \sum_{\varrho=1}^{g_r} \sum_{\sigma=1}^{g_s} |q_k^{(r,s)}(\varrho, \sigma)|^2$$

sind sämtlich reell und positiv. Sind speziell alle Eigenwerte einfach, so reduzieren sie sich auf die quadrierten Beträge $|q_k(n, m)|^2$ aller Elemente von q_k .

Ist allgemeiner $f(q, p)$ irgend eine Funktion der q_k, p_k , so kann man auch diese in rechteckige Matrizen entsprechend den Vielfachheiten g_r der Eigenwerte zerlegen:

$$(11) \quad f(q, p) = (f^{(r,s)}(q, p)).$$

Dann transformieren sich die $f^{(r,s)}$, wie die Teilmatrizen der Koordinaten und Impulse selber, nach (8). Also ist

$$(12) \quad Sp f^{(r,s)}$$

eine *Relativinvariante*.

Ist $\varphi(q, p)$ eine zweite solche Funktion, so gilt dasselbe für

$$(13) \quad Sp(f^{(r,s)} \varphi^{(s,r)}) = \sum_{\varrho=1}^{g_r} \sum_{\sigma=1}^{g_s} f_{\varrho\sigma}^{(r,s)} \varphi_{\sigma\varrho}^{(s,r)}.$$

Ist insbesondere $f = \varphi$, so wollen wir für diese Größe ein besonderes Zeichen benutzen:

$$(14) \quad \|f^{(r,s)}\|^2 = Sp(f^{(r,s)} f^{(s,r)}) = \sum_{\varrho=1}^{g_r} \sum_{\sigma=1}^{g_s} |f_{\varrho\sigma}^{(r,s)}|^2.$$

Sie ist immer reell und positiv. *Alle physikalisch beobachtbaren Eigenschaften müssen durch willkürfreie Relativinvarianten, wie insbesondere (12), (13), (14) darstellbar sein.*

Bei fehlender Entartung der Energie ($g_r = 1$) gehen sie in

$$f(n, m) \varphi(m, n) \quad \text{bzw.} \quad |f(n, m)|^2$$

über.

Die einfachste Anwendung hiervon bekommen wir, wenn wir an den Ausgangspunkt HEISENBERGS bei der Aufstellung der Quantenmechanik denken (s. § 2). Er ordnet den Komponenten des *elektrischen Moments* \mathfrak{P} des mechanischen Systems je eine Matrix mit den Elementen

$$(15) \quad \mathfrak{P}_x(n, m) = X_{nm}, \quad \mathfrak{P}_y(n, m) = Y_{nm}, \quad \mathfrak{P}_z(n, m) = Z_{nm}$$

zu; sind e_k die Ladungen, x_k, y_k, z_k die rechtwinkligen Koordinaten der Teilchen ($k = 1, 2 \dots$), so ist

$$(16) \quad \mathfrak{P}_x = X = \sum_k e_k x_k, \dots,$$

wo mit den x_k, \dots auch die $\mathfrak{P}_x = X, \dots$ Matrizen sind. Die pro Zeiteinheit emittierte Energie beim Übergang $n \rightarrow m$ soll dann nach § 2, (1)

$$(17) \quad J_{nm} = \frac{\Delta E_{nm}^{(e)}}{\Delta t} = \frac{64 \pi^4}{3 c^3} \nu_{n,m}^4 |\mathfrak{P}_{nm}|^2$$

sein, wo

$$(18) \quad |\mathfrak{P}_{nm}|^2 = |X_{nm}|^2 + |Y_{nm}|^2 + |Z_{nm}|^2$$

ist.

Diese Annahme ist zulässig, wenn das System nicht entartet ist; denn dann sind die Größen $|\mathfrak{P}_{nm}|^2$ eindeutig durch das mechanische System bestimmt. Im Falle der Entartung aber ist das nicht der Fall; offenbar hat man dann an die Stelle von $|\mathfrak{P}_{nm}|^2$ die Größe

$$(19) \quad \|\mathfrak{P}^{(r,s)}\|^2 = \|X^{(r,s)}\|^2 + \|Y^{(r,s)}\|^2 + \|Z^{(r,s)}\|^2$$

zu setzen; dann gehört zu jedem *Übergang zwischen Zuständen verschiedener Energie*, $r \rightarrow s$, eine *eindeutig bestimmte Lichtintensität*

$$(20) \quad J^{(r,s)} = \frac{64 \pi^4}{3 c^3} \nu^{(r,s)^4} \|\mathfrak{P}^{(r,s)}\|^2.$$

Dieser Umstand hat aber eine einfache physikalische Bedeutung. Denkt man sich die Entartung durch irgend eine kleine Störung, im Falle einer *Richtungsentartung* (vgl. § 28) z. B. durch ein

äußeres elektrisches oder magnetisches Feld, aufgehoben, so wird jeder g_r -fache Term in g_r Einzelterme aufgespalten; zwischen zwei solchen Termsystemen r und s gibt es dann maximal $g_r \cdot g_s$ Übergänge.

Die relativen Intensitäten der Einzellinien (Komponenten) einer solcherart aufgespaltenen Spektrallinie sind nun in verschiedenen Fällen vollkommen verschieden, je nach der qualitativen Natur der Störung, welche die Aufspaltung verursacht. Man muß aber offenbar vom Standpunkte der Physik folgende Gesetzmäßigkeit erwarten, die von HEISENBERG¹ schon vor Aufstellung der Quantenmechanik als *Satz von der spektroskopischen Stabilität* ausgesprochen worden ist:

Die Summe der Intensitäten der Komponenten einer aufgespaltenen Spektrallinie strebt bei abnehmender Größe der Aufspaltung einem von der Art der Aufspaltung unabhängigen Grenzwert zu; dieser ist gleich der Intensität der unaufgespaltenen Linie des ungestörten (entarteten) Systems.

Wäre es nicht so, dann könnte man an einer unaufgespaltenen Spektrallinie durch Erzeugung einer geeigneten unendlich kleinen Aufspaltung eine endliche Intensitätsänderung erhalten.

Dieser Satz ergibt sich also als Folgerung der exakten deduktiven Theorie², nämlich daraus, daß im entarteten Grenzfall

$$\| \mathfrak{P}^{(r, s)} \|^2 = \sum_{\rho=1}^{g_r} \sum_{\sigma=1}^{g_s} | \mathfrak{P}_{(\rho, \sigma)}^{(r, s)} |^2$$

eine Relativinvariante, also unabhängig ist von der Verschiedenheit der Lösungen des entarteten Problems, die sich als Grenzfälle qualitativ verschiedener gestörter Lösungen ergeben.

Den Zusammenhang dieses Satzes mit den „Summenregeln“ der Intensitäten von Multipletts, ZEEMAN-Aufspaltungen usw., die gleichfalls schon vor der Begründung der Quantenmechanik bekannt waren und als Leitfaden eine wichtige Rolle gespielt haben, werden wir später [§ 31] erörtern.

Der HEISENBERGSche Ansatz für die Ausstrahlung muß natürlich selbst wieder als Folge der Quantentheorie erwiesen werden, indem die Wechselwirkung eines Atoms mit dem elektromagne-

¹ HEISENBERG, W.: Z. Phys. Bd. 31, S. 617. 1925.

² BORN, M., W. HEISENBERG u. P. JORDAN: Z. Phys. Bd. 35, S. 557. 1926.

tischen Strahlungsfelde quantenmechanisch untersucht wird. Wir werden dies später tun und zeigen, daß tatsächlich die Größen (19) in erster Näherung für die Emission und Absorption einer g, g_s -fachen Spektrallinie maßgebend sind, sofern sie nicht identisch verschwinden. In letzterem Falle hat man die Absorption bzw. Emission des betreffenden Übergangs $r \rightarrow s$ aus der Matrix des Quadrupolmoments (gegebenenfalls noch höherer Momente) in analoger Weise [gemäß (12)] zu berechnen.

Wir wollen endlich hervorheben, daß unsere Ergebnisse bezüglich der teilweisen Unbestimmtheit der Lösung eines quantenmechanischen Problems bei entarteter Energie korrespondenzmäßig vollkommen analog sind den in Bd. I, § 14, § 15 besprochenen Lehren der BOHRSchen Theorie. In der BOHRSchen Theorie konnte die Bewegungsform eines Systems von f Freiheitsgraden nur im Falle nicht entarteter Energie durch f Quantenbedingungen bis auf Phasenkonstanten eindeutig festgelegt werden. Bei Entartung der Energie konnte nur eine entsprechend geringere Anzahl von Quantenbedingungen vorgeschrieben werden (Bd. I, § 15), so daß zwar die Energiewerte eindeutig festgelegt, aber die Bewegungsform teilweise willkürlich wurde. Entsprechend sind in der Quantenmechanik die Matrizen q_k, p_k , welche die Kinematik des Systems beschreiben, nur im Falle nichtentarteter Energie bis auf Phasenkonstanten festgelegt und besitzen im Entartungsfall eine gewisse, oben untersuchte Unbestimmtheit.

§ 21. Integrale der Bewegungsgleichungen und kontinuierliche Gruppen. Wenn bei einem abgeschlossenen System eine selbstadjungierte Funktion der Koordinaten und Impulse $F(q, p)$ die Eigenschaft

$$(1) \quad \dot{F} = 0$$

besitzt, so nennen wir F (entsprechend dem Sprachgebrauch der klassischen Theorie) ein *Integral der Bewegungsgleichungen*. Nach § 4, (31) und § 19, (8) kann die Definition (1) eines Integrals auch durch

$$(2) \quad [H, F] = 0$$

ersetzt werden, wiederum in formaler Übereinstimmung mit der klassischen Theorie [s. Anhang 2, (19)]. Das Analogon des *Satzes* von POISSON (Anhang 2) ist hier trivial; denn wenn F und G

mit H vertauschbar sind, so ist selbstverständlich auch die daraus *rational* gebildete Größe $[F, G]$ mit H vertauschbar. Übrigens folgt das auch aus der Formel § 18, (4).

Während nun bei der Energie H nicht nur die Erhaltungsgleichung $\dot{H} = 0$, sondern auch die schärfere Beziehung $H = \text{Diagonalmatrix}$ allgemein gültig ist, müssen wir uns bei einem beliebigen Integrale F mit der schwächeren Aussage begnügen [vgl. § 5, (13)]:

Satz 1: Schreibt man die Matrizen so, daß H als geordnete Diagonalmatrix erscheint, so werden alle Integrale der Bewegungsgleichungen Stufenmatrizen, deren Stufengrößen den statistischen Gewichten g_r gleich sind.

Wenn insbesondere die Energie *nicht entartet* ist, dann (und nur dann) werden die Integrale sämtlich Diagonalmatrizen. Da aber zwei Diagonalmatrizen stets vertauschbar sind, so ergibt sich umgekehrt:

Satz 2: Gibt es zwei (oder mehr) Integrale F_1, F_2 , die nicht vertauschbar sind,

$$(3) \quad [F_1, F_2] \neq 0,$$

so ist die Energie notwendig entartet.

Die Kenntnis von Integralen erleichtert wesentlich die Untersuchung der Lösungen eines quantenmechanischen Problems. Hat man nämlich ein oder mehrere Integrale F_1, F_2, \dots, F_l , die sämtlich *vertauschbar* sind,

$$(4) \quad [F_\lambda, F_\mu] = 0, \quad \lambda, \mu = 1, 2, \dots, l$$

so kann man, auch bei entarteter Energie, zunächst eine Lösung suchen, bei der alle F_1, \dots, F_l *Diagonalmatrizen* sind. Denn nach § 13, Satz 1 können alle F_λ und H gleichzeitig auf Hauptachsen transformiert werden. Die allgemeinste Lösung des Problems erhält man dann nach den Ergebnissen des § 20, indem man die so gefundene besondere Lösung mit einer unitären Stufenmatrix der Stufengrößen g_r transformiert.

Für ein System von nicht durchweg vertauschbaren Integralen F_1, F_2, \dots, F_l kann man allgemeiner zeigen:

Satz 3: Es gibt stets eine spezielle Lösung der Bewegungsgleichungen, bei welcher das Matrizenystem F_1, F_2, \dots, F_l voll-

ständig ausreduziert ist, also als direkte Summe seiner einfachen Bestandteile erscheint.

Denn da, wie wir sahen, die F_1, \dots, F_l notwendig Stufenmatrizen mit den Stufengrößen g_r sind, kann man ihre Reduktion zu Ende führen mit Hilfe einer unitären Stufenmatrix \hat{U} mit denselben Stufengrößen; diese ist mit der Energiematrix W vertauschbar, also sind die mit \hat{U} transformierten Matrizen q_k, p_k wiederum Lösungen der Bewegungsgleichungen.

Aus einer hermiteschen Matrix F kann man nach § 7, (7) und (10) die unitären Matrizen

$$(5) \quad U = e^{iF}$$

und

$$(6) \quad V = \frac{1 - iF}{1 + iF}$$

bilden. Wenn F ein Integral, also mit H vertauschbar ist, so gilt dasselbe für U und V . Man kann aber auch eine Art Umkehrung davon beweisen, indem man statt (5) und (6) die Formeln

$$(5a) \quad U(\lambda) = e^{i\lambda F},$$

$$(6a) \quad V(\lambda) = \frac{1 - i\lambda F}{1 + i\lambda F}$$

ansetzt, wobei λ eine reelle Zahl ist. U und V sind dann für jedes λ zugleich mit F Integrale; man hat zwei einparametrische Scharen unitärer Integrale. Wenn nun bekannt ist, daß U und V für alle reellen Werte von λ Integrale sind, so muß es auch F sein. Denn man kann dann λ insbesondere klein annehmen, so daß λ^2 vernachlässigt werden kann; dann ist $HU = UH$ gleichbedeutend mit

$$H(1 + i\lambda F) = (1 + i\lambda F)H$$

oder

$$[H, F] = 0,$$

und das analoge gilt für V .

Man kann diese Verhältnisse besonders klar überblicken vom Standpunkte der Theorie der kontinuierlichen Gruppen. Wir werden auf diese Theorie später ausführlich einzugehen haben; hier ge-

nügt es, die Begriffsbestimmung einer Gruppe und einige einfache Folgerungen anzuführen.

Als *Gruppe* bezeichnet man ein System von endlich oder unendlich vielen Elementen, das folgende Eigenschaften hat:

1. Zu je zwei, in bestimmter Reihenfolge gegebenen, gleichen oder verschiedenen Elementen A, B ist eindeutig ein Element C des Systems zugeordnet, das man das *Produkt* nennt; in Zeichen

$$A B = C .$$

2. Es gilt das Assoziativgesetz

$$(A B) C = A (B C) .$$

(Nicht zu gelten braucht das Kommutativgesetz $A B = B A$.)

3. Es gibt ein *Einheitselement* E , das für jedes Element A der Gruppe der Formel

$$A E = E A = A$$

genügt.

4. Zu jedem Element A gibt es ein inverses Element $X = A^{-1}$, das der Gleichung

$$X A = A X = E$$

genügt.

Bildet eine kleinere Menge unter den Elementen einer Gruppe wieder für sich eine Gruppe, so nennt man diese eine *Untergruppe* der umfassenderen Gruppe.

In unserm Falle werden wir als Elemente der Gruppe lineare, insbesondere unitäre Transformationen bzw. die zugehörigen Matrizen betrachten. Die Gesamtheit aller gleichartigen Matrizen, die eine Reziproke besitzen, bilden offenbar eine Gruppe. Sie enthält als Untergruppe die Gesamtheit aller gleichartigen *unitären* Matrizen [nach § 6, (15)]; hier wird es sich um Untergruppen dieser unitären Gruppe handeln.

Die durch (5a) definierte Schar von unitären Matrizen,

$$(7) \quad U(\lambda) = e^{i\lambda F} ,$$

ist offenbar eine Gruppe; da sie von *einem* kontinuierlich veränderlichen Parameter abhängt, nennt man sie eine *eingliedrige kontinuierliche Gruppe*. Das Produkt zweier Matrizen der Form (7), $U(\lambda_1)$ und $U(\lambda_2)$, ist nämlich

$$U(\lambda_1) U(\lambda_2) = e^{i\lambda_1 F} e^{i\lambda_2 F} = e^{i(\lambda_1 + \lambda_2) F} = U(\lambda_1 + \lambda_2) ,$$

gehört also wieder zu den Matrizen der Form (7) mit $\lambda = \lambda_1 + \lambda_2$; ferner ist die Identität in der Schar vorhanden (für $\lambda = 0$) und auch die zu $U(\lambda)$ inverse Matrix $U(-\lambda) = e^{-i\lambda F} = U(\lambda)^{-1}$.

Die Bedeutung des Gruppenbegriffs für das mechanische Problem erkennt man durch folgende Betrachtung.

Transformiert man die Koordinaten und Impulse mit Hilfe der zunächst beliebigen unitären Matrix V in neue,

$$(8) \quad p'_k = V^{-1} p_k V, \quad q'_k = V^{-1} q_k V,$$

so geht irgend eine Funktion $\varphi(p, q)$ über in

$$(9) \quad \varphi(p', q') = V^{-1} \varphi(p, q) V.$$

Es sei nun V eine Funktion der p_k, q_k ; dann liefern die Gleichungen (8) die p'_k, q'_k ebenfalls als Funktionen der p_k, q_k . Setzt man diese in $\varphi(p', q')$ ein, so entsteht eine Funktion $\Phi(p, q)$, die im allgemeinen von $\varphi(p, q)$ verschieden ist. Man kann nun fragen, wann $\Phi(p, q)$ mit $\varphi(p, q)$ identisch ist, d. h. wann φ in sich transformiert wird oder *invariant* ist. Die Antwort liest man aus (9) ab: Es ist notwendig und hinreichend, daß

$$V \varphi(p, q) = \varphi(p, q) V$$

oder

$$(10) \quad [\varphi, V] = 0.$$

Satz 4: *Eine Funktion $\varphi(p, q)$ ist dann und nur dann invariant gegen eine unitäre Transformation (8):*

$$p'_k = V^{-1} p_k V, \quad q'_k = V^{-1} q_k V,$$

wenn die unitäre Matrix V mit φ vertauschbar ist.

Die Formel (7) liefert nun zu jedem Integral F eine Gruppe unitärer Matrizen, die H in sich transformieren; man sagt dann, H ist eine *Invariante der Gruppe*.

Da man die Entwicklung von (7)

$$(11) \quad U(\lambda) = 1 + i\lambda F + \dots$$

hat, so bestimmt F die *infinitesimale Transformation* der Gruppe. Denkt man sich nämlich U als Transformation der Vektoren des HILBERTSchen Raumes,

$$(12) \quad x = U(\lambda) x_0,$$

so hat man für kleine λ nach (11)

$$(13) \quad x = (1 + i \lambda F) x_0$$

oder

$$(13a) \quad \frac{x - x_0}{\lambda} = i F x_0.$$

Im $\lim \lambda \rightarrow 0$ gibt das die Differentialgleichung

$$(14) \quad \frac{dx}{d\lambda} = i F x,$$

welche die infinitesimale Änderung des Vektors x infolge der Transformation mit der Matrix $i F d\lambda$ bestimmt.

Ist F als beliebige hermitesche Matrix gegeben, so führt die Lösung der Gleichung (14) notwendig wieder auf die Formel (7) zurück. Denn denkt man sich das Intervall der λ -Achse von 0 bis λ in n gleiche Teile der Länge $\frac{\lambda}{n}$ geteilt, so hat man nach (13) für die den Teilpunkten entsprechenden Vektoren x_0, x_1, \dots, x_n approximativ

$$\begin{aligned} x_1 &= \left(1 + i \frac{\lambda}{n} F \right) x_0, \\ x_2 &= \left(1 + i \frac{\lambda}{n} F \right) x_1, \\ &\dots\dots\dots \\ x_n &= \left(1 + i \frac{\lambda}{n} F \right) x_{n-1} \end{aligned}$$

und daraus

$$x_n = \left(1 + i \frac{\lambda}{n} F \right)^n x_0.$$

Das gibt aber im $\lim n \rightarrow \infty$

$$(15) \quad x = \lim_{n \rightarrow \infty} x_n = e^{i\lambda F} x_0,$$

woraus im Hinblick auf (12) die Behauptung hervorgeht.

Ist umgekehrt irgend eine kontinuierliche, eingliedrige Gruppe unitärer Transformationen $U(\lambda)$ gegeben, so kann man λ so normiert denken, daß $\lambda = 0$ der Einheitsmatrix entspricht und $U(\lambda_1) U(\lambda_2) = U(\lambda_1 + \lambda_2)$ ist. Dann ist die Matrix der infinitesimalen Transformation F zu definieren durch den Koeffizienten

der Entwicklung (11) von $U(\lambda)$. F ist dann notwendig hermitisch; denn aus $U U^\dagger = 1$ folgt

$$(1 + i \lambda F + \dots)(1 - i \lambda F^\dagger + \dots) = 1,$$

also

$$(16) \quad F - F^\dagger = 0.$$

Wir fassen das Ergebnis so zusammen:

Satz 5: Jedes Integral F der Bewegungsgleichungen bestimmt als Matrix der infinitesimalen Transformation eine eingliedrige Gruppe $U(\lambda) = e^{i\lambda F}$ unitärer Matrizen, die Integrale sind; die zu $U(\lambda)$ gehörige kanonische Transformation führt die Energiefunktion H in sich über. Umgekehrt gehört zu jeder Gruppe kanonischer Transformationen $U(\lambda)$, die H invariant läßt, ein von λ unabhängiges Integral F , das der infinitesimalen Transformation der Gruppe entspricht und aus dem sich die Gruppenelemente nach der Formel $U(\lambda) = e^{i\lambda F}$ ableiten.

In vielen Fällen ist eine Gruppe von Transformationen, die H zur Invariante hat, auf Grund der geometrischen und physikalischen Eigenschaften des Systems durch Gleichungen der Form

$$(17) \quad q'_k = q'_k(p, q; \lambda), \quad p'_k = p'_k(p, q; \lambda)$$

gegeben. Wir haben die Frage zu beantworten, wie daraus die Matrix F der infinitesimalen Transformation der Gruppe und damit auch die endliche Transformationsmatrix $U(\lambda)$ gewonnen werden kann.

Die Gleichungen (17) müssen sich mit Hilfe von $U(\lambda)$ so schreiben lassen:

$$(18) \quad q'_k = U(\lambda)^{-1} q_k U(\lambda), \quad p'_k = U(\lambda)^{-1} p_k U(\lambda).$$

Entwickelt man nun $U(\lambda)$ nach (11), so erhält man

$$(19) \quad \begin{cases} q'_k = q_k + i \lambda (q_k F - F q_k) + \dots, \\ p'_k = p_k + i \lambda (p_k F - F p_k) + \dots. \end{cases}$$

Entwickelt man andererseits die Ausdrücke (17),

$$(20) \quad \begin{cases} q'_k = q_k + \lambda f_k(p, q) + \dots, & f_k = \left(\frac{\partial q'_k}{\partial \lambda} \right)_{\lambda=0}, \\ p'_k = p_k + \lambda g_k(p, q) + \dots, & g_k = \left(\frac{\partial p'_k}{\partial \lambda} \right)_{\lambda=0}, \end{cases}$$

so erhält man durch Vergleich von (19) und (20)

$$(21) \quad \begin{cases} \varkappa [F, q_k] = i f_k, \\ \varkappa [F, p_k] = i g_k, \end{cases}$$

oder nach § 18, (3) wegen $\varkappa = \frac{h}{2\pi i}$:

$$(22) \quad \begin{cases} \frac{\partial F}{\partial p_k} = -\frac{2\pi}{h} f_k, \\ \frac{\partial F}{\partial q_k} = \frac{2\pi}{h} g_k. \end{cases}$$

Dies sind zwei partielle „Differentialgleichungen“, aus denen F zu bestimmen ist. Beispiele hierfür werden wir später kennen lernen.

§ 22. Systeme mit zeitabhängiger Energie. In manchen Fällen ist es zweckmäßig, statt einer HAMILTON-Funktion $H(q, p)$ der Form (1), § 19, die nur von den q, p abhängt, eine allgemeinere HAMILTON-Funktion

$$(1) \quad H(q, p; t) = H(q_1, q_2, \dots, q_f; p_1, p_2, \dots, p_f; t)$$

zu betrachten, die explizit von der „Zeit“ t abhängt. Dabei wird t als ein *Zahlparameter* aufgefaßt. Den physikalischen Sinn dieses Vorgehens werden wir nachher näher erläutern.

Für die Matrizen des Systems (1), die wieder ein kanonisches System bilden sollen, wird man entsprechend § 19, (2) die Gültigkeit der Bewegungsgleichungen

$$(2) \quad \dot{q}_k = \frac{\partial H}{\partial p_k}, \quad \dot{p}_k = -\frac{\partial H}{\partial q_k}$$

fordern. Es ist aber im allgemeinen nicht mehr möglich, die Forderungen § 19, (3) bzw. analoge Forderungen aufrecht zu erhalten; vielmehr muß man jetzt die Art der Abhängigkeit der Matrizen-elemente der q, p von dem Zeitparameter t zunächst offen lassen. Der Verzicht auf Gleichungen, welche den Forderungen § 19, (3) für abgeschlossene Systeme analog sind, bedingt eine weitgehende Willkür der Lösung unseres Problems. Offenbar erhält man aus einer bestimmten Lösung von (2) stets wieder eine Lösung, wenn man mit einer *beliebigen* unitären Matrix transformiert, die lediglich *zeitunabhängig* sein muß. Die genauere Festlegung der Lösung

muß in jedem einzelnen Falle besonders, durch sinngemäße Berücksichtigung der gestellten physikalischen Frage, geschehen. In diesem Paragraphen haben wir zunächst nur das durch die Gleichungen (2) gestellte mathematische Problem zu betrachten.

Wir zeigen zunächst: Ist

$$(3) \quad F(q, p, t) = F(q_1, \dots, q_f; p_1, \dots, p_f; t)$$

eine beliebige Funktion der Koordinaten und Impulse sowie des Zeitparameters, so gilt für die zeitliche Ableitung auf Grund von (2):

$$(4) \quad \dot{F} = [H, F] + \frac{\partial F}{\partial t}.$$

Diese Formel ist jedenfalls richtig für die Sonderfälle

$$(5) \quad F = q_k, \quad F = p_k, \quad F = t.$$

Wir zeigen weiter: wenn (4) für zwei Funktionen F_1, F_2 richtig ist, gilt sie auch für die Summe und das Produkt

$$(6) \quad F_a = F_1 + F_2, \quad F_b = F_1 F_2.$$

In der Tat ist nach § 4, (26)

$$(7) \quad \begin{cases} \dot{F}_a = \dot{F}_1 + \dot{F}_2 \\ \dot{F}_b = \dot{F}_1 F_2 + F_1 \dot{F}_2 \end{cases}$$

und nach § 18, (2)

$$(8) \quad \begin{cases} [H, F_a] + \frac{\partial F_a}{\partial t} = [H, F_1] + [H, F_2] + \frac{\partial F_1}{\partial t} + \frac{\partial F_2}{\partial t}, \\ [H, F_b] + \frac{\partial F_b}{\partial t} = [H, F_1] F_2 + F_1 [H, F_2] + \frac{\partial F_1}{\partial t} F_2 + F_1 \frac{\partial F_2}{\partial t} \\ \quad = \left([H, F_1] + \frac{\partial F_1}{\partial t} \right) F_2 + F_1 \left([H, F_2] + \frac{\partial F_2}{\partial t} \right). \end{cases}$$

Der Vergleich von (7) und (8) zeigt die Richtigkeit der Behauptung.

Damit ist aber zugleich die Formel (4) für beliebige Funktionen bewiesen.

Als spezielle Folgerung ergibt sich aus (4) für $F = H$ analog zur klassischen Theorie

$$(9) \quad \dot{H} = \frac{\partial H}{\partial t};$$

im Spezialfall eines abgeschlossenen Systems ergibt sich wieder der Energiesatz $\dot{H} = 0^1$.

Wir denken uns wieder ein beliebiges (einfaches) kanonisches Matrizenystem q_k^0, p_k^0 gewählt, dessen Elemente wir aber — abweichend von den in § 20 gemachten Annahmen — als zeitlich konstant ansehen wollen. Dann suchen wir eine Lösung q_k, p_k der Bewegungsgleichungen (2) durch den Ansatz

$$(10) \quad \begin{cases} q_k = U^{-1} q_k^0 U, \\ p_k = U^{-1} p_k^0 U \end{cases}$$

zu gewinnen, wo U eine von t abhängige unitäre Matrix ist:

$$(11) \quad U = U(t)$$

Durch Differentiation folgt aus (10) mit Rücksicht auf die vorausgesetzte zeitliche Konstanz der q_k^0, p_k^0 :

$$(12) \quad \begin{cases} \dot{q}_k = -U^{-1}(\dot{U} U^{-1} q_k^0 - q_k^0 \dot{U} U^{-1}) U \\ \quad = -\kappa U^{-1} [\dot{U} U^{-1}, q_k^0] U, \\ \dot{p}_k = -U^{-1}(\dot{U} U^{-1} p_k^0 - p_k^0 \dot{U} U^{-1}) U \\ \quad = -\kappa U^{-1} [\dot{U} U^{-1}, p_k^0] U; \end{cases}$$

hierbei ist Gebrauch gemacht von

$$\frac{d}{dt}(U^{-1}) = -U^{-1} \dot{U} U^{-1},$$

[Vgl. § 4, (27)]. Nun ist nach (10) andererseits

$$(13) \quad \begin{cases} \frac{\partial H(q, p; t)}{\partial p_k} = U^{-1} \frac{\partial H(q^0, p^0; t)}{\partial p_k^0} U, \\ \frac{\partial H(q, p; t)}{\partial q_k} = U^{-1} \frac{\partial H(q^0, p^0; t)}{\partial q_k^0} U. \end{cases}$$

Führen wir aber (12) und (13) in die Bewegungsgleichungen (2) ein, so wird

$$(14) \quad \begin{cases} -\kappa [\dot{U} U^{-1}, q_k^0] = [H(q^0, p^0; t), q_k^0], \\ -\kappa [\dot{U} U^{-1}, p_k^0] = [H(q^0, p^0; t), p_k^0]. \end{cases}$$

¹ Jedoch kann ohne Heranziehung der Forderungen (3), § 19 nicht gefolgert werden, daß $H = W = \text{Diagonalmatrix}$ werden muß.

Aus diesen Gleichungen, die auch in der Form

$$(14') \quad \begin{cases} [H(q^0, p^0; t) + \varkappa \dot{U} U^{-1}, q_k^0] = 0, \\ [H(q^0, p^0; t) + \varkappa \dot{U} U^{-1}, p_k^0] = 0 \end{cases}$$

geschrieben werden können, folgt aber nach § 15, Satz 1

$$(15) \quad H(q^0, p^0; t) + \varkappa \dot{U} U^{-1} = a(t) \cdot E,$$

wo $a(t)$ eine reine (reelle) Zahl und E die Einheitsmatrix ist.

Die q_k, p_k im Ansatz (10) bleiben ungeändert, wenn wir U mit einer reinen Zahl (vom Betrage 1) multiplizieren. Durch Anbringung eines Faktors

$$e^{\varkappa \int_0^t a(\tau) d\tau}$$

an U kann aber in (15) die rechte Seite $a(t)E$ aufgehoben werden; wir können und wollen daher statt (15) stets

$$H(q^0, p^0; t) + \varkappa \dot{U} U^{-1} = 0$$

oder

$$(16) \quad H(q^0, p^0; t) U + \varkappa \dot{U} = 0$$

fordern. Rückwärts ergibt sich wieder aus (16) — falls diese Gleichung wirklich mit $U^\dagger = U^{-1}$ lösbar ist — die Erfüllung der Bewegungsgleichungen durch den Ansatz (10).

Es kann nun aber behauptet werden: *Ist eine Lösung $U(t)$ der Gleichung (16) für $t = 0$ unitär, so ist sie für alle t unitär.* (Vgl. den analogen Satz am Schluß von § 9.) Schreiben wir nämlich mit der Abkürzung $H(q^0, p^0; t) = H_0$ die Gleichung (16):

$$(16') \quad \varkappa \frac{d}{dt} U + H_0 U = 0$$

und die ihr äquivalente Gleichung

$$(16'') \quad -\varkappa \frac{d}{dt} U^\dagger + U^\dagger H_0 = 0$$

an (beachte $\varkappa^* = -\varkappa$), so ergibt sich

$$(17) \quad \varkappa \frac{d}{dt} (U^\dagger U) = \varkappa \frac{dU^\dagger}{dt} U + \varkappa U^\dagger \frac{dU}{dt} = U^\dagger H_0 U - U^\dagger H_0 U = 0.$$

Wir sehen also, daß aus $U^\dagger = U^{-1}$ für $t = 0$ gewiß auch $U^\dagger U = E$

für alle Zeiten folgt. Die Relation $UU^\dagger = E$ für alle t ergibt sich andererseits aus der Erwägung, daß die Reziproke $U(t)^{-1}$, nachdem sie für $t = 0$ existiert, aus Stetigkeitsgründen auch wenigstens in einer Umgebung von $t = 0$ existieren muß; wegen $U^\dagger U = E$ (für alle t) muß sie dann aber stets mit U^\dagger identisch sein.

Die Lösung der Gleichung (16) kann nun auf die Integration einer Vektor-Differentialgleichung der in § 9 besprochenen Art zurückgeführt werden. Ist y^0 irgend ein konstanter Vektor, so setze man

$$(18) \quad y(t) = U(t)y^0;$$

dann genügt $y(t)$ der Gleichung

$$(19) \quad x\dot{y} + H(p^0, q^0; t)y = 0;$$

fordert man überdies noch die Anfangsbedingung, welche U zur Zeit $t = 0$ einer beliebig vorgegebenen unitären Matrix U_0 gleichsetzt, so wird $y(0) = U_0 y^0$.

Wählt man nun für y^0 der Reihe nach das vollständige System orthogonaler Einheitsvektoren

$$(20) \quad (1, 0, 0, \dots), \quad (0, 1, 0, \dots), \dots$$

so sieht man, daß die entsprechenden Vektoren $y(t)$ jeweils die *Elemente einer Spalte* von $U(t)$ zu Komponenten haben. Umgekehrt kann man die Matrix U spaltenweise bilden, indem man ein vollständiges Orthogonalsystem y_1^0, y_2^0, \dots wählt und die Lösungen $y_k(t)$ bestimmt, die für $t = 0$ in die entsprechenden Vektoren y_k^0 übergehen (s. § 9).

Wir können also behaupten: Auf dem geschilderten Wege [über (16)] ist tatsächlich eine Lösung der Bewegungsgleichungen zu erhalten. Sie ist — wie von vornherein zu erkennen war — willkürlich derart, daß sie eine beliebige zeitlich konstante unitäre Transformation gestattet. *Im übrigen* jedoch ist sie vollständig bestimmt.

Den Zusammenhang der hier durchgeführten Überlegungen mit unserer früheren Theorie der abgeschlossenen Systeme bekommen wir durch folgende Bemerkung: Die bei den abgeschlossenen Systemen gestellte Forderung, daß die zeitliche Ableitung einer Matrix gemäß (13), § 3 bzw. (3), § 19 zu bestimmen sei, kann, wie wir wissen, formal durch Einführung eines Zeitparameters t befriedigt werden, indem die einzelnen Matrixelemente

Zeitfaktoren $e^{2\pi i \nu_{nm} t}$ erhalten. Wegen der BOHRschen Frequenzbedingung können wir diese Forderung auch so formulieren: Jede Matrix q_k, p_k [oder allgemein $F(q, p)$] soll aus einer Matrix mit zeitlich konstanten Elementen hervorgehen durch Transformation mit

$$(21) \quad U_1 = e^{-\frac{Wt}{\hbar}}; \quad W = (\delta_{nm} W_n).$$

Wir haben also, um im Spezialfall $\frac{\partial H}{\partial t} = 0$ von (16) zu den Formulierungen von § 19 zurückzukommen, für die Matrix U in (10), (16) die spezielle Gestalt

$$(22) \quad U = U' \cdot U_1 = U' e^{-\frac{Wt}{\hbar}}$$

zu fordern, wo U' zeitlich konstant ist. Dann geht (16) über in

$$(23) \quad H(q^0, p^0) U' - U' W = 0,$$

d. h. wir kommen wieder, wie in § 20, auf das Hauptachsenproblem der Energiematrix.

Mit Hilfe des Zeitparameters t können wir jetzt eine Anwendung der in § 21 besprochenen gruppentheoretischen Begriffe machen. Setzt man in (22), § 21 für F die Energiefunktion H selber, multipliziert mit $-\frac{2\pi}{\hbar}$, ein und nimmt als Parameter λ der Gruppe die Zeit t , so bekommt man

$$(24) \quad \frac{\partial H}{\partial p_k} = \left(\frac{dq_k'}{dt} \right)_{t=0}, \quad \frac{\partial H}{\partial q_k} = - \left(\frac{dp_k'}{dt} \right)_{t=0}$$

Das sind aber die kanonischen Bewegungsgleichungen für den Zeitpunkt $t = 0$, der im übrigen willkürlich ist. Man gelangt daher, in voller Übereinstimmung mit der klassischen Mechanik, zu folgender Auffassung des Bewegungsvorgangs:

Satz: Die Matrizen p_k, q_k als Funktionen der Zeit t gehen aus den Matrizen p_k^0, q_k^0 zur Zeit $t = 0$ hervor durch unitäre Transformationen, die eine eingliedrige Gruppe mit dem Parameter t bilden; die infinitesimale Matrix der Gruppe ist $-\frac{2\pi}{\hbar} H$.

Wenn H die Zeit nicht explizite enthält, so ist die endliche, unitäre Matrix der Gruppe also

$$(25) \quad U(t) = e^{-\frac{2\pi i}{\hbar} t H} = e^{-\frac{t H}{\hbar}}.$$

Für die kanonischen Bewegungsgleichungen unabgeschlossener Systeme haben wir in (16) einen gleichwertigen Ersatz gefunden, der wiederum keinen wesentlichen Bezug mehr nimmt auf die speziellen kanonischen Variablen, die als Argumente von H gebraucht werden. Es folgt daraus wieder, wie früher in § 19, die Invarianz der kanonischen Gleichungen gegenüber kanonischen Transformationen, welche nicht explizit den Zeitparameter t enthalten.

Wenn aber zwei Systeme von kanonischen Variablen q_k, p_k und Q_k, P_k durch eine *zeitabhängige* unitäre Transformation $V(t)$ auseinander hervorgehen:

$$(26) \quad \begin{cases} Q_k(t) = V^{-1}(t) q_k(t) V(t) \\ P_k(t) = V^{-1}(t) p_k(t) V(t), \end{cases}$$

und wenn das System q, p die kanonischen Gleichungen mit der HAMILTON-Funktion $H(q, p, t)$ erfüllt:

$$\dot{q}_k = [H, q_k], \quad \dot{p}_k = [H, p_k],$$

so folgt daraus, daß das System Q, P die kanonischen Gleichungen

$$\dot{Q}_k = [K, Q_k], \quad \dot{P}_k = [K, P_k]$$

erfüllt mit der HAMILTON-Funktion:

$$(27) \quad K(Q, P, t) = H(Q, P, t) - \varkappa V^{-1} \dot{V}.$$

Es ist nämlich:

$$\begin{aligned} \dot{Q}_k &= -V^{-1} \dot{V} V^{-1} q_k V + V^{-1} \dot{q}_k V + V^{-1} q_k V V^{-1} \dot{V} = \\ &= V^{-1} [H(q, p, t), q_k] V - \varkappa [V^{-1} \dot{V}, Q_k] = \\ &= [H(Q, P, t), Q_k] - \varkappa [V^{-1} \dot{V}, Q_k]. \end{aligned}$$

Wir müssen endlich einige Worte sagen über die *physikalische Bedeutung* der hier benutzten Auffassung der Zeit t als eines Zahlparameters. Zu betonen ist dabei: Die Beschreibung eines physikalischen Systems durch eine zeitabhängige HAMILTON-Funktion, in welcher die Zeit als Zahlparameter aufgefaßt wird, kann *nicht* eine *exakte* Darstellung der physikalischen Verhältnisse geben, sondern ist lediglich — unter besonderen, sogleich zu erläuternden Umständen — als ein approximatives Rechenverfahren zu betrachten, das grundsätzliche Vernachlässigungen in sich schließt. Exakte Bedeutung kann nur unsere *Theorie der*

abgeschlossenen Systeme beanspruchen; und wir haben deshalb schon früher Wert darauf gelegt, in der Formulierung dieser Theorie keinen Gebrauch von einem Zahlparameter t zu machen: Die zeitliche Ableitung \dot{F} einer Matrix $F = (F(nm))$ wurde *definiert* durch $\dot{F} = (2\pi i \nu(nm) F(nm))$, so daß eine explizite Verwendung der Zeitfaktoren $e^{2\pi i \nu(nm)t}$ überflüssig war. Die exakte, allgemein gültige Theorie der abgeschlossenen Systeme muß deshalb auch benutzt werden zur Begründung der bedingungsweise und näherungsweise gültigen, hier entwickelten Theorie der ungeschlossenen Systeme.

Die *klassische* Theorie lehrt nun, daß bei einem abgeschlossenen Systeme die Energie $H = W$ und die Zeit t *kanonisch konjugiert* zueinander sind. Bei sinngemäßer korrespondenzmäßiger Übertragung in die Quantenmechanik müssen also die W, t analog den p, q durch gewisse *nichtkommutative* Rechen-symbole dargestellt werden und Rechenregeln unterworfen sein, die den kanonischen Vertauschungsregeln der p, q analog sind. Die Theorie der „diskreten“ Matrizen, die wir hier benutzen, erlaubt uns freilich noch nicht die vollständige mathematische Durchführung dieser Auffassung, die erst im zweiten Teil dieses Bandes im Zusammenhang mit der Operatorenrechnung gebracht werden kann. Hier genüge die Feststellung, daß Energie und Zeit miteinander nicht vertauschbar sind.

Betrachten wir nun ein ungeschlossenes System, so können wir es zu einem abgeschlossenen ergänzen durch Hinzunahme der darauf einwirkenden Systeme. Dann ist die Zeit mit der *Gesamtenergie* dieser Systeme nicht vertauschbar (kanonisch konjugiert zu ihr). Wenn aber das betrachtete *Teilsystem* so *schwach* ist gegenüber den einwirkenden Systemen, daß die *Rückwirkung* auf diese *vernachlässigt* werden kann, so ist es plausibel, daß die Koordinaten und Impulse und damit auch die Energie des Teilsystems näherungsweise als mit der zur Gesamtenergie kanonisch konjugierten Zeit *vertauschbar* angenommen werden dürfen. Dadurch kann der Gebrauch von t als Zahlparameter gerechtfertigt werden.

Auch ein für sich schon abgeschlossenes System der Energie $H^{(1)}$ kann natürlich durch Hinzunahme eines anderen abgeschlossenen Systems $H^{(2)}$ formal zu einem umfangreicheren System $H^{(1)} + H^{(2)}$ erweitert werden. Verwenden wir das zweite System $H^{(2)}$ als

„Uhr“, benutzen also statt der zu $H^{(1)}$ kanonisch konjugierten Zeit $t^{(1)}$ die durch $H^{(2)}$ gemessene (zu $H^{(2)}$ kanonisch konjugierte) Zeit $t^{(2)}$, so bekommen wir auch für das abgeschlossene System $H^{(1)}$ eine Zeit $t = t^{(2)}$, die mit allen auf das System $H^{(1)}$ bezüglichen Matrizen vertauschbar ist und deshalb, solange wir nur $H^{(1)}$ untersuchen, als *Zahlparameter* t behandelt werden kann. Damit bekommen wir auch eine exakte Rechtfertigung für den Gebrauch eines *Parameters* t bei *abgeschlossenen* Systemen, wie er durch die Zeitfaktoren $e^{2\pi i \nu^{(nm)} t}$ eingeführt wird.

Diese hier nur flüchtig berührten Verhältnisse werden wir später (Teil II) erneut und ausführlich behandeln.

§ 23. Der lineare harmonische Oszillator. Existenz- und Eindeutigkeitsuntersuchung der kanonischen Matrizenysteme. In der klassischen Mechanik bezeichnet man als *linearen harmonischen Oszillator* ein System, dessen Energiefunktion

$$(1) \quad H = \frac{1}{2\mu} p^2 + \frac{a}{2} q^2$$

ist¹; μ ist die Masse, a die quasielastische Konstante. Die Lösung der Bewegungsgleichungen liefert harmonische Schwingungen mit der konstanten Frequenz

$$(2) \quad \nu_0 = \frac{1}{2\pi} \sqrt{\frac{a}{\mu}}.$$

Wir können uns auch in der Quantenmechanik die Frage vorlegen, wie die kanonischen Matrizen p, q beschaffen sein müssen, wenn die HAMILTONSche Funktion des Systems, ausgedrückt in den nicht vertauschbaren p, q , gerade die Gestalt (1) besitzt.

Die korrespondenzmäßige Analogie der klassischen und der neuen Mechanik würde es leicht erlauben, eine Lösung des Problems zu erraten. Wir ziehen es jedoch vor, die Lösung des Problems durch eine systematische Untersuchung zu bestimmen, um so zugleich die eindeutige Bestimmtheit der Lösung (bis auf die Phasenkonstanten und eine Permutation der Diagonalelemente) zu erweisen.

Die kanonischen Bewegungsgleichungen

$$(3) \quad \dot{q} = \frac{\partial H}{\partial p} = \frac{p}{\mu}, \quad \dot{p} = -\frac{\partial H}{\partial q} = -aq$$

¹ S. Bd. I, § 7, S. 39.

oder, mit der Abkürzung (2):

$$(3a) \quad \ddot{q} = - (2\pi\nu_0)^2 q,$$

lauten wie in der klassischen Theorie. Definiert man „komplexe Amplituden“

$$(4) \quad \begin{cases} b = C(p - 2\pi i\nu_0\mu q), \\ b^\dagger = C(p + 2\pi i\nu_0\mu q) \end{cases}$$

mit einer vorläufig beliebigen Konstanten C , so gehen die Gleichungen (3) über in

$$(5) \quad \dot{b} = -2\pi i\nu_0 b, \quad \dot{b}^\dagger = 2\pi i\nu_0 b^\dagger.$$

Gleichzeitig wird

$$\begin{aligned} bb^\dagger &= C^2(p^2 + 4\pi^2\nu_0^2\mu^2q^2 + 2\pi i\nu_0\mu(pq - qp)), \\ b^\dagger b &= C^2(p^2 + 4\pi^2\nu_0^2\mu^2q^2 - 2\pi i\nu_0\mu(pq - qp)). \end{aligned}$$

Hieraus sieht man, daß die Vertauschungsrelation

$$[p, q] = \frac{1}{\kappa}(pq - qp) = 1$$

sich in

$$(6) \quad bb^\dagger - b^\dagger b = 1$$

verwandelt, wenn man

$$(7) \quad C = \frac{1}{\sqrt{2\hbar\nu_0\mu}}$$

setzt. Zugleich erhält man für die Energie (1) die beiden Darstellungen

$$(8) \quad \begin{cases} H = \hbar\nu_0 b b^\dagger - \frac{\hbar\nu_0}{2}, \\ \quad = \hbar\nu_0 b^\dagger b + \frac{\hbar\nu_0}{2}. \end{cases}$$

Die Gleichungen (5) lassen sich zu

$$(5a) \quad Wb - bW = -\hbar\nu_0 b$$

zusammenfassen und besagen für die Elemente einzeln:

$$(9) \quad (W_n - W_m)b_{nm} = -\hbar\nu_0 b_{nm}.$$

Hieraus kann man schließen: Bei einer *einfachen* Lösung müssen

die *verschiedenen* Eigenwerte der Energie eine *lückenlose arithmetische Reihe* bilden mit der Differenz $h\nu_0$.

Ist nämlich W' irgend ein Eigenwert, so können wir b als Übermatrix $(b_{\kappa\lambda}^{(k,l)})$ schreiben mit $k = 1, 2; l = 1, 2$ derart, daß zu den Diagonal-Untermatrizen mit $k = l = 1$ alle diejenigen Eigenwerte $W_k^{(1)}$ gehören, die eine *nach beiden Seiten möglichst lange*, den Wert W' enthaltende, lückenlose arithmetische Reihe der Differenz $h\nu_0$ bilden; zu $k = l = 2$ dagegen alle andern Eigenwerte (also alle, die sich irgendwo etwa zwischen die der Reihe $W_k^{(1)}$ einschieben oder — nach einer Lücke — außerhalb der Reihe liegen). Da dann stets $W_{\kappa}^{(1)} - W_{\lambda}^{(2)} \neq \pm h\nu_0$ ist, folgt aus (9), daß $b_{\kappa\lambda}^{(1,2)} = 0$, $b_{\kappa\lambda}^{(2,1)} = 0$ ist, so daß b und damit zugleich b^\dagger Stufenmatrizen sind. Das widerspricht der vorausgesetzten Einfachheit. Also können außer einer lückenlosen arithmetischen Reihe keine andern Eigenwerte vorhanden sein.

Nun ist H als Summe zweier Quadrate (nach § 10, S. 53) eine positiv definite Matrix, hat also nur positive Eigenwerte. Folglich muß jene arithmetische Reihe ein *Anfangsglied* besitzen. Wir ordnen die Zeilen und Spalten der Matrizen so, daß $H = W$ eine *geordnete* Diagonalmatrix wird, und schreiben W , b und b^\dagger nunmehr als Übermatrizen entsprechend der Vielfachheit der Eigenwerte:

$$(10) \quad \begin{cases} W = (W^{(r)} \delta_{rs}), \\ b = (b^{(r,s)}), \quad b^\dagger = (b^{\dagger(r,s)}). \end{cases} \quad r, s = 0, 1, 2, \dots$$

Indem wir die Zählung der Stufen mit 0 beginnen, nennen wir g_r die Länge der $(r + 1)$ -ten Stufe (die wir ausdrücklich als endlich voraussetzen). Dann ist $b^{(r,s)}$ eine Matrix von g_r Zeilen und g_s Spalten. $W^{(r)}$ ist nach (9) das Produkt der Zahl

$$(11) \quad W_r = r \cdot h\nu_0 + W_0$$

mit der Einheitsmatrix von g_r Zeilen und Spalten $E^{(r)}$:

$$(11a) \quad W^{(r)} = W_r \cdot E^{(r)}.$$

Für die Teilmatrizen in b und b^\dagger hat man bei der neuen Indizierung nach (5)

$$\begin{aligned} (W_r - W_s) b^{(r,s)} &= -h\nu_0 b^{(r,s)}, \\ (W_r - W_s) b^{\dagger(r,s)} &= h\nu_0 b^{\dagger(r,s)}; \end{aligned}$$

also wegen (11):

$$(12) \quad \begin{cases} b^{(r,s)} = 0, & \text{wenn } r - s \neq -1, \\ b^{\dagger(r,s)} = 0, & \text{wenn } r - s \neq +1. \end{cases}$$

Hieraus folgt, daß

$$(13) \quad \begin{cases} b b^{\dagger} = (b^{(r,r+1)} b^{\dagger(r+1,r)} \delta_{r_s}), \\ b^{\dagger} b = (b^{\dagger(r,r-1)} b^{(r-1,r)} \delta_{r_s}) \end{cases}$$

ist, und da der Index r keine negativen Werte annehmen kann, so verschwindet die Teilmatrix von $b^{\dagger} b$ für $r = 0$. Somit ergibt sich aus der zweiten Gleichung (8) zunächst für $r = 0$:

$$(14) \quad W^{(0)} = \frac{\hbar \nu_0}{2} E^{(0)}, \quad W_0 = \frac{\hbar \nu_0}{2};$$

sodann allgemein

$$\begin{aligned} W^{(r)} &= \hbar \nu_0 b^{(r,r+1)} b^{\dagger(r+1,r)} - \frac{\hbar \nu_0}{2} E^{(r)}, \\ W^{(r+1)} &= \hbar \nu_0 b^{\dagger(r+1,r)} b^{(r,r+1)} + \frac{\hbar \nu_0}{2} E^{(r+1)}. \end{aligned}$$

Nach (11), (11a) und (14) wird daher

$$(15) \quad \begin{cases} \text{a) } b^{(r,r+1)} b^{\dagger(r+1,r)} = (r+1) E^{(r)}, \\ \text{b) } b^{\dagger(r+1,r)} b^{(r,r+1)} = (r+1) E^{(r+1)}. \end{cases}$$

Bilden wir nun die Spuren dieser (endlichen) Matrizen, so ergibt sich in (15a) und (15b) linker Hand [nach § 4, (22)] dieselbe Größe, während rechts $(r+1) \cdot g_r$ und $(r+1) \cdot g_{r+1}$ herauskommt. Also muß

$$(16) \quad g_{r+1} = g_r$$

sein; d. h. *alle Eigenwerte besitzen dieselbe Vielfachheit g_0 .*

Wir können jetzt weiter zeigen: Für eine einfache (irreduzible) Lösung ist bei allen Termen die Vielfachheit gleich 1; eine Lösung der Vielfachheit $g_0 > 1$ kann in g_0 einfache Lösungen zerfällt werden. Wegen (16) ist nämlich $b^{(r,r+1)}$ eine *quadratische* Matrix und $b^{\dagger(r+1,r)}$ sodann ihre Adjungierte im gewöhnlichen Sinne¹:

$$(17) \quad b^{\dagger(r+1,r)} = [b^{(r,r+1)}]^{\dagger}.$$

¹ D. h. nicht in dem verallgemeinerten Sinne für rechteckige Matrizen, gemäß § 6, (22).

Folglich ist nach (15a)

$$(18) \quad b^{(r, r+1)} = U_r \cdot \sqrt{r+1},$$

wo U_r eine g_0 -reihige unitäre Matrix ist; denn für $U_r = \frac{b^{(r, r+1)}}{\sqrt{r+1}}$ gilt $U_r U_r^\dagger = E^{(r)}$. Wenn wir b mit einer unitären Stufenmatrix

$$(19) \quad V = (\delta_{rs} V^{(r)})$$

transformieren, entsteht eine Übermatrix

$$(20) \quad B = V^{-1} b V,$$

in der wieder nur die Teilmatrizen $B^{(r, r+1)}$ von Null verschieden sind, und zwar ist nach (18)

$$(20a) \quad \begin{aligned} B^{(r, r+1)} &= V^{(r)-1} b^{(r, r+1)} V^{(r+1)} \\ &= \sqrt{r+1} V^{(r)-1} U_r V^{(r+1)}. \end{aligned}$$

Man kann nun leicht V so wählen, daß die $B^{(r, r+1)}$ Vielfache der Einheitsmatrix $E^{(0)}$ werden; man setze

$$(21) \quad U_r V^{(r+1)} = V^{(r)}, \quad V^{(0)} = E^{(0)},$$

d. h.

$$V^{(r+1)} = U_r^{-1} V^{(r)},$$

also

$$(21a) \quad V^{(r+1)} = (U_0 U_1 \dots U_r)^{-1}.$$

Dann wird

$$(22) \quad B^{(r, r+1)} = \sqrt{r+1} E^{(0)},$$

und entsprechendes gilt für die Adjungierte $B^\dagger = V b^\dagger V^{-1}$.

Damit ist aber die Zerfällung durchgeführt: B und B^\dagger sind nach (22) in g_0 übereinstimmende Lösungen mit nur einfachen Eigenwerten gespalten.

Wir können uns also jetzt auf eine einfache Lösung ($g_0 = 1$) beschränken, deren Matrixelemente wir mit b_{nm} bezeichnen. Dann ist in (18) U_r eine einreihige unitäre Matrix, d. h. eine komplexe Zahl vom Betrage 1, etwa $e^{i\varphi_r}$. Unser Resultat lautet also:

$$(23) \quad \left\{ \begin{array}{l} \text{a) } \begin{cases} b_{nm} = 0, & \text{wenn } n - m \neq -1, \\ b_{nm}^\dagger = 0, & \text{wenn } n - m \neq +1; \end{cases} \\ \text{b) } \begin{cases} b_{n,n+1} = \sqrt{n+1} e^{i\varphi_n}, \\ b_{n+1,n}^\dagger = \sqrt{n+1} e^{-i\varphi_n}; \end{cases} \\ \text{c) } W_n = h\nu_0 \left(n + \frac{1}{2} \right); \quad n = 0, 1, 2, \dots \end{array} \right.$$

Durch (23) wird offenbar die Gleichung $H = W$ und die Vertauschungsregel (6) tatsächlich befriedigt. Daß diese Lösung mit einfachen Eigenwerten einfach (irreduzibel) ist, ergibt sich daraus, daß in der arithmetischen Reihe der Eigenwerte keine Lücke auftreten darf, daß der kleinste Eigenwert nach (14) eindeutig bestimmt ist und daß unendlich viele Eigenwerte vorhanden sein müssen.

Vergleichen wir nun die gewonnene Lösung mit der des entsprechenden klassischen Problems (Bd. 1, §§ 9, 10), so ist zunächst auffällig das Auftreten einer „Nullpunktsenergie“ $W_0 = \frac{h\nu_0}{2}$; wir werden später auf ihre physikalische Bedeutung zurückkommen. Bilden wir ferner durch Auflösen der Gleichungen (4) die Matrizen p und q , so erhalten wir mit Rücksicht auf (7)

$$(24) \quad \begin{cases} p = \sqrt{\frac{h\nu_0\mu}{2}} (b^\dagger + b), \\ q = \frac{1}{2\pi i} \sqrt{\frac{h}{2\nu_0\mu}} (b^\dagger - b); \end{cases}$$

also nach (23)

$$(25) \quad \left\{ \begin{array}{l} \text{a) } p_{nm} = 0, \quad q_{nm} = 0, \quad \text{wenn } n - m \neq \pm 1, \\ \text{b) } \begin{cases} p_{n,n+1} = \sqrt{\frac{h\nu_0\mu}{2}} b_{n,n+1} = \frac{1}{2} \sqrt{2h\nu_0\mu(n+1)} e^{i\varphi_n}, \\ p_{n+1,n} = \sqrt{\frac{h\nu_0\mu}{2}} b_{n+1,n}^\dagger = \frac{1}{2} \sqrt{2h\nu_0\mu(n+1)} e^{-i\varphi_n}, \end{cases} \\ \text{c) } \begin{cases} q_{n,n+1} = \frac{1}{2\pi i} \sqrt{\frac{h}{2\nu_0\mu}} b_{n,n+1} = \frac{1}{2} \sqrt{\frac{h(n+1)}{2\pi^2\mu\nu_0}} e^{i\left(\varphi_n + \frac{\pi}{2}\right)}, \\ q_{n+1,n} = \frac{1}{2\pi i} \sqrt{\frac{h}{2\nu_0\mu}} b_{n+1,n}^\dagger = \frac{1}{2} \sqrt{\frac{h(n+1)}{2\pi^2\mu\nu_0}} e^{-i\left(\varphi_n + \frac{\pi}{2}\right)}. \end{cases} \end{array} \right.$$

Vergleicht man das mit den Formeln Bd. 1, § 9, (9) und (10), so sieht man, daß das Verschwinden aller Matrixelemente außer für $m = n \pm 1$ dem Auftreten eines einzigen FOURIER-Gliedes in den klassischen Formeln entspricht; dabei ist zu beachten, daß die beiden Matrixelemente $p_{n, n+1}$ und $p_{n+1, n}$ zusammen *einem* cos-Gliede der klassischen FOURIER-Reihe für p entsprechen, wodurch sich der Faktor $\frac{1}{2}$ in (25) erklärt. Sieht man von diesem ab, so stimmen die Amplituden in (25) mit den klassischen überein, wenn man dort $J = h(n+1)$ setzt, anstatt, wie in der älteren Quantentheorie, $J = hn$. Endlich besteht zwischen q und p eine Phasendifferenz von $\frac{\pi}{2}$, genau wie in den klassischen Formeln.

Wir haben mit der ausführlichen Untersuchung des harmonischen Oszillators als des einfachsten matrizenmechanischen Systems überhaupt auch die bisher noch verbliebenen Lücken unserer allgemeinen Theorie ausgefüllt. Erstens haben wir jetzt die mathematische *Existenz* von Matrizen p, q mit der kanonischen Eigenschaft $[p, q] = 1$ erwiesen. Zweitens aber können wir behaupten: *Es seien p, q und P, Q zwei einfache kanonische Systeme: $[p, q] = 1$; $[P, Q] = 1$, und es sei möglich, die Matrizen $p^2 + q^2$ und $P^2 + Q^2$ zu bilden und auf Hauptachsen zu transformieren* (ohne daß Konvergenzschwierigkeiten oder sonstige Singularitäten eintreten). *Dann kann man p, q unitär in P, Q transformieren:*

$$(26) \quad P = U^{-1} p U, \quad Q = U^{-1} q U.$$

Denn man kann in diesem Falle zunächst V_1 und V_2 so bestimmen, daß die durch

$$p_0 = V_1^{-1} p V_1, \quad q_0 = V_1^{-1} q V_1, \\ P_0 = V_2^{-1} P V_2, \quad Q_0 = V_2^{-1} Q V_2$$

definierten p_0, q_0, P_0, Q_0 so beschaffen sind, daß $p_0^2 + q_0^2$ und $P_0^2 + Q_0^2$ Diagonalmatrizen sind. Dann bilden aber p_0, q_0 sowohl als auch P_0, Q_0 eine einfache Lösung der Bewegungsgleichungen für einen harmonischen Oszillator der Energie $\frac{1}{2}(p_0^2 + q_0^2)$ bzw. $\frac{1}{2}(P_0^2 + Q_0^2)$, und müssen nach unseren soeben erhaltenen Ergebnissen bei geeigneter Einrichtung der Phasen und Anordnung der Diagonalglieder übereinstimmen :

$$p_0 = P_0, \quad q_0 = Q_0.$$

Also wird (26) erfüllt mit

$$(27) \quad U = V_1 V_2^{-1}.$$

Die Forderung, $p^2 + q^2$ solle singularitätenfrei auf Hauptachsen transformierbar sein, gibt allerdings zunächst nur eine etwas äußerliche Kennzeichnung der in Betracht zu ziehenden Klasse kanonischer Matrixensysteme. Das gerade diese (und nur diese) wirklich für die physikalische Anwendung etwa auf rechtwinklige Koordinaten $x, y, z = q_1, q_2, q_3$ in Betracht kommen, werden wir später¹ bei der tieferen mathematischen Untersuchung dieser Matrizen begründen können. (Vgl. auch die Schlußbemerkung in Kap. VI, § 52.)

Aus Satz 5, § 15 kann nun ohne weiteres gefolgert werden: *Zwei kanonische Systeme p_k, q_k und P_k, Q_k von je f Freiheitsgraden mit der Eigenschaft, daß alle $p_k^2 + q_k^2$ und $P_k^2 + Q_k^2$ singularitätenfrei unitär auf Hauptachsen transformierbar sind, können unitär ineinander transformiert werden.* Denn für jedes einzelne Paar p_k, q_k ist nach dem soeben Bewiesenen die Voraussetzung (b) des Satzes 5, § 15 erfüllt, daß die einfachen Bestandteile sämtlich unitär ineinander transformierbar sind. Man kann also die p_k, q_k sowohl als auch die P_k, Q_k unitär transformieren in die Gestalt einer Verschmelzung von f Lösungen der Bewegungsgleichungen des harmonischen Oszillators.

§ 24. Systeme harmonischer Oszillatoren. Mit Hilfe der Theorie des linearen harmonischen Oszillators und der in § 16 entwickelten Sätze läßt sich auch das allgemeinere Problem gekoppelter Oszillatoren mit beliebig vielen Freiheitsgraden beherrschen.

Es seien

$$(1) \quad \left\{ \begin{array}{l} T(\xi, \xi) = \sum_{k, l=1}^f a_{kl} \xi_k \xi_l, \quad a_{kl} = a_{lk}; \\ \Phi(\eta, \eta) = \sum_{k, l=1}^f b_{kl} \eta_k \eta_l, \quad b_{kl} = b_{lk} \end{array} \right.$$

zwei reelle quadratische Formen, von denen die erste T positiv definit sei. Dann kann man in der klassischen Mechanik $T(p, p)$ als kinetische und $\Phi(q, q)$ als potentielle Energie eines Massensystems mit den Koordinaten und Impulsen q_k, p_k ansehen, und

¹ In der geplanten Fortsetzung des Buches.

zwar beschreibt dieser Ansatz kleine Schwingungen um eine Gleichgewichtslage.

Bilden wir nun in der Quantenmechanik durch Einsetzen von konjugierten Matrizen q_k, p_k für ξ_k, η_k die Energiefunktion

$$(2) \quad H(p, q) = T(p, p) + \Phi(q, q),$$

so werden wir diese einem System gekoppelter „Quantenoszillatoren“ von f Freiheitsgraden zuschreiben können. Solche Systeme treten in der Physik häufig auf, wenigstens als Näherungen.

Nach § 16, Satz 2 gibt es eine Matrix S von f Zeilen und Spalten, derart, daß durch

$$(3) \quad \xi = \check{S} \mathcal{E}, \quad \eta = S H,$$

d. h. ausführlich

$$(3a) \quad \begin{cases} \xi_k = \sum_l \check{S}_{kl} \mathcal{E}_l, \\ \eta_k = \sum_l S_{kl} H_l \end{cases}$$

die Formen $T(\xi, \xi)$ und $\Phi(\eta, \eta)$ gleichzeitig auf die Diagonalform gebracht werden; und zwar kann man auf Grund der Schlußbemerkung von § 16 die Matrix S *reell* wählen. In (3) kann man nun die Zahlen ξ_k, η_k durch die Matrizen p_k, q_k ersetzen und entsprechend die Zahlen \mathcal{E}_k, H_k durch transformierte Matrizen P_k, Q_k , so daß

$$(4) \quad q = S Q, \quad p = \check{S} P$$

wird; dann sind nach § 18, Satz 4, Q_k, P_k wieder kanonisch konjugierte Variable.

Die Energiefunktion hat dann die Form

$$(5) \quad H = \sum_{k=1}^f \left\{ \frac{1}{2 \mu_k} P_k^2 + \frac{a_k}{2} Q_k^2 \right\}.$$

Sie erscheint als Summe von Funktionen je eines Variablenpaars P_k, Q_k , deren jede mit der Energiefunktion des einzelnen linearen Oszillators § 23, (1) übereinstimmt. Das Verfahren der Verschmelzung (§ 15) erlaubt dann, die lösenden Matrizen für das System von f ungekoppelten Oszillatoren aus den bekannten Matrizen für den einzelnen Oszillator abzuleiten.

Die Grundfrequenzen, die man nach der Regel § 23, (2) aus (5) bildet, nämlich

$$(6) \quad \nu_0^{(k)} = \frac{1}{2\pi} \sqrt{\frac{a_k}{\mu_k}}$$

sind nach § 16, Satz 2 unabhängig von der Wahl der Transformationsmatrix S in (3) durch die Energiefunktion (2) eindeutig bestimmt. Die Energie des Gesamtsystems ist dann und nur dann *entartet*, wenn zwischen den Grundfrequenzen eine oder mehrere lineare Beziehungen

$$(7) \quad \sum_{k=1}^f \tau_k \nu_0^{(k)} = 0$$

mit *ganzzahligen* (positiven oder negativen) τ_k bestehen.

Aus den P_k, Q_k kann man entsprechend § 23, (4) *komplexe Amplituden*

$$(8) \quad \left\{ \begin{array}{l} b_k = C_k (P_k - 2\pi i \nu_0^{(k)} \mu_k Q_k) \\ b_k^\dagger = C_k (P_k + 2\pi i \nu_0^{(k)} \mu_k Q_k) \end{array} \right\}, \quad C_k = \frac{1}{\sqrt{2h\nu_0^{(k)} \mu_k}}$$

bilden, für welche die Vertauschungsregeln

$$(9) \quad \left\{ \begin{array}{l} b_k b_l - b_l b_k = b_k^\dagger b_l^\dagger - b_l^\dagger b_k^\dagger = 0, \\ b_l b_k^\dagger - b_k^\dagger b_l = \delta_{lk} \end{array} \right.$$

gelten. Die Energie (5) verwandelt sich nach § 23, (8) bei der Einführung dieser b_k, b_k^\dagger in

$$(10) \quad \left\{ \begin{array}{l} H = \sum_{k=1}^f h \nu_0^{(k)} \left(b_k^\dagger b_k + \frac{1}{2} \right) \\ = \sum_{k=1}^f h \nu_0^{(k)} \left(b_k b_k^\dagger - \frac{1}{2} \right). \end{array} \right.$$

Wir wollen uns die Frage vorlegen, wieweit die so eingeführten komplexen Amplituden b_k durch das Problem (2) festgelegt sind. Die Antwort ergibt sich aus § 16, und zwar aus der an Satz 2 anknüpfenden Bemerkung, wonach die Matrix S bis auf den Faktor $T = \alpha^{-\frac{1}{2}} U \gamma^{\frac{1}{2}}$ bestimmt ist; dabei sind α, γ reelle Diagonalmatrizen und U eine orthogonale (reelle unitäre) Matrix. Bei Multiplikation von S mit einer Diagonalmatrix $(\beta_k \delta_{kl})$ geht nach (4) Q_k über in $Q'_k = \beta_k Q_k, P_k$ in $P'_k = \beta_k^{-1} P_k$; da aber H invariant ist, gehen die Größen μ_k und α_k in (5) über in $\mu'_k = \beta^{-2} \mu_k,$

$a'_k = \beta^{-2} a_k$, während $\nu_0^{(k)}$ nach (6) ungeändert bleibt. Nach (8) wird $C'_k = \beta C_k$, daher bleiben die Größen b_k und b_k^\dagger invariant. Die beiden Faktoren $\alpha^{-\frac{1}{2}}$ und $\gamma^{\frac{1}{2}}$ von T haben also keinen Einfluß auf b_k , wohl aber der Faktor U ; und zwar wird in symbolischer Schreibweise $b = Ub'$. Bei dieser Transformation aber muß die Diagonalmatrix $\left(\frac{1}{2} \frac{a_k}{\mu_k} \frac{1}{2} \delta_{kl}\right) = (\pi^2 \nu_0^{(k)^2} \delta_{kl})$, also auch $(h \nu_0^{(k)} \delta_{kl})$ invariant sein (s. § 16, Satz 2), also U eine Stufenmatrix, deren Stufengrößen den Vielfachheiten der $\nu_0^{(k)}$ entsprechen. So sehen wir: *Die b_k sind in demselben Maße durch die Energiefunktion (5) des Oszillatorensystems bestimmt, wie die Eigenvektoren einer quadratischen Form mit den Eigenwerten $h \nu_0^{(k)}$.*

Viertes Kapitel.

Die Sätze über den Drehimpuls¹.

§ 25. Vertauschungsregeln. In der klassischen Mechanik gelten bei frei beweglichen Systemen außer dem Energiesatz die Schwerpunkts- und Flächensätze. Die ersteren lassen sich in die Quantenmechanik der bisher entwickelten Form nicht ohne weiteres übertragen; denn die Matrizendarstellung der mechanischen Größen erlaubt eine adäquate Beschreibung nur von periodischen Vorgängen, nicht aber von translativen Schwerpunktsbewegungen. Wir werden die Schwerpunktsintegrale also zurückstellen, bis wir die Methoden zur Behandlung aperiodischer Vorgänge entwickelt haben.

Dagegen können wir hier das Analogon der Flächensätze vollständig behandeln, da es sich dabei um Drehungen, also um Vorgänge periodischen Charakters, handelt. Die kinematischen und dynamischen Eigenschaften des Drehimpulses eines Atoms oder einer Molekel sind von großer Bedeutung für die Struktur der Spektren; es ist daher angebracht, sie ausführlich zu untersuchen.

¹ Vgl. BORN, M., W. HEISENBERG u. P. JORDAN: Z. Phys. Bd. 35, S. 597. 1926. DIRAC, P. A. M.: Proc. Roy. Soc. Lond. (A) Bd. 111, S. 281. 1926. Die hier gegebenen vereinfachten Ableitungen rühren zum großen Teil von W. PAULI jun. her.

Wir betrachten ein System von f' Massenpunkten mit den $f = 3f'$ kartesischen Koordinaten

$$\begin{aligned} x_1, x_2, \dots, x_{f'}, \\ y_1, y_2, \dots, y_{f'}, \\ z_1, z_2, \dots, z_{f'} \end{aligned}$$

und den Impulsen

$$\begin{aligned} p_{1x}, p_{2x}, \dots, p_{f'x}, \\ p_{1y}, p_{2y}, \dots, p_{f'y}, \\ p_{1z}, p_{2z}, \dots, p_{f'z}. \end{aligned}$$

Wir definieren als *Drehimpuls des k -ten Teilchens* um den Koordinaten-Nullpunkt, wie in der klassischen Mechanik, den Vektor \mathfrak{M}_k mit den Komponenten

$$(1) \quad \begin{cases} M_{kx} = y_k p_{ky} - z_k p_{kz}, \\ M_{ky} = z_k p_{kx} - x_k p_{kz}, \\ M_{kz} = x_k p_{ky} - y_k p_{kx}. \end{cases}$$

Die Eigenschaften dieses Vektors haben wir zu untersuchen, und zwar zunächst seine *Multiplikationseigenschaften*.

Man findet nach den Regeln $[F, q_k] = \frac{\partial F}{\partial p_k}$, $[F, p_k] = -\frac{\partial F}{\partial q_k}$ [s. § 18, (3)]:

$$(2) \quad [M_{kx}, x_k] = 0, \dots, [M_{kx}, p_{kx}] = 0, \dots$$

$$(3) \quad \begin{cases} [M_{kz}, x_k] = -y_k, \dots \\ [M_{kz}, y_k] = x_k, \dots \end{cases}$$

$$(4) \quad \begin{cases} [M_{kz}, p_{kx}] = -p_{ky}, \dots \\ [M_{kz}, p_{ky}] = p_{kx}, \dots \end{cases}$$

Dabei ist jedesmal von drei Gleichungen, die durch zyklische Vertauschung der Koordinatenrichtungen x, y, z auseinander hervorgehen, nur eine angeschrieben, und dasselbe soll auch im folgenden geschehen.

Wir werden von jetzt an eine Koordinatenrichtung, etwa z , vor den beiden andern auszeichnen. So setzen wir

$$(5) \quad \begin{cases} \xi_k = x_k + i y_k, \\ \eta_k = p_{kx} + i p_{ky}. \end{cases}$$

Dann kann man statt (3) und (4) auch schreiben

$$(3a) \quad [M_{kz}, \xi_k] = i \xi_k,$$

$$(4a) \quad [M_{kz}, \eta_k] = i \eta_k.$$

Man hat ferner

$$\begin{aligned} [M_{ky}, M_{kz}] &= [M_{ky}, x_k p_{ky} - y_k p_{kx}] \\ &= [M_{ky}, x_k] p_{ky} - y_k [M_{ky}, p_{kx}], \end{aligned}$$

also nach (3) und (4)

$$(6) \quad \begin{cases} [M_{ky}, M_{kz}] = -M_{kx}, \\ [M_{kz}, M_{kx}] = -M_{ky}, \\ [M_{kx}, M_{ky}] = -M_{kz}. \end{cases}$$

Übrigens gelten alle diese Gleichungen formal ebenso in der klassischen Mechanik [s. Anhang 2, (21) bis (27)].

Mit der Abkürzung

$$(7) \quad \mathbf{M}_k = M_{kx} + i M_{ky}, \quad \mathbf{M}_k^\dagger = M_{kx} - i M_{ky}$$

kann man (6) umschreiben in

$$(6a) \quad [M_{kz}, \mathbf{M}_k] = i \mathbf{M}_k,$$

$$(6b) \quad [\mathbf{M}_k, \mathbf{M}_k^\dagger] = 2i M_{kz}.$$

Durch Summation über die f' Massenpunkte bilden wir das *gesamte Impulsmoment*

$$(8) \quad \mathfrak{M} = \sum_k \mathfrak{M}_k;$$

ferner setzen wir

$$(9) \quad \mathfrak{M} = \sum_k \mathbf{M}_k = M_x + i M_y.$$

Da nun die zu *verschiedenen* Teilchen gehörigen Koordinaten und Impulse stets vertauschbar sind, gewinnen wir unmittelbar aus (6) die Beziehungen (die wir ihrer Wichtigkeit wegen, neben der laufenden Numerierung, noch mit römischen Ziffern versehen):

$$(I) \quad (10) \quad \begin{cases} [M_y, M_z] = -M_x, \\ [M_z, M_x] = -M_y, \\ [M_x, M_y] = -M_z \end{cases}$$

oder aus (6a), (6b) die gleichwertigen

$$(I') \quad \begin{cases} (10a) & [M_z, M] = i M, \\ (10b) & [M, M^\dagger] = 2 i M_z. \end{cases}$$

Ferner entnehmen wir noch aus (2), (3a), (4a):

$$(11) \quad [M_z, z_k] = 0, \quad [M_z, p_{kz}] = 0,$$

$$(12) \quad \begin{cases} [M_z, \xi_k] = i \xi_k, \\ [M_z, \eta_k] = i \eta_k. \end{cases}$$

Für die Wechselwirkung des Systems mit Strahlung (und für manche andere Wirkungen) ist in erster Linie maßgebend das *elektrische Dipolmoment* \mathfrak{P} mit den Komponenten

$$(13) \quad \begin{cases} \mathfrak{P}_x = \sum_k e_k x_k = X, \\ \mathfrak{P}_y = \sum_k e_k y_k = Y, \\ \mathfrak{P}_z = \sum_k e_k z_k = Z, \end{cases}$$

wo e_k die elektrische Ladung des k -ten Teilchens ist. Setzen wir abkürzend

$$(14) \quad II = X + i Y,$$

so wird nach (11) und (12)

$$(II) \quad (15) \quad \begin{cases} [M_z, Z] = 0, \\ [M_z, II] = i II. \end{cases}$$

Daneben gelten natürlich die durch zyklische Vertauschung der Koordinatenrichtungen entstehenden Gleichungen. Wir heben noch folgende hervor:

$$(II') \quad (16) \quad \begin{cases} [II, M] = 0, \\ [II, M^\dagger] = 2 i Z. \end{cases}$$

§ 26. Vertauschbarkeit des Drehimpulses mit drehinvarianten Größen. In der klassischen Mechanik gelten die bekannten *Erhaltungssätze* für den Drehimpuls bei frei drehbaren Systemen: eine Komponente des Drehimpulses ist immer dann ein Integral, wenn das System eine Drehung um die entsprechende Achse „zuläßt“, d. h. wenn die Energiefunktion bei der Drehung invariant ist. Genau dasselbe gilt nun auch in der Quantenmechanik,

und zwar folgt das aus den in § 21 bewiesenen Sätzen, insbesondere Satz 5.

Um diese Sätze anwenden zu können, ist nur zu überlegen, wie eine Drehung in der Form

$$(1) \quad p'_k = U^{-1} p_k U, \quad q'_k = U^{-1} q_k U$$

dargestellt werden kann. Das kann einfach mit Hilfe der Formeln § 21, (22) geschehen. Eine infinitesimale Drehung um die z -Achse wird nämlich bekanntlich dargestellt durch

$$(2) \quad \begin{cases} x'_k = x_k - \alpha y_k, \\ y'_k = y_k + \alpha x_k, \\ z'_k = z_k, \end{cases}$$

wo α der infinitesimale Drehwinkel ist (der den Gruppenparameter λ von § 21 vertritt). Da diese Transformation *linear* ist, transformieren sich die zugehörigen Impulskomponenten nach § 18, Satz 4, kontragredient, d. h. bei Vernachlässigung von α^2 :

$$(3) \quad \begin{cases} p'_{kx} = p_{kx} - \alpha p_{ky}, \\ p'_{ky} = p_{ky} + \alpha p_{kx}, \\ p'_{kz} = p_{kz}. \end{cases}$$

Die Differentialgleichungen § 21, (22) lauten also hier:

$$(4) \quad \begin{cases} \frac{\partial F}{\partial p_{kx}} = \frac{2\pi}{h} y_k, & \frac{\partial F}{\partial x_k} = -\frac{2\pi}{h} p_{ky}, \\ \frac{\partial F}{\partial p_{ky}} = -\frac{2\pi}{h} x_k, & \frac{\partial F}{\partial y_k} = \frac{2\pi}{h} p_{kx}, \\ \frac{\partial F}{\partial p_{kz}} = 0, & \frac{\partial F}{\partial z_k} = 0; \end{cases}$$

sie werden gelöst durch

$$(5) \quad F = \frac{2\pi}{h} \sum_k (y_k p_{kx} - x_k p_{ky}) = -\frac{2\pi}{h} M_z.$$

Die Matrix, die der infinitesimalen Drehung um die z -Achse entspricht, ist also der mit $-\frac{2\pi}{h}$ multiplizierte Drehimpuls um diese Achse. Die gesamte Gruppe der endlichen Drehungen um die

z-Achse wird dann geliefert durch die unitäre Matrix

$$(6) \quad U(\alpha) = e^{-\frac{2\pi i}{\hbar} \alpha M_z},$$

wo α den Drehwinkel bedeutet¹.

Es ist nicht überflüssig, dieses Resultat durch eine kleine Rechnung unmittelbar zu bestätigen. Wir schreiben die Transformationsgleichungen für eine *endliche Drehung* an; das geht am bequemsten mit Hilfe der in § 25, (5) definierten Größen ξ_k, η_k , nämlich:

$$(7) \quad \begin{cases} \xi'_k = e^{i\alpha} \xi_k, & \eta'_k = e^{i\alpha} \eta_k, \\ z'_k = z_k, & p'_{kz} = p_{kz}, \end{cases}$$

wo α eine reelle Zahl, der Drehwinkel, ist. Nun gilt nach § 25, (12)

$$M_z \xi_k = \xi_k \left(M_z + \frac{\hbar}{2\pi} \right)$$

oder

$$\xi_k^{-1} M_z \xi_k = M_z + \frac{\hbar}{2\pi};$$

daraus folgt nach § 7, Satz 1

$$\xi_k^{-1} e^{i\alpha' M_z} \xi_k = e^{i\alpha' M_z} e^{i\alpha},$$

wo die Abkürzung

$$(8) \quad \alpha' = \frac{2\pi}{\hbar} \alpha$$

gebraucht ist. Dafür kann man schreiben:

$$e^{i\alpha' M_z} \xi_k e^{-i\alpha' M_z} = e^{i\alpha} \xi_k;$$

das ist aber nach (7) gleich ξ'_k . Man hat somit, indem man dieselbe Betrachtung auch für η anstellt und die Vertauschbarkeit von M_z mit z_k, p_{kz} beachtet, die Transformation (7) in

$$(9) \quad \begin{cases} \xi'_k = e^{i\alpha' M_z} \xi_k e^{-i\alpha' M_z}, \\ \eta'_k = e^{i\alpha' M_z} \eta_k e^{-i\alpha' M_z} \end{cases}$$

verwandelt, also in die Gestalt (1) gebracht, und zwar ist die Transformationsmatrix mit Rücksicht auf die Bedeutung (8)

¹ Dieser Drehwinkel hat nämlich die in § 21 für den Parameter λ geforderte Eigenschaft, daß der Zusammensetzung zweier Transformationen die *Addition* der Parameterwerte entspricht.

von α' gerade die durch (6) gegebene Funktion der Drehimpuls-komponente M_z .

Jede drehinvariante Funktion ist also mit der Matrix (6) vertauschbar; ist insbesondere das System um die z -Achse frei drehbar, also H gegen diese Drehungen invariant, so ist H mit M_z vertauschbar, M_z ist ein Integral. Dasselbe gilt für Drehungen um die x - bzw. y -Achse. Übrigens folgt aus § 25, (10) sofort, wie in der klassischen Mechanik, daß entweder nur eine Komponente des Drehimpulses konstant ist, oder alle drei; denn wenn zwei, etwa M_x und M_y , mit H vertauschbar sind, ist es notwendig auch $M_z = -[M_x, M_y]$. Der innere Grund hierfür liegt offenbar darin, daß eine Funktion, die invariant bei Drehungen um die x - und um die y -Achse ist, notwendig auch bei Drehungen um die z -Achse ungeändert bleibt.

Die Nichtvertauschbarkeit der Größen M_x, M_y, M_z hat gemäß § 21, Satz 2 zur Folge, daß ein System notwendig entartet ist, wenn alle drei Drehimpulssätze gelten; wir sprechen dann von Richtungs-entartung.

Unabhängig von der Gültigkeit der Erhaltungssätze hat man für jedes System das Ergebnis, daß die drehinvarianten Funktionen

$$(10) \quad r_k^2 = x_k^2 + y_k^2 + z_k^2,$$

$$(11) \quad p_k^2 = p_{kx}^2 + p_{ky}^2 + p_{kz}^2$$

und auch

$$(12) \quad r_{kl}^2 = (x_k - x_l)^2 + (y_k - y_l)^2 + (z_k - z_l)^2$$

mit M_z vertauschbar sind; natürlich kann das auch auf Grund der Formeln von § 25 durch einfaches Nachrechnen bestätigt werden.

Das *Quadrat des Gesamtdrehimpulses* ist definiert durch

$$(13) \quad \mathfrak{M}^2 = M_x^2 + M_y^2 + M_z^2;$$

da diese Größe offenbar drehinvariant ist, folgt die wichtige Tatsache, daß

$$(14) \quad [\mathfrak{M}^2, M_z] = 0$$

ist — was man natürlich wiederum auf Grund von § 25, (6) bestätigen kann. Da die entsprechende Gleichung für M_x und M_y gelten muß, hat man auch

$$(15) \quad [\mathfrak{M}^2, M] = 0.$$

Endlich haben wir noch eine Folgerung aus (6) bzw. (9) zu ziehen. Wählt man $\alpha = 2\pi$, also $\alpha' = \frac{4\pi^2}{h}$, so wird (7) die *identische* Substitution; die Gleichungen (9) besagen dann, daß

$$(16) \quad \left[e^{2\pi i \frac{2\pi}{h} M_z}, \xi_k \right] = 0, \quad \left[e^{2\pi i \frac{2\pi}{h} M_z}, \eta_k \right] = 0.$$

Wegen der stets vorausgesetzten Einfachheit des Systems aller Koordinaten- und Impulsmatrizen folgt daraus nach § 15, Satz 1, daß $e^{2\pi i \frac{2\pi}{h} M_z}$ bis auf einen Zahlenfaktor gleich der Einheitsmatrix ist. Man kann also offenbar setzen:

$$(17) \quad e^{2\pi i \frac{2\pi}{h} M_z} = e^{2\pi i \alpha_0} E.$$

Das bedeutet aber nach § 14, (1):

Alle Eigenwerte $M_z(n)$ von M_z haben bei einem bestimmten System die Form

$$(18) \quad M_z(n) = (\alpha_0 + \bar{m}) \frac{h}{2\pi}$$

mit ganzzahligem \bar{m} und festem α_0 .

§ 27. Zerlegung der Drehimpulsmatrizen in einfache Bestandteile. Durch die Vertauschungsregeln (I), (I') [§ 25, (10), (10a), (10b)] für die Komponenten des Drehimpulses und die daraus folgenden Gleichungen § 26, (14), (15) ist die mathematische Struktur dieser Matrizen schon weitgehend festgelegt, unabhängig von der Beschaffenheit der Energiefunktion. Dies werden wir jetzt durch Entwicklung der aus (I), (I') fließenden Folgerungen zeigen.

Natürlich ist nicht zu erwarten, daß das Matrixsystem M_x, M_y, M_z für sich einfach ist. Es kann jedoch nach § 15 derart unitär transformiert werden, daß es in die direkte Summe seiner irreduziblen Bestandteile übergeht; einen dieser einfachen Bestandteile wollen wir mit $\frac{2\pi}{h}$ multiplizieren und dann mit m_x, m_y, m_z bezeichnen.

Es sei also m_x, m_y, m_z ein irreduzibler Bestandteil des Matrixsystems $\frac{2\pi}{h} M_x, \frac{2\pi}{h} M_y, \frac{2\pi}{h} M_z$.

Die bei der Definition der irreduzibeln Bestandteile noch willkürliche unitäre Transformation [s. § 15] wählen wir so, daß m_z eine *Diagonalmatrix* wird. Die Eigenwerte von m_z bezeichnen wir mit m , so daß

$$(1) \quad m_z(r, s) = m \delta_{rs}$$

ist.

Bei unitärer Transformation bleiben die Vertauschungsgleichungen (I), § 25, (10) erhalten. Denkt man sich dabei die M_x , M_y , M_z in ihrer reduzierten Form als gleichstufige Stufenmatrizen, so müssen dieselben Gleichungen für jede Stufe einzeln gültig sein; somit gelten sie — abgesehen von einem Faktor $\frac{h}{2\pi}$ — auch für die Matrizen m_x , m_y , m_z . Mit den Bezeichnungen

$$(2) \quad \begin{cases} \mu = m_x + i m_y, \\ m^2 = m_x^2 + m_y^2 + m_z^2 \end{cases}$$

erhält man also aus § 25, (10a), (10b):

$$(3) \quad \begin{cases} \text{a) } m_z \mu - \mu m_z = \mu, \\ \text{b) } \mu \mu^\dagger - \mu^\dagger \mu = 2 m_z, \end{cases}$$

und aus § 26, (14):

$$(4) \quad [m^2, m_x] = [m^2, m_y] = [m^2, m_z] = 0.$$

Da nun m_x , m_y , m_z ein einfaches System bilden, folgt nach § 15, Satz 1, aus (4), daß m^2 proportional der Einheitsmatrix wird:

$$(5) \quad m^2 = \omega^2 \cdot E.$$

Dann ergibt sich der zu dem betrachteten einfachen Bestandteil $m = (m_x, m_y, m_z)$ gehörige Eigenwert von \mathfrak{M}^2 als

$$(6) \quad \mathfrak{M}^2(n) = \frac{h^2}{4\pi^2} \omega^2.$$

Die Gleichung (3a) ist bis auf die Bezeichnung dieselbe wie beim harmonischen Oszillator die Gleichung (5a), § 23; nur stand dort b statt μ und $-W/h\nu_0$ statt m_z . Es folgt also wörtlich wie in § 23: *Die verschiedenen Eigenwerte m von m_z bilden eine lückenlose arithmetische Reihe mit der Differenz 1.* Da die Matrix $m^2 - m_z^2$

$= m_x^2 + m_y^2$ semi-definit ist, folgt ferner, daß ihre Eigenwerte $\omega^2 - m^2$ nicht negativ sind, daß also

$$(7) \quad |m| \leq \omega$$

ist. Folglich gibt es für jeden einfachen Bestandteil m nur *endlich viele* verschiedene Werte m .

Unser nächstes Ziel ist, zu zeigen, daß alle Eigenwerte m *einfach* sind. Hierzu denken wir uns m_z als *geordnete* Diagonalmatrix geschrieben und m_x, m_y als Übermatrizen entsprechend den (noch unbekanntenen) Vielfachheiten der Eigenwerte von m_z (vgl. die analoge Überlegung beim harmonischen Oszillator, § 23), also

$$(8) \quad \mu = (\mu^{(r,s)}), \quad \mu^\dagger = (\mu^{\dagger(r,s)}).$$

Wir können als Indizes r, s einfach die betreffenden m -Werte wählen; dann bedeutet die erste Gleichung (3):

$$(9) \quad \begin{cases} \mu^{(r,s)} = 0, & \text{wenn } r - s \neq 1, \\ \mu^{\dagger(r,s)} = 0, & \text{wenn } r - s \neq -1; \end{cases}$$

[entsprechend den Gleichungen § 23, (12)]. Nun ist nach (2)

$$(10) \quad \mu \mu^\dagger + \mu^\dagger \mu = 2(m_x^2 + m_y^2) = 2(m^2 - m_z^2),$$

also nach (3 b) und (10):

$$(11) \quad \begin{cases} \mu^\dagger \mu = m^2 - m_z(m_z + 1), \\ \mu \mu^\dagger = m^2 - m_z(m_z - 1). \end{cases}$$

Das gibt wegen (9):

$$(12) \quad \begin{cases} \mu^{\dagger(m,m+1)} \mu^{(m+1,m)} = \{\omega^2 - m(m+1)\} E^{(m)}, \\ \mu^{(m+1,m)} \mu^{\dagger(m,m+1)} = \{\omega^2 - m(m+1)\} E^{(m+1)}, \end{cases}$$

wo $E^{(m)}, E^{(m+1)}$ die entsprechenden Einheitsmatrizen sind.

Ebenso wie in § 23, (15) schließt man *jetzt* durch Bildung der *Spuren* in (12), daß alle Eigenwerte m die *gleiche* Vielfachheit g besitzen. Danach kann man — wörtlich wie beim Oszillator — aus (12) weiter erkennen, daß im Falle $g > 1$ das System m_x, m_y, m_z zerlegbar wäre. Folglich muß $g = 1$ sein.

Die Einfachheit der Eigenwerte m läßt uns aber aus der Gleichung (3 b) einen wichtigen Schluß ziehen: Bilden wir auf beiden Seiten die Spur, so ergibt sich links (wegen der Endlich-

keit der Matrizen) der Wert Null; also ist die Summe aller Eigenwerte m von m_z gleich Null, die arithmetische Reihe der m muß *symmetrisch zur Null* liegen und dafür gibt es nur zwei Möglichkeiten: entweder sind alle m -Werte ganze Zahlen oder sie fallen in die Mitte zwischen zwei ganzen Zahlen. Man sagt kurz: m ist *ganzzahlig oder halbzahlig*. Sei $j \geq 0$ der größte der endlich vielen möglichen m -Werte, dann sind diese durch die Zahlenreihe

$$(13) \quad m = -j, -j + 1, \dots, j - 1, j$$

gegeben, wo j eine Zahl aus den Reihen

$$(13a) \quad \left\{ \begin{array}{l} j = 0, 1, 2, \dots \\ \text{oder} \\ j = \frac{1}{2}, \frac{3}{2}, \frac{5}{2}, \dots \end{array} \right.$$

ist.

Die Gleichungen (9) und (12) gelten wegen der Einfachheit der Eigenwerte m unmittelbar für die Matrixelemente von μ, μ^\dagger , von denen wir die nicht verschwindenden jetzt mit $\mu(m+1, m)$, $\mu^\dagger(m, m+1)$ bezeichnen wollen. Man sieht nun, daß in der Matrix $\mu^\dagger \mu$ das Element

$$\mu^\dagger \mu(j, j) = 0$$

wird, da es kein von Null verschiedenes Element $\mu(m, j)$ in μ gibt.

Aus (12) folgt deshalb:

$$(14) \quad \omega^2 = j(j+1).$$

Damit haben wir die Eigenwerte von \mathfrak{M}^2 zu

$$(15) \quad \mathfrak{M}^2(n) = \frac{\hbar^2}{4\pi^2} j(j+1)$$

bestimmt.

Endlich ergibt sich noch aus (12)

$$(16) \quad |\mu(m+1, m)|^2 = j(j+1) - m(m+1), \quad -j \leq m < j.$$

Es muß also bei geeigneter Wahl der Phasenkonstanten

$$(16a) \quad \mu(m+1, m) = \sqrt{j(j+1) - m(m+1)}$$

werden; und damit sind umgekehrt alle Bedingungsgleichungen befriedigt.

Zur Erläuterung geben wir die Matrizen m_z und μ für die Fälle $j = 0$, $j = \frac{1}{2}$, $j = 1$ an:

$$17) \quad \begin{cases} j = 0: m_z^{(0)} = (0), & \mu^{(0)} = (0); \\ j = \frac{1}{2}: m_z^{(\frac{1}{2})} = \begin{pmatrix} -\frac{1}{2} & 0 \\ 0 & \frac{1}{2} \end{pmatrix}, & \mu^{(\frac{1}{2})} = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}; \\ j = 1: m_z^{(1)} = \begin{pmatrix} -1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}, & \mu^{(1)} = \begin{pmatrix} 0 & 0 & 0 \\ \sqrt{2} & 0 & 0 \\ 0 & \sqrt{2} & 0 \end{pmatrix}. \end{cases}$$

Damit haben wir die Matrixensysteme, die als einfache Bestandteile von M_x , M_y , M_z auftreten können, vollständig bestimmt. Es sei noch hervorgehoben, daß bei ein und demselben quantenmechanischen System entweder *alle* vorkommenden m ganzzahlig oder *alle* halbzahlig sein müssen (in *allen* einfachen Bestandteilen); das folgt z. B. aus dem am Ende von § 26 in Formel (18) ausgesprochenen Satze.

§ 28. Auswahlregeln für die Drehimpuls-Quantenzahlen. Während die vorstehenden Betrachtungen rein „kinematischer“ Natur waren, wenden wir uns jetzt zu ihrer Anwendung auf die *Dynamik*. Dann werden die ganz- oder halbzahlgigen Eigenwerte m und ihre Maximalwerte j zu „Quantenzahlen“ im Sinne der älteren BOHRschen Theorie.

Wir betrachten zunächst ein System, in welchem der eine Drehimpuls-Erhaltungssatz $\dot{M}_z = 0$ gültig ist. Es gibt dann nach den Bemerkungen von § 21 gewiß eine Lösung der Bewegungsgleichungen, bei der M_z eine Diagonalmatrix ist; sie hat also die durch § 27, (1) angegebene Form mit ganz- oder halbzahlgigen m -Werten. Da dieser Fall praktisch häufig z. B. in der Weise realisiert ist, daß ein mechanisches System unter einem konstanten Magnetfeld parallel zur z -Richtung steht, heißt m *magnetische Quantenzahl*; sie korrespondiert offenbar mit der in Bd. I, § 17 und § 34 ebenso bezeichneten Größe.

Wir wollen nun für diese Lösung das Verhalten des elektrischen Dipolmoments X, Y, Z oder der beiden Größen $\Pi = X + iY$ und Z [s. § 25, (14)] untersuchen.

Sie lassen sich als Übermatrizen nach dem Index m schreiben; die Teilmatrizen von einer Matrix Q seien also $Q^{(m, m')}$. Insbesondere

$$\text{ist } M_z = (M_z^{(m, m')}) \text{ mit } M_z^{(m, m')} = \frac{\hbar}{2\pi} m \delta_{m m'}.$$

Dann folgt aus den Gleichungen § 25, (15):

$$(1) \quad \begin{cases} (m - m') Z^{(m, m')} = 0, \\ (m - m' - 1) II^{(m, m')} = 0, \end{cases}$$

und das bedeutet

$$(1a) \quad \begin{cases} Z^{(m, m')} = 0 & \text{für } m - m' \neq 0, \\ II^{(m, m')} = 0 & \text{für } m - m' \neq 1; \end{cases}$$

für II^\dagger erhält man also

$$(1b) \quad II^\dagger{}^{(m, m')} = 0 \quad \text{für } m - m' \neq -1.$$

Das kann man so aussprechen:

In dem Spektrum des Systems, soweit es von der Dipolstrahlung herrührt, gibt es nur solche Frequenzen, die zu Änderungen der magnetischen Quantenzahl m um

$$(2) \quad \Delta m = 0 \quad \text{oder} \quad \Delta m = \pm 1$$

gehören. Dabei bedeutet der Übergang

$$(3) \quad \begin{cases} m \rightarrow m & \text{parallel zur } z\text{-Achse linear schwingende} \\ & (II = 0, Z \neq 0), \\ m \rightarrow m \pm 1 & \text{um die } z\text{-Achse rechts bzw. links zirkular} \\ & \text{schwingende } (II \text{ bzw. } II^\dagger \neq 0, Z = 0) \end{cases}$$

Dipolstrahlung.

Man sieht, daß hier die neue Quantenmechanik vollständig mit der älteren (Bd. 1, § 17, § 34) übereinstimmt.

Wir wollen jetzt ferner annehmen, daß auch $\dot{M}_x = 0$, $\dot{M}_y = 0$ sind, wie es bei einem *frei drehbaren* System der Fall ist. Dann gibt es gemäß § 21, Satz 3 eine Lösung der mechanischen Gleichungen, bei welcher M_z und \mathfrak{M}^2 als Diagonalmatrizen erscheinen, während das Matrixsystem M_x, M_y, M_z vollständig ausreduziert ist, also als direkte Summe seiner in § 27 untersuchten einfachen Bestandteile erscheint. Auf diese spezielle Lösung wollen wir uns jetzt beziehen.

Wir haben jetzt neben der magnetischen Quantenzahl m noch ihren Maximalwert j , der nach § 27, (15) die Eigenwerte von \mathfrak{M}^2 bestimmt; j heißt *innere Quantenzahl* nach einer, in der älteren Quantentheorie von SOMMERFELD von spektroskopischen Gesichtspunkten aus eingeführten Bezeichnungweise. (In Bd. 1, § 17,

§ 34 wurde das Zeichen j in der korrespondierenden Bedeutung gebraucht, die Benennung allerdings nicht.)

Wir haben in § 27 gesehen, daß die einzelnen irreduzibeln Bestandteile des Systems M_x, M_y, M_z nur von dem Werte j abhängen; aber es können mehrere gleiche Bestandteile (mit demselben j) vorkommen (entsprechend dem physikalischen Umstande, daß es mehrere Zustände verschiedener Energie mit gleichem Gesamtdrehimpuls gibt). Daher ist zur Numerierung der einfachen Bestandteile ein weiterer Index n nötig. Sodann kann man (anders als oben!) alle Matrizen als Übermatrizen mit dem Indexpaar n, j schreiben, wobei die Elemente der Teilmatrizen durch den Index m gekennzeichnet sind: $Q = (Q^{(n,j;n',j')}(m, m'))$. Die Indizes n, n' werden wir gewöhnlich weglassen, also $Q^{(j,j)}(m, m')$ schreiben. Bei dieser Schreibweise werden M_x, M_y, M_z Stufenmatrizen mit den einfachen Bestandteilen als Stufen. Die Aussagen (1) bzw. (2), (3) hat man jetzt so zu schreiben:

$$(4) \quad \begin{cases} Z^{(j,j)}(m, m') = 0 & \text{für } m - m' \neq 0, \\ II^{(j,j)}(m, m') = 0 & \text{für } m - m' \neq 1. \end{cases}$$

Alle Komponenten von \mathfrak{M} sind (wegen der vorausgesetzten Ausreduzierung) Diagonalmatrizen in bezug auf j . Nach § 27, (16a) hat man

$$(5) \quad \begin{cases} M^{(j,j)}(m, m') = 0 & \text{für } m - m' \neq 1, \\ M^{(j,j)}(m + 1, m) = \frac{\hbar}{2\pi} \sqrt{j(j+1) - m(m+1)}. \end{cases}$$

Daher folgt aus der ersten Gleichung § 25, (16):

$$(6) \quad M^{(j,j)}(m + 1, m) II^{(j,j')}(m, m - 1) \\ = II^{(j,j')}(m + 1, m) M^{(j',j')}(m, m - 1).$$

Da aber j der Maximalwert von m ist, so ist hier einer der Faktoren der linken Seite durch Null zu ersetzen, sobald eine der Ungleichungen

$$(6a) \quad -j \leq m \leq j - 1, \quad -j' + 1 \leq m \leq j' + 1$$

nicht erfüllt ist; und ebenso ist einer der Faktoren der rechten Seite durch Null zu ersetzen, sobald

$$(6b) \quad -j - 1 \leq m \leq j - 1, \quad -j' + 1 \leq m \leq j'$$

nicht erfüllt ist.

Wir suchen nun solche Wertssysteme von j, j', m auf, für die zwar (6a), aber nicht (6b) erfüllt ist, oder umgekehrt zwar (6b) aber nicht (6a). In einem solchen Falle haben wir statt (6) entweder — bei Verletzung von (6a):

$$(7a) \quad II^{(j, j')}(m+1, m) = 0,$$

oder — bei Verletzung von (6b):

$$(7b) \quad II^{(j, j')}(m, m-1) = 0.$$

Ein solcher Fall tritt *nicht* ein, wenn $\Delta j = j - j'$ einen der Werte $0, \pm 1$ besitzt. Denn für $\Delta j = 0, \pm 1$ bedeutet das Ungleichungspaar (6a) genau denselben Wertebereich für m wie das Ungleichungspaar (6b), nämlich

$$\text{für } j' = j: \quad -j + 1 \leq m \leq j - 1,$$

$$\text{für } j' = j + 1: \quad -j \leq m \leq j - 1,$$

$$\text{für } j' = j - 1: \quad -j + 2 \leq m \leq j - 1.$$

Wir betrachten nun aber die Fälle $\Delta j \leq -2$ und $\Delta j \geq +2$. Ist $j' = j + s$ mit $s \geq 2$, so ist für $m = -j - 1$ zwar (6b) erfüllt, aber nicht (6a). Nach (7a) wird also

$$II^{(j, j+s)}(-j, -j-1) = 0;$$

und auf Grund des rekursiven Charakters der Formel (6) in bezug auf m folgt nun der Reihe nach für alle Werte $m+1 = -j, -j+1, \dots, j-1, j$ die Gleichung

$$(8a) \quad II^{(j, j+s)}(m+1, m) = 0 \quad \text{für } s \geq 2.$$

Ist andererseits $j' = j - s$ oder $j = j' + s$ mit $s \geq 2$, so ist für $m = j' + 1$ zwar (6a), aber nicht (6b) erfüllt. Also ist nach (7b)

$$II^{(j'+s, j')}(j'+1, j') = 0$$

und wie oben folgt aus (6) durch sukzessives Einsetzen der Werte $m-1 = j', j'-1, \dots, -j'+1, -j'$:

$$(8b) \quad II^{(j'+s, j')}(m, m-1) = 0 \quad \text{für } s \geq 2.$$

Man kann (8a) und (8b) zusammenfassen in die Aussage:

$$(9) \quad II^{(j, j')}(m+1, m) = 0, \quad \text{außer für } \Delta j = 0, \pm 1.$$

Nun kann man Z aus II nach § 25, (16) berechnen:

$$(10) \quad 2iZ = [II, M^\dagger],$$

und da M bezüglich j Diagonalmatrix ist, so folgt aus (9) auch

$$(11) \quad Z^{(i,j)}(m, m) = 0, \quad \text{au\ss}er \text{ f\"ur } \Delta j = 0, \pm 1.$$

Wir erhalten also (f\"ur Dipolstrahlung) die *Auswahlregel f\"ur die innere Quantenzahl j* :

$$(12) \quad \Delta j = 0 \quad \text{oder} \quad \pm 1.$$

Auch dieses Resultat korrespondiert genau mit einem der \u00e4lteren Quantentheorie. Dort (s. Bd. 1, § 17) war bei einem frei drehbaren System der gesamte Drehimpuls $|\mathfrak{M}|$ zu quanteln und wurde gleich $\frac{h}{2\pi}j$ gesetzt, so da\ss $\mathfrak{M}^2 = \frac{h^2}{4\pi^2}j^2$ war. In der neuen Theorie hat man f\"ur die Eigenwerte von \mathfrak{M}^2 nach § 27, (15) den Ausdruck $\frac{h^2}{4\pi^2}j(j+1)$, der f\"ur gro\ss e j in den klassischen \u00fcbergeht. Die Abweichung (Ersetzung der Potenz j^2 durch das Produkt $j(j+1)$) ist f\"ur die Quantenmechanik charakteristisch. Bereits vor ihrer systematischen Entwicklung hat LAND\u00c9¹ festgestellt, da\ss in spektroskopischen Formeln (Multipletts und ZEEMAN-Effekte) an Stellen, wo man korrespondenzm\u00e4\ss igh das Auftreten von $\frac{4\pi^2}{h^2}\mathfrak{M}^2$ erwarten mu\ss te, die Erfahrung nicht j^2 , sondern $j(j+1)$ fordert.

§ 29. Schwingungsamplituden beim ZEEMAN-Effekt. Die Beeinflussung von Atomen oder Molekeln durch konstante \u00e4u\ss ere Felder (ZEEMAN-Effekt, STARK-Effekt) l\u00e4\ss t sich unter folgendem Ansatz fassen:

Die Energiefunktion H sei von einem Parameter λ abh\u00e4ngig:

$$(1) \quad H = H(p, q; \lambda),$$

und zwar derart, da\ss f\"ur einen bestimmten Wert von λ , etwa $\lambda = 0$, alle drei Drehimpulss\u00e4tze gelten, da\ss aber f\"ur $\lambda \neq 0$ nur noch M_z ein Integral ist.

Die innere Quantenzahl j , die zun\u00e4chst nur f\"ur den Fall $\lambda = 0$ definiert war, kann nun auch f\"ur $\lambda \neq 0$ erkl\u00e4rt werden. Die allgemeine L\u00f6sung des durch (1) gegebenen mechanischen Problems, als Funktion von λ betrachtet, mu\ss n\u00e4mlich f\"ur $\lambda \rightarrow 0$ stetig in eine L\u00f6sung des Problems mit $\lambda = 0$ \u00fcbergehen; wir ordnen je-

¹ LAND\u00c9, A.: Z. Phys. Bd. 5, S. 231. 1921; Bd. 15, S. 189. 1923; Bd. 19, S. 112. 1923.

dem Eigenwert der allgemeinen Energiefunktion (1) denjenigen Wert von j zu, der ihm in diesem Grenzfall $\lambda = 0$ zukommt.

Gemäß den Bemerkungen in § 21 ergibt sich für $\lambda = 0$ eine Entartung der Energie von solcher Art, daß jedenfalls zu einem irreduzibeln Bestandteil von M_x, M_y, M_z , d. h. zu einem Indexpaar n, j , ein und derselbe Energiewert (unabhängig von m) gehört. Für $\lambda \neq 0$ tritt dann im allgemeinen eine Aufspaltung dieses mehrfachen Eigenwerts ein. Die Intensitätsverhältnisse der aufgespaltenen Spektrallinien können für kleine Werte von λ entnommen werden aus den Grenzwerten, denen die Matrixelemente

$$(2) \quad II^{(n, j; n', j')} (m, m'), \quad Z^{(n, j; n', j')} (m, m')$$

für den Grenzfall $\lambda \rightarrow 0$ zustreben.

Wir wollen deshalb jetzt die Matrixelemente (2) für diese Lösung des Problems mit $\lambda = 0$ — für welche also M_z Diagonalmatrix ist — bezüglich ihrer Abhängigkeit von m wirklich ausrechnen. Hierfür benutzen wir die Formeln (4) bis (12) des § 28, die sich auf eben diese Lösung des entarteten Problems beziehen. Dabei haben wir die 3 Fälle $\Delta j = 0, +1, -1$ gesondert zu betrachten.

α) Elemente mit $\Delta j = 0$. Für $j = j'$ wird die Gleichung § 28, (6) eine einfache Proportionalität, aus der man

$$(3) \quad II^{(j, j)} (m + 1, m) = M^{(j, j)} (m + 1, m) \cdot a \frac{2\pi}{h}$$

schließen kann, wo $a = a^{(n, j; n', j)}$ (oder kurz: $a^{(j, j)}$) von m unabhängig ist. Mit dem in (5) gegebenen Werte des Elements von M wird

$$(4) \quad II^{(j, j)} (m + 1, m) = \sqrt{j(j+1) - m(m+1)} \cdot a.$$

Z läßt sich aus (II') [§ 25, (16)] berechnen; da M bezüglich j Diagonalmatrix ist, kommen für $Z^{(j, j)}$ bei II auch nur die Teilmatrizen $II^{(j, j)}$ in Betracht, deren Elemente nach (3) denen von $M^{(j, j)}$ proportional sind. Es wird also

$$2i Z^{(j, j)} = \left([M, M^\dagger] a \frac{2\pi}{h} \right)^{(j, j)}$$

und das ist nach (I') [§ 25, (10b)] gleich $2i (M_z a)^{(j, j)} \frac{2\pi}{h}$.

Nach § 27, (1) erhält man

$$(5) \quad Z^{(j, j)} (m, m) = m \cdot a.$$

Wegen $Z = Z^\dagger$ muß auch $a = a^\dagger$ sein.

β) Elemente mit $\Delta j = +1$. Für $j' = j - 1$ wird die Gleichung § 28, (6) mit Rücksicht auf § 28, (5):

$$\begin{aligned} & \sqrt{j(j+1) - m(m+1)} II^{(j,j-1)}(m, m-1) \\ &= \sqrt{(j-1)j - (m-1)m} II^{(j,j-1)}(m+1, m). \end{aligned}$$

Um dies in eine Proportionalität zu verwandeln, zerlegen wir die Radikanden zunächst in Faktoren:

$$(6) \quad \begin{cases} j(j+1) - m(m+1) = (j+m+1)(j-m), \\ (j-1)j - (m-1)m = (j+m-1)(j-m). \end{cases}$$

Der Faktor $\sqrt{j-m}$ fällt also aus der obigen Gleichung heraus. Statt seiner fügen wir beiderseits den Faktor $\sqrt{j+m}$ hinzu:

$$\begin{aligned} & \sqrt{(j+m+1)(j+m)} II^{(j,j-1)}(m, m-1) \\ &= \sqrt{(j+m-1)(j+m)} II^{(j,j-1)}(m+1, m). \end{aligned}$$

Nunmehr geht die Wurzel der rechten Seite aus der Wurzel der linken Seite durch Vertauschung von m mit $m-1$ hervor. Daher kann man schließen:

$$II^{(j,j-1)}(m+1, m) = \sqrt{(j+m)(j+m+1)} b.$$

Aus (II') [§ 25, (16)] folgt nun weiter:

$$\begin{aligned} 2i Z^{(j,j-1)}(m, m) &= \frac{2\pi i}{h} \{ II^{(j,j-1)}(m, m-1) \mathbf{M}^{\dagger(j-1,j-1)}(m-1, m) \\ &\quad - \mathbf{M}^{\dagger(j,j)}(m, m+1) II^{(j,j-1)}(m+1, m) \} \\ &= i \sqrt{j^2 - m^2} ((j+m-1) - (j+m+1)) b \\ &= -2i \sqrt{j^2 - m^2} b, \end{aligned}$$

also

$$(7) \quad Z^{(j,j-1)}(m, m) = -\sqrt{j^2 - m^2} b.$$

γ) Elemente mit $\Delta j = -1$. In genau derselben Weise erhält man für $j' = j + 1$ aus § 28, (5), (6):

$$\begin{aligned} & \sqrt{(j-1)j - m(m+1)} II^{(j-1,j)}(m, m-1) \\ &= \sqrt{j(j+1) - (m-1)m} II^{(j-1,j)}(m+1, m). \end{aligned}$$

Nun hat man die Faktorenerlegungen

$$(8) \quad \begin{cases} (j-1)j - m(m+1) = (j+m)(j-m-1), \\ j(j+1) - (m-1)m = (j+m)(j-m+1); \end{cases}$$

der Faktor $\sqrt{j+m}$ fällt fort, an seiner Stelle wird $\sqrt{j-m}$ hinzugefügt:

$$\begin{aligned} & \sqrt{(j-m)(j-m-1)} II^{(j-1, j)}(m, m-1) \\ &= \sqrt{(j-m+1)(j-m)} II^{(j-1, j)}(m+1, m). \end{aligned}$$

Daraus folgt

$$(9) \quad II^{(j-1, j)}(m+1, m) = \sqrt{(j-m)(j-m-1)} b'.$$

Ferner liefert (II') [§ 25, (16)]:

$$\begin{aligned} 2i Z^{(j-1, j)}(m, m) &= \frac{2\pi i}{h} (II^{(j, j-1)}(m, m-1) \mathbf{M}^{\dagger(j, j)}(m-1, m) \\ &\quad - \mathbf{M}^{\dagger(j-1, j-1)}(m, m+1) II^{(j-1, j)}(m+1, m)) \\ &= i \sqrt{j^2 - m^2} ((j-m+1) - (j-m-1)) b' \\ &= 2i \sqrt{j^2 - m^2} b', \end{aligned}$$

$$(10) \quad Z^{(j-1, j)}(m, m) = \sqrt{j^2 - m^2} b'.$$

Da nun für die Elemente von Z zwei unabhängige Darstellungen (7) und (10) vorliegen, kann man daraus eine Beziehung zwischen den Matrizen b und b' herleiten. Z muß hermitisch sein, also $Z^{(n, j; n', j-1)}(m, m)$ konjugiert zu $Z^{(n', j-1; n, j)}(m, m)$; daraus folgt

$$(11) \quad b' = -b^\dagger.$$

Damit haben wir alle nicht verschwindenden Elemente von II und Z in ihrer Abhängigkeit von m bestimmt. Wir fassen das Ergebnis zusammen:

$$(12) \quad \begin{cases} II^{(j, j)}(m+1, m) = \sqrt{(j+m+1)(j-m)} a, \\ Z^{(j, j)}(m, m) = m a; \quad a = a^\dagger. \end{cases}$$

$$(13) \quad \begin{cases} II^{(j, j-1)}(m+1, m) = \sqrt{(j+m+1)(j+m)} b, \\ II^{(j-1, j)}(m+1, m) = -\sqrt{(j-m-1)(j-m)} b^\dagger, \\ Z^{(j, j-1)}(m, m) = -\sqrt{j^2 - m^2} b, \\ Z^{(j-1, j)}(m, m) = -\sqrt{j^2 - m^2} b^\dagger. \end{cases}$$

Die adjungierten Matrizen lassen sich hieraus ohne weiteres bilden.

Bevor wir aus den Amplituden (12), (13) die Intensitäten beim ZEEEMAN-Effekt berechnen, haben wir einige allgemeine Überlegungen über Drehinvarianten anzustellen.

§ 30. Dreh- und Spiegelungsinvarianten von Raumvektoren und Raumtensoren. Bei einem System mit entarteter Energie ist die Lösung q_k, p_k der Bewegungsgleichungen nach § 20 nur bis auf eine Transformation mit einer unitären Stufenmatrix entsprechend den Vielfachheiten der Eigenwerte der Energie festgelegt. Diejenigen Matrixelemente oder Funktionen von Matrixelementen, die bestimmten meßbaren Eigenschaften des Systems (z. B. Strahlungsintensitäten) entsprechen, müssen aber *invariant* gegenüber den willkürlichen unitären Transformationen sein. Es ergibt sich also in jedem solchen Falle die Aufgabe, die fraglichen Invarianten festzustellen. Die Auffindung solcher Invarianten geschieht besonders leicht für den Fall der *Richtungsentartung*. Es sei nämlich eine Entartung der Energie dadurch veranlaßt, daß die HAMILTON Funktion $H(q, p)$ invariant gegen Drehungen der dreidimensionalen Raumkoordinaten (und entsprechender Transformation der zugehörigen Impulse; vgl. § 18) ist. Man kann dann aus einer bestimmten Lösung $q_k = q_k^0, p_k = p_k^0$ der Bewegungsgleichungen eine neue Lösung $q_k = q'_k, p_k = p'_k$ ableiten, indem man die q'_k, p'_k als dreidimensionale orthogonale Transformation der q_k^0, p_k^0 ansetzt. Als *notwendige* Bedingung für die oben geforderte Invarianz ergibt sich also die Invarianz gegenüber räumlichen Drehungen des Koordinatensystems. Solche Invarianten sind aber leicht zu erzeugen mit Hilfe von Funktionen der q_k, p_k , die selber Komponenten dreidimensionaler Vektoren bzw. Tensoren sind.

In der klassischen Mechanik sind die q_k, p_k bei einem mehrfach periodischen System als FOURIERSche Reihen nach den Winkelvariablen w_k darstellbar, deren Koeffizienten von den Wirkungsvariablen J_k abhängen. Dasselbe gilt für irgend eine Funktion der p_k, q_k :

$$f(p, q) = \sum_{\tau} f_{\tau}(J) e^{2\pi i(w\tau)}.$$

Ist nun z. B. ein Vektor $\mathfrak{A}(p, q)$ gegeben, so ist bekanntlich das Quadrat seiner Länge \mathfrak{A}^2 eine Drehinvariante; diese ist aber Funktion der w_k, J_k (also auch der Zeit) und repräsentiert daher tatsächlich unendlich viele konstante Drehinvarianten. Als solche

kann man etwa die Quadrate der FOURIER-Koeffizienten $\mathfrak{A}_\tau^2(J)$ nehmen. Allgemeiner, hat man zwei Vektoren $\mathfrak{A}, \mathfrak{B}$ mit den FOURIER-Koeffizienten $\mathfrak{A}_\tau(J), \mathfrak{B}_\tau(J)$, so ist das skalare Produkt $\mathfrak{A}(J)\mathfrak{B}(J')$ eine Drehinvariante, also auch sämtliche FOURIER-Koeffizienten dieses Produkts, nämlich die Größen $\mathfrak{A}_\tau(J)\mathfrak{B}_{\sigma-\tau}(J')$, die von den Werten der Wirkungsvariablen in zwei Zuständen, J_k und J'_k , abhängen. Man kann solche Invarianten gewinnen, indem man aus Vektoren oder Tensoren, die Funktionen der p_k, q_k sind, und aus konstanten Hilfsvektoren Skalare bildet und diese in FOURIERSche Reihen entwickelt; die Koeffizienten sind dann Invarianten, die noch von den J_k eines oder mehrerer Zustände abhängen. Wenn man gemäß der BOHRschen Quantentheorie $J_k = \hbar n_k$ setzt, gewinnt man auf diesem Wege Invarianten, die den durch die n_k bestimmten Quantenzuständen zugeordnet sind. Dieser Gedankengang läßt sich nun vollständig auf die Quantenmechanik übertragen. Der FOURIERSchen Reihe korrespondiert dabei die Matrixdarstellung einer Größe. Es handelt sich darum, aus den Matrixelementen Drehinvarianten zu gewinnen, die den Zuständen, d. h. den Diagonalelementen, zugeordnet sind. Dabei ist noch darauf Rücksicht zu nehmen, daß im Entartungsfalle mehrere Diagonalelemente demselben Zustande entsprechen; man wird daher alle Matrizen in Teilmatrizen gemäß der Vielfachheit der Energiewerte aufspalten. Ist $f(p, q)$ eine Matrixfunktion, so seien die Teilmatrizen, wie früher, mit $f^{(rs)} = (f_{\sigma\sigma}^{(r,s)})$ bezeichnet. Wir erinnern nun an das in § 20 abgeleitete Resultat, daß die Größe

$$(1) \quad Sp f^{(r,r)} = \sum_{\sigma} f_{\sigma\sigma}^{(r,r)}$$

eine dem Zustande r zugeordnete Relativinvariante ist, d. h. unabhängig von der Willkür in der Wahl desjenigen kanonischen Bezugssystems, in dem die Energie als Diagonalmatrix erscheint. Hat man zwei Funktionen $f(p, q), \varphi(p, q)$, so gilt dasselbe für

$$(2) \quad Sp (f_{\sigma\sigma}^{(r,s)} \varphi_{\sigma\sigma}^{(s,r)}) = \sum_{\sigma} f_{\sigma\sigma}^{(r,s)} \varphi_{\sigma\sigma}^{(s,r)};$$

diese Invariante ist den beiden Zuständen r, s zugeordnet, und zwar offenbar den „Übergängen“ $r \rightarrow s$ und $s \rightarrow r$ in gleicher Weise. (Das entsprechende ist aus der klassischen Analogie nicht ohne weiteres zu ersehen.)

Sobald nun die Energiefunktion eine Gruppe von Drehungen gestattet, so entsprechen diesen nach § 26 kanonische Transformationen (unitäre Transformationen im HILBERTSchen Raume), durch die die Diagonalf orm der Energiematrix nicht geändert wird. Die Größen (1), (2) sind dann also Drehinvarianten. Im Falle freier Drehbarkeit um einen Punkt ist dabei die in (1) und (2) berücksichtigte Entartung vor allem die der Richtung; ist außer dieser keine andere Entartung vorhanden, so bedeuten ϱ, σ Werte der magnetischen Quantenzahl m .

Wir können nun das oben für die klassische Theorie erläuterte Verfahren zur Bildung von Invarianten ohne weiteres auf die Quantenmechanik übertragen. Aus den Matrixvektoren und Tensoren konstruieren wir mit Hilfe von Zahlenvektoren Skalare und erzeugen aus diesen nach den Regeln (1) und (2) die Invarianten der Zustände und Übergänge.

Ist \mathfrak{A} ein Matrixvektor und \mathfrak{S} ein Zahlenvektor, so ist

$$(3) \quad f = \mathfrak{A} \mathfrak{S}$$

die einzige Funktion, die in den Komponenten von \mathfrak{A} linear ist. Man hat die Invariante

$$(4) \quad Sp \mathfrak{A}^{(r,r)} \mathfrak{S} = \sum_x Sp \mathfrak{A}_x^{(r,r)} \cdot S_x.$$

Ist nun das System um eine Achse frei drehbar, so mache man diese etwa zur z -Achse; dann lege man \mathfrak{S} parallel zur x - oder y -Achse und drehe das System um die z -Achse durch 180° . Dabei muß die Größe (4) ungeändert bleiben, und das ist nur möglich, wenn

$$(5) \quad Sp \mathfrak{A}_x^{(r,r)} = 0, \quad Sp \mathfrak{A}_y^{(r,r)} = 0$$

ist. Besteht überdies noch Drehbarkeit um eine weitere Achse (also um einen festen Punkt), so gilt überdies:

$$(6) \quad Sp \mathfrak{A}_z^{(r,r)} = 0.$$

Das Quadrat der Länge des Vektors \mathfrak{A} ist jedenfalls drehinvariant, also auch

$$(7) \quad Sp (\mathfrak{A}^2)^{(r,r)}.$$

Wir betrachten nun die Simultaninvarianten zweier Matrixvektoren $\mathfrak{A}, \mathfrak{B}$. Mit Hilfe zweier Zahlenvektoren $\mathfrak{S}, \mathfrak{T}$ setzen wir

$$(8) \quad f = \mathfrak{A} \mathfrak{S}, \quad \varphi = \mathfrak{B} \mathfrak{T};$$

dann ist

$$(9) \quad Sp \{ (\mathfrak{A}^{(r,s)} \mathfrak{C}) \cdot (\mathfrak{B}^{(s,r)} \mathfrak{I}) \} = \sum_{xy} Sp (A_x^{(r,s)} B_y^{(s,r)}) S_x T_y$$

eine Drehinvariante, wenn \mathfrak{C} und \mathfrak{I} einen festen Winkel bilden.

Hat man nun Drehbarkeit um die z -Achse, so wähle man zunächst \mathfrak{C} parallel zur x -Achse, \mathfrak{I} parallel zur z -Achse; dann sind alle Glieder der Summe (9) gleich Null außer dem mit $S_x T_z = 1$. Dreht man nun das Vektorenpaar \mathfrak{C} , \mathfrak{I} so, daß \mathfrak{I} parallel zur z -Achse bleibt, um den Winkel π , so wird $S_x T_z = -1$, und da (9) dabei invariant sein soll, muß der Koeffizient von $S_x T_z$ verschwinden. Genau dasselbe gilt für die Koeffizienten von $S_z T_x$, $S_y T_z$, $S_z T_y$. Nun wählen wir \mathfrak{C} parallel der x -Achse, \mathfrak{I} parallel der y -Achse und drehen das Vektorpaar um $\frac{\pi}{2}$; dann geht $S_x T_y$ vom Werte 1 in den Wert 0, $S_y T_x$ vom Werte 0 in den Wert -1 über. Folglich müssen die Koeffizienten von $S_x T_y$ und $S_y T_x$ entgegengesetzt gleich sein, und dasselbe gilt für $T_x S_y$ und $T_y S_x$. Wählt man endlich $\mathfrak{C} = \mathfrak{I}$ und legt diesen Vektor parallel zur x - oder y -Achse, so sieht man durch Drehung um die z -Achse, daß die Koeffizienten von S_x^2 und S_y^2 einander gleich sein müssen. Man erhält also bei *Drehbarkeit um die z -Achse*:

$$(10) \quad \begin{cases} Sp (A_x^{(r,s)} B_x^{(s,r)}) = Sp (A_y^{(r,s)} B_y^{(s,r)}), \\ Sp (A_x^{(r,s)} B_z^{(s,r)}) = Sp (A_y^{(r,s)} B_z^{(s,r)}) = 0, \\ Sp (A_z^{(r,s)} B_x^{(s,r)}) = Sp (A_z^{(r,s)} B_y^{(s,r)}) = 0, \\ Sp (A_x^{(r,s)} B_y^{(s,r)}) = - Sp (A_y^{(r,s)} B_x^{(s,r)}). \end{cases}$$

Mit dieser Drehbarkeit sind vereinbar erstens Drehbarkeit um die x -Achse, also freie Drehbarkeit, zweitens Spiegelung an einer Ebene durch die z -Achse, drittens Spiegelung an einer Ebene senkrecht zur z -Achse. Man findet, daß dann außer den Relationen (10) noch die folgenden bestehen:

Freie Drehbarkeit:

$$(11) \quad \begin{cases} \text{a) } Sp (A_x^{(r,s)} B_x^{(s,r)}) = Sp (A_y^{(r,s)} B_y^{(s,r)}) \\ \quad \quad \quad = Sp (A_z^{(r,s)} B_z^{(s,r)}) = \frac{1}{3} Sp (\mathfrak{A}^{(r,s)} \mathfrak{B}^{(s,r)}), \\ \text{b) } Sp (A_x^{(r,s)} B_y^{(s,r)}) = Sp (A_y^{(r,s)} B_x^{(s,r)}) = 0. \end{cases}$$

Eine Spiegelebene durch die z -Achse liefert nur (11 b); eine Spiegelebene senkrecht zur z -Achse keine anderen Relationen außer (10). Wird ein System, das im freien Zustand irgend eine Spiegelebene hat, in ein elektrisches Feld gebracht, so entspricht das dem Fall (11 b).

Wir ziehen eine Folgerung über die Realitätsverhältnisse der Größe (11a) für hermitesche Matrizen bei freier Drehbarkeit des Systems. Offenbar ist

$$(12) \quad Sp(\mathfrak{A} \mathfrak{B})^{(r,r)}$$

eine Drehinvariante; für $\mathfrak{A} = \mathfrak{B}$ hat man die Invariante (7). Nun ist

$$(13) \quad \begin{aligned} Sp(\mathfrak{A} \mathfrak{B})^{(r,r)} &= Sp\{(A_x B_x)^{(r,r)} + (A_y B_y)^{(r,r)} + (A_z B_z)^{(r,r)}\} \\ &= Sp\left\{\sum_s A_x^{(r,s)} B_x^{(s,r)} + \sum_s A_y^{(r,s)} B_y^{(s,r)} + \sum_s A_z^{(r,s)} B_z^{(s,r)}\right\} \\ &= \sum_s Sp(\mathfrak{A}^{(r,s)} \mathfrak{B}^{(s,r)}). \end{aligned}$$

Da nun für irgend eine Matrix f

$$Sp f^\dagger = (Sp f)^*$$

ist, so gilt

$$Sp((\mathfrak{A} \mathfrak{B})^\dagger)^{(r,r)} = Sp(\mathfrak{B}^\dagger \mathfrak{A}^\dagger)^{(r,r)} = [Sp(\mathfrak{A} \mathfrak{B})^{(r,r)}]^*,$$

und wenn $\mathfrak{A}, \mathfrak{B}$ hermitisch sind:

$$(14) \quad Sp(\mathfrak{B} \mathfrak{A})^{(r,r)} = \{Sp(\mathfrak{A} \mathfrak{B})^{(r,r)}\}^*.$$

Sind endlich noch $\mathfrak{A}, \mathfrak{B}$ vertauschbar, so folgt der Satz:

Für vertauschbare, hermitesche Matrizen $\mathfrak{A}, \mathfrak{B}$ ist die Invariante

$$(15) \quad Sp(\mathfrak{A} \mathfrak{B})^{(r,r)} = \sum_s Sp(\mathfrak{A}^{(r,s)} \mathfrak{B}^{(s,r)})$$

reell.

Wir gehen nun zu höheren Tensoren über. Für die meisten Zwecke genügt die Betrachtung von symmetrischen und antisymmetrischen Tensoren zweiter Ordnung (A_{xy}), deren Komponenten $A_{xy} = \pm A_{yx}$ hermitesche Matrizen sind, die sich wie die Quadrate und Produkte der Koordinaten transformieren. Wir beschränken uns auf den Fall *freier Drehbarkeit*.

Sind $\mathfrak{S}, \mathfrak{T}$ zwei Zahlenvektoren, die einen festen Winkel miteinander bilden, so ist der Skalar

$$(16) \quad Sp\left(\sum_{xy} A_{xy}^{(r,r)} S_x T_y\right) = \sum_{xy} Sp(A_{xy}^{(r,r)}) S_x T_y$$

eine Drehinvariante. Daraus folgen, wie oben, die *drehinvarianten Relationen*:

$$(17) \quad \begin{cases} Sp(A_{xx}^{(r,r)}) = Sp(A_{yy}^{(r,r)}) = Sp(A_{zz}^{(r,r)}), \\ Sp(A_{yz}^{(r,r)}) = Sp(A_{zx}^{(r,r)}) = Sp(A_{xy}^{(r,r)}) = 0. \end{cases}$$

Um die Simultaninvarianten eines Vektors \mathfrak{A} und eines Tensors (B_{xy}) zu bilden, betrachten wir die Invariante

$$(18) \quad Sp(\sum_{xyz} A_x^{(r,s)} B_{yz}^{(s,r)} \mathfrak{S}_x \mathfrak{I}_y \mathfrak{R}_z),$$

wo $\mathfrak{S}, \mathfrak{I}, \mathfrak{R}$ ein starres Tripel von Zahlenvektoren ist. Läßt man alle drei oder je zwei von diesen zusammenfallen, so sieht man, daß die entsprechenden Tensorkomponenten mit drei oder zwei gleichen Indizes verschwinden müssen; nur die mit drei verschiedenen Indizes sind nicht Null und einander gleich bzw. entgegengesetzt gleich:

$$(19) \quad \begin{cases} Sp(A_x^{(r,s)} B_{xx}^{(s,r)}) = Sp(A_y^{(r,s)} B_{yy}^{(s,r)}) = Sp(A_z^{(r,s)} B_{zz}^{(s,r)}) = 0, \\ Sp(A_x^{(r,s)} B_{yy}^{(s,r)}) = Sp(A_y^{(r,s)} B_{zz}^{(s,r)}) = Sp(A_z^{(r,s)} B_{xx}^{(s,r)}) = 0, \\ Sp(A_x^{(r,s)} B_{zz}^{(s,r)}) = Sp(A_y^{(r,s)} B_{xx}^{(s,r)}) = Sp(A_z^{(r,s)} B_{yy}^{(s,r)}) = 0; \end{cases}$$

dagegen

$$(20) \quad \begin{cases} Sp(A_x^{(r,s)} B_{yz}^{(s,r)}) = -Sp(A_x^{(r,s)} B_{zy}^{(s,r)}) \\ = Sp(A_y^{(r,s)} B_{zx}^{(s,r)}) = -Sp(A_y^{(r,s)} B_{xz}^{(s,r)}) \\ = Sp(A_z^{(r,s)} B_{yx}^{(s,r)}) = -Sp(A_z^{(r,s)} B_{xy}^{(s,r)}). \end{cases}$$

Ist der Tensor B symmetrisch, so verschwinden auch die Größen (20); im allgemeinen sind sie aber von Null verschieden. Führt man den „axialen Vektor“ \mathfrak{M} mit den Komponenten

$$(21) \quad M_x = \frac{1}{2}(B_{yz} - B_{zy}), \quad M_y = \frac{1}{2}(B_{zx} - B_{xz}), \quad M_z = \frac{1}{2}(B_{xy} - B_{yx})$$

ein, so gehen die Beziehungen (20) über in

$$(22) \quad Sp(A_x^{(r,s)} M_x^{(s,r)}) = Sp(A_y^{(r,s)} M_y^{(s,r)}) = Sp(A_z^{(r,s)} M_z^{(s,r)}) \\ = \frac{1}{3} Sp(\mathfrak{A}^{(r,s)} \mathfrak{M}^{(s,r)}),$$

und man zeigt, wie oben, daß die Invariante

$$(23) \quad Sp(\mathfrak{A} \mathfrak{M})^{(r,s)} = \sum_s Sp(\mathfrak{A}^{(r,s)} \mathfrak{M}^{(s,r)})$$

reell ist, sobald $\mathfrak{A}, \mathfrak{M}$ hermitisch und vertauschbar sind.

Besitzt das System außer der freien Drehbarkeit noch eine *Spiegelebene*, so verschwinden die Größen (20) und (22).

Zum Schluß stellen wir noch die Simultaninvarianten zweier Tensoren 2. Stufe (A_{xy}) , (B_{xy}) auf. Man bilde mit Hilfe von vier starr verbundenen Zahlenvektoren \mathfrak{S} , \mathfrak{X} , \mathfrak{S}' , \mathfrak{X}' die Drehinvariante

$$(24) \quad \begin{aligned} Sp \{ (\sum_{xy} A_{xy}^{(r,s)} S_x T_y) \cdot (\sum_{xy} B_{xy}^{(s,r)} S_x' T_y') \} \\ = \sum_{xyx'y'} Sp (A_{xy}^{(r,s)} B_{x'y'}^{(s,r)}) S_x T_y S_{x'} T_{y'}, \end{aligned}$$

wobei die vierfache Summe so zu verstehen ist, daß x, y, x', y' unabhängig voneinander die Koordinatenrichtungen durchlaufen.

Läßt man nun zwei der vier Vektoren \mathfrak{S} , \mathfrak{X} , \mathfrak{S}' , \mathfrak{X}' zusammenfallen und wählt die beiden andern senkrecht auf diesem Vektor und aufeinander, so folgen aus der Invarianz gegen Drehungen um die Koordinatenachsen Relationen vom Typus:

$$(25) \quad \begin{cases} Sp (A_{xx}^{(r,s)} B_{yz}^{(s,r)}) = Sp (A_{xx}^{(r,s)} B_{zy}^{(s,r)}) = \dots = 0, \\ Sp (A_{yz}^{(r,s)} B_{xx}^{(s,r)}) = Sp (A_{zy}^{(r,s)} B_{xx}^{(s,r)}) = \dots = 0, \\ Sp (A_{xz}^{(r,s)} B_{yz}^{(s,r)}) = Sp (A_{zx}^{(r,s)} B_{yz}^{(s,r)}) = \dots = 0, \\ Sp (A_{zx}^{(r,s)} B_{zy}^{(s,r)}) = Sp (A_{xz}^{(r,s)} B_{zy}^{(s,r)}) = \dots = 0, \end{cases}$$

wo die Punkte diejenigen Größen bedeuten, die man durch zyklische Vertauschung der Achsen bekommt.

Läßt man ferner je zwei Paare der vier Vektoren zusammenfallen, entweder $\mathfrak{S} = \mathfrak{X}$, $\mathfrak{S}' = \mathfrak{X}'$, oder $\mathfrak{S} = \mathfrak{S}'$, $\mathfrak{X} = \mathfrak{X}'$, und wählt diese beiden Vektoren parallel zu zwei verschiedenen Achsen, so erhält man

$$(26) \quad \begin{cases} Sp (A_{yy}^{(r,s)} B_{zz}^{(s,r)}) = Sp (A_{zz}^{(r,s)} B_{xx}^{(s,r)}) = Sp (A_{xx}^{(r,s)} B_{yy}^{(s,r)}) \\ = Sp (A_{zz}^{(r,s)} B_{yy}^{(s,r)}) = Sp (A_{xx}^{(r,s)} B_{zz}^{(s,r)}) = Sp (A_{yy}^{(r,s)} B_{xx}^{(s,r)}) \end{cases}$$

und

$$(27) \quad \begin{cases} Sp (A_{yz}^{(r,s)} B_{yz}^{(s,r)}) = Sp (A_{zx}^{(r,s)} B_{zx}^{(s,r)}) = Sp (A_{xy}^{(r,s)} B_{xy}^{(s,r)}), \\ Sp (A_{yz}^{(r,s)} B_{zy}^{(s,r)}) = Sp (A_{zx}^{(r,s)} B_{xz}^{(s,r)}) = Sp (A_{xy}^{(r,s)} B_{yx}^{(s,r)}). \end{cases}$$

Läßt man drei der vier Vektoren zusammenfallen und wählt den vierten senkrecht dazu, so bekommt man

$$(28) \quad Sp (A_{xx}^{(r,s)} B_{xy}^{(s,r)}) = Sp (A_{xx}^{(r,s)} B_{yx}^{(s,r)}) = \dots = 0.$$

Läßt man endlich alle vier Vektoren \mathfrak{S} , \mathfrak{L} , \mathfrak{S}' , \mathfrak{L}' zusammenfallen, so folgt

$$(29) \quad Sp(A_{xx}^{(r,s)} B_{xx}^{(s,r)}) = Sp(A_{yy}^{(r,s)} B_{yy}^{(s,r)}) = Sp(A_{zz}^{(r,s)} B_{zz}^{(s,r)}).$$

Alle physikalischen Eigenschaften frei oder um eine Achse drehbarer Systeme können nur von diesen Invarianten abhängen.

§ 31. Intensitäten beim ZEEMAN-Effekt. Wir kehren nun zu den Formeln am Ende von § 29 für die Amplituden des elektrischen Moments unter der Wirkung eines äußeren Feldes zurück und berechnen die Ausstrahlung; und zwar wollen wir diese im Verhältnis zur Strahlung des Systems ohne Feld bestimmen. Letztere ist nach § 20, (19) proportional mit

$$(1) \quad \|\mathfrak{P}^{(r,s)}\|^2 = Sp(\mathfrak{P}^{(r,s)} \mathfrak{P}^{(s,r)}),$$

wo r, s je eine Kombination der Werte von n, j bedeuten. Indem wir den Index n , wie bisher, weglassen, schreiben wir einfach j, j' statt r, s . Nach der Formel § 30, (11) gilt

$$(2) \quad \begin{cases} \|X^{(j,j')}\|^2 = \|Y^{(j,j')}\|^2 = \|Z^{(j,j')}\|^2 = \frac{1}{3} \|\mathfrak{P}^{(j,j')}\|^2, \\ Sp(Y^{(j,j')} Z^{(j',j)}) = Sp(Z^{(j,j')} X^{(j',j)}) = Sp(X^{(j,j')} Y^{(j',j)}) = 0. \end{cases}$$

Es ist wohl nicht überflüssig, diese Relationen durch direkte Ausrechnung auf Grund der Formeln § 29, (12), (13) zu bestätigen. Dazu wird man sie auf die entsprechenden Relationen zwischen Π, Π^\dagger und Z zurückführen. Man hat

$$(3) \quad X = \frac{1}{2}(\Pi + \Pi^\dagger), \quad Y = \frac{1}{2i}(\Pi - \Pi^\dagger).$$

Da nun in Z nach § 29 (12), (13) nur die Diagonale, in Π nur eine Nebendiagonale bezüglich m besetzt ist, so sind alle Diagonalelemente von

$$\begin{aligned} &\Pi^{(j,j')} \Pi^{(j',j)}, \quad \Pi^\dagger{}^{(j,j')} \Pi^\dagger{}^{(j',j)}, \quad \Pi^{(j,j')} Z^{(j',j)}, \\ &\Pi^\dagger{}^{(j,j')} Z^{(j',j)}, \quad Z^{(j,j')} \Pi^{(j',j)}, \quad Z^{(j,j')} \Pi^\dagger{}^{(j',j)} \end{aligned}$$

gleich Null, also auch die Spuren. Ferner wird

$$(4) \quad \begin{aligned} &Sp(X^{(j,j')} Y^{(j',j)}) \\ &= \frac{1}{4i} Sp\{\Pi^\dagger{}^{(j,j')} \Pi^{(j',j)} - \Pi^{(j,j')} \Pi^\dagger{}^{(j',j)}\} \end{aligned}$$

und das verschwindet nach § 5, (3); nach demselben Satze hat man

$$(5) \quad Sp(X^{(j,j')} X^{(j',j)}) = Sp(Y^{(j,j')} Y^{(j',j)}) = \frac{1}{2} Sp(\Pi^{(j,j')} \Pi^\dagger{}^{(j',j)}).$$

Aus § 29, (12), (13) erhält man schließlich auf Grund der Formel

$$(6) \quad \sum_{m=-j}^j m^2 = \frac{1}{3} (2j+1) j(j+1)$$

folgende Werte:

$$(7) \quad \left\{ \begin{aligned} Sp(II^{(i,j)} II^{\dagger(i,j)}) &= 2 \cdot \frac{2j+1}{3} j(j+1) \cdot |a_{nn'}^{(i,j)}|^2, \\ Sp(II^{(i,j-1)} II^{\dagger(j-1,i)}) &= 2 \cdot \frac{2j+1}{3} j(2j-1) \cdot |b_{nn'}^{(i,j-1)}|^2, \\ Sp(II^{(i,j+1)} II^{\dagger(j+1,i)}) &= 2 \cdot \frac{2j+1}{3} (j+1)(2j+3) \cdot |b_{nn'}^{(i,j+1)}|^2; \end{aligned} \right.$$

$$(8) \quad \left\{ \begin{aligned} Sp(Z^{(i,j)} Z^{(i,j)}) &= \frac{2j+1}{3} j(j+1) \cdot |a_{nn'}^{(i,j)}|^2, \\ Sp(Z^{(i,j-1)} Z^{(j-1,i)}) &= \frac{2j+1}{3} j(2j-1) \cdot |b_{nn'}^{(i,j-1)}|^2, \\ Sp(Z^{(i,j+1)} Z^{(j+1,i)}) &= \frac{2j+1}{3} (j+1)(2j+3) \cdot |b_{nn'}^{(i,j+1)}|^2; \end{aligned} \right.$$

$$(9) \quad \left\{ \begin{aligned} Sp(II^{(i,j')} II^{(j',i)}) &= Sp(II^{\dagger(i,j')} II^{\dagger(j',i)}) = 0, \\ Sp(II^{(i,j')} Z^{(j',i)}) &= Sp(II^{\dagger(i,j')} Z^{(j',i)}) = 0. \end{aligned} \right.$$

Aus diesen Formeln zusammen mit (5) geht die Richtigkeit der Behauptung (2) hervor, und zwar wird

$$(10) \quad \|\mathfrak{P}^{(i,j')}\|^2 = \frac{3}{2} Sp(II^{(i,j')} II^{\dagger(j',i)}),$$

also

$$(11) \quad \left\{ \begin{aligned} \|\mathfrak{P}^{(i,j)}\|^2 &= (2j+1) \cdot j(j+1) \cdot |a_{nn'}^{(i,j)}|^2, \\ \|\mathfrak{P}^{(i,j-1)}\|^2 &= (2j+1) \cdot j(2j-1) \cdot |b_{nn'}^{(i,j-1)}|^2, \\ \|\mathfrak{P}^{(i,j+1)}\|^2 &= (2j+1) \cdot (j+1)(2j+3) \cdot |b_{nn'}^{(i,j+1)}|^2. \end{aligned} \right.$$

Die Lichtintensität der Emission im feldfreien Zustande ist

$$(12) \quad J^{(i,j)} = \frac{64\pi^4}{3c^3} \nu^{(i,j)4} \|\mathfrak{P}^{(i,j')}\|^2.$$

Wir wollen nun die Lichtintensitäten der durch das Feld aufgespaltenen Linien (der ZEEMAN-Komponenten) berechnen und durch die Intensitäten ohne Feld auszudrücken suchen.

Betrachtet man etwa den transversalen ZEEMAN-Effekt, so sind die Intensitäten der parallel und senkrecht zum Felde schwingenden Komponenten

diese *Summenregeln* besagen, daß die Summe der Strahlungsenergien, die in den drei Linien enthalten sind, deren Anfangs- oder Endterme übereinstimmen, unabhängig von m , also für alle ZEEMAN-Terme des unteren bzw. des oberen Zustandes dieselbe ist. (Der Faktor 2 bei den \perp -Komponenten rührt davon her, daß, wie schon oben erwähnt, die beiden senkrecht zum Felde schwingenden Komponenten des Moments je den Betrag J_{\perp} ausstrahlen). Ferner hat man:

$$(19) \quad \sum_m J_{\parallel}^{(j, j')}(m, m) = 2 \sum_m J_{\perp}^{(j, j')}(m, m+1) \\ = 2 \sum_m J_{\perp}^{(j, j')}(m, m-1) = \frac{1}{3} J^{(j, j')}, \\ (j' = j-1, j, j+1).$$

Die Summe der Energien der \parallel -Komponenten und der \perp -Komponenten sind also gleich stark, oder: durch die Aufspaltung der Linie wird diese in Summa nicht polarisiert.

Diese Regeln sind zuerst auf Grund empirischen Materials von ORNSTEIN und BURGER aufgestellt worden¹. Sie lassen eine einfache korrespondenzmäßige Deutung zu. In dem der klassischen Theorie entsprechenden Grenzfall (j groß) führt das Atom im Magnetfeld eine Präzession des Drehimpulsvektors um die Feldrichtung aus. Dabei wird jede schon vorhandene harmonische Komponente in drei Schwingungen aufgespalten, eine lineare parallel dem Felde und zwei entgegengesetzt zirkulare um die Feldrichtung. Die Summenregeln besagen dann, daß je drei solche Schwingungen zusammen dieselbe Intensität haben müssen wie die ursprüngliche, unzerlegte Schwingung; das ist ein einfacher Fall des bereits besprochenen HEISENBERGSchen Satzes von der „spektroskopischen Stabilität“ (s. § 20)². Auch daß die aufgespaltene Linie im ganzen unpolarisiert sein muß, gilt schon für das klassische Modell. Dieses liefert ferner bestimmte Ausdrücke für die Amplituden bzw. Intensitäten der Schwingungen im Felde als Funktion des Winkels α , den die Achse des Drehimpulses mit der Feldrichtung bildet. Setzt man nach den Regeln der BOHRschen Quantentheorie

$$\cos \alpha = \frac{m}{j}$$

¹ ORNSTEIN, L. S. u. H. C. BURGER: Z. Phys. Bd. 29, S. 241. 1924.

² HEISENBERG, W.: Z. Phys. Bd. 31, S. 617. 1925.

(Richtungsquantelung, s. Bd. 1, § 17, § 34), so erhält man Formeln für die Intensitäten, die mit (10), (11), (12) korrespondenzmäßig (für große j) übereinstimmen. Umgekehrt sind die halbklassischen Ausdrücke benützt worden, um die wirklichen Intensitäten abzuschätzen¹. Übernimmt man von den halbklassischen Formeln den Umstand, daß die Intensitäten quadratische Funktionen von m sind und postuliert die Gültigkeit der Summen- und Polarisationsregeln, so ergeben sich zwangsläufig die exakten Formeln (10), (11), (12); auf diesem Wege sind sie, wie in § 2 erwähnt wurde, bereits vor der Entwicklung der systematischen Quantenmechanik gewonnen worden². Dabei haben die von ORNSTEIN und seiner Schule³ gewonnenen empirischen Ergebnisse als Wegweiser eine wesentliche Rolle gespielt und zur Bestätigung der theoretischen Formeln gedient.

§ 32. Ganzzahligkeit der inneren Quantenzahl bei Systemen von Massenpunkten. Wir haben für die in §§ 27 bis 29 bewiesenen Sätze und Formeln neben dem allgemeinen Satz § 26 ausschließlich *die Vertauschungsregeln* (I), (II) [s. § 25, (10), (10a, b), (15), (16)] benutzt und sind dabei zu dem Ergebnis gekommen, daß j und m ganz- oder halbzahlig sein müssen. Wir haben dagegen *nicht* unmittelbar Gebrauch gemacht von der *Definition des Drehimpulses* \mathfrak{M} nach § 25, (1). Außerdem haben wir nicht untersucht, wie die Matrizen a , b in § 29, (12), (13) beschaffen sein müssen, damit X, Y, Z oder Π, Π^\dagger, Z vertauschbar werden. Eine Ergänzung unserer Betrachtungen in diesen beiden Punkten führt zu einer grundsätzlich wichtigen Erkenntnis.

Wir betrachten zunächst einen einzigen Massenpunkt, der sich in einem Zentralfeld bewegen möge, so daß $\mathfrak{M} = 0$ ist. Es gilt

¹ SOMMERFELD, A. u. W. HEISENBERG: Z. Phys. Bd. 11, S. 191. 1922.

² KRONIG, R. DE L.: Z. Phys. Bd. 31, S. 885. 1925. GOUDSMIT, S. u. R. KRONIG: Naturwissensch. Bd. 13, S. 90. 1925. HÖNL, H.: Z. Phys. Bd. 31, S. 340. 1925.

³ Eine Übersicht über die empirischen Arbeiten findet man bei DORGELO, H. B.: Phys. Z. Bd. 26, S. 756. 1925; Bd. 27, S. 182. 1926. Bezüglich des Standes der Theorie vor der Quantenmechanik vgl. man GEIGER u. SCHEEL: Handbuch d. Physik Bd. 23, Artikel von W. PAULI jr., Quantentheorie; Ziff. 12, 41. Alle dort gemachten Angaben sind auch heute noch unverändert richtig.

Satz 1: Für einen Massenpunkt im Zentralfeld verschärft sich die Auswahlregel für die innere Quantenzahl j zu $\Delta j = \pm 1$.

Für einen Massenpunkt wird nämlich, wenn wir mit \mathbf{r} den Ortsvektor x, y, z bezeichnen, auf Grund der Definition § 25, (1)

$$(1) \quad \mathfrak{M} \cdot \mathbf{r} = \mathbf{r} \cdot \mathfrak{M} = x M_x + y M_y + z M_z = 0.$$

In der klassischen Theorie bedeutet das — zusammen mit dem Erhaltungssatz des Drehimpulses \mathfrak{M} —, daß der Massenpunkt sich in einer festen Ebene bewegt, die senkrecht zu \mathfrak{M} durch den Koordinaten-Nullpunkt geht. In der Quantenmechanik muß man auf eine solche anschauliche Deutung der Gleichung (1) verzichten; doch hat sie, wie wir sogleich zeigen werden, die in Satz 1 ausgesprochene Verschärfung der Auswahlregel für j zu $\Delta j = \pm 1$, also den Ausschluß von $\Delta j = 0$ zur Folge, und dies kann korrespondenzmäßig so erläutert werden, daß keine Schwingungen des Teilchens parallel \mathfrak{M} auftreten.

Zum Beweise unseres Satzes bilden wir durch Multiplikation von \mathbf{r} mit der Ladung e das elektrische Moment \mathfrak{P} mit den Komponenten X, Y, Z bzw. $II = X + iY, Z$; dann können wir statt (1) schreiben

$$(2) \quad II M^\dagger + II^\dagger M + 2 Z M_z = 0.$$

Es wird also insbesondere

$$\begin{aligned} & II^{(j, j)}(m, m-1) M^{\dagger(j, j)}(m-1, m) \\ & + II^{\dagger(j, j)}(m, m+1) M^{(j, j)}(m+1, m) \\ & + 2 Z^{(j, j)}(m, m) \cdot \frac{h}{2\pi} m = 0 \end{aligned}$$

und nach § 28, (5) und § 29, (12)

$$\begin{aligned} & \{(j+m)(j-m+1) + (j+m+1)(j-m) + 2m^2\} a^{(j, j)} = 0 \\ \text{oder} & \quad 2j(j+1) a^{(j, j)} = 0. \end{aligned}$$

Daraus folgt für $j \neq 0$

$$(3) \quad a^{(j, j)} = 0,$$

d. h. die $\Delta j = 0$ entsprechenden Matrixelemente § 29, (12) verschwinden sämtlich. Ist aber $j = 0$, so gehört dazu als einziger m -Wert $m = 0$; dann sind aber auch die Matrixelemente § 29, (12) gleich Null.

In jedem Falle kommen also nur die $\Delta j = \pm 1$ entsprechenden Matrixelemente § 29, (13) vor.

Daraus ergibt sich der weitere

Satz 2: Für einen einzelnen Massenpunkt ist die innere Quantenzahl j notwendig ganzzahlig, und ihr kleinster vorkommender Wert ist $j_{\min} = 0^1$.

Wir bilden nämlich die Matrizen $II II^\dagger$ und $II^\dagger II$, in denen wegen Satz 1 nach den Formeln § 29, (13) nur folgende Teilmatrizen von Null verschieden sind:

$$(4) \quad \begin{aligned} II^{(j, j-1)}(m, m-1) II^{\dagger(j-1, j)}(m-1, m) \\ + II^{(j, j+1)}(m, m-1) II^{\dagger(j+1, j)}(m-1, m) \\ = (j+m)(j+m-1) b^{(j, j-1)} b^{\dagger(j-1, j)} \\ + (j-m+1)(j-m+2) b^{\dagger(j, j+1)} b^{(j+1, j)} \end{aligned}$$

bzw.

$$(5) \quad \begin{aligned} II^{\dagger(j, j-1)}(m, m+1) II^{(j-1, j)}(m+1, m) \\ + II^{\dagger(j, j+1)}(m, m+1) II^{(j+1, j)}(m+1, m) \\ = (j-m)(j-m-1) b^{(j, j-1)} b^{\dagger(j-1, j)} \\ + (j+m+1)(j+m+2) b^{\dagger(j, j+1)} b^{(j+1, j)}. \end{aligned}$$

Die Forderung

$$(6) \quad [X, Y] = \frac{i}{2} [II, II^\dagger] = 0$$

liefert also

$$(7) \quad \frac{\kappa}{2} [II, II^\dagger]^{(j, j)}(m, m) = m \{ (2j-1) b^{(j, j-1)} b^{\dagger(j-1, j)} - (2j+3) b^{\dagger(j, j+1)} b^{(j+1, j)} \} = 0.$$

Diese Gleichungen (4) bis (7) gelten unter der Annahme $j > j_{\min}$; ist aber $j = j_{\min}$, so sind überall die Glieder mit dem Index $j-1$ zu streichen; insbesondere gilt dann statt (7):

$$(7') \quad m(2j+3) b^{\dagger(j, j+1)} b^{(j+1, j)} = 0, \quad j = j_{\min}.$$

Jetzt führt die Annahme $j_{\min} > 0$ auf einen Widerspruch. Denn für $j > 0$ nimmt der Faktor m in (7) bzw. (7') von Null verschiedene

¹ Die zum Beweis gebrauchten Formeln stehen im wesentlichen bei MENSING, L.: Z. Phys. Bd. 36, S. 814. 1926.

Werte an, so daß der andere Faktor in (7) bzw (7') verschwinden muß. Es würde also für $j_{\min} > 0$ aus (7') die Gleichung

$$b^{\dagger(j, j+1)} b^{(j+1, j)} = 0, \quad j = j_{\min}$$

folgen, die mit

$$b^{\dagger(j, j+1)} = b^{(j+1, j)} = 0, \quad j = j_{\min}$$

gleichwertig ist. Aus der rekursiven Gleichung (7) würde sich dann dasselbe für alle j ergeben, d. h. es wäre $X = Y = Z = 0$.

Also haben wir notwendig

$$(8) \quad j_{\min} = 0.$$

Nach den Schlußbemerkungen von §§ 26, 27 müssen aber, wenn *ein* ganzzahliger Wert j vorhanden ist, *alle* Werte j ganzzahlig sein, so daß Satz 2 bewiesen ist.

Die Vertauschbarkeit von X und Y (oder von x und y) ist dann nach (7) ausgedrückt in

$$(7'') \quad (2j - 1) b^{(j, j-1)} b^{\dagger(j-1, j)} = (2j + 3) b^{\dagger(j, j+1)} b^{(j+1, j)}$$

für alle $j > 0$.

Aus Satz 2 aber ergibt sich sofort:

Satz 3: *Auch für Systeme von mehreren Massenpunkten kann die innere Quantenzahl nur ganzzahlige Werte annehmen, wobei wieder $j_{\min} = 0$ wird.*

Das sieht man so ein. Nimmt man zunächst an, daß keine Wechselwirkungen zwischen den Teilchen bestehen, so ist wegen der Additivität der Drehimpulskomponenten [§ 25, (8)] die magnetische Quantenzahl m des Gesamtsystems offenbar die Summe der Quantenzahlen m_k für die einzelnen Teilchen, also gleichfalls ganzzahlig, womit auch die innere Quantenzahl j des Gesamtsystems ganzzahlig wird. Ferner sieht man leicht, daß z. B. für einen Zustand, bei dem die inneren Quantenzahlen j_k aller einzelnen Massenpunkte Null sind, auch die innere Quantenzahl j des Gesamtsystems gleich Null wird, womit $j_{\min} = 0$ bewiesen ist.

Dies muß dann aber allgemein auch bei Wechselwirkung zwischen den Teilchen gelten. Denn die Eigenwerte von \mathfrak{M}^2 oder \mathfrak{M}_z müssen offenbar unabhängig von der Gestalt der Energiefunktion des Systems sein; die Lösung q_k, p_k der Bewegungsgleichungen bei Wechselwirkung geht ja aus der Lösung q_k^0, p_k^0

der Bewegungsgleichungen ohne Wechselwirkung durch eine unitäre Transformation hervor. (Vgl. § 23.)

Wir heben noch ausdrücklich hervor:

Satz 4: *Die bei einem beliebigen System von Massenpunkten auftretenden j -Werte müssen eine lückenlose arithmetische Reihe $j = 0, 1, 2, \dots$ bilden.*

Denn die Gleichungen II [§ 25, (15), (16)] für die Vertauschung von M_x, M_y, M_z mit dem Polarisationsvektor X, Y, Z gelten ebenso für die Vertauschung von M_x, M_y, M_z mit jedem einzelnen Koordinatentripel x_k, y_k, z_k oder Impulstripel p_{kx}, p_{ky}, p_{kz} [s. § 25, (2), (3), (4) bzw. (3a), (4a)]. Also gelten die in § 28, 29 für X, Y, Z abgeleiteten Formeln, insbesondere die Auswahlregel $\Delta j = 0, \pm 1$, auch für die einzelnen r_k und p_k ($= p_{kx}, p_{ky}, p_{kz}$). Folglich könnte das Matrizenystem r_k, p_k nicht *einfach* sein, wenn in der Reihe der j -Werte eine Lücke eintrete. (Vgl. wieder die ähnliche Überlegung bezüglich der Lückenlosigkeit der Reihe der Energiewerte beim harmonischen Oszillator.)

Aus dem obigen Ergebnis, daß für Systeme von Massenpunkten m und j ganzzahlig sind, geht hervor, daß das *Elektron nicht als eine einfache elektrische Punktladung* behandelt werden kann. Denn die Erfahrung lehrt, daß auch halbzahlige m und j auftreten, und zwar liegt es im einzelnen so: Jeder optische Term eines freien Atoms spaltet im Magnetfeld, wie wir sahen, in $2j + 1$ Terme auf. Erfahrungsgemäß haben nun die Atome (oder Ionen) mit gerader Anzahl von Elektronen eine ungerade Anzahl von magnetischen Termkomponenten, also ganzzahliges j ; die Atome (Ionen) mit ungerader Elektronenzahl aber eine gerade Anzahl von Komponenten, also halbzahliges j . *Folglich muß das einzelne Elektron halbzahlige m und j besitzen*; eine Tatsache, die zuerst von PAULI¹ klar erkannt wurde. Andererseits zeigt die Erfahrung², daß trotzdem die in § 25 bis § 29 abgeleiteten Auswahl-, Polarisations- und Intensitätsregeln bei den ZEEMAN-Effekten der wirklichen Atome erfüllt sind, und zwar für ganzzahlige und für halbzahlige m, j . Man wird so zu der Vermutung geführt, daß die Formeln (I) und (II) [§ 25, (10), (15), (16)], aus denen diese empirisch be-

¹ PAULI, W. jr.: Z. Phys. Bd. 16, S. 155. 1923.

² Vgl. etwa den Bericht von DORGELO, H. B.: Z. Phys. Bd. 26, S. 756. 1925; Bd. 27, S. 182. 1926.

stättigten Regeln abgeleitet wurden, auch beim Elektron richtig sind, während die zur Gleichung $r\mathfrak{M} = 0$ führende *Definition des Drehimpulses* \mathfrak{M} nach § 25, (1) aufgegeben bzw. abgeändert werden muß. Man wird also dazu genötigt, dem Elektron neben dem aus Orts- und Impulsmatrizen r, p gebildeten Drehimpuls im gewöhnlichen Sinne noch ein besonderes zusätzliches „*Eigenmoment*“ zuzuschreiben, welches das Auftreten „halbzahliger“ irreduzibler Bestandteile in \mathfrak{M} veranlaßt. Wir haben die einfachsten dieser irreduziblen Bestandteile in § 27, (17) angeführt, nämlich die für $j = \frac{1}{2}$; man erhält wegen $m_x = \frac{1}{2}(\mu + \mu^\dagger)$, $m_y = \frac{1}{2i}(\mu - \mu^\dagger)$:

$$(9) \quad m_x^{(\frac{1}{2})} = \frac{1}{2} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad m_y^{(\frac{1}{2})} = \frac{1}{2} \begin{pmatrix} 0 & i \\ -i & 0 \end{pmatrix}, \quad m_z^{(\frac{1}{2})} = \frac{1}{2} \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}.$$

Später werden wir sehen, daß die *Hypothese des Magnetelektrons*¹, auf welche uns die letzten Betrachtungen hingeführt haben, unter Verwendung der Matrizen (9) in der Tat in sinngemäßer und der Erfahrung entsprechender Weise quantenmechanisch formuliert werden kann, und ferner daß dabei die für die §§ 27 bis 29 grundlegenden Gleichungen (I), (II) unverändert bleiben.

Für die Anwendung auf Atome mit mehr als einem einzigen Elektron bedarf übrigens die Theorie noch einer andern Ergänzung. Bei einem wirklichen Atom ist es überhaupt nicht möglich, Koordinaten und Impulse den Elektronen *einzeln* zuzuordnen, weil diese einander völlig gleich, nicht einzeln unterscheidbar sind. Man darf vielmehr nur *symmetrische Funktionen* der Koordinaten und Impulse aller Elektronen als physikalisch sinnvolle Größen betrachten. Dadurch wird naturgemäß die Frage nach der *Einfachheit* einer Lösung des Problems einschneidend verändert. Aber auch dann bleiben noch, wie wir später sehen werden, die Gleichungen (I) und (II) ungeändert.

Zusammenfassend stellen wir also fest, daß die Gleichungen (I) und (II) und die daraus ableitbaren Gesetzmäßigkeiten auch bei den später vorzunehmenden Erweiterungen der Theorie unverändert erhalten bleiben.

¹ GOUDSMIT, S. u. G. E. UHLENBECK: Naturwissensch. Bd. 13, S. 953. 1925. BICHOWSKY, F. R. u. H. C. UREY: Proc. nat. Acad. Sci. U. S. A. Bd. 12, S. 80. 1926.

§ 33. Separation der Variablen bei der Zentralbewegung eines Massenpunkts¹. Die Ergebnisse der letzten Paragraphen ermöglichen es, das Problem

$$(1) \quad H = \frac{1}{2m_0} p^2 + U(r),$$

wo p wieder der Impulsvektor und

$$(2) \quad r = \sqrt{r^2} = \sqrt{x^2 + y^2 + z^2}$$

ist, in einer Weise zu behandeln, die mit der klassischen Lösung durch Separation der Variablen korrespondiert. (Vgl. Bd. I, § 14.)

Wir berechnen

$$(3) \quad p_r = m_0 \dot{r},$$

wo p_r zunächst nur ein abkürzendes Zeichen sein soll, von dem wir aber sogleich zeigen werden, daß es die Bedeutung eines zu r konjugierten Impulses hat. Es wird

$2 p_r = 2 m_0 \dot{r} = 2 m_0 [H, r] = [p^2, r] = p \cdot [p, r] + [p, r] \cdot p$;
nach den Differentiationsregeln § 18, (3) ist z. B.

$$[p_x, r] = \frac{\partial r}{\partial x} = \frac{x}{r},$$

ferner

$$\left[p_x, \frac{x}{r} \right] = \frac{\partial}{\partial x} \frac{x}{r} = \frac{1}{r} - \frac{x^2}{r^3},$$

also

$$(4) \quad \begin{aligned} 2 p_r = p \frac{r}{r} + \frac{r}{r} p &= 2 p \frac{r}{r} - 2 \kappa \frac{1}{r} \\ &= 2 \frac{r}{r} p + 2 \kappa \frac{1}{r} \end{aligned} \quad \left(\kappa = \frac{h}{2\pi i} \right).$$

Daraus ergibt sich zunächst die Beziehung

$$(5) \quad p_r r + r p_r = p r + r p,$$

(die man auch aus $r^2 = r^2$ unmittelbar durch Differentiation erhält), sodann

$$p_r r - r p_r = p r - r p - 2 \kappa$$

oder

$$(6) \quad [p_r, r] = 1.$$

¹ PAULI W. jr.: Z. Phys. Bd. 36, S. 336. 1926. MENSING, L.: Z. Phys. Bd. 36, S. 814. 1926.

In der Tat ist also $p_r = m_0 \dot{r}$, wie in der klassischen Theorie, kanonischer Impuls zu r . Weiter wird nach (4) und (6)

$$(7) \quad r^2 p_r^2 = (r p_r)^2 - \kappa r p_r = r p_r (r p_r - \kappa) = (\tau \nu + \kappa) \cdot \tau \nu.$$

Andererseits ergibt direktes Ausrechnen:

$$\begin{aligned} \mathfrak{M}^2 &= (y p_z - z p_y)^2 + (z p_x - x p_z)^2 + (x p_y - y p_x)^2 \\ &= r^2 \nu^2 - \tau \nu (\tau \nu + \kappa), \end{aligned}$$

also nach (7)

$$\mathfrak{M}^2 = r^2 \nu^2 - r^2 p_r^2$$

oder

$$(8) \quad \nu^2 = p_r^2 + \frac{\mathfrak{M}^2}{r^2}.$$

Daher geht die Energiefunktion (1) über in

$$(9) \quad H = \frac{1}{2 m_0} p_r^2 + \frac{1}{2 m_0} \frac{\mathfrak{M}^2}{r^2} + U(r).$$

Damit ist unser Problem auf einen Freiheitsgrad reduziert. Denn die Größen r und $p_r = m_0 \dot{r}$ sind dreihinvariant, also nach den Ergebnissen des § 26 mit den Komponenten von \mathfrak{M} vertauschbar (wodurch auch die Schreibweise $\frac{\mathfrak{M}^2}{r^2}$ für $r^{-2} \mathfrak{M}^2 = \mathfrak{M}^2 r^{-2}$ gerechtfertigt wird); sie sind also Stufenmatrizen, nämlich Diagonalmatrizen bezüglich j , und jede Stufe ist mit dem zugehörigen irreduzibeln Bestandteil von M_x, M_y, M_z vertauschbar, also nach § 15, Satz 1 ein Zahlenvielaches der zur Stufe j gehörigen Einheitsmatrix, d. h. von m unabhängig. Man bekommt also die Lösung von (1) folgendermaßen:

Man löse für jeden Wert

$$(10) \quad j = 0, 1, 2, \dots$$

das Problem

$$(11) \quad H_j = \frac{1}{2 m_0} p_r^{(j)2} + \frac{1}{2 m_0} \frac{h^2}{4 \pi^2} \frac{j(j+1)}{r^{(j)2}} + U(r^{(j)})$$

durch kanonische Matrizen

$$(12) \quad p_r^{(j)} = (p_r^{(j)}(n_r, n'_r)), \quad r^{(j)} = (r^{(j)}(n_r, n'_r)).$$

Die Gesamtheit der Eigenwerte der Energie (1) stimmt überein mit der Gesamtheit der Energiewerte aller Probleme (11) mit $j = 0, 1, 2, \dots$

Die Matrizen r , p_r erhält man aus (12) als Stufenmatrizen

$$(13) \quad \begin{cases} p_r = (p_r^{(j)}(n_r, n'_r) \delta_{m m'} \delta_{j j'}) , \\ r = (r^{(j)}(n_r, n'_r) \delta_{m m'} \delta_{j j'}) . \end{cases}$$

Man hat dann aus (13) noch die Matrizen für r (und p) selbst abzuleiten. Zu diesem Zwecke kann man die Gleichung

$$(14) \quad \frac{1}{2} (II II^\dagger + II^\dagger II) + Z^2 = e^2 \cdot r^2 = e^2 \cdot r^2$$

benutzen. Unter Verwendung der Formeln § 32, (4), (5) für $II II^\dagger$ und $II^\dagger II$ und der Formeln § 29, (12), (13) für Z werden auf der linken Seite in (14) die Teilmatrizen mit $\Delta j = \Delta m = 0$ gleich

$$(15) \quad \left\{ \begin{aligned} & \frac{1}{2} \{ (j+m)(j+m-1) + (j-m)(j-m-1) \\ & \quad + 2(j+m)(j-m) \} b^{(j, j-1)} b^{\dagger(j-1, j)} \\ & + \frac{1}{2} \{ (j-m+1)(j-m+2) + (j+m+1)(j+m+2) \\ & \quad + 2(j+m+1)(j-m+1) \} b^{\dagger(j, j+1)} b^{(j+1, j)} \\ & = j(2j-1) b^{(j, j-1)} b^{\dagger(j-1, j)} \\ & \quad + (j+1)(2j+3) b^{\dagger(j, j+1)} b^{(j+1, j)} . \end{aligned} \right.$$

Wenn wir hierin die Bedingung § 32, (7'') für die Vertauschbarkeit von x und y einführen, so bekommen wir aus (14) endlich die beiden Beziehungen

$$(16) \quad \left\{ \begin{aligned} b^{\dagger(j, j+1)} b^{(j+1, j)} &= \frac{e^2 r^{(j)^2}}{(2j+1)(2j+3)} , \\ b^{(j+1, j)} b^{\dagger(j, j+1)} &= \frac{e^2 r^{(j+1)^2}}{(2j+1)(2j+3)} ; \end{aligned} \right.$$

aus diesen hat man die $b^{(j+1, j)}$ zu bestimmen, sobald die $r^{(j)}$ bekannt sind, und damit ist der Ortsvektor r selbst gefunden.

Ordnet man die Matrizen $p_r^{(j)}$, $r^{(j)}$ so, daß die Energien W_j mit der Quantenzahl n_r wachsen, wenn dieser die Werte $n_r = 0, 1, 2, \dots$ erteilt werden, so nennt man (nach SOMMERFELD) n_r die *radiale Quantenzahl*. Unsere Konstruktion zeigt, daß durch Angabe der drei Quantenzahlen m, j, n_r je eine Zeile oder Spalte der Matrizen des Problems (1) eindeutig gekennzeichnet wird.

Die Durchführung dieser Rechnungen soll für einen speziellen Fall im nächsten Paragraphen gegeben werden. Zuvor wollen wir hier noch eine später gebrauchte Hilfsformel ableiten. Ganz ähnlich wie oben bei der Bestimmung von \dot{r} erhalten wir die zeitliche Ableitung von $\frac{x}{r}$ durch:

$$\begin{aligned} 2m_0 \frac{d}{dt} \frac{x}{r} &= 2m_0 \left[W, \frac{x}{r} \right] = \left[p^2, \frac{x}{r} \right] \\ &= p \left[p, \frac{x}{r} \right] + \left[p, \frac{x}{r} \right] p \\ &= \{p_x (y^2 + z^2) - p_y x y - p_z x z\} \frac{1}{r^3} \\ &\quad + \frac{1}{r^3} \{(y^2 + z^2) p_x - x y p_y - x z p_z\}, \end{aligned}$$

oder

$$(17) \quad 2m_0 \frac{d}{dt} \frac{x}{r} = (M_y z - M_z y) \frac{1}{r^3} + \frac{1}{r^3} (z M_y - y M_z).$$

Führen wir nun neben dem schon benützten *skalaren Produkt*

$$(18) \quad \mathbf{a} \cdot \mathbf{b} = a_x b_x + a_y b_y + a_z b_z$$

zweier Vektoren \mathbf{a} , \mathbf{b} das durch

$$(19) \quad \begin{cases} (\mathbf{a} \times \mathbf{b})_x = a_y b_z - a_z b_y, \\ (\mathbf{a} \times \mathbf{b})_y = a_z b_x - a_x b_z, \\ (\mathbf{a} \times \mathbf{b})_z = a_x b_y - a_y b_x \end{cases}$$

definierte *Vektorprodukt* $\mathbf{a} \times \mathbf{b}$ ein¹, so geht aus der Formel (17) wegen der Vertauschbarkeit der Drehinvariante r mit den Komponenten von \mathfrak{M} folgende Vektorformel hervor:

$$(20) \quad \frac{d}{dt} \frac{\mathbf{r}}{r} = \left[W, \frac{\mathbf{r}}{r} \right] = \frac{1}{2m_0 r^3} \{(\mathfrak{M} \times \mathbf{r}) - (\mathbf{r} \times \mathfrak{M})\}.$$

§ 34. Die zweiatomige Molekel². Ein Beispiel, bei dem die auf einen Freiheitsgrad reduzierte Bewegungsgleichung § 33, (11)

¹ Das hier und später in §§ 35ff. für das Vektorprodukt gebrauchte Multiplikationszeichen \times ist natürlich nicht zu verwechseln mit dem in § 10 gebrauchten Zeichen für das direkte Produkt.

² MENSING, L.: Z. Phys. Bd. 36, S. 814. 1926. Vgl. ferner OPPENHEIMER, J. R.: Proc. Cambridge Philos. Soc. Bd. 23, III, S. 327. 1926. LANDAU, L.: Z. Phys. Bd. 40, S. 621. 1927. DENNISON, D. M.: Phys. Rev. Bd. 28, S. 318. 1926.

durch ein Näherungsverfahren gelöst werden kann, liefert die *zweiatomige Molekel*, aufgefaßt als System zweier Massenpunkte, der Kerne, im Abstände r , zwischen denen eine potentielle Energie $U(r)$ wirkt¹. Man hat dann

$$(1) \quad H = \frac{1}{2m_1} p_1^2 + \frac{1}{2m_2} p_2^2 + U(r), \quad r = \sqrt{(r_1 - r_2)^2}$$

wo m_1, m_2 die Massen, r_1, r_2 die Orts- und p_1, p_2 die Impulsvektoren der beiden Kerne sind. Diese Energiefunktion läßt sich dann, genau wie in der klassischen Theorie, in die Schwerpunkts- und die Relativbewegung separieren; man mache die Transformation (s. Bd. 1, § 20)

$$(2) \quad \begin{cases} \mathfrak{R} = \frac{m_1 r_1 + m_2 r_2}{m_1 + m_2}, \\ \mathfrak{r} = r_1 - r_2; \end{cases}$$

da sie linear ist, transformieren sich nach § 18, Satz 4, die konjugierten Impulse kontragredient in

$$(3) \quad \begin{cases} p_{\mathfrak{R}} = p_1 + p_2, \\ p = \frac{m_2 p_1 - m_1 p_2}{m_1 + m_2}, \end{cases}$$

oder aufgelöst

$$(3') \quad \begin{cases} p_1 = \frac{m_1}{m_1 + m_2} p_{\mathfrak{R}} + p, \\ p_2 = \frac{m_2}{m_1 + m_2} p_{\mathfrak{R}} - p. \end{cases}$$

Dann erhält man

$$(4) \quad H = \frac{1}{2} \frac{1}{m_1 + m_2} p_{\mathfrak{R}}^2 + \frac{1}{2} \left(\frac{1}{m_1} + \frac{1}{m_2} \right) p^2 + U(r).$$

Läßt man den Anteil der Schwerpunktsbewegung weg, der sich mit den Hilfsmitteln der elementaren Matrizenmechanik nicht behandeln läßt, so bleibt

$$(5) \quad H = \frac{1}{2m_0} p^2 + U(r), \quad \frac{1}{m_0} = \frac{1}{m_1} + \frac{1}{m_2},$$

¹ Die Frage, wie dieser Ansatz vom Standpunkte des Elektronenaufbaus der Molekel zu rechtfertigen ist, wird später erörtert werden.

also die Energiefunktion einer Zentralbewegung [s. § 33, (1)]. Auf diese kann man dann die in den §§ 32, 33 angegebenen Überlegungen anwenden und erhält die Gleichung § 33, (11)

$$(6) \quad H^{(j)} = \frac{1}{2} m_0 p_r^{(j)2} + \frac{1}{2} m_0 \frac{\hbar^2}{4\pi^2} \frac{j(j+1)}{r^{(j)2}} + U(r^{(j)}).$$

Eine *stabile Molekel* ist dadurch gekennzeichnet, daß die Zahlenfunktion $U(r)$ ein ausgeprägtes Minimum für einen positiven Wert von r hat. Dieser entspräche in der klassischen Theorie dem Gleichgewicht der Kerne; in der Quantenmechanik gibt es diesen Begriff nicht, wie das Beispiel des Oszillators zeigt, bei dem der niederste Quantenzustand noch Energie (Nullpunktsenergie, s. § 23) besitzt. Trotzdem läßt sich das Problem (6) nach einem der klassischen Theorie nachgebildeten Verfahren lösen¹.

Wir entwickeln die Zahlenfunktion $U(r)$ in der Umgebung einer zunächst beliebigen Stelle r_0 in eine Potenzreihe nach der Differenz

$$(7) \quad r - r_0 = \varrho :$$

$$(8) \quad U(r) = U_0 + U'_0 \varrho + \frac{1}{2} U''_0 \varrho^2 + \dots;$$

dabei ist in U_0, U'_0, \dots als Argument die Zahl r_0 einzusetzen.

Nun setzen wir für r die korrespondierende Matrix ein, von der in § 33 gezeigt wurde, daß sie eine Stufenmatrix mit den Stufen $r^{(j)}$ ist; zugleich wählen wir r_0 als Diagonalmatrix, bei der immer die zu einem j -Werte gehörigen Diagonalelemente gleich sind, also

$$(9) \quad r_0 = (r_0^{(j)} E^{(j)}),$$

wo $E^{(j)}$ die zur Stufe j gehörige, $(2j+1)$ -zeilige Einheitsmatrix ist. Dann ist auch ϱ eine Stufenmatrix mit den Stufen

$$(10) \quad \varrho^{(j)} = r^{(j)} - r_0^{(j)} E^{(j)},$$

und aus (8) folgt für die einzelne Stufe

$$(11) \quad U(r^{(j)}) = U_0^{(j)} + U_0^{(j)'} \varrho^{(j)} + \frac{1}{2} U_0^{(j)''} \varrho^{(j)2} + \dots;$$

dabei sind $U_0^{(j)}, U_0^{(j)'}, \dots$ gleich den Werten der Zahlenfunktionen U_0, U'_0, \dots für die Zahlenargumente $r_0^{(j)}$.

¹ Dieses Verfahren ist für die ältere Quantentheorie von M. BORN und E. HÜCKEL: Phys. Z. Bd. 24, S. 1. 1923 ausgeführt worden und ist in Bd. 1, § 20, dargestellt.

Ferner setzen wir die Binomialreihe

$$(12) \quad \frac{1}{r^{(j)2}} = \frac{1}{(r_0^{(j)} + \varrho^{(j)})^2} = \frac{1}{r_0^{(j)2}} \left(1 + \frac{\varrho^{(j)}}{r_0^{(j)}}\right)^{-2} \\ = \frac{1}{r_0^{(j)2}} \left(1 - 2 \frac{\varrho^{(j)}}{r_0^{(j)}} + 3 \frac{\varrho^{(j)2}}{r_0^{(j)2}} - + \dots\right)$$

an.

Da $r_0^{(j)} E^{(j)}$ mit $r^{(j)}$ vertauschbar ist, so ist $p_\varrho^{(j)} = p_r^{(j)}$ konjugierter Impuls zu $\varrho^{(j)}$. Daher wird die Gleichung (6), geordnet nach steigenden Potenzen von $\varrho^{(j)}$:

$$(13) \quad H^{(j)} = \frac{1}{2m_0} p_\varrho^{(j)2} + \frac{1}{2m_0 r_0^{(j)2}} \frac{\hbar^2}{4\pi^2} j(j+1) + U_0^{(j)} \\ + \left(-\frac{1}{m_0 r_0^{(j)3}} \frac{\hbar^2}{4\pi^2} j(j+1) + U_0^{(j)'}\right) \varrho^{(j)} \\ + \left(\frac{3}{2m_0 r_0^{(j)4}} \frac{\hbar^2}{4\pi^2} j(j+1) + \frac{1}{2} U_0^{(j)''}\right) \varrho^{(j)2} \\ + \dots$$

Hier sind die Faktoren der Potenzen von $\varrho^{(j)}$ Zahlen. Man kann nun die Glieder 1. Ordnung in $\varrho^{(j)}$ zum Verschwinden bringen, indem man

$$(14) \quad -\frac{1}{m_0 r_0^{(j)3}} \frac{\hbar^2}{4\pi^2} j(j+1) + U_0^{(j)'} = 0$$

setzt; das ist nämlich eine Gleichung zur Bestimmung von $r_0^{(j)}$. Sie korrespondiert offenbar mit der Gleichung der klassischen Theorie, die man erhält, wenn man für die gleichförmig (ohne Schwingungen) rotierende Molekel das Gleichgewicht zwischen Zentrifugalkraft und Anziehungskraft fordert:

$$\frac{\mathfrak{M}^2}{m_0 r_0^3} = U'(r_0);$$

und zwar geht (14) einfach hieraus hervor, wenn man für \mathfrak{M}^2 einen Eigenwert $\frac{\hbar^2}{4\pi^2} j(j+1)$ einsetzt.

Sodann kann man (13) so schreiben:

$$(15) \quad H^{(j)} = U_0^{(j)} + W_{\text{rot}}^{(j)} + W_{\text{osz}}^{(j)} + \dots,$$

wo

$$(16) \quad W_{\text{rot}}^{(j)} = \frac{1}{2 m_0 r_0^{(j)2}} \frac{\hbar^2}{4 \pi^2} j(j+1)$$

die *Rotationsenergie* und

$$(17) \quad W_{\text{osz}}^{(j)} = \frac{1}{2 m_0} p_e^{(j)2} + \frac{\kappa^{(j)}}{2} \varrho^{(j)2}$$

die *Oszillationsenergie* bedeutet, mit

$$(18) \quad \kappa^{(j)} = \frac{3}{m_0 r_0^{(j)4}} \frac{\hbar^2}{4 \pi^2} j(j+1) + U_0^{(j)''}.$$

Der Ausdruck (17) ist aber die Energiefunktion eines Oszillators mit der *Frequenz*

$$(19) \quad \nu_0^{(j)} = \frac{1}{2 \pi} \sqrt{\frac{\kappa^{(j)}}{m_0}},$$

hat also nach § 23, (23c) den Wert

$$(20) \quad W_{\text{osz}}^{(j)} = W_{\text{osz}}^{(j,n)} = \hbar \nu_0^{(j)} \left(n + \frac{1}{2} \right), \quad n = 0, 1, 2, \dots$$

wo n die Schwingungsquantenzahl ist.

Damit ist die Matrix H näherungsweise (bis auf Glieder 2. Ordnung der Schwingungsamplitude einschließlich) in eine Diagonalmatrix verwandelt, deren Elemente von den Quantenzahlen j, n abhängen. Als charakteristische Konstanten treten dabei die durch (19) gegebene Frequenz $\nu_0^{(j)}$ und das *Trägheitsmoment*

$$(21) \quad A^{(j)} = m_0 r_0^{(j)2}$$

auf; beide hängen noch von j , also vom Rotationszustand ab.

Die Formeln (16), (17) unterscheiden sich von den entsprechenden klassischen nur durch das Eingehen von $j(j+1)$ statt j^2 im Drehmoment und durch $n + \frac{1}{2}$ statt n bei der Schwingung [s. Bd. 1, § 20, (12)]. Die höheren Näherungen können mit Hilfe der Störungsrechnung (s. § 40) gefunden werden.

Über die Anwendbarkeit dieser Theorie auf Bandenspektren verweisen wir auf das im Bd. 1, § 20, gesagte.

Für heteropolare Molekeln (Ladungen $\pm e$ im Gleichgewichtsabstande r_0) kann man die Intensitäten der ultraroten Banden ausrechnen; dazu hat man aus den Gleichungen § 33, (16), die hier

$$(22) \quad \begin{cases} b^{\dagger(j-1, j)} b^{(j, j-1)} = \frac{e^2 (r_0^{(j-1)} + \varrho^{(j-1)})^2}{(2j-1)(2j+1)}, \\ b^{(j, j-1)} b^{\dagger(j-1, j)} = \frac{e^2 (r_0^{(j)} + \varrho^{(j)})^2}{(2j-1)(2j+1)} \end{cases}$$

lauten, die Matrizen $b^{(j, j-1)}$ auszurechnen und in § 29, (12), (13) einzusetzen.

Diese Rechnung wird einfach, wenn man sich auf den Fall kleiner Werte der Quantenzahl j , also kleiner Drehmomente beschränkt. Dann kann man in (14) das Glied mit $j(j+1)$ fortlassen und erhält ein von j unabhängiges r_0 aus der Gleichung

$$(14') \quad U'(r_0) = 0;$$

r_0 entspricht also dem Minimalwert der potentiellen Energie, d. h. dem Gleichgewicht nach der gewöhnlichen Statik. Ebenso wird man dann die Zentrifugaleffekte in den höheren Gliedern der Entwicklung (13) weglassen, insbesondere ein von j unabhängiges

$$(18') \quad \kappa = U''(r_0)$$

statt $\kappa^{(j)}$ einführen, so daß die Frequenz (19) eine Konstante

$$(19') \quad \nu_0 = \frac{1}{2\pi} \sqrt{\frac{\kappa}{m_0}}$$

wird. Dann tritt auch an die Stelle von $\varrho^{(j)}$ eine von j unabhängige Matrix ϱ , und an Stelle von $r^{(j)} = r_0 + \varrho^{(j)}$ die Matrix $r = r_0 + \varrho$.

In diesem Falle kann man offenbar die Gleichungen (22) einfach dadurch erfüllen, daß man

$$(23) \quad b^{(j, j-1)} = b^{\dagger(j-1, j)} = \frac{e(r_0 + \varrho)}{\sqrt{(2j-1)(2j+1)}}$$

setzt. Für ϱ ist dabei die in § 23, (25c) angegebene Matrix q des linearen Oszillators einzusetzen. Die Intensitäten der Bandenlinien ergeben sich dann aus den Formeln § 31, (11), (12)

$$(24) \quad \begin{cases} J^{(n, j; n', j)} = 0, \\ J^{(n, j; n', j-1)} = \frac{64 \pi^4 \nu_0^4 e^2}{3 c^3} j |Q_{nn'}|^2, \\ J^{(n, j-1; n', j)} = \frac{64 \pi^4 \nu_0^4 e^2}{3 c^3} j |Q_{nn'}|^2; \end{cases}$$

und mit den Oszillatorformeln erhält man

$$(25) \quad \left\{ \begin{array}{l} J(n, j; n', j') = 0, \quad \text{wenn } n - n' \neq \pm 1 \\ \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \text{oder } j - j' \neq \pm 1; \\ J(n, j; n-1, j) = J(n, j-1; n-1, j) \\ = J(n-1, j; n, j-1) = J(n-1, j-1; n, j) = \frac{8\pi^2 \nu_0^3 e^2 \hbar}{3\mu c^3} j \cdot n. \end{array} \right.$$

Die Intensitäten der aufgespaltenen Linie im Magnetfeld sind dann aus den Formeln § 31, (15), (16), (17) abzulesen.

§ 35. Integrale der Bewegungsgleichungen beim Wasserstoffatom¹. Wir besprechen zum Schluß dieses Kapitels einige Anwendungen unserer Theorie auf das *Wasserstoffatom*, oder allgemeiner auf ein Ion mit einer festen Punktladung als Kern und einem einzigen Elektron. An dieser Stelle sollen aber nur die größeren Gesetzmäßigkeiten der Spektren dieser Systeme behandelt werden; die Untersuchung der mit der Feinstruktur zusammenhängenden Effekte wird auf einen späteren Abschnitt des Buchs verschoben.

Bezüglich der *Wahl der Energiefunktion* müssen wir eine *Hypothese* machen; als *einfachsten*, aber sonst durchaus nicht von vornherein eindeutig bestimmten Ansatz übernehmen wir aus der BOHRschen Theorie die Funktion (s. Bd. 1, § 21, 22)

$$(1) \quad H = \frac{1}{2m_0} p^2 - \frac{\varepsilon}{r}, \quad \varepsilon = Ze^2.$$

Dabei ist e das elektrische Elementarquantum, Z die Kernladungszahl, und m_0 (s. die in § 34 ausgeführte Überlegung) die „effektive“ Masse, die mit der Kernmasse M und der Elektronenmasse durch die Formel

$$(1') \quad \frac{1}{m_0} = \frac{1}{M} + \frac{1}{\mu}$$

verbunden ist. Der *Erfolg* zeigt nachträglich, daß (1) in der gewünschten Annäherung wirklich richtig ist.

Nach § 33, (11) kann (1) auf das Problem mit *einem* Freiheitsgrad

$$(1'') \quad H_j = \frac{1}{2m_0} p_r^{(j)2} + \frac{1}{2m_0} \cdot \frac{\hbar^2}{4\pi^2} \frac{j(j+1)}{r^{(j)2}} - \frac{\varepsilon}{r^{(j)}}$$

¹ PAULI, W. jr.: Z. Phys. Bd. 36, S. 336. 1926.

zurückgeführt werden. Diese Gleichung (1'') können wir jedoch erst im 2. Teile dieses Buches untersuchen, nachdem wir tieferliegende mathematische Hilfsmittel zur Integration der quantenmechanischen Bewegungsgleichungen gewonnen haben (SCHRÖDINGERSche Eigenfunktionen).

An dieser Stelle wollen wir zunächst folgende Frage behandeln: In der klassischen Theorie besitzt das Problem (1) einige einfache *Integrale*, die zum Ausdruck bringen, daß die *Bahnkurve die Gestalt eines Kegelschnitts hat. Was ist das Analogon dieser Integrale in der Quantenmechanik?*

Mit Hilfe des am Ende von § 33 eingeführten Vektorprodukts [§ 33, (19)] lassen sich die Definition und die Vertauschungsregeln (I) von \mathfrak{M} [§ 25, (10)] in der Form

$$(2) \quad \begin{cases} \mathfrak{M} = \mathbf{r} \times \mathfrak{p}, \\ \mathfrak{M} \times \mathfrak{M} = -\kappa \mathfrak{M}, \end{cases} \quad \kappa = \frac{h}{2\pi i}$$

schreiben.

Mit Hilfe der Formel § 33, (20), die für jedes Zentralfeld gilt, folgt sofort, daß *der Vektor*

$$(3) \quad \mathfrak{A} = \frac{1}{m_0 \varepsilon} \frac{1}{2} \{ \mathfrak{M} \times \mathfrak{p} - \mathfrak{p} \times \mathfrak{M} \} + \frac{\mathbf{r}}{r}$$

zeitlich konstant ist.

Denn die kanonischen Bewegungsgleichungen lauten

$$m_0 \dot{\mathbf{r}} = \mathfrak{p}, \quad \dot{\mathfrak{p}} = -\frac{\varepsilon}{r^3} \mathbf{r};$$

und da ferner $\dot{\mathfrak{M}} = 0$ ist, so hat man

$$(4) \quad \dot{\mathfrak{A}} = \frac{1}{m_0 \varepsilon} \frac{1}{2} \{ \mathfrak{M} \times \dot{\mathfrak{p}} - \dot{\mathfrak{p}} \times \mathfrak{M} \} + \frac{d}{dt} \frac{\mathbf{r}}{r} = 0$$

wegen § 33, (20).

Die Konstanz des Vektors \mathfrak{A} ist als quantenmechanisches Analogon zu der Tatsache anzusehen, daß nach der klassischen Theorie das Elektron einen Kegelschnitt beschreibt. Es besteht nämlich die Gleichung

$$\mathfrak{A} \mathbf{r} + \mathbf{r} \mathfrak{A} = -\frac{2}{m_0 \varepsilon} \left\{ \mathfrak{M}^2 + \frac{3}{2} |\kappa|^2 \right\} + 2r,$$

deren Ableitung wir hier allerdings nicht ausführlich vorrechnen wollen, da wir sie nicht weiter benutzen werden. Für den Grenz-

fall der klassischen Theorie ($\kappa \rightarrow 0$) geht diese Gleichung in $\mathfrak{A}r = -\frac{\mathfrak{M}^2}{m_0 \varepsilon} + r$ über, und das ist wegen der Konstanz von \mathfrak{A} und \mathfrak{M}^2 die Gleichung eines Kegelschnitts. (Man zähle in der auf \mathfrak{M} senkrechten Bahnebene, in der auch \mathfrak{A} liegt, das Azimut φ von der Richtung \mathfrak{A} aus, so daß $\mathfrak{A}r = |\mathfrak{A}| r \cos \varphi$; dann wird

$$r(1 - |\mathfrak{A}| \cos \varphi) = \frac{\mathfrak{M}^2}{m_0 \varepsilon}$$

die Polargleichung des Kegelschnitts).

Die Betrachtung des Vektors \mathfrak{A} liegt übrigens auch der in Bd. 1, § 38 erläuterten Integrationsmethode von LENZ und KLEIN für das Wasserstoffatom der klassischen Mechanik zugrunde. (Dem Vektor \mathfrak{A} entspricht bis auf einen Faktor der dort betrachtete Mittelwert \bar{r} des Ortsvektors r .)

Wir haben jetzt die *Multiplikationsbeziehungen zwischen \mathfrak{A} , \mathfrak{M} und $H = W$* zu untersuchen. Zunächst betrachten wir den Vektor

$$(5) \quad 2\mathfrak{B} = \mathfrak{M} \times \mathfrak{p} - \mathfrak{p} \times \mathfrak{M};$$

dabei gebrauchen wir auch die Bezeichnung

$$(6) \quad B_x + iB_y = \mathfrak{B}.$$

Nach § 25, (2), (4) folgt aus

$$(\mathfrak{M} \times \mathfrak{p})_x + (\mathfrak{p} \times \mathfrak{M})_x = \kappa[M_y, p_z] - \kappa[M_z, p_y]$$

die Gleichung

$$\mathfrak{M} \times \mathfrak{p} + \mathfrak{p} \times \mathfrak{M} = -2\kappa \mathfrak{p};$$

man kann also \mathfrak{B} auch durch

$$(5') \quad \mathfrak{B} = (\mathfrak{M} \times \mathfrak{p}) + \kappa \mathfrak{p} = -(\mathfrak{p} \times \mathfrak{M}) - \kappa \mathfrak{p}$$

definieren. Es wird dann

$$\begin{aligned} [M_z, B_z] &= [M_z, M_x p_y - M_y p_x] + \kappa [M_z, p_z] \\ &= [M_z, M_x] p_y + M_x [M_z, p_y] \\ &\quad - [M_z, M_y] p_x - M_y [M_z, p_x], \end{aligned}$$

und nach (I) [§ 25, (10)] und § 25, (4) folgt:

$$(7) \quad [M_z, B_z] = 0.$$

Ferner wird

$$\begin{aligned} [M_z, B_x] &= [M_z, M_y p_z - M_z p_y] + \kappa [M_z, p_x] \\ &= M_x p_z - M_z p_x - \kappa p_y \end{aligned}$$

oder

$$(8) \quad [M_z, B_x] = -B_y;$$

ebenso

$$(9) \quad [M_z, B_y] = B_x.$$

Wir schreiben (7), (8), (9) zusammenfassend:

$$(10) \quad \begin{cases} [M_z, B_z] = 0, \\ [M_z, \mathfrak{B}] = i\mathfrak{B}. \end{cases}$$

Daneben gelten die hieraus durch zyklische Vertauschung der Koordinatenachsen entstehenden Relationen.

Ferner hat man

$$\begin{aligned} \mathfrak{M} \mathfrak{B} &= M_x (M_y p_z - M_z p_y) + M_y (M_z p_x - M_x p_z) \\ &\quad - M_z (M_x p_y - M_y p_x) + \kappa \mathfrak{M} \mathfrak{p} \\ &= -\kappa \mathfrak{M} \mathfrak{p} + \kappa \mathfrak{M} \mathfrak{p} = 0, \end{aligned}$$

also

$$(11) \quad \mathfrak{M} \mathfrak{B} = \mathfrak{B} \mathfrak{M} = 0.$$

Endlich beweisen wir die Formel

$$(12) \quad \mathfrak{B} \times \mathfrak{B} = \kappa \mathfrak{M} \cdot \mathfrak{p}^2.$$

Es ist nämlich

$$\begin{aligned} \mathfrak{B} \times \mathfrak{B} &= ((\mathfrak{M} \times \mathfrak{p}) \times (\mathfrak{M} \times \mathfrak{p})) \\ &\quad + \kappa ((\mathfrak{M} \times \mathfrak{p}) \times \mathfrak{p}) + \kappa (\mathfrak{p} \times (\mathfrak{M} \times \mathfrak{p})). \end{aligned}$$

Durch Ausrechnung erhält man

$$\begin{aligned} ((\mathfrak{M} \times \mathfrak{p}) \times (\mathfrak{M} \times \mathfrak{p}))_x &= \kappa [M_z p_x - M_x p_z, M_x p_y - M_y p_x] \\ &= \kappa M_x p^2; \end{aligned}$$

ferner wird

$$\begin{aligned} ((\mathfrak{M} \times \mathfrak{p}) \times \mathfrak{p})_x + (\mathfrak{p} \times (\mathfrak{M} \times \mathfrak{p}))_x &= \kappa [M_z p_x - M_x p_z, p_z] - \kappa [M_x p_y - M_y p_x, p_y] \\ &= -\kappa [M_x, p_z] p_z - \kappa [M_x, p_y] p_y = 0, \end{aligned}$$

und damit ist Formel (12) bewiesen.

Nummehr wenden wir uns zur Betrachtung von \mathfrak{A} , das nach (3) und (5) sich so durch \mathfrak{B} ausdrücken läßt:

$$(13) \quad \begin{cases} \mathfrak{A} = \frac{1}{\varepsilon m_0} \mathfrak{B} + \frac{\mathfrak{r}}{r}, \\ \mathbf{A} = A_x + i A_y. \end{cases}$$

Da r mit \mathfrak{M} vertauschbar ist, so erhält man aus (10), (11) sowie (II) [§ 25, (15)] und § 32, (1) die wichtigen Beziehungen:

$$(III) (14) \quad \begin{cases} [M_z, A_z] = 0, \\ [M_z, A] = iA; \end{cases}$$

$$(IV) (15) \quad \mathfrak{M}\mathfrak{A} = \mathfrak{A}\mathfrak{M} = 0.$$

Ferner beweisen wir

$$(V) (16) \quad \mathfrak{A} \times \mathfrak{A} = \kappa \frac{2}{m_0 \varepsilon^2} \mathfrak{M} \cdot H.$$

Nach (13) wird nämlich

$$\mathfrak{A} \times \mathfrak{A} = \frac{1}{\varepsilon^2 m_0^2} (\mathfrak{B} \times \mathfrak{B}) + \frac{1}{\varepsilon m_0} \left\{ \left(\mathfrak{B} \times \frac{\mathbf{r}}{r} \right) + \left(\frac{\mathbf{r}}{r} \times \mathfrak{B} \right) \right\},$$

also wegen (12)

$$(17) \quad \mathfrak{A} \times \mathfrak{A} = \kappa \cdot \frac{2}{m_0 \varepsilon^2} \mathfrak{M} \cdot \frac{1}{2m_0} \mathfrak{p}^2 + \frac{1}{\varepsilon m_0} \mathfrak{C},$$

wo

$$\mathfrak{C} = \left(\mathfrak{B} \times \frac{\mathbf{r}}{r} \right) + \left(\frac{\mathbf{r}}{r} \times \mathfrak{B} \right).$$

Nun ist die x -Komponente von \mathfrak{C}

$$\begin{aligned} C_x &= \kappa \left[\frac{y}{r}, B_z \right] - \kappa \left[\frac{z}{r}, B_y \right] \\ &= \kappa \frac{1}{r} [y, B_z] + \kappa \left[\frac{1}{r}, B_z \right] y \\ &\quad - \kappa \frac{1}{r} [z, B_y] + \kappa \left[\frac{1}{r}, B_y \right] z. \end{aligned}$$

Unter Benutzung von (5') wird das

$$\begin{aligned} C_x &= -\frac{\kappa}{r} \frac{\partial B_z}{\partial p_y} + \frac{\kappa}{r} \frac{\partial B_y}{\partial p_z} \\ &\quad + \kappa \left[\frac{1}{r}, M_x p_y - M_y p_x \right] y - \kappa \left[\frac{1}{r}, M_z p_x - M_x p_z \right] z \\ &\quad - \kappa^2 \frac{\partial \frac{1}{r}}{\partial z} \cdot y + \kappa^2 \frac{\partial \frac{1}{r}}{\partial y} \cdot z \end{aligned}$$

oder

$$C_x = -\frac{\kappa}{r} \frac{\partial B_z}{\partial p_y} + \frac{\kappa}{r} \frac{\partial B_y}{\partial p_z} \\ - \kappa M_x \left(\frac{\partial \frac{1}{r}}{\partial y} y + \frac{\partial \frac{1}{r}}{\partial z} z \right) + \kappa (M_y y + M_z z) \frac{\partial \frac{1}{r}}{\partial x}.$$

Wegen $\mathfrak{M}t = 0$ ist aber die letzte Zeile

$$- \kappa M_x \left(x \frac{\partial \frac{1}{r}}{\partial x} + y \frac{\partial \frac{1}{r}}{\partial y} + z \frac{\partial \frac{1}{r}}{\partial z} \right) = - \frac{\kappa}{r} M_x,$$

und man hat ferner

$$-\frac{\partial B_z}{\partial p_y} + \frac{\partial B_y}{\partial p_z} = -\frac{\partial}{\partial p_y} (M_x p_y) - \frac{\partial}{\partial p_z} (M_x p_z) \\ = -2 M_x - \left(p_y \frac{\partial}{\partial p_y} + p_z \frac{\partial}{\partial p_z} \right) M_x = -3 M_x,$$

also insgesamt

$$(18) \quad C_x = -2 \frac{\kappa}{r} M_x.$$

Setzt man das in (17) ein, so wird

$$\mathfrak{A} \times \mathfrak{A} = \kappa \frac{2}{m_0 \varepsilon^2} \mathfrak{M} \cdot \left(\frac{1}{2m_0} \mathfrak{p}^2 - \frac{\varepsilon}{r} \right),$$

und damit ist nach (1) die Gleichung (V), (16) bewiesen.

Die drei Komponentengleichungen, in die (V) zerlegt werden kann, sind übrigens nicht unabhängig voneinander, wenn man voraussetzt, daß \mathfrak{M} ein Integral ist und daß (I) und (III) gelten: aus einer von ihnen folgen die beiden andern. Z. B. folgt aus der Gleichung für die z -Komponenten:

$$[A_x, A_y] = \frac{2}{m_0 \varepsilon^2} H \cdot M_z,$$

nach der Rechenregel § 18, (4) und nach (I), (III):

$$0 = [M_y, [A_x, A_y]] + [A_x, [A_y, M_y]] + [A_y, [M_y, A_x]] \\ = -\frac{2}{m_0 \varepsilon^2} H \cdot M_x + [A_y, A_z],$$

und das ist die Gleichung (V) für die x -Komponente. Endlich haben wir noch die Gleichung

$$(VI) \quad (19) \quad \mathfrak{A}^2 - 1 = -\frac{2}{m_0 \varepsilon^2} (\mathfrak{M}^2 + |\kappa|^2) H$$

abzuleiten. Diese wird bewiesen sein, wenn wir die Richtigkeit von

$$(20) \quad \mathfrak{B}^2 = (\mathfrak{M}^2 + |\kappa|^2) p^2,$$

$$(21) \quad \mathfrak{B} \frac{r}{r} + \frac{r}{r} \mathfrak{B} = -2 (\mathfrak{M}^2 + |\kappa|^2) \frac{1}{r}$$

zeigen können.

Nun folgt (21) aus (5'):

$$\begin{aligned} \mathfrak{B} r = \kappa p r + (M_y p_z - M_z p_y) x + (M_z p_x - M_x p_z) y \\ + (M_x p_y - M_y p_x) z = \kappa p r + \mathfrak{M}^2, \end{aligned}$$

ebenso

$$r \mathfrak{B} = -\kappa r p - \mathfrak{M}^2;$$

also

$$\begin{aligned} \mathfrak{B} \frac{r}{r} + \frac{r}{r} \mathfrak{B} &= -2 \frac{\mathfrak{M}^2}{r} + \kappa^2 \left[p, \frac{r}{r} \right] \\ &= -2 \frac{\mathfrak{M}^2}{r} - |\kappa|^2 \left([p, r] \frac{1}{r} + \left[p, \frac{1}{r} \right] r \right) \\ &= -2 (\mathfrak{M}^2 + |\kappa|^2) \frac{1}{r}. \end{aligned}$$

Die Gleichung (20) ergibt sich aus (5') durch einfaches Nachrechnen, wenn man noch die zu $r \mathfrak{M} = 0$ analoge Beziehung $p \mathfrak{M} = 0$ beachtet; man bilde $\mathfrak{B}^2 = \mathfrak{B}^2 + (p \mathfrak{M})^2$ nach (5') und schiebe alle Faktoren p_x, p_y, p_z nach rechts.

Aus der Berechnung von C_x entnehmen wir endlich als Nebenresultat, daß die x -Komponente von $(\mathfrak{B} \times r) + (r \times \mathfrak{B})$ gleich

$$-\kappa \frac{\partial B_z}{\partial p_y} + \kappa \frac{\partial B_y}{\partial p_z} = -3 \kappa M_x$$

ist. Daraus ergibt sich die weitere Formel

$$(19) \quad (\mathfrak{A} \times r) + (r \times \mathfrak{A}) = -\frac{3\kappa}{\varepsilon \mu} \mathfrak{M}.$$

§ 36. **Ableitung der BALMER-Formel¹.** Die Feststellungen von § 35 führen unmittelbar zur BALMERSchen Formel. Wir können nämlich die *irreduzibeln Bestandteile* des Matrixsystems \mathfrak{A} , \mathfrak{M} bestimmen; sie seien (in vorläufig unbestimmter Numerierung) mit \mathfrak{A}_n , \mathfrak{M}_n bezeichnet. Die Gleichungen (I) (§ 25) und (III) bis (VI) (§ 35) gelten auch für \mathfrak{A}_n , \mathfrak{M}_n statt \mathfrak{A} , \mathfrak{M} ; dabei ist für H jetzt eine gewöhnliche Zahl, nämlich der zu diesem irreduzibeln Bestandteil gehörige Energiewert W_n zu setzen. (Vgl. § 21, Satz 1.) Entsprechend den allgemeinen Erörterungen von § 3 und § 18 müssen wir uns hier beschränken auf die *diskreten* Eigenwerte W_n der Energie; wir behaupten: *Die diskreten Eigenwerte der Energie W_n sind sämtlich negativ.* Das folgt aus dem am Ende von § 19 bewiesenen Satze (Virialsatz), der hier mit $\varrho = -1$ anwendbar ist; nach § 19, (17) gilt für die diskreten Energiewerte

$$(1) \quad \bar{T} = -\frac{1}{2} \bar{\Phi},$$

also

$$(2) \quad H = W = \bar{T} + \bar{\Phi} = -\bar{T} = -\frac{1}{2m_0} \bar{p}^2.$$

Nun sind notwendig, z. B. in p_x^2 alle Diagonalelemente positiv; denn wäre $p_x^2(n, n) = \sum_{n'} |p_x(n, n')|^2 = 0$ für irgend ein n , so wäre für dieses n und *alle* n' notwendig $p_x(n, n') = 0$, was mit $[p_x, x] = 1$ unverträglich ist. Also wird in der Tat $W_n < 0$.

Erste Ableitung der BALMER-Terme. Statt der \mathfrak{A}_n , \mathfrak{M}_n betrachten wir

$$(3) \quad \begin{cases} 2 \mathfrak{S}_n^{(1)} = \frac{2\pi}{h} \mathfrak{M}_n + \sqrt{\frac{R h Z^2}{|W_n|}} \mathfrak{A}_n, \\ 2 \mathfrak{S}_n^{(2)} = \frac{2\pi}{h} \mathfrak{M}_n - \sqrt{\frac{R h Z^2}{|W_n|}} \mathfrak{A}_n, \end{cases}$$

wo R eine Abkürzung für die RYDBERG-Konstante

$$(4) \quad R = \frac{2\pi^2 m_0 e^4}{h^3}$$

¹ Daß die Quantenmechanik die BALMERSchen Formeln richtig ergibt, ist zuerst von PAULI jr., W.: Z. Phys. Bd. 36, S. 336. 1926 gezeigt worden. Seine Ableitung ist die *zweite* der beiden hier vorgeführten. Die erste der hier gegebenen Ableitungen stützt sich auf Formeln, die gleichfalls in der genannten PAULISchen Arbeit erhalten worden sind.

ist. [Vgl. Bd. I, § 23, (2)]. Die $\mathfrak{M}_n, \mathfrak{N}_n$ drücken sich umgekehrt durch die $\mathfrak{J}_n^{(1)}, \mathfrak{J}_n^{(2)}$ aus als

$$(5) \quad \mathfrak{M}_n = \frac{h}{2\pi} (\mathfrak{J}_n^{(1)} + \mathfrak{J}_n^{(2)}), \quad \mathfrak{N}_n = \sqrt{\frac{|W_n|}{R h Z^2}} (\mathfrak{J}_n^{(1)} - \mathfrak{J}_n^{(2)}).$$

Die Gleichungen (I) und (III) bis (VI) gehen jetzt über in:

$$(Ia) \quad [I_{nx}^{(1)}, I_{nx}^{(2)}] = [I_{nx}^{(1)}, I_{ny}^{(2)}] = \dots = 0;$$

d. h. in Worten: jede der drei Komponenten von $\mathfrak{J}_n^{(1)}$ ist mit jeder Komponente von $\mathfrak{J}_n^{(2)}$ vertauschbar. Ferner:

$$(IIa) \quad \mathfrak{J}_n^{(1)} \times \mathfrak{J}_n^{(1)} = i \mathfrak{J}_n^{(1)}; \quad \mathfrak{J}_n^{(2)} \times \mathfrak{J}_n^{(2)} = i \mathfrak{J}_n^{(2)};$$

und endlich:

$$(IIIa) \quad \mathfrak{J}_n^{(1)2} = \mathfrak{J}_n^{(2)2} = \frac{1}{4} \left\{ \frac{R h Z^2}{|W_n|} - 1 \right\}.$$

Die Gleichungen (IIa) wiederholen für $\mathfrak{J}_n^{(1)}$ sowohl als auch für $\mathfrak{J}_n^{(2)}$, die schon in §§ 25 bis 27 gefundenen und untersuchten Vertauschungsregeln von $\frac{2\pi}{h} \mathfrak{M}$. [Vgl. auch § 35, (2).]

Die Ergebnisse von § 27 lassen uns also behaupten: Die Eigenwerte von $\mathfrak{J}_n^{(1)2}$ (und auch $\mathfrak{J}_n^{(2)2}$) sind notwendig von der Form

$$(6) \quad \omega^2 = \iota(\iota + 1); \quad \iota = 0, \frac{1}{2}, 1, \frac{3}{2}, 2, \dots$$

Statt ι gebrauchen wir hier besser die ganzzahlig fortschreitende Zahl

$$n = 2\iota + 1, \quad (n = 1, 2, 3, \dots),$$

die

$$(6') \quad \omega^2 = \frac{n-1}{2} \cdot \frac{n+1}{2} = \frac{1}{4} (n^2 - 1)$$

ergibt. Bei entsprechender Numerierung der Energieeigenwerte wird dann gemäß (IIIa):

$$\frac{1}{4} \left\{ \frac{R h Z^2}{|W_n|} - 1 \right\} = - \frac{1}{4} \left\{ \frac{R h Z^2}{W_n} + 1 \right\} = \frac{1}{4} (n^2 - 1)$$

oder

$$(7) \quad W_n = - \frac{R h Z^2}{n^2}.$$

Damit haben wir die gesuchte Energieformel. *Die Zahl n entspricht der Hauptquantenzahl.*

Wir wollen noch überlegen, wie groß die *Vielfachheit* eines Zustands mit der Hauptquantenzahl n ist.

Wir erhalten eine irreduzible Lösung der Gleichungen (Ia) bis (IIIa), indem wir die *Verschmelzung* bilden aus zwei irreduziblen Systemen, deren jedes äquivalent ist mit dem in § 27 beschriebenen einfachen System

$$m_x^{(\iota)}, m_y^{(\iota)}, m_z^{(\iota)}; \quad \iota = \frac{n-1}{2}.$$

Die so entstehenden Matrizen sind (bis auf eine beliebige unitäre Transformation) den Matrizen $\mathfrak{S}_n^{(1)}$ bzw. $\mathfrak{S}_n^{(2)}$ gleichzusetzen. Der Satz 5 des § 15, Kap. II, dessen Voraussetzungen hier erfüllt sind, versichert uns, daß dies die *einzig* Möglichkeit einer irreduziblen Lösung ist. Nun besteht jedes einzelne der beiden Systeme $m_x^{(\iota)}, \dots$ aus Matrizen mit $2\iota + 1 = n$ Diagonalelementen (vgl. § 27); in der Verschmelzung beider treten also n^2 Diagonalelemente auf: *Die Vielfachheit des Zustands mit der Hauptquantenzahl n ist gleich n^2 .*

Zweite Ableitung der BALMER-Terme. Wir wollen der soeben vorgeführten Ableitung der BALMER-Formel noch eine zweite gegenüberstellen, die zwar etwas langwieriger ist, aber dafür den Vorteil bietet, uns auch über die zur Hauptquantenzahl n gehörigen Werte der Drehimpulsquantenzahl j Auskunft zu verschaffen. Wir schreiben diesmal die einfachen Bestandteile von $\mathfrak{A}_n, \mathfrak{M}_n$ in solcher Form, daß das Matrixsystem \mathfrak{M}_n , d. h. M_{xn}, M_{yn}, M_{zn} für sich allein ausreduziert ist; es wird also \mathfrak{M}_n^2 eine *Diagonalmatrix*. Dann können wir so schließen:

Die Gleichungen (III), (IV) für den Vektor \mathfrak{A}_n sind dieselben wie (II) und § 32, (1) für den Vektor $\mathfrak{P} = (X, Y, Z)$; alle aus letzteren Formeln gezogenen Folgerungen über \mathfrak{P} gelten also ebenso für \mathfrak{A}_n . Insbesondere können in A_n und A_{zn} nur die Teilmatrizen

$$\begin{aligned} A_n^{(j, j-1)}(m+1, m), & \quad A_n^{(j-1, j)}(m+1, m), \\ A_{zn}^{(j, j-1)}(m, m), & \quad A_{zn}^{(j-1, j)}(m, m) \end{aligned}$$

von Null verschieden sein; und diese sind — entsprechend den Formeln § 29, (13) für I und Z — gegeben durch

$$(8) \quad \begin{cases} A_n^{(j, j-1)}(m+1, m) = \sqrt{(j+m+1)(j+m)} c_n^{(j, j-1)}, \\ A_n^{(j-1, j)}(m+1, m) = -\sqrt{(j-m-1)(j-m)} c_n^{\dagger(j-1, j)}, \\ A_{zn}^{(j, j-1)}(m, m) = -\sqrt{(j+m)(j-m)} c_n^{(j, j-1)}, \\ A_{zn}^{(j-1, j)}(m, m) = -\sqrt{(j+m)(j-m)} c_n^{\dagger(j-1, j)}. \end{cases}$$

Wir haben nun weiter die Gleichung (V), § 35, (16) für die z -Komponenten:

$$(9) \quad \frac{\kappa}{2} [A_n, A_n^\dagger] = -\frac{2i\kappa}{m_0 \varepsilon^2} W_n \cdot M_{zn}$$

zu erfüllen (nach der in § 35, S. 182 gemachten Bemerkung ist sie dann für die x - und y -Komponenten von selbst erfüllt). Wir

erinnern uns nun, daß wir in § 32, (7) schon die zu $\frac{\kappa}{2} [A, A^\dagger]$

analoge Größe $\frac{\kappa}{2} [II, II^\dagger]$ berechnet haben; danach können wir (9) unmittelbar umschreiben in

$$(10) \quad m \{ (2j-1) c_n^{(j, j-1)} c_n^{\dagger(j-1, j)} - (2j+3) c_n^{\dagger(j, j+1)} c_n^{(j+1, j)} \} \\ = -\frac{2|\kappa|^2}{m_0 \varepsilon^2} W_n \cdot m.$$

Das gilt wieder nur dann, wenn j nicht der kleinste, zu diesem irreduzibeln Bestandteil $\mathfrak{A}_n, \mathfrak{M}_n$ gehörige Wert $j_{\min}^{(n)}$ von j ist; sonst müssen in (10) die Glieder mit dem Index $j-1$ gestrichen werden. Dann bekommt aber m auf der linken Seite von (10) einen gewiß negativen Faktor, während auf der rechten Seite wegen $W_n < 0$ der Faktor von m sicher positiv ist. Also kann in diesem Falle nur $m = 0$ sein, d. h. es ist $j_{\min}^{(n)} = 0$. Bei einem aus endlichen Matrizen bestehenden einfachen Bestandteil $\mathfrak{A}_n, \mathfrak{M}_n$ müssen also die zugehörigen j -Werte eine endliche arithmetische Reihe mit dem Anfang 0 bilden; wir wählen die Numerierung n der einfachen Bestandteile nunmehr so, daß es die arithmetische Reihe

$$(11) \quad j = 0, 1, 2, \dots, n-1 \quad (n = 1, 2, \dots)$$

wird. Wir suchen nun eine solche Lösung von (10), bei der jeder Wert (11) von j nur einfach auftritt; man kann ähnlich, wie in den früheren Beispielen, einsehen, daß $\mathfrak{A}_n, \mathfrak{M}_n$ wirklich nur in

diesem Falle einfach ist. Die Formel (10) gibt dann für die Zahlen $C_{n,j} = |c_n^{(j,j-1)}|^2$ die Gleichungen

$$(12) \quad \begin{cases} (2j-1)C_{n,j} - (2j+3)C_{n,j+1} = C_n, & 0 < j < n-1; \\ (2n-3)C_{n,n-1} & = C_n; \\ & C_n = \frac{2|\kappa|^2}{m_0 \varepsilon^2} |W_n|. \end{cases}$$

Die Auflösung dieses rekursiven Systems liefert

$$(13) \quad C_{n,j} = C_n \frac{n^2 - j^2}{(2j-1)(2j+1)}.$$

Man bestätigt nämlich, daß für $j = n-1$ der Faktor von C_n in (13) gleich $\frac{1}{2n-3}$ wird; und mit

$$C_{n,j+1} = C_n \frac{n^2 - (j+1)^2}{(2j+1)(2j+3)}$$

wird auch die erste Gleichung (12) erfüllt:

$$\frac{n^2 - j^2}{2j+1} - \frac{n^2 - (j+1)^2}{2j+1} = 1.$$

Nunmehr bekommt man den Energiewert W_n selbst aus (VI), § 35, (19). Wir erinnern uns an die Berechnung der zu \mathfrak{A}_n^2 analogen Größe $\mathfrak{P}^2 = X^2 + Y^2 + Z^2$ in § 33, (14), (15) und schreiben statt (VI) die Gleichung

$$\begin{aligned} j(2j-1)c_n^{(j,j-1)}c_n^{\dagger(j-1,j)} + (j+1)(2j+3)c_n^{\dagger(j,j+1)}c_n^{(j+1,j)} \\ = 1 - \frac{2|\kappa|^2}{m_0 \varepsilon^2} |W_n| \{j(j+1) + 1\} \\ = 1 - C_n(j^2 + j + 1). \end{aligned}$$

Die linke Seite ist nach (12) und nach (13):

$$\begin{aligned} j(2j-1)C_{n,j} + (j+1)(2j+3)C_{n,j+1} \\ = j(2j-1)C_{n,j} + (j+1)\{(2j-1)C_{n,j} - C_n\} \\ = (2j-1)(2j+1)C_{n,j} - (j+1)C_n \\ = C_n(n^2 - j^2 - j - 1). \end{aligned}$$

Also erhält man für C_n die Gleichung

$$C_n(n^2 - j^2 - j - 1) = 1 - C_n(j^2 + j + 1),$$

$$(14) \quad C_n = \frac{1}{n^2}.$$

Daher wird nach (12)

$$(15) \quad |W_n| = \frac{m_0 \varepsilon^2}{2|\kappa|^2} \cdot \frac{1}{n^2} = \frac{2\pi^2 m_0 Z^2 e^4}{h^2} \cdot \frac{1}{n^2}.$$

Führt man wieder die RYDBERG-Konstante (4) ein, so erhält man erneut die Formel (7):

$$(16) \quad W_n = - \frac{R h Z^2}{n^2}.$$

In einem Punkte bedarf aber unsere Ableitung einer Ergänzung: Die Gleichungen (12) bestehen nur im Falle $n > 1$ (weil in der zweiten Gleichung der Index $n - 1$ auftritt); also ist auch (16) bisher nur für $n > 1$ bewiesen. Für $n = 1$ hat j nur den Wert $j = 0$, und \mathfrak{M}_1 stimmt überein mit dem einfachen Bestandteil

$$m_x^{(0)} = (0), \quad m_y^{(0)} = (0), \quad m_z^{(0)} = (0)$$

von \mathfrak{M} [vgl. § 27, (17)]. Wegen (III) [§ 35, (14)] wird daher auch $\mathfrak{A}_1 = 0$, und (VI) [§ 35, (19)] ergibt

$$-1 = \frac{2|\kappa|^2}{m_0 \varepsilon^2} W_1,$$

also

$$(16') \quad W_1 = - R h Z^2,$$

so daß (16) auch für $n = 1$ bewiesen ist. — Die Quantenmechanik liefert also im Falle des Wasserstoffatoms genau das Ergebnis, welches BOHR zuerst mit seiner halbklassischen Methode gewonnen hat, und das stets als wichtigste Stütze der quantentheoretischen Grundsätze gegolten hat.

Was aber die zu bestimmter Hauptquantenzahl n gehörigen Werte der Impulsquantenzahl j betrifft, so zeigt sich die neue Theorie der alten *überlegen*. In der BOHRschen Theorie waren nämlich auf Grund der Quantenbedingungen für die Impulsquantenzahl, die gewöhnlich als „Nebenquantenzahl“ k bezeichnet wurde, bei vorgegebenem n zunächst alle Werte $k = 0, 1, 2, \dots, n$

zulässig. (Vgl. Bd. I, § 21.) Die Anzahl $n + 1$ dieser Werte ist aber gegenüber der experimentellen Erfahrung um 1 zu groß. Man hat deshalb den Fall $k = 0$ (der einer „Pendelbahn“ mit Durchgang des Elektrons durch den Kern entsprechen hätte) durch ein besonderes „Zusatzverbot“ ausgeschlossen; dabei ergaben sich jedoch anderweitige Schwierigkeiten. Die Quantenmechanik dagegen liefert zwangsläufig nur n verschiedene Werte j , nämlich $j = 0, 1, 2, \dots, n - 1$.

Ein Zustand mit bestimmten n und j besitzt noch eine Vielfachheit $2j + 1$. Durch Summation bekommt man wieder die Vielfachheit der BALMER-Terme:

$$(17) \quad \sum_{j=0}^{n-1} (2j + 1) = n^2.$$

§ 37. Verhalten des H -Atoms in gekreuzten Feldern¹. Der Ausschluß der Pendelbahnen ($k = 0$) und weitere, ähnliche Zusatzverbote, die sich in der BOHRschen Theorie nicht vermeiden ließen, bedingten insbesondere beim Problem der gekreuzten Felder (vgl. Bd. I, § 38, S. 269 bis 276) tiefliegende Schwierigkeiten. Diese Schwierigkeiten sind schon in Bd. I besprochen, mögen aber hier noch einmal kurz erläutert werden. Bei einem H -Atom in gekreuzten Feldern treten in der BOHRschen Theorie neben der Hauptquantenzahl n zwei weitere adiabatisch invariante Quantenzahlen (vgl. Bd. I, § 16) n_1, n_2 auf (die Quantenzahl k kommt hier *nicht* vor), durch welche sich die durch die Felder verursachten Zusatzenergien der betreffenden stationären Zustände (in erster Näherung) darstellen lassen als

$$(1) \quad W^{(1)} = \left(\frac{n}{2} - n_1\right) \cdot h\nu_1 + \left(\frac{n}{2} - n_2\right) \cdot h\nu_2,$$

wo die ν_1, ν_2 gewisse nicht-negative Konstanten sind, die von der Richtung und Stärke der Felder abhängen. Sind $\mathfrak{E}, \mathfrak{H}$ die elektrische und magnetische Feldstärke, so setzen wir

$$(2) \quad \left\{ \begin{array}{ll} w_E = \frac{3}{2} \cdot \frac{e a_1}{h} n \cdot \mathfrak{E}, & w_H = \frac{e}{4\pi m_0 c} \cdot \mathfrak{H}; \\ w_1 = w_H + w_E, & w_2 = w_H - w_E; \\ \nu_E = |w_E|, & \nu_H = |w_H|, \end{array} \right.$$

¹ PAULI jr., W.: Z. Phys. Bd. 36, S. 336. 1926.

wo als Abkürzung

$$(3) \quad a_1 = \frac{\hbar^2}{4\pi^2 Z e^2 m_0}$$

sein soll¹. Wir nehmen jetzt übrigens den *Kern* unendlich schwer an, so daß m_0 gleich der Masse des Elektrons ist. Die Größe ν_H bedeutet die *LARMOR-Frequenz*. Dann wird in (1):

$$(4) \quad \nu_1 = |w_1| = |w_H + w_E|, \quad \nu_2 = |w_2| = |w_H - w_E|.$$

Für die n_1, n_2 erlauben die Quantenbedingungen — *ohne* Berücksichtigung von Zusatzverboten — beliebige ganzzahlige Wertkombinationen aus den Intervallen

$$(5) \quad 0 \leq n_1 \leq n, \quad 0 \leq n_2 \leq n.$$

Wir wollen nun die zunächst beliebigen Feldrichtungen $\mathfrak{E}, \mathfrak{H}$ in adiabatischer Weise (vgl. Bd. I, § 10) *parallel* werden lassen. Für parallele Felder haben wir dann als Spezialfall von (1) die Energieformel

$$(1') \quad W^{(1)} = (s\nu_E + m\nu_H)h.$$

Darin ist m die magnetische Quantenzahl; s ist die „Starkeffektquantenzahl“². Da in diesem Falle $\nu_2 = |\nu_H - \nu_E|$ wird, so berechnen sich die Zahlen s, m aus den n_1, n_2 nach

$$(6) \quad \begin{aligned} m &= n - (n_1 + n_2), \quad s = n_2 - n_1 \quad \text{für } \nu_H > \nu_E; \\ m &= n_2 - n_1, \quad s = n - (n_1 + n_2) \quad \text{für } \nu_H < \nu_E. \end{aligned}$$

(Im Falle $\nu_H = \nu_E$ ergibt sich mit $\nu_2 = 0$ eine Entartung der Energie.) Denken wir uns nunmehr ein H -Atom zunächst parallelen Feldern $\mathfrak{E}, \mathfrak{H}$ ausgesetzt, wobei etwa $\nu_H > \nu_E$ sein möge; dann sollen $\mathfrak{E}, \mathfrak{H}$ gegeneinander verdreht werden, und danach möge durch Änderung der Absolutwerte der Feldstärken $\nu_E < \nu_H$ gemacht werden; endlich sind $\mathfrak{E}, \mathfrak{H}$ wieder parallel zu machen. Alle diese Feldänderungen können adiabatisch ausgeführt werden, unter Vermeidung einer Entartung der Energie (denn $\nu_H = \nu_E$ bewirkt bei *nicht* parallelen $\mathfrak{E}, \mathfrak{H}$ im allgemeinen *keine* Entartung).

¹ In der BOHRschen Theorie bedeutet a_1 den Radius der einquantigen Wasserstoffbahn (für $Z = 1$; s. Bd. I, § 23, (8)).

² Ihre anschauliche Bedeutung in der BOHRschen Theorie ergibt sich daraus, daß die Projektion des Abstandes des elektrischen Mittelpunktes der Elektronenbahn vom Kern auf die Feldrichtung gegeben ist durch $\zeta = \frac{2}{3} a_1 n s$ (s. Bd. I, § 35, (1), (13), (16) und § 37, (3), (5), (6)).

Es bleiben also die n_1, n_2 dauernd ungeändert; aber wegen des Übergangs von $\nu_H > \nu_E$ zu $\nu_H < \nu_E$ vertauschen die Quantenzahlen s und m ihre Werte.

Unter diesen Umständen ist es nun nicht möglich, in widerspruchsfreier Weise Zusatzverbote einzuführen, um die gegenüber den empirischen Verhältnissen zu große Termmannigfaltigkeit einzuschränken. BOHR hat gezeigt, daß das Verbot der Pendelbahnen $k = 0$ im Falle beliebiger achsensymmetrischer Kraftfelder ein Verbot der Werte $m = 0$ der magnetischen Quantenzahl nach sich ziehen müßte. Im Falle des Starkeffekts allein ($\mathfrak{S} = 0$) hat man sich ferner genötigt gesehen, den Fall $s = n$ auszuschließen. Es geht aber bei unserer obigen Vertauschung der Werte von s und m der erlaubte Zustand $s = 0, m = n$ in den (doppelt) verbotenen $s = n, m = 0$ adiabatisch über.

Die Quantenmechanik beseitigt diese Schwierigkeiten. Wir werden in Kap. V, § 44 die Methoden der Störungstheorie auf das Problem des H -Atoms in gekreuzten Feldern anwenden, und dann für das hier betrachtete Modell zu (1), (5) ähnliche Ergebnisse erhalten; als Unterschied gegen die ältere Theorie ergibt sich jedoch auch hier ein Auftreten der Zahl $n - 1$ an Stelle von n :

$$(7) \quad \begin{cases} W^{(1)} = \left(\frac{n-1}{2} - n_1\right) h\nu_1 + \left(\frac{n-1}{2} - n_2\right) h\nu_2, \\ 0 \leq n_1 \leq n-1, \quad 0 \leq n_2 \leq n-1. \end{cases}$$

Danach gilt für *parallele* Felder wieder Formel (1') mit

$$(6') \quad \begin{cases} m = n - (n_1 + n_2 + 1), & s = n_2 - n_1 & \text{für } \nu_H > \nu_E, \\ m = n_2 - n_1, & s = n - (n_1 + n_2 + 1) & \text{für } \nu_E > \nu_H. \end{cases}$$

Insbesondere wird für den *Starkeffekt* ($\mathfrak{S} = 0$):

$$(7') \quad W^{(1)} = \frac{3}{2} e |\mathfrak{E}| a_1 n s; \quad 0 \leq s \leq n-1.$$

Dabei haben die Zahlen n_1, n_2 , wie wir später bei der Durchführung der Störungsrechnung begründen werden, folgende Bedeutung: Wir bestimmen zunächst für das ungestörte Problem eine solche Lösung, für welche die zu w_1 parallele Komponente $I_n^{(1)\parallel}$ des Vektors $\mathfrak{S}_n^{(1)}$ und die zu w_2 parallele Komponente $I_n^{(2)\parallel}$ von $\mathfrak{S}_n^{(2)}$ *Diagonalmatrizen* sind. Das ist nach unseren obigen Be-

merkungen bezüglich der Gleichungen (Ia) bis (IIIa) (§ 36) ohne weiteres zu erzielen: Wir sollten ja $\mathfrak{S}_n^{(1)}, \mathfrak{S}_n^{(2)}$ bilden durch *Verschmelzung* von zwei Systemen

$$(8) \quad m_x^{(i)}, \quad m_y^{(i)}, \quad m_z^{(i)}.$$

In § 27 hatten wir ein System (8) derart angeschrieben, daß $m_z^{(i)}$ Diagonalmatrix wird; jetzt wählen wir statt dessen einmal die zu w_1 parallele und das andere Mal die zu w_2 parallele Komponente des Vektors (8) als Diagonalmatrix. Dann werden in der Verschmelzung wirklich $I_n^{(1)\parallel}, I_n^{(2)\parallel}$ Diagonalmatrizen. Die Eigenwerte m_1, m_2 dieser Diagonalmatrizen bilden (wie die von $m_z^{(i)}$) je eine arithmetische Reihe der Differenz 1:

$$- \iota \leq m_1 \leq \iota, \quad - \iota \leq m_2 \leq \iota;$$

für

$$(9) \quad \begin{cases} n_1 = \iota - m_1 = \frac{n-1}{2} - m_1, \\ n_2 = \iota - m_2 = \frac{n-1}{2} - m_2 \end{cases}$$

gibt das in der Tat die in (7) angegebene Wertemannigfaltigkeit.

Bezüglich des Vergleichs mit der Erfahrung ist jedoch noch zu bemerken: In Wirklichkeit kommt man nicht aus mit dem hier zugrunde gelegten Modell des „Punktelektrons“, sondern muß zur Gewinnung einer exakten Theorie des H -Spektrums die Annahme eines *Eigenmomentes* des Elektrons (Magnelektron) quantenmechanisch formulieren; wir haben davon schon in § 32 gesprochen. *Dann wird jeder hier als einfach auftretende Zustand zweifach.* Statt n^2 (verschiedenen oder übereinstimmenden) Termen mit der Hauptquantenzahl n bekommt man also $2n^2$ Terme. Insbesondere ist der *Grundzustand* $n = 1$ *zweifach* und nicht einfach.

Sehen wir ab von einer magnetischen Störung des Atoms (ZEEMAN-Effekt), so werden durch die vom Magnelektron herührenden Effekte lediglich Änderungen (bzw. Aufspaltungen) der Terme von der *Größenordnung der Feinstruktur* (vgl. Bd. I, § 33) bewirkt. Insbesondere bleiben die BALMER-Formel und auch die Formel (7') des linearen Starkeffekts bis auf Korrekturen dieser Größenordnung richtig. (Man kann also (7') anwenden für Starkeffektaufspaltungen, die groß gegen die Feinstruktur sind.)

Dagegen ergibt sich eine Änderung bezüglich des ZEEMAN-Effekts auch bei Aufspaltungen, die groß gegen die Feinstruktur sind¹.

Fünftes Kapitel.

Störungstheorie.

§ 38. Störungsrechnung für abgeschlossene nicht entartete Systeme. Wenn die Energiefunktion eines abgeschlossenen mechanischen Systems von einem Parameter λ abhängt, $H(p, q, \lambda)$, so kann es vorkommen, daß für einen bestimmten Wert von λ , etwa $\lambda = 0$, die Lösung bekannt ist. Dann kann man auch für kleine Werte von λ die Lösung durch eine Reihenentwicklung — wenigstens formal — gewinnen; die Konvergenz des Verfahrens soll hier nicht untersucht werden². Es sei also

$$(1) \quad H(p, q, \lambda) = H^0(p, q) + \lambda H^{(1)}(p, q) + \lambda^2 H^{(2)}(p, q) + \dots$$

und man wisse, daß für bestimmte, bekannte Matrizen p_k^0, q_k^0

$$(2) \quad H^0(p^0, q^0) = \hat{W}$$

eine Diagonalmatrix ist. Man wird dann diese p_k^0, q_k^0 in $H(p, q, \lambda)$ einsetzen und kann das mechanische Problem nach § 19 zurückführen auf die Bestimmung einer unitären, natürlich von λ abhängigen Matrix U , für die

$$(3) \quad U(\lambda)^{-1} H(p^0, q^0, \lambda) U(\lambda) = W(\lambda)$$

eine Diagonalmatrix ist. Aus der Gleichung (3) selbst folgt nicht, daß $U(\lambda)$ unitär sein muß, sondern nur, daß $U^\dagger U$ mit $W(\lambda)$ vertauschbar ist. Das ergibt sich aus § 6, Satz d), (18 b) mit Rücksicht darauf, daß H und W hermitisch sind. Man hat also

$$(4) \quad [W, U^\dagger U] = 0.$$

Wir wollen nun zunächst den Fall behandeln, daß die Energie der ungestörten Bewegung nicht entartet ist. Dann gilt dasselbe auch für die gestörte Bewegung, solange λ hinreichend klein ist. Aus (4) folgt, daß $U^\dagger U$ eine Diagonalmatrix ist. Soll U unitär sein, so

¹ Also in der üblichen Bezeichnungsweise: bezüglich des vollendeten PASCHEN-BACK-Effektes.

² Vgl. diesbezüglich: WILSON, A. H.: Proc. Roy. Soc. A, Bd. 122, S. 589. 1929; Bd. 124, S. 176. 1929.

genügt es, zu fordern, daß die Diagonalelemente

$$(5) \quad (U^\dagger U)_{nn} = 1$$

sind.

Nun suchen wir die Bedingungen (3) und (5) durch den Ansatz

$$(6) \quad \begin{cases} U(\lambda) = 1 + \lambda \overset{(1)}{U} + \lambda^2 \overset{(2)}{U} + \dots, \\ W(\lambda) = \overset{\circ}{W} + \lambda \overset{(1)}{W} + \lambda^2 \overset{(2)}{W} + \dots \end{cases}$$

zu befriedigen. Man hat ferner nach (1) und (2)

$$(7) \quad H(\lambda) = \overset{\circ}{H} + \lambda \overset{(1)}{H} + \lambda^2 \overset{(2)}{H} + \dots,$$

wo überall die Argumente p_k^0, q_k^0 eingesetzt zu denken sind.

Um die Entwicklung von U^{-1} zu vermeiden, schreiben wir statt (3)

$$(8) \quad H(\lambda) U(\lambda) - U(\lambda) W(\lambda) = 0$$

und erhalten durch Nullsetzen der Glieder k -ter Ordnung in λ :

$$(9) \quad \sum_{l=0}^k \overset{(l)}{H} \overset{(k-l)}{U} - \overset{(k-l)}{U} \overset{(l)}{W} = 0, \quad k = 0, 1, 2, \dots,$$

wobei $\overset{\circ}{U} = 1$ gebraucht ist. Indem man nun immer das Glied $l = 0$ auf die eine, den Rest auf die andere Seite der Gleichung schreibt, erhält man das rekursive Gleichungssystem:

$$(10) \quad \begin{cases} 1) & \overset{\circ}{W} \overset{(1)}{U} - \overset{(1)}{U} \overset{\circ}{W} = \overset{(1)}{W} - \overset{(1)}{H}, \\ 2) & \overset{\circ}{W} \overset{(2)}{U} - \overset{(2)}{U} \overset{\circ}{W} = \overset{(2)}{W} - \overset{(2)}{H} + \overset{(1)}{U} \overset{(1)}{W} - \overset{(1)}{H} \overset{(1)}{U} \\ & \dots \\ k) & \overset{\circ}{W} \overset{(k)}{U} - \overset{(k)}{U} \overset{\circ}{W} = \overset{(k)}{W} - \overset{(k)}{H} + \sum_{l=1}^{k-1} (\overset{(k-l)}{U} \overset{(l)}{W} - \overset{(l)}{H} \overset{(k-l)}{U}) \\ & \dots \end{cases}$$

Andrerseits folgt aus $U^\dagger U = 1$

$$(11) \quad \sum_{l=0}^k \overset{(l)}{U}^\dagger \overset{(k-l)}{U} = 0, \quad k = 1, 2, 3, \dots$$

Trennt man hier die Glieder $l = 0$ und $l = k$ von der Summe ab,

so hat man

$$(11a) \quad U^{(k)} + U^{\dagger (k)} + \sum_{l=1}^{k-1} U^{\dagger (l)} U^{(k-l)} = 0.$$

Für alle Elemente außerhalb der Diagonale ist diese Matrixgleichung auf Grund von (4) von selbst erfüllt; die Bedingung (5) für die Diagonalglieder liefert:

$$(12) \quad U_{nn}^{(k)} + U_{nn}^{\dagger (k)} = - \sum_{l=1}^{k-1} \sum_j U_{nj}^{\dagger (l)} U_{jn}^{(k-l)}.$$

Hieraus sieht man, daß die Kenntnis aller Matrizen $U^{(1)}, U^{(2)}, \dots, U^{(k)}$ sofort die Realteile der Diagonalelemente von $U^{(k)}$ bestimmt. Es genügt also, wenn die Gleichungen (10) die außerhalb der Diagonale stehenden Elemente von $U^{(k)}$ zu berechnen erlauben, und das ist, wie wir sogleich sehen werden, gerade der Fall.

Die Gleichung (10,1) ist nämlich nach dem Satz des § 5, S. 30, dann und nur dann lösbar, wenn die Diagonalglieder der rechten Seite verschwinden:

$$(13) \quad W_n^{(1)} = H_{nn}^{(1)};$$

die Lösung lautet dann:

$$(14) \quad U_{nm}^{(1)} = - \frac{H_{nm}^{(1)}}{W_n^{(1)} - W_m^{(1)}} = - \frac{H_{nm}^{(1)}}{h \nu_{nm}^{(1)}}, \quad n \neq m,$$

während $U_{nn}^{(1)}$ unbestimmt bleibt. Andererseits folgt aus (12) für $k=1$, daß

$$(15) \quad U_{nn}^{(1)} + U_{nn}^{* (1)} = 0$$

sein muß; man könnte also $U_{nn}^{(1)}$ als beliebige rein imaginäre Zahl wählen. Es genügt aber

$$(14a) \quad U_{nn}^{(1)} = 0$$

zu setzen; denn wir brauchen ja nur *irgend eine* Lösung zu konstruieren und können nachträglich durch Multiplikation von $U(\lambda)$ mit einer Phasenmatrix die allgemeinste Lösung gewinnen.

Man verifiziert nun leicht, daß die durch (14), (14a) definierte Matrix U die Unitaritätsbedingung (11a) für $k = 1$, nämlich

$$U + U^\dagger = 0 \quad \text{oder} \quad U_{nm} + U_{mn}^* = 0$$

erfüllt; das folgt aus $H_{nm} = H_{mn}^*$, $\dot{v}_{nm} = -\dot{v}_{mn}$.

Nummehr wenden wir uns zur Gleichung (10, 2); die Lösbarkeitsbedingung lautet:

$$W_n = \left(H - U W + H U \right)_{nn},$$

oder nach (13), (14), (14a)

$$(16) \quad W_n = H_{nn} + \sum_k' \frac{H_{nk} H_{kn}}{h \dot{v}_{nk}}.$$

wo der Strich am Summenzeichen bedeutet, daß der Wert $k = n$ auszulassen ist.

Die Formeln (13) und (16) sind uns schon in § 2 (15), (16) begegnet (wobei $H = 0$ angenommen und $k = n - \tau$ geschrieben war); sie wurden dort durch korrespondenzmäßige Übertragung der entsprechenden klassischen Gleichungen gewonnen. Diese provisorischen Betrachtungen werden also von der systematischen Quantenmechanik vollauf bestätigt.

Man kann nun (10, 2) lösen und in derselben Weise fortfahren. Wir wollen sogleich den k -ten Schritt ins Auge fassen. Für die Lösbarkeit von (10, k) ist hinreichend und notwendig, daß

$$(17) \quad W_n = H_{nn} - \sum_{l=1}^{k-1} \left(U W - H U \right)_{nn}^{(k-l)l}.$$

Sodann wird

$$(18) \quad U_{nm}^{(k)} = - \frac{1}{h \dot{v}_{nm}} \left(H_{nm} - \sum_{l=1}^{k-1} \left(U W - H U \right)_{nm}^{(k-l)l} \right), \quad n \neq m;$$

ferner können wir $U_{nn}^{(k)}$ reell wählen und haben nach (12)

$$(18a) \quad U_{nn}^{(k)} = - \frac{1}{2} \sum_{l=1}^{k-1} \left(U^\dagger U \right)_{nn}^{(l)(k-l)}.$$

Die Gleichungen (17), (18), (18a) bilden ein vollständiges rekur-

sives System, aus dem man U, W berechnen kann, wenn man $U, \dots U$ und $W, \dots W$ kennt. Da wir U, W bereits explizite angegeben haben [(13), (14), (14a)], so ist damit das Störungsproblem — wenigstens formal — vollständig gelöst.

Will man die Entwicklung der p_k, q_k selbst, z. B.

$$(19) \quad q_k = \hat{q}_k + \lambda q_k^{(1)} + \lambda^2 q_k^{(2)} + \dots$$

gewinnen, so schreibt man die Transformationsgleichungen am besten in der Form

$$U q_k = \hat{q}_k U,$$

setzt die Reihen für U und q_k ein und vergleicht die Koeffizienten der Potenzen von λ . Man erhält

$$\sum_{l=0}^j U^{(l)(j-l)} q_k = \hat{q}_k U^{(j)}.$$

Trennt man hier immer das Glied $l = 0$ ab, so bekommt man die Rekursionsformeln

$$(20) \quad q_k^{(j)} = \hat{q}_k U^{(j)} - \sum_{l=1}^j U^{(l)(j-l)} q_k^{(l)},$$

deren erste lauten:

$$(21) \quad \left\{ \begin{array}{l} 1) \quad q_k^{(1)} = \hat{q}_k U^{(1)} - U \hat{q}_k, \\ 2) \quad q_k^{(2)} = \hat{q}_k U^{(2)} - U \hat{q}_k^{(1)} - U \hat{q}_k^{(1)}, \\ \dots \end{array} \right.$$

Viel gebraucht wird der vollständige Ausdruck von $q_k^{(1)}$, den man aus (21, 1) erhält, wenn man für U den Wert (14), (14a) benutzt:

$$(22) \quad q_k^{(1)}(n, m) = \sum_s' \left(\frac{H^{(1)}(n, s) \hat{q}_k(s, m)}{h \hat{\nu}^{(1)}(n, s)} - \frac{\hat{q}_k(n, s) H^{(1)}(s, m)}{h \hat{\nu}^{(1)}(s, m)} \right);$$

der Akzent am Summenzeichen bedeutet, daß man die Glieder, deren Nenner verschwinden würde, fortlassen soll. Ein genau entsprechender Ausdruck gilt für $p_k^{(1)}$.

und $\hat{f}'(n, m) = \overset{(1)}{H}(n, m)$; daher wird für $n \neq m$

$$\overset{(1)}{W}(n, m) = \frac{\overset{(1)}{H}(n, m) \overset{(1)}{W}_m}{\hbar \overset{(1)}{\nu}(n, m)} - \frac{\overset{(1)}{W}_n \overset{(1)}{H}(n, m)}{\hbar \overset{(1)}{\nu}(n, m)} + \overset{(1)}{H}(n, m) = 0,$$

dagegen für $n = m$

$$\overset{(1)}{W}(n, n) = \overset{(1)}{H}(n, n)$$

in Übereinstimmung mit (13).

Wir erwähnen noch den Fall, der in der Theorie des Magnetismus vorkommt (s. § 43), daß die Störungsfunktion 1. Ordnung $\overset{(1)}{H}$ ein Integral der ungestörten Bewegung ist. Dann wird $\overset{(1)}{H}(\overset{(1)}{p}, \overset{(1)}{q}) = \overset{(1)}{W}$ eine Diagonalmatrix; daher wird man in diesem Falle $\overset{(1)}{W} + \lambda \overset{(1)}{W} = \overset{(1)}{W}(\lambda)$ als ungestörte Energie ansehen und

$$(30) \quad H(\lambda) = \overset{(2)}{W}(\lambda) + \lambda^2 \overset{(2)}{H} + \dots$$

schreiben; es fehlt also die Störungsfunktion 1. Ordnung. Dann treten offenbar an die Stelle von (13), (14) die Formeln

$$(31) \quad \overset{(2)}{W}_n = \overset{(2)}{H}_{nn}, \quad \overset{(2)}{U}_{nm} = - \frac{\overset{(2)}{H}_{nm}}{\hbar \overset{(2)}{\nu}_{nm}},$$

und die höheren Näherungen werden entsprechend ein wenig modifiziert. Wenn nun eine Funktion $f(p, q, \lambda)$ berechnet werden soll, so fällt in (29) die Summe fort (weil ja kein $\overset{(1)}{H}$ vorhanden ist) und man hat einfach

$$(32) \quad \overset{(1)}{f} = \overset{(1)}{f}' = \left(\frac{\partial f(\overset{(1)}{p}, \overset{(1)}{q}, \lambda)}{\partial \lambda} \right)_{\lambda=0};$$

d. h. man erhält die Störung von f als Koeffizienten von λ in der Entwicklung von f für die ungestörte Bewegung.

§ 39. Der anharmonische Oszillator. Als allgemeinen *linearen Oszillator* bezeichnet man ein System mit der Energiefunktion

$$(1) \quad H = \frac{1}{2\mu} p^2 + U(q),$$

wo die potentielle Energie $U(q)$ als Funktion der Zahl q bei q_0 ein Minimum hat. Man kann $q_0 = 0$ annehmen. Entwickelt man U in der Umgebung dieser Stelle nach Potenzen von q , so kann man schreiben

$$(2) \quad H = H_0 + \lambda H_1 + \lambda^2 H_2 + \dots,$$

wo

$$(3) \quad H_0 = U(0) + \frac{1}{2\mu} p^2 + \frac{1}{2} U''(0) q^2$$

die Energiefunktion des harmonischen Oszillators (s. § 23) ist und die Störungsfunktionen der verschiedenen Ordnungen die Ausdrücke

$$(4) \quad \left\{ \begin{array}{l} H_1 = a q^3, \quad a = \frac{1}{3!} U'''(0), \\ H_2 = b q^4, \quad b = \frac{1}{4!} U''''(0), \\ \dots \dots \dots \end{array} \right.$$

haben; dabei ist der Größenordnungsfaktor λ gleich 1 gesetzt.

Man kann dieses System nach der im vorigen Paragraphen angegebenen Methode behandeln und erhält nach § 38, (13) und (16) für die Störungsenergien 1. und 2. Ordnung

$$(5) \quad \left\{ \begin{array}{l} W_n^{(1)} = a q^3(n, n), \\ W_n^{(2)} = b q^4(n, n) + \alpha^2 \sum_k' \frac{|q^3(n, k)|^2}{h \nu(n, k)}. \end{array} \right.$$

Man hat also Potenzen der in § 23, (25) angegebenen Matrix q auszurechnen. Dazu benutzt man am besten die Darstellung § 23, (24)¹:

$$(6) \quad q = A(b^\dagger - b), \quad A = \frac{1}{2\pi i} \sqrt{\frac{\hbar}{2\nu_0\mu}}.$$

Dann hat man zunächst

$$(7) \quad \left\{ \begin{array}{l} q^2 = A^2(b^{\dagger 2} + b^2 - b^\dagger b - b b^\dagger), \\ q^3 = A^3(b^{\dagger 3} - b^3 + b^2 b^\dagger + b^\dagger b^2 - b^{\dagger 2} b - b b^{\dagger 2} \\ \quad \quad \quad + b b^\dagger b - b^\dagger b b^\dagger); \end{array} \right.$$

¹ Die hier auftretende Matrix b ist nicht mit der Konstanten b in (4), (5) zu verwechseln.

nun sind nach § 23, (23 b)

$$(8) \quad b b^\dagger = ((n + 1) \delta_{n,m}), \quad b^\dagger b = (n \delta_{n,m})$$

Diagonalmatrizen, dagegen alle Potenzen von b und b^\dagger Matrizen ohne Diagonalelemente, z. B.

$$(9) \quad \begin{cases} b_{n,m} = \sqrt{n+1} e^{i q_n} \delta_{n,m-1}, \\ b_{n,m}^2 = \sqrt{(n+1)(n+2)} e^{i(q_n+q_{n+1})} \delta_{n,m-2}, \\ b_{n,m}^3 = \sqrt{(n+1)(n+2)(n+3)} e^{i(q_n+q_{n+1}+q_{n+2})} \delta_{n,m-3}, \\ \dots \end{cases}$$

Daher haben auch alle Produkte einer Potenz von b oder b^\dagger mit $b b^\dagger$ oder $b^\dagger b$ keine Diagonalelemente, und daraus folgt sofort $q^3(n, n) = 0$, also

$$(10) \quad W_n^{(1)} = 0.$$

Quadriert man nun den Ausdruck (7) für q^2 und behält nur die Diagonalelemente bei, so findet man leicht

$$(11) \quad q^4(n, n) = A^4 (6n^2 + 6n + 3).$$

Ferner wird nach (7), (8) und (9)

$$(12) \quad \begin{aligned} \frac{1}{A^3} q^3(n, m) &= 3(n+1) \sqrt{n+1} e^{i q_n} \delta_{n,m-1} \\ &\quad - 3n \sqrt{n} e^{-i q_{n-1}} \delta_{n,m+1} \\ &\quad - \sqrt{(n+1)(n+2)(n+3)} e^{i(q_n+q_{n+1}+q_{n+2})} \delta_{n,m-3} \\ &\quad + \sqrt{n(n-1)(n-2)} e^{-i(q_{n-1}+q_{n-2}+q_{n-3})} \delta_{n,m+3}. \end{aligned}$$

Die in $W_n^{(2)}$ auftretende Summe wird also mit Rücksicht auf § 23, (23 c)

$$\begin{aligned} \sum_k' \frac{|q^3(n, k)|^2}{h^2 \nu(n, k)} &= \frac{|A|^6}{h \nu_0} \left\{ -9(n+1)^3 + 9n^3 \right. \\ &\quad \left. - \frac{1}{3}(n+1)(n+2)(n+3) + \frac{1}{3}n(n-1)(n-2) \right\} \\ &= - \frac{|A|^6}{h \nu_0} (30n^2 + 30n + 11). \end{aligned}$$

Daher wird schließlich

$$(13) \quad \overset{(2)}{W}_n = \frac{3}{2} b \frac{h^2 (n^2 + n + \frac{1}{2})}{(2\pi)^4 \nu_0^3 \mu^2} - \frac{15}{4} a^2 \frac{h^2 (n^2 + n + \frac{11}{30})}{(2\pi)^6 \nu_0^4 \mu^3}.$$

Vergleicht man das mit der klassischen Formel Bd. 1, § 42, (6), so sieht man, daß für große Werte von $J = hn$ beide Ausdrücke übereinstimmen.

Man kann nun auch die Korrektur der Koordinate $q^{(1)}$ durch die anharmonischen Energien (5) nach § 38, (22) berechnen; nach längeren Reduktionen erhält man:

$$(14) \quad q^{(1)}(n, m) = -a \frac{h}{(2\pi)^4 \cdot \nu_0^3 \mu^2} \left\{ \delta_{nm} \frac{3}{2} (2n + 1) \right. \\ \left. + \delta_{n, m-2} \frac{1}{2} \sqrt{(n+1)(n+2)} e^{i(q_n + q_{n+1})} \right. \\ \left. + \delta_{n-2, m} \frac{1}{2} \sqrt{(m+1)(m+2)} e^{-i(q_m + q_{m+1})} \right\}.$$

Dies korrespondiert offenbar mit der FOURIER-Reihe

$$q^{(1)} = -a \frac{J}{(2\pi)^4 \cdot \nu_0^3 \mu^2} (3 + \cos 4\pi w),$$

wo $J = hn$ gesetzt und w die zugehörige Winkelvariable ist. Der Vergleich mit Bd. 1, § 42, (8) zeigt die völlige Übereinstimmung.

§ 40. Koppelung von Rotationen und Schwingungen bei der zweiatomigen Molekel. In § 34 wurde die Energie der zweiatomigen Molekel mit Hilfe eines Näherungsverfahrens berechnet, das in der Zurückführung auf die Energie eines, durch die Rotation modifizierten linearen Oszillators bestand. Dort wurde aber die Näherung nur so weit getrieben, als sie auf einen harmonischen Oszillator führt. Nunmehr können wir die höheren Näherungen, die anharmonischen Schwingungen entsprechen, mit Hilfe der Formeln § 39, (10), (13) berechnen. Wir hatten dort, § 34, (13) die HAMILTONsche Funktion nach Potenzen der Abstandsänderung $q^{(j)}$ entwickelt, wobei die Koeffizienten Funktionen der Rotationsquantenzahl j waren, die Reihe aber mit $q^{(j)^2}$ abgebrochen. Setzen wir sie um zwei Glieder fort, so erhalten wir

$$(1) \quad H^{(j)} = \tilde{W}^{(j)} + a q^{(j)^3} + b q^{(j)^4} + \dots;$$

dabei ist

$$(2) \quad \hat{W}^{(j)} = U^{(j)} + W_{\text{rot}}^{(j)} + W_{\text{osz}}^{(j)}$$

schon in § 34, (16), (17) berechnet und kann hier als ungestörte Energie gelten. Die Koeffizienten a , b der Störungsglieder lassen sich nach § 34, (11), (12) leicht bilden:

$$(3) \quad \begin{cases} a^{(j)} = -\frac{2}{m_0 r_0^{(j)5}} \frac{\hbar^2}{4\pi^2} j(j+1) + \frac{1}{6} U_0^{(j)\text{III}}, \\ b^{(j)} = \frac{5}{2m_0 r_0^{(j)6}} \frac{\hbar^2}{4\pi^2} j(j+1) + \frac{1}{24} U_0^{(j)\text{IV}}. \end{cases}$$

Nun kann man auf Grund von § 39, (13) sofort die Gesamtenergie als Funktion der Quantenzahlen n , j hinschreiben:

$$(4) \quad W_n^{(j)} = \hat{W}_n^{(j)} + W_n^{(j)(2)} + \dots,$$

oder ausführlich nach § 34, (16), (20)

$$(5) \quad W_n^{(j)} = U_0^{(j)} + \frac{\hbar^2 j(j+1)}{8\pi^2 m_0 r_0^{(j)2}} + \hbar \nu_0^{(j)} \left(n + \frac{1}{2} \right) \\ + \frac{3 b^{(j)} \hbar^2 (n(n+1) + \frac{1}{2})}{2 (2\pi)^4 \nu_0^{(j)2} m_0^2} \\ - \frac{15 a^{(j)2} \hbar^2 (n(n+1) + \frac{11}{30})}{4 (2\pi)^6 \nu_0^{(j)4} m_0^3} + \dots$$

Dabei ist noch $r_0^{(j)}$ aus der Gleichung § 34, (14)

$$(6) \quad \frac{\hbar^2 j(j+1)}{4\pi^2 m_0 r_0^{(j)3}} = U'(r_0^{(j)})$$

zu berechnen und daraus dann die Konstanten $U_0^{(j)}$, $\nu_0^{(j)}$, $a_0^{(j)}$, $b_0^{(j)}$. Für kleine Rotationsquantenzahlen kann das durch eine Reihenentwicklung nach $r_1^{(j)} = r_0^{(j)} - r_0$ geschehen, wo r_0 das Minimum der potentiellen Energie ist, also der Gleichung

$$(7) \quad U'(r_0) = 0$$

genügt. Begnügt man sich mit der ersten Näherung bezüglich $r_1^{(j)}$, so kann man aus (6) schließen:

$$\frac{\hbar^2 j(j+1)}{4\pi^2 m_0} = r_0^{(j)3} U'(r_0^{(j)}) = r_1^{(j)} \left(\frac{d}{dr} [r^3 U'(r)] \right)_{r=r_0} = r_1^{(j)} \cdot r_0^3 U''_0,$$

also

$$(8) \quad r_1^{(j)} = \frac{\hbar^2 j(j+1)}{4\pi^2 m_0 r_0^3 U_0''},$$

wo nunmehr r_0 und $U_0'' = U''(r_0)$ Konstanten der Molekel sind.

Nun kann man alle vorkommenden Funktionen von j nach $r_1^{(j)}$ oder, was dasselbe ist, nach $\hbar^2 j(j+1)$ entwickeln und erhält

$$(9) \quad \begin{cases} U_0^{(j)} = U_0 + \hbar^2 j(j+1) U_1 + \dots, \\ v_0^{(j)} = v_0 + \hbar^2 j(j+1) v_1 + \dots, \\ a^{(j)} = a_0 + \hbar^2 j(j+1) a_1 + \dots, \\ b^{(j)} = b_0 + \hbar^2 j(j+1) b_1 + \dots, \end{cases}$$

wo sämtliche Koeffizienten Zahlen sind, die von $r_0, U_0, U_0', U_0'', \dots$ abhängen; z. B. wird

$$(9a) \quad \begin{cases} U_1 = \frac{U_0'}{4\pi^2 m_0 r_0^3 U_0''}, & v_0 = \frac{1}{2\pi} \sqrt{\frac{U_0''}{m_0}}, \\ v_1 = \frac{3}{16\pi^2 r_0^4 U_0'' m_0^{\frac{3}{2}}} \left(1 + \frac{r_0 U_0'''}{3 U_0''} \right), \end{cases}$$

usw. Durch Einführung der Ausdrücke (9) in (5) erhält man die Energie als explizite Funktion der Quantenzahlen n und j . Die entstehende Formel ist mit der entsprechenden der älteren Theorie¹ [Bd. 1, § 20, (12)] bis auf additive Konstanten völlig identisch, wenn man j^2 durch $j(j+1)$ und n^2 durch $n(n+1)$ ersetzt. Die in Bd. 1, § 20 durchgeführte Diskussion der Schwingungszahlen und die Beschreibung der Banden bleibt also auch hier gültig; nur die quantitativen Verhältnisse verschieben sich ein wenig².

§ 41. Störungstheorie für abgeschlossene Systeme mit entarteter Energie. Wir knüpfen jetzt an die Ausführungen des § 38 an und lassen die Voraussetzung fallen, daß das ungestörte System nicht entartet ist. Sind einige der \hat{W}_n einander gleich, also einige der \hat{v}_{nm} gleich Null, so versagt bereits der erste Schritt der Approximation, dargestellt durch die Formel § 38, (14).

¹ Die hier mit j bezeichnete Quantenzahl wurde in Bd. 1, § 20 m genannt.

² Über die Anwendung dieser Bandentheorie auf die Messungen und die dabei gewonnenen Ergebnisse s. *Molecular spectra in gases* (Bull. of the Nat. Research Council, Vol. II, Part 3, No. 37), herausgeg. von KEMBLE, BIRGE, COLBY, LOOMIS, PAGE.

Wir denken uns \hat{W} als geordnete Diagonalmatrix geschrieben mit den Vielfachheiten g_r und zerlegen entsprechend alle Matrizen in rechteckige Teilmatrizen, z. B. die Gesamtenergie

$$(1) \quad H = (H^{(r,s)}), \quad H^{(r,s)} = (H_{\rho,\sigma}^{(r,s)}), \quad \begin{matrix} \rho = 1, 2, \dots, g_r \\ \sigma = 1, 2, \dots, g_s \end{matrix}$$

Die Aufgabe besteht wiederum in der Lösung der Gleichungen § 38, (10), die wir hier nochmals hinschreiben:

$$\left\{ \begin{array}{l} (2) \quad \overset{\circ}{W} \overset{(1)}{U} - \overset{(1)}{U} \overset{\circ}{W} = \overset{(1)}{W} - \overset{(1)}{H} \\ (3) \quad \overset{\circ}{W} \overset{(2)}{U} - \overset{(2)}{U} \overset{\circ}{W} = \overset{(2)}{W} - \overset{(2)}{H} + \overset{(1)}{U} \overset{(1)}{W} - \overset{(1)}{H} \overset{(1)}{U} \\ (4) \quad \overset{\circ}{W} \overset{(3)}{U} - \overset{(3)}{U} \overset{\circ}{W} = \overset{(3)}{W} - \overset{(3)}{H} + \overset{(1)}{U} \overset{(2)}{W} - \overset{(2)}{H} \overset{(1)}{U} + \overset{(2)}{U} \overset{(1)}{W} - \overset{(1)}{H} \overset{(2)}{U} \\ \dots \end{array} \right.$$

Dabei dürfen wir, wie in § 38, die Diagonalelemente der Matrizen $U^{(k)}$ reell wählen, sie lassen sich rekursiv durch die Gleichungen § 38, (12) oder

$$(5) \quad 2 \overset{(k)}{U}_{\rho\rho}^{(rr)} = - \sum_{l=1}^{k-1} \sum_{j,\tau} \overset{(l)}{U}_{\rho\tau}^{\dagger(r,i)} \overset{(k-l)}{U}_{\tau,\rho}^{(j,r)}$$

bestimmen. Nach dem Satz des § 5, S. 30 ist nun die Gleichung (2) dann und nur dann lösbar, wenn

$$\left(\overset{(1)}{W} - \overset{(1)}{H} \right)^{(r,r)} = 0$$

d. h. wenn $\overset{(1)}{H}^{(r,r)}$ eine Diagonalmatrix ist. Das wird im allgemeinen nicht von vornherein der Fall sein, was dem Umstand entspricht, daß die willkürlich gewählte Lösung \hat{q}_k, \hat{p}_k des ungestörten Problems im allgemeinen nicht stetig, für verschwindende λ , aus einer Lösung des gestörten Problems hervorgeht, und der Ansatz (6), § 38:

$$U(\lambda) = 1 + \lambda (\dots)$$

folglich für den entarteten Fall zu speziell war. Wir können aber $\overset{(1)}{H}$ in die gewünschte Form bringen, ohne die Diagonalfom von $\overset{\circ}{H} = \overset{\circ}{W}$ zu ändern, indem wir vorher alle \hat{q}_k, \hat{p}_k mit einer unitären Stufenmatrix $\overset{(1)}{u} = \left(\overset{(1)}{u}^{(r)} \delta_{rs} \right)$ transformieren in

$$(6) \quad \begin{matrix} (01) & (1) & \dots & (1) \\ q_k = u^{-1} & q_k & u, & p_k = u^{-1} & p_k & u \end{matrix}$$

derart, daß die unitären Stufen $u^{(1)}$ die einzelnen Untermatrizen $H^{(r,r)}$ auf Diagonalf orm bringen:

$$(7) \quad u^{(r)-1} H^{(r,r)} u^{(r)} = W^{(r)}$$

Dies ist ein Hauptachsenproblem im Raume von endlich vielen (g_r) Dimensionen, das nach den Methoden von § 11 gelöst werden kann. Die ersten Näherungen für die Störungen der Energiewerte sind also die Eigenwerte $W_0^{(r)}$ der Matrizen $H^{(r,r)}$; man erhält sie als Wurzeln der zu (7) gehörigen Determinantengleichung¹ g_r -ten Grades.

Man kann dies auch so ausdrücken: Es muß eine der Störung „angepaßte“ Lösung des ungestörten Systems gesucht werden, für welche der Zeitmittelwert der Störungsenergie $H^{(1)}$ die Diagonalf orm hat. Dies entspricht offenbar korrespondenzmäßig genau der BOHRschen Auffassung der entarteten Störungsprobleme in der früheren Theorie. (Vgl. Bd. I, § 43, sowie auch die Schlußbemerkung zu § 20 dieses Bandes).

Es seien also q_k, p_k so bestimmt, daß $H^{(r,r)}$ als Funktion der q_k, p_k die Diagonalf orm hat; dann besagt der Satz von § 5, S. 30, daß die Lösung von (2) durch

$$(8) \quad U^{(r,s)} = \frac{H^{(r,s)}}{W^{(r)} - W^{(s)}} = \frac{H^{(r,s)}}{h \dot{p}^{(r,s)}} \quad (r \neq s)$$

gegeben ist, wobei $U^{(r,r)}$ unbestimmt bleibt. Im Gegensatz zum § 38 genügt also hier infolge der Entartung die erste Näherungsgleichung nicht mehr, um die erste Näherungsmatrix $U^{(1)}$ restlos zu bestimmen; wir müssen jetzt wenigstens noch die zweite Nähe-

¹ Man nennt diese auch in Anlehnung an die Himmelsmechanik „Säkulargleichung“. Aus den $W_\epsilon^{(r)}$ leiten sich nämlich Frequenzen $\frac{\lambda}{h} (W_\epsilon^{(r)} - W_\sigma^{(r)})$ ab, die klein von der ersten Ordnung in λ sind und daher langsame, „säkulare“ Schwingungen darstellen.

rungsgleichung (3) mitbehandeln. Dies ist offenbar dem Umstand analog, daß zur Gewinnung der *angepaßten* Lösung der *ungestörten* Bewegungsgleichungen bereits die ersten Näherungsgleichungen herangezogen werden müssen.

Die Gleichung (3) ist ebenfalls nur dann lösbar, wenn

$$\overset{(2)}{W}^{(r)} - \overset{(2)}{H}^{(rr)} + \overset{(1)}{U}^{(rr)} \overset{(1)}{W}^{(r)} - \sum_t \overset{(1)}{H}^{(rt)} \overset{(1)}{U}^{(tr)} = 0$$

ist. Beachtet man, daß jetzt $\overset{(1)}{H}^{(rr)} = \overset{(1)}{W}^{(r)}$ ist, so wird diese Gleichung

$$\overset{(2)}{W}^{(r)} - \left\{ \overset{(2)}{H}^{(rr)} + \sum_t \overset{(1)}{H}^{(rt)} \overset{(1)}{U}^{(tr)} \right\} = \overset{(1)}{W}^{(r)} \overset{(1)}{U}^{(rr)} - \overset{(1)}{U}^{(rr)} \overset{(1)}{W}^{(r)},$$

wo der Strich am Summationszeichen bedeutet, daß der Wert $t = r$ auszulassen ist. Nach (8) lautet diese Bedingung

$$(9) \quad \overset{(2)}{W}^{(r)} - \left\{ \overset{(2)}{H}^{(rr)} + \sum_t' \frac{\overset{(1)}{H}^{(rt)} \overset{(1)}{H}^{(tr)}}{h \dot{\nu}^{(rt)}} \right\} = \overset{(1)}{W}^{(r)} \overset{(1)}{U}^{(rr)} - \overset{(1)}{U}^{(rr)} \overset{(1)}{W}^{(r)}.$$

Wenn nun alle $\overset{(1)}{W}_\rho^{(r)}$ verschieden sind, so bekommen wir aus (9) durch abermalige Anwendung des Satzes von § 5 (diesmal zur Auflösung nach $\overset{(1)}{U}^{(rr)}$):

$$(10) \quad \left\{ \begin{array}{l} \overset{(2)}{W}_\rho^{(r)} = \overset{(2)}{H}_{\rho\rho}^{(rr)} + \sum_t' \frac{\sum_\tau \overset{(1)}{H}_{\rho\tau}^{(rt)} \overset{(1)}{H}_{\tau\rho}^{(tr)}}{h \dot{\nu}^{(rt)}} = \overset{(2)}{H}_{\rho\rho}^{(rr)} + \sum_t' \frac{\sum_\tau \left| \overset{(1)}{H}_{\rho\tau}^{(rt)} \right|^2}{h \dot{\nu}^{(rt)}} \\ \overset{(1)}{U}_{\rho\sigma}^{(rr)} = -\frac{1}{h \dot{\nu}_{\rho\sigma}^{(r)}} \left[\overset{(2)}{H}_{\rho\sigma}^{(rr)} + \sum_t' \frac{\sum_\tau \overset{(1)}{H}_{\rho\tau}^{(rt)} \overset{(1)}{H}_{\tau\sigma}^{(tr)}}{h \dot{\nu}^{(rt)}} \right], \quad \rho \neq \sigma, \end{array} \right.$$

mit den Säkularfrequenzen

$$(11) \quad h \dot{\nu}_{\rho\sigma}^{(1)(r)} = \overset{(1)}{W}_\rho^{(r)} - \overset{(1)}{W}_\sigma^{(r)}.$$

Für die nicht-diagonalen Elemente liefert dann Gleichung (3) weiter

$$(12) \quad \overset{\circ}{W}^{(r)} \overset{(2)}{U}^{(rs)} - \overset{(2)}{U}^{(rs)} \overset{\circ}{W}^{(s)} \quad (r \neq s) \\ = -\overset{(2)}{H}^{(rs)} + \overset{(1)}{U}^{(rs)} \overset{(1)}{W}^{(s)} - \overset{(1)}{W}^{(r)} \overset{(1)}{U}^{(rs)} - \sum_t \overset{(1)}{H}^{(rt)} \overset{(1)}{U}^{(ts)},$$

woraus man

$$(13) \quad U_{\rho\sigma}^{(2)(rs)} = -\frac{1}{\hbar \gamma^{(rs)}} \left[H_{\rho\sigma}^{(2)(rs)} - U_{\rho\sigma}^{(1)(rs)} \left(W_{\sigma}^{(1)(s)} - W_{\rho}^{(1)(r)} \right) \right. \\ \left. + \sum_t \sum_{\tau} H_{\rho\tau}^{(1)(rt)} U_{\tau\sigma}^{(1)(ts)} \right] \quad (r \neq s)$$

ableitet (immer dem Satz von § 5 zufolge): dadurch ist aber nach (8) $U_{\rho\sigma}^{(2)(rs)}$ vollständig bestimmt. Die Elemente von $U^{(2)(rr)}$ bestimmen sich aus der dritten Näherung, und das sukzessive Näherungsverfahren läßt sich ohne Schwierigkeit beliebig weit verfolgen.

Wir fassen jetzt den Fall ins Auge, wo einige der $W_{\rho}^{(1)}$ einander gleich sind, d. h. wo die Entartung durch die Störung erster Näherung nicht oder nur teilweise aufgehoben wird. Die Behandlungsmethode dieses Falles ist genau dieselbe, die wir eben an Hand der Gleichung (2) erörtert haben. Wir wollen die Rechnung noch einen Schritt weiter führen.

Wir unterteilen also die Untermatrizen nach den Vielfachheiten der $W_{\rho}^{(1)}$ und schreiben von jetzt an:

$$A_{\rho\sigma}^{(rs)} = (A_{\alpha\beta}^{(r,\rho; s,\sigma)}).$$

Die Bedingung für die Lösbarkeit der Gleichung (9) lautet jetzt:

$$(14) \quad W^{(2)(r,\rho)} = H^{(2)(r,\rho; r,\rho)} + \sum_t \sum_{\tau} \frac{H^{(1)(r,\rho; t,\tau)} H^{(1)(t,\tau; r,\rho)}}{\hbar \gamma^{(rt)}};$$

d. h. die rechten Seiten müssen auf Diagonalform gebracht werden. Das leisten Transformationen $u^{(2)(r,\rho)}$. Wir müssen also alle Matrizen der Transformation

$$u = (u^{(2)(r,\rho)} \delta_{rs} \delta_{\rho\sigma})$$

unterwerfen, d. h. als Funktionen der neuen Variablen

$$(15) \quad q_k = u^{-1} q_k u, \quad p_k = u^{-1} p_k u$$

betrachten. Nun erhebt sich sofort die Frage, ob diese neue Transformation mit der durch die vorige $u^{(1)}$ befriedigten Lösbarkeits-

bedingung der Gleichung (2) verträglich ist. Dazu ist zu bemerken, daß die Transformation $\overset{(1)}{u}$ nach § 11 nur bis auf eine Stufenmatrix von der Form $(\varphi^{(r,\varrho)} \delta_{rs} \delta_{\varrho\sigma})$ bestimmt ist, so daß eine weitere Transformation von dieser Form die durch $\overset{(1)}{u}$ erreichte Lösbarkeit von Gleichung (2) bestehen läßt: unsere Transformation $\overset{(2)}{u}$ ist aber gerade von der zulässigen Form, so daß, wie schon oben erwähnt, die aus Gleichung (2) abgeleiteten Konsequenzen (6) und (8) auch nach Transformation mit $\overset{(2)}{u}$ ihre Gültigkeit behalten. Die Transformationen $\overset{(1)}{u}$ und $\overset{(2)}{u}$ muß man ausführen, bevor man die Störungsrechnung anfängt. Das ist möglich, da man sie aus den Diagonalforderungen (7) und (14) bestimmt, welche nur von vornherein bekannte Größen enthalten.

Die Gleichung (9) liefert sodann

$$(16) \quad \overset{(1)}{u}^{(r,\varrho;r\sigma)} = - \frac{1}{\hbar \overset{(1)}{\gamma}_{\varrho\sigma}^{(r)}} \left[\overset{(2)}{H}^{(r,\varrho;r\sigma)} + \sum_t' \frac{\sum_{\tau} \overset{(1)}{H}^{(r,\varrho;t,\tau)} \overset{(1)}{H}^{(t,\tau;r\sigma)}}{\hbar \overset{(1)}{\gamma}^{(r,t)}} \right] (\varrho \neq \sigma).$$

Es bleiben noch die Elemente $\overset{(1)}{U}_{\alpha\beta}^{(r,\varrho;r,\varrho)}$ zu bestimmen; diese ergeben sich erst aus der dritten Näherungsgleichung (4).

Vorher schreiben wir noch (13) in der neuen Indizierung hin:

$$(17) \quad \overset{(2)}{U}^{(r,\varrho;s,\sigma)} = - \frac{1}{\hbar \overset{(2)}{\gamma}^{(r,s)}} \left[\overset{(2)}{H}^{(r,\varrho;s,\sigma)} \quad (r \neq s) \right. \\ \left. + \sum_t' \sum_{\tau} \overset{(1)}{H}^{(r,\varrho;t,\tau)} \overset{(1)}{U}^{(t,\tau;s,\sigma)} - \overset{(1)}{U}^{(r,\varrho;s,\sigma)} \left(\overset{(1)}{W}^{(s,\sigma)} - \overset{(1)}{W}^{(r,\varrho)} \right) \right].$$

Dann betrachten wir die Lösbarkeitsbedingungen der dritten Näherungsgleichung (4):

$$\overset{(3)}{W}^{(r)} - \overset{(3)}{H}^{(r,r)} + \overset{(1)}{U}^{(r,r)} \overset{(2)}{W}^{(r)} - \sum_t' \overset{(2)}{H}^{(r,t)} \overset{(1)}{U}^{(t,r)} + \overset{(2)}{U}^{(r,r)} \overset{(1)}{W}^{(r)} \\ - \sum_t \overset{(1)}{H}^{(r,t)} \overset{(2)}{U}^{(t,r)} = 0.$$

Dies schreiben wir zunächst um in

$$(18) \quad \overset{(1)}{W}^{(r)} \overset{(2)}{U}^{(r,r)} - \overset{(2)}{U}^{(r,r)} \overset{(1)}{W}^{(r)} = \overset{(3)}{W}^{(r)} - \overset{(3)}{H}^{(r,r)} + \overset{(1)}{U}^{(r,r)} \overset{(2)}{W}^{(r)} \\ - \sum_t \overset{(2)}{H}^{(r,t)} \overset{(1)}{U}^{(t,r)} - \sum_t' \overset{(1)}{H}^{(r,t)} \overset{(2)}{U}^{(t,r)}.$$

Die Lösbarkeitsbedingung für diese letzte Gleichung ist nun

$$\begin{aligned}
 0 = & \overset{(3)}{W}(r, \varrho) - \overset{(3)}{H}(r, \varrho; r, \varrho) + \overset{(1)}{U}(r, \varrho; r, \varrho) \overset{(2)}{W}(r, \varrho) - \overset{(2)}{H}(r, \varrho; r, \varrho) \overset{(1)}{U}(r, \varrho; r, \varrho) \\
 & - \sum_t \sum_\tau \overset{(2)}{H}(r, \varrho; t, \tau) \overset{(1)}{U}(t, \tau; r, \varrho) \\
 & - \sum_t \frac{1}{h \hat{\nu}(r t)} \sum_\tau \overset{(1)}{H}(r, \varrho; t, \tau) \overset{(1)}{H}(t, \tau; r, \varrho) \overset{(1)}{U}(r, \varrho; r, \varrho) \\
 & - \sum_t \frac{1}{h \hat{\nu}(r t)} \sum_\tau \overset{(1)}{H}(r, \varrho; t, \tau) \left[\overset{(2)}{H}(t, \tau; r, \varrho) + \sum_{\sigma, \alpha}'' \overset{(1)}{H}(t, \tau; s, \sigma) \overset{(1)}{U}(s, \sigma; r, \varrho) \right. \\
 & \left. - \overset{(1)}{U}(t, \tau; r, \varrho) \left(\overset{(1)}{W}(r, \varrho) - \overset{(1)}{W}(t, \tau) \right) \right],
 \end{aligned}$$

wobei der Doppelstrich am Summenzeichen $\sum_{\sigma, \alpha}''$ bedeutet: die Werte $s = t$ und $s = r$, $\sigma = \varrho$ sind auszulassen.

Unter Beachtung von (14) läßt sich diese Gleichung auf die Form bringen:

$$\begin{aligned}
 (19) \quad & \overset{(2)}{W}(r, \varrho) \overset{(1)}{U}(r, \varrho; r, \varrho) - \overset{(1)}{U}(r, \varrho; r, \varrho) \overset{(2)}{W}(r, \varrho) \\
 & = \overset{(3)}{W}(r, \varrho) - \overset{(3)}{H}(r, \varrho; r, \varrho) - \sum_t \sum_\tau \left\{ \overset{(2)}{H}(r, \varrho; t, \tau) \overset{(1)}{U}(t, \tau; r, \varrho) \right. \\
 & \quad \left. + \frac{1}{h \hat{\nu}(r t)} \overset{(1)}{H}(r, \varrho; t, \tau) \left[\overset{(2)}{H}(t, \tau; r, \varrho) + \sum_{s, \alpha}'' \overset{(1)}{H}(t, \tau; s, \sigma) \overset{(1)}{U}(s, \sigma; r, \varrho) \right. \right. \\
 & \quad \left. \left. - \overset{(1)}{U}(t, \tau; r, \varrho) \left(\overset{(1)}{W}(r, \varrho) - \overset{(1)}{W}(t, \tau) \right) \right] \right\}.
 \end{aligned}$$

Wenn alle $\overset{(2)}{W}(r, \varrho)$ voneinander verschieden sind, dann ist (19) die Bestimmungsgleichung für $\overset{(3)}{W}_\alpha(r, \varrho)$ und für die noch unbekanntes $\overset{(1)}{U}_{\alpha\beta}(r, \varrho; r, \varrho)$. Wenn die Entartung noch nicht völlig aufgehoben ist, so ist vorher eine dritte Unterteilung der Matrizen vorzunehmen und eine dritte Transformation $\overset{(3)}{u}$ anzuwenden, die die diagonalen Untermatrizen der rechten Seite auf die Diagonalf orm bringt. Das Verfahren ist ohne weiteres beliebig fortsetzbar.

Nur dann, wenn die Entartung bei einer gewissen Näherung völlig aufgehoben wird, gelangt man, wenn man bis zu dieser

Näherung fortschreitet, zur restlosen Bestimmung der *ersten* Näherung (und dann auch natürlich der weiteren Näherungen) der Transformationsmatrix $U(\lambda)$.

Es kann aber vorkommen, daß die Entartung bei keiner Näherung aufgehoben wird. Das ist der Fall, wenn die gestörte Energiefunktion $H(p, q, \lambda)$ für alle Werte von λ entartet ist. Unterteilt man die Matrizen entsprechend der nicht-aufheb-
baren Entartung, so bleiben die Nicht-Diagonalelemente der Diagonaluntermatrizen von U ⁽¹⁾ unbestimmt (und entsprechen-
des gilt für die höheren Näherungen); die Diagonalelemente können, nach § 38, (12), beliebig rein imaginär gewählt werden. Man darf also z. B. diese Diagonaluntermatrizen null setzen, was im Einklang mit der Normierung (5) ist. Diese Willkür entspricht dem Umstand, daß $U(\lambda)$ in diesem Falle nach § 20 nur bis auf eine Stufenmatrix bestimmt ist. Physikalisch sinn-
voll sind nur Größen, welche bezüglich dieser Willkür invariant sind (Relativinvarianten).

Als Beispiel für eine nicht völlig aufheb-
bare Entartung führen wir den im nächsten Paragraphen behandelten Fall eines Systems im elektrischen Feld an. Die Entartung beruht auf dem Umstand, daß das gestörte System die Gruppe der Drehungen um die Feld-
richtung und der Spiegelungen an Ebenen durch die Feldrichtung zuläßt.

Schließlich läßt sich aus U ⁽¹⁾, wie in § 38, die erste Näherung der gestörten Koordinaten und jeder Funktion von ihnen berechnen; es ist

$$(20) \quad q_k = q_k^{(1)} U - U q_k^{(os)}$$

und entsprechend für p_k ⁽¹⁾; dabei ist s die Ordnung derjenigen Näherung, bei der die Entartung so vollständig wie überhaupt möglich aufgehoben wird. Ein Beispiel behandeln wir im nächsten Paragraphen.

§ 42. Starkeffekt und elektrische Polarisierbarkeit. Man kann die Störungsrechnung verwenden, um den Einfluß eines konstanten elektrischen Feldes auf atomare Systeme zu untersuchen. Beim Einschalten des Feldes \mathfrak{E} wird [s. § 2, (12)] die Arbeit $A = \mathfrak{P}\mathfrak{E}$ geleistet, wo \mathfrak{P} den Vektor des elektrischen Moments

bedeutet; bezeichnen wir seine Komponenten, wie in § 25, (13), mit X , Y , Z , und nehmen an, daß das Feld \mathcal{E} parallel zur z -Achse gerichtet ist und die Stärke E hat, so wird.

$$(1) \quad A = E Z.$$

Dabei ist E eine Zahl, Z eine Matrix, nämlich

$$(2) \quad Z = \sum_k e_k z_k.$$

Ist nun das Feld schwach, so kann man die negative Arbeit als Störungsenergie betrachten und hat

$$(3) \quad \overset{(1)}{H} = - E Z, \quad \overset{(2)}{H} = \overset{(3)}{H} = \dots = 0.$$

Wenn es sich um freie Atome oder Molekeln handelt, liegt Entartung vor (s. Kap. IV, § 26); daher sind die Formeln von § 41 anzuwenden. Um die Störung 1. Ordnung zu finden, hat man die Gleichungen § 41, (7) zu bilden, welche hier lauten:

$$(4) \quad - \overset{(1)}{u}^{(r)-1} E Z^{(r,r)} \overset{(1)}{u}^{(r)} = \overset{(1)}{W}^{(r)};$$

dabei sind die Indizes r so gewählt, daß zu jedem ein verschiedener Eigenwert $\overset{(1)}{W}^{(r)}$ des ungestörten Systems gehört. Liegt keine andere Entartung vor als die durch die freie Drehbarkeit des Systems bedingte, so sind die Elemente der in (4) vorkommenden Matrizen nach den Werten der magnetischen Quantenzahl m zu numerieren; ist noch eine weitere Entartung vorhanden, wie z. B. beim Wasserstoffatom, so sind weitere Indizes hinzuzufügen.

Die Gleichungen (4) stellen den *linearen Starkeffekt* dar, nämlich eine Aufspaltung jedes Terms in Komponenten, deren Abstände proportional der Feldstärke E wachsen. Dieser Effekt ist aber, wie die Erfahrung lehrt, nur beim Wasserstoffatom und bei sehr wasserstoffähnlichen Termen anderer Atome vorhanden. Der weitere Ausbau der Theorie wird hiervon Rechenschaft geben; das liegt daran, daß infolge der Symmetrie der Elektronenanordnungen in Atomen und Molekeln die einem Eigenwert des ungestörten Atoms entsprechende Untermatrix von \mathfrak{P} , also auch $\overset{\circ}{Z}^{(r,r)} = Z^{(r,r)} (\overset{\circ}{p}, \overset{\circ}{q})$ in (4), im allgemeinen verschwindet, nämlich immer dann, wenn keine andere Entartung (wie beim Wasserstoffatom) vorhanden ist. Wir werden daher hier vom linearen

Starkeffekt absehen und unsere Betrachtung auf solche Atome beschränken, für die $\overset{(1)}{W}^{(r)} = 0$ ist.

In erster Näherung wird die Entartung also gar nicht aufgehoben; die Unterteilung der Matrizen bleibt ungeändert und wir sind auf Formeln (14) u. ff. des § 41 angewiesen. Wir müssen zunächst eine Lösung $\overset{(02)}{q}_k, \overset{(02)}{p}_k$ des ungestörten Problems ausfindig machen, für welche die Matrix

$$\sum_t' \frac{\overset{(1)}{H}^{(r,t)} \overset{(1)}{H}^{(t,r)}}{\hbar \overset{\circ}{\psi}^{(r,t)}} = E^2 \sum_t' \frac{\overset{\circ}{Z}^{(r,t)} \overset{\circ}{Z}^{(t,r)}}{\hbar \overset{\circ}{\psi}^{(r,t)}}$$

diagonal ist. Es ist zu erwarten, daß die in § 28 gewählte Lösung diese Eigenschaft hat; in der Tat wird [wegen § 28, (1a)], nach der hier getroffenen Unterteilung der Matrizen:

$$(\overset{\circ}{Z}^{(r,t)} \overset{\circ}{Z}^{(t,r)})_{mm'} = \overset{\circ}{Z}_{mm}^{(r,t)} \overset{\circ}{Z}_{mm}^{(t,r)} \delta_{mm'} = |\overset{\circ}{Z}_{mm}^{(r,t)}|^2 \delta_{mm'}.$$

Nach § 41, (14) ist die Störung 2. Ordnung der Energie:

$$(5) \quad \overset{(2)}{W}_m^{(r)} = E^2 \sum_t' \frac{|\overset{\circ}{Z}_{mm}^{(r,t)}|^2}{\hbar \overset{\circ}{\psi}^{(r,t)}}.$$

Diese E^2 proportionale Energiestörung liefert eine Verschiebung und Aufspaltung der Terme, und zwar wird

$$(5a) \quad \overset{(2)}{W}_m^{(r)} = (A^{(r)} + B^{(r)} m^2) E^2.$$

Das folgt aus den Formeln § 29, (12) und (13). Ersetzt man nämlich den Index r durch das Paar n, j , wo j die innere Quantenzahl ist, so wird

$$\begin{aligned} \overset{(2)}{W}_m^{(n,j)} &= \sum_{n'}' \left(\frac{|\overset{\circ}{Z}_{mm}^{(n,j;n',j-1)}|^2}{\hbar \overset{\circ}{\psi}^{(n,j;n',j-1)}} + \frac{|\overset{\circ}{Z}_{mm}^{(n,j;n',j+1)}|^2}{\hbar \overset{\circ}{\psi}^{(n,j;n',j+1)}} + \frac{|\overset{\circ}{Z}_{mm}^{(n,j;n',j)}|^2}{\hbar \overset{\circ}{\psi}^{(n,j;n',j)}} \right) E^2 \\ &= \left\{ (j^2 - m^2) \sum_{n'}' \frac{|b^{(n,j;n',j-1)}|^2}{\hbar \overset{\circ}{\psi}^{(n,j;n',j-1)}} \right. \\ &\quad \left. + ((j+1)^2 - m^2) \sum_{n'}' \frac{|b^{\dagger(n,j;n',j+1)}|^2}{\hbar \overset{\circ}{\psi}^{(n,j;n',j+1)}} + m^2 \sum_{n'}' \frac{|a^{(n,j;n',j)}|^2}{\hbar \overset{\circ}{\psi}^{(n,j;n',j)}} \right\} E^2, \end{aligned}$$

woraus die Behauptung hervorgeht.

Dieser STARK-Effekt 2. Ordnung ist äußerst klein; doch ist er an Linien des He und Li zuerst von J. STARK¹ beobachtet worden. Auf den prinzipiellen Unterschied dieses quadratischen Effekts von dem beim Wasserstoffatom auftretenden linearen Effekt hat LADENBURG² zuerst hingewiesen, dem es gelang, eine Verschiebung der D-Linien des Natriums in hohen elektrischen Feldern zu beobachten.

Aus (5a) sehen wir, daß $W_m^{(r)} = W_{-m}^{(r)}$ ist: d. h. in dieser Näherung ist (außer für $m = 0$) die Energie noch 2-fach entartet. Wie in § 41 gesagt, ist diese Entartung nicht aufhebbar. In den meßbaren Größen (z. B. in der Dielektrizitätskonstante) treten dementsprechend nur Spuren der 2- bzw. 1-reihigen Untermatrizen auf, welche dieser Entartung entsprechen.

Wir wollen jetzt die Störung \mathfrak{P} des elektrischen Moments \mathfrak{P} berechnen. Dazu müssen wir zunächst U bestimmen; die Formeln § 41 (8), (19) liefern uns

$$(6) \quad U_{mm'}^{(1)(rs)} = -\frac{H_{mm'}^{(rs)}}{h \nu^{(rs)}} = E \cdot \frac{\tilde{Z}_{mm}^{(rs)}}{h \nu^{(rs)}} \delta_{mm'}; \quad r \neq s$$

und mit

$$h \nu_{mm'}^{(2)(r)} = W_m^{(2)(r)} - W_{m'}^{(2)(r)}$$

$$U_{mm'}^{(1)(rs)} = -\frac{1}{h \nu_{mm'}^{(2)(r)}} \sum_t' \sum_\tau \frac{H_{m\tau}^{(1)(rt)}}{h \nu^{(rt)}} \cdot \sum_{s,\sigma}'' H_{\tau\sigma}^{(1)(ts)} U_{\sigma m'}^{(1)(sr)};$$

($m' \neq m$ und $m' \neq -m$)

in der Summation $\sum_{s,\sigma}''$ sind jetzt einfach $s = t$ und $s = r$ auszulassen; da $H^{(1)(rs)}$ und [nach (6)] auch $U^{(1)(rs)}$ in bezug auf m diagonal

¹ STARK, J.: Ann. Physik Bd. 48, S. 210. 1915; ferner LÜSSEM, H.: Ann. Physik Bd. 49, S. 879, 1916; SIEBERT, G.: Ann. Physik Bd. 56, S. 587 u. 593. 1918.

² LADENBURG, R.: Phys. Z. Bd. 22, S. 549. 1921; s. ferner FOSTER J. ST.: Phys. Rev. Bd. 23, S. 667. 1924 usw.; WALLER, J.: Z. Phys. Bd. 38, S. 641. 1926; RAUSCH v. TRAUBENBERG, H. und R. GEBAUER: Z. Phys. Bd. 54, S. 304; Bd. 56, S. 255. 1929. Weitere Literatur findet sich in dem Bericht von LADENBURG: Die STARK-Effekte höherer Atome. Phys. Z. Bd. 30, S. 369. 1929.

sind, so vereinfacht sich die letzte Formel zu

$$U_{mm'}^{(1)(rr)} = - \frac{1}{\hbar \dot{\nu}_{m m'}} \sum_t' \sum_s'' \frac{1}{\hbar \dot{\nu}_{\tau m'}} \overset{(1)}{H}_{m m}^{(rt)} \overset{(1)}{H}_{m m}^{(ts)} \overset{(1)}{U}_{m m}^{(sr)} \delta_{m m'},$$

($m' \neq m, -m$)

d. h.

$$(6a) \quad U_{m m'}^{(1)(rr)} = 0 \quad \text{für} \quad m' \neq m, -m.$$

Entsprechend der Bemerkung am Ende von § 41 darf man auch

$$(6b) \quad U_{m, m}^{(1)(rr)} = 0, \quad U_{m, -m}^{(1)(rr)} = 0$$

setzen. Nach § 41, (20) haben wir dann wegen (6a), (6b):

$$\begin{aligned} \mathfrak{P}_{m m'}^{(1)(rs)} &= \sum_{t, \tau} \left(\overset{\circ}{\mathfrak{P}}_{m \tau}^{(rt)} \overset{(1)}{U}_{\tau m'}^{(ts)} - \overset{(1)}{U}_{m \tau}^{(rt)} \overset{\circ}{\mathfrak{P}}_{\tau m'}^{(ts)} \right) \\ &= \sum_t'' \left(\overset{\circ}{\mathfrak{P}}_{m m'}^{(rt)} \overset{(1)}{U}_{m' m'}^{(ts)} - \overset{(1)}{U}_{m m}^{(rt)} \overset{\circ}{\mathfrak{P}}_{m m'}^{(ts)} \right), \end{aligned}$$

und nach (6)

$$(7) \quad \mathfrak{P}_{m m'}^{(1)(rs)} = E \cdot \sum_t'' \left(\frac{\overset{\circ}{\mathfrak{P}}_{m m'}^{(rt)} \overset{\circ}{Z}_{m' m'}^{(ts)}}{\hbar \dot{\nu}_{\tau m'}} - \frac{\overset{\circ}{Z}_{m m}^{(rt)} \overset{\circ}{\mathfrak{P}}_{m m'}^{(ts)}}{\hbar \dot{\nu}_{\tau m'}} \right).$$

Aus dieser Formel schließt man zunächst, daß das induzierte Moment $\overset{(1)}{\mathfrak{P}}$ im Zeitmittel parallel der Feldrichtung liegt, d. h. daß

$$\overset{(1)}{H}_{m m}^{(rr)} = \overset{(1)}{X}_{m m}^{(rr)} + i \overset{(1)}{Y}_{m m}^{(rr)} = 0$$

ist; das folgt aus der Eigenschaft § 28, (1a) von $\overset{\circ}{H}$. Wir können übrigens auch aus (7) den Mittelwert $\overset{(1)}{Z}_{m m}^{(r r)}$ von \mathfrak{P} sofort entnehmen. Indem wir für den Index r (bzw. s, t) das in § 29 mit n, j bezeichnete Indexpaar schreiben, wo j die innere Quantenzahl und n ein weiterer, alle übrigen Quantenzahlen zusammenfassender Index ist, so haben wir

$$(8) \quad \begin{cases} \overset{(1)}{H}_{m, m}^{(n, j; n, j)} = 0, \\ \overset{(1)}{Z}_{m, m}^{(n, j; n, j)} = \alpha(n, j, m) E \end{cases}$$

mit

$$(9) \quad \alpha(n, j, m) = \sum_{n', j'}' \frac{2}{h \nu(n', j'; n, j)} \cdot |Z_{mm}^{(n, j; n', j')}|^2.$$

Man erkennt durch Vergleich von (5) und (9), daß

$$(10) \quad Z_{mm}^{(1)(n, j; n, j)} = -\frac{\partial W^{(2)(n, j)}}{\partial E}, \quad W^{(2)(n, j)} = -\frac{1}{2} \alpha(n, j, m) E^2$$

ist.

Außerdem aber folgt aus (7) eine Änderung der Intensitäten von Spektrallinien, die besonders dann sehr auffällig wird, wenn einem Übergang ohne Feld die Intensität Null (d. h. $\mathfrak{P}_{m m}^{(r, s)} = 0$) entspricht. Dann erscheinen „neue“ Linien bei Erregung des Feldes¹. Man kann Intensität und Polarisationszustand dieser Linien berechnen oder wenigstens abschätzen, indem man aus den Emissionslinien ohne Feld die ν und \mathfrak{P} empirisch bestimmt².

Aus der Größe $\alpha(n, j, m)$ läßt sich die *Dielektrizitätskonstante* ε eines Gases, das aus N Atomen oder Molekeln pro Volumeneinheit besteht, in der Weise berechnen, daß man das mittlere Moment pro Volumeneinheit

$$(11) \quad \overline{\overline{\mathfrak{P}}} = N \alpha \mathfrak{E} \quad \text{oder} \quad Z = N \alpha E, \quad \overline{\overline{\Pi}} = 0$$

bildet; dabei bedeutet das doppelte Überstreichen den *statistischen Mittelwert* von \mathfrak{P} über alle Zustände des Systems. Dann wird bekanntlich

$$(12) \quad \alpha = \frac{3}{4\pi N} \frac{\varepsilon - 1}{\varepsilon + 2}.$$

Wie wir später sehen werden, verlangt die Quantenmechanik strenggenommen eine Abänderung der gewöhnlichen statistischen Mechanik (BOSE-EINSTEINsche bzw. FERMI-DIRACsche Statistik). Wenn wir uns aber auf genügend hohe Temperaturen beschränken, können wir mit der MAXWELL-BOLTZMANNschen Statistik rechnen, nach der das Gewicht eines Zustandes der Energie $W^{(r)}$

¹ Systematische Beobachtungen sind ausgeführt von J. KOCH: Ann. Physik Bd. 48, S. 98. 1915, J. STARK ebenda S. 210, J. ST. FOSTER: a. a. O.; s. auch zit. Bericht von R. LADENBURG (Fußnote 2, S. 215).

² Dies wurde am Beispiel des Hg ausgeführt von W. PAULI: Math.-Phys. Mitt. d. dan. Ges. d. Wiss., Bd. 7, Nr. 3. 1925.

bei der Temperatur T proportional $e^{-\frac{W^{(r)}}{kT}}$ ist. Machen wir nun die Annahme, daß den völlig aufgespaltenen Zuständen $r = (n, j, m)$ gleiche a priori-Wahrscheinlichkeit zukommt, so erhalten wir

$$(13) \quad \alpha = \frac{\sum_n \sum_j \sum_m \alpha(n, j, m) e^{-\frac{W_m^{(n, j)}}{kT}}}{\sum_n \sum_j \sum_m e^{-\frac{W_m^{(n, j)}}{kT}}}.$$

In dieser Summe erhalten $\alpha(n, j, m)$ und $\alpha(n, j, -m)$ dieselben Gewichte; die Größe α ist also [im Gegensatz zu $\alpha(n, j, m)$ (9)] relativinvariant bezüglich der nicht-aufhebbaren Entartung (vgl. S. 215).

Da sich nun, auf Grund unserer Annahmen (S. 214), $W_m^{(n, j)}$ von $\dot{W}^{(n, j)}$ nur um Glieder der Ordnung E^2 [Gl. (5a)] unterscheidet, so können wir $W_m^{(n, j)}$ durch $\dot{W}^{(n, j)}$ ersetzen und die Summation nach m ausführen:

$$(14) \quad \alpha = \frac{\sum_n \sum_j \alpha(n, j) e^{-\frac{\dot{W}^{(n, j)}}{kT}}}{\sum_n \sum_j (2j+1) e^{-\frac{\dot{W}^{(n, j)}}{kT}}},$$

wo nach (9)

$$\alpha(n, j) = \sum_{m=-j}^j \alpha(n, j, m) = \sum_{n, j} \frac{2 \|\dot{Z}^{(n, j; n', j')}\|^2}{h \dot{\nu}^{(n', j'; n, j)}},$$

und nach § 31, (2)

$$(15) \quad \alpha(n, j) = \frac{2}{3} \sum_{n', j'} \frac{\|\mathfrak{P}^{(n, j; n', j')}\|^2}{h \dot{\nu}^{(n', j'; n, j)}}.$$

Wir nehmen nun an, daß die Energie im „Normalzustande“ $n = 0$ so überwiegt, daß in (14) dieser allein berücksichtigt werden muß. Schreiben wir statt $W^{(o, j)}$ kurz $\dot{W}^{(j)}$ und

$$(16) \quad \alpha(j) = \alpha(o, j) = \frac{2}{3} \sum_{n', j'} \frac{\|\mathfrak{P}^{(o, j; n', j')}\|^2}{h \dot{\nu}^{(n', j'; o, j)}},$$

so wird

$$(17) \quad \alpha = \frac{1}{S} \sum_j \alpha(j) e^{-\frac{\dot{W}^{(j)}}{kT}},$$

wo die Abkürzung

$$(18) \quad S = \sum_j (2j + 1) e^{-\frac{\dot{W}^{(j)}}{kT}}$$

gebraucht ist.

Wir können nun die Glieder der Summe (16) in zwei Klassen teilen¹, solche mit $n' = 0$ und solche mit $n' \neq 0$, die sich ganz verschieden verhalten. Im allgemeinen wird nämlich die Rotation nur eine Termänderung erzeugen, die klein ist gegen die von den inneren Vorgängen bewirkte; d. h. die Funktion $\dot{W}^{(n, j)}$ hängt wesentlich nur von n , sehr wenig aber von j ab. Daher ist für $n' \neq 0$ die Rotation nur eine kleine Störung gegenüber den durch Sprünge $0 \rightarrow n'$ erzeugten Änderungen des Systems; für $n' = 0$ aber ist die Rotation der Haupteffekt.

Wir nehmen also an, daß für $n' \neq 0$ die Energie $\dot{W}^{(n', j')}$ von $\dot{W}^{(n' j)}$ nicht wesentlich verschieden sei und ersetzen daher $\dot{\nu}^{(n', j'; 0, j)}$ einfach durch $\dot{\nu}^{(n', 0)}$. In dem Gliede $n' = 0$ schreiben wir $\dot{W}^{(j)}$ statt $\dot{W}^{(0, j)}$ und $\dot{\nu}^{(j, j')}$ statt $\dot{\nu}^{(0, j; 0, j')}$. Vorausgesetzt wird, daß zwar $\dot{W}^{(j)}$ von derselben Größenordnung ist wie die Energie der Wärmebewegung, aber die Differenz $\dot{W}^{(j)} - \dot{W}^{(j')} = h \dot{\nu}^{(j, j')}$ dem Betrage nach klein gegen kT .

Nach diesen Annahmen wird:

$$(19) \quad \alpha(j) = \frac{2}{3} \sum_{j'} \frac{\|\mathfrak{P}^{(0, j; 0, j')}\|^2}{h \dot{\nu}^{(j', j)}} + \frac{2}{3} \sum_{n'} \frac{1}{h \dot{\nu}^{(n', 0)}} \sum_{j'} \|\mathfrak{P}^{(0, j; n', j')}\|^2.$$

Setzen wir das in (17) ein, so zerfällt α in zwei Anteile,

$$(20) \quad \alpha = \alpha_0 + \alpha_1,$$

wobei α_0 nur von den Frequenzen $\dot{\nu}^{(n', 0)}$ der „schnellen“ Schwingungen, α_1 nur von denen der „langsamen“ $\dot{\nu}^{(j, j')}$ abhängt, nämlich

¹ Nach J. H. VAN VLECK: Nature Bd. 118, S. 226. 1926; Phys. Rev. Bd. 29, S. 727. 1927; Bd. 30, S. 31. 1927; Bd. 31, S. 587. 1928. Dort findet man auch ausführliche Literaturangaben. S. auch L. PAULING: Proc. Roy. Soc. Bd. 114A, S. 181. 1927.

$$(21) \quad \alpha_0 = \frac{2}{3} \cdot \frac{1}{S} \sum_{n'} \frac{1}{\hbar \hat{\nu}(n', o)} \sum_{j, j'} \|\hat{\mathfrak{P}}^{(o, j; n', j')}\|^2 e^{-\frac{\hat{W}^{(o)}}{kT}},$$

$$(22) \quad \alpha_1 = \frac{2}{3} \cdot \frac{1}{S} \sum_{j, j'} \frac{1}{\hbar \hat{\nu}(j', j)} \|\hat{\mathfrak{P}}^{(o, j; o, j')}\|^2 e^{-\frac{\hat{W}^{(j)}}{kT}}.$$

Wir wollen diese Summen zunächst für den Fall der *zwei-atomigen heteropolaren Molekel* in der Idealisierung der §§ 34, 40 ausrechnen¹. Nach § 34, (15), (16), (20) ist die Energie

$$(23) \quad \hat{W}^{(n, j)} = U_0 + \frac{\hbar^2 j(j+1)}{8\pi^2 A_0} + \hbar \nu_0 \left(n + \frac{1}{2} \right) + \dots,$$

wobei die Energie U_0 , das Trägheitsmoment $A_0 = m_0 r_0^2$ [s. § 34, (21)] und die Frequenz ν_0 in erster Näherung als Konstante behandelt werden. Man hat daher näherungsweise

$$(24) \quad \left\{ \begin{array}{l} \hat{\nu}^{(n, o)} = \nu_0 n, \quad n \neq 0 \\ \hat{\nu}^{(j+1, j)} = \frac{\hbar}{4\pi^2 A_0} (j+1), \\ \hat{\nu}^{(j-1, j)} = -\frac{\hbar}{4\pi^2 A_0} j; \end{array} \right.$$

ferner kann man

$$(25) \quad \hat{W}^{(j)} = \frac{\hbar^2 j(j+1)}{8\pi^2 A_0}$$

setzen, weil die von j unabhängigen Glieder von $\hat{W}^{(o, j)}$ aus α_0 und α_1 herausfallen. Ferner benutzen wir für die $\|\hat{\mathfrak{P}}\|^2$ die Ausdrücke § 31, (11); in diesen können wir die in § 34, (23) gegebenen Näherungswerte

$$b_{o, n}^{(j, j-1)} = b_{o, n}^{\dagger(j-1, j)} = \frac{e(r_0 \delta_{o, n} + \varrho(0, n))}{\sqrt{(2j-1)(2j+1)}}$$

einsetzen; ferner ist $a = 0$, wie für jedes Einkörperproblem. Die Größe $\varrho(0, n)$ bedeutet die Amplitude des harmonischen Oszilla-

¹ MENSING, S. L. und W. PAULI jr.: Phys. Z. Bd. 27, S. 509. 1926. Eine wellenmechanische Behandlung desselben Problems bei C. MANNEBACK: Phys. Z. Bd. 27, S. 563. 1926; Bd. 28, S. 72. 1927. GANS, R.: Phys. Z. Bd. 28, S. 309. 1927; eine Richtigstellung hierzu von C. MANNEBACK: ebenda S. 514.

tors, ist also nur für $n = 1$ von Null verschieden, und zwar nach § 23, (25 c)

$$(26) \quad |\varrho(0, 1)|^2 = \frac{h}{8\pi^2 m_0 v_0}.$$

Daher hat man

$$(27) \quad \begin{cases} b_{0,0}^{(j,j-1)} = b_{0,0}^{\dagger(j-1,j)} = \frac{e r_0}{\sqrt{(2j-1)(2j+1)}}, \\ b_{0,1}^{(j,j-1)} = b_{0,1}^{\dagger(j-1,j)} = \frac{e \varrho(0, 1)}{\sqrt{(2j-1)(2j+1)}}. \end{cases}$$

Somit wird

$$(28a) \quad \begin{cases} \|\hat{\mathfrak{P}}^{(0,j;0,j-1)}\|^2 = e^2 r_0^2 j, \\ \|\hat{\mathfrak{P}}^{(0,j;0,j+1)}\|^2 = e^2 r_0^2 (j+1), \end{cases}$$

und

$$(28b) \quad \begin{cases} \|\hat{\mathfrak{P}}^{(0,j;1,j-1)}\|^2 = e^2 |\varrho(0, 1)|^2 j, \\ \|\hat{\mathfrak{P}}^{(0,j;1,j+1)}\|^2 = e^2 |\varrho(0, 1)|^2 (j+1). \end{cases}$$

Mithin wird aus (21)

$$\alpha_0 = \frac{2}{3} \frac{1}{S h v_0} e^2 |\varrho(0, 1)|^2 \sum_j (2j+1) e^{-\frac{\hat{w}^{(j)}}{kT}},$$

oder nach der Definition (18) von S mit Rücksicht auf (26)

$$(29) \quad \alpha_0 = \frac{1}{3} \frac{e^2}{4\pi^2 m_0 v_0^2}.$$

In (22) werden die beiden Quotienten

$$\begin{aligned} \frac{\|\hat{\mathfrak{P}}^{(0,j;0,j-1)}\|^2}{h \hat{v}^{(j-1,j)}} &= - \frac{e^2 r_0^2 4\pi^2 A_0}{h^2}, \\ \frac{\|\hat{\mathfrak{P}}^{(0,j;0,j+1)}\|^2}{h \hat{v}^{(j+1,j)}} &= + \frac{e^2 r_0^2 4\pi^2 A_0}{h^2}, \end{aligned}$$

heben sich also in der Summe sämtlich fort außer für $j = 1$, bei dem nur der positive übrig bleibt. Also wird

$$(30) \quad \alpha_1 = \frac{2}{3} \frac{1}{S} \frac{4\pi^2 e^2 r_0^2 A_0}{h^2}.$$

Setzt man nun

$$(31) \quad \sigma = \frac{\hbar^2}{8\pi^2 A_0 k T}, \quad \mu_0 = e r_0,$$

so kann man (30) schreiben

$$(32) \quad \alpha_1 = \frac{\mu_0^2}{3kT} \frac{1}{f(\sigma)}$$

mit

$$(33) \quad f(\sigma) = \sigma \sum_{j=0}^{\infty} (2j+1) e^{-\sigma j(j+1)}.$$

Unser Resultat ist also, daß die *Polarisierbarkeit eines Gases, das aus zweiatomigen, idealisierten heteropolaren Molekeln besteht*, die Form

$$(34) \quad \alpha = \alpha_0 + \frac{\mu_0^2}{3kT} \frac{1}{f(\sigma)}$$

hat. Der Unterschied gegen die von der klassischen Theorie gelieferte DEBYESche Formel¹

$$(35) \quad \alpha = \alpha_0 + \frac{\mu_0^2}{3kT}$$

besteht in dem Faktor $f(\sigma)^{-1}$; denn das erste Glied in (34) stimmt nach (29) genau mit dem entsprechenden in (35) überein, wo es die mittlere Polarisierbarkeit eines *linearen*, quasielastischen Oszillators bedeutet².

Wir zeigen nun, daß für hohe Temperaturen, also kleine σ , dieser Faktor gegen 1 strebt. Für $\sigma \ll 1$ kann man nämlich in (33) die Summe durch ein Integral ersetzen und hat asymptotisch

$$f(\sigma) \rightarrow \int_0^{\infty} \sigma (2j+1) e^{-\sigma j(j+1)} dj = \int_0^{\infty} e^{-x} dx = 1.$$

Daher geht die Formel (34) für $T \rightarrow \infty$ genau in die korrespondierende klassische Formel (35) über. Diese Tatsache war einer der ersten „praktischen“ Erfolge der neuen Quantenmechanik. Denn

¹ DEBYE, P.: Phys. Z. Bd. 13, S. 97. 1912.

² Bildet das elektrische Feld den Winkel ϑ mit der Molekelachse, so ist die Vergrößerung ϱ des Atomabstandes gegeben durch $m_0 (2\pi\nu_0)^2 \varrho = eE \cos \vartheta$; die zum Felde parallele Komponente des elektrischen Moments ist im Mittel

$$e \varrho \cos \vartheta = \cos^2 \vartheta \frac{e^2}{4\pi^2 \nu_0^2 m_0} E = \frac{1}{3} \frac{e^2}{4\pi^2 \nu_0^2 m_0} E.$$

die Methoden der älteren Quantentheorie¹, bei denen eine „Richtungsquantelung“ der Molekelachsen vorgenommen wurde, führen bei hohen Temperaturen auf einen falschen Zahlenfaktor.

Man kann nun zeigen, daß für hohe Temperaturen die DEBYE'sche Formel nicht nur für das vereinfachte Modell der zweiatomigen Molekel, sondern für beliebige Systeme gilt, denen ein „permanentes“ elektrisches Moment zugeschrieben werden kann.

Wir kehren also wieder zu den allgemeinen Formeln (21), (22) zurück und entwickeln die Summe in (22) für große T in folgender Weise: Zunächst vertauschen wir die Indizes j, j' und nehmen die halbe Summe der beiden entstehenden Ausdrücke. Da $\|\hat{\mathfrak{P}}\|^2$ in j, j' symmetrisch und $\hat{v}^{(j', j)}$ antisymmetrisch ist, erhält man als Faktor die Differenz

$$e^{-\frac{\hat{W}^{(j')}}{kT}} - e^{-\frac{\hat{W}^{(j)}}{kT}},$$

die man nach Potenzen von $\frac{1}{kT}(\hat{W}^{(j)} - \hat{W}^{(j')}) = \frac{h\hat{v}^{(j, j')}}{kT}$ entwickeln

kann; sie ist also in erster Näherung durch

$$e^{-\frac{\hat{W}^{(j)}}{kT}} \cdot \frac{h\hat{v}^{(j, j')}}{kT}$$

zu ersetzen. Dann wird

$$(36) \quad \alpha_1 = \frac{1}{3SkT} \sum'_{j, j'} \|\hat{\mathfrak{P}}^{(0, j; 0, j')}\|^2 e^{-\frac{\hat{W}^{(j)}}{kT}}.$$

Definiert man nun allgemein:

$$(37) \quad \mu_n(T) = \frac{1}{S} \sum'_{j, j'} \|\hat{\mathfrak{P}}^{(0, j; n, j')}\|^2 e^{-\frac{\hat{W}^{(j)}}{kT}},$$

wobei für $n = 0$ das Glied $j' = j$ fortzulassen ist, so wird

$$(38) \quad \left\{ \begin{array}{l} \alpha_0 = \frac{2}{3} \sum'_n \frac{\mu_n(T)^2}{h\hat{v}^{(n, 0)}}, \\ \alpha_1 = \frac{\mu_0(T)^2}{3kT}. \end{array} \right.$$

¹ PAULI, W. jr.: Z. Phys. Bd. 6, S. 319. 1921.

Die Ableitung der DEBYESchen Formel (35) läuft also darauf hinaus, die Bedingungen anzugeben, unter denen die Größe (37) von T merklich unabhängig wird. Dazu brauchen wir die Größen $\|\mathfrak{K}\|^2$, die in § 31, (11) angegeben sind. Es ergibt sich:

$$(39) \quad \mu_n(T)^2 = \frac{1}{S} \sum_j (2j+1) e^{-\frac{W(j)}{kT}} \{ j(j+1) |a_{o,n}^{(j,j)}|^2 \\ + j(2j-1) |b_{o,n}^{(j,j-1)}|^2 + (j+1)(2j+3) |b_{o,n}^{(j,j+1)}|^2 \},$$

wobei für $n=0$ das Glied mit a fortzulassen ist.

Man sieht leicht, daß im Falle der idealisierten zweiatomigen Molekel $\mu_n(T)$ von T unabhängig wird. Setzt man nämlich die Ausdrücke (27) nebst $a=0$ ein, so wird die Klammer gleich $e^2 r_0^2$ für $n=0$, gleich $e^2 |Q(0,1)|^2$ für $n=1$ und gleich Null für $n=2, 3, \dots$; die übrigbleibende Summe aber ist gerade gleich S , also wird

$$\mu_0^2 = e^2 r_0^2, \quad \mu_1^2 = e^2 |Q(0,1)|^2, \quad \mu_2 = \mu_3 = \dots = 0.$$

Setzt man das in (38) ein, so erhält man wieder die Formel (35) mit denselben Konstanten, die oben in (29), (31) angegeben wurden.

Hieraus erkennt man, daß allgemein μ_n^2 von T unabhängig wird, wenn

$$(40) \quad \begin{cases} j(j+1) |a_{o,n}^{(j,j)}|^2 = \alpha_n^2, & n = 1, 2, \dots \\ j(2j-1) |b_{o,n}^{(j,j-1)}|^2 + (j+1)(2j+3) |b_{o,n}^{(j,j+1)}|^2 = \beta_n^2, \\ n = 0, 1, 2, \dots \end{cases}$$

näherungsweise von j unabhängig sind.

Das bedeutet offenbar korrespondenzmäßig, daß die Komponenten des elektrischen Moments parallel und senkrecht zur „Rotationsachse“ durch die Rotation nicht merklich verändert werden — eine Annahme, die auch den Näherungsformeln (27) bei der zweiatomigen Molekel zugrunde lag. In diesem Falle kann man von der Existenz eines *permanenten elektrischen Moments* sprechen, und zwar ergibt sich für dies aus (39) und (40)

$$(41a) \quad \mu_0^2 = \beta_0^2;$$

ferner wird

$$(41b) \quad \mu_n^2 = \alpha_n^2 + \beta_n^2.$$

Setzt man das in (38) ein, so ergibt sich wieder die DEBYESche Formel (35).

§ 43. ZEEMAN-Effekt und Magnetisierung. Wir haben bereits früher (§ 32) gesehen, daß die Übertragung der klassischen Formeln für den Einfluß eines Magnetfeldes auf Atome zu falschen Ergebnissen bezüglich der Zahl und Anordnung der Linien im ZEEMAN-Effekt führt und haben angedeutet, daß diese Schwierigkeit durch die Hypothese des Magnetelektrons überwunden werden kann. Wir wollen hier die elementare Theorie des ZEEMAN-Effektes und der damit zusammenhängenden magnetischen Erscheinungen als Beispiel zur Störungsrechnung durchführen, obwohl die Resultate durch den späteren Ausbau der Theorie wesentlich modifiziert werden.

Hierzu übernehmen wir den Ausdruck der Energiefunktion im Magnetfelde aus der klassischen Theorie, wie er in Bd. 1, § 34 angegeben worden ist. Dort betrachteten wir Atome, deren Kern als *fest* angenommen wird, bezeichneten die Ladung der beweglichen Elektronen mit $-e$, ihre Masse mit μ . Das Magnetfeld \mathfrak{H} werde aus dem Vektorpotential \mathfrak{A} abgeleitet nach der Formel

$$(1) \quad \mathfrak{H} = \text{rot } \mathfrak{A},$$

wobei \mathfrak{A} die Bedingung

$$(2) \quad \text{div } \mathfrak{A} = 0$$

erfüllt.

Die Energiefunktion der klassischen Theorie, ausgedrückt durch Koordinaten und Impulse, ist nach Bd. 1, § 34, S. 239:

$$(3) \quad H = H_0 + \sum_k \left(\frac{e}{c\mu} p_k \mathfrak{A}_k + \frac{e^2}{2\mu c^2} \mathfrak{A}_k^2 \right),$$

wo

$$(4) \quad H_0 = \sum_k \frac{1}{2\mu} p_k^2 + U$$

die Energie des ungestörten Systems (U seine potentielle Energie) ist und \mathfrak{A}_k den Wert des Vektors \mathfrak{A} an der Stelle des k -ten Teilchens bedeutet.

Wir übernehmen den Ausdruck (3) in die Quantenmechanik, wobei die Koordinaten x_k, y_k, z_k (Vektor \mathfrak{r}_k) und Impulse p_{kx}, p_{ky}, p_{kz} (Vektor \mathfrak{p}_k) den kanonischen Vertauschungsregeln ge-

nügen sollen; allerdings müssen wir beachten, daß das skalare Produkt $\mathfrak{p}_k \mathfrak{A}_k$ nicht unabhängig von der Reihenfolge der Faktoren zu sein scheint. Tatsächlich ist aber die Reihenfolge gleichgültig; denn man hat nach § 18, (3)

$$p_{kx} \mathfrak{A}_{kx} - \mathfrak{A}_{kx} p_{kx} = \kappa \frac{\partial \mathfrak{A}_{kx}}{\partial x_k},$$

also wegen (2)

$$(5) \quad \mathfrak{p}_k \mathfrak{A}_k - \mathfrak{A}_k \mathfrak{p}_k = \kappa \operatorname{div}_k \mathfrak{A}_k = 0.$$

Die kanonischen Bewegungsgleichungen liefern erstens

$$\dot{x}_k = v_{kx} = \frac{\partial H}{\partial p_{kx}} = \frac{1}{\mu} p_{kx} + \frac{e}{c\mu} \mathfrak{A}_{kx},$$

oder

$$(6) \quad \mu v_k = \mathfrak{p}_k - \frac{e}{c} \mathfrak{A}_k;$$

zweitens

$$\dot{p}_{kx} = - \frac{\partial U}{\partial x_k} - \frac{e}{c} \left(v_{kx} \frac{\partial \mathfrak{A}_{kx}}{\partial x_k} + v_{ky} \frac{\partial \mathfrak{A}_{ky}}{\partial x_k} + v_{kz} \frac{\partial \mathfrak{A}_{kz}}{\partial x_k} \right)$$

und daraus folgt in Verbindung mit (6)

$$\mu \dot{v}_k = - \operatorname{grad}_k U + \frac{e}{c} (v_k \times \mathfrak{H}_k).$$

Der Ansatz (3) für die HAMILTONSche Funktion ist also äquivalent mit der Annahme, daß die mechanische Wirkung des Magnetfeldes durch die LORENTZsche Kraft $\frac{e}{c} (v \times \mathfrak{H})$ dargestellt wird. Das *magnetische Moment* ist in der Elektrodynamik¹ definiert als der Vektor

$$m = - \frac{e}{2c} \sum_k (\mathbf{r}_k \times \mathbf{v}_k)$$

oder, wenn man mit (6) v_k durch \mathfrak{p}_k ersetzt und das mechanische Impulsmoment $\mathfrak{M} = \sum_k (\mathbf{r}_k \times \mathfrak{p}_k)$ einführt:

$$(7) \quad m = - \frac{e}{2c\mu} \mathfrak{M} + \frac{e^2}{2\mu c^2} \sum_k (\mathbf{r}_k \times \mathfrak{A}_k).$$

¹ S. etwa ABRAHAM, M.: Theorie der Elektrizität Bd. 2, 3. Aufl. § 25, Formel (137e), S. 207. Leipzig: B. G. Teubner 1914.

Insbesondere soll nun ein homogenes Magnetfeld untersucht werden; es sei parallel zur z -Achse und habe die Stärke $|\mathfrak{H}|$. Dann ist (1) erfüllt für

$$(8) \quad \mathfrak{A}_{kx} = -\frac{1}{2} y_k |\mathfrak{H}|, \quad \mathfrak{A}_{ky} = \frac{1}{2} x_k |\mathfrak{H}|, \quad \mathfrak{A}_{kz} = 0.$$

Führt man ferner das BOHRsche Magneton

$$(9) \quad M_0 = \frac{e}{2\mu c} \frac{\hbar}{2\pi} = 9,21 \cdot 10^{-21} \text{ Gauß} \times \text{cm}$$

ein¹, so wird

$$(10) \quad H = H_0 + M_0 \frac{2\pi}{\hbar} M_z \cdot |\mathfrak{H}| + \frac{1}{8\mu c^2} \Theta \cdot \mathfrak{H}^2,$$

wo

$$M_z = \sum_k (x_k p_{ky} - y_k p_{kx})$$

die dem Felde parallele Komponente des Drehimpulses \mathfrak{M} und

$$(11) \quad \Theta = \sum_k e^2 (x_k^2 + y_k^2)$$

das „elektrische Trägheitsmoment“ um die Feldrichtung ist.

Das magnetische Moment wird nach (7) und (8)

$$(12) \quad \begin{cases} m_x = -\frac{e}{2\mu c} M_x - \frac{e^2}{4\mu c^2} \sum_k x_k z_k |\mathfrak{H}|, \\ m_y = -\frac{e}{2\mu c} M_y - \frac{e^2}{4\mu c^2} \sum_k y_k z_k |\mathfrak{H}|, \\ m_z = -\frac{e}{2\mu c} M_z + \frac{1}{4\mu c^2} \Theta |\mathfrak{H}|. \end{cases}$$

Bei der Anwendung der Störungstheorie auf die Energiefunktion (10) hat man zu beachten, daß M_z ein Integral der ungestörten Bewegung (mit H_0 vertauschbar) ist; man hat also den am Ende von § 38 behandelten Fall vor sich.

Man wird

$$(13) \quad \overset{\circ}{H} = H_0 + M_0 \frac{2\pi}{\hbar} M_z \cdot |\mathfrak{H}|$$

¹ Durch Multiplikation mit der LOSCHMIDT'schen Zahl $L = 6,06 \cdot 10^{23}$ für das Mol erhält man das Molmagneton

$$LM_0 = 5584 \text{ Gauß} \times \text{cm pro Mol.}$$

als Energie der ungestörten Bewegung ansehen und

$$(14) \quad H = \frac{\Theta}{8\mu c^2} \mathfrak{H}^2$$

als Störungsfunktion zweiter Ordnung, während $H = 0$ ist. Bringt man \hat{H} auf die Diagonalform, so erhält man

$$(15) \quad \hat{H} = \hat{W}(n, j, m) = W_0(n, j) + M_0 m |\mathfrak{H}|,$$

wo j die innere, m die magnetische Quantenzahl und n eine Zusammenfassung aller übrigen Quantenzahlen ist.

Die Formel (15) zeigt, daß bei Annahme eines punktförmigen Elektrons (also ganzzahligen m , s. § 32) bei schwachen Feldern stets der sogenannte *normale ZEEMAN-Effekt* herauskommt, im Widerspruch zur Erfahrung. Nach der Störungsrechnung § 38, (31) wird

$$(16) \quad \overset{(2)}{W}(n, j, m) = \overset{(2)}{H}(n, j; n j)(m, m) = \frac{\Theta(n, j, m)}{8\mu c^2} \mathfrak{H}^2,$$

wo $\Theta(n, j, m)$ das Diagonalelement der durch (11) definierten Matrix ist; dabei ist vorausgesetzt, daß keine Entartung des Systems im Magnetfelde mehr vorhanden ist, da andernfalls auf (10) die Methoden des § 41 anzuwenden wären.

Man hat also die *Gesamtenergie*

$$(17) \quad W(n, j; n, j) = W_0(n, j) + M_0 m |\mathfrak{H}| + \frac{\Theta(n, j, m)}{8m_0 c^2} \mathfrak{H}^2.$$

Wir berechnen nun die zeitlichen Mittelwerte (Diagonalelemente) der Komponenten des magnetischen Moments (12); nach § 38, (32) erhält man die Störungen dieser Größen, indem man einfach die Glieder erster Ordnung in $|\mathfrak{H}|$ für die ungestörte Bewegung nimmt. Das bedeutet aber, daß man in (12) die der Störung durch ein Feld parallel zur z -Achse angepaßten ungestörten Koordinaten einzusetzen und von den entstehenden Matrizen die Diagonalelemente zu nehmen hat. Nun sind die Diagonalelemente von M_x und M_y gleich Null. Dasselbe gilt nach den Sätzen von § 3) von den Diagonalelementen der Größen $\sum_k x_k z_k$, $\sum_k y_k z_k$, da auch das gestörte System um die z -Achse freidrehbar ist. Also verschwinden die Zeitmittelwerte von m_x und m_y , wie es der Symmetrie des Vorgangs entspricht. *Das Diagonal-*

element von m_z bezeichnen wir mit $M(n, j, m)$ und erhalten aus (12) mit Rücksicht auf (9)

$$(18) \quad M(n, j, m) = -M_0 m - \frac{\Theta(n, j, m)}{4\mu c^2} |\mathfrak{H}|.$$

Der Vergleich mit (17) zeigt, daß

$$(19) \quad M(n, j, m) = -\frac{\partial W(n, j, m)}{\partial |\mathfrak{H}|}$$

ist, genau wie in der klassischen Theorie des Magnetismus.

Ein Gas, das aus solchen Atomen zusammengesetzt ist, hat (bei hinreichend hohen Temperaturen — damit die BOLTZMANNsche Statistik anwendbar ist) ein mittleres magnetisches Moment pro Atom

$$(20) \quad \bar{M} = \frac{\sum M(n, j, m) e^{-\frac{W(n, j, m)}{kT}}}{\sum e^{-\frac{W(n, j, m)}{kT}}}.$$

In allen praktisch wichtigen Fällen, sogar bei Feldstärken bis etwa 10^6 Gauß, ist in der Energie (17) der mit \mathfrak{H}^2 proportionale Anteil immer gegen die niederen Glieder zu vernachlässigen; zerlegt man sodann \bar{M} in die beiden Anteile, die der Formel (18) entsprechen,

$$(21) \quad M = M_p + M_a,$$

so wird \bar{M}_a , das den Faktor $|\mathfrak{H}|$ hat, relativ klein sein, man kann also bei der Berechnung von \bar{M}_a in der Energie auch die Glieder mit $|\mathfrak{H}|$ fortlassen. So erhält man (indem man m durch $-m$ ersetzt):

$$(22) \quad \bar{M}_p = M_0 \frac{\sum m e^{-\frac{W_0 - M_0 |\mathfrak{H}| m}{kT}}}{\sum e^{-\frac{W_0 - M_0 |\mathfrak{H}| m}{kT}}},$$

$$(23) \quad \bar{M}_a = \chi_a |\mathfrak{H}|, \quad \chi_a = -\frac{1}{4\mu c^2} \frac{\sum \Theta e^{-\frac{W_0}{kT}}}{\sum e^{-\frac{W_0}{kT}}}.$$

Hier bedeutet \bar{M}_p das paramagnetische, \bar{M}_a das diamagnetische Moment. Wir behandeln diese jetzt einzeln.

I. Paramagnetismus.

Setzt man

$$(24) \quad S(j) = \sum_{m=-j}^j e^{\beta m} = \frac{e^{\beta j} - e^{-\beta(j+1)}}{1 - e^{-\beta}}, \quad \beta = \frac{M_0 |\mathfrak{S}|}{kT},$$

ferner

$$(25) \quad M(j) = M_0 \frac{\sum_{m=-j}^j m e^{\beta m}}{\sum_{m=-j}^j e^{\beta m}} = M_0 \frac{d \ln S(j)}{d\beta},$$

so hat man

$$(26) \quad \bar{M}_p = \frac{\sum_{j,n} M(j) S(j) e^{-\frac{W_0(n,j)}{kT}}}{\sum_{j,n} S(j) e^{-\frac{W_0(n,j)}{kT}}}.$$

In praxi genügt es wohl stets, für n den Wert zu nehmen, der dem Normalzustande entspricht, etwa $n = 0$; in vielen Fällen kann man auch bei j sich auf den niedersten Wert j_0 beschränken (wenn der nächste j -Wert bereits einer wesentlich höheren Energie entspricht). Man hat dann näherungsweise

$$(27) \quad \bar{M}_p = M(j_0)$$

und nennt j_0 die *Magnetonenzahl des Normalzustandes*.Die Ausrechnung der Formel (25) gibt¹

$$(28) \quad M(j) = M_0 \left\{ \frac{j e^{\beta j} + (j+1) e^{-\beta(j+1)}}{e^{\beta j} - e^{-\beta(j+1)}} - \frac{1}{e^{\beta} - 1} \right\}.$$

Ist j groß, so daß $j+1$ durch j ersetzt werden kann, so hat man näherungsweise

$$M(j) \rightarrow M_0 \left\{ j \mathfrak{C} \operatorname{tg} \beta j - \frac{1}{e^{\beta} - 1} \right\};$$

ist zugleich β klein, so wird

$$(29) \quad M(j) \rightarrow M_0 j \left\{ \mathfrak{C} \operatorname{tg} \beta j - \frac{1}{\beta j} \right\}.$$

Das ist die bekannte *Formel von LANGEVIN*², die dieser aus¹ BRILLOUIN, L.: J. Phys. Bd. 8, S. 74. 1927.² LANGEVIN, P.: Ann. de chim. et Phys. (8), Bd. 5, S. 70. 1905.

der klassischen Theorie erhalten hat, wo sie streng für alle β gilt; dann würde das paramagnetische Moment mit wachsender Feldstärke, d. h. $\beta \rightarrow \infty$, dem Grenzwert $M_0 j$, der „Sättigung“, zustreben. Dasselbe Verhalten liefert auch die strenge quantentheoretische Formel (28), und zwar mit demselben *Sättigungswert*

$$(30) \quad |\mathfrak{H}| \rightarrow \infty: M(j)_\infty \rightarrow M_0 j.$$

Im Sättigungszustand verhält sich also das magnetische Moment so, als wären j BOHRsche Magnetonen M_0 pro Atom gleichgerichtet.

Für *kleine Feldstärken*, $\beta \rightarrow 0$, hat man angenähert

$$(31) \quad |\mathfrak{H}| \rightarrow 0: M(j) \rightarrow \chi_p(j) \cdot |\mathfrak{H}|,$$

mit der *paramagnetischen Anfangssuszeptibilität*

$$(32) \quad \chi_p(j) = \frac{M_0^2 j(j+1)}{kT} \frac{1}{3} = \frac{M(j)_\infty^2 j+1}{3kT} \frac{1}{j}.$$

Dies ist das CURIESche Gesetz:

$$\chi_p = \frac{C}{kT},$$

wobei die CURIESche *Konstante* C den Wert

$$(32a) \quad C = \frac{M_0^2 j(j+1)}{3} = \frac{M(j)_\infty^2}{3} \cdot \frac{j+1}{j}$$

erhält.

Für große j geht dieser Ausdruck in den korrespondierenden $\frac{M(j)_\infty^2}{3}$ der klassischen Theorie über. Eine experimentelle Entscheidung darüber, ob der exakte Ausdruck (32a) für kleine Magnetonenzahlen des Normalzustandes (kleine j_0) dem klassischen überlegen ist, hat sich bisher noch nicht treffen lassen¹.

Der Paramagnetismus tritt niemals rein auf, sondern wird immer vom Diamagnetismus überlagert. Dieser ist allerdings bei

¹ Messungen über einen großen Temperaturbereich liegen nur für Gadoliniumsulfat vor (vgl. MARX' Handbuch der Radiologie, Bd. 6, S. 683). Sie zeigen, daß das CURIESche Gesetz hier genau bis zu sehr tiefen Temperaturen gilt; doch scheint die Genauigkeit nicht auszureichen, um den Faktor $\frac{j+1}{j}$ sicher von 1 zu unterscheiden.

normalen Feldstärken sehr schwach; bei hohen Feldstärken (Größenordnung 10^7 Gauß) aber würde sich der Diamagnetismus bemerkbar machen, und zwar auf zwei Weisen: einmal direkt durch den additiven Anteil \overline{M}_d , der in $|\mathfrak{H}|$ von erster Ordnung ist, sodann durch die vom Faktor $e^{-\frac{\Theta \mathfrak{H}^2}{8\mu c^2 k T}}$ hervorgerufene Korrektur, die in $|\mathfrak{H}|$ von zweiter Ordnung ist. Wegen dieses Faktors können daher die Formeln (21), (22), (23) nicht den Anspruch erheben, den Verlauf der Magnetisierung paramagnetischer Substanzen für *alle* Feldstärken in Strenge darzustellen; doch reicht sie in allen praktisch vorkommenden Fällen aus. Der Diamagnetismus hat somit zur Folge¹, daß bei paramagnetischen Körpern keine wirkliche Sättigung zu erwarten ist, sondern bei sehr großen Feldstärken eine Abnahme der Magnetisierung — ein Effekt, der noch nicht beobachtet ist (erforderliche Feldstärke $\sim 10^7$ Gauß).

II. Diamagnetismus.

Wenn bei einem Körper der Paramagnetismus fehlt, d. h. wenn der Wert der inneren Quantenzahl im Normalzustand $j_0 = 0$ ist, so tritt der Diamagnetismus rein zutage; die diamagnetische Suszeptibilität ist nach (23)

$$(33) \quad \chi_d = \frac{\sum_{n,j} \chi_d(n,j) e^{-\frac{W_0(n,j)}{k T}}}{\sum_{n,j} (2j+1) e^{-\frac{W_0(n,j)}{k T}}}$$

mit

$$(34) \quad \chi_d(n,j) = -\frac{1}{4\mu c^2} \sum_{m=-j}^j \Theta(n,j,m).$$

Im allgemeinen ist χ_d sehr wenig von der Temperatur abhängig, was darin seinen Grund hat, daß der Energiewert im Normalzustand $W_0(0,0)$ sehr viel tiefer liegt als für alle anderen Zustände; dann stellt $\chi_d(0,0)$ direkt die diamagnetische Suszeptibilität dar. Es gibt aber auch Ausnahmen von diesem Verhalten²; bei diesen Substanzen wird man zwar $n = 0$ setzen, aber

¹ Dies wurde von R. GANS bemerkt. Gött. Nachr. 1910, S. 197; 1911, S. 118.

² S. etwa Handbuch der Physik von GEIGER und SCHEEL, Bd. 15, A, Kap. 3. STEINHAUS, W.: Magnetische Eigenschaften der Körper, Ziff. 10, S. 155.

verschiedene j -Werte berücksichtigen. Dann hat man

$$(35) \quad \chi_a = \frac{\sum_j \chi_a(j) e^{-\frac{W_0(j)}{kT}}}{\sum_j (2j+1) e^{-\frac{W_0(j)}{kT}}}$$

$$(36) \quad \chi_a(j) = -\frac{1}{4\mu c^2} \sum_{m=-j}^j \Theta(0, j, m).$$

Da $\Theta(0, j, m)$ das Diagonalglied der Matrix Θ bedeutet, so ist die in (36) auftretende Summe gleich $Sp \Theta^{(j,j)}$. Nach (11) aber ist Θ eine Komponente des Tensors „elektrisches Trägheitsmoment“

$$\Theta_{xx} = \sum_k e^2 (y_k^2 + z_k^2), \dots \Theta_{yz} = \sum_k e_k^2 y_k z_k, \dots$$

Nach § 30, (11) ist also

$$\begin{aligned} Sp \Theta^{(j,j)} &= \frac{1}{3} Sp (\Theta_{xx}^{(j,j)} + \Theta_{yy}^{(j,j)} + \Theta_{zz}^{(j,j)}) \\ &= \frac{2}{3} Sp \sum_k e^2 (x_k^2 + y_k^2 + z_k^2)^{(j,j)} = \frac{2e^2}{3} Sp \left(\sum_k r_k^2 \right)^{(j,j)}. \end{aligned}$$

Mithin wird

$$(37) \quad \chi_a(j) = -\frac{e^2}{6\mu c^2} Sp \left(\sum_k r_k^2 \right)^{(j,j)}.$$

Diese Formel ist die quantenmechanische Übertragung der von PAULI¹ auf Grund der klassischen Theorie abgeleiteten.

Nimmt man an, daß im Grundzustand $j = 0$ die Größe $(\sum_k r_k^2)^{(0,0)}$ von der Ordnung des Quadrats der atomaren Dimensionen 10^{-8} cm ist, so kann man die Größenordnung von χ_a abschätzen. Multipliziert man χ_a mit der LOSCHMIDTSchen Zahl $L = 6,06 \cdot 10^{23}$ für das Mol, so erhält man für die Suszeptibilität pro Mol die Größenordnung $3 \cdot 10^{-6}$; während ältere Messungen² für *He* und *A* 10 bis 100 mal so große Werte ergaben, sind neuer-

¹ PAULI, W. jr.: Z. Phys. Bd. 2, S. 201. 1920.

² TÄNZLER, P.: Ann. Phys. Bd. 24, S. 931. 1907. TAKÉ SONÉ: Phil. Mag. Bd. 39, S. 305. 1920.

dings von WILLS und HECTOR¹ folgende Zahlen gefunden worden:

$$\text{für He } 1,88 \cdot 10^{-6} \text{ und für } H_2 \text{ } 3,94 \cdot 10^{-6}.$$

Diese stimmen sehr gut mit der berechneten Größenordnung überein.

§ 44. Verhalten des H -Atoms in gekreuzten Feldern. Wir wollen die in § 41 entwickelte Störungsrechnung für entartete Systeme nunmehr dazu anwenden, das Verhalten des in §§ 35 bis 37 betrachteten H -Atommodells in gekreuzten Feldern zu untersuchen und die schon in § 37 gemachten diesbezüglichen Angaben zu rechtfertigen.

Ist \mathfrak{E} die elektrische Feldstärke und \mathfrak{A} das Vektorpotential des magnetischen Feldes, so wird die HAMILTON-Funktion unseres Systems entsprechend § 35, (1) und § 43, (3) gleich

$$(1) \quad \begin{cases} H = H_0 - e(\mathfrak{r} \mathfrak{E}) + \frac{e}{c\mu} (\mathfrak{p} \mathfrak{A}) \\ = \frac{1}{2\mu} \mathfrak{p}^2 - \frac{\varepsilon}{r} - e(\mathfrak{r} \mathfrak{E}) + \frac{e}{c\mu} (\mathfrak{p} \mathfrak{A}); \quad \varepsilon = e^2 Z. \end{cases}$$

Wir haben hier die Kernmasse als unendlich und demnach $m_0 = \mu$ angenommen [vgl. § 35, (1')]; ferner haben wir den mit \mathfrak{A}^2 proportionalen Teil der Störungsenergie *vernachlässigt*. Wir nehmen die Feldstärken \mathfrak{E} , \mathfrak{H} als räumlich konstant (homogen) an; gemäß (10), § 43 können wir dann schreiben:

$$(1') \quad H = H_0 - e(\mathfrak{r} \mathfrak{E}) + \frac{e}{2\mu c} (\mathfrak{M} \mathfrak{H}).$$

Nach den Feststellungen von § 41 haben wir nun zur Durchführung der Störungsrechnung eine solche Lösung des *ungestörten* Problems zu suchen, für welche die *Zeitmittelwerte* der zu \mathfrak{E} parallelen Komponente des Ortsvektors \mathfrak{r} und der zu \mathfrak{H} parallelen Komponente des Drehimpulses \mathfrak{M} *Diagonalmatrizen* sind.

Nach § 36 gehört je ein irreduzibles Teilsystem $\mathfrak{U}_n, \mathfrak{M}_n$ des Matrixsystems $\mathfrak{U}, \mathfrak{M}$ (der Matrixvektor \mathfrak{U} ist nicht zu verwechseln mit dem Vektorpotential des Magnetfeldes) zu einem Eigenwert W_n der ungestörten Energie H_0 . Schreiben wir nun auch \mathfrak{r}

¹ WILLS, A. P. and L. G. HECTOR: Phys. Rev. Bd. 23, 209. 1926.
HECTOR, L. G.: ebenda Bd. 24, S. 418. 1926.

entsprechend den Vielfachheiten der Eigenwerte W_n als Übermatrix, so wird der Zeitmittelwert \bar{r} eine *Stufenmatrix*. Es sei \bar{r}_n die zur Stufe $\mathfrak{A}_n, \mathfrak{M}_n$ von $\mathfrak{A}, \mathfrak{M}$ gehörige Stufe von \bar{r} ; und $\bar{x}_n, \bar{y}_n, \bar{z}_n$ seien die Komponenten von \bar{r}_n . Es stimmt aber jedes in \bar{r}_n überhaupt enthaltene Matrixelement mit dem entsprechenden Element von r überein, und wir können deshalb die für die Vertauschung von x, y, z mit M_x, M_y, M_z geltenden Gleichungen § 25, (2), (3) und die zwischen den Vertauschungen von x, y, z mit A_x, A_y, A_z bestehende Beziehung § 35, (19) auch für die $\bar{x}_n, \bar{y}_n, \bar{z}_n$ und $\mathfrak{M}_n, \mathfrak{A}_n$ anwenden. Also ist mit $\bar{\xi}_n = \bar{x}_n + i \bar{y}_n$:

$$(2) \quad \begin{cases} [M_z, \bar{z}_n] = 0, \\ [M_z, \bar{\xi}_n] = i \bar{\xi}_n, \end{cases}$$

und ferner, unter Berücksichtigung von Formel (V), (16), § 35:

$$(3) \quad (\mathfrak{A}_n \times \bar{r}_n) + (\bar{r}_n \times \mathfrak{A}_n) = 3 \frac{\varepsilon}{2 |W_n|} (\mathfrak{A} \times \mathfrak{A}).$$

Endlich folgt aus $(r\mathfrak{M}) = 0$ auch

$$(4) \quad (\bar{r}_n \mathfrak{M}_n) = 0.$$

Die Beziehungen (2) und (4) für \bar{r}_n sind nun dieselben, wie § 35, (III), (14) und (IV), (15) für \mathfrak{A}_n . Zusammen mit (3) liefern sie ausreichend viele Gleichungen, um die Übereinstimmung¹

$$(5) \quad \bar{r}_n = \frac{3}{2} a_n \mathfrak{A}_n; \quad a_n = \frac{\varepsilon}{2 |W_n|}$$

zu begründen; das zeigt leicht eine Betrachtung der in § 36 entwickelten expliziten Formeln für die $\mathfrak{A}_n, \mathfrak{M}_n$.

Mit der Gewinnung der Formel (5) ist unsere Aufgabe aber im wesentlichen erledigt. Wir bekommen für den *Zeitmittelwert der Störungsenergie* (1), (1') den Ausdruck

$$(6) \quad \overline{H}_1(n) = \frac{3}{2} e a_n (\mathfrak{E} \mathfrak{A}_n) + \frac{e}{2 \mu c} (\mathfrak{S} \mathfrak{M}_n),$$

oder, mit den Bezeichnungen von §§ 36 bis 37:

$$(7) \quad \overline{H}_1(n) = (\mathfrak{S}_1 w_1) h + (\mathfrak{S}_2 w_2) h.$$

¹ Zur korrespondenzmäßigen Erläuterung: In der klassischen Theorie ist a_n die *große Halbachse der KEPLER-Ellipse*.

Die Eigenwerte der zu w_1 bzw. w_2 parallelen Komponenten von $\mathfrak{F}_1, \mathfrak{F}_2$ sind uns aber schon aus § 36 bekannt und liefern gemäß der Störungstheorie von § 41 in der Tat die Energieformel (7), § 37.

In der zweiten Ableitung der BALMER-Formel (§ 36) hatten wir eine solche Lösung des ungestörten H -Atomproblems betrachtet, bei welcher \mathfrak{M}_z und \mathfrak{M}^2 Diagonalmatrizen waren. Diese Lösung ist offenbar eine „angepaßte Lösung“ (vgl. § 41) zu einer Störung durch ein zusätzliches, *zentralsymmetrisches* (nicht COULOMBSches) *Kraftfeld*, oder auch durch die *relativistische Massenveränderlichkeit*. Sie ist deshalb zu verwenden für eine Untersuchung der *Feinstruktur* des Wasserstoffspektrums¹.

§ 45. Störungstheorie nicht abgeschlossener Systeme. Ein nicht abgeschlossenes System kann man mit den Methoden der Störungstheorie behandeln, wenn sich der Vorgang auffassen läßt als zeitabhängige Störung eines abgeschlossenen Systems, wenn also die Energiefunktion sich auf die Gestalt

$$H = \dot{H}(p, q) + \lambda H'(p, q; t)$$

bringen läßt, wobei die Störungsenergie $\lambda H'$ relativ schwach ist. Man wird dann zunächst solche Koordinaten \hat{p}, \hat{q} einführen, die \dot{H} auf die Diagonalform \dot{W} bringen, so daß

$$(1) \quad H = \dot{W} + \lambda H'(\hat{q}, \hat{p}; t)$$

ist. Die Aufgabe besteht darin, die mit diesem H gebildete Gleichung § 22, (16)

$$(2) \quad HU + \varkappa \dot{U} = 0$$

durch eine unitäre Matrix U zu lösen. Da H zeitabhängig ist, gibt es keine allgemeine Vorschrift, um $U(t)$ näher festzulegen (vgl. § 22). Wir nehmen aber an, daß die Störung zur Zeit t_0 anfängt, daß also $H'(p, q; t) = 0$ für $t \leq t_0$. Dann ist für $t \leq t_0$ die Lösung durch die Forderung: $H(p, q) = \dot{W}$ (Diagonalmatrix) festgelegt (bis auf eine zeitunabhängige unitäre *Stufenmatrix*, welche der Entartung von \dot{W} entspricht und welche wir uns schon in \hat{q}, \hat{p} aufgenommen denken), und zwar ist für $t \leq t_0$:

¹ Vgl. HEISENBERG, W. u. P. JORDAN: Z. Phys. Bd. 37, S. 263. 1926.

$U(t) = e^{-\frac{\hat{W}}{\varkappa}t}$. Macht man für alle t den Ansatz:

$$(3) \quad U(t) = e^{-\frac{\hat{W}}{\varkappa}t} V(t),$$

(wo $V(t)$ ebenfalls unitär ist), so ist es also sinngemäß zu fordern, daß $V(t_0)$ die Einheitsmatrix ist, wodurch die Unbestimmtheit aufgehoben wird. Es gilt:

$$\dot{U} = -\frac{\hat{W}}{\varkappa}U + e^{-\frac{\hat{W}}{\varkappa}t} \dot{V}$$

und die Gleichung (2) lautet auf Grund von (1):

$$\lambda H' e^{-\frac{\hat{W}}{\varkappa}t} V + \varkappa e^{-\frac{\hat{W}}{\varkappa}t} \dot{V} = 0.$$

Wir multiplizieren vorn mit $e^{\frac{\hat{W}}{\varkappa}t}$ und setzen

$$(4) \quad K = e^{\frac{\hat{W}}{\varkappa}t} H' e^{-\frac{\hat{W}}{\varkappa}t} = (H'_{nm} e^{2\pi i \nu_{nm} t}).$$

Dann lautet die Gleichung für V :

$$(5) \quad \varkappa \dot{V} + \lambda K V = 0,$$

hat also wieder die ursprüngliche Gestalt (2), nur daß K allein von der Störungsfunktion H' abhängt.

Setzen wir die Lösung von (5) als Potenzreihe nach λ an:

$$(6) \quad V = V^0 + \lambda V^{(1)} + \lambda^2 V^{(2)} + \dots,$$

so ergeben sich aus (5) durch Koeffizientenvergleich die Näherungsgleichungen:

$$(7) \quad \begin{cases} \varkappa \dot{V}^0 = 0 \\ \varkappa \dot{V}^{(1)} = -K V^{(0)} \\ \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \\ \varkappa \dot{V}^{(l)} = -K V^{(l-1)}. \end{cases}$$

Wegen $V(t_0) = E$ folgen aus (6) die Anfangsbedingungen:

$$(8) \quad V^0(t_0) = E, \dots, \quad V^{(l)}(t_0) = 0, \dots$$

Die Lösung, welche (7) und (8) befriedigt, lautet:

$$(9) \quad V^0(t) = E, \quad V^{(l)}(t) = -\frac{1}{\varkappa} \int_{t_0}^t K V^{(l-1)} d\tau;$$

das ist eine Rekursionsformel, aus der man $V^{(1)}, V^{(2)}, \dots$ der Reihe nach berechnen kann. Nach den Ergebnissen von § 22 ist dann die Matrix (6) für alle t unitär, da sie für $t = t_0$ in die Einheitsmatrix übergeht.

Als Beispiel betrachten wir den Fall, wo

$$H'(\tau) = \begin{cases} 0 & \text{für } \tau < 0 \text{ und } \tau > t, \\ H' = \text{konst.} & \text{für } 0 < \tau < t. \end{cases}$$

Dann wird für $\tau = t$

$$(10) \quad \begin{cases} V_{nm}^{(1)} = -\frac{1}{\hbar \dot{\nu}_{nm}} H'_{nm} (e^{2\pi i \dot{\nu}_{nm} t} - 1), \\ V_{nm}^{(2)} = \sum_k \frac{1}{\hbar \dot{\nu}_{km}} H'_{nk} H'_{km} \left(\frac{e^{2\pi i \dot{\nu}_{nm} t} - 1}{\hbar \dot{\nu}_{nm}} - \frac{e^{2\pi i \dot{\nu}_{nk} t} - 1}{\hbar \dot{\nu}_{nk}} \right). \end{cases}$$

Besteht H' selbst aus Gliedern verschiedener Größenordnung in λ :

$$H' = H^{(1)} + \lambda H^{(2)} + \dots,$$

so ändern sich die Formeln (7), (9), (10) dementsprechend und man erhält schließlich für U nach (3)

$$(10') \quad \begin{cases} U_{nm}^{(1)} = -\frac{e^{-\frac{\dot{W}_n}{\kappa} t}}{\hbar \dot{\nu}_{nm}} H_{nm}^{(1)} (e^{2\pi i \dot{\nu}_{nm} t} - 1), \\ U_{nm}^{(2)} = -\frac{e^{-\frac{\dot{W}_n}{\kappa} t}}{\hbar \dot{\nu}_{nm}} H_{nm}^{(2)} (e^{2\pi i \dot{\nu}_{nm} t} - 1) \\ \quad + e^{-\frac{\dot{W}_n}{\kappa} t} \sum_k \frac{1}{\hbar \dot{\nu}_{km}} H_{nk}^{(1)} H_{km}^{(1)} \left(\frac{e^{2\pi i \dot{\nu}_{nm} t} - 1}{\hbar \dot{\nu}_{nm}} - \frac{e^{2\pi i \dot{\nu}_{nk} t} - 1}{\hbar \dot{\nu}_{nk}} \right). \end{cases}$$

Von diesen Formeln wird später (Kap. VII) Gebrauch gemacht werden.

Will man eine *rein periodische Störung* mit Hilfe der Formel (9) behandeln, so kann man sie als Grenzfall einer langsam anklingenden Störung ansehen, die für $t_0 = -\infty$ verschwindet. Wir setzen

$$(11) \quad H' = (A(\dot{p}, \dot{q}) e^{2\pi i \nu t} + A^\dagger(\dot{p}, \dot{q}) e^{-2\pi i \nu t}) e^{2\pi \nu t}$$

mit reellem ν und positivem σ ; dann ist H' für eine willkürliche Matrix A hermitisch. Nach (4) wird

$$(12) \quad K_{nm} = A_{nm} e^{2\pi i[(\nu_{nm} + \nu) - i\sigma]t} + A_{nm}^\dagger e^{2\pi i[(\nu_{nm} - \nu) - i\sigma]t}.$$

Aus (9) folgt nun:

$$(13) \quad V_{nm}^{(1)} = -\frac{A_{nm}}{\hbar [(\nu_{nm} + \nu) - i\sigma]} (e^{2\pi i[(\nu_{nm} + \nu) - i\sigma]t} - e^{2\pi i[(\nu_{nm} + \nu) - i\sigma]t_0}) \\ - \frac{A_{nm}^\dagger}{\hbar [(\nu_{nm} - \nu) - i\sigma]} (e^{2\pi i[(\nu_{nm} - \nu) - i\sigma]t} - e^{2\pi i[(\nu_{nm} - \nu) - i\sigma]t_0}),$$

wobei $t_0 = -\infty$ ist. Die Glieder mit t_0 fallen also fort. Nun kann man σ beliebig klein nehmen, dann wird für nicht zu große t :

$$(14) \quad V_{nm}^{(1)} = -\frac{A_{nm}}{\hbar (\nu_{nm} + \nu)} e^{2\pi i(\nu_{nm} + \nu)t} - \frac{A_{nm}^\dagger}{\hbar (\nu_{nm} - \nu)} e^{2\pi i(\nu_{nm} - \nu)t}.$$

Die höheren Näherungen enthalten Glieder mit

$$e^{2\pi i(\nu_{nm} + l\nu)t},$$

und zwar geht in $V^{(k)}$ der Index l von $-k$ bis $+k$ ein. Man kann leicht Rekursionsformeln für die Koeffizienten dieser Glieder aufstellen, am bequemsten durch Einsetzen in die Gleichungen (9). Doch soll hier von der Aufstellung dieser Formeln abgesehen werden, da sie praktisch nicht gebraucht werden.

Wir berechnen jetzt die Störung der Koordinatenmatrix q . Nach (3) ist

$$(15) \quad q = U^{-1}q^0 U = V^{-1}Q^0 V,$$

wo

$$(16) \quad Q^0 = e^{\frac{\hat{W}}{\hbar}t} q^0 e^{-\frac{\hat{W}}{\hbar}t} = (q_{nm}^0 e^{2\pi i\nu_{nm}t}).$$

Die Entwicklung von q setzen wir in der Form

$$(17) \quad q = Q^0 + \lambda q^{(1)} + \dots$$

an und bekommen, wie in § 38, (21, 1)

$$(18) \quad q^{(1)} = Q^0 V^{(1)} - V^{(1)}Q^0, \\ q_{nm}^{(1)} = \sum_k (q_{nk}^0 V_{km}^{(1)} e^{2\pi i\nu_{nk}t} - V_{nk}^{(1)} q_{km}^0 e^{2\pi i\nu_{km}t})$$

oder nach (14) mit Rücksicht auf $\nu_{nk} + \nu_{km} = \nu_{nm}$:

$$(19) \quad q_{nm}^{(1)} = e^{2\pi i(\nu_{nm} + \nu)t} \sum_k \left(\frac{A_{nk} q_{km}^0}{\hbar(\nu_{nk} + \nu)} - \frac{q_{nk}^0 A_{km}}{\hbar(\nu_{km} + \nu)} \right) \\ + e^{2\pi i(\nu_{nm} - \nu)t} \sum_k \left(\frac{A_{nk}^\dagger q_{km}^0}{\hbar(\nu_{nk} - \nu)} - \frac{q_{nk}^0 A_{km}^\dagger}{\hbar(\nu_{km} - \nu)} \right).$$

Ist das ungestörte System entartet und schreibt man alle Matrizen als Übermatrizen in der Form $(q_{\rho, \sigma}^{(r, s)})$, so wird

$$(20) \quad q_{\rho, \sigma}^{(1)} = e^{2\pi i(\nu^{(r, s)} + \nu)t} \sum_k \left(\frac{\sum_{\tau} A_{\rho, \tau}^{(r, k)} \overset{\circ}{q}_{\tau, \sigma}^{(k, s)}}{\hbar(\nu^{(r, k)} + \nu)} - \frac{\sum_{\tau} \overset{\circ}{q}_{\rho, \tau}^{(r, k)} A_{\tau, \sigma}^{(k, s)}}{\hbar(\nu^{(k, s)} + \nu)} \right) \\ + e^{2\pi i(\nu^{(r, s)} - \nu)t} \sum_k \left(\frac{\sum_{\tau} A_{\rho, \tau}^{\dagger(r, k)} \overset{\circ}{q}_{\tau, \sigma}^{(k, s)}}{\hbar(\nu^{(r, k)} - \nu)} - \frac{\sum_{\tau} \overset{\circ}{q}_{\rho, \tau}^{(r, k)} A_{\tau, \sigma}^{\dagger(k, s)}}{\hbar(\nu^{(k, s)} - \nu)} \right).$$

Genau entsprechende Ausdrücke gelten für die Störung irgend einer Funktion $f(p, q)$; man hat in (19) oder (20) einfach $\overset{\circ}{q}$ durch $\overset{\circ}{f} = f(\overset{\circ}{p}, \overset{\circ}{q})$ und q durch f zu ersetzen [s. § 38, (23), (24)].

Bemerkenswert ist, daß in (20) keine „säkularen“ Zusatzglieder auftreten, wie bei zeitunabhängiger Störung (§ 41); das gilt jedoch nur, solange weder $\nu = 0$ (statischer Fall) noch gleich einer der Eigenfrequenzen $\nu^{(r, s)}$ (Resonanzfall) ist. Denn nur dann verschwindet keiner der Nenner in (19) bzw. (20).

Die beiden Ausnahmefälle, besonders den physikalisch sehr wichtigen Resonanzfall, werden wir noch ausführlich erörtern. Vorläufig lassen wir sie unberücksichtigt; die Resultate, die wir in den folgenden Paragraphen gewinnen werden, gelten also nur mit der Einschränkung, daß die betreffenden Frequenzen ν nicht $= 0$ und weit von den Resonanzstellen entfernt sind.

§ 46. Dispersion des Lichts. Von den Formeln § 45, (19), (20) machen wir eine Anwendung auf den Fall, daß eine monochromatische Lichtwelle das System trifft. Von dieser setzen wir voraus, daß ihre Wellenlänge sehr groß ist gegen die linearen Abmessungen des Systems; dann kann man von den Verschiebungen der Phase in den einzelnen Punkten des Systems absehen und von dem Werte des Feldes „am Orte des Systems“ schlechtweg sprechen. Im nächsten Paragraphen werden wir diese Voraussetzung fallen lassen. Das elektrische Feld der Welle sei durch

$$(1) \quad \lambda \mathfrak{E}_x = \lambda a_x \cos(2\pi\nu t + \varphi_x), \quad \lambda \mathfrak{E}_y = \lambda a_y \cos(2\pi\nu t + \varphi_y), \\ \lambda \mathfrak{E}_z = \lambda a_z \cos(2\pi\nu t + \varphi_z)$$

mit reellen Amplituden und Phasen dargestellt. Statt dessen kann man auch die komplexen Amplituden

$$(2) \quad \mathfrak{C}_x = a_x e^{i\varphi_x}, \quad \mathfrak{C}_y = a_y e^{i\varphi_y}, \quad \mathfrak{C}_z = a_z e^{i\varphi_z}$$

einführen und schreiben:

$$(3) \quad \mathfrak{C} = \frac{1}{2} (\mathfrak{C} e^{2\pi i \nu t} + \mathfrak{C}^* e^{-2\pi i \nu t}),$$

wo \mathfrak{C}^* der konjugiert komplexe Vektor zu (2) ist.

Wir werden hier die Rückwirkung des getroffenen Systems auf das Licht (erzwungene Ausstrahlung) vernachlässigen; eine Berücksichtigung dieser Wechselwirkung kann erst später erfolgen (s. Kap. VII).

Ist \mathfrak{P} das elektrische Moment des Systems (eine hermitesche Matrix), so ist die Störungsenergie

$$(4) \quad H' = -\mathfrak{P}\mathfrak{C} = -\frac{1}{2} (\mathfrak{P}\mathfrak{C} e^{2\pi i \nu t} + \mathfrak{P}\mathfrak{C}^* e^{-2\pi i \nu t});$$

sie hat wegen $\mathfrak{P} = \mathfrak{P}^\dagger$ die Form des Ausdrucks § 45, (11) für $\sigma = 0$ mit

$$(5) \quad A = -\frac{1}{2} \mathfrak{P}\mathfrak{C}.$$

Ist x_r eine rechtwinklige Koordinate eines Teilchens des Systems und entwickelt man sie nach λ :

$$(6) \quad x_r(n, m) = x_r^0(n, m) e^{2\pi i \nu n m t} + \lambda x_r^{(1)}(n, m) + \dots,$$

so erhält man aus § 45, (19) für die Störung

$$(7) \quad x_r^{(1)}(n, m) = \frac{1}{2} e^{2\pi i (\nu n m + \nu) t} \sum_k \left(\frac{x_r^0(n, k) (\mathfrak{P}(k, m) \cdot \mathfrak{C})}{\hbar(\nu_{km} + \nu)} - \frac{(\mathfrak{P}(n, k) \cdot \mathfrak{C}) x_r^0(k, m)}{\hbar(\nu_{nk} + \nu)} \right) + \frac{1}{2} e^{2\pi i (\nu n m - \nu) t} \sum_k \left(\frac{x_r^0(n, k) (\mathfrak{P}(k, m) \cdot \mathfrak{C}^*)}{\hbar(\nu_{km} - \nu)} - \frac{(\mathfrak{P}(n, k) \cdot \mathfrak{C}^*) x_r^0(k, m)}{\hbar(\nu_{nk} - \nu)} \right)$$

und hieraus für die Störung des elektrischen Moments

$$(8) \quad \mathfrak{P}(n, m) = \mathfrak{P}(n, m) e^{2\pi i \nu n m t} + \lambda \mathfrak{P}^{(1)}(n, m) + \dots$$

in erster Näherung:

$$(9) \quad \mathfrak{P}^{(1)}(n, m) = \\ \frac{1}{2} e^{2\pi i(\nu_{nm} + \nu)t} \sum_k \left(\frac{\mathfrak{P}(n, k) \cdot \mathfrak{P}(k, m) \cdot \mathfrak{G}}{\hbar(\nu_{km} + \nu)} - \frac{(\mathfrak{P}(n, k) \cdot \mathfrak{G}) \mathfrak{P}(k, m)}{\hbar(\nu_{nk} + \nu)} \right) \\ + \frac{1}{2} e^{2\pi i(\nu_{nm} - \nu)t} \sum_k \left(\frac{\mathfrak{P}(n, k) \cdot \mathfrak{P}(k, m) \cdot \mathfrak{G}^*}{\hbar(\nu_{km} - \nu)} - \frac{(\mathfrak{P}(n, k) \cdot \mathfrak{G}^*) \cdot \mathfrak{P}(k, m)}{\hbar(\nu_{nk} - \nu)} \right).$$

Diese Formeln enthalten die Grundlage für die Erklärung der Dispersion des Lichts. Wir haben uns ja mit HEISENBERG auf den Standpunkt gestellt (der erst später tiefer begründet werden wird), daß das elektrische Moment \mathfrak{P} die vom System emittierte Strahlung bestimmt. Durch das einfallende Licht wird nun ein Zusatzmoment (9) erzeugt, dessen Matrixelemente Frequenz und Intensität des „gestreuten“ Lichts bestimmen. In der klassischen Theorie wird gezeigt, daß die Streuwellen der einzelnen Atome sich durch Interferenz miteinander und mit der einfallenden ebenen Welle so beeinflussen, daß eine veränderte ebene Welle (mit anderer Wellenlänge) in dem Körper zu laufen scheint. Wir übernehmen diese Erklärung der Farbenzerstreuung vorläufig in die Quantentheorie.

Ehe wir aber auf den physikalischen Inhalt der Formeln näher eingehen, wollen wir den Grenzfall hoher Lichtfrequenzen ν untersuchen¹. Es sei also ν groß gegen alle Eigenfrequenzen ν_{nm} des Systems, also die mechanische Bindung der Teilchen des Systems relativ schwach gegen die Lichteinwirkung. Wir behaupten, daß dann die einzelnen Teilchen so schwingen, als wären sie frei und gehorchten den Gesetzen der klassischen Theorie.

Entwickelt man nämlich die Summen in (7) nach Potenzen von $\frac{\nu_{nm}}{\nu}$, so erhält man

$$(10) \quad x_r^{(1)}(n, m) = \frac{e^{2\pi i \nu_{nm} t}}{\hbar \nu} \kappa \{ \mathfrak{F}_x [x_r^0, X^0] + \mathfrak{F}_y [x_r^0, Y^0] + \mathfrak{F}_z [x_r^0, Z^0] \}_{nm} \\ - \frac{e^{2\pi i \nu_{nm} t}}{\hbar \nu} \frac{\kappa}{2\pi i \nu} \{ \mathfrak{G}_x [x_r^0, \dot{X}^0] + \mathfrak{G}_y [x_r^0, \dot{Y}^0] + \mathfrak{G}_z [x_r^0, \dot{Z}^0] \}_{nm},$$

¹ Die folgenden Betrachtungen stammen von KRAMERS, H. A.: *Physica* Bd. 5, S. 369. 1925.

Bedingung dafür darstellen, daß das System sich im Grenzfall schwacher Kräfte $\left(\frac{\nu_n m}{\nu} \ll 1\right)$ wie ein klassisches System freier Teilchen verhält. Nun wurden, wie bereits in der Einleitung, § 2, gesagt, Mitschwingungsformeln der Gestalt (7) bereits vor der systematischen Entwicklung der Quantenmechanik aufgestellt, zuerst nicht ganz vollständig von LADENBURG¹, dann in vollkommener Form von KRAMERS². Der Gedanke war, aus den klassischen Formeln durch Korrespondenzbetrachtungen solche Quantenformeln herzuleiten, daß die Resonanzstellen mit den Emissionsfrequenzen ν_{nm} zusammenfallen. Die Zähler der Resonanzbrüche blieben dabei weitgehend unbestimmt. Doch erkannten THOMAS³ und KUHN⁴, daß sich diese Willkür einschränken lasse durch die Forderung, daß sich das System im Grenzfall verschwindender Bindung klassisch verhalte. Die von ihnen als „Summensätze“ aufgestellten Gleichungen sind im wesentlichen die kanonischen Vertauschungsrelationen, die also hierzum erstmal in der Literatur auftreten. Ihre prinzipielle Bedeutung ist aber erst durch die HEISENBERGSche Theorie ans Licht gekommen.

Wir kehren nun zu der Formel (9) zurück. Die mit der einfallenden Welle gleichfarbigen Schwingungen ($n = m$) werden mit der einfallenden Welle und untereinander interferieren. Wenden wir nun auf die ausgesandten Wellen die Gesetze der gewöhnlichen elektromagnetischen Lichttheorie an, so werden wir schließen, daß durch die Interferenz der einfallenden Welle mit den von allen Atomen des Gases ausgehenden Sekundärwellen die Erscheinung der „Brechung“ zustande kommt: Es entsteht eine ebene Überlagerungswelle mit abgeänderter Wellenlänge. Vorausgesetzt ist dabei eine Anordnung der Atome in genauer Gitterstruktur⁵

¹ LADENBURG, R.: Z. Phys. Bd. 4, S. 451. 1921; ferner LADENBURG, R. u. F. REICHE: Naturwissensch. Bd. 11, S. 584. 1923.

² KRAMERS, H. A.: Nature (Lond.) Bd. 113, S. 673. 1924; Bd. 114, S. 310. 1924. KRAMERS, H. A. u. W. HEISENBERG: Z. Phys. Bd. 31, S. 681. 1925.

³ THOMAS, W.: Naturwissensch. Bd. 13, S. 627. 1925. REICHE, F. u. W. THOMAS: Z. Phys. Bd. 34, S. 510. 1925.

⁴ KUHN, W.: Z. Phys. Bd. 33, S. 408. 1925.

⁵ EWALD, P. P.: Diss. München 1912; Ann. Physik Bd. 49, S. 1, 117. 1916; Bd. 54, S. 519, 557. 1917; Phys. Z. Bd. 21, S. 617. 1920; Z. Phys. Bd. 2, S. 332. 1920. BORN, M.: Dynamik der Kristallgitter. Teil II Leipzig 1915. Atomtheorie des festen Zustandes. Nr. 41 ff. Leipzig 1923.

oder völlige räumliche Unordnung¹. Das Resultat dieser Interferenz läßt sich bekanntlich mit Hilfe der MAXWELLSCHEN Theorie für kontinuierliche Medien beschreiben: Die Gleichungen (9) stellen für $n = m$ das vom Licht erregte mittlere Moment des Systems als lineare Vektorfunktion der elektrischen Feldstärke dar; verbindet man diese Relationen mit der aus den MAXWELLSCHEN Gleichungen folgenden Formel für das mittlere Moment pro Volumeinheit \mathfrak{P} :

$$(15) \quad \mathfrak{P} - n^2 \mathfrak{s}(\mathfrak{s} \mathfrak{P}) = \frac{n^2 - 1}{4\pi} \mathfrak{E},$$

wo n der Brechungsindex, \mathfrak{s} der Einheitsvektor der Wellennormale ist, so erhält man Bestimmungsgleichungen für die Wellenamplituden und den Brechungsindex. Im allgemeinen, bei nicht entarteten Systemen, wird sich dabei *Doppelbrechung* ergeben; ist z. B. die Richtungsentartung der Atome eines Gases durch ein elektrisches oder magnetisches Feld aufgehoben, so erhält man den *KERR-Effekt* bzw. *FARADAY-Effekt*. Ehe wir darauf eingehen (§§ 48, 49), betrachten wir den Fall eines sich selbst überlassenen Gases mit richtungsentarteten Atomen. Nach § 20 kann

man aus dem Vektor \mathfrak{P} Größen bilden, die unabhängig von der Orientierung des Atoms sind, nämlich die Komponenten von $Sp \mathfrak{P}^{(r,r)} = \sum_m^{(1)} \mathfrak{P}_{mm}^{(r,r)}$, wo r alle Quantenzahlen (n, j) außer der äquatorialen m zusammenfaßt. Ist g_r das Gewicht des Zustandes r , so führen wir das *mittlere induzierte Moment eines einfachen Zustandes* (vom Gewicht 1) ein durch

$$(16) \quad \mathfrak{P}^{(r,r)} = \frac{1}{g_r} Sp \mathfrak{P}^{(1)} = \frac{1}{g_r} \sum_m^{(1)} \mathfrak{P}_{mm}^{(r,r)}.$$

¹ LORENTZ, H. A.: Verh. d. k. Akad. Amsterdam Bd. 18. 1879; Arch. néerl. Bd. 25, S. 363. 1892; Wied. Ann. Bd. 9, S. 657. 1880; Versuch einer Theorie der elektr. u. opt. Erscheinungen (Leiden 1895); The Theory of Electrons (Leipzig 1916), Chap. IV. PLANCK, M.: Berl. Ber. Bd. 24, S. 470. 1902. ESMARCH, W.: Ann. Physik Bd. 42, S. 1257. 1913. NATANSON, L.: Krak. Anz. (Bull. internat.) A., S. 1 u. 335. 1914; S. 221. 1916. BOTHE, W.: Diss., Berlin 1914; Ann. Physik Bd. 64, S. 693. 1921. OSEEN, C. W.: Ann. Physik Bd. 48, S. 1. 1915; Phys. Z. Bd. 16, S. 404. 1915. FAXÉN, H.: Z. Phys. Bd. 2, S. 218. 1920. REICHE, F.: Ann. Physik Bd. 50, S. 1. 1916. DARWIN, C. G.: Trans. Cambr. phil. Soc. Bd. 23, S. 137. 1924.

Diese Größe ist offenbar in die MAXWELLSchen Gleichungen als mittleres Moment pro Teilchen einzusetzen. Man erkennt leicht, daß sie dem Feldvektor \mathfrak{E} proportional ist. Aus (9) erhält man nämlich:

$$g_r \bar{\mathfrak{P}}(r, r) = \frac{1}{2} e^{2\pi i r t} \sum_s \left(\frac{Sp[\hat{\mathfrak{P}}^{(r,s)}(\hat{\mathfrak{P}}^{(s,r)} \cdot \hat{\mathfrak{E}})]}{\hbar(\nu^{(s,r)} + \nu)} - \frac{Sp[(\hat{\mathfrak{P}}^{(r,s)} \cdot \hat{\mathfrak{E}}) \hat{\mathfrak{P}}^{(s,r)}]}{\hbar(\nu^{(r,s)} + \nu)} \right) \\ + \frac{1}{2} e^{-2\pi i r t} \sum_s \left(\frac{Sp[\hat{\mathfrak{P}}^{(r,s)}(\hat{\mathfrak{P}}^{(s,r)} \hat{\mathfrak{E}}^*)]}{\hbar(\nu^{(s,r)} - \nu)} - \frac{Sp[(\hat{\mathfrak{P}}^{(r,s)} \hat{\mathfrak{E}}^*) \hat{\mathfrak{P}}^{(s,r)}]}{\hbar(\nu^{(r,s)} - \nu)} \right).$$

Beachtet man, daß $\nu^{(r,s)} = -\nu^{(s,r)}$ ist und die Relationen § 30, (11) gelten, so erhält man

$$g_r \bar{\mathfrak{P}}^{(r,r)} = \frac{1}{2} (\hat{\mathfrak{E}} e^{2\pi i r t} + \hat{\mathfrak{E}}^* e^{-2\pi i r t}) \\ \cdot \frac{1}{3} \sum_s Sp(\hat{\mathfrak{P}}^{(r,s)} \cdot \hat{\mathfrak{P}}^{(s,r)}) \left(\frac{1}{\hbar(\nu^{(s,r)} - \nu)} + \frac{1}{\hbar(\nu^{(s,r)} + \nu)} \right) \\ = \mathfrak{E} \frac{2}{3\hbar} \sum_s \|\hat{\mathfrak{P}}^{(r,s)}\|^2 \frac{\nu^{(s,r)}}{\nu^{(s,r)^2} - \nu^2}.$$

Führt man nun, wie früher in § 42, die *Deformierbarkeit* $\alpha(r)$ für ein harmonisches Wechselfeld ein durch

$$(17) \quad \bar{\mathfrak{P}}^{(r,r)} = \alpha_r \mathfrak{E},$$

so hat man

$$(18) \quad \alpha_r = \frac{2}{3\hbar g_r} \sum_s \|\hat{\mathfrak{P}}^{(r,s)}\|^2 \frac{\nu^{(s,r)}}{\nu^{(s,r)^2} - \nu^2}.$$

Hieraus ergibt sich der Brechungsindex n auf Grund von (15); aus (15) folgt nämlich durch skalare Multiplikation mit \mathfrak{s}

$$\mathfrak{s} \mathfrak{P} = -\frac{1}{4\pi} \mathfrak{s} \mathfrak{E};$$

das bedeutet aber die Transversalität des Verschiebungsvektors $\mathfrak{D} = \mathfrak{E} + 4\pi \mathfrak{P}$: $\mathfrak{D} \mathfrak{s} = 0$. Ist nun \mathfrak{P} parallel zu \mathfrak{E} , so ist auch $\mathfrak{E} \mathfrak{s} = 0$. Ist N_r die Anzahl Teilchen pro Volumeneinheit im Zustande r , so ist das \mathfrak{P} in (15) mit $\sum_r N_r \bar{\mathfrak{P}}^{(r,r)}$ gleichbedeutend.

Aus (15) und (17) folgt nun

$$(19) \quad n^2 = 1 + 4\pi \sum_r N_r \alpha_r.$$

Setzt man hier (18) ein, so hat man die *Dispersionsformel*, die den Brechungsindex als Funktion der Frequenz darstellt.

Man pflegt nun in dem Ausdruck (18) verschiedene Schreibweisen zu benützen, die seine Beziehung zu andern Erscheinungen sinnfällig machen. Man schreibt statt (18)

$$(20) \quad \alpha_r = \frac{e^2}{4\pi^2 m_0} \sum_s \frac{f^{(s,r)}}{\nu^{(s,r)^2 - \nu^2}$$

und nennt $f^{(s,r)}$ die *Stärke der Eigenschwingung* $\nu^{(s,r)}$. Jedes Glied der Summe hat nämlich bis auf den Faktor $f^{(s,r)}$ die Form der Resonanzformel für einen harmonischen Oszillator der Masse m_0 , Ladung e und Frequenz $\nu^{(s,r)}$. Durch Vergleich von (18) und (20) ergibt sich

$$(21) \quad f^{(s,r)} = \frac{8\pi^2 m_0 \nu^{(s,r)}}{3e^2 \hbar} \left| \mathfrak{P}^{(s,r)} \right|^2.$$

Diese Stärkefaktoren sind positiv oder negativ, je nachdem $\nu^{(s,r)} > 0$ oder < 0 ist, d. h. je nachdem das Energieniveau r tiefer oder höher liegt als s . Letzteres ist nur dann möglich, wenn r nicht das tiefste Energieniveau des betrachteten Systems ist; die entsprechenden Glieder spielen also nur dann eine Rolle, wenn das Gas auch angeregte Molekeln in beträchtlicher Anzahl enthält. Wie schon bemerkt, wurde die erste quantentheoretische Dispersionsformel, deren Resonanzstellen die Emissionsfrequenzen $\nu^{(s,r)} > 0$ sind, von LADENBURG¹ aufgestellt; er hat auch die sogleich zu besprechenden Beziehungen der $f^{(s,r)}$ zu den Übergangswahrscheinlichkeiten entdeckt. Die Glieder „negativer Dispersion“ ($f^{(s,r)} < 0$) wurden von KRAMERS² aus Gründen des Korrespondenzprinzips hinzugefügt.

Wir wollen jetzt die $f^{(s,r)}$ durch andere beobachtbare Größen ausdrücken. Wie wir früher (§ 20) gesehen haben, ist die beim Übergang $s \rightarrow r$ ($\overset{\circ}{W}_s > \overset{\circ}{W}_r$) spontan emittierte Lichtenergie ebenfalls den Größen $\left| \mathfrak{P}^{(s,r)} \right|^2$ proportional, nämlich

$$(22) \quad J^{(s,r)} = \frac{64\pi^4 \nu^{(s,r)^4}{3c^3} \left| \mathfrak{P}^{(s,r)} \right|^2.$$

¹ LADENBURG, R.: Z. Phys. Bd. 4, S. 451. 1921.

² KRAMERS, H. A.: Nature Bd. 113, S. 673. 1924; ebenda Bd. 114, S. 310. 1924; vgl. auch KRAMERS, H. A. u. W. HEISENBERG: Z. Phys. Bd. 31, S. 681. 1925.

Man bekommt also

$$(23) \quad f^{(s, r)} = \frac{m_0 c^3}{8 \pi^2 e^2 \hbar} \frac{1}{\nu^{(s, r)^3} g_r} J^{(s, r)}.$$

Statt $J^{(s, r)}$ gebraucht man gewöhnlich die Übergangswahrscheinlichkeit beim Sprunge $s \rightarrow r$. Diese definieren wir vorläufig (wir gehen später ausführlich auf diesen Begriff ein) durch folgende, dem Gedankenkreis der „Lichtquanten“ entnommene Vorstellung: Denkt man sich die pro Zeiteinheit emittierte Energie $J^{(s, r)}$ in Form von Lichtquanten der Größe $\hbar \nu^{(s, r)}$, so ist deren Zahl pro Zeiteinheit $\frac{J^{(s, r)}}{\hbar \nu^{(s, r)}}$. Diese Quanten gehen vom Zustande s aus, der durch ein Feld in g_s Einzelterme aufgespalten werden kann; die *spontane Übergangswahrscheinlichkeit pro Zeiteinheit* von dem Einzelterm ist also

$$(24) \quad A^{(s, r)} = \frac{J^{(s, r)}}{g_s \hbar \nu^{(s, r)}}, \quad \nu^{(s, r)} > 0;$$

daraus folgt

$$(24a) \quad A^{(r, s)} = -\frac{g_s}{g_r} \frac{J^{(s, r)}}{g_s \hbar \nu^{(s, r)}}, \quad \nu^{(s, r)} < 0.$$

Statt dieser gebraucht man oft auch die *Wahrscheinlichkeit des erzwungenen Übergangs* $s \rightarrow r$, definiert durch

$$(25) \quad \left\{ \begin{array}{l} B^{(s, r)} = \frac{c^3}{8 \pi \hbar \nu^{(s, r)^3}} A^{(s, r)} \\ g_r B^{(r, s)} = g_s B^{(s, r)} \end{array} \right\} \nu^{(s, r)} > 0.$$

Diese Formeln werden wir bald ausführlich begründen (s. §§ 63, 74). Nach (25) ist

$$(26) \quad B^{(s, r)} = \frac{g_r}{g_s} B^{(r, s)} = \frac{c^3}{8 \pi \hbar \nu^{(r, s)^3}} A^{(r, s)} \frac{g_r}{g_s}, \quad \nu^{(s, r)} < 0.$$

Nach (24), (24a) wird

$$(27) \quad \left\{ \begin{array}{l} f^{(s, r)} = \frac{m_0 c^3}{8 \pi^2 e^2 \nu^{(s, r)^2}} \cdot \frac{g_s}{g_r} A^{(s, r)}, \quad \nu^{(s, r)} > 0, \\ f^{(s, r)} = -\frac{m_0 c^3}{8 \pi^2 e^2 \nu^{(r, s)^2}} A^{(r, s)}, \quad \nu^{(s, r)} < 0; \end{array} \right.$$

oder nach (25), (26)

$$(28) \quad f^{(s, r)} = \frac{m_0 \hbar \nu^{(s, r)}}{\pi e^2} \frac{g_s}{g_r} B^{(s, r)}, \quad \nu^{(s, r)} \geq 0.$$

Der Faktor in (27) hat eine einfache Bedeutung. Nach der klassischen Elektrodynamik ist nämlich

$$(29) \quad \tau^{(s, r)} = \frac{3 c^3 m_0}{8 \pi^2 e^2 \nu^{(s, r)^2}}$$

die *Dämpfungskonstante* eines elektrischen Dipols der Frequenz $\nu^{(s, r)}$, dessen Energie nach dem Gesetz

$$W = W_0 e^{-\frac{t}{\tau^{(s, r)}}}$$

abklingt¹. Also wird

$$(30) \quad \begin{cases} f^{(s, r)} = \frac{1}{3} \frac{g_s}{g_r} \tau^{(s, r)} A^{(s, r)}, & \nu^{(s, r)} > 0, \\ f^{(s, r)} = -\frac{1}{3} \tau^{(s, r)} A^{(r, s)}, & \nu^{(s, r)} < 0. \end{cases}$$

Die Stärke der Dispersion ist also (bis auf den Gewichtsfaktor $\frac{g_s}{g_r}$) die Anzahl² der Sprünge $s \rightarrow r$ in der Zeit $\frac{\tau^{(s, r)}}{3}$.

Die Dispersionsformel selbst wird nach (28) und (30) oft in folgenden Formen geschrieben:

$$(31) \quad \alpha_r = \frac{1}{4 \pi^3} \sum_s \frac{g_s}{g_r} \frac{\hbar \nu^{(s, r)} B^{(s, r)}}{\nu^{(s, r)^2} - \nu^2}$$

oder

$$\alpha_r = \frac{e^2}{4 \pi^2 m_0} \left\{ \sum_{\nu^{(s, r)} > 0} \frac{\tau^{(s, r)} g_s}{3 g_r} \frac{A^{(s, r)}}{\nu^{(s, r)^2} - \nu^2} - \sum_{\nu^{(r, s)} < 0} \frac{\tau^{(s, r)} A^{(r, s)}}{3 \nu^{(r, s)^2} - \nu^2} \right\}.$$

In der letzten Form treten die positive und negative Dispersion getrennt in Erscheinung. Der in den Formeln (30) ausgespro-

¹ S. etwa ABRAHAM, M.: Theorie der Elektrizität Bd. 2. 3. Aufl., § 9, (56e), (56f), S. 66. Leipzig 1914.

² Der Faktor $\frac{1}{3}$ rührt daher, daß die Übergangswahrscheinlichkeiten $A^{(r, s)}$ auf ein frei drehbares System bezogen sind (entsprechend einer unpolarisierten Lichtemission), während $\tau^{(s, r)}$ die Abklingungszeit des linearen Oszillators bedeutet.

chene Zusammenhang zwischen Stärke der Spektrallinie und spontaner Übergangswahrscheinlichkeit ist experimentell eingehend von LADENBURG untersucht und bestätigt worden¹.

Die Dispersionsformeln versagen, wenn die erregende Frequenz ν mit einer Eigenfrequenz zusammenfällt (Resonanz). Das liegt natürlich an der Vernachlässigung der Rückwirkung des Lichts auf das System, d. h. der Strahlungsdämpfung. Im Resonanzfalle ist kein periodisches Mitschwingen möglich, die Amplitude wächst vielmehr mit der Zeit an. Dadurch werden Quantensprünge erzeugt, und es sei hier gleich bemerkt, daß die Übergangswahrscheinlichkeiten sich ohne Berücksichtigung der Strahlungsdämpfung berechnen lassen.

§ 47. Optische Aktivität. Bisher wurde vorausgesetzt, daß die Wellenlänge des Lichts unendlich lang sei gegen die Dimensionen des mechanischen Systems. In Wirklichkeit ist das nicht der Fall; daher wirkt das Licht auf die verschiedenen Teile des Systems mit verschiedener Phase ein, und hieraus resultieren Erscheinungen, die man als *optische Aktivität* bezeichnet. Diese Auffassung wurde in der klassischen Optik zuerst von BORN² und OSEEN³ (unabhängig voneinander) entwickelt; hier soll sie auf die Quantenmechanik übertragen werden⁴.

Die im vorigen Paragraphen verwendete Methode, bei der die Störungsfunktion als Arbeit des elektrischen Feldes angesetzt ist, erweist sich hier als unbequem. Das rührt daher, daß bei Berücksichtigung der Phasendifferenzen auch die magnetische Feldstärke des Lichtfeldes an dem Vorgange merklich mitwirkt. Daher empfiehlt sich eine Darstellung des elektromagnetischen Feldes mit Hilfe des *Vektorpotentials* \mathcal{U} , aus dem sich die elektrische Feldstärke \mathcal{E} und die magnetische Induktion \mathcal{B} nach den

¹ LADENBURG, R.: Z. Phys. Bd. 48, S. 15. 1928. KOPFERMANN, H. u. R. LADENBURG: Ebenda S. 26, 51. CARST, A. u. R. LADENBURG: Ebenda S. 192. LADENBURG, R.: Naturwissensch. Bd. 14, S. 1208. 1926. CARST, A., H. KOPFERMANN, R. LADENBURG: Sitzungsber. preuß. Akad. Wiss., Physik.-math. Kl. S. 255. 1926.

² BORN, M.: Phys. Z. Bd. 16, S. 251. 1915.

³ OSEEN, C. W.: Ann. Physik Bd. 48, S. 1. 1915.

⁴ Die erste Darstellung dieser Theorie hat im Anschluß an die Ausarbeitung dieses Buchs L. ROSENFELD gegeben (Z. Phys. Bd. 52, S. 161. 1928).

Regeln

$$(1) \quad \mathfrak{E} = -\frac{1}{c} \frac{\partial \mathfrak{A}}{\partial t}, \quad \mathfrak{B} = \text{rot } \mathfrak{A}$$

berechnen. \mathfrak{A} selbst genügt der Bedingung

$$(2) \quad \text{div } \mathfrak{A} = 0.$$

Die Störungsfunktion der klassischen Theorie ist nach Bd. 1, § 34, (2)

$$(3) \quad H' = -\sum_{\tau} \frac{e_{\tau}}{m_{\tau} c} (\mathfrak{p}_{\tau} \cdot \mathfrak{A}(\mathbf{r}_{\tau})),$$

wo \mathbf{r}_{τ} der Ortsvektor des τ -ten Teilchens mit den Komponenten x_{τ} , y_{τ} , z_{τ} und \mathfrak{p}_{τ} der Impulsvektor, $p_{\tau x}$, $p_{\tau y}$, $p_{\tau z}$ ist.

Die Funktion (3) ist scheinbar nicht sich selbst adjungiert, weil die Komponenten von \mathfrak{p}_{τ} und die von $\mathfrak{A}(\mathbf{r}_{\tau})$ nicht sämtlich vertauschbar sind. Wir machen daher vorläufig den naheliegenden Ansatz

$$(4) \quad H' = -\sum_{\tau} \frac{e_{\tau}}{m_{\tau} c} \frac{1}{2} \{ \mathfrak{p}_{\tau} \mathfrak{A}(\mathbf{r}_{\tau}) + \mathfrak{A}(\mathbf{r}_{\tau}) \mathfrak{p}_{\tau} \};$$

wir werden aber sogleich zeigen, daß dieser sich wegen der Eigenschaft (2) des Vektorpotentials wieder auf die Form (3) reduziert, also keine Willkür bedeutet. Eine ebene Lichtwelle wird, analog wie in § 46, (1), (2), (3) durch

$$(5) \quad \mathfrak{A} = \frac{1}{2} \left(\hat{\mathfrak{A}} e^{2\pi i \nu \left(t - \frac{\mathbf{r} \cdot \mathfrak{S}}{c} \right)} + \hat{\mathfrak{A}}^* e^{-2\pi i \nu \left(t - \frac{\mathbf{r} \cdot \mathfrak{S}}{c} \right)} \right)$$

dargestellt, wobei

$$(6) \quad \mathfrak{S} = n \hat{\mathfrak{s}}$$

den Vektor in Richtung der *Wellennormalen* $\hat{\mathfrak{s}}$ ($|\hat{\mathfrak{s}}| = 1$) bedeutet, dessen Länge gleich dem *Brechungsindex* n ist ($\hat{\mathfrak{A}}$ und \mathfrak{S} sind *Zahlenvektoren*). Dann gilt nach (2)

$$(7) \quad \hat{\mathfrak{A}} \mathfrak{S} = 0.$$

Ferner wird nach (1)

$$(8) \quad \left\{ \begin{aligned} \mathfrak{E} &= -\frac{2\pi i \nu}{c} \frac{1}{2} \left(\hat{\mathfrak{A}} e^{2\pi i \nu \left(t - \frac{\mathbf{r} \cdot \mathfrak{S}}{c} \right)} - \hat{\mathfrak{A}}^* e^{-2\pi i \nu \left(t - \frac{\mathbf{r} \cdot \mathfrak{S}}{c} \right)} \right), \\ \mathfrak{B} &= -\frac{2\pi i \nu}{c} \frac{1}{2} \left((\mathfrak{S} \times \hat{\mathfrak{A}}) e^{2\pi i \nu \left(t - \frac{\mathbf{r} \cdot \mathfrak{S}}{c} \right)} - (\mathfrak{S} \times \hat{\mathfrak{A}}^*) e^{-2\pi i \nu \left(t - \frac{\mathbf{r} \cdot \mathfrak{S}}{c} \right)} \right). \end{aligned} \right.$$

Wir wollen nun voraussetzen, daß die Atomdimensionen r_τ zwar nicht vollständig gegen die Wellenlänge $\lambda = \frac{c}{\nu n}$ vernachlässigt werden können, daß aber das Verhältnis $\frac{r_\tau}{\lambda}$ klein ist und es genügt, die Größen 1. Ordnung in $\frac{r_\tau}{\lambda}$ mitzunehmen. Indem wir die Exponentialfunktionen in (5) entwickeln und das Ergebnis in (4) einsetzen, erhalten wir

$$(9) \quad H' = A e^{2\pi i \nu t} + A^\dagger e^{-2\pi i \nu t}$$

mit

$$(10) \quad A = - \sum_\tau \frac{e_\tau}{2 m_\tau c} \left\{ (\mathfrak{p}_\tau \mathfrak{A}) - \frac{2\pi i \nu}{c} \cdot \frac{1}{2} [(\mathfrak{p}_\tau \mathfrak{A})(\mathfrak{r}_\tau \mathfrak{S}) + (\mathfrak{r}_\tau \mathfrak{S})(\mathfrak{A} \mathfrak{p}_\tau)] \right\}.$$

Nun ist nach den kanonischen Vertauschungsregeln

$$(11) \quad (\mathfrak{A} \mathfrak{p}_\tau)(\mathfrak{r}_\tau \mathfrak{S}) - (\mathfrak{r}_\tau \mathfrak{S})(\mathfrak{A} \mathfrak{p}_\tau) = \frac{\hbar}{2\pi i} (\mathfrak{A} \mathfrak{S}) = 0$$

wegen (7). Daher hat man statt (10) einfacher

$$(12) \quad A = - \sum_\tau \frac{e_\tau}{2 m_\tau c} \left\{ (\mathfrak{p}_\tau \mathfrak{A}) - \frac{2\pi i \nu}{c} (\mathfrak{p}_\tau \mathfrak{A})(\mathfrak{r}_\tau \mathfrak{S}) \right\}.$$

Wir führen nun das elektrische Moment

$$(13) \quad \mathfrak{P} = \sum_\tau e_\tau \mathfrak{r}_\tau$$

und den Tensor

$$(14) \quad \Theta_{xy} = \sum_\tau \frac{e_\tau}{m_\tau c} x_\tau p_{\tau y}$$

ein. Dann wird die in (12) auftretende Summe

$$(15) \quad \sum_\tau \frac{e_\tau}{m_\tau} \mathfrak{p}_\tau = \sum_\tau e_\tau \dot{\mathfrak{r}}_\tau = \dot{\mathfrak{P}}.$$

Ferner ist

$$(16) \quad M_z = \frac{1}{2} (\Theta_{xy} - \Theta_{yx}) = \sum_\tau \frac{e_\tau}{2 m_\tau c} (x_\tau p_{\tau y} - y_\tau p_{\tau x})$$

die z -Komponente des magnetischen Moments

$$(17) \quad \mathfrak{M} = \sum_\tau \frac{e_\tau}{2 m_\tau c} (\mathfrak{r}_\tau \times \mathfrak{p}_\tau).$$

Mit diesen Begriffen wird (12)

$$(18) \quad A = -\frac{1}{2c} (\mathfrak{P} \mathfrak{A}) + \frac{2\pi i \nu}{c} \frac{1}{2} \sum_{xy} \Theta_{xy} \mathfrak{E}_x \mathfrak{A}_y.$$

Die Störungsfunktion (9) erlaubt nun in genau derselben Weise wie im vorigen Paragraphen das durch das elektromagnetische Feld induzierte mittlere Moment zu berechnen. Dabei sind überall in \mathfrak{P} und Θ_{xy} die Koordinaten des ungestörten Zustandes eingesetzt zu denken; für diesen hat man aber

$$(19) \quad \mathfrak{P}^{(r,s)} = 2\pi i \nu^{(r,s)} \mathfrak{P}^{(r,s)}.$$

Somit erhält man

$$g_r \overline{\mathfrak{P}^{(r,r)}} = e^{2\pi i \nu t} \frac{2\pi i}{c} \frac{1}{2} \sum_s \left\{ \frac{Sp[\mathfrak{P}^{(r,s)} ((\mathfrak{P}^{(s,r)} \mathfrak{A}) \nu^{(s,r)} - \sum_{xy} \hat{\Theta}_{xy}^{(s,r)} \mathfrak{E}_x \mathfrak{A}_y \nu)]}{h(\nu^{(s,r)} + \nu)} - \frac{Sp[(\mathfrak{P}^{(r,s)} \mathfrak{A}) \nu^{(r,s)} - \sum_{xy} \hat{\Theta}_{xy}^{(r,s)} \mathfrak{E}_x \mathfrak{A}_y \nu] \mathfrak{P}^{(s,r)}}{h(\nu^{(r,s)} + \nu)} \right\} \\ + e^{-2\pi i \nu t} \frac{2\pi i}{c} \frac{1}{2} \sum_s \left\{ \frac{Sp[\mathfrak{P}^{(r,s)} ((\mathfrak{P}^{(s,r)} \mathfrak{A}^*) \nu^{(s,r)} + \sum_{xy} \hat{\Theta}_{xy}^{(s,r)} \mathfrak{E}_x \mathfrak{A}_y^* \nu)]}{h(\nu^{(s,r)} - \nu)} - \frac{Sp[(\mathfrak{P}^{(r,s)} \mathfrak{A}^*) \nu^{(r,s)} + \sum_{xy} \hat{\Theta}_{xy}^{(r,s)} \mathfrak{E}_x \mathfrak{A}_y^* \nu] \mathfrak{P}^{(s,r)}}{h(\nu^{(r,s)} - \nu)} \right\}.$$

Jetzt ziehen wir die Relationen § 30 heran und finden

$$g_r \overline{\mathfrak{P}^{(r,r)}} = \frac{2\pi i}{c} \frac{1}{2} \mathfrak{A} e^{2\pi i \nu t} \frac{1}{3h} \sum_s \|\mathfrak{P}^{(r,s)}\|^2 \left(\frac{\nu^{(s,r)}}{\nu^{(s,r)} + \nu} - \frac{\nu^{(r,s)}}{\nu^{(r,s)} + \nu} \right) \\ - \frac{2\pi i \nu}{c} \frac{1}{2} (\mathfrak{E} \times \mathfrak{A}) e^{2\pi i \nu t} \frac{1}{3h} \sum_s \left(\frac{Sp(\mathfrak{P}^{(r,s)} \mathfrak{M}^{(s,r)})}{\nu^{(s,r)} + \nu} - \frac{Sp(\mathfrak{M}^{(r,s)} \mathfrak{P}^{(s,r)})}{\nu^{(r,s)} + \nu} \right) \\ + \frac{2\pi i}{c} \frac{1}{2} \mathfrak{A}^* e^{-2\pi i \nu t} \frac{1}{3h} \sum_s \|\mathfrak{P}^{(r,s)}\|^2 \left(\frac{\nu^{(s,r)}}{\nu^{(s,r)} - \nu} - \frac{\nu^{(r,s)}}{\nu^{(r,s)} - \nu} \right) \\ + \frac{2\pi i \nu}{c} \frac{1}{2} (\mathfrak{E} \times \mathfrak{A}^*) e^{-2\pi i \nu t} \frac{1}{3h} \sum_s \left(\frac{Sp(\mathfrak{P}^{(r,s)} \mathfrak{M}^{(s,r)})}{\nu^{(s,r)} - \nu} - \frac{Sp(\mathfrak{M}^{(r,s)} \mathfrak{P}^{(s,r)})}{\nu^{(r,s)} - \nu} \right).$$

Nun ist, wie leicht zu sehen,

$$(20) \quad Sp(\mathfrak{P}^{(r,s)} \mathfrak{M}^{(s,r)}) = [Sp(\mathfrak{M}^{(r,s)} \mathfrak{P}^{(s,r)})]^*.$$

Wir setzen

$$(21) \quad Sp(\mathring{\mathfrak{P}}^{(r,s)} \mathring{\mathfrak{M}}^{(s,r)}) = R^{(r,s)} + i J^{(r,s)};$$

dann wird

$$(21a) \quad Sp(\mathring{\mathfrak{M}}^{(r,s)} \mathring{\mathfrak{P}}^{(s,r)}) = R^{(r,s)} - i J^{(r,s)}.$$

Nunmehr ergibt eine einfache Rechnung

$$(22) \quad g_r \overline{\mathfrak{P}^{(r,r)}} = \\ - \frac{2\pi i \nu}{c} \frac{1}{2} (\mathring{\mathfrak{U}} e^{2\pi i \nu t} - \mathring{\mathfrak{U}}^* e^{-2\pi i \nu t}) \frac{2}{3\hbar} \sum_s \|\mathring{\mathfrak{P}}^{(r,s)}\|^2 \frac{\nu^{(s,r)}}{\nu^{(s,r)^2} - \nu^2} \\ - \frac{2\pi i \nu}{c} \frac{1}{2} ((\mathring{\mathfrak{C}} \times \mathring{\mathfrak{U}}) e^{2\pi i \nu t} - (\mathring{\mathfrak{C}} \times \mathring{\mathfrak{U}}^*) e^{-2\pi i \nu t}) \frac{2}{3\hbar} \sum_s R^{(r,s)} \frac{\nu^{(s,r)}}{\nu^{(s,r)^2} - \nu^2} \\ + \frac{2\pi i \nu}{c} \frac{1}{2} ((\mathring{\mathfrak{C}} \times \mathring{\mathfrak{U}}) e^{2\pi i \nu t} + (\mathring{\mathfrak{C}} \times \mathring{\mathfrak{U}}^*) e^{-2\pi i \nu t}) \frac{2i}{3\hbar} \sum_s J^{(r,s)} \frac{\nu}{\nu^{(s,r)^2} - \nu^2}.$$

Wir setzen nun

$$(23) \quad \left\{ \begin{array}{l} \alpha = \frac{2}{3\hbar g_r} \sum_s \|\mathring{\mathfrak{P}}^{(r,s)}\|^2 \frac{\nu^{(s,r)}}{\nu^{(s,r)^2} - \nu^2}, \\ \beta = \frac{2}{3\hbar g_r} \sum_s R^{(r,s)} \frac{\nu^{(s,r)}}{\nu^{(s,r)^2} - \nu^2}, \\ \gamma = \frac{1}{2\pi \nu} \cdot \frac{2}{3\hbar g_r} \sum_s J^{(r,s)} \frac{\nu}{\nu^{(s,r)^2} - \nu^2}. \end{array} \right.$$

Dann wird nach (8)

$$(24) \quad \overline{\mathfrak{P}^{(r,r)}} = \alpha \mathring{\mathfrak{C}} + \beta \mathring{\mathfrak{B}} - \gamma \mathring{\mathfrak{A}},$$

wo $\mathring{\mathfrak{C}}$, $\mathring{\mathfrak{B}}$, $\mathring{\mathfrak{A}}$ die Werte der Feldgrößen im Nullpunkt bedeuten.

Es genügt aber nicht, das mittlere elektrische Moment zu berechnen, das durch die Lichtwelle erregt wird; von gleicher Größenordnung ist das vom Licht erzeugte magnetische Moment. Dabei haben wir die Substanz als unmagnetisch vorausgesetzt, d. h. angenommen, daß

$$(25) \quad Sp \mathring{\mathfrak{M}}^{(r,r)} = 0$$

ist. In der Störungsrechnung für $\mathring{\mathfrak{M}}^{(1)}$ genügt es offenbar bei der

hier angestrebten Annäherung, von der Störungsfunktion nur den Anteil zu berücksichtigen, der im Grenzfall unendlicher Wellenlänge übrig bleibt, also nach (18)

$$A = -\frac{1}{2c} (\mathfrak{P} \mathfrak{M})$$

zu setzen. Dann wird

$$g_r \overline{\mathfrak{M}^{(r,r)}} = e^{2\pi i \nu t} \frac{2\pi i}{c} \frac{1}{2} \sum_s \left\{ \frac{Sp [\mathfrak{M}^{(r,s)} (\mathfrak{P}^{(s,r)} \mathfrak{U}) \nu^{(s,r)}]}{h(\nu^{(s,r)} + \nu)} - \frac{Sp [(\mathfrak{P}^{(r,s)} \mathfrak{U}) \mathfrak{M}^{(s,r)} \nu^{(r,s)}]}{h(\nu^{(r,s)} + \nu)} \right\} + e^{-2\pi i \nu t} \frac{2\pi i}{c} \frac{1}{2} \sum_s \left\{ \frac{Sp [\mathfrak{M}^{(r,s)} (\mathfrak{P}^{(s,r)} \mathfrak{U}^*) \nu^{(s,r)}]}{h(\nu^{(s,r)} - \nu)} - \frac{Sp [(\mathfrak{P}^{(r,s)} \mathfrak{U}^*) \mathfrak{M}^{(r,s)} \nu^{(r,s)}]}{h(\nu^{(r,s)} - \nu)} \right\}.$$

Nach § 30 gibt das

$$(26) \quad g_r \overline{\mathfrak{M}^{(r,r)}} = -\frac{2\pi i \nu}{c} \frac{1}{2} (\mathfrak{U} e^{2\pi i \nu t} - \mathfrak{U}^* e^{-2\pi i \nu t}) \frac{2}{3h} \sum_s R^{(r,s)} \frac{\nu^{(s,r)}}{\nu^{(s,r)^2} - \nu^2} - \frac{2\pi i \nu}{c} \frac{1}{2} (\mathfrak{U} e^{2\pi i \nu t} + \mathfrak{U}^* e^{-2\pi i \nu t}) \frac{2i}{3h\nu} \sum_s J^{(r,s)} \frac{\nu^{(s,r)^2}}{\nu^{(s,r)^2} - \nu^2}.$$

Nun ist nach § 30, (23) die Invariante

$$(27) \quad Sp (\mathfrak{P} \mathfrak{M})^{(r,r)} = \sum_s Sp (\mathfrak{P}^{(r,s)} \mathfrak{M}^{(s,r)}) = \sum_s R^{(r,s)} + i \sum_s J^{(r,s)}$$

reell, weil \mathfrak{P} und \mathfrak{M} hermitisch und vertauschbar sind.

Mithin muß

$$(28) \quad \sum_s J^{(r,s)} = 0$$

sein. Daher hat man für die zweite, in (26) vorkommende Summe

$$(29) \quad \sum_s J^{(r,s)} \frac{\nu^{(s,r)^2}}{\nu^{(s,r)^2} - \nu^2} = \nu^2 \sum_s J^{(r,s)} \frac{1}{\nu^{(s,r)^2} - \nu^2}.$$

Unter Benutzung der Definition (23) wird also:

$$(30) \quad \overline{\mathfrak{M}^{(r,r)}} = \beta \mathfrak{E} + \gamma \mathfrak{E}^{\dot{}}$$

Die Größen (24) und (30) sind in den MAXWELLSchen Gleichungen als mittlere Polarisation und Magnetisierung einzuführen; man hat für den Verschiebungs- und Induktionsvektor

$$(31) \quad \begin{cases} \mathfrak{D} = \mathfrak{E} + 4\pi N \overline{\mathfrak{P}} = \varepsilon \mathfrak{E} + \sigma \mathfrak{H} - \frac{\rho}{2c} \mathfrak{H}, \\ \mathfrak{H} = \mathfrak{H} + 4\pi N \overline{\mathfrak{M}} = \mathfrak{H} + \sigma \mathfrak{E} + \frac{\rho}{2c} \mathfrak{E}, \end{cases}$$

wo N die Anzahl Molekeln pro Volumeneinheit bedeutet und

$$(32) \quad \varepsilon = 1 + 4\pi N \alpha, \quad \sigma = 4\pi N \beta, \quad \frac{\rho}{2c} = 4\pi N \gamma$$

gesetzt ist. Die MAXWELLSchen Gleichungen

$$(33) \quad \text{rot } \mathfrak{H} = \frac{1}{c} \dot{\mathfrak{D}}, \quad \text{rot } \mathfrak{E} = -\frac{1}{c} \dot{\mathfrak{H}}$$

liefern für eine ebene Welle

$$(34) \quad -\mathfrak{E} \times \mathfrak{H} = \mathfrak{D}, \quad \mathfrak{E} \times \mathfrak{E} = \mathfrak{H},$$

und das gibt zusammen mit (31) in der angestrebten Näherung

$$(35) \quad \begin{cases} -\mathfrak{E} \times \mathfrak{H} = \varepsilon \mathfrak{E} + \left(\sigma - \frac{2\pi i \nu \rho}{2c} \right) \mathfrak{H}, \\ \mathfrak{E} \times \mathfrak{E} = \mathfrak{H} + \left(\sigma + \frac{2\pi i \nu \rho}{2c} \right) \mathfrak{E}. \end{cases}$$

Multipliziert man diese Gleichungen skalar mit \mathfrak{E} , so erhält man

$$\begin{aligned} \varepsilon \mathfrak{E} \mathfrak{E} + \left(\sigma - \frac{2\pi i \nu \rho}{2c} \right) \mathfrak{H} \mathfrak{E} &= 0, \\ \mathfrak{H} \mathfrak{E} + \left(\sigma + \frac{2\pi i \nu \rho}{2c} \right) \mathfrak{E} \mathfrak{E} &= 0, \end{aligned}$$

und da die Determinante nicht verschwindet, folgt die Transversalität der Welle:

$$(36) \quad \mathfrak{E} \mathfrak{E} = 0, \quad \mathfrak{H} \mathfrak{E} = 0.$$

Eliminiert man nun \mathfrak{H} aus (35), so bekommt man nach einfacher

Rechnung wegen $\mathfrak{E} = n \mathfrak{s}$, $\lambda \nu = \frac{c}{n}$:

$$(37) \quad \mathfrak{E} (n^2 - \varepsilon) = i \cdot \frac{2\pi \rho}{\lambda} (\mathfrak{E} \times \mathfrak{s}).$$

Diese Beziehung bedeutet aber optische Aktivität, d. h. zirkulare Doppelbrechung. Denn pflanzt sich die Welle etwa in der z -Rich-

tung fort, so besagt (37), daß

$$E_x(n^2 - \varepsilon) = i \frac{2\pi\rho}{\lambda} E_y,$$

$$E_y(n^2 - \varepsilon) = -i \frac{2\pi\rho}{\lambda} E_x,$$

woraus folgt

$$n^2 = \varepsilon \pm \frac{2\pi\rho}{\lambda}, \quad \frac{E_y}{E_x} = \mp i.$$

Die *zirkulare Doppelbrechung* beträgt also, wenn mit λ_0 die Vakuumwellenlänge bezeichnet wird,

$$(38) \quad \Delta n = \frac{2\pi\rho}{\lambda n} = \frac{2\pi\rho}{\lambda_0}$$

und die *Drehung der Polarisationssebene pro Längeneinheit*

$$(39) \quad \vartheta = \frac{2\pi\nu}{c} \cdot \frac{\Delta n}{2} = \frac{2\pi^2}{n^2 \lambda^2} \rho = \frac{2\pi^2}{\lambda_0^2} \rho.$$

Die durch (32) und (23) definierte Größe ρ ist also der *Drehungsparameter*. Die Größe σ bestimmt, wie man leicht sieht, das Amplitudenverhältnis und die relative Phase von elektrischem und magnetischem Vektor und ist der Beobachtung wohl nicht zugänglich.

Man sieht, daß ρ nur von Null verschieden ist, wenn die Molekel keine Spiegelung gestattet; denn wenn das der Fall ist, so verschwinden nach der zu § 30, (22) gemachten Bemerkung die Größen

$$Sp(\mathfrak{R}^{(r,s)} \mathfrak{M}^{(s,r)}) = R^{(r,s)} + i J^{(r,s)}$$

und damit auch γ und ρ .

Hier erhebt sich der Einwand: Da jede Molekel aus punktförmigen Kraftzentren (Kerne und Elektronen) besteht, die nach dem COULOMBSchen Gesetz aufeinander wirken, ist die Energiefunktion

$$H = \sum_k \frac{1}{2m_k} p_k^2 + \frac{1}{2} \sum_{kl} \frac{e_k e_l}{r_{kl}}$$

immer invariant gegen eine Spiegelung (Vertauschung aller Vektoren r_k, p_k mit $-r_k, -p_k$). Folglich könnte es gar keine optisch aktiven Molekeln geben, was der Erfahrung widerspricht. Die

Lösung dieser Paradoxie beruht auf einer Überlegung, die wir erst viel später ausführlich durchführen werden und hier nur andeuten können.

Das Studium der Quantenmechanik von Molekeln zeigt, daß die elementare Vorstellung zu recht besteht, wonach die Kerne kleine Schwingungen um Gleichgewichtslagen ausführen, die durch das Minimum der Energie der Elektronenbewegung bei festgedachten Kernen bestimmt sind. Es kann nun vorkommen, daß es mehrere solche Gleichgewichts-Konfigurationen gibt, und hier interessiert besonders der Fall, daß zwei Gleichgewichts-Konfigurationen durch Spiegelung an einer Ebene auseinander hervorgehen (Rechts- und Linksmodifikation).

In der klassischen Mechanik würden dann die Kerne Schwingungen um die eine oder die andere Konfiguration ausführen können, es würde aber ohne Zufuhr von beträchtlicher Energie nie möglich sein, daß das System von der einen Lage in die andere überschnappt; denn dazu muß eine bestimmte Energieschwelle überwunden werden. Anders in der Quantenmechanik¹: Hier ist eine endliche Energieschwelle kein Hindernis für die Umformung der Rechts- in die Linksmodifikation; diese vollzieht sich vielmehr stets von selbst hin und zurück, wobei die Wahrscheinlichkeit eine periodische Funktion der Zeit ist (s. hierzu die Ausführungen in Kap. 6, § 64). Über lange Zeiten existiert also in der Tat die Molekel stets nur im „razemischen Gemisch“, in dem beide Formen gleich stark vertreten sind und das daher optisch inaktiv ist, wie es die Spiegelsymmetrie der Energiefunktion verlangt. Aber oft ist die Umwandlungsperiode äußerst lang, und dann kann man die eine Form experimentell isolieren und beobachten; sie wird natürlich optisch aktiv sein. Im Laufe der Zeit entsteht von selbst die andere Modifikation und die Aktivität sinkt zu Null herab².

Ein solcher aktiver Zustand der Molekel kann also beschrieben werden durch eine feste Konfiguration der Kerne ohne Symmetrieebene; denn da alle Elektronen gleich sind, kann die Asymmetrie nur auf den Kernen beruhen. Daraus geht auch sofort hervor, daß die Mindestzahl der Kerne für die Existenz der optischen

¹ HUND, J. F.: Z. Phys. Bd. 43, S. 806. 1927.

² Die genauere Durchführung dieses Gedankens ist von L. ROSENFELD gegeben worden: Z. Phys. Bd. 52, S. 161. 1929.

Aktivität *vier* beträgt. Die bekanntesten aktiven Verbindungen sind solche mit mehr als *vier* Kernen, nämlich einem Kohlenstoffatom, umgeben von vier verschiedenen Atomen oder Atomgruppen.

§ 48. Elektrische Doppelbrechung (KERR-Effekt)¹. Wir haben schon in § 46 darauf hingewiesen, daß die dort angegebene Grundformel (9) für das von einer Lichtwelle in einem atomaren System erregte elektrische Moment im allgemeinen Doppelbrechung ergibt. Diese tritt z. B. dann tatsächlich ein, wenn eine an sich frei drehbare Molekel in ein konstantes homogenes elektrisches Feld gebracht wird; wir haben dann in der Formel § 46, (9) unter $\hat{\mathfrak{P}}$ das Moment der Molekel unter der Wirkung des konstanten elektrischen Feldes zu verstehen. Wir wollen die *z*-Achse des Koordinatensystems in die Richtung dieses konstanten Feldes *E* legen. Mit $\hat{\mathfrak{p}}$ bezeichnen wir das Moment der Molekel im feldfreien Zustande (früher mit $\hat{\mathfrak{P}}$ identisch). Dann können wir einmal die Wirkung der Lichtwelle auf die kohärente Strahlung durch die Formel § 46, (9) ausdrücken, wenn wir darin $n = m$ setzen:

$$(1) \quad \hat{\mathfrak{P}}_{nk}^{(1)} = \frac{1}{2} e^{2\pi i \nu t} \sum_k \left(\frac{\hat{\mathfrak{P}}_{nk} (\hat{\mathfrak{P}}_{kn} \hat{\mathfrak{C}})}{\hbar (\nu_{kn} + \nu)} - \frac{(\hat{\mathfrak{P}}_{nk} \hat{\mathfrak{C}}) \hat{\mathfrak{P}}_{kn}}{\hbar (\nu_{nk} + \nu)} \right) + \frac{1}{2} e^{-2\pi i \nu t} \sum_k \left(\frac{\hat{\mathfrak{P}}_{nk} (\hat{\mathfrak{P}}_{kn} \hat{\mathfrak{C}}^*)}{\hbar (\nu_{kn} - \nu)} - \frac{(\hat{\mathfrak{P}}_{nk} \hat{\mathfrak{C}}^*) \hat{\mathfrak{P}}_{kn}}{\hbar (\nu_{nk} - \nu)} \right),$$

sodann $\hat{\mathfrak{P}}$ als Funktion von *E* angeben auf Grund der Betrachtungen von § 42; man hat eine Entwicklung

$$(2) \quad \hat{\mathfrak{P}} = \hat{\mathfrak{p}} + \hat{\mathfrak{p}}^{(1)} E + \hat{\mathfrak{p}}^{(2)} E^2 + \dots,$$

wobei sich $\hat{\mathfrak{p}}^{(1)}$, $\hat{\mathfrak{p}}^{(2)}$, ... durch $\hat{\mathfrak{p}}$ ausdrücken lassen. Der Ausdruck von $\hat{\mathfrak{p}}^{(1)}$ ist in § 42, (7) angegeben (wobei statt $\hat{\mathfrak{p}}$, $\hat{\mathfrak{p}}^{(1)}$ dort natürlich $\hat{\mathfrak{P}}$, $\hat{\mathfrak{P}}$ geschrieben ist). Wir werden von diesen Ausdrücken nur benutzen, daß

$$(3) \quad \begin{aligned} &\text{in } \hat{\mathfrak{p}}^{(1)} \text{ Produkte von 2 Komponenten von } \hat{\mathfrak{p}}, \\ &\text{in } \hat{\mathfrak{p}}^{(2)} \text{ Produkte von 3 Komponenten von } \hat{\mathfrak{p}} \end{aligned}$$

¹ Zu §§ 48, 49 vergl. auch die Arbeiten von R. DE L. KRONIG: Z. f. Phys. Bd. 45, S. 458, 508. 1927; Bd. 47, S. 702. 1927.

vorkommen. Ferner haben wir zu beachten, daß $\overset{\circ}{\mathfrak{P}}$ zu einem im Raume frei drehbaren Systeme gehört, $\overset{\circ}{\mathfrak{P}}$ und damit $\overset{(1)}{\mathfrak{P}}, \overset{(2)}{\mathfrak{P}}, \dots$ aber zu einem System, daß nur um die z -Achse drehbar ist und außerdem eine durch diese gehende Ebene, etwa die xz -Ebene, zur Spiegelebene hat.

Die Energie im elektrischen Felde E ist unter der Voraussetzung, daß die Molekel nur den quadratischen STARK-Effekt zeigt, nach § 42, (5a)

$$(4) \quad W_m^{(r)} = \overset{\circ}{W}^{(r)} + (A^{(r)} + B^{(r)} m^2) E^2 = \overset{\circ}{W}^{(r)} + \overset{(2)}{W}_m^{(r)} E^2,$$

wobei m die magnetische Quantenzahl bedeutet, r alle übrigen Quantenzahlen vertritt. Wie schon in § 42 bemerkt wurde, sieht man hieraus, daß je zwei Zustände, die sich nur durch das Vorzeichen von m unterscheiden, gleiche Energie haben; das System bleibt also auch im Felde entartet.

Wie in § 46, (16) bilden wir den Mittelwert des Momentes

$$(5) \quad \overline{\mathfrak{P}}_m^{(r)} = \frac{1}{2} s p_m \overset{(1)}{\mathfrak{P}}_{m,m}^{(r)} = \frac{1}{2} (\overset{(1)}{\mathfrak{P}}_{+m}^{(r)} + \overset{(1)}{\mathfrak{P}}_{-m}^{(r)});$$

dabei bedeutet $s p_m$ hier und im folgenden stets die Summe über die beiden Werte $+m$ und $-m$; diese Größe (5) ist relativ-invariant. Nach § 30, (10) folgt aus den oben angegebenen Symmetrieelementen für $\overset{\circ}{\mathfrak{P}}$, daß der Ausdruck (1) für die Komponenten $\overline{X}_m^{(r)}, \overline{Y}_m^{(r)}, \overline{Z}_m^{(r)}$ von $\overline{\mathfrak{P}}_m^{(r)}$ die einfache Form

$$(6) \quad \overline{X}_m^{(r)} = \alpha_m^{(r)} \mathfrak{E}_x, \quad \overline{Y}_m^{(r)} = \alpha_m^{(r)} \mathfrak{E}_y, \quad \overline{Z}_m^{(r)} = \gamma_m^{(r)} \mathfrak{E}_z$$

hat, wobei die Formel § 46, (3) berücksichtigt ist und die Abkürzungen

$$(7) \quad \left\{ \begin{array}{l} \alpha_m^{(r)} = \frac{1}{h} \sum_{r', m'} s p_m (\overset{\circ}{X}_{m, m'}^{(r, r')} \overset{\circ}{X}_{m', m}^{(r', r)}) \frac{\gamma_{m', m}^{(r', r)}}{\gamma_{m', m}^{(r', r) 2} - \gamma^2}, \\ \gamma_m^{(r)} = \frac{1}{h} \sum_{r', m} s p_m (\overset{\circ}{Z}_{m, m'}^{(r, r')} \overset{\circ}{Z}_{m', m}^{(r', r)}) \frac{\gamma_{m', m}^{(r', r)}}{\gamma_{m', m}^{(r', r) 2} - \gamma^2} \end{array} \right.$$

gebraucht sind.

Ist \mathfrak{P} das mittlere Moment der Volumeneinheit, gemittelt über alle im Gase vorkommenden Zustände, so hat man

$$(8) \quad \mathfrak{P} = \sum_{r, m} N_m^{(r)} \overline{\mathfrak{P}}_m^{(r)}$$

zu setzen, wo $N_m^{(r)}$ die Anzahl der Molekel im Zustande (r, m) ist (die Summation über m soll von $-j$ bis $+j$ erstreckt werden). Während diese in der Dispersionstheorie (§ 46) nur von der Temperatur abhing, ist sie jetzt auch von der Feldstärke E abhängig, und zwar hat man unter Annahme des BOLTZMANNschen Verteilungsgesetzes:

$$(9) \quad N_m^{(r)} = N \frac{e^{-\frac{W_m^{(r)}}{kT}}}{\sum_{r,m} e^{-\frac{W_m^{(r)}}{kT}}}.$$

Nach (6) sind die Komponenten von \mathfrak{P} :

$$(10) \quad X = \alpha \mathfrak{E}_x, \quad Y = \alpha \mathfrak{E}_y, \quad Z = \gamma \mathfrak{E}_z,$$

mit

$$(11) \quad \alpha = \sum_{r,m} N_m^{(r)} \alpha_m^{(r)}, \quad \gamma = \sum_{r,m} N_m^{(r)} \gamma_m^{(r)}.$$

Die Formeln (10) verbinden wir mit der optischen Grundgleichung § 46, (15); aus dieser folgt zunächst durch skalare Multiplikation mit \mathfrak{s} :

$$\mathfrak{s} \mathfrak{P} = -\frac{1}{4\pi} (\mathfrak{s} \mathfrak{E}),$$

sie läßt sich also schreiben

$$4\pi \mathfrak{P} = (n^2 - 1) \mathfrak{E} - n^2 \mathfrak{s} (\mathfrak{E} \mathfrak{s}),$$

oder, wenn wir beiderseits \mathfrak{E} addieren und den Verschiebungsvektor $\mathfrak{D} = \mathfrak{E} + 4\pi \mathfrak{P}$ einführen:

$$(12) \quad \mathfrak{D} = n^2 \{ \mathfrak{E} - \mathfrak{s} (\mathfrak{E} \mathfrak{s}) \}.$$

Wir können ohne Beschränkung der Allgemeinheit annehmen, daß die Wellennormale in der xz -Ebene liegt; bildet sie den Winkel ϑ mit der z -Achse, so haben wir

$$(13) \quad \mathfrak{s}_x = s = \sin \vartheta, \quad \mathfrak{s}_y = 0, \quad \mathfrak{s}_z = c = \cos \vartheta.$$

Nun wird nach (10)

$$(14) \quad \mathfrak{D}_x = n_o^2 \mathfrak{E}_x, \quad \mathfrak{D}_y = n_n^2 \mathfrak{E}_y, \quad \mathfrak{D}_z = n_e^2 \mathfrak{E}_z,$$

wo die Abkürzungen

$$(15) \quad n_o^2 = 1 + 4\pi\alpha, \quad n_e^2 = 1 + 4\pi\gamma$$

eingeführt sind. Aus (12) und (14) ergibt sich mit Rücksicht auf (13)

$$(16) \quad \begin{cases} \mathfrak{D}_x = cn^2 \left(\mathfrak{D}_x \frac{c}{n_o^2} - \mathfrak{D}_z \frac{s}{n_e^2} \right), \\ \mathfrak{D}_y = \frac{n^2}{n_o^2} \mathfrak{D}_y, \\ \mathfrak{D}_z = -sn^2 \left(\mathfrak{D}_x \frac{c}{n_o^2} - \mathfrak{D}_z \frac{s}{n_e^2} \right). \end{cases}$$

Daher pflanzen sich zwei Wellen fort, eine *ordentliche*, die senkrecht zur Einfallsebene schwingt ($\mathfrak{D}_x = \mathfrak{D}_z = 0, \mathfrak{D}_y \neq 0$) mit dem Brechungsindex n_o , und eine *außerordentliche*, die parallel zur Einfallsebene schwingt ($\mathfrak{D}_x + \mathfrak{D}_z \neq 0, \mathfrak{D}_y = 0$); für diese ergibt sich zunächst

$$s \mathfrak{D}_x + c \mathfrak{D}_z = 0,$$

d. h. Transversalität der Schwingung von \mathfrak{D} , sodann das Gesetz für den Brechungsindex

$$(17) \quad \frac{n^2}{n_o^2} c^2 + \frac{n^2}{n_e^2} s^2 = 1.$$

Ist α von γ , also nach (15) n_o von n_e nur wenig verschieden, so folgt hieraus für die Doppelbrechung

$$(18) \quad \Delta n = \frac{s^2 n_e^2 - n_o^2}{2n} = \frac{2\pi}{n} (\gamma - \alpha) \sin^2 \vartheta = \frac{2\pi}{n} K \sin^2 \vartheta,$$

wo \bar{n} ein mittlerer Brechungsindex ist. Es kommt also darauf an, die Größe

$$(19) \quad K = \gamma - \alpha$$

als Funktion des Feldes E zu berechnen; sie ist ein Maß des KERR-Effekts. Hierzu bilden wir zunächst

$$(20) \quad K_m^{(r)} = \gamma_m^{(r)} - \sigma_m^{(r)}$$

und dann nach (9) und (11)

$$(21) \quad K = \sum_{r,m} N_m^{(r)} K_m^{(r)} = N^{\gamma,m} \frac{\sum_r K_m^{(r)} e^{-\frac{W_m^{(r)}}{kT}}}{\sum_r e^{-\frac{W_m^{(r)}}{kT}}}.$$

Hier sind $K_m^{(r)}$ und $W_m^{(r)}$ Funktionen von E , die wir bis zu Gliedern mit E^2 entwickeln. Den Nenner schreiben wir kurz

$$(22) \quad S = \sum_{r,m} e^{-\frac{W_m^{(r)}}{kT}},$$

und wir werden sehen, daß wir nur seinen Wert für $E = 0$ benötigen.

Nach (21) und (4) wird

$$(23) \quad K = \frac{N}{S} \left\{ \sum_r e^{-\frac{\hat{W}^{(r)}}{kT}} \sum_n K_m^{(r)} - \frac{E^2}{kT} \sum_r e^{-\frac{\hat{W}^{(r)}}{kT}} \sum_n W_m^{(r)} K_m^{(r)} \right\}.$$

Wir haben nach (20) und (7)

$$(24) \quad \sum_m K_m^{(r)} = \frac{2}{h} \sum_{r', m, m'} (\hat{Z}_{m m'}^{(r, r')} \hat{Z}_{m' m}^{(r', r)} - \hat{X}_{m m'}^{(r, r')} \hat{X}_{m' m}^{(r', r)}) \frac{\nu_{m m'}^{(r', r)}}{\nu_{m m'}^{(r, r)^2} - \nu^2};$$

hier sind nicht nur die Komponenten von \mathfrak{B} , sondern auch die Frequenzen $\nu_{m m'}^{(r, r')}$ Funktionen von E . Man hat nach (4)

$$(25) \quad \frac{\nu_{m m'}^{(r', r)}}{\nu_{m m'}^{(r', r)^2} - \nu^2} = \frac{\hat{\nu}_{m m'}^{(r', r)}}{\hat{\nu}_{m m'}^{(r', r)^2} - \nu^2} (1 - \lambda_{m m'}^{(r', r)} E^2),$$

wo

$$(26) \quad \lambda_{m m'}^{(r', r)} = (W_{m'}^{(r')} - W_m^{(r)}) \frac{1}{h} \frac{\hat{\nu}_{m m'}^{(r', r)^2} + \nu^2}{\hat{\nu}_{m m'}^{(r', r)^2} - \nu^2}.$$

Ferner setzen wir

$$(27) \quad \begin{aligned} & \hat{Z}_{m m'}^{(r, r')} \hat{Z}_{m' m}^{(r', r)} - \hat{X}_{m m'}^{(r, r')} \hat{X}_{m' m}^{(r', r)} \\ &= A_{m m'}^{(r, r')} + B_{m m'}^{(r, r')} E + C_{m m'}^{(r, r')} E^2, \end{aligned}$$

wo

$$(28) \quad \left\{ \begin{aligned} A_{m m'}^{(r, r')} &= \hat{z}_{m m'}^{(r, r')} \hat{z}_{m' m}^{(r', r)} - \hat{x}_{m m'}^{(r, r')} \hat{x}_{m' m}^{(r', r)}, \\ B_{m m'}^{(r, r')} &= \hat{z}_{m m'}^{(r, r')} \hat{z}_{m' m}^{(1)(r', r)} + \hat{z}_{m m'}^{(1)(r, r')} \hat{z}_{m' m}^{(r', r)} \\ &\quad - \hat{x}_{m m'}^{(r, r')} \hat{x}_{m' m}^{(1)(r', r)} - \hat{x}_{m m'}^{(1)(r, r')} \hat{x}_{m' m}^{(r', r)}, \\ C_{m m'}^{(r, r')} &= \hat{z}_{m m'}^{(r, r')} \hat{z}_{m' m}^{(2)(r', r)} + \hat{z}_{m m'}^{(1)(r, r')} \hat{z}_{m' m}^{(1)(r', r)} + \hat{z}_{m m'}^{(2)(r, r')} \hat{z}_{m' m}^{(r', r)} \\ &\quad - \hat{x}_{m m'}^{(r, r')} \hat{x}_{m' m}^{(2)(r', r)} - \hat{x}_{m m'}^{(1)(r, r')} \hat{x}_{m' m}^{(1)(r', r)} - \hat{x}_{m m'}^{(2)(r, r')} \hat{x}_{m' m}^{(r', r)}. \end{aligned} \right.$$

Nach (3) kommen vor

$$(29) \quad \left\{ \begin{array}{l} \text{in } A \text{ Produkte von 2 Komponenten von } \hat{p}, \\ \text{in } B \text{ Produkte von 3 Komponenten von } \hat{p}, \\ \text{in } C \text{ Produkte von 4 Komponenten von } \hat{p}. \end{array} \right.$$

Da \hat{p} zu einem im Raume frei drehbaren System gehört, ergibt sich aus den Überlegungen von § 30, daß

$$Sp(\hat{z}^{(r, r')} \hat{z}^{(r', r)}) = Sp(\hat{x}^{(r, r')} \hat{x}^{(r', r)}),$$

wo jetzt das Zeichen Sp sich auf die Summation über alle m -Werte bezieht. Das gibt aber

$$(30) \quad \sum_{m m'} A_{m m'}^{(r, r')} = Sp(\hat{z}^{(r, r')} \hat{z}^{(r', r)} - \hat{x}^{(r, r')} \hat{x}^{(r', r)}) = 0,$$

und ebenso wird

$$(31) \quad \sum_{m m'} B_{m m'}^{(r, r')} = 0.$$

Also erhalten wir nach (24), (25), (27)

$$(32) \quad \sum_m K_m^{(r)} = C^{(r)} \cdot E^2,$$

mit der Abkürzung

$$(33) \quad C^{(r)} = \frac{2}{h} \sum_{r'} \frac{\hat{v}^{(r', r)}}{\hat{v}^{(r', r)^2} - v^2} \sum_{m, m'} (C_{m m'}^{(r, r')} - \lambda_{m m'}^{(r, r')} A_{m m'}^{(r, r')}).$$

Jetzt berechnen wir die zweite, in (23) auftretende Summe, wobei wir uns auf die von E unabhängigen Glieder beschränken können:

$$(34) \quad \sum_m \overset{(2)}{W}_m^{(r)} K_m^{(r)} = \frac{2}{h} \sum_{r'} \frac{\hat{v}^{(r', r)}}{\hat{v}^{(r', r)^2} - v^2} \sum_{m m'} \overset{(2)}{W}_m^{(r)} A_{m m'}^{(r, r')} = D^{(r)}.$$

Dann wird endlich

$$(35) \quad K = E^2 \frac{N}{S_0} \sum_r e^{-\frac{\hat{W}^{(r)}}{kT}} \left(C^{(r)} - \frac{1}{kT} D^{(r)} \right),$$

wo S_0 den Wert von S für $E = 0$ bedeutet.

Wir diskutieren nun die Temperaturabhängigkeit von K genau wie in § 42. Dazu zerlegen wir die Quantenzahl r in n, j und nehmen an, daß nur der Wert $n = 0$ einen merklichen Beitrag liefert. Dann ist $\hat{W}^{(r)}$ durch $\hat{W}^{(n)}$ zu ersetzen und die Eigenschwingungen zerfallen in schnelle $\hat{v}^{(0, n')}$, $n' \neq 0$, und lang-

same $\nu^{(j,j')}$. Die Größe $C^{(r)}$ hat die Form

$$(36) \quad C^{(r)} = C^{(j)} = \sum_{n'j'} c_{n'}^{(j',j)}$$

mit

$$(37) \quad c_{n'}^{(j',j)} = \frac{\nu^{(n',j';0,j)}}{\nu^{(n',j';0,j)^2} - \nu^2} \sum_{mm'} (C_{mm'}^{(0,j;n',j')} - \lambda_{mm'}^{(0,j;n',j')} A_{mm'}^{(0,j;n',j')})$$

und man sieht, daß für $n' = 0$

$$(38) \quad c_0^{(j',j)} = -c_0^{(j,j')}$$

ist. Ganz analoges gilt für $D^{(r)}$. Nun wird

$$\frac{N}{S_0} \sum_r e^{-\frac{\hat{W}^{(r)}}{kT}} C^{(r)} = \frac{N}{S_0} \sum_j e^{-\frac{\hat{W}^{(j)}}{kT}} \sum_{n'j'} c_{n'}^{(j',j)} + \frac{N}{S_0} \sum_{j'} e^{-\frac{\hat{W}^{(j')}}{kT}} c_0^{(j',j)}.$$

Vertauscht man in der zweiten Summe j mit j' und bildet das arithmetische Mittel beider Ausdrücke, so erhält man wegen (38)

$$\begin{aligned} \frac{N}{S_0} \sum_{jj'} e^{-\frac{\hat{W}^{(j)}}{kT}} c_0^{(j',j)} &= \frac{N}{2S_0} \sum_{jj'} \left(e^{-\frac{\hat{W}^{(j)}}{kT}} - e^{-\frac{\hat{W}^{(j')}}{kT}} \right) c_0^{(j',j)} \\ &= \frac{N}{2S_0} \cdot \frac{1}{kT} \cdot \sum_{jj'} h \nu^{(j',j)} c_0^{(j',j)} e^{-\frac{\hat{W}^{(j')}}{kT}}. \end{aligned}$$

Daher wird

$$(39) \quad \frac{N}{S_0} \sum_r e^{-\frac{\hat{W}^{(r)}}{kT}} C^{(r)} = C_0 + \frac{1}{kT} C_1,$$

wo

$$(40) \quad \begin{cases} C_0 = N \sum_{n'} \frac{\nu^{(n',0)}}{\nu^{(n',0)^2} - \nu^2} \frac{2}{S_0} \sum_{jj'} \sum_{mm'} e^{-\frac{\hat{W}^{(j')}}{kT}} (C_{mm'}^{(0,j;n',j')} + \lambda_{mm'}^{(0,j;n',j')} A_{mm'}^{(0,j;n',j')}), \\ C_1 = \frac{Nh}{S_0} \sum_{jj'} \frac{\nu^{(j',j)^2}}{\nu^{(j',j)^2} - \nu^2} e^{-\frac{\hat{W}^{(j')}}{kT}} \sum_{mm'} (C_{mm'}^{(0,j;0,j')} + \lambda_{mm'}^{(0,j;0,j')} A_{mm'}^{(0,j;0,j')}). \end{cases}$$

Diese Größen sind nur wenig von der Temperatur T abhängig und können mit derselben Annäherung als Konstanten behandelt werden, mit der man der Molekel ein permanentes elektrisches Moment und eine konstante Polarisierbarkeit zuschreibt. Ganz

ebenso hat man

$$(41) \quad \frac{N}{S_0} \sum_r e^{-\frac{\tilde{W}^{(r)}}{kT}} D^{(r)} = D_0 + \frac{1}{kT} D_1$$

mit nahezu temperaturunabhängigen D_0, D_1 . Daher wird schließlich

$$(42) \quad K = E^2 \left(C_0 + \frac{1}{kT} (C_1 + D_0) + \frac{1}{k^2 T^2} D_1 \right).$$

Diese Formel gibt die Abhängigkeit der *KERR-Konstante* von Temperatur und Frequenz. Um die anschauliche Bedeutung der 4 Glieder zu übersehen, bedenken wir folgendes: Für C_0 und D_0 sind maßgebend die Sprünge der Quantenzahl n , welche nur von der relativen Lage der Elektronen in der Molekel abhängt; C_1 und D_1 dagegen kommen durch die Sprünge der Quantenzahl j zustande, die zum Gesamtdrehimpuls gehört und mit der Orientierung der Molekel zusammenhängt. Andererseits ist $C = C_0 + \frac{1}{kT} C_1$ von den Energiedifferenzen bei wechselnder Orientierung unabhängig, während $D = D_0 + \frac{1}{kT} D_1$ davon abhängt. Daher entsprechen die einzelnen Glieder folgendem Mechanismus:

1. Das Feld erzeugt an jeder einzelnen Molekel eine Anisotropie, gleichgültig, ob sie an sich isotrop ist oder nicht, bzw. ob sie ein permanentes Moment hat oder nicht; das ist das temperaturunabhängige Glied C_0 .

2. Wenn die Molekel dielektrisch anisotrop ist (sie braucht kein permanentes Moment zu haben), so wird durch das Feld ein Moment induziert, das nicht dem Felde parallel ist, sich aber im Felde einzustellen sucht; dem entspricht die durch $\frac{D_0}{kT}$ dargestellte Anisotropie der Verteilung.

3. Wenn die Molekel ein permanentes Moment hat (sie kann dabei dielektrisch isotrop sein), so strebt dieses sich im Felde einzustellen und es entsteht die durch $\frac{C_1}{kT}$ dargestellte Anisotropie der Verteilung.

4. Ist die Molekel dielektrisch anisotrop und hat zugleich ein Moment, das gegen die dielektrischen Hauptachsen geneigt ist, so entsteht im Felde eine Verlagerung der Richtung dieses Moments; dem Bestreben des resultierenden Moments, sich im Felde einzustellen, entspricht das Glied $\frac{D_2}{k^2 T^2}$

Hieraus geht hervor, daß bei *Atomen* nur der Term C_0 , also eine temperaturunabhängige KERR-Konstante, zu erwarten ist; die andern Effekte können nur bei *Molekeln* vorkommen.

§ 49. Magnetische Doppelbrechung (FARADAY-Effekt). Wir untersuchen jetzt die Wirkung eines konstanten äußeren Magnetfeldes, das parallel zur z -Richtung liegen möge und dessen Betrag H sei, auf die Dispersion des Lichts. Es sei, wie in § 48, $\hat{\mathfrak{P}}$ das elektrische Moment der Molekel im feldfreien Zustande, $\hat{\mathfrak{P}}^{(1)}$ das Moment unter der Wirkung des Magnetfeldes H und $\hat{\mathfrak{P}}$ das vom Lichte erzeugte Zusatzmoment, also wie früher

$$(1) \quad \hat{\mathfrak{P}}_{nn}^{(1)} = \frac{1}{2} e^{2\pi i \nu t} \sum_k \left(\frac{\hat{\mathfrak{P}}_{nk} (\hat{\mathfrak{P}}_{kn} \hat{\mathfrak{G}})}{h(\nu_{kn} + \nu)} - \frac{(\hat{\mathfrak{P}}_{nk} \hat{\mathfrak{G}}) \hat{\mathfrak{P}}_{kn}}{h(\nu_{kn} + \nu)} \right) + \frac{1}{2} e^{-2\pi i \nu t} \sum_k \left(\frac{\hat{\mathfrak{P}}_{nk} (\hat{\mathfrak{P}}_{kn} \hat{\mathfrak{G}}^*)}{h(\nu_{kn} - \nu)} - \frac{(\hat{\mathfrak{P}}_{nk} \hat{\mathfrak{G}}^*) \hat{\mathfrak{P}}_{kn}}{h(\nu_{kn} - \nu)} \right).$$

Der Unterschied gegen den Fall des elektrischen Feldes liegt in den andern Symmetrieeigenschaften, die ein System im magnetischen Felde hat. Während im elektrischen Felde freie Drehbarkeit um die Feldrichtung (z -Achse) und außerdem Spiegelsymmetrie zu einer durch die z -Achse gehenden Ebene (etwa die xz -Ebene) bestehen, gilt im Magnetfeld folgendes:

1. Freie Drehbarkeit um die Feldachse (z -Achse);

2. Spiegelsymmetrie an der zum Felde senkrechten Ebene (der xy -Ebene). Wenn nicht irgend eine zufällige Entartung vorliegt (die wir ausschließen wollen), bewirkt das Magnetfeld eine vollständige Aufhebung der Richtungsentartung. Denn die Energie im Magnetfeld ist z. B. für Atome, auf die wir uns später beschränken werden, nach § 43, (17):

$$(2) \quad W_m^{(r)} = \hat{W}^{(r)} + M_0 m H + \frac{\mathfrak{C}_m^{(r)}}{8 m_0 c^2} H^2 + \dots$$

Benützen wir die Formeln § 30, (10) für Systeme im Magnetfeld, so dürfen wir überall das Zeichen Sp weglassen, da die Spur immer nur aus einem Gliede besteht. Für die Komponenten \hat{X} , \hat{Y} , \hat{Z} von $\hat{\mathfrak{P}}$ folgt also

$$(3) \quad \begin{cases} \hat{X}_{m m'}^{(r, r')} \hat{Y}_{m' m}^{(r', r)} = - \hat{Y}_{m m'}^{(r, r')} \hat{X}_{m' m}^{(r', r)}, \\ \hat{X}_{m m'}^{(r, r')} \hat{Z}_{m' m}^{(r', r)} = \hat{Z}_{m m'}^{(r, r')} \hat{X}_{m' m}^{(r', r)} = 0, \\ \hat{Y}_{m m'}^{(r, r')} \hat{Z}_{m' m}^{(r', r)} = \hat{Z}_{m m'}^{(r, r')} \hat{Y}_{m' m}^{(r', r)} = 0. \end{cases}$$

Daher nimmt (1) die Form an

$$(4) \quad \begin{cases} \hat{X}_{m m}^{(r r)} = \alpha_m^{(r)} \mathfrak{E}_x + \beta_m^{(r)} \mathfrak{F}_y, \\ \hat{Y}_{m m}^{(r r)} = \alpha_m^{(r)} \mathfrak{E}_y - \beta_m^{(r)} \mathfrak{F}_x, \\ \hat{Z}_{m m}^{(r r)} = \gamma_m^{(r)} \mathfrak{E}_z, \end{cases}$$

wo außer dem Lichtvektor

$$5) \quad \mathfrak{E} = \frac{1}{2} (\hat{\mathfrak{E}} e^{2\pi i \nu t} + \hat{\mathfrak{E}}^* e^{-2\pi i \nu t}), \quad \mathfrak{E}_x = A_x \cos(2\pi \nu t + \varphi_x), \dots$$

noch

$$(5') \quad \mathfrak{F} = \frac{1}{2i} (\hat{\mathfrak{E}} e^{2\pi i \nu t} - \hat{\mathfrak{E}}^* e^{-2\pi i \nu t}), \quad \mathfrak{F}_x = A_x \sin(2\pi \nu t + \varphi_x), \dots$$

eingeführt ist und die Konstanten folgende Werte haben:

$$(6) \quad \begin{cases} \alpha_m^{(r)} = \frac{2}{h} \sum_{r', m'} \hat{X}_{m m'}^{(r, r')} \hat{X}_{m' m}^{(r', r)} \frac{\nu_{m' m}^{(r', r)}}{\nu_{m' m}^{(r', r)^2} - \nu^2}, \\ \beta_m^{(r)} = \frac{2}{h i} \sum_{r', m'} \hat{X}_{m m'}^{(r, r')} \hat{Y}_{m' m}^{(r', r)} \frac{\nu}{\nu_{m' m}^{(r', r)^2} - \nu^2}, \\ \gamma_m^{(r)} = \frac{2}{h} \sum_{r', m'} \hat{Z}_{m m'}^{(r, r')} \hat{Z}_{m' m}^{(r', r)} \frac{\nu_{m' m}^{(r', r)}}{\nu_{m' m}^{(r', r)^2} - \nu^2}. \end{cases}$$

Diese Größen sind reell; für $\alpha_m^{(r)}$, $\gamma_m^{(r)}$ folgt das daraus, daß die Matrizen \hat{X} und \hat{Z} hermitisch sind, also

$$(7) \quad \begin{cases} \hat{X}_{m m'}^{(r, r')} \hat{X}_{m' m}^{(r', r)} = |\hat{X}_{m m'}^{(r, r')}|^2 > 0, \\ \hat{Z}_{m m'}^{(r, r')} \hat{Z}_{m' m}^{(r', r)} = |\hat{Z}_{m m'}^{(r, r')}|^2 > 0; \end{cases}$$

für $\beta_m^{(r)}$ folgt wegen:

$$\hat{X} = \frac{1}{2} (\hat{H} + \hat{H}^\dagger), \quad \hat{Y} = \frac{1}{2i} (\hat{H} - \hat{H}^\dagger)$$

aus § 28, (4) (hier anwendbar, da M_z diagonal ist):

$$(7') \quad i \dot{X}_{mm'}^{(r r')} \dot{Y}_{m'm}^{(r' r)} = (m' - m) |\dot{X}_{m'm}^{(r' r)}|^2$$

und daraus dann nach (6) die Realität von $\beta_m^{(r)}$.

Das elektrische Moment pro Volumeneinheit erhält man aus (4) durch Mittelbildung und Multiplikation mit der LOSCHMIDT-schen Zahl N ; also wird

$$(8) \quad \begin{cases} \bar{X} = \alpha \mathfrak{E}_x + \beta \mathfrak{F}_y, \\ \bar{Y} = \alpha \mathfrak{E}_y - \beta \mathfrak{F}_x, \\ \bar{Z} = \gamma \mathfrak{E}_z \end{cases}$$

mit

$$(9) \quad \alpha = N \frac{\sum_{r,m} \alpha_m^{(r)} e^{-\frac{W_m^{(r)}}{k T}}}{\sum_{r,m} e^{-\frac{W_m^{(r)}}{k T}}}, \dots$$

Um die optische Bedeutung der Parameter zu erkennen, verbinden wir die Gleichungen (8) mit der optischen Grundgleichung [s. § 48, (12)]

$$(10) \quad \mathfrak{D} = n^2 (\mathfrak{E} - \bar{s} (\bar{s} \mathfrak{E})).$$

Dabei ist $\mathfrak{D} = \mathfrak{E} + 4\pi \bar{\mathfrak{P}}$, wo $\bar{\mathfrak{P}}$ der Vektor mit den Komponenten (8) ist, also

$$(11) \quad \begin{cases} \mathfrak{D}_x = (1 + 4\pi \alpha) \mathfrak{E}_x + \beta \mathfrak{F}_y, \\ \mathfrak{D}_y = (1 + 4\pi \alpha) \mathfrak{E}_y - \beta \mathfrak{F}_x, \\ \mathfrak{D}_z = 4\pi \gamma \mathfrak{E}_z. \end{cases}$$

Man kann nun natürlich statt \mathfrak{E} ebensogut auch \mathfrak{F} als elektrischen Lichtvektor betrachten, zu dem statt \mathfrak{D} der Verschiebungsvektor

$$(12) \quad \mathfrak{E} = n^2 (\mathfrak{F} - \bar{s} (\bar{s} \mathfrak{F}))$$

gehört; dann treten an die Stelle von (11) die Gleichungen

$$(13) \quad \begin{cases} \mathfrak{E}_x = (1 + 4\pi \alpha) \mathfrak{F}_x - \beta \mathfrak{E}_y, \\ \mathfrak{E}_y = (1 + 4\pi \alpha) \mathfrak{F}_y + \beta \mathfrak{E}_x, \\ \mathfrak{E}_z = 4\pi \gamma \mathfrak{F}_z. \end{cases}$$

Die Formeln (10), (11), (12), (13) enthalten die optischen Gesetze eines Gases im Magnetfeld für beliebige Richtung der Wellennormalen. Wir beschränken die Diskussion auf zwei Spezialfälle:

I. Longitudinale Wellennormale, $\hat{s}_x = \hat{s}_y = 0$, $\hat{s}_z = 1$. Dann erhält man durch Elimination von \mathfrak{C} und \mathfrak{D} :

$$(14) \quad \begin{cases} (1 + 4\pi\alpha)\mathfrak{C}_x + 4\pi\beta\mathfrak{F}_y = n^2\mathfrak{C}_x, \\ (1 + 4\pi\alpha)\mathfrak{C}_y - 4\pi\beta\mathfrak{F}_x = n^2\mathfrak{C}_y, \\ \mathfrak{C}_z = 0, \\ (1 + 4\pi\alpha)\mathfrak{F}_x - 4\pi\beta\mathfrak{C}_y = n^2\mathfrak{F}_x, \\ (1 + 4\pi\alpha)\mathfrak{F}_y + 4\pi\beta\mathfrak{C}_x = n^2\mathfrak{F}_y, \\ \mathfrak{F}_z = 0. \end{cases}$$

Die Lösung lautet

$$(15) \quad n^2 - 1 = 4\pi(\alpha \pm \beta), \quad \mathfrak{F}_y = \pm \mathfrak{C}_x, \quad \mathfrak{C}_y = \mp \mathfrak{F}_x.$$

Hieraus ergibt sich für die Amplituden und Phasen

$$(16) \quad A_x = A_y, \quad \varphi_y - \varphi_x = \pm \frac{\pi}{2}.$$

Es pflanzen sich also zwei zirkulare, entgegengesetzt rotierende Wellen fort, und die *zirkulare Doppelbrechung* ist nach (15) näherungsweise

$$(17) \quad n_+ - n_- = \frac{4\pi\beta}{\bar{n}},$$

wo \bar{n} ein mittlerer Brechungsindex (näherungsweise $\bar{n} = \sqrt{1 + 4\pi\alpha}$) ist.

II. Transversale Wellennormale, $\hat{s}_x = \hat{s}_z = 0$, $\hat{s}_y = 1$. Die Elimination von \mathfrak{D} und \mathfrak{C} liefert

$$(18) \quad \begin{cases} (1 + 4\pi\alpha)\mathfrak{C}_x + 4\pi\beta\mathfrak{F}_y = n^2\mathfrak{C}_x, \\ (1 + 4\pi\alpha)\mathfrak{C}_y - 4\pi\beta\mathfrak{F}_x = 0, \\ (1 + 4\pi\gamma)\mathfrak{C}_z = n^2\mathfrak{C}_z, \\ (1 + 4\pi\alpha)\mathfrak{F}_x - 4\pi\beta\mathfrak{C}_y = n^2\mathfrak{F}_x, \\ (1 + 4\pi\alpha)\mathfrak{F}_y + 4\pi\beta\mathfrak{C}_x = 0, \\ (1 + 4\pi\gamma)\mathfrak{F}_z = n^2\mathfrak{F}_z. \end{cases}$$

Es pflanzen sich zwei Wellen fort:

a) Schwingung senkrecht zum Feld:

$$(19) \quad n_{\perp}^2 - 1 = 4\pi\alpha - \frac{(4\pi\beta)^2}{1 + 4\pi\alpha}, \quad \mathfrak{E}_z = \mathfrak{F}_z = 0,$$

$$\mathfrak{E}_y = \delta \mathfrak{F}_x, \quad \mathfrak{F}_y = -\delta \mathfrak{E}_x, \quad \delta = \frac{4\pi\beta}{1 + 4\pi\alpha};$$

diese Welle hat also eine longitudinale \mathfrak{E} -Komponente von der Größenordnung β .

b) Schwingung parallel zum Feld:

$$(20) \quad n_{\parallel}^2 - 1 = 4\pi\gamma, \quad \mathfrak{E}_x = \mathfrak{E}_y = \mathfrak{F}_x = \mathfrak{F}_y = 0.$$

Die lineare Doppelbrechung beträgt in erster Näherung

$$(21) \quad n_{\perp} - n_{\parallel} = \frac{2\pi(\alpha - \gamma)}{\bar{n}}.$$

Wir untersuchen nun die Abhängigkeit der beiden Doppelbrechungs-Parameter β und $\alpha - \gamma$ vom magnetischen Felde und der Temperatur, und zwar, wie stets im vorangehenden, in Spektralbereichen, die von den Absorptionslinien hinreichenden Abstand haben. Vor allem interessiert uns der Drehungsparameter β .

Bei der Bildung der statistischen Mittelwerte (9) dürfen wir uns wieder (s. § 43) auf den Wert $n = 0$ der Hauptquantenzahl beschränken, da die höheren Zustände bei normalen Temperaturen nicht angeregt sind. Außerdem betrachten wir *von jetzt ab nur Atome*, und können uns dann in vielen Fällen auch auf den Wert j_0 der Quantenzahl j für den tiefsten Term (Grundzustand) beschränken. j_0 ist die „Magnetonzahl“ im Normalzustand. (Wenn $j_0 = 0$ ist, hat man ein diamagnetisches Atom, sonst ein paramagnetisches.) Wir ersetzen also r durch (o, j_0) , r' durch (n', j') und wollen jetzt β nach Potenzen von H entwickeln bis zu Gliedern erster Ordnung einschließlich. Zunächst erhält man nach (2):

$$(22) \quad \frac{1}{v_{mm'}^{(o, j_0; n', j')^2} - v^2} = \frac{1}{v^{(o, j_0; n', j')^2} - v^2} \left(1 - M_0 \lambda_{mm'}^{(o, j_0; n', j')} H + \dots \right),$$

mit

$$(23) \quad \lambda_{mm'}^{(o, j_0; n', j')} = \frac{2 \hat{v}^{(o, j_0; n', j')} (m - m')}{\hbar (\hat{v}^{(o, j_0; n', j')^2} - v^2)}.$$

In § 43 wurde gezeigt, daß die Koordinaten $\hat{x}_k, \hat{y}_k, \hat{z}_k$

der Elektronen im feldfreien Zustande des Systems bis auf Glieder zweiter Ordnung in H auch die Koordinaten im Magnetfelde sind. Folglich kann man bei der hier angestrebten Näherung die Elemente der Matrizen \hat{X} , \hat{Y} , \hat{Z} durch die von \hat{x} , \hat{y} , \hat{z} ersetzen, die zu dem frei drehbaren, ungestörten Atom gehören.

Durch Einsetzen von (22) in (6) unter Beachtung von (7') bekommen wir:

$$\beta_m^{(o, j_0)} = \frac{2\nu}{\hbar} \sum_{n', j', m'} \frac{(m - m') \left| \hat{x}_{m'm}^{(n', j'; o, j_0)} \right|^2}{\nu^{(o, j_0; n', j')^2} - \nu^2} \left(1 - M_0 \lambda_{m'm}^{(o, j_0; n', j')} H \right).$$

Wir zerlegen $\beta_m^{(o, j_0)}$ in zwei Bestandteile:

$$(24) \quad \beta_m^{(o, j_0)} = \beta'_m{}^{(o, j_0)} + \beta''_m{}^{(o, j_0)}$$

mit

$$(25) \quad \beta'_m{}^{(o, j_0)} = \frac{2\nu}{\hbar} \sum_{n', j', m'} \frac{(m - m') \left| \hat{x}_{m'm}^{(n', j'; o, j_0)} \right|^2}{\nu^{(o, j_0; n', j')^2} - \nu^2}$$

und

$$(25a) \quad \beta''_m{}^{(o, j_0)} = -H \frac{2\nu M_0}{\hbar} \sum_{n', j', m'} \frac{(m - m') \left| \hat{x}_{m'm}^{(n', j'; o, j_0)} \right|^2 \lambda_{m'm}^{(o, j_0; n', j')}}{\nu^{(o, j_0; n', j')^2} - \nu^2}$$

Entsprechend setzen wir nach (9) und (2)

$$(26) \quad \beta = \beta' + \beta''$$

mit

$$(27) \quad \beta' = N \frac{\sum_m \beta'_m{}^{(o, j_0)} e^{-\frac{m M_0 H}{k T}}}{\sum_m e^{-\frac{m M_0 H}{k T}}},$$

$$(27a) \quad \beta'' = N \frac{\sum_m \beta''_m{}^{(o, j_0)} e^{-\frac{m M_0 H}{k T}}}{\sum_m e^{-\frac{m M_0 H}{k T}}}.$$

Wir beschäftigen uns zunächst mit β'' . Durch Einsetzen von (23) in (25a) bekommt man, da $(m' - m) \hat{x}_{m'm}^2$ nur für $(m' - m)^2 = 1$

von Null verschieden ist:

$$\beta_m''^{(o, i_0)} = H \cdot \frac{4 \nu M_0}{h^2} \sum_{n', j'} \frac{\hat{\nu}^{(n', j'; o, i_0)}}{(\hat{\nu}^{(n', j'; o, i_0)^2} - \nu^2)^2} \sum_{m'} \left| \hat{x}_{m' m}^{(n', j'; o, i_0)} \right|^2.$$

Da $\beta_m''^{(o, i_0)}$ schon H als Faktor enthält und wir höhere Potenzen von H vernachlässigen, so darf man in (27a) die Faktoren $e^{-\frac{m M_0 H}{k T}}$ durch 1 ersetzen. Der Anteil β'' ist also in dieser Näherung (d. h. bei nicht zu tiefen Temperaturen und nicht sehr hohen Feldstärken) nahezu temperaturunabhängig.

Die Formel (27a) reduziert sich auf

$$\beta'' = H N \frac{4 \nu M_0}{h^2} \sum_{n', j'} \frac{\hat{\nu}^{(n', j'; o, i_0)}}{(\hat{\nu}^{(n', j'; o, i_0)^2} - \nu^2)^2} I^{(n', j'; o, i_0)}$$

mit

$$I^{(n', j'; o, i_0)} = \frac{1}{g(o, j_0)} \left\| \hat{x}^{(n', j'; o, i_0)} \right\|^2 = \frac{1}{3g(o, j_0)} \left\| \hat{p}^{(n', j'; o, i_0)} \right\|^2,$$

d. h. mit der im § 46, (21) eingeführten Bezeichnung

$$I^{(n', j'; o, i_0)} = \frac{e^2 h}{8 \pi^2 m_0} \cdot \frac{f^{(n', j'; o, i_0)}}{\nu^{(n', j'; o, i_0)}}.$$

Es wird also schließlich

$$(28) \quad \beta'' = \beta_0'' \cdot H,$$

wobei, nach dem Werte § 43, (9) des BOHRschen Magnetons,

$$(28a) \quad \beta_0'' = \frac{N e^3}{8 \pi^3 m_0^2 c} \nu \sum_{n', j'} \frac{f^{(n', j'; o, i_0)}}{(\nu^{(n', j'; o, i_0)^2} - \nu^2)^2}.$$

Die Drehung der Polarisationssebene pro Längeneinheit, die vom Anteil β'' der zirkularen Doppelbrechung herrührt, ist also nach (17), wenn λ die Vakuumwellenlänge darstellt,

$$\Theta'' = \frac{\pi}{\lambda} (n_+ - n_-)'' = \frac{4 \pi^2}{\lambda \bar{n}} \beta'' = \Theta_0'' \cdot H,$$

mit der sogenannten VERDETSchen Konstanten

$$\Theta_0'' = \frac{4 \pi^2}{\lambda \bar{n}} \cdot \beta_0''.$$

Durch Vergleich mit der Dispersionsformel § 46, (19), (20) bekommt man nun durch eine elementare Rechnung die zuerst von H. BECQUEREL¹ aufgestellte Beziehung

$$(29) \quad \Theta_0'' = \frac{e}{2 m_0 c^2} \cdot \lambda \frac{dn}{d\lambda}.$$

Wir werden sogleich sehen, daß bei *diamagnetischen* Körpern nur der Anteil Θ'' vorhanden ist. Wenn man also bei einer diamagnetischen Substanz die VERDETSche Konstante und den Verlauf $n(\lambda)$ der Dispersionskurve kennt, gestattet diese Formel eine Bestimmung² von $e : m_0$. Die PLANCKSche Konstante h ist dagegen aus den Formeln (28a), (29) herausgefallen. Aus der Formel (28a) geht ferner hervor, daß die Bestimmung der VERDETSchen Konstanten in ihrer Abhängigkeit von der Wellenlänge zur empirischen Bestimmung der Stärken $f^{(n', j'; o, j_0)}$ von Absorptionslinien dieselben Dienste leisten kann wie die Beobachtung des Brechungsindex selber³.

Nun gehen wir über zur Auswertung des Anteils β' . Aus den Formeln § 31, (15), (16), (17) ergibt sich leicht

$$\sum_{m'} (m - m') \left| \hat{x}_{m' m}^{(n', j_0+1; o, j_0)} \right|^2 = \frac{m j_0}{2 j_0 (j_0 + 1) (2 j_0 + 1)} \cdot \left\| \hat{\mathfrak{p}}^{(n', j_0+1; o, j_0)} \right\|^2,$$

$$\sum_{m'} (m - m') \left| \hat{x}_{m' m}^{(n', j_0; o, j_0)} \right|^2 = \frac{m}{2 j_0 (j_0 + 1) (2 j_0 + 1)} \left\| \hat{\mathfrak{p}}^{(n', j_0; o, j_0)} \right\|^2,$$

$$\sum_{m'} (m - m') \left| \hat{x}_{m' m}^{(n', j_0-1; o, j_0)} \right|^2 = \frac{m (j_0 + 1)}{2 j_0 (j_0 + 1) (2 j_0 + 1)} \left\| \hat{\mathfrak{p}}^{(n', j_0-1; o, j_0)} \right\|^2.$$

Ersetzen wir wieder die $\left\| \hat{\mathfrak{p}}^{(n', j'; o, j_0)} \right\|^2$ durch die entsprechenden Stärken $f^{(n', j'; o, j_0)}$ unter Beachtung von $g(o, j_0) = 2 j_0 + 1$ und setzen

¹ BECQUEREL, H.: Comptes Rendus Bd. 125, S. 679. 1897.

² Die Messungen haben nur bei H_2 annähernd den normalen Wert $e : m_0 = 1,7 \cdot 10^7$ ergeben, sonst kleinere Werte, s. LADENBURG, R.: Z. Phys. Bd. 34, S. 898. 1925; das liegt vermutlich daran, daß Formel (29) nur für Atome (auch bei Berücksichtigung des Elektronendralls), s. ROSENFELD, L.: Z. Phys. Bd. 57, S. 835. 1929), nicht ohne weiteres für Molekeln gilt

³ LADENBURG, R.: Z. Phys. Bd. 4, S. 451. 1921. LADENBURG, R. und R. MINKOWSKI: Z. Phys. Bd. 6, S. 153. 1921.

$$(30) \quad \beta'_0 = -\frac{N e^3 h \nu}{32 \pi^3 m_0^2 c} \sum_{n'} \left[\frac{f(n', j_0; o, j_0)}{(\nu(n', j_0; o, j_0)^2 - \nu^2) \nu(n', j_0; o, j_0)} \right. \\ \left. j_0 \frac{f(n', j_0+1; o, j_0)}{(\nu(n', j_0+1; o, j_0)^2 - \nu^2) \nu(n', j_0+1; o, j_0)} + (j_0+1) \frac{f(n', j_0-1; o, j_0)}{(\nu(n', j_0-1; o, j_0)^2 - \nu^2) \nu(n', j_0-1; o, j_0)} \right],$$

so bekommen wir nach (25) und (27)

$$\beta' = \beta'_0 \cdot \frac{\sum_m (-m) e^{-\frac{m M_0 H}{k T}}}{\sum_m e^{-\frac{m M_0 H}{k T}}} \cdot \frac{3}{M_0 j_0 (j_0 + 1)}.$$

Nun ist aber nach § 43, (25), (32a) der paramagnetische Anteil des magnetischen Moments $M_p(j_0)$ bzw. die CURIESche Konstante C

$$(31) \quad \left\{ \begin{aligned} M_p(j_0) &= M_0 \cdot \frac{\sum_{m=-j_0}^{j_0} (-m) e^{-\frac{m M_0 H}{k T}}}{\sum_{m=-j_0}^{j_0} e^{-\frac{m M_0 H}{k T}}}, \\ C &= \frac{M_0^2 j_0 (j_0 + 1)}{3}, \end{aligned} \right.$$

so daß

$$(32) \quad \beta' = \beta'_0 \cdot \frac{M_p(j_0)}{C}$$

mit dem Wert (30) von β'_0 wird. Hieraus sieht man, daß β' nur für *paramagnetische Körper* von Null verschieden ist: man nennt deshalb den durch β' gegebenen Anteil den „paramagnetischen FARADAY-Effekt“, während der durch β'' gegebene Anteil als der „diamagnetische FARADAY-Effekt“ bezeichnet wird.

Die Temperaturabhängigkeit von β' ist also dieselbe wie die des magnetischen Moments. Bei gewöhnlichen Temperaturen und Feldstärken hat man nach § 43, (31)

$$(33) \quad M_p(j_0) = C \frac{H}{k T},$$

folglich

$$(32a) \quad \beta' = \beta'_0 \cdot \frac{H}{k T}.$$

Auf die Existenz dieses paramagnetischen FARADAY-Effekts hat LADENBURG¹ aufmerksam gemacht; die dem CURIESchen Gesetz (s. § 43) analoge Temperaturabhängigkeit ist an Kristallen von Salzen seltener Erden beobachtet worden. Neuerdings ist es J. BECQUEREL und W. J. DE HAAS² gelungen, den Verlauf von μ_p bis zu tiefen Temperaturen und hohen Feldstärken zu verfolgen.

Fassen wir das Ergebnis für den Bereich üblicher Temperaturen und Feldstärken zusammen, so sehen wir: Der FARADAY-Effekt besteht zunächst aus einem temperaturunabhängigen Anteil, der bei diamagnetischen Substanzen allein vorhanden ist, und überdies bei paramagnetischen Substanzen aus einem in erster Näherung der Temperatur umgekehrt proportionalen Anteil. Beide Bestandteile unterscheiden sich ferner noch durch ihre Abhängigkeit von der Frequenz: Nach (28 a) ist der Verlauf des diamagnetischen Parameters β''_0 auf beiden Seiten einer Eigenfrequenz symmetrisch, dagegen nach (30) der Verlauf des paramagnetischen Parameters β'_0 unsymmetrisch. Dies hat neuerdings zur Auffindung des paramagnetischen FARADAY-Effekts bei Gasen (Caesium-Dampf) durch MINKOWSKI³ geführt.

In analoger Weise kann man den Transversaleffekt behandeln, indem man die Größe $\alpha - \gamma$ nach H entwickelt. Der Effekt ist sehr klein und nur in der Nähe einer Spektrallinie nachweisbar; er ist von VOIGT vorhergesagt und auch experimentell gefunden worden. Doch wollen wir darauf verzichten, die Formeln im einzelnen zu entwickeln.

§ 50. Intensität der Streustrahlung. In § 46 haben wir gesehen, daß das elektrische Moment eines Atoms oder einer Molekel im Felde einer Lichtwelle Schwingungen ausführt, die zum Teil mit denjenigen des Lichtes gleichfarbig und kohärent, zum Teil andersfarbig und inkohärent sind. Die von den ersten Schwingungen sekundär ausgesandten Kugelwellen setzen sich untereinander und mit der einfallenden Welle zu der gebrochenen Welle zu-

¹ LADENBURG, R.: Z. Phys. Bd. 34, S. 898. 1925; Bd. 46, S. 168. 1927. Dort findet man Literaturangaben über empirische Bestätigungen des Effektes.

² Z. Phys. Bd. 52, S. 678. 1928.

³ MINKOWSKI, R.: Naturwissensch. Bd. 17, S. 567. 1929.

sammen, wie sie in der geometrischen Optik betrachtet wird, vorausgesetzt, daß die Atome gleichmäßig verteilt sind. Berücksichtigt man aber die Dichteschwankungen des Mediums, so entsteht außerdem noch eine Streustrahlung des kohärenten Anteils, die von LORD RAYLEIGH entdeckt und behandelt worden ist¹, und die andersfarbige inkohärente Streustrahlung, die SMEKAL² unter Benutzung der Lichtquantenvorstellung vorhergesagt hat. Diese ist dann im Rahmen der BOHRschen Quantentheorie von KRAMERS und HEISENBERG³ durch Korrespondenzbetrachtungen theoretisch näher untersucht worden; die experimentelle Entdeckung gelang fast gleichzeitig RAMAN und KRISHNAN⁴ an Flüssigkeiten und LANDSBERG und MANDELSTAM⁵ an Kristallen. Wir wollen hier die Intensitäts- und Polarisationsverhältnisse der Streustrahlung beider Art betrachten. Dabei machen wir die Annahme, daß die Moleküle des streuenden Mediums alle voneinander unabhängig sind.

Klassisch ist die Intensität des seitlich gestreuten, parallel einer Richtung q polarisierten Lichtes proportional dem durch die Inhomogenität des Mediums bedingten mittleren Schwan-

kungsquadrat $(\Delta \mathfrak{R}_q)^2$ der q -Komponente der induzierten elektrischen Polarisierung \mathfrak{R} . Wenn p das in einer Molekel induzierte Moment und N die Anzahl der Molekel in der Volumeinheit ist, so ist

$$\mathfrak{R} = \sum_{i=1}^{N(1)} p_i^{(1)}$$

die geometrische Summe aller N Einzelmomente in der Volumeinheit. Es sind nun zweierlei Schwankungen zu berücksich-

¹ LORD RAYLEIGH: Phil. Mag. Bd. 41, S. 107 u. 274. 1871; Bd. 47, S. 375. 1899. Bezüglich der nicht vollständigen Polarisierung des zerstreuten Lichtes vgl. GANS, R.: Ann. Physik Bd. 37, S. 896. 1912; Bd. 62, S. 331. 1920; Bd. 65, S. 97. 1921; Z. Phys. Bd. 17, S. 353. 1923. LORD RAYLEIGH: Phil. Mag. (6) Bd. 35, S. 373. 1918. BORN, M.: Verh. d. D. phys. Ges. 1917, S. 243; 1918, S. 16.

² SMEKAL, A.: Naturwissensch. Bd. 11, S. 873. 1923.

³ KRAMERS, H. A. und W. HEISENBERG: Z. Phys. Bd. 31, S. 681. 1925.

⁴ RAMAN, C. V. u. K. S. KRISHNAN: Indian J. Phys. Bd. 2, Part. IV, Juni 1928; Nature Bd. 121, S. 501. 1928; Bd. 122, S. 12. 1928.

⁵ LANDSBERG, G. u. L. MANDELSTAM: Naturwissensch. Bd. 16, S. 558. 1928; Z. Phys. Bd. 50, S. 769. 1928.

tigen: 1. die Schwankung der Richtung des Einzelmoments:

$$\overline{p_i^{(1)}} = \overline{p} + \Delta \overline{p_i^{(1)}}$$

2. die Schwankung der Anzahl N der Moleküle in der Volumeneinheit:

$$N = \overline{N} + \Delta N.$$

Daher ist

$$\begin{aligned} \overline{\mathfrak{R}_q^{(1)}} &= \sum_{i=1}^{\overline{N} + \Delta N} \left(\overline{p_q^{(1)}} + \Delta \overline{p_{qi}^{(1)}} \right) = \overline{p_q^{(1)}} (\overline{N} + \Delta N) + \sum_{i=1}^{\overline{N} + \Delta N} \Delta \overline{p_{qi}^{(1)}}, \\ \Delta \overline{\mathfrak{R}_q^{(1)}} &= \overline{p_q^{(1)}} \cdot \Delta N + \sum_{i=1}^{\overline{N} + \Delta N} \Delta \overline{p_{qi}^{(1)}}, \\ \left(\Delta \overline{\mathfrak{R}_q^{(1)}} \right)^2 &= \overline{p_q^{(1)2}} (\Delta N)^2 + \left(\sum_{i=1}^{\overline{N} + \Delta N} \Delta \overline{p_{qi}^{(1)}} \right)^2 \\ &= \overline{p_q^{(1)2}} (\Delta N)^2 + N \left(\Delta \overline{p_q^{(1)}} \right)^2 \end{aligned}$$

Da nun nach bekannten statistischen Sätzen¹

$$(\Delta N)^2 = N$$

und

$$\overline{p_q^{(1)2}} + \left(\Delta \overline{p_q^{(1)}} \right)^2 = \overline{p_q^2}$$

ist, so finden wir für die gesuchte Intensität

$$(1) \quad \left(\Delta \overline{\mathfrak{R}_q^{(1)}} \right)^2 = \overline{N} \cdot \overline{p_q^2}.$$

Es zeigt sich also, daß die Intensität der Streustrahlung proportional ist der Intensität, die hervorgerufen wird von einer Molekel mit dem elektrischen Moment p . Quantenmechanisch wird diese durch Formel (20), § 20 gegeben.

Die Matrix des induzierten Moments hat nach § 46, (9) die Form:

$$(2) \quad \mathfrak{P}^{(rs)} = \mathfrak{P}_+^{(rs)} \cdot e^{2\pi i(\nu + \nu^{(rs)})t} + \mathfrak{P}_-^{(rs)} \cdot e^{-2\pi i(\nu - \nu^{(rs)})t}$$

¹ Vgl. z. B. LORENTZ, H. A.: Les Théories statistiques en Thermodynamique S. 38. Leipzig 1916.

mit

$$(3) \quad \mathfrak{P}_+^{(1)(rs)} = \frac{1}{2h} \sum_k \left[\frac{\mathfrak{P}^{(rk)} \mathfrak{P}^{(ks)} \cdot \mathfrak{C}}{\nu^{(ks)} + \nu} - \frac{(\mathfrak{P}^{(rk)} \cdot \mathfrak{C}) \mathfrak{P}^{(ks)}}{\nu^{(rk)} + \nu} \right]$$

und wie man leicht kontrolliert,

$$(4) \quad \mathfrak{P}_-^{(1)(rs)} = \mathfrak{P}_+^{(1)(sr)\dagger}.$$

Zu jedem Übergang $r \leftrightarrow s$ ($\nu^{(rs)} > 0$) gibt es also zwei SMEKAL-linien: die *normale* mit der der „STOKESSchen Regel“ genügenden Frequenz $\nu - \nu^{(rs)}$ und die *antistokessche* mit der Frequenz $\nu + \nu^{(rs)}$. Als \mathfrak{p} ist für diese Linien $\mathfrak{P}_-^{(1)(rs)}$ (oder, was auf dasselbe herauskommt $\mathfrak{P}_+^{(1)(sr)}$ bzw. $\mathfrak{P}_+^{(1)(rs)}$ anzusehen. Ihre Relativintensitäten in der Polarisationsrichtung q sind also nach § 20, (20), wenn man für die Verteilung der Molekel in den Zuständen r, s bei der Temperatur T das BOLTZMANNsche Gesetz annimmt,

$$(5) \quad \left\{ \begin{array}{l} J_{q,N}^{(rs)} \sim e^{-\frac{W^{(s)}}{kT}} Sp \left[\left(\mathfrak{P}_+^{(1)(sr)} \cdot q \right) \left(\mathfrak{P}_+^{(1)(sr)\dagger} \cdot q \right) \right] \cdot (\nu - \nu^{(rs)})^4 \\ \text{bzw. } J_{q,A}^{(rs)} \sim e^{-\frac{W^{(r)}}{kT}} Sp \left[\left(\mathfrak{P}_+^{(1)(rs)} \cdot q \right) \left(\mathfrak{P}_+^{(1)(rs)\dagger} \cdot q \right) \right] \cdot (\nu + \nu^{(rs)})^4. \end{array} \right.$$

Da $W^{(r)} > W^{(s)}$ ist, so ist unter gewöhnlichen Umständen $J_{q,A}^{(rs)} \ll J_{q,N}^{(rs)}$.

Die Entstehung der SMEKAL-Linien kann man nämlich folgendermaßen beschreiben:

1. *Normale Linie.* Die leuchtende Molekel ist im Zustande s (etwa im Grundzustand); von der einfallenden Energie $h\nu$ benutzt sie $h\nu^{(rs)}$, um sich in den oberen (angeregten) Zustand r zu heben und die übrigbleibende $h(\nu - \nu^{(rs)})$ wird wieder ausgestrahlt.

2. *Antistokessche Linie.* Die Molekel ist ursprünglich angeregt (im Zustande r). Unter dem Einfluß der einfallenden Strahlung fällt sie in den unteren (Grund-)Zustand s zurück. Die dabei frei werdende Energie $h\nu^{(rs)}$ wird mit der einfallenden Energie $h\nu$ als Licht von der Frequenz $\nu + \nu^{(rs)}$ gestreut.

Wir wollen nun die Faktoren

$$\begin{aligned} \dot{i}_{q,N}^{(rs)} &= Sp \left[\left(\mathfrak{P}_+^{(1)(sr)} \cdot q \right) \left(\mathfrak{P}_+^{(1)(sr)\dagger} \cdot q \right) \right] \\ \dot{i}_{q,A}^{(rs)} &= \dot{i}_{q,N}^{(sr)} \end{aligned}$$

näher betrachten.

Nun ist

$$(6) \quad \mathfrak{P}_+^{(sr)} \cdot \mathfrak{q} = \frac{1}{2\hbar} \sum_{xy} \mathfrak{q}_x \mathfrak{E}_y \Theta_{xy}^{(sr)}$$

mit

$$(7) \quad \Theta_{xy}^{(sr)} = \sum_k \left[\frac{X^{(sk)} Y^{(kr)}}{\nu^{(kr)} + \nu} - \frac{Y^{(sk)} X^{(kr)}}{\nu^{(sk)} + \nu} \right].$$

Also ist

$$(8) \quad \mathfrak{q}_{\mathfrak{q}, \mathfrak{N}}^{(rs)} = \frac{1}{4\hbar^2} \sum_{xy} \sum_{x'y'} \mathfrak{q}_x \mathfrak{q}_{x'} \mathfrak{E}_y \mathfrak{E}_{y'}^* Sp(\Theta_{xy}^{(sr)} \Theta_{x'y'}^{(sr)\dagger}).$$

Im Falle freier Drehbarkeit (Gasmolekel) läßt sich diese Gleichung vereinfachen. Dann sind von den rechts auftretenden Spuren wegen des Tensorcharakters von Θ nur die folgenden von Null verschieden [§ 30, Gl. (25) bis (29)]:

$$(9) \quad \begin{cases} A^{(rs)} = Sp(\Theta_{xx}^{(sr)} \Theta_{xx}^{(sr)\dagger}) = Sp(\Theta_{yy}^{(sr)} \Theta_{yy}^{(sr)\dagger}) \\ \quad = Sp(\Theta_{zz}^{(sr)} \Theta_{zz}^{(sr)\dagger}), \\ B^{(rs)} = Sp(\Theta_{xx}^{(sr)} \Theta_{yy}^{(sr)\dagger}) = \dots \\ C^{(rs)} = Sp(\Theta_{xy}^{(sr)} \Theta_{xy}^{(sr)\dagger}) = \dots \\ D^{(rs)} = Sp(\Theta_{xy}^{(sr)} \Theta_{yx}^{(sr)\dagger}) = \dots \end{cases}$$

wobei jedesmal Spuren, welche durch irgend eine Permutation von (x, y, z) auseinander hervorgehen, einander gleich sind. Ferner besteht für zwei beliebige Raumtensoren A_{xy} und B_{xy} im Falle freier Drehbarkeit die Beziehung:

$$(10) \quad Sp(A_{xx}^{(r,s)} B_{xx}^{(s,r)}) = Sp(A_{xx}^{(r,s)} B_{yy}^{(s,r)}) + Sp(A_{xy}^{(r,s)} B_{xy}^{(s,r)}) \\ + Sp(A_{xy}^{(r,s)} B_{yx}^{(s,r)}).$$

Es genügt dies zu beweisen für den Fall, daß A_{xy} und B_{xy} Produkte zweier Vektoren sind. Für diesen Fall reduziert sich die Beziehung (10) auf:

$$(11) \quad Sp(X_1 X_2 X_3 X_4) = Sp(X_1 X_2 Y_3 Y_4) + Sp(X_1 Y_2 X_3 Y_4) \\ + Sp(X_1 Y_2 Y_3 X_4),$$

wobei die Indizes r, s weggelassen sind. Letztere Gleichung folgt, unter Beachtung von § 30, (25) bis (29), aus der Invarianz der lin-

ken Seite z. B. bezüglich der Substitution $X \rightarrow \frac{1}{\sqrt{2}}(X + Y)$ (Drehung um 45° um die Z -Achse).

In unserem Fall lautet diese Beziehung (10):

$$(12) \quad A^{(r,s)} = B^{(r,s)} + C^{(r,s)} + D^{(r,s)}.$$

Mit Hilfe von (12) reduziert sich (8) auf:

$$(13) \quad i_{q,N}^{(r,s)} = \frac{1}{4\hbar^2} \{C^{(r,s)}|\mathfrak{E}|^2 + (A^{(r,s)} - C^{(r,s)})|\mathfrak{E}q|^2\}.$$

Die Streustrahlung besteht somit aus einem unpolarisierten Anteil $\frac{1}{4\hbar^2}C|\mathfrak{E}|^2$ und einem polarisierten Anteil $\frac{1}{4\hbar^2}(A - C)|\mathfrak{E}q|^2$. Und zwar hängt, für linear polarisiertes einfallendes Licht, die in einer bestimmten Richtung durch einen Analysator von der Stellung q (senkrecht zur Beobachtungsrichtung) beobachtete Intensität der Streustrahlung in derselben Weise von dem Winkel zwischen \mathfrak{E} und q ab, wie in der klassischen Theorie¹; z. B. ist für einen Beobachter in der Richtung des Vektors \mathfrak{E} für jede Stellung q der polarisierte Anteil Null.

Sobald man für einen speziellen Fall Auswahlregeln für \mathfrak{P} kennt, kann man aus (7) und (8) Auswahlregeln für $i^{(r,s)}$ entnehmen. Für eine zweiatomige Gasmolekel ergibt sich z. B. folgendes: Spaltet man von der Quantenzahl r die Bestandteile j (Rotations-) und n (Oszillationsquantenzahl) ab, so sind die Komponenten von $\mathfrak{P}^{(n,j;n',j')}$ nur dann von null verschieden, wenn $\Delta j = j' - j = \pm 1$ [vgl. § 29, (12), (13) und § 32, (3)] und $\Delta n = n' - n = 0, \pm 1$ ist [vgl. § 34 (23)]; berücksichtigt man die Koppelung von Rotation und Oszillation (§ 40), so kommen auch andere Werte von Δn vor]. Aus (7) entnimmt man folglich, daß $\Theta_{xy}^{(n,j;n',j')}$ und also auch $i^{(n,j;n',j')}$ nur dann von null verschieden sind, wenn $\Delta j = 0, \pm 2$ und $\Delta n = 0, \pm 1, \pm 2$ ist². Das ist in Übereinstimmung mit den Beobachtungen³. — Für einen rein-harmo-

¹ Vgl. BORN, M.: Verh. dtsch. phys. Ges. Bd. 19, S. 243. 1917; Bd. 20, S. 16. 1918.

² Vgl. HILL, E. L. u. E. C. KEMBLE: Proc. nat. Acad. Sci. U. S. A. Bd. 15, S. 387. 1929 und MANNEBACK, C.: Naturwissensch. Bd. 17, S. 365. 1929.

³ Näheres s. WOOD, R. W.: Nature Bd. 123, S. 279. 1929; Phil. Mag. Bd. 7, S. 744. 1929; McLENNAN, J. u. J. McLEOD: Nature Bd. 123, S. 160. 1929.

nischen (räumlichen) Oszillator (isotrop oder anisotrop) findet man aus (7) und (8) (am einfachsten, indem man die Koordinatenachsen parallel zu den drei Hauptschwingungsrichtungen legt), daß nur die RAYLEIGHsche Streuung auftritt, aber nicht die RAMANSche. Daß trotzdem bei Kristallen RAMAN-Strahlung auftritt, welche einer Schwingung der Bestandteile des Kristallgitters gegeneinander entspricht, ist also eine Folge der Anharmonizität dieser Schwingung¹.

§ 51. Intensität und Polarisation des Fluoreszenzlichts. Wir haben in § 50 die Streuung des Lichts an Molekeln berechnet unter der Voraussetzung, daß die Lichtfrequenz mit keiner der Eigenfrequenzen der Molekel übereinstimmt (RAYLEIGH-, RAMAN-Effekt); wenn aber dieser „Resonanzfall“ eintritt, versagten die Formeln [§ 50, (3)] wegen des Auftretens verschwindender Nenner. Man wird erwarten, daß gerade bei Resonanz eine besonders intensive Streuung auftritt; um diese zu berechnen, müssen wir aber auf die Formeln des § 45 zurückgreifen.

Denkt man sich die Molekel vom Zeitpunkt $\tau = 0$ an mit monochromatischem Licht der Frequenz $\nu = \nu_{ab}$ bestrahlt, so ergibt das Verfahren von § 45, wie man leicht nachrechnet, daß $V_{ab}^{(1)}$, neben periodischen Gliedern, ein additives Glied enthält, das proportional mit t wächst; und daraus folgt durch Einsetzen in Formel (18), § 45, daß $\mathfrak{P}_{am}^{(1)}$ ein Glied enthält, das eine Schwingung der Frequenz ν_{bm} darstellt mit einer Amplitude, welche proportional t ist. Korrespondenzmäßig entspricht das einer Ausstrahlung des Fluoreszenzlichtes der Frequenz ν_{bm} , deren Intensität für kleine Zeiten (welche aber groß sind im Vergleich zu $\frac{1}{\nu_{bm}}$) wie t^2 wächst. (Für größere Werte von t müßte man auch die höheren Näherungen berücksichtigen).

Ein wesentlich anderes Resultat erhält man, wenn das eingestrahelte Licht nicht rein-monochromatisch ist, sondern *natürliches* Licht ist, d. h. eine endliche Spektralbreite um die Resonanzfrequenz ν_{ab} besitzt, wobei benachbarte Frequenzen mit allen möglichen Phasen vertreten sind. In diesem Falle ist die Intensität der Fluoreszenzstrahlung für kleine Zeiten proportional mit t , wie wir sehen werden.

¹ Vgl. auch SCHÄFER, Cl.: Z. Phys. Bd. 54, S. 153. 1929.

Wir wollen allgemein den Fall betrachten, daß *natürliches* Licht einer *beliebigen* Spektralverteilung eingestrahlt wird. Im Intervall $0 \leq \tau \leq t$ sei der elektrische Lichtvektor dargestellt durch das FOURIER-Integral:

$$(1) \quad \mathfrak{E}(\tau) = \int_{-\infty}^{+\infty} e(\nu) e^{2\pi i \nu \tau} d\nu$$

mit $e(-\nu) = e^*(\nu)$. Die FOURIER-Koeffizienten, welche von der Länge t des betrachteten Zeitabschnittes abhängen, sind durch

$$(1') \quad e(\nu) = \int_0^t \mathfrak{E}(\tau) e^{-2\pi i \nu \tau} d\tau$$

gegeben und sind mit der mittleren Energiedichte $\varrho(\nu)$ (für das Zeitintervall von 0 bis t) durch die Formel

$$|e(\nu)|^2 = 2\pi t \varrho(\nu)$$

verknüpft (vgl. Anhang I). Im Falle *natürlichen* Lichtes ist $\varrho(\nu)$ für Frequenzen ν , welche im Lichte *stark* vertreten sind, nahezu unabhängig von t , d. h. $|e(\nu)|^2$ ist proportional mit t .

Nach § 45, (18) wird allgemein das gestörte Moment gegeben durch

$$(2) \quad \mathfrak{P}^{(1)}(nm) = \sum_k \mathfrak{P}^{(1)}(nk) \mathfrak{V}^{(1)}(km) e^{2\pi i \nu_{nk} t} - \mathfrak{V}^{(1)}(nk) \mathfrak{P}^{(1)}(km) e^{2\pi i \nu_{km} t}.$$

Hierbei ist nach § 45, (9) (mit $t_0 = 0$)

$$\mathfrak{V}^{(1)}(nm) = -\frac{1}{\varkappa} \int_0^t K^{(nm)} d\tau.$$

In unserem Falle ist

$$K^{(nm)} = -\mathfrak{P}^{(1)}(nm) e^{2\pi i \nu_{nm} \tau} \cdot \mathfrak{E}(\tau),$$

folglich wird

$$\mathfrak{V}^{(1)}(nm) = \frac{1}{\varkappa} \left(\mathfrak{P}^{(1)}(nm) \cdot \int_0^t \mathfrak{E}(\tau) e^{2\pi i \nu_{nm} \tau} d\tau \right) = \frac{1}{\varkappa} \left(\mathfrak{P}^{(1)}(nm) \cdot e^*(\nu_{nm}) \right),$$

und das gestörte Moment in der Polarisationsrichtung \mathfrak{a}

$$(3) \quad \begin{aligned} \mathfrak{P}^{(1)}(nm) \cdot \mathfrak{a} &= \frac{1}{\varkappa} \sum_k \{ (\mathfrak{P}^{(1)}(nk) \cdot \mathfrak{a}) (\mathfrak{P}^{(1)}(km) \cdot e^*(\nu_{km})) e^{2\pi i \nu_{nk} t} \\ &\quad - (\mathfrak{P}^{(1)}(nk) \cdot e^*(\nu_{nk})) (\mathfrak{P}^{(1)}(km) \cdot \mathfrak{a}) e^{2\pi i \nu_{km} t} \}. \end{aligned}$$

Das Glied

$$(4) \quad \frac{1}{\varkappa} (\mathfrak{P}^{(nk)} \cdot e^*(\nu_{nk})) (\mathfrak{P}^{(km)} \cdot \mathfrak{q}) e^{2\pi i \nu_{km} t}$$

hat im wesentlichen die Frequenz ν_{km} und eine Amplitude, welche wegen des Faktors $e^*(\nu_{nk})$ wie \sqrt{t} wächst (vorausgesetzt, daß die betreffende Resonanzfrequenz ν_{kn} im eingestrahlten natürlichen Lichte stark vertreten ist¹). Korrespondenzmäßig entspricht dieses Glied also der Ausstrahlung einer Fluoreszenzlinie der Frequenz ν_{km} , hervorgerufen durch die Anwesenheit der Resonanzfrequenz ν_{kn} im eingestrahlten Licht, also in klassischer Deutung einem Übergange $n \rightarrow k \rightarrow m$: Durch Absorption eines Lichtquants $h \nu_{kn}$ wird eine Molekel aus dem Zustand n in den Zustand k gehoben und fällt dann unter spontaner Emission des Lichtquants $h \nu_{km}$ in den Zustand m . In Analogie zu dem HEISENBERGSchen Ansatz über die Intensitäten bei spontaner Emission [§ 20, (20)] machen wir hier den Ansatz, daß die Intensität (in der Polarisationsrichtung \mathfrak{q}) der Fluoreszenzstrahlung, welche dem Übergang $n \rightarrow k \rightarrow m$ entspricht, gegeben ist durch den Koeffizienten von $e^{2\pi i \nu_{km} t}$ im Ausdruck (4), gemäß:

$$(5) \quad J_{\mathfrak{q}}(n \rightarrow k \rightarrow m) = \frac{64 \pi^4}{3 c^3} \nu_{km}^4 \cdot \frac{1}{|\varkappa|^2} \| (\mathfrak{P}^{(nk)} \cdot e^*(\nu_{nk})) (\mathfrak{P}^{(km)} \cdot \mathfrak{q}) \|^2,$$

sie wächst also proportional mit t . Das gilt natürlich nur für kleine Werte von t , für große t strebt $J_{\mathfrak{q}}(n \rightarrow k \rightarrow m)$ gegen einen Grenzwert. Hierauf können wir aber nicht näher eingehen, solange wir uns auf die erste Näherung beschränken. Wir wollen uns damit begnügen, die Intensität, so wie sie durch Formel (5) gegeben ist (also für kleine t), näher zu diskutieren.

Ist das System *nicht entartet*, so reduziert sich (5) auf

$$(6) \quad \begin{aligned} J_{\mathfrak{q}}(n \rightarrow k \rightarrow m) &= \frac{64 \pi^4}{3 c^3} \nu_{km}^4 \cdot \frac{1}{|\varkappa|^2} | \mathfrak{P}^{(nk)} \cdot e^*(\nu_{nk}) |^2 | \mathfrak{P}^{(km)} \cdot \mathfrak{q} |^2 \\ &= \Psi^{(nk)} \cdot J_{\mathfrak{q}}^{(km)}. \end{aligned}$$

Hierbei ist $J_{\mathfrak{q}}^{(km)} = \frac{64 \pi^4}{3 c^3} \nu_{km}^4 | \mathfrak{P}^{(km)} \cdot \mathfrak{q} |^2$ die Intensität der spontanen Emission der Frequenz ν_{km} , was man deuten kann als die

¹ Wenn man dagegen z. B. monochromatisches Licht hat einer Frequenz ν , welche von allen Resonanzfrequenzen verschieden ist, so folgt aus (1') daß $e(\nu_{nk})$ selber eine *periodische* Funktion von t ist.

Wahrscheinlichkeit pro Zeiteinheit für den spontanen Übergang $k \rightarrow m$. Und $\Psi^{(n k)} = \frac{1}{|\mathfrak{x}|^2} |\mathfrak{P}^{(n k)} \cdot e^*(\nu_{nk})|^2$ ist, wie wir in § 63 sehen werden, die Wahrscheinlichkeit dafür, daß ein Atom, das sich zur Zeit $\tau = 0$ im Zustand n befand, unter Einfluß des eingestrahlten Lichts, innerhalb des Zeitintervalls $0 - t$ vom Zustand n in den Zustand k gehoben wird, und sich im Zeitpunkt t noch dort befindet; und zwar gilt auch diese Formel (Proportionalität mit $t!$) nur für kleine Zeiten t .

Die im Zeitintervall $t, t + dt$ ausgestrahlte Energie ist nach (6) gegeben durch:

$$J_q(n \rightarrow k \rightarrow m) dt = \Psi^{(n k)} J_q^{(k m)} dt.$$

Dasselbe Resultat erhält man, wenn man den Mechanismus des Vorganges als: Absorption und Reemission beschreibt und darauf eine klassische Wahrscheinlichkeitsüberlegung anwendet: die im Zeitintervall $t, t + dt$ ausgestrahlte Energie ist bestimmt durch die Wahrscheinlichkeit dafür, daß ein Atom, das sich zur Zeit $\tau = 0$ im Zustand n befand, sich zur Zeit t im Zustand k befindet, und im Intervall $t, t + dt$ spontan in den Zustand m übergeht; diese Wahrscheinlichkeit ist klassisch das Produkt der Wahrscheinlichkeiten für das Zutreffen der beiden Bedingungen, d. h.

$$\Psi^{(n k)} \cdot J_q^{(k m)} dt.$$

Wenn das System dagegen *entartet* ist, so kann man im allgemeinen nicht mehr nach diesem klassischen Schema rechnen: wir werden sehen, daß die Intensität des Fluoreszenzlichts dann nicht mehr eine Funktion der Wahrscheinlichkeiten für Absorption und Emission ist, sondern es tritt sogenannte „Interferenz der Wahrscheinlichkeiten“ auf (dieser Begriff wird in § 60 näher erläutert). Das hat zur Folge, daß kleine Störungen, welche die Entartung aufheben, große Änderungen der Polarisations- und Intensitätsverhältnisse erzeugen können. Das wollen wir am Beispiel der Richtungsentartung (Gasmolekel) erläutern. Wir denken uns diese zunächst durch ein schwaches Magnetfeld parallel der z -Richtung aufgehoben. Wir betrachten den Fall, daß das eingestrahlte Licht fast monochromatisch ist und nur die Resonanzfrequenzen $\nu_{m m}^{(a b)}$ enthält, in welche eine Eigenfrequenz $\nu_{a b}$ des entarteten Systems durch das Magnetfeld aufgespalten

wird; und zwar sei die Aufspaltung so klein, daß annähernd $e(\nu_{mm'}^{(ab)}) = e(\nu_{ab})$ ist. Dann geht (3) über in (für $n \neq a, b$)

$$(7) \quad (\hat{\mathfrak{P}}_{mm'}^{(1)(an)} \cdot q) = -\frac{1}{\varkappa} \sum_{m''} (\hat{\mathfrak{P}}_{mm''}^{(ab)} \cdot e^*(\nu_{ab})) (\hat{\mathfrak{P}}_{m''m'}^{(bn)} \cdot q) e^{2\pi i \nu_{m''m'}^{(bn)} t},$$

(für $n = a$ kommen noch mehr ähnliche Glieder hinzu.) Da das System nicht entartet ist, so gilt Formel (6):

$$(8) \quad i_q((a, m) \rightarrow (b, m'') \rightarrow (n, m')) \\ = |\hat{\mathfrak{P}}_{mm''}^{(ab)} \cdot e^*(\nu_{ab})|^2 |\hat{\mathfrak{P}}_{m''m'}^{(bn)} \cdot q|^2,$$

wobei von jetzt an der Faktor $\frac{64 \pi^4}{3 c^3} \nu^4 \cdot \frac{1}{|\varkappa|^2}$ weggelassen werden soll.

Die totale Intensität für den Übergang $(a, m) \rightarrow (n, m')$ ist also

$$(9) \quad i_q\left((a, m) \rightarrow \left(b, \begin{matrix} m_1 \\ m_2 \\ \vdots \end{matrix}\right) \rightarrow (n, m')\right) \\ = \sum_{m''} |\hat{\mathfrak{P}}_{mm''}^{(ab)} \cdot e^*(\nu_{ab})|^2 |\hat{\mathfrak{P}}_{m''m'}^{(bn)} \cdot q|^2.$$

Dies, ebenso wie Formel (6) und (8), gilt aber nur, solange die Aufspaltung nicht zu klein ist; im Grenzfall, wenn $\nu_{m''m'}^{(bn)} = \nu_{bn}$ wird, wird in (7) der Koeffizient von $e^{2\pi i \nu_{bn} t}$

$$\frac{1}{\varkappa} \sum_{m''} (\hat{\mathfrak{P}}_{mm''}^{(ab)} \cdot e^*(\nu_{ab})) (\hat{\mathfrak{P}}_{m''m'}^{(bn)} \cdot q)$$

und man wird zu dem Ansatz

$$(10) \quad i_q\left[(a, m) \rightarrow \left(b, \begin{matrix} m_1 \\ m_2 \\ \vdots \end{matrix}\right) \rightarrow (n, m')\right] \\ = \left| \sum_{m''} (\hat{\mathfrak{P}}_{mm''}^{(ab)} \cdot e^*(\nu_{ab})) (\hat{\mathfrak{P}}_{m''m'}^{(bn)} \cdot q) \right|^2$$

geführt, wobei also eine „Interferenz der Wahrscheinlichkeiten“ auftritt. Um die totale Intensität beim Übergang $(a \rightarrow b \rightarrow n)$ zu erhalten, muß man noch über m und m' summieren, was wieder auf die Größe (5) führt, welche in der Tat eine Relativinvariante ist.

Der Unterschied zwischen (9) und (10) erklärt die plötzliche Änderung der Polarisationsverhältnisse der Fluoreszenzlinien beim Einschalten eines Magnetfeldes, was zuerst experimentell

von WOOD und ELLET¹ entdeckt und später von HANLE² und anderen untersucht und korrespondenzmäßig gedeutet wurde. — Daß diese Änderung in Wirklichkeit trotzdem stetig ist, liegt daran, daß man schon dann von (9) zu (10) übergehen muß, wenn die Frequenzen $\nu_{mm'}^{(ab)}$ so nahe beieinander liegen, daß sie miteinander interferieren, d. h. so nahe, daß ihr Unterschied klein ist gegenüber der reziproken Dauer des Ausstrahlungsprozesses (Linienbreite). (Darauf können wir hier nicht näher eingehen.) Es ist hierbei zu beachten, daß, wenn man zu einer anderen Lösung für das (*nicht-entartete!*) System übergeht, d. h. $\hat{\mathfrak{P}}$ mit einer beliebigen Phasenmatrix transformiert, dabei alle Summanden eines bestimmten Matrixelementes $\hat{\mathfrak{P}}_{mm'}^{(1)(an)}$ [Gl. (7)] mit *einem und demselben* Phasenfaktor multipliziert werden; sie sind also in gewissem Sinne untereinander kohärent, obwohl sie verschiedene Frequenzen haben; der Ausdruck (10) ist, ebenso wie (9), für *nicht-entartete* Systeme relativinvariant.

Es gibt einige Fälle, wo (9) und (10) miteinander übereinstimmen, nämlich dann, wenn die Summation über m'' auf Grund von Auswahlregeln fortfällt. Führen wir die Bezeichnung

$$e_x + i e_y = \eta, \quad e_x - i e_y = \eta'$$

ein, so geht (10) über in:

$$\begin{aligned} (10') \quad & \dot{i}_q [(a, m) \rightarrow (b, \dots) \rightarrow (n, m')] \\ & = \left| e_z^* Z_{mm}^{(ab)} (\hat{\mathfrak{P}}_{mm'}^{(bn)} \cdot q) \right. \\ & \quad \left. + \frac{i}{2} \eta' \dot{I}I_{m, m-1}^{(ab)} (\hat{\mathfrak{P}}_{m-1, m'}^{(bn)} \cdot q) + \frac{i}{2} \eta \dot{I}I_{m, m+1}^{\dagger(ab)} (\hat{\mathfrak{P}}_{m+1, m'}^{(bn)} \cdot q) \right|^2. \end{aligned}$$

Hieraus sieht man, daß in den drei Fällen: einfallendes Licht linear polarisiert parallel der z -Richtung ($\eta = \eta' = 0$), rechts- oder linkszirkular polarisiert in der xy -Ebene ($e_z = \eta = 0$ bzw. $e_z = \eta' = 0$), die Ausdrücke (10') und (9) identisch sind, und folglich das Einschalten eines schwachen Magnetfeldes parallel der z -Richtung nur kleine Änderungen der Polarisationsverhält-

¹ WOOD, R. W. und A. ELLET: Proc. roy. Soc. Bd. 103 A, S. 396. 1923.

² HANLE, W.: Naturwissensch. Bd. 11, S. 690, 1923. Weitere Literaturangaben bei W. HANLE: Ergebnisse d. ex. Naturwissensch. Bd. 6. PRINGSHEIM, P.: Fluoreszenz und Phosphoreszenz, diese Sammlung Bd. 6.

nisse erzeugt; da die Polarisations-eigenschaften der aufgespaltenen Fluoreszenzlinien im Magnetfelde, entsprechend der Bemerkung zu (6), dieselben sind wie die der ZEEMAN-Komponenten, welche zu denselben Übergängen gehören, so ergibt sich die zuerst von HEISENBERG auf Grund von korrespondenzmäßigen Überlegungen formulierte *Regel*¹: Bei Einstrahlung mit linear bzw. zirkular polarisiertem Licht erhält man im Falle von Richtungsentartung die Polarisation und Intensität der Fluoreszenzstrahlung, indem man die Intensitäten der ZEEMAN-Komponenten in einem Magnetfeld, das man sich parallel dem linearpolarisierten bzw. senkrecht zur Ebene des zirkular polarisierten Lichtes eingeschaltet denkt, mit den betreffenden Anregungswahrscheinlichkeiten multipliziert und zueinander addiert. [Siehe Formel (10'), vgl. auch § 63, (10), (11).]

Sechstes Kapitel.

Statistische Deutung der Quantenmechanik².

§ 52. Meßbare Größen und Eigenwerte. Überblicken wir das Voranstehende, so können wir feststellen, daß eine große Zahl von Erscheinungen mit Hilfe der aus HEISENBERGS Gedanken entwickelten Theorie abgeleitet werden konnten. Erinnern wir uns an ihre Grundbegriffe: Sie übernimmt sämtliche Vorstellungen der BOHRschen Quantentheorie, wie stationäre Zustände, Quantensprünge usw., und ergänzt sie durch eine mathematische Darstellung, von der wir hier folgende Züge hervorheben: Jede physi-

¹ BOHR, N.: Naturwissensch. Bd. 12, S. 1115. 1924. HEISENBERG, W.: Z. Phys. Bd. 31, S. 617. 1925. Quantenmechanisch behandelt von OPPENHEIMER, J. R.: Z. Phys. Bd. 43, S. 27. 1927. HOYT, F. C.: Phys. Rev. Bd. 32, S. 377. 1928.

² DIRAC, P. A. M.: Proc. Roy. Soc. London (A) Bd. 113, S. 621. 1927. JORDAN, P.: Z. Phys. Bd. 40, S. 809. 1927; Bd. 41, S. 797. 1927; Bd. 44, S. 1. 1927. Ferner NEUMANN, J. v.: Göttinger Nachrichten S. 1, 1927; S. 245, 1927. HEISENBERG, W.: Z. Phys. Bd. 43, S. 172. 1927. BOHR, N.: Naturwissensch. Bd. 16, S. 245. 1928. Die ersten Ansätze zur statistischen Deutung der Quantenmechanik stammen von: BORN, M.: Z. Phys. Bd. 37, S. 863. 1926; Bd. 38, S. 803. 1926; Bd. 40, S. 167. 1926. HEISENBERG, W.: Z. Phys. Bd. 40, S. 501. 1926. JORDAN, P.: Z. Phys. Bd. 40, S. 661. 1926. PAULI jr., W.: Z. Phys. Bd. 41, S. 81. 1927. Wesentliche Anregungen hat ferner die vorquantenmechanische Arbeit von BOHR, N., H. A. KRAMERS und J. C. SLATER: Z. Phys. Bd. 24, S. 69. 1924 gegeben.

kalische Größe wird durch eine Matrix beschrieben. Die Werte, welche die Energie eines (abgeschlossenen) Systems annehmen kann, sind die Eigenwerte ihrer Matrix. Die Matrixelemente außerhalb der Diagonalen sind den Quantensprüngen zugeordnet; die Diagonalelemente der Matrix irgend einer Größe stellen die Zeitmittel der Größe in den stationären Zuständen dar. Als beobachtbar werden in erster Linie angesehen: der jeweils verwirklichte Wert der Energie (ein Eigenwert der Energiematrix) und die Quadrate der Beträge der außerhalb der Diagonalen stehenden Elemente solcher Matrizen, die für die Lichtemission, Absorption und Dispersion maßgebend sind (elektrisches Dipolmoment; in höherer Näherung elektrisches Quadrupolmoment usw.); sodann aber auch die Zeitmittelwerte (Diagonalelemente der Matrix) irgend einer Größe. So kann man z. B. die Mittelwerte des elektrischen Dipolmoments eines Atoms aus den Änderungen der Energiewerte im schwachen elektrischen Felde (Starkeffekt) entnehmen (s. § 42); ferner geben Dielektrizitätskonstante, Magnetisierbarkeit und ähnliche Größen über solche Mittelwerte Auskunft.

Dieses Begriffssystem scheint aber als Grundlage einer umfassenden neuen Mechanik trotz des Umfangs seiner Anwendungsmöglichkeiten noch unzulänglich und unbefriedigend; es enthält eine sehr spezielle, einseitige Auswahl der als beobachtbar geltenden Größen. Zweifellos ist diese Auswahl durch die natürliche Entwicklung der Theorie und der experimentellen Erfahrung an die Hand gegeben worden; daß sie aber nicht endgültig sein kann, zeigt sich an zahlreichen experimentellen Fragestellungen. Besonders auffällig ist das Beispiel des freien, nicht atomar gebundenen Elektrons. Dabei wird selbstverständlich die Frage gestellt und experimentell beantwortet, an welchem *Orte* das Elektron sich gerade befindet; aber Fragen solcher Art stehen in keinem unmittelbar erkennbaren Zusammenhang mit den Begriffsbildungen, die der ursprünglichen HEISENBERG'schen Theorie zugrunde liegen. Unsere Theorie der *diskreten* Matrizen ist zwar schon deswegen nicht geeignet zur Behandlung aperiodischer Bewegungen, weil sie sich eben ausdrücklich auf die *periodischen* Bewegungen (im quantenmechanischen Sinn) beschränkt. Doch haben wir bereits angedeutet, daß durch Zulassung kontinuierlich veränderlicher Indizes die nötige Verallgemeinerung für

aperiodische Bewegungen formal sehr wohl durchführbar ist. Die Unzulänglichkeit der bisher dargestellten Theorie liegt also weniger auf der mathematischen als auf der *begrifflichen* Seite: Der Begriff des *Ortes* eines Elektrons ist uns zunächst verloren gegangen bei der ausschließlichen Betrachtung der „FOURIER-Koeffizienten“ (Matrizelemente) der Bewegung. Offenbar erfassen also die oben in kurzer Wiederholung erläuterten Begriffe nur einen kleinen (wenn auch sehr wichtigen) Teil dessen, was man für die vollständige Mechanik der Elektronen und Atome braucht. Vor allem müssen noch andere Größen als „direkt beobachtbar“ zugelassen werden, und es entsteht die Aufgabe, den Begriff der Beobachtbarkeit im Zusammenhang mit dem Formalismus der Theorie zu analysieren.

Hier fällt nun auf, daß einige wichtige mathematische Hilfsbegriffe der Theorie bislang ohne physikalische Deutung geblieben sind: Vor allem die Vektoren im HILBERTSchen Raume, die bisher nur als bequemes Mittel zur Handhabung der physikalisch deutbaren Matrizen dienten, ferner die unitären Matrizen, durch welche Abbildungen des HILBERTSchen Raumes auf sich dargestellt werden. Es wird sich zeigen, daß gerade diese Dinge für die begriffliche Fortentwicklung der Theorie von größter Bedeutung sind. Die dargelegten Unvollkommenheiten werden behoben durch einen einfachen Gedankengang¹, der die ursprünglichen Ansätze HEISENBERGS in natürlicher Weise fortsetzt und gleich ihnen eine folgerichtige Fortentwicklung der ursprünglichen Ideen BOHRS bedeutet. Er beruht auf einer engen Verschmelzung von Mechanik und Statistik zu einer neuen, einheitlichen Theorie. Die ersten Schritte in dieser Richtung sollen hier getan werden, wobei wir uns möglichst an bereits bewährte Begriffe halten und nur allmählich neue Annahmen einführen.

In zwei Punkten aber wollen wir, der inneren Logik unseres Formalismus vertrauend, von vornherein eine radikale Erweiterung und Verallgemeinerung unserer Grundannahmen wagen. Erstlich wollen wir die Vorstellung als erlaubt ansehen, daß *jede*

¹ DIRAC, P. A. M.: Proc. Roy. Soc. London (A) Bd. 113, S. 621. 1927. JORDAN, P.: Z. Phys. Bd. 40, S. 809. 1927; Bd. 44, S. 1. 1927. Vgl. auch den diesbezüglichen Bericht von D. HILBERT, J. v. NEUMANN und L. NORDHEIM: Math. Ann. Bd. 98, S. 1. 1927.

durch eine hermitesche Matrix dargestellte Größe F an einem (abgeschlossenen) System in demselben Sinne eine „meßbare“ Größe ist wie die Energie des Systems. Das heißt: es soll bestimmte Zahlenwerte geben, welche für diese Größe die Rolle von „möglichen Werten“ spielen, und es soll ein physikalisches Meßverfahren denkbar sein, durch das entschieden wird, welchen dieser Werte die Größe F in einem konkreten Falle gerade angenommen hat — ähnlich, wie man physikalisch feststellen kann, daß ein Atom sich gerade z. B. im Zustand kleinster Energie befindet.

Dabei soll ferner allgemein angenommen werden: Ist die einer Matrix F entsprechende physikalische Größe gemessen, so ist damit ohne weiteres auch jede physikalische Größe gemessen, deren Matrix von der Form $g(F)$, also eine Funktion von F ist (in dem Sinne von § 14); und zwar ist dann der Wert von $g(F)$ gleich der Funktion $g(f_k)$ mit dem gerade gefundenen Wert f_k von F als Argument.

Zweitens nehmen wir an: *Die möglichen Werte einer physikalischen Größe F sind* (wie im Beispiel der Energie) *die Eigenwerte der zugehörigen Matrix.* Daß diese Eigenwerte nicht nur diskrete Zahlen, sondern auch stetige Wertebereiche umfassen können, wurde schon früher gesagt (s. § 18).

Zur anschaulichen Erläuterung seien etwa die rechtwinkligen Ortskoordinaten x , y , z des Elektrons im Wasserstoffatom betrachtet, oder auch die Koordinate q eines harmonischen Oszillators. Unsere erste Hypothese fordert, daß es möglich sei, eine Koordinate x oder q selbst zu messen. Welche Art von Experimenten hierzu in Frage kommen, werden wir allerdings erst später erörtern können. Die „möglichen Werte“ müssen aber in diesen Fällen alle reellen Zahlen von $-\infty$ bis $+\infty$ umfassen. Die zweite Annahme schließt daher die Forderung ein, daß die Matrizen der Größen x , y , z oder q ein kontinuierliches Eigenwertspektrum von $-\infty$ bis $+\infty$ haben. Wir werden später (Teil II) zeigen, daß sich das tatsächlich aus unseren Ansätzen ergibt; nämlich aus den kanonischen Vertauschungsregeln und der in § 23 besprochenen Forderung, daß $p^2 + q^2$ in nicht-singulärer Weise auf Diagonalform transformierbar ist.

Im folgenden beschränken wir uns auf Größen mit diskreten Eigenwerten.

§ 53. **Wahrscheinlichkeiten quantenmechanischer Größen in stationären Zuständen.** Ein abgeschlossenes System, das wir vorläufig als nicht entartet voraussetzen, habe die Energie

$$(1) \quad H(p, q) = W,$$

wo W Diagonalmatrix ist. Ferner sei $F(p, q)$ irgend eine (hermitesche) Matrixgröße, von der nur angenommen werden soll, daß sie ein diskretes Eigenwertspektrum besitzt. Um ein anschauliches Beispiel zu haben, denke man sich etwa das gegebene System aus zwei locker gekoppelten Teilsystemen zusammengesetzt, und verstehe unter $F(p, q)$ die Energiefunktion eines der Teile bei Lösung der Koppelung.

Diese Matrix F möge durch die unitäre Matrix U auf Hauptachsen transformiert werden:

$$(2) \quad F = U f U^{-1}, \quad f = (f_k \delta_{kl}).$$

Der Zeitmittelwert von F im n -ten stationären Zustande des Systems wird also (wegen $U^{-1} = U^\dagger$)

$$(3) \quad F_{nn} = \sum_k U_{nk} f_k U_{kn}^{-1} = \sum_k f_k |U_{nk}|^2,$$

wo

$$(4) \quad \sum_k |U_{nk}|^2 = 1$$

ist.

Stellen wir uns nun vor, wir wüßten von dem betrachteten System (auf Grund einer Messung), daß es im n -ten Quantenzustande ist. Nun soll durch eine andere Messung der Wert der Größe F bestimmt werden; die Tatsache einer *zeitlichen Veränderlichkeit* von F wird sich darin zeigen müssen, daß wir bei mehrfacher Wiederholung des Experiments *verschiedene* der möglichen Werte f_1, f_2, \dots verwirklicht finden.

Die Formeln (3), (4) legen nun für den Fall, daß f_k ein einfacher Eigenwert ist, folgende Vermutung nahe: *Die Zahl $|U_{nk}|^2$ gibt die Wahrscheinlichkeit an, daß bei der Beobachtung des im n -ten Quantenzustande befindlichen Systems die Größe F gerade den Wert f_k annimmt.* Denn mit dieser Annahme verstehen wir den Sinn der beiden Formeln: Gleichung (3) liefert den Zeitmittelwert F_{nn} als statistischen Mittelwert der Eigenwerte f_k mit den Gewichten $|U_{nk}|^2$ oder als „Erwartungswert“, und zugleich zeigt (4), daß die Summe der Wahrscheinlichkeiten aller möglichen Fälle den Wert 1 hat.

Diese physikalische Deutung der Formel (3) bewährt sich auch dann, wenn wir statt der Größe F irgend eine Funktion $g(F)$ betrachten (z. B. F^2). Wenn für F bei einer Messung der Wert f_k gefunden wird, so muß dabei die Größe $g(F)$ den Wert $g(f_k)$ angenommen haben; der Erwartungswert von $g(F)$ im Zustande n muß daher nach der Wahrscheinlichkeitsrechnung

$$\sum_k |U_{nk}|^2 g(f_k)$$

sein. Das ist aber in der Tat wieder gleich dem n -ten Diagonalelement der Matrix

$$g(F) = U g(f) U^{-1}.$$

Es folgt aber umgekehrt, daß die fraglichen Wahrscheinlichkeiten *notwendig* gleich $|U_{nk}|^2$ angenommen werden müssen, damit die aus ihnen berechneten *statistischen Mittelwerte* einer beliebigen Größe stets gleich den entsprechenden *Zeitmittelwerten*, also jeweils gleich dem betreffenden Diagonalelement der zu der Größe gehörenden Matrix werden. Um das zu beweisen, wählen wir für die Funktion $g(x)$ speziell die dem Eigenwert f_k zugehörige Einzelmatrix $F^{(k)}$ in der Zerlegung der Einheit von F . Sie ist eine Funktion $F^{(k)}(F)$ und hat die Eigenschaft, den Wert 1 anzunehmen, wenn F den Wert f_k annimmt, und den Wert Null, wenn F einen von f_k verschiedenen Wert annimmt. Bildet man also von einer, im n -ten Zustande des Systems ausgeführten Beobachtungsreihe der Größe F , bei der sich der Reihe nach $f_{l_1}, f_{l_2}, f_{l_3}, \dots$ als Messungsergebnisse einstellten, die zugehörigen $F^{(k)}(f_{l_1}), F^{(k)}(f_{l_2}), F^{(k)}(f_{l_3}), \dots$, so sind diese so oft 1, als f_k beobachtet worden ist, sonst Null; ihr Mittelwert, dargestellt durch $[F^{(k)}(F)]_{nn}$, ist also die relative Häufigkeit der „günstigen Fälle“, d. h. die Wahrscheinlichkeit, daß f_k gefunden wurde. Andererseits ist aber

$$(5) \quad [F^{(k)}(F)]_{nn} = (U F^{(k)} U^{-1})_{nn} = |U_{nk}|^2,$$

und damit ist die Notwendigkeit unserer Annahme bewiesen.

Ist f_k nicht einfacher, sondern mehrfacher Eigenwert, so wird offenbar die Wahrscheinlichkeit gleich der Summe

$$(6) \quad \sum_{f_l = f_k} |U_{nl}|^2$$

derjenigen $|U_{nl}|^2$ zu setzen sein, für welche $f_l = f_k$ ist; das sieht man auf genau dieselbe Weise ein.

Die unitäre Matrix U , welche F auf die Diagonalform transformiert, ist durch diese Forderung nicht eindeutig bestimmt. Wir haben uns zu überzeugen, daß trotzdem die als Wahrscheinlichkeiten gedeuteten $|U_{nk}|^2$ bzw. ihre Summen (6) eindeutig festgelegt sind. Das ist tatsächlich auf Grund unserer früheren Ergebnisse (Kap. III, § 20) sofort ersichtlich; wir kommen darauf in den nächsten Paragraphen in allgemeinerem Zusammenhange zurück.

§ 54. Allgemeine Relativwahrscheinlichkeiten zweier Größen¹.

Die Fortentwicklung der im vorigen Paragraphen ausgeführten Betrachtungen veranlaßt uns, die bisherige Darstellung physikalischer Größen durch Matrizen einer Kritik zu unterziehen und zu verallgemeinern. Wir hatten uns bis jetzt durchgehend an die von HEISENBERG eingeführte Auffassung gehalten, nach welcher die Matrixdarstellung das korrespondenzmäßige Analogon der klassischen Entwicklung in (einfache oder mehrfache) Fourierreihen nach der Zeit bedeutete. Dementsprechend war die *Energie* als eine *Diagonalmatrix* darzustellen. Hierdurch und durch die kanonischen Vertauschungsregeln waren bei nicht-entarteter Energie die Matrizen q_k , p_k bis auf Phasenkonstanten eindeutig festgelegt. Bei bestimmter Wahl dieser Phasenkonstanten war also einer jeden, als Funktion der q_k , p_k definierten physikalischen Größe $A(q, p)$ eine *eindeutig bestimmte Matrix* zugeordnet. Bei entarteter Energie gab es dagegen zwar verschiedene Lösungen der Bewegungsgleichungen; doch ergab sich auch dort Eindeutigkeit der Lösung (bis auf die Phasen), wenn man das entartete Problem in bestimmter Weise als Grenzfall eines nichtentarteten betrachtete.

Statt dessen wollen wir nunmehr die *hermiteschen Tensoren des HILBERT-Raums als Darstellungen der physikalischen Größen* gebrauchen. (Vgl. Kap. II, § 10.) Es soll also jede der Größen q_k , p_k durch einen HILBERT-Tensor dargestellt werden; für diese Tensoren sollen die invarianten Tensorgleichungen (kanonische Vertauschungsregeln)

$$(1) \quad [p_k, q_l] = \delta_{kl}; \quad [p_k, p_l] = [q_k, q_l] = 0$$

¹ DIRAC, P. A. M.: Proc. Roy. Soc. (A) Bd. 113, S. 621. 1927. JORDAN, P.: Z. Phys. Bd. 40, S. 809. 1927; Bd. 41, S. 797. 1927; Bd. 44, S. 1. 1927.

gelten. Danach entspricht auch jeder Größe $A(q, p)$ ein bestimmter HILBERT-Tensor, da die Prozesse, durch welche die Funktion $A(q, p)$ aus ihren Argumenten q_k, p_k aufgebaut ist (im wesentlichen Multiplikationen und Additionen), *kovariante Tensoroperationen* sind. (Vgl. Kap. II, § 10.)

In einem bestimmten Koordinatensystem $\Sigma^{(1)}$ des HILBERT-Raumes wird der Tensor A durch ein *Komponentenschema*, also eine *Matrix*, dargestellt, die jetzt $A^{(1)}$ genannt werden möge. Ist dann $\Sigma^{(2)}$ ein anderes Koordinatensystem, und $A^{(2)}$ das zugehörige Komponentenschema von A , so besteht die Beziehung

$$(2) \quad A^{(2)} = U^{-1} A^{(1)} U,$$

wo U diejenige unitäre Matrix ist, welche den Übergang von $\Sigma^{(1)}$ nach $\Sigma^{(2)}$ vermittelt.

Die Matrix $A^{(1)}$ eines Tensors A in $\Sigma^{(1)}$ ist dann und nur dann die zur physikalischen Größe A gehörige Matrix in dem früheren HEISENBERG'schen Sinne, wenn $\Sigma^{(1)}$ so gewählt ist, daß der HILBERT-Tensor $H(q, p)$ der *Energie* in $\Sigma^{(1)}$ durch eine *Diagonalmatrix*

$$(3) \quad H^{(1)} = W$$

dargestellt wird.

Die Matrix $A^{(1)}$ eines Tensors A in $\Sigma^{(1)}$ kann natürlich überestimmen mit der Matrix $B^{(2)}$ eines *anderen* Tensors B in einem *anderen* Koordinatensystem $\Sigma^{(2)}$; man muß darauf achten, daß dieser Umstand nicht zu Verwechslungen Anlaß gibt. Schon in Kap. III, § 18 haben wir auf die zweifache Bedeutung der unitären Transformationen

$$q_k = U^{-1} q_k^0 U, \quad p_k = U^{-1} p_k^0 U$$

hingewiesen. Betrachten wir sie in dem dort erläuterten Sinne mit $U = U(q^0, p^0)$ als *kanonische Transformation* (z. B. orthogonale Transformation der dreidimensionalen rechtwinkligen Koordinaten und der zugehörigen Impulse), so bezieht sie sich auf die Matrizen $q_k = Q_k^{(1)}, q_k^0 = q_k^{(1)}$ usw. *verschiedener Größen* Q_k, q_k in *demselben Koordinatensystem* $\Sigma^{(1)}$ des HILBERT-Raums. Sie stellt dann eine *invariante* Beziehung zwischen den Tensoren Q_k, q_k dar. Betrachten wir sie aber als *Transformation im HILBERT-Raum*, so bezieht sie sich auf die verschiedenen Matrixschemata $q_k = q_k^{(2)}, q_k^0 = q_k^{(1)}$ *derselben Größen* q_k in *verschiedenen Koordinatensystemen* $\Sigma^{(1)}, \Sigma^{(2)}$ des HILBERT-Raums.

Diese verallgemeinerte Beschreibung der physikalischen Größen durch HILBERT-Tensoren (statt durch die früher betrachteten speziellen Matrizen) führt uns nun ganz von selbst zu einer Beantwortung der folgenden Frage: Es sei eine nichtentartete Größe A gemessen und dabei ihr Eigenwert a_n verwirklicht gefunden. Danach soll eine andere nichtentartete Größe B gemessen werden. Wie groß ist die Wahrscheinlichkeit, daß von ihren Eigenwerten gerade b_k gefunden wird? Dies ist offenbar eine naturgemäße Verallgemeinerung der im vorigen Paragraphen betrachteten Frage, bei welcher die als gemessen vorausgesetzte Größe immer speziell die *Energie* des Systemes war. Unser Formalismus legt eine bestimmte, einfache Antwort nahe. Diese ist allerdings als eine nicht aus den bisher formulierten Gesetzen ableitbare, sondern wesentlich darüber hinausgehende Hypothese anzusehen, über deren Richtigkeit nur empirisch, an Hand konkreter Anwendungen, zuverlässig entschieden werden kann.

Wir bemerken zunächst, daß unsere Frage keinerlei Bezug mehr nimmt auf die Energie des Systems, die durch einen dritten Tensor $H(p, q)$ dargestellt ist. Die Antwort ist dementsprechend *von der Gestalt dieser Energiefunktion ganz unabhängig*. Die Energie verliert in dieser Hinsicht die Ausnahmestellung, die sie bisher besaß, so daß es auch nicht mehr sinnvoll erscheint, in der Erörterung dieser Fragen die spezielle HEISENBERGSche Matrixdarstellung vor anderen Matrixdarstellungen (in anderen HILBERTschen Koordinatensystemen) zu bevorzugen.

Wir können nun den Ansatz für die relativen Wahrscheinlichkeiten von § 53 in folgender Weise verallgemeinern. Zunächst hat man ein solches Koordinatensystem $\Sigma^{(1)}$ im HILBERT-Raume aufzusuchen, in dem A als Diagonalmatrix

$$(4) \quad A^{(1)} = (\delta_{nm} a_n)$$

erscheint; sodann haben wir ebenso ein solches Koordinatensystem $\Sigma^{(2)}$ im HILBERT-Raume zu suchen, in welchem B als Diagonalmatrix

$$(5) \quad B^{(2)} = (\delta_{kl} b_l)$$

erscheint. Dann sei wieder U die unitäre Matrix, die den Übergang von $\Sigma^{(1)}$ nach $\Sigma^{(2)}$ vermittelt; es sind dann die Darstellungen von A in $\Sigma^{(2)}$ und von B in $\Sigma^{(1)}$ gegeben durch

$$(6) \quad \begin{cases} A^{(2)} = U^{-1} A^{(1)} U = \left(\sum_k U_{nk}^{-1} a_k U_{km} \right), \\ B^{(1)} = U B^{(2)} U^{-1} = \left(\sum_l U_{nl} b_l U_{lm}^{-1} \right). \end{cases}$$

Indem wir die Ergebnisse des § 53 verallgemeinern, machen wir die Annahme, daß die Größen $|U_{nk}|^2$ die gesuchten Relativwahrscheinlichkeiten sind. Wir formulieren das so:

Wenn bekannt ist, daß eine Größe A den Wert a_n besitzt, so ist die Wahrscheinlichkeit, einen bestimmten Eigenwert b_k der Größe B zu finden, gegeben durch

$$(7) \quad w_{nk} = |U_{nk}|^2,$$

wo U die unitäre Matrix ist, die das Hauptachsensystem $\Sigma^{(1)}$ von A in das $\Sigma^{(2)}$ von B überführt¹.

Diese Wahrscheinlichkeit w_{nk} kann auch ohne explizite Bezugnahme auf die Koordinatensysteme $\Sigma^{(1)}$ und $\Sigma^{(2)}$ folgendermaßen invariant ausgedrückt werden². Es seien

$$(8) \quad \begin{cases} A = \sum_n a_n F^{(n)}, \\ B = \sum_k b_k G^{(k)} \end{cases}$$

die invarianten Zerlegungen von A und B in hermitesche Einzeltensoren; es gilt also (s. Kap. II, § 11)

$$(9) \quad \begin{cases} F^{(n)} F^{(m)} = \delta_{nm} F^{(n)}, & G^{(k)} G^{(l)} = \delta_{kl} G^{(k)}; \\ \sum_n F^{(n)} = \sum_k G^{(k)} = 1. \end{cases}$$

Dann ist die Wahrscheinlichkeit gegeben durch die invariante Größe

$$(10) \quad w_{nk} = Sp (F^{(n)} G^{(k)}).$$

Denn da nach Voraussetzung die Tensoren A , B nicht entartet sind, besteht die Matrix jedes Einzeltensors in dem betreffenden Hauptachsensystem aus lauter Nullen, ausgenommen eine einzige 1 in der Diagonale:

$$(11) \quad \begin{cases} F^{(n)}(n', m') = \delta_{nn'} \delta_{nm'}, \\ G^{(k)}(k', l') = \delta_{kk'} \delta_{kl'}; \end{cases}$$

¹ DIRAC, P. A. M.: a. a. O. JORDAN, P.: a. a. O.

² NEUMANN, J. v.: a. a. O.

also ist nach (6):

$$(11') \quad G^{(k)}(n', m') = U_{n'k} U_{km'}^{-1},$$

und (10) bedeutet

$$(12) \quad w_{nk} = Sp(F^{(n)} G^{(k)}) = U_{nk} U_{kn}^{-1} = |U_{nk}|^2.$$

Die Formel (10) zeigt unmittelbar, daß die Wahrscheinlichkeiten durch die Tensoren A und B *eindeutig bestimmt* sind. Man kann das natürlich auch an der ursprünglichen Definition (7) mit Hilfe der unitären Matrix U sehen: Das Koordinatensystem $\Sigma^{(1)}$ ist bei nicht entartetem A festgelegt bis auf eine beliebige Permutation der Achsen (Numerierung der Eigenwerte) und Transformation mit einer Phasenmatrix. Entsprechendes gilt von $\Sigma^{(2)}$. Also kann man — abgesehen von den belanglosen Permutationen — die Matrix U im allgemeinsten Falle ersetzen durch U' mit

$$U'_{nk} = e^{i\varphi_n} U_{nk} e^{-i\psi_k},$$

wo φ_n, ψ_k beliebige reelle Phasenkonstanten sind. Dabei bleibt aber

$$|U'_{nk}|^2 = |U_{nk}|^2 = w_{nk}.$$

Wir heben ein aus unserer Formulierung unmittelbar ersichtliches *Symmetriegesetz der Wahrscheinlichkeiten* hervor: *Die Größe w_{nk} ist auch die Wahrscheinlichkeit für den Wert a_n von A bei gegebenem Werte b_k von B .*

Wenden wir das auf den in § 53 betrachteten Fall an, wo die *Energie* die als bekannt vorausgesetzte Größe war, so ergibt sich folgendes: Ist eine beliebige Größe F gemessen worden, so ist hernach im allgemeinen die Energie W nicht mehr fest bestimmt, sondern zeigt ihrerseits eine Statistik gemäß w_{nk} . Man hat das so zu verstehen, daß der *Prozeß der Messung von F* unvermeidlicherweise bestimmte Wahrscheinlichkeiten für das Eintreten eines Quantensprunges der Energie erzeugt, daß also ein *Festhalten* eines einmal gemessenen Energiewerts bei Messung einer anderen Größe F im allgemeinen nicht möglich ist. Wir kommen darauf noch zurück. (Vgl. § 56.)

Die jetzt gewonnene Darstellung der Wahrscheinlichkeiten ist sehr bequem für die Durchführung einer nunmehr zu besprechenden Verallgemeinerung. Bisher hatten wir vorausgesetzt,

daß (ebenso wie die Größe A) die betrachtete Größe B nicht entartet ist. Wir lassen nun diese Annahme für die Größe B fallen. Um zum richtigen Ansatz für diesen Fall zu kommen, denken wir uns zunächst die Eigenwerte von B alle verschieden und betrachten dann den Spezialfall, daß eine Gruppe von Eigenwerten zusammenrückt. Seien also zunächst $b_j, b_k, \dots b_l$ irgendwelche einfachen Eigenwerte von B ; welches ist die Wahrscheinlichkeit, daß bei einer Messung entweder b_j oder b_k oder \dots oder b_l gefunden wird? Wir wollen dafür kurz sagen: Der Wert von B soll im „Intervall“ $I = (b_j, b_k, \dots b_l)$ liegen — auch dann, wenn die Eigenwerte $b_j, b_k, \dots b_l$ keineswegs benachbart, sondern durch andere getrennt sind.

Die gesuchte Wahrscheinlichkeit, daß B im Intervall $I = (b_j, b_k, \dots b_l)$ liegt, ist der Erwartungswert derjenigen Größe, die gleich 1 ist, wenn B einen der Eigenwerte von I annimmt, und sonst gleich Null; diese Größe ist offenbar der Einzeltensor

$$(13) \quad G(I) = G^{(j)} + G^{(k)} + \dots + G^{(l)};$$

Wir nennen ihn den zum Intervall I gehörigen Einzeltensor.

Die gesuchte Wahrscheinlichkeit ist dann

$$(14) \quad w_I = Sp(F^{(n)} G(I)).$$

Wie man sieht, hat also die Wahrscheinlichkeit für ein Intervall I dieselbe Gestalt wie für einen einzigen einfachen Eigenwert, wobei nur an die Stelle des zu diesem gehörigen Einzeltensors der Einzeltensor des Intervalles tritt.

Unter Benutzung der Matrix U an Stelle der Einzeltensoren $F^{(n)}, G^{(k)}$ können wir für (14) auch

$$(14') \quad w_I = |U_{nj}|^2 + |U_{nk}|^2 + \dots + |U_{nl}|^2$$

schreiben.

Es ist nun ohne weiteres klar, daß diese Formeln gültig bleiben müssen, wenn die Eigenwerte des Intervalles I zusammenrücken, wenn es sich also um die Messung eines mehrfachen Eigenwertes handelt: dann geht $G(I)$ in den diesem mehrfachen Eigenwert entsprechenden Einzeltensor in der zu B gehörenden Zerlegung der Einheit über. Wir haben damit das Resultat:

Die Formel

$$w_{nk} = Sp(F^{(n)} G^{(k)})$$

gilt auch dann, wenn die Größe $B = \sum_k G^{(k)} b_k$ mehrfache Eigenwerte hat; und allgemeiner gilt für ein Intervall I von Eigenwerten die Formel (14), (14').

Dagegen müssen wir, wie nochmals betont sei, an der Voraussetzung festhalten, daß die ursprünglich gemessene Größe A nicht-entartet (oder mindestens ihr fraglicher Eigenwert a_n einfach) ist.

§ 55. **Wahrscheinlichkeitsvektoren**¹. Wir betrachten wieder zwei Größen A, B und setzen, wie oben, eine Messung von A mit dem Ergebnis a_n voraus. Zur Berechnung der Wahrscheinlichkeiten für das Auftreten der Werte b_1, b_2, \dots einer Größe B brauchen wir dann von allen ∞^2 Elementen U_{mi} der Matrix U nur die Zeile $m = n$, die zu a_n gehört, also die Größen U_{nk} ($k = 1, 2 \dots$), oder, was auf dasselbe hinausläuft,

$$(1) \quad U_{n1}^*, \quad U_{n2}^*, \quad U_{n3}^*, \quad \dots$$

Wir wollen nun für B der Reihe nach verschiedene Größen nehmen und demgemäß auch das Koordinatensystem $\Sigma^{(2)}$ als *veränderlich* betrachten, während $\Sigma^{(1)}$ *festgehalten* wird. Dann behaupten wir:

Jede Zeile (1) der Transformationsmatrix U^ verhält sich bei Änderung von $\Sigma^{(2)}$ wie ein HILBERT-Vektor.*

Ist x ein beliebiger HILBERT-Vektor, so geht er definitionsgemäß bei der Transformation U in

$$(2) \quad x' = U^{-1}x$$

mit den Komponenten

$$(2a) \quad x'_m = \sum_k U_{mk}^{-1} x_k = \sum_k U_{km}^* x_k$$

über. Lassen wir nun $\Sigma^{(2)}$ mit $\Sigma^{(1)}$ zusammenfallen, so ist U die Einheitsmatrix, die Zahlenreihe (1) lautet also

$$(3) \quad 0, 0, 0, \dots 0, 1, 0, \dots,$$

wo die 1 an n -ter Stelle steht. Derjenige HILBERT-Vektor der Länge 1, der diese Zahlen (3) in $\Sigma^{(1)}$ zu Komponenten hat, d. h. der zum Eigenwert a_n von A gehörige Eigenvektor, hat aber nach (2a) in $\Sigma^{(2)}$ gerade die Komponenten (1). Nennen wir diesen

¹ NEUMANN, J. v.: a. a. O.

Vektor c , und seine Komponenten in $\Sigma^{(2)}$, die bisher U_{nk}^* hießen, einfach c_k , so ist $w_{nk} = |c_k|^2$ die Wahrscheinlichkeit, daß B den einfachen Eigenwert b_k annimmt, nachdem zuvor der zum Eigenvektor c gehörige Eigenwert a_n von A gemessen ist. Statt dessen kann man auch die invariante Schreibweise

$$w_{nk} = |c_k|^2 = (G^{(k)} c, c)$$

benutzen, da die Einzelform $G^{(k)}$ in $\Sigma^{(2)}$ die Darstellung § 54, (11) hat.

Wir wollen nun folgende Ausdrucksweise benutzen: *Ein reiner Fall¹ liegt vor, wenn durch Messung festgestellt ist, daß eine nichtentartete Größe A einen bestimmten Eigenwert a annimmt; dieser läßt sich kennzeichnen durch Angabe des zu a gehörigen Eigenvektors c vom Betrage $|c|^2 = 1$, den wir Wahrscheinlichkeitsvektor nennen. Man berechnet dann die Wahrscheinlichkeit w_k für den Eigenwert b_k einer Größe $B = \sum_k b_k G^{(k)}$ nach der Formel*

$$(4) \quad w_k = (G^{(k)} c, c).$$

Der Vektor c ist also nicht für die ganze Größe A charakteristisch, sondern nur für die zum Eigenwert a_n gehörige Einzelgröße $F^{(n)}$ von A . Die Angabe von c oder von $F^{(n)}$ ist die natürlichste Beschreibung des Meßergebnisses, da alle von $F^{(n)}$ verschiedenen Einzeltensoren $F^{(m)}$ in A für die aus der Messung zu ziehenden Schlüsse belanglos sind. Zwischen $F^{(n)}$ und c besteht der invariante Zusammenhang [s. § 14, (11)]

$$(5) \quad F^{(n)} = c \times c ;$$

man kann auch wegen dieses Zusammenhanges auf Grund der Tensorformeln § 9, (14), § 10, (9), Kap. II die Formel (10), § 54 unmittelbar in (4) überführen. Für die Wahrscheinlichkeit, daß der Wert von B in ein Intervall I fallen werde, bekommt man entsprechend die mit (14), (14'), § 54 gleichwertige Formel

$$(6) \quad w_I = (G(I) c, c).$$

Der *statistische Mittelwert* \overline{B} von B in dem reinen Fall, der durch den Wahrscheinlichkeitsvektor c beschrieben wird, ist nach

¹ Dieser Begriff wurde unabhängig von H. WEYL: Z. Phys. Bd. 46, S. 1. 1927 und J. v. NEUMANN: Gött. Nachr. 1927, S. 245 eingeführt.

(4) nichts anderes als die mit c gebildete zu B gehörige hermitesche Form, die dadurch eine physikalische Bedeutung bekommt; denn es ist

$$(7) \quad \overline{B} = \sum_k b_k w_k = \sum_k b_k (G^{(k)} c, c) = (B c, c).$$

Hieraus ergibt sich eine wichtige Folgerung: *Der Mittelwert der Summe zweier Größen B_1, B_2 ist gleich der Summe der Mittelwerte der Summanden:*

$$(8) \quad \overline{B_1 + B_2} = \overline{B_1} + \overline{B_2};$$

denn dies bedeutet

$$((B_1 + B_2) c, c) = (B_1 c, c) + (B_2 c, c).$$

In diesem Satze kann man geradezu die *physikalische Definition* der bislang nur formal (durch die Addition der darstellenden Matrizen) definierten Addition quantenmechanischer Größen sehen: *Die Summe B von B_1 und B_2 ist diejenige, eindeutig definierte Größe, deren Mittelwert unter allen Umständen (für jeden Wahrscheinlichkeitsvektor c) die Summe der Mittelwerte von B_1 und B_2 ist.*

§ 56. Grenzen der Meßbarkeit. Mit Hilfe der im Voranstehenden bestimmten Relativwahrscheinlichkeiten ergeben sich neue Aufschlüsse über das Problem der *Meßbarkeit physikalischer Größen* an einem quantenmechanischen System. Zwar kann im Prinzip jede Größe A eines solchen Systems gemessen werden; aber es ist *im allgemeinen nicht möglich, eine Größe A und zugleich eine andere Größe B (die natürlich für sich allein auch meßbar ist) zu messen.* Haben wir nämlich für A einen Wert a_n und damit den zugehörigen Wahrscheinlichkeitsvektor c gefunden, so wird andererseits eine Messung von B , die b_k ergibt, einen Wahrscheinlichkeitsvektor c' liefern, und die Eigenschaften der physikalischen Statistik des Systems, die wir auf Grund von c' vorhersagen können, stimmen nur dann mit den aus c abgeleiteten widerspruchslos überein, wenn die beiden Einzeltensoren $F^{(n)} = c \times c$ und $G^{(k)} = c' \times c'$ übereinstimmen, so daß also $c = e^{i\varphi} \cdot c'$ ist.

Wenn aber $c \times c \neq c' \times c'$ ist, so zerstört die zu c' führende Messung den vorher bestehenden, zu c gehörigen Zustand. Für zwei spezielle Eigenwerte a_n, b_k kann die Bedingung $F^{(n)} = G^{(k)}$ erfüllt sein, obwohl zu A und B im übrigen ganz verschiedene

Einzelntensoren gehören. Betrachten wir aber die *Gesamtheit der möglichen Messungsergebnisse*, so gilt folgendes:

Zwei nichtentartete physikalische Größen A und B sind dann und nur dann zugleich meßbar, wenn sie vertauschbar sind.

Ist nämlich diese Voraussetzung $[A, B] = 0$ erfüllt, so gibt es nach § 13 ein Koordinatensystem im HILBERT-Raum, in welchem die Matrixdarstellungen von A und B gleichzeitig Diagonalform haben. Also ist die (genauer: *eine*) Transformation, die von A auf B überführt, die Identität, $U = 1$; mithin ist $w_{nm} = |U_{nm}|^2 = \delta_{nm}$. Die Wahrscheinlichkeit, bei einer Messung von A , die a_n liefert, den Wert b_n von B zu finden, ist also 1, irgend einen anderen Wert b_m ($m \neq n$) zu finden, aber Null. Also ist B mit A zugleich meßbar. Umgekehrt ist aber die Voraussetzung $[A, B] = 0$ auch notwendig. Sind nämlich A und B zugleich meßbar, so muß nach der obigen Betrachtung, wenn für A ein bestimmter Wert a_n gefunden wird, notwendig zugleich ein durch a_n eindeutig bestimmter Eigenwert b_k von B gefunden werden, nämlich derjenige, dessen $G^{(k)}$ gleich $F^{(n)}$ ist. Indem wir die den a_n zugeordneten Eigenwerte von B mit jeweils denselben Indizes versehen, erhalten wir also, während $A = \sum_n a_n F^{(n)}$ ist, andererseits $B = \sum_n b_n F^{(n)}$; also ist $[A, B] = 0$.

Wir stoßen hier auf den wesentlichsten Unterschied der Quantenmechanik und der klassischen Theorie. Während in dieser angenommen wird, daß jede Größe unabhängig von allen anderen gemessen werden kann, wird man hier in der Quantentheorie durch den Formalismus selbst dazu geführt, diese Annahme fallen zu lassen. *Man kann nur solche Größen zugleich messen, die miteinander vertauschbar sind.*

Diese quantenmechanische Tatsache, daß nicht irgend zwei Messungen miteinander verträglich sind, hängt offenbar eng damit zusammen, daß *jede Messung einen Eingriff in das betrachtete System bedeutet*, der selbst den zuvor bestehenden Zustand stört¹. In der klassischen Theorie wird die stillschweigende, oft fälschlich für selbstverständlich gehaltene Voraussetzung gemacht, daß es immer möglich (oder wenigstens denkbar) ist, die Rückwirkung des Meßapparates auf das zu betrachtende System be-

¹ Vgl. HEISENBERG, W.: Z. Phys. Bd. 43, S. 172. 1927. BOHR, N.: Naturwissensch. Bd. 16, S. 245. 1928; Bd. 17, S. 483. 1929.

liebig klein zu machen. Dies ist im Bereiche der makroskopischen Körper erlaubt, da man die vom Meßinstrument ausgehenden störenden Kraftwirkungen beliebig abschwächen kann, etwa durch Wahl entsprechend *feiner* Instrumente. In der Atomphysik dagegen können die Instrumente nicht beliebig verfeinert werden: Zur Abtastung und Ausmessung atomarer Gebilde hat man immer wieder nur atomare Gebilde zur Verfügung, und man begreift, daß es hier kein eigentliches Messen gibt, sondern ein Koppeln mit als Meßwerkzeug betrachteten Systemen, die, wie das zu beobachtende, selbst quantenmechanische Gebilde sind.

Nach dem Festgestellten kann zugleich mit einer *nichtentarteten* Größe A nur eine solche Größe B gemessen werden, die eine *Funktion von A* ist, deren besondere Messung bei Messung von A sich also erübrigt. Es kann trotzdem sinnvoll sein, mehrere *vertauschbare* Größen A_1, A_2, \dots *zugleich* zu messen, nämlich dann, wenn sie *entartet* sind. Denn es kann durch geeignete Kombination der Messungen mehrerer vertauschbarer entarteter Größen die Messung einer einzigen nichtentarteten ersetzt werden. Wenn nämlich $F_1^{(n_1)}, F_1^{(n_2)}, \dots$ Einzeltensoren sind, die zu gewissen mehrfachen Eigenwerten $a_{n_1}^{(1)}, a_{n_2}^{(2)}, \dots$ von A_1, A_2, \dots gehören, und wenn das Produkt

$$F^{(n)} = F_1^{(n_1)} F_2^{(n_2)} \dots$$

ein *unzerlegbarer* Einzeltensor ist (d. h. mit nur *einem* Eigenwert 1), dann ist (trivialerweise) die gleichzeitige Feststellung der Eigenwerte $a_{n_1}^{(1)}, a_{n_2}^{(2)}, \dots$ äquivalent mit der Feststellung des zu $F^{(n)}$ gehörigen einfachen Eigenwertes a_n einer nichtentarteten Größe A . Wir können sagen: *Die Messung der vertauschbaren entarteten Matrizen A_1, A_2, \dots ersetzt die Messung einer nichtentarteten Größe, wenn alle „Eigenwerte des Matrixsystems“ (nach der Definition von Kap. II, § 14, S. 71) voneinander verschieden sind.*

Diese bei Beschränkung auf Größen mit diskreten Eigenwerten rein triviale Erwägung wird bedeutungsvoll bei der Ausdehnung der Theorie auf Größen mit stetigen Spektralgebieten; denn bei diesen kann man die gleichzeitige Messung mehrerer vertauschbarer Größen nicht mehr ersetzen durch Messung einer einzigen Größe. Man hat deshalb dann nicht mehr Wahrscheinlichkeitsrelationen zwischen den verschiedenen Eigenwerten nicht-

entarteter Matrizen, sondern zwischen den Eigenwerten solcher vertauschbarer *Matrizensysteme*, bei denen keine zwei übereinstimmenden Eigenwerte vorkommen. Ein solches Matrizen-system bildet für den Fall eines Massenpunktes z. B. die rechtwinkligen Koordinaten x, y, z ; ein Eigenwert dieses Systems bedeutet einen möglichen *Ort* des Massenpunktes.

Die Tatsache, daß die Ortskoordinaten nicht vertauschbar sind mit ihren zugehörigen Impulsen p_x, p_y, p_z , besagt nach dem oben Festgestellten, daß *man Ort und Impuls des Massenpunktes nicht zugleich messend exakt bestimmen kann*. Wir werden später sehen (Teil II), daß die *kanonischen Vertauschungsregeln* sich in einer einfachen statistischen Beziehung zwischen den Eigenwerten des Ortes und des Impulsvektors widerspiegeln.

§ 57. Statistik quantenmechanischer Gemenge¹. Es möge an einer Anzahl (gleichartiger) Atome die nichtentartete Größe A gemessen werden, und es möge an einem Teil der Atome der Wert a_1 von A gefunden sein, an anderen dagegen der Eigenwert a_2 , an noch anderen a_3 usw. Greifen wir aus dieser Menge von Atomen ein beliebiges heraus, ohne daß wir wissen, welcher Eigenwert von A bei ihm verwirklicht war, und versuchen wir dann Voraussagen über die Reaktionsweise dieses Atoms bei Vornahme neuer Experimente zu machen, so kommt zu der Unsicherheit, die durch die statistische Natur der quantenmechanischen Gesetze selbst bestimmt ist, noch eine auf *Unvollständigkeit unserer Kenntnis* beruhende Unsicherheit in der Voraussage der künftigen Reaktionen unseres Systems hinzu.

Es ist jedoch auch in diesem Falle möglich, *bestimmte statistische Aussagen* zu machen, sobald wir nur wissen, *an welchem Bruchteil w_n der Atomgesamtheit* ($\sum_n w_n = 1$) *der Eigenwert a_n gefunden war* (während wir also *nicht* wissen, welcher Eigenwert von A bei einem bestimmten *einzelnen* Atom gemessen worden ist). Diese statistischen Aussagen sind zu gewinnen durch eine einfache Kombination der uns bekannten charakteristischen quantenmechanisch-statistischen Gesetze mit der gewöhnlichen *Wahrscheinlichkeitsrechnung*.

Erinnern wir uns, unter welchen Umständen diese letztere anwendbar ist. Es wird ein Versuch ausgeführt, der nicht not-

¹ NEUMANN, J. v.: Gött. Nachr. 1927, S. 245.

wendig zu einem bestimmten Ergebnis führt, sondern eine „zufällige“ Auswahl trifft aus einer Anzahl von möglichen Ergebnissen. Wiederholt man denselben Versuch außerordentlich oft, so kann es vorkommen, daß die relativen Häufigkeiten der verschiedenen möglichen Ergebnisse mit steigender Zahl der Versuche einem Grenzwert zustreben; wenn überdies noch die Bedingung erfüllt ist, daß irgend eine, vorweg bestimmte Teilmenge der Versuche, deren Zahl ebenfalls wächst, dieselben Grenzwerte der relativen Häufigkeit liefert wie die Menge aller Versuche, so nennt man diese Menge ein *Kollektiv*¹ und die Grenzwerte der relativen Häufigkeiten der Merkmale ihre *Wahrscheinlichkeiten*. Diese sind also in einem konkret vorgegebenen Falle durch solche langen Versuchsreihen mit beliebiger Annäherung empirisch bestimmbar.

Die Wahrscheinlichkeitsrechnung analysiert solche Wahrscheinlichkeiten unter Zurückführung auf „unabhängige“ Wahrscheinlichkeiten und untersucht die logischen Beziehungen zwischen den verschiedenen, am gleichen Objekt auszuführenden statistischen Feststellungen.

Um also die gewöhnliche Wahrscheinlichkeitsrechnung anwenden zu können, wird man eine häufige Wiederholung von geeigneten Experimenten vornehmen; entweder zu wiederholten Malen am selben mechanischen System, welches immer wieder bestimmten Bedingungen ausgesetzt wird, oder je einmal an vielen Exemplaren desselben Systems, die insgesamt bestimmten Einflüssen unterworfen waren. Letzteres Verfahren ist anschaulicher und soll daher hier bevorzugt werden. Wir denken uns also eine (praktisch unendlich große) Menge \mathfrak{S} von genau gleichen, mechanisch völlig voneinander isolierten Systemen; ein solches ist beispielsweise gegeben durch ein *klassisches exakt ideales Gas*, dessen Atome keinerlei Wechselwirkungen aufeinander ausüben. Dieses Gas kann sich zu der betrachteten Zeit in irgend einem thermodynamischen Gleichgewichts- oder Nichtgleichgewichtszustand befinden; es bestehen also unendlich viele verschiedene Möglichkeiten der statistischen Verteilung in \mathfrak{S} .

Ist A irgend eine (entartete oder nichtentartete) Größe an einem Einzelsystem (Atom) von \mathfrak{S} , so kann man an einer hinrei-

¹ Vgl. dazu R. v. MISÈS: *Wahrscheinlichkeit, Statistik und Wahrheit*. Berlin 1928. Darin findet sich weitere Literatur angegeben.

chend großen, aber gegenüber \mathfrak{S} sehr kleinen Teilmenge von \mathfrak{S} denjenigen Bruchteil w_n der Systeme bestimmen ($\sum_n w_n = 1$), bei denen der Eigenwert a_n von A angenommen wird, wenn man A einer Messung unterwirft. Da wir voraussetzen, daß \mathfrak{S} ein Kollektiv ist bezüglich dieser Messungsergebnisse, so stellt die Zahl w_n einen Näherungswert dar für die Wahrscheinlichkeit, daß an einem beliebig herausgegriffenen Atom bei Messung von A der Wert a_n gefunden wird. Denkt man sich eine entsprechende „statistische Erhebung“ für *alle* meßbaren Größen A ausgeführt, so gelangt man zu einer *erschöpfenden statistischen Beschreibung des Systems* \mathfrak{S} .

Für tatsächliche empirische Feststellungen statistischer Verteilungen bietet die Atomphysik unzählige Beispiele. Man denke an die Auszählung von α - und β -Strahlen oder Molekularstrahlen; dabei werden die einzelnen Teilchen als isolierte Systeme betrachtet. Dabei gelingt es häufig, die *Anzahl* der Teilchen, die eine bestimmte Eigenschaft haben, auch dann zu messen, wenn diese Eigenschaft beim *Einzelteilchen* unter der praktisch erreichbaren Meßgenauigkeit liegt.

Es ist freilich zu beachten, daß ein Schwarm von Teilchen auch dann, wenn keine energetischen Wechselwirkungen vorliegen, nicht immer als eine Menge \mathfrak{S} isolierter Systeme betrachtet werden darf, sondern nur im Grenzfall hinreichend kleiner Dichten und hoher Temperaturen, kurz dann, wenn die *klassische Theorie der idealen Gase* anwendbar ist. Wir werden später sehen, daß ein ideales Gas unter Umständen (hohe Drucke, tiefe Temperaturen) bezüglich eines Teils seiner Eigenschaften keineswegs als ein Haufen von unabhängigen Molekeln aufgefaßt werden kann (Gasentartung). In den praktisch vorkommenden Beispielen von Molekularstrahluntersuchungen usw. ist diese genannte Bedingung jedoch immer erfüllt.

Wir nennen ein solches Kollektiv \mathfrak{S} quantenmechanischer Systeme ein *Gemenge*, im Gegensatz zum schon behandelten *reinen Fall*, wo für eine gewisse Größe A bei allen Systemen in \mathfrak{S} (oder bei allen an einem System vorgenommenen Experimenten) ein und derselbe Eigenwert a_n gemessen wurde. Auch beim reinen Fall ist das Verhalten irgend einer mit A nicht vertauschbaren Größe nur statistisch bestimmt, aber die Wahrscheinlichkeiten

sind (durch einen HILBERT-Vektor) gesetzmäßig völlig bestimmt, so daß empirische Auszählungen eines Kollektivs nur noch als Prüfung bzw. Bestätigung der Aussagen der Theorie Bedeutung besitzen. Man hat hier „a-priori“-Wahrscheinlichkeiten. Beim Gemenge dagegen treten daneben auch „a-posteriori“-Wahrscheinlichkeiten w_n auf, die in jedem konkreten Beispiel zunächst empirisch neu bestimmt werden müssen. (Natürlich können sie jedoch unter Umständen auch durch theoretische Erwägungen bestimmt werden, wenn man den *Prozeß* kennt, durch welchen das betreffende Gemenge \mathfrak{S} erzeugt worden ist.)

Zu Beginn dieses Paragraphen betrachteten wir ein Gemenge der folgenden Art: Eine nichtentartete Größe A wird an *allen* einzelnen Systemen gemessen, und \mathfrak{S} entsteht durch Mischung der den Eigenwerten a_n entsprechenden reinen Fälle in den relativen Mengen w_n (mit $\sum_n w_n = 1$). Für ein so hergestelltes Gemenge

können wir *trivialerweise* sofort die Wahrscheinlichkeit w bestimmen, daß an einem herausgegriffenen Atom bei Messung von B der Wert b_k gefunden wird. Unter Benutzung der in § 55 bestimmten Relativwahrscheinlichkeiten w_{nk} des reinen Falles erhalten wir offenbar

$$(1) \quad w = \sum_n w_n w_{nk}.$$

Es ist also in diesem Falle bei Kenntnis der zu A gehörigen Größen w_n das statistische Verhalten von \mathfrak{S} bezüglich *beliebiger* Größen B bereits völlig bekannt.

Wir wollen jedoch nunmehr, ohne die in § 55 entwickelte Theorie des reinen Falles als schon bekannt vorauszusetzen, die Theorie der Gemenge deduktiv mittels einiger einfacher Postulate begründen¹. Da sich hierbei zuletzt als Spezialfall wieder zwangsläufig die Theorie des reinen Falles ergeben wird, so erhalten unsere Annahmen von § 55 eine neue theoretische Stütze. Andererseits werden wir beweisen können (§ 59), daß der durch die Formel (1) beherrschte Typus von Gemengen bereits das allgemeinste überhaupt mögliche Gemenge darstellt.

Die statistischen Eigenschaften von \mathfrak{S} sind vollständig bekannt, wenn für jede Größe A der *statistische Mittelwert oder Erwartungswert* in \mathfrak{S} bekannt ist; denn auch die Wahrscheinlichkeiten für das

¹ NEUMANN, J. v.: a. a. O.

Eintreten bestimmter Eigenwerte einer Größe können, wie wir schon in § 53 sahen, durch solche Mittelwerte dargestellt werden, nämlich durch die Erwartungswerte der betreffenden Einzelgrößen.

Wir wollen den Erwartungswert von A mit \bar{A} bezeichnen; natürlich darf man ihn nicht verwechseln mit den Zeitmittelwerten an einem einzelnen, abgeschlossenen System, die wir früher durch die Diagonalelemente der Matrix im Koordinatensystem der Energie dargestellt und ebenso bezeichnet haben. Dann nehmen wir zwischen den Erwartungswerten der verschiedenen Größen folgende Beziehungen an:

1. Sind λ, μ zwei reelle Zahlen, so ist

$$\overline{\lambda A + \mu B} = \lambda \bar{A} + \mu \bar{B}.$$

Daß diese Relation im reinen Fall tatsächlich erfüllt ist, wissen wir aus § 55, (8); daher muß sie, damit den Gesetzen der gewöhnlichen Wahrscheinlichkeitsrechnung Genüge geschieht, auch für beliebige Gemenge gefordert werden.

Wir sprechen ferner noch die selbstverständlichen Forderungen aus:

2. Ist eine Größe P keiner negativen Werte fähig, so ist auch

$$\bar{P} \geq 0.$$

3. Der Erwartungswert einer unveränderlichen Zahl ist gleich dieser Zahl selbst.

Diese Postulate kombinieren wir mit der Grundannahme der Quantenmechanik, daß die meßbaren Größen an einem System durch Matrizen bzw. HILBERT-Tensoren darzustellen seien. Dabei wird insbesondere eine negativer Werte nicht fähige Größe durch einen semidefiniten Tensor dargestellt, und eine unveränderliche Zahl durch den mit ihrem Werte multiplizierten Einheitstensor.

Wir behaupten nun: aus 1. folgt, daß für ein vorgegebenes Kollektiv \mathfrak{S} der Erwartungswert jeder beliebigen Größe A mit Hilfe eines festen, von A unabhängigen, hermiteschen Tensors Φ in der Form

$$(2) \quad (I) \quad \bar{A} = Sp(\Phi A)$$

ausgedrückt werden kann. Dieser Tensor Φ kennzeichnet also vollkommen den statistischen Zustand der Menge \mathfrak{S} .

Zum Beweise denken wir uns A durch eine hermitesche Matrix $A = (A_{nm})$ dargestellt. Zerlegt man diese in $A = A' + iA''$, wo A', A'' reelle Matrizen sind, so gilt für deren Elemente

$$(3) \quad A'_{nm} = A'_{mn}, \quad A''_{nm} = -A''_{mn};$$

also sind von diesen Größen nur die mit $n \geq m$ unabhängig, und A ist zu schreiben in der Form:

$$A = \sum_{n \geq m} (A'_{nm} D'^{(nm)} - A''_{nm} D''^{(nm)}),$$

wobei $D'^{(nm)}$ und $D''^{(nm)}$ hermitesche Matrizen sind, welche nicht von A abhängen (für $n > m$ sei nämlich $D'^{(nm)}$ bzw. $D''^{(nm)}$ die Matrix mit den Elementen $d'_{nm} = d'_{mn} = 1$, sonst lauter Nullen, bzw. $d''_{nm} = -d''_{mn} = -i$, sonst Nullen; während $D''^{(nn)} = 0$ und $D'^{(nn)}$ nur an der n -ten Diagonalstelle eine 1 hat, sonst lauter Nullen). Da die Koeffizienten A'_{nm} und A''_{nm} reell sind, folgt nach der Forderung 1. (wenn man die Erwartungswerte der Größen, welche den Matrizen $\frac{1}{2} D'^{(nm)}$ bzw. $\frac{1}{2} D''^{(nm)}$ zugeordnet sind, mit Φ'_{nm} bzw. Φ''_{nm} bezeichnet):

$$(4) \quad \bar{A} = 2 \sum_{n \geq m} (\Phi'_{nm} A'_{nm} - \Phi''_{nm} A''_{nm})$$

wo Φ'_{nm}, Φ''_{nm} reelle Zahlen sind, die für $n \geq m$ definiert sind. Wir definieren sie auch für $n < m$ durch die Formeln

$$(5) \quad \Phi'_{nm} = \Phi'_{mn}, \quad \Phi''_{nm} = -\Phi''_{mn}.$$

Dann kann man offenbar (4) auch schreiben als:

$$(6) \quad \bar{A} = \sum_{n,m} (\Phi'_{nm} A'_{nm} - \Phi''_{nm} A''_{nm}),$$

wo über alle n, m summiert wird. Andererseits ist wegen (3) und (5):

$$\sum_{n,m} (\Phi'_{nm} A''_{nm} + \Phi''_{nm} A'_{nm}) = 0;$$

fügt man diese Summe, mit i multipliziert, zu (6) hinzu, so kommt:

$$(7) \quad \bar{A} = \sum_{n,m} (\Phi'_{nm} + i\Phi''_{nm}) (A'_{nm} + iA''_{nm}).$$

Nun ist

$$(8) \quad \Phi = (\Phi'_{nm} + i\Phi''_{nm})$$

nach (5) eine hermitesche Matrix; sehen wir sie als Repräsentant eines hermiteschen Tensors Φ an, so ist (7) nichts anderes als die zu beweisende Formel (2), (I).

Nun folgern wir aus der Forderung 2.: *Der Tensor Φ ist semi-definit*; es ist also

$$(9) \text{ (II)} \quad (\Phi x, x) \geq 0$$

für jeden HILBERT-Vektor x .

Denn der aus x gebildete Tensor

$$P = x \times x$$

ist semi-definit, da er lauter verschwindende Eigenwerte hat, außer einem, der gleich 1 ist; er erfüllt also die Voraussetzung von 2., also ist $\bar{P} \geq 0$.

Andererseits ist

$$(\Phi x, x) = Sp(\Phi P) = \bar{P} \geq 0,$$

womit (9), (II) bewiesen ist.

Endlich ergibt sich aus der Forderung 3.: *Die Spur von Φ ist gleich 1*:

$$(10) \text{ (III)} \quad Sp \Phi = 1.$$

Zum Beweise hat man nur in (I) für A den Einheitstensor E einzusetzen, dessen Erwartungswert nach Forderung 3. gleich 1 sein soll:

$$\bar{E} = Sp(\Phi E) = Sp \Phi = 1.$$

Diese Formeln (I), (II), (III), die sich unmittelbar aus den Forderungen 1., 2., 3. zusammen mit der Darstellung physikalischer Größen durch HILBERT-Tensoren ergeben, sind umgekehrt auch hinreichend, um die Forderungen 1., 2., 3. zu gewährleisten. Denn aus (I) und (III) folgt unmittelbar 1. und 3. Um 2. zu beweisen, bemerken wir, daß ein semidefiniter Tensor P die Zerlegung der Einheit $P = \sum_n p_n (x_n \times x_n)$ mit $p_n \geq 0$ zuläßt, so daß aus (II) folgt:

$$Sp(\Phi P) = \sum_n p_n (\Phi x_n, x_n) \geq 0.$$

Wir können daher behaupten:

Jeder den Gleichungen (II), (III) genügende hermitesche Tensor Φ liefert gemäß (I) eine mögliche statistische Zustandsverteilung in der Menge \mathfrak{S} .

Im § 52 haben wir die Annahme eingeführt, daß die möglichen Werte einer physikalischen Größe lediglich ihre Eigenwerte sind. Auch diese Annahme ist bereits eine notwendige Folgerung der hier formulierten Postulate.

Als möglicher Wert einer Größe wird dabei eine solche Zahl bezeichnet, die ein „scharfer“ Wert dieser Größe sein kann (d. h. an allen Systemen einer statistischen Gesamtheit \mathfrak{S} soll für A derselbe Wert a gemessen sein). Dann gilt nämlich für die Wahrscheinlichkeit einen Eigenwert $a^{(k)}$ von A bei der Messung eines Systems in \mathfrak{S} zu finden, d. h. für den Erwartungswert des zu $a^{(k)}$ gehörigen Einzeltensors $F^{(k)}$:

$$Sp(\Phi F^{(k)}) = \begin{cases} 0, & \text{wenn } a^{(k)} \neq a, \\ 1, & \text{wenn } a^{(k)} = a. \end{cases}$$

Der Mittelwert von A ist aber

$$a = \bar{A} = Sp(\Phi A) = \sum_k a^{(k)} Sp(\Phi F^{(k)}).$$

Es muß somit einen Eigenwert $a^{(k)}$ von A geben, der gleich a ist, (oder es ist $a = 0$, in diesem Fall betrachte man aber die Größe $A + 1$).

§ 58. Der reine Fall als spezielles Gemenge¹. Jetzt beweisen wir, daß die in § 55 entwickelte Statistik der reinen Fälle sich als Sonderfall der in den Formeln (I), (II), (III) ausgedrückten allgemeinen Statistik beliebiger Gemenge ergibt. Wir wollen hierfür zunächst eine Hilfsformel ableiten.

Angenommen, wir wissen daß an allen Systemen der Menge \mathfrak{S} die Größe A in einem gewissen, die Eigenwerte $a_{n_1}, a_{n_2}, \dots, a_{n_j}$ enthaltenden Intervalle liegt; es sei $F^{(n)}$ der zu diesem Intervall gehörige Einzeltensor (also die Summe der zu $a_{n_1}, a_{n_2}, \dots, a_{n_j}$ gehörigen Einzeltensoren von A), und Φ sei der Wahrscheinlichkeitstensor, der die unserer Annahme entsprechende Statistik von \mathfrak{S} beschreibt; dann gilt mit Rücksicht auf § 57, (I)

$$(1) \quad \overline{F^{(n)}} = Sp(\Phi F^{(n)}) = 1.$$

Wir behaupten nun: Aus (1) folgt mit Rücksicht auf § 57, (II) und (III):

$$(2) \quad \Phi = F^{(n)} \Phi = \Phi F^{(n)}.$$

¹ NEUMANN, J. v.: Gött. Nachr. a. a. O.

Zum Beweise wählen wir einen Einheitsvektor z , der Eigenvektor zum Eigenwert 0 von $F^{(n)}$ ist, für den also

$$(3) \quad F^{(n)} z = 0$$

gilt. Aus z bilden wir den Einzeltensor

$$(4) \quad P = z \times z,$$

für den nach (3) gilt:

$$(5) \quad F^{(n)} P = F^{(n)} z \times z = 0.$$

Nach § 8, δ sind daher nicht nur $F^{(n)}$ und P Einzeltensoren, sondern auch $F^{(n)} + P$, und nach § 8, β gilt dasselbe auch von $E - F^{(n)} - P$. Da somit P und $E - F^{(n)} - P$ keine negativen Eigenwerte haben, sind auch ihre Mittelwerte (nach der Forderung 2. bzw. der Formel (II) des § 57) positiv:

$$(6) \quad 0 \leq \overline{P} \leq \overline{P} + \overline{(E - F^{(n)} - P)} = \overline{E - F^{(n)}}.$$

Nun ist aber nach (1) und § 57 (III)

$$(7) \quad \overline{E - F^{(n)}} = 1 - Sp(\Phi F^{(n)}) = 0.$$

Also wird nach (6) auch

$$(8) \quad \overline{P} = Sp(\Phi P) = (\Phi z, z) = 0$$

und nach § 10 (14) und § 57 (II)

$$(9) \quad \Phi z = 0.$$

Das gilt für jeden Vektor z , der (3) befriedigt. Wenn y ein beliebiger Vektor ist, so genügt

$$z = (E - F^{(n)}) y$$

der Gleichung (3), weil

$$F^{(n)} z = (F^{(n)} - F^{(n)2}) y = 0$$

ist.

Folglich gilt für jeden HILBERT-Vektor y

$$\Phi(E - F^{(n)}) y = 0;$$

daraus folgt

$$(10) \quad \Phi F^{(n)} = \Phi.$$

Geht man beiderseits zu den adjungierten Tensoren über, so folgt, weil Φ und $F^{(n)}$ beide hermitisch sind:

$$(11) \quad F^{(n)} \Phi = \Phi.$$

Damit ist die Hilfsformel (2) bewiesen. Man schließt aus ihr weiter:

$$(12) \quad \Phi = F^{(n)} \Phi F^{(n)}.$$

Wir nehmen jetzt an, A habe den *einfachen* Eigenwert a_n . Dann ist $F^{(n)} = c \times c$, also nach (1):

$$(13) \quad (\Phi c, c) = 1.$$

Das Element der Matrixdarstellung des Tensors (12) ist

$$\Phi_{nm} = \sum_{kl} c_n c_k^* \Phi_{k,l} c_l c_m^* = c_n c_m^* \sum_{kl} \Phi_{k,l} c_k^* c_l;$$

also hat man

$$\Phi = F^{(n)} (\Phi c, c),$$

und wegen (13) ergibt sich hieraus

$$(14) \quad \Phi = F^{(n)}.$$

Folglich kann man unter der Voraussetzung, daß man eine Größe A scharf gemessen und den Eigenwert a_n mit dem zugehörigen Einzeltensor $F^{(n)} = c \times c$ gefunden hat, den Mittelwert einer beliebigen Größe B nach der Formel

$$(15) \quad \bar{B} = Sp(\Phi B) = Sp(F^{(n)} B) = (B c, c)$$

berechnen, und damit ist in der Tat die in § 55 entwickelte Statistik des reinen Falles wiedergewonnen.

Man kann den reinen Fall unter allen Möglichkeiten von statistischen Gesamtheiten durch eine einfache Eigenschaft charakterisieren. Dazu dient die folgende Begriffsbildung.

Zwei statistische Gesamtheiten, bestimmt durch die Tensoren $\overset{(1)}{\Phi}$ und $\overset{(2)}{\Phi}$, kann man zu einer einzigen vereinigen mit dem Tensor

$$(16) \quad \Phi = \alpha \overset{(1)}{\Phi} + \beta \overset{(2)}{\Phi}, \quad \alpha, \beta > 0, \quad \alpha + \beta = 1.$$

die man eine *Mischung* der beiden ersten nennt. Unter allen Gesamtheiten kann man nun den reinen *Fall* kennzeichnen als ein solches Gemenge Φ , welches auf keine Weise durch Mischung

aus zwei Gesamtheiten erzeugt werden kann, wenn diese nicht beide untereinander (und folglich auch mit Φ) *übereinstimmen*. Wir wollen also beweisen:

Satz: *Eine Gesamtheit ist dann und nur dann ein reiner Fall, wenn sie auf keine andere Weise durch Mischung erzeugt werden kann als aus Gesamtheiten der gleichen Art.* Also in Formeln:

1. Wenn für ein gewisses Φ die Beziehung (16) nur mit $\Phi^{(1)} = \Phi^{(2)}$ möglich ist, so gibt es einen Einheitsvektor c derart, daß

$$\Phi = \Phi^{(1)} = \Phi^{(2)} = c \times c .$$

ist.

2. Wenn $\Phi = c \times c$ ist, so ist eine Darstellung (16) nur mit $\Phi^{(1)} = \Phi^{(2)} = \Phi$ möglich.

Um die erste Behauptung zu beweisen, wählen wir einen Einheitsvektor x irgendwie derart, daß $\Phi x \neq 0$ ist; dann gilt nach § 10, (14) die Ungleichung $(x, \Phi x) > 0$. Wir bilden nun den Tensor

$$(17) \quad \Phi^{(1)} = \frac{(\Phi x) \times (\Phi x)}{\alpha (x, \Phi x)} ,$$

wobei der reelle Faktor α so bestimmt werden kann, daß $Sp \Phi^{(1)} = 1$ und $0 < \alpha < 1$ ist. Denn aus der Forderung $Sp \Phi^{(1)} = 1$ folgt

$$\alpha = Sp \left\{ \frac{(\Phi x) \times (\Phi x)}{(x, \Phi x)} \right\} = \frac{(\Phi x, \Phi x)}{(x, \Phi x)} = \frac{|\Phi x|^2}{(x, \Phi x)} ;$$

denkt man sich nun Φ auf Hauptachsen transformiert mit den Eigenwerten w_k , so wird

$$(18) \quad \alpha = \frac{\sum_k w_k^2 |x_k|^2}{\sum_k w_k |x_k|^2} \leq 1 ,$$

weil die w_k positive echte Brüche oder Null sind, und es ist niemals $\alpha < 0$.

Definieren wir nun den Tensor $\Phi^{(2)}$ durch

$$(19) \quad \Phi^{(2)} = \frac{\Phi - \alpha \Phi^{(1)}}{1 - \alpha} ,$$

so gilt die Mischungsregel (16), und $\overset{(1)}{\Phi}$, $\overset{(2)}{\Phi}$ sind hermitische Tensoren der Spur 1, wie aus der Definition ohne weiteres hervorgeht. Sie sind überdies auch semidefinit; denn man hat

$$\overset{(1)}{(\Phi y, y)} = \frac{1}{\alpha} \frac{|(\Phi x, y)|^2}{(x, \Phi x)} \geq 0,$$

$$\overset{(2)}{(\Phi y, y)} = \frac{1}{1-\alpha} \frac{(\Phi x, x)(\Phi y, y) - |(y, \Phi x)|^2}{(x, \Phi x)} \geq 0;$$

letzteres auf Grund von § 10, (13). Also kann man $\overset{(1)}{\Phi}$ und $\overset{(2)}{\Phi}$ als Wahrscheinlichkeitstensoren ansehen, aus denen Φ gemäß (4) gemischt ist:

$$\Phi = \alpha \overset{(1)}{\Phi} + (1-\alpha) \overset{(2)}{\Phi}.$$

Nach Voraussetzung ist also $\overset{(1)}{\Phi} = \overset{(2)}{\Phi}$, mithin beide gleich Φ . Setzt man nun

$$(20) \quad c = \frac{\Phi x}{\sqrt{\alpha(x, \Phi x)}} = \frac{\Phi x}{|\Phi x|},$$

so liefert (17):

$$\Phi = \overset{(1)}{\Phi} = \overset{(2)}{\Phi} = c \times c,$$

wie unter 1. behauptet wurde.

Wir beweisen nun die Behauptung 2. Vorausgesetzt ist, daß sich $\Phi = c \times c$ gemäß (16) zerlegen läßt, wobei $\overset{(1)}{\Phi}$, $\overset{(2)}{\Phi}$ hermitische, semidefinite Tensoren der Spur 1 sind.

Wenn nun für einen Vektor x die Gleichung $\Phi x = 0$ gilt, so gilt auch $\overset{(1)}{\Phi} x = 0$ oder $\overset{(2)}{\Phi} x = 0$; denn sei etwa $\alpha > 0$, so folgt aus

$$(x, \Phi x) = \alpha (x, \overset{(1)}{\Phi} x) + \beta (x, \overset{(2)}{\Phi} x) \geq \alpha (x, \overset{(1)}{\Phi} x) \geq 0$$

für $\Phi x = 0$ auch $(x, \overset{(1)}{\Phi} x) = 0$ und daraus nach § 10, (14) weiter $\overset{(1)}{\Phi} x = 0$.

Ist aber $\alpha = 0$, so muß $\beta > 0$ sein, und dann folgt ebenso $\overset{(2)}{\Phi} x = 0$. Wir können annehmen, daß $\alpha > 0$ ist.

Ist nun x auf c normal, so ist $\Phi x = (x, c) c = 0$, also auch $\Phi x = 0$. Daher gilt für jeden Vektor y

$${}^{(1)}(\Phi x, y) = (x, \Phi y) = 0,$$

d. h. Φy ist zu jedem Vektor x normal, der zu c normal ist, mit hin müssen die Vektoren Φy und c parallel sein; insbesondere gilt

$$\Phi c = \gamma c,$$

wo γ eine Zahl ist. Jeden Vektor x kann man nun in eine Komponente parallel zu c und eine senkrecht zu c zerlegen, und zwar hat man wegen $(c, c) = 1$:

$$x = (x, c) c + x', \quad (x', c) = 0;$$

für diesen Vektor x' gilt wie für jeden auf c normalen: $\Phi x' = 0$. Also wird für einen beliebigen Vektor x

$$\Phi x = (x, c) \Phi c = (x, c) \gamma c = \gamma \Phi x,$$

woraus $\Phi = \gamma \Phi$ folgt. Wegen $S_p \Phi = S_p \Phi = 1$ ist $\gamma = 1$, und damit ist die Behauptung $\Phi = \Phi = \Phi$ bewiesen.

§ 59. Das allgemeine Gemenge als Mischung reiner Fälle. Es sei wieder A eine nichtentartete Größe mit den Eigenwerten a_r und den zugehörigen unzerlegbaren Einzeltensoren F_r . Wir wollen nun eine Mischung herstellen aus den verschiedenen reinen Gesamtheiten, welche den verschiedenen Eigenwerten a_r entsprechen. Ist $w_r \geq 0$ der zur Herstellung der Mischung aus dem reinen Fall $\Phi = F_r$ zu entnehmende Bruchteil ($\sum_r w_r = 1$), so ist die Mischung gegeben durch

$$(1) \quad \Phi = \sum_r w_r F_r.$$

Wir behaupten nun: Wenn man für A beliebige nichtentartete Größen zuläßt, so erhält man durch (1) das allgemeinste überhaupt mögliche Gemenge. Anders ausgedrückt: Für jedes Gemenge gibt es eine nichtentartete Größe A derart, daß es als Mischung der verschiedenen reinen Fälle bezüglich der Messung von A beschrieben werden kann.

Die Richtigkeit dieser Behauptung ergibt sich unmittelbar daraus, daß man den Wahrscheinlichkeitstensor Φ eines beliebigen Gemisches, da er hermitisch ist, in *unzerlegbare* Einzeltensoren F_r zerlegen kann, womit eine Darstellung der Form (1) erzielt wird; denn da Φ semidefinit ist und die Spur 1 besitzt, so sind die obigen Bedingungen $w_r \geq 0$, $\sum_r w_r = 1$ gewiß erfüllt. Für A ist dann eine beliebige hermitesche Größe

$$(2) \quad A = \sum_r a_r F_r$$

mit lauter verschiedenen Eigenwerten a_r zu wählen.

Die *unzerlegbaren* F_r in (1) sind aber durch Φ nur dann sämtlich eindeutig bestimmt, wenn Φ selber ein nichtentarteter Tensor ist. Es sei dagegen nunmehr Φ entartet, und

$$(3) \quad \Phi = \sum_n w^{(n)} F^{(n)}$$

die *invariante* Einzeltensorenzerlegung von Φ . Entsprechend den verschiedenen Möglichkeiten einer Zerlegung der zu *mehrfachen* Eigenwerten $w^{(n)}$ von Φ gehörigen $F^{(n)}$ in *unzerlegbare* F_r gemäß (1) besteht dann eine weitergehende Willkür in der Wahl von A . *Notwendig und hinreichend* dafür, daß eine nichtentartete Größe A für ein Gemenge Φ die eben beschriebene Rolle spielen kann, ist die *Vertauschbarkeit von A mit Φ* :

$$(4) \quad [\Phi, A] = 0.$$

Sei etwa g_n die Vielfachheit des Eigenwertes $w^{(n)}$ von Φ , und

$$(5) \quad F^{(n)} = \sum_{k=1}^{g_n} F_k^{(n)}$$

eine Zerlegung von $F^{(n)}$ in unzerlegbare F_r , und endlich sei

$$(2') \quad A = \sum_{n,k} a_k^{(n)} F_k^{(n)}$$

eine entsprechend gewählte nichtentartete Größe A . Nach (3) sind nun in dem Gemenge Φ diejenigen g_n reinen Fälle, die zu den $a_k^{(n)}$ mit bestimmten n gehören, in derselben Häufigkeit $w^{(n)}$ vertreten: Wir haben eine *gleichmäßige Mischung* in bezug auf alle zu festem n gehörigen reinen Fälle $F_k^{(n)}$. *Diese Gleichmäßigkeit der Mischung besteht aber nicht nur für eine spezielle Größe A , sondern für alle der Bedingung (4) genügenden Größen.*

Als Anwendung betrachten wir denjenigen $(2j+1)$ -fachen Term eines richtungsentarteten (räumlich frei drehbaren) Atoms (oder einer Molekel), welcher einem bestimmten einfachen Bestandteil $m_x^{(j)}, m_y^{(j)}, m_z^{(j)}$ der Drehimpulsmatrizen entspricht. (Vgl. Kap. IV, § 27). Ist eine Menge \mathfrak{S} von Atomen, die sich in diesem Zustand befinden, eine gleichmäßige Mischung reiner Fälle in bezug auf die $2j+1$ verschiedenen Eigenwerte der Drehimpulskomponente \mathfrak{M}_z , so ist sie auch eine gleichmäßige Mischung reiner Fälle in bezug auf die $2j+1$ Eigenwerte der Drehimpulskomponente $\mathfrak{M}_{z'}$, in einer beliebigen Richtung z' . Dies ist der präzise Ausdruck für die völlige statistische Symmetrie des Gemenges bezüglich aller Raumrichtungen.

Wir betrachten noch den Grenzfall einer gleichmäßigen Verteilung der Systeme von \mathfrak{S} über *alle* Eigenwerte $a_k^{(n)}$ von A . In diesem Falle werden alle $w_k^{(n)}$ einander gleich; um den Grenzübergang ausführen zu können, ohne daß jedes einzelne $w_k^{(n)}$ zu Null wird, müssen wir die Anzahl der Systeme in der Gesamtheit \mathfrak{S} als *unendlich* annehmen und *auf die Normierung* $\sum_n w^{(n)} g_n = \sum_{n,k} w_k^{(n)} = 1$ *verzichten*; statt von Wahrscheinlichkeiten sprechen wir dann lieber von „*relativen Häufigkeiten*“. Wir können die $w_k^{(n)}$ jetzt z. B. gleich 1 annehmen und erhalten aus (3)

$$(6) \quad \Phi = \sum_{nk} F_k^{(n)} = E.$$

Der statistische Tensor Φ wird also *völlig unabhängig* von der Wahl der Größe A : *Herrscht in einer Gesamtheit \mathfrak{S} gleichmäßige Verteilung in bezug auf die Eigenwerte irgend einer (nicht entarteten) Größe A , so besteht auch gleichmäßige Verteilung in bezug auf jede andere Größe.* Wir bezeichnen diesen Fall als *vollkommene Gleichverteilung*.

Besonders wichtig sind diejenigen Gemenge \mathfrak{S} , welche *verwirklicht* werden durch *klassische ideale Gase*, also Gase, deren einzelne Atome man (in erster Annäherung) als energetisch ungekoppelt betrachten kann. Betrachten wir zunächst ein Gas von solchen Atomen, die eine *nichtentartete Energie* besitzen (wozu insbesondere eine Beseitigung von Richtungsentartung etwa durch ein äußeres Magnetfeld vorauszusetzen ist), so gilt nach bekannten Sätzen für die *normale statistische Verteilung bei einer*

absoluten Temperatur T die BOLTZMANN-GIBBSsche Formel

$$(7) \quad w_r = e^{-\frac{W_r}{kT}}$$

für die relative Häufigkeit derjenigen Atome von \mathfrak{S} , die sich im r -ten stationären Zustand mit der Energie W_r befinden. Wir nehmen diese Formel hier einfach als gegeben an; erst später (Teil II) werden wir untersuchen, wie die zu ihrer Begründung führenden Überlegungen mit unserer quantenmechanisch-statistischen Theorie in inneren Zusammenhang gebracht werden können¹.

Der Tensor Φ ist gemäß (4) vertauschbar mit der Energie H des Atoms; und da die Eigenwerte von Φ durch (7) gegeben sind, erhalten wir in invarianter Schreibweise

$$(8) \quad \Phi = e^{-\frac{H}{kT}} = \sum_r w_r F_r,$$

wo F_r die Einzeltensoren zu den Eigenwerten W_r der Energie sind.

Tritt nun eine *Entartung* der Energie H ein, so werden gleichzeitig mit zwei Eigenwerten W_r, W_s von H auch die zugehörigen w_r, w_s einander gleich. Hat also die entartete Energie H die invariante Einzelgrößenzerlegung

$$(9) \quad H = \sum_n W^{(n)} F^{(n)},$$

so wird

$$(10) \quad \Phi = e^{-\frac{H}{kT}} = \sum_n e^{-\frac{W^{(n)}}{kT}} \cdot F^{(n)}.$$

Betrachten wir nun irgend eine *zeitlich konstante Größe* A , die also mit H vertauschbar ist:

$$(11) \quad \dot{A} = [H, A] = 0,$$

[vgl. Kap. III, § 21, (2)], und welche die Eigenschaft haben möge, daß zu einem bestimmten Eigenwert $W^{(n)}$ der Energie lauter verschiedene Eigenwerte $a_k^{(n)}$ von A gehören; dann hat nach dem oben Festgestellten jeder Eigenwert $a_k^{(n)}$ eine von k unabhängige Wahrscheinlichkeit

$$(12) \quad w_k^{(n)} = w^{(n)} = e^{-\frac{W^{(n)}}{kT}}.$$

¹ Vgl. NEUMANN, J. v.: Gött. Nachr. 1927, S. 273.

Denn eine mit H vertauschbare Größe A ist zugleich auch mit Φ vertauschbar.

Ein ideales Gas mit richtungsentarteten Atomen im thermodynamisch-statistischen Normalzustand zeigt also in der Tat die oben besprochene gleichmäßige statistische Verteilung bezüglich der Eigenwerte jeder Drehimpulskomponente.

Fragen wir noch nach der relativen Häufigkeit eines Eigenwertes $W^{(n)}$ der Energie selbst, so ergibt sich

$$(13) \quad Sp(\Phi F^{(n)}) = g_n w^{(n)},$$

wo das *statistische Gewicht* g_n gleich der *Vielfachheit des Eigenwertes* $W^{(n)}$ ist.

Die Formel (7) ist allerdings auch für ideale Gase als eine nur für hinreichend hohe Temperaturen gültige Annäherungsformel zu betrachten, welche auf die sog. „Gasentartung“ keine Rücksicht nimmt. Die exakte Formel (nach BOSE-EINSTEIN oder FERMI-DIRAC) ändert jedoch nichts an der Gültigkeit der soeben gezogenen Folgerungen, welche lediglich den einen qualitativen Umstand benutzen, daß in (7) die Häufigkeit w_r bei gegebener Temperatur T ausschließlich vom Energiewert W_r abhängt.

Wir gewinnen deshalb mit diesen Ergebnissen die allgemeine Rechtfertigung der in Kap. III, § 20 gegebenen Formeln für die Intensitäten der Lichtemissionen entarteter Atome. Betrachten wir zunächst wieder den Fall nichtentarteter Atome, so sind die Intensitätsfaktoren

$$(14) \quad \frac{64 \pi^4}{3 c^3} \nu_{rs}^4 |\mathfrak{P}(r, s)|^2; \quad W_s < W_r,$$

der vom r -ten Zustand ausgehenden Emissionen noch zu multiplizieren mit den relativen Häufigkeiten w_r der verschiedenen Ausgangszustände. Ist dagegen die Energie W_s im Endzustand ein mehrfacher Eigenwert: $W_s = W_l^{(m)}$, $l = 1, 2, \dots, g_m$, so sind die Intensitäten der g_m gleichen Frequenzen $(W_r - W_k^{(m)})/h$ zu addieren. Wird nun auch W_r mehrfach: $W_r = W_k^{(n)}$, $k = 1, 2, \dots, g_n$, so ist die Addition der Intensitätsfaktoren (14) *dann und nur dann* berechtigt, wenn auch die $w_k^{(n)}$ von k unabhängig sind, was nach den obigen Bemerkungen insbesondere bei thermisch-statistischer Verteilung der Atome gewiß der Fall ist. *Auf diesen praktisch fast ausschließlich vorkommenden Fall beziehen sich*

also die in Kap. III §20 angegebenen Intensitätsformeln. Sie sind aber *nicht* anwendbar z. B. auf die Atome eines Teilstrahls aus einem STERN-GERLACH-Versuch, den wir (zur Wiederherstellung der Entartung) in einen feldfreien Raum eintreten lassen, ohne jedoch den Atomen Zeit oder Gelegenheit zu geben, sich infolge ungeordneter Wechselwirkungen wieder thermisch-statistisch einzustellen. Vielmehr wäre in einem solchen Falle lediglich über l und nicht über k zu summieren.

Endlich weisen wir noch darauf hin, daß wir die *vollkommene Gleichverteilung* nach (9), (10) im *Grenzfall unendlich hoher Temperatur* verwirklicht finden.

§ 60. Quantenmechanik und Determinismus. Wir wollen nun das Verhältnis der statistischen Deutung unseres Formalismus, zu der wir nahezu zwangsläufig gelangt sind, zu den Fragen der *Kausalität* und des *Determinismus* in der Mikrophysik erörtern.

Nach den Vorstellungen der klassischen Theorie ist es möglich, den physikalischen Zustand eines abgeschlossenen Systems zu einer gewissen Anfangszeit durch ideal genaue Messungen derart zu bestimmen, daß für jede spätere Zeit das Verhalten des Systems mit absoluter Genauigkeit vorausberechnet werden kann. Aus praktischen Gründen sind zwar die Messungen, und folglich auch die Voraussagen, in Wirklichkeit stets nur mit endlicher Genauigkeit ausführbar, doch ist es grundsätzlich denkbar, diese Genauigkeit beliebig weit zu steigern.

Diesen *vollständigen Determinismus* der klassischen Theorie auch bei den Atomvorgängen nachzuweisen konnte jedoch von den Experimentatoren nicht versucht werden, weil eine vollständige Bestimmung des Anfangszustandes auch die Kenntnis inneratomarer Vorgänge in allen Einzelheiten einschließt. Betrachten wir z. B. ein bestimmtes einzelnes Radiumatom, so ist es nach dem gegenwärtigen Stande unserer experimentellen Erfahrung nicht möglich, durch eine Untersuchung des Atoms vorauszubestimmen, ob es innerhalb der nächsten Sekunde oder erst zu einer späteren Zeit radioaktiv zerfallen wird; man kann nur statistische Aussagen darüber machen. Ganz Entsprechendes gilt für einen Quantensprung mit Lichtemission und für alle anderen atomaren Elementarprozesse. Die experimentellen Methoden der Atomphysik haben sich von selbst, durch die Erfahrung geleitet,

ausschließlich auf statistische Fragestellungen eingestellt. Die Quantenmechanik, welche die systematische Theorie der so beobachteten Gesetzmäßigkeiten liefert, entspricht vollkommen dem gegenwärtigen Stande der Experimentalphysik, indem sie sich gleichfalls von vornherein auf statistische Fragen und Antworten beschränkt.

Daß die Quantenmechanik dem Prinzip des Determinismus nicht entspricht, ist aus den vorstehenden Entwicklungen schon deutlich hervorgegangen. Beim reinen Fall — der in der Quantenmechanik den Idealfall *möglichst* genauer Kenntnis des betrachteten System darstellt — beruht die Unmöglichkeit der vollständigen Determinierung darauf, daß es zu einer (nichtentarteten) Größe A , deren Messung den reinen Fall bestimmt, immer andere Größen gibt, die mit A nicht vertauschbar, also auch nicht zugleich mit A meßbar sind. Es kann deshalb nicht möglich sein, zu erreichen, daß die Reaktionsweise eines Systems gegenüber beliebigen künftigen Messungsprozessen eindeutig vor auszusehen ist.

Während die klassische Theorie behauptete: Aus der Kenntnis von Ort und Impuls jedes Massenpunktes des abgeschlossenen Systems zur Zeit $t = t_0$ kann man den Verlauf der Bewegung für alle Zeiten vollständig berechnen, so leugnet die Quantenmechanik nicht etwa ganz primitiv die Möglichkeit solcher Folgerungen, sondern behauptet in viel radikalerer Weise, daß die *Voraussetzung* einer derart vollständigen Ausmessung des Systems *unerfüllbar* ist¹. Der Zustand eines atomaren Systems zeigt eine charakteristische quantenphysikalische *Unbestimmtheit*, die sich in der nur statistischen Determinierung seiner Reaktionen auswirkt.

Die einschneidende *Beeinflussung*, welche ein System notwendigerweise durch die Ausführung eines *Messungsprozesses* erleidet, und mit welcher, wie schon erläutert, die Unmöglichkeit einer Messung mehrerer beliebiger Größen zugleich unmittelbar verknüpft ist, wird verdeutlicht durch folgende Überlegung. Es seien wieder A und B zwei beliebige (im allgemeinen nichtvertauschbare) nichtentartete Größen, und A' eine dritte solche Größe. Auch sei wieder $\Sigma^{(1)}$ das (genauer: *ein*) Hauptachsensystem von A ,

¹ Vgl. W. HEISENBERG: Z. Phys. Bd. 43, S. 172. 1927.

$\Sigma^{(2)}$ das von B , ferner Σ' das von A' . Die zwischen diesen vermittelnden unitären Transformationen seien:

$$\begin{aligned}\Sigma^{(1)} \rightarrow \Sigma' &: U' = (U'_{nm}), \\ \Sigma' \rightarrow \Sigma^{(2)} &: U'' = (U''_{mk}), \\ \Sigma^{(1)} \rightarrow \Sigma^{(2)} &: U = (U_{nk}).\end{aligned}$$

Dann besteht die Gleichung $U = U' U''$, oder

$$(1) \quad U_{nk} = \sum_m U'_{nm} U''_{mk}.$$

Wenn wir nun in einer Menge \mathfrak{S} von Atomen, an denen A gemessen und a_n gefunden war (so daß ein reiner Fall vorliegt), die Größe B messen, so finden wir ihren Eigenwert b_k verwirklicht bei einem Bruchteil $w_{nk} = |U_{nk}|^2$ aller Atome, wo also

$$(2) \quad w_{nk} = \left| \sum_m U'_{nm} U''_{mk} \right|^2.$$

Wenn wir aber vor der Messung von B eine Messung von A' an *jedem* Atom von \mathfrak{S} einschalten, so verwandelt sich dabei \mathfrak{S} in ein Gemenge, bestehend aus reinen Fällen mit definierten Werten von A' , wobei der Bruchteil w'_{nm} derjenigen Atome von \mathfrak{S} , bei welchen der m -te Eigenwert von A' vorliegt, gleich der Relativwahrscheinlichkeit $w'_{nm} = |U'_{nm}|^2$ ist. Wenn wir hernach an den Atomen dieses Gemenges die Größe B messen, so bekommen wir den Wert b_k mit der Wahrscheinlichkeit

$$(3) \quad \bar{w}_{nk} = \sum_m |U'_{nm} U''_{mk}|^2.$$

Diese Wahrscheinlichkeit \bar{w}_{nk} ist im allgemeinen durchaus verschieden von der obigen w_{nk} ; der Messungsprozeß von A' hat das System einschneidend verändert.

Man sieht hier deutlich, daß der Prozeß der Messung in der Quantenmechanik von ganz anderer Bedeutung ist als in der makroskopischen Physik: Es handelt sich hier nicht mehr um die einfache *Kenntnisnahme* von einem Tatbestand, der auch vor der Messung und unabhängig von der Messung vorhanden war, sondern um die *Schaffung* eines neuen Tatbestandes durch einen physikalischen *Eingriff* in das System. Ist zuvor an einem Atom der Wert a_n von A , gehörend zum (unzerlegbaren) Einzeltensor $F^{(n)}$, gemessen, und mißt man nun eine mit $F^{(n)}$ nicht vertauschbare Größe B , so wird ein definierter Wert b_k von B erst durch

den Messungsprozeß hervorgerufen. Erst die Messung zwingt das Atom, sich für einen bestimmten der Eigenwerte von B zu entscheiden; dabei verliert andererseits A seinen vorherigen definierten Wert. Man sieht also, daß das Durchdenken der durch die Quantenmechanik aufgedeckten Gesetze der Mikrophysik zwangsläufig zu radikalen Folgerungen führt: Nicht nur die klassisch-makroskopischen Vorstellungen von Raum, Zeit und Kausalität erweisen sich als unzulänglich im Gebiete der atomaren Reaktionen, sondern sogar die herkömmlichen Begriffe von Subjekt, Objekt und Realität verlieren allmählich ihre gewohnte Bedeutung¹.

Die beiden Formeln (2), (3) haben Veranlassung gegeben, die Zahlen U_{nk} als die *Amplituden* der Wahrscheinlichkeiten $w_{nk} = |U_{nk}|^2$ oder kurz als *Wahrscheinlichkeitsamplituden* zu bezeichnen. Ferner nennt man die Formel (1) das Gesetz von der *Interferenz der Wahrscheinlichkeiten*². Der Sinn dieser Bezeichnungen ist dieser: Man kann die klassischen, aus der Wellentheorie des Lichtes berechneten, stetig verteilten *Lichtintensitäten* in einem gewissen, später genauer zu formulierenden Sinne (Teil II) auffassen als Wahrscheinlichkeiten oder (relative Häufigkeiten) des unstetigen Auftreffens eines korpuskularen Lichtquants. Treffen sich nun an einem Orte mehrere *inkohärente Strahlen*, so addieren sich die Intensitäten, also die Wahrscheinlichkeiten. Treffen sich aber *kohärente Wellenzüge*, so sind es die *Wellenamplituden*, und nicht die aus ihnen quadratisch gebildeten Intensitäten oder Wahrscheinlichkeiten, welche sich addieren; es zeigt sich *Interferenz*.

Die Formel (2) für den *reinen Fall* verhält sich nun zur Formel (3) für das *Gemenge* ähnlich wie die Intensitätsformel bei *Kohärenz* zur Intensitätsformel bei *Inkohärenz* der sich treffenden Wellenzüge: In (2) werden die Amplituden $U'_{nm} U''_{mk}$ erst summiert (nach m) und dann absolut quadriert; in (3) dagegen erst absolut quadriert und dann summiert. Man könnte deshalb auch von der *kohärenten Statistik* eines reinen Falles gegenüber der *inkohärenten Statistik* eines Gemenges sprechen. Die späteren Entwicklungen (Teil II) werden deutlich machen, daß es sich

¹ Vgl. hierzu N. BOHR: Naturwissensch. Bd. 17, S. 483. 1929.

² PAULI jr. W.: Z. Phys. Bd. 41, S. 81. 1927 (s. insbes. die Fußnote S.83).
JORDAN, P.: Z. Phys. Bd. 40, S. 809. 1917; Bd. 41, S. 797. 1927.

hier nicht nur um äußerliche Ähnlichkeiten handelt, sondern daß die Interferenz des Lichtes der allgemeinen quantenmechanischen Statistik einzuordnen ist.

Machen wir uns aber zum Schluß noch einmal den fundamentalen Unterschied von kohärenter und inkohärenter Wahrscheinlichkeit (von reinem Fall und Gemenge) an einem konkreten Beispiel deutlich. Wenn Quecksilberdampf, dessen Atome zunächst im Grundzustand $1\ ^1S_0$ sein mögen, durchstrahlt wird mit Licht der Resonanzfrequenz $1\ ^1S_0 - 2\ ^3P_1$, so wird ein gewisser Bruchteil der Atome angeregt: Wir bekommen ein *Gemenge* von Atomen in den Zuständen $1\ ^1S_0$ und $2\ ^2P_1$. Man weiß also von einem beliebig herausgegriffenen Atom nicht ohne weiteres, ob es in $1\ ^1S_0$ oder in $2\ ^2P_1$ ist, sondern muß das erst feststellen. Aber doch besitzt jedes einzelne Atom einen *definierten* Energiewert, zu dessen Feststellung also kein verändernder Eingriff nötig ist. Bei genügender Hg-Dampfdichte ergibt sich aber an der Oberfläche eine *selektive Reflexion der Resonanzlinie*¹; die in der Oberfläche enthaltenen Atome senden also eine *kohärente Resonanzstrahlung* aus; und das bedeutet, daß die Energie des einzelnen Atoms *nicht mehr definiert ist*: An ihrer Stelle hat jetzt eine andere, mit H nicht vertauschbare Größe (die mit der *Schwingungsphase* zusammenhängt) einen definierten Wert angenommen. Ein einzelnes Atom dieser reflektierenden Oberfläche entspricht also einem *reinen Fall mit statistisch unbestimmter Energie*; was offenbar etwas gänzlich anderes ist als ein aus dem zuvor besprochenen Gemenge entnommenes Atom, das einen *definierten Energiewert* besitzt, den wir nur noch *nicht kennen*.

§ 61. Übergangswahrscheinlichkeit bei äußerer Einwirkung.

Eine der wichtigsten Anwendungen der im voranstehenden entwickelten Begriffe besteht in der genauen Definition und Berechnung der *Übergangswahrscheinlichkeiten*. In der älteren Quantentheorie gab es kein exaktes Maß für die Wahrscheinlichkeit eines „Quantensprunges“, der von irgendwelchen äußeren Einwirkungen induziert ist; BOHR'S Korrespondenzprinzip lieferte nur approximative Bestimmungen solcher Wahrscheinlichkeiten, die nur in Sonderfällen zu quantitativen Aussagen verschärft

¹ Betreffs der experimentellen Tatsachen vgl. etwa P. PRINGSHEIM: Fluoreszenz und Phosphoreszenz (diese Sammlung Nr. VI).

werden konnten. In der Quantenmechanik läßt sich aber der Begriff der Übergangswahrscheinlichkeit streng auf den der relativen Wahrscheinlichkeit (§ 54) zurückführen.

Wir rechnen hier wieder mit einem Zeitparameter t . Es ist schon in § 22 darauf hingewiesen worden, daß die zur Energie eines abgeschlossenen Systems kanonisch konjugierte Zeit *nicht* durch einen Zahlparameter dargestellt werden kann, sondern selbst eine mit den übrigen meßbaren Größen an 'diesem System im allgemeinen *nicht vertauschbare* Größe ist. Deshalb kann man, mit Bezug auf die so definierte Zeit, im allgemeinen auch nicht davon sprechen, daß irgendeine Größe $A(p, q)$ *in einem bestimmten Zeitpunkt gemessen* wird; denn das würde bedeuten, daß *zugleich* einerseits der Wert von A und andererseits der Wert einer mit A (im allgemeinen) nicht vertauschbaren Größe, nämlich der Zeit, festgelegt würde. Das Verhältnis der Messung einer Größe A zur Zeitmessung bedarf also genauerer Untersuchung; in den allgemeinen Erörterungen der §§ 54 bis 60 haben wir absichtlich diese Frage zunächst *unberührt* gelassen. Auch hier wollen wir diese Frage noch nicht ausführlicher untersuchen, sondern uns — den im Folgenden zu machenden Anwendungen entsprechend — beschränken auf solche Probleme, in welchen man praktisch mit der Darstellung der Zeit t als Zahlparameter auskommt. Wir können, wie in § 22 kurz erläutert wurde, diese Zeit t als durch ein mit dem betrachteten System gar nicht oder ganz schwach gekoppeltes anderes System (eine „Uhr“) definiert auffassen.

Da dieser Zeitparameter t mit jeder zu betrachtenden Größe $A(p, q)$ vertauschbar ist, können wir jetzt von einer Messung von A zu einem bestimmten Zeitpunkt $t = t_0$ sprechen. Für die Anwendung der Theorie der Relativwahrscheinlichkeiten aus § 54 ist dabei zu beachten: Die Größen p_k, q_k sind jetzt als *Funktionen der Zeit* t aufzufassen, die den Bewegungsgleichungen

$$\dot{q}_k(t) = [H, q_k], \quad \dot{p}_k = [H, p_k]; \quad H = H(p(t), q(t); t)$$

genügen, und gemäß § 22 (10), (16) durch

$$(1) \quad q_k(t) = U^{-1}(t) q_k(0) U(t), \quad p_k(t) = U^{-1}(t) p_k(0) U(t)$$

aus den Größen $q_k(0), p_k(0)$ für $t = 0$ zu erhalten sind, wobei sich $U(t)$ durch

$$(2) \quad \dot{\ast} U + H(p(0), q(0); t) U = 0, \quad U(0) = E$$

bestimmt. Folglich wird jede Größe $A(p, q)$ jetzt eine Funktion $A(t)$ des Zeitparameters. Wenn sie *nicht* auch noch *explizit* von t abhängt, so ist:

$$(1') \quad A(t) = A(p(t), q(t)) = U^{-1}(t) \cdot A(p(0), q(0)) \cdot U(t).$$

Es sind also, wenn wir die *Terminologie* von § 54 anwenden wollen, $A(t_1)$ und $A(t_2)$, mit $t_1 \neq t_2$, (im allgemeinen) zwei *verschiedene Tensoren* oder zwei *verschiedene physikalische Größen* mit den gleichen Eigenwerten, aber verschiedenen Hauptachsensystemen; eine Messung von $A(p, q)$ zur Zeit t wäre nach der Terminologie von § 54 einfach eine *Messung der Größe $A(t)$* .

Es sei $A(0)$ (und folglich auch $A(t)$) eine nichtentartete Größe. Bezeichnen wir das Koordinatensystem des Hilbertraumes, in dem der Tensor $A(0)$ bzw. der Tensor $A(t)$ die Diagonalform hat bzw. mit $\Sigma^{(0)}$, $\Sigma^{(t)}$, und verstehen wir unter $\overset{(0)}{A}(0)$ bzw. $\overset{(t)}{A}(0)$ die Matrix des Tensors $A(0)$ im Koordinatensystem $\Sigma^{(0)}$ bzw. $\Sigma^{(t)}$, so sind $\overset{(0)}{A}(0)$ und $\overset{(t)}{A}(0)$ identische Diagonalmatrizen, also

$$\overset{(0)}{A}(t) = U^{-1}(t) \overset{(0)}{A}(0) U(t) = U^{-1}(t) \overset{(t)}{A}(0) U(t)$$

und folglich

$$(3) \quad \overset{(t)}{A}(t) = U(t) \overset{(0)}{A}(0) U^{-1}(t);$$

hierin ist $U(t)$ die Lösung von (2), wenn für $H(p(0), q(0); t)$ die Matrix $\overset{(0)}{H}(p(0), q(0); t)$ dieses Tensors *im Koordinatensystem* $\Sigma^{(0)}$ eingesetzt wird, d. h. die Matrix $H(\overset{\circ}{p}, \overset{\circ}{q}, t)$, wobei $\overset{\circ}{p}_k, \overset{\circ}{q}_k$ Matrizen sind, welche den kanonischen Vertauschungsregeln und der Bedingung $A(\overset{\circ}{p}, \overset{\circ}{q}) = \text{Diagonalmatrix}$ genügen. Der Vergleich von (3) mit § 54 (6), (7) ergibt jetzt:

Wenn bekannt ist, daß $A(p, q)$ zur Zeit $t = 0$ den scharfen Wert a_n besitzt, so ist die Wahrscheinlichkeit, zur Zeit t für $A(p, q)$ den Wert a_k zu finden, gegeben durch:

$$(4) \quad \Psi_{nk} = |U_{nk}^{-1}(t)|^2 = |U_{kn}(t)|^2.$$

Dieses Ergebnis wollen wir nun speziell zur Berechnung der *Übergangswahrscheinlichkeiten* anwenden.

Ein abgeschlossenes mechanisches System mit der Energiefunktion $H_0(p, q)$ werde dem Einfluß einer äußeren Störung aus-

gesetzt, so daß die Gesamtenergie durch

$$(5) \quad H(p, q, t) = H_0(p, q) + H'(p, q; t)$$

gegeben ist. Fragt man nach der Wahrscheinlichkeit dafür, daß zur Zeit t für die *ungestörte Energie* H_0 der Wert W_k^0 gefunden wird, wenn zur Zeit $t = 0$ der Wert W_n^0 gemessen wurde, so können wir unser soeben erhaltenes Ergebnis anwenden, da H_0 nach Voraussetzung nicht explizit von t abhängt. Mit $A = H_0$ finden wir:

Die „Übergangswahrscheinlichkeit“ $n \rightarrow k$ in der Zeit t ist gleich:

$$(6) \quad \Psi_{nk}(t) = |U_{kn}(t)|^2,$$

wobei die Matrix U die Lösung von:

$$\kappa \dot{U} + H(\dot{p}, \dot{q}; t) U = 0, \quad U(0) = E$$

ist; die *Matrizen* \dot{p}_k, \dot{q}_k müssen hier neben den kanonischen Vertauschungsregeln die Bedingung $H_0(\dot{p}, \dot{q}) = \text{Diagonalmatrix}$ erfüllen.

Diese Lösung U wurde schon in § 45 mathematisch untersucht. Für den wichtigen Spezialfall, daß auch die Störungsfunktion H' die Zeit nicht explizit enthält, findet man¹ aus § 45, (10') in erster Näherung:

$$(7) \quad \Psi_{nk}^{(1)} = |U_{kn}^{(1)}|^2 = \frac{|H'_{kn}|^2}{\hbar^2 \nu_{kn}^2} \cdot 4 \sin^2 \pi \nu_{kn} t; \quad k \neq n.$$

Auch die zweite Näherung wird später (Kap. VII, § 75) Verwendung finden.

Einige besondere Worte erfordern die Verhältnisse bei *Entartung* der Energie H_0 . Nach dem Verfahren des § 59 denken wir uns zur Charakterisierung der reinen Fälle, aus denen sich das entartete Gemisch zusammensetzt, eine kleine Störung angebracht (dort Messung der Größe A !), die die Entartung aufhebt. Die Zustände, zwischen denen Übergänge stattfinden, seien durch den Doppelindex (r, ρ) numeriert und in bezug auf ρ entartet. Für die einzelnen Übergänge $(r, \rho) \rightarrow (s, \sigma)$, die bei wirklicher Auf-

¹ Man erhält offenbar formal dasselbe Resultat, wenn man $H'(\dot{p}, \dot{q}; \tau)$ als eine *zeitabhängige* Matrix betrachtet, welche aber nur im Intervall $0 < \tau < t$ von Null verschieden und dort konstant ist.

hebung der Entartung stattfinden würden, hat man die Wahrscheinlichkeiten $|U_{\varrho, \sigma}^{(r, s)}|^2$. Die Wahrscheinlichkeit des beobachtbaren Überganges $r \rightarrow s$ ist die Summe all dieser Wahrscheinlichkeiten für alle möglichen Werte von ϱ und σ :

$$(6') \quad \Psi^{(r, s)} = \sum_{\varrho, \sigma} |U_{\varrho, \sigma}^{(r, s)}|^2 = S p (U^{(r, s)} U^{\dagger(s, r)}).$$

Wir werden insbesondere später sehen, daß die Intensität einer Lichtemission proportional ist einer gewissen Übergangswahrscheinlichkeit, für welche die Matrix U proportional dem elektrischen Moment ist; daher ist die Lichtintensität proportional

$$S p (\mathfrak{P}^{(rs)} \mathfrak{P}^{(sr)}) = \|\mathfrak{P}^{(rs)}\|^2.$$

Hierdurch findet der bisher nur durch Invarianzbetrachtungen plausibel gemachte Ansatz von § 20 eine strenge Begründung.

Bis jetzt haben wir den Fall betrachtet, daß *die zu messende Größe nicht explizit von der Zeit abhängt*. Im folgenden Paragraphen (Adiabatensatz) wird aber die zu messende Größe die *totale* Energie $H(p, q; t)$ selber sein, wobei es gerade darauf ankommen wird, daß sie explizit von t abhängt.

Sei $H(\overset{\circ}{p}, \overset{\circ}{q}; 0) = W(0)$ Diagonalmatrix, und $U(t)$ die Lösung von

$$(8) \quad \varkappa \dot{U} + H(\overset{\circ}{p}, \overset{\circ}{q}; t) U = 0, \quad U(0) = E.$$

Sei weiter $V(t)$ eine unitäre Matrix mit der Eigenschaft:

$$(9) \quad V^{-1} H(\overset{\circ}{p}, \overset{\circ}{q}; t) V = W(t) = \text{Diagonalmatrix.}$$

Bezeichnen wir für den Augenblick allgemein mit $A^{(0; 0)}$ bzw. $A^{(t; 0)}$

bzw. A die Matrix eines Tensors A im Hauptachsensystem von $H(p(0), q(0); 0)$ bzw. $H(p(t), q(t); 0)$ bzw. $H(p(t), q(t); t)$, so ist

$$(10) \quad \overset{(0; 0)}{H}(p(t), q(t); t) = U^{-1}(t) \overset{(0; 0)}{H}(p(0), q(0); 0) U(t)$$

$$(10') \quad \overset{(t; t)}{H}(p(t), q(t); t) = W(t) = V^{-1} H(\overset{\circ}{p}, \overset{\circ}{q}; t) V \\ = V^{-1} \overset{(0; 0)}{H}(p(0), q(0); 0) V.$$

Wenn man die Matrix

$$(11) \quad X = V^\dagger U$$

eingeführt, so wird also:

$$(12) \quad H^{(t;t)}(p(t), q(t); t) = X(t) H^{(0;0)}(p(t), q(t); t) X^{-1}(t).$$

Der Vergleich von (12) mit § 54 (6), (7) ergibt: Die Wahrscheinlichkeit, daß das System, wenn es zur Zeit $t = 0$ die Energie $W_n(0)$ hatte, zur Zeit t die Energie $W_k(t)$ besitzt, ist gegeben durch:

$$(13) \quad \Psi_{nk} = |X_{nk}^{-1}(t)|^2 = |X_{kn}(t)|^2.$$

Es handelt sich nun darum, eine Differentialgleichung für X aufzustellen. Aus (11) folgt

$$(14) \quad \dot{X} = \dot{V}^\dagger U + V^\dagger \dot{U},$$

und wenn man hier \dot{U} aus (8) einsetzt:

$$(15) \quad \begin{aligned} \varkappa \dot{X} &= -V^\dagger H(\hat{p}, \hat{q}, t) U + \varkappa \dot{V}^\dagger U \\ &= -(V^\dagger H V - \varkappa \dot{V}^\dagger V) X. \end{aligned}$$

Setzen wir zur Abkürzung

$$(16) \quad Q = \varkappa \dot{V}^\dagger V$$

und berücksichtigen (9), so bekommen wir die gesuchte Gleichung

$$(17) \quad \varkappa \dot{X} = (Q - W(t)) X.$$

Sind die instantanen Eigenwerte $W(t)$ und die zugehörige Transformation V bekannt, so kennt man auch Q aus (16) und kann dann X aus der Differentialgleichung (17) berechnen. Dabei ist aber zu bedenken, daß die unitäre Matrix V durch das Hauptachsenproblem (9) nicht eindeutig festgelegt ist, sondern nur bis auf eine Phasenmatrix (wobei wir fehlende Entartung voraussetzen); da aber in (9) die Zeit t als Parameter vorkommt, sind die Phasen im allgemeinen Funktionen der Zeit $\varphi_n(t)$. Ist V_0 eine bestimmte Lösung von (9), so ist die allgemeine Lösung

$$(18) \quad V = V_0 e^{-i\varphi}, \quad \varphi = (\varphi_n(t) \delta_{nm});$$

dann wird

$$(19) \quad Q = e^{i\varphi} \cdot Q_0 e^{-i\varphi},$$

wo

$$(20) \quad Q_0 = \varkappa \dot{V}_0^\dagger V_0 + i \varkappa \dot{\varphi}$$

ist. Man kann nun die willkürlichen Phasen $\varphi_n(t)$ dadurch festlegen, daß man fordert, die Diagonalelemente von $P = Q - W$ sollen verschwinden; das gibt nach (20) die Bedingung

$$(21) \quad (Q_0)_{nn} - W_n = \varkappa (\dot{V}_0^\dagger V_0)_{nn} - W_n + i \varkappa \dot{\varphi}_n = 0.$$

Hier sind alle Glieder reell; man bekommt also die reelle Lösung

$$(22) \quad \varphi_n(t) = \frac{2\pi}{h} \int \{W_n - \varkappa (\dot{V}_0^\dagger V_0)_{nn}\} dt,$$

die bis auf eine additive Konstante eindeutig bestimmt ist. Dann lautet die Differentialgleichung (17) einfach

$$(23) \quad \varkappa \dot{X} = P X,$$

wo

$$(24) \quad \begin{cases} P_{nn} = 0, \\ P_{nm} = Q_{nm} = (Q_0)_{nm} e^{i[\varphi_n(t) - \varphi_m(t)]}, \quad n \neq m. \end{cases}$$

Für diese Elemente läßt sich noch eine andere Darstellung angeben. Differenziert man nämlich (9) nach dem Parameter t , so kommt

$$\dot{W} = \dot{V}^{-1} H V + V^{-1} H \dot{V} + V^{-1} \dot{H} V;$$

nun ist nach (9)

$$H V = V W, \quad V^{-1} H = W V^{-1},$$

ferner (wegen $V^\dagger V = 1$)

$$\dot{V}^\dagger V = -V^\dagger \dot{V};$$

also wird nach (16)

$$(25) \quad \dot{W} = -\frac{1}{\varkappa} (W Q - Q W) + R,$$

wo

$$(26) \quad R = V^{-1} \dot{H} V$$

gesetzt ist. Für die außerhalb der Diagonale stehenden Elemente von (25) folgt

$$(W_n - W_m) Q_{nm} = \varkappa R_{nm}, \quad (n \neq m),$$

also

$$(27) \quad Q_{nm} = \frac{\varkappa R_{nm}}{W_n - W_m}, \quad (n \neq m).$$

woraus nach (18), (19)

$$(28) \quad (Q_0)_{nm} = \frac{\kappa (R_0)_{nm}}{W_n - W_m}, \quad R_0 = V_0^{-1} \dot{H} V_0.$$

Aus der ursprünglichen Definition von Q durch (16) geht hervor, daß Q_{nm} auch dann endlich bleibt, wenn für einen speziellen Wert von t die Differenz $W_n(t) - W_m(t)$ verschwindet („vorübergehende Entartung“).

§ 62. Der Adiabatenatz¹. Es ist nun zu zeigen, daß für die durch § 61, (13) definierten Übergangswahrscheinlichkeiten der Adiabatenatz gilt.

In der alten Theorie besagte dieser von EHRENFEST aufgestellte Satz (s. Bd. 1, § 10, § 16), daß die Wirkungsvariablen J gegenüber unendlich langsamen (adiabatischen) Änderungen von Parametern des mechanischen Systems invariant sind. Daher war es möglich, Quantenzahlen n durch $J = \hbar n$ einzuführen, ohne im Grenzfalle unendlich langsamer Änderungen mit der klassischen Mechanik im Widerspruch zu kommen (EHRENFESTS *Adiabatenhypothese*); ein und dieselbe ganze Zahl n beschreibt den Zustand vor und nach dem adiabatischen Prozeß, obwohl die Energie eine endliche Änderung erfahren haben mag.

Ein analoger Satz gilt nun auch in der neuen Theorie. An Stelle der Quantenzahlen n hat man hier die Indizes n der Energieniveaus. Der Adiabatenatz behauptet: Wenn das System sich anfangs in einem Zustand mit bestimmtem Index n befand, so ist bei einem unendlich langsamen (adiabatischen) Prozesse die Wahrscheinlichkeit des Übergangs in einen Zustand mit einer anderen Nummer unendlich klein, trotzdem das Energieniveau nach der Änderung sich von seinem Anfangswert um einen endlichen Betrag unterscheiden kann.

Um die Behauptung mathematisch zu formulieren, betrachten wir ein sehr großes Zeitintervall T und sehen in der Energiefunktion und allen aus ihr abgeleiteten Funktionen als unabhängige Variable statt t die dimensionslose Größe $\varkappa = \frac{t}{T}$ an. *Wir setzen voraus,*

¹ BORN, M.: Z. Phys. Bd. 40, S. 167. 1927. FERMI E. und F. PERSICO: Lincei Rend. (6) Bd. 4, S. 452. 1926. BORN, M. und V. FOCK: Z. Phys. Bd. 51, S. 165. 1928.

daß die Ableitungen von $H(p, q, s)$ nach s endlich bleiben, wenn T beliebig wächst. Dasselbe gilt dann von der Lösung $V(s)$ der Gleichung § 61 (9). Ersetzt man nun überall in den Formeln des § 61 t durch $s = \frac{t}{T}$, so verwandelt sich nach § 61 (16) Q in $\frac{1}{T}Q$, nach § 61, (22) $\varphi_n(t)$ in $T\varphi_n(s)$ und \dot{X} in $\frac{1}{T} \frac{dX}{ds}$. Daher wird die Differentialgleichung § 61, (23)

$$(1) \quad \kappa \frac{dX}{ds} = PX$$

mit

$$P_{nn} = 0$$

$$(2) \quad P_{nm} = Q_{nm} = (Q_0)_{nm} e^{iT[\varphi_n(s) - \varphi_m(s)]}.$$

Die P_{nm} werden also bei großem T schnell oszillierende Funktionen von s .

Die Behauptung des Adiabatensatzes lautet nun:

$$(3) \quad \lim_{T \rightarrow \infty} \Psi_{mn} = \lim_{T \rightarrow \infty} |X_{nm}|^2 = 0, \quad n \neq m,$$

oder einfach

$$(4) \quad \lim_{T \rightarrow \infty} X_{nm} = 0, \quad n \neq m.$$

Für den vorliegenden Zweck brauchen wir eine strengere Behandlung der Gleichung (1) als die im vorigen Paragraphen durchgeführte.

Die Differentialgleichung (1) läßt sich genau so wie die allgemeine quantenmechanische Grundgleichung § 22, (16) behandeln. Wir ersetzen die unitäre Matrix X durch einen Vektor x und haben

$$(5) \quad \kappa \frac{dx}{ds} = Px.$$

Dann liefert jedes vollständige, orthogonale Lösungssystem dieser Gleichung (5) auch eine Lösung von (1); sind nämlich $x_1(s)$, $x_2(s)$, ... die Vektoren, die dieses Lösungssystem bilden, so hängen diese mit ihren Anfangswerten \hat{x}_k durch eine unitäre Transformation

$$(6) \quad x_k(s) = X(s) \hat{x}_k$$

zusammen; dann ist $X(s)$ die Lösung von (1), die für $s = 0$ in die Einheitsmatrix übergeht.

Das System gewöhnlicher Differentialgleichungen (5) läßt sich leicht durch sukzessive Approximationen lösen. Setzt man rechter Hand zunächst einen Vektor x_k ein, so folgt aus (5) als erste Näherung

$$x_k^{(1)} = \frac{1}{\kappa} \int_0^s P(\sigma) \hat{x}_k d\sigma,$$

und allgemein, wenn $x_k^{(l-1)}$ gefunden ist, erhält man

$$x_k^{(l)} = \frac{1}{\kappa} \int_0^s P(\sigma) x_k^{(l-1)} d\sigma.$$

Die Lösung von (5) ist also

$$(7) \quad x_k(s) = \hat{x}_k + \sum_{l=1}^{\infty} \kappa^{-l} \int_0^s ds_l \int_0^{s_l} ds_{l-1} \cdots \int_0^{s_2} ds_1 P(s_l) \cdots P(s_1) \hat{x}_k,$$

und damit nach (6) die Lösung von (1)

$$(8) \quad X(s) = E + \sum_{l=1}^{\infty} \kappa^{-l} \int_0^s ds_l \int_0^{s_l} ds_{l-1} \cdots \int_0^{s_2} ds_1 P(s_l) \cdots P(s_1).$$

Um die Konvergenz des Verfahrens zu sichern, muß man eine Voraussetzung über die Beschaffenheit der Matrix $P(s)$ machen: $P(s)$ soll für alle s absolut beschränkt¹ sein und eine konstante beschränkte Matrix M als Majorante zulassen:

$$(9) \quad |P_{nm}(s)| = |Q_{nm}(s)| \leq M_{nm}; \quad M = (M_{nm}) \text{ beschränkt.}$$

Das majorante Gleichungssystem

$$(10) \quad \frac{db}{ds} = Mb$$

hat die Lösungen

$$(11) \quad b_k = e^{Ms} \hat{b}_k,$$

¹ Eine Matrix A heißt beschränkt, wenn für alle Einheitsvektoren x, y die Bilinearform (Ax, y) konvergiert und ihr absoluter Betrag unterhalb einer von der Wahl von x und y unabhängigen Schranke bleibt. Eine Matrix A heißt absolut beschränkt, wenn die Matrix $(|A_{nm}|)$ beschränkt ist.

wo die $\overset{\circ}{b}_k$ irgendein vollständiges System orthogonaler Vektoren bilden; und man kann zeigen, daß die Potenzreihe

$$(12) \quad e^{Ms} = \sum_{l=0}^{\infty} \frac{s^l}{l!} M^l$$

beständig konvergiert¹.

Indem man in (8) P durch M ersetzt, sieht man, daß der absolute Betrag jedes Gliedes der Reihe (8) nicht größer ist als das entsprechende Glied der Reihe (12). Daher ist die Bedingung (9) für die absolute Konvergenz der Reihe (8) hinreichend. Auf die Frage, in welchen Einzelfällen man über die absolute Beschränktheit der Matrix P etwas aussagen kann, kommen wir nachher zurück.

Zum Beweise des Adiabatensatzes sind noch einige weitere *Voraussetzungen* nötig; wir stellen sie hier einschließlich der schon angegebenen zusammen:

Im Intervalle $0 \leq s \leq \sigma$ soll

$$1. \quad |(Q_0)_{nm}(s)| = |P_{nm}(s)| \leq M_{nm} \quad (n \neq m)$$

sein;

2. die Funktionen (Frequenzen)

$$\frac{d\varphi_n}{ds} - \frac{d\varphi_m}{ds} = 2\pi\nu_{nm}(s),$$

gebildet aus den in (2) auftretenden Phasen der Matrixelemente P_{nm} , sollen im Intervall höchstens N_1 Nullstellen von höchstens r -ter Ordnung („vorübergehende“ Entartungszustände) besitzen, und in der Nähe jeder Nullstelle s_0 soll die Abschätzung

$$|2\pi\nu_{nm}(s)| < \frac{1}{|s - s_0|^r} A$$

gelten;

3. der reelle und imaginäre Teil der Funktion

$$\frac{(Q_0)_{nm}(s)}{\nu_{nm}(s)}$$

sollen stückweise monoton sein; die Anzahl der Strecken, wo sie monoton sind, sei höchstens gleich N_2 .

¹ HART, W. L.: Amer. Journ. Bd. 39, S. 407. 1917.

Unter diesen Voraussetzungen gilt ein *Hilfssatz*, dessen Beweis in Anhang 3 mitgeteilt wird; er besagt, daß die Ungleichung

$$(13) \quad \left| \int_0^{\sigma} P_{nm}(s) ds \right| \\ = \left| \int_0^{\sigma} (Q_0)_{nm}(s) e^{i T(q_n - q_m)} ds \right| < 4 M_{nm} (N_1 + N_2) \sqrt{\frac{4A}{T}}^{r+1}$$

gilt.

Jetzt läßt sich der Adiabatenatz leicht beweisen. Im l -ten Gliede der Reihe (8) hat man für das Element X_{nm} das Integral

$$\sum_k \int_0^s ds_l \int_0^{s_l} ds_{l-1} \cdots \int_0^{s_3} ds_2 (P(s_l) \cdots P(s_2))_{nk} \int_0^{s_2} ds_1 P_{km}(s_1).$$

Hier schätzen wir das innerste Integral mit Hilfe der Formel (13) ab (wobei zu beachten ist, daß $P_{nn} = 0$ ist); die übrigen Integrationen führen wir aus, indem wir P durch die Majorante M ersetzen. Dann wird der Betrag der Summe

$$\leq \sum_k (M^{l-1})_{nk} \frac{s^{l-1}}{(l-1)!} 4 M_{km} (N_1 + N_2) \sqrt{\frac{4A}{T}}^{r+1}.$$

Also erhält man aus (8)

$$|X_{nm} - \delta_{nm}| \leq 4 (N_1 + N_2) \sqrt{\frac{4A}{T}} \sum_{l=1}^{\infty} \frac{s^{l-1}}{(l-1)!} (M^l)_{nm},$$

oder nach (12)

$$(14) \quad |X_{nm} - \delta_{nm}| \leq 4 (N_1 + N_2) \sqrt{\frac{4A}{T}}^{r+1} (M e^{Ms})_{nm}.$$

Diese Gleichung zeigt, daß $X_{nm} - \delta_{nm}$ mit wachsendem T gegen Null konvergiert, und zwar hat man die Abschätzung¹

$$(15) \quad X_{nm} = \delta_{nm} + O\left(T^{-\frac{1}{r+1}}\right).$$

Für die *Übergangswahrscheinlichkeit von einem Niveau n zu einem anderen m* hat man nach § 61, (13):

$$(16) \quad \Psi_{nm} = |X_{mn}|^2 = O\left(T^{-\frac{2}{r+1}}\right), \quad (n \neq m).$$

¹ Die Schreibweise $x = O(\alpha)$ bedeutet: $|x| \leq c\alpha$, wo c eine von α unabhängige Zahl ist, und wird gelesen: x ist von der Größenordnung α .

Wenn z. B. keine Frequenz $\nu_{nm}(s)$ während der adiabatischen Änderung verschwindet, so ist $r = 0$ und es wird $\Psi_{nm} = O(T^{-2})$. Für die Wahrscheinlichkeit Ψ_{nn} , daß das System in dem ursprünglichen Zustande n bleibt, erhält man wegen der Relationen $\sum_m |X_{nm}|^2 = 1$

$$(17) \quad \Psi_{nn} = 1 - \sum'_m \Psi_{nm} = 1 - O\left(T^{-\frac{2}{r+1}}\right).$$

Diese Wahrscheinlichkeit weicht also von der Einheit um eine Größe derselben Ordnung ab wie die Ψ_{nm} für $n \neq m$.

Damit ist der Adiabatensatz unter den angegebenen Voraussetzungen bewiesen.

Der Begriff der Übergangswahrscheinlichkeit bezieht sich auf eine solche Anordnung, bei der der Anfangszustand ein reiner Fall mit vorgegebenem Energiewert ist.

Es liege nun irgend ein anderer reiner Fall vor, dargestellt durch den Vektor x im HILBERTSchen Raume; die Komponenten von x in dem Koordinatensystem, für das zur Zeit $t = 0$ die Energiefunktion eine Diagonalmatrix ist, seien x_1, x_2, \dots . Dann sind $|x_1|^2, |x_2|^2, \dots$ die Wahrscheinlichkeiten dafür, zur Zeit $t = 0$ die Energiewerte $W_1(0), W_2(0), \dots$ zu finden. Die Komponenten von x im Hauptachsensystem der Energie zur Zeit t sind $(Xx)_1, (Xx)_2, \dots$ [vgl. § 61, (12)]; also ist

$$(18) \quad |(Xx)_n|^2 = \left| \sum_m X_{nm} x_m \right|^2$$

die Wahrscheinlichkeit, das System zur Zeit t im Zustand mit der Energie $W_n(t)$ zu finden.

Da nun nach (15)

$$X_{nn} = 1 + O\left(T^{-\frac{1}{r+1}}\right), \quad X_{nm} = O\left(T^{-\frac{1}{r+1}}\right)$$

ist, so erhält man

$$(19) \quad |(Xx)_n|^2 = \begin{cases} |x_n|^2 + O\left(T^{-\frac{1}{r+1}}\right), & \text{wenn } x_n \neq 0, \\ O\left(T^{-\frac{2}{r+1}}\right), & \text{wenn } x_n = 0. \end{cases}$$

Die Abweichung der Wahrscheinlichkeit des Zustandes n von ihrem Anfangswert $|x_n|^2$ ist also von verschiedener Größenordnung, je nachdem der Anfangswert Null ist oder nicht, und zwar ist

im ersteren Falle die Abweichung im allgemeinen [nämlich außer im reinen Fall bestimmter Energie, s. (16), (17)] kleiner, d. h. von höherer Ordnung in $\frac{1}{T}$.

Wir haben jetzt noch zu diskutieren, ob die Voraussetzungen des Beweises in bestimmten Fällen als erfüllt nachgewiesen werden können. Es handelt sich vor allem um die Voraussetzung 1., die absolute Beschränktheit der Matrix P . Ob diese erfüllt ist, läßt sich in einigen Fällen mit Hilfe der folgenden (hinreichenden) Kriterien beurteilen. Nach einem Satze von SCHUR¹ ist es sicher der Fall, wenn die Reihe

$$z_n = \sum_k |P_{nk}|$$

konvergiert und unterhalb einer von n unabhängigen Schranke liegt. Nach § 61, (24), (27) ist

$$z_n = \sum_k' \frac{|R_{nk}|}{|W_n - W_k|},$$

wo der Strich am Summenzeichen bedeutet, daß das Glied $k = n$ fortzulassen ist.

Nehmen wir nun an, daß die Reihe

$$(20) \quad \alpha_n = \sum_k' \frac{1}{(W_n - W_k)^2}$$

konvergiert und bezeichnen mit β_n den Ausdruck

$$(21) \quad \beta_n = \sum_k' |R_{nk}|^2 = \sum_k' |(V^{-1} \dot{H} V)_{nk}|^2,$$

so können wir z mit Hilfe der SCHWARZSchen Ungleichung § 10, (10) abschätzen, nämlich

$$z_n \leq \sqrt{\alpha_n \beta_n}.$$

Somit erhalten wir die folgende Bedingung für die absolute Beschränktheit der Matrix P : Das Produkt $\alpha_n \beta_n$ soll für alle n unterhalb einer von n unabhängigen Schranke A liegen:

$$(22) \quad \alpha_n \beta_n \leq A.$$

Wenn die Eigenwerte W_n (wie im Falle des harmonischen Oszillators) proportional n wachsen, so konvergiert die Reihe (20),

¹ SCHUR, I.: Crelles Journ. Bd. 140, S. 1. 1911; (Satz 1).

und ihre Summe α_n bleibt kleiner als eine von n unabhängige Zahl. Dann genügt für die absolute Beschränktheit der Matrix P die Endlichkeit des Ausdrucks β_n (21), und das wird immer der Fall sein, wenn die Ableitung der Störungsenergie nach der Zeit eine beschränkte Funktion ist.

Außer bei Systemen vom Typus des harmonischen Oszillators ist die Annahme von nur diskreten Energiewerten erfüllt für Systeme, die in eine feste Hülle eingeschlossen sind. Solche werden wir später behandeln und sehen, daß die Eigenwerte W_n (im Falle eines Freiheitsgrades) proportional n^2 wachsen. Dann nehmen die Größen α_n wie $\frac{1}{n^2}$ ab, und es genügt vorauszusetzen, daß die Größen β_n nicht schneller als proportional n^2 wachsen, was bei sehr allgemeinen Voraussetzungen über die Störungsenergie zutrifft.

Zum Schluß sei noch bemerkt, daß der Adiabatensatz auch in Fällen gelten kann, für die er hier nicht bewiesen ist¹. Vermutlich bleibt der Adiabatensatz im wesentlichen auch dann gültig, wenn neben dem Punktspektrum auch ein Streckenspektrum vorhanden ist.

§ 63. Wahrscheinlichkeit der Lichtabsorption. Der Begriff der Übergangswahrscheinlichkeit wurde in der älteren Quantentheorie faßt ausschließlich auf den Fall angewandt, daß ein System durch Emission oder Absorption von Licht in einen andern Zustand übergeht. Streng genommen hat man solche Vorgänge in der Weise zu behandeln, daß man das Atom oder die Molekel und das umgebende Strahlungsfeld als ein abgeschlossenes System ansieht; das werden wir auch später (Kap. VII) so machen. Die Schwierigkeit dabei besteht darin, daß das Lichtfeld selbst als quantenmechanisches System behandelt werden muß. Man kann diese Unbequemlichkeit im Falle der *Absorption* umgehen, indem man versucht, ob es nicht genügt, das elektrische Feld der Lichtwelle als äußere, eingeprägte Kraft aufzufassen, so wie wir es bereits in der Dispersionstheorie (s. § 46) erfolgreich getan haben. Wir zeigen jetzt, daß man in der Tat auf diesem Wege schon zu einer

¹ So hat V. FOCK: Z. Phys. Bd. 49, S. 323, 1928 gezeigt, daß der Adiabatensatz für den harmonischen Oszillator auch gilt, wenn die Störungsenergie so beschaffen ist, daß die Matrix P nicht beschränkt ist.

in erster Näherung befriedigenden Darstellung der durch Licht-
einwirkung erzwungenen Übergänge gelangen kann.

Das elektrische Feld einer beliebigen Lichtwelle sei an der
Stelle des Atoms¹ im Zeitintervall Δt dargestellt durch das
FOURIER-Integral²

$$(1) \quad \mathfrak{E}(t) = \int_{-\infty}^{\infty} e(\omega) e^{2\pi i \omega t} d\omega, \quad 0 \leq t \leq \Delta t.$$

Dann gilt für den FOURIER-Koeffizienten die Umkehrformel

$$(2) \quad e(\omega) = \int_0^{\Delta t} \mathfrak{E}(t) e^{-2\pi i \omega t} dt$$

und die Beziehung

$$(3) \quad e^*(\omega) = e(-\omega).$$

Die mittlere Energiedichte der Lichtwelle ist gegeben durch

$$\begin{aligned} \bar{u} &= \frac{1}{8\pi} (\overline{\mathfrak{E}^2} + \overline{\mathfrak{H}^2}) = \frac{1}{4\pi} \overline{\mathfrak{E}^2} \\ &= \frac{1}{4\pi \Delta t} \int_0^{\Delta t} \mathfrak{E}^2(t) dt \\ &= \frac{1}{4\pi \Delta t} \int_0^{\Delta t} \mathfrak{E}(t) \int_{-\infty}^{\infty} e(\omega) e^{2\pi i \omega t} d\omega dt. \end{aligned}$$

Durch Umkehrung der Integrationsfolge erhält man wegen (3)

$$\begin{aligned} \bar{u} &= \frac{1}{4\pi \Delta t} \int_{-\infty}^{\infty} e(\omega) e(-\omega) d\omega \\ &= \frac{1}{4\pi \Delta t} \int_{-\infty}^{\infty} |e(\omega)|^2 d\omega \\ &= \frac{1}{2\pi \Delta t} \int_0^{\infty} |e(\omega)|^2 d\omega. \end{aligned}$$

¹ Wir vernachlässigen hier die räumliche Ausdehnung des Atoms.

² Die folgenden Überlegungen sind in vieler Hinsicht den entsprechenden der klassischen Theorie ähnlich, wie sie im Anhang 1 ausgeführt sind.

Die mittlere Energiedichte $\varrho(\omega)$ im Frequenzintervall $\omega, \omega + d\omega$, definiert durch

$$(4) \quad \bar{u} = \int_0^{\infty} \varrho(\omega) d\omega,$$

hängt aber mit dem FOURIER-Koeffizienten $e(\omega)$ so zusammen:

$$(5) \quad |e(\omega)|^2 = 2\pi \Delta t \cdot \varrho(\omega).$$

Jetzt können wir die Wahrscheinlichkeit dafür, daß ein Atom, das sich im Zeitpunkt $t = 0$ im Zustand n befand, im Zeitpunkt $t = \Delta t$ im Zustand m gefunden wird, aus der Formel § 61, (6) berechnen, indem wir für die Störungsfunktion

$$H' = - \mathfrak{P} \mathfrak{E}(t)$$

setzen. Das ergibt nach § 45, (3), (4), (9)

$$(6) \quad U_{nm}^{(1)} = - \frac{e^{-\frac{\tilde{W}_n}{\kappa} t}}{\kappa} \mathfrak{P}_{nm} \int_0^{\Delta t} \mathfrak{E}(t) e^{2\pi i \nu_{nm} t} dt$$

$$= \frac{2\pi i}{\hbar} e^{-\frac{\tilde{W}_n}{\kappa} t} \mathfrak{P}_{nm} \cdot e^*(\nu_{nm}) \quad n \neq m.$$

Die Übergangswahrscheinlichkeit ist also in erster Näherung

$$(7) \quad \Psi_{nm}^{(1)} = |U_{nm}^{(1)}|^2 = \frac{4\pi^2}{\hbar^2} \left| \mathfrak{P}_{mn} \cdot e(\nu_{nm}) \right|^2;$$

ausführlicher im Falle mehrerer Quantenzahlen

$$(8) \quad \Psi_{\varrho\sigma}^{(r,s)} = \frac{4\pi^2}{\hbar^2} \left| \mathfrak{P}_{\sigma\varrho}^{(s,r)} \cdot e(\nu_{\varrho\sigma}^{(r,s)}) \right|^2.$$

Liegt Entartung vor, ist also etwa

$$\nu_{\varrho\sigma}^{(r,s)} = \nu^{(r,s)}$$

für alle möglichen Werte von ϱ, σ , so ist die Wahrscheinlichkeit des Überganges $r \rightarrow s$ nach § 61, (6')

$$\Psi^{(r,s)} = \frac{4\pi^2}{\hbar^2} \left\| \mathfrak{P}^{(sr)} \cdot e(\nu^{(r,s)}) \right\|^2.$$

Wenn das Licht in der z -Richtung linear polarisiert ist ($e_z \neq 0$, $e_x = e_y = 0$), so ist

$$(9) \quad \Psi_{\rho\sigma}^{(r,s)} \parallel = \frac{4\pi^2}{\hbar^2} \left| \overset{\circ}{Z}_{\rho\sigma}^{(sr)} \right|^2 \cdot \left| e_z(\overset{\circ}{\nu}_{\rho\sigma}^{(r,s)}) \right|^2.$$

Wenn das Licht in der xy -Ebene rechts bzw. links zirkular polarisiert ist ($e_x = \pm i e_y$ für $\omega > 0$, $e_z = 0$), so ist für die Absorption ($r < s$)

$$(10) \quad \Psi_{\rho\sigma}^{(r,s)\perp} = \begin{cases} \frac{4\pi^2}{\hbar^2} \left| \overset{\circ}{H}_{\rho\sigma\varrho}^{(sr)} \right|^2 \cdot \left| e_x(\overset{\circ}{\nu}_{\rho\sigma}^{(r,s)}) \right|^2 \text{ (rechtszirkular, } e_x = i e_y) \\ \frac{4\pi^2}{\hbar^2} \left| \overset{\circ}{H}_{\rho\sigma\varrho}^{\dagger(sr)} \right|^2 \cdot \left| e_x(\overset{\circ}{\nu}_{\rho\sigma}^{(r,s)}) \right|^2 \text{ (linkszirkular, } e_x = -i e_y) \end{cases}$$

und umgekehrt für die erzwungene Emission ($r > s$)

Den allgemeinen Ausdruck (7) bzw. (9) für die Übergangswahrscheinlichkeit kann man mit der Energiedichte des Lichts in Beziehung setzen.

Es sei \hat{e} ein Einheitsvektor in der Richtung von e , dann ist nach (7)

$$\Psi_{nm} = \frac{4\pi^2}{\hbar^2} \left| \overset{\circ}{\mathfrak{F}}_{mn} \cdot \hat{e} \right|^2 \cdot \left| e(\overset{\circ}{\nu}_{nm}) \right|^2,$$

also nach (5)

$$(11) \quad \Psi_{nm} = \frac{8\pi^3}{\hbar^2} \Delta t \left| \overset{\circ}{\mathfrak{F}}_{mn} \cdot \hat{e} \right|^2 \cdot \varrho(\overset{\circ}{\nu}_{nm})^*.$$

Die *Übergangswahrscheinlichkeit pro Zeiteinheit*¹ ist daher

$$(12) \quad w_{nm} = \frac{1}{\Delta t} \Psi_{nm} = \frac{8\pi^3}{\hbar^2} \left| \overset{\circ}{\mathfrak{F}}_{mn} \hat{e} \right|^2 \cdot \varrho(\overset{\circ}{\nu}_{nm}),$$

bzw. bei Entartung

$$(13) \quad w^{(r,s)} = \frac{8\pi^3}{\hbar^2} \left\| \overset{\circ}{\mathfrak{F}}^{(sr)} \cdot \hat{e} \right\|^2 \varrho(\overset{\circ}{\nu}^{(r,s)}).$$

¹ Die Proportionalität mit Δt [Formel (11)] kann nur für kleine Werte von Δt gelten, da ein Teil der durch Absorption angeregten Atome später durch spontane Emission zurückfällt; tatsächlich gelten unsere Betrachtungen nur für kleine Werte von Δt , wegen der Beschränkung auf die erste Näherung und der Nichtberücksichtigung der Rückwirkung des Atoms auf das Strahlungsfeld.

Im Falle einer gleichmäßig über alle Richtungen verteilten Strahlung (Hohlraumstrahlung) ist im Mittel

$$(14) \quad \begin{cases} |\hat{e}_x|^2 = |\hat{e}_y|^2 = |\hat{e}_z|^2 = \frac{1}{3}. \\ \hat{e}_y \hat{e}_z^* + \hat{e}_z \hat{e}_y^* = 0, \dots \end{cases}$$

und infolgedessen

$$(15) \quad w_{nm} = B_{nm} \varrho(\hat{\nu}_{nm}),$$

wo

$$(16) \quad B_{nm} = B_{mn} = \frac{8\pi^3}{3\hbar^2} |\mathfrak{P}_{nm}|^2.$$

Bei Entartung betrachtet man die Übergangswahrscheinlichkeit von den aufgespaltenen Einzeltermen; es gilt dann

$$(17) \quad g_r B^{(r,s)} = g_s B^{(s,r)} = \frac{8\pi^3}{3\hbar^2} \|\mathfrak{P}^{(r,s)}\|^2.$$

Damit haben wir die in der Einleitung [§ 2, (10)] durch Korrespondenzbetrachtungen gewonnene Formel für den EINSTEINSCHEN Wahrscheinlichkeitskoeffizienten B der Einstrahlung streng aus der Quantenmechanik abgeleitet.

Man überzeuge sich, daß die Formel (17) mit den Formeln § 46, (25), (24), (22) übereinstimmt.

§ 64. Innere Resonanz entarteter Systeme. Wir wollen an dieser Stelle eine Lücke ausfüllen, die wir früher, in § 45, bei der Behandlung der Störung nicht abgeschlossener Systeme gelassen haben, nämlich den Fall der „Resonanz“, dem wir schon in § 51 und § 63 begegnet sind, wo wir uns aber auf die erste Näherung beschränkten.

Wir haben in § 45 gesehen, daß sich für jede Störung, die zu irgend einer Zeit $t = t_0$ einsetzt, die Lösung durch iterierte Integrationen nach der Zeit [§ 45, (9)] gewinnen läßt, aber auch in dem Falle, daß die Störung rein periodisch (von $t = -\infty$ bis $t = +\infty$) verläuft, konnten wir die Aufgabe dadurch lösen, daß wir den periodischen Störungsablauf als Grenzfall einer langsam anklingenden Störung auffaßten. Die so gewonnenen Formeln § 45, (13) enthielten jedoch Nenner $\hat{\nu}_{nm} \pm \nu$, wo $\hat{\nu}_{nm}$ die Eigenfrequenzen des ungestörten Systems und ν die Frequenz der Störung bedeuten; sie versagen also für $\nu = |\hat{\nu}_{nm}|$, und man spricht dann von

Resonanz zwischen der periodischen Störung und dem System. Die Störungsfunktion ist jetzt

$$(1) \quad H' = A e^{2\pi i \nu_{ab} t},$$

wo $A = A(n, m)$ eine konstante Matrix und a, b irgend zwei bestimmte stationäre Zustände des ungestörten Systems sind; und die für die Störungsrechnung maßgebende Matrix K hat nach § 45, (4) die Elemente

$$(2) \quad K_{nm} = A_{nm} e^{2\pi i (\nu_{nm} + \nu_{ab}) t},$$

die für $n = b, m = a$ konstant sind. Setzt eine solche Störung zur Zeit $t = t_0$ ein, so werden die entsprechenden Elemente der Matrizen $V^{(1)}, V^{(2)}, \dots$, aus welchen sich nach § 45, (6) die Lösung aufbaut, Polynome 1., 2., ... Grades in $(t - t_0)$. Wir sehen also, daß das Näherungsverfahren von § 45 im Falle der Resonanz nur für kleine Werte von $(t - t_0)$ brauchbar ist.

In diesem und den folgenden Paragraphen werden wir diese Resonanzfälle nach einer andern Methode behandeln, welche für beliebige Werte von $(t - t_0)$ brauchbar ist.

Und zwar wollen wir einen Sonderfall betrachten, den man als *innere Resonanz* bezeichnen kann. Es hindert uns nämlich nichts, in Formel (1) $a = b$ zu setzen; dann tritt „Resonanz“ ein für $\nu = 0$, d. h. bei zeitunabhängiger Störung. In der Tat, wenn eine solche im Augenblick t_0 eingeschaltet wird, werden nach (2) die Diagonalglieder von K konstant und bewirken, daß die Diagonalglieder der Lösung V Polynome der Zeit werden. Wir haben nun bereits früher (§ 38) Verfahren kennen gelernt, bei zeitunabhängiger Störungsfunktion die Lösung durch Matrizen mit *periodischen* Elementen darzustellen; dabei handelt es sich aber immer um eine wirklich von $t = -\infty$ bis $t = +\infty$ bestehende konstante Störung, während wir hier den Fall ins Auge fassen, daß die Störung in einem bestimmten Augenblicke t_0 einsetzt¹. Wir werden durch ein ganz analoges Verfahren wie früher zeigen, daß auch dann die Lösung durch periodische Funktionen (ohne Polynome in t) dargestellt werden kann. Wenn das ungestörte System bezüglich der Energie nicht entartet ist, bieten die Resultate kein erhebliches physikalisches Interesse; wohl aber ist der Fall ent-

¹ Vgl. Fußnote S. 329.

arteter Energie zum Verständnis eines ausgedehnten Bereiches physikalischer und chemischer Tatsachen von Wichtigkeit.

Es ist zweckmäßig, von der Transformationsmatrix U zum Wahrscheinlichkeitsvektor x überzugehen mittels

$$(3) \quad x(t) = U x(0);$$

dann läßt sich aus der Gleichung

$$(4) \quad \kappa \dot{x} + \{H_0(\dot{p}, \dot{q}) + H'(\dot{p}, \dot{q})\} x = 0$$

der Zeitablauf von x bestimmen, und zwar, wenn $U(0) = 1$, und $H_0(\dot{p}, \dot{q}) = W^0$ Diagonalmatrix ist, so sind $x(t)_n$ die Komponenten eines Wahrscheinlichkeitsvektors im Hauptachsensystem von $H_0(p(t), q(t))$, wenn $x(0)_k$ seine Komponenten im Hauptachsensystem von $H_0(p(0), q(0))$ waren [vgl. § 61, (3)]. $|x(t)_n|^2$ ist dann die Wahrscheinlichkeit bei Messung zur Zeit t für die ungestörte Energie den Wert W_n^0 zu finden.

Wie in § 45, führen wir durch den Ansatz

$$(5) \quad x = e^{-\frac{W^0}{\kappa} t} \cdot y,$$

die Gleichung (4) in die Form

$$(6) \quad \kappa \dot{y} + K y = 0, \quad K = e^{-\frac{W^0 t}{\kappa}} H' e^{-\frac{W^0}{\kappa} t}$$

über; dabei soll H' von der Zeit unabhängig sein.

Ist die Energie H_0 entartet, so zerlegen wir alle Matrizen in Übermatrizen mit den Indizes (r, ρ) . Sodann ziehen wir von K die Matrix

$$(7) \quad \overline{\overline{H}} = (K^{(r,r)} \delta_{rs}) = (H'^{(r,r)} \delta_{rs})$$

ab, welche aus allen Diagonal-Untermatrizen von H' besteht, setzen also

$$(8) \quad K' = K - \overline{\overline{H}}.$$

$\overline{\overline{H}}$ ist von der Zeit unabhängig und in K' gibt es nur Elemente mit nichtverschwindenden Frequenzen $\dot{p}^{(r,s)}$. Wir schreiben (6) so:

$$(9) \quad \kappa \dot{y} + \overline{\overline{H}} y + K' y = 0.$$

Die Gleichungen (9) lassen sich durch ein Näherungsverfahren lösen, das sich von dem des § 45 dadurch unterscheidet, daß das

wo die c_{nm} gewisse Funktionen der H'_{kl} , ν_{kl} und $x_k(0)$ sind, die wir nicht explizite hinschreiben brauchen.

Durch Einsetzen in (15) und (5) erhält man also

$$(17) \quad x_n = e^{-\frac{\bar{W}_n + H'_{nn}}{\kappa} t} \left(x_n(0) + \sum'_m c_{nm} (e^{2\pi i \nu_{nm} t} - 1) \right).$$

Die Komponenten des Wahrscheinlichkeitsvektors x führen demnach „säkulare“ Schwingungen aus mit den Frequenzen $\frac{1}{\hbar} (\bar{W}_n + H'_{nn})$ und mit Amplituden, die aus einem konstanten Gliede und überlagerten Schwingungen mit den Eigenfrequenzen ν_{nm} des Systems bestehen. Wenn die Energiewerte nicht sehr nahe beieinander liegen, haben diese Schwingungen keine physikalische Bedeutung, da man nur langsame Änderungen von $|x_n(t)|^2$ beobachten kann.

Wenn aber das ungestörte System *entartet* ist, so zerfällt die erste Gleichung (11)

$$(18) \quad \kappa \dot{y}^0 + \bar{H} y^0 = 0$$

in eine Folge von Säkularproblemen entsprechend den Stufen der Stufenmatrix $\bar{H} = (H'^{(r,r)} \delta_{rs})$.

Man bestimme nun die unitäre Stufenmatrix $u = (u^{(r)} \delta_{rs})$, $u^{(r)} = (u^{(r)}_{\sigma})$ als Lösung des Eigenwertproblems

$$(19) \quad \bar{H} u = u w, \quad w = (w^{(r)}_{\sigma} \delta_{r\sigma}),$$

und setze dann

$$(20) \quad y^0 = u z^0.$$

Dann wird nach (18)

$$\kappa u \dot{z}^0 + u w z^0 = 0,$$

also

$$\kappa \dot{z}^0 + w z^0 = 0$$

mit der Lösung

$$z^0 = e^{-\frac{wt}{\kappa}} z^0(0)$$

und

$$(21) \quad y^0 = u e^{-\frac{wt}{\kappa}} z^0(0).$$

Die Anfangsbedingung $y^0(0) = x(0) = u z^0(0)$ bestimmt $z^0(0) = u^\dagger x(0)$. Die Lösung lautet also nach (5) in nullter Näherung

$$(22) \quad x^0 = e^{-\frac{\bar{W}t}{\kappa}} u e^{-\frac{wt}{\kappa}} u^\dagger x(0)$$

mit den Komponenten

$$(23) \quad x_e^{(r)} = \sum_{\sigma\tau} u_{e\sigma}^{(r)} u_{\tau\sigma}^{(r)*} x_\tau^{(r)}(0) e^{-\frac{2\pi i}{\hbar} (\bar{W}^{(r)} + w_\sigma^{(r)})t}$$

Die folgenden Näherungsgleichungen (11) haben alle dieselbe Form

$$(24) \quad \kappa \dot{y} + \bar{H} y = c,$$

wo $c(t)$ ein durch die vorangehenden Näherungen bereits bekannter Vektor ist. Setzt man analog zu (20)

$$(25) \quad y = uz,$$

so erhält man

$$(26) \quad \kappa \dot{z} + wz = u^\dagger c$$

mit der Lösung

$$(27) \quad z = \frac{1}{\kappa} \int_0^t e^{-\frac{w(t-t')}{\kappa}} u^\dagger c(t') dt',$$

die für $t = 0$ verschwindet. In erster Näherung ist hier

$$(28) \quad c(t) = -K' y^0 = -K' u e^{-\frac{wt}{\kappa}} z^0(0)$$

einzusetzen; das gibt

$$(29) \quad z^{(1)} = -\frac{1}{\kappa} e^{-\frac{wt}{\kappa}} \int_0^t e^{\frac{wt'}{\kappa}} u^\dagger K' u e^{-\frac{wt'}{\kappa}} z^0(0) dt',$$

mit der Komponente

$$z_e^{(1)(r)} = -e^{-\frac{w^{(r)}t}{\kappa}} \sum_{s,\sigma} (u^\dagger K' u)_{e\sigma}^{(r,s)} z_\sigma^{(s)}(0) \frac{e^{2\pi i \nu_{e\sigma}^{(r,s)} t} - 1}{\hbar \nu_{e\sigma}^{(r,s)}},$$

wo

$$(30) \quad \nu_{e\sigma}^{(r,s)} = \frac{1}{\hbar} \{ (\bar{W}^{(r)} + w_\sigma^{(r)}) - (\bar{W}^{(s)} + w_\sigma^{(s)}) \}$$

gesetzt ist.

Analoges gilt für alle Störungen höherer Ordnung; sie haben das gemeinsame Kennzeichen, mit abnehmendem Betrage der Störungsfunktion selbst zu verschwinden. *Dagegen bleibt die säkulare*

Störung (22) auch bei beliebig kleiner Störung endlich, nur die Zusatzenergien $w_\sigma^{(r)}$ in den Exponenten gehen gegen Null. Es erhebt sich nun die Frage, was diese säkulare Störung eigentlich physikalisch bedeutet.

In der klassischen Mechanik gibt es auch innere Resonanz; am bekanntesten ist der Fall, den wir im folgenden Paragraphen quantenmechanisch ausführlich diskutieren werden, daß zwei gleiche Systeme miteinander gekoppelt werden. Ganz allgemein gilt: Hat ein mechanisches System eine mehrfache Eigenschwingung und wird diese durch eine λ proportionale Störung aufgespalten, so entsteht eine langsame Schwebung, deren Frequenz proportional λ ist, deren Amplitudenverhältnisse aber von λ unabhängig sind. Ganz das analoge gilt also hier auch in der Quantenmechanik, wie die Gleichung (23) zeigt. Während aber in der klassischen Theorie diese Schwebung ohne weiteres beobachtbar ist, müssen wir hier in der Quantenmechanik besonders überlegen, ob und wie die durch innere Resonanz erzeugte Schwebung zur Beobachtung gelangen kann.

Wir behaupten, daß dies durch folgende Versuchsanordnung möglich ist: Man denke sich eine Gesamtheit \mathfrak{S} gleicher Systeme mit entarteter Energie; durch Messung einer geeigneten (mit der Energie vertauschbaren) Größe A werde die Gesamtheit zum reinen Fall mit dem Wahrscheinlichkeitsvektor $x(0)$ gemacht. Dann werde zur Zeit $t = 0$ eine konstante Störung eingeschaltet; das Ausschalten der Störung soll für verschiedene Teile der Gesamtheit \mathfrak{S} zu verschiedenen Zeiten t vorgenommen werden. Für jeden dieser Teile kann man dann durch Messung der Energie und der Größe A bestimmen, welcher Bruchteil der Systeme im Zustande (r, ρ) ist; man erhält daher diese Wahrscheinlichkeit $\Psi_\rho^{(r)} = |x_\rho^{(r)}|^2$ als Funktion der Zeit. Betrachtet man insbesondere den Grenzfall *unendlich schwacher Störung*, so bleibt die Gesamtzahl der zu einem Energiewert des ungestörten Systems gehörigen Systeme stets konstant; denn die Differentialgleichung (18) gilt für jeden g_r -dimensionalen Unterraum mit den Vektoren $\hat{y}_\rho^{(r)}$, und in § 9, S. 48 wurde allgemein bewiesen, daß dann $|\hat{y}_\rho^{(r)}|^2 = \sum_\rho |\hat{y}_\rho^{(r)}|^2$ zeitlich konstant ist. Dasselbe gilt daher auch für die Summe der Wahrscheinlichkeiten in nullter Näherung, $\sum_\rho \Psi_\rho^{(r)} = |\hat{x}^{(r)}|^2$. Man kann daher in diesem Falle von Energiesprüngen ganz ab-

sehen und gelangt zu folgender Auffassung: *Bei entarteter Energie des ungestörten Systems bewirkt eine unendlich schwache, zeitlich konstante Störung, daß der Wert einer Größe A , durch deren Messung ein reiner Fall erzeugt wird, dauernd Sprünge macht, die wir „Resonanzsprünge“ nennen* (im Gegensatz zu den „Energiesprüngen“); ihre Häufigkeit innerhalb einer Zeit t läßt sich (durch Messung von A vor dem Einschalten und nach dem Ausschalten der Störung) bestimmen und hat nach (23) den Ausdruck

$$(31) \quad \Psi_{\varrho}^{(r)} = |\dot{x}_{\varrho}^{(r)}|^2 = \sum_{\sigma, \tau} \sum_{\sigma', \tau'} u_{\varrho\sigma}^{(r)} u_{\tau\sigma'}^{(r)*} u_{\varrho\sigma'}^{(r)*} u_{\tau\sigma}^{(r)} x_{\tau}^{(r)}(0) x_{\tau'}^{(r)}(0)^* e^{-2i\pi\nu_{\sigma\sigma'}^{(r)}t},$$

$$\nu_{\sigma\sigma'}^{(r)} = \frac{1}{h} (w_{\sigma}^{(r)} - w_{\sigma'}^{(r)}).$$

Die Wahrscheinlichkeit ist also ein Aggregat periodischer Zeitfunktionen mit den Frequenzen $\nu_{\sigma\sigma'}^{(r)}$, die mit verschwindender Störung gegen Null gehen: Das ist das quantenmechanische Analogon der Resonanz-Schwebung.

Setzt man in (31) $x_{\tau}^{(r)}(0) = \delta_{\tau\varrho}$, so bekommt man die *Übergangswahrscheinlichkeit vom Zustande (r, ϱ) zum Zustande (r, ϱ')* :

$$(32) \quad \Psi_{\varrho\varrho'}^{(r)} = \sum_{\sigma, \sigma'} u_{\varrho\sigma}^{(r)} u_{\varrho'\sigma'}^{(r)*} u_{\varrho\sigma'}^{(r)*} u_{\varrho'\sigma}^{(r)} e^{-2i\pi\nu_{\sigma\sigma'}^{(r)}t}.$$

Als *Beispiel* für einen solchen Schwebungsvorgang führen wir etwa die *optisch aktiven Molekeln* an, von denen wir in § 47 gesprochen haben.

Wenn eine Molekel in einer Rechts- und einer Linksform existiert, die spiegelbildlich gleich sind, so wird man bei der Überführung der einen Konfiguration in die andere eine hohe Schwelle der potentiellen Energie zu überschreiten haben. Denkt man sich diese Schwelle unendlich hoch, so hat man den Grenzfall, den man naturgemäß als „ungestörtes“ System behandelt; für dieses gibt es ja eine Rechts- und eine Linksmolekel von *gleicher* Energie. Hat man nun eine Gesamtheit voneinander unabhängiger Molekeln, die zur Zeit $t = 0$ in der Rechtsform sind, so bedeutet das einen reinen Fall, wo nicht die (entartete) Gesamtenergie, sondern eine andere Größe A , etwa das Vorzeichen des optischen Drehungsvermögens, gemessen ist. Nun werden diese Rechtsmolekeln allmählich in Linksmolekeln umspringen; die Anzahl Rechtsmolekeln, die zur Zeit t noch vorhanden sind, wird dann durch

eine Formel von der Art (31) dargestellt, und zwar für zweifache Entartung. Der vorliegende Fall ist noch dadurch gekennzeichnet, daß die Störungsfunktion H' als Funktion der Koordinaten des ungestörten Systems bei Vertauschung der Rechts- und Linkskonfiguration ungeändert bleiben muß; hieraus folgt, daß in der Matrixdarstellung, wo nicht nur die Gesamtenergie (des ungestörten Systems), sondern auch die Energien der beiden Systeme einzeln die Diagonalform haben, gilt

$$(33) \quad \begin{cases} H'_{11}(r, r) = H'_{22}(r, r) = a(r), \\ H'_{12}(r, r) = H'_{21}(r, r) = b(r), \end{cases}$$

wo $a(r)$, $b(r)$ reelle Zahlen sind. Daher führen die Gleichungen (19) hier auf das Eigenwertproblem

$$\begin{aligned} a(r) y_1^{(r)} + b(r) y_2^{(r)} &= w(r) y_1^{(r)}, \\ b(r) y_1^{(r)} + a(r) y_2^{(r)} &= w(r) y_2^{(r)}, \end{aligned}$$

mit den Lösungen

$$\begin{aligned} w_1^{(r)} &= a(r) + b(r), & y_1^{(r)} &= y_2^{(r)}, \\ w_2^{(r)} &= a(r) - b(r), & y_1^{(r)} &= -y_2^{(r)}. \end{aligned}$$

Es ist also die Matrix u zusammengesetzt aus den Bestandteilen

$$u^{(r)} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}.$$

Sind nun zu Anfang alle Systeme der Gesamtheit im Zustande 1, also $x_1^{(r)}(0) = 1$, $x_2^{(r)}(0) = 0$, so erhält man aus (23)

$$\begin{aligned} \hat{x}_1^{(r)} &= e^{-\frac{2\pi i}{\hbar} \bar{W}^{(r)} t} \frac{1}{2} \left(e^{-\frac{2\pi i}{\hbar} w_1^{(r)} t} + e^{-\frac{2\pi i}{\hbar} w_2^{(r)} t} \right), \\ \hat{x}_2^{(r)} &= e^{-\frac{2\pi i}{\hbar} \bar{W}^{(r)} t} \frac{1}{2} \left(e^{-\frac{2\pi i}{\hbar} w_1^{(r)} t} - e^{-\frac{2\pi i}{\hbar} w_2^{(r)} t} \right) \end{aligned}$$

und daraus die Wahrscheinlichkeiten, nach der Zeit t die Zustände $(r, 1)$ bzw. $(r, 2)$ zu finden:

$$(34) \quad \begin{cases} \Psi_{11}^{(r)} = |\hat{x}_1^{(r)}|^2 = \cos^2 \frac{\pi}{\hbar} (w_1^{(r)} - w_2^{(r)}) t = \cos^2 \frac{2\pi}{\hbar} b^{(r)} t, \\ \Psi_{12}^{(r)} = |\hat{x}_2^{(r)}|^2 = \sin^2 \frac{\pi}{\hbar} (w_1^{(r)} - w_2^{(r)}) t = \sin^2 \frac{2\pi}{\hbar} b^{(r)} t. \end{cases}$$

Die Formeln zeigen das Hin- und Herschwanen der Wahrscheinlichkeiten, deren Summe dabei immer gleich 1 bleibt. Die Frequenz der Schwebung hängt von der Größe $b^{(c)}$ ab, die nach (33) die Koppelung zwischen den beiden Konfigurationen darstellt.

In Wirklichkeit wird sich bei optisch aktiven Substanzen schwerlich ein solcher periodischer Zustand ausbilden, weil dazu nötig wäre, daß jede Molekel ungestört, unabhängig von allen andern ist; das ist natürlich niemals der Fall, es finden Zusammenstöße statt, und so wird der Prozeß im allgemeinen irreversibel so laufen, daß am Schluß ein „razemisches Gemisch“ der beiden spiegelbildlichen Formen vorhanden ist. Auf die quantenmechanische Beschreibung solcher einseitiger Vorgänge gehen wir im nächsten Paragraphen kurz ein.

Hier soll noch zum Schluß auf den Einfluß der bisher vernachlässigten Glieder höherer Ordnung kurz eingegangen werden; sie bewirken natürlich Quantensprünge zwischen den Zuständen *verschiedener* Energie des ungestörten Systems, deren Wahrscheinlichkeit einen viel kleineren Betrag hat, als die der Sprünge zwischen den Entartungszuständen, dafür aber um so schneller oszilliert. Wir haben diese Formeln auf anderm Wege schon im § 61, (7) abgeleitet, wollen sie aber hier noch einmal beweisen, indem wir die Formeln für ein nicht entartetes ungestörtes System benutzen; nach (12) und (13) ist

$$y = y^{(0)} + y^{(1)} = e^{-\frac{\bar{H}t}{\varkappa}} \left(x(0) - \frac{1}{\varkappa} \int_0^t K' x(0) dt \right);$$

nun war nach (3)

$$x = U \cdot x(0),$$

wo die Matrix U also ist:

$$U = e^{-\frac{1}{\varkappa}(\bar{W} + \bar{H})t} \left(1 - \frac{1}{\varkappa} \int_0^t K' dt \right).$$

Ihr Element ist für $m \neq n$

$$U_{nm} = - e^{-\frac{1}{\varkappa}(\bar{W}_n + \bar{H}_{nm})t} H'_{nm} \frac{e^{2\pi i \nu_{nm} t} - 1}{\hbar \nu_{nm}}.$$

Also ist die Übergangswahrscheinlichkeit $m \rightarrow n$:

$$(35) \quad \Psi_{mn} = |U_{nm}|^2 = \frac{4 |H'_{nm}|^2}{\hbar^2 \dot{\nu}_{nm}^2} \sin^2 \pi \nu_{nm} t,$$

in voller Übereinstimmung mit § 61, (7).

Vergleicht man das mit (34) und bezeichnet die Größenordnung von H' mit λ , so sieht man: Während die Wahrscheinlichkeiten (34) mit verschwindendem λ endlich bleiben, aber eine λ proportionale (also kleine) Frequenz haben, sind die Wahrscheinlichkeiten (35) von der Ordnung λ^2 und haben eine endliche Frequenz.

Für ein entartetes ungestörtes System tritt im Grenzfall verschwindender Störung (bei gleichzeitiger Messung der Energie und der Größe A) an die Stelle von (35)

$$(36) \quad \Psi_{e\sigma}^{(r,s)} = \frac{4 |H'_{e\sigma}{}^{(r,s)}|^2}{\hbar^2 \dot{\nu}^{(r,s)2}} \sin^2 \pi \dot{\nu}^{(r,s)} t.$$

Man hat nämlich nach (5), (25) und (29) für diesen Grenzfall ($w = 0$)

$$\begin{aligned} x &= e^{-\frac{1}{\kappa} \hat{W} t} y = e^{-\frac{1}{\kappa} \hat{W} t} u z \\ &= e^{-\frac{1}{\kappa} \hat{W} t} u (z^0 - z^{(1)}) \\ &= e^{-\frac{1}{\kappa} \hat{W} t} u \left\{ 1 - \frac{1}{\kappa} \int_0^t u^\dagger K' u dt \right\} u^\dagger x(0) \\ &= e^{-\frac{1}{\kappa} \hat{W} t} \left\{ 1 - \frac{1}{\kappa} \int_0^t K' dt \right\} x(0). \end{aligned}$$

Ist nun $x^{(r)}$ die Projektion des Vektors x in den zu $\hat{W}^{(r)}$ gehörigen Teilraum, so wird

$$x^{(r)} = \sum_s U^{(r,s)} x^{(s)}(0),$$

wobei für $r \neq s$

$$\begin{aligned} U^{(r,s)} &= -\frac{1}{\kappa} e^{-\frac{1}{\kappa} \hat{W}^{(r)} t} \int_0^t K'^{(r,s)} dt \\ &= -e^{-\frac{1}{\kappa} \hat{W}^{(r)} t} \frac{H'^{(r,s)} e^{2\pi i \hat{\nu}^{(r,s)} t} - 1}{\hbar \dot{\nu}^{(r,s)}} \end{aligned}$$

ist. Nun ergibt sich für die Übergangswahrscheinlichkeit $(s, \sigma) \rightarrow (r, \rho)$

$$\Psi_{\rho\sigma}^{(r,s)} = |U_{\rho\sigma}^{(r,s)}|^2$$

unmittelbar der Wert (36). Wird nur die Energie gemessen, so erhält man für die Wahrscheinlichkeit des Sprunges vom Energiewert $W^{(r)}$ zum Energiewert $W^{(s)}$

$$\begin{aligned} (37) \quad \Psi^{(r,s)} &= \sum_{\rho\sigma} |U_{\rho\sigma}^{(r,s)}|^2 = \|U^{(r,s)}\|^2 \\ &= \frac{4}{\hbar^2 \hat{\nu}^{(r,s)^2}} \|H'^{(r,s)}\|^2 \sin^2 \pi \hat{\nu}^{(r,s)} t. \end{aligned}$$

Bemerkenswert ist noch, daß für Werte von t , die klein sind gegenüber den Schwebungsperioden, die so verschiedenen aussehenden Formeln für die Resonanzübergänge (32) und für die Sprünge zwischen verschiedenen Energieniveaus (36) übereinstimmen. Entwickelt man nämlich die Exponentialfunktion in (32) nach t , so erhält man:

$$\begin{aligned} (38) \quad \Psi_{\rho\sigma}^{(r)} &= \sum_{\sigma\sigma'} u_{\rho\sigma}^{(r)} u_{\rho'\sigma'}^{(r)*} u_{\rho\sigma'}^{(r)*} u_{\rho'\sigma'}^{(r)} (1 + 2\pi i \nu_{\sigma\sigma'}^{(r)} t - 2\pi^2 \nu_{\sigma\sigma'}^{(r)2} t^2 + \dots). \end{aligned}$$

Wegen der Unitarität von $u^{(r)}$ gibt das erste Glied der Klammer für $\rho \neq \rho'$ Null; das zweite verschwindet wegen $\nu_{\sigma\sigma'}^{(r)} = -\nu_{\sigma'\sigma}^{(r)}$. Im dritten Gliede hat man

$$\nu_{\sigma\sigma'}^{(r)2} = \frac{1}{\hbar^2} (w_{\sigma}^{(r)2} + w_{\sigma'}^{(r)2} - 2w_{\sigma}^{(r)} w_{\sigma'}^{(r)});$$

beachtet man nun die Relation (9), aus der

$$\bar{H} = u w u^{-1} = u w u^\dagger$$

folgt, so erhält man aus (38)

$$(39) \quad \Psi_{\rho\sigma}^{(r)} = \frac{4\pi^2}{\hbar^2} |H'_{\rho\sigma}^{(r,r)}|^2 t^2.$$

Aus (36) aber ergibt sich für kleine t

$$(40) \quad \Psi_{\rho\sigma}^{(r,s)} = \frac{4\pi^2}{\hbar^2} |H'_{\rho\sigma}^{(r,s)}|^2 t^2,$$

was für $r = s$ mit (39) übereinstimmt. Man sieht also, daß eine kurz dauernde Störung auf die Resonanzübergänge ebenso wirkt wie auf die Energiesprünge; erst bei lang andauernder Störung

bildet sich der Unterschied aus, indem die Wahrscheinlichkeiten der Energiesprünge klein bleiben, die der Resonanzsprünge aber endliche Beträge (unabhängig von der Stärke der Störung) annehmen.

§ 65. Resonanz zwischen gekoppelten Systemen. Ein besonders wichtiges Beispiel für die innere Resonanz entarteter Systeme ist folgendes: Zwei (nicht entartete) Systeme Σ' , Σ'' mit den Energien \mathring{W}' , \mathring{W}'' seien durch eine lockere Koppelung $\lambda H'$ zu einem Gesamtsystem Σ vereinigt. In nullter Näherung (bei verschwindender Koppelung) ist die Energie des Gesamtsystems

$$\mathring{W} = \mathring{W}' + \mathring{W}''.$$

Es kann nun geschehen, daß diese entartet ist, nämlich wenn eine

Eigenfrequenz $\mathring{\nu}_{n'm'} = \frac{1}{h} (\mathring{W}'_n - \mathring{W}'_{m'})$ von Σ' mit einer Eigenfrequenz $\mathring{\nu}_{n''m''} = \frac{1}{h} (\mathring{W}''_{n''} - \mathring{W}''_{m''})$ von Σ'' übereinstimmt. Die

beiden Systeme stehen dann in „Resonanz“, und zwar ist \mathring{W} mindestens zweifach entartet; denn man hat

$$\mathring{W}'_n + \mathring{W}''_{m''} = \mathring{W}'_{m'} + \mathring{W}''_{n''},$$

oder

$$\mathring{W}_{n'm'} = \mathring{W}_{m'n''}.$$

Wie bereits im allgemeinen Falle des vorigen Paragraphen auseinander gesetzt wurde, wäre in der klassischen Mechanik das Verhalten der in Resonanz stehenden Systeme folgendermaßen zu beschreiben: Bei Anlegung der schwachen, dem kleinen Parameter λ proportionalen Koppelung entsteht ein langsamer, stetiger Energieaustausch zwischen den beiden Systemen durch das Auftreten „säkularer“ Schwebungen; die *Amplituden* (Fourierkoeffizienten) dieser Schwebungen sind in erster Annäherung unabhängig von der *Stärke* der Koppelung (d. h. von λ), aber natürlich abhängig von der *Art* der Koppelung; das *Tempo* der Schwebungen ist der Stärke der Koppelung *proportional* (d. h. ihre Frequenz ist proportional λ). Wie wir gesehen haben, bleibt in der Quantenmechanik der Formalismus weitgehend unverändert; die Deutung ist aber eine ganz andere. Der Energieaustausch zwischen den Teilsystemen beruht auf *unstetigen Quantensprüngen*; das System Σ' geht z. B. vom Zustande \mathring{W}'_n in den Zustand $\mathring{W}'_{m'}$ über, wäh-

rend zugleich Σ'' von $\bar{W}_{m''}$ nach $\bar{W}_{n''}$ springt. In einer Gesamtheit \mathcal{S} solcher gekoppelter Systeme Σ erfährt durch diese Sprünge die *mittlere* Anzahl der Teilsysteme Σ' bzw. Σ'' in den betreffenden Zuständen eine zeitliche Änderung, und zwar mit einer, der klassischen „Schwungsperiode“ analogen, λ proportionalen Frequenz; die Amplitude der Schwebung, d. h. der Bruchteil der in einem bestimmten Zustande befindlichen Systeme, hängt nicht von der Stärke, wohl aber von der Art der Koppelung ab. Die quantentheoretische Beschreibung des Vorgangs hat also wieder durchaus statistischen Charakter; über den Zeitpunkt eines einzelnen Quantensprungs läßt sich keine Aussage machen, nur über ihre Häufigkeit.

Aber die beschriebenen Verhältnisse scheinen zu einem *Paradoxon* zu führen, wenn man sich überlegt, daß irgend ein System niemals vollständig von der Außenwelt isoliert werden kann; es wird also immer eine sehr schwache Koppelung zwischen dem betrachteten System Σ' und der übrigen Welt Σ'' bestehen, und da in dieser natürlich alle denkbaren Frequenzen vorkommen, wird immer irgend eine Resonanz zwischen Σ' und der Umwelt Σ'' vorhanden sein. Nun ist aber die Schwebungsamplitude von der *Stärke* der Koppelung *unabhängig*. Daher scheint es, als ob es in der Quantenmechanik prinzipiell unmöglich wäre, ein System als isoliert zu denken. Es ist wichtig, sich klar zu machen, daß dies doch erlaubt ist. Hierzu wollen wir beweisen, daß das betrachtete System Σ' bei schwacher Koppelung mit Σ'' sich statistisch so verhält, als sei es isoliert, und zwar genauer in folgender Weise: Das System Σ' führt entsprechend den Grundvorstellungen der Quantentheorie infolge der Koppelung mit Σ'' unregelmäßige Sprünge aus, und zwar brauchen wir bei ganz schwacher Koppelung nur die Resonanzsprünge zu berücksichtigen. In jedem Zustande verweilt Σ' eine gewisse Zeit; die einzelnen Verweilzeiten aber sind regellos, nur die Wahrscheinlichkeit, das System nach einer gewissen Zeit in einem bestimmten Energiezustand anzutreffen, folgt den im vorigen Paragraphen abgeleiteten Gesetzen. Wir werden nun zeigen, daß folgende zwei Dinge sich hinsichtlich der Wahrscheinlichkeit wie unabhängige Ereignisse verhalten:

1. Das Antreffen des Systems Σ' in einem bestimmten der in Resonanz stehenden Energiezustände;

2. das Antreffen eines bestimmten Werts f_k einer beliebigen, zu Σ' gehörigen Größe F .

Damit sind wir dann sicher, daß jeder Mittelwert an Σ' trotz der Koppelung mit Σ'' so berechnet werden kann, als wäre Σ' allein vorhanden.

Um auszudrücken, daß eine Größe F nur vom System Σ' abhängt, müssen wir auf die Koordinatendarstellung zurückgehen. Es seien

$$(1) \quad Q' = (Q'_{n'm'}), \quad P' = (P'_{n'm'})$$

die kanonischen Variablen, für welche die (nichtentartete) Energiefunktion $\hat{H}'(p', q')$ des isolierten Systems Σ' zur Diagonalmatrix \hat{W}' wird.

$$(2) \quad Q'' = (Q''_{n''m''}), \quad P'' = (P''_{n''m''})$$

seien dasselbe für Σ'' , so daß $\hat{H}''(p'', q'') = \hat{W}''$. Fügen wir beide Systeme, zunächst ohne mechanische Koppelung, zu einem System zusammen, so können wir nach § 17 Matrizen angeben, die

$$\hat{H}'(p', q') + \hat{H}''(p'', q'')$$

zur Diagonalmatrix machen, nämlich die Verschmelzung

$$(3) \quad \begin{cases} \hat{q}'_{n',n'';m',m''} = Q'_{n'm'} \delta_{n''m''} \\ \hat{q}''_{n',n'';m',m''} = \delta_{n'm'} Q''_{n''m''} \end{cases}$$

und entsprechend für die p .

Da nun Resonanz bestehen soll, so wird die Energie des Gesamtsystems ohne Koppelung

$$(4) \quad \hat{H}'(\hat{p}', \hat{q}') + \hat{H}''(\hat{p}'', \hat{q}'') = \hat{W}' + \hat{W}'' = \hat{W}$$

mehrfache Eigenwerte haben. Wir schreiben \hat{W} als *geordnete* Diagonalmatrix mit den *verschiedenen* Eigenwerten $\hat{W}^{(r)}$; ist dieser g_r -fach, so gibt es g_r verschiedene Darstellungen von $\hat{W}^{(r)}$ als Summe der Teilenergien:

$$(5) \quad \hat{W}^{(r)} = \hat{W}^{(r)}_{\varrho} + \hat{W}^{(r)}_{\varrho}, \quad \varrho = 1, 2, \dots, g_r.$$

Sei $F(p', q')$ irgend eine Größe des Systems Σ' . Dann ist ihre Matrixdarstellung für das Gesamtsystem ohne Koppelung nach (3)

$$(6) \quad F_{n',n'';m',m''} = \hat{F}_{n',m'} \delta_{n'',m''},$$

wo $\hat{F} = F(P', Q')$ ist.

Wir haben nun für die zu einem Energiewert $\overset{\circ}{W}^{(r)}$ gehörigen Zustände der Teilsysteme zwei Indizierungen: einmal $\varrho = 1, 2, \dots, g_r$, sodann solche Wertsysteme n', n'' , für die $\overset{\circ}{W}'_{n'} + \overset{\circ}{W}''_{n''}$ denselben Wert $\overset{\circ}{W}^{(r)}$ annimmt. Einem Paar von Zuständen derselben Energie $\overset{\circ}{W}^{(r)}$ entsprechen also entweder die Indizes (r, ϱ) und (r, σ) oder alle solche Indizes $n', n''; m', m''$, für welche

$$\overset{\circ}{W}'(n') + \overset{\circ}{W}''(n'') = \overset{\circ}{W}'(m') + \overset{\circ}{W}''(m'') = \overset{\circ}{W}^{(r)}$$

ist. Ist insbesondere $\varrho \neq \sigma$, so handelt es sich um *verschiedene* Aufteilungen von $\overset{\circ}{W}^{(r)}$; also muß $\overset{\circ}{W}'(n') \neq \overset{\circ}{W}'(m')$ sein, somit auch $\overset{\circ}{W}''(n'') \neq \overset{\circ}{W}''(m'')$, mithin $n'' \neq m''$. Nun zeigt (6), daß dann

$$F_{n', n''; m', m''} = 0$$

ist; ist aber $\varrho = \sigma$, so ist $n' = m', n'' = m''$ und $F_{n', n''; n', n''} = \overset{\circ}{F}_{n', n'}$.

Dieses Resultat läßt sich auch so formulieren: schreibt man $\overset{\circ}{F}$ als Übermatrix ($\overset{\circ}{F}_{\varrho\sigma}^{(r, \varrho)}$) entsprechend den Werten von $\overset{\circ}{W}^{(r)}$, so ist nach (6)

$$(7) \quad \overset{\circ}{F}_{\varrho\sigma}^{(r, \varrho)} = \overset{\circ}{F}_{\varrho}^{(r)} \delta_{\varrho\sigma};$$

die Untermatrix $\overset{\circ}{F}^{(r, r)}$ ist Diagonalmatrix, und die Diagonalelemente sind die Mittelwerte von F für das isolierte System Σ' :

$$(8) \quad \overset{\circ}{F}_{\varrho}^{(r)} = \overset{\circ}{F}_{n'n'}.$$

Nun sei zur Zeit $t = 0$ das Gesamtsystem Σ im Zustande (r, ϱ) , d. h. seine Energie sei $\overset{\circ}{W}^{(r)}$ und die des Teilsystems Σ' sei $\overset{\circ}{W}_{\varrho}^{(r)} = \overset{\circ}{W}'_{n_1}$. Zur Zeit t ist dann (bei hinreichend schwacher Koppelung) ein anderer Zustand vorhanden, mit derselben Energie $\overset{\circ}{W}^{(r)}$ von Σ , der aber im allgemeinen kein reiner Fall mehr ist bezüglich der Energie von Σ' .

In dem zu $\overset{\circ}{W}^{(r)}$ gehörigen, g_r -dimensionalen Teilraum des HILBERT-Raumes ist der Anfangszustand durch den Vektor $x^{(\varrho)}(0)$ mit den Komponenten $x_{\tau}^{(\varrho)}(0) = \delta_{\varrho\tau}$ gegeben. Die Wahrscheinlichkeit, bei Messung zur Zeit t für die Energie von Σ' den Wert $\overset{\circ}{W}'_{\sigma}$ zu finden, ist

$$(9) \quad w_{\varrho\sigma}^{(r)} = |x_{\sigma}^{(\varrho)}(t)|^2$$

Der Mittelwert der Größe F im Endzustande ist andererseits

$$(10) \quad \bar{F}^{(r)} = (\overset{\circ}{F}^{(r,r)} x^{(e)}(t), x^{(e)}(t)),$$

und da nach (7) $\overset{\circ}{F}^{(r,r)}$ Diagonalmatrix ist, erhält man

$$\bar{F}^{(r)} = \sum_{\tau} \overset{\circ}{F}_{\tau}^{(r)} |x_{\tau}^{(e)}|^2$$

und wegen (9)

$$(11) \quad F^{(r)} = \sum_{\tau} \overset{\circ}{F}_{\sigma}^{(r)} w_{\sigma}^{(r)}(t).$$

Diese Formel enthält unsere Behauptung: Der Mittelwert einer Größe F des Systems Σ' im Zustande (r, σ) des Gesamtsystems Σ setzt sich zusammen aus den Mittelwerten derselben Größe für das isolierte System Σ' , soweit sie zum selben Energiewert $\overset{\circ}{W}^{(r)}$ des Gesamtsystems Σ gehören, mit Gewichten, die gleich den Wahrscheinlichkeiten sind, das Teilsystem Σ' in den verschiedenen, zu $\overset{\circ}{W}^{(r)}$ gehörigen Zuständen zu finden.

Man kann dasselbe auch folgendermaßen ausdrücken: Die Formel (11) gilt offenbar auch für jede Funktion von F ; wir wenden sie an auf den Einzeltensor F_k , der zum Eigenwert f_k von F gehört. Dann ist der Mittelwert $\bar{F}^{(r)} = w^{(r)}(f_k)$ die Wahrscheinlichkeit dafür, daß im Zustande r des Gesamtsystems Σ die Größe F den Wert f_k annimmt; $\overset{\circ}{F}_{\sigma}^{(r)} = w'_{n'}(f_k)$ ist ebenso die Wahrscheinlichkeit, daß im Zustande n' des (isolierten) Teilsystem Σ' die Größe F den Wert f_k annimmt. Schreiben wir noch statt $w_{\sigma}^{(r)}$ entsprechend $w_{n'}^{(r)}$, als die Wahrscheinlichkeit, daß im Zustande r von Σ das Teilsystem Σ' im Zustande n' ist, so wird aus (11)

$$(12) \quad w^{(r)}(f_k) = \sum_{n'} w_{n'}^{(r)} \cdot w'_{n'}(f_k).$$

Die Wahrscheinlichkeiten setzen sich also so zusammen, als wenn es sich um unabhängige Ereignisse handelte.

Wenn man noch bedenkt, daß die Verweilzeit in einem Zustande, d. h. die Zeit, worin die $w_{n'}^{(r)}(t)$ sich nicht merklich geändert haben, bei lockerer Koppelung sehr lang ist, so bedeuten die vorstehenden Betrachtungen die quantenmechanische *Rechtfertigung des Grundprinzips jedes Experimentierens*: man darf hinreichend lose mit der übrigen Welt gekoppelte Systeme als isoliert betrachten. Die für streng abgeschlossene Systeme gültigen Be-

griffe und Vorstellungen sind auch auf *beinahe* isolierte Systeme anwendbar. Zu gleicher Zeit gewinnen wir auch eine nachträgliche Rechtfertigung unseres ganzen Aufbaues der Quantenmechanik. Wir haben ja als empirische Grundlage der Theorie die Eigenschaften des von den Atomen ausgesandten Lichts genommen; das bedeutet aber eine Koppelung der Atome mit dem umgebenden elektromagnetischen Felde. Solange diese hinreichend schwach ist, kann man trotzdem diese Atomzustände so berechnen, als wären die Atome isoliert; nachträglich kann man dann die Rückwirkung der Strahlung auf das Atom in Rechnung setzen. Das entspricht gerade den Entwicklungen in diesem Buche, dessen letztes Kapitel die Erörterung der Strahlungswirkung bringen wird.

Als Vorbereitung hierzu wollen wir nun ganz allgemein den Fall betrachten, daß die Koppelung der Systeme Σ' , Σ'' etwas enger ist, so daß nicht nur Resonanzsprünge, sondern auch Sprünge der Gesamtenergie (des ungestörten Systems) vorkommen. Es soll jetzt angenommen werden, daß das System Σ'' ein ungeheuer dichtes Energiespektrum besitzt, was z. B. der Fall ist, wenn Σ'' ein Σ' umgebendes Strahlungsfeld oder gar die ganze „Außenwelt“ bedeutet. Dabei wollen wir wieder das Störungsverfahren von § 45 anwenden, und müssen dementsprechend (vgl. den Anfang von § 64) erwarten, daß die erste Näherung im allgemeinen nur für nicht zu große Werte von t mit der exakten Übergangswahrscheinlichkeit übereinstimmt.

Wir gehen von der Formel (36) des vorigen Paragraphen [s. auch § 61, (7)] für die Übergangswahrscheinlichkeiten aus, die hier für einen Sprung $n' \rightarrow m'$, $n'' \rightarrow m''$ so lautet:

$$(13) \quad \Psi_{n'n'',m'm''} = \frac{4 |H'_{n'n'',m'm''}|^2}{\hbar^2 (\hat{\nu}'_{n'm'} + \hat{\nu}''_{n''m''})^2} \sin^2 \pi (\hat{\nu}'_{n'm'} + \hat{\nu}''_{n''m''}) t.$$

Wir fragen nach der Wahrscheinlichkeit der Übergänge $n' \rightarrow m'$ des Systems Σ' , wenn der Anfangszustand n' , n'' scharf bestimmt ist, d. h. nach der Summe

$$(14) \quad \Psi'_{n'm'} = \sum_{m''} \Psi_{n'n'',m'm''}.$$

Wir teilen nun den Frequenzbereich $-\infty, +\infty$ in kleine Intervalle $\Delta \nu_k$ ($k = -\infty, \dots, \infty$) und setzen

$$(15) \quad \Phi_k \Delta \nu_k = \sum_{\Delta \nu_k} |H'_{n'n'',m'm''}|^2,$$

wobei die Summe über alle m'' zu erstrecken ist, für welche $\dot{\nu}_{m'',n'} = -\dot{\nu}'_{n',m'}$ im Intervall $\Delta \nu_k$ liegt. Dann wird (14)

$$\begin{aligned} \Psi'_{n'm'} &= \sum_{k=-\infty}^{\infty} \sum_{\Delta \nu_k} \Psi_{n',n'',m'm'} \\ &= \frac{4}{\hbar^2} \sum_{k=-\infty}^{\infty} \frac{\sin^2 \pi (\dot{\nu}'_{n'm'} - \bar{\nu}_k) t}{(\dot{\nu}'_{n'm'} - \bar{\nu}_k)^2} \Phi_k \Delta \nu_k, \end{aligned}$$

wo $\bar{\nu}_k$ ein Mittelwert im Intervall $\Delta \nu_k$ ist. Läßt man nun die Intervalllänge gegen Null konvergieren und nimmt an, daß dabei die Werte Φ_k einer stetigen Funktion von ν sich annähern, so erhält man den Näherungswert

$$(16) \quad \Psi'_{n'm'} = \frac{4}{\hbar^2} \int_{-\infty}^{\infty} \frac{\sin^2 \pi (\dot{\nu}'_{n'm'} - \nu) t}{(\dot{\nu}'_{n'm'} - \nu)^2} \Phi(\nu) d\nu.$$

Nun sei

$$(17) \quad \xi = \pi (\dot{\nu}'_{n'm'} - \nu) t,$$

dann wird

$$\Psi'_{n'm'} = \frac{4\pi t}{\hbar^2} \int_{-\infty}^{\infty} \frac{\sin^2 \xi}{\xi^2} \Phi\left(\dot{\nu}'_{n'm'} - \frac{\xi}{\pi t}\right) d\xi.$$

Für Werte von t , die groß sind gegen die Schwingungsperiode $\frac{1}{\dot{\nu}'_{n'm'}}$ des betrachteten Überganges, geht dies im Limes über in

$$\Psi'_{n'm'} = \frac{4\pi t}{\hbar^2} \Phi(\dot{\nu}'_{n'm'}) \int_{-\infty}^{\infty} \frac{\sin^2 \xi}{\xi^2} d\xi.$$

Das Integral¹ hat den Wert π ; also erhalten wir

$$(18) \quad \Psi'_{n'm'} = \frac{4\pi^2}{\hbar^2} \Phi(\dot{\nu}'_{n'm'}) t.$$

Es existiert also eine *Übergangswahrscheinlichkeit pro Zeiteinheit* für den Sprung $n' \rightarrow m'$

$$(19) \quad w'_{n'm'} = \frac{4\pi^2}{\hbar^2} \Phi(\dot{\nu}'_{n'm'}).$$

¹ S. etwa R. COURANT und D. HILBERT: Methoden der mathematischen Physik I, 2. Kap., § 5, S. 55: Berlin 1924.

Die Funktion $\Phi(\nu)$ ist nach (15) durch die Frequenzen des Systems Σ'' definiert; die Formel (19) zeigt, daß für die Anregung von Sprüngen des Systems Σ' die Funktion $\Phi(\nu)$ nur für solche Frequenzen ν in Betracht kommt, die mit den Eigenfrequenzen von Σ' in „Resonanz“ stehen. Auch die Energiesprünge eines an die Außenwelt gekoppelten Systems beruhen also im wesentlichen auf Resonanz.

Im allgemeinen wird $\Phi(-\nu)$ von $\Phi(\nu)$ verschieden sein; so kann z. B. Φ für negative ν -Werte Null sein, dann gibt es in Σ' nur Sprünge von höheren zu tieferen Energiewerten (z. B. bei der spontanen Ausstrahlung, die wir im nächsten Kapitel untersuchen werden).

Wir denken uns nun eine Mannigfaltigkeit \mathfrak{S} von Systemen Σ , deren jedes an ein anderes System mit unendlich dicht liegenden Eigenfrequenzen gekoppelt ist¹. Sei $N_k(t)$ die Anzahl der Systeme Σ im Zustande k ; dann läßt sich die zeitliche Änderung von N_k mit Hilfe der durch (19) definierten Übergangswahrscheinlichkeiten w_{kl} für den Sprung $k \rightarrow l$ berechnen, wenn man aus der Lehre vom radioaktiven Zerfall die Gleichung

$$(20) \quad \frac{dN_k}{dt} = \sum_l w_{lk} N_l - N_k \sum_l w_{kl}$$

übernimmt. Aus dieser folgt sofort der Erhaltungssatz

$$(21) \quad \frac{d}{dt} \sum_k N_k = 0.$$

Ferner kann man eine Bedingung dafür aufstellen, daß solche quantenmechanischen Gemenge sich irreversibel verhalten: das äußere System, mit dem das betrachtete System Σ gekoppelt ist, möge einen beliebig großen Energievorrat aufnehmen können; das bewirkt einen gerichteten Ablauf aller Prozesse in Σ derart, daß ständig Energie abgegeben, aber keine aufgenommen wird. Man kann das auch formal sehen durch Betrachtung einer Größe, die der thermodynamischen Entropie entspricht²:

$$(22) \quad S = - \sum_k N_k \ln N_k.$$

¹ Wir lassen also von jetzt ab den Strich an den Größen des betrachteten Systems fort.

² Vgl. hierzu den Aufsatz von W. PAULI jun. in der SOMMERFELD-Festschrift (Probleme der modernen Physik), S. 30; Leipzig 1928.

Man hat

$$\begin{aligned} \frac{dS}{dt} &= -\frac{d}{dt} \sum_k N_k \ln N_k = -\sum_k (1 + \ln N_k) \frac{dN_k}{dt} \\ &= -\sum_{kl} (1 + \ln N_k) (w_{lk} N_l - w_{kl} N_k). \end{aligned}$$

Vertauscht man die Summationsindizes und bildet die halbe Summe beider Ausdrücke, so erhält man

$$\frac{dS}{dt} = \frac{1}{2} \sum_{kl} (\ln N_k - \ln N_l) (w_{kl} N_k - w_{lk} N_l).$$

Die Glieder dieser Summe sind symmetrisch in k und l ; also kann man auch schreiben

$$\frac{dS}{dt} = \sum_{k>l} (\ln N_k - \ln N_l) (w_{kl} N_k - w_{lk} N_l).$$

Wir machen nun die *Annahme*, daß die Sprünge, bei denen die Energie des betrachteten Systems wächst, nicht häufiger sind als die, wo die Energie abnimmt: d. h. in Formeln, wenn die Indizes k den nach der Größe geordneten Energiewerten zugeordnet sind:

$$(23) \quad w_{kl} \geq w_{lk} \quad \text{für} \quad k > l.$$

Dann folgt

$$\frac{dS}{dt} \geq \sum_{k>l} w_{kl} (\ln N_k - \ln N_l) (N_k - N_l).$$

Da $\ln x$ eine monoton wachsende Funktion ist, folgt hieraus

$$(24) \quad \frac{dS}{dt} \geq 0.$$

Die Größe S nimmt also niemals ab.

Die Voraussetzung (23) aber bedeutet, daß

$$(25) \quad \Phi(v) \geq \Phi(-v) \quad \text{für} \quad v > 0$$

ist, und damit ist die Irreversibilität auf eine Eigenschaft der Störungsfunktion (15) zurückgeführt.

Siebentes Kapitel.

Einleitung in die Quantentheorie des Lichtes.

§ 66. Einleitung. In diesem letzten Kapitel des vorliegenden Bandes sollen einige Anwendungen der im vorangehenden entwickelten Theorie gemacht werden auf die Dynamik des elektromagnetischen Feldes und seine Wechselwirkung mit Atomen, oder anders ausgedrückt, auf die quantenmechanische Theorie des Lichtes.

Was die *Wechselwirkung mit Atomen* betrifft, so haben wir zunächst einige schon früher gemachte Ansätze jetzt streng zu begründen: Die Annahme, daß bei einem Atom aus dem Matrixelement des elektrischen Dipolmoments \mathfrak{P}_{nm} die Übergangswahrscheinlichkeit bei spontaner Emission der Frequenz ν_{nm} sich nach den Formeln § 46, (22), (24) bestimme, war von HEISENBERG nur durch korrespondenzmäßige Analogie zur klassischen Theorie gestützt worden. Dann hatten wir in Kap. V, § 46 ff. die durch ein Strahlungsfeld erzeugte Polarisation und in Kap. VI, § 63 Absorptionswahrscheinlichkeiten eines Atoms in einem Strahlungsfeld derart berechnet, daß wir den Einfluß dieses Strahlungsfeldes durch eine zeitlich variierende HAMILTONSche Funktion¹ des Atoms darstellten. Dabei ergab sich keine Möglichkeit, auch umgekehrt die Rückwirkung des Atoms auf das Strahlungsfeld zu bestimmen; um das in exakter Weise durchführen zu können, muß man eben, wie schon gesagt, berücksichtigen, daß das Strahlungsfeld *selbst ein quantenmechanisch reagierendes physikalisches Gebilde ist*. Das soll nunmehr geschehen, und wir werden damit die früheren, halb korrespondenzmäßig gewonnenen Ergebnisse bezüglich der Emission und Absorption der Atome von einem höheren Standpunkte aus rechtfertigen und miteinander in engere Verbindung bringen können. Entsprechendes gilt für die mit der Dispersion verknüpften Prozesse. Darüber hinaus erhalten wir durch die quantenmechanisch exakte Behandlung der Wechselwirkung von Atom und Lichtfeld die Unterlagen einer in feinere Einzelheiten der Erscheinungen eingehenden (aber hier noch nicht darzustellenden) Theorie, nämlich der *natürlichen Linienbreite* und der damit zusammenhängenden Effekte.

¹ Sei es mit Hilfe eines variabeln skalaren (§ 46) oder eines variabeln Vektorpotentials (§ 47).

Aber auch das *Strahlungsfeld allein*, ohne Rücksicht auf seine Wechselwirkung mit Atomen, bietet eine Reihe tiefliegender Probleme dar, deren Untersuchung für die geschichtliche Entwicklung der Quantentheorie nicht weniger bedeutungsvoll gewesen ist, als die der atomphysikalischen Aufgaben. Zunächst handelt es sich hier um das *PLANCKsche Strahlungsgesetz*, das aus einer exakten Theorie des Strahlungshohlraums auch ohne explizite Berücksichtigung der Wechselwirkungsgesetze des Lichts mit Atomen folgen muß. Der Weg, der hierzu führt, ist freilich schon lange vor der Entstehung der Quantenmechanik von DEBYE vorgezeichnet worden (1910); der exakten quantenmechanischen Theorie bleibt in diesem Punkte keine andere Aufgabe als der Nachweis, daß die von DEBYE angenommenen *Voraussetzungen* seiner Ableitung tatsächlich erfüllt sind.

Demgegenüber bieten sich in den *Schwankungserscheinungen* der Strahlung Effekte dar, deren Aufklärung und einheitliche theoretische Beherrschung vor der Quantenmechanik unmöglich war und erst durch Einführung der charakteristisch quantenmechanischen Begriffe gelingt.

EINSTEIN hat im Jahre 1905 durch die Untersuchung dieser Schwankungserscheinungen gezeigt, daß die klassische Wellentheorie bezüglich dieser Effekte zu falschen Aussagen führt, also nicht der Wirklichkeit entsprechen kann. Er wies ferner insbesondere Effekte nach, die für eine Erneuerung der NEWTONschen Emissionstheorie des Lichtes sprechen, und begründete durch diese Feststellungen seine berühmte *Lichtquantenhypothese*, nach der jede monochromatische Strahlung der Frequenz ν aus Korpuskeln der Energie $E = h\nu$ besteht.

Die Interferenzgesetze des Lichtes hatte man ehemals als entscheidend für die Wellentheorie und gegen die Emissionstheorie angesehen. Die kühne Vorstellung, daß trotz dieser Interferenz die Strahlung „andererseits“ eine korpuskulare Struktur besitze, ist von großer Fruchtbarkeit gewesen, trotz der Unklarheit, die zunächst darüber bestand, wie man diese beiden nach klassischen Begriffen ganz entgegengesetzten Vorstellungsweisen miteinander in Verbindung und Einklang bringen könnte. Diese Fruchtbarkeit hat sich besonders dann entfaltet, als L. DE BROGLIE umgekehrt Interferenzeffekte in einer materiellen Korpuskularstrahlung voraussagte (1924). Er schuf damit die Unterlagen für

SCHRÖDINGERS „Wellenmechanik“, die einen neuen, unabhängigen Zugang eröffnete zu den Begriffen und Gesetzen der Quantenmechanik, die wir in diesem Bande vom BOHRschen Korrespondenzprinzip aus erreichten. Aber diese Entwicklungen werden uns erst im nächsten Bande beschäftigen. Dort werden wir ausführlich eingehen auf die neuen Gesichtspunkte und Anregungen, die sich aus der Lichtquantenhypothese für die Gewinnung eines einheitlichen und exakten quantenmechanischen Begriffssystems ergeben haben.

Hier soll es unsere Aufgabe sein, umgekehrt zu zeigen, was die in diesem Bande entwickelte Theorie beizutragen vermag zur Aufklärung der physikalischen Problematik, die der Lichtquantenhypothese zugrunde liegt. Es kann nun gezeigt werden, daß das Versagen der klassischen Wellentheorie gegenüber den Strahlungsschwankungen aufgefaßt werden kann als eine Parallelerscheinung zu dem Versagen der klassischen Kinematik und Mechanik in der Atomtheorie, und daß dieselben Begriffsbildungen hier wie dort zu einer Lösung der Schwierigkeiten führen. Wir werden nämlich, ganz entsprechend unserem Vorgehen in der Atomtheorie, die modellmäßigen Vorstellungen der klassischen Theorie — Darstellung der Strahlungsvorgänge als *Wellenbewegungen eines schwingungsfähigen Mediums* — ungeändert übernehmen. Aber auf dieses, durch eine bestimmte HAMILTON-Funktion gekennzeichnete Modell wenden wir statt der klassischen die quantenmechanische Kinematik an. Dann zeigt sich, daß sich aus dieser quantenmechanisch modifizierten Wellentheorie genau die richtigen Schwankungsgesetze ergeben. Man gewinnt also auch die Effekte, die sich mit der Vorstellung der Lichtquanten veranschaulichen lassen, von selbst, ohne ausdrückliche Heranziehung der Korpuskularhypothese, als notwendige Folge der quantenmechanisch durchgeführten Wellentheorie. In diesen Ergebnissen zeigt sich — was wir später (folgender Band) von andern Seiten her noch deutlicher erkennen werden — daß die *Quantenmechanik in Übereinstimmung mit der physikalischen Wirklichkeit den klassischen Unterschied von Wellenstrahlung und Korpuskularstrahlung aufhebt*.

§ 67. Das klassische Modell des Strahlungsfeldes. Die klassische Elektrodynamik lehrt, die Schwingungen des elektromagnetischen Feldes in einem endlichen Raumgebiet („Hohl-

raum“), das z. B. durch reflektierende Wände begrenzt sein mag, darzustellen als *Superposition von Eigenschwingungen*. Wir betrachten statt des Hohlraums zunächst das späterhin noch zu benutzende Beispiel einer idealen *eindimensionalen elastischen Saite der Länge l* , deren seitliche Auslenkung

$$u = u(x, t), \quad 0 \leq x \leq l$$

der *Differentialgleichung*¹

$$(1) \quad u_{xx} = u_{tt}$$

und der *Randbedingung*

$$(2) \quad u(0, t) = u(l, t) = 0$$

genügt.

Die *Energie* einer solchen Saite ist (wenn wir die lineare Massendichte längs der Saite gleich 1 annehmen):

$$(3) \quad H = \frac{1}{2} \int_0^l (u_t^2 + u_x^2) dx.$$

Wir schreiben nun, (2) berücksichtigend, die Lösung $u(x, t)$ in der Form

$$(4) \quad u = \sum_{s=1}^{\infty} u_s = \sum_{s=1}^{\infty} q_s(t) \cdot \sin s \frac{\pi}{l} x,$$

indem wir (FOURIERScher Lehrsatz)

$$q_s(t) = \frac{2}{l} \int_0^l u(x, t) \sin s \frac{\pi}{l} x \cdot dx$$

setzen. Aus (4) erhalten wir für u_{xx} und u_{tt} wieder zwei FOURIER-Reihen in bezug auf x , und ihre Übereinstimmung gemäß (1) verlangt, daß die von t abhängigen FOURIER-Koeffizienten von u_{xx} und u_{tt} entsprechend *einzelnen* gleich sein müssen. Anders ausgedrückt muß also jeder einzelne Summand u_s in (4) für sich eine Lösung der Bewegungsgleichung (1) sein. Diese Summanden u_s werden *Eigenschwingungen* genannt. Jede Lösung (4) von (1), (2) ist also eine *Superposition von Eigenschwingungen*.

¹ Mit u_x , u_{xx} usw. bezeichnen wir die Ableitungen $\frac{\partial u}{\partial x}$, $\frac{\partial^2 u}{\partial x^2}$, ...

Soll u_s der Gleichung (1) genügen, so muß für $q_s(t)$ gelten:

$$(5) \quad \begin{cases} \left(s \frac{\pi}{l}\right)^2 q_s + \ddot{q}_s = 0, \\ q_s(t) = a_s \cos 2\pi(\nu_s t + \varphi_s), \quad \nu_s = \frac{s}{2l}. \end{cases}$$

Die ν_s sind die *Eigenfrequenzen*.

Einsetzen von (4) in (3) liefert

$$(6) \quad H = \frac{l}{4} \sum_{s=1}^{\infty} \{ \dot{q}_s(t)^2 + 4\pi^2 \nu_s^2 \cdot q_s(t)^2 \}.$$

Mit

$$(7) \quad \mu \dot{q}_s = p_s, \quad \mu = \frac{l}{2}$$

ergibt sich dann

$$(8) \quad H = \frac{1}{2} \sum_{s=1}^{\infty} \left\{ \frac{p_s^2}{\mu} + \mu \cdot 4\pi^2 \nu_s^2 \cdot q_s^2 \right\}.$$

Die durch (4) und (7) definierten unendlich vielen Variablen q_s, p_s bringen also die Energie H der Saite in die Gestalt einer Summe von Energien unendlich vieler *ungekoppelter harmonischer Oszillatoren*; jeder Oszillator entspricht einer Eigenschwingung. In der Tat sind *die kanonischen Bewegungsgleichungen*

$$(9) \quad \dot{p}_s = -\frac{\partial H}{\partial q_s} = -\mu(2\pi\nu_s)^2 q_s; \quad \dot{q}_s = \frac{\partial H}{\partial p_s} = \frac{p_s}{\mu},$$

wie man aus (5), (7) entnimmt, *äquivalent mit der Bewegungsgleichung* (1).

Wir wenden uns nunmehr zu den *MAXWELLSchen Gleichungen im Vakuum*:

$$(10) \quad \text{rot } \mathfrak{E} + \frac{1}{c} \dot{\mathfrak{H}} = 0, \quad \text{div } \mathfrak{H} = 0;$$

$$(11) \quad \text{rot } \mathfrak{H} - \frac{1}{c} \dot{\mathfrak{E}} = 0, \quad \text{div } \mathfrak{E} = 0.$$

Die Gleichungen (10) bedingen in bekannter Weise die Möglichkeit, \mathfrak{E} und \mathfrak{H} durch einen Vektor \mathfrak{A} und einen Skalar Φ in der Form

$$(12) \quad \mathfrak{E} = -\text{grad } \Phi - \frac{1}{c} \dot{\mathfrak{A}}, \quad \mathfrak{H} = \text{rot } \mathfrak{A}$$

darzustellen¹. Umgekehrt werden sie durch den Ansatz (12) stets erfüllt. Wir können insbesondere \mathfrak{A} derart wählen, daß $\Phi = 0$ wird. Denn ist (12) erfüllt mit $\Phi = \Phi_0$, $\mathfrak{A} = \mathfrak{A}_0$, so definieren wir $\mathfrak{A}_1(x, y, z, t)$ durch $\dot{\mathfrak{A}}_1 = c \cdot \text{grad } \Phi_0$ und $\mathfrak{A}_1(x, y, z, 0) = 0$. Dann ist $\text{rot } \mathfrak{A}_1 = 0$, und folglich werden die Gleichungen (12) auch durch $\Phi = 0$, $\mathfrak{A} = \mathfrak{A}_0 + \mathfrak{A}_1$ erfüllt:

$$(13) \quad \mathfrak{E} = -\frac{1}{c} \dot{\mathfrak{A}}, \quad \mathfrak{H} = \text{rot } \mathfrak{A}.$$

Bei vorgegebenen \mathfrak{E} , \mathfrak{H} ist \mathfrak{A} durch (13) eindeutig bestimmt bis auf einen additiven zeitlich konstanten Gradienten; dieser kann noch so eingerichtet werden, daß für $t = 0$

$$(14) \quad \text{div } \mathfrak{A}(x, y, z, 0) = 0$$

wird.

Die Gleichungen (11) bedeuten nun:

$$(15) \quad \Delta \mathfrak{A} - \frac{1}{c^2} \ddot{\mathfrak{A}} = 0; \quad \text{div } \mathfrak{A} = 0;$$

mit Δ ist der LAPLACESche Operator

$$\Delta = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}$$

gemeint.

Die *Energie* des Hohlraums ist jetzt gleich

$$(16) \quad H = \frac{1}{8\pi} \iiint_V \{ \mathfrak{E}^2 + \mathfrak{H}^2 \} dx dy dz \\ = \frac{1}{8\pi} \iiint_V \left\{ \frac{1}{c^2} \dot{\mathfrak{A}}^2 + (\text{rot } \mathfrak{A})^2 \right\} dx dy dz.$$

Wir wollen die Eigenschwingungen eines würfelförmigen Raumstücks

¹ Das Vektorfeld \mathfrak{H} ist nach (10) quellenfrei und kann folglich als Rotation eines Vektorfeldes \mathfrak{A} dargestellt werden. Mit $\mathfrak{H} = \text{rot } \mathfrak{A}$ geht aber die erste Gleichung (10) über in $\text{rot} \left(\mathfrak{E} + \frac{1}{c} \dot{\mathfrak{A}} \right) = 0$; das Vektorfeld $\mathfrak{E} + \frac{1}{c} \dot{\mathfrak{A}}$ ist also wirbelfrei und kann folglich als Gradient zu einem Skalarfeld Φ aufgefaßt werden.

$$(17) \quad \begin{cases} 0 \leq x \leq l \\ 0 \leq y \leq l \\ 0 \leq z \leq l \end{cases}$$

mit dem Volum $V = l^3$ betrachten, und zwar bei der *Randbedingung spiegelnder Wände*: d. h. wir haben auf den Wandflächen das Verschwinden der tangentiellen Komponenten von \mathfrak{E} und der normalen Komponente von \mathfrak{H} zu fordern. Für \mathfrak{A} bedeutet das: Verschwinden der tangentiellen Komponenten.

Die Eigenschwingungen unseres Problems sind nun folgendermaßen zu erhalten. Es sei \mathfrak{s} ein Vektor mit positiv-ganzzahligen Komponenten s_1, s_2, s_3 und

$$(18) \quad \mathfrak{k}_s = \frac{\mathfrak{s}}{2l}.$$

Ferner seien $\mathfrak{e}_s^{(1)}, \mathfrak{e}_s^{(2)}$ zwei zu \mathfrak{k}_s und zueinander senkrechte Einheitsvektoren. Die FOURIER-Entwicklung der gesuchten Lösung \mathfrak{A} unter Berücksichtigung der Randbedingungen und der zweiten Gleichung (15) liefert dann

$$(19a) \quad \begin{cases} \mathfrak{A}_x = \sum_s \mathfrak{D}_{s_x} \cos 2\pi \mathfrak{k}_{s_x} x \cdot \sin 2\pi \mathfrak{k}_{s_y} y \cdot \sin 2\pi \mathfrak{k}_{s_z} z, \\ \mathfrak{A}_y = \sum_s \mathfrak{D}_{s_y} \sin 2\pi \mathfrak{k}_{s_x} x \cdot \cos 2\pi \mathfrak{k}_{s_y} y \cdot \sin 2\pi \mathfrak{k}_{s_z} z, \\ \mathfrak{A}_z = \sum_s \mathfrak{D}_{s_z} \sin 2\pi \mathfrak{k}_{s_x} x \cdot \sin 2\pi \mathfrak{k}_{s_y} y \cdot \cos 2\pi \mathfrak{k}_{s_z} z; \end{cases}$$

$$(19b) \quad \mathfrak{D}_s = q_s^{(1)}(t) \mathfrak{e}_s^{(1)} + q_s^{(2)}(t) \mathfrak{e}_s^{(2)}.$$

Wir werden die Gleichungen (19a) später auch abkürzend schreiben als

$$(19c) \quad \mathfrak{A} = \sum_s \left(q_s^{(1)} \mathfrak{A}^{(1)s} + q_s^{(2)} \mathfrak{A}^{(2)s} \right).$$

Die erste Gleichung (15), die wiederum für jeden einzelnen FOURIER-Summanden gelten muß, liefert

$$(20) \quad \begin{cases} 4\pi^2 \mathfrak{k}_s^2 c^2 \mathfrak{D}_s + \ddot{\mathfrak{D}}_s = 0; \\ q_s^{(1)} = a_s^{(1)} \cos 2\pi (\nu_s t + \varphi_s^{(1)}), \\ q_s^{(2)} = a_s^{(2)} \cos 2\pi (\nu_s t + \varphi_s^{(2)}); \quad \nu_s = c |\mathfrak{k}_s|. \end{cases}$$

Damit haben wir auch für das hier betrachtete Schwingungsproblem nachgewiesen, daß jede Lösung \mathfrak{A} eine *Superposition von Eigenschwingungen ist*; und zwar entsprechen jedem Wertesystem s_1, s_2, s_3 je zwei durch $q_s^{(1)}$ und $q_s^{(2)}$ gekennzeichnete Eigenschwingungen mit der gemeinsamen Eigenfrequenz ν_s .

Die Energie H ist wiederum die Summe der Energien der einzelnen Eigenschwingungen. Es ergibt sich nach (16) und (19b):

$$(21) \quad H = \frac{V}{64\pi c^2} \sum_s \{ \dot{\mathfrak{D}}_s^2 + 4\pi^2 c^2 [\mathfrak{f}_s^2 \mathfrak{D}_s^2 - (\mathfrak{f}_s \mathfrak{D}_s)^2] \}.$$

Darin ist aber $\mathfrak{f}_s \mathfrak{D}_s = 0$, weil $e_s^{(1)}, e_s^{(2)}$ auf \mathfrak{f}_s senkrecht stehen.

Es folgt also

$$(21') \quad H = \frac{V}{64\pi c^2} \sum_s \{ \dot{q}_s^{(1)2} + \dot{q}_s^{(2)2} + 4\pi^2 \nu_s^2 (q_s^{(1)2} + q_s^{(2)2}) \},$$

womit die Zerlegung nach den Eigenschwingungen durchgeführt ist. Führen wir endlich statt der Geschwindigkeiten $\dot{q}_s^{(1)}, \dot{q}_s^{(2)}$ entsprechende Impulse

$$(22) \quad p_s^{(1)} = \mu \dot{q}_s^{(1)}, \quad p_s^{(2)} = \mu \dot{q}_s^{(2)}; \quad \mu = \frac{V}{32\pi c^2}$$

ein, so erhalten wir die HAMILTON-Funktion

$$(23) \quad H = \frac{1}{2} \sum_s \left\{ \frac{p_s^{(1)2}}{\mu} + \frac{p_s^{(2)2}}{\mu} + \mu \cdot 4\pi^2 \nu_s^2 (q_s^{(1)2} + q_s^{(2)2}) \right\}.$$

Wiederum gelten die HAMILTONSchen Gleichungen

$$(24) \quad \dot{p}_s^{(1)} = -\frac{\partial H}{\partial q_s^{(1)}}, \quad \dot{p}_s^{(2)} = -\frac{\partial H}{\partial q_s^{(2)}}; \quad \dot{q}_s^{(1)} = \frac{\partial H}{\partial p_s^{(1)}}, \quad \dot{q}_s^{(2)} = \frac{\partial H}{\partial p_s^{(2)}}.$$

Sie sind äquivalent mit der ersten Gleichung (15), da sie wie diese zu (20) führen.

§ 68. Asymptotische Verteilung der Eigenschwingungen. Es sei $\mathfrak{N}(\nu_0, V)$ die Anzahl derjenigen Eigenfrequenzen eines Hohlraums vom Volumen V , die kleiner als ν_0 sind. Für die Statistik und Thermodynamik des Hohlraums braucht man die Kenntnis des asymptotischen Wertes von $\mathfrak{N}(\nu_0, V)$ für große Werte von ν_0

(oder für große Dimensionen des Volums V), d. h. man wünscht eine Funktion $N(\nu_0, V)$ zu kennen, die sich bei hinreichend großen ν_0 (bzw. V) relativ beliebig wenig von $\mathfrak{N}(\nu_0, V)$ unterscheidet:

$$\lim_{\nu_0 \rightarrow \infty} \frac{\mathfrak{N}(\nu_0, V) - N(\nu_0, V)}{N(\nu_0, V)} = 0.$$

Daß ein solcher asymptotischer Wert besteht, und zwar *unabhängig von der Gestalt des Hohlraums* sowohl als auch von der *Wahl der Randbedingungen*, ist zuerst von LORENTZ¹ als ein beweisbedürftiger mathematischer Satz ausgesprochen worden. Der erste Beweis des Satzes ist von WEYL² gegeben. Der asymptotische Wert $N(\nu_0, V)$ ist *dem Volum V proportional*.

Für unseren würfelförmigen spiegelnden Hohlraum ist $\mathfrak{N}(\nu_0)$ die doppelte Anzahl derjenigen Punkte mit ganzzahligen Koordinaten, welche der Bedingung

$$(1) \quad \nu_s^2 = c^2 t_s^2 = \frac{c^2}{4l^2} (s_1^2 + s_2^2 + s_3^2) < \nu_0^2$$

genügen, also innerhalb eines Kugeloktanten mit dem Radius

$$(2) \quad R = \frac{2l\nu_0}{c}$$

liegen. Diese doppelte Anzahl ist aber asymptotisch gleich

$$N(\nu_0, V) = 2 \cdot \frac{1}{8} \cdot \frac{4\pi}{3} R^3 = \frac{8\pi}{3c^3} l^3 \nu_0^3.$$

Nach dem LORENTZ-WEYLSchen Satze besteht also für *jeden* beliebig geformten Hohlraum vom Volum V eine asymptotische Verteilung

$$(3) \quad N(\nu_0, V) = \frac{8\pi}{3c^3} \cdot V \nu_0^3.$$

Als asymptotische Anzahl der Eigenschwingungen im Frequenzintervall $\nu, \nu + \Delta\nu$ ergibt sich daraus

$$(4) \quad \frac{dN(\nu, V)}{d\nu} \Delta\nu = \frac{8\pi}{c^3} \cdot V \nu^2 \cdot \Delta\nu.$$

¹ LORENTZ, H. A.: Phys. Zs. Bd. 11, S. 1234. 1910.

² WEYL, H.: Math. Ann. Bd. 71, S. 441. 1912. Crelles: J. Bd. 141, 1, S. 163. 1912; Bd. 143, S. 177. 1913. Ein anderer Beweis ist von COURANT durchgeführt worden. Vgl. COURANT-HILBERT: Kap. 6, § 4.

Bei der *schwingenden Saite* haben wir offenbar an Stelle von (3) die Formel

$$(5) \quad N(\nu_0, l) = 2l\nu_0.$$

§ 69. Die PLANCKSche Formel. Der elektromagnetische Hohlraum, aufgefaßt als ein System ungekoppelter harmonischer Oszillatoren, muß den Gesetzen der statistischen Mechanik unterworfen sein. Die erste Aufgabe der statistischen Theorie ist die, die Strahlungsdichte ϱ_ν einer statistisch normalen Strahlung (schwarze Strahlung) als Funktion der Frequenz ν und der Temperatur T zu bestimmen. Dabei ist ϱ_ν definiert durch

$$(1) \quad V \cdot \varrho_\nu \cdot \Delta \nu = E(\nu, \nu + \Delta \nu),$$

wo V wieder das Volum des Hohlraums und $E(\nu, \nu + \Delta \nu)$ derjenige Anteil der Hohlraumenergie ist, der zu Eigenfrequenzen ν_s im Intervall $\nu \leq \nu_s < \nu + \Delta \nu$ gehört. Ist also ε_ν der statistische Mittelwert der Energie eines harmonischen Oszillators (Eigenschwingung) mit der Frequenz ν , so wird ϱ_ν gegeben sein durch

$$(2) \quad \varrho_\nu = \varepsilon_\nu \cdot \frac{dN(\nu, V)}{V d\nu} = \varepsilon_\nu \cdot \frac{8\pi}{c^3} \nu^2.$$

Wenn man die *klassische* Theorie anwendet, so müßte man nach dem sogenannten *Gleichverteilungssatz* der klassischen statistischen Mechanik

$$(3) \quad \varepsilon_\nu = kT$$

unabhängig von der Frequenz ν setzen. Das führt mit (2) zur klassischen *RAYLEIGH-JEANSSchen Strahlungsformel*

$$(4) \quad \varrho_\nu = \frac{8\pi}{c^3} \nu^2 \cdot kT.$$

Diese ist jedoch nur für hinreichend kleine Werte des Verhältnisses $\frac{\nu}{T}$ richtig. PLANCK hat durch Interpolation zwischen dieser

und der für große $\frac{\nu}{T}$ gültigen *WIENSchen Strahlungsformel*

$$(5) \quad \varrho_\nu = \frac{8\pi h}{c^3} \nu^3 \cdot e^{-\frac{h\nu}{kT}}$$

die empirisch richtige *PLANCKsche Strahlungsformel*

$$(6) \quad \varrho_\nu = \frac{8\pi h}{c^3} \cdot \frac{\nu^3}{e^{h\nu/T} - 1}$$

gefunden; diese hat ihn, wie man weiß, zur Entdeckung der Quanten geführt, mit deren Hilfe er eine theoretische Begründung der Strahlungsformel (6) geben konnte.

Diese *PLANCKsche* Ableitung kann etwa folgendermaßen wiedergegeben werden. *PLANCK* hatte, im Anschluß an die klassische Definition der „schwarzen Strahlung“ als derjenigen Strahlungsverteilung, welche in thermodynamischem Gleichgewicht mit einem „absolut schwarzen“ Körper sein kann, zunächst durch Rechnungen auf Grund der klassischen Mechanik und Elektrodynamik festgestellt: Ein materieller elektrisch geladener Oszillator der Frequenz ν ist dann und nur dann mit dem ihn umgebenden elektromagnetischen Strahlungsfelde im statistischen Gleichgewicht, wenn zwischen der mittleren Energie η_ν des elektrischen Oszillators und der Strahlungsdichte ϱ_ν die Beziehung

$$(7) \quad \varrho_\nu = \eta_\nu \cdot \frac{8\pi}{c^3} \nu^2$$

besteht. Er nahm an, daß diese Beziehung trotz der Unzuverlässigkeit der zu ihrer Ableitung benutzten klassischen Theorien auch in der Quantentheorie gültig bleiben müsse. Diese Annahme erscheint nachträglich ganz überzeugend, wenn man sie vom Standpunkt der Eigenschwingungstheorie des Hohlraums betrachtet; wegen (2) bedeutet sie offenbar einfach:

$$(7') \quad \eta_\nu = \varepsilon_\nu;$$

d. h. *die statistisch mittlere Energie einer Hohlraumeigenschwingung ist bei einer bestimmten Temperatur ebenso groß wie die eines materiellen Oszillators derselben Frequenz*¹.

PLANCK bestimmte dann η_ν für den materiellen Oszillator durch seine kühne Quantenhypothese, nach der ein Oszillator der Frequenz ν nur Energiewerte $n h \nu$; $n = 0, 1, 2, \dots$ besitzen kann, und zwar alle diese Werte mit a priori gleicher Wahrscheinlichkeit. Nach dieser Hypothese ist die statistische Häufigkeit, mit

¹ Ohne Berücksichtigung der quantenmechanisch bedingten Nullpunktsenergie.

welcher ein Oszillator bei der Temperatur T die Energie $n h \nu$ hat, nach dem bekannten BOLTZMANNschen Verteilungsgesetz gleich

$$(8) \quad W_n = \frac{e^{-n \frac{h\nu}{kT}}}{\sum_{m=0}^{\infty} e^{-m \frac{h\nu}{kT}}};$$

wegen

$$e^{-\frac{h\nu}{kT}} = x < 1$$

ergibt das

$$(8') \quad W_n = \left(1 - e^{-\frac{h\nu}{kT}}\right) e^{-n \frac{h\nu}{kT}} = (1-x)x^n.$$

Der Energiemittelwert wird also gleich

$$\begin{aligned} \eta_\nu &= \sum_{n=0}^{\infty} n h \nu \cdot W_n = h \nu \cdot (1-x) \sum_{n=0}^{\infty} n x^n \\ &= h \nu \cdot \frac{x}{1-x}; \end{aligned}$$

oder:

$$(9) \quad \eta_\nu = \frac{h \nu}{e^{\frac{h\nu}{kT}} - 1}.$$

Für kleine Werte von $\frac{h \nu}{kT}$ ergibt das

$$(10) \quad \left\{ \begin{aligned} \eta_\nu &= \frac{h \nu}{\frac{h \nu}{kT} + \frac{1}{2} \left(\frac{h \nu}{kT}\right)^2 + \dots} = kT \left(1 - \frac{1}{2} \frac{h \nu}{kT}\right) \\ &+ \frac{h^2 \nu^2}{kT} (\dots) \sim kT - \frac{h \nu}{2}; \end{aligned} \right.$$

also bis auf den Subtrahenden $\frac{h \nu}{2}$ dasselbe, wie die klassische Formel (3). Allgemein gibt Einsetzen von (9) in (7) die PLANCKsche Formel (6).

An Hand der Beziehung (7') ist es aber nunmehr unausweichlich, anzunehmen, daß die von PLANCK für die materiellen harmonischen Oszillatoren eingeführte *Quantelung auch für die Hohl-*

raumeigenschwingungen gültig sei. Man erhält dann die Berechtigung, von vornherein für ϵ_r dieselben Formeln anzuwenden, wie für η_r ; damit kommt man ohne Umwege unmittelbar zum PLANCKschen Gesetz.

Diese Annahme einer Quantelung der Hohlraumschwingungen ist von DEBYE 1910 eingeführt worden¹. Im übernächsten Paragraphen werden wir sie in quantenmechanisch exakter Weise formulieren. Zuvor soll jedoch im nächstfolgenden Paragraphen die PLANCKschen Formel noch einmal ausführlicher abgeleitet werden aus der Hypothese der gequantelten Eigenschwingungen, indem wir nicht die fertige BOLTZMANN-GIBBSsche Verteilungsformel anwenden, sondern auf die Berechnung der Entropie zurückgehen.

§ 70. Andere Ableitungen des PLANCKschen Gesetzes. Wir wollen den ganzen Hohlraum mit seinen sämtlichen Eigenschwingungen als ein einziges gequanteltes System betrachten. Wenn wir die Eigenschwingungen numerieren mit $r = 0, 1, 2, 3, \dots$, so gehört zu jedem r also eine Eigenfrequenz ν_r und eine Quantenzahl n_r des betreffenden Oszillators (Eigenschwingung); und

$$(1) \quad E = \sum_r h \nu_r \cdot n_r$$

ist die Energie des ganzen Hohlraums. Wir wollen nun die Frequenzskala $0 < \nu < \infty$ einteilen in schmale Bereiche der Größe $\Delta \nu$, die wir der Reihe nach $\Delta_1, \Delta_2, \dots, \Delta_k, \dots$ nennen. Zum Intervall $\Delta_x = (\bar{\nu}_x, \bar{\nu}_x + \Delta \nu)$ mögen jeweils Z_x verschiedene Eigenschwingungen gehören; diese Anzahl ist nach § 68, (4) proportional dem Volum V des Hohlraums:

$$(2) \quad Z_x = z_x \cdot V; \quad z_x = \frac{8 \pi}{c^3} \bar{\nu}_x^{-3} \cdot \Delta \nu.$$

Ferner bezeichnen wir mit E_x die zum Intervall Δ_x gehörige Energie

$$(3) \quad \left\{ \begin{aligned} E_x &= \sum_{(\bar{\nu}_x \leq \nu_r < \bar{\nu}_x + \Delta \nu)} h \nu_r \cdot n_r \\ &\approx h \bar{\nu}_x \cdot \sum_{(\bar{\nu}_x \leq \nu_r < \bar{\nu}_x + \Delta \nu)} n_r = h \bar{\nu}_x \cdot N_x. \end{aligned} \right.$$

¹ DEBYE, P.: Ann. Phys. Bd. 33, S. 1427. 1910.

Wir können den Quotienten

$$(4) \quad \frac{E_\kappa}{h \bar{\nu}_\kappa \cdot Z_\kappa} = \bar{n}_\kappa = \frac{N_\kappa}{Z_\kappa}$$

als *mittlere Quantenzahl* der Oszillatoren in Δ_κ bezeichnen. Sind die Anzahlen Z_κ hinreichend große Zahlen — wie auch bei beliebig schmalen Frequenzbereichen Δ_κ der Fall sein wird, wenn nur V hinreichend groß ist — so werden die mittleren Quantenzahlen \bar{n}_κ unabhängig von der Breite $\Delta \nu$ der gewählten Bereiche Δ_κ sein; sie kennzeichnen den *makroskopischen Zustand* des betrachteten Hohlraums. Als *mikroskopischen Zustand* bezeichnen wir demgegenüber einen durch Angabe aller Quantenzahlen n_r exakt definierten stationären Zustand des Systems.

Die *Entropie* des *makroskopischen Zustands* ist nach dem *BOLTZMANNNSCHEN Prinzip* gleich

$$(5) \quad S = k \log W,$$

wo W die Anzahl derjenigen *mikroskopischen* Zustände ist, durch welche der makroskopische Zustand realisiert werden kann¹. Wir bekommen also die gesuchte Größe W , wenn wir fragen: auf wieviele Weisen können wir die n_r so wählen, daß für jedes Intervall Δ_κ die Summe

$$(6) \quad \sum_{\bar{\nu}_\kappa \leq \nu_r < \bar{\nu}_\kappa + \Delta \nu} n_r = \frac{E_\kappa}{h \bar{\nu}_\kappa} = N_\kappa$$

gleich einer vorgegebenen Zahl wird.

Betrachten wir zunächst nur ein einziges Intervall $\Delta \nu$, so kann die Frage folgendermaßen beantwortet werden. Sei $Z_\kappa = a$, $N_\kappa = b$. Man denke sich eine Reihe von $a + b - 1$ Symbolen: $\varepsilon_1, \varepsilon_2, \dots, \varepsilon_{a+b-1}$, wovon man $a - 1$ herauswählt und durch Symbole ω ersetzt, z. B.

$$\varepsilon_1 \varepsilon_2 \omega_1 \omega_2 \varepsilon_5 \dots \omega_{a-1} \varepsilon_{a+b-1}.$$

Versteht man unter n_r die Anzahl der Symbole ε , welche zusammen (nicht getrennt durch ein ω) vor ω_r bzw. hinter $\omega_{r=1}$ stehen², so entspricht jeder Auswahl der $a - 1$ Symbole eine

¹ An dieser Stelle wird davon Gebrauch gemacht, daß alle mikroskopischen Zustände gleich wahrscheinlich sind, d. h. daß jeder Oszillator alle Quantenzustände n_r a priori mit gleicher Wahrscheinlichkeit annehmen kann.

² Im obigen Beispiel ist also: $n_1 = 2, n_2 = 0, \dots, n_a = 1$.

und nur eine Verteilung der Zahl b in a ganze nicht-negative Summanden in bestimmter Reihenfolge und umgekehrt. Die gesuchte Anzahl $W(a, b)$ der fraglichen mikroskopischen Zustände ist also die Anzahl der Möglichkeiten, aus $a + b - 1$ Dingen $a - 1$ Dinge herauszugreifen:

$$W(a, b) = \binom{a + b - 1}{a - 1} = \binom{a + b - 1}{b} = \frac{(a + b - 1)!}{(a - 1)! b!}.$$

Insgesamt (für alle Intervalle) erhalten wir dann:

$$(7) \quad W = \prod_{\kappa} \binom{N_{\kappa} + Z_{\kappa} - 1}{N_{\kappa}};$$

also

$$(8) \quad S = k \sum_{\kappa} \{ \log(N_{\kappa} + Z_{\kappa} - 1) - \log N_{\kappa}! - \log(Z_{\kappa} - 1)! \}.$$

Bei hinreichend großem Volum V , also gleichfalls großen Z_{κ} , N_{κ} können wir mit der STIRLINGSchen Formel¹⁾

$$\log N! \approx N(\log N - 1)$$

übergehen zu

$$(9) \quad S = k \sum_{\kappa} \{ (N_{\kappa} + Z_{\kappa}) \log(N_{\kappa} + Z_{\kappa}) - N_{\kappa} \log N_{\kappa} - Z_{\kappa} \log Z_{\kappa} \}.$$

Die Energie ist andererseits

$$(10) \quad E = \sum_{\kappa} E_{\kappa} = \sum_{\kappa} Z_{\kappa} \cdot h \bar{\nu}_{\kappa} \cdot \bar{n}_{\kappa}.$$

Indem wir von der Summe in (9) zum Integral übergehen und dabei \bar{n}_{κ} durch n_{ν} ersetzen, erhalten wir mit (2) als Schlussergebnis unserer Abzählungsstatistik die Entropieformel

$$(11) \quad \left\{ \begin{aligned} S &= V \cdot k \int_0^{\infty} \frac{8 \pi \nu^2}{c^3} d\nu \cdot \{ (n_{\nu} + 1) \log(n_{\nu} + 1) - n_{\nu} \log n_{\nu} \} \\ &= \int_0^{\infty} \varphi_{\nu}(n_{\nu}) d\nu; \end{aligned} \right.$$

die Energie wird gleichzeitig

$$(12) \quad E = V \cdot \int_0^{\infty} \frac{8 \pi \nu^2}{c^3} d\nu \cdot h \nu \cdot n_{\nu} = \int_0^{\infty} \gamma_{\nu} \cdot n_{\nu} \cdot d\nu.$$

Von hier aus bekommen wir die PLANCKSche Formel durch Lösung des Variationsproblems, n_ν als Funktion von ν so zu bestimmen, daß bei vorgegebener Energie E die Entropie S ihren größten Wert erhält.

Es muß also

$$(13) \quad \delta(S - k\lambda E) = 0$$

werden mit einem LAGRANGESchen Faktor $k\lambda$; also nach (11), (12):

$$(14) \quad \frac{d}{dn_\nu} \{ \varphi_\nu(n_\nu) - k\lambda \gamma_\nu n_\nu \} = \varphi'_\nu(n_\nu) - k\lambda \gamma_\nu = 0;$$

oder

$$(15) \quad \frac{d}{dn_\nu} \{ (n_\nu + 1) \log(n_\nu + 1) - n_\nu \log n_\nu \} - \lambda h \nu \\ = \log \frac{n_\nu + 1}{n_\nu} - \lambda h \nu = 0;$$

und daraus

$$(16) \quad 1 + \frac{1}{n_\nu} = e^{\lambda h \nu}; \quad n_\nu = \frac{1}{e^{\lambda h \nu} - 1}.$$

Dabei wird

$$(17) \quad \lambda = \frac{1}{kT}.$$

Wenn nämlich der Hohlraum bei festem Volum V eine Temperaturerhöhung ΔT erfährt, so gilt für die entsprechenden Änderungen ΔS und ΔE von Entropie und Energie einerseits

$$(18) \quad \Delta S = \frac{\Delta E}{T},$$

während andererseits nach (13)

$$(19) \quad \Delta S = k\lambda \Delta E$$

sein muß.

Wir können statt (14) auch

$$(20) \quad \varphi'_\nu(n_\nu) = \frac{\gamma_\nu}{T}$$

schreiben.

Für n_ν erhalten wir als endgültige Formel

$$(21) \quad n_\nu = \frac{1}{e^{\frac{h\nu}{kT}} - 1},$$

woraus mit

$$(22) \quad \varrho_\nu = \frac{8\pi}{c^3} \nu^2 \cdot h\nu \cdot n_\nu = \frac{8\pi h}{c^3} \nu^3 \cdot n_\nu$$

wieder das PLANCKSche Gesetz [§ 69, (6)] folgt. Es ist übrigens für viele Überlegungen bequemer, mit der Größe n_ν statt ϱ_ν zu rechnen; unabhängig von allen statistisch-modellmäßigen Vorstellungen können wir n_ν durch (22) definieren.

Die eben erläuterte Ableitung des PLANCKSchen Gesetzes entspricht im wesentlichen dem Wege, den DEBYE 1910, gestützt auf ältere Überlegungen PLANCKS, eingeschlagen hat.

Die der PLANCKSchen Formel äquivalente Formel (9) des § 69 ist mit dem soeben erhaltenen Ergebnis aus der Betrachtung der Entropie hergeleitet. Wir können unsere Überlegungen noch verfeinern derart, daß wir auch die Verteilungsformel (8), (8'), § 69 aus der Entropiebetrachtung ableiten können¹.

Von den zu einem Bereich Δ_κ gehörenden Oszillatoren mögen

$$(23) \quad P_\kappa^{(n)} = Z_\kappa p_\kappa^{(n)}$$

gerade die Quantenzahl n besitzen. Es ist dann also

$$(24) \quad N_\kappa = 0 \cdot P_\kappa^{(0)} + 1 \cdot P_\kappa^{(1)} + 2 \cdot P_\kappa^{(2)} + \dots$$

bzw.

$$(24') \quad \bar{n}_\kappa = 0 \cdot p_\kappa^{(0)} + 1 \cdot p_\kappa^{(1)} + 2 \cdot p_\kappa^{(2)} + \dots$$

Während wir bislang nur die Summe \bar{n}_κ untersucht haben, wollen wir jetzt auch die statistischen Normalwerte der einzelnen $p_\kappa^{(n)}$ bestimmen.

Wir haben vor allem wieder auszurechnen die Anzahl W von möglichen Verteilungen der mikroskopischen Quantenzahlen n_r

¹ Die folgenden Überlegungen nach N. S. BOSE: Z. Phys. Bd. 26, S. 178. 1924. BOSE behandelte die Statistik eines Lichtquantengases; doch können seine Ergebnisse formal ohne weiteres auf den gequantelten Hohlraum angewandt werden, wie von M. BORN, W. HEISENBERG und P. JORDAN, Z. Phys. Bd. 35, S. 557. 1926 und E. SCHRÖDINGER: Phys. Zs. Bd. 27, S. 95. 1926 bemerkt wurde.

bei vorgegebenen Werten der makroskopischen Größen $p_x^{(n)}$. Für ein bestimmtes Intervall Δ_x haben wir aus den Z_x vorhandenen Oszillatoren $P_x^{(0)}$ Oszillatoren herauszugreifen, denen wir die Quantenzahl $n_r = 0$ Null erteilen wollen; und dann $P_x^{(1)}$ Oszillatoren, bei denen $n_r = 1$ werden soll, usw. Im ganzen muß dabei

$$(25) \quad \sum_m P_x^{(m)} = Z_x$$

werden. Die Anzahl der verschiedenen Möglichkeiten solcher Zerlegungen ist aber

$$(26) \quad W_x = \frac{Z_x!}{P_x^{(0)}! P_x^{(1)}! P_x^{(2)}! \dots}$$

und insgesamt wird

$$(27) \quad W = \prod_x W_x = \prod_x \frac{Z_x!}{\prod_m P_x^{(m)}!}.$$

Für die Entropie bekommen wir also:

$$(28) \quad \left\{ \begin{aligned} S &= k \sum_x \left\{ \log Z_x! - \sum_n \log P_x^{(n)}! \right\} \\ &\approx k \sum_x \left\{ Z_x \log Z_x - \sum_n P_x^{(n)} \log P_x^{(n)} \right\} \\ &= -k \sum_x Z_x \sum_n p_x^{(n)} \log p_x^{(n)}. \end{aligned} \right.$$

Die Energie ist andererseits gleich

$$(29) \quad E = \sum_x Z_x \sum_n h \bar{\nu}_x \cdot n \cdot p_x^{(n)}.$$

Wir fordern wieder, daß S ein Maximum sei bei den Nebenbedingungen

$$(30) \quad E = \text{const}; \quad \sum_n p_x^{(n)} = 1.$$

Das heißt:

$$(31) \quad \delta \left(S - k \lambda E + k \sum_x \lambda'_x \sum_n p_x^{(n)} \right) = 0;$$

also

$$\frac{\partial}{\partial p_x^{(n)}} \left\{ -p_x^{(n)} \log p_x^{(n)} - \lambda h \bar{\nu}_x n p_x^{(n)} + \lambda'_x p_x^{(n)} \right\} = 0,$$

oder

$$\lambda'_x - 1 - \log p_x^{(n)} - \lambda \hbar \bar{\nu}_x \cdot n = 0,$$

also

$$(32) \quad p_x^{(n)} = B_x \cdot e^{-\lambda \hbar \bar{\nu}_x \cdot n}.$$

Für B_x ergibt sich dabei aus

$$(33) \quad \sum_n p_x^{(n)} = 1$$

die Formel

$$B_x = (1 - e^{-\lambda \hbar \bar{\nu}_x});$$

wenn wir sogleich wieder für λ den Wert $(kT)^{-1}$ einführen¹, wird also endlich

$$(34) \quad p_\nu^{(n)} = \left(1 - e^{-\frac{\hbar \nu}{kT}}\right) e^{-n \frac{\hbar \nu}{kT}}.$$

Das ist in der Tat wieder die Formel (8), (8'), § 9, aus der wir mit der Summation

$$(35) \quad n_\nu = \sum_n n p_\nu^{(n)}$$

das PLANCKSche Gesetz erhalten.

Alle in den §§ 69, 70 ausgeführten Überlegungen sind, wie wir zum Schluß hervorheben wollen, ohne weiteres auch für die *Eigenschwingungen eines Kristallgitters* anwendbar. Der einzige Unterschied liegt darin, das für N (ν_0 , V) an Stelle von (3), § 68 bei Kristallgittern andere Formeln eintreten.

§ 71. Quantenmechanik der Eigenschwingungen. Wir haben in den vorangehenden Paragraphen die Annahme diskreter Energiedifferenzen der stationären Zustände der einzelnen Eigenschwingungen eines Hohlraums als hinreichend für die Ableitung der PLANCKSchen Formel erkannt. Nunmehr wollen wir diese Annahme ihrerseits ableiten aus einer matrizenmechanischen Theorie des Hohlraums, die der in § 67 dargestellten klassischen Theorie korrespondenzmäßig analog ist.

¹ Aus (31) entnimmt man für Entropie- und Energiezunahme ΔS , ΔE bei einer kleinen Temperaturänderung (da $\Delta \sum_n p_x^{(n)} = 0$) wieder $\Delta S = k \lambda \Delta E$, was mit $\Delta S = \frac{\Delta E}{T}$ auch hier auf (17) führt.

Mit den Formeln (8) und (23) des § 67 haben wir die Energie der schwingenden Saite sowohl als auch des elektromagnetischen Hohlraums in die Form der HAMILTON-Funktion eines Systems ungekoppelter harmonischer Oszillatoren gebracht: *HAMILTON-Funktion der schwingenden Saite:*

$$(1) \quad H = \frac{1}{2} \sum_{s=1}^{\infty} \left\{ \frac{p_s^2}{\mu} + \mu \cdot 4\pi^2 \nu_s^2 q_s^2 \right\};$$

HAMILTON-Funktion des schwingenden Hohlraums:

$$(2) \quad H = \frac{1}{2} \sum_s \left\{ \frac{p_s^{(1)2}}{\mu} + \frac{p_s^{(2)2}}{\mu} + \mu \cdot 4\pi^2 \nu_s^2 (q_s^{(1)2} + q_s^{(2)2}) \right\}.$$

Wir haben ferner gesehen, daß die *kanonischen Gleichungen* zu diesen HAMILTON-Funktionen *gleichwertig* mit den Schwingungsgleichungen der Saite bzw. des elektromagnetischen Feldes sind, falls wir

$$(3) \quad u = \sum_{s=1}^{\infty} q_s \cdot \sin s \frac{\pi}{l} x$$

bzw.

$$(4) \quad \left\{ \begin{aligned} \mathfrak{A}_x &= \sum_s \mathfrak{D}_{sx} \cdot \cos 2\pi \mathfrak{k}_{sx} x \cdot \sin 2\pi \mathfrak{k}_{sy} y \cdot \sin 2\pi \mathfrak{k}_{sz} z, \\ \mathfrak{A}_y &= \sum_s \mathfrak{D}_{sy} \cdot \sin 2\pi \mathfrak{k}_{sx} x \cdot \cos 2\pi \mathfrak{k}_{sy} y \cdot \sin 2\pi \mathfrak{k}_{sz} z, \\ \mathfrak{A}_z &= \sum_s \mathfrak{D}_{sz} \cdot \sin 2\pi \mathfrak{k}_{sx} x \cdot \sin 2\pi \mathfrak{k}_{sy} y \cdot \cos 2\pi \mathfrak{k}_{sz} z; \\ \mathfrak{D}_s &= q_s^{(1)} c_s^{(1)} + q_s^{(2)} c_s^{(2)}; \\ \mathfrak{E} &= -\frac{1}{c} \dot{\mathfrak{A}}, \quad \mathfrak{H} = \text{rot } \mathfrak{A} \end{aligned} \right.$$

setzen.

Danach werden wir den als notwendig erkannten Übergang von der klassischen Theorie zur Quantenmechanik in der Weise vollziehen, daß wir unter Beibehaltung der HAMILTON-Funktionen (1), (2) die p, q durch *Matrizen* ersetzen, die den kanonischen Vertauschungsregeln genügen: Numerieren wir die p, q fortlaufend mit Indizes k, l , so muß

$$(5) \quad [p_k, q_l] = \delta_{kl}; \quad [p_k, p_l] = [q_k, q_l] = 0$$

werden. Wir werden ferner auch die Gleichungen (3), (4) beibehalten, um aus den Matrizen q_k, p_k die Matrix u an einer Stelle x bzw. die Matrixvektoren $\mathfrak{E}, \mathfrak{S}$ an einer Stelle x, y, z herzustellen (x, y, z sind dabei gewöhnliche *Zahlen*, nicht etwa gleichfalls Matrizen). Die in der letzten Gleichung (4) benötigte zeitliche Ableitung ist dabei in der bekannten Weise gegeben durch

$$\mathfrak{U} = [H, \mathfrak{U}].$$

Indem wir diese Gleichungen (2), (4), (5) unserer Theorie des elektromagnetischen Hohlraums zugrunde legen, machen wir eine *physikalische Hypothese*, die zwar durch die korrespondenzmäßige Analogie zur klassischen Theorie sehr nahe gelegt ist, aber im übrigen nur durch den Vergleich ihrer Folgerungen mit der *Erfahrung* geprüft und gerechtfertigt werden kann. Ihrem Charakter nach ist diese Hypothese offenbar ganz analog, z. B. zu der Hypothese, daß die HAMILTON-Funktion eines Wasserstoffatoms (bei Vernachlässigung der Feinstruktureffekte) durch

$$H = \frac{p^2}{2m} - \frac{\varepsilon}{r}$$

gegeben sei (vgl. Kap. IV, § 35); sie hat sich ebensogut wie diese Hypothese empirisch bewährt.

Zunächst gibt die durch (2), (4), (5) begründete quantenmechanische Theorie des Hohlraums in der Tat die gewünschte diskrete Quantelung der Energien der Eigenschwingungen. Nach Kap. III, § 23 kann die Energie der k -ten Eigenschwingung die Werte

$$(6) \quad W_k = h\nu_k(n_k + \frac{1}{2}); \quad n_k = 0, 1, 2, \dots$$

annehmen; der *Überschuß über die Energie im Grundzustand* beträgt also

$$(7) \quad W_k - \frac{1}{2}h\nu_k = h\nu_k \cdot n_k.$$

Insgesamt wird die Energie des Hohlraums gleich

$$(8) \quad \sum_k W_k = \sum_k \{W_k - \frac{1}{2}h\nu_k\} + \frac{1}{2}h \sum_k \nu_k = \sum_k \{W_k - \frac{1}{2}h\nu_k\} + W_0.$$

Die *Nullpunktenergie* W_0 wird wegen der unendlichen Anzahl der Eigenschwingungen unendlich groß; ihre physikalische Bedeutung ist zur Zeit noch ungewiß. Die Temperaturstatistik der *Überschußenergie* $\sum_k \{W_k - \frac{1}{2}h\nu_k\}$ führt in der in §§ 69, 70 er-

läuterten Weise zur PLANCKSchen Formel, die sich somit als Folgerung der Quantenmechanik der Eigenschwingungen erweist.

Die Matrizen q_k dieses Systems von ungekoppelten linearen Oszillatoren sind nach dem Verfahren der Verschmelzung § 15 darstellbar als

$$(9) \quad q_k(n_1 \dots n_k \dots; m_1 \dots m_k \dots) = \Delta_k q_k(n_k, m_k),$$

wobei $(q_k(n_k, m_k))$ die gewöhnlichen Oszillatormatrizen [§ 23, (25)] bedeuten, also wenn man den Wert von μ aus § 67, (22) einsetzt

$$(10) \quad \begin{cases} q_k(n_k, n_k + 1) = \sqrt{\frac{4\hbar c^2}{\pi \nu_k V}} (n_k + 1) e^{i\left(\varphi_{n_k} + \frac{\pi}{2}\right)}, \\ q_k(n_k + 1, n_k) = \sqrt{\frac{4\hbar c^2}{\pi \nu_k V}} (n_k + 1) e^{-i\left(\varphi_{n_k} + \frac{\pi}{2}\right)}, \\ q_k(n_k, m_k) = 0, \quad \text{wenn } (n_k - m_k) \neq \pm 1. \end{cases}$$

§ 72. EINSTEINS thermodynamische Schwankungsgesetze der Strahlung. Die im letzten Paragraphen erläuterte Quantelung des schwingenden Kontinuums lieferte uns die gewünschten diskreten Energiestufen der einzelnen Eigenschwingungen, aus denen schon in §§ 69, 70 die PLANCKSche Formel erhalten worden war. Wir haben aber damit zunächst nur Überlegungen bestätigen können, die schon vor der Quantenmechanik ausgeführt waren, und welche aus der exakten Quantentheorie des harmonischen Oszillators lediglich die eine Tatsache der diskreten Energiestufen im Abstand $h\nu$ benutzen. Es erhebt sich die Frage, ob darüber hinaus die quantenmechanische Behandlung des schwingenden Hohlraums auch tiefere Probleme der Strahlungstheorie zur Lösung führen vermag. Das ist in der Tat der Fall.

Der Strahlungshohlraum verhält sich nicht nur darin unklassisch, daß die PLANCKSche Strahlungsformel statt der RAYLEIGH-JEANSchen gilt, sondern zeigt auch in anderer Beziehung Effekte, die durch die klassische Theorie nicht zu erklären sind und für deren Deutung auch die ältere Quantentheorie nicht ausreichte. Es sind dies vor allem die von EINSTEIN¹ untersuchten, durch Interferenzschwebungen veranlaßten *Intensitätsschwankungen* der Hohlraumstrahlung. Zu ihrer quantentheoretischen Erklärung

¹ EINSTEIN, A.: Ann. Phys. (4), Bd. 17, S. 132. 1905; Phys. Z. Bd. 10, S. 185. 1909.

genügt nicht mehr — wie beim PLANCKSchen Gesetz — die einfache Tatsache, daß die Oszillatoren quantenhafte Energiestufen $h\nu$ besitzen, sondern man bedarf dazu einer vollständigeren Kenntnis der quantentheoretischen Kinematik des harmonischen Oszillators. Wir wollen zunächst die EINSTEINSchen Überlegungen wiedergeben und dann feststellen, wieweit die im vorangehenden entwickelten Methoden der Quantenmechanik die dadurch gestellten Aufgaben zu erledigen vermögen.

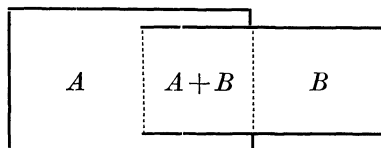
Wir haben in § 70 das PLANCKSche Gesetz abgeleitet aus der dort statistisch begründeten Entropieformel

$$(1) \quad S = V k \int_0^{\infty} \frac{8\pi\nu^2}{c^3} d\nu \{ (n_\nu + 1) \log (n_\nu + 1) - n_\nu \log n_\nu \}$$

[vgl. § 70, Gl. (11)]. Wir behaupten nun: Diese Formel (1) ist umgekehrt ableitbar aus der empirisch gesicherten PLANCKSchen Formel, ohne Heranziehung irgendwelcher modellmäßiger oder statistischer Voraussetzungen. Zunächst nämlich muß sich die Entropie irgend einer Strahlungsverteilung n_ν *additiv zusammensetzen* aus den Entropien der verschiedenen Frequenzintervalle $d\nu$, so daß man wie in § 70, (11)

$$(2) \quad S = \int_0^{\infty} \varphi_\nu(n_\nu) d\nu$$

ansetzen kann. Dies entspricht einem bekannten Satze aus der Thermodynamik der Gase: Wenn sich in einem Volum V zwei verschiedene, ideale, keine Wechselwirkungen ausübende Gase befinden, so ist (nicht nur die Energie, sondern auch) die Entropie ebenso groß, als wenn die Gase einzeln in zwei getrennten Gefäßen vom Volum V eingeschlossen sind (bei derselben Temperatur). Denn man kann mit Hilfe von semipermeablen Wänden, welche nur die A -Atome bzw. nur die B -Atome hindurchlassen, die beiden getrennten Gefäße ohne Arbeitsleistung reversibel ineinanderschieben (vgl. das Schema).



In ganz gleicher Weise ist die durch die obige Formel ausgedrückte Additivität der Entropien der einzelnen monochromatischen Partialstrahlungen der Hohlraumstrahlung zu begründen. Man braucht nur statt der semipermeablen Wände zwei ideale Lichtfilter, von denen der eine etwa nur das Licht im Frequenzbereich $\nu_0, \nu_0 + \Delta\nu$ durchläßt und alles andere reflektiert, während der andere umgekehrt den Frequenzbereich $\Delta\nu$ reflektiert.

Der durch die Forderung maximaler Entropie gegebene Zusammenhang zwischen der Funktion $\varphi_\nu(n_\nu)$ und der Formel¹

$$(3) \quad n_\nu = f_\nu(T) = \frac{1}{e^{\frac{h\nu}{kT}} - 1}$$

der stationären Strahlungsverteilung gestattet aber nicht nur die Berechnung von $f_\nu(T)$ bei gegebenem $\varphi_\nu(n_\nu)$, sondern auch umgekehrt die Bestimmung von $\varphi_\nu(n_\nu)$ bei bekanntem $f_\nu(T)$. Denn diese Maximumforderung drückt sich mit der Bezeichnungswiese der Formeln (11), (12), § 70 aus durch (20), § 70:

$$(4) \quad \varphi'_\nu(n_\nu) = \frac{\gamma_\nu}{T}.$$

Kennt man nun $n_\nu = f_\nu(T)$, so daß man als Umkehrung

$$(5) \quad T = g_\nu(n_\nu)$$

bestimmen kann, so erhält man wegen (4):

$$(6) \quad \varphi'_\nu(n_\nu) = \frac{\gamma_\nu}{g_\nu(n_\nu)},$$

oder

$$(7) \quad \varphi_\nu(x) = \int_0^x \frac{\gamma_\nu}{g_\nu(y)} dy.$$

Es ist also $\varphi_\nu(n_\nu)$ durch die PLANCKSche Formel eindeutig bestimmt und somit notwendig übereinstimmend mit unserem früheren, in (1) wiedergegebenen Ausdruck: Die Entropieformel (1) kann also ebenso, wie die PLANCKSche Formel, als ein von jeder speziellen Theorie unabhängiges *empirisches Gesetz* angesehen werden.

¹ Hierbei ist n_ν entsprechend der Bemerkung in § 70, S. 381 lediglich als Abkürzung für $n_\nu = \frac{\rho_\nu}{\gamma_\nu}$ aufzufassen, wo $\gamma_\nu = \frac{8\pi h}{c^3} \nu^3$ ist.

Ist überall $n_\nu \ll 1$ (wie es dem WIENSCHEN Strahlungsgesetz entspricht), so können wir (1) unter Ersetzung von $(n_\nu + 1)$ durch n_ν vereinfachen zu

$$(8) \quad S = V \cdot k \int_0^\infty \frac{8 \pi \nu^2}{c^3} d\nu \cdot n_\nu \{ 1 - \log n_\nu \}.$$

Aus dieser Formel hat EINSTEIN im Jahre 1905 eine sehr überraschende Folgerung gezogen. Wir wollen einen Hohlraum betrachten, bei welchem die ganze darin vorhandene Strahlung zu einem schmalen Frequenzintervall $\Delta \nu$ gehört. Da aber immerhin noch viele verschiedene Eigenschwingungen zu diesem Intervall gehören sollen, so wird infolge von Interferenzen und Schwebungen die Energie im Hohlraum räumlich hin- und herschwanken, wobei sich jedoch die Frequenz der Strahlung *nicht* ändert. Es kann z. B. in seltenen Fällen einmal vorkommen, daß zu einem bestimmten Zeitpunkt praktisch die ganze Strahlungsenergie in einem Teilstück des Volums V angesammelt ist. Die Wahrscheinlichkeit für ein solches Vorkommnis kann aber nach dem BOLTZMANN'SCHEN Prinzip

$$(9) \quad S = k \log W$$

berechnet werden: Ist zunächst die Energie $E = V \rho_\nu \Delta \nu$ gleichmäßig über V verteilt, so liefert (9) für die Entropie den Wert:

$$(10) \quad \begin{aligned} S &= V \cdot k \frac{8 \pi \nu^2}{c^3} \Delta \nu \cdot n_\nu \{ 1 - \log n_\nu \} \\ &= k \frac{E}{h \nu} \left\{ 1 - \log \frac{E}{V \cdot \alpha \nu^3 \Delta \nu} \right\}; \quad \alpha = \frac{8 \pi h}{c^3}. \end{aligned}$$

(Vgl. Fußnote S. 388.) Ist aber E auf V_0 zusammengezogen, so wird die Entropie gleich

$$(10') \quad S_0 = k \frac{E}{h \nu} \left\{ 1 - \log \frac{E}{V_0 \cdot \alpha \nu^3 \cdot \Delta \nu} \right\}.$$

Das gibt als relative Wahrscheinlichkeit des fraglichen Ausnahmezustandes gegenüber dem Normalzustand den Ausdruck

$$(11) \quad \frac{W_0}{W} = e^{\frac{S_0 - S}{k}} = \left(\frac{V_0}{V} \right)^{\frac{E}{h \nu}}.$$

Diese Formel zeigt nun mit anschaulicher Deutlichkeit, daß die Strahlung eine wesentlich andere Struktur besitzt, als ihr durch die klassische Wellentheorie zugeschrieben wurde. Wir sehen, daß jedenfalls im Grenzfall des WIENSchen Gesetzes die Strahlung sich völlig *analog einem idealen Gase aus Lichtkorpuskeln* $h\nu$ *verhält*; in einem idealen Gase von n Atomen ist ja gleichfalls die relative Wahrscheinlichkeit, daß alle n Atome in einem Teilvolum V_0 versammelt sind, gegeben durch

$$(12) \quad \frac{W_0}{W} = \left(\frac{V_0}{V}\right)^n;$$

und das stimmt überein mit (11'), wenn wir

$$(13) \quad n = \frac{E}{h\nu}$$

annehmen.

Die Quantenstruktur der Strahlung erweist sich auch¹ bei Betrachtung des mittleren Schwankungsquadrates

$$(14) \quad \overline{(\Delta E_0)^2} = \overline{(E_0 - \bar{E}_0)^2},$$

wo E_0 den in V_0 enthaltenen Anteil von E und die Überstreichung die Bildung des statistischen oder zeitlichen Mittelwerts bedeuten soll. Das Teilvolum V_0 möge jetzt sehr klein gegenüber V sein.

Schwankt die Energie E_0 in V_0 (während die totale Energie in V und der Frequenzbereich, wie oben, ungeändert bleibt), so hängen die Temperatur $T_0(E_0)$ in V_0 , die Temperatur $T_1(E_0)$ im übrigen Raume V_1 , die Entropie $S_0(E_0)$ von V_0 und die Entropie $S_1(E_0)$ von V_1 nur von E_0 ab. Und zwar ist:

$$(15) \quad T_0(\bar{E}_0) = T_1(\bar{E}_0) = T.$$

Ist $S = S_0 + S_1$ die Entropie des ganzen Raumes V , so entspricht also einem Überschuß ΔE_0 von E_0 über \bar{E}_0 die Entropieschwankung

$$(16) \quad \Delta S = \Delta(S_0 + S_1) = \frac{dS}{dE_0} \Delta E_0 + \frac{1}{2} \frac{d^2 S_0}{dE_0^2} (\Delta E_0)^2 \\ + \frac{1}{2} \frac{d^2 S_1}{dE_0^2} (\Delta E_0)^2.$$

¹ EINSTEIN, A.: Phys. Z. Bd. 10, S. 185. 1909.

Die Koeffizienten in dieser Entwicklung sind für den Wert $E_0 = \bar{E}_0$ zu nehmen. Also ist

$$\frac{dS}{dE_0} = 0; \quad \frac{d^2 S_0}{dE_0^2} = \frac{d}{dE_0} \left(\frac{1}{T_0} \right) = -\frac{1}{c T^2},$$

wo

$$\frac{1}{c} = \frac{dT_0}{dE_0};$$

dagegen ist $\frac{d^2 S_1}{dE_0^2} = -\frac{d}{dE_0} \left(\frac{1}{T_1} \right)$ zu vernachlässigen, falls V_0 sehr klein gegenüber V ist. Also wird

$$(17) \quad \Delta S = -\frac{(\Delta E_0)^2}{2 c T^2}.$$

Die *Wahrscheinlichkeit* einer solchen Entropieschwankung wird aber nach dem BOLTZMANNschen Prinzip gleich:

$$(18) \quad W = \text{const. } e^{-\frac{(\Delta E_0)^2}{2 k c T^2}}.$$

Daraus ergibt sich

$$(\Delta E_0)^2 = \frac{\int_{-\infty}^{+\infty} (\Delta E_0)^2 \cdot e^{-\frac{(\Delta E_0)^2}{2 k c T^2}} \cdot d\Delta E_0}{\int_{-\infty}^{+\infty} e^{-\frac{(\Delta E_0)^2}{2 k c T^2}} \cdot d\Delta E_0},$$

oder¹

$$(19) \quad \overline{(\Delta E_0)^2} = k c T^2 = k T^2 \frac{dE_0(T)}{dT}.$$

Mit

$$\bar{E}_0 = Z_\nu^0 n_\nu \cdot h \nu = Z_\nu^0 \frac{h \nu}{e^{\frac{h \nu}{k T}} - 1}$$

(Z_ν^0 = Anzahl der Eigenschwingungen in $\Delta \nu$ für V_0) wird also

$$(20) \quad \overline{(\Delta E_0)^2} = h \nu \cdot \bar{E}_0 + \frac{\bar{E}_0^2}{Z_\nu^0}.$$

¹ Es sei betont, daß die Gl. (19) im allgemeinen *nicht* auf ein *materielles Gas* angewandt werden darf, da bei diesem neben T noch ein zweiter unabhängiger Zustandsparameter (z. B. das spezifische Volum) vorhanden ist. Allerdings ist (19) *zufällig* auch für ein *klassisches ideales Gas* richtig.

Bei großer Strahlungsdichte — im Gültigkeitsgebiet der RAYLEIGH-JEANSschen Strahlungsformel — verbleibt lediglich

$$(20') \quad (\Delta E_0)^2 = \frac{\overline{E_0^2}}{Z_\nu^0};$$

das ist also die der *klassischen* Lichttheorie entsprechende Formel. Bei kleiner Strahlungsdichte (WIENSchem Strahlungsgesetz) dagegen wird¹, wenn wir noch mit $(h\nu)^2$ dividieren:

$$(20'') \quad \left(\frac{\Delta E_0}{h\nu} \right)^2 = \frac{\overline{E_0}}{h\nu}.$$

Das entspricht wieder genau einem klassischen idealen Gase mit Korpuskeln der Energie $h\nu$, deren Anzahl in V_0 also $\frac{\overline{E_0}}{h\nu}$ beträgt.

Wir werden im folgenden (11) als das erste und (20) als das *zweite* EINSTEINSche *Schwankungsgesetz* bezeichnen.

Es sei endlich hervorgehoben, daß auch die hier vorgeführten Betrachtungen ohne weiteres auf die thermische Schwingungsenergie eines *Kristallgitters* übertragen werden können. Dabei ist nur die schon in § 70 erwähnte Abänderung der Verteilungsfunktion $N(\nu, V)$ zu berücksichtigen.

§ 73. Wellentheoretische Ableitung des zweiten EINSTEINSchen Schwankungsgesetzes². Wir wollen nun der thermodynamischen, nur auf die PLANCKsche Formel gegründeten und von jeder modellmäßigen Vorstellung unabhängigen Ableitung der Schwankungsgesetze eine umgekehrt von Thermodynamik und thermodynamischer Statistik freie wellentheoretische Berechnung von $(\Delta E_0)^2$ gegenüberstellen, wobei wir unsere in § 71 formulierte quantenmechanische Beschreibung der Wellen zu benutzen haben. Der

¹ Die Bedingung für RAYLEIGH-JEANSches bzw. WIENSches Strahlungsgesetz ist offenbar

$$\frac{h\nu}{kT} \ll 1 \quad \text{bzw.} \quad \frac{h\nu}{kT} \gg 1;$$

oder

$$n_\nu \gg 1 \quad \text{bzw.} \quad n_\nu \ll 1.$$

² BORN, M., W. HEISENBERG und P. JORDAN: Z. Phys. Bd. 35, S. 557. 1926.

Einfachheit halber führen wir die Rechnung nicht am Beispiel des elektromagnetischen Hohlraums, sondern am Beispiel der schwingenden Saite vor, bei der sich genau entsprechende Verhältnisse zeigen¹.

Die Energie H_0 einer Teilstrecke $0 \leq x \leq l_0$ einer eingespannten Saite der Länge l ist entsprechend § 67, (3) gegeben durch

$$(1) \quad H_0 = \frac{1}{2} \int_0^{l_0} (u_t^2 + u_x^2) dx;$$

setzen wir für u den Ausdruck § 67, (4) ein, so wird mit § 67, (7):

$$(2) \quad \left\{ \begin{aligned} H_0 &= \frac{1}{2} \int_0^{l_0} \sum_{j,k=1}^{\infty} \{ \dot{q}_j \dot{q}_k \sin 2\pi \nu_j x \sin 2\pi \nu_k x \\ &\quad + q_j q_k (2\pi)^2 \nu_j \nu_k \cos 2\pi \nu_j x \cos 2\pi \nu_k x \} dx \\ &= \frac{2}{l^2} \int_0^{l_0} \sum_{j,k=1}^{\infty} \{ p_j p_k \sin 2\pi \nu_j x \sin 2\pi \nu_k x \\ &\quad + \pi^2 l^2 \nu_j \nu_k q_j q_k \cos 2\pi \nu_j x \cos 2\pi \nu_k x \} dx; \nu_j = \frac{j}{2l}. \end{aligned} \right.$$

Nehmen wir nun in der Summe (2) nur diejenigen Summanden, in denen $j = k$ ist, so erhalten wir unter der ausdrücklichen Voraussetzung, daß alle in Betracht fallenden *Wellenlängen klein gegen l_0* sind, gerade den *Zeitmittelwert* $\bar{H}_0 = \frac{l_0}{l} H$ von H_0 . Man

sieht daraus: Die Differenz

$$(3) \quad \Delta = H_0 - \bar{H}_0$$

ist (unter dieser Voraussetzung) gegeben durch den Ausdruck (2) *unter Beschränkung auf Summanden mit $j \neq k$* . Durch Ausführung der Integration erhalten wir dann:

$$(4) \quad \Delta = \frac{1}{l^2} \sum_{j \neq k} \{ p_j p_k \kappa_{jk} + 4\pi^2 \mu^2 \nu_j \nu_k q_j q_k \kappa'_{jk} \},$$

¹ Die Rechnung wurde für die *klassische* Theorie ausgeführt (am Beispiel der elektromagnetischen Wellen) von H. A. LORENTZ: *Théories statistiques en Thermodynamique*. Leipzig 1916.

wo nach § 67, (7) die Bezeichnung $\mu = \frac{l}{2}$ gebraucht ist und ferner

$$(5) \quad \begin{aligned} \varkappa_{jk} &= \varkappa_{kj} = \frac{\sin 2\pi(\nu_j - \nu_k)l_0}{2\pi(\nu_j - \nu_k)} - \frac{\sin 2\pi(\nu_j + \nu_k)l_0}{2\pi(\nu_j + \nu_k)}, \\ \varkappa'_{jk} &= \varkappa'_{kj} = \frac{\sin 2\pi(\nu_j - \nu_k)l_0}{2\pi(\nu_j - \nu_k)} + \frac{\sin 2\pi(\nu_j + \nu_k)l_0}{2\pi(\nu_j + \nu_k)}. \end{aligned}$$

Zerlegen wir Δ entsprechend den beiden Summanden in (4) in $\Delta = \Delta_1 + \Delta_2$, so bekommen wir

$$(6) \quad \Delta^2 = \Delta_1^2 + \Delta_2^2 + \Delta_1 \Delta_2 + \Delta_2 \Delta_1$$

mit

$$(7) \quad \left\{ \begin{aligned} \Delta_1^2 + \Delta_2^2 &= \frac{1}{l^4} \sum_{j \neq k} \sum_{i \neq \kappa} \{ p_j p_k p_i p_\kappa \varkappa_{jk} \varkappa_{i\kappa} \\ &\quad + (4\pi^2 \mu^2)^2 q_j q_k q_i q_\kappa \nu_j \nu_k \nu_i \nu_\kappa \varkappa_{jk} \varkappa_{i\kappa} \}; \\ \Delta_1 \Delta_2 + \Delta_2 \Delta_1 &= \frac{4\pi^2 \mu^2}{l^4} \sum_{j \neq k} \sum_{i \neq \kappa} \{ p_j p_k q_i q_\kappa \nu_i \nu_\kappa \varkappa_{jk} \varkappa'_{i\kappa} \\ &\quad + q_j q_k p_i p_\kappa \nu_j \nu_k \varkappa'_{jk} \varkappa_{i\kappa} \}. \end{aligned} \right.$$

Nun ist der *Zeitmittelwert* von $p_j p_k p_i p_\kappa$ mit $j \neq k, i \neq \kappa$ nach unseren quantenmechanischen Matrizenformeln (ebenso wie in der klassischen Theorie) gleich

$$\overline{p_j p_k p_i p_\kappa} = (\delta_{ji} \delta_{k\kappa} + \delta_{j\kappa} \delta_{ki}) \overline{p_j^2} \cdot \overline{p_k^2};$$

und entsprechend verhalten sich die andern in (7) auftretenden Glieder. Denn es ist auf Grund der Formeln § 15, (7) für die Verschmelzung $\overline{p_j p_i} = \overline{p_j} \cdot \overline{p_i}$, wenn $j \neq i$; beim harmonischen Oszillator ist aber nach § 23 stets $\overline{p_j} = 0$.

Es wird deshalb

$$(8) \quad \left\{ \begin{aligned} \overline{\Delta_1^2 + \Delta_2^2} &= \frac{2}{l^4} \sum_{j \neq k} \{ \overline{p_j^2} \overline{p_k^2} \varkappa_{jk}^2 + (4\pi^2 \mu^2)^2 \overline{q_j^2} \overline{q_k^2} \nu_j^2 \nu_k^2 \varkappa_{jk}^2 \}; \\ \overline{\Delta_1 \Delta_2 + \Delta_2 \Delta_1} &= \frac{8\pi^2 \mu^2}{l^4} \sum_{j \neq k} \{ \overline{p_j q_j} \cdot \overline{p_k q_k} \\ &\quad + \overline{q_j p_j} \cdot \overline{q_k p_k} \} \nu_j \nu_k \varkappa_{jk} \varkappa'_{jk}. \end{aligned} \right.$$

In allen Fällen nun, wo ein Impuls p proportional der Geschwindigkeit q ist, folgt aus der Beziehung $p q - q p = \frac{\hbar}{2\pi i}$, daß

$$(9) \quad \overline{p q} = -\overline{q p} = \frac{1}{2} \cdot \frac{\hbar}{2\pi i}$$

ist. Denn es wird dann

$$\begin{aligned} (p q)(n n) &= C \sum_k v(n k) q(n k) q(k n) \\ &= -C \sum_k q(n k) v(k n) q(k n) = -(q p)(n n). \end{aligned}$$

Ferner ist

$$(10) \quad \overline{p_k^2} = 4\pi^2 \mu^2 v_k^2 \overline{q_k^2}$$

[vgl. § 67, (7)]. Also wird nach (8):

$$(11) \quad \overline{\Delta^2} = \frac{4}{l^4} \sum_{j \neq k} \left\{ \frac{\overline{p_j^2} \overline{p_k^2} \overline{\varkappa_{jk}^2} + \overline{\varkappa'_{jk}{}^2}}{2} - \frac{1}{4} \hbar^2 \mu^2 v_j v_k \overline{\varkappa_{jk} \varkappa'_{jk}} \right\}.$$

In (11) ist die Diagonalmatrix $\overline{\Delta^2}$ als Funktion der verschiedenen Diagonalmatrizen $\overline{p_k^2}$ dargestellt. Jedes einzelne, einem bestimmten Wertsystem n_1, n_2, n_3, \dots der Quantenzahlen aller Eigenschwingungen entsprechende *Diagonalelement* $\overline{\Delta^2}(n n)$ von $\overline{\Delta^2}$ ist also dieselbe Funktion der entsprechenden *Diagonalelemente* in den $\overline{p_k^2}$. Diese Diagonalelemente sind aber gleich

$$(12) \quad \overline{p_k^2}(n n) = \mu \hbar v_k \left(n_k + \frac{1}{2} \right) = \mu \left(\varepsilon_k + \frac{\hbar v_k}{2} \right),$$

wo ε_k den Energieüberschuß des k -ten Oszillators über die Nullpunktsenergie bezeichnet. Es wird also nach (11):

$$(13) \quad \left[\begin{aligned} \overline{\Delta^2}(n n) &= \frac{4\mu^2}{l^4} \sum_{j \neq k} \left\{ \left(\varepsilon_j + \frac{\hbar v_j}{2} \right) \left(\varepsilon_k + \frac{\hbar v_k}{2} \right) \frac{\overline{\varkappa_{jk}^2} + \overline{\varkappa'_{jk}{}^2}}{2} \right. \\ &\quad \left. - \frac{\hbar^2}{4} v_j v_k \overline{\varkappa_{jk} \varkappa'_{jk}} \right\} \\ &= \frac{4\mu^2}{l^4} \sum_{j \neq k} \hbar^2 v_j v_k \left\{ \left(n_j + \frac{1}{2} \right) \left(n_k + \frac{1}{2} \right) \frac{\overline{\varkappa_{jk}^2} + \overline{\varkappa'_{jk}{}^2}}{2} \right. \\ &\quad \left. - \frac{1}{4} \overline{\varkappa_{jk} \varkappa'_{jk}} \right\}. \end{aligned} \right.$$

Lassen wir nun die Saitenlänge l sehr groß gegen l_0 werden, also $\frac{l_0}{l} \ll 1$, so rücken die Argumente $\nu_k l_0 = \frac{k}{2} \cdot \frac{l_0}{l}$ immer enger zusammen, und die Summen (13) können durch Integrale ersetzt werden. Dabei haben wir für die ε_j bzw. n_j ihre Mittelwerte ε_ν , n_ν in den infinitesimalen Frequenzintervallen $d\nu$ zu setzen:

$$(14) \quad \left\{ \begin{aligned} \overline{\Delta^2}(n n) &= 4 \int_0^\infty \int_0^\infty d\nu \cdot d\nu' \left\{ \left(\varepsilon_\nu + \frac{h\nu}{2} \right) \left(\varepsilon_{\nu'} + \frac{h\nu'}{2} \right) \frac{\varkappa_{\nu\nu'}^2 + \varkappa'_{\nu\nu'}^2}{2} \right. \\ &\quad \left. - \frac{h^2}{4} \nu \nu' \varkappa_{\nu\nu'} \varkappa'_{\nu\nu'} \right\} \\ &= 4 \int_0^\infty \int_0^\infty d\nu \cdot d\nu' h^2 \nu \nu' \left\{ \left(n_\nu + \frac{1}{2} \right) \left(n_{\nu'} + \frac{1}{2} \right) \right. \\ &\quad \left. - \frac{\varkappa_{\nu\nu'}^2 + \varkappa'_{\nu\nu'}^2}{2} - \frac{1}{4} \varkappa_{\nu\nu'} \varkappa'_{\nu\nu'} \right\}. \end{aligned} \right.$$

Wir wollen nun auch l_0 sehr groß annehmen (bei ungeändertem Verhältnis $\frac{l_0}{l}$), was physikalisch bedeutet, daß wir die in Betracht kommenden *Wellenlängen* als *klein gegen l_0* voraussetzen. Wir können dann Gebrauch machen von der Beziehung

$$(15) \quad \lim_{l_0 \rightarrow \infty} \frac{1}{l_0} \int_{-\Omega}^{\Omega'} \frac{\sin^2 2\pi \nu l_0}{(2\pi \nu)^2} f(\nu) d\nu = \frac{1}{2} f(0) \text{ für } \Omega, \Omega' > 0.$$

Man sieht, daß jetzt nur die Anteile

$$\frac{\sin 2\pi(\nu - \nu') l_0}{2\pi(\nu - \nu')}$$

in $\varkappa_{\nu\nu'}$ und $\varkappa'_{\nu\nu'}$ in Betracht kommen; und zwar wird

$$(16) \quad \left\{ \begin{aligned} \overline{\Delta^2}(n n) &= 2 l_0 \int_0^\infty d\nu \left\{ \left(\varepsilon_\nu + \frac{h\nu}{2} \right)^2 - \left(\frac{h\nu}{2} \right)^2 \right\} \\ &= 2 l_0 \int_0^\infty d\nu \cdot (h\nu)^2 \left\{ \left(n_\nu + \frac{1}{2} \right)^2 - \frac{1}{4} \right\}; \end{aligned} \right.$$

oder

$$(16') \quad \overline{\Delta^2}(n n) = 2 l_0 \int_0^\infty d\nu \cdot \varepsilon_\nu (\varepsilon_\nu + \hbar \nu) = 2 l_0 \int_0^\infty d\nu \cdot (\hbar \nu)^2 \cdot n_\nu (n_\nu + 1).$$

Nach § 68, (5) ist hier

$$(17) \quad 2 l_0 = Z_\nu^0.$$

Die in § 72 betrachtete Größe $\overline{(\Delta E_0)^2}$ ist der auf $d\nu$ bezügliche Anteil von $\overline{\Delta^2}(n n)$; also wird

$$(18) \quad \begin{aligned} \overline{(\Delta E_0)^2} &= Z_\nu^0 \cdot \varepsilon_\nu (\varepsilon_\nu + \hbar \nu) \\ &= Z_\nu^0 \cdot (\hbar \nu)^2 n_\nu (n_\nu + 1). \end{aligned}$$

Andererseits ist

$$(19) \quad \overline{E_0} = Z_\nu^0 \cdot \varepsilon_\nu = Z_\nu^0 \cdot \hbar \nu \cdot n_\nu.$$

Vergleich von (18) und (19) gibt nun in der Tat die EINSTEINSche Formel

$$(20) \quad \overline{(\Delta E_0)^2} = \hbar \nu \overline{E_0} + \frac{\overline{E_0^2}}{Z_\nu^0}$$

in Übereinstimmung mit § 72, (20).

Der physikalisch anschauliche Sinn des Umstandes, daß bei Beschränkung auf Wellenlängen $\lambda \ll l_0$ gemäß (15) die Doppelintegrale (14) in einfache Integrale (16) übergehen, ist offenbar folgender: Die Schwankungen der Teilenergie H_0 kommen wellentheoretisch zustande durch *Interferenzschwebungen* zwischen den verschiedenen Frequenzpaaren ν_j, ν_k . Dabei können aber nur solche Schwebungswellen eine Schwankung der Gesamtenergie der Strecke l_0 bewirken, welche selbst eine Wellenlänge von mindestens der Größenordnung l_0 besitzen.

Wird also l_0 sehr groß gegenüber den Wellenlängen λ der betrachteten Eigenschwingungen, so liefern nur noch die Schwebungen mit sehr großen Wellenlängen gegenüber λ eine Wirkung; und es kommt demgemäß lediglich die Interferenz jeder Frequenz ν mit den ihr unmittelbar benachbarten Frequenzen in Betracht.

Dieselbe Rechnung, die wir hier für die *quantenmechanische Wellentheorie* ausgeführt haben, läßt sich auch in der *klassischen Wellentheorie* durchführen. Dabei sind wörtlich die obigen Überlegungen zu wiederholen, abgesehen davon, daß wir erstens an

Stelle von (9) die Gleichung $\overline{p_k q_k} = 0$ zu benutzen haben (womit $\overline{\Delta_1 \Delta_2 + \Delta_2 \Delta_1} = 0$ wird), und zweitens statt (12) die Beziehung $p_k^2 = \mu \varepsilon_k$ (ohne die zusätzliche Nullpunktenergie $\frac{1}{2} \hbar \nu_k$). Indem wir weiterhin je in der *oberen* Gleichung (13), (14), (16), (16') die entsprechenden Vereinfachungen ausführen, bekommen wir statt (18):

$$(18') \quad \overline{(\Delta E_0)^2} = Z_\nu^0 \cdot \varepsilon_\nu^2,$$

was mit (19) zu

$$(20') \quad \overline{(\Delta E_0)^2} = \frac{\overline{E_0^2}}{Z_\nu^0}$$

führt. Das ist aber die Formel (20') aus § 72.

Wir sehen also: Die *klassische Wellentheorie* führt — ohne irgendwelche besondere Hypothesen — notwendig zur Formel (20'), § 72¹. Sie ist also *unverträglich* mit den aus dem PLANCKschen Gesetz fließenden thermodynamischen Folgerungen und verträgt sich nur mit dem RAYLEIGH-JEANSschen Gesetz². Die Betrachtung des WIENSchen Grenzfalles der PLANCKschen Strahlungsformel legt nach EINSTEIN die Vermutung nahe, daß

¹ Dieser Umstand ist früher in langwierigen Diskussionen durch sehr komplizierte und schwierige Überlegungen teils bestritten, teils zu beweisen versucht worden. (Vgl. A. EINSTEIN u. L. HOPF: Ann. Phys. Bd. 33, S. 1096. 1910. M. v. LAUE: Ann. Phys. Bd. 47, S. 853. 1915. A. EINSTEIN: Ann. Phys. Bd. 47, S. 879. 1915. M. v. LAUE: Ann. Phys. Bd. 48, S. 668. 1915. M. PLANCK: Ann. Phys. Bd. 73, S. 272. 1924.) Der Grund dieser Unsicherheit war der, daß LORENTZ (vgl. die Fußnote S. 393) wellentheoretisch einen *statistischen* Mittelwert berechnet hatte vermittle einer *Phasenmittelung*, deren Zuverlässigkeit hernach angezweifelt wurde. Diese Diskussionen sind jedoch gegenstandslos geworden durch die Feststellung (BORN-HEISENBERG-JORDAN a. a. O.), daß man den Mittelwert hypothesenfrei als *Zeitmittelwert* berechnen kann.

² Man kann sogar aus (20') das RAYLEIGH-JEANSsche Gesetz ableiten, indem man die thermodynamische Schwankungsformel (19), § 72, benutzt:

$$\overline{(\Delta E_0)^2} = k T^2 \cdot \frac{d \overline{E_0}}{dT} = \frac{\overline{E_0^2}}{Z_\nu^0}$$

ergibt durch Integration die RAYLEIGH-JEANSsche Formel

$$\overline{E_0} = Z_\nu^0 \cdot k T;$$

den *Gleichverteilungssatz* braucht man also bei dieser Ableitung *nicht*.

die Wellenvorstellung durch die Korpuskularvorstellung ersetzt bzw. ergänzt werden müsse. Die Quantenmechanik aber macht es möglich, *ohne* Verzicht auf die Wellenvorstellung die Übereinstimmung mit dem PLANCKSchen Gesetz und allen seinen Folgerungen wiederherzustellen: Es genügt, die Modellvorstellung der klassischen Wellentheorie mit der exakten quantentheoretischen Kinetik und Mechanik durchzuführen. Die charakteristischen Lichtquanteneffekte ergeben sich dann von selbst, ohne Hinzuziehung neuer Hypothesen, als notwendige Folgerungen der Wellentheorie.

Gezeigt ist das hier freilich zunächst für das *zweite* EINSTEINSche Schwingungsgesetz. Daß auch das *erste* EINSTEINSche Schwingungsgesetz sich aus der quantenmechanischen Wellentheorie richtig ergibt, werden wir erst an späterer Stelle (Teil II dieses Buches) beweisen.

Daß es möglich ist, mit Hilfe der Quantenmechanik die Schwingungsprobleme auf Grund der Wellenvorstellung befriedigend zu erledigen, kann nicht überraschend erscheinen, wenn man bedenkt, daß in der Dynamik der Kristallgitter ganz dieselben Probleme auftreten: Die *klassische* Punktmechanik, angewandt auf die Massenpunkte eines Kristallgitters, führt durch Zeitmittelbildung über die Eigenschwingungen des Gitters zur Formel (20') für das Schwingungsquadrat der Energie in einem kleinen Teilvolumen des Kristalls. Die Thermodynamik dagegen verlangt, wie schon in § 72 hervorgehoben wurde, auch für den Kristall die Gültigkeit der Formel (20) statt (20')¹. Wenn also die Quantenmechanik für die *Mechanik der Massenpunkte* allgemein zu richtigen Ergebnissen führt, so muß insbesondere die Quantelung der Eigenschwingungen für das Kristallgitter die richtigen Schwingungsformeln liefern; und dann gilt dasselbe für den Hohlraum, der in diesen Eigenschaften dem Kristallgitter durchaus analog ist.

Unter Benutzung der Größe n_ν können wir (20') in die Form

$$\overline{(\Delta E_0)^2} = (h \nu)^2 \cdot Z_\nu^0 n_\nu^2$$

setzen, analog zur Formel (18):

$$\overline{(\Delta E_0)^2} = (h \nu)^2 \cdot Z_\nu^0 \cdot n_\nu (n_\nu + 1).$$

¹ Man muß also auch den Schwingungsquanten des Kristallgitters im EINSTEINSchen Sinne einen korpuskularen Charakter zuschreiben.

Die quantenmechanische Formel unterscheidet sich also von der klassischen durch das Auftreten von n_v , $(n_v + 1)$ statt n_v^2 ähnlich, wie sich etwa die auf den Drehimpuls eines Atoms bezüglichen quantenmechanischen Formeln von den klassischen durch das Auftreten von $j(j + 1)$ an Stelle von j^2 unterscheiden.

§ 74. Absorption und Emission von Strahlung durch Atome¹.

Aus der bei diesen Untersuchungen der Hohlraumstrahlung benützten Darstellung des elektromagnetischen Feldes durch gequantelte Eigenschwingungen erhalten wir auch die Möglichkeit, die *Wechselwirkung zwischen Materie und Strahlung* in exakter Weise zu behandeln. Wir denken uns in den gequantelten Strahlungshohlraum ein *Atom* hineingesetzt. Der Kern des Atoms sei unendlich schwer und liege fest an einem Orte $x = x_0$, $y = y_0$, $z = z_0$. Richten wir unsere Aufmerksamkeit zunächst auf diejenigen Eigenschwingungen, deren Wellenlängen groß gegen die mittleren Abstände der Atomelektronen vom Kern sind, so können wir der Berechnung der Wechselwirkung von Atom und Strahlungsfeld die Werte der Feldstärken *am Ort* $x_0 y_0 z_0$ *des Kerns* zugrunde legen. Ist also H_S die Energie des Hohlraums allein, H_A die des isolierten Atoms, so haben wir für das Gesamtsystem Hohlraum plus Atom mit Einschluß ihrer Wechselwirkung eine HAMILTON-Funktion

$$(1) \quad H = H_S + H_A + H',$$

wobei

$$(1') \quad H' = -(\mathfrak{P} \mathfrak{E}) = \frac{1}{c} (\mathfrak{P} \mathfrak{H})$$

ist (vgl. § 46 (4), § 67 (13)); \mathfrak{A} ist am Ort x_0, y_0, z_0 zu nehmen.

Als *ungestörtes* System betrachten wir die Verschmelzung der beiden Systeme: Atom und Strahlungsfeld. Als Funktion der Koordinaten dieser Verschmelzung enthält die Störungsfunktion die *Zeit nicht* explizit. Wir können daher die Formel § 45 (10') und die Betrachtungen des § 65, S. 361f. anwenden.

Nach § 65, (19) ist die Übergangswahrscheinlichkeit des Atoms vom Zustand n' in den Zustand m' (bei gegebenem Anfangs-

¹ Die Überlegungen dieses und des folgenden Paragraphen rühren von P. A. M. DIRAC her (Proc. Roy. Soc. Bd. 114, S. 243 u. 710. 1927).

zustand n'' der Strahlung) in der Zeiteinheit

$$(2) \quad w_{n' m'} = \frac{4\pi^2}{h^2} \Phi_{n', n''; m'} (\nu_{m' n'}),$$

wobei

$$\Phi_{n', n''; m'} (\nu) \cdot \Delta\nu = \sum_{\substack{m'' \\ \nu < \nu_{n' m''} < \nu + \Delta\nu}} |H'_{n', n''; m', m''}|^2.$$

In diesem Fall ist ein Zustand n'' des Hohlraums gekennzeichnet durch die Besetzungszahlen n''_k sämtlicher Oszillatoren. Ein Matricelement von

$$H' = \frac{1}{c} \sum_s \left(\dot{q}^s \left(\mathfrak{P} \mathfrak{Q}^s \right) + \dot{q}^s \left(\mathfrak{P} \mathfrak{Q}^s \right) \right)$$

wird dann nach § 71:

$$(3) \quad H'_{n', n_1'' \dots n_s'' \dots; m', m_1'' \dots m_s'' \dots} = \frac{1}{c} \left(\mathfrak{P}_{n' m'} \mathfrak{Q}^s \right) \dot{q}^s_{n_1'' \dots n_s'' \dots; m_1'' \dots m_s'' \dots}$$

Hierbei ist die Summe über s fortzulassen, da bei jedem Indexsystem $n''_1 \dots n''_k \dots, m''_1 \dots m''_k \dots$ höchstens eins der q von Null verschieden sein kann.

Dann ist

$$|H'_{n', n''; m', m''}|^2 = \frac{1}{c^2} \left| \mathfrak{P}_{n' m'} \mathfrak{Q}^s \right|^2 \left| \dot{q}^s_{n_1'' \dots n_s'' \dots; m_1'' \dots m_s'' \dots} \right|^2,$$

und durch Einsetzen von q aus § 71 ergibt sich mit dem in § 15, S. 75 eingeführten Symbol Δ_s :

$$(4) \quad \left\{ \begin{aligned} |H'_{n', n''; m', m''}|^2 &= \frac{16\pi h \nu_s}{V} n''_s \left| \mathfrak{P}_{n' m'} \mathfrak{Q}^s \right|^2 \cdot \Delta_s \\ &\quad \text{für } m''_s = n''_s - 1, \\ &\quad \text{d. h. } \nu_{n' m''} = \nu_s > 0; \\ &= \frac{16\pi h \nu_s}{V} (n''_s + 1) \left| \mathfrak{P}_{n' m'} \mathfrak{Q}^s \right|^2 \cdot \Delta_s \\ &\quad \text{für } m''_s = n''_s + 1, \\ &\quad \text{d. h. } \nu_{n' m''} = -\nu_s < 0; \\ &= 0 \text{ sonst.} \end{aligned} \right.$$

$\Phi_{n', n''; m'} (\nu)$ wird also verschieden sein, je nachdem

$$\nu \gtrless 0.$$

Wir betrachten zuerst den Fall $\nu > 0$. Dann wird

$$\Phi_{n', m'; n''}(\nu) \Delta \nu = \frac{16\pi h}{V} \sum_{\substack{\nu < \nu_s < \nu + \Delta \nu \\ s_1 s_2 s_3}} \nu_s \left(n_s^{(1)} | \mathfrak{P}_{n' m'} \mathfrak{A}_s^{(1)} |^2 + n_s^{(2)} | \mathfrak{P}_{n' m'} \mathfrak{A}_s^{(2)} |^2 \right).$$

Wir nehmen nun an, daß alle Eigenschwingungen gleicher Frequenz etwa gleiche Besetzungszahlen haben.

Führt man dann die Funktion $\varrho(\nu)$ ein durch:

$$(5) \quad \sum_{\nu < \nu_s < \nu + \Delta \nu} h \nu_s (n_s^{(1)} + n_s^{(2)}) = V \varrho(\nu) \Delta \nu,$$

so ist $\varrho(\nu)$ die gesamte monochromatische Strahlungsdichte (entsprechend Anhang I) und $\varrho(\nu) \frac{d\omega_s}{8\pi}$ ist die monochromatische Strahlungsdichte der Eigenschwingungen, welche gekennzeichnet sind durch die Vektoren \mathfrak{s} und \mathfrak{e} (vgl. § 67, S. 371), wobei \mathfrak{s} in einem Winkelement $d\omega_s$ liegt. (Die Strahlungsdichte für \mathfrak{e} ist, nach Voraussetzung, gleich groß.) Damit erhält man

$$(6) \quad \Phi(\nu) = 2\varrho(\nu) \int \left(| \mathfrak{P}_{n' m'} \mathfrak{A}_s^{(1)} |^2 + | \mathfrak{P}_{n' m'} \mathfrak{A}_s^{(2)} |^2 \right) d\omega_s.$$

Mittelt man noch über die unbekannte Lage x_0, y_0, z_0 des Atoms im Hohlraum¹, so erhält man

$$\begin{aligned} \Phi(\nu) &= \frac{\varrho(\nu)}{4} \int \left\{ | X_{n' m'} e_x^{(1)} |^2 + | X_{n' m'} e_x^{(2)} |^2 + | Y_{n' m'} e_y^{(1)} |^2 \right. \\ &\quad \left. + | Y_{n' m'} e_y^{(2)} |^2 + | Z_{n' m'} e_z^{(1)} |^2 + | Z_{n' m'} e_z^{(2)} |^2 \right\} d\omega_s \\ &= \frac{2\pi\varrho(\nu)}{3} | \mathfrak{P}_{n' m'} |^2. \end{aligned}$$

Daher ergibt sich als Wert der Übergangswahrscheinlichkeit $w_{n' m'}$ im Falle $\nu_{n' m'} = -\nu_{n'' m''} < 0$, d. h. der *Übergangswahrscheinlichkeit* (für die Zeiteinheit) *durch Absorption*:

$$(7) \quad \begin{aligned} w_{n' m'}^{(a)} &= \frac{8\pi^3}{3h^2} | \mathfrak{P}_{n' m'} |^2 \varrho(\nu_{n' m'}) \\ &= B_{n' m'} \varrho(\nu_{n' m'}), \end{aligned}$$

¹ Man beachte, daß $\int_V \mathfrak{A}_x^{(1)} \mathfrak{A}_y^{(1)} dx dy dz = 0$ und $\int_V | \mathfrak{A}_x^{(1)} |^2 dx dy dz = \frac{V}{8} | e_x^{(1)} |^2$ ist.

wobei der Faktor

$$(8) \quad B_{n' m'} = \frac{8\pi^3}{3h^2} |\mathfrak{P}_{n' m'}| ^2$$

in völliger Übereinstimmung ist mit dem schon in § 2, (10) korrespondenzmäßig abgeleiteten EINSTEINSchen Absorptionskoeffizienten [s. auch § 63, (16)].

Es bleibt noch die *Emission* zu untersuchen, nämlich die Übergangswahrscheinlichkeiten von n' nach m' bei $\nu_{n' m'} = -\nu_{n'' m''} > 0$. In diesen Fällen ist

$$\begin{aligned} \Phi_{n', n''; m'}(\nu) \Delta \nu &= \frac{16\pi h}{V} \sum_{\substack{\nu < \nu_s < \nu + \Delta \nu \\ s_1 s_2 s_3}} \nu_s \left((n_s^{(1)} + 1) |\mathfrak{P}_{n' m'}^{(1)} \mathfrak{U}^s|^2 \right. \\ &\quad \left. + (n_s^{(2)} + 1) |\mathfrak{P}_{n' m'}^{(2)} \mathfrak{U}^s|^2 \right) \end{aligned}$$

Außer der in (5) definierten Funktion $\varrho(\nu)$ müssen wir hier noch eine weitere Funktion $\sigma(\nu)$ einführen:

$$(9) \quad \sum_{\nu < \nu_s < \nu + \Delta \nu} h \nu_s = V \sigma(\nu) \Delta \nu.$$

Dann wird

$$(6a) \quad \Phi(\nu) = 2(\varrho(\nu) + \sigma(\nu)) \int \left(|\mathfrak{P}_{n' m'}^{(1)} \mathfrak{U}^s|^2 + |\mathfrak{P}_{n' m'}^{(2)} \mathfrak{U}^s|^2 \right) d\omega_s.$$

$V \cdot \sigma(\nu) \cdot \Delta \nu$ ist gerade die mit $h\nu$ multiplizierte Anzahl der Eigenschwingungen, für die ν_s im Intervall ν bis $\nu + \Delta \nu$ liegt. Nun wurde aber in § 68 die asymptotische Anzahl der Eigenschwingungen im Frequenzintervall ν bis $\nu + \Delta \nu$ berechnet zu

$$\frac{dN(\nu, V)}{d\nu} \cdot \Delta \nu = \frac{8\pi \nu^2}{c^3} V \Delta \nu,$$

und daher wird

$$(10) \quad \sigma(\nu) = \frac{8\pi h \nu^3}{c^3}.$$

Dann erhält man durch völlig analoge Rechnung wie bei der Absorption die *Wahrscheinlichkeit der Emission*

$$\begin{aligned} (7a) \quad w_{n' m'}^{(e)} &= \frac{8\pi^3}{3h^2} |\mathfrak{P}_{n' m'}|^2 (\varrho(\nu_{n' m'}) + \sigma(\nu_{n' m'})) \\ &= B_{n' m'} \varrho(\nu_{n' m'}) + A_{n' m'}, \end{aligned}$$

wobei $B_{n' m'}$, dieselbe Größe wie in (7), den Koeffizienten der erzwungenen Emission bedeutet.

$$(8a) \quad A_{n' m'} = \frac{64\pi^4 \nu_{n' m'}^3}{3\hbar c^3} |\mathfrak{P}_{n' m'}|^2$$

ist derjenige Bestandteil der Übergangswahrscheinlichkeit, der noch übrigbleibt bei verschwindender Strahlungsdichte, also die Wahrscheinlichkeit der *spontanen Emission*, in voller Übereinstimmung mit der korrespondenzmäßigen Formel § 2, (9).

Damit sind wir zum Ausgangspunkt dieses ganzen Buches zurückgekehrt. Wir erinnern uns, daß HEISENBERG bei der Grundlegung seiner Theorie die Annahme machte, daß außer den Energiewerten der stationären Zustände die Elemente der Momentenmatrix \mathfrak{P}_{nm} beobachtbare Größen seien, indem sie die emittierte Energie beim Übergang $n \rightarrow m$ gemäß der Formel (8a) bestimmten. Was dort Hypothese war, ist nunmehr nach dem Ausbau der Theorie zu einem ableitbaren Resultat geworden.

§ 75. Zerstreuung und Dispersion. Wir wollen nun in ähnlicher Weise wie im vorigen Paragraphen zeigen, daß auch die Streuung und damit die Dispersion des Lichts mit Hilfe der gequantelten Hohlraumstrahlung abgeleitet werden kann.

Nach § 45, (10') bestimmt sich ein Element der Matrix U in zweiter Näherung durch

$$(1) \quad U_{n' n''; m' m''}^{(2)} = e^{-\frac{\tilde{W}_n t}{\hbar}} \sum_{k' k''} \frac{H'_{n', n''; k', k''} H_{k', k''; m', m''}}{\hbar (\nu_{k' m'} + \nu_{n'' m''})} \left(\frac{e^{2\pi i (\nu_{n' m'} + \nu_{n'' m''}) t} - 1}{\hbar (\nu_{n' m'} + \nu_{n'' m''})} - \frac{e^{2\pi i (\nu_{n' k'} + \nu_{n'' k''}) t} - 1}{\hbar (\nu_{n' k'} + \nu_{n'' k''})} \right).$$

Für H' ist der Ausdruck § 74, (1') einzusetzen. Es ist

$$\mathfrak{A} = \sum_s \left(q^s \mathfrak{A}^s + q^s \mathfrak{A}^s \right),$$

und wenn wir hier zur Abkürzung die q durchnumerieren, also schreiben

$$\mathfrak{A} = \sum_\alpha q^\alpha \mathfrak{A}^\alpha,$$

so wird

$$U_{n', n''; m', m''}^{(2)} = \frac{1}{c^2} e^{-\frac{W_n}{\kappa} t} \sum_{\alpha\beta} \sum_{k''} \dot{q}_{n'', k''}^\alpha \dot{q}_{k'', m''}^\beta \sum_{k'} \left\{ \frac{(\mathfrak{P}_{n' k'} \mathfrak{A}^\alpha) (\mathfrak{P}_{k' m''} \mathfrak{A}^\beta)}{h (\nu_{k' m'} + \nu_{k'' m''})} \right. \\ \left. \left(\frac{1 - e^{2\pi i (\nu_{n' m'} + \nu_{n'' m''}) t}}{h (\nu_{n' m'} + \nu_{n'' m''})} - \frac{1 - e^{2\pi i (\nu_{k' m'} + \nu_{k'' m''}) t}}{h (\nu_{k' m'} + \nu_{k'' m''})} \right) \right\}.$$

Dabei ist zu beachten, daß bei einem bestimmten Indexsystem $m'' = m''_1 \dots m''_\alpha \dots m''_\beta \dots$, $k'' = k''_1 \dots k''_\alpha \dots k''_\beta \dots$, wenn $m'' \neq n''$, höchstens ein Element eines der q^α und der q^β von 0 verschieden sein kann, nämlich nur dann, wenn $k''_\alpha = n''_\alpha \pm 1$, $k'' = n''_\gamma$ für $\gamma \neq \alpha$; $m''_\beta = k''_\beta \pm 1$ und $k''_\delta = m''_\delta$ für $\delta \neq \beta$.

$U_{n', n''; m', m''}^{(2)}$ kann also nur dann überhaupt von 0 verschieden sein, wenn $m''_\alpha = n''_\alpha \pm 1$, $m''_\beta = n''_\beta \pm 1$, $m''_\gamma = n''_\gamma$ für $\gamma \neq \alpha, \beta$. Das bedeutet, der Zustand m'' geht aus n'' hervor durch Absorption bzw. Emission eines Lichtquants ν_α und gleichzeitiger Absorption bzw. Emission eines Quants ν_β . Da aber für solche Übergänge,

wie wir sahen, die erste Näherung $U_{n', n''; m', m''}^{(1)} = 0$ ist, so ist die Wahrscheinlichkeit eines Übergangs $n' \rightarrow m'$, $n'' \rightarrow m''$

$$\Psi_{n', n''; m', m''}^{(2)} = |U_{n', n''; m', m''}^{(2)}|^2.$$

Uns interessiert hier speziell die Wahrscheinlichkeit dafür, daß bei einem Übergang $n' \rightarrow m'$ des Atoms ein Lichtquant einer Frequenz ν_α absorbiert und ein Lichtquant ν_β emittiert wird, wobei ν_α, ν_β nicht Eigenfrequenzen des Atoms sind. Wesentlichen Beitrag werden dabei nur solche Frequenzen ν_α, ν_β geben, für die einer der drei Nenner sehr klein ist, und wegen der Beschränkung der ν_α, ν_β auf Frequenzen ungleich $\nu_{n' k'}$ kommt von den drei Nennern nur der eine

$$\nu_{n' m'} + \nu_{n'' m''} = \nu_{n' m'} + \nu_\alpha - \nu_\beta$$

in Betracht. Für diese Fälle ist daher das zweite Glied unter der Summe über k' gegen das erste zu vernachlässigen, und es wird

$$\Psi_{n', n''; m', m''}^{(2)} = \frac{16 \cdot 16 \pi^2 h^2}{V^2} \cdot n''_\alpha (n''_\beta + 1) \nu_\alpha \nu_\beta \\ (2) \cdot \left| \sum_{k'} \frac{(\mathfrak{P}_{n' k'} \mathfrak{A}^\alpha) (\mathfrak{P}_{k' m''} \mathfrak{A}^\beta)}{h (\nu_{k' m'} - \nu_\beta)} + \frac{(\mathfrak{P}_{n' k'} \mathfrak{A}^\beta) (\mathfrak{P}_{k' m''} \mathfrak{A}^\alpha)}{h (\nu_{k' m'} + \nu_\alpha)} \right|^2 \\ \cdot \frac{4 \sin^2 \pi (\nu_{n' m'} + \nu_\alpha - \nu_\beta) t}{h^2 (\nu_{n' m'} + \nu_\alpha - \nu_\beta)^2}.$$

Beschränken wir uns jetzt auf den Fall, daß der Hohlraum im ungestörten Zustand nur Strahlung einer Eigenschwingung \mathfrak{A} bestimmter Frequenz ν_r enthält, d. h. daß alle $n''_\alpha = 0$ sind außer einem einzigen n''_r , so ist der einzig mögliche Prozeß der betrachteten Art derjenige, bei dem ein Quantum ν_r absorbiert und ein Quantum ν_β emittiert wird, das in der Gegend der Frequenz

$$(3) \quad \nu_s = \nu_r + \nu_{n' m'}$$

liegt¹.

Um die Wahrscheinlichkeit eines solchen Prozesses, d. h. der Aussendung eines solchen Lichtquants ν_s zu erhalten, müssen wir $\Psi_{n' n''; m' m''}$ über einen kleinen Frequenzbereich $\nu_s - \Delta \nu$ bis $\nu_s + \Delta \nu$ summieren.

Gehen wir analog § 65 von der Summe zum Integral über, so wird die Wahrscheinlichkeit der Emission von ν_s beim Sprung $n' \rightarrow m'$ in der Zeiteinheit

$$w(\nu_s) = \frac{4\pi^2}{h^2} \Phi(\nu_s),$$

wobei

$$\Phi_{n', n''; m' m''}(\nu) \Delta \nu = \sum_{\substack{m'' \\ \nu < \nu_\beta < \nu + \Delta \nu}} \frac{16 \cdot 16 \pi^2 h^2}{V^2} n''_r \nu_r \nu_\beta \left| \sum_{k'} \frac{(\mathfrak{P}_{n' k'} \mathfrak{A}^r) (\mathfrak{P}_{k' m'} \mathfrak{A}^\beta)}{h(\nu_{k' m'} - \nu_\beta)} + \frac{(\mathfrak{P}_{n' k'} \mathfrak{A}^\beta) (\mathfrak{P}_{k' m'} \mathfrak{A}^r)}{h(\nu_{k' m'} + \nu_\alpha)} \right|^2.$$

Da die Anzahl der Eigenschwingungen im betrachteten Frequenzintervall $\frac{8\pi^2 \nu^2}{c^3} V \Delta \nu$ ist, so ergibt sich, ganz wie in § 74

$$(4) \quad w(\nu_s) = \frac{16 \cdot 64 \cdot \pi^4 \nu_s^3 \nu_r}{c^3 V} \int d\omega_s \left| \sum_{k'} \frac{(\mathfrak{P}_{n' k'} \mathfrak{A}^r) (\mathfrak{P}_{k' m'} \mathfrak{A}^s)}{h(\nu_{k' n'} - \nu_r)} + \frac{(\mathfrak{P}_{n' k'} \mathfrak{A}^s) (\mathfrak{P}_{k' m'} \mathfrak{A}^r)}{h(\nu_{k' m'} + \nu_r)} \right|^2,$$

¹ Außer den betrachteten Prozessen kann im Fall $\nu_{n' m'} > \nu_r > 0$ noch ein anderer mit gleicher Größenordnung der Intensität auftreten, nämlich eine erzwungene Emission von ν_r zugleich mit einer Emission der Frequenz

$$\nu_t = \nu_{n' m'} - \nu_r.$$

Die Rechnungen sind ganz analog (s. M. GÖPPERT: Naturwiss. Bd. 17, S. 932. 1929).

wobei rechts außerdem noch über die verschiedenen Polarisationsrichtungen \mathfrak{A}^s und \mathfrak{A}^r zu summieren ist. Es ist zweckmäßig, an Stelle der Wahrscheinlichkeit der Emission der Frequenz ν_s zu berechnen, wie groß das Matrixelement \mathfrak{P}^s eines Atoms sein würde, das diese Frequenz mit der angegebenen Intensität spontan emittiert.

Durch Vergleich mit Formel (10) und (6a) aus § 74 ergibt sich

$$\int d\omega_s (|\mathfrak{P}^s \mathfrak{A}^s|^{(1)2} + |\mathfrak{P}^s \mathfrak{A}^r|^{(2)2}) = \frac{c^3 h}{64 \pi^3 \nu_s^3} w(\nu_s),$$

d. h.

$$|\mathfrak{P}^s \mathfrak{A}^s|^2 = \frac{16 \pi h n_r'' \nu_r}{V}$$

$$\cdot \left| \sum_{k'} \frac{(\mathfrak{P}_{n'k'} \mathfrak{A}^r) (\mathfrak{P}_{k'm'} \mathfrak{A}^s)}{h(\nu_{k'n'} - \nu_r)} + \frac{(\mathfrak{P}_{n'k'} \mathfrak{A}^s) (\mathfrak{P}_{k'm'} \mathfrak{A}^r)}{h(\nu_{k'm'} + \nu_r)} \right|^2$$

oder

$$(5) \quad |\mathfrak{P}^s| = \sqrt{\frac{16 \pi n_r'' h \nu_r}{V}} \left| \sum_{k'} \frac{(\mathfrak{P}_{n'k'} \mathfrak{A}^r) \mathfrak{P}_{k'm'}}{h(\nu_{k'n'} - \nu_r)} + \frac{\mathfrak{P}_{n'k'} (\mathfrak{P}_{k'm'} \mathfrak{A}^r)}{h(\nu_{k'm'} + \nu_r)} \right|.$$

Um dieses Ergebnis mit § 46ff. vergleichen zu können, müssen wir den Vektor \mathfrak{E} der mittleren eingestrahlten elektrischen Feldstärke an der Stelle $x_0 y_0 z_0$ des Atoms einführen.

Da

$$\mathfrak{E} = -\frac{1}{c} \dot{\mathfrak{A}} = -\frac{1}{c} \sum_{\alpha} \dot{q}_{\alpha} \mathfrak{A}^{\alpha},$$

so ist das Zeitmittel des Quadrats von \mathfrak{E}^r

$$(6) \quad \begin{aligned} 2|\bar{\mathfrak{E}}|^2 &= |\mathfrak{E}_{n'r'}^r|^2 = \frac{1}{c^2} |\dot{q}^r \mathfrak{A}^r|^2 = \frac{16 \pi h \nu_r}{V} (n_r'' + n_r'' + 1) |\mathfrak{A}^r|^2 \\ &= \frac{32 \pi}{V} \cdot h \nu_r \left(n_r'' + \frac{1}{2} \right) |\mathfrak{A}^r|^2. \end{aligned}$$

Da n_r'' groß ist, so ist $\frac{1}{2}$ dagegen zu vernachlässigen, und da außerdem \mathfrak{E} die Richtung von \mathfrak{A}^r hat, so wird

$$(7) \quad |\mathfrak{P}^s| = \left| \sum_{k'} \frac{(\mathfrak{P}_{n'k'} \mathfrak{E}) \mathfrak{P}_{k'm'}}{h(\nu_{k'n'} - \nu_r)} + \frac{\mathfrak{P}_{n'k'} (\mathfrak{P}_{k'm'} \mathfrak{E})}{h(\nu_{k'm'} + \nu_r)} \right|.$$

Die ausgestrahlte Frequenz war dabei nach (3)

$$\nu_s = \nu_r + \nu_{n'm'}.$$

Ist $\nu_{n' m'} < 0$, d. h. wird das Atom vom Zustand n' zum Zustand m' gehoben, so haben wir für $\nu_r > \nu_{m' n'}$ den STOKESSchen Fall des RAMAN-Effekts vor uns, die ausgestrahlte Frequenz

$$\nu_s = \nu_r - |\nu_{n' m'}|$$

ist kleiner als die eingestrahlte.

Ist $\nu_{n' m'} > 0$, d. h. sinkt das Atom zu einem tieferen Zustand m' , so liegt der anti-STOKESSche Fall vor, die ausgestrahlte Frequenz ist

$$\nu_s = \nu_r + |\nu_{n' m'}|.$$

Das in Formel (7) ausgedrückte Resultat stimmt mit § 50, (3) überein. Damit sind die in § 46 ff. gemachten Annahmen bezüglich der Intensität der sekundären Strahlung von einem höheren Standpunkt gerechtfertigt.

§76. Thermisches Gleichgewicht zwischen Strahlung und Atomen.

Die Formeln § 74, (7), (7a), (8a) oder:

$$(1) \quad w^{(a)} = B \varrho_\nu; \quad w^{(e)} = B \left(\varrho_\nu + \frac{8\pi h}{c^3} \nu^3 \right)$$

sind, wie schon erwähnt, nichts anderes als die berühmten von EINSTEIN 1916/17 angegebenen¹ Wahrscheinlichkeitsgesetze der Emission und Absorption der Quantenatome. Gebrauchen wir statt ϱ_ν wieder die *mittlere Quantenzahl* n_ν der Oszillatoren der Frequenz ν , so erhalten sie die Form:

$$(2) \quad w^{(a)} = C n_\nu; \quad w^{(e)} = C (n_\nu + 1),$$

$$\text{mit} \quad n_\nu = \frac{\varrho_\nu c^3}{8\pi h \nu^3}, \quad C = \frac{8\pi h \nu^3}{c^3} B.$$

Sie liefern in bekannter Weise das PLANCKSche Gesetz als Bedingung für das thermische Gleichgewicht zwischen Strahlung und statistisch normal verteilten Atomen (vgl. auch Band I, § 2, S. 11, 12). Ist nämlich N_1 die Anzahl der Atome im unteren und N_2 die im oberen Zustand, so gilt für die normale Verteilung bei der Temperatur T :

$$(3) \quad \frac{N_2}{N_1} = e^{-\frac{W_2 - W_1}{kT}} = e^{-\frac{h\nu}{kT}};$$

¹ EINSTEIN, A.: Verh. d. D. Phys. Ges. Bd. 18, S. 318. 1916. Phys. Zs. Bd. 18, S. 121. 1917.

und die in

$$(4) \quad N_1 w^{(a)} = N_2 w^{(e)}$$

ausgedrückte Gleichgewichtsforderung, daß im Mittel in der Zeiteinheit ebenso viele Atome der Energie W_1 absorbieren, wie Atome der Energie W_2 emittieren, führt mit (2) zu

$$(5) \quad \frac{N_2 w^{(e)}}{N_1 w^{(a)}} = e^{-\frac{h\nu}{kT}} \frac{n_r + 1}{n_r} = e^{-\frac{h\nu}{kT}} \left(1 + \frac{1}{n_r} \right) = 1$$

oder

$$(6) \quad n_r = \frac{1}{e^{\frac{h\nu}{kT}} - 1}.$$

Die Einführung von n_r statt ϱ_r in den EINSTEINSCHEN Formeln (1) läßt auch eine zunächst nicht ganz sichtbare *Symmetrie* der Gleichungen für Absorption und Emission erkennen. Die quantenmechanischen Gleichungen, welche den Energieaustausch zwischen Atom und Hohlraumeigenschwingungen bestimmen, sind offenbar *invariant* gegen eine *Umkehr der Zeitrichtung*; diese Invarianz muß in den Formeln (2) wiederzufinden sein. In der Tat ist die scheinbare Unsymmetrie der Gleichungen (2) für $w^{(a)}$ und $w^{(e)}$ leicht aufzuklären. Die Formel für $w^{(a)}$ bedeutet: Die Wahrscheinlichkeit, daß das Atom aus einer bestimmten Eigenschwingung von passender Frequenz, etwa der r -ten, ein Quantum $h\nu$ herausabsorbiert, ist proportional der Quantenzahl n_r , die diese Eigenschwingung *vor* dem Prozeß besitzt; *nach* dem Prozeß ist ja diese Quantenzahl um 1 vermindert. In vollkommener zeitlicher Symmetrie ist dagegen die Abgabe eines Quantums vom Atom an eine bestimmte Eigenschwingung proportional der Quantenzahl dieser Eigenschwingung *nach* Durchführung des Emissionsprozesses; also proportional mit $n_r + 1$, wenn n_r der Wert der Quantenzahl *vor* dem Prozesse war. Bei der statistischen Mittelbildung über alle in Betracht kommenden Eigenschwingungen ist nun zu beachten, daß n_r dem Mittelwert von n_r *ohne* Berücksichtigung des Prozesses entspricht, also *vor* der etwaigen Änderung von n_r durch den Prozeß zu bilden ist. Damit ergibt sich für Absorption Proportionalität mit n_r und für Emission Proportionalität mit $n_r + 1$.

Man kann sich weiterhin auch überzeugen, daß die Absorption und Emission von Strahlung durch Atome nicht nur die PLANCK-

sche Verteilung n_ν , ungeändert läßt, sondern auch¹ die statistischen Normalwerte der $p_\nu^{(n)}$ in § 70.

Wir können nämlich die Formeln (2) offenbar folgendermaßen verschärfen: Sei $w_1^{(n)}$ die Wahrscheinlichkeit, daß das Atom ein Quant $h\nu$ absorbiert aus einer solchen Eigenschwingung, welche (vor dem Prozeß) die Quantenzahl $n_r = n$ besitzt; und sei $w_2^{(n)}$ die Wahrscheinlichkeit, daß das Atom emittiert in eine Eigenschwingung hinein, welche (vor der Emission) die Quantenzahl $n_r = n$ besitzt. Dann gilt

$$(7) \quad w_1^{(n)} = C \cdot n p_\nu^{(n)}; \quad w_2^{(n)} = C \cdot (n + 1) p_\nu^{(n)}.$$

Mit den Formeln (33), (35) aus § 70, nämlich

$$\sum_n n p_\nu^{(n)} = n_\nu; \quad \sum_n p_\nu^{(n)} = 1,$$

ergibt sich aus (7) durch Summation wieder (2). Die statistischen Gleichgewichtsbedingungen aber lauten:

$$(8) \quad N_1 w_1^{(n+1)} = N_2 w_2^{(n)};$$

und sie sind wirklich erfüllt mit der Atomverteilung (3) und der Formel (34), § 70 für die statistisch normalen Werte der $p_\nu^{(n)}$:

$$p_\nu^{(n)} = \left(1 - e^{-\frac{h\nu}{kT}} \right) e^{-n \frac{h\nu}{kT}},$$

aus der

$$\frac{w_1^{(n+1)}}{w_2^{(n)}} = \frac{p_\nu^{(n+1)}}{p_\nu^{(n)}} = e^{-\frac{h\nu}{kT}}$$

folgt.

Endlich haben wir nachzuweisen, daß sich auch bei *Dispersionsprozessen* das statistische Gleichgewicht zwischen Strahlung und Materie erhält². Aus § 75, Formel (2) ff. sieht man: Die Wahrscheinlichkeit eines Prozesses, bei dem ein Atom im Zustand W_1 in ein Atom im Zustand W_2 übergeht (wobei jetzt insbesondere auch $W_2 = W_1$ sein kann), während ein Quantum $h\nu'$ verschwindet und ein Quantum $h\nu''$ neu erzeugt wird, ist gleich

$$(9) \quad w_1 = D \cdot n_{\nu'} (n_{\nu''} + 1).$$

¹ Vgl. BOTHE, W.: Z. Phys. Bd. 41, S. 345. 1927.

² Vgl. dazu W. PAULI JR.: Z. Phys. Bd. 18, S. 272. 1923.

Die Wahrscheinlichkeit des *inversen* Prozesses — Übergang des Atoms von der Energie W_2 zur Energie W_1 , Verschwinden eines Quants $h\nu''$ und Erzeugung eines Quants $h\nu'$ — ist entsprechend gleich

$$(9') \quad w_2 = D \cdot n_{\nu''} \cdot (n_{\nu'} + 1).$$

Die Gleichgewichtsforderung $N_1 w_1 = N_2 w_2$ ergibt also diesmal

$$(10) \quad N_1 \frac{n_{\nu''} + 1}{n_{\nu''}} = N_2 \frac{n_{\nu'} + 1}{n_{\nu'}}.$$

Da aber bei thermischer Normalverteilung

$$(11) \quad \frac{N_2}{N_1} = e^{-\frac{W_2 - W_1}{kT}} = e^{-\frac{h\nu' - h\nu''}{kT}}$$

ist, so ist (10) gleichbedeutend mit

$$(12) \quad e^{-\frac{h\nu''}{kT}} \frac{n_{\nu''} + 1}{n_{\nu''}} = e^{-\frac{h\nu'}{kT}} \frac{n_{\nu'} + 1}{n_{\nu'}}.$$

Diese Gleichung ist aber wirklich erfüllt, wenn die $n_{\nu'}$, $n_{\nu''}$ dem PLANCKSchen Gesetz entsprechen; denn nach (5) sind dann beide Seiten der Gleichung (12) gleich 1.

Es ist nun ohne weiteres zu sehen, wie diese Betrachtungen zu verallgemeinern¹ sind für noch verwickeltere Strahlungsreaktionen, bei denen das von W_1 nach W_2 übergehende Atom beliebige Anzahlen von Lichtquanten $h\nu'_1, h\nu'_2, \dots$ bzw. $h\nu''_1, h\nu''_2, \dots$ verschwinden läßt bzw. neu erzeugt. Ein besonderes Beispiel hierfür, Emission von zwei Lichtquanten zugleich, haben wir schon in § 75 erwähnt (vgl. die Fußnote S. 406). Noch kompliziertere Fälle sind zwar praktisch-experimentell von geringerer Bedeutung gegenüber den gewöhnlichen Absorptions-, Emissions- und Dispersionsprozessen, ergeben sich aber als notwendige Folgerung der Theorie, wenn man die Störungsrechnungen von §§ 74, 75 zu noch höheren Näherungen weiter treibt. Die Wahrscheinlichkeit eines solchen Prozesses wird allgemein die Gestalt

$$(13) \quad w_1 = G \cdot n_{\nu'_1} n_{\nu'_2} \dots (n_{\nu''_1} + 1) (n_{\nu''_2} + 1) \dots$$

erhalten, während die des *inversen* Prozesses gleich

$$(13') \quad w_2 = G \cdot n_{\nu''_1} n_{\nu''_2} \dots (n_{\nu'_1} + 1) (n_{\nu'_2} + 1) \dots$$

¹ Vgl. A. EINSTEIN und P. EHRENFEST: Z. Phys. Bd. 19, S. 301. 1923.

wird. Bedenkt man nun, daß bei normaler Verteilung

$$(14) \quad \frac{N_2}{N_1} = e^{-\frac{W_2 - W_1}{kT}} = e^{-\frac{(\hbar\nu_1' + \hbar\nu_2' + \dots) - (\hbar\nu_1'' + \hbar\nu_2'' + \dots)}{kT}}$$

wird, so erhält man die Gleichgewichtsbedingung $N_1 w_1 = N_2 w_2$ in der Form

$$(15) \quad \left\{ e^{-\frac{\hbar\nu_1''}{kT} \frac{n_{\nu_1''} + 1}{n_{\nu_1''}}} \right\} \cdot \left\{ e^{-\frac{\hbar\nu_2''}{kT} \frac{n_{\nu_2''} + 1}{n_{\nu_2''}}} \right\} \dots \\ = \left\{ e^{-\frac{\hbar\nu_1'}{kT} \frac{n_{\nu_1'} + 1}{n_{\nu_1'}}} \right\} \cdot \left\{ e^{-\frac{\hbar\nu_2'}{kT} \frac{n_{\nu_2'} + 1}{n_{\nu_2'}}} \right\} \dots;$$

und wiederum ist sie auf Grund des PLANCKSchen Gesetzes erfüllt, indem jeder der eingeklammerten Faktoren der rechten und linken Seite von (15) gleich 1 wird.

In allen Erörterungen dieses Paragraphen ist, ebenso wie in §§ 74, 75, das *Atom* (oder der *Atomkern*) als *unendlich schwer* angenommen worden, so daß die Impulse $\frac{\hbar\nu}{c}$ der Lichtquanten $\hbar\nu$ nicht berücksichtigt zu werden brauchten. Die quantenmechanische Untersuchung ihrer Einflüsse auf die Strahlungsreaktionen wird zu den Aufgaben der Fortsetzung dieses Buches gehören.

Wir wollen zum Schluß einen Rückblick auf den Gedankengang des ganzen Buches werfen. Dieser beruhte auf der schrittweisen induktiven Verallgemeinerung der einfachen HEISENBERG'schen Ansätze. Der Ausgangspunkt war die Annahme, daß die Emissionen und Absorptionen der Atome gemäß der BOHR'schen Frequenzbedingung und den EINSTEIN'schen Wahrscheinlichkeitsgesetzen durch die Energiewerte der stationären Zustände (hinsichtlich der Frequenzen) und durch die Matricelemente des elektrischen Moments \mathfrak{P} (hinsichtlich der Intensitäten) bestimmt seien. Das war ursprünglich eine reine Hypothese und nur durch korrespondenzmäßige Analogien zur klassischen Theorie gerechtfertigt. Hieraus entstand in natürlicher Weise der HEISENBERG'sche Formalismus. Aber die dabei gemachten Annahmen und Vorstellungen bezüglich der Verknüpfung der Strahlungsreaktion eines Atoms mit der matrisenmäßig beschriebenen Kinematik und

Mechanik dieses Atoms hatten zunächst den Charakter logisch selbständiger Hypothesen; und diese waren unentbehrliche Elemente im induktiven Aufbau der Theorie.

Eine *deduktive* Darstellung, welche die logische Einheit und Widerspruchslosigkeit der Theorie erweisen will, muß jedoch ausgehen von möglichst *einfachen*, *umfassenden* und *exakten* Hypothesen; nicht dagegen von Annahmen bezüglich eines konkreten Einzelproblems, wie es die Strahlungsreaktionen darbieten, und noch dazu von Annahmen, die statt grundsätzlich exakter nur eine gewisse angenäherte Gültigkeit beanspruchen können.

Formulierungen, die in diesem Sinne geeignet sind, als Ausgangspunkt einer deduktiven Entwicklung sowohl der formalen Theorie als auch ihres physikalischen Inhalts zu dienen, haben wir — unseren induktiven Weg verfolgend — mit den Feststellungen über die „Relativwahrscheinlichkeiten“ erreicht (Kap. VI). Es mußte also gezeigt werden, daß aus diesen Feststellungen die Zuverlässigkeit der ursprünglichen Hypothesen (in erster Annäherung) deduktiv abgeleitet werden kann. Das ist nunmehr geschehen, so daß der Kreis unserer Betrachtungen geschlossen ist.

Die weiteren Ergebnisse, die wir in §§ 72, 73 bezüglich des Lichtquantenproblems erhalten haben, zeigen überdies, daß die aufklärende Kraft der aus dem Korrespondenzprinzip erwachsenen Quantenmechanik auch auf solche Gebiete hinüberreicht, an die in den ersten Entwürfen HEISENBERGS wohl noch nicht gedacht war.

Anhang I.

Emission und Absorption elektromagnetischer Wellen nach der klassischen Theorie. Die Emission eines atomaren Systems hängt nach der klassischen Theorie von HEINRICH HERTZ von dem elektrischen Moment \mathfrak{P} ab, und zwar ist die in der Zeit Δt ausgesandte Energie

$$(1) \quad \Delta E^{(e)} = -\frac{2}{3c^3} \overline{\mathfrak{P}^2} \cdot \Delta t,$$

wo c die Lichtgeschwindigkeit ist, die Punkte Differentiationen nach der Zeit und der Strich das Zeitmittel bedeutet¹. Dabei ist angenommen, daß das Atom ohne Ausstrahlung eine mehrfach periodische Bewegung ausführen würde und daß die ausgestrahlte Energie während einer großen Zahl von Grundperioden noch relativ klein bleibt. Nun kann man in (1) für \mathfrak{P} die Reihe § 1, (3) bzw. (4) des Textes setzen:

$$(2) \quad \mathfrak{P} = \sum_{\tau} \mathfrak{P}_{\tau} e^{2\pi i (w \nu)} = \sum_{\tau} \mathfrak{P}_{\tau} e^{2\pi i [(\tau \nu) t + (\tau \delta)]}.$$

Dann erhält man

$$\ddot{\mathfrak{P}}^2 = 16 \pi^4 \sum_{\tau} \sum_{\sigma} (\tau \nu)^2 (\sigma \nu)^2 \mathfrak{P}_{\tau} \mathfrak{P}_{\sigma} e^{2\pi i [(\tau + \sigma, \nu) t + (\tau + \sigma, \delta)]}$$

und mit Rücksicht auf $\mathfrak{P}_{-\tau} = \mathfrak{P}_{\tau}^*$ [s. § 1, (5)]:

$$\overline{\ddot{\mathfrak{P}}^2} = 16 \pi^4 \sum_{\tau} (\tau \nu)^4 |\mathfrak{P}_{\tau}|^2.$$

Erstreckt man hier die Summation anstatt über alle Wertsysteme der τ nur über solche, für die $\tau \nu > 0$ ist, so hat man rechts den Faktor 2 hinzuzufügen und erhält durch Einsetzen in (1)

$$(3) \quad \Delta E^{(e)} = -\frac{64 \pi^4}{3 c^3} \sum_{(\tau \nu) > 0} (\tau \nu)^4 |\mathfrak{P}_{\tau}|^2 \cdot \Delta t.$$

Die einzelnen Glieder dieser Summe entsprechen den monochromatischen Emissionen

$$(3a) \quad \Delta E_{(\tau \nu)}^{(e)} = -\frac{64 \pi^4}{3 c^3} (\tau \nu)^4 |\mathfrak{P}_{\tau}|^2 \cdot \Delta t,$$

in Übereinstimmung mit § 1, (8).

In analoger Weise kann man die Absorption einer Lichtwelle durch das Atom berechnen². Das elektrische Feld der Welle kann im Bereiche des Atoms als räumlich konstant angesehen werden und werde (als kleine Größe) mit $\lambda \mathfrak{E}(t)$ bezeichnet. In einem Zeitabschnitt $\Delta t = t_1 - t_0$ sei $\lambda \mathfrak{E}(t)$ dargestellt durch das FOURIER-

¹ S. etwa ABRAHAM, M.: Theorie der Elektrizität, Bd. 2, 3. Aufl. § 9, S. 63, Formel (56c), Leipzig 1914.

² Wir folgen hier der Methode von BORN, M. u. P. JORDAN: Z. Phys. Bd. 33, S. 479. 1925. S. auch NIESSEN, K. F.: Ann. Phys. Bd. 75, S. 743. 1924 und VAN VLECK, J. H.: Phys. Rev. Bd. 24, S. 330. 1924; Journ. Opt. Soc. Amer. Bd. 9, S. 27. 1924.

sche Integral

$$(4) \quad \lambda \mathfrak{E}(t) = \int_{-\infty}^{\infty} e(\omega) e^{2\pi i \omega t} d\omega, \quad t_0 \leqq t \leqq t_1.$$

Dann ist der Amplitudenvektor gegeben durch

$$(4a) \quad e(\omega) = \lambda \int_{t_0}^{t_1} \mathfrak{E}(t) e^{-2\pi i \omega t} dt.$$

Zu der Eigenenergie H_0 des Atoms tritt im Felde die Störungsenergie λH_1 , wo

$$(5) \quad H_1 = - \mathfrak{P} \mathfrak{E}(t) = - \sum_{\tau} \mathfrak{P}_{\tau} \mathfrak{E}(t) e^{2\pi i (w\tau)}.$$

Da diese von der Zeit explizite abhängt, tritt nach Bd. 1, § 10, S. 64 als HAMILTONSche Funktion nicht die Energie $H = H_0 + \lambda H_1$ auf, sondern

$$(6) \quad W = H + \frac{\partial S}{\partial t},$$

wo S die Erzeugende der kanonischen Transformation ist, welche die Winkel- und Wirkungsvariabeln w_k^0, J_k^0 des ungestörten Systems in die des gestörten w_k, J_k verwandelt, und zwar auf Grund der Formeln

$$(7) \quad J_k^0 = \frac{\partial S}{\partial w_k^0}, \quad w_k = \frac{\partial S}{\partial J_k}.$$

Man hat nun S so zu bestimmen, daß W in eine Funktion der J_k allein übergeht. Da H als Funktion der w_k^0, J_k^0 des ungestörten Systems gegeben ist, so hat man nach (6) und (7) die HAMILTON-JACOBISCHE Differentialgleichung

$$(8) \quad H\left(w^0, \frac{\partial S}{\partial w^0}\right) + \frac{\partial S}{\partial t} = W(J)$$

zu lösen. Dies geschieht durch den Ansatz:

$$(9) \quad \begin{cases} W = W_0 + \lambda W_1 + \lambda^2 W_2 + \dots \\ S = \sum_k w_k^0 J_k + \lambda S_1 + \lambda^2 S_2 + \dots, \end{cases}$$

der auf die Rekursionsformeln führt:

$$(10) \quad \begin{cases} W_0 = H_0(J), \\ W_1 = \sum_k v_k \frac{\partial S_1}{\partial w_k^0} + \frac{\partial S_1}{\partial t} + H_1(J, w^0) \\ \dots \end{cases}$$

Dabei bedeuten die ν_k die Frequenzen des *ungestörten* Systems: $\nu_k = \frac{\partial H_0}{\partial J_k}$. Die erste von diesen Gleichungen liefert nichts Neues, nämlich die Energie des ungestörten Systems als Funktion der J_k . Die zweite Gleichung lautet nach (5) ausführlich:

$$W_1 = \sum_k \nu_k \frac{\partial S_1}{\partial w_k^0} + \frac{\partial S_1}{\partial t} - \sum_\tau \mathfrak{P}_\tau \mathfrak{E}(t) e^{2\pi i (w^0 \tau)}.$$

Durch den Ansatz¹

$$(11) \quad S_1 = - \sum_\tau \mathfrak{P}_\tau \mathfrak{S}_\tau(t) e^{2\pi i (w^0 \tau)}$$

geht sie über in

$$W_1 = - \sum_\tau (2\pi i (\nu \tau) \mathfrak{S}_\tau + \dot{\mathfrak{S}}_\tau + \mathfrak{E}) \mathfrak{P}_\tau e^{2\pi i (w^0 \tau)}.$$

Mittelt man über die w_k^0 , so erhält man

$$(12) \quad W_1 = - \mathfrak{E} \mathfrak{P}_0.$$

Sodann ergibt sich für \mathfrak{S}_τ ($\tau \neq 0$) die Differentialgleichung

$$(13) \quad \dot{\mathfrak{S}}_\tau + 2\pi i (\nu \tau) \mathfrak{S}_\tau = - \mathfrak{E}(t).$$

Ihre für $t = t_0$ verschwindende Lösung ist

$$(14) \quad \mathfrak{S}_\tau(t) = - \int_{t_0}^t e^{-2\pi i (\nu \tau) (t-t')} \mathfrak{E}(t') dt'.$$

Das Näherungsverfahren läßt sich fortsetzen, doch genügt hier die erste Näherung.

Aus der Entwicklung (9) von S erhält man nach (7)

$$J_k^0 = J_k + \lambda \frac{\partial S_1}{\partial w_k^0} + \dots, \quad w_k = w_k^0 + \lambda \frac{\partial S_1}{\partial J_k} + \dots,$$

oder aufgelöst in erster Näherung

$$J_k = J_k^0 - \lambda \frac{\partial S_1}{\partial w_k^0}, \quad w_k = w_k^0 + \lambda \frac{\partial S_1}{\partial J_k^0}.$$

Daher wird das elektrische Moment unter dem Einfluß der Störung

$$(15) \quad \mathfrak{P}(w, J) = \mathfrak{P}_0 + \lambda \mathfrak{P}_1 + \dots$$

¹ Diese Methode stammt von JORDAN, P.: Z. Phys. Bd. 33, S. 506. 1925.

mit

$$(15a) \quad \mathfrak{F}_1 = \sum_k \left(\frac{\partial \mathfrak{F}}{\partial w_k^0} \frac{\partial S_1}{\partial J_k^0} - \frac{\partial \mathfrak{F}}{\partial J_k^0} \frac{\partial S_1}{\partial w_k^0} \right).$$

Setzt man hier die Ausdrücke (2) und (11) für \mathfrak{F} und S_1 ein, so erhält man

$$\mathfrak{F}_1 = -2\pi i \sum_{\tau, \sigma} \left\{ \mathfrak{F}_\tau \cdot \left(\tau, \frac{\partial \mathfrak{F}_\sigma \mathfrak{S}_\sigma}{\partial J} \right) - \left(\sigma, \frac{\partial \mathfrak{F}_\tau}{\partial J} \right) \cdot \mathfrak{F}_\sigma \mathfrak{S}_\sigma \right\} e^{2\pi i (\tau + \sigma, w)},$$

wobei der Index 0 bei w_k, J_k fortgelassen ist. Da nach (14) $\mathfrak{S}_\tau(t_0) = 0$ ist, so stellt (16) in der Tat eine zur Zeit t_0 einsetzende Störung dar.

Die in der Zeit Δt absorbierte Energie ist die vom Felde $\lambda \mathfrak{E}$ geleistete Arbeit:

$$(17) \quad \Delta E^{(a)} = -\lambda \int_{t_0}^{t_1} \mathfrak{F} \mathfrak{E} dt;$$

sie hängt noch von den Phasenkonstanten ab, und da man diese nicht kennt, hat man über sie zu mitteln. Nun ist das Phasenmittel von \mathfrak{F}_0 gleich Null, und da offenbar Phasenmittelung und Zeitdifferentiation vertauschbare Operationen sind, erhält man:

$$(17a) \quad \Delta E^{(a)} = -\lambda^2 \int_{t_0}^{t_1} \overline{\mathfrak{F}_1} \mathfrak{E} dt.$$

Aus (16) folgt zunächst

$$\overline{\mathfrak{F}_1} = -2\pi i \sum_\tau \left\{ \mathfrak{F}_\tau \cdot \left(\tau, \frac{\partial \mathfrak{F}_{-\tau} \mathfrak{S}_{-\tau}}{\partial J} \right) + \left(\tau, \frac{\partial \mathfrak{F}_\tau}{\partial J} \right) \cdot \mathfrak{F}_{-\tau} \mathfrak{S}_{-\tau} \right\}.$$

Hier hängt \mathfrak{S}_τ noch von der Zeit ab, und man hat nach (13)

$$\dot{\mathfrak{S}}_{-\tau} = -\mathfrak{E} + 2\pi i (\nu \tau) \mathfrak{S}_{-\tau}.$$

Daher wird

$$\begin{aligned} -\overline{\mathfrak{F}_1} \mathfrak{E} &= 2\pi i \sum_\tau \left\{ \mathfrak{F}_\tau \mathfrak{E} \cdot \left(\tau, \frac{\partial \mathfrak{F}_{-\tau} \dot{\mathfrak{S}}_{-\tau}}{\partial J} \right) + \left(\tau, \frac{\partial \mathfrak{F}_\tau \mathfrak{E}}{\partial J} \right) \cdot \mathfrak{F}_{-\tau} \dot{\mathfrak{S}}_{-\tau} \right\} \\ &= 2\pi i \sum_\tau \left(\tau, \frac{\partial}{\partial J} [\mathfrak{F}_\tau \mathfrak{E} \cdot \mathfrak{F}_{-\tau} \dot{\mathfrak{S}}_{-\tau}] \right) \\ &= -2\pi i \sum_\tau \left(\tau, \frac{\partial |\mathfrak{F}_\tau \mathfrak{E}|^2}{\partial J} \right) \\ &\quad + 4\pi^2 \sum_\tau \left(\tau, \frac{\partial}{\partial J} [(\nu \tau) \cdot \mathfrak{F}_\tau \mathfrak{E} \cdot \mathfrak{F}_{-\tau} \mathfrak{S}_{-\tau}] \right). \end{aligned}$$

Die erste Summe verschwindet, da sich immer je zwei Glieder τ , $-\tau$ fortheben. In die zweite setzen wir den Ausdruck (14) von \mathfrak{S}_τ ein und finden dann nach (17a)

$$\Delta E^{(a)} = 4\pi^2 \lambda^2.$$

$$\int_{t_0}^{t_1} dt \int_{t_0}^t dt' \sum_{\tau} \left(\tau, \frac{\partial}{\partial J} [(\nu\tau) \cdot \mathfrak{F}_\tau \mathfrak{E}(t) \cdot \mathfrak{F}_{-\tau} \mathfrak{E}(t')] \right) e^{2\pi i (\nu\tau)(t-t')}.$$

Der Integrand ist symmetrisch in t und t' ; denn vertauscht man t und t' , so kann man den Vorzeichenwechsel im Exponenten durch Vertauschung von τ mit $-\tau$ rückgängig machen. Man kann daher im innern Integral bis t_1 integrieren, wenn man zugleich den Faktor $\frac{1}{2}$ zufügt. Dann erhält man unter Benutzung der Definition (4a):

$$\Delta E^{(a)} = 2\pi^2 \sum_{\tau} \left(\tau, \frac{\partial}{\partial J} [(\nu\tau) \mathfrak{F}_\tau e(\nu\tau) \cdot \mathfrak{F}_{-\tau} e(-\nu\tau)] \right)$$

oder, indem man nur über die Werte von τ summiert, für die $(\nu\tau) > 0$ ist:

$$(18) \quad \Delta E^{(a)} = 4\pi^2 \sum_{(\nu\tau) > 0} \left(\tau, \frac{\partial}{\partial J} [(\nu\tau) \cdot |\mathfrak{F}_\tau e(\nu\tau)|^2] \right).$$

Die Stellung des Vektors \mathfrak{F} im Raume ist willkürlich, sofern nicht durch ein äußeres Kraftfeld eine ausgezeichnete Richtung hergestellt wird; man wird daher noch über alle Stellungen zu mitteln haben. Statt dessen kann man auch das Atom festgehalten denken und über alle Richtungen des Feldvektors oder seines monochromatischen Anteils $e(\nu\tau)$ mitteln. Diese Art der Mittelung ist auch dann erlaubt, wenn für die Atome zwar eine bestimmte Richtung der Einstellung ausgezeichnet, aber das Strahlungsfeld völlig isotrop ist. Nun ist

$$\begin{aligned} |\mathfrak{F}_\tau e(\nu\tau)|^2 &= \mathfrak{F}_\tau e(\nu\tau) \cdot \mathfrak{F}_{-\tau} e(-\nu\tau) \\ &= \mathfrak{F}_{\tau x} \mathfrak{F}_{-\tau x} e_x(\nu\tau) e_x(-\nu\tau) + \dots \\ &+ \mathfrak{F}_{\tau y} \mathfrak{F}_{-\tau y} e_y(\nu\tau) e_y(-\nu\tau) + \dots \\ &+ \mathfrak{F}_{\tau z} \mathfrak{F}_{-\tau z} e_z(\nu\tau) e_z(-\nu\tau) + \dots \end{aligned}$$

Die Mittelwerte der Produkte $e_y(\nu\tau) e_z(-\nu\tau)$, ... sind Null, die der drei Quadrate $|e_x(\nu\tau)|^2$, ... einander gleich, und zwar

gleich $\frac{1}{3} |e(\nu\tau)|^2$. Also wird im Mittel

$$(19) \quad |\mathfrak{A}_\tau e(\nu\tau)|^2 = \frac{1}{3} |\mathfrak{A}_\tau|^2 \cdot |e(\nu\tau)|^2.$$

Man kann nun $|e(\nu\tau)|^2$ durch die mittlere monochromatische Energiedichte des Feldes ausdrücken.

Die gesamte Energiedichte im elektromagnetischen Felde $\mathfrak{E}, \mathfrak{H}$ beträgt

$$u = \frac{1}{8\pi} (\mathfrak{E}^2 + \mathfrak{H}^2),$$

Bei stationärer Strahlung sind die Zeitmittelwerte von \mathfrak{E}^2 und \mathfrak{H}^2 gleich, also die mittlere Energiedichte

$$(20) \quad \bar{u} = \frac{1}{4\pi} \frac{1}{\Delta t} \int_{t_0}^{t_1} \mathfrak{E}(t)^2 dt.$$

Nun ist nach (4) und (4a)

$$\begin{aligned} \int_{t_0}^{t_1} \mathfrak{E}(t)^2 dt &= \int_{t_0}^{t_1} \mathfrak{E}(t) dt \cdot \int_{-\infty}^{\infty} e(\omega) e^{2\pi i \omega t} d\omega \\ &= \int_{-\infty}^{\infty} e(\omega) \int_{t_0}^{t_1} \mathfrak{E}(t) e^{2\pi i \omega t} dt d\omega \\ &= \int_{-\infty}^{\infty} |e(\omega)|^2 d\omega = 2 \int_0^{\infty} |e(\omega)|^2 d\omega. \end{aligned}$$

Setzt man das in (20) ein, so kann man schreiben

$$(21) \quad \bar{u} = \int_0^{\infty} \varrho(\omega) d\omega, \quad \varrho(\omega) = \frac{|e(\omega)|^2}{2\pi \Delta t}$$

und $\varrho(\omega)$ hat dann die Bedeutung der mittleren monochromatischen Energiedichte. Aus (19) wird jetzt

$$(22) \quad |\mathfrak{A}_\tau e(\nu\tau)|^2 = \frac{2\pi}{3} |\mathfrak{A}_\tau|^2 \cdot \varrho(\nu\tau) \cdot \Delta t,$$

und für die absorbierte Energie folgt aus (18):

$$(23) \quad \Delta E^{(a)} = \frac{8\pi^3}{3} \sum_{(\nu\tau) > 0} \left(\tau, \frac{\partial}{\partial J} [(\nu\tau) \cdot |\mathfrak{A}_\tau|^2 \cdot \varrho(\nu\tau)] \right) \cdot \Delta t.$$

Das einzelne Glied dieser Summe ist die im Texte § 1, (8b) angegebene *monochromatische Absorption*:

$$(24) \quad \Delta E_{(\tau\nu)}^{(a)} = \frac{8\pi^3}{3} \sum_k \tau_k \frac{\partial}{\partial J_k} [(\nu\tau) \cdot |\mathfrak{A}_\tau|^2 \cdot \varrho(\nu\tau)] \cdot \Delta t.$$

Anhang II.

Die JACOBI-POISSONSCHEN KLAMMERSYMBOLS in der klassischen Mechanik¹. In Bd. 1, § 7 wurde gezeigt, wie man auf verschiedene Weisen alle kanonischen Transformationen von f Paaren kanonisch konjugierter Variablen q_k, p_k mit Hilfe einer willkürlichen Funktion von $2f$ Veränderlichen (außer der Zeit t) darstellen kann. Eine dieser Darstellungen [Bd. 1, § 7, (2)] schreiben wir hier für den Spezialfall, daß die Zeit nicht explizite vorkommt, mit etwas anderer Bezeichnung an:

$$(1) \quad \begin{cases} p_k = \frac{\partial}{\partial q_k} V(q, P), \\ Q_k = \frac{\partial}{\partial P_k} V(q, P); \end{cases}$$

dabei ist $V(q, P)$ statt $V(q_1, \dots, q_f, P_1, \dots, P_f)$ geschrieben.

In Bd. 1, § 8 wurde sodann eine Reihe von f mehrfachen Integralen angegeben, die gegenüber kanonischen Transformationen invariant sind, die sogenannten POINCARÉschen Integralinvarianten. Es gibt aber auch *kanonische Differentialinvarianten*, und von diesen soll hier die Rede sein.

Es seien $x(p, q), y(p, q)$ irgend welche (differenzierbare) Funktionen der p_k, q_k . Dann nennt man die Größen

$$(2) \quad [x, y] = \sum_k \begin{vmatrix} \frac{\partial x}{\partial p_k} & \frac{\partial x}{\partial q_k} \\ \frac{\partial y}{\partial p_k} & \frac{\partial y}{\partial q_k} \end{vmatrix}$$

JACOBIsche oder POISSONSche Klammersymbole.

Wir zeigen nun, daß diese kanonisch invariant (besser: kovariant) sind, und zwar nach einer Methode, die der in Bd. 1, § 7 benützten, von BRODY stammenden nachgebildet ist.

Man denke sich zunächst in x und y statt p_k, q_k als unabhängige Variable p_k, Q_k , wobei die q_k und die Q_k durch die zweite der Gleichungen (1) (bei fest gehaltenen P_k) verbunden sind.

¹ Vgl. auch die Darstellungen bei NORDHEIM, L. u. E. FUES: Handb. d. Physik, Bd. V, Kap. 3; FRANK, PH.: Phys. Z. Bd. 30, S. 209. 1929.

Dann hat man

$$\begin{aligned}
 [x, y] &= \sum_k \begin{vmatrix} \frac{\partial x}{\partial p_k} & \sum_l \frac{\partial x}{\partial Q_l} \frac{\partial^2 V}{\partial P_l \partial q_k} \\ \frac{\partial y}{\partial p_k} & \sum_l \frac{\partial y}{\partial Q_l} \frac{\partial^2 V}{\partial P_l \partial q_k} \end{vmatrix} = \sum_{kl} \frac{\partial^2 V}{\partial P_l \partial q_k} \begin{vmatrix} \frac{\partial x}{\partial p_k} & \frac{\partial x}{\partial Q_l} \\ \frac{\partial y}{\partial p_k} & \frac{\partial y}{\partial Q_l} \end{vmatrix} \\
 &= \sum_l \begin{vmatrix} \sum_k \frac{\partial x}{\partial p_k} \frac{\partial^2 V}{\partial q_k \partial P_l} & \frac{\partial x}{\partial Q_l} \\ \sum_k \frac{\partial y}{\partial p_k} \frac{\partial^2 V}{\partial q_k \partial P_l} & \frac{\partial y}{\partial Q_l} \end{vmatrix};
 \end{aligned}$$

denkt man sich nunmehr in x und y statt p_k, Q_k unabhängige Variable P_k, Q_k , wobei die p_k und die P_l vermöge der ersten Gleichung (1) (bei festgehaltenen q_k) zusammenhängen, so erhält man

$$(3) \quad [x, y] = \sum_l \begin{vmatrix} \frac{\partial x}{\partial P_l} & \frac{\partial x}{\partial Q_l} \\ \frac{\partial y}{\partial P_l} & \frac{\partial y}{\partial Q_l} \end{vmatrix}.$$

Die $[x, y]$ hängen also tatsächlich nicht von der Wahl der zu ihrer Berechnung benützten kanonischen Variablen ab.

Nimmt man nun speziell für x und y eine Koordinate oder einen Impuls selbst, so folgt aus (2) unmittelbar

$$(4) \quad \begin{cases} [p_r, q_s] = \delta_{rs}, \\ [p_r, p_s] = 0, \quad [q_r, q_s] = 0. \end{cases}$$

Das liefert aber nach (3) notwendige Bedingungen dafür, daß die Variablen P_k, Q_k aus den p_k, q_k durch kanonische Transformationen hervorgehen, nämlich

$$(5) \quad \left\{ \begin{array}{l} \sum_k \begin{vmatrix} \frac{\partial p_r}{\partial P_k} & \frac{\partial p_r}{\partial Q_k} \\ \frac{\partial q_s}{\partial P_k} & \frac{\partial q_s}{\partial Q_k} \end{vmatrix} = \delta_{rs}, \\ \sum_k \begin{vmatrix} \frac{\partial p_r}{\partial P_k} & \frac{\partial p_r}{\partial Q_k} \\ \frac{\partial p_s}{\partial P_k} & \frac{\partial p_s}{\partial Q_k} \end{vmatrix} = 0, \quad \sum_k \begin{vmatrix} \frac{\partial q_r}{\partial P_k} & \frac{\partial q_r}{\partial Q_k} \\ \frac{\partial q_s}{\partial P_k} & \frac{\partial q_s}{\partial Q_k} \end{vmatrix} = 0. \end{array} \right.$$

Diese Bedingungen sind aber auch hinreichend. Denn damit zugleich mit p_k, q_k auch P_k, Q_k kanonische Variable sind, ist notwendig und hinreichend, daß

$$(6) \quad \sum_k (p_k dq_k - P_k dQ_k)$$

ein totales Differential ist; denkt man sich hier Q_k als Funktion von p_k, q_k ausgedrückt, so soll also

$$(6a) \quad \sum_r \left\{ \left(p_r - \sum_k P_k \frac{\partial Q_k}{\partial q_r} \right) dq_r - \sum_k P_k \frac{\partial Q_k}{\partial p_r} dp_r \right\}$$

ein totales Differential sein. Hierfür aber wieder ist notwendig und hinreichend, daß die Integrationsbedingungen

$$\begin{aligned} \frac{\partial}{\partial p_r} \left(p_s - \sum_k P_k \frac{\partial Q_k}{\partial q_s} \right) + \frac{\partial}{\partial q_s} \sum_k P_k \frac{\partial Q_k}{\partial p_r} &= 0, \\ \frac{\partial}{\partial p_r} \sum_k P_k \frac{\partial Q_k}{\partial p_s} - \frac{\partial}{\partial p_s} \sum_k P_k \frac{\partial Q_k}{\partial p_r} &= 0, \\ \frac{\partial}{\partial q_r} \left(p_s - \sum_k P_k \frac{\partial Q_k}{\partial q_s} \right) - \frac{\partial}{\partial q_s} \left(p_r - \sum_k P_k \frac{\partial Q_k}{\partial q_r} \right) &= 0 \end{aligned}$$

erfüllt sind; diese vereinfachen sich aber zu

$$(7) \quad \left\{ \begin{array}{l} \sum_k \left| \begin{array}{cc} \frac{\partial P_k}{\partial p_r} & \frac{\partial P_k}{\partial q_s} \\ \frac{\partial Q_k}{\partial p_r} & \frac{\partial Q_k}{\partial q_s} \end{array} \right| = \delta_{rs}, \\ \sum_k \left| \begin{array}{cc} \frac{\partial P_k}{\partial p_r} & \frac{\partial P_k}{\partial p_s} \\ \frac{\partial Q_k}{\partial p_r} & \frac{\partial Q_k}{\partial p_s} \end{array} \right| = 0, \quad \sum_k \left| \begin{array}{cc} \frac{\partial P_k}{\partial q_r} & \frac{\partial P_k}{\partial q_s} \\ \frac{\partial Q_k}{\partial q_r} & \frac{\partial Q_k}{\partial q_s} \end{array} \right| = 0. \end{array} \right.$$

Die Summen linker Hand sind Spezialfälle des LAGRANGESchen Klammersymbols:

$$(8) \quad \{x, y\} = \sum_k \left| \begin{array}{cc} \frac{\partial P_k}{\partial x} & \frac{\partial P_k}{\partial y} \\ \frac{\partial Q_k}{\partial x} & \frac{\partial Q_k}{\partial y} \end{array} \right|,$$

die man für irgend eine zweidimensionale Fläche $P_k(x, y), Q_k(x, y)$ des PQ -Raumes bilden kann. Man kann also (7) schreiben:

$$(9) \quad \begin{cases} \{p_r, q_s\} = \delta_{rs}, \\ \{p_r, p_s\} = 0, \quad \{q_r, q_s\} = 0. \end{cases}$$

Sind u_1, u_2, \dots, u_{2f} irgend $2f$ unabhängige Funktionen der p_k, q_k , so sind die Matrizen mit den Elementen $[u_r, u_s]$ bzw. $\{u_s, u_r\}$ *reziprok*, d. h. es gilt

$$(10) \quad \sum_{t=1}^{2f} [u_t, u_r] \cdot \{u_t, u_s\} = \delta_{rs}.$$

Die linke Seite ist nämlich auf Grund der Multiplikationsregel der Determinanten nach (3) und (8) gleich

$$\begin{aligned} & \sum_{kl} \sum_t \begin{vmatrix} \frac{\partial u_t}{\partial P_l} & \frac{\partial u_t}{\partial Q_l} \\ \frac{\partial u_r}{\partial P_l} & \frac{\partial u_r}{\partial Q_l} \end{vmatrix} \cdot \begin{vmatrix} \frac{\partial P_k}{\partial u_t} & \frac{\partial P_k}{\partial u_s} \\ \frac{\partial Q_k}{\partial u_t} & \frac{\partial Q_k}{\partial u_s} \end{vmatrix} \\ &= \sum_{kl} \sum_t \begin{vmatrix} \frac{\partial u_t}{\partial P_l} \frac{\partial P_k}{\partial u_t} + \frac{\partial u_t}{\partial Q_l} \frac{\partial Q_k}{\partial u_t} & \frac{\partial u_t}{\partial P_l} \frac{\partial P_k}{\partial u_s} + \frac{\partial u_t}{\partial Q_l} \frac{\partial Q_k}{\partial u_s} \\ \frac{\partial u_r}{\partial P_l} \frac{\partial P_k}{\partial u_t} + \frac{\partial u_r}{\partial Q_l} \frac{\partial Q_k}{\partial u_t} & \frac{\partial u_r}{\partial P_l} \frac{\partial P_k}{\partial u_s} + \frac{\partial u_r}{\partial Q_l} \frac{\partial Q_k}{\partial u_s} \end{vmatrix}. \end{aligned}$$

Hier ist das erste Glied der Diagonale, summiert über t , gleich $2\delta_{kl}$; daher wird der Ausdruck gleich:

$$\begin{aligned} & 2 \sum_k \left(\frac{\partial u_r}{\partial P_k} \frac{\partial P_k}{\partial u_s} + \frac{\partial u_r}{\partial Q_k} \frac{\partial Q_k}{\partial u_s} \right) \\ & - \sum_{kl} \left(\frac{\partial P_k}{\partial u_s} \frac{\partial u_r}{\partial P_l} \sum_t \frac{\partial P_k}{\partial u_t} \frac{\partial u_t}{\partial P_l} + \frac{\partial P_k}{\partial u_s} \frac{\partial u_r}{\partial Q_l} \sum_t \frac{\partial Q_k}{\partial u_t} \frac{\partial u_t}{\partial P_l} \right. \\ & \quad \left. + \frac{\partial Q_k}{\partial u_s} \frac{\partial u_r}{\partial P_l} \sum_t \frac{\partial P_k}{\partial u_t} \frac{\partial u_t}{\partial Q_l} + \frac{\partial Q_k}{\partial u_s} \frac{\partial u_r}{\partial Q_l} \sum_t \frac{\partial Q_k}{\partial u_t} \frac{\partial u_t}{\partial Q_l} \right) \\ &= 2\delta_{rs} - \sum_k \left(\frac{\partial P_k}{\partial u_s} \frac{\partial u_r}{\partial P_k} + \frac{\partial Q_k}{\partial u_s} \frac{\partial u_r}{\partial Q_k} \right) \\ &= 2\delta_{rs} - \delta_{rs} = \delta_{rs}, \end{aligned}$$

womit (10) bewiesen ist.

Aus (10) folgt zunächst, daß die LAGRANGESchen Klammer-
symbole ebenso wie die JACOBI-POISSONSchen kanonisch invariant
sind.

Nimmt man ferner für die u_1, \dots, u_{2f} die Variablen $q_1, \dots, q_f,$
 p_1, \dots, p_f , so zeigt (10) ohne weiteres, daß die Gleichungssysteme
(4) und (9) gleichbedeutend sind. Da nun die Bedingungen (9)
für die kanonische Invarianz *hinreichend* sind, gilt dasselbe für (4).

Von diesen Formeln (4) wird im Texte (§ 17) für den Fall Ge-
brauch gemacht, daß P_k, Q_k die Wirkungs- und Winkelvariablen
 J_k, w_k sind.

Die Klammersymbole $[x, y]$ genügen folgenden trivialen
Identitäten:

$$(11) \quad \begin{cases} [x, y] = -[y, x], \\ [x + y, z] = [x, z] + [y, z], \\ [xy, z] = [x, z]y + x[y, z]. \end{cases}$$

Ferner hat man

$$(12) \quad \frac{\partial x}{\partial q_k} = -[x, p_k], \quad \frac{\partial x}{\partial p_k} = [x, q_k].$$

Man kann die Klammersymbole auch als Operatoren auffassen,
nämlich als *infinitesimale Berührungstransformationen*. Eine in-
finitesimale Transformation, durch die eine beliebige Funktion
 $z(p, q)$ in $z + \delta z$ mit

$$(13) \quad \delta z = \sum_k \left(\frac{\partial z}{\partial p_k} \delta p_k + \frac{\partial z}{\partial q_k} \delta q_k \right)$$

verwandelt wird, heißt Berührungstransformation, wenn $\delta p_k, \delta q_k$
sich aus einer willkürlich gegebenen Funktion $x(p, q)$, der Er-
zeugenden, durch die Gleichungen

$$(14) \quad \delta p_k = -\frac{\partial x}{\partial q_k}, \quad \delta q_k = \frac{\partial x}{\partial p_k}$$

ableiten lassen; wir schreiben $\delta_x z$. Dann wird aber

$$(15) \quad \delta_x z = \sum_k \left(\frac{\partial x}{\partial p_k} \frac{\partial z}{\partial q_k} - \frac{\partial x}{\partial q_k} \frac{\partial z}{\partial p_k} \right) = [x, z].$$

Bildet man nun mit Hilfe einer zweiten Erzeugenden $y(p, q)$
die infinitesimale Transformation δ_y und dann

$$\delta_x \delta_y z - \delta_y \delta_x z,$$

so ist das wieder eine infinitesimale Berührungstransformation, und zwar gehörig zu der erzeugenden Funktion $[x, y]$. Denn dieser Ausdruck ist jedenfalls von der Form

$$\sum_k \left(P_k \frac{\partial z}{\partial p_k} + Q_k \frac{\partial z}{\partial q_k} \right).$$

Zur Bestimmung von P_k und Q_k setze man zunächst $z = p_l$; dann wird offenbar

$$\delta_x \delta_y p_l - \delta_y \delta_x p_l = P_l,$$

oder, da nach (15) $\delta_x p_l = -\frac{\partial x}{\partial q_l}$ ist:

$$\begin{aligned} P_l &= -\delta_x \frac{\partial y}{\partial q_l} + \delta_y \frac{\partial x}{\partial q_l} \\ &= \left[\frac{\partial y}{\partial q_l}, x \right] + \left[y, \frac{\partial x}{\partial q_l} \right] \\ &= -\frac{\partial}{\partial q_l} [x, y]. \end{aligned}$$

In derselben Weise zeigt man, indem man $z = q_l$ setzt, daß

$$Q_l = \frac{\partial}{\partial p_l} [x, y]$$

ist. Somit wird

$$(16) \quad \delta_x \delta_y z - \delta_y \delta_x z = \sum_k \left(\frac{\partial [x, y]}{\partial p_k} \frac{\partial z}{\partial q_k} - \frac{\partial [x, y]}{\partial q_k} \frac{\partial z}{\partial p_k} \right) = \delta_{[x, y]} z.$$

Hierfür kann man aber auch schreiben:

$$[x, [y, z]] - [y, [x, z]] = [[x, y], z]$$

oder in symmetrischer Form

$$(17) \quad [x, [y, z]] + [y, [z, x]] + [z, [x, y]] = 0.$$

Dies ist die sogenannte **JACOBIsche Identität**.

Die kanonischen Bewegungsgleichungen

$$(18) \quad \dot{q}_k = \frac{\partial H}{\partial p_k}, \quad \dot{p}_k = -\frac{\partial H}{\partial q_k}$$

kann man wegen (12) in der Form

$$(18a) \quad \dot{q}_k = [H, q_k], \quad \dot{p}_k = [H, p_k]$$

schreiben. Ist $F(p, q)$ ein Integral dieser Gleichungen, das die Zeit nicht explizite enthält, d. h. $\dot{F} = 0$, so gilt nach (18)

$$(19) \quad [H, F] = 0.$$

Ist G ein zweites solches Integral, also auch

$$[H, G] = 0,$$

so folgt aus (17) und (18)

$$(20) \quad [H, [F, G]] = \frac{d}{dt}[F, G] = 0,$$

d. h. auch $[F, G]$ ist ein Integral der kanonischen Gleichungen. Dieser *Satz von POISSON* erlaubt es manchmal, aus bekannten Integralen neue abzuleiten.

Hat man ein System von f' Massenpunkten mit den $f = 3f'$ kartesischen Koordinaten und Impulsen

$$\begin{aligned} x_1, x_2, \dots, x_{f'}, & \quad p_{1x}, p_{2x}, \dots, p_{f'x}, \\ y_1, y_2, \dots, y_{f'}, & \quad p_{1y}, p_{2y}, \dots, p_{f'y}, \\ z_1, z_2, \dots, z_{f'}, & \quad p_{1z}, p_{2z}, \dots, p_{f'z}, \end{aligned}$$

so definiert man als *Drehmoment des k -ten Teilchens* den Vektor \mathfrak{M}_k mit den Komponenten

$$(21) \quad \begin{cases} M_{kx} = y_k p_{kz} - z_k p_{ky}, \\ M_{ky} = z_k p_{kx} - x_k p_{kz}, \\ M_{kz} = x_k p_{ky} - y_k p_{kx}. \end{cases}$$

Bildet man nun die JACOBI-POISSONSCHEN Klammersymbole für die Kombination dieser Größen mit den Koordinaten und Impulsen und untereinander, so bekommt man nach (2):

$$(22) \quad [M_{kx}, x_k] = 0, \dots, [M_{kx}, p_{kx}] = 0, \dots$$

$$(23) \quad \begin{cases} [M_{kz}, x_k] = -y_k, \dots \\ [M_{kz}, y_k] = x_k, \dots \end{cases}$$

$$(24) \quad \begin{cases} [M_{kz}, p_{kx}] = -p_{ky}, \dots \\ [M_{kz}, p_{ky}] = p_{kx}, \dots \end{cases}$$

und

$$(25) \quad [M_{ky}, M_{kz}] = -M_{kx}, \dots$$

Auch für die Gesamtdrehimpulse

$$(26) \quad M_x = \sum_{k=1}^f M_{kx}, \dots$$

gelten entsprechende Relationen, insbesondere

$$(27) \quad [M_y, M_z] = -M_x, \dots$$

Alle diese Relationen stimmen formal mit den entsprechenden quantenmechanischen Formeln der §§ 17, 18, 21, 25 überein.

Anhang III¹.

Hilfssatz zum Adiabatensatz.

Um das Integral

$$\int_0^\sigma (Q_0)_{mn}(s) e^{iT(\varphi_m - \varphi_n)} ds$$

abzuschätzen, bezeichnen wir abkürzend den reellen (oder den imaginären) Teil der Funktion $(Q_0)_{mn}(s)$ mit $f(s)$ und die Differenz $\varphi_m(s) - \varphi_n(s)$ mit $g(s)$ und betrachten das Integral

$$J = \int_0^\sigma f(s) e^{iTg(s)} ds.$$

Wir teilen das Integrationsintervall $(0, \sigma)$ in zwei Gruppen, E_1 und E_2 , von Teilintervallen, nämlich erstens (E_1) die Umgebungen

$$\alpha_k - \varepsilon < s < \alpha_k + \varepsilon$$

der Nullstellen α_k der Ableitung $g'(s)$ und zweitens (E_2) den übrigen Teil von $(0, \sigma)$.

Das Integral über E_1

$$J_1 = \int_{E_1} f(s) e^{iTg(s)} ds$$

genügt offenbar der Ungleichung

$$|J_1| < M \int_{E_1} ds = 2M N_1 \varepsilon,$$

wo N_1 die Anzahl der Nullstellen α_k von $g'(s)$ und M das Maximum des absoluten Betrages von $f(s)$ ist.

¹ BORN, M. und V. FOCK: Z. Phys. Bd. 51, S. 165, 1928.

Das Integral über E_2 schreiben wir in der Form

$$J_2 = \int_{E_2} \frac{f(s)}{g'(s)} e^{i T g(s)} g'(s) ds.$$

In E_2 ist $\frac{1}{g'(s)}$ endlich, und zwar kann man wegen der in der Nähe der Nullstellen gültigen Abschätzungen

$$\frac{1}{|g'(s)|} < \frac{A}{\varepsilon^r}$$

annehmen.

Indem wir den zweiten Mittelwertsatz der Integralrechnung

$$\int_{\alpha}^{\beta} \varphi(s) \psi(s) ds = \varphi(\alpha) \int_{\alpha}^{\vartheta} \psi(s) ds + \varphi(\beta) \int_{\vartheta}^{\beta} \psi(s) ds,$$

$$\alpha \leq \vartheta \leq \beta$$

für

$$\varphi(s) = \frac{f(s)}{g'(s)}$$

und

$$\psi(s) = g'(s) \cos[T g(s)]$$

oder

$$\psi(s) = g'(s) \sin[T g(s)]$$

auf jedes der N_2 Teilintervalle anwenden, wo $\frac{f(s)}{g'(s)}$ monoton ist, gelangen wir wegen

$$\frac{f(s)}{g'(s)} < \frac{MA}{\varepsilon^r}$$

und

$$\left| \int_{\beta_1}^{\beta_2} g'(s) \sin_{\cos}[T g(s)] ds \right| = \left| \int_{g_1}^{g_2} \sin_{\cos}[T g] dg \right| < \frac{2}{T}$$

zur Ungleichung

$$|J_2| < \frac{8MA}{\varepsilon^r T} N_2.$$

Zusammen mit der Ungleichung für das erste Integral ergibt das

$$|J| < 2MN_1 \varepsilon + \frac{8MAN_2}{\varepsilon^r T}.$$

Die Wahl von ε blieb bisher willkürlich (nur daß es klein sein sollte). Wählen wir nun

$$\varepsilon = \left(\frac{4A}{T} \right)^{\frac{1}{r+1}},$$

so erhalten wir die gewünschte Abschätzung

$$|J| < 2M(N_1 + N_2) \sqrt[r+1]{\frac{4A}{T}}.$$

Der imaginäre Teil von $(Q_0)_{mn}(s)$ kann einfach durch Multiplikation der Abschätzung mit 2 berücksichtigt werden. Somit ist die Formel § 62, (13) bewiesen.

Sachverzeichnis.

Die Zahlen geben die Seiten an.

- Ableitung einer Matrix** s. Differentiation.
- Absorption**, klassisch 4, **413ff.**, korrespondenzmäßig 9ff., Wahrscheinlichkeit der — 340ff., 400ff., negative — s. erzwungene Emission.
- Adiabatenpostulat** 6.
- Adiabatensatz** 333ff.
- Adjungierte Matrix** 31, 34, Funktion 36, —s Polynom 35.
- Aktivität**, optische 250ff., 351.
- Anharmonischer Oszillator** 200ff.
- Antistokessche Linie** 279, 408.
- Asymptotische Verteilung der Eigenfrequenzen** **372ff.**, 375, 377, 397, 403.
- Auswahlregeln für die Drehimpuls-Quantenzahlen** 143ff., 162ff., für den harmonischen Oszillator 125ff.
- Axialer Vektor** 156.
- BALMER-Formel** 184ff.
- Bewegungsgleichungen für abgeschlossene Systeme** 97ff., für nicht abgeschlossene Systeme 114ff.
- Bilinearform**, zu einer Matrix gehörige 44.
- Brechungsindex** 246.
- CURIESche Konstante** 231, 275.
- CURIESches Gesetz** 231.
- Dämpfungskonstante eines elektrischen Dipols** 249.
- Definite Form** 53, 61.
- Deformierbarkeit**, elektrische 246.
- Determinante einer Matrix** 24.
- Determinismus** 322ff.
- Diagonalelemente** 17, 18.
- Diagonalmatrix** 18, 24, geordnete — 29, 30.
- Diamagnetisches Moment** 229, —r FARADAY-Effekt 275.
- Diamagnetismus** 232ff.
- Dielektrizitätskonstante** 217.
- Differentiation einer Matrix nach der Zeit** 18, 27, nach einem Parameter 26, partielle — 38, 91.
- Dipolemission**, klassisch 4.
- Dipolmoment**, elektrisches 105.
- Direkte Multiplikation** 50, 51.
- Dispersion des Lichtes** 240ff., 404ff., negative — 247.
- Doppelbrechung** 245, elektrische 259ff., magnetische 267ff.
- Drehbarkeit**, freie 138, 144, 154ff.
- Drehimpuls**, Vertauschungsregeln 132ff., Vertauschbarkeit mit drehinvarianten Größen 135ff., Eigenwerte 139, irreduzible Bestandteile 139ff., Auswahlregeln 143ff.
- Drehinvariante Größen** 151ff., Vertauschbarkeit mit Drehimpuls 135ff.
- Drehung**, infinitesimale 136.
- Drehungsparameter** 257, 271.
- Drehungsvermögen** s. optische Aktivität.
- Eigenschwingungen**, einer Saite, klassisch 368, eines Hohlraums, klassisch 369ff., asymptotische Verteilung der — **372ff.**, 375, 377, 397, 403, Quantelung der — 376ff., quantenmechanische Begründung der Quantelung 383ff., eines Kristallgitters 383.

- Eigenvektor 58.
 Eigenwerte einer Matrix 56, Extremaleigenschaften der — 62ff., eines Matrixsystems 71, 304, einer physikalischen Größe 288ff.
 Einfache Matrixensysteme s. irreduzible.
 Einheitsmatrix 23.
 EINSTEINSche Übergangskoeffizienten 10, 248, 344, 403.
 EINSTEINS Schwankungsgesetze 386ff., 392ff.
 Einzelform 60.
 Einzelmatrix 40, 293, Eigenwerte einer — 72.
 Einzeltensor 51, zu einem Eigenwert eines hermiteschen Tensors gehöriger 61, 297, zu einem Intervall gehöriger 299.
 Emission, klassisch 4, korrespondenzmäßig, 8, bei Entartung der Energie 105f., 321, quantenmechanisch 400ff., erzwungene und spontane — 10, 248, 404.
 Entartung der Energie 103ff., Intensität der spontanen Emission bei — 105f., Störungstheorie bei — 205ff., 344ff., nicht aufhebbare — 212, 215.
 Entropie 363, der Strahlung 378ff., 387ff.
 Erwartungswert 292.
 Erzwungene Emission 10, 248, 404.
 Exponentialfunktion 36.
 Extremaleigenschaften der Eigenwerte 62ff.
 FARADAY-Effekt 245, 267ff., diamagnetischer und paramagnetischer 275.
 Feinstruktur 193, 236.
 Fluoreszenz 282ff.
 Form, hermitesche 44.
 Frequenzbedingung, BOHRsche 5, 98.
 Funktionen, Matrizen- 35ff., vertauschbarer Matrizen 68ff.
 Ganzzahligkeit der inneren Quantenzahl 162ff.
 Gekreuzte Felder 190ff., 234ff.
 Gemenge 305ff., als Mischung reiner Fälle 317ff.
 Geordnete Diagonalmatrix 29, 30.
 Gestürzte Matrix s. transponierte.
 Gewicht eines Eigenwertes 103, 321.
 Gleichartige Matrizen 21.
 Gruppe 110.
 Harmonischer Oszillator 122ff., Systeme —er —en 129ff.
 Hauptachsensystem 57.
 Hauptachsentransformation 54ff., vertauschbarer Matrizen 65ff.
 Hauptquantenzahl 186.
 HEISENBERGsche Regel (Fluoreszenz) 288.
 Hermitesche Matrix 19, 31, — Form 44, Eigenwerte derselben 56.
 HERTZsche Formel für Dipolemission 4.
 HILBERT-Raum 45ff., 294ff.
 Hohlraum-Strahlung s. Strahlung.
 Infinitesimale Transformation 111, 424, der Drehgruppe 136.
 Innere Quantenzahl s. Quantenzahl.
 Integrale der Bewegungsgleichungen 107ff., 426, beim Wasserstoffatom 177ff.
 Interferenz der Wahrscheinlichkeiten 285, 325.
 Invariante einer Gruppe 111, s. auch Relativinvariante.
 Irreduzible Matrixensysteme 73ff., 93, Bestandteile 77, — der Drehimpulsmatrizen 139ff.
 Irreversibilität 363f.
 JACOBI-POISSONSches Klammersymbol 89f., 420ff.
 JACOBISCHE Identität 92, 425.
 Kanonische Bewegungsgleichungen 97ff., 114ff., 425f., Transformationen 94, 420ff., Vertauschungsregeln 86ff., 294, Variable 90,

- 420ff., Matrizenysteme, Existenz und Eindeutigkeit 128.
KERR-Effekt 245, 259ff.
 Klammersymbol 23, 26, **JACOBI-POISSONS**ches — 89f., 420ff., **LAGRANGES**ches — 422.
 Kollektiv 306.
 Kombinationsgesetz 5, 17.
 Kontinuierliche Gruppen 109ff.
 Kontragrediente Matrix 32, Transformation 48.
 Koppelung von Rotationen und Schwingungen bei der zweiatomigen Molekel 203ff.
 Korrespondenzprinzip, 7ff.
 Kristallgitter, Eigenschwingungen eines —s 383, Schwankungerscheinungen an einem — 399.
 Länge eines Vektors im **HILBERT**-Raum 44.
LARMOR-Frequenz 191.
Magnetelektron 167, 193.
 Magnetische Quantenzahl 143.
 Magnetisierung 225ff.
 Magneton, **BOHR**sches 227, 231.
 Magnetonenanzahl 230, 271.
 Matrix 16, hermitesche 19, 31, quadratische 21ff., rechteckige 27ff., schiefssymmetrische 31, unitäre 32.
 Matrix-Matrix s. Übermatrix.
 Matrizenanalyse 37ff.
 Matrizenfunktionen 35ff., 68ff.
 Meßbarkeit 303ff.
 Molekel, zweiatomige 171ff., Koppelung von Rotationen und Schwingungen 203ff., Polarisierbarkeit 220ff.
 Moment eines elektrischen Dipols 4, 105, magnetisches — 226ff., diamagnetisches und paramagnetisches — 229f.
 Nebenquantenzahl 189.
 Nullpunktenergie 127, 385, 395, 398.
 Nullteiler 22.
Optische Aktivität 250ff., 351.
 Orthogonal s. unitär.
 Orthogonalsystem, vollständiges normiertes 46.
 Oszillationsenergie einer zweiatomigen Molekel 175, 204.
 Oszillator, harmonischer 122ff., anharmonischer 200ff., Hohlraum-, s. Eigenschwingung.
Paramagnetisches Moment 229, —r **FARADAY**-Effekt 275.
 Paramagnetismus 230.
 Partielle Differentiation 38, 91.
 Pendelbahn 190, 192.
 Permutationsmatrix 32.
 Phasenkonstante 3, 9, 17.
 Phasenmatrix 24, 33, 67.
PLANCKsches Strahlungsgesetz 366, 374ff., 377ff., 408ff.
 Polarisierbarkeit, elektrische 212ff.
 Positiv definit s. definit.
Quantenzahl, Haupt- 186, innere — 144, Auswahlregel für die innere 147, 163, Ganzzahligkeit der inneren 162ff., magnetische 143, Neben- 189, radiale 170.
Radiale Quantenzahl 170.
RAMAN-Strahlung s. Streustrahlung,
RAYLEIGH-JEANSsche Strahlungsformel 374, 392, 398.
 Razemisches Gemisch 258, 353.
 Reduzible Matrizenysteme 73.
 Reiner Fall 301, als spezielles Gemenge 312ff.
 Relativinvarianten 103ff., 152.
 Relativwahrscheinlichkeiten 294ff., bei Entartung 299f.
 Resonanz 282, innere — entarteter Systeme 344ff., zwischen gekoppelten Systemen 356ff.
 Resonanzschwebung 351, 356ff.
 Reziproke Matrix 24.
 Richtungsentartung 138, 151.
 Rotation, Koppelung von — und Schwingung 203ff.

Rotationsenergie 175, 204.
 Rotationsquantenzahl 203.
 RYDBERG-Konstante 184.

 Säkulare Störung 348ff.
 Säkulargleichung 207, 348.
 Sättigung, paramagnetische 231.
 Schiefsymmetrische Matrix 31.
 Schwankungserscheinungen der
 Strahlung 366.
 Schwankungsgesetze, EINSTEINS
 386ff., 392ff.
 Schwarze Strahlung 374ff.
 SCHWARZSche Ungleichung 52.
 Selbstadjungierte Matrix s. hermi-
 tische, — Polynome 35.
 Selektive Reflexion der Resonanz-
 linie 326.
 Semidefinit 53, 61.
 Skalares Produkt 43.
 Spektroskopische Stabilität 13, 106,
 161.
 Spiegelungsinvarianten 151ff.
 Spontane Übergangswahrschein-
 lichkeit 11, 248, 404.
 Spur einer Matrix, Definition 25,
 relativinvariante — 103ff., 152ff.
 Stabilität, spektroskopische 13, 106,
 161.
 Stärkefaktor der Dispersion 247ff.,
 273f.
 STARK-Effekt 192, 212ff., linearer,
 bei Wasserstoff 213, quadrati-
 scher 215ff.
 STARK-Effektquantenzahl 191.
 Statistische Deutung der Quanten-
 mechanik 288ff.
 STIRLINGSche Formel 379.
 STOKESSche Regel 279, 408.
 Störungsrechnung für abgeschlos-
 sene nicht entartete Systeme 194ff.
 für abgeschlossene entartete Sy-
 steme 205ff., 345ff., für nicht
 abgeschlossene Systeme 236, im
 Resonanzfall 282, 342, 344f.
 Strahlung in einem Hohlraum,
 klassisch 367ff., quantenmecha-
 nisch 383ff., s. auch Eigen-

 Born-Jordan, Quantenmechanik.

schwingung, in Wechselwirkung
 mit Materie 400ff., Schwan-
 kungsgesetze der — 386ff., 392ff.
 Strahlungsformel, PLANCKSche 374ff.,
 377ff., 408, RAYLEIGH-JEANS-
 sche 374, 392, 398, WIENSche 374,
 392.
 Streustrahlung 276ff., 404ff.
 Stufenmatrix 29, 59, 103.
 Summenregeln von ORNSTEIN und
 BURGER 161.
 Summensatz von THOMAS und
 KUHN 12, 85, 244.
 Suszeptibilität 231ff.

 Teilmatrix, s. Untermatrix.
 Thermisches Gleichgewicht zwischen
 Strahlung und Atomen 408ff.
 THOMAS und KUHN, Summensatz
 von 12, 85, 244.
 THOMSONSche Formel 243.
 Trägheitsmoment, elektrisches 227,
 233.
 Transformation, Hauptachsen- 54ff.,
 65ff., kanonische 94, 420ff., kon-
 tragrediente 48, unitäre 45.
 Transponierte Matrix 30.

 Übergangskoeffizienten, EINSTEIN-
 sche 10, 248, 344, 403.
 Übergangswahrscheinlichkeit 326ff.,
 pro Zeiteinheit 362, bei Absorp-
 tion 340ff., 402, bei Emission 403,
 bei Streuung 406.
 Übermatrix 22, 28.
 Unitäre Matrix 32, Transformation
 45.
 Untergruppe 110.
 Untermatrix 28.

 VERDETSche Konstante 273f.
 Verschmelzung 75ff., 87, 99.
 Vertauschbare Matrizen 23, Haupt-
 achsentransformation — r — 65ff.,
 Funktionen — r — 68ff.
 Vertauschungsregeln, kanonische
 86ff., 294, für den Drehimpuls
 132ff.

Verweilzeit 357, 360.

Virialsatz 100, 184.

Wahrscheinlichkeit der Absorption

s. Absorption, der Emission s.

Emission, —en quantenmechanischer Größen in stationären Zuständen 292ff., s. auch Relativwahrscheinlichkeiten, Interferenz der — 285, 325, Symmetriengesetz der — 298, 409.

Wahrscheinlichkeitsamplitude 325.

Wahrscheinlichkeitsvektor 300ff.

Wasserstoffatom 177, in gekreuzten Feldern 190ff., 234ff., linearer STARK-Effekt beim — 213.

Wechselwirkung zwischen Materie und Strahlung 400ff.

WIENSches, Strahlungsgesetz 374, 389f. 392, 398.

Zeemaneffekt 147ff., 158ff., 225ff.

Zeit als Zahlparameter 114, 120ff., 327.

Zeitmittelwert einer Matrix 18, 293.

Zentralbewegung 168ff.

Zerlegbare Matrixensysteme s. reduzible.

Zerlegung der Einheit, die zu einem hermiteschen Tensor gehörige 61, 297.

Zweiatomige Molekel s. Molekel.