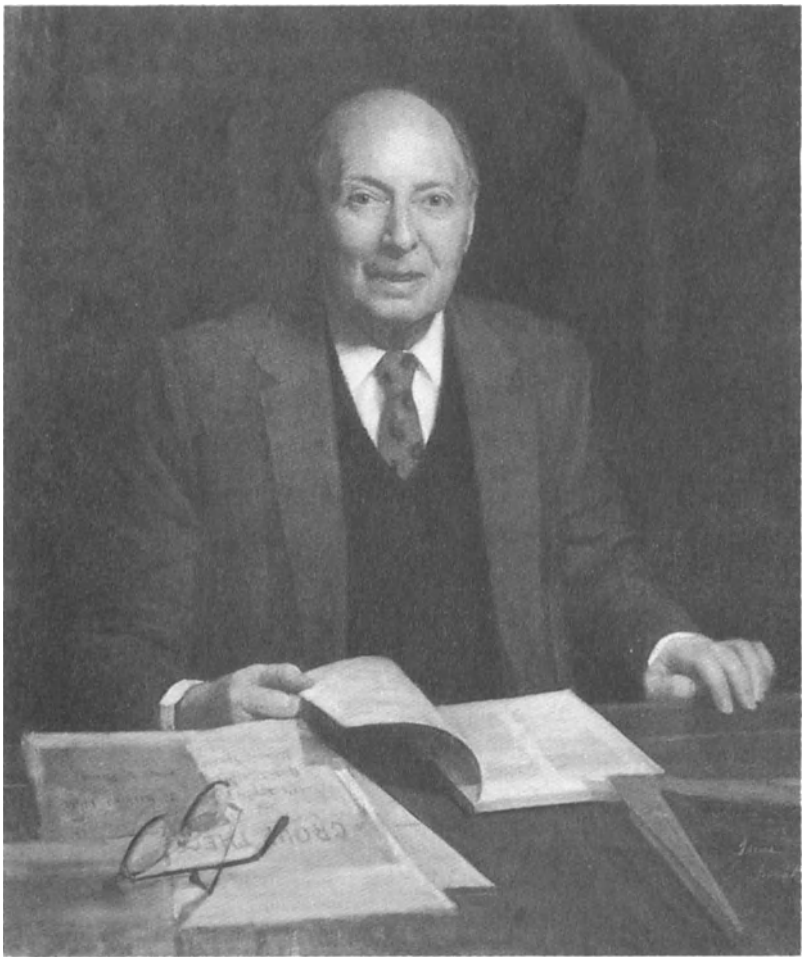
The Collected Works of Eugene Paul Wigner

Volume I

The Collected Works of Eugene Paul Wigner

Part A · Volume I



Eugene Paul Wigner

The Collected Works of Eugene Paul Wigner

Part A

The Scientific Papers

Editor: Arthur Wightman

Annotated by

Nandor Balazs Herman Feshbach Brian Judd Walter Kohn
George Mackey Jagdish Mehra Abner Shimony Alvin Weinberg
Arthur Wightman

Part B

Historical, Philosophical, and Socio-Political Papers

Editor: Jagdish Mehra

Annotated by

Conrad Chester Gérard Emch Jagdish Mehra

The Collected Works of Eugene Paul Wigner

Part A

The Scientific Papers

Volume I
Part I: Eugene Paul Wigner - A Biographical Sketch
Part II: Applied Group Theory 1926–1935
Part III: The Mathematical Papers

Volume II
Nuclear Physics

Volume III
Part I: Particles and Fields
Part II: Foundations of Quantum Mechanics

Volume IV
Part I: Physical Chemistry
Part II: Solid State Physics

Volume V
Nuclear Energy:
Part I: Eugene Wigner and Nuclear Energy
Part II: Memoir of the Uranium Project
Part III: Articles, Reports, and Memoranda on Nuclear Energy
Part IV: The Wigner Patents

Part B

Historical, Philosophical, and Socio-Political Papers

Volume VI
Philosophical Reflections and Syntheses

Volume VII
Historical and Biographical Reflections and Syntheses

Volume VIII
Socio-Political Reflections and Civil Defense

The Collected Works of

Eugene Paul Wigner

Part A

The Scientific Papers

Volume I

Part I: Eugene Paul Wigner – A Biographical Sketch
By Jagdish Mehra

Part II: Applied Group Theory 1926–1935
Annotated by Brian R. Judd

Part III: The Mathematical Papers
Annotated by George W. Mackey

Edited by Arthur S. Wightman

Springer-Verlag Berlin Heidelberg GmbH

Arthur S. Wightman
Department of Physics, Princeton University
Joseph Henry Laboratories, Jadwin Hall
Princeton, NJ 08544, USA

Jagdish Mehra
7830 Candle Lane
Houston, TX 77071, USA

Brian R. Judd
The Henry A. Rowland Department of Physics and Astronomy
The Johns Hopkins University
Homewood Campus
Baltimore, MD 21218, USA

George W. Mackey
Department of Mathematics, Harvard University
1 Oxford Street
Cambridge, MA 02138, USA

The photograph of E. P. Wigner on page II is taken by R. P. Matthews
from an oil painting by June Hecht 1991

ISBN 978-3-642-08154-5 ISBN 978-3-662-02781-3 (eBook)
DOI 10.1007/978-3-662-02781-3

Library of Congress Cataloging-in-Publication Data
(Revised for pt. A, v. 1, pt. 1-3)
Wigner, Eugene Paul, 1902- The collected works of Eugene Paul Wigner.
Includes bibliographical references.
Contents: pt. A. The scientific papers - v. 1, pt. 1. A biographical sketch/by Jagdish Mehra.
pt. 2. Applied group theory 1926-35 annotated/by Brian Judd.
pt. 3. Mathematical papers annotated/by George Mackey -
- v. 5. Nuclear energy. 1. Mathematical physics. I. Weinberg, Alvin Martin, 1915- . II. Title
QC19.3.W54 1992 530 92-38376

This work is subject to copyright. All rights are reserved, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilm or in any other way, and storage in data banks. Duplication of this publication or parts thereof is permitted only under the provisions of the German Copyright Law of September 9, 1965, in its current version, and permission for use must always be obtained from Springer-Verlag Berlin Heidelberg GmbH.
Violations are liable for prosecution under the German Copyright Law.

© Springer-Verlag Berlin Heidelberg 1993
Originally published by Springer-Verlag Berlin Heidelberg New York in 1993

The use of general descriptive names, registered names, trademarks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.

Typesetting of the Annotations, Bibliographies to Parts II and III, and of the article on p. 453ff.:
Springer TEX in-house system
55/3140-543210 - Printed on acid-free paper

Editors' Preface

The papers have been divided, necessarily somewhat arbitrarily, into two parts

Part A: Scientific Papers
Part B: Historical, Philosophical, and Socio-Political Papers

Within each part, the papers have been divided by subject, and within each subject printed chronologically. With some exceptions, every scientific paper is reprinted in its original form. One class of exceptions consists of papers that are simply translations into Hungarian from German or English; they are omitted, but listed in the bibliographies. Scientific papers originally in Hungarian have been translated into English. Some of the papers of Volume V/ Part III, Articles, Reports, and Memoranda on Nuclear Energy, have been reset and the figures redrawn. The originals were declassified reports, some in nearly illegible shape. Some reports and patents in Volume V/ Part III and Part IV are listed by title only. In contrast to the scientific papers where the coverage is essentially complete, in Part B, a selection has been made. We believe it is representative of Wigner's far ranging concerns. The five books in which Wigner was involved as author, co-author, or lecturer are not reprinted in the Collected Works, but are noted in the annotations and bibliographies.

Jagdish Mehra
Arthur S. Wightman

Contents

PART I

Eugene Paul Wigner – A Biographical Sketch
By Jagdish Mehra 3

PART II

Applied Group Theory 1926–1935

Annotation by Brian R. Judd 17

On Noncombining Terms in the New Quantum Theory. Part I / Über nicht kombinierende Terme in der neueren Quantentheorie. Erster Teil 34

On Noncombining Terms in the New Quantum Theory. Part II / Über nicht kombinierende Terme in der neueren Quantentheorie. Zweiter Teil 43

Some Consequences of the Schrödinger Theory for the Term Structure / Einige Folgerungen aus der Schrödingerschen Theorie für die Termstrukturen 53

Correction to the Article: Some Consequences of the Schrödinger Theory for the Term Structure / Berichtigung zu der Arbeit: Einige Folgerungen aus der Schrödingerschen Theorie für die Termstrukturen 82

On the Conservation Laws in Quantum Mechanics / Über die Erhaltungssätze in der Quantenmechanik 84

On the Explanation of Some Properties of Spectra in the Quantum Mechanics of the Spinning Electron. Part I / Zur Erklärung einiger Eigenschaften der Spektren aus der Quantenmechanik des Drehelektrons. Erster Teil (with J. von Neumann) 91

On the Pauli Exclusion Principle / Über das Paulische Äquivalenzverbot (with P. Jordan) 109

On the Explanation of Some Properties of Spectra in the Quantum Mechanics of the Spinning Electron. Part II / Zur Erklärung einiger Eigenschaften der Spektren aus der Quantenmechanik des Drehelektrons. Zweiter Teil (with J. von Neumann) 130

On the Explanation of Some Properties of Spectra in the Quantum Mechanics of the Spinning Electron. Part III / Zur Erklärung einiger Eigenschaften der Spektren aus der Quantenmechanik des Drehelektrons. Dritter Teil (with J. von Neumann) 152
On the Structure of Diatomic Molecular Spectra in Quantum Mechanics / Über die Struktur der zweiatomigen Molekelspektren nach der Quantenmechanik (with E. E. Witmer) 167
On the Elastic Eigenvibrations of Symmetric Systems / Über die elastischen Eigenschwingungen symmetrischer Systeme 195
On a Sharpened Sum Rule / Über eine Verschärfung des Summensatzes 209
On the Operation of Time Reversal in Quantum Mechanics / Über die Operation der Zeitumkehr in der Quantenmechanik 213
On the Experimental Test of the Conservation Law of Spin / Über die experimentelle Prüfung des Spinerhaltungssatzes (with R. Döpel and K. Gailer) 227
Separation of Rotational Coordinates from the Schrödinger Equation for N Particles (with J. O. Hirschfelder) 229

Bibliography of Papers on Applied Group Theory 1926–1935 236

PART III

The Mathematical Papers

Annotation by George W. Mackey 241
On Remarkable Discrete Eigenvalues / Über merkwürdige diskrete Eigenwerte (with J. von Neumann) 291
On the Behavior of Eigenvalues in Adiabatic Processes / Über das Verhalten von Eigenwerten bei adiabetischen Prozessen (with J. von Neumann) 294
On an Algebraic Generalization of the Quantum Mechanical Formalism (with P. Jordan and J. von Neumann) 298
On Unitary Representations of the Inhomogeneous Lorentz Group 334
Minimally Almost Periodic Groups (with J. von Neumann) 390
On Representations of Certain Finite Groups 395
Localized States for Elementary Systems (with T. D. Newton) 402
On a Class of Analytic Functions from the Quantum Theory of Collisions 409
On the Connection Between the Distribution of Poles and Residues for an R Function and Its Invariant Derivative 441
Derivative Matrix and Scattering Matrix 453
Representations of the Galilei Group (with E. Inönü) 459
Simplified Derivation of the Properties of Elementary Transcendentals .. 473

On the Contraction of Groups and Their Representations (with E. Inönü) 488
On a Particular Type of Convergence to a Singular Matrix (with E. Inönü) 503
Application of the Rayleigh-Schrödinger Perturbation Theory to the Hydrogen Atom 506
Significance of Loewner's Theorem in the Quantum Theory of Collisions (with J. von Neumann) 508
Characteristic Vectors of Bordered Matrices with Infinite Dimensions I 524
Characteristic Vectors of Bordered Matrices with Infinite Dimensions II 541
On a Theorem of Magnus (with Edwin Hewitt) 546
Normal Form of Antiunitary Operators 551
Phenomenological Distinction Between Unitary and Antiunitary Symmetry Operators 556
On Weakly Positive Matrices 559
Unitary Representations of the Inhomogenous Lorentz Group Including Reflections 564
On the Matrices Which Reduce the Kronecker Products of Representations of S. R. Groups 608
Unimodular Matrices Homomorphic to Lorentz Transformations in $n \geq 2$ Spacelike Dimensions (with A. W. Sáenz) 655
On a Generalization of Euler's Angles 662
Condition That the Irreducible Representations of a Group, Considered as Representations of a Subgroup, Do Not Contain Any Representation of the Subgroup More Than Once 673
Orthogonalization Methods (with H. C. Schweinler) 679
Restriction of Irreducible Representations of Groups to a Subgroup 681
Condition That All Irreducible Representations of a Compact Lie Group, If Restricted to a Subgroup, Contain No Representation More Than Once (with F. E. Goldrich) 690
Reduction of Direct Products and Restriction of Representations to Subgroups: The Everyday Tasks of the Quantum Theorists 697

Bibliography of Mathematical Papers 714

PART I

Eugene Paul Wigner: A Biographical Sketch

Eugene Paul Wigner: A Biographical Sketch

By Jagdish Mehra

Family Background and Upbringing

Eugene Paul Wigner (Jenö Pál Wigner) was born in Budapest, Hungary, on 17 November 1902. His father Anthony (Antal) Wigner and mother Elizabeth (Erzsébet) née Einhorn, were both descended from Austrian Jewish families. Anthony's family originally had the name 'Wiegner', meaning 'candlemaker', but his grandfather dropped the 'e' in the name. Anthony's father died when he was very young, which made it difficult for him to get a regular education or to have 'great fantasies for his life, and if he did, he did not share them'. Anthony Wigner became a director in a leather-tanning factory.

Anthony and Elizabeth Wigner had three children: the eldest, a daughter, Bertha; Eugene, and a daughter, two years younger than Eugene, Margit (who married P.A.M. Dirac in 1934).

Eugene had his early education at home between the ages of five and ten, from a lady whom his mother employed to teach him and his sisters. She taught them reading, writing and arithmetic, including the multiplication table. When Eugene was ten, the teacher left, and he was enrolled in the Lutheran High School, a private and prestigious gymnasium. His teachers were excellent there, eager to impart knowledge.

At the age of thirteen, Eugene fell ill with lung disease and had to spend some time in a sanatorium in a small place called Breitenstein in Austria. During that fall he had to take a special examination covering the material he had missed during his illness, and it was then that he first encountered his mathematics teacher László Rátz, from whom he apparently learned the most. 'Rátz did unbelievably much to arouse the attention of students toward mathematics. He founded the *Mathematics Journal for Secondary Schools* and wrote books which elucidated the simplest solutions of the problems posed there.'

Rátz gave private lessons in mathematics to Eugene Wigner's friend John von Neumann, who was also a student in the Lutheran High School, but one class junior, although he was two years ahead in mathematics. 'Johnny was a most unusual person, a marvellously quick thinker, and was recognized as such in high school.' Wigner admired von Neumann greatly, 'but I was not intimidated by him. I did not want to be compared with him or to compete with him. It was clear that he was a much better mathematician than I was, and a better scientist. But I knew more physics.'

Sandor Miklós was the physics teacher in high school, and author of 'an excellent physics book', while Andras Kubascka taught natural science, mostly

elementary zoology. Miklós, the physics teacher, 'showed nice experiments and the experiments were always successful. He taught solidly and well. After two years of study of physics with him, the first two years of physics either at the Budapest Technical University or the Berlin Technische Hochschule seemed to be almost mere repetition.'

Eugene had his *Bar Mitzvah* on 17 November 1915, his thirteenth birthday.

Soon after the end of World War I, the communists took over Hungary. 'My father and I were deeply against Communism.' Evidently, it was already in his boyhood that Wigner learned to abhor Communism. Anthony Wigner took his family to Austria for a while, to a place south of Vienna, where he had a cousin. But the communists were soon overthrown and in 1919 the family returned to Budapest. Anthony decided that the family should change its religion to Lutheran Protestant, more as an anti-Communist gesture than as a result of compulsion or religious conviction.

Higher Studies

Eugene Wigner completed high school in 1920. He found that he was principally interested in physics and mathematics, but that it was not easy to find a job as a scientist. He could have obtained a reasonably good job in the tannery where his father worked. So he thought he would study chemical engineering, and his father agreed, but Eugene told him: 'To be honest, I want to become a scientist.' He had in mind a real scientist who works on theoretical and experimental development of physics and not just a teacher of physics. It was then decided that he would study chemical engineering and later on work in the tannery. He enrolled in the technical university at Budapest for the year 1920–21 to study chemical engineering, and the next year switched to the Technische Hochschule in Berlin. There was much activity in theoretical physics in Berlin but it was primarily concentrated at the University of Berlin where Albert Einstein, Max Planck, Max von Laue, Walther Nernst, among others, were active. The Technische Hochschule was primarily devoted to applications, and among these chemistry was considered important and fundamental. But even at the Technische Hochschule Wigner chose to spend most of his time in trying to learn as much theoretical physics as possible. In his regular course work Wigner mostly pursued lectures and laboratory work in inorganic chemistry. For his *Diplomarbeit* he worked on the crystal structure of rhombic sulphur under the supervision of the chemist Hermann Mark, completing his degree in the spring of 1924.

Wigner then began doctoral work for a degree in chemical engineering ('Dr. Ing.') under Michael Polanyi. His thesis, entitled '*Bildung und Zerfall von Molekülen*' [1] dealt with chemical reaction rates and the formation of molecules. 'They have skill enough to collide with the right energy. Polanyi accepted my proposal that angular momenta are quantized. Atoms collide with angular momentum in a proportion consistent with Planck's constant.' On studying the equilibrium between the electron and the excited atom, 'we

can predict that the energy levels are not absolutely precise, that they have a spread. However, with angular momentum, this was not so. Hence, we said that the electron's angular momentum had its definite value. This, of course, was a very feeble expression of the uncertainty principle.' Wigner obtained the 'Dr. Ing.' degree in 1925.

Return to Hungary and Call to Berlin

Upon the completion of his degree, Wigner returned to Budapest and went to work in the tannery. He believed he would not have the opportunity to continue his science there. However, he did have one link to the physics he had left behind: a subscription to *Zeitschrift der Physik*. It was there that he first read the article by Born and Jordan on matrix mechanics [2]. 'This paper convinced me that the human being had enough imagination to understand microscopic processes.' Wigner knew matrices quite well as a result of his interest in crystals; the symmetries of a crystal were described in terms of matrices. So he understood the Born-Jordan article on matrix mechanics right away, as well as the following article by Born, Heisenberg and Jordan [3]. From these papers Wigner learned how to calculate energy levels and to describe atomic transitions.

Meanwhile, Michael Polanyi, Wigner's doctoral thesis advisor, had been active in his behalf. Polanyi had been quite impressed by Wigner, and recommended him to Richard Becker who had just become Professor of Theoretical Physics at the Technische Hochschule in Berlin. Becker invited Wigner to be his research assistant, a post that Wigner accepted with alacrity. Wigner's duties included attending Becker's lectures; this was in response to Becker's request that Wigner point out the obscurities in his lectures. 'But this was not necessary because Becker spoke clearly and gave excellent explanations.' Wigner was also asked to work with G. Wessenberg, a *Privatdozent* at the Technische Hochschule. Wessenberg remarked to Wigner, '...it's a miracle that in a crystal the atoms are often arranged along the axis or plane of symmetry, why?' 'I noticed that if you have an axis of symmetry, the potential is an extremum (a maximum or minimum), but the probability of its being a maximum is much larger, and a similar consideration applied to the plane of symmetry. Wessenberg said, "Perhaps you are right, but it is important to give a more elegant demonstration!" ' This sent Wigner back to his study of group theory, in particular, to the classic *Lehrbuch der Algebra* by Heinrich Weber, of which the second volume deals with group theory. Wigner then produced a group-theoretical proof, but it was not published.

Wigner then took up the problem of the symmetry of atomic eigenfunctions under permutation of electrons. He generalized an article of Heisenberg's [4], which treated the case of two electrons, first to three electrons and then to four or more. All this was for electrons without spin. Then, in three joint articles with John von Neumann, the whole theory was generalized to the case in which the electron's spin is described by the Pauli theory [5, 6].

All these results were given a systematic and pedagogical exposition in Wigner's book *Gruppentheorie und ihre Anwendung in der Theorie der Atomspektren* [7]. He was urged to write this book by Max von Laue who was also helpful in persuading the publisher, Vieweg, that it should be published. The methods of group representation theory used by Wigner had been developed by Frobenius and Schur in the context of the theory of finite groups. They were sufficient for Wigner's application to the permutation group of the electrons. For the rotation group Wigner needed results and methods of I. Schur and H. Weyl. What Wigner offered as a preliminary to his thoroughgoing practical application of these group-theoretical methods to atomic physics was a fundamental analysis of the notion of symmetry in quantum mechanics. He showed how it reduces to a problem in group representation theory.

Edward Teller and Leo Szilard in Berlin

Eugene Wigner had not known either Edward Teller or Leo Szilard, his compatriots from Hungary, in Budapest. He first met them when he was in the Technische Hochschule in Berlin. Szilard was then working in Berlin, while Teller was working under Werner Heisenberg in Leipzig. Teller used to visit Berlin and Wigner occasionally went to Leipzig. In matters of politics they agreed, so they mainly discussed problems of theoretical physics with each other. 'All of us realized that there was a danger of submission of all countries to the Communist threat. Fascism at that time was still not very serious.' And, of course, the evil of Fascism', especially the threat of anti-semitism, was a great shock to both Wigner and Teller and to many, many other people.

Göttingen

In 1927 Eugene Wigner was invited to Göttingen to work as an assistant to David Hilbert. This came about because for many years Arnold Sommerfeld used to arrange for a young physicist to go to Göttingen to help Hilbert keep abreast of current developments in physics. But Hilbert was very sick when Wigner arrived in Göttingen; he suffered from pernicious anemia, and Wigner saw him only a few times and did not have close contact with him.

However, he worked hard on physics, and became acquainted with the other people who were there. He got to know Max Born whom he liked a great deal. Among the younger physicists were Pascual Jordan and Walter Heitler. John von Neumann visited Göttingen frequently. Von Neumann and Lothar Nordheim had attended lectures by Hilbert on quantum mechanics during the winter semester of 1926–27 and had written them up for publication. Von Neumann was working on the spectral theory of self-adjoint operators, a subject fundamental both for mathematics and quantum theory. Among the experimental physicists whom Wigner got to know was James Franck, who in the famous

Franck-Hertz experiment, had shown the inelastic collisions of electrons with atoms yield excitation energies in agreement with those predicted from spectroscopy according to Bohr's theory. Wigner came into close contact with Jordan and together they published a paper on Pauli's exclusion principle [8], which turned out to be very important in the history of quantum field theory. This work was inspired by previous work by Pascual Jordan and Jordan and Oskar Klein, who gave an algebraic treatment (second quantization) for nonrelativistic charged bosons. 'But it was Jordan who first thought of the idea which we (Jordan and Wigner) published together. However, the mathematical proof, that this was the only way to explain Pauli's exclusion principle, was mine, and it was based on the theory of group representations – but the principal idea was Jordan's. ... He was deeply impressed by my idea of unitarity, but that was not difficult; it did not require imagination.'

Jordan and Wigner tried to write down a relativistic equation for the electron, just as several other people had done at that time; they did not succeed. One day Born received a letter from Paul Dirac with a description of his equation [9]. Born showed the letter to Jordan and 'Jordan told me, " It's a *marvellous* equation. It's a pity that we didn't invent it, but still it is very good it has been found." '

Wigner continued to work with von Neumann. One product was a joint paper, 'Über das Verhalten von Eigenwerten bei adiabatischen Prozessen' [10] 'that was a very interesting question: How do energy levels change if the atom is subjected to a slow transition? It was written by me (Wigner).'

Another product of Wigner's stay in Göttingen was a pair of influential papers written with Viktor Weisskopf [11]. These treated the problem of line width in Dirac's theory of atoms and radiation.

Back to Berlin

After the year in Göttingen, Wigner returned to Berlin, where he was again associated with Richard Becker. Upon his return in 1928 he was appointed *Privatdozent* (lecturer) at the Technische Hochschule and, in 1930, was given the title (untenured) 'nicht beamtlicher außerordentlicher Professor'.

It was during this period that Wigner wrote the above-mentioned book on group theory and its applications to atomic spectroscopy. The book invites comparison with two other classic books on group theory and quantum mechanics: that of Hermann Weyl in two German editions of 1928 and 1931 and its English translation of 1931 [12], and the book of B. L. van der Waerden [13]. These are of a flavour quite different from Wigner's. There is little doubt that most spectroscopists would prefer Wigner's book, and most mathematicians Weyl's or van der Waerden's. Mathematical physicists are well advised to be acquainted with all three.

Offer from Princeton University

One day in the fall of 1929, late in October or November, Eugene Wigner received a cable from Princeton University, offering him a job for half a year (one semester) in the spring of 1930. The salary was a considerable increase (seven times) over his salary in Berlin. John von Neumann also received an offer from Princeton University at the same time. (The University was using the income from the endowment funds of a chair, the Thomas D. Jones Professorship of Mathematical Physics, established in the mid-1920s to offer two temporary half-time positions. Hermann Weyl had been appointed to this Chair in 1928, but had resigned to return to Europe in 1929. It was to this Chair that Wigner was permanently appointed ten years later.) Their appointments had been recommended by Paul Ehrenfest of Leyden.

Upon arrival in Princeton Wigner was offered a continuation of this arrangement for the next five years (1930–1935). At the same time the Technische Hoschschule in Berlin offered him half-yearly appointments for the same period. Thus he could divide the academic year between Princeton and Berlin, and spend about three months each year with his parents in Hungary.

Princeton

Wigner prospered scientifically in Princeton. He ranged widely across theoretical physics and chemistry. Some of his European friends and coworkers came to North America. John von Neumann was there of course and, eventually, Leo Szilard and Edward Teller came, as did many other Hungarians and they got together from time to time. He was also fortunate in his students: Frederick Seitz, John Bardeen, and Conyers Herring worked with him and went on to become outstanding physicists. Their work set a high standard in theoretical physics.

Soon after Wigner came to Princeton dramatic developments took place in physics. In 1932, James Chadwick discovered the neutron and Carl Anderson the positron. Werner Heisenberg put forward the proposal that the atomic nucleus is composed of neutrons and protons; theoretical nuclear physics in a modern sense was born. Wigner contributed significantly to the development of nuclear physics through such papers as [14].

In 1933, the National Socialists seized power in Germany, and Wigner's position at the Technische Hochschule vanished. His job at Princeton was converted into a full-time position. Meanwhile, the Institute for Advanced Study had been created, and its newly appointed professors including Albert Einstein, Hermann Weyl and John von Neumann, were temporarily housed in Fine Hall, the University Mathematics Department's new building.

When Wigner first came to Princeton, he saw the Physics Department as 'a little behind times. Quantum mechanics was known to a very few. During my times here, first H. P. Robertson and then E. U. Condon were added to the department, and they were interested, but not as intensely as Johnny (von

Neumann) and I. Occasionally I went to Columbia to discuss matters with I. I. Rabi and with Breit. Gregory Breit, whom I admire very much, was at New York University. Rabi was at Columbia. Breit spent a year at the Institute (for Advanced Study) and we wrote a couple of papers together. One of them was on what people call the Breit-Wigner formula' [15].

When Wigner's contract as Visiting Lecture in Mathematical Physics expired in 1935, he was offered a three year term as Visiting Professor of Mathematical Physics. This was a promotion but only to a temporary position. He was bitterly disappointed, and sought a position elsewhere. With Gregory Breit as an advocate, he received an offer of a tenured position from the University of Wisconsin in Madison, Wisconsin, which he accepted.

Madison

Wigner was very happy in Madison, where he found the members of the Department very friendly and interested in his work. He continued to be extraordinarily productive, especially in nuclear physics.

Another notable event that occured about this time was that Wigner sent the manuscript of his paper 'On Unitary Representations of the Inhomogenous Lorentz Group' to the *American Journal of Mathematics*. It was rejected with the remark that 'this work is not interesting for mathematics'. Wigner mentioned the fact that it had been rejected to John von Neumann, who saw to it that the article was published in the *Annals of Mathematics* of which he was an Editor; so what is possibly Wigner's most famous paper saw the light of day [16].

In Madison, Wigner met Amelia Zipora Frank, who had been a student of physics. They were married on 23 December 1936, but she had heart disease and died on 16 August 1937.

When Princeton finally decided to offer Wigner an endowed chair with a permanent appointment, the Thomas D. Jones Professorship of Mathematical Physics, beginning in the fall of 1938, he accepted. (They had first offered it to J. H. Van Vleck, but he, newly appointed at Harvard, refused, and recommended the appointment of Eugene Wigner to the Chair.)

Consequences of the Discovery of Nuclear Fission

In 1939, Otto Hahn and Fritz Strassmann discovered nuclear fission. Niels Bohr brought the news of its discovery to the United States and spent the spring term in Princeton, working on the theory of nuclear fission with John Archibald Wheeler. The phenomenon contradicted the ideas of Enrico Fermi, whose previous explanation had been interpreted in terms of neutron absorption without fission but with the formation of transuranic elements. Fermi 'told us that there existed great danger that a nuclear explosion was possible. We also immediately understood the possibility of using the fission process for producing

energy.' Szilard had discussed the possibility of using a chain reaction involving neutrons for such purposes several years before but then he had no idea how to get the required neutrons; now fission might supply them. Wigner and Szilard were concerned that the Germans under Adolf Hilter's regime might develop nuclear weapons and with them conquer the world.

Wigner and Szilard were convinced that the security of the United States and of the other western democracies depended on having such weapons first, if indeed they could be made. They convinced Einstein to write a letter to President Franklin Delano Roosevelt explaining the situation. Einstein's letter was conveyed to Washington by Alexander Sachs, who saw to it that it got to the President. The result was a commitment by the Army and the Navy. The initial support was small, $6,000 mainly to buy graphite to slow Fermi's neutrons down at Columbia University. However, after the entry of the United States into the war, the Manhattan Project, as it was called, grew to a billion dollar program.

The Metallurgical Laboratory and the Atomic Bomb Project

In the initial stages, 'Fermi did the actual (experimental) work at Columbia University, and we were in close contact with each other. I decided we should investigate one of the factors, the so-called resonance absorption. Robert R. Wilson, later Director of Fermilab, and Ed. Creutz, later Associate Director of the National Science Foundation, were two experimental physicists whom I persuaded to work on this. They worked together beautifully with a skill and understanding I admire. We had constant conferences with Fermi and one day, literally the day after Pearl Harbor, Arthur Compton turned up in Princeton and said, "We want to organize a laboratory." '

In the meantime, Wigner had married again. On 4 June 1941, he married Annette Mary Wheeler, a physicist from Vassar College. 'My wife and I moved in April 1942 and I began to work at the University of Chicago in the so-called "Metallurgical Laboratory", where I was in charge of the theoretical physics section. Of course, it wasn't really a metallurgical laboratory; it was only given a name which does not tell you what it is about. The Metallurgical Laboratory was really about establishing a chain reaction. And not only a chain reaction but a chain reaction of great power which is a very different thing. A tiny chain reaction shows only that a chain reaction is possible. To make a giant chain reaction and to control it effectively you have to extract the best that it makes, you have to measure its power, and so many other essential things very correctly.'

The Metallurgical Laboratory occupied practically all of Eckart Hall, a large building at the University of Chicago. The group which Wigner headed had two or three moderately large rooms on the fourth floor. There were about twenty other theoretical physicists. 'If I had to pick one person who was most

indispensable at the Metallurgical Lab, I would pick Fermi, for two reasons: first, he was an excellent experimental physicist, so he knew how to build the chain reaction, how to take care of it, and many such things; second, he was very good in theoretical work.'

A sustained nuclear chain reaction was first achieved by Fermi's group on 2 December 1942 in a squash court under the Stagg Field stands at the University of Chicago.

Wigner's scientific collaborators were recruited by Compton. 'Again, I had unbelievable luck with the collaborators. One of them was Alvin M. Weinberg; the other one, who was technically equally able and far sighted, was Gale Young. I was enormously lucky. We designed the so-called Hanford reactor. We were firmly convinced right from the beginning that the chain reaction would go and the problem was to get it to go on a scale that would produce significant amounts of plutonium.'

Wigner's background in chemical engineering was a great help to him in the design process. 'I remembered engineering. I remembered, for instance, that corrosion can become a very important problem. I knew how to calculate the water flow through the tubes.'

Wigner's group produced a design for a plutonium-producing reactor, but the blueprints for what was actually constructed were produced by du Pont's engineers, and checked by Wigner's group. 'We did not disperse, and that was Compton's desire. Compton's and Harry Smyth's. They felt it was important for us to keep together and review the detailed du Pont plans, and in the meantime they told us to work on power reactors. That was a sort of trick to keep us together, and a successful trick. When du Pont drawings came, it was very important – with due respect to du Pont – that we review them.' (For a more detailed account of the relations between the theoretical group at the Metallurgical Lab and the du Pont Company, see Volume 5: Memoir of the Uranium Project.)

The Hanford reactor was built and worked well, producing plutonium. The Oak Ridge separation plant produced U^{235}. The Los Alamos Laboratory was created to design and produce bombs using these materials. Some members of the Metallurgical Lab went on to Los Alamos, Enrico Fermi in particular, Eugene Wigner did not. After some further work on power reactors, he returned to Princeton.

'Was it good or bad to have worked on the atom bomb? The debate has been going on for decades. We were mistaken about the Germans and their progress toward an atomic bomb. Some people say we should never have created this weapon. I have often been asked whether I would again help to create the first atomic bomb. And I have often asked myself this important question. In a way, I would like (to be able to say) that I regret having helped to build the bomb, because this response would please many people who ask me this question. But I really do not regret helping to build the bomb, either as an intellectual thing or an emotional one. The principle of building the bomb would have been discovered by someone. It was much better that it was done by America than by Hitler's Germany.

'In fact, my regret is that it was not done sooner. If we had begun trying seriously to control fission in 1939, we might have had an atomic bomb by the Winter of 1943–1944. At that time Stalin's army was still bottled up in Stalingrad. By the middle of 1945, when we first used the bomb, they had already overrun much of Central Europe. The Yalta Conference would have produced a document much less favourable to Russia, and even Communist China might have been set back. So I do not regret helping to build the atomic bomb.'

Clinton Laboratory, Oak Ridge

During the academic year 1945–46 Wigner was back at Princeton and his scientific activity resumed in full force, both in nuclear physics where he began a systematic study of nuclear reactions, and in relativistic quantum mechanics, but 1946–47 was spent at Oak Ridge. 'It was essentially (because of) Charlie Thomas of Monsanto Company. He told me that it was very important to have a good strong leadership for uranium power development. So, during 1946–47 I became Director of Research and Development for one year at a new laboratory, called the Clinton Laboratory (in Oak Ridge, Tennessee). It is now called the Oak Ridge National Laboratory. All my collaborators from Chicago had also gone there, especially Gale Young and Alvin Weinberg .... But Oak Ridge at that time was so bureaucratized that I am sorry to say I couldn't stand it, and we came back to Princeton after a year. The person who took over was Alvin Weinberg and he slowly, slowly improved things. I would not have had the patience to do what he did.'

However, Wigner continued to take a very close interest in the work at Oak Ridge National Laboratory. He wrote papers and, together with Alvin Weinberg, a book on nuclear reactor theory. He also developed there an interest in questions of Civil Defense.

Civil Defense

'I define Civil Defense as a passive defense which makes nuclear weapons much less effective. I want different kinds of governments to persist on earth – not only Communism. During the war, I knew Civil Defense should be undertaken, and I talked about that, but my work was on nuclear reactors. After World War II, I supported Civil Defense very vigorously as a political issue, but was still not working on it scientifically.'

'The danger was that the U.S. will be threatened and will be told: "Unless you do this and that which will make you powerless, tomorrow we will destroy all big cities and a large part of your population will perish." This threat had a name ("nuclear blackmail"), and I was very much after it and against it.'

'A country can defend itself against this kind of threat in two ways. The principal way ... is to have good Civil Defense so that the threat is not serious.

The second is also to have nuclear weapons so that it can retaliate to some degree.'

'There were two proposals (for Civil Defense): first, that we organize an evacuation of the cities, so that people can go to larger distances where the effect of nuclear weapons exploding over cities will be much less. But also: in the cities we should have shelters – good, heavy shelters, well-protected. This was a necessity then, and I consider it even now a necessity.'

Wigner continued to participate actively in questions of Civil Defense and other programs as a consultant to Oak Ridge National Laboratory until well into the 1980s, and always held on to the belief that 'our neglect of Civil Defense may become a true disaster.' In this context, he also supported SDI (the Strategic Defense Initiative, which proposed a defense against intercontinental ballistic missles using anti-missle missles). He considered it another form of Civil Defense, suggesting that 'Civil Defense is based on shelters and SDI on having a shield. SDI means you should protect against aggressive nuclear weapons, and that's a very good idea.'

Post World War II Themes

In the three decades following the second world war, Wigner continued to work in all the fields he had cultivated before the war except solid state physics, but he broadened his interests to include philosophical and semi-philosophical subjects. 'I also become interested in what is called the interpretation and epistemology of quantum mechanics. It is a weakness of old people that their interests spread out, and they know as the Germans say "nothing about everything". There is a saying, "A full professor knows nothing about everything, an associate professor knows something about a little, but an assistant professor knows it better." ' Wigner's many post-war students tell quite a different story about Wigner's incisiveness, his broad insight into physics, and his love of the subject.

Wigner retired as of 1 July 1971 but continued to work, actively most of the time until later years, for another two decades. Mary Wigner, his second wife, with whom he had a son David and a daughter Martha, died on 8 November 1977. On 29 December 1979, he married Eileen (Pat) Hamilton, the widow of a friend and colleague at Princeton, and, together, they continued to provide gentle companionship to each other.

Acknowledgement. I am grateful to Arthur S. Wightman for a critical reading of this biographical sketch and helpful suggestions. All unnumbered quotations are from the author's conversations and interviews with Eugene P. Wigner.

References

[1] M. Polanyi and E. Wigner: Bildung und Zerfall von Molekülen. Published part of Wigner's doctoral dissertation. Z. Phys. **33**, 429–434 (1925)
[2] M. Born and P. Jordan: Zur Quantenmechanik. Z. Phys. **34**, 858–888 (1925)
[3] M. Born, W. Heisenberg and P. Jordan: Zur Quantenmechanik II. Z. Phys. **35**, 557–615 (1926)
[4] W. Heisenberg: Über die Spektra von Atomsystemen mit zwei Elektronen. Z. Phys. **39**, 499–518 (1926)
[5] J. von Neumann and E. Wigner: Zur Erklärung einiger Eigenschaften der Spektren aus der Quantenmechanik des Drehelektrons I. Z. Phys. **47**, 203–220 (1928)
[6] J. von Neumann and E. Wigner: Zur Erklärung einiger Eigenschaften der Spektren aus der Quantenmechanik des Drehelektrons II, III. Z. Phys. **49**, 73–94 (1928) and **51**, 844–858 (1928)
[7] E. Wigner: Gruppentheorie und ihre Anwendung auf die Quantenmechanik der Atomspektren. Vieweg, Braunschweig 1931. (English translation: Group Theory and Its Applications to the Quantum Mechanics of Atomic Spectra. Academic Press, New York 1959)
[8] P. Jordan and E. Wigner: Über das Paulische Äquivalenzverbot. Z. Phys. **47**, 631–651 (1928)
[9] P. A. M. Dirac: The Quantum Theory of the Electron. Proc. Roy. Soc. **A117**, 610–624 (1928)
[10] J. von Neumann and E. Wigner: Über das Verhalten von Eigenwerten bei Adiabatischen Prozessen. Phys. Z. **30**, 467–470 (1929)
[11] V. Weisskopf and E. Wigner: Berechnung der natürlichen Linienbreite auf Grund der Diracschen Lichttheorie. Z. Phys. **63**, 54–73 (1930); Über die natürliche Linienbreite in der Strahlung des harmonischen Oszillators. Z. Phys. **65**, 18–29 (1930)
[12] H. Weyl: Gruppentheorie und Quantenmechanik. Hirzel, Leipzig 1928. Second German edition 1931; Group Theory and Quantum Mechanics. Methuen, London 1931
[13] B. L. van der Waerden: Die gruppentheoretische Methode in der Quantenmechanik. Springer, Berlin 1932
[14] E. Wigner: Über die Streuung von Neutronen an Protonen. Z. Phys. **83**, 253–258 (1933); On the Saturation of Exchange Forces. Proc. Natl. Acad. Sci. (USA) **22**, 662–666 (1936)
[15] G. Breit and E. P. Wigner: Capture of Slow Neutrons. Phys. Rev. **49**, 519–531 (1936)
[16] E. P. Wigner: On the Unitary Representations of the Inhomogeneous Lorentz Group. Ann. Math. **40**, 149–204 (1939)

PART II

Applied Group Theory 1926–1935

Applied Group Theory 1926–1935

Annotation by Brian R. Judd

1. Introduction

The articles written before 1931 can be regarded to a large extent as precursors to Wigner's seminal work "Gruppentheorie und ihre Anwendung auf die Quantenmechanik der Atomspektren", [1] which appeared in that year. The importance of this book can scarcely be overstated. In the preface to the English edition [2], which was published 28 years later, Wigner remarks that there was a great reluctance among physicists toward accepting group-theoretical arguments and the group-theoretical point of view. He also comments that this antipathy has subsequently vanished. One of the reasons for this is, of course, the influence that his book had on a generation of physicists. However, at the time of its publication, and for some years afterwards, opposition to group theory was real enough. In the rush to fit theory to experiment, simplicity in the mathematical treatment counted for more than elegance and generality. The uncompromising tone of Weyl's treatise on group theory and quantum mechanics [3] scarcely made the subject more accessible to physicists. They were more impressed by the success Slater [4] achieved in obtaining expressions for the energies of spectroscopic terms by elementary mathematical methods. Condon and Shortley [5] explicitly rejected the group-theoretical approach. Their algebraic methods were extended by Racah [6], and is was not until he published his fourth article [7] on complex spectra that group theory became an accepted part of theoretical atomic spectroscopy.

2. Groups in Atomic Theory

In his 1931 preface [1], Wigner stressed the importance of taking full advantage of the symmetry of a system before calculations are attempted. He had made similar arguments some four years earlier, when he pointed out that group theory could lead to exact results not just for the hydrogen atom but for more complicated systems [8]. The basic principles on which the book is based are contained in that article. The groups that are introduced involve the permutations of the electrons of the atom and coordinate rotations in 3-dimensional space. In the absence of external fields these are symmetry groups of the Hamiltonian, and their irreducible representations can be used to label

the eigenstates. As Wigner recognized in the first sentence of his article [8], the theory of representations is at the core of the analysis. In equation (10) he introduced matrix representations D^l_{jk} of the rotation group – what we would now call $SO(3)$ – without derivation and without an explicit construction. In particular, he did not follow what would be normal practice today and state that they could be derived by taking matrix elements of the operator

$$\exp(i\alpha L_z)\exp(i\beta L_x)\exp(i\gamma L_z)\ , \tag{1}$$

where $\mathbf{L}$ is an appropriate angular-momentum vector (measured in units of $\hbar$). In the absence of spin effects (which Wigner had in any case temporarily set aside) $\mathbf{L}$ can be taken to be the total orbital angular momentum of the electrons.

The matrices d^l_{jk} play a key role in putting group theory to work. For, in calculating an integral of the type

$$\int z\psi_i\psi_k\ , \tag{2}$$

which we find in Sect. 15 of [8], we do not seek explicit forms for ψ_i and ψ_k in order to perform the integration: rather, we only need their transformation properties under rotations, with the result that the integral reduces to a product over three D functions. In a subsequent article with von Neumann [9], a formula was given for an integrand of the type $D^A_{a0}D^B_{b0}D^C_{c0}$; but several years had to elapse before the general result for an integrand of the form $D^A_{ae}D^B_{bf}D^C_{cg}$ appeared (Ref [1], p. 204). For the special case (2), Wigner was able to avoid such complications by writing z as $r\cos\beta$ and using some recursion relations for the D functions. The inadvertent omission of some terms from these relations (an error that was quickly rectified [10]) led to a selection rule for electric-dipole radiation of the form $\Delta L = \pm 1$, thus overlooking the possibility $\Delta L = 0$. This oversight, and the fact that the reader could scarcely have been expected to detect it without being provided with an explicit form for the D functions, may have been an contributing factor in Wigner's decision, as stated in the 1931 preface to his book [1], to spell out all derivations in full detail.

Wigner recognized that integrals of the type (2) also occur in the evaluation of the Stark effect, and he was able to confirm (in equation (17) of [8]) the general expression (that was already known) for the displacements of the components of an atomic level. He was also able to make some general comments on the effect on an atom of a weak magnetic field. The paper [8] includes a calculation of the spectroscopic terms that occur in the configuration $p'p^2$. Today we would first evaluate the terms of p^2 as ^3P, ^1S and ^1D; then we would couple these states to the third p electron (which is inequivalent to the others) in all possible ways, thereby getting ^4S, ^4P, ^4D, ^2S, ^2P and ^2D (from ^3P), ^2P (from ^1S) and the three terms ^2P, ^2D and ^2F (from ^1D). Wigner, however, obtained these terms (with ^2F represented in his notation as ^2B) by examining the possible assignments of the magnetic quantum numbers m_s and m_l to the three electrons, subject to the constraint that not more than two

electrons could have identical sets of labels. Pauli's Exclusion Principle is implicit here, of course. The details of the assignments are set out in Table 1 of [8]. This procedure is certainly lengthier than that based on coupling; in fact, the advantages of using coupling techniques to combine inequivalent particles can be extended by regarding spin-up ($m_s = 1/2$) and spin-down ($m_s = -1/2$) electrons as distinct particles, as Shudeman [11] later showed so strikingly.

The paper [8] ends with a description of how spectroscopic terms can be described as normal or reflected. This terminology did not become widely used, mainly because the criterion for making the distinction involves both parity and the quantum number L. The latter ceases to be a good quantum number in the presence of spin-orbit interaction. Parity by itself provides a much better label, as Wigner himself recognized later [12].

3. Collaboration with von Neumann

In 1928 a collaboration between von Neumann and Wigner bore fruit in three lengthy articles [13, 12, 9]. Their aim was to give a more rigorous basis for those aspects of atomic spectra having to do with electron spin. Instead of postulating a vector $\mathbf{S}$ whose commutation properties paralleled those of the orbital angular momentum $\mathbf{L}$, they developed the kinematics of electron spin in terms of a two-dimensional irreducible representation of $SO(3)$. Much of the first article is taken up by the proof that the transformation matrix for a collection of electron spins can be formed by taking products of the individual matrices, a result that follows immediately from the expression (1) above if we make the replacements

$$\mathbf{L} \to \mathbf{S} = \mathbf{s}_1 + \mathbf{s}_2 + \ldots + \mathbf{s}_n \tag{3}$$

and then use the commutation of different $\mathbf{s}_i$ to separate out the rotation operators for each electron i. Instead of arguing along these lines, von Neumann and Wigner constructed an inductive proof that makes extensive use of the theory of groups. This put them in the position, at the end of that article [13], to calculate the number of times N_S a given total spin quantum number S occurs in a system of n electrons, these electrons being unrestricted by statistical constraints. There are $\binom{n}{x}$ ways of picking x spin-up electrons, each of which corresponds to a product state that contributes $\exp(i\alpha M_S)$, that is, $\exp[i(2x-n)\alpha/2]$, to the character. The total has now to be decomposed into the characters of irreducible representations of $SO(3)$. As successive characters are extracted from the sum, a pattern emerges that leads to the result

$$N_S = \binom{n}{\frac{1}{2}n - S} - \binom{n}{\frac{1}{2}n - S - 1} \,. \tag{4}$$

This agrees with the last equation of von Neumann and Wigner ([13], p. 220) provided we make the identification $q = 2S+1$. On putting $n = 3$, we find from equation (4) that the 2^3 product states provide a quartet and two doublets.

This is a special case of the system that Wigner had previously studied [14, 15] with regard to the permutation symmetry of product states. It is interesting to note that the second of the two articles uses group theory while the first does not. The reference to "non-combining" terms in this early work refers to the impossibility that a perturbation, symmetric with respect to the three electrons, could connect states of different symmetry types.

Just how permutation symmetry relates to the $SO(3)$ classification is a central topic of the second of Wigner's papers with von Neumann [12]. The starting point is taken to be an m-fold degenerate solution of the spin-independent Schrödinger equation for n electrons. The 2^n spin functions associated with this solution transform among themselves under permutations of the electrons; however, M_S (or $\tau/2$ in the notation of the authors) necessarily remains constant and thus can be taken to distinguish the representations. The decomposition of the representation labeled by M_S into irreducible parts is given in Sect. 4 of [12], and is equivalent to the sequence

$$[n,0] + [n-1,1] + [n-2,2] + \ldots + \left[\tfrac{1}{2}n + |M_S|, \tfrac{1}{2}n - |M_S|\right] \tag{5}$$

of Young tableaux, where $[n-z,z]$ specifies a tableau with $n-z$ cells in its first row and z in its second. Today we would probably calculate the dimension of such a tableau by asking for the number of "standard" tableaux that can be found for the given shape $[n-z,z]$; that is, the number of ways the numbers 1 to n can be arranged in the cells such that the numbers increase in going from left to right along the rows and from top to bottom down the columns. Young's general formula is given by Rutherford [16], with the aid of which we find

$$\mathrm{Dim}[n-z,z] = \binom{n}{z} - \binom{n}{z-1}, \tag{6}$$

in agreement with von Neumann and Wigner ([12], Sect. 5). When the functions associated with the tableau $[n-z,z]$ are multiplied by the m spin-independent functions, we are faced with the problem of decomposing the $m\,\mathrm{Dim}[n-z,z]$ products into their irreducible components. It is here that von Neumann and Wigner introduce the Pauli exclusion principle. The only components of interest are the totally antisymmetric ones. If the m-fold degeneracy derives solely from permutation symmetry (and not from rotational or other types of symmetry), then the analysis reveals that there is only one antisymmetric component, and all others can be set aside. From a modern perspective the tableaux in the sequence (5) correspond to the various values of S for which $S \geq M_S$; in fact, we have $z = n/2 - S$, and equations (4) and (6) coalesce. The single antisymmetric component found by von Neumann and Wigner corresponds to a given S, a given M_S, and a given label describing the spin-independent part of the problem. The symmetry of this last part is specified by the tableau adjoint to $[n-z,z]$, that is, to $[22\ldots211\ldots1]$ in which z twos and $n-z$ ones appear. This is the content of equation (17b) of von Neumann and Wigner [12].

The need to cope with this procedure analytically led von Neumann and Wigner, in the third of their articles [9], to introduce what are now usually called

Clebsch-Gordan coefficients (though they are sometimes referred to as Wigner coefficients or vector-coupling coefficients). They appear as s functions in the definition of the wavefunction Ψ ([9], equation (9)). The attendant labels, in modern notation, would be written as $s^J_{M_J M_S}(S, L)$. The coefficients for which $L = 1$ had already been worked out by Wigner in connection with his recursion formulas for the D functions [10], though those corresponding to $J = L - 1$ possess opposite signs to what Wigner's general formula, presented later in his book ([1], p. 206) as $s^{(S,L)}_{JM_S M_L}$, would give.

Von Neumann and Wigner ended their three articles with two applications: the derivations of the Landé g factor and the Burger-Dorgelo sum rule. Two years before this analysis, Heisenberg and Jordan [17] had shown that

$$g = 1 + \frac{J(J+1) - L(L+1) + S(S+1)}{2J(J+1)}, \tag{7}$$

thus slightly amending the original expression of Back and Landé [18], in which each product of the type $X(X+1)$ appears as $X^2 - 1/4$. To find the factor g that makes the crucial operator $\mathbf{L}+2\mathbf{S}$ equivalent to $g\mathbf{J}$, Heisenberg and Jordan wrote $\mathbf{L}+2\mathbf{S} = \mathbf{J}+\mathbf{S}$ and then projected $\mathbf{S}$ on to $\mathbf{J}$ by making the replacement $\mathbf{S} \to \mathbf{J}(\mathbf{J}^2)^{-1}\mathbf{J}\cdot\mathbf{S}$. By using

$$\mathbf{J}\cdot\mathbf{S} = \mathbf{S}^2 + \mathbf{L}\cdot\mathbf{S} = \mathbf{S}^2 + \left(\mathbf{J}^2 - \mathbf{L}^2 - \mathbf{S}^2\right)/2, \tag{8}$$

the eigenvalues of $(\mathbf{J}^2)^{-1}\mathbf{J}\cdot\mathbf{S}$ immediately yield the second term on the right-hand side of equation (7). Instead of replacing $\mathbf{S}$ by a more tractable operator, von Neumann and Wigner decided to pick its z component and directly evaluate its expectation value for the eigenstate Ψ. This is a more obvious and technically correct approach, but it led, in equation (24) of [9], to a sum over a product of two Clebsch-Gordan coefficients weighted by the factor M_S. Faced with such a sum today, we would probably convert M_S to a third Clebsch-Gordan coefficient and reduce the sum to the product of a 6-j symbol and a 3-j symbol by means of equation (6.2.8) of Edmonds [19] or its equivalent. Von Neumann and Wigner could not resort to such a procedure in 1928. Nevertheless, they were able to arrive at the required result by a strategy of considerable ingenuity, in which $g - 1$ appears as the difference between two infinite series.

The sum rule of Burger and Dorgelo [20] states that, for a multiplet characterized by a given initial term ${}^{2S+1}L$ and a given final term ${}^{2S+1}L'$, the sum of the intensities of all lines originating (or terminating) on a particular level J is proportional to its degeneracy, namely $2J + 1$. That is, for isotropic radiation, each component M_J of the level contributes equally. Von Neumann and Wigner prove this result for electric-dipole radiation by using two sum rules for products of Clebsch-Gordan coefficients. The calculation is reasonably straightforward. Nowadays we would probably express their transition matrix element as a product of a 3-j and a 6-j symbol, in which case the sums would appear in another guise.

4. The Wigner-Eckart Theorem

Although von Neumann and Wigner were able to put theoretical atomic spectroscopy on a more rigorous footing, the price was a level of abstraction that few experimentalists were willing to accept. When writing his book [1], Wigner aimed at greater accessibility. By giving a formal treatment of Kronecker products, he was able to provide a sound theoretical basis for the vector-addition model of the atom, and with it the concept of coupling that Hund [21] had used so extensively. A more suggestive notation was adopted: thus S is used instead of $\frac{1}{2}n - z$ and L for l. In addition to such technical improvements, Wigner also made a major advance in stating what has become known as the Wigner-Eckart theorem ([1], p. 264). This generalizes the treatment of vector operators to those whose transformation properties with respect to $SO(3)$ are specified by an arbitrary irreducible representation of that group. It extracts (by means of a Clebsch-Gordan coefficient) the geometric aspects of a matrix element from those features that depend on physical magnitudes. More remarkably, the theorem is not special to $SO(3)$, and can be extended to other groups, as Wigner, in noting some work of Eckart [22] (to judge from a footnote on page 264 of [1]), realized. Some 18 years later, applications of the theorem to $SO(7)$ and G_2 were made to the spectroscopy of f electrons by Racah [7], and a new chapter in theoretical atomic physics was opened.

The establishing of the Wigner-Eckart theorem enabled Wigner to simplify some of the derivations worked out earlier in his papers with von Neumann. The relations between Clebsch-Gordan coefficients given on pp. 297–299 of [1] presage Wigner's introduction of 6-j symbols, which was to come almost a decade later [23]. With their aid, the formula for the Landé g factor emerges without the previous intricacies ([1], p. 302). Furthermore, Wigner was able to subsume the earlier analysis of the Burger-Dorgelo sum rule in a treatment of the Hönl-Kronig intensity formulas [24, 25], which specify the relative intensities of the component lines of a given multiplet ([1], p. 299) and not merely the sums of the relative intensities of certain groups of lines.

5. Second Quantization

Early in 1928, Jordan and Wigner submitted for publication an article dealing with a new technique for imposing the Pauli Exclusion Principle on an ensemble of electrons [26]. It arrived at the offices of the Zeitschrift für Physik less than a month after the first of the three articles in which Wigner collaborated with von Neumann [13]. Jordan and Klein [27] had already extended to many-boson systems the work of Dirac [28] on field quantization, and the question arose as to whether the formalism could be repeated for fermions. As Jordan and Wigner showed in Sect. 6 of their article, the key difference lies in replacing the commutation relations between the $b^\dagger$ and b operators by anticommutation relations between their analogues, the operators $a^\dagger$ and a. During the early years of their use, when analyses involving explicit matrix representations were

common, they were often referred to as the Jordan-Wigner matrices (see, for example, Kramers [29]); today they are usually called creation and annihilation operators. Equation (66a) of Jordan and Wigner [26] gives the contribution to the Hamiltonian of an operator that acts on one electron at a time; equations (66b) and (66c) generalize this result to two-electron operators and then to n-electron operators. It is to be noted that there are no references to specific electrons in these formulas: indeed, the operators are independent of the number of electrons of the system under study.

The general method outlined above has come to be referred to as second quantization. Its advantages were appreciated only slowly. Heisenberg applied the method to atomic shell theory, realizing that $a^\dagger$ and a for electrons could be replaced by a' and $a'^\dagger$ for holes, thereby finding formulas for the Coulomb energies of electrons in nearly closed shells [30]. As a result of his using a misidentified ^1S term of TiI $3d^2$, his prediction of $27\,500 \pm 150$ cm^{-1} for the energy for ^1S of NiI $3d^8$ has turned out to differ markedly from the experimental figure of 51457 [31]. This inconsequential error did not affect the work of Johnson [32], who extended Heisenberg's calculations to find expressions for the electrostatic energies of d^9p and d^9d' and to confirm those of p^5p', which were already known. Apart from this initial flurry of activity, the method of second quantization, as applied to shell theory, lay dormant for some years. Wigner made no use of it in the writing of his book [1]; nor did Condon and Shortley [5] or Racah [6, 7] in their work. In 1956, Brink and Satchler [33] revived the hole-particle connection in the context of nuclear shell theory, and recognized that a creation operator could be regarded as a spherical tensor of the type studied by Racah [6]. This allowed them to use the Wigner-Eckart theorem to identify a fractional parentage coefficient as the reduced matrix element of a creation operator. However, this result as well as the method of second quantization as whole was disregarded by de-Shalit and Talmi in their 1963 book on nuclear shell theory [34].

Two developments have brought the method of second quantization to the fore in recent years. First, the creation and annihilation operators can be coupled to produce operators that change the number of particles, thereby enormously increasing the possibilities for forming the generators of Lie groups. Operators that conserve particle number can now be classified by the irreducible representations of these groups, and the Wigner-Eckart theorem can be used to relate matrix elements in different configurations [35, 36]. Secondly, the creation and annihilation operators are ideally suited for representing within a configuration the effects of perturbations from other configurations. A good example of this in atomic physics is the third-order and fourth-order analyses of the two-photon absorption of gadolinium ions in various environments [37]. Such applications of the method of second quantization are a striking justification for its use. As late as 1968, Slater [38], writing in the third person, stated that he "read the papers on second quantization when they first came out, and decided at that time that they seemed to him like a more complicated method of treating matters which could be more easily taken up without their use". It is difficult to see how that attitude could be plausibly maintained today.

6. Correlation Rules for Diatomic Molecules

A month after von Neumann and Wigner submitted their third article [9] to the Zeitschrift für Physik, Wigner and Witmer sent to the same journal an article that has had widespread use in molecular spectroscopy [39]. By 1928, the year of the submission of these articles, much was already known about the kinematic structure of simple molecules; indeed, the same issue of the Zeitschrift für Physik in which articles [9] and [39] were published also contains the last of four articles by Hund on molecular spectra [40]. By bringing group theory into play, Wigner and Witmer were able to cover all possible descriptions of the molecular states that derive from symmetry considerations. The first nine sections of their article describe how group theory can be used to deduce the selection rules and some relative intensities for the lines making up the rotational bands of diatomic molecules. They began by noting that, for a molecule built from two different nuclei, the internal symmetry of the molecule comprises the two-dimensional rotations about the inter-nuclear axis together with a reflection in the plane containing this axis. This last symmetry operation is sometimes called the Kronig reflection after its originator [41]. If we follow current usage and take Λ to denote the eigenvalue $\langle L_\zeta \rangle$ of the axial component of the electrons' orbital angular momentum $\mathbf{L}$ (rather than the symbol λ used by Wigner and Witmer), the Kronig reflection converts a state labeled by Λ to one labeled by $-\Lambda$. When $\Lambda \neq 0$, a phase choice has to be made in relating a wavefunction to its reflected counterpart. Wigner and Witmer, in their equation (1a) of [39], took it to be $+1$. The spherical harmonic $Y_{k,q}$, on the other hand, goes over into $(-1)^q Y_{k,-q}$ when its usual definition is adopted (Edmonds [19]). If the many-electron wavefunctions of the molecule are built from superpositions of atomic wavefunctions, it would be more natural to use $(-1)^{\Sigma q}$, that is, $(-1)^\Lambda$ instead of $+1$. If this choice is made, the factors $(-1)^\lambda$ appearing in equation (7) of Wigner and Witmer [39] disappear. When $\Lambda = 0$, corresponding to Σ states, no phase ambiguity occurs. Current usage requires that we write Σ^+ or Σ^- to indicate the behavior of a particular state under the Kronig reflection rather than the labels Σ and Σ' used by Wigner and Witmer. We should also note that the rotational angular-momentum quantum number l appearing in equations (7) of [39] would be written today as N (if the recommendations of the 1963 Joint Commission for Spectroscopy [42] are followed, as is often the case).

The phase difference stemming from the factor $(-1)^\Lambda$ is innocuous enough. For example, the crucial alternation of the parities of the molecular states as N runs over successive values holds whether or not this factor is included. However, a rather more serious difference was produced when Wigner and Witmer turned to homonuclear diatomics. Instead of taking an inversion in the midpoint of the internuclear axis as a symmetry operation, they used a reflection in the plane passing through this midpoint and perpendicular to the axis. Today, the intrinsic parity of the states of homonuclear diatomics is universally represented by the symbols g and u, corresponding to the former of the two symmetry operations. Wigner and Witmer, on the other hand, used the sub-

scripts + and − to distinguish the properties of states under their reflection. Since a reflection can be formed from an inversion and a rotation by π, the difference between g and +, and also between u and −, is given by the expectation value of $\exp(i\pi L_\zeta)$, that is, by $(-1)^\Lambda$. With these changes, the selection rules for electric-dipole radiation, as set out in the second table of Wigner and Witmer on p. 867 of [39], can be subsumed into $g \leftrightarrow u$, $\Delta\Lambda = 0, \pm 1$ (with $0 \leftrightarrow 0$ allowed only between two Σ^+ states or between two Σ^- states). As for the rotational quantum numbers, the allowed transitions satisfy $\Delta N = +1$ or -1 (for P and R bands) or 0 (for Q bands), except for $\Sigma_g \leftrightarrow \Sigma_u$, in which case the Q band vanishes.

In the sections that follow a statement of these selection rules, Wigner and Witmer [39] derived the intensity alternations that characterize successive lines of homonuclear diatomics, and they extended their analysis to include the effects of electron spin. Depending on the strength of the spin-orbit interaction, the analysis breaks up into two parts: for light atoms, for which the spin is almost free (corresponding to Hund's case (b) coupling); and for heavy atoms, for which it often happens that the spin is strongly coupled to the inter-nuclear axis (corresponding to Hund's case (a)). As in the papers written in collaboration with von Neumann, the spin S appears somewhat incongruously as $\frac{1}{2}n - z$.

Wigner and Witmer next turned their attention to the Aufbau principle: that is, a description of how molecular states can be built up from the states of the constituent atoms. It is this part of their article that is remembered and widely used today. Their method consists in imagining two separated atoms, whose states are known, and bringing them together to form the molecule of interest. In the limit of close approach, an atom is formed from the two coalesced nuclei, and the molecular states go over into the states of the united atom. Wigner and Witmer were not the only ones to consider this model: in fact, Hund, in the same issue of the Zeitschrift für Physik, gave correlation schemes that represent the coalescing of C + H to give N, C + N (or N^+ + N) to give Al, C + O (and N + N) to give Si, N + O to give P, as well as a sequence of diagrams relating Li + H, LiH, Be, He_2 and He + He [40]. The only example given by Wigner and Witmer ([39], pp. 884 and 885) treats the coalescing of two hydrogen atoms to form He. What makes their analysis more profound is their use of labels derived from reflection operations. In addition to showing that a state L of the united atom provides $L + 1$ states for which $\Lambda = 0, \pm 1, \pm 2, \ldots, \pm L$, Wigner and Witmer demonstrated that the single Σ state in that sequence is Σ^+ or Σ^- according as $L + p$ is an even or an odd number, where p is the parity of the atomic state (and which is given by the sum of the azimuthal quantum numbers of the constituent electrons). Hund, who was aware of the analysis of Wigner and Witmer, wondered whether the Kronig reflection would prove useful, believing that most Σ states, like those occuring in the H + H problem studied by Wigner and Witmer, would be of the type Σ^+ ([40] p. 780). However, Σ^- states frequently appear in many-electron configurations, the classic example being the ground state $^3\Sigma_g^-$ of O_2.

The states of H_2 deriving from five low-lying levels of He are set out on p. 885 of [39]. In modern notation we would write Σ_g^+ for Σ_+, Σ_u^+ for Σ_-, and Π_u for Π_+. We note that states of a homonuclear diatomic springing from the configuration of the united atom characterized by a parity p (g or u) also carry that same parity label. In addition to exemplifying how molecular states can be derived from the united atom, Wigner and Witmer gave a table of the possible molecular states that certain specified states of the two separated H atoms can provide ([39], p. 884). The general procedure for determining molecular states from the states of separated atoms is sometimes referred to as the second part of the Wigner-Witmer rules; it is slightly more intricate to codify than the previous part. A number of tables and several examples have been provided by Bingel to cover most cases of physical interest [43].

7. Conservation Laws

The act of bringing two atoms together so that their nuclei coincide was also used by Wigner to exemplify a conservation law [44], a topic that he was considering early in 1928. Having pointed out that the commutation of displacement and rotation operators with the Hamiltonian of a system leads to the conservation of the position of the centre of mass and the conservation of the total angular momentum of the system, Wigner turned to the inversion operation. He stated the corresponding conservation law in the folliwing way: The probability for a state to be labelled g or u (Sp_+ or Sp_- in Wigner's notation) does not change with time provided the Hamiltonian commutes with the parity operation. To illustrate its significance, Wigner considered bringing an H nucleus (a proton) towards an H atom, ultimately forming He^+. The pattern of the energy levels as the motion takes place had already been considered by Hund [45]. The initial state of the system, with respect to its center of mass, has a probability of 1/2 of being of type g and 1/2 of being of type u; this distribution of probabilities must persist as the two nuclei are brought together, with the result that the final state of He^+ must have an equal probability of being of type S (the lowest g state) or type P (the lowest u state). As Wigner pointed out, this example has no classical analogue.

The question of the conservation of spin was considered a few years later by Döpel, Gailer, and Wigner [46]. The motivation for this article was some experimental work in which canal rays of neutral helium atoms were found capable of exciting part of the triplet spectrum of neutral mercury [47]. Among the more prominent lines were those corresponding to transitions between the levels $6s6p\,^3P_J$ and $6s7s\,^3S_1$. The extreme feebleness of any triplet spectrum of helium was interpreted to mean that the helium atoms remained in the singlet ground state; so the question arose as to how triplet states of mercury could be excited from the singlet ground state without violating the conservation of the total spin of the combined system He + Hg. Wigner and his colleagues realized that there is substantial spin-orbit coupling in an atom as heavy as mercury, and the triplet states contain significant admixtures of singlet states. This can be

seen immediately by noting that the energy ratio $[E(^3P_2) - E(^3P_1)]/[E(^3P_1) - E(^3P_0)]$, from Döpel and Gailer's own measurements [47], is 2.63. This number shows an appreciable deviation from 2.00, which the Landé interval rule predicts in the limit of small spin-orbit coupling. Thus ^3S and ^3P are only nominally triplets, and there is nothing to prevent their excitation (either directly or through intermediate states) by He-Hg collisions. Gailer subsequently turned his attention to lighter atoms, where the interval rule works better; but, here again, triplet spectra were observed [48]. As a result of communications from Wigner, it was finally realized by Gailer that only a fraction of the helium atoms in the canal rays were in their singlet ground state [49]. Many existed in the metastable state $1s2s^3S$, which, although undetectable spectroscopically, could excite the target atoms from singlet states to higher triplets, thereby producing the observed triplet spectra. Thus the experimental arrangement for testing the conservation of total spin turned out to ineffective.

8. Extension to the Thomas-Reiche-Kuhn Sum Rule

The sum rule usually associated today with the names of Thomas, Reiche, and Kuhn [50, 51] was of considerable importance in the early days of quantum mechanics because it is independent of Planck's constant and has a direct correspondence with its classical counterpart. It fell to Wigner to see that it could be extended without any loss of generality [52]. That this is possible stems from the fact that the sum rule, as usually stated, refers to isotropic radiation, whereas equally valid sums can be constructed for the x, y, and z directions separately. These are set out in equations (2) of [52]. Wigner used them to investigate transitions between states labelled by the total orbital angular momentum L of an atomic system comprising n electrons. The selection rule $\Delta L = +1, 0$ or -1 leads to the introduction of the three oscillator strengths f_{+1}, f_0 or f_{-1} for the transitions originating in a term Q (with a given L) and terminating in various terms P (labelled by quantum numbers $L+1$, L, and $L-1$). By the ingenious device of selecting the maximum component M_L for the term Q, Wigner was able to augment the condition $f_{+1} + f_0 + f_{-1} = n$ with another, which is given in his equation (9). With the angular-momentum techniques available today, we would probably take an arbitrary M_L, in which case Wigner's equation (9) would appear as

$$\sum_q \begin{pmatrix} L & 1 & L+q \\ -M_L & 0 & M_L \end{pmatrix}^2 f_q = \frac{n}{3(2L+1)} . \tag{9}$$

It is straightforward to check that the above equation is a superposition of equations (5) and (9) of Wigner [52], the coefficients of superposition being functions of M_L.

Since we have three unknowns and only two equations, general expressions for the f_q cannot be determined. The information that Wigner was able to extract from these expressions is in the form of various inequalities that the f_q

must satisfy. However, for the special case of $n = 1$, parity considerations force f_0 to be zero. The concomitant solutions for f_{+1} and f_{-1} are given in equations (10) of [52]. It is for these formulas that Wigner's article is chiefly referred to today. Although strictly relevant only for the one-electron case, they are often applied to atoms for which the condition $n = 1$ is interpreted as meaning one electron outside closed shells. As Fock pointed out, exchange effects can lead to corrections to the oscillator strengths of the order of a few percent [53]; but this adjustment is sufficiently small not to cause serious difficulties. Perhaps the most striking feature is that f_{-1} is negative, implying the existence of at least one state lying at lower energies than the term Q. For many-electron atoms, such states are often occupied by the electrons in the closed shells, so the transitions cannot be directly observed; but the hypothetical intensities of these transitions can be very much larger than that of the first allowed transition to the lowest excited state of the corresponding type.

9. Normal Modes

One of the most remarkable of Wigner's early papers concerns his application of group theory to the normal modes of vibration of molecules [54]. Wigner showed, first, that there are as many types of normal modes as there are irreducible representations of the group G whose operations send the equilibrium configuration of the molecule into itself. Secondly, he devised a method for finding how many times a given irreducible representation occurs in a complete listing of the normal modes. This is done by using characters, the crucial equation being (9) of [54]. An atom of the molecule that is displaced from its position after an operation of G contributes nothing to the character; one that remains in the same position after a (proper) rotation of the molecule through an angle φ contributes $(1 + 2\cos\varphi)$, an expression that is familiar to us today and which appears here for the first time. The formula for the number of occurrences of a given non-degenerate mode ([54], equation (11a)) needs some minor corrections, as was noted by Knox and Gold in their English translation [55]. As an example of selection rules, the infra-red active modes are determined. For us this is an example of the Wigner-Eckart theorem applied to the finite group G. The paper ends with a listing of the basic character tables from which those for all 32 point groups can be constructed. Much of this article has become second nature to us today, and it is easy to overlook its fundamental nature. Its significance was not lost on the writers of the time, as is clear from the review of Rosenthal and Murphy [56].

Wigner used the methane molecule (CH_4) as an example of his techniques [54]. A normal-mode analysis had already been carried out by Dennison, who showed that there are just four normal frequencies [57]. The group G consists of the 24 elements that rotate and reflect the molecule into itself; it is denoted by T_d. The characters corresponding to rotations (proper and improper) belonging to the five classes must be corrected by withdrawing the contributions representing rotations and displacements of the molecule as a whole. When

the residual characters are expressed as linear combinations of those for the irreducible representations, it is found that just four of them are required, one of the three-dimensional ones occurring twice. Diagrams of the corresponding displacements were provided later by Jahn [58] when he wrote up his Leipzig thesis for publication.

10. Time-Reversal

In addition to his direct interest in methane, Wigner was influential in the work on polyatomic molecules of Tisza, who extended the selection rules to include those for the Raman effect and went beyond the 32 point groups to study the icosahedral group as well as a variety of cyclic groups [59]. Tisza also considered irreducible representations with complex characters. These occur in pairs in the normal-mode analysis, since the characters on which the analysis depends are always real. Tisza was aware of Wigner's work on time reversal [60], which had yet to be published; and he recognized that this concept could be used to relate one member of such a pair to the other. In general, an eigenfunction ψ of a real Hamiltonian H that is labeled by an irreducible representation whose characters are complex numbers possesses a companion ψ^* corresponding to the same energy E. In the simple Schrödinger picture, the connection between time reversal and complex conjugation can be seen by inspecting the factor $\exp(-iEt/\hbar)$ that describes the time dependence of a wavefunction for a stationary state. This is the starting point for Wigner's article [60]. The development of the theory was revised and cast in a more expansive mathematical form in the 1959 edition of his book on group theory [2]. The time-reversal operator K of the earlier article is shown to be antiunitary and the concept of corepresentations is developed. In reading that earlier work, it is clear that much of the motivation stemmed from some work of Kramers on the Faraday effect for crystals whose active atoms possess an odd number of electrons [61]. Kramers showed that, in the absence of a magnetic field, all the eigenstates of such atoms must be at least doubly degenerate. For Wigner, such degeneracy could only come from some symmetry property of the Hamiltonian, and this he found in its invariance under time reversal. In fact, Kramers came very close to making this point himself, since he compared the two Hamiltonians that differ only with respect to the directions of the spins of the electrons, and these, of course, are related by time reversal.

In his later discussion of time-reversal symmetry, Wigner remarked that it has no far-reaching consequences in the theory of atomic spectra ([2], p. 333). An interesting analogue can nevertheless be found. We can define an angular-momentum vector $\mathbf{Q}$, the quasispin, whose components, all of which are scalars with respect to $\mathbf{S}$ and $\mathbf{L}$, allow us to change the electron number n in steps of 2. A conjugation operator C analogous to that of Bell [62] can be introduced that parallels the time-reversal operator; it is antiunitary and, by reversing the sign of Q_0, converts electron states into hole states. Matrix elements of an operator invariant with respect to C must thus be the same (apart from

a possible phase) in the configuration l^{4l+2-n} as in l^n. We can also use C to relate the coefficients of fractional parentage for particles to those for holes, thereby obtaining a formula of Racah [63] in a direct and transparent way [36].

11. Rotating Molecules

A central problem of molecular spectroscopy is how best to separate the rotations of the molecule as a whole from the internal motions of the electrons and nuclei that comprise it. The solution proposed by Hirschfelder and Wigner [64] is based on a variational principle. The center of mass is taken as the origin of the internal coordinates, and the orientation of the molecule is specified by the directions of the instantaneous principal axes of inertia. At about the same time, Eckart, working directly with the Lagrangian form of the kinetic energy and using a similar coordinate scheme, obtained a slightly different form for the Hamiltonian [65]. Both solutions possess anomalous coefficients for N_1^2, N_2^2, and N_3^2, the squares of the total angular momenta about the three principal axes. In terms of the moments of inertia I_i, the coefficient of N_i^2 is more nearly equal to $I_i/2(I_{i+1} - I_{i+2})^2$ than to the conventional value $1/2I_i$, as Van Vleck pointed out [66]. In that same article Van Vleck also performed a detailed analysis showing that the interaction between rotation and vibration produced terms in the Hamiltonian that were larger than those coming from rotation alone, and that, if perturbation calculations were carried through with either the Eckart or the Hirschfelder-Wigner Hamiltonians, the anomalous coefficients reverted to the conventional forms.

The difficulty arises because, surprising as it may seem, the directions of the instantaneous principal axes of inertia are not the best choice in determining the orientation of a molecule. We have some discretion in this matter when the Euler angles are being introduced, since we are faced with the problem of defining what we mean by the rotation of a non-rigid body. Van Vleck was aware of the difficulties in referring internal motions to the principal axes. He suggested a rotating frame defined by the positions of individual particles ([66], p. 494), evidently inspired by Wigner's approach in atomic spectra [8]. Little came of this, though something similar was explored by Hirschfelder and his colleagues some years later [67]. The solution that is almost always followed today was devised by Eckart [68], who showed that the coupling between vibration and rotation could be reduced by a special choice of coordinate axes. When referred to these axes, the so-called vibrational angular momentum of the nuclei is minimized; in fact, we can use that condition, together with the approximation that displacements of the nuclei from their equilibrium positions are small, to derive what have become known as the Eckart equations. A description of the procedure for determining the Eckart axes, as well as a detailed example for the water molecule, have been provided by Bunker [69]. From a modern perspective, the work of Hirschfelder and Wigner [64] represents an early first step in understanding the kinetics of molecules; but a change of direction was necessary to get us where we are today.

References

[1] E. Wigner: Gruppentheorie und ihre Anwendungen auf die Quantenmechanik der Atomspektren. Vieweg, Braunschweig 1931

[2] E. P. Wigner: Group Theory. Academic Press, New York 1959

[3] H. Weyl: Gruppentheorie und Quantenmechanik. Hirzel, Leipzig 1928

[4] J. C. Slater: The Theory of Complex Spectra. Phys. Rev. **34**, 1293–1322 (1929)

[5] E. U. Condon and G. H. Shortley: The Theory of Atomic Spectra. Cambridge University Press, Cambridge 1935

[6] G. Racah: Theory of Complex Spectra I–III. Phys. Rev. **61**, 186–197 (1942); ibid. **62**, 438–462 (1942); ibid. **63**, 367–382 (1943)

[7] G. Racah: Theory of Complex Spectra IV. Phys. Rev. **76**, 1352–1365 (1949)

[8] E. Wigner: Einige Folgerungen aus der Schrödingerschen Theorie für die Termstrukturen. Z. Phys. **43**, 624–652 (1927); volume I, part II

[9] J. v. Neumann and E. Wigner: Zur Erklärung einiger Eigenschaften der Spektren aus der Quantenmechanik des Drehelektrons, III. Z. Phys. **51**, 844–858 (1928); volume I, part II

[10] E. Wigner: Berichtigung zu der Arbeit: Einige Folgerungen aus der Schrödingerschen Theorie für die Termstrukturen. Z. Phys. **45**, 601–602 (1927); volume I, part II

[11] C. L. B. Shudeman: Equivalent Electrons and their Spectroscopic Terms. J. Franklin Inst. Philadelphia **224**, 501–518 (1937)

[12] J. v. Neumann and E. Wigner: Zur Erklärung einiger Eigenschaften der Spektren aus der Quantenmechanik des Drehelektrons, II. Z. Phys. **49**, 73–94 (1928); volume I, part II

[13] J. v. Neumann and E. Wigner: Zur Erklärung einiger Eigenschaften der Spektren aus der Quantenmechanik des Drehelektrons, I. Z. Phys. **47**, 203–220 (1928); volume I, part II

[14] E. Wigner: Über nicht kombinierende Terme in der neueren Quantentheorie, I. Z. Phys. **40**, 492–500 (1926); volume I, part II

[15] E. Wigner: Über nicht kombinierende Terme in der neueren Quantentheorie, II. Z. Phys. **40**, 883–892 (1927); volume I, part II

[16] D. E. Rutherford: Substitutional Analysis. Edinburgh University Press, Edinburgh 1948, equation (16.1)

[17] W. Heisenberg and P. Jordan: Anwendung der Quantenmechanik auf das Problem der anomalen Zeemaneffekte. Z. Phys. **37**, 263–277 (1926)

[18] E. Back and A. Landé: Zeemaneffekt und Multipletstruktur der Spektrallinien. Springer, Berlin 1925, p. 79

[19] A. R. Edmonds: Angular Momentum in Quantum Mechanics. Princeton University Press, Princeton 1957

[20] H. C. Burger and H. B. Dorgelo: Beziehung zwischen inneren Quantenzahlen und Intensitäten von Mehrfachlinien. Z. Phys. **23**, 258–268 (1924)

[21] F. Hund: Linienspektren und periodisches System der Elemente. Springer, Berlin 1927

[22] C. Eckart: The Application of Group Theory to the Quantum Dynamics of Monatomic Systems. Rev. Mod. Phys. **2**, 305–380 (1930)

[23] E. P. Wigner: On the Matrices which Reduce the Kronecker Products of Representations of S. R. [Simply Reducible] Groups, unpublished manuscript of 1940, printed in Quantum Theory of Angular Momentum, eds. L. C. Biedenharn and H. van Dam. Academic Press, New York 1965; volume I, part III

[24] H. Hönl: Die Intensitäten der Zeemankomponenten. Z. Phys. **31**, 340–354 (1925)

[25] R. de L. Kronig: Über die Intensität der Mehrfachlinien und ihrer Zeemankomponenten. Z. Phys. **31**, 885–897 (1925)

[26] P. Jordan and E. Wigner: Über das Paulische Äquivalenzverbot. Z. Phys. **47**, 631–651 (1928); volume I, part II
[27] P. Jordan and O. Klein: Zum Mehrkörperproblem der Quantentheorie. Z. Phys. **45**, 751–765 (1927)
[28] P. A. M. Dirac: On the Theory of Quantum Mechanics. Proc. Roy. Soc. London A **112**, 661–677 (1927)
[29] H. A. Kramers: Quantum Mechanics. North-Holland, Amsterdam 1958, p. 328
[30] W. Heisenberg: Zum Paulischen Ausschliessungsprinzip. Ann. Phys. **10**, 888–904 (1931)
[31] J. Sugar and C. Corliss: Atomic Energy Levels of the Iron-Period Elements: Potassium through Nickel. J. Phys. Chem. Reference Data **14**, Supp. no. 2 (1985)
[32] M. H. Johnson: The Vector Model and the Pauli Principle. Phys. Rev. **43**, 627–631 (1933); Note on Almost Closed Shells. Phys. Rev. **43**, 632–635 (1933)
[33] D. M. Brink and G. R. Satchler: Holes and Particles in Shell Models. Nuovo Cimento **4**, 549–557 (1956)
[34] A. de-Shalit and I. Talmi: Nuclear Shell Theory. Academic Press, New York 1963
[35] R. D. Lawson: Theory of the Nuclear Shell Model. Clarendon, Oxford 1980
[36] B. R. Judd: Second Quantization and Atomic Spectroscopy. Johns Hopkins, Baltimore 1967
[37] M. C. Downer and A. Bivas: Third- and Fourth-Order Analysis of the Intensities and Polarization Dependence of Two-Photon Absorption Lines of Gd^{3+} in LaF_3 and Aqueous Solutions. Phys. Rev. B **28**, 3677–3696 (1983)
[38] J. C. Slater: Book Review (of Ref. [36]). Am. J. Phys. **36**, 69 (1968)
[39] E. Wigner and E. E. Witmer: Über die Struktur der zweiatomigen Molekelspektren nach der Quantenmechanik. Z. Phys. **51**, 859–886 (1928); volume I, part II
[40] F. Hund: Zur Deutung der Molekelspektren, IV. Z. Phys. **51**, 759–795 (1928)
[41] R. de L. Kronig: Zur Deutung der Bandspektren. Z. Phys. **46**, 814–825 (1928)
[42] F. A. Jenkins: Report of Subcommittee *f* (Notation for the Spectra of Diatomic Molecules) of the Joint Commission for Spectroscopy. J. Opt. Soc. Am. **43**, 425–426 (1953)
[43] W. A. Bingel: Theorie der Molekülspektren. Verlag Chemie, Weinheim 1967
[44] E. Wigner: Über die Erhaltungssätze in der Quantenmechanik. Nachrichten d. Gesell. d. Wissenschaften z. Göttingen, Math.-Phys. Klasse 375–381 (1927); volume I, part II
[45] F. Hund: Zur Deutung der Molekelspektren, I. Z. Phys. **40**, 742–764 (1927), figure 11
[46] R. Döpel, K. Gailer, and E. Wigner: Über die experimentelle Prüfung des Spinerhaltungssatzes. Phys. Z. **35**, 336–337 (1934); volume I, part II
[47] R. Döpel and K. Gailer: Experimentelle Prüfung des Spin-Erhaltungssatzes beim Atomstoss. Physik. Z. **34**, 827–831 (1933)
[48] K. Gailer: Über die wechselseitige Lichtanregung bei den Atomstössen He bzw. H auf Be, Mg, Ca, Sr, Ba. Z. Phys. **108**, 580–591 (1938)
[49] K. Gailer: Experimenteller Beitrag zur Frage der Gültigkeit des Spinerhaltungssatzes. Physik. Z. **39**, 407–409 (1938)
[50] W. Kuhn: Über die Gesamtstärke der von einem Zustande ausgehenden Absorptionslinien. Z. Phys. **33**, 408–412 (1925)
[51] F. Reiche and W. Thomas: Über die Zahl der Dispersionselektronen, die einem stationären Zustand zugeordnet sind. Z. Phys. **34**, 510–525 (1925)
[52] E. Wigner: Über eine Verschärfung des Summensatzes. Phys. Z. **32**, 450–453 (1931); volume I, part II
[53] V. Fock: Über die Anwendbarkeit des quantenmechanischen Summensatzes. Z. Phys. **89**, 744–749 (1934)

[54] E. Wigner: Über die Eigenschwingungen symmetrischer Systeme. Nachrichten d. Gesell. d. Wissenschaften z. Göttingen, Math.-Phys. Klasse 133–146 (1930); volume I, part II

[55] R. S. Knox and A. Gold: Symmetry in the Solid State. Benjamin, New York 1964, p. 177

[56] J. E. Rosenthal and G. M. Murphy: Group Theory and the Vibrations of Polyatomic Molecules. Rev. Mod. Phys. **8**, 317–346 (1936)

[57] D. M. Dennison: Molecular Structure of Methane. Astrophys. J. **62**, 84–103 (1925)

[58] H. A. Jahn: Rotation und Schwingung des Methanmoleküls. Ann. Phys. **23**, 529–556 (1935)

[59] L. Tisza: Zur Deutung der Spektren mehratomiger Moleküle. Z. Phys. **82**, 48–72 (1933)

[60] E. Wigner: Über die Operation der Zeitumkehr in der Quantenmechanik. Nachrichten d. Gesell. d. Wissenschaften z. Göttingen, Math.-Phys. Klasse 546–559 (1932); volume I, part II

[61] H. A. Kramers: Théorie Générale de la Rotation Paramagnétique dans les Cristaux. Proc. K. Ned. Akad. Wet. **33**, 959–972 (1930)

[62] J. S. Bell: Particle-Hole Conjugation in the Shell Model. Nucl. Phys. **12**, 117–124 (1959)

[63] G. Racah: Theory of Complex Spectra. Phys. Rev. **63**, 371 (1943), equation (19)

[64] J. O. Hirschfelder and E. Wigner: Separation of Rotational Coödinates from the Schrödinger Equation for N particles. Proc. Nat. Acad. Sci. **21**, 113–119 (1935); volume I, part II

[65] C. Eckart: The Kinetic Energy of Polyatomic Molecules. Phys. Rev. **46**, 383–387 (1934)

[66] J. H. Van Vleck: The Rotational Energy of Polyatomic Molecules. Phys. Rev. **47**, 487–494 (1935)

[67] C. F. Curtiss, J. O. Hirschfelder, and F. T. Adler: The Separation of the Rotational Coordinates from the N-Particle Schroedinger Equation. J. Chem. Phys. **18**, 1638–1641 (1950)

[68] C. Eckart: Some Studies Concerning Rotating Axes and Polyatomic Molecules. Phys. Rev. **47**, 552–558 (1935)

[69] P. R. Bunker: Molecular Symmetry and Spectroscopy. Academic Press, New York 1979, pp. 137–153

Über nicht kombinierende Terme in der neueren Quantentheorie. Erster Teil

E. P. Wigner

Zeitschrift für Physik *40*, 492–500 (1926)

(Eingegangen am 12. November 1926.)

Es wird in einem System mit drei Elektronen der Zerfall der Terme in mehrere, untereinander nichtkombinierende Gruppen bewiesen. Die Eigenrotation des Elektrons ist nicht berücksichtigt.

1. In einer Arbeit über Mehrkörperproblem und Resonanz in der Quantenmechanik [1]) gibt Heisenberg an, daß wenn n gleichwertige Teile in einem System in Wechselwirkung stehen (z. B. die beiden Elektronen eines He-Atoms), eine Einteilung der Terme in $n!$ Gruppen in der Weise möglich ist, daß ein Atom im Zustand der einen Gruppe auf keine Weise in ein Atom im Zustand der anderen Gruppe übergehen kann. Er gibt weiter an, daß man aus diesen $n!$ Gruppen eine herauswählen kann, in der sowohl Paulis Verbot der äquivalenten Bahnen, wie auch die Einstein-Bosesche Statistik gilt.

Im Gegensatz zu letzterem kommt Dirac [2]) auf Grund von etwas anderen Überlegungen zu dem Resultat, daß man zwei Termsysteme unterscheiden kann, von denen das eine dem Paulischen Verbot, das andere der Bose-Einsteinschen Statistik genügt.

In der vorliegenden Arbeit soll gezeigt werden, daß die Terme tatsächlich [3]) in mehrere Gruppen zerfallen. Diese Gruppen können miteinander nicht interkombinieren und es befinden sich unter ihnen zwei Gruppen, die folgende Eigenschaften haben: Beide enthalten etwa den $1/n!$ten Teil der Terme des ganzen Systems; das eine System genügt dem Paulischen Verbot, die Gewichtsverteilung des anderen ist durch die Bose-Einsteinsche Statistik gegeben. Die übrigen Systeme sind alle entartet, die Entartung ist aber so, daß sie durch keine Störung, die symmetrisch in den einzelnen gleichwertigen Teilchen ist, aufgehoben werden kann.

---

[1]) ZS. f. Phys. **38**, 411, 1926.

[2]) Proc. Roy. Soc. **112**, 661, 1926.

[3]) Heisenberg erwähnt, daß ihm ein Beweis seiner Behauptung bisher nicht gelungen ist.

Wir betrachten also ein System mit n gleichwertigen Teilchen. Im folgenden schreibe ich immer $\mathfrak{r}_i$ an Stelle des Koordinatentripels X_i, Y_i, Z_i des $i$ten Teilchens. Es erschien mir zweckmäßig, zu der Rechnung die Schrödingersche Methode zu benutzen.

Die Ladungsdichte an der Stelle $\mathfrak{r}$ wird nach Schrödinger durch folgenden Ausdruck gegeben:

$$\varrho(\mathfrak{r}) = \sum_j \int \Psi(\mathfrak{r}_1, \mathfrak{r}_2 \ldots \mathfrak{r}_{j-1}, \mathfrak{r}, \mathfrak{r}_{j+1} \ldots \mathfrak{r}_n)$$
$$\Psi^*(\mathfrak{r}_1 \ldots \mathfrak{r}_{j-1}, \mathfrak{r}, \mathfrak{r}_{j+1} \ldots \mathfrak{r}_n)\, d\mathfrak{r}_1 \ldots d\mathfrak{r}_{j-1}\, d\mathfrak{r}_{j+1} \ldots d\mathfrak{r}_n. \tag{1}$$

Dabei ist Ψ die Schrödingersche Wellenfunktion

$$\Psi = \sum_k c_k \Psi_k(\mathfrak{r}_1 \ldots \mathfrak{r}_n)\, e^{\frac{2\pi i E_k t}{h} + \varphi_k}, \tag{2}$$

worin E_k die Energie des mit k bezeichneten Zustandes, Ψ_k die zugehörige Eigenfunktion und c_k die Anregungsgröße dieses Zustandes ist.

Bekanntlich ist die Übergangswahrscheinlichkeit nach Heisenberg-Schrödinger in erster Näherung proportional dem Quadrat der Summe der Matrizengrößen oder, was auf dasselbe herausläuft, der Polarisation des Atoms. Will man auch die Quadrupol- und höheren Momente berücksichtigen, so muß man nach Heisenberg symmetrische Funktionen der p und q bilden, die die entsprechenden Momente der klassischen Theorie repräsentieren. Es scheint mir aber zweckmäßiger, als hinreichende Bedingung dafür, daß die Übergangswahrscheinlichkeit zwischen zwei Zuständen Null ist, die Tatsache anzusehen, daß die Ladungsdichte in jedem Punkte des Raumes zeitlich konstant ist, wenn nur diese zwei Zustände angeregt sind. In diesem Falle ist also

$$\varrho(\mathfrak{r}, t) = \sum_j \int \Big[c_1 \Psi_1(\mathfrak{r}_1 \ldots \mathfrak{r}_{j-1}\, \mathfrak{r}, \mathfrak{r}_{j+1} \ldots \mathfrak{r}_n)\, e^{\frac{2\pi i E_1 t}{h}}$$
$$+ c_2 \Psi_2(\ldots \mathfrak{r}_{j-1}\, \mathfrak{r}, \mathfrak{r}_{j+1} \ldots)\, e^{\frac{2\pi i E_2 t}{h}}\Big] \cdot \Big[c_1 \Psi_1(\ldots \mathfrak{r}_{j-1}\, \mathfrak{r}\, \mathfrak{r}_{j+1} \ldots)\, e^{-\frac{2\pi i E_1 t}{h}}$$
$$+ c_2 \Psi_2(\ldots \mathfrak{r}_{j-1}\, \mathfrak{r}\, \mathfrak{r}_{j+1} \ldots)\, e^{-\frac{2\pi i E_2 t}{h}}\Big]\, d\mathfrak{r}_1 \ldots d\mathfrak{r}_{j-1}\, d\mathfrak{r}_{j+1} \ldots d\mathfrak{r}_n \tag{3}$$

unabhängig von t. Oder, was dasselbe ist,

$$\sum_j \int \Psi_1(\mathfrak{r}_1\, \mathfrak{r}_2 \ldots \mathfrak{r}_{j-1}\, \mathfrak{r}\, \mathfrak{r}_{j+1} \ldots \mathfrak{r}_n)\, \Psi_2(\mathfrak{r}_1 \ldots \mathfrak{r}_{j-1}\, \mathfrak{r}\, \mathfrak{r}_{j+1} \ldots \mathfrak{r}_n)$$
$$d\mathfrak{r}_1\, d\mathfrak{r}_2 \ldots d\mathfrak{r}_{j-1}\, d\mathfrak{r}_{j+} \ldots d\mathfrak{r}_n = 0. \tag{4}$$

2. Der Vollständigkeit halber behandeln wir zuerst das Zweielektronenproblem, obwohl es von Heisenberg und Dirac vollkommen erledigt ist und dabei nichts Neues mehr herauskommt.

Haben wir eine Eigenfunktion $\Psi(\mathfrak{r}_1 \mathfrak{r}_2)$, so können wir drei Fälle unterscheiden: a) Es ist für beliebige Werte von $\mathfrak{r}_1$ und $\mathfrak{r}_2$ $\Psi(\mathfrak{r}_1 \mathfrak{r}_2) = \Psi(\mathfrak{r}_2 \mathfrak{r}_1)$; in diesem Falle sagen wir, die Eigenfunktion ist symmetrisch. b) Es ist für beliebige Werte von $\mathfrak{r}_1$ und $\mathfrak{r}_2$ $\Psi(\mathfrak{r}_1 \mathfrak{r}_2) = -\Psi(\mathfrak{r}_2 \mathfrak{r}_1)$; in diesem Falle sagen wir, die Eigenfunktion ist antisymmetrisch. c) Weder a) noch b) trifft zu. In diesem Falle ist weder $\Psi(\mathfrak{r}_1 \mathfrak{r}_2) + \Psi(\mathfrak{r}_2 \mathfrak{r}_1)$ noch $\Psi(\mathfrak{r}_1 \mathfrak{r}_2) - \Psi(\mathfrak{r}_2 \mathfrak{r}_1)$ identisch Null. Wegen der Symmetrie der potentiellen Energie und der Linearität der Differentialgleichung repräsentieren beide Lösungen, und zwar zueinander orthogonale Lösungen des Eigenwertproblems, die zu demselben Eigenwert gehören, wie der ursprüngliche Ψ. Dieser Eigenwert ist also entartet, man kann jedoch die Entartung, wie dies Heisenberg gezeigt hat, durch eine kleine, symmetrische Störung aufheben, wodurch aus dem entarteten Eigenwert zwei einfache, aus der entarteten Eigenfunktion zwei unentartete entstehen, die dann natürlich (da sie einfach sind) entweder in die Kategorie a) oder b) gehören. Wir können demnach sagen, daß der Fall c) unstabil ist, da er bei einer kleinen Störung in a) oder b) übergeht und legen ihm keine weitere Bedeutung bei. Die Zustände mit den Eigenfunktionen der Sorte a) ($\underline{\Psi}$) kombinieren nicht mit den Zuständen der Sorte b) ($\overline{\Psi}$). Es ist nämlich

$$\int \underline{\Psi}(\mathfrak{r}\,\mathfrak{r}_2)\,\overline{\Psi}(\mathfrak{r}\,\mathfrak{r}_2)\,d\mathfrak{r}_2 + \int \underline{\Psi}(\mathfrak{r}_1\,\mathfrak{r})\,\overline{\Psi}(\mathfrak{r}_1\,\mathfrak{r})\,d\mathfrak{r}_1$$
$$= -\int \underline{\Psi}(\mathfrak{r}_2\,\mathfrak{r})\,\overline{\Psi}(\mathfrak{r}_2\,\mathfrak{r})\,d\mathfrak{r}_2 - \int \underline{\Psi}(\mathfrak{r}\,\mathfrak{r}_1)\,\overline{\Psi}(\mathfrak{r}\,\mathfrak{r}_1)\,d\mathfrak{r}_1.$$

Wenn wir jetzt auf der rechten Seite die Rolle der Integrationsvariablen $\mathfrak{r}_1$ und $\mathfrak{r}_2$ vertauschen, so sehen wir unmittelbar, daß der ganze Ausdruck verschwindet, also die Bedingung (4) erfüllt ist.

Wir gehen nun zu dem Falle von drei Elektronen über. Wir werden als potentielle Energien $V(\mathfrak{r}_1 \mathfrak{r}_2 \mathfrak{r}_3)$ und Störungen nur in $\mathfrak{r}_1 \mathfrak{r}_2 \mathfrak{r}_3$ symmetrische Funktionen zulassen, d. h. solche, in denen sich der Funktionswert nicht ändert, wenn wir die Werte der Argumente beliebig vertauschen. Es ist also für beliebige Werte von $\mathfrak{a}\,\mathfrak{b}\,\mathfrak{c}$ die potentielle Energie

$$V(\mathfrak{a}\,\mathfrak{b}\,\mathfrak{c}) = V(\mathfrak{a}\,\mathfrak{c}\,\mathfrak{b}) = V(\mathfrak{b}\,\mathfrak{a}\,\mathfrak{c}) = V(\mathfrak{b}\,\mathfrak{c}\,\mathfrak{a}) = V(\mathfrak{c}\,\mathfrak{a}\,\mathfrak{b}) = V(\mathfrak{c}\,\mathfrak{b}\,\mathfrak{a}). \quad (5)$$

Wir gehen in Anlehnung an Heisenberg von einer „ungestörten Bewegung" aus, in der die potentielle Energie als Summe von Funktionen darstellbar ist, die nur von dem einen Koordinatentripel, also entweder von $\mathfrak{r}_1$ oder von $\mathfrak{r}_2$ oder von $\mathfrak{r}_3$ abhängen, also wegen (5)

$$V_0(\mathfrak{r}_1 \mathfrak{r}_2 \mathfrak{r}_3) = v(\mathfrak{r}_1) + v(\mathfrak{r}_2) + v(\mathfrak{r}_3).$$

Die Gleichung für die ungestörte Eigenfunktion

$$\Delta \Psi(\mathfrak{r}_1 \mathfrak{r}_2 \mathfrak{r}_3) + k^2 \big(E - V_0(\mathfrak{r}_1 \mathfrak{r}_2 \mathfrak{r}_3)\big) \Psi(\mathfrak{r}_1 \mathfrak{r}_2 \mathfrak{r}_3) = 0 \quad (6)$$

wird also gelöst durch

$$\Psi_{ijk}(\mathfrak{r}_1\,\mathfrak{r}_2\,\mathfrak{r}_3) = \psi_i(\mathfrak{r}_1)\,\psi_j(\mathfrak{r}_2)\,\psi_k(\mathfrak{r}_3), \quad E_{ijk} = e_i + e_j + e_k, \tag{7}$$

worin für ψ und e die Lösungen der Gleichung

$$\Delta\,\psi(\mathfrak{r}) + k^2\,[e - v(\mathfrak{r})]\,\psi(\mathfrak{r}) = 0 \tag{8}$$

gesetzt sind. Wir nehmen an, daß die Lösungen von (8) nicht entartet sind.

Es wird sich folgendes zeigen: Wenn $i = j = k$, also alle Elektronen auf derselben Bahn sind, so ist der Eigenwert einfach, kann also nicht aufspalten; ist $i = j \neq k$, so ist er dreifach und spaltet in einen einfachen und einen doppelten Eigenwert auf; ist $i \neq j \neq k$, so ist er [wie Formel (9) zeigt] sechsfach und gibt zwei entartete und zwei unentartete Terme. Der eine hiervon bildet mit den unentarteten der vorigen beiden Fälle eine Gruppe (die Bose-Einsteinsche), der andere unentartete bildet eine Gruppe für sich (die Paulische), während die entarteten die dritte Gruppe bilden.

3. Wir betrachten zunächst den Fall $i \neq j \neq k$. Wie zu sehen, ist jeder Eigenwert sechsfach, da die sechs orthogonalen Eigenfunktionen

$$\left.\begin{array}{lll} \psi_i(\mathfrak{r}_1)\,\psi_j(\mathfrak{r}_2)\,\psi_k(\mathfrak{r}_3) & \psi_i(\mathfrak{r}_1)\,\psi_k(\mathfrak{r}_2)\,\psi_j(\mathfrak{r}_3) & \psi_j(\mathfrak{r}_1)\,\psi_i(\mathfrak{r}_2)\,\psi_k(\mathfrak{r}_3) \\ \psi_j(\mathfrak{r}_1)\,\psi_k(\mathfrak{r}_2)\,\psi_i(\mathfrak{r}_3) & \psi_k(\mathfrak{r}_1)\,\psi_i(\mathfrak{r}_2)\,\psi_j(\mathfrak{r}_3) & \psi_k(\mathfrak{r}_1)\,\psi_j(\mathfrak{r}_2)\,\psi_i(\mathfrak{r}_3) \end{array}\right\} \tag{9}$$

dazu gehören. Wir behandeln dieses System nach dem Vorgang von Schrödinger[1]) und setzen der Einfachheit halber $i = 1$, $j = 2$, $k = 3$. Die Störungsfunktion, die also auch in $\mathfrak{r}_1\,\mathfrak{r}_2\,\mathfrak{r}_3$ symmetrisch sein muß, bezeichnen wir mit $V'(\mathfrak{r}_1\,\mathfrak{r}_2\,\mathfrak{r}_3)$.

Eine gestörte Eigenfunktion ist, wenn die Störung klein ist, eine lineare Kombination[2]) derjenigen ungestörten, aus denen sie hervorgegangen ist. In unserem Falle der Ψ_{ijk} in (9)

$$\left.\begin{array}{l} b_{123}\,\psi_1(\mathfrak{r}_1)\,\psi_2(\mathfrak{r}_2)\,\psi_3(\mathfrak{r}_3) + b_{132}\,\psi_1(\mathfrak{r}_1)\,\psi_3(\mathfrak{r}_2)\,\psi_2(\mathfrak{r}_3) \\ + b_{213}\,\psi_2(\mathfrak{r}_1)\,\psi_1(\mathfrak{r}_2)\,\psi_3(\mathfrak{r}_3) + b_{231}\,\psi_2(\mathfrak{r}_1)\,\psi_3(\mathfrak{r}_2)\,\psi_1(\mathfrak{r}_3) \\ + b_{312}\,\psi_3(\mathfrak{r}_1)\,\psi_1(\mathfrak{r}_2)\,\psi_2(\mathfrak{r}_3) + b_{321}\,\psi_3(\mathfrak{r}_1)\,\psi_2(\mathfrak{r}_2)\,\psi_1(\mathfrak{r}_3). \end{array}\right\} \tag{10}$$

Die Gleichung für die Energiestörung ΔE ist sechsten Grades,

$$\begin{vmatrix} 123-\Delta E & 213 & 231 & 321 & 312 & 132 \\ 213 & 123-\Delta E & 132 & 312 & 321 & 231 \\ 312 & 132 & 123-\Delta E & 213 & 231 & 321 \\ 321 & 231 & 213 & 123-\Delta E & 132 & 312 \\ 231 & 321 & 312 & 132 & 123-\Delta E & 213 \\ 132 & 312 & 321 & 231 & 213 & 123-\Delta E \end{vmatrix} = 0, \tag{11}$$

$$b_{123} \qquad b_{213} \qquad b_{231} \qquad b_{321} \qquad b_{312} \qquad b_{132}$$

[1]) Ann. d. Phys. 80, 437, 1926.

[2]) l. c. S. 453.

worin

$$l\,m\,n = \int V'(\mathfrak{a}\,\mathfrak{b}\,\mathfrak{c})\,\psi_1(\mathfrak{a})\,\psi_2(\mathfrak{b})\,\psi_3(\mathfrak{c})\,\psi_l(\mathfrak{a})\,\psi_m(\mathfrak{b})\,\psi_n(\mathfrak{c})\,d\mathfrak{a}\,d\mathfrak{b}\,d\mathfrak{c}.$$

Sie ist die Bedingung für die Lösbarkeit der sechs linearen homogenen Gleichungen für die b_{123} usw. Die $b - s$ sind unter die entsprechenden Kolonnen eingetragen. (Es ist $2\,3\,1 = 3\,1\,2$.)

Gleichung (11) hat zwei einfache und zwei doppelte [1]) Wurzeln.

$$\left.\begin{aligned}
\Delta E_1 &= 1\,2\,3 + 1\,3\,2 + 2\,3\,1 + 2\,1\,3 + 3\,1\,2 + 3\,2\,1,\\
\Delta E_2 &= 1\,2\,3 - 1\,3\,2 + 2\,3\,1 - 2\,1\,3 + 3\,1\,2 - 3\,2\,1,\\
\Delta E_3 = \Delta E_4 &= 1\,2\,3 + 2\,3\,1\\
&+ \sqrt{1\,3\,2^2 + 3\,2\,1^2 + 2\,1\,3^2 - 2\,1\,3.1\,3\,2 - 3\,2\,1.1\,3\,2 - 3\,2\,1.2\,1\,3},\\
\Delta E_5 = \Delta E_6 &= 1\,2\,3 + 2\,3\,1\\
&- \sqrt{1\,3\,2^2 + 3\,2\,1^2 + 2\,1\,3^2 - 2\,1\,3.1\,3\,2 - 3\,2\,1.1\,3\,2 - 3\,2\,1.2\,1\,3}.
\end{aligned}\right\} \quad (12)$$

4. Die zu diesen gestörten Eigenwerten gehörenden Eigenfunktionen haben gewisse Transformationseigenschaften, die wir nun herleiten wollen.

Einfache Eigenwerte:

Wenn $\Psi_{123}(\mathfrak{r}_1\,\mathfrak{r}_2\,\mathfrak{r}_3)$ eine Lösung ist, so ist auch $\Psi_{ijk}(\mathfrak{r}_1\,\mathfrak{r}_2\,\mathfrak{r}_3) \equiv \Psi_{123}(\mathfrak{r}_i\,\mathfrak{r}_j\,\mathfrak{r}_k)$, wo $(i\,j\,k)$ eine Permutation von $(1\,2\,3)$ ist, eine Lösung. Da der Eigenwert einfach ist und also keine Entartung vorliegt, so ist es von vornherein klar, daß diese Funktionen bis auf einen konstanten Faktor, der wegen der Normierung nur $+1$ oder -1 sein kann, identisch sind. Und zwar sind entweder alle einander gleich, oder alle gehen durch jede Transposition der Argumente in den negativen Wert über. Die weiteren Fälle, daß eine Transposition den Wert ungeändert läßt, während eine andere das Vorzeichen wechselt, führen, wie leicht zu sehen, zu einem Widerspruch, z. B.

$$\begin{aligned}
\Psi(\mathfrak{r}_1\,\mathfrak{r}_2\,\mathfrak{r}_3) &= +\Psi(\mathfrak{r}_1\,\mathfrak{r}_3\,\mathfrak{r}_2) = -\Psi(\mathfrak{r}_3\,\mathfrak{r}_1\,\mathfrak{r}_2) = -\Psi(\mathfrak{r}_3\,\mathfrak{r}_2\,\mathfrak{r}_1)\\
&= +\Psi(\mathfrak{r}_2\,\mathfrak{r}_3\,\mathfrak{r}_1) = \Psi(\mathfrak{r}_2\,\mathfrak{r}_1\,\mathfrak{r}_3) = -\Psi(\mathfrak{r}_1\,\mathfrak{r}_2\,\mathfrak{r}_3).
\end{aligned}$$

Die Eigenfunktionen dieser beiden Typen nennen wir mit Dirac [2]) symmetrische bzw. antisymmetrische Eigenfunktionen.

Doppelte Eigenwerte:

Ich greife zwei linear unabhängige Ψ heraus und unterscheide zwei Fälle. a) Bei der Vertauschung des ersten mit dem letzten Argument

[1]) Man findet sie, indem man aus der Determinante (11) die zwei trivialen Lösungen ΔE_1 und ΔE_2 durch zweckmäßige Addition von Zeilen und Spalten abspaltet.

[2]) l. c.

bleiben entweder beide Funktionswerte ungeändert oder beide wechseln das Vorzeichen. Dann hat jede Eigenfunktion diese Eigenschaft. b) Die eine der beiden (Ψ_1) bleibt durch die angegebene Operation ungeändert, die andere (Ψ_2) wechselt ihr Vorzeichen. Den Fall, daß wenigstens eine Eigenfunktion (Ψ') bei dieser Vertauschung auch ihren absoluten Betrag ändert, kann man auf b) zurückführen, indem man andere Eigenfunktionen herausgreift:

$$\Psi_1(\mathfrak{r}_1\,\mathfrak{r}_2\,\mathfrak{r}_3) = \Psi'(\mathfrak{r}_1\,\mathfrak{r}_2\,\mathfrak{r}_3) + \Psi'(\mathfrak{r}_3\,\mathfrak{r}_2\,\mathfrak{r}_1),$$
$$\Psi_2(\mathfrak{r}_1\,\mathfrak{r}_2\,\mathfrak{r}_3) = \Psi'(\mathfrak{r}_1\,\mathfrak{r}_2\,\mathfrak{r}_3) - \Psi'(\mathfrak{r}_3\,\mathfrak{r}_2\,\mathfrak{r}_1).$$

Ψ_1 ist orthogonal zu Ψ_2. Ich leite hier nur für b) die Transformationseigenschaften ab.

Zunächst können wir nach Schema (9) vier weitere Lösungen angeben:

$$\left.\begin{aligned}
\Psi_3(\mathfrak{r}_1\,\mathfrak{r}_2\,\mathfrak{r}_3) &\equiv \Psi_1(\mathfrak{r}_2\,\mathfrak{r}_1\,\mathfrak{r}_3) = \Psi_1(\mathfrak{r}_3\,\mathfrak{r}_1\,\mathfrak{r}_2),\\
\Psi_4(\mathfrak{r}_1\,\mathfrak{r}_2\,\mathfrak{r}_3) &\equiv \Psi_1(\mathfrak{r}_1\,\mathfrak{r}_3\,\mathfrak{r}_2) = \Psi_1(\mathfrak{r}_2\,\mathfrak{r}_3\,\mathfrak{r}_1),\\
\Psi_5(\mathfrak{r}_1\,\mathfrak{r}_2\,\mathfrak{r}_3) &\equiv \Psi_2(\mathfrak{r}_2\,\mathfrak{r}_1\,\mathfrak{r}_3) = -\Psi_2(\mathfrak{r}_3\,\mathfrak{r}_1\,\mathfrak{r}_2),\\
\Psi_6(\mathfrak{r}_1\,\mathfrak{r}_2\,\mathfrak{r}_3) &\equiv \Psi_2(\mathfrak{r}_1\,\mathfrak{r}_3\,\mathfrak{r}_2) = -\Psi_2(\mathfrak{r}_2\,\mathfrak{r}_3\,\mathfrak{r}_1).
\end{aligned}\right\} \tag{13}$$

Da wir es aber mit einem nur doppelten Eigenwert zu tun haben, müssen die Ψ_3, Ψ_4, Ψ_5, Ψ_6 linear von den $\Psi_1(\mathfrak{r}_1\,\mathfrak{r}_2\,\mathfrak{r}_3)$ und $\Psi_2(\mathfrak{r}_1\,\mathfrak{r}_2\,\mathfrak{r}_3)$ abhängen. Wir setzen also an:

$$\left.\begin{aligned}
\Psi_3(\mathfrak{r}_1\,\mathfrak{r}_2\,\mathfrak{r}_3) &= \Psi_1(\mathfrak{r}_2\,\mathfrak{r}_1\,\mathfrak{r}_3) = \lambda_1\,\Psi_1(\mathfrak{r}_1\,\mathfrak{r}_2\,\mathfrak{r}_3) + \nu_1\,\Psi_2(\mathfrak{r}_1\,\mathfrak{r}_2\,\mathfrak{r}_3),\\
\Psi_4(\mathfrak{r}_1\,\mathfrak{r}_2\,\mathfrak{r}_3) &= \Psi_1(\mathfrak{r}_1\,\mathfrak{r}_3\,\mathfrak{r}_2) = \lambda_2\,\Psi_1(\mathfrak{r}_1\,\mathfrak{r}_2\,\mathfrak{r}_3) + \nu_2\,\Psi_2(\mathfrak{r}_1\,\mathfrak{r}_2\,\mathfrak{r}_3),\\
\Psi_5(\mathfrak{r}_1\,\mathfrak{r}_2\,\mathfrak{r}_3) &= \Psi_2(\mathfrak{r}_2\,\mathfrak{r}_1\,\mathfrak{r}_3) = \lambda_3\,\Psi_1(\mathfrak{r}_1\,\mathfrak{r}_2\,\mathfrak{r}_3) + \nu_3\,\Psi_2(\mathfrak{r}_1\,\mathfrak{r}_2\,\mathfrak{r}_3),\\
\Psi_6(\mathfrak{r}_1\,\mathfrak{r}_2\,\mathfrak{r}_3) &= \Psi_2(\mathfrak{r}_1\,\mathfrak{r}_3\,\mathfrak{r}_2) = \lambda_4\,\Psi_1(\mathfrak{r}_1\,\mathfrak{r}_2\,\mathfrak{r}_3) + \nu_4\,\Psi_2(\mathfrak{r}_1\,\mathfrak{r}_2\,\mathfrak{r}_3).
\end{aligned}\right\} \tag{14}$$

Diese Gleichungen geben also an, wie sich Ψ_1 und Ψ_2 transformiert, wenn man die Werte der ersten bzw. letzten beiden Variablen vertauscht.

Durch mehrfache Anwendung der angeschriebenen Regeln, sowie der Symmetrieeigenschaft von Ψ_1 und Ψ_2 in der ersten zur letzten Variable, lassen sich die λ und ν bestimmen. Wir erhalten dann folgende Gleichungen:

$$\left.\begin{aligned}
\Psi_1(\mathfrak{r}_1\,\mathfrak{r}_3\,\mathfrak{r}_2) &= \Psi_1(\mathfrak{r}_2\,\mathfrak{r}_3\,\mathfrak{r}_1) = -\frac{1}{2}\Psi_1(\mathfrak{r}_1\,\mathfrak{r}_2\,\mathfrak{r}_3) \pm \frac{\sqrt{3}}{2}\Psi_2(\mathfrak{r}_1\,\mathfrak{r}_2\,\mathfrak{r}_3),\\
\Psi_1(\mathfrak{r}_2\,\mathfrak{r}_1\,\mathfrak{r}_3) &= \Psi_1(\mathfrak{r}_3\,\mathfrak{r}_1\,\mathfrak{r}_2) = -\frac{1}{2}\Psi_1(\mathfrak{r}_1\,\mathfrak{r}_2\,\mathfrak{r}_3) \mp \frac{\sqrt{3}}{2}\Psi_2(\mathfrak{r}_1\,\mathfrak{r}_2\,\mathfrak{r}_3),\\
\Psi_2(\mathfrak{r}_1\,\mathfrak{r}_3\,\mathfrak{r}_2) &= -\Psi_2(\mathfrak{r}_2\,\mathfrak{r}_3\,\mathfrak{r}_1) = \pm\frac{\sqrt{3}}{2}\Psi_1(\mathfrak{r}_1\,\mathfrak{r}_2\,\mathfrak{r}_3) + \frac{1}{2}\Psi_2(\mathfrak{r}_1\,\mathfrak{r}_2\,\mathfrak{r}_3),\\
\Psi_2(\mathfrak{r}_2\,\mathfrak{r}_1\,\mathfrak{r}_3) &= -\Psi_2(\mathfrak{r}_3\,\mathfrak{r}_1\,\mathfrak{r}_2) = \mp\frac{\sqrt{3}}{2}\Psi_1(\mathfrak{r}_1\,\mathfrak{r}_2\,\mathfrak{r}_3) + \frac{1}{2}\Psi_2(\mathfrak{r}_1\,\mathfrak{r}_2\,\mathfrak{r}_3),
\end{aligned}\right\} \tag{14a}$$

wo entweder in allen Gleichungen das obere oder in allen das untere Vorzeichen gilt.

5. Wir wenden uns nun zurück zu 3. Zunächst sehen wir durch direkte Berechnung der Beziehungen zwischen den b_{123}, b_{132}, b_{213} usw., daß tatsächlich bei den doppelten Eigenwerten der im Punkt 4 mit b) bezeichnete Fall vorliegt. Wollen wir direkt die dort gewählten Ψ_1 und Ψ_2 erhalten, so ergibt sich für

$$\Psi_1 b^1_{123} = b^1_{321},\ b^1_{132} = b^1_{312},\ b^1_{231} = b^1_{312},\ b^1_{123} + b^1_{132} + b^1_{231} = 0,$$
$$\Psi_2 \sqrt{3}\, b^2_{123} = b^1_{132} - b^1_{213},\quad \sqrt{3}\, b^2_{132} = b^1_{123} - b^1_{213},$$
$$\sqrt{3}\, b^2_{213} = b^1_{132} - b^1_{123},\quad b^2_{123} = -b^2_{321},\quad b^2_{132} = -b^2_{312},$$
$$b^2_{231} = -b^2_{312},$$

die zusammen mit einer der Gleichungen, deren Determinante in (11) steht, die $b^1_{123}, b^1_{132}, \ldots$ und $b^2_{123}, b^2_{132}, \ldots$ eindeutig bestimmen.

Bringen wir nun noch eine weitere in $\mathfrak{r}_1 \mathfrak{r}_2 \mathfrak{r}_3$ symmetrische Störung $V''(\mathfrak{r}_1 \mathfrak{r}_2 \mathfrak{r}_3)$ an, so zeigt es sich, daß die Entartung dieser Zustände nicht mehr aufgehoben wird. Wir bezeichnen im folgenden die symmetrischen Eigenfunktionen mit Ψ, die antisymmetrischen mit $\overline{\Psi}$, die entarteten behalten die Bezeichnung Ψ_1 und Ψ_2 bei. Für die Erhaltung der Entartung ist nach Schrödinger zunächst notwendig, daß die Funktionaldeterminante, als Gleichung für ΔE aufgefaßt,

$$\begin{vmatrix} \int V''(\mathfrak{r}_1 \mathfrak{r}_2 \mathfrak{r}_3)\, \Psi_1(\mathfrak{r}_1 \mathfrak{r}_2 \mathfrak{r}_3)^2\, d\mathfrak{r}_1\, d\mathfrak{r}_2\, d\mathfrak{r}_3 - \Delta E, & \int V'' \Psi_1 \Psi_2, \\ \int V'' \Psi_1 \Psi_2, & \int V'' \Psi_2^2 - \Delta E, \end{vmatrix} = 0 \qquad (15)$$

doppelte Wurzeln habe. Nun ist wegen (14a) und (5)

$$\begin{aligned} A &= \int V''(\mathfrak{r}_1 \mathfrak{r}_2 \mathfrak{r}_3)\, \Psi_1(\mathfrak{r}_1 \mathfrak{r}_2 \mathfrak{r}_3)\, \Psi_2(\mathfrak{r}_1 \mathfrak{r}_2 \mathfrak{r}_3)\, d\mathfrak{r}_1\, d\mathfrak{r}_2\, d\mathfrak{r}_3 \\ &= -\int V''(\mathfrak{r}_3 \mathfrak{r}_2 \mathfrak{r}_1)\, \Psi_1(\mathfrak{r}_3 \mathfrak{r}_2 \mathfrak{r}_1)\, \Psi_2(\mathfrak{r}_3 \mathfrak{r}_2 \mathfrak{r}_1)\, d\mathfrak{r}_1\, d\mathfrak{r}_2\, d\mathfrak{r}_3 \end{aligned}$$

und

$$\begin{aligned} &\int V''(\mathfrak{r}_1 \mathfrak{r}_2 \mathfrak{r}_3)\, \Psi_1^2(\mathfrak{r}_1 \mathfrak{r}_2 \mathfrak{r}_3)\, d\mathfrak{r}_1\, d\mathfrak{r}_2\, d\mathfrak{r}_3 \\ &= \int V''(\mathfrak{r}_2 \mathfrak{r}_1 \mathfrak{r}_3) \left[-\tfrac{1}{2} \Psi_1(\mathfrak{r}_2 \mathfrak{r}_1 \mathfrak{r}_3) \mp \frac{\sqrt{3}}{2} \Psi_2(\mathfrak{r}_2 \mathfrak{r}_1 \mathfrak{r}_3) \right]^2 d\mathfrak{r}_1\, d\mathfrak{r}_2\, d\mathfrak{r}_3 \\ &= \tfrac{1}{4} \int V''(\mathfrak{r}_2 \mathfrak{r}_1 \mathfrak{r}_3)\, \Psi_1^2(\mathfrak{r}_2 \mathfrak{r}_1 \mathfrak{r}_3)\, d\mathfrak{r}_1\, d\mathfrak{r}_2\, d\mathfrak{r}_3 \\ &\quad + \tfrac{3}{4} \int V''(\mathfrak{r}_2 \mathfrak{r}_1 \mathfrak{r}_3)\, \Psi_2(\mathfrak{r}_2 \mathfrak{r}_1 \mathfrak{r}_3)^2\, d\mathfrak{r}_1\, d\mathfrak{r}_2\, d\mathfrak{r}_3 \pm \frac{\sqrt{3}}{2} A. \end{aligned}$$

Nun können wir die Rolle der Integrationsvariablen $\mathfrak{r}_1$ und $\mathfrak{r}_3$ bzw. $\mathfrak{r}_2$ vertauschen und erhalten

$$A = \int V'' \Psi_1 \Psi_2 = 0, \quad \int V'' \Psi_1^2 = \int V'' \Psi_2^2. \qquad (16)$$

Gleichung (15) hat also die doppelte Wurzel $\Delta E = \int V'' \Psi_1^2\, d\mathfrak{r}_1\, d\mathfrak{r}_2\, d\mathfrak{r}_3$.

Außer diesen Gleichungen (16), deren erste sich gemäß ihrer Ableitung nicht nur auf zwei, aus derselben ungestörten Eigenfunktion entstandene Ψ_1 und Ψ_2 beziehen, sondern für ein beliebiges Paar Ψ_1, Ψ_2 gültig sind, brauchen wir noch folgende, die sich ganz analog herleiten lassen:

$$\left.\begin{aligned} \int V'' \Psi_1 \underline{\Psi}\, d\mathfrak{r}_1\, d\mathfrak{r}_2\, d\mathfrak{r}_3 &= \int V'' \Psi_1 \overline{\Psi}\, d\mathfrak{r}_1\, d\mathfrak{r}_2\, d\mathfrak{r}_3 = 0, \\ \int V'' \Psi_2 \underline{\Psi}\, d\mathfrak{r}_1\, d\mathfrak{r}_2\, d\mathfrak{r}_3 &= \int V'' \Psi_2 \overline{\Psi}\, d\mathfrak{r}_1\, d\mathfrak{r}_2\, d\mathfrak{r}_3 = 0, \\ \int V'' \underline{\Psi}\, \overline{\Psi}\, d\mathfrak{r}_1\, d\mathfrak{r}_2\, d\mathfrak{r}_3 &= 0. \end{aligned}\right\} \tag{16a}$$

Diese Gleichungen (16) und (16a) drücken die Tatsache aus, daß das Integral über das Produkt einer beliebigen symmetrischen Störungsfunktion mit zwei $\Psi - s$ verschiedener Gattung immer verschwindet. Diese Integrale sind aber nach Schrödinger nichts anderes als die Entwicklungskoeffizienten einer gestörten Eigenfunktion der einen Gattung nach den ungestörten Eigenfunktionen einer anderen Gattung. Aus dem Verschwinden dieser Koeffizienten ergibt sich, daß für Ψ_1 und Ψ_2 die Gleichungen (14a) auch bei einer endlichen Störung richtig sind und auch $\underline{\Psi}$ und $\overline{\Psi}$ bei einer solchen ihre Symmetrie- bzw. Antisymmetrieeigenschaft beibehält.

Um den Zerfall des Termsystems in der in 2. angedeuteten Weise zu beweisen, müssen wir noch das Verschwinden des Ausdrucks (4) als Folge der Transformationseigenschaften der $\underline{\Psi}, \overline{\Psi}, \Psi_1, \Psi_2$ nachweisen. Es ergibt sich tatsächlich z. B.

$$\begin{aligned} &\int \underline{\Psi}(\mathfrak{r}\,\mathfrak{r}_2\,\mathfrak{r}_3)\,\overline{\Psi}(\mathfrak{r}\,\mathfrak{r}_2\,\mathfrak{r}_3)\,d\mathfrak{r}_2\,d\mathfrak{r}_3 + \int \underline{\Psi}(\mathfrak{r}_1\,\mathfrak{r}\,\mathfrak{r}_3)\,\overline{\Psi}(\mathfrak{r}_1\,\mathfrak{r}\,\mathfrak{r}_3)\,d\mathfrak{r}_1\,d\mathfrak{r}_3 \\ &+ \int \underline{\Psi}(\mathfrak{r}_1\,\mathfrak{r}_2\,\mathfrak{r})\,\overline{\Psi}(\mathfrak{r}_1\,\mathfrak{r}_2\,\mathfrak{r})\,d\mathfrak{r}_1\,d\mathfrak{r}_2 = -\int \underline{\Psi}(\mathfrak{r}\,\mathfrak{r}_3\,\mathfrak{r}_2)\,\overline{\Psi}(\mathfrak{r}\,\mathfrak{r}_3\,\mathfrak{r}_2)\,d\mathfrak{r}_2\,d\mathfrak{r}_3 \\ &- \int \underline{\Psi}(\mathfrak{r}_3\,\mathfrak{r}\,\mathfrak{r}_1)\,\overline{\Psi}(\mathfrak{r}_3\,\mathfrak{r}\,\mathfrak{r}_1)\,d\mathfrak{r}_1\,d\mathfrak{r}_3 - \int \underline{\Psi}(\mathfrak{r}_2\,\mathfrak{r}_1\,\mathfrak{r})\,\overline{\Psi}(\mathfrak{r}_2\,\mathfrak{r}_1\,\mathfrak{r})\,d\mathfrak{r}_1\,d\mathfrak{r}_2 = 0. \end{aligned}$$

und ähnlich die anderen.

6. Wenn wir von einem Zustand ausgegangen wären, der alle drei Elektronen in demselben Zustand enthält, so hätten wir als ungestörte Eigenfunktion $\psi_i(\mathfrak{r}_1)\,\psi_i(\mathfrak{r}_2)\,\psi_i(\mathfrak{r}_3)$, als gestörte nur ein symmetrisches $\underline{\Psi}$ erhalten.

Wenn nur zwei Teilchen auf derselben Bahn, das dritte dagegen auf einer anderen Bahn war, wenn also $i = j \neq k$, so ergibt sich eine symmetrische und eine entartete Eigenfunktion. Wir erhalten also mit Heisenberg das System der antisymmetrischen Eigenfunktion als das einzige, in dem das Paulische Verbot gilt.

Aus dieser Aufzählung ersehen wir auch, daß allen Zuständen, die mit der Angabe der Zellen bestimmt sind, in denen ein Elektron vorhanden ist, nur eine und immer eine symmetrische Eigenfunktion zugeordnet werden kann. Wenn wir allen symmetrischen Eigenfunktionen das gleiche statistische Gewicht beilegen, so haben wir, wie dies schon Dirac bemerkt hat, den Fall der Bose-Einsteinschen Statistik vor uns.

In einer zweiten Mitteilung hoffe ich die gleiche Aufgabe für mehrere Elektronen zu lösen.

Berlin, Institut für theoretische Physik der Technischen Hochschule.

Über nicht kombinierende Terme in der neueren Quantentheorie. Zweiter Teil

E. P. Wigner

Zeitschrift für Physik *40*, 883–892 (1927)

(Eingegangen am 26. November 1926.)

Es wird der Termzerfall bei dem Mehrelektronenproblem behandelt. Die Eigenrotation des Elektrons bleibt unberücksichtigt.

1. Im ersten Teil konnte man drei verschiedene Arten von Eigenfunktionen (symmetrische, antisymmetrische und entartete) unterscheiden, die in der Heisenberg-Schrödingerschen Theorie einem System mit drei absolut gleichen Teilchen zugeordnet waren. Die Rechnung war ziemlich umständlich, sie erforderte z. B. die Lösung einer (reduziblen) Gleichung sechsten Grades und dergleichen. Es ist klar, daß man diese elementaren Methoden schon bei vier Elektronen kaum anwenden kann, da die rechnerischen Schwierigkeiten zu groß werden. Es existiert aber eine wohl ausgebildete mathematische Theorie, die man hier verwenden kann: die Theorie der mit der symmetrischen Gruppe (der Gruppe der Permutationen) isomorphen Transformationsgruppen, die am Ende des vorigen Jahrhunderts von Frobenius[1]) begründet und später unter anderem von W. Burnside[2]) und J. Schur[3]) ausgebildet worden ist. Auf diese Arbeiten hat mich Herr J. von Neumann freundlichst hingewiesen, der mir auch, als ich ihm das Resultat der Rechnung für $n = 3$ mitteilte, das allgemeine Resultat richtig voraussagte.

Da die erwähnte Theorie wohl nicht allgemein bekannt sein dürfte, werde ich im nächsten Abschnitt (2) die Resultate mitteilen und erst weiter unten einen Beweis erbringen, dessen Gedankengang sich an die Rechnungen im 1. Teil anschließen wird, nur daß die erwähnte Theorie hier weitgehend verwendet wird.

2. Liegt ein System mit n Elektronen vor, so schreibe man die Zahl n in Form einer solchen Summe auf, bei der jeder Summand größer

---

[1]) Berl. Ber. 1896, S. 985, 1343; 1897, S. 994; 1899, S. 482.

[2]) Acta mathematica **28**, 369, 1904. Proc. London Math. Soc. (2) **1**, 117, 1904.

[3]) Berl. Ber. 1905, S. 406; 1908, S. 664 (diese beiden Veröffentlichungen enthalten alles, was wir brauchen).

oder gleich dem vorangehenden ist. Jeder solchen Zerlegung ist dann ein Termsystem zugeordnet. Die Anzahl der möglichen Zerlegungen, die zugleich die Anzahl der untereinander nicht kombinierenden Termsysteme angibt, bezeichnet man als „partitio numerorum" von n. Es sind z. B. für $n = 2$ die Zerlegungen: $1 + 1$, 2, möglich, für $n = 3$ haben wir drei: $1 + 1 + 1$, $1 + 2$, 3, für $n = 4$: $1 + 1 + 1 + 1$, $1 + 1 + 2$, $1 + 3$, $2 + 2$, 4. Nach dem letzten Beispiel ist die partitio numerorum für $n = 4$ gleich 5, also existieren bei dem Vierelektronenproblem 5 untereinander nicht kombinierende Termsysteme.

Die Anzahl der linear unabhängigen Lösungen, die zu einem Term gehören (der Grad der Entartung), ist für alle Terme des Termsystems dieselbe. Um sie zu bestimmen, geht man von jener Zerlegung von n in Summanden aus, welcher das Termsystem zugeordnet ist. Ist diese $\lambda_1 + \lambda_2 + \cdots + \lambda_i$, so ist der Term

$$N_{\lambda_1 \lambda_2 \ldots \lambda_i} = \frac{n!}{\lambda_1!\,(\lambda_2 + 1)! \ldots (\lambda_i + i - 1)!} \prod_{j<l} (\lambda_l - \lambda_j + l - j) \quad (1)$$

fach entartet. Für N gilt die Rekursionsformel

$$N_{\lambda_1 \lambda_2 \ldots \lambda_i} = N_{\lambda_1 - 1 \lambda_2 \ldots \lambda_i} + N_{\lambda_1 \lambda_2 - 1 \ldots \lambda_i} + \cdots + N_{\lambda_1 \lambda_2 \ldots \lambda_i - 1}. \quad (2)$$

Dabei sind alle N gleich Null zu setzen, bei denen in den λ eine absteigende Folge da ist; ist $\lambda_1 = 0$, so ist einfach $N_{\lambda_2 \lambda_3 \ldots \lambda_i}$ zu schreiben; $N_1 = 1$. Es ist also z. B. bei dem Term der zu der Zerlegung $2 + 2$ gehört,

$$\begin{aligned} N_{22} &= N_{12} + N_{21}, & N_{21} &= 0, \\ N_{12} &= N_{02} + N_{11}, & N_{02} &= N_2, \quad N_2 = N_1 = 1, \\ N_{11} &= N_{01} + N_{10}, & N_{10} &= 0, \quad N_{01} = N_1 = 1, \end{aligned}$$

also ist $N_{22} = 2$. Ebenso ist

$$N_{1111} = 1, \quad N_4 = 1, \quad N_{13} = 3, \quad N_{112} = 3.$$

In einem System mit n gleichen Massenpunkten, zwischen denen zunächst keine Wechselwirkungsenergie da ist, ist jeder Eigenwert (sofern der betreffende Zustand keine äquivalenten Bahnen enthält) $n!$ fach entartet. Läßt man eine Wechselwirkung entstehen, so spaltet jeder Eigenwert in mehrere Werte auf. Von diesen gehören so viele einem Termsystem $\lambda_1 \lambda_2 \ldots \lambda_i$ an, als das $N_{\lambda_1 \lambda_2 \ldots \lambda_i}$ dieser Gruppe beträgt [$N_{\lambda_1 \lambda_2 \ldots \lambda_i}$ gibt gemäß (1) zugleich auch den Grad der Entartung des betreffenden Termsystems an]. Also entsteht auf diese Weise bei $n = 4$ ein einfacher ($N_{1111} = 1$), 3 dreifach entartete ($N_{112} = 3$), noch 3 dreifach

entartete ($N_{13} = 3$), 2 zweifach entartete ($N_{22} = 2$), noch ein einfacher ($N_4 = 1$) Term.

Wenn wir bedenken, daß ein $N_{\lambda_1 \lambda_2 \ldots \lambda_i}$ fach entarteter Term eigentlich $N_{\lambda_1 \lambda_2 \ldots \lambda_i}$ fach zu zählen ist, so erhalten wir für die so gezählte Anzahl der Terme, die aus einem ungestörten Term entstanden sind,

$$\sum_{\lambda_1 \ldots \lambda_i} (N_{\lambda_1 \lambda_2 \ldots \lambda_i})^2 = n! \tag{3}$$

wenn $n = 4$, z. B.

$$N_{1111}^2 + N_{112}^2 + N_{13}^2 + N_{22}^2 + N_4^2 = 1 + 9 + 9 + 4 + 1 = 24 = 4!$$

wie es ja auch sein muß.

Die Terme, die verschiedenen Zerlegungen von n entsprechen, kombinieren nicht miteinander, dagegen werden Kombinationen zwischen verschiedenen Termen, die zu derselben Zerlegung von n in Summanden gehören, im allgemeinen vorkommen.

3. Zum Verständnis des folgenden ist die Kenntnis der Grundlagen der Frobenius-Schurschen Theorie notwendig.

Wir gehen wieder, wie im 1. Teil, von einer „ungestörten Bewegung" aus, die keine äquivalenten Bahnen enthält, und bringen eine kleine Störung an. Die $\psi(\mathfrak{r})$ sind die Lösungen der Gleichung (8) im 1. Teil, zur Abkürzung sei die Bezeichnung

$$\varphi_R(\mathfrak{r}_1 \mathfrak{r}_2 \ldots \mathfrak{r}_n) = \psi_1(\mathfrak{r}_{\alpha_1}) \cdot \psi_2(\mathfrak{r}_{\alpha_2}) \ldots \psi_n(\mathfrak{r}_{\alpha_n}) \tag{4}$$

eingeführt, wobei R die Permutation $\alpha_1 \alpha_2 \ldots \alpha_n$ ist.

Die gestörten Ψ sind lineare Kombinationen der $\varphi_R(\mathfrak{r}_1 \ldots \mathfrak{r}_n)$, und zwar sei

$$\sum_{R=E}^{H_{h-1}} b_R^k \varphi_R = \Psi_k, \tag{5}$$

wo die Summation über alle Permutationen, also über alle Elemente der symmetrischen Gruppe ($E, H_1, H_2 \ldots H_{h-1}$; es ist $h = n!$) zu erstrecken ist (E ist die identische Permutation $1\,2 \ldots n$). Den Index k wählen wir zur Unterscheidung der verschiedenen Ψ, da ja, wie wir aus dem Fall von drei Elektronen wissen, mehrere entstehen werden. Für das Produkt zweier Permutationen R und S schreiben wir einfach RS.

Wir bezeichnen weiter mit J das Integral

$$J_R = \int V'(\mathfrak{r}_1 \ldots \mathfrak{r}_n)\, \varphi_E(\mathfrak{r}_1 \ldots \mathfrak{r}_n)\, \varphi_R(\mathfrak{r}_1 \ldots \mathfrak{r}_n)\, d\mathfrak{r}_1 \ldots d\mathfrak{r}_n, \tag{6}$$

worin V' die zusätzliche potentielle Energie, die von der Störung herrührt, $d\mathfrak{r}_1 \ldots d\mathfrak{r}_n$ die Integration über alle Koordinaten andeutet.

Die Gleichung für die Energiestörung $\Delta\varepsilon$ ist $n! = h$ ten Grades.

$$\begin{vmatrix} J_E - \Delta\varepsilon & J_{H_1} & J_{H_2} & J_{H_3} & \dots & J_{H_{h-1}} \\ J_{H_1^{-1}} & J_E - \Delta\varepsilon & J_{H_2 H_1^{-1}} & J_{H_3 H_1^{-1}} & \dots & J_{H_{h-1} H_1^{-1}} \\ J_{H_2^{-1}} & J_{H_1 H_2^{-1}} & J_E - \Delta\varepsilon & J_{H_3 H_2^{-1}} & \dots & J_{H_{h-1} H_2^{-1}} \\ \dots & \dots & \dots & \dots & \dots & \dots \\ J_{H_{h-1}^{-1}} & J_{H_1 H_{h-1}^{-1}} & J_{H_2 H_{h-1}^{-1}} & J_{H_3 H_{h-1}^{-1}} & \dots & J_E - \Delta\varepsilon \end{vmatrix} = 0. \tag{7}$$

Es ist hierin zu beachten, daß $J_R = J_{R^{-1}}$ ist, wie man durch Vertauschung der Integrationsvariablen in (6) ohne weiteres sieht, daß also die Determinante (7) symmetrisch ist, und demgemäß lauter reelle Lösungen hat.

Nun betrachten wir eine k-dimensionale irreduzible Darstellung der symmetrischen Gruppe. Ihr, der Permutation R entsprechendes Element, bezeichnen wir mit

$$\begin{pmatrix} a_{11}^R & a_{12}^R & \dots & a_{1k}^R \\ a_{21}^R & a_{22}^R & \dots & a_{2k}^R \\ \dots & \dots & \dots & \dots \\ a_{k1}^R & a_{k2}^R & \dots & a_{kk}^R \end{pmatrix} \tag{8}$$

Es ist $a_{ij}^E = \delta_{ij}$; (a_{ij}^E) ist die Einheitsmatrix. Wir bilden folgende Gleichung

$$\begin{vmatrix} \sum\limits_R a_{11}^R J_R - \Delta\varepsilon & \sum\limits_R a_{12}^R J_R & \dots & \sum\limits_R a_{1k}^R J_R \\ \sum\limits_R a_{21}^R J_R & \sum\limits_R a_{22}^R J_R - \Delta\varepsilon & \dots & \sum\limits_R a_{2k}^R J_R \\ \vdots & \vdots & \ddots & \vdots \\ \sum\limits_R a_{k1}^R J_R & \sum\limits_R a_{k2}^R J_R & \dots & \sum\limits_R a_{kk}^R J_R - \Delta\varepsilon \end{vmatrix} = 0. \tag{9}$$

Gleichung (9) ist nichts anderes als die Gruppenmatrix unserer irreduziblen Darstellung mit den Variablen $J_E - \Delta\varepsilon$, J_{H_1}, $J_{H_2} \dots J_{H_{h-1}}$ und ist also[1]) eine irreduzible Gleichung, z. B. für $J_E - \Delta\varepsilon$, also auch für $\Delta\varepsilon$. Insbesondere hat sie also keine doppelten Wurzeln.

Eine ihrer Wurzeln sei $\Delta\varepsilon_1$. Dann bilden wir folgendes Gleichungssystem für die b_E^j, als Unbekannte:

$$\sum_j \sum_R a_{\varkappa j}^R J_R b_E^j - \Delta\varepsilon_1 b_E^\varkappa = 0 \quad (\varkappa = 1\,2 \dots k) \tag{10}$$

Da seine Determinante verschwindet, hat es eine Lösung und wir bestimmen die b_E^j. Wir setzen nun

$$b_R^\varkappa = \sum_{j=1}^{k} a_{\varkappa j}^{R^{-1}} b_E^j. \tag{11}$$

[1]) J. Schur, Berl. Ber. 413, 1905. Bei uns ist zwar $J_R = J_{R^{-1}}$, doch ist dies, wie man leicht einsieht belanglos.

Wir wollen nun beweisen, daß die so bestimmten $b_R^\varkappa$ die Gleichungen, die der Determinante (7) zugrunde liegen, befriedigen, und zwar für $\Delta\varepsilon = \Delta\varepsilon_1$. Es ist nämlich

$$\sum_R b_R^\varkappa J_{RS^{-1}} - b_S^\varkappa \Delta\varepsilon_1 = \sum_R b_R^\varkappa J_{SR^{-1}} - b_S^\varkappa \Delta\varepsilon_1$$

$$= \sum_{Rj} a_{\varkappa j}^{S^{-1}} {}^{SR^{-1}} b_E^j J_{SR^{-1}} - \sum_l a_{\varkappa l}^{S^{-1}} b_E^l \Delta\varepsilon_1$$

$$= \sum_{Rjl} a_{\varkappa l}^{S^{-1}} \left(a_{lj}^{SR^{-1}} b_E^j J_{SR^{-1}} - b_E^l \Delta\varepsilon_1\right) = 0$$

Also befriedigen die in (11) definierten $b_R^\varkappa$ im Verein mit dem $\Delta\varepsilon$ von (9) die Gleichungen. Wenn wir in (11) für $\varkappa$ der Reihe nach $1, 2 \ldots k$ setzen, so haben wir k Lösungen, die alle zu demselben $\Delta\varepsilon$ gehören. Ich behaupte nun, die k Funktionen

$$\sum_R b_R^\varkappa y_R \quad (\varkappa = 1, 2 \ldots k) \tag{12}$$

sind linear unabhängig. Denn sollten unter ihnen die Gleichungen

$$\sum_{\varkappa=1}^{k} \alpha_\varkappa \sum_R b_R^\varkappa y_R = 0$$

identisch in den y bestehen, so könnte ich unsere Darstellung in eine äquivalente überführen, in der aber in der ersten Zeile jeder Matrix nur das erste Element von Null verschieden ist. Mithin wäre unsere Darstellung reduzibel. Die Rechnung, da sie ziemlich weitläufig ist, sei hier übergangen.

[Die Transformationsmatrix ist

$$\begin{vmatrix} \alpha_1 & \alpha_2 & \alpha_3 & \ldots & \alpha_k \\ 0 & 1 & 0 & \ldots & 0 \\ 0 & 0 & 1 & \ldots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \ldots & 1 \end{vmatrix} \quad \text{ihre reziproke} \quad \begin{vmatrix} \frac{1}{\alpha_1} & -\frac{\alpha_2}{\alpha_1} & \ldots & -\frac{\alpha_k}{\alpha_1} \\ 0 & 1 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & 1 \end{vmatrix}$$

alle b_R^1 werden zu Null, die übrigen b_R^j ändern sich nicht. Ähnlich verfahren wir, wenn mehrere Gleichungen der Art (12) bestehen.]

Nun können wir leicht sehen, daß die zu einer k-dimensionalen irreduziblen Darstellung gehörende $\Delta\varepsilon$ eine k-fache Wurzel der Gleichung (7) ist. Da nämlich das zu (7) gehörende Gleichungssystem bei diesem $\Delta\varepsilon$ k linear unabhängige Lösungssysteme hat (die b_R^j für $j = 1, 2 \ldots k$), müssen alle $h - k$-reihigen Minoren verschwinden, insbesondere auch die Hauptminoren. Die Summe der $h - l$-reihigen Hauptminoren gibt aber den l-ten Differentialquotienten des Ausdrucks (7)

nach $\Delta\varepsilon$ an. Da diese, wenn $l < k$, verschwindet, ist unser $\Delta\varepsilon$ tatsächlich eine k-fache Wurzel.

Zu jeder k-dimensionalen irreduziblen Darstellung gehört also durch (9) eine k-fache Wurzel. Andererseits sind die Gleichungen (9) lauter irreduzible algebraische Gleichungen[1]) und für nicht äquivalente Darstellungen verschieden[2]), so daß wir wirklich durch jede neue Darstellung neue Wurzeln der Gleichung (7) bekommen. Im ganzen bekommen wir auf diese Weise, wie aus der Theorie der Darstellungen ohne weiteres hervorgeht, eben h. Da Gleichung (7) nur h-ten Grades ist, haben wir eben alle Wurzeln.

Aus den Sätzen über die irreduziblen Darstellungen einer symmetrischen Gruppe geht nun ohne weiteres hervor, daß wir in erster Näherung eben diejenigen Termsysteme erhalten, die im Abschnitt 2 angegeben sind.

In der Tat stellen die dort angegebenen Termsysteme nichts anderes dar, als die Resultate der Theorie der irreduziblen Darstellungen der symmetrischen Gruppe[3]), soweit sie sich auf die Anzahl derselben beziehen. Ich brauche sie deshalb hier nicht zu wiederholen.

4. Unser nächster Schritt, um zu zeigen, daß hieran sich auch bei einer endlichen Störung nichts ändert, ist nun, wie im 1. Teil, die Angabe, wie sich die Eigenfunktionen transformieren, wenn wir eine Permutation der Argumente (der $\mathfrak{r}$) vornehmen. Zu diesem Zwecke schreiben wir zunächst folgende Formel hin:

$$\left.\begin{aligned} b^{\varkappa}_{RS} &= \sum_{j=1}^{k} a^{S^{-1}R^{-1}}_{\varkappa j}\, b^{j}_{E} = \sum_{j=1}^{k}\sum_{l=1}^{k} a^{S^{-1}}_{\varkappa l}\, a^{R^{-1}}_{lj}\, b^{j}_{E} \\ &= \sum_{l=1}^{k} a^{S^{-1}}_{\varkappa l}\, b^{l}_{R}. \end{aligned}\right\} \tag{13}$$

Daraus folgt, da [siehe (4), (5); S ist die Permutation $\beta_1\beta_2\ldots\beta_n$]

$$\left.\begin{aligned} &\Psi_{\varkappa}(\mathfrak{r}_1\ldots\mathfrak{r}_2) = \sum_{R} b^{\varkappa}_{R}\,\varphi_R(\mathfrak{r}_1\ldots\mathfrak{r}_n) \\ &\Psi_{\varkappa}(\mathfrak{r}_{\beta_1}\mathfrak{r}_{\beta_2}\ldots\mathfrak{r}_{\beta_n}) = \sum_{R} b^{\varkappa}_{R}\,\varphi_R(\mathfrak{r}_{\beta_1}\mathfrak{r}_{\beta_2}\ldots\mathfrak{r}_{\beta_n}) \\ &= \sum_{R} b^{\varkappa}_{R}\,\varphi_{RS}(\mathfrak{r}_1\ldots\mathfrak{r}_n) = \sum_{R} b^{\varkappa}_{RS^{-1}}\,\varphi_R = \sum_{R}\sum_{l} a^{S}_{\varkappa l}\, b^{l}_{R}\,\varphi_R \\ &= \sum_{l} a^{S}_{\varkappa l}\,\Psi_l(\mathfrak{r}_1\ldots\mathfrak{r}_n). \end{aligned}\right\} \tag{14}$$

---

[1]) J. Schur, l. c. 1905, S. 413.
[2]) S. 409.
[3]) Berl. Ber. 1908, S. 665, 666.

Dies Resultat besagt, daß, wenn wir eine Permutation der Argumente an einem Ψ vornehmen, die so entstandene Funktion sich als lineares Aggregat der k linear unabhängigen Ψ darstellen läßt, und zwar sind die Koeffizienten Elemente derjenigen Matrix der irreduziblen Darstellung die zu der betreffenden Permutation gehört.

Dies ist nun in keiner Weise verwunderlich. Wenn $\Psi_\varkappa(\mathfrak{r}_1 \dots \mathfrak{r}_n)$ eine Lösung unseres Eigenwertproblems ist, so ist auch $\Psi_\varkappa(\mathfrak{r}_{\beta_1} \dots \mathfrak{r}_{\beta_n})$ eine Lösung, da aber nur die Lösungen $\Psi_1(\mathfrak{r}_1 \dots \mathfrak{r}_n)$, $\Psi_2(\mathfrak{r}_1 \dots \mathfrak{r}_n) \dots$ $\Psi_k(\mathfrak{r}_1 \dots \mathfrak{r}_n)$ voneinander linear unabhängig sind, lassen sich alle anderen Lösungen z. B. $\Psi_\varkappa(\mathfrak{r}_{\beta_1} \dots \mathfrak{r}_{\beta_n})$ durch diese ausdrücken. Auch sieht man leicht ein, daß die Koeffizienten eine Darstellung der Permutationsgruppe bilden müssen. Wenn wir nämlich nacheinander zwei Permutationen R und S der Argumente vornehmen, so erhalten wir bei leicht verständlicher Bezeichnung [es ist $F(S)$ diejenige Funktion, die aus der Funktion $F(E) = F(x_1 x_2 \dots x_n)$ durch die Permutation S der Argumente hervorgeht: $F(S) = F(x_{\beta_1}, x_{\beta_2}, \dots x_{\beta_n})$]

$$\left.\begin{aligned} \Psi_\varkappa(R) &= \sum_l a^R_{\varkappa l}\, \Psi_l(E) \\ \Psi_\varkappa(RS) &= \sum_l a^R_{\varkappa l}\, \Psi_l(S) = \sum_l a^R_{\varkappa l} \sum_j a^S_{lj}\, \Psi_j(E) \end{aligned}\right\} \tag{15}$$

andererseits ist

$$\Psi_\varkappa(RS) = \sum_j a^{RS}_{\varkappa j}\, \Psi_j(E), \tag{15a}$$

woraus wegen der linearen Unabhängigkeit der $\Psi_j(E)$

$$a^{RS}_{\varkappa j} = \sum_l a^R_{\varkappa l}\, a^S_{lj} \tag{15b}$$

hervorgeht, was eben die Tatsache ausdrückt, daß, wenn $RS = U$, so folgt $(a^R_{ik}) \cdot (a^S_{ik}) = (a^U_{ik})$.

Daß die hier vorkommenden Darstellungen irreduzibel sind, geht daraus hervor, daß der Eigenwert, wenn die in (14) vorkommenden Koeffizienten einer reduziblen Darstellung angehören würden, bei einer weiteren Störung im allgemeinen aufspalten würde, während dies hier, wie wir sehen werden, nie der Fall sein wird.

Die in (14) ausgedrückte Substitutionsformel ist also in gewissem Sinne trivial und wir hatten auch unsere bisherigen Ableitungen nur dazu nötig, um festzustellen, wieviel k-fach entartete Eigenwerte im Termsystem vorkommen werden. Daß die Eigenfunktionen, die zu einem k-fach entarteten Eigenwert gehören, wirklich der Formel (14) genügen, geht ebensogut aus den drei Zeilen (15), (15a), (15b) hervor. Für die weiteren Schlüsse (den Fall endlicher Störungen) werden wir aber nur

die Formel (14) benutzen, da nur sie bei einer endlichen Störung richtig bleibt, während die Formeln (4) und (5) selbstverständlich versagen.

5. Nun suchen wir unter unseren k-linear unabhängigen Ψ gewisse heraus, wie wir dies schon bei dem Dreielektronenfall mit der Wahl von Ψ_1 und Ψ_2 getan haben. Hier wenden wir folgende Vorschrift an.

Wir betrachten die irreduzible Darstellung, die zu unserem speziellen $\Psi_1, \Psi_2 \ldots \Psi_k$ gehört, und deren Koeffizienten auch in (14) vorkommen[1]). Wir nehmen nun diejenige Transformation (T), durch welche diese Darstellung orthoganilisiert wird. Bekanntlich existiert eine solche immer[2]). Dann bilden wir folgende linear unabhängige Funktionen [t_{ik} sind die Elemente von (T), τ_{ik} von $(T)^{-1}$, $|T| = |t_{ik}| \neq 0$)]

$$\Phi_i = \sum_{\varkappa=1}^{k} t_{i\varkappa} \Psi_\varkappa.$$

Wir finden nun

$$\Phi_i(R) = \sum_{\varkappa} t_{i\varkappa} \Psi_\varkappa(R) = \sum_{\varkappa l} t_{i\varkappa} a_{\varkappa l}^R \Psi_l(E), \tag{16}$$

da aber

$$\Psi_l(E) = \sum_j \tau_{lj} \Phi_j(E), \tag{16a}$$

so finden wir

$$\Phi_i(R) = \sum_{\varkappa l} \sum_j t_{i\varkappa} a_{\varkappa l}^R \tau_{lj} \Phi_j(E) = \sum_j A_{ij}^R \Phi_j(E), \tag{16b}$$

worin (A_{ik}^R) die durch (T) transformierte Matrix (a_{ik}^R), also eine orthogonale Matrix ist. Insbesondere ist[2]) $A_{ik}^R = A_{ki}^{R^{-1}}$.

6. Wir wollen nun im Anschluß an Teil I sowohl die Erhaltung der Transformationseigenschaft in der Form (16b) bei einer weiteren Störung, wie das Verschwinden des Ausdrucks (4) des Teil I nachweisen. Hierzu brauchten wir zwei Sätze, aus diesen ging alles hervor: die Gleichheit der Integrale

$$\int V'' \Phi_\varkappa^2 \, d\mathfrak{r}_1 \ldots d\mathfrak{r}_n = \int V'' \Phi_\lambda^2 \, d\mathfrak{r}_1 \ldots d\mathfrak{r}_n, \tag{17}$$

worin V'' eine in den Variablen $\mathfrak{r}_1 \ldots \mathfrak{r}_n$ [es ist $V''(R) = V''(E)$] symmetrische Störungsfunktion, $\Phi_\varkappa$ und Φ_λ zwei zu demselben Eigenwert gehörende Eigenfunktionen sind; außerdem das Verschwinden des Integrals

$$\int V'' \Phi_i \Phi_j \, d\mathfrak{r}_1 \ldots d\mathfrak{r}_n = 0, \tag{18}$$

[1]) Zu einer anderen äquivalenten Darstellung in (8) hätten wir andere b erhalten. Geht die andere Darstellung durch Transformation der Matrizen (8) mit T aus ihnen hervor, so gehen dazu gehörige b durch Substitution mit T aus den alten b hervor. Vgl. das folgende.

[2]) Vgl. z. B. A. Speiser, Theorie der Gruppen von endlicher Ordnung, S. 107, Satz 100.

worin Φ_i und Φ_j zwei beliebige, zu verschiedenen Gruppen gehörende Eigenfunktionen sind, oder zwar derselben Gruppe angehören, aber verschiedenen Zeilen der Matrix (A_{ik}^R) entsprechen: verschiedene Transformationseigenschaften haben.

(17) und (18) bedingen, daß die Entartung keines Eigenwertes aufgehoben wird und daß auch die Transformationseigenschaften der Φ in (16b) erhalten bleiben. Da man den Ausdruck (4), Teil I, als ein Integral (18) auffassen kann, worin V'' überall Null ist, nur in der Nähe der Stellen $\mathfrak{r} = \mathfrak{r}_1$, $\mathfrak{r} = \mathfrak{r}_2 \ldots \mathfrak{r} = \mathfrak{r}_n$ einen endlichen, und zwar sehr großen Wert annimmt, so bedingt (18) auch das Verschwinden jeglicher Interkombination zwischen den beiden Zuständen. Dies ist alles, was wir beweisen wollen.

Um (17) und (18) zu beweisen, zerlegen wir das Integrationsgebiet in $n!$ Teile. Im ersten Teil, den wir mit G_E bezeichnen, ist

$$\mathfrak{r}_1 < \mathfrak{r}_2 < \mathfrak{r}_3 < \cdots < \mathfrak{r}_n \tag{19}$$

(worin unter $<$ etwa der absolute Betrag zu verstehen ist), die anderen Teile G_R entstehen durch die Permutation R der $\mathfrak{r}$ in der Bedingung (19).

Für (17) haben wir

$$\left.\begin{aligned}
&\int V''(E)\,\Phi_\varkappa^2(E)\,d\mathfrak{r}_1 \ldots d\mathfrak{r}_n = \sum_R \int_{G_R} V''(E)\,\Phi_\varkappa^2(E)\,d\mathfrak{r}_1 \ldots d\mathfrak{r}_n \\
&= \sum_R \int_{G_E} V''(R^{-1})\,\Phi_\varkappa^2(R^{-1})\,d\mathfrak{r}_1 \ldots d\mathfrak{r}_n \\
&= \int_{G_E} V''(E) \sum_R \Phi_\varkappa^2(R)\,d\mathfrak{r}_1 \ldots d\mathfrak{r}_n \\
&= \int_{G_E} V''(E) \sum_R \Big(\sum_j A_{\varkappa j}^R\,\Phi_j(E)\Big)^2 d\mathfrak{r}_1 \ldots d\mathfrak{r}_n \\
&= \int_{G_E} V''(E) \sum_{Rjl} A_{\varkappa j}^R\,A_{\varkappa l}^R\,\Phi_j(E)\,\Phi_l(E)\,d\mathfrak{r}_1 \ldots d\mathfrak{r}_n
\end{aligned}\right\} \tag{17a}$$

Nun ist[1]).

$$\left.\begin{aligned}
&\sum_R A_{\varkappa j}^R\,A_{\varkappa l}^R = \sum_R A_{j\varkappa}^{R^{-1}}\,A_{\varkappa l}^R = \frac{h}{k}\,\delta_{jl}\,\delta_{\varkappa\varkappa} = \frac{h}{k}\,\delta_{jl} \\
&\int V''(E)\,\Phi_\varkappa^2(E)\,d\mathfrak{r}_1 \ldots d\mathfrak{r}_n = \frac{h}{k}\int_{G_E} V''(E) \sum_j \Phi_j^2(E)\,d\mathfrak{r}_1 \ldots d\mathfrak{r}_n
\end{aligned}\right\} \tag{17b}$$

[1]) J. Schur, Berl. Ber. 1905, S. 411.

also unabhängig von $\varkappa$, womit (17) bewiesen ist. Für

$$\left.\begin{aligned} &\int V''(E)\,\Phi_i(E)\,\Phi_j(E)\,d\mathfrak{r}_1 \ldots d\mathfrak{r}_n \\ &= \sum_R \int_{G_R} V''(E)\,\Phi_i(E)\,\Phi_j(E)\,d\mathfrak{r}_1 \ldots d\mathfrak{r}_n \\ &= \sum_R \int_{G_E} V''(E)\,\Phi_i(R)\,\Phi_j(R)\,d\mathfrak{r}_1 \ldots d\mathfrak{r}_n \end{aligned}\right\} \tag{18a}$$

Es ist nun

$$\sum_R \Phi_i(R)\,\Phi_j(R) = \sum_{R\varkappa\lambda} A^R_{i\varkappa}\,A^R_{j\lambda}\,\Phi_\varkappa(E)\,\Phi_\lambda(E) \tag{18b}$$

und[1])

$$\sum_R A^R_{i\varkappa}\,A^R_{j\lambda} = \sum_R A^{R^{-1}}_{\varkappa i}\,A^R_{j\lambda} = \frac{h}{k}\,\delta_{\varkappa\lambda}\,\delta_{ij}; \tag{18c}$$

da aber voraussetzungsgemäß $i \neq j$ ist, ist (18 c) $= 0$ und auch (18) $= 0$. Hiermit haben wir den Zerfall des Termsystems in im 2. Abschnitt beschriebene, untereinander nicht kombinierende Gruppen bewiesen.

7. Vom physikalischen Standpunkt aus könnte es zunächst scheinen, daß diese Untersuchung der Termsysteme ohne Belang ist, da ja praktisch nach Dirac[2]) nur das eine, unentartete Termsystem in Betracht kommt. Dieses System entspricht derjenigen Darstellung, bei der zu jeder Permutation der alternierenden Gruppe die Matrix (1) zu jeder in der Nebengruppe enthaltenen (— 1) gehört[2]). (In diesem Fall haben die Matrizen natürlich nur je ein Element, da ja keine Entartung vorliegt.)

Es zeigt sich aber, daß wegen der Eigenrotation des Elektrons auch den übrigen Gruppen eine Bedeutung zukommt, wie dies Heisenberg[3]) auch explizite am Helium gezeigt hat. (Zum Beispiel sind bei Li die normalen optischen Terme den entarteten Eigenfunktionen des Teil I zugeordnet.)

Auf alle Fälle ist zu erwarten, daß ein, der Ortho-Paraeinteilung der Terme beim Helium analoger, näherungsweiser Zerfall auch bei den anderen Elementen existiert.

Bei dem strengen Zerfall in mehrere Termsysteme ist folgendes bemerkenswert. Bei jeder symmetrischen Störung lassen sich die neuen Eigenwerte und Eigenfunktionen eines Termsystems aus den alten Eigenwerten und Eigenfunktionen desselben Termsystems allein berechnen, d. h. die Eigenfunktionen und Eigenwerte anderer Termsysteme gehen in die Störungsrechnung gar nicht ein.

Berlin, Institut f. theoret. Physik d. Techn. Hochschule, Nov. 1926.

---

[1]) J. Schur, Berl. Ber. 1905, S. 411.

[2]) L. c.

[3]) ZS. f. Phys. **38**, 411, 1926; **39**, 499, 1926.

Einige Folgerungen aus der Schrödingerschen Theorie für die Termstrukturen

E. P. Wigner

Zeitschrift für Physik *43*, 624–652 (1927)

Mit 2 Abbildungen. (Eingegangen am 5. Mai 1927.)

Es wird versucht, aus der Form der Schrödingerschen Differentialgleichung einige strukturelle Eigenschaften der Spektren abzuleiten. Es wird die Aufspaltung im elektrischen und magnetischen Feld, das Aufbauprinzip der Serienspektren und einiges Verwandte behandelt. Es ergibt sich — soweit das rotierende Elektron nicht in Betracht zu ziehen ist — Übereinstimmung mit der Erfahrung.

1. Die einfache Gestalt der Schrödingerschen Differentialgleichung gestattet die Anwendung einiger Methoden der Gruppen, genauer gesagt, der Darstellungstheorie. Diese Methoden haben den Vorteil, daß man mit ihrer Hilfe beinahe ganz ohne Rechnung Resultate erhalten kann, die nicht nur für das Einkörperproblem (Wasserstoffatom), sondern auch für beliebig komplizierte Systeme exakt gültig sind. Der Nachteil der Methode ist, daß sie keine Näherungsformeln abzuleiten gestattet. Es ist auf diese Weise möglich, einen großen Teil unserer qualitativen spektroskopischen Erfahrung zu erklären. Die Methode ist dabei so allgemein, daß sie vielfach gar nicht an die spezielle Gestalt der Differentialgleichung gebunden ist. So z. B. kann die Frage nach der Anzahl der Aufspaltungskomponenten im Magnetfeld behandelt werden, ohne daß man mehr voraussetzen müßte, als daß die Terme Eigenwerte einer linearen homogenen Differentialgleichung sind, in die physikalisch gleichwertige Dinge (z. B. Richtungen im Raume, solange keine äußeren Felder da sind) in gleicher Weise eingehen. Die relativen Intensitäten der Komponenten können noch bei schwachen Feldern berechnet werden, ohne daß man eine Annahme über die Differentialgleichung mit Feld machen müßte.

Viele hier abgeleitete Beziehungen sind schon bekannt. Andererseits kenne ich außer der Londonschen Ableitung[1]) des Kuhn-Reiche-Thomasschen Summensatzes keine streng gültige Beziehung, die hier nicht vorkäme, es sei denn, daß sie sich auf das Wasserstoffatom bezieht.

Um das Resultat vorwegzunehmen, sei schon hier gesagt, daß in allen Fällen Übereinstimmung mit den spektroskopischen Erfahrungen besteht [wie sie zuletzt von F. Hund[2]) so schön zusammengefaßt worden

[1]) ZS. f. Phys. **39**, 322, 1926.

[2]) F. Hund, Linienspektren und periodisches System der Elemente. Berlin 1927. Dortselbst auch weitere Literatur.

sind], soweit hierbei das „rotierende Elektron" keine Rolle spielt. Wenn man also aus dem Hundschen Buche alle Kapitel über das rotierende Elektron streichen könnte, so würde das übrigbleibende Material sich unseren Ergebnissen fügen.

In der vorliegenden Arbeit ist das rotierende Elektron zumeist nicht mit berücksichtigt und die Resultate sind dementsprechend vielfach nur für Singulettsysteme gültig.

Es wird allerdings bei vielen Ableitungen gar nicht angenommen, daß die Differentialgleichung nur die Lagenkoordinaten der Elektronen enthält; sie könnte noch beliebig viele, sich auf die Rotation der Elektronen beziehende Koordinaten enthalten. Da bei Nichtsingulettsystemen auch in diesen Fällen keine Übereinstimmung mit der Erfahrung erzielt werden konnte, so würde ich glauben, daß man ohne einen neuen Gedanken, lediglich durch Einführung neuer Koordinaten zur Beschreibung der Elektronenmagnete, nicht auskommen kann.

Im folgenden allgemeinen Teil werden die mathematischen Hilfsmittel bereitgestellt, die dann im speziellen Teile verwertet werden.

Allgemeiner Teil.

2. Im folgenden ist die Determinante aller Matrizen ungleich Null, alle Integrationen, wo nichts anderes angegeben ist, sind über den gesamten Bereich der Variablen zu erstrecken, $\delta_{jk} = 0, 1$, je nachdem $j \neq k$ oder $j = k$.

Ist $\psi(x_1, x_2, \ldots x_n)$ eine Funktion, R eine lineare Substitution in n Variablen

$$\left.\begin{aligned} x_1' &= \alpha_{11} x_1 + \alpha_{12} x_2 + \cdots \alpha_{1n} x_n, \\ x_2' &= \alpha_{21} x_1 + \alpha_{22} x_2 + \cdots \alpha_{2n} x_n, \\ &\cdots\cdots \\ x_n' &= \alpha_{n1} x_1 + \alpha_{n2} x_2 + \cdots \alpha_{nn} x_n, \end{aligned}\right\} \tag{1}$$

so verstehe ich unter der Funktion $\psi\big(R(x_1, x_2 \ldots x_n)\big)$ die Funktion, für die identisch in den $x_1, x_2 \ldots x_n$ gilt

$$\psi\big(R(x_1, x_2 \ldots x_n)\big) \equiv \psi(x_1', x_2' \ldots x_n'), \tag{2}$$

wobei für $x_1', x_2' \ldots x_n'$ die Werte aus (1) eingesetzt gedacht werden müssen. Wir werden für $\psi\big(R(x_1, x_2 \ldots x_n)\big)$ auch kurz $\psi(R)$ schreiben. Da E immer die Einheitssubstitution ist ($\alpha_{ik} = \delta_{ik}$), ist $\psi(E) = \psi$.

Es sei nun eine lineare homogene Differentialgleichung, ein Eigenwertproblem gegeben: $H(\psi, \varepsilon) = 0$. Hat die Differentialgleichung, bzw.

die Substitution R die Eigenschaft, daß mit jeder $\psi(x_1, x_2 \dots x_n)$ auch $\psi(R(x_1, x_2 \dots x_n))$ eine Lösung der Differentialgleichung bei demselben ε ist, so sagen wir „R ist in der Substitutionsgruppe der Differentialgleichung $H(\psi, \varepsilon)$ enthalten". Es ist nämlich klar, daß die Substitutionen, für die dies gilt, eine Gruppe bilden.

Zumeist ist ihre Substitutionsgruppe der Differentialgleichung direkt anzusehen. Zum Beispiel wenn ich in der Schrödingerschen Gleichung unter x_i, y_i, z_i die Lagenkoördinaten des i-ten Teilchens verstehe, unter R dagegen die Substitution

$$\left.\begin{array}{lll} x_1' = \alpha_1 x_1 + \beta_1 y_1 + \gamma_1 z_1, & y_1' = \alpha_2 x_1 + \beta_2 y_1 + \gamma_2 z_1, & z_1' = \alpha_3 x_1 + \beta_3 y_1 + \gamma_3 z_1, \\ \vdots & \vdots & \vdots \\ x_i' = \alpha_1 x_i + \beta_1 y_i + \gamma_1 z_i, & y_i' = \alpha_2 x_i + \beta_2 y_i + \gamma_2 z_i, & z_i' = \alpha_3 x_i + \beta_3 y_i + \gamma_3 z_i, \\ \vdots & \vdots & \vdots \end{array}\right\} \quad (A)$$

wobei die Matrix

$$\begin{pmatrix} \alpha_1 & \beta_1 & \gamma_1 \\ \alpha_2 & \beta_2 & \gamma_2 \\ \alpha_3 & \beta_3 & \gamma_3 \end{pmatrix}$$

orthogonal ist, so verifiziert man leicht, daß R in der Substitutionsgruppe der Schrödingerschen Gleichung enthalten ist. In einer vorangehenden Mitteilung [1]) wurde zur Ableitung des Heisenberg-Diracschen Termzerfalls nur die Tatsache benutzt, daß z. B. bei dem He die Substitution

$$x_1' = x_2, \; y_1' = y_2, \; z_1' = z_2, \; x_2' = x_1, \; y_2' = y_1, \; z_2' = z_1$$

in die Substitutionsgruppe gehört.

3. Ist mir umgekehrt die Substitutionsgruppe E, R_2, $R_3, \dots$ einer Differentialgleichung bekannt, so kann ich aus jeder Lösung sofort andere angeben, nämlich aus $\psi(x_1, x_2 \dots x_n)$, die weiteren $\psi(R_2(x_1, x_2 \dots x_n))$, $\psi(R_3(x_1, x_2 \dots x_n))$ usw., die alle zum Eigenwert ε gehören. Diese Funktionen werden im allgemeinen nicht alle linear unabhängig voneinander sein. $\psi_1, \psi_2, \dots \psi_l$ sei ein linear unabhängiges Aggregat, durch welches sich alle zu diesem Eigenwert ε gehörende Eigenfunktionen ausdrücken lassen. Da auch $\psi_\lambda(R_i)$ eine solche Eigenfunktion ist, ist

$$\psi_\lambda(R_i) = \sum_{\varkappa=1}^{l} a_{\lambda\varkappa}^{R_i} \psi_\varkappa(E). \tag{3}$$

[1]) E. Wigner, Über nichtkombinierende Terme in der neueren Quantentheorie, 2. Teil, ZS. f. Phys. **40**, 883, 1927.

Führen wir in diese Gleichung durch die Substitution R_j neue Variable ein, so erhalten wir

$$\psi_\lambda(R_i R_j) = \sum_{\varkappa=1}^{l} a_{\lambda\varkappa}^{R_i} \psi_\varkappa(R_j) = \sum_{\varkappa=1}^{l} \sum_{\mu=1}^{l} a_{\lambda\varkappa}^{R_i} a_{\varkappa\mu}^{R_j} \psi_\mu(E).$$

Andererseits ist

$$\psi_\lambda(R_i R_j) = \sum_{\mu=1}^{l} a_{\lambda\mu}^{R_i R_j} \psi_\mu(E).$$

Wegen der linearen Unabhängigkeit der ψ_μ folgt hieraus

$$a_{\lambda\mu}^{R_i R_j} = \sum_{\varkappa=1}^{l} a_{\lambda\varkappa}^{R_i} a_{\varkappa\mu}^{R_j}. \tag{4}$$

Mit anderen Worten: die Matrizen $\left(a_{ik}^{E}\right)$, $\left(a_{ik}^{R_2}\right)$, $\left(a_{ik}^{R_3}\right)\ldots$ usw. bilden eine zur Substitutionsgruppe E, R_2, $R_3\ldots$ der Differentialgleichung isomorphe Substitutionsgruppe, oder wie man das in der Gruppentheorie sagt, sie bilden eine Darstellung dieser Gruppe [1]).

Diese Darstellung nennen wir dann zur Kürze die Darstellung des betreffenden Terms ε.

Durch eine andere Wahl der linear unabhängigen $\psi_1, \psi_2 \ldots \psi_l$, wenn wir z. B.

$$\begin{aligned} \psi_1' &= \alpha_{11}\psi_1 + \alpha_{12}\psi_2 + \cdots \alpha_{1l}\psi_l \\ &\vdots \\ \psi_l' &= \alpha_{l1}\psi_1 + \alpha_{l2}\psi_2 + \cdots \alpha_{ll}\psi_l \end{aligned}$$

gewählt hätten, erfährt die Darstellung nur eine „Ähnlichkeitstransformation", indem aus den Matrizen $\left(a_{ik}^{E}\right)$, $\left(a_{ik}^{R_2}\right)$, $\left(a_{ik}^{R_3}\right)$ usw. die Matrizen $(\alpha_{ik})\left(\alpha_{ik}^{E}\right)(\alpha_{ik})^{-1}$, $(\alpha_{ik})\left(a_{ik}^{R_2}\right)(\alpha_{ik})^{-1}$, $(\alpha_{ik})\left(a_{ik}^{R_3}\right)(\alpha_{ik})^{-1}$ usw. entstehen. Solche Darstellungen, die sich nur durch eine Ähnlichkeitstransformation unterscheiden, wollen wir als nicht verschieden voneinander ansehen. In diesem Sinne hat also jeder Term eine Darstellung.

Umgekehrt entspricht natürlich einer Ähnlichkeitstransformation der Darstellung lediglich eine andere Wahl der linear unabhängigen $\psi_1, \psi_2 \ldots \psi_l$.

---

[1]) Siehe z. B. A. Speiser, Theorie der Gruppen von endlicher Ordnung. Berlin, 1923. Eine sehr schöne und leicht verständliche Darstellung findet man bei J. Schur, Neue Begründung der Theorie der Gruppencharaktere. Berl. Ber. 1905, S. 406 in den ersten 6 bis 8 Seiten. Die hier gebrauchten Sätze sind entweder hier abgeleitet oder wenigstens [(6), I, II und Punkt 11], ausführlich besprochen.

4. Nehmen wir insbesondere die $\psi_1, \psi_2 \ldots \psi_l$ orthogonal normiert, bzw. im komplexen Falle hermetisiert an, so daß $\int \psi_i \widetilde{\psi_j} = \delta_{ij}$, so folgt

$$\int \psi_\varkappa(R_i) \widetilde{\psi_\lambda(R_i)} = \sum_{\mu\nu} \int a^{R_i}_{\varkappa\mu} \psi_\mu(E) \widetilde{a^{R_i}_{\lambda\nu}} \widetilde{\psi_\nu(E)} = \sum_{\mu\nu} a^{R_i}_{\varkappa\mu} \widetilde{a^{R_i}_{\lambda\nu}} \delta_{\mu\nu}.$$

Führen wir nun auf der linken Seite die Transformation R_i^{-1} der Integrationsvariablen aus, so können wir dafür einfach $\int \psi_\varkappa(E) \widetilde{\psi_\lambda(E)} = \delta_{\varkappa\lambda}$ schreiben, und wir sehen unmittelbar, daß

$$\sum_\mu a^{R_i}_{\varkappa\mu} \widetilde{a^{R_i}_{\lambda\mu}} = \delta_{\varkappa\lambda}, \tag{5}$$

d. h. die Darstellung hat eine orthogonale, bzw. hermitische Form. Es folgt hieraus bekanntlich, da $\left(a^{R_i^{-1}}_{jk}\right) = \left(a^{R_i}_{jk}\right)^{-1}$ ist, daß

$$\widetilde{a^{R_i}_{jk}} = a^{R_i^{-1}}_{kj}. \tag{5a}$$

Eine Umkehrung dieses Satzes — die für „irreduzible" Darstellungen gilt — wird uns später begegnen.

5. Irreduzibel nennt man eine Darstellung, wenn es nicht möglich ist, die Matrix (t_{ik}) der Ähnlichkeitstransformation so zu wählen, daß alle Matrizen $(t_{ik})\,(a^E_{ik})\,(t_{ik})^{-1}$, $(t_{ik})\,(a^{R_2}_{ik})\,(t_{ik})^{-1}$, $(t_{ik})\,(a^{R_3}_{ik})\,(t_{ik})^{-1}$ usw. die Gestalt haben

$$\begin{pmatrix} & 0 & 0 \ldots 0 \\ & \vdots & \vdots \\ & 0 & 0 \ldots 0 \\ 0 \ldots 0 & & \\ 0 \ldots 0 & & \\ \vdots \quad \vdots & & \\ 0 \ldots 0 & & \end{pmatrix}$$

wobei die leer gelassenen quadratförmigen Stellen bei allen $(t_{ik})\,(a^R_{ik})\,(t_{ik})^{-1}$ beliebig besetzt sein können, während die ausgeschriebenen Nullen bei allen dieselben Stellen einnehmen. Es würden etwa bei allen in den ersten l'-Zeilen nur die ersten l'-Spalten besetzt sein, in den letzten $l - l'$-Zeilen nur die letzten $l - l'$-Spalten. In den ersten l'-Zeilen stehen in den letzten $l - l'$-Spalten, in den letzten $l - l'$-Zeilen in den ersten l'-Spalten lauter Nullen. Ist eine solche Bestimmung des (t_{ik}) nicht möglich, ist also die Darstellung des Terms eine irreduzible, so sagen wir, die Entartung (d. h. das Zugehören mehrerer Eigenfunktionen zu einem Eigenwert) ist normal. Sonst sprechen wir von einer zufälligen Entartung.

In dem Falle der zufälligen Entartung, wo also die Darstellung des Terms keine irreduzible ist, können wir diese durch eine zweckmäßig gewählte Ähnlichkeitstransformation $(t_{ik}) \ldots (t_{ik})^{-1}$ „ausreduzieren", d. h. alle Matrizen (a_{ik}^{E}), $(a_{ik}^{R_2})$, $(a_{ik}^{R_3}) \ldots$ in die Form bringen:

$$\left(\begin{array}{ccccccccc}
 & & 0 & \cdots & 0 & \cdots & \cdots & \cdots & 0 \\
 & J_1 & \vdots & & \vdots & & & & \vdots \\
 & & 0 & \cdots & 0 & \cdots & \cdots & \cdots & 0 \\
0 & \cdots & 0 & & 0 & \cdots & \cdots & \cdots & 0 \\
\vdots & & \vdots & J_2 & \vdots & & & & \vdots \\
0 & \cdots & 0 & & 0 & \cdots & \cdots & \cdots & 0 \\
\vdots & & \vdots & 0 & \cdots & 0 & \ddots & & \vdots \\
\vdots & & \vdots & \vdots & & \vdots & & 0 & \cdots 0 \\
\vdots & & \vdots & \vdots & & \vdots & & 0 & \\
\vdots & & \vdots & \vdots & & \vdots & & \vdots & J_s \\
0 & \cdots & 0 & 0 & \cdots & 0 & \cdots & 0 &
\end{array}\right) \tag{6}$$

wo an den quadratförmigen Stellen $J_1, J_2, \ldots J_s$ Matrizen irreduzibler Darstellungen, sonst lauter Nullen stehen. Führen wir nun mit Hilfe der t_{ik} neue Eigenfunktionen ein

$$\psi_i' = \sum_k t_{ik} \psi_k,$$

so zerfallen diese Eigenfunktionen in s-Gruppen. Die $\psi_i'(R)$ für $R = E_1, R_2, R_3 \ldots$ usw. der einen Gruppe lassen sich durch lineare Kombination der ψ' dieser selben Gruppe allein ausdrücken. Dies geht aus der Gestalt (6) der Darstellung hervor, die für die ψ' gilt. Daß der Eigenwert der einen dieser Gruppen mit dem Eigenwert der anderen Gruppen übereinstimmt, ist quasi zufällig, woher auch das Wort zufällige Entartung genommen wurde. Im entgegengesetzten Falle der normalen Entartung sind — bei beliebiger Wahl des ψ — alle linear unabhängigen Eigenfunktionen, die zu diesem Eigenwert gehören, unter den $\psi(E)$, $\psi(R_2)$, $\psi(R_3) \ldots$ usw. enthalten. Im Falle der zufälligen Entartung können wir auch sagen, daß die s-Terme normaler Entartung (mit den Darstellungseigenschaften $J_1, J_2, \ldots J_s$) zusammenfallen. Wir werden einen solchen Eigenwert s-fach zählen, indem wir jeder der s-Gruppen von ψ' einen Eigenwert zuordnen. In diesem Sinne ist also die Darstellung eines jeden Terms eine irreduzible. Wir wollen an dieser Nomenklatur festhalten.

6. Für die Matrixelemente einer irreduziblen Darstellung gelten zwei wichtige Sätze. Sind (a_{ik}) und (b_{ik}) zwei nicht ähnliche irreduzible Darstellungen derselben Gruppe, so gilt[1])

$$\sum_R a_{k\varkappa}^R b_{\lambda l}^{R^{-1}} = 0 \tag{II}$$

und

$$\sum_R a_{k\varkappa}^R a_{\lambda l}^{R^{-1}} = \delta_{kl}\delta_{\varkappa\lambda}c, \tag{I}$$

wobei die Summation über alle Elemente der Gruppe zu erstrecken ist und c eine Zahl ist, deren Größe nur von den Eigenschaften der Gruppe und der Darstellung abhängt, von k, l, $\varkappa$, λ dagegen unabhängig ist.

Kombinieren wir diese Gleichung mit (5a), indem wir annehmen daß die Form der Darstellung eine hermitische ist, so erhalten wir

$$\sum_R a_{k\varkappa}^R \widetilde{b_{l\lambda}^R} = 0 \tag{IIa}$$

und

$$\sum_R a_{k\varkappa}^R a_{l\lambda}^R = \delta_{kl}\delta_{\varkappa\lambda}c. \tag{Ia}$$

Für kontinuierliche Gruppen (Drehgruppe usw.) sind an Stelle der Summen Integrale zu setzen[2]).

$$\int_R a_{k\varkappa}^R \widetilde{b_{l\lambda}^R} = 0, \tag{IIb}$$

$$\int_R a_{k\varkappa}^R \widetilde{a_{l\lambda}^R} = \delta_{kl}\delta_{\varkappa\lambda}c. \tag{Ib}$$

Jede Gruppe, also auch die Substitutionsgruppe unserer Differentialgleichung, besitzt offenbar eine Darstellung — die „identische Darstellung" — bei der jedem R die Matrix (1) zugeordnet ist. Funktionen, die diese Darstellungseigenschaft haben, für die also gilt

$$f(R) = f(E)$$

nennen wir in Anlehnung an Dirac[3]) symmetrische Funktionen.

Haben wir nun zwei Eigenwerte mit den Darstellungseigenschaften (a_{ik}^R) bzw. (b_{ik}^R) (die voneinander verschieden sind), und den Eigenfunktionen $\psi_1, \psi_2, \ldots$ bzw. $\psi_1', \psi_2', \ldots$, die wir unter sich orthogonal-hermitisch annehmen wollen und ist die Funktion f symmetrisch, so gilt

$$\int \psi_k \widetilde{\psi_l'} f = A_{kl} = 0. \tag{7}$$

[1]) J. Schur, Berl. Ber. 1905, S. 411.

[2]) Vgl. J. Schur, Neue Anwendungen der Integralrechnung auf Probleme der Invariantentheorie. Berl. Ber. S. 199, 1924.

[3]) Proc. Roy. Soc. **112**, 661, 1926.

Es ist nämlich

$$A_{kl} = \int \psi_k(E)\,\overline{\psi_l'(E)}\,f(E) = \int \psi_k(R)\,\overline{\psi_l'(R)}\,f(R),$$

wie bei (5) ausgeführt ist. Wir haben nun

$$\begin{aligned} A_{kl} &= \int \psi_k(R)\,\overline{\psi_l'(R)}\,f(R) \\ &= \sum_{\varkappa\lambda} a_{k\varkappa}^{R}\,\widetilde{b_{l\lambda}^{R}} \int \psi_\varkappa(E)\,\overline{\psi_\lambda'(E)}\,f\,E = \sum_{\varkappa\lambda} a_{k\varkappa}^{R}\,\widetilde{b_{l\lambda}^{R}}\,A_{\varkappa\lambda}. \end{aligned}$$

Summieren wir diese Gleichung über alle R der Substitutionsgruppe der Differentialgleichung, so erhalten wir wegen (IIa) $A_{kl} = 0$.

7. Wir haben im Punkt 4 folgendes gesehen: Nehmen wir die Eigenfunktionen $\psi_1, \psi_2, \psi_3, \ldots$, die zu einem Eigenwert gehören, orthogonal-hermitisch an, so ist die zugehörige Form der Darstellung des Eigenwertes ebenfalls orthogonal-hermitisch. Es gilt nun für irreduzible Darstellungen auch die Umkehrung: Ist die Form der Darstellung eine hermitische, was durch eine Ähnlichkeitstransformation immer zu erreichen ist, ist also

$$\sum_{\mu} a_{\varkappa\mu}^{R}\,\widetilde{a_{\lambda\mu}^{R}} = \delta_{\varkappa\lambda}\ (R = E, R_2, R_3 \ldots) \tag{5}$$

oder, was damit gleichbedeutend ist,

$$a_{\varkappa\lambda}^{R} = \overline{a_{\lambda\varkappa}^{R^{-1}}}\quad (R = E, R_2, R_3 \ldots), \tag{5a}$$

so sind die Eigenfunktionen $\psi_1, \psi_2, \psi_3, \ldots$ auch orthogonal-hermitisch. Es ist sogar allgemeiner (f wieder eine symmetrische Funktion)

$$\int \psi_k(E)\,\overline{\psi_l(E)}\,f(E) = \delta_{kl}\,C.$$

Setzen wir nämlich

$$\int \psi_k(E)\,\overline{\psi_l(E)}\,f(E) = \int \psi_k(R)\,\overline{\psi_l(R)}\,f(R)$$

und

$$\int \psi_k(R)\,\overline{\psi_l(R)}\,f(R) = \sum_{\varkappa,\lambda} a_{k\varkappa}^{R}\,\widetilde{a_{l\lambda}^{R}} \int \psi_k(E)\,\overline{\psi_l(E)}\,f(E),$$

dies ergibt über R summiert wegen (Ia)

$$\int \psi_k(E)\,\overline{\psi_l(E)}\,f(E) = \delta_{kl}\,C, \tag{7a}$$

wo C eine von k, l unabhängige, von der Funktion f dagegen abhängige Konstante ist.

Man verifiziert auch leicht folgende Verallgemeinerung: Gehören zu zwei Termen derselben Darstellungseigenschaft die Eigenfunktionen $\psi_1, \psi_2, \psi_3, \ldots$ bzw. $\psi_1', \psi_2' \ldots$, wobei diese so gewählt sind, daß auch die Form der Darstellung dieselbe ist, so gilt ebenso wie vorher

$$\int \psi_k\,\widetilde{\psi_l'}\,f = C\,\delta_{kl}. \tag{7b}$$

8. Wir haben nun alles in der Hand, um die Störungstheorie von Schrödinger[1]) zu behandeln. Aus seinen Formeln in § 2 und den Formeln (7), (7 a), (7 b) geht ohne weiteres hervor, daß eine normale Entartung, solange die Störung eine symmetrische Funktion ist, also die Substitutionsgruppe der Differentialgleichung sich nicht ändert — niemals aufgehoben wird und der Term seine Darstellung beibehält. Halten wir bei zufälligen Entartungen an der Nomenklatur des Punktes 5 fest, so ändert sich hieran auch in diesem Falle nichts: Es rücken höchsten die zufällig zusammengefallenen Terme etwas auseinander, oder es rücken einige bisher getrennte zusammen, um sich eventuell bei einer weiteren Störung wieder zu trennen.

Dieses Bild ändert sich ganz, wenn die Störung nicht „symmetrisch" ist, die neue Differentialgleichung also nicht mehr die Substitutionsgruppe der alten hat. Hier interessiert uns namentlich der Fall, daß die Substitutionsgruppe der gestörten Differentialgleichung eine Untergruppe der ungestörten ist. Dies ist z. B. der Fall, wenn man eine elektrische oder magnetische Kraft auf das System einwirken läßt.

Die Behandlung dieses Falles ist auch sehr einfach. Die Substitutionsgruppe der alten Differentialgleichung bezeichne ich mit $\mathfrak{R}$, die der neuen mit $\mathfrak{R}'$, die Darstellung des betrachteten Terms (a_{ik}^{R}) sei eine irreduzible Darstellung von $\mathfrak{R}$. Sie ist selbstverständlich auch eine Darstellung ihrer Untergruppe $\mathfrak{R}'$, aber als solche wahrscheinlich nicht irreduzibel. Wir können sie aber mit Hilfe einer Ähnlichkeitstransformation $(t_{ik})\,(a_{ik}^{R})\,(t_{ik})^{-1}$ ausreduzieren. Sind E, R_2', R_3', ... die Elemente von $\mathfrak{R}'$ (sie sind unter den E, R_2, R_3, ... enthalten), so haben die $(t_{ik})\,(a_{ik}^{R'})\,(t_{ik})^{-1}$ die Form (Punkt 5)

$$\left(\begin{array}{cccccccc}
 & 0 & \cdots & \cdots & \cdots & \cdots & \cdots & 0 \\
J_1' & \vdots & & & & & & \vdots \\
 & 0 & \cdots & \cdots & \cdots & \cdots & \cdots & 0 \\
0 \cdots 0 & & 0 & \cdots & \cdots & \cdots & \cdots & 0 \\
\vdots \quad \vdots & J_2' & \vdots & & & & & \vdots \\
 & & 0 & \cdots & \cdots & \cdots & \cdots & 0 \\
\vdots \quad \vdots & 0 \cdots 0 & \vdots & \vdots & \ddots & & & \vdots \\
\vdots \quad \vdots & \vdots & \vdots & \vdots & & \ddots & & \vdots \\
\vdots \quad \vdots & \vdots & \vdots & \vdots & & & 0 & 0 \cdots 0 \\
\vdots \quad \vdots & \vdots & \vdots & \vdots & & 0 & & \\
\vdots \quad \vdots & \vdots & \vdots & \vdots & & \vdots & & J_s' \\
0 \quad 0 & 0 \cdots & 0 & \cdots & \cdots & 0 & &
\end{array}\right) \tag{8}$$

[1]) Ann. d. Phys. **80**, 437, 440—457, 1926.

für alle Elemente R' der Gruppe $\mathfrak{R}'$, wo J_1', J_2', ... J_s' irreduzible Darstellungen von $\mathfrak{R}'$ sind. Bei den anderen Elementen von $\mathfrak{R}$ werden natürlich andere Formen auftreten und insbesondere nicht überall, wo hier, lauter Nullen stehen.

Wir werden dieser Form der Matrix entsprechend auch neue Eigenfunktionen einführen $\left(\psi_i' = \sum_k t_{ik}\psi_k\right)$, für die die Darstellung die Form (8) hat. Wenn wir dann der ursprünglichen Differentialgleichung auch nur die Gruppe $\mathfrak{R}'$ als Substitutionsgruppe zuschreiben, ist der betrachtete Term s-fach „zufällig" entartet und enthält eigentlich einen Term der Darstellungseigenschaft J_1', einen weiteren der Darstellungseigenschaft J_2' usw. bis J_s'. Verändern wir etwas die Differentialgleichung, wobei nun die jetzt betrachtete Substitutionsgruppe nicht mehr verändert wird, so wird die zufällige Entartung im allgemeinen aufgehoben, und es entstehen die s-Terme mit den Darstellungseigenschaften J_1', J_2', ... J_s' aus den betrachteten.

Um die Aufspaltung eines Terms mit der Darstellung (a_{ik}^R) bei einer Störung, die die Substitutionsgruppe $\mathfrak{R}$ der Differentialgleichung ändert, zu erhalten, muß man also (a_{ik}^R) als Darstellung der neuen Substitutionsgruppe $\mathfrak{R}'$ ausreduzieren. Hat sie dann die Gestalt (8), so spaltet der ursprüngliche Term in s Terme mit den Darstellungseigenschaften J_1', J_2', ... J_s' auf.

Natürlich hätten wir all dies auch durch direkte Rechnung mit Hilfe der Formeln von Schrödinger erhalten können.

9. Im Punkte 5 haben wir jedem Term eine irreduzible Darstellung der Gruppe $\mathfrak{R}$ zugeordnet. Es erhebt sich die Frage, ob auch umgekehrt alle irreduziblen Darstellungen der Gruppe $\mathfrak{R}$ gebraucht werden.

Es sei $H\,[\psi(x_1, x_2, \ldots x_n), \varepsilon]$ eine Differentialgleichung und $\left(a_{\varkappa\lambda}^R\right)$ eine Darstellung ihrer Substitutionsgruppe. Existieren nun beliebige Funktionen $f_1(x_1, x_2, \ldots x_n)$, $f_2(x_1, x_2, \ldots x_n)$, ... $f_l(x_1, x_2, \ldots x_n)$, die diese Darstellungseigenschaft haben, daß also

$$f_\varkappa(R) = \sum_\lambda a_{\varkappa\lambda}^R f_\lambda(E),$$

so gibt es (im allgemeinen unendlich viele) Eigenwerte mit der Darstellung $(a_{\varkappa\lambda}^R)$. (Im entgegengesetzten Falle gibt es natürlich keinen, da ja sonst schon seine Eigenfunktionen ψ_1, ψ_2, ... ψ_l die für f_1, f_2, ... f_l verlangte Eigenschaft hätten, diese also nicht unerfüllbar wäre.)

Um dies einzusehen, muß man nur eine der Funktionen $f_\varkappa$ nach dem vollständigen System von Eigenfunktionen der Differentialgleichung $H[\psi(x_1, x_2, \ldots x_n), \varepsilon]$ entwickeln. Die Entwicklungskoeffizienten nach

Eigenfunktionen, deren Darstellung von $(a^R_{\varkappa\lambda})$ verschieden ist, verschwinden hierbei wegen (7). Hieraus folgt unser Satz.

Die Voraussetzungen dieses Satzes sind — wie wir sehen werden — im Falle von wenigstens vier Teilchen für alle irreduziblen Darstellungen der in Betracht kommenden Substitutionsgruppen erfüllt.

10. Die Substitutionsgruppen, die für uns in Betracht kommen, sind im wesentlichen Verknüpfungen von symmetrischen Gruppen, zwei- oder dreidimensionalen Dreh- oder Drehspiegelungsgruppen, eventuell noch mit einer Spiegelung. Die Darstellungen dieser Gruppen einzeln sind bekannt[1]), wir brauchen nur einen Satz für die Verknüpfung.

Unter der Dimension einer Darstellung versteht man die Anzahl Zeilen oder Spalten in ihren Matrizen, diese ist also gleich der Anzahl der linear unabhängigen Eigenfunktionen, die zu diesem Term gehören.

Habe ich eine Gruppe $\mathfrak{A}$ vom Grade $\mathfrak{a}$ mit den Elementen $E, A_2, A_3, \ldots A_\mathfrak{a}$ und eine andere $\mathfrak{B}$ vom Grade $\mathfrak{b}$ mit den Elementen $E, B_2, B_3, \ldots B_\mathfrak{b}$ so, daß sie außer E kein gemeinsames Element haben, so sei das direkte Produkt $\mathfrak{C} = \mathfrak{A}.\mathfrak{B} = \mathfrak{B}.\mathfrak{A}$ der Gruppen $\mathfrak{A}$ und $\mathfrak{B}$ vom Grade $\mathfrak{a}\mathfrak{b}$ und enthalte die Elemente $E, A_2, A_3, \ldots A_\mathfrak{a}, B_2, B_2 A_2, B_2 A_3, \ldots B_2 A_\mathfrak{a}, \ldots B_\mathfrak{b}, B_\mathfrak{b} A_2, B_\mathfrak{b} A_3, \ldots B_\mathfrak{b} A_\mathfrak{a}$. Alle A sind mit allen B vertauschbar, $A_n B_m = B_m A_n$. Es ist klar, daß $\mathfrak{C}$ die Gruppeneigenschaften besitzt.

Sind nun die α irreduziblen Darstellungen der Gruppe $\mathfrak{A}$ die Matritzen

$$\left({}^1a^{A_n}_{ik}\right), \left({}^2a^{A_n}_{ik}\right), \ldots \left({}^\alpha a^{A_n}_{ik}\right)$$

mit den Dimensionen $u_1, u_2, \ldots u_\alpha$; die β irreduziblen Darstellungen der Gruppe $\mathfrak{B}$ die Matrizen

$$\left({}^1b^{B_m}_{ik}\right), \left({}^2b^{B_m}_{ik}\right), \ldots \left({}^\beta b^{B_m}_{ik}\right)$$

mit den Dimensionen $v_1, v_2, \ldots v_\beta$, so hat die Gruppe $\mathfrak{C}$ genau $\alpha\beta$ irreduzible Darstellungen, und zwar sind diese

$$\begin{array}{l}
\left({}^{1,1}c^{A_n B_m}_{i_1, i_2;\, k_1, k_2}\right), \left({}^{1,2}c^{A_n B_m}_{i_1, i_2;\, k_1, k_2}\right), \ldots \left({}^{1,\beta}c^{A_n B_m}_{i_1, i_2;\, k_1, k_2}\right) \\
\left({}^{2,1}c^{A_n B_m}_{i_1, i_2;\, k_1, k_2}\right), \left({}^{2,2}c^{A_n B_m}_{i_1, i_2;\, k_1, k_2}\right), \ldots \left({}^{2,\beta}c^{A_n B_m}_{i_1, i_2;\, k_1, k_2}\right) \\
\qquad\vdots \qquad\qquad\qquad\qquad \vdots \\
\left({}^{\alpha,1}c^{A_n B_m}_{i_1, i_2;\, k_1, k_2}\right), \left({}^{\alpha,2}c^{A_n B_m}_{i_1, i_2;\, k_1, k_2}\right), \ldots \left({}^{\alpha,\beta}c^{A_n B_m}_{i_1, i_2;\, k_1, k_2}\right)
\end{array} \tag{9}$$

dabei ist

$${}^{\nu,\mu}c^{A_n B_m}_{i_1, i_2;\, k_1, k_2} = {}^\nu a^{A_n}_{i_1, k_1} \cdot {}^\mu b^{B_m}_{i_2, k_2} \tag{9a}$$

[1]) Vgl. A. Speiser, Theorie der Gruppen usw. Für die symmetrische Gruppe insbesondere J. Schur, Berl. Ber. 1908, S. 664. Für die Drehgruppen: J. Schur, Neue Anwendungen der Integralrechnung auf Probleme der Invariantentheorie I., II. und III. Berl. Ber. 1924, S. 189, 297 und 346. H. Weyl, Zur Theorie der Darstellung der einfachen kontinuierlichen Gruppen. Berl. Ber. 1924, S. 338. Für den Hinweis auf diese Arbeiten bin ich wiederum Herrn J. v. Neumann zu Danke verpflichtet.

mit den Dimensionen $u_1 v_1$, $u_1 v_2$, ... $u_1 v_\beta$, $u_2 v_1$, $u_2 v_2$, ... $u_2 v_\beta$, ... $u_\alpha v_1$, $u_\alpha v_2$, ... $u_\alpha v_\beta$. i_1 und i_2 beziehen sich auf die Zeilen, k_1 und k_2 beziehen sich auf die Spalten.

Man beweist zunächst leicht, daß diese wirklich Darstellungen von $\mathfrak{C}$ sind, dann mit Hilfe der Gleichung (I) von J. Schur[1]), daß alle Funktionen der y

$$\sum_{B_m = E}^{B_\mathfrak{b}} \sum_{A_n = E}^{A_\mathfrak{a}} {}^{\nu,\mu}c^{A_n\, B_m}_{i_1,\, i_2;\; k_1\, k_2}\, y_{n,\, m}$$

linear unabhängig sind, woraus die Irreduzibilität folgt. Die Anzahl der Klassen konjugierter Elemente von $\mathfrak{A}$ ist α, von $\mathfrak{B}$ ist β[2]). Die von $\mathfrak{C}$ bestimmt sich zu $\alpha\beta$. Dies muß auch die Anzahl der irreduziblen Darstellungen sein, folglich existieren außer diesen keine weiteren.

Wir werden im folgenden die Terme nach ihren Darstellungen unterscheiden und bezeichnen. Unser (trivialer) Satz sagt also folgendes aus: besteht die Substitutionsgruppe einer Differentialgleichung aus dem direkten Produkt mehrerer Substitutionsgruppen (die eine Gruppe sei z. B. die Vertauschung der Elektronen untereinander, die andere etwa die der Wasserstoffkerne), läßt also die Differentialgleichung zwei vertauschbare Substitutionsgruppen zu, so definiert jedes Paar von irreduziblen Darstellungen der beiden Gruppen eine irreduzible Darstellung der ganzen Gruppe. Diese Darstellung, sowie die Terme mit diesen Darstellungen, benennt man dann zweckmäßig so, daß man die Zeichen beider Darstellungen nebeneinander fügt.

Aus den Formeln (9), (9 a) liest man nun auch leicht ab — da alle $({}^{\nu}a^{E}_{ik})$ sowie $({}^{\mu}b^{E}_{ik})$ Einheitsmatrizen sind —, daß die ψ sich bei einer Substitution der einen Untergruppe so transformieren, wie wenn diese Untergruppe allein da wäre.

Spezieller Teil.

11. Wir werden jetzt die Substitutionsgruppe unseres Eigenwertproblems aufsuchen.

Es gehören folgende Substitutionen in diese Gruppe:

I. Auf alle Fälle das Vertauschen aller Koordinaten aller Elektronen.

II. Ebenfalls auf alle Fälle das Vertauschen anderer gleichwertiger Teilchen.

---

[1]) l. c.

[2]) Ebenda, Satz XIII.

Diese beiden Gruppen sind Permutationsgruppen oder, wie man sie auch nennt, symmetrische Gruppen. Ihre irreduziblen Darstellungen sind bekannt[1]).

III a[1]). Im feldlosen Falle ist der Raum isotrop. Ich kann also das ganze System von einem anderen, zum ursprünglichen beliebig verdrehten, Koordinatensystem betrachten und kann dieses Koordinatensystem auch noch spiegeln. Diese Gruppe ist also die dreidimensionale Drehspiegelungsgruppe, ihre Substitutionen stehen in (A). Ihre irreduziblen Darstellungen sind ebenfalls bekannt; es existieren zwei Darstellungen von der Dimension $2l+1$ (wo l eine nicht negative ganze Zahl ist), die in den Matrizen, die reinen Drehungen zugeordnet sind, übereinstimmen. Der reinen Spiegelung durch den Ursprungspunkt entspricht bei der einen Darstellung eine Diagonalmatrix mit lauter Gliedern -1^l, wir nennen diese die „normale" Darstellung, bei der zweiten eine Diagonalmatrix mit lauter Gliedern -1^{l+1}, wir nennen diese die „gespiegelte" Darstellung[2]).

III b. Im Falle des elektrischen Feldes wird die Isotropie des Raumes aufgehoben. Das Koordinatensystem kann bloß um die Achse des elektrischen Feldes gedreht werden und außerdem in einer zum Felde parallelen Ebene gespiegelt. Es bleibt uns also nur eine zweidimensionale Drehspiegelungsgruppe. Sie hat zwei Darstellungen der Dimension 1, unendlich viele der Dimension 2. Der reinen Verdrehung um einen Winkel α entsprechen dabei die Matrizen

$$(1)^{3}),\quad (1)^{4}),\quad \begin{pmatrix} e^{i\alpha} & 0 \\ 0 & e^{-i\alpha} \end{pmatrix},\ \begin{pmatrix} e^{2i\alpha} & 0 \\ 0 & e^{-2i\alpha} \end{pmatrix},\ \begin{pmatrix} e^{3i\alpha} & 0 \\ 0 & e^{-3i\alpha} \end{pmatrix}\cdots,$$

der Spiegelung in der Ebene durch die Feldachse

$$(1)^{3}),\ (-1)^{4}),\quad \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix},\quad \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix},\quad \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}\cdots$$

III c. Im Falle des magnetischen Feldes kann man das Koordinatensystem um die Feldachse drehen und in der Ebene senkrecht zu dieser spiegeln. Dies ist ein direktes Produkt einer zweidimensionalen Drehgruppe mit einer Spiegelung.

In den irreduziblen Darstellungen der zweidimensionalen Drehgruppe entsprechen einer Verdrehung um den Winkel α die Matrizen

$$(1),\ (e^{i\alpha}),\ (e^{-i\alpha}),\ (e^{2i\alpha});\ (e^{-2i\alpha})\ldots$$

[1]) Vgl. J. Schur, l. c. H. Weyl, l. c.

[2]) Diese Unterscheidung wird sich als zweckmäßig erweisen.

[3]) Diese Darstellung werden wir normal nennen.

[4]) Diese Darstellung werden wir gespiegelt nennen.

In den beiden Darstellungen der Spiegelung entsprechen einer solchen die Matrizen

$$(1),\ (-1).$$

In den Darstellungen der Gruppe III c entsprechen also (Punkt 10) einer Drehung um den Winkel α

$$(1),\ (1),\ (e^{i\alpha}),\ (e^{i\alpha}),\ (e^{-i\alpha}),\ (e^{-i\alpha}),\ (e^{2i\alpha}),\ (e^{2i\alpha}),\ (e^{-2i\alpha})\ldots,$$

der Spiegelung in der zur Achse des Feldes senkrechten Ebene die Matrizen

$$(1),\ (-1),\ (1),\ (-1),\ (1),\ (-1),\ (1),\ (-1),\ (1)\ldots$$

Der Fall, daß beide Felder vorhanden sind, wird sich als nicht mehr wesentlich erweisen.

Da nun die Gruppen I und II sowohl untereinander, wie mit jeder der Gruppen III a, III b, III c vertauschbar sind, können wir die Substitutionsgruppe der Differentialgleichung erhalten, indem wir das direkte Produkt der Gruppen I und II mit derjenigen Gruppe III bilden, die in Frage kommt.

12. Zunächst betrachten wir den Fall gekreuzter elektrischer und magnetischer Felder, um die Wirkung der Gruppen III ganz auszuschließen.

Wäre nur die Gruppe I vorhanden, so würde im Falle von n Elektronen jeder Aufteilung der Zahl n in ganzzahlige Summanden (wobei jeder Summand größer oder gleich dem vorangehenden sein muß) eine Darstellung entsprechen. Sind außerdem n_1 Teilchen der Sorte (1) (etwa Protonen), n_2 Teilchen der Sorte (2) usw. da, so muß man im Sinne des Punktes 10 jede Aufteilung der Zahl n mit jeder Aufteilung der Zahl n_1 und jeder Aufteilung der Zahl n_2 usw. kombinieren, um alle irreduziblen Darstellungen zu erhalten. Wegen 9. werden unendlich viele Eigenwerte jeder Darstellungseigenschaft existieren. Daß keine Interkombinationen zwischen Termen mit verschiedenen Darstellungen vorkommen, ist einfach eine Folge von (7).

Es werden aber nicht alle Eigenwerte in der Natur als Terme realisiert sein. Vielmehr ist z. B. bei Elektronen am einfachsten zu fordern, daß bei adiabatischem Verschwindenlassen der Wechselwirkung zwischen den Elektronen nur solche Zustände entstehen sollen, die höchstens zwei Elektronen auf einer Bahn enthalten. Dies sind diejenigen Termsysteme, bei denen die Zahl n in lauter 1 und 2 zerlegt wird [1]).

[1]) W. Heisenberg (ZS. f. Phys. **41**, 239, 1927) leitet diese Termsysteme ab, indem er davon ausgeht, daß die Wirkung der Elektronenrotation durch neue

Wir entnehmen nun mit Heisenberg[1]) der Erfahrung, daß ein Term bei einem Atom zu einer Multiplizität gehört, die um 1 größer ist als die Anzahl der 1 in der Zerlegung von n in Summanden.

Bei H	$1 = 1$	Dublett
„ He	$2 = 2$	Singulett
„ „	$2 = 1 + 1$	Triplett
„ Li	$3 = 1 + 2$	Dublett
„ „	$3 = 1 + 1 + 1$	Quartett
„ Be	$4 = 2 + 2$	Singulett
„ „	$4 = 1 + 1 + 2$	Triplett
„ „	$4 = 1 + 1 + 1 + 1$	Quintett

usw.

Wie die Auswahl der in der Natur vorkommenden Termsysteme bei Protonen usw. ist, kann vorläufig nur durch Vergleich mit der Erfahrung entschieden werden, wie insbesondere in der Diskussion dieses Themas bei F. Hund[2]) auseinandergesetzt ist[3]).

13. Wir gehen nun zu dem Falle über, wo auf unser System keine Kraft wirkt. Dann sind die Darstellungen des III a mit den Darstellungen I und II zu kombinieren.

Wir müssen nun die Darstellungen der dreidimensionalen Drehgruppe etwas genauer betrachten. Dazu führen wir die Parameterdarstellung von Euler ein. Die Verdrehung, die den Bogen ▬ in den

---

Eigenfunktionen, und zwar entsprechend den zwei Einstellungsmöglichkeiten des Elektrons, zwei Eigenfunktionen beschrieben wird. Dann sucht er diejenigen Termsysteme aus, die durch zwei Funktionen antisymmetrisiert werden können. Diese Vorschrift führt genau zu den im Text angegebenen Termsystemen. Ich kann davon absehen näheres hierüber mitzuteilen, da Herr F. Hund unabhängig von mir nach einer etwas anderen Methode zu demselben Ergebnis gekommen ist und seine Resultate bald in dieser Zeitschrift erscheinen.

[1]) l. c.

[2]) ZS. f. Phys. **42**, 93, 1927.

[3]) Ich möchte an dieser Stelle noch eine Bemerkung zur Frage der Bose-Einsteinschen Statistik machen, soweit sie die Gasentartung betrifft. Ob in einem Gase die Bose-Einsteinsche oder die Fermi-Diracsche Theorie zu Recht besteht, hängt lediglich davon ab (W. Pauli jr., ZS. f. Phys. **41**, 81, 1927; siehe auch R. H. Fowler, Proc. Roy. Soc. (A) **113**, 432, 1926), ob bei der Vertauschung zweier Atome bzw. Moleküle die Eigenfunktion ungeändert bleibt oder ihr Vorzeichen umkehrt. Dies läßt sich mit großer Wahrscheinlichkeit für Moleküle, die aus zwei gleichen Atomen bestehen, angeben. Wie immer nämlich das allgemeine Gesetz sein mag, das den Vorzeichenwechsel bei dem Austausch von Elektronen und etwa von Stickstoffkernen regelt, wenn man zwei N_2 miteinander vertauscht, kann sich am Ende das Vorzeichen nicht geändert haben, weil ja jede Operation zweimal ausgeführt worden ist. In solchen Fällen wenigstens besteht also die Einsteinsche Theorie der Gasentartung höchstwahrscheinlich zu Recht. Natürlich gilt dies nicht für Elektronen.

Bogen —·—·— überführt, kennzeichnen wir mit den drei Eulerschen Winkeln α, β, γ. Sie ist aus einer Drehung um die Z-Achse mit dem Winkel γ, um die X-Achse mit β und aus noch einer Drehung um die Z-Achse mit dem Winkel α zusammengesetzt. Das Matrizenelement, das in der $2l+1$-dimensionalen Darstellung in der j-ten Zeile und der k-ten Spalte steht, bezeichnen wir mit D^l_{jk} (j und k laufen von $-l$ bis $+l$). Wir betrachten diejenige Form der Darstellung, die als Darstellung der zweidimensionalen Drehgruppe um die Z-Achse ausreduziert und außerdem hermitisch ist. D^l_{jk} hat die Form

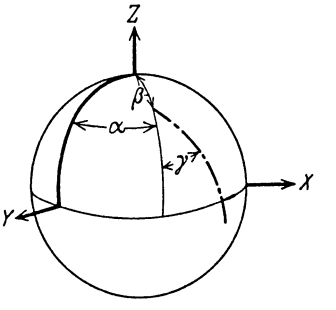


Fig. 1.

$$D^l_{jk}(\alpha, \beta, \gamma) = e^{ij\alpha} d^l_{jk}(\beta)\, e^{ik\gamma}, \tag{10}$$

$d^l_{jk}(\beta)$ ist reell, und es gilt

$$\begin{aligned} d^l_{jk}(\beta) &= (-1)^{j+k} d^l_{jk}(-\beta) = (-1)^{j+k} d^l_{kj}(\beta) \\ &= d^l_{kj}(-\beta) = (-1)^{j+k} d_{-j,-k}(\beta). \end{aligned} \tag{11}$$

Für später merken wir uns die Rekursionsformeln an:

$$\begin{aligned} &D^{l+1}_{jk} \frac{\sqrt{(l+1)^2 - j^2}\sqrt{(l+1)^2 - k^2}}{(2l+1)(l+1)} - \\ &\qquad - D^l_{jk} \cos\beta + D^{l-1}_{jk} \frac{\sqrt{l^2 - j^2}\sqrt{l^2 - k^2}}{(2l+1)\,l} = 0, \end{aligned} \tag{12}$$

$$\begin{aligned} &D^{l+1}_{j-1,k} \frac{\sqrt{l-j+1}\sqrt{l-j+2}\sqrt{(l+1)^2 - k^2}}{(2l+1)(l+1)} - \\ &- D^l_{jk} \sin\beta\, e^{-i\alpha} - D^{l-1}_{j-1,k} \frac{\sqrt{l+j}\sqrt{l+j-1}\sqrt{l^2 - k^2}}{(2l+1)\,l} = 0, \end{aligned} \tag{12 a}$$

$$\begin{aligned} &D^{l+1}_{j+1,k} \frac{\sqrt{l+j+1}\sqrt{l+j+2}\sqrt{(l+1)^2 - k^2}}{(2l+1)(l+1)} + \\ &+ D^l_{jk} \sin\beta\, e^{i\alpha} - D^{l-1}_{j+1,k} \frac{\sqrt{l-j-1}\sqrt{l-j}\sqrt{l^2 - k^2}}{(2l+1)\,l} = 0. \end{aligned} \tag{12 b}$$

Durch (11) gehen aus ihnen weitere hervor.

Es seien noch die ersten zwei Darstellungen für reine Drehung explizite hergeschrieben. Die „normalen" und „gespiegelten" Darstellungen der Drehspiegelungsgruppe ergeben sich daraus gemäß Punkt 11.

$$(1),\quad \begin{pmatrix} e^{-i\alpha}\frac{1+\cos\beta}{2}e^{-i\gamma}, & e^{-i\alpha}\frac{\sin\beta}{\sqrt{2}}, & e^{-i\alpha}\frac{1-\cos\beta}{2}e^{i\gamma}, \\ -\frac{\sin\beta}{\sqrt{2}}e^{-i\gamma}, & \cos\beta, & \frac{\sin\beta}{\sqrt{2}}e^{i\gamma}, \\ e^{i\alpha}\frac{1-\cos\beta}{2}e^{-i\gamma}, & e^{i\alpha}\frac{\sin\beta}{\sqrt{2}}, & e^{i\alpha}\frac{1+\cos\beta}{2}e^{i\gamma}. \end{pmatrix}$$

14. Ich will nun $2l+1$ Funktionen finden:

$$f_{-l}(x_1, y_1, z_1 \dots x_n, y_n, z_n),\quad f_{-l+1}(x_1, y_1, z_1 \dots x_n, y_n, z_n) \dots,$$
$$f_l(x_1, y_1, z_1 \dots x_n, y_n, z_n),$$

die sich bei der Substitution (A) nach Maßgabe der $2l+1$-dimensionalen irreduziblen Darstellung der dreidimensionalen Drehgruppe transformieren, für die also gilt

$$f_\varkappa(x_1', y_1', z_1' \dots x_n', y_n', z_n') = \sum_\lambda D^l_{\varkappa\lambda} f_\lambda(x_1, y_1, z_1 \dots x_n, y_n, z_n), \quad (13)$$

worin für $x_1', y_1', z_1' \dots x_n', y_n', z_n'$ die Werte aus (A) zu entnehmen sind.

Ich betrachte die Funktionen an der Überfläche $x_1 = y_1 = x_2 = 0$. An dieser Überfläche sind die Funktionen durch die gegenseitige Konfiguration der Teilchen bestimmt, während die Lage der ganzen Konfiguration im Raume vorgeschrieben ist, ich nenne sie die Lage B. Für die Koordinaten, die die gegenseitige Konfiguration beschreiben, schreibe ich g. Die Werte der Funktionen f_{-l} bis f_l an dieser Überfläche bezeichne ich mit $\chi_{-l}(g)$, $\chi_{-l+1}(g)$, $\chi_{-l+2}(g) \dots \chi_l(g)$. Die Funktion $f_\varkappa(x_1, y_2, z_1 \dots x_n, y_n, z_n)$ ist dann wegen (13)

$$f_\varkappa(x_1, y_1, z_1 \dots x_n, y_n, z_n) = \sum_\lambda D^l_{\varkappa\lambda}(\alpha, \beta, \gamma)\chi_\lambda(g), \quad (13\,\mathrm{a})$$

wo α das Azimut des ersten Elektrons, β sein Polabstand und γ der Winkel zwischen den Ebenen Zr_1 und $r_1 r_2$ ist[1]). Umgekehrt bestimme ich die $\chi_\lambda(g)$ ganz beliebig, so haben die in (13a) definierten Funktionen $f_\varkappa$ die verlangte Transformationseigenschaft (13).

Sollen die $f_\varkappa$ die normale Darstellungseigenschaft der Drehspiegelungsgruppe haben, so muß

$$f_\varkappa(-x_1, -y_1, -z_1 \cdots -x_n, -y_n, -z_n) = (-1)^l f_\varkappa(x_1, y_1, z_1 \cdots x_n, y_n, z_n) \quad (14)$$

sein, soll die Darstellungseigenschaft die gespiegelte sein, so ist

$$f_\varkappa(-x_1, -y_1, -z_1 \cdots -x_n, -y_n, -z_n) = -(-1)^l f_\varkappa(x_1, y_1, z_1 \cdots x_n, y_n, z_n). \quad (14\,\mathrm{a})$$

[1]) Es geht hieraus hervor, daß die $D(\alpha, \beta, \gamma)$ ein vollständiges Funktionensystem bilden müssen.

Ist die Anzahl der Teilchen, die in die Differentialgleichung eingehen, mindestens gleich 3 (die gesamte Anzahl der Teilchen also mindestens gleich 4, da ein Teilchen wegen des Schwerpunktsatzes abseparierte werden muß: z. B. gehen in die Gleichung des H-Atoms nur drei Koordinaten einem Teilchen entsprechend ein. Vgl. z. B. E. Fues, Ann. d. Phys. **80**, 367, 1926), so kann man beides erreichen, indem man etwa die χ in der Lage B nur dort willkürlich definiert, wo $x_3 > 0$ und die Werte für $x_3 < 0$ durch die Gleichung (14) bzw. (14a) bestimmt. Im Falle von nur drei Teilchen (z. B. He) ist dies jedoch nicht möglich; dort kommen nur die normalen Darstellungen in Betracht.

Will man Funktionen bilden [1]), die sich bei der Vertauschung von Elektronen so benehmen, wie eine bestimmte in 11. I beschriebene Darstellung dies verlangt, so hat man lediglich die χ nur dort willkürlich zu definieren, wo $r_1 < r_2 < \cdots < r_n$ ist (r_i Abstand des i-ten Teilchens vom Schwerpunkt) und an den anderen Stellen die Funktionen mit Hilfe ihrer Darstellungseigenschaften (9), (9a) zu bestimmen.

Wegen 9. können wir hieraus schließen: Vom Li ab kommen in jedem Multiplettsystem Terme jeder in Punkt 11, IIIa beschriebenen Darstellungseigenschaft vor. Bei He und H kommen nur normale Terme vor.

Die Eigenfunktionen sind durch die Ausdrücke (13a) gegeben, die $\chi_\lambda(g)$ hängen nur von der gegenseitigen Konfiguration der Teilchen ab.

15. In diesem Punkte möchte ich die Auswahlregeln für die azimutale Quantenzahl ableiten [2]). Unter Intensität ist hier die „Stärke des Oszillators“ zu verstehen, also die lineare Polarisation

$$\left| \int x \, \psi_i \, \psi_k \right|,$$

die höheren Momente von Heisenberg [3]) sind nicht in Betracht gezogen.

[1]) Es gilt dies nur, wenn ein Teilchen (bei uns der Kern) da ist, von dem nur ein Exemplar vorhanden ist und dessen Koordinaten man mit Hilfe des Schwerpunktsatzes abseparierte.

[2]) M. Born, W. Heisenberg und P. Jordan erhielten bereits in ihrer Arbeit (ZS. f. Phys. **35**, 557, 1926) Resultate, die sich in vielen Punkten mit den hier folgenden decken. Insbesondere findet sich dort die Auswahlregel für die magnetische Quantenzahl m und auch die Intensitätsformel (18). Auch bei P. A. M. Dirac (Proc. Roy. Soc. **111**, 281, 1926) finden sich diese Formeln. Unsere Betrachtungen führen insofern etwas weiter, als wir auch einen Übergang mit $l_1 = l_2$ (dort $j_1 = j_2$) ausschließen können, wodurch die Deutung des l als azimutale Quantenzahl gesichert erscheint. Außerdem kann l (j) dort ganz- oder halbzahlige Werte annehmen; bei uns sind nur ganzzahlige Werte möglich. Für den Fall eines einzigen Elektrons findet sich dies auch bei L. Brillouin, Journ. de phys. et le Rad. **8**, 74, 1927.

[3]) ZS. f. Phys. **38**, 411, 1926.

Da wir die Wirkung des rotierenden Elektrons nicht berücksichtigen konnten, werden alle unsere „Multiplettsysteme" (siehe Punkt 12) die Eigenschaften haben, die experimentell bei Singulettsystemen bekannt sind.

Die Dimensionen der Darstellungen der Drehgruppe sind $2l+1$. Die Zahl l nennen wir azimutale Quantenzahl, da folgendes gilt: l ändert sich bei einem Übergang um $+1$ oder -1. Alle anderen Übergänge sind verboten.

Es genügt, nachzuweisen, daß bei solchen verbotenen Übergängen die Polarisation, die vom ersten Elektron herrührt, verschwindet. Bezeichnen wir mit r_1, r_2 den Radiusvektor des ersten bzw. zweiten Elektrons, mit α, β Azimut und Polabstand von der Z-Achse des ersten, γ ist der Winkel der Ebenen Zr_1, $r_1 r_2$, so ist die $\varkappa$-te Eigenfunktion, die zu einem Term mit der azimutalen Quantenzahl l gehört, wegen (13a)

$$\sum_{\lambda} D^{l}_{\varkappa\lambda}(\alpha, \beta, \gamma)\,\chi_{\lambda}(g),$$

die $\varkappa'$-te Eigenfunktion mit der azimutalen Quantenzahl l'

$$\sum_{\lambda'} D^{l'}_{\varkappa'\lambda'}(\alpha, \beta, \gamma)\,\chi'_{\lambda'}(g).$$

Die Polarisation in der Z-Achse, die vom ersten Elektron herrührt,

$$\sum_{\lambda,\lambda'} \int D^{l}_{\varkappa\lambda}(\alpha, \beta, \gamma)\,\chi_{\lambda}(g)\,\overline{D^{l'}_{\varkappa'\lambda'}(\alpha, \beta, \gamma)\,\chi'_{\lambda'}(g)}\,r_1 \cos\beta$$

oder wegen (12)

$$\sum_{\lambda,\lambda'} \int (D^{l+1}_{\varkappa\lambda}\,u_{\varkappa\lambda} + D^{l-1}_{\varkappa\lambda}\,v_{\varkappa\lambda})\,\chi_{\lambda}(g)\,\overline{D^{l'}_{\varkappa'\lambda'}\,\chi'_{\lambda'}(g)}\,r_1,$$

$$u_{\varkappa\lambda} = \frac{\sqrt{(l+1)^2-\varkappa^2}\,\sqrt{(l+1)^2-\lambda^2}}{(2l+1)(l+1)}, \quad v_{\varkappa\lambda} = \frac{\sqrt{l^2-\varkappa^2}\,\sqrt{l^2-\lambda^2}}{(2l+1)\,l},$$

dies verschwindet schon bei der Integration über α, β, γ wegen (IIb), wenn weder $l+1=l'$ noch $l-1=l'$. Ähnlich die Polarisation in der X- und Y-Richtung.

Terme, zu denen eine normale Darstellung der Drehspiegelungsgruppe gehört, bezeichnen wir als normal, Terme, zu denen eine gespiegelte Darstellung gehört, als gespiegelt. Übergänge zwischen normalen und gespiegelten Termen kommen nicht vor.

Es genügt, dies für eine Änderung von l um ± 1 nachzuweisen, da andere Übergänge überhaupt nicht vorkommen. Die Polarisation in der Z-Richtung vom ersten Elektron ist (ψ^l normal, ψ^{l+1} gespiegelt)

$$\int \psi^l\,\overline{\psi^{l+1}}\,z_1.$$

Führen wir für x_i, y_i, z_i neue Integrationsvariable $-x_i$, $-y_i$, $-z_i$ ein, so ändert sich ψ^l mit dem Faktor $(-1)^l$; ψ^{l+1} mit dem Faktor $-(-1)^{l+1}$, z mit -1. Es ist also

$$\int \psi^l \widetilde{\psi^{l+1}} z_1 = -\int \psi^l \widetilde{\psi^{l+1}} z_1 = 0.$$

Dieser Satz wird sich als identisch mit der Laporteschen Regel erweisen. [Punkt 25][1]).

16. Wirkung des elektrischen Feldes. Durch ein elektrisches Feld in der Z-Achse wird die Substitutionsgruppe unserer Differentialgleichung verkleinert. Wir müssen also im Sinne des achten Punktes verfahren und die Darstellung der dreidimensionalen Drehgruppe als Darstellung der zweidimensionalen (um die Z-Achse) ausreduzieren. In (11) ist die Darstellung schon in dieser Form angenommen worden. [Es gilt nämlich für $\beta = 0$

$$d^l_{\varkappa\lambda}(0) = \delta_{\varkappa\lambda},$$

wie schon daraus hervorgeht, daß $D^l_{\varkappa\lambda}(\alpha, \beta, \gamma)$ für $\alpha = \beta = \gamma = 0$ die Einheitsmatrix sein muß.] In den irreduziblen Bestandteilen von $D^l_{\varkappa\lambda}(\alpha, 0, 0)$, als Darstellung der zweidimensionalen Drehspiegelungsgruppe um die Z-Achse, sind also

$$(1), \quad \begin{pmatrix} e^{i\alpha} & 0 \\ 0 & e^{-i\alpha} \end{pmatrix}, \quad \begin{pmatrix} e^{2i\alpha} & 0 \\ 0 & e^{-2i\alpha} \end{pmatrix}, \ldots, \quad \begin{pmatrix} e^{il\alpha} & 0 \\ 0 & e^{-il\alpha} \end{pmatrix}$$

die Matrizen für reine Drehung um den Winkel α; für die reine Spiegelung in der YZ-Ebene (nicht mehr durch den Koordinatenursprungspunkt!) haben wir, wenn die Darstellung normal war,

$$(1), \quad \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \ldots, \quad \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}^{2)},$$

bei gespiegelter Darstellung

$$(-1), \quad \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \ldots, \quad \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}^{2)\,3)}.$$

Wir finden also: im homogenen elektrischen Felde spaltet ein Term mit der azimutalen Quantenzahl l in $l+1$ Terme auf. Ein neuer Term

---

[1]) Vgl. F. Hund, Linienspektren usw., S. 132. Die Laportesche Regel soll also vom Li ab gelten.

[2]) Vgl. Fußnote 3 und 4, S. 636.

[3]) Die Matrizen $\begin{pmatrix} e^{i\varkappa\alpha} & 0 \\ 0 & e^{-i\varkappa\alpha} \end{pmatrix}$ und $\begin{pmatrix} 0 & -1 \\ -1 & 0 \end{pmatrix}$ gehen durch eine Ähnlichkeitstransformation in die Matrizen $\begin{pmatrix} e^{i\varkappa\alpha} & 0 \\ 0 & e^{-i\varkappa\alpha} \end{pmatrix}$ und $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ über.

ist einfach, und je nachdem der ursprüngliche Term normal oder gespiegelt war, normal oder gespiegelt, die anderen haben noch zwei Eigenfunktionen und unterscheiden sich in den beiden Fällen nicht in ihren Darstellungseigenschaften.

17. Größe der Aufspaltung und Auswahlregeln. Ein Term mit der Darstellung $\begin{pmatrix} e^{im\alpha} & 0 \\ 0 & e^{-im\alpha} \end{pmatrix}$ hat die elektrische Quantenzahl m. Bei einer Strahlung, bei der der elektrische Vektor parallel zum Felde ist, ändert sich m nicht (ein normaler Term geht in einen normalen, ein gespiegelter in einen gespiegelten über). Ist der elektrische Vektor der Strahlung senkrecht zum Feld, so kann sich m nur mit $+1$ oder -1 ändern. Wir sehen, daß die Auswahlregeln für schwache Felder durchbrochen sind: es kommt gar nicht mehr darauf an, aus was für einem Term der neue entstanden ist, nur die elektrische Quantenzahl m spielt eine Rolle.

18. Mehr Interesse beanspruchen die Verhältnisse in schwachen Feldern [1]), wo wir noch die Störungstheorie von Schrödinger anwenden können. Das Störungspotential ist

$$\sum_{i=1}^{n} \mathfrak{E}\, e\, z_i.$$

Ist keine zufällige Entartung vorhanden, so wird sich der Starkeffekt immer als quadratisch erweisen; wir müssen also, um zu einer Aufspaltung zu gelangen, bis zur zweiten Näherung vorschreiten.

Als linear unabhängige unter den $2l+1$ Eigenfunktionen mit der azimutalen Quantenzahl l nehmen wir diejenigen heraus, deren Darstellungseigenschaft sich in der Form (11) schreiben läßt: sie haben die Gestalt (13a). Die Eigenfunktionen numerieren wir so, daß wir die azimutale Quantenzahl und die laufende Nummer des Terms mit dieser azimutalen Quantenzahl als oberen Index, die Nummer der Eigenfunktion, die zu diesem Term gehört, als unteren Index hinzufügen. Die oberen Indizes sind also bei allen Eigenfunktionen, die zu demselben Term gehören, dieselben, während der untere Index von $-l$ bis $+l$ läuft.

Für die Störungstheorie brauchen wir Integrale der Gestalt

$$(l_1 s_1 m_1;\, l_2 s_2 m_2)_1 = \int z_1 \Psi_{m_1}^{l_1 s_1} \overline{\Psi_{m_2}^{l_2 s_2}} = \int r_1 \cos\beta\, \Psi_{m_1}^{l_1 s_1} \overline{\Psi_{m_2}^{l_2 s_2}} \quad (15)$$

(da das i-te Elektron von dem ersten in keiner Weise bevorzugt ist, rechnen wir immer mit dem ersten).

---

[1]) Vgl. auch A. Unsöld, Beiträge zur Quantenmechanik der Atome. Ann. d. Phys. 82, 355, 1927, § 10.

Zur Berechnung setzen wir (13a) in (15) ein:

$$(l_1 s_1 m_1;\ l_2 s_2 m_2)_1 = \sum_{\lambda_1, \lambda_2} \int D^{l_1}_{m_1 \lambda_1} \chi^{l_1 s_1}_{\lambda_1}(g)\, \widetilde{D^{l_2}_{m_2 \lambda_2}}\, \widetilde{\chi^{l_2 s_2}_{\lambda_2}}(g)\, r_1 \cos\beta \tag{15a}$$

und beachten (12), wodurch wir $D^{l_2}_{m_2 \lambda_2}(\alpha, \beta, \gamma) \cos\beta$ durch $D^{l_2+1}_{m_2 \lambda_2}(\alpha, \beta, \gamma)$ und $D^{l_2-1}_{m_2 \lambda_2}(\alpha, \beta, \gamma)$ ausdrücken. Dann führen wir die Integration über α, β und γ (d. h. über alle Lagen im Raume bei gleicher gegenseitiger Konfiguration der Teilchen) mit Hilfe von Ib und IIb aus, indem wir

$$\int D^{l_1}_{m_1 \lambda_1}(\alpha, \beta, \gamma)\, D^{k}_{m_2 \lambda_2}(\alpha, \beta, \gamma) = c_{l_1} \delta_{l_1 k} \delta_{m_1 m_2} \delta_{\lambda_1 \lambda_2}$$

setzen und erhalten für (15a)

$$\sum_{\lambda_1, \lambda_2} \{u\, c_{l_1} \delta_{l_1, l_2+1} \delta_{m_1 m_2} \delta_{\lambda_1 \lambda_2} + v\, c_{l_1} \delta_{l_1, l_2-1} \delta_{m_1 m_2} \delta_{\lambda_1 \lambda_2}\}\, \chi^{l_1 s_1}_{\lambda_1}(g)\, \widetilde{\chi^{l_2 s_2}_{\lambda_2}}(g)\, r_1\, dg,$$

$$u = \frac{\sqrt{(l_2+1)^2 - m_2^2}\,\sqrt{(l_2+1)^2 - \lambda_2^2}}{(2 l_2 + 1)(l_2 + 1)}, \quad v = \frac{\sqrt{l_2^2 - m_2^2}\,\sqrt{l_2^2 - \lambda_2^2}}{(2 l_2 + 1)\, l_2}.$$

Dies ist für $l_1 = l_2 + 1$, $m_1 = m_2$

$$(l_2 + 1, s_1, m_2;\ l_2 s_2 m_2)_1 = \sum_{\lambda_2} u\, c_{l_2+1} \int \chi^{l_2+1, s_1}_{\lambda_2}(g)\, \widetilde{\chi^{l_2 s_2}_{\lambda_2}}(g)\, r_1\, dg.$$

für $l_1 = l_2 - 1$, $m_1 = m_2$

$$(l_2 - 1, s_1 m_2;\ l_2 s_2 m_2)_1 = \sum_{\lambda_2} v\, c_{l_2-1} \int \chi^{l_2-1, s_1}_{\lambda_2}(g)\, \widetilde{\chi^{l_2 s_2}_{\lambda_2}}(g)\, r_1\, dg.$$

Wir können noch die Abhängigkeit von m — auf die es uns allein ankommt — besser zum Ausdruck bringen, indem wir für $l_1 = l_2 + 1$, $m_2 = m_1$

$$(l_1 s_1 m_1;\ l_1 - 1, s_2, m_1)_1 = \sqrt{l_1^2 - m_1^2}\, C_{l_1 s_1 s_2} \tag{16a}$$

für $l_1 = l_2 - 1$, $m_2 = m_1$

$$(l_1 s_1 m_1;\ l_1 + 1, s_2, m_1)_1 = \sqrt{(l_1+1)^2 - m_1^2}\, C'_{l_1 s_1 s_2} \tag{16b}$$

für $m_1 \neq m_2$ und für $|l_1 - l_2| \neq 1$

$$(l_1 s_1 m_1;\ l_2 s_2 m_2)_1 = 0, \quad m_1 \neq m_2 \tag{16c}$$

$$(l_1 s_1 m_1;\ l_2 s_2 m_1)_1 = 0, \quad |l_1 - l_2| \neq 1. \tag{16d}$$

Ist der eine Term normal, der andere gespiegelt, so gilt offenbar auf alle Fälle

$$(l_1 s_1 m_1;\ l_2, s_2, m_2)_1 = 0. \tag{16e}$$

In diesen Formeln hängt $C_{l_1 s_1 s_2}$ und $C'_{l_1 s_1 s_2}$ nicht mehr von m_1 ab. Da diese Formeln (16a), (16b), (16c), (16d) und (16e) nicht nur für das erste, sondern für alle Elektronen gelten, gelten sie auch für die Summe $\sum_i (l_1 s_1 m_1;\ l_2 s_2 m_2)_i$, nur daß die Zahlen $C_{l_1 s_1 s_2}$ und $C'_{l_1 s_1 s_2}$ eine etwas andere Bedeutung haben.

Jetzt können wir die Schrödingersche Störungstheorie anwenden. Für den linearen Effekt kommen hierbei — solange keine zufällige Entartung auftritt — nur Ausdrücke der Art

$$(l_1 s_1 m_1;\ l_1 s_2 m_2)$$

in Frage. Diese verschwinden aber wegen (16 d). Es ist nur ein quadratischer Effekt zu erwarten.

Die Existenz eines linearen Effekts hängt also (z. B. bei dem Wasserstoffatom) mit einer zufälligen Entartung zusammen. Auch in der älteren Theorie war dies ähnlich der Fall.

Wir können noch das quadratische Glied berechnen. Die Verschiebung des Eigenwerts, der zu der Eigenfunktion $\Psi_{m_1}^{l_1 s_1}$ gehört, ist

$$\Delta E_{l_1 s_1 m_1} = \mathfrak{E}^2 e^2 \sum_{l_2 s_2 m_2} \frac{(l_1 s_1 m_1;\ l_2 s_2 m_2)^2}{E_{l_1 s_1} - E_{l_2 s_2}}$$

oder wegen (16 c), (16 d) und (16 a), (16 b)

$$\Delta E_{l_1 s_1 m_1} = \mathfrak{E}^2 \left\{ (l_1^2 - m_1^2) A' + \left((l_1 + 1)^2 - m_1^2\right) B' \right\}$$

wo A' und B' nicht mehr von m abhängen. Anders geschrieben

$$\Delta E_{l s m} = \mathfrak{E}^2 (A_{l s} - B_{l s} m^2). \tag{17}$$

Die Zusatzenergie ist für die m-te und $-m$-te Eigenfunktion dieselbe, was ja auch natürlich ist, da diese Entartung selbst bei einem starken Feld (Punkt 16) nicht aufgehoben wird. Die absolute Größe der Verschiebung und Aufspaltung (A und B) läßt sich nicht allgemein angeben, jedoch gibt unsere Formel (17) das Bild der Aufspaltung an.

Dieselbe Abhängigkeit von m findet sich auch bei R. Becker[1]), dessen Theorie auch eine sehr gute Abschätzung von A und B ermöglicht. Auf Grund der Quantenmechanik hat seine Rechnung A. Unsöld[2]) wiederholt, der im wesentlichen zu derselben Formel gelangt.

Unsere Formel (17) muß — wenigstens für Singulettsysteme — etwa ebenso genau gelten wie z. B. die Formeln für die Aufspaltung bei dem Zeemaneffekt. Leider ist sie am experimentellen Material vorläufig nicht zu prüfen, da man bisher eine wirkliche Aufspaltung (außer natürlich bei H) kaum erreicht hat.

In der Fig. 2 ist das Bild eines solchen Effekts schematisch aufgezeichnet. Bei den vier aufgezeichneten Termen (ein S-, P-, D- und B-Term) sind die Linien mit $m = 0$ untereinander gezeichnet, und die $B_{l s}$ sind bei allen vieren gleich groß angenommen, was in Wirklichkeit

[1]) ZS. f. Phys. 9, 332, 1922.
[2]) l. c.

natürlich nicht der Fall sein wird. Die gestrichelten Linien bedeuten Übergänge, die Zahlen auf ihnen sind Intensitäten.

19. Ich möchte noch kurz auf die Verhältnisse im magnetischen Felde eingehen. Zu diesem Zwecke muß man wieder die Überlegungen des Punktes 8 heranziehen.

Ein Term habe die azimutale Quantenzahl l und gehöre zu einem Multiplettsystem M. Bis kein Feld da ist, ist die Substitutionsgruppe der Differentialgleichung (vgl. Punkt 11) das direkte Produkt der dreidimensionalen Drehspiegelungsgruppe IIIa mit der symmetrischen Gruppe I. Die Darstellung unseres Termes entsteht, indem wir die Darstellung, die zum Multiplettsystem M gehört, mit der $2l+1$ dimensionalen Darstellung der Drehgruppe kombinieren. Ich nehme die letztere Darstellung in der Form (11) an und das entstehende magnetische Feld in der Richtung der Z-Achse. Unsere Darstellung ist dann als Darstellung des direkten Produkts der symmetrischen Gruppe I mit der Gruppe IIIc bereits ausreduziert. Sie hat $2l+1$ irreduzible Bestandteile. Diese sind zusammengesetzt aus derjenigen Darstellung der Gruppe I, die zu unserem Multiplettsystem gehört, mit einer der folgenden Darstellungen von IIIc

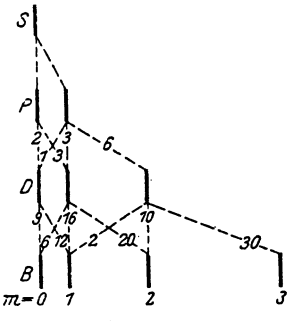


Fig. 2.

$$(e^{-il\alpha}),\ (e^{-i(l-1)\alpha})\ \ldots\ (e^{-i\alpha}),\ (1),\ (e^{i\alpha})\ \ldots\ (e^{il\alpha})$$

für eine Drehung. Diese $2l+1$ Darstellungen sind die Darstellungseigenschaften der entstandenen $2l+1$ Terme.

Dieses Resultat ist von der speziellen Gestalt der Schrödingerschen Gleichung sowie von der Darstellungseigenschaft der Gruppe I, die zu unserem Multiplettsystem M gehört, wesentlich unabhängig. Trotzdem zeigt es sich, daß die Aufspaltung immer in eine ungerade Anzahl von Komponenten eintritt, was bekanntlich der Erfahrung bei Dublett-, Quartett- usw. Systemen widerspricht. Diese Schwierigkeit tritt natürlich nicht nur an dieser Stelle auf; sie ist aber hier am krassesten zu formulieren. Es scheint mir, daß hier zur Erklärung der Verhältnisse noch ein wesentlicher Punkt fehlt.

20. Wir könnten nun — ebenso wie bei dem Starkeffekt — durch Berücksichtigung der speziellen Gestalt der Schrödingerschen Gleichung (in der allerdings das rotierende Elektron auf keine Weise

berücksichtigt ist) eine Aufspaltungsformel erhalten. Es würde sich[1]) — wie zu erwarten — die für Singulettsysteme gültige Formel des normalen Zeemaneffekts ergeben (Aufspaltungsfaktor $g = 1$). Ebenso könnten wir die Auswahlregel im beliebig starken Feld berechnen, was allerdings ein etwas komplizierteres Resultat ergibt. Für schwache Felder stehen die Integrale für die Intensitäten der π-Komponenten fertig in (16a) oder (16b) da, die Intensitäten der σ-Komponenten ergeben sich ebenso einfach durch Anwendung der Formel (12a) und (12b). Wenn wir die magnetische Quantenzahl mit m bezeichnen, so gilt für einen Übergang

$$\left.\begin{array}{lll} & & \text{Relative Intensitäten} \\ l,\ m \to l+1.\ m & (\pi\text{-Komponente}) & (l+1)^2 - m^2 \\ l,\ m \to l+1,\ m+1 & (\sigma\text{-Komponente}) & \frac{1}{2}(l+1+m)\,(l+2+m) \\ l,\ m \to l+1,\ m-1 & (\sigma\text{-Komponente}) & \frac{1}{2}(l+1-m)\,(l+2-m) \end{array}\right\} \quad (18)$$

Auch diese Formeln gelten für Singulettsysteme.

21. Als vorletzten Punkt unserer Überlegungen behandeln wir das sukzessive Einfangen von Elektronen von einem Ion. Wir werden hierbei von dem im wesentlichen von Heisenberg (l. c.) gegebenen Schema der Zuordnung der Darstellungen zu Multiplettsystemen ausgehen, die in Punkt 12 beschrieben ist.

Unsere Aufgabe besteht eigentlich aus zwei Teilen. Im ersten Teil[2]) gehen wir von einem „ungestörten" System aus, in dem keine Wechselwirkung zwischen den Elektronen vorhanden ist und diese nur unter der Einwirkung des Kerns stehen. Dann sind die Eigenfunktionen Produkte von Funktionen der Koordinaten nur je eines Elektrons, oder lineare Kombinationen solcher Produkte. Es läßt sich dann jedem Elektron i eine Hauptquantenzahl N_i und eine Azimutalquantenzahl l_i zuordnen. Wir fragen nun, was für Terme (Multiplettsystem und Azimutalquantenzahl l des Atoms) aus diesem Term entstehen, wenn wir die Wechselwirkung zwischen den Elektronen langsam, als Störung angewachsen denken. Wir haben also zu jedem Elektron eine Azimutalquantenzahl zugeordnet und fragen, wie sich diese zu der gesamten azimutalen Quantenzahl des Atoms zusammensetzen. Dieser Teil der Aufgabe läßt sich auch ganz formal erledigen; ich werde indessen hier eine etwas andere Behandlung wählen.

---

[1]) Siehe auch M. Born und P. Jordan, ZS. f. Phys. **34**, 858, 1925, und E. Schrödinger, Ann. d. Phys. **81**, 139, 1926.

[2]) Vgl. F. Hund, Linienspektren usw. §§ 24, 25.

Die zweite Frage[1]) betrifft die Zuordnung der einzelnen Seriengrenzen des Atoms zu Termen des Ions. Es ist mir leider nicht gelungen, diese Aufgabe in dieser Allgemeinheit zu lösen, obwohl es mit den hier entwickelten Methoden durchaus möglich sein muß. Ich verzichte daher auf die Behandlung dieser Frage.

22. Wir müssen hier eine kleine Hilfsbetrachtung einschalten. Es sei ein System von n gleichen Teilchen gegeben, die alle unter der Wirkung einer Kraft stehen (man denke an die n Elektronen des Atoms, wobei die Wechselwirkung vernachlässigt ist). Die Eigenfunktionen dieses Systems sind Produkte von Eigenfunktionen der einzelnen Teilchen, oder Summen solcher Produkte. Insbesondere betrachten wir einen Zustand, in dem d Bahnen doppelt besetzt sind, alle anderen einfach. Zu einem Term gehören dann $\frac{n!}{2^d}$ Eigenfunktionen. Lassen wir nun die Wechselwirkung zwischen den Teilchen eintreten, so spalten die Terme im allgemeinen auf, und es entstehen solche, deren Darstellungen für Substitutionsgruppen der Art I irreduzibel sind. Die Frage ist nun, wieviel Terme einer bestimmten Darstellungseigenschaft aus einem bestimmten, „ungestörten" Term entstehen. Das Problem wurde für den Fall $d = 0$ bereits in einer vorangehenden Mitteilung gelöst[2]), die $N_{\lambda_1 \lambda_2 \ldots \lambda_\varrho}$ geben die Zahlen der Terme an, die zu der Darstellung λ_1 $\lambda_2, \ldots \lambda_\varrho$ gehören ($\lambda_1 + \lambda_2 + \cdots + \lambda_\varrho = n$). Im allgemeinen ist das Problem etwas verwickelter, glücklicherweise interessieren uns aber nur diejenigen Termsysteme (vgl. Punkt 12), bei denen in der partitio nur die Zahlen 1 und 2 vorkommen. Die Anzahl der Zweier bezeichnen wir mit z, die Anzahl der Einser ist dann $n - 2z$, die Multiplizität des Terms $n - 2z + 1$.

Das Resultat der Rechnung, die ich nicht anführen möchte (sie ist mit der für $d = 0$ vollkommen analog, nur erfordert sie etwas mehr Rechenarbeit), ist folgendes: Aus einem ungestörten Term mit d doppelt besetzten Bahnen entstehen

$$\binom{n-2d}{z-d} - \binom{n-2d}{z-d-1} \tag{19}$$

Terme der Multiplizität $n - 2z + 1$.

---

[1]) Vgl. F. Hund, Linienspektren usw. § 39.
[2]) E. Wigner, ZS. f. Phys. **40**, 883, 1927.

Diese Formel gilt natürlich auch für $d = 0$ und ist in diesem Falle – wie man sich leicht überzeugt — mit der Formel für $N_{\lambda_1 \lambda_2 \ldots \lambda_\varrho}$ identisch, wenn nämlich alle λ gleich 1 oder 2 sind.

23. Wir betrachten nunmehr ein „ungestörtes" Atom mit n Elektronen mit den Hauptquantenzahlen $N_1, N_2, \ldots, N_n$ und den Azimutalquantenzahlen $l_1, l_2, \ldots, l_n$. Die direkte Berechnung der Aufspaltung der Eigenwerte ist zwar durchaus möglich, doch kommt man einfacher zum Ziele, wenn man sich ein schwaches magnetisches Feld angelegt denkt und so den $2l_i + 1$fachen Term des $i$ten Elektrons in $2l_i + 1$ Terme zerlegt. Dann ist unsere Aufgabe einfach die: Ein Elektron ist auf einer der $2l_1 + 1$ Bahnen N_1, l_1, m_1 (m_1 läuft von $-l_1$ bis $+l_1$), ein zweites auf einer der $2l_2 + 1$ Bahnen N_2, l_2, m_2 usw. und schließlich ein Elektron auf einer der $2l_n + 1$ Bahnen N_n, l_n, m_n (m_n läuft von $-l_n$ bis $+l_n$). Es ist nun zu beachten, daß die magnetische Quantenzahl m des Atoms aufgebaut aus Elektronen mit den magnetischen Quantenzahlen $m_1, m_2, \ldots, m_n$ gleich $m_1 + m_2 + \cdots m_n$ ist. Die Eigenfunktion des Atoms ist nämlich das Produkt der Eigenfunktionen der Elektronen, diese multiplizieren sich bei einer Drehung mit dem Winkel α um die magnetische Feldachse mit $e^{im_1\alpha}, e^{im_2\alpha}, \ldots, e^{im_n\alpha}$, die Eigenfunktion des Atoms also mit $e^{i(m_1+m_2+\cdots m_n)\alpha}$.

Hiernach können wir alle etwa $(2l_1 + 1)(2l_2 + 1)\ldots(2l_n + 1)$ Zustände separat betrachten, die durch spezielle Wahl der $m_1, m_2, \ldots, m_n$ entstehen und auf alle separat das Ergebnis des vorangehenden Punktes anwenden. Dieses genügt offenbar, da Fälle, wo mehr als zwei Elektronen dieselben drei Zahlen N, l, m haben, für uns überhaupt ausscheiden (Punkt 12).

24. Ich möchte dies an einem Beispiel illustrieren. Es seien ein $1p$- und zwei $2p$-Elektronen vorhanden ($N_1 = 1$, $N_2 = N_3 = 2$, $l_1 = l_2 = l_3 = 1$). Die erste Spalte in Tabelle 1 repräsentiert durch die drei Striche $|||$ die drei Bahnen $N = 1$, $l = 1$, $m = -1, 0, +1$, in der zweiten Spalte sind die drei Striche die drei Zustände $N = 2$, $l = 1$, $m = -1, 0, +1$. Die Punkte auf den Strichen bedeuten Elektronen, die in diesem Zustand sind, d die Anzahl der doppelt besetzten Bahnen, m sind magnetische Quantenzahlen des Atoms, also $m_1 + m_2 + m_3$. Darunter stehen die Ausdrücke

$$\binom{n-2d}{z-d} - \binom{n-2d}{z-d-1}, \tag{19}$$

Tabelle 1.

	$N=1,\ l=1$			$N=2,\ l=1$			d	$z=0$, Quartett					$z=1$, Dublett						
$m=$	-1	0	1	-1	0	1		-2	-1	0	1	2	-3	-2	-1	0	1	2	3
				↕	|	|	1						1					
				↑	↑	|	0	1							2				
				↑	|	↑	0		1						2				
	↑	|	|	|	↕	|	1							1				
				|	↑	↑	0			1						2			
				|	|	↕	1									1		
				↕	|	|	1						1	1				
				↑	↑	|	0		1	1					2	2			
	|	↑	|	↑	|	↑	0			1	1					2	2		
	und			|	↕	|	1									1	1	
	|	|	↑	|	↑	↑	0				1	1					2	2	
				|	|	↕	1											1

die angeben, wieviel solche Terme entstehen. Es entsteht also ein Quartetterm mit $m = -2$, zwei mit $m = -1$, drei mit $m = 0$, zwei mit $m = 1$ und einer mit $m = 2$. Diese kann man nur in ein ${}^4D$, ${}^4P$, ${}^4S$ zusammenfassen. Ähnlich erhalten wir für die Dubletterme ${}^2B$, ${}^2D$, ${}^2D$, ${}^2P$, ${}^2P$, ${}^2P$, ${}^2S$. Es sind dies offenbar durchweg dieselben Terme, die nach dem vorhandenen empirischen Material erwartet werden müssen. (Vgl. z. B. F. Hund, l. c. § 24, 25, wo hierfür viele Beispiele gegeben sind.)

Auf die empirische Bedeutung dieses Aufbauprinzipes hinzuweisen, ist kaum notwendig, kann man doch bekanntlich mit seiner Hilfe einen großen Teil des Baues der Serienspektren sowie des periodischen Systems erklären.

25. Wir untersuchen noch die Frage, welche Terme normal und welche gespiegelt sind. Da die Eigenfunktionen eines Elektrons immer normal sind (Punkt 14), so geht die Eigenfunktion des $i$ten Elektrons bei der Spiegelung durch den Koordinatenursprungspunkt in den $(-1)^{l_i}$fachen Wert über. Die ganze Eigenfunktion des Atoms geht in den $(-1)^{l_1+l_2+\cdots l_n}$fachen Wert über. Ist also l die azimutale Quantenzahl des ganzen Atoms, so ist der zugehörige Term normal oder gespiegelt, je nachdem $(-1)^{l_1+l_2+\cdots l_n-l}$ gleich $+1$ oder -1 ist. Nun kommen Übergänge nur von normalen in normale, von gespiegelten in gespiegelte Terme vor. $l_1+l_2+\cdots l_n-l$ kann sich also nur um eine gerade Zahl ändern. Da sich l nur um $+1$ oder -1 ändert, darf sich $l_1+l_2+\cdots l_n$ nur um eine ungerade Zahl ändern. Dies ist der Inhalt

der Laporteschen Regel. Es ist merkwürdig, wie wenig anschaulich diese Deutung ist, so daß man die in der älteren Quantentheorie von W. Heisenberg[1]) gegebene Deutung kaum wiedererkennen kann.

26. Zum Schluß möchte ich noch darauf hinweisen, daß es leicht ist — nicht nur die Kugelsymmetrie aller S-Terme [die ja von unserem Standpunkte aus trivial ist][2]) — zu beweisen, sondern das Verschwinden der Dipolmomente aller stationären Zustände zu zeigen. Auch die Quadrupolmomente lassen sich berechnen. Ich möchte aber hierauf nicht eingehen, da die vorliegende Note ohnehin schon allzu lang geworden ist. Die Bedeutung dieser Beziehungen besteht hauptsächlich in der zuerst von A. Unsöld[3]) bemerkten Tatsache, daß man bei jeder Störungsrechnung in erster Näherung so rechnen kann, wie wenn die elektrische Ladung nach der Schrödingerschen Formel über den ganzen Raum „verschmiert" wäre und nach den Regeln der klassischen Elektrostatik wirken würde.

27. Zweck der vorliegenden Arbeit war, zu zeigen, daß man durch ganz einfache Symmetriebetrachtungen über die Schrödingersche Differentialgleichung schon einen wesentlichen Teil der rein qualitativen spektroskopischen Erfahrung erklären kann.

Wie die in Punkt 19 gestreifte Schwierigkeit zu beheben ist, vermag ich nicht zu sagen. Es erscheint einem zunächst recht schwierig, mit einer Differentialgleichung mit eindeutigem ψ eine Aufspaltung in eine gerade Anzahl von Komponenten im magnetischen Felde herauszuholen. Auch zeigen sich noch einige andere Schwierigkeiten in solchen Fällen, wo die rotierenden Elektronen eine Rolle spielen.

Berlin, Institut für theoretische Physik der Techn. Hochschule.

---

[1]) ZS. f. Phys. **32**, 841, 1925.

[2]) Zu diesem Ergebnis kommt auch A. Unsöld (l. c.) für S-Terme wasserstoffähnlicher Spektren und für abgeschlossene Schalen (die natürlich in unserem Sinne auch S-Terme sind), indem er sich für die Wechselwirkung der Elektronen auf die erste bzw. zweite Näherung beschränkt. Vgl. auch A. Sommerfeld, Phys. ZS. **28**, 231, 1927.

[3]) l. c.

Berichtigung zu der Arbeit: Einige Folgerungen aus der Schrödingerschen Theorie für die Termstrukturen

E. P. Wigner

Zeitschrift für Physik 45, 601–602 (1927)

(Eingegangen am 8. September 1927.)

1. Bei der Ableitung der Rekursionsformeln (12), (12a), (12b) in zitierter Arbeit[1] ist mir ein Fehler unterlaufen. Genannte Formeln lauten richtigerweise:

$$\begin{aligned}
D^l_{jk}\cos\beta &= u_j\, u_k D^{l-1}_{j,k} + v_j\, v_k D^l_{j,k} + w_j\, w_k D^{l+1}_{j,k},\\
D^l_{jk}\sin\beta\, e^{-i\alpha} &= u^-_j\, u_k D^{l-1}_{j-1,k} + v^-_j\, v_k D^l_{j-1,k} + w^-_j\, w_k D^{l+1}_{j-1,k},\\
D^l_{jk}\sin\beta\, e^{i\alpha} &= u^+_j\, u_k D^{l-1}_{j+1,k} + v^+_j\, v_k D^l_{j+1,k} + w^+_j\, w_k D^{l+1}_{j+1,k}
\end{aligned}$$

mit

$$u_\lambda = \frac{\sqrt{l^2-\lambda^2}}{\sqrt{l}\sqrt{2l+1}},\quad v_\lambda = \frac{\lambda}{\sqrt{l}\sqrt{l+1}},\quad w_\lambda = \frac{\sqrt{(l+1)^2-\lambda^2}}{\sqrt{l+1}\sqrt{2l+1}},$$

$$u^-_\lambda = -\frac{\sqrt{l+\lambda}\sqrt{l+\lambda-1}}{\sqrt{l}\sqrt{2l+1}},\quad v^-_\lambda = \frac{\sqrt{l-\lambda+1}\sqrt{l+\lambda}}{\sqrt{l}\sqrt{l+1}},\quad w^-_\lambda = \frac{\sqrt{l-\lambda+1}\sqrt{l-\lambda+2}}{\sqrt{l+1}\sqrt{2l+1}},$$

$$u^+_\lambda = \frac{\sqrt{l-\lambda-1}\sqrt{l-\lambda}}{\sqrt{l}\sqrt{2l+1}},\quad v^+_\lambda = \frac{\sqrt{l+\lambda+1}\sqrt{l-\lambda}}{\sqrt{l}\sqrt{l+1}},\quad w^+ = -\frac{\sqrt{l+\lambda+1}\sqrt{l+\lambda+2}}{\sqrt{l+1}\sqrt{2l+1}}.$$

L. c. fehlten die Glieder mit den v. Man leitet diese Formeln ab, indem man die Darstellung $D^l_{jk} D^1_{i\varkappa}$ ausreduziert.

2. Hierdurch ändert sich Punkt 15. Außer den dort erwähnten Übergängen sind nämlich noch Übergänge mit $\Delta l = 0$ zwischen normalen und gespiegelten Termen erlaubt.

Die Auswahlregel für die azimutale Quantenzahl lautet also:

Ist ein Zustand aufgebaut aus einem l_1-, einem l_2- usw. und einem l_n-Elektron und hat er dabei die azimutale Quantenzahl l, so ist er normal oder gespiegelt, je nachdem $l_1 + l_2 + \cdots l_n - l$ eine gerade oder ungerade Zahl ist[2].

Es kommen Übergänge vor: mit $\Delta l = \pm 1$ von normalen in normale, von gespiegelten in gespiegelte Terme; mit $\Delta l = 0$, wenn $l \neq 0$ ist,

[1] ZS. f. Phys. 43, 624, 1927.
[2] l. c. Punkt 25.

von normalen in gespiegelte, gespiegelten in normale Terme[1]. Alle anderen Übergänge sind verboten.

Man beweist dies ganz nach Punkt 15. Diese Auswahlregel (wie auch das Verbot von Interkombination zwischen verschiedenen Multiplettsystemen) gilt zunächst nur — wie l. c. auseinandergesetzt — unter Vernachlässigung der Elektronenrotation, d. h. bei kleiner Multiplettaufspaltung. Die Elemente niedriger Ordnungszahl (z. B. Stickstoff) fügen sich diesen Verboten noch gut, während sie bei den höheren Elementen durchbrochen werden.

3. Auch in Punkt 18 muß man noch einige Glieder in Betracht ziehen, wodurch jedoch das Resultat nicht geändert wird.

4. In Punkt 20 kommen zu den Formeln (18) noch folgende hinzu:

	Relative Intensität
$l, m \rightarrow l, m$ (π-Komponente)	m^2
$l, m \rightarrow l, m+1$ (σ-Komponente) . .	$\frac{1}{2}(l+1+m)(l-m)$

Auch diese Formeln stehen bei Born, Heisenberg und Jordan[2].

Es ist zu beachten, daß solche Übergänge nur von normalen in gespiegelte Terme oder umgekehrt vorkommen.

Alles andere bleibt ungeändert.

Berlin, Inst. f. theoret. Phys. d. Techn. Hochschule.

---

[1] Diese Auswahlregel ist empirisch zuerst von O. Laporte (ZS. f. Phys. **23**, 135, 1924) und H. N. Russell (Science **59**, 512, 1924) mit Bezug auf „ungestrichene“ und „gestrichene“ Terme gefunden worden, welche Einteilung nach W. Heisenberg (ZS. f. Phys. **32**, 841, 1925) mit der Einteilung in „normale“ und „gespiegelte“ Terme identisch ist. Hierauf machte mich freundlichst Herr W. Pauli jr. aufmerksam.

[2] ZS. f. Phys. **35**, 557, 1926.

Über die Erhaltungssätze in der Quantenmechanik

E. P. Wigner

Nachrichten der Gesellschaft der Wissenschaften zu Göttingen
Mathematisch-Physikalische Klasse *1927*, 375–381

Vorgelegt von Max Born in der Sitzung vom 10. Februar 1928.

1. Durch die „statistische Deutung der Quantenmechanik“ [1]) mußten viele unserer gewohnten physikalischen Begriffe einer weitgehenden Revision unterzogen werden. Ob das Geschehen selber akausal ist, soll hier nicht untersucht werden, es sollen nur die Erhaltungssätze der nunmehr modifizierten Begriffe Energie, Impuls usw. besprochen werden.

Bekanntlich kann man im Sinne der Quantenmechanik niemals die Frage aufstellen: „Wie groß ist die X-Koordinate oder etwa die Energie dieses Körpers?“ Die vernünftige Fragestellung ist: Wie groß ist die Wahrscheinlichkeit, daß ein Versuch zur Bestimmung der X-Koordinate (oder Energie) diesen oder diesen Wert ergibt? In diesem Sinne müssen wir auch die Erhaltungssätze formulieren. Sie lauten dann z. B.: Die Wahrscheinlichkeit, daß die Energie den Wert E hat, ändert sich im Laufe der Zeit nicht. Dies ist also so zu verstehen, daß man bei einer Bestimmung der Energie eines Systems mit derselben Wahrscheinlichkeit den Wert E erhält, gleichgültig, ob man den Versuch zur Zeit 0 oder t ausführt. Ist speziell die Wahrscheinlichkeit für alle Energiewerte, abgesehen von E_0 gleich Null, für E_0 also 1, so können wir sagen, das System hat die Energie E_0. Dann wird dies auch zu jedem späteren Zeitpunkt der Fall sein.

1. Die analytische Bedingung dafür, daß für eine Größe, der der hermiteische Operator Q entspricht, der Erhaltungssatz gelten soll, ist die Vertauschbarkeit von Q mit dem Energieoperator H [2]).

---

1) M. Born, Zs. f. Phys., **37**, 863, **38**, 803, 1926; P. A. M. Dirac, Proc. Roy. Soc. A., **112**, 661, 1926, **113**, 621, 1927; W. Pauli jr., Zs. f. Phys., **41**, 81, 1927; P. Jordan, Zs. f. Phys., **40**, 809, **44**, 81, 1927.

2) Siehe M. Born, W. Heisenberg u. P. Jordan, Zs. f. Phys., **35**, 557, 1926; H. Weyl, Zs. f. Phys., **46**, 1, 1027.

Ist nämlich Q mit H vertauschbar, so existiert ein vollständiges Orthogonalsystem, dessen Funktionen χ_{rs} sowohl von Q, als von H Eigenfunktionen sind. Insbesondere gehöre χ_{rs} zum Eigenwert ε_r von H und λ_s von Q.

Es läßt sich nun jede Wellenfunktion ψ, durch die der Zustand beschrieben ist, in folgender Form schreiben:

$$\psi(0) = \sum_{r,s} a_{rs} \chi_{rs}.$$

Wegen der Schrödingerschen zeitabhängigen Differentialgleichung wird die Wellenfunktion zur Zeit t

$$\psi(t) = \sum_{r,s} a_{rs} \chi_{rs} e^{\frac{\varepsilon_r t}{2\pi i h}}$$

sein. Die allgemein statistischen Ansätze ergeben für die Wahrscheinlichkeit für den Wert λ_s von Q zur Zeit $t = 0$ den Wert

$$\int \left| \sum_r a_{rs} \chi_{rs} \right|^2 = \sum_r |a_{rs}|^2 \int |\chi_{rs}|^2,$$

zur Zeit t erhalten wir

$$\int \left| \sum_r a_{rs} \chi_{rs} e^{\frac{\varepsilon_r t}{2\pi i h}} \right|^2 = \sum_r |a_{rs}|^2 \int |\chi_{rs}|^2,$$

das letztere wegen der Orthogonalität der χ_{rs}. Die beiden Ausdrücke sind gleich, die Wahrscheinlichkeit hat sich nicht geändert.

Es gilt aber auch die Umkehrung, daß nämlich der Erhaltungssatz für eine Größe nur gelten kann, wenn ihr Operator mit H vertauschbar ist.

3. Wenn man bemerkt, daß mit einem Operator Q auch alle seine Funktionen mit H vertauschbar sind, so liegt es nahe nach denjenigen mit H vertauschbaren Operatoren zu fragen, die nicht mehr als Funktionen von weniger Operatoren dargestellt werden können. Diese Operatoren werden gewissermaßen die voneinander unabhängigen Erhaltungssätze liefern. (So z. B. folgt in der klassischen Mechanik der Drehimpulssatz um jeden Punkt aus dem Drehimpulssatz um einen Punkt und dem Schwerpunktssatz.) Unsere Absicht ist die Aufsuchung dieser, voneinander unabhängigen Größen.

Man kann jedem hermiteischen Operator Q im Allgemeinen in umkehrbarer Weise einen unitären Operator zuordnen, etwa $O = e^{2\pi i Q}$ oder $O = \frac{1 - iQ}{1 + iQ}$, der gleichzeitig mit Q mit dem Energieoperator H vertauschbar ist. Wir können also an Stelle

der voneinander unabhängigen hermiteischen die voneinander unabhängigen unitären, mit H vertauschbaren Operatoren O betrachten. Ist ψ Eigenfunktion von H, so ist auch $O\psi$ Eigenfunktion u. zwar zum selben Eigenwert und gleichzeitig mit ψ normiert:

$$H\psi = \varepsilon\psi, \quad OH\psi = H(O\psi) = \varepsilon(O\psi).$$

Die Operatoren O haben also die anschauliche Eigenschaft, daß sie eine Eigenfunktion von H wieder in eine Eigenfunktion überführen [1]).

4. Es werden im folgenden nicht alle Größen betrachtet, für die der Erhaltungssatz gilt, sondern nur die, für die er unter ziemlich allgemeinen, von der speziellen Gestalt des H unabhängigen Bedingungen gilt.

Zu ihrer Aufsuchung werden wir uns des Resultates des vorigen Punktes bedienen. Wir nehmen an, daß keine äußeren Kräfte da sind.

a) Dann ist es klar, daß man die Eigenfunktionen beliebig im Raume verschieben kann, ohne, daß ihr Charakter, Eigenfunktion zu bleiben, verloren gehen würde. Dem entspricht der orthogonale Operator: Ersetzen von x_i, y_i, z_i durch $x_i + a$, $y_i + b$, $z_i + c$.

b) Weiter kann man die Eigenfunktionen im Raume um einen beliebigen Punkt, etwa um den Nullpunkt des Koordinatensystems, beliebig verdrehen. Dem entspricht der orthogonale Operator: Ersetzen von x_i durch $\alpha_{11}x_i + \alpha_{12}y_i + \alpha_{13}z_i$, von y_i durch $\alpha_{21}x_i + \alpha_{22}y_i + \alpha_{23}z_i$, von z_i durch $\alpha_{31}x_i + \alpha_{32}y_i + \alpha_{33}z_i$, wo (α_{nk}) eine orthogonale Matrix, $|\alpha_{nk}| = 1$ ist. Wenn die Teilchen magnetische Momente haben, so sind die „Spinkoordinaten" mit zu transformieren [2]).

c) Zuletzt kann man die Eigenfunktion in einem beliebigen Punkt des Raumes spiegeln. Dem entspricht das Ersetzen von x_i, y_i, z_i durch $-x_i$, $-y_i$, $-z_i$.

Weiter könnte man die Koordinaten von zwei gleichen Teilchen vertauschen. Das gibt aber wegen der Heisenberg-Diracschen Forderung der Antisymmetrie der Eigenfunktionen nichts Neues.

Im Ganzen haben wir also 6 von einander unabhängige Operationen. Diese sind etwa: Verschiebung in Richtung der X-Achse, in der Y-Achse, in der Z-Achse; Drehung um die X-Achse, um die Y-Achse (jede Drehung läßt sich aus diesen zusammensetzen), und Spiegelung im Ursprungspunkt. Alle übrigen erwähnten Operationen lassen sich aus diesen zusammensetzen. Man überzeugt

1) Auf diesen Zusammenhang der Erhaltungssätze mit den Invarianzeigenschaften von H wurde von Herrn Prof. Born hingewiesen.

2) W. Pauli jr., Zs. f. Phys., **43**, 601, 1902, siehe auch J. v. Neumann und E. Wigner. Zs. f. Phys, **47**, 203, 1928.

sich leicht durch zurückgehen zu den hermiteischen Operatoren, daß aus a) der Schwerpunktssatz, aus b) der Drehimpulssatz folgt. Von drei Drehimpulssätzen sind hier also nur zwei unabhängig, worauf schon M. Born, W. Heisenberg und P. Jordan[1]) hingewiesen haben.

Etwas anders verhält es sich mit dem letzten Satz: er hat kein Analogon in der klassischen Mechanik. Ich möchte gleich bemerken, daß (wie die anderen) auch er eine wesentliche Bedeutung für die Spektren hat, er liefert nämlich die sog. Laportesche Regel, die experimentell vollkommen bestätigt ist.

Um einen hermiteischen Operator aus dem unitären Operator c) zu bilden, bestimmen wir zunächst seine Eigenwerte. Es ist

$$f(\ldots, -x_i, -y_i, -z_i, \ldots) = \lambda f(\ldots, x_i, y_i, z_i, \ldots)$$

das Eigenwertproblem. Es hat offenbar zwei Eigenwerte, $+1$ und -1. Der Operator c) ist also schon hermiteisch und wir nennen einen Zustand, für den $f(-x_i, -y_i, -z_i) = f(x_i, y_i, z_i)$ gilt, einen Sp_+ Zustand, einen für den $f(-x_i, -y_i, -z_i) = -f(x_i, y_i, z_i)$ gilt einen Sp_- Zustand. Unser Erhaltungssatz sagt also aus: die Wahrscheinlichkeit für einen Sp_+ oder Sp_- Zustand ändert sich im Laufe der Zeit nicht. (Man kann offenbar diese Wahrscheinlichkeiten berechnen, indem man die Wellenfunktion ψ in einen geraden ψ_+ und ungeraden ψ_- Teil zerlegt und $\int|\psi_+|^2$ und $\int|\psi_-|^2$ bildet.)

Um eine einigermaßen anschauliche Vorstellung über den Sinn dieses Erhaltungssatzes zu gewinnen, betrachten wir eine Anwendung. Denken wir uns, um ein möglichst einfaches Beispiel zu nehmen[2]) ein Wasserstoffatom, das einem Wasserstoffkern genähert wird. Am Ende lassen wir die beiden Kerne zusammenfallen, so daß sich ein Heliumion bildet. Die beiden entstehenden Terme sind im wesentlichen der Grundzustand und der erste P-Term des He-Jons. Ursprünglich war das System (vom Schwerpunkt aus betrachtet) mit der Wahrscheinlichkeit $\frac{1}{2}$ im Sp_+ und mit derselben Wahrscheinlichkeit im Sp_- Zustand. Folglich muß dies auch am Ende des Prozesses der Fall sein. Der Grundzustand ist ein Sp_+ Zustand, der P-Term ein Sp_- Zustand. Beide sollen also mit gleicher Wahrscheinlichkeit entstehen. Aber auch wenn mehrere Zustände entstehen, muß die Wahrscheinlichkeit der Zustände mit gerader Azimutalquantenzahl (Sp_+) gleich sein der

1) l. c.

2) Siehe F. Hund, Zs. f. Phys., **40**, 742, 1927.

Wahrscheinlichkeit der Zustände mit ungerader Azimutalquantenzahl (Sp_-).

Wir haben damit ein Beispiel für die Anwendung dieses Gesetzes gesehen. Man wird ihn aber nur selten gebrauchen können, da er nur zwei Eigenwerte (± 1) hat und so zu wenig auszusagen vermag. Die anderen beiden Gesetze (Impuls und Drehimpuls) haben eine viel weitere Anwendungsmöglichkeit[1]).

Die soeben besprochenen Erhaltungssätze müssen in der Quantenmechanik ebenso genau gelten, wie ähnliche Gesetze in der klassischen Mechanik. Wir müssen noch den Energiesatz erwähnen, der aber offenbar in der jetzigen Formulierung der Quantenmechanik eine ausgezeichnete Rolle hat.

Wir werden nun noch einige Gesetze besprechen, die nicht exakt gültig sind. Diese hängen mit orthogonalen Operatoren zusammen, die bei kleiner Wechselwirkung der magnetischen Momente der Elektronen eine Eigenfunktion in eine Eigenfunktion überführen. Diese Größen werden, solange keine zu großen Geschwindigkeiten auftreten und man sich auf kurze Zeiten beschränkt, dem Erhaltungssatz noch gut gehorchen.

Die beiden hierfür in Frage kommenden Operatoren sind folgende.

d) Verdrehung der Schwerpunktskoordinaten der Teilchen um einen beliebigen Punkt (wie in b) ohne Mitverdrehung der Spinkoordinaten.

e) Vertauschen der Ortskoordinaten gleicher Teilchen, ohne Vertauschung der Spinkoordinaten.

Es dürfte hierzu noch ein dritter Satz kommen, der sich auf die Translation bezieht, doch dürfte er nicht besonders bedeutend sein.

Während nun die vorher gegebenen Erhaltungssätze eher prinzipiell von Interesse waren, lassen diese eine Anzahl von interessanten Anwendungen zu.

Um uns über den Gültigkeitsbereich dieser Erhaltungssätze zu orientieren, erinnern wir uns daran, daß der erste die Auswahlregel für die Azimutalquantenzahl, die zweite das Verbot der Interkombinationen verschiedener Multiplettsysteme regelte.

Die physikalische Größe und das klassische Analogon von d) sind klar[2]), die von e) müssen wir etwas genauer betrachten.

Es erweist sich dabei nicht zweckmäßig zu einem hermiteischen Operator zurück zu kehren und für die Eigenwerte desselben den

1) Siehe z. B. F. Hund, l. c.

2) Siehe auch Punkt 4.

Erhaltungssatz (der also nur näherungsweise gilt) zu formulieren, da wir uns — e) hat kein klassisches Analogon — doch nichts darunter vorstellen könnten. Die verschiedenen Zustände, in denen das System sein kann, sind ja durch die Zerlegung der Zahl n in lauter 1 und 2 gekennzeichnet [1]). Wir bezeichnen den Zustand also zweckmäßig mit der Anzahl der Zweier $z(e)$ in dieser Zerlegung. (Das Argument e bedeutet, daß es sich um Elektronen handelt.) Es läuft dann $z(e)$ von 0 bis $\frac{n}{2}$ bzw. $\frac{n-1}{2}$ über alle ganzen Zahlen. Wenn unser System ein Atom ist, so ist $n-2z(e)+1$ das Multiplettsystem und unsere Regel sagt aus, daß sich das Multiplettsystem nicht ändert.

6. Wir betrachten jetzt einige Anwendungen dieser letzten Regel auf Stoßprozesse. Der einfachste Fall ist offenbar der, daß wir das $z(e)$ des Systems von vornherein kennen, wie z. B. bei dem Stoß eines Teilchens, das gar keine Elektronen hat gegen ein beliebiges Atom. In diesem Fall ist nämlich das $z(e)$ des Atoms auch das $z(e)$ des ganzen Systems. Da sich dies durch den Prozeß nicht ändern darf, kommen wir zum Resultat, daß das $z(e)$ des Atoms bei einem solchen Stoß nicht geändert wird.

Diese Folgerung könnte einer unmittelbaren experimentellen Prüfung zugänglich sein. Als stoßendes Teilchen, das keine Elektronen enthält, kommen zunächst α-Strahlen in Frage. Man hätte nur darauf zu achten, daß die Sekundärelektronen keinen wesentlichen Beitrag zur Lichtanregung geben sollen. Alsdann hätte man zu erwarten, daß kein anderes Multiplettsystem angeregt wird, als das Multiplettsystem des Normalzustandes.

Wenn wir nun zu dem Fall übergehen, wo beide stoßenden Teile Elektronen enthalten, entsteht die Aufgabe, die möglichen $z(e)$-Werte des Gesamtsystems aus den Werten dieser Größen bei den einzelnen Atomen ($z_1(e)$ und $z_2(e)$) zu bestimmen. Diese Aufgabe wurde zu einem etwas anderen Zwecke von F. London [2]) gelöst. Das Resultat ist, daß das $z(e)$ des Gesamtsystems zwischen $z_1(e)+z_2(e)$ und der kleineren der beiden Zahlen $n_1+z_2-z_1$ und $n_2+z_1-z_2$ liegen muß.

Nehmen wir etwa ein He- und ein H-Atom. He sei etwa im Grundzustand $z_1(e) = 1$. H hat $z_2(e) = 0$. Das $z(e)$ des ganzen Systems ist also 1. Nach dem Stoß kann $z_1(e)$ sowohl 0, wie auch

1) W. Heisenberg, Zs. f. Phys., **41**, 239, 1927; E. Wigner, Zs. f. Phys., **43**, 624, 1927; siehe auch F. Hund, Zs. f. Phys., **43**, 788, 1927.

2) Zs. f. Phys., **46**, 455, 1928. Siehe auch W. Heitler, ebenda, im Erscheinen.

1 sein, da beide Möglichkeiten ein $z(e) = 1$ ergeben können. Das He-Atom kann also durch Stoß mit einem H-Atom (oder Elektron) in den Triplett ($z_1(e) = 0$) oder auch in den Singulettzustand angeregt werden.

Betrachten wir dagegen zwei He-Atome im Grundzustand, die gegeneinander stoßen, so haben wir

$$n_1 = n_2 = 2, \quad z_1(e) = z_2(e) = 1.$$

Es ist also $z(e) = 2$. Nun ist es nicht mehr möglich, daß das eine He-Atom in einen Triplettzustand angeregt wird, während das andere unangeregt bleiben würde. Dies würde nämlich einen Zustand des Gesamtsystems mit $z(e) = 1$ bedeuten.

Ähnlich überlegt man sich leicht, daß durch den Stoß eines einzelnen Elektrons das $z(e)$ eines Atoms sich im Gültigkeitsbereich unseres Erhaltungssatzes um ± 1 oder 0 ändern kann. Es ist also z. B. eine Anregung von Singulett- in den Quintettzustand bei niedrigen Atomen durch Elektronenstoß unwahrscheinlich. Diese Beispiele ließen sich offenbar noch beträchtlich vermehren.

Zusammenfassung: Nach einer allgemeinen Betrachtung über den Sinn der Erhaltungssätze in der Quantenmechanik wird der Zusammenhang von Darstellungstheorie und Erhaltungssätzen berührt. Die Erhaltungssätze werden aus unitären Transformationen gewonnen, die den Energieoperator invariant lassen. Es ergeben sich so die drei Impulssätze, zwei Drehimpulssätze und noch ein Spiegelungsgesetz (sowie der Erhaltungssatz) als voneinander unabhängig.

Wenn man die Wechselwirkung der Drehmomente der Elektronen usw. vernachlässigt, ergeben sich noch zwei Erhaltungssätze.

Es werden einige Anwendungen der gewonnenen Resultate auf Stoßprozesse gegeben.

Herrn Prof. M. Born bin ich für viele wertvolle Ratschläge und Anregungen herzlich dankbar.

Göttingen, Februar 1828.

Zur Erklärung einiger Eigenschaften der Spektren aus der Quantenmechanik des Drehelektrons. Erster Teil

J. von Neumann und E. P. Wigner

Zeitschrift für Physik 47, 203–220 (1928)

(Eingegangen am 28. Dezember 1927.)

Die kinematischen Eigenschaften beliebiger Systeme von Drehelektronen werden aus dem Paulischen Bilde des Spins (ohne weitere Annahmen) mit Hilfe der Dirac-Jordanschen Transformationstheorie hergeleitet. — Die Fortsetzung soll die Anwendung auf die optischen Spektren bringen.

Erster Teil.

Einleitung. Vorliegende Note hat den Zweck, die von W. Pauli gegebene Beschreibung des Drehelektrons* zur Erklärung gewisser spektroskopischer Regelmäßigkeiten heranzuziehen; mit einer Methode die bereits von einem der Verfasser angewandt wurde, um einen Teil der spektroskopischen Erfahrungen aus der Quantenmechanik (ohne Drehelektron!) herzuleiten**. Diese Methode beruht auf der Ausnutzung der elementaren Symmetrieeigenschaften aller atomaren Systeme, nämlich der Gleichheit aller Elektronen und der Gleichwertigkeit aller Richtungen des Raumes (diese letztere wird hauptsächlich benutzt werden); sie findet in der sogenannten Darstellungstheorie ihr adäquates mathematisches Werkzeug.

Die Einteilung der Arbeit ist diese: im vorliegenden ersten Teile wird die Kinematik des Drehelektrons (bzw. eines Systems mehrerer Drehelektronen) aus der qualitativen Grundidee Paulis hergeleitet, sodann im zweiten Teile die Anwendung auf diejenigen Gesetzmäßigkeiten der Spektroskopie, die mit Berücksichtigung der magnetischen Wechselwirkungen der Elektronen-Spins streng gelten, gegeben. Ein dritter Teil soll schließlich die erst bei Vernachlässigung dieser Wechselwirkungen gültigen Regelmäßigkeiten behandeln.

Daß die bekannte*** Kinematik des Drehelektrons hier nochmals in extenso behandelt wird, sei damit entschuldigt, daß sie bisher stets durch spezielle Annahmen gewonnen wurde. Pauli entwickelt sie durch Analogisieren der Heisenberg-Born-Jordanschen Formeln für Dreh-

---

* ZS. f. Phys. **43**, 601, 1927.

** E. Wigner, ebenda **43**, 624, 1927; **45**, 601, 1927.

*** Vgl. Anmerkung 1, sowie P. Jordan, ZS. f. Phys. **44**, 21—25, 1927, § 6.

impulse*, und Jordan durch Anwendung seiner Theorie der kanonisch konjugierten Größen. Es soll nun gezeigt werden, daß sie sich ohne alle weiteren Annahmen aus der Dirac-Jordanschen Transformationstheorie** ergeben — wenn die eingangs erwähnten Prinzipien (Gleichheit aller Elektronen und aller Raumrichtungen) konsequent berücksichtigt werden.

Im Anschluß daran ist es vielleicht nicht unnütz, darauf hinzuweisen, welche Spürkraft diesen (und ähnlichen) Symmetrie- (d. h. Invarianz-) Prinzipien beim Aufsuchen von Naturgesetzen innewohnt: so wird es z. B. in diesem Falle vom qualitativen Bilde Paulis vom Drehelektron eindeutig und zwingend zu den Regelmäßigkeiten der Atomspektren führen. Es ist ähnlich wie in der allgemeinen Relativitätstheorie, wo ein Invarianzprinzip die Auffindung der universellen Naturgesetze ermöglichte.

I. Vorbereitungen.

1. Der Rahmen der folgenden Überlegungen ist, wie bereits erwähnt, die Dirac-Jordansche Transformationstheorie der Quantenmechanik. Nach derselben erfolgt die Ermittlung des gegenwärtigen Zustandes eines Systems durch einen „maximalen Versuch", d. h. durch eine Anzahl simultaner Experimente, die so geartet sind, daß jedes weitere Experiment (welches eine, durch die bereits gemessenen Größen noch nicht bestimmte Größe mißt) die Resultate derselben stören muß***. Solche „maximalen Versuche" sind z. B. (bei einem Massenpunkt) die Messungen aller kartesischen Koordinaten, oder aller dazu konjugierten Impulse, oder (wenn sie unentartet ist) der Energie allein, usw. Die bei diesem Versuch gemessenen Größen bilden das „Koordinatensystem", ihre möglichen Werte den „Zustandsraum". Ein beliebiger „Zustand" des Systems wird also beschrieben, indem man jedem Punkte (d. h. jedem möglichen Resultat des maximalen Versuchs) eine komplexe Zahl zuordnet: diese (komplexwertige) Funktion im „Zustandsraum" ist die „Wellenfunktion", und ihr Absolutwertquadrat gibt die Wahrscheinlichkeitsdichte dafür an, daß das

* ZS. f. Phys. **35**, 557, 1926, § 4.

** P. A. M. Dirac, Proc. Roy. Soc. **112**, 661, 1926 und **113**, 621—641, 1927; P. Jordan, ZS. f. Phys. **40**, 809, 1927 (I) und **44**, 1, 1927 (II). Gewisse Teile dieser Theorie sind schon bei F. London, ZS. f. Phys. **37**, 915, 1926 und **40**, 193, 1926 zu finden. Vgl. auch J. v. Neumann, Gött. Nachr., Sitzung vom 20. Mai 1927. — Der physikalische Sinn der Theorie tritt bei W. Heisenberg, ZS. f. Phys. **43**, 172, 1927, am klarsten hervor.

*** Die Frage, ob zwei Messungen simultan ausführbar sind (ohne sich gegenseitig zu stören) oder nicht, ist bekanntlich in der Quantenmechanik fundamental.

Resultat des erwähnten „maximalen Versuchs“ das System als im betreffenden Punkte des „Zustandsraumes“ befindlich erweisen wird. Die Phasen der „Wellenfunktion“ haben eine weniger unmittelbare Bedeutung, die erst bei Messung nicht vom „Koordinatensystem“ bestimmter — und nicht gleichzeitig mit ihm meßbarer — Größen, sowie bei der zeitlichen Fortentwicklung des Systems zutage tritt*. Übrigens bestimmen zwei Wellenfunktionen, die sich bloß um einen konstanten Faktor unterscheiden, denselben Zustand des Systems; sonst aber gehören zu verschiedenen Wellenfunktionen verschiedene Zustände.

Zu jeder (meßbaren) physikalischen Größe gehört ein (hermiteisch) symmetrischer und linearer Funktionaloperator: der, auf eine „Wellenfunktion“ angewandt, aus ihr eine andere macht**. Die zu den Koordinatengrößen gehörigen Operatoren multiplizieren jede „Wellenfunktion“ mit der entsprechenden Koordinate des „Zustandsraumes“.

Eine physikalische Größe kann die und nur die Werte annehmen, die Eigenwerte ihres Operators sind.

2. Wir betrachten nunmehr ein von n Elektronen gebildetes System***, deren Goudsmit-Uhlenbeckschen Spin wir (im Paulischen Sinne) berücksichtigen wollen. Wie finden wir in diesem System einen „maximalen Versuch“?

Zunächst können wir allenfalls alle kartesischen Koordinaten x_1, y_1, z_1, ..., x_n, y_n, z_n der n Elektronen messen, sodann aber nach Pauli ein sehr starkes Magnetfeld in Richtung der $+Z$-Achse einschalten: dies bewirkt eine Orientierung aller Elektronen in der $+Z$- oder $-Z$-Richtung. Die Feststellung, welche Elektronen sich nach $+Z$ und welche sich nach $-Z$ eingestellt haben, macht den Versuch zu einem maximalen. Er besteht somit aus einer Messung aller kartesischen Koordinaten sowie

* Auch ist die Bestimmung dieses Teiles der Wellenfunktion mehrdeutig, solange nicht weitere (über die Wahl des „Koordinatensystems“ hinausgehende) Festsetzungen getroffen werden. Vgl. P. Jordan, l. c. (II) und J. v. Neumann, l. c.

** Linear ist ein Operator T (der die Funktion f in die Funktion Tf überführt), wenn

$$T(af) = aTf, \quad T(f+g) = Tf + Tg$$

(a eine Konstante, f, g „Wellenfunktionen“) ist; zur Symmetrie vgl. das erste Zitat von Jordan oder auch das von v. Neumann in Anmerkung **, S. 204. Wie die Kenntnis des Operators T die Statistik der zugehörigen Größe vermittelt, sei hier nicht erörtert, vgl. näheres a. a. O.

*** Wobei natürlich beliebige potentielle und Wechselwirkungsenergien vorhanden sein dürfen: wir treiben zunächst nur Kinematik.

Spins in der $+Z$-Richtung für alle Elektronen: $x_1, y_1, z_1, s_1, \ldots, x_n, y_n, z_n, s_n$*.

Die $x_1, \ldots, z_n$ können alle Werte von $-\infty$ bis $+\infty$ annehmen, die $s_1, \ldots, s_n$ die Werte ± 1. Die „Wellenfunktionen" haben somit die folgende Form:

$$\varphi(x_1, \ldots, z_n; s_1, \ldots, s_n),$$

wobei hinter dem Semikolon lauter $\pm$ stehen, sie werden also von 2^n, auch mit

$$\varphi_{s_1, \ldots, s_n}(x_1 \ldots, z_n)$$

zu bezeichnenden, Funktionen der kartesischen Koordinaten allein („spinfreien Wellenfunktionen") gebildet.

3. Sei ξ irgendeine räumliche Richtung, so ist der Spin des νten Elektrons ($\nu = 1, 2, \ldots, n$) in der ξ-Richtung eine physikalische Größe, zu der also ein Operator gehören muß, wir nennen denselben $S_\xi^{(\nu)}$. $S_{+Z}^{(\nu)}$ gehört zum „Koordinatensystem" und ist daher unmittelbar angebbar: er bewirkt das Multiplizieren der „Wellenfunktion" mit der ihm entsprechenden Koordinate des Zustandsraumes, s_ν. Wenn also $S_{+Z}^{(\nu)} \varphi = \psi$ ist, so gilt

$$\varphi_{s_1, \ldots, s_n} = \pm \psi_{s_1, \ldots, s_n},$$

je nach dem $s_\nu = \pm 1$ (φ, ψ „Wellenfunktionen", $\varphi_{s_1, \ldots, s_n}, \psi_{s_1, \ldots, s_n}$ deren „spinfreie Wellenfunktionen").

Um aber die übrigen Operatoren $S_\xi^{(\nu)}$ zu bestimmen (die im allgemeinen gar nicht mit unserem „Koordinatensystem" gleichzeitig beobachtet werden können), sind etwas eingehendere Betrachtungen notwendig: ein Zurückgreifen auf die physikalische Natur unseres Systems.

Sei $\mathfrak{R}$ irgendeine Drehung (um den Nullpunkt) des dreidimensionalen Raumes, der x, y, z in x', y', z'

$$\begin{aligned} x' &= \alpha_{11} x + \alpha_{12} y + \alpha_{13} z, \\ y' &= \alpha_{21} x + \alpha_{22} y + \alpha_{23} z, \\ z' &= \alpha_{31} x + \alpha_{32} y + \alpha_{33} z \end{aligned}$$

überführt. Jede „Wellenfunktion" φ beschreibt einen gewissen Zustand unseres Systems; wenn wir auf diesen Zustand die Drehung $\mathfrak{R}$ ausüben,

* Wir normieren der Einfachheit halber die Spins mit $s = \pm 1 \left(\text{statt} \pm \frac{h m}{4 \pi e}\right)$.

so entsteht ein neuer, der die Wellenfunktion $\varphi(\mathfrak{R})$ haben möge*. Die Natur der Zuordnung

$$\varphi \rightarrow \varphi(\mathfrak{R})$$

läßt sich aus physikalischen Überlegungen leicht ermitteln. Zunächst muß sie das Absolutwertquadrat des inneren Produkts invariant** lassen ($|(\varphi, \psi)|^2$ ist ja die Übergangswahrscheinlichkeit aus dem Zustand φ in den Zustand ψ). Hieraus folgert man leicht, daß es einen linearen orthogonalen Operator $O_{\mathfrak{R}}$ gibt***, so daß stets

$$\varphi(\mathfrak{R}) = O_{\mathfrak{R}}\varphi \quad \text{oder} \quad = \overline{O_{\mathfrak{R}}\varphi}$$

(— bedeutet: komplexkonjugierte) ist****.

Man sieht leicht ein, daß nur der erste Fall in Frage kommt†.

---

* Man beachte, daß $\varphi(\mathfrak{R})$ bloß bis auf einen konstanten Faktor vom Absolutwert 1 definiert ist!

** Das innere Produkt von zwei „Wellenfunktionen" f und g ist das über den ganzen „Zustandsraum" erstreckte Integral des Produktes von f mit dem komplexkonjugierten von g. Es werde mit (f, g) bezeichnet.

*** Orthogonal heißt ein linearer Operator, wenn er das innere Produkt invariant läßt: $(f, g) = (Of, Og)$. Man beachte, daß $\varphi \rightarrow \varphi(\mathfrak{R})$ erstens kein eindeutiger Operator ist (da $\varphi(\mathfrak{R})$ bloß bis auf einen konstanten Faktor Sinn hat) und zweitens nicht linear sein müßte.

**** Man wähle nämlich ein vollständiges normiertes Orthogonalsystem $\varphi_1, \varphi_2, \ldots$ aus, dann sind die $\varphi_1(\mathfrak{R}), \varphi_2(\mathfrak{R}), \ldots$ vollständig und orthogonal und können normiert gewählt werden. Weiter schließt man unschwer aus der Invarianz des Absolutwertquadrates des inneren Produktes: wenn

$$\psi = a_1\varphi_1 + a_2\varphi_2 + \ldots, \quad \psi(\mathfrak{R}) = b_1\varphi_1(\mathfrak{R}) + b_2\varphi_2(\mathfrak{R}) + \ldots$$

ist, so ist $|a_1| = |b_1|$, $|a_2| = |b_2|, \ldots$; für ein ψ können wir also (durch geeignete Wahl der $\varphi_1(\mathfrak{R}), \varphi_2(\mathfrak{R}), \ldots$, unbeschadet ihrer Normierung) $a_1 = b_1$, $a_2 = b_2, \ldots$ machen, und zwar seien alle $|a_n| > 0$.

Durch Betrachten der übrigen

$$\psi' = a_1'\varphi_1 + a_2'\varphi_2 + \ldots, \quad \psi'(\mathfrak{R}) = b_1'\varphi_1(\mathfrak{R}) + b_2'\varphi_2(\mathfrak{R}) + \ldots$$

zeigt man nunmehr leicht, daß stets (wenn $\psi'(\mathfrak{R})$ mit einer geeigneten Konstanten multipliziert wird)

$$a_1' = b_1', \; a_2' = b_2', \ldots \quad \text{oder} \quad a_1' = \overline{b_1'}, \; a_2' = \overline{b_2'}, \ldots$$

ist. Das ist aber unsere Behauptung. Hierdurch ist auch gezeigt, daß O wirklich bis auf einen konstanten Faktor eindeutig bestimmt ist.

† Wenn H der Energieoperator ist, so variiert die Wellenfunktion φ_t (t ist die Zeit) nach Schrödinger gemäß der Differentialgleichung

$$\frac{h}{2\pi i}\frac{\partial}{\partial t}\varphi_t = H\varphi_t.$$

Hieraus folgt sofort für $\varphi_t(\mathfrak{R})$, je nachdem, ob

$$\varphi(\mathfrak{R}) = O_{\mathfrak{R}}\varphi \quad \text{oder} \quad = \overline{O_{\mathfrak{R}}\varphi}$$

gilt

$$\frac{h}{2\pi i}\frac{\partial}{\partial t}\varphi_t(\mathfrak{R}) = H'\varphi_t(\mathfrak{R}), \quad H' = O_{\mathfrak{R}} H O_{\mathfrak{R}}^{-1} \text{ bzw. } = -O_{\mathfrak{R}} H O_{\mathfrak{R}}^{-1}.$$

Also entspricht der Energie im um $\mathfrak{R}$ verdrehten Raume der Operator H'. H' hat im ersten Falle offenbar dieselben Eigenwerte wie H, im zweiten entgegengesetzte: daher scheidet der letztere aus.

Für das Folgende ist es von grundlegender Wichtigkeit, daß der lineare orthogonale Operator $O_{\mathfrak{R}}$ bloß bis auf einen konstanten Faktor (vom Absolutwert 1) definiert ist — dies wird (wie sich im nächsten Paragraphen zeigt) dazu führen, daß $O_{\mathfrak{R}}$ im wesentlichen eine sogenannte mehrdeutige Darstellung der Drehgruppe ist (vgl. Anmerkung *, S. 210), was für die spektroskopischen Anwendungen entscheidend wird.

Wir gehen nunmehr daran, durch gleichzeitige Betrachtung zweier Drehungen, $\mathfrak{R}, \mathfrak{S}$, den Darstellungscharakter von $O_{\mathfrak{R}}$ zu gewinnen.

4. Wenn wir $\varphi(x_1', \ldots, z_n'; s_1, \ldots, s_n)$ durch $\varphi(x_1, \ldots, z_n; s_1, \ldots, s_n)$ ersetzen (wobei die $x_1', \ldots, z_n'$ und $x_1, \ldots, z_n$ durch

$$\begin{aligned} x_\nu' &= \alpha_{11} x_\nu + \alpha_{12} y_\nu + \alpha_{13} z_\nu, \\ y_\nu' &= \alpha_{21} x_\nu + \alpha_{22} y_\nu + \alpha_{23} z_\nu, \\ z_\nu' &= \alpha_{31} x_\nu + \alpha_{32} y_\nu + \alpha_{33} z_\nu \end{aligned}$$

verknüpft sind), so üben wir offenbar eine linear orthogonale Operation aus, ihr Operator heiße $P_{\mathfrak{R}}^{-1}$. Betrachten wir nunmehr den (gleichfalls linear orthogonalen) Operator $P_{\mathfrak{R}}^{-1} O_{\mathfrak{R}} = Q_{\mathfrak{R}}$.

Wenn $\mathfrak{R}, \mathfrak{S}$ zwei Drehungen sind, und $\mathfrak{R}\mathfrak{S}$ ihr Produkt, so ist $\varphi(\mathfrak{R}\mathfrak{S})$ bis auf einen konstanten Faktor definiert und somit

$$O_{\mathfrak{R}} O_{\mathfrak{S}} = \text{const } O_{\mathfrak{R}\mathfrak{S}}.$$

Ferner gilt offenbar

$$P_{\mathfrak{R}}^{-1} P_{\mathfrak{S}}^{-1} = P_{\mathfrak{R}\mathfrak{S}}^{-1}.$$

Wenn T der Operator irgend einer Größe ist, in die der Spin überhaupt nicht eingeht, so ist es klar, daß für diese Größe $\varphi(\mathfrak{R})$ derselbe Zustand ist, wie derjenige, der aus φ durch die Koordinatentransformation $x_1, \ldots, z_n \to x_1', \ldots, z_n'$ entsteht. Das heißt: für diese Größe ist ψ nicht von $P_{\mathfrak{R}}^{-1} O_{\mathfrak{R}} \psi = Q_{\mathfrak{R}} \psi$ verschieden. Aber wenn alle ψ in $Q_{\mathfrak{R}} \psi$ übergehen, so geht T in $Q_{\mathfrak{R}} T Q_{\mathfrak{R}}^{-1}$ über, es ist also

$$T = Q_{\mathfrak{R}} T Q_{\mathfrak{R}}^{-1}, \quad T Q_{\mathfrak{R}} = Q_{\mathfrak{R}} T.$$

Da in die Größe T der Spin nicht eingeht, entspricht ihr ein Operator T', der bereits auf die „spinfreien Wellenfunktionen" $\varphi_{s_1, \ldots, s_n}$ anwendbar ist, und wir können $T\varphi = \psi$ durch

$$T' \varphi_{s_1, \ldots, s_n} = \psi_{s_1, \ldots, s_n}$$

definieren. Dabei ist T' ganz willkürlich. Da $Q_{\mathfrak{R}}$ mit allen diesen T vertauschbar ist, muß es die folgende Form haben:

$$\begin{aligned} Q_{\mathfrak{R}} \varphi &= \psi \\ \psi_{s_1, \ldots, s_n} &= \sum_{t_1, \ldots, t_n = \pm 1} a^{(\mathfrak{R})}_{s_1, \ldots; s_n; t_1, \ldots, t_n} \varphi_{t_1, \ldots, t_n} \end{aligned}$$

(die $a^{(\mathfrak{R})}_{s_1, \ldots, s_n; t_1, \ldots, t_n}$ sind irgendwelche, bloß von $\mathfrak{R}$ und $s_1, \ldots, s_n$, $t_1, \ldots, t_n$ abhängige Konstanten*).

Weil $Q_{\mathfrak{R}}$ orthogonal ist, ist die Matrix

$$\left\{a^{(\mathfrak{R})}_{s_1, \ldots, s_n; t_1, \ldots, t_n}\right\}$$

(die $s_1, \ldots, s_n$ numerieren die Zeilen, die $t_1, \ldots, t_n$ die Spalten) orthogonal**. Weiter ist offenbar $Q_{\mathfrak{R}}$ mit allen $P_{\mathfrak{S}}^{-1}$ vertauschbar, also auch diese mit $O_{\mathfrak{R}}$, daher gilt

$$Q_{\mathfrak{R}}\, Q_{\mathfrak{S}} = \text{const}\; Q_{\mathfrak{R}, \mathfrak{S}},$$

$$\left\{a^{(\mathfrak{R})}_{s,t}\right\} \left\{a^{(\mathfrak{S})}_{t,r}\right\} = \text{const}\; \left\{a^{(\mathfrak{R}\mathfrak{S})}_{s,r}\right\}.$$

Schließlich ist $Q_{\mathfrak{R}}$ bloß bis auf einen konstanten Faktor vom Absolutwert 1 definiert (weil $O_{\mathfrak{R}}$ so ist!), wir können daher die Determinante von $\left\{a^{(\mathfrak{R})}_{s,t}\right\}$ gleich 1 machen (ihr Absolutwert ist ja wegen der Orthogonalität $= 1$). Dann muß die Konstante unserer obigen Gleichung eine 2^n-te Einheitswurzel sein.

5. Die letzte Gleichung des vorigen Paragraphen gibt uns eine Handhabe, die Matrizen $\left\{a^{(\mathfrak{R})}_{s,t}\right\}$, und damit die Operatoren $Q_{\mathfrak{R}}$ und $O_{\mathfrak{R}} = P_{\mathfrak{R}}\, Q_{\mathfrak{R}}$ zu bestimmen: $\left\{a^{(\mathfrak{R})}_{s,t}\right\}$ ist ja eine (höchstens 2^n-fach) vieldeutige 2^n-dimensionale Darstellung der dreidimensionalen Drehgruppe***.

Die Bestimmung der Spinoperatoren $S^{(\nu)}_{\xi}$ ist dann leicht: wenn die Drehung $\mathfrak{R}$ die $+Z$-Richtung in die ξ-Richtung überführt, ist

$$S^{(\nu)}_{\xi} = O_{\mathfrak{R}}\, S^{(\nu)}_{+Z}\, O_{\mathfrak{R}}^{-1}\text{****}.$$

---

* Es ist jedenfalls

$$\psi_{s_1, \ldots, s_n} = \sum_{t_1, \ldots, t_n = \pm 1} A^{(\mathfrak{R})}_{s_1, \ldots, s_n; t_1, \ldots, t_n}\, \varphi_{t_1, \ldots, t_n},$$

wobei die $A^{(\mathfrak{R})}_{s_1, \ldots, s_n; t_1, \ldots, t_n}$ bloß von $\mathfrak{R}$ und $s_1, \ldots, s_n$; $t_1, \ldots, t_n$ abhängige, auf „spinfreie Wellenfunktionen" anwendbare Operatoren sind. Da diese mit allen T^{ν} vertauschbar sein müssen, sind sie alle von der Form $\alpha \,.\, 1$, woraus die Behauptung folgt.

** Im komplexen Sinne:

$$\sum_s a^{(\mathfrak{R})}_{s,t}\, \overline{a^{(\mathfrak{R})}_{s,r}} = \begin{cases} 1, & \text{wenn alle } t_\nu = r_\nu \text{ sind,} \\ 0, & \text{sonst,} \end{cases}$$

$$\sum_s a^{(\mathfrak{R})}_{t,s}\, \overline{a^{(\mathfrak{R})}_{r,s}} = \begin{cases} 1, & \text{wenn alle } t_\nu = r_\nu \text{ sind,} \\ 0, & \text{sonst.} \end{cases}$$

*** Man beachte, daß, im Gegensatz zur zitierten Arbeit Wigners, eine mehrdeutige Darstellung herausgekommen ist. Dies wird es uns ermöglichen, die geradzahligen Multiplizitäten der Atomspektren zu erfassen — die eindeutigen Darstellungen der Drehgruppe haben stets ungerade Dimensionszahl und führen nur zu ungeraden Multiplizitäten (vgl. a. a. O.).

**** Da φ in $O_{\mathfrak{R}}\, \varphi$ übergeht, ist jeder Operator T durch $O_{\mathfrak{R}}\, T\, O_{\mathfrak{R}}^{-1}$ zu ersetzen.

Hieraus folgt aber, daß

$$S_\xi^{(\nu)} f = g$$

$$g_{s_1, \ldots, s_n} = \sum_{t_1, \ldots, t_n = \pm 1} \beta^{\nu(\xi)}_{s_1, \ldots, s_n; t_1, \ldots, t_n} f_{t_1, \ldots, t_n}$$

ist: dies wissen wir bereits für $\xi = +Z$, mit

$$\beta^{\nu(+Z)}_{s_1, \ldots, s_n; t_1, \ldots, t_n} = \begin{cases} 1 & \text{wenn alle } s_\mu = t_\mu \text{ sind, und } s_\nu = 1 \text{ ist,} \\ -1 & \text{„ „ } s_\mu = t_\mu \text{ „ „ } s_\nu = -1 \text{ ist,} \\ 0 & \text{sonst} \end{cases}$$

oder kurz

$$\beta^{\nu(+Z)}_{s,t} = s_\nu \cdot \delta_{s,t};$$

und für andere ξ folgt es aus unserer Gleichung, mit

$$\left\{\beta^{\nu(\xi)}_{s,t}\right\} = \left\{a^{(\mathfrak{R})}_{s,u}\right\} \left\{\beta^{\nu(+Z)}_{u,v}\right\} \left\{a^{(\mathfrak{R})}_{t,v}\right\}^{-1}.$$

II. Der Fall eines Elektrons.

1. Wir setzen zunächst $n = 1$, da dieser Fall am leichtesten zu erledigen ist. In diesem Falle ist die Matrix $\left\{Q^{(\mathfrak{R})}_{s,t}\right\}$ zweidimensional, und wir können alle zweidimensionalen (endlichvieldeutigen) Darstellungen der Drehgruppe ohne weiteres angeben*. Es gilt nämlich:

Entweder sind alle Matrizen $\left\{a^{(\mathfrak{R})}_{s_1,t_1}\right\}$ gleich der Einheit $\left\{\begin{smallmatrix}1, & 0,\\ 0, & 1,\end{smallmatrix}\right\}$, oder sie entstehen alle durch Transformieren mit einer konstanten Matrix aus der folgenden:

$$\begin{Bmatrix} e^{-\frac{1}{2}i\alpha}, & 0 \\ 0 & , e^{\frac{1}{2}i\alpha} \end{Bmatrix} \begin{Bmatrix} \cos \frac{1}{2}\beta, & \sin \frac{1}{2}\beta \\ -\sin \frac{1}{2}\beta, & \cos \frac{1}{2}\beta \end{Bmatrix} \begin{Bmatrix} e^{-\frac{1}{2}i\gamma}, & 0 \\ 0 & , e^{\frac{1}{2}i\gamma} \end{Bmatrix},$$

wobei α, β, γ die Eulerschen Drehwinkel von $\mathfrak{R}$ sind**

---

* Vgl. z. B. H. Weyl, Math. ZS. (2) **24**, 342—353, 1925, insbesondere Satz 6, S. 353. Freilich ist für unsere gegenwärtigen Zwecke viel weniger als Weyls tief liegende Resultate notwendig. Eine Zusammenstellung der für uns wichtigen Begriffe findet sich übrigens bei Wigner, a. a. O., auf S. 628—631. Der besseren Übersicht halber sei hier noch erwähnt: Die dreidimensionale Drehgruppe hat irreduzible Darstellungen aller Grade $l = 1, 2, \ldots$, und zwar für jedes q genau eine. Diese ist ein- bzw. zweideutig, wenn q ungerade bzw. gerade ist. Wenn α, β, γ die Eulerschen Drehwinkel der Drehung $\mathfrak{R}$ sind (vgl. Anmerkung ** unten), so lautet das Element in der νten Zeile und μten Spalte ($\nu, \mu = -q+1, -q+3 \ldots q-3, q-1$)

$$e^{i\nu\frac{\alpha}{2}} e^{i\mu\frac{\gamma}{2}} f_{\nu\mu}(\beta),$$

wo $f_{\nu\mu}(\beta)$ ein homogenes Polynom $q-1$ten Grades in $\cos \frac{1}{2}\beta$ und $\sin \frac{1}{2}\beta$ ist. (Die Bezeichnung ist hier etwas anders gewählt als l. c.)

** Die Bedeutung der Eulerschen Drehwinkel ersieht man aus Fig. 1 bei Wigner, ZS. f. Phys. **43**, 639, 1927, Nr. 9/10 ($\mathfrak{R}$ führt den Quadranten —— in ----- über).

Der erste Fall scheidet aus: wegen $O_{\mathfrak{R}} = 1$ wären dann alle Spinoperatoren $S_{\xi}^{(1)}$ dem $S_{+Z}^{(1)}$ gleich, also auch die Spins, was unsinnig ist. Im zweiten Falle ist zu beachten, daß die Drehungen $\mathfrak{R}$ um die Z-Achse ($\beta = 0$) $+Z$ in sich selbst überführen, also $O_{\mathfrak{R}}$ mit $S_{+Z}^{(1)}$ vertauschbar ist, d. h. $\left\{a_{s_1, s_1}^{(\mathfrak{R})}\right\}$ mit $\begin{Bmatrix} 1, & 0 \\ 0, & -1 \end{Bmatrix}$. Somit muß es eine Diagonalmatrix sein. Die konstante Matrix, mit der transformiert wird, muß somit $\begin{Bmatrix} e^{-1/2 i(\alpha+\gamma)}, & 0 \\ 0 & e^{1/2 i(\alpha+\gamma)} \end{Bmatrix}$ in eine Diagonalmatrix überführen.

Daher hat sie die Form

$$\begin{Bmatrix} \varepsilon, & 0 \\ 0, & \eta \end{Bmatrix} \quad \text{oder} \quad \begin{Bmatrix} 0, & -\varepsilon \\ \eta, & 0 \end{Bmatrix},$$

wobei sie orthogonal sein muß (weil $\left\{a_{s_1, t_1}^{(\mathfrak{R})}\right\}$ es ist), also

$$|\varepsilon| = |\eta| = 1.$$

Das Ersetzen der „Wellenfunktion" φ durch ψ, wobei

$$\varphi_1 = \varepsilon\,\psi_1,\ \varphi_{-1} = \eta\,\psi_{-1} \quad \text{bzw.} \quad \varphi_1 = -\varepsilon\,\psi_{-1},\ \varphi_{-1} = \eta\,\psi_1$$

ist, ist eine linear orthogonale Operation, ihr Operator sei A. Wenn wir alle φ durch $A\varphi$ ersetzen, so geht jeder Operator T in ATA^{-1} über. Die spinfreien Operatoren (vgl. Ende von § 4 in I) ändern sich dabei nicht, ebensowenig $S_{+Z}^{(1)}$, d. h. unser „Koordinatensystem" bleibt ungeändert. Wir können daher diese Operatoren zugrunde legen, wodurch $O_{\mathfrak{R}}$ in $A\,O_{\mathfrak{R}}\,A^{-1}$ übergeht. Das neue $O_{\mathfrak{R}}$ hat daher die Matrix

$$\left\{a_{s_1 t_1}^{(\mathfrak{R})}\right\} = \begin{Bmatrix} e^{-1/2 i\alpha}, & 0 \\ 0, & e^{1/2 i\alpha} \end{Bmatrix} \begin{Bmatrix} \cos 1/2\,\beta, & \sin 1/2\,\beta \\ -\sin 1/2\,\beta, & \cos 1/2\,\beta \end{Bmatrix} \begin{Bmatrix} e^{-1/2 i\alpha}, & 0 \\ 0, & e^{1/2 i\alpha} \end{Bmatrix}.$$

2. Es bereitet nun keine Schwierigkeit, die Matrizen $\left\{\beta_{s_1 t_1}^{1(\xi)}\right\}$ auszurechnen: wenn ξ die Richtungscosinuse u, v, w hat ($u^2 + v^2 + w^2 = 1$), so findet man (die Eulerschen Drehwinkel von $\mathfrak{R}$ bestimmen sich aus $\cos\beta = w$, $e^{i\alpha}\sin\beta = u + iv$, $\gamma = 0$)

$$\left\{\beta_{s_1 t_1}^{1(\xi)}\right\} = \begin{Bmatrix} \cos\beta, & e^{-i\alpha}\sin\beta \\ e^{i\alpha}\sin\beta, & -\cos\beta \end{Bmatrix} = \begin{Bmatrix} w, & u - iv \\ u + iv, & -w \end{Bmatrix}.$$

Insbesondere ergibt sich für die $+X$- bzw. $+Y$- bzw. $+Z$-Richtung die Matrix

$$\begin{Bmatrix} 0, & 1 \\ 1, & 0 \end{Bmatrix} \text{ bzw. } \begin{Bmatrix} 0, & -i \\ i, & 0 \end{Bmatrix} \text{ bzw. } \begin{Bmatrix} 1, & 0 \\ 0, & -1 \end{Bmatrix},$$

im Einklang mit Pauli. Somit ist

$$S_{\xi}^{(1)} = u\,S_{+X}^{(1)} + v\,S_{+Y}^{(1)} + w\,S_{+Z}^{(1)}.$$

Der Spin des Elektrons hat ganz bestimmt die Richtung ξ, wenn seine „Wellenfunktion" φ Eigenfunktion von $S_\xi^{(1)}$ vom Eigenwerte 1 ist:

$$S_\xi^{(1)}\varphi = \varphi,$$

d. h.

$$\left.\begin{aligned} w\varphi_1 + (u - iv)\varphi_{-1} &= \varphi_1 \\ (u + iv)\varphi_1 - w\varphi_{-1} &= \varphi_{-1} \end{aligned}\right\}$$

oder auch

$$\varphi_1 : \varphi_{-1} = (u - iv) : (1 - w) = (1 + w) : (u + iv)$$

(die zwei letzten Verhältnisse sind wegen $u^2 + v^2 + w^2 = 1$ gleich). Dies kann erreicht werden, wenn $\varphi_1 : \varphi_{-1}$ konstant ist, also wenn z. B. φ außerhalb eines kleinen Bereiches verschwindet, und in diesem $\varphi_1 : \varphi_{-1}$ konstant ist*. Dann hat also der Spin eine wohldefinierte Richtung**.

III. Der allgemeine Fall.

1. Sei nunmehr n beliebig. Wir schreiben für $\left\{a_{s,t}^{(\mathfrak{R})}\right\}$ (s, t steht für $s_1, \ldots, s_n, t_1, \ldots, t_n$) der größeren Klarheit halber $\left\{{}^n a_{s,t}^{(\mathfrak{R})}\right\}$. $\left\{{}^1 a_{s_1,t_1}^{(\mathfrak{R})}\right\}$ haben wir bereits bestimmt; wir wollen zeigen, daß

$${}^n a_{s,t}^{(\mathfrak{R})} = {}^1 a_{s_1,t_1}^{(\mathfrak{R})}, \ldots, {}^1 a_{s_n,t_n}^{(\mathfrak{R})}$$

ist. Zu diesem Zwecke nehmen wir an, daß die Behauptung bereits für $n - 1$ bewiesen ist — wenn wir sie hieraus auf n übertragen können, so gilt sie allgemein.

Betrachten wir denjenigen Zustand der n Elektronen, in dem alle ihre Spins in Richtung der $\pm Z$-Achse orientiert sind, und zwar mögen sie

* Dies ist im wesentlichen die Umschreibung dafür, daß die Lage des Elektrons genau bekannt ist.

** Es ist vielleicht nicht ganz uninteressant, zu bemerken, daß im Falle des Spins auch der komplexe Teil der „Wahrscheinlichkeitsamplitude" eine unmittelbar anschauliche Deutung zuläßt. Es sei nämlich $\frac{\varphi_1}{\varphi_{-1}}$ konstant, und wir betrachten die Wahrscheinlichkeit W dafür, daß der Spin in der $+Z$-Richtung $= 1$ ist. Diese ist, wie man leicht berechnet, $W = \frac{|\varphi_1|^2}{|\varphi_1|^2 + |\varphi_{-1}|^2}$, und sie steht mit dem Cosinus w des Winkels zwischen $+Z$-Achse und Spinrichtung, $w = \frac{|\varphi_1|^2 - |\varphi_{-1}|^2}{|\varphi_1|^2 + |\varphi_{-1}|^2}$, im einfachen Zusammenhang:

$$W = 2w + 1.$$

In beiden kommt nur $\left|\frac{\varphi_1}{\varphi_{-1}}\right|$ zum Ausdruck. Der Arcus $\frac{\varphi_1}{\varphi_{-1}}$ hingegen ist, wie man sich leicht überzeugt, gleich dem Winkel zwischen der $+X$-Achse und der Projektion der Spinrichtung in der XY-Ebene.

nach $\sigma_1, \sigma_2, \ldots, \sigma_n$ $(= \pm 1)$ weisen. Dann ist also $\varphi_{s_1, \ldots, s_n} \equiv 0$, wenn nicht $s_1 = \sigma_1, \ldots, s_n = \sigma_n$ ist. Wie wahrscheinlich ist es nun, daß (beim Einschalten eines Magnetfeldes in der ξ-Richtung) sich die $n-1$ ersten Elektronen nach $\tau_1, \ldots, \tau_{n-1}$ $(= \pm 1$, in bezug auf $\xi)$ einstellen?

Wenn $\mathfrak{R}$ eine Drehung ist, welche $+Z$ in ξ überführt, so ist diese Wahrscheinlichkeit, an $n-1$ Elektronen formuliert*,

$$\int \left| {}^{n-1}a^{(\mathfrak{R})}_{\tau_1, \ldots, \tau_{n-1}, \sigma_1, \ldots, \sigma_{n-1}} \psi_{\sigma_1, \ldots, \sigma_n} \right|^2 = \left| {}^{n-1}a^{(\mathfrak{R})}_{\tau_1, \ldots, \tau_{n-1}, \sigma_1, \ldots, \sigma_{n-1}} \right|^2$$
$$= \cos^{2i} {}^1/_2\, \beta \sin^{2j} {}^1/_2\, \beta$$

(es gilt i-mal $\tau_\nu = \sigma_\nu$ und j-mal $\tau_\nu \neq \sigma_\nu$, $\nu = 1, \ldots, m-1$; α, β, γ sind die Eulerschen Drehwinkel von $\mathfrak{R}$); und an n Elektronen

$$\int \left| {}^{n}a^{(\mathfrak{R})}_{\tau_1, \ldots, \tau_{n-1}, 1, \sigma_1, \ldots, \sigma_n} \psi_{\sigma_1, \ldots, \sigma_n} \right|^2 + \int \left| {}^{n}a^{(\mathfrak{R})}_{\tau_1, \ldots, \tau_{n-1}, -1, \sigma_1, \ldots, \sigma_n} \psi_{\sigma_1, \ldots, \sigma_n} \right|^2$$
$$= \left| {}^{n}a^{(\mathfrak{R})}_{\tau_1, \ldots, \tau_{n-1}, 1, \sigma_1, \ldots, \sigma_n} \right|^2 + \left| {}^{n}a^{(\mathfrak{R})}_{\tau_1, \ldots, \tau_{n-1}, -1, \sigma_1, \ldots, \sigma_n} \right|^2.$$

Hieraus folgt:

$$\left| {}^{n}a^{(\mathfrak{R})}_{\tau_1, \ldots, \tau_{n-1}, 1, \sigma_1, \ldots, \sigma_n} \right|^2 + \left| {}^{n}a^{(\mathfrak{R})}_{\tau_1, \ldots, \tau_{n-1}, -1, \sigma_1, \ldots, \sigma_n} \right|^2$$
$$= \cos^{2i} {}^1/_2\, \beta \sin^{2j} {}^1/_2\, \beta. \qquad \text{(I.)}$$

Nun sind alle Matrizenelemente irreduzibler Darstellungen der Drehgruppe homogene Polynome in $\cos {}^1/_2\, \beta$, $\sin {}^1/_2\, \beta$, somit auch alle Matrizenelemente beliebiger Darstellungen (vgl. Anmerkung *, S. 210), also auch die ${}^{n}a^{(\mathfrak{R})}_{\tau_1, \ldots, \tau_{n-1}, \pm 1, \sigma_1, \ldots, \sigma_n}$. Daher ist ${}^{n}a^{(\mathfrak{R})}_{\tau_1, \ldots, \tau_{n-1}, \pm 1, \sigma_1, \ldots, \sigma_n}$ durch $\cos^i {}^1/_2\, \beta$ und $\sin^j {}^1/_2\, \beta$ teilbar. Die Rolle des n kann aber in unseren Überlegungen jede andere Zahl $1, \ldots, n$ übernehmen, wählen wir diese so, daß i bzw. j möglichst groß wird, so ergibt sich:

${}^{n}a^{(\mathfrak{R})}_{\tau_1, \ldots, \tau_n, \sigma_1, \ldots, \sigma_n}$ ist ein homogenes Polynom in $\cos {}^1/_2\, \beta$, $\sin {}^1/_2\, \beta$, und zwar ist es durch $\cos^i {}^1/_2\, \beta$ und $\sin^j {}^1/_2\, \beta$ teilbar, wo i bzw. j die Anzahl der $\tau_\nu = \sigma_\nu$ bzw. $\tau_\nu \neq \sigma_\nu$, $\nu = 1, \ldots, n$ ist (ausgenommen, wenn diese Anzahl $= n$ ist, d. h. wenn stets $\tau_\nu = \sigma_\nu$ oder stets $\tau_\nu = -\sigma_\nu$ ist — dann ist es um 1 kleiner).

2. Für $\beta = 0$ ist $\mathfrak{R}$ eine Drehung um die $+Z$-Achse, läßt also alle Spins $S^{(\nu)}_{+Z}$ invariant, d. h. es ist $\left\{\beta^{\nu(+Z)}_{s,t}\right\}$ mit $\left\{{}^{n}a^{(\mathfrak{R})}_{s,t}\right\}$ vertauschbar, und somit $\left\{{}^{n}a^{(\mathfrak{R})}_{s,t}\right\}$ eine Diagonalmatrix (die $\left\{\beta^{\nu(+Z)}_{s,t}\right\}$ sind von I. § 5 her bekannt). Es ist eine Darstellung der zweidimensionalen Drehgruppe (Drehwinkel $\alpha + \gamma$), daher haben seine Diagonalelemente, wie man leicht zeigt, die Form $e^{i\mu\, {}^1/_2 (\alpha + \gamma)}$. Es gilt also für $\beta = 0$

$${}^{n}a^{(\mathfrak{R})}_{s,t} = e^{i\mu_s\, {}^1/_2 (\alpha + \gamma)} \delta_{s,t}.$$

* Wir müßten hier, und immer wieder im folgenden, über alle räumlichen Koordinaten integrieren; wir schreiben dafür kurz $\int$.

Andererseits entsteht die Matrix $\{{}^n a_{s,t}^{(\mathfrak{R})}\}$ durch Transformation (also linear) aus einer Aneinanderreihung irreduzibler Darstellungen der Drehgruppe (vgl. Anmerkung *, S. 210). Unter den dabei vorkommenden irreduziblen Darstellungen möge die höchstgradige den Grad m haben.

Es sei also

$$\{{}^n a_{s,t}^{(\mathfrak{R})}\} = \{T_{s,u}\}\ \{J_{u,v}^{(\mathfrak{R})}\}\ \{T_{t,v}\}^{-1}$$

($\{T_{s,t}\}$ konstant, $\{J_{u,v}^{(\mathfrak{R})}\}$ eine Aneinanderreihung irreduzibler Matrizen), d. h.

$${}^n a_{s,t}^{(\mathfrak{R})} = \sum_{\substack{u_1,\ldots,u_n \\ v_1,\ldots,v_n}\} = \pm 1} J_{u,v}^{(\mathfrak{R})}\, T_{s,u}\, \overline{T}_{t,v}.$$

Insbesondere ist für ein Diagonalelement

$${}^n a_{s,s}^{(\mathfrak{R})} = \sum_{u,v} J_{u,v}^{(\mathfrak{R})}\, T_{s,u}\, \overline{T}_{s,v}.$$

Die irreduzible Darstellung höchsten Grades, die in $J_{u,v}^{(\mathfrak{R})}$ vorkommt, sei vom Grade m. Ihr letztes Diagonalelement ist $e^{i(m-1)\,1/2\,(\alpha+\gamma)}$ $f_{m-2,\,m-1}(\beta) = e^{i(m-1)\,1/2\,(\alpha+\gamma)} \cos^{m-1} {}^1/_2\,\beta$ (wie dies etwa aus den Rekursionsformeln bei E. Wigner, l. c., hervorgeht).

In der irreduziblen Darstellung höchsten Grades, $q = m$, kommt $e^{i(m-1)\,1/2\,(\alpha+\gamma)} \cos^{m-1} {}^1/_2\,\beta$ vor, etwa als $J_{\bar{u}\bar{u}}^{\mathfrak{R}}$, es sei $T_{\bar{s},\bar{u}} \neq 0$, und wir betrachten ${}^n a_{\bar{s},\bar{s}}^{(\mathfrak{R})}$. Dann hat für $\beta = 0$ $J_{\bar{u}\bar{u}}^{(\mathfrak{R})}$ den Koeffizienten $|T_{\bar{s},\bar{u}}|^2$, und $e^{ip\,1/2\,(\alpha+\gamma)}$ den Koeffizienten $\sum^1 |T_{\bar{s},u}|^2$, $\sum'$ erstreckt über alle u mit $J_{u,u}^{(\mathfrak{R})} = e^{ip\,1/2\,(\alpha+\gamma)} \cos^{p-1} {}^1/_2\,\beta$. Daher hat $e^{i(m-1)\,1/2\,(\alpha+\gamma)}$ einen Koeffizienten $\neq 0$, es ist $\mu_{\bar{s}} = m-1$, und alle anderen $e^{ip\,1/2\,(\alpha+\gamma)}$ haben den Koeffizienten 0. Daher ist $T_{\bar{s},u} = 0$, wenn nicht $J_{u,u}^{(\mathfrak{R})}$ das Diagonalelement mit $e^{i(m-1)\,1/2\,(\alpha+\gamma)}$ (d. h. das letzte) einer irrreduziblen Darstellung höchsten Grades ($= m$) ist.

Wenn nun nicht mehr $\beta = 0$ ist, so wird

$${}^n a_{\bar{s},\bar{s}}^{(\mathfrak{R})} = \sum J_{u,v}^{(\mathfrak{R})}\, T_{\bar{s},u}\, \overline{T}_{\bar{s},v}.$$

Hier ist $T_{\bar{s},u} = 0$ bzw. $T_{\bar{s},v} = 0$, wenn nicht u, v von der oben genannten Art sind; $J_{u,v}^{(\mathfrak{R})} = 0$, wenn sie es sind, aber $u \neq v$*. Daher haben wir ($\sum'$ erstreckt sich auf die genannten u)

$${}^n a_{\bar{s},\bar{s}}^{(\mathfrak{R})} = \sum{}' J_{u,u}^{(\mathfrak{R})} |T_{\bar{s},u}|^2 = \text{const.}\, e^{i(m-1)\frac{\alpha+\gamma}{2}} \cos^{m-1} {}^1/_2\,\beta.$$

Da für $\alpha = \beta = \gamma = 0$ $\{{}^n a_{s,t}^{(\mathfrak{R})}\}$ die Einheit ist, ist diese Konstante $= 1$.

* Dann geht die Zeile u durch eine andere irreduzible Darstellung in $\{J_{u,v}^{(\mathfrak{R})}\}$ als die Spalte v, so daß $J_{u,v}^{(\mathfrak{R})}$ außerhalb der irreduziblen Darstellung liegt und verschwindet.

3. Aus der Gleichung (I.) in § 1 und der Schlußgleichung von § 2 folgt:

$$\left|{}^{n}a^{(\mathfrak{R})}_{\bar{s}_1,\ldots,\bar{s}_n;\bar{s}_1,\ldots,-\bar{s}_n}\right|^2 = \cos^{2(m-2)}{}^1/_2\,\beta$$

$$\left|{}^{n}a^{(\mathfrak{R})}_{\bar{s}_1,\ldots,\bar{s}_n;\bar{s}_1,\ldots,-\bar{s}_n}\right|^2 + \left|{}^{n}a^{(\mathfrak{R})}_{\bar{s}_1,\ldots,\bar{s}_n;\bar{s}_1,\ldots,\bar{s}_n}\right|^2 = \cos^{2(n-1)}{}^1/_2\,\beta,$$

$$\left|{}^{n}a^{(\mathfrak{R})}_{\bar{s}_1,\ldots,\bar{s}_{n-1},s_n;\bar{s}_1,\ldots,-\bar{s}_{n-1},-\bar{s}_n}\right|^2 + \left|{}^{n}a^{(\mathfrak{R})}_{\bar{s}_1,\ldots,\bar{s}_{n-1},\bar{s}_n;\bar{s}_1,\ldots,-\bar{s}_{n-1},\bar{s}_n}\right|^2$$
$$= \cos^{2(n-2)}{}^1/_2\,\beta\,\sin^2{}^1/_2\,\beta,$$

also:

$$\cos^{2(m-1)}{}^1/_2\,\beta - \left|{}^{n}a^{(\mathfrak{R})}_{\bar{s}_1,\ldots\bar{s}_{n-1},\bar{s}_n;\bar{s}_1,\ldots,-\bar{s}_{n-1},-\bar{s}_n}\right|^2$$
$$= \cos^{2(n-1)}{}^1/_2\,\beta - \cos^{2(n-2)}{}^1/_2\,\beta\,\sin^2{}^1/_2\,\beta,$$

d. h., wenn wir $\cos^2{}^1/_2\,\beta = x$ setzen:

$$\left|{}^{n}a^{(\mathfrak{R})}_{\bar{s}_1,\ldots,\bar{s}_{n-1},\bar{s}_n;\bar{s}_1,\ldots,-\bar{s}_{n-1},-\bar{s}_n}\right|^2 = x^{m-1} - 2\,x^{n-1} + x^{n-2}.$$

Die rechte Seite muß also für alle $0 \leqq x \leqq 1 \geqq 0$ sein, woraus, wie man leicht einsieht, $m \leqq n+1$ folgt*.

Somit hat jede in $\{J^{(\mathfrak{R})}_{u,v}\}$ vorkommende irreduzible Darstellung einen Grad $\leqq n+1$. Die $\{{}^{n}a^{(\mathfrak{R})}_{s,t}\}$ sind also (in ihrer Abhängigkeit von β) Polynome von $\cos{}^1/_2\,\beta$, $\sin{}^1/_2\,\beta$ vom Grade $\leqq n$. Daher ist (vgl. das Schlußresultat von § 1)

$$ {}^{n}a^{(\mathfrak{R})}_{s,t} = C_{s,t}(\alpha,\gamma)\,.\,\cos^{i}{}^1/_2\,\beta\,\sin^{j}{}^1/_2\,\beta \qquad \text{(II.)}$$

(i bzw. j ist die Anzahl der $\nu = 1, \ldots, n$ mit $s_\nu = t_\nu$ bzw. $s_\nu \neq t_\nu$), ausgenommen für $s_1 = t_1, \ldots, s_n = t_n$ und $s_1 = -t_1, \ldots, s_n = -t_n$. In diesen Fällen ist

$$ {}^{n}a^{(\mathfrak{R})}_{s,s} = C_{s,s}(\alpha,\gamma)\,.\,\cos^{n-1}{}^1/_2\,\beta \quad (d_s\cos{}^1/_2\,\beta + e_s\sin{}^1/_2\,\beta)$$
$$ {}^{n}a^{(\mathfrak{R})}_{s,-s} = C_{s,s}(\alpha,\gamma)\,.\,\sin^{n-1}{}^1/_2\,\beta \quad (f_s\cos{}^1/_2\,\beta + g_s\sin{}^1/_2\,\beta)$$

Durch Anwendung (I.) in § 1 ergibt sich aber $e_s = 0$, $f_s = 0$, so daß unsere Gleichung (II.) stets gilt.

4. Für $\alpha = \gamma = 0$ hat $\{{}^{n}a^{(\mathfrak{R})}_{s,t}\}$ die Spur, d. h. Summe der Diagonalelemente

$$\text{const}\,.\,\cos^{n}{}^1/_2\,\beta.$$

Für $\beta = 0$ ist es die Einheitsmatrix, seine Spur 2^n, also ist die Konstante 2^n. Dieselbe Spur hat nun (für $\alpha = \gamma = 0$) die Matrix

$$\left\{{}^{1}a^{(\mathfrak{R})}_{s^1 s^1},\ \ldots,\ {}^{1}a^{(\mathfrak{R})}_{s_n t_n}\right\},$$

---

* $x^{m-1} - 2\,x^{n-1} + x^{n-2}$ ist für $x = 1$ gleich 0, daher muß seine Derivierte daselbst $\leqq 0$ sein:

$$m - 1 - 2\,(n-1) + n - 2 \leqq 0,\ m \leqq n+1.$$

mit der wir Übereinstimmen beweisen wollen. Beide Spuren sind Linearkombinationen der Spuren der in den betreffenden Darstellungen enthaltenen irreduziblen Darstellungen, beide Darstellungen enthalten somit dieselben irreduziblen Darstellungen (und können daher ineinander transformiert werden), wenn die Spuren der irreduziblen Darstellung der Drehgruppe auch für $\alpha = \gamma = 0$ voneinander linear unabhängig sind. Dies ist aber der Fall: die irreduzibe Darstellung vom Grade q hat für $\alpha = \gamma = 0$ die Spur $q \cos^{q-1} {}^1/_2 \beta$ (vgl. Anmerkung *, S. 210). Damit steht bereits die orthogonale Transformierbarkeit (wenn auch noch nicht die Gleichheit) unserer Matrix in die gewünschte fest.

Für $\beta = 0$ ist (vgl. § 2)

$$ {}^n a_{s,t}^{(\mathfrak{R})} = e^{i \mu_s {}^1/_2 (\alpha + \gamma)} \delta_{s,t} $$

und die Drehung mit den Drehwinkeln α, β, γ ist das Produkt solcher mit den Drehwinkeln $\alpha, 0, 0;\ 0, \beta, 0;\ \gamma, 0, 0$*; daher ist allgemein

$$ \begin{aligned} {}^n a_{s,t}^{(\mathfrak{R})} &= e^{i \mu_s {}^1/_2 \alpha} \,.\, C_{s,t}(0, 0) \cos^i {}^1/_2 \beta \sin^j {}^1/_2 \beta \,.\, e^{i \mu_t {}^1/_2 \gamma} \\ &= C_{s,t}\, e^{i(\mu_s {}^1/_2 \alpha + \mu_t {}^1/_2 \gamma)} \cos^i {}^1/_2 \beta \sin^j {}^1/_2 \beta, \end{aligned} $$

und dabei ist $C_{s,t}^{(\mathfrak{R})}$ konstant. Wenn wir $\alpha = \beta = \gamma = 0$ setzen, so ist $\left\{ {}^n a_{s,t}^{(\mathfrak{R})} \right\}$ die Einheitsmatrix: daher sind die $C_{s,s}^{(\mathfrak{R})} = 1$.

Die Spur von $\left\{ {}^n a_{s,t}^{(\mathfrak{R})} \right\}$ ist somit

$$ \sum_s e^{i \mu_s \frac{\alpha + \gamma}{2}} \,.\, \cos^n {}^1/_2 \beta, $$

und dies muß gleich der Spur von $\left\{ {}^1 a_{s_1 t_1}^{(\mathfrak{R})}, \ldots, {}^1 a_{s_n t_n}^{(\mathfrak{R})} \right\}$,

$$ \sum_s e^{i \left(\sum_\nu s_\nu\right) \frac{\alpha + \gamma}{2}} \,.\, \cos^n {}^1/_2 \beta $$

sein. Daher stimmen die μ_s mit den $\sum_\nu s_\nu$ bis auf die Reihenfolge überein; d. h. es gibt eine eindeutige und umkehrbare Zuordnung $s \to \tilde{s}$ aller Komplexe $s = s_1, \ldots, s_n$ $(s_1, \ldots, \pm 1)$ zu sich selbst, so daß $\mu_s = \sum_\nu \tilde{s}_\nu$ ist. Wir sehen also: $\left\{ {}^n a_{s,t}^{(\mathfrak{R})} \right\}$ kann sich von $\left\{ {}^1 a_{s_1 t_1}^{(\mathfrak{R})}, \ldots, {}^1 a_{s_n t_n}^{(\mathfrak{R})} \right\}$ nur in konstanten Faktoren vor den Elementen, sowie einer (gemeinsamen) Permutation aller Zeilen und Spalten unterscheiden. Die dafür noch übrigen Möglichkeiten gilt es nun weiter einzuengen.

5. Betrachten wir einen Zustand der n Elektronen, in dem alle Spins in Richtung der $\pm Z$-Achse orientiert sind, aber zwei Orientierungs-

* Das heißt: Um die Z-Achse mit α, um die X-Achse mit β, um die Z-Achse mit γ; das ist ja die Definition der Eulerschen Drehwinkel.

kombinationen möglich sind: $s_1, \ldots, s_n$ und $r_1, \ldots, r_n$. Die „Wellenfunktion" φ sei also den Bedingungen

$$\varphi_{t_1, \ldots, t_n} \equiv 0 \quad \text{für} \quad t_1, \cdots, t_n \neq \begin{cases} s_1, \ldots, s_n \\ r_1, \ldots, r_n \end{cases}$$

unterworfen. Wir nehmen weiter vorläufig $s_n = t_n = 1$ an, dann hat das n-te Elektron sicher den Spin in Richtung der $+Z$-Achse — wir können also von ihm absehen, und auch die $n-1$ ersten Elektronen durch eine „Wellenfunktion" φ^* mit

$$\varphi^*_{t_1, \ldots, t_{n-1}} \equiv 0 \quad \text{für} \quad t_1, \ldots, t_{n-1} \neq \begin{cases} s_1, \ldots, s_{n-1} \\ r_1, \ldots, s_{n-1} \end{cases}$$

charakterisieren. Wie wahrscheinlich ist es nun, daß (beim Einschalten eines Magnetfeldes in der ξ-Richtung) sich die $n-1$ ersten Elektronen $\tau_1, \ldots, \tau_{n-1}$ ($= \pm 1$ in bezug auf ξ) einstellen?

Wenn $\mathfrak{R}$ eine Drehung ist, welche $+Z$ in ξ überführt, so ist diese Wahrscheinlichkeit, an $n-1$ Elektronen formuliert†,

$$\int^* \Big| {}^{n-1}a^{(\mathfrak{R})}_{\tau_1, \ldots, \tau_{n-1}; s_1, \ldots, s_{n-1}} \varphi^*_{s_1, \ldots, s_{n-1}} + {}^{n-1}a^{(\mathfrak{R})}_{\tau_1, \ldots, \tau_{n-1}, r_1, \ldots, r_{n-1}} \varphi^*_{r_1, \ldots, r_{n-1}} \Big|^2,$$

und in n-Elektronen

$$\begin{aligned} \int \Big| & {}^{n}a^{(\mathfrak{R})}_{\tau_1, \ldots, \tau_{n-1}\,1, s_1, \ldots, s_{n-1}\,1} \varphi_{s_1, \ldots, s_{n-1}\,1} \\ & + {}^{n}a^{(\mathfrak{R})}_{\tau_1, \ldots, \tau_{n-1}\,1, r_1, \ldots, r_{n-1}\,1} \varphi_{r_1, \ldots, r_{n-1}\,1} \Big|^2 \\ & + \int \Big| {}^{n}a^{(\mathfrak{R})}_{\tau_1, \ldots, \tau_{n-1}-1, s_1, \ldots, s_{n-1}\,1} \varphi_{s_1, \ldots, s_{n-1}\,1} \\ & + {}^{n}a^{(\mathfrak{R})}_{\tau_1, \ldots, \tau_{n-1}-1, r_1, \ldots, r_{n-1}\,1} \varphi_{r_1, \ldots, r_{n-1}\,1} \Big|^2. \end{aligned}$$

Diese beiden Ausdrücke sollen nun gleich sein. Um überflüssige Rechnungen zu vermeiden, setzen wir $\alpha = 0$, $\beta = \pi/2$ und betrachten die Abhängigkeit von γ:

$${}^{n-1}a^{(\mathfrak{R})}_{u_1, \ldots, u_{n-1}; v_1, \ldots, v_{n-1}} = \pm \frac{1}{\sqrt{2^n}} e^{i \sum v_r \cdot {}^1/_2 \gamma},$$

$${}^{n}a^{(\mathfrak{R})}_{u_1, \ldots, u_n, v_1, \ldots, v_n} = c_{u,v} \frac{1}{\sqrt{2^n}} e^{i \mu_r \cdot {}^1/_2 \gamma}.$$

Man sieht sofort, daß der erste Ausdruck darum ein lineares Aggregat von $1, e^{\pm i (\sum s_r - \sum t_r) {}^1/_2 \gamma}$ ist, der zweite aber eines von $1, e^{\pm i (u_s - \mu_r) {}^1/_2 \gamma}$. Daher ist

$$\mu_s - \mu_r = \pm \left(\sum s_r - \sum r_r\right). \tag{III.}$$

† Die Integration über alle räumlichen Koordinaten werde durch $\int$, die über diejenigen der $n-1$ ersten Elektronen durch $\int^*$ andedeutet.

Dieses für $s_n = r_n = 1$ bewiesene Resultat gilt natürlich ebenso für $s_n = r_n = -1$; also braucht man nur $s_n = r_n$. Ferner kann die Rolle von n von jeder Zahl $1, \ldots, n$ übernommen werden, daher gilt (III.) immer, wenn es ein $\nu = 1, \ldots, n$ mit $s_\nu = r_\nu$ gibt. D. h. solange nicht $s_1 = -r_1, \ldots, s_n = -r_n$ ist.

Hieraus und aus der Feststellung am Schlusse von § 4 folgt aber, daß stets $\mu_s = \sum s_\nu$ oder stets $\mu_s = -\sum s_\nu$ ist*. Es gilt also:

$$ {}^n a_{s,t}^{(\mathfrak{R})} = C_{s,t}\, e^{\pm i\left(\Sigma s_\nu \cdot \frac{\alpha}{2} + \Sigma t_\nu \cdot \frac{\gamma}{2}\right)} \cos^i {}^1/_2\, \beta \sin^j {}^1/_2\, \beta $$

(stets + oder stets —), d. h. stets

$$ {}^n a_{s,t}^{(\mathfrak{R})} = C'_{s,t}\, {}^1 a_{s,t}^{(\mathfrak{R})}, \ldots, {}^1 a_{s_n, t_n}^{(\mathfrak{R})}, $$

oder stets

$$ {}^n a_{s,t}^{(\mathfrak{R})} = C'_{s,t}\, {}^1 a_{-s_1, -t_1}^{(\mathfrak{R})}, \ldots, {}^1 a_{-s_n, -t_n}^{(\mathfrak{R})}. $$

Da $\left\{{}^n a_{s,t}^{(\mathfrak{R})}\right\}$ durch Transformation aus $\left\{{}^1 a_{s_1, t_1}^{(\mathfrak{R})}, \ldots, {}^1 a_{s_n, t_n}^{(\mathfrak{R})}\right\} = \left\{A_{s,t}^{(\mathfrak{R})}\right\}$ entsteht, etwa mittels der konstanten orthogonalen Matrix $\{T_{su}\}$, so ist

$$ {}^n a_{s,t}^{(\mathfrak{R})} = \sum_{uv} A_{uv}^{(\mathfrak{R})}\, T_{su}\, \overline{T}_{tv}, $$

also

$$ C'_{s,t}\, A_{s,t}^{(\mathfrak{R})} \text{ bzw. } C'_{s,t}\, A_{-s,-t}^{(\mathfrak{R})} = \sum_{u,v} A_{u,v}^{(\mathfrak{R})}\, T_{su}\, \overline{T}_{tv}. $$

Nun gibt es unter den $A_{u,v}^{(\mathfrak{R})}$ zwar gleiche**, aber die (bis aufs Vorzeichen) voneinander verschiedenen sind offenbar linear unabhängig; daher bleibt die obige Gleichung bestehen, wenn wir alle $A_{u,v}^{(\mathfrak{R})}$ durch 1 ersetzen:

$$ C'_{s,t} = \sum_{u,v} T_{s,u}\, \overline{T}_{t,v} = \left(\sum_u T_{s,u}\right)\left(\overline{\sum_v T_{t,v}}\right) = d_s\, \bar{d}_t. $$

---

* Dieser Schluß gelingt in der Tat mühelos für $n > 2$, für $n = 2$ wäre aber zunächst auch die Möglichkeit

$$ \mu_{1,1} = \mu_{-1,-1} = 0,\ \mu_{1,-1} = \pm 2,\ \mu_{-1,1} = \mp 2 $$

zu erwägen. Indessen folgte hieraus für $n = 3$ mit der vorhin benutzten Methode

$$ \mu_{1,-1,1} - \mu_{-1,1,1} = \mu_{1,-1} - \mu_{-1,1} = \pm 4, $$

und durch zykliche Vertauschung der drei Indizes

$$ \mu_{-1,1,1} - \mu_{1,1,-1} = \pm 4,\ \mu_{1,1,-1} - \mu_{1,-1,1} = \pm 4 $$

(diese drei $\pm$ sind voneinander unabhängig). Durch Addition wird hieraus $0 = \pm 4 \pm 4 \pm 4$, was offenkundig unmöglich ist.

** Aus der Formel für ${}^1 a_{s_1, t_1}^{(\mathfrak{R})}$ folgt:

$$ A_{s,t}^{(\mathfrak{R})} = \pm\, e^{i\left(\Sigma s_\nu \cdot \frac{\alpha}{2} + \Sigma t_\nu \cdot \frac{\gamma}{2}\right)} \cos^i {}^1/_2\, \beta \sin^j {}^1/_2\, \beta, $$

Unser Schlußresultat ist also:

$$^{n}a^{(\mathfrak{R})}_{s,t} = d_s \bar{d}_t \,.\, {}^{1}a^{(\mathfrak{R})}_{s_1, t_1}, \ldots, {}^{1}a^{(\mathfrak{R})}_{s_n, t_n}, \tag{IV.}$$

oder stets

$$= d_s \bar{d}_t \,.\, {}^{1}a^{(\mathfrak{R})}_{-s_1, -t_1}, \ldots, {}^{1}a^{(\mathfrak{R})}_{-s_n, -t_n}. \tag{IV.}$$

Wegen der Orthogonalität muß außerdem

$$|d_1| = \cdots = |d_n| = 1$$

sein.

6. Weiter als bis zur Gleichung (IV.) können wir direkt nicht gelangen: auch für $n = 1$ (II, § 1) blieb eine entsprechende Ungewißheit zurück. Hier wie dort, genügt aber eine Transformation mit der Matrix

$$\{d_s \delta_{s,t}\} \quad \text{bzw.} \quad \{d_s \delta_{s,-t}\},$$

um zur gewünschten Matrix zu kommen. Wenn wir daher alle Wellenfunktion φ durch ψ ersetzen

$$\varphi_s = d_s \psi_s \quad \text{bzw.} \quad \varphi_s = d_s \psi_{-s}{}^{*},$$

so gewinnt erstens $\{{}^{n}a^{(\mathfrak{R})}_{s,t}\}$ die gewünschte Form, zweitens gehen alle $S^{(\nu)}_{+Z}$ in sich selbst über, und drittens alle spinfreien Operatoren gleichfalls. D. h.: wir können bei ungeändertem „Koordinatensystem" die gewünschte Form von $\{{}^{n}a^{(\mathfrak{R})}_{s,t}\}$ erreichen.

Damit sind wir am Ziele: $\{{}^{n}a^{(\mathfrak{R})}_{s,t}\}$ ist für alle n bestimmt.

IV. Folgerungen.

1. Wir bestimmen zunächst die $S^{(\nu)}_{\xi}$ für beliebiges n, d. h. die Matrix $\{\beta^{\nu(\xi)}_{s,t}\}$. Wenn die Anweisung am Schlusse von I, § 1 befolgt wird, so ergibt sich für die Richtung ξ mit den Richtungskosinusen u, v, w ohne weiteres der Wert dieser Matrix; es ist aber zweckmäßiger, gleich die Wirkung des Operators $S^{(\nu)}_{\xi}$ hinzuschreiben (die hervorgehobene Stelle ist die ν-te):

$$S^{(\nu)}_{\xi}\, \varphi = \psi,$$

$$\psi_{s_1, \ldots, 1, \ldots, s_n} = w\, \varphi_{s_1, \ldots, 1, \ldots, s_n} + (u - iv)\, \varphi_{s_1, \ldots, -1, \ldots, s_n},$$

$$\psi_{s_1, \ldots, -1, \ldots, s_n} = (u + iv)\, \varphi_{s_1, \ldots, 1, \ldots, s_n} - w_{s_1, \ldots, -1, \ldots, s_n}.$$

---

wobei i bzw. j die Zahl der ν mit $s_\nu = t_\nu$ bzw. $s_\nu \neq t_\nu$ ist und $+$ bzw. $-$ zu wählen ist, je nachdem für eine gerade oder ungerade Anzahl von ν

$$s_\nu = 1,\ t_\nu = -1$$

gilt. (Eine Verwechslung der imaginären Einheit i mit unserem i ist wohl nicht zu befürchten.)

* Für $s = s_1, \ldots, s_n$ ist $-s = -s_1, \ldots, -s_n$; weiter setzen wir $\delta_{s,t} = 1$ für $s_1 = t_1, \ldots, s_n = t_n$, sonst $= 0$.

Somit benimmt sich das ν-te Elektron genau so wie in II, d. h. als ob die $n-1$ anderen nicht da wären; insbesondere ist wieder

$$S_{\xi}^{(\nu)} = u\, S_{+X}^{(\nu)} + v\, S_{+Y}^{(\nu)} + w\, S_{+Z}^{(\nu)}.$$

2. Für die Anwendungen wird es von Wichtigkeit sein, die in $\left\{{}^n a_{s,t}^{(\mathfrak{R})}\right\}$ enthaltenen irreduziblen Darstellungen der Drehgruppe zu kennen; dies gelingt leicht, da sich seine Spur aus deren Spuren additiv zusammensetzt. Wir setzen sogar $\beta = \gamma = 0$, dann hat $\left\{{}^n a_{s,t}^{(\mathfrak{R})}\right\}$ die Spur

$$\sum_s e^{i \sum s_\nu \cdot \frac{\alpha}{2}} = \sum_{x=0}^{n} \binom{n}{x} e^{i(2x-n)\frac{\alpha}{2}},$$

während die irreduzible Darstellung vom Grade q die Spur

$$\sum_{x=0}^{q-1} e^{i(2x-q+1)\frac{\alpha}{2}}$$

hat. Wenn somit in $\left\{{}^n a_{s,t}^{(\mathfrak{R})}\right\}$ die irreduzible Darstellung vom Grade q genau N_q-mal vorkommt, so ist

$$\sum_{x=0}^{n} \binom{n}{x} e^{i(2x-n)\frac{\alpha}{2}} = \sum_q N_q \cdot \sum_{x=0}^{q-1} e^{i(2x-q+1)\frac{\alpha}{2}}.$$

Hieraus folgt sofort, daß nur Darstellungen der Grade $q = 1, \ldots, n+1$ vorkommen (was wir von III, § 3 her sowieso wußten), und daß dann

$$\begin{aligned} N_q &= 0, \quad \text{für gerades } n-q, \\ &= \binom{n}{\frac{n-q+1}{2}} - \binom{n}{\frac{n-q-1}{2}}, \quad \text{für ungerades } n-q \end{aligned}$$

ist. Die vorkommenden irreduziblen Darstellungen haben also die Grade $n+1,\ n-1,\ n-3, \ldots$ und die Vielfachheiten $1,\ n-1,\ \frac{n(n-3)}{2}, \cdots$

Man beachte, daß für gerade n nur eindeutige irreduzible Darstellungen vorkommen, und für ungerade n nur zweideutige.

Göttingen und Berlin, Dezember 1927.

Über das Paulische Äquivalenzverbot

P. Jordan und E. P. Wigner

Zeitschrift für Physik 47, 631–651 (1928)

(Eingegangen am 26. Januar 1928.)

Die Arbeit enthält eine Fortsetzung der kürzlich von einem der Verfasser vorgelegten Note „Zur Quantenmechanik der Gasentartung", deren Ergebnisse hier wesentlich erweitert werden. Es handelt sich darum, ein ideales oder nichtideales, dem Paulischen Äquivalenzverbot unterworfenes Gas zu beschreiben mit Begriffen, die keinen Bezug nehmen auf den abstrakten Koordinatenraum der Atomgesamtheit des Gases, sondern nur den gewöhnlichen dreidimensionalen Raum benutzen. Das wird ermöglicht durch die Darstellung des Gases vermittelst eines gequantelten dreidimensionalen Wellenfeldes, wobei die besonderen nichtkommutativen Multiplikationseigenschaften der Wellenamplitude gleichzeitig für die Existenz korpuskularer Gasatome und für die Gültigkeit des Paulischen Äquivalenzverbots verantwortlich sind. Die Einzelheiten der Theorie besitzen enge Analogien zu der entsprechenden Theorie für Einsteinsche ideale oder nichtideale Gase, wie sie von Dirac, Klein und Jordan ausgeführt wurde.

§ 1. Schon bei den ersten Untersuchungen zur systematischen Ausbildung der Matrizentheorie der Quantenmechanik ergaben sich Hinweise darauf, daß die bekannten Schwierigkeiten der Strahlungstheorie überwunden werden könnten, indem man nicht nur auf die materiellen Atome, sondern auch auf das Strahlungsfeld die quantenmechanischen Methoden anwendet*. In diesem Sinne sind durch mehrere Arbeiten** der letzten Zeit Fortschritte erzielt worden einerseits bezüglich einer quantenmechanischen Beschreibung des elektromagnetischen Feldes, andererseits bezüglich einer Formulierung der Quantenmechanik materieller Teilchen, welche die Wellendarstellung im abstrakten Koordinatenraum vermeidet zugunsten einer Darstellung durch quantenmechanische Wellen im gewöhnlichen dreidimensionalen Raume, und welche die Existenz materieller Teilchen in ähnlicher Weise zu erklären sucht, wie durch die Quantelung der elektromagnetischen Wellen die Existenz von Lichtquanten bzw. jeder durch die Annahme von Lichtquanten zu deutende physikalische Effekt erklärt wird.

Man verfährt bei dieser Beschreibung so, daß man diejenige, als q-Zahl aufzufassende Größe N_r, welche in korpuskulartheoretischer Um-

* M. Born, W. Heisenberg und P. Jordan, ZS. f. Phys. **35**, 557, 1926.

** P. A. M. Dirac, Proc. Roy. Soc. London (A) **114**, 243, 719, 1927; P. Jordan, ZS. f. Phys. **44**, 473, 1927 (im folgenden als A bezeichnet); P. Jordan und O. Klein, ZS. f. Phys. **45**, 751, 1927; P. Jordan, ZS. f. Phys. **45**, 766, 1927 (im folgenden als B bezeichnet); P. Jordan und W. Pauli jr., ZS. f. Phys. (im Erscheinen).

deutung die Anzahl von Atomen (etwa innerhalb eines Kastens) im r-ten Quantenzustande mißt, in zwei Faktoren

$$N_r = b_r^\dagger b_r \tag{1}$$

der Form

$$\left.\begin{aligned} b_r &= e^{-\frac{2\pi i}{h}\Theta_r} N_r^{1/2}, \\ b_r^\dagger &= N_r^{1/2} e^{\frac{2\pi i}{h}\Theta_r} \end{aligned}\right\} \tag{2}$$

zerlegt, wobei man fordert, daß N_r, Θ_r kanonisch konjugiert seien. Legt man nun als Definition kanonisch konjugierter Größen diejenige zugrunde, die von einem der Verfasser kürzlich vorgeschlagen wurde*, so erhält man die Möglichkeit, in dieser Form nicht nur die Einsteinsche Statistik darzustellen, bei der die Eigenwerte N_r' von N_r durch

$$N_r' = 0, 1, 2, 3, \ldots \tag{3}$$

gegeben sind, sondern auch die Paulische, bei der nur

$$N_r' = 0, 1 \tag{4}$$

in Frage kommt. Man erhält dann ferner sofort neben (2) weitere Gleichungen, und zwar im Einsteinschen Falle

$$\left.\begin{aligned} b_r &= (1 + N_r)^{1/2} e^{-\frac{2\pi i}{h}\Theta_r}, \\ b_r^\dagger &= e^{\frac{2\pi i}{h}\Theta_r} (1 + N_r)^{1/2}; \end{aligned}\right\} \tag{5'}$$

aber statt dessen im Paulischen Falle:

$$\left.\begin{aligned} b_r &= (1 - N_r)^{1/2} e^{-\frac{2\pi i}{h}\Theta_r}, \\ b_r^\dagger &= e^{\frac{2\pi i}{h}\Theta_r} (1 - N_r)^{1/2}, \end{aligned}\right\} \tag{5''}$$

wie in A gezeigt wurde.

Diese Formeln stützen bereits sehr die Überzeugung, daß diese Darstellungsweise des Paulischen Äquivalenzverbotes dem Wesen der Sache entspricht und in ihrer weiteren Verfolgung zu richtigen Ergebnissen führen wird. Die Formeln (5′), (5″) stehen nämlich in enger Beziehung einerseits zu den Problemen der stoßartigen Wechselwirkungen von Korpuskeln, und andererseits zu den Dichteschwankungen quantenmechanischer Gase.

---

* P. Jordan, ZS. f. Phys. **44**, 1, 1927.

§ 2. Was zunächst die Wechselwirkungen betrifft, so mag es erlaubt sein, aus einer früheren Note folgendes zu wiederholen*: In einem abgeschlossenen Kasten mögen (endlich oder unendlich viele) Arten verschiedener Teilchen (materielle oder Lichtquanten) vorhanden sein. Die Dichte der l-ten Teilchenart pro Zelle im Phasenraum sei $n^{(l)}(E)$, wo E die zu den betrachteten Zellen gehörige Energie bedeutet. Die Gesamtzahlen $N^{(l)}$ [Integrale der $n^{(l)}(E)$ über den Phasenraum] seien beliebig vielen ($j = 1, 2, \ldots$) linearen Nebenbedingungen

$$\sum_l C_l^j N^{(l)} = C^j = \text{const} \tag{6}$$

unterworfen (Beispiele dazu a. a. O.), wo die C_l^j ganze positive oder negative Zahlen (oder Null) sind. Dann wird im statistischen Gleichgewicht

$$n^{(l)}(E) = \frac{1}{e^{\sum_j C_l^j a_j(T) + \frac{E}{kT}} \pm 1}, \tag{7}$$

wobei das positive oder negative Vorzeichen in ± 1 zu wählen ist, je nachdem, ob die l-te Teilchenart dem Pauliprinzip oder der Einsteinstatistik gehorcht.

Als Wechselwirkungsprozesse sind nun natürlich nur solche zuzulassen, bei denen die Forderungen (6) nicht verletzt werden. Eine bestimmte Form eines solchen Elementaraktes ist zu beschreiben durch Angabe der Indizes l und der Geschwindigkeiten der vor dem Elementarakt vorhandenen und der nach dem Prozeß vorhandenen mitwirkenden Teilchen. Es seien $n_1^+, n_2^+, \ldots, n_r^+$ die zugehörigen Dichten $n^{(l)}(E)$ für diejenigen vor dem Prozeß vorhandenen Teilchen, welche der Einsteinstatistik folgen; und $n_1^-, n_2^-, \ldots, n_s^-$ die $n^{(l)}(E)$ für die vor dem Prozeß vorhandenen mitwirkenden Teilchen Paulischer Art. Entsprechend sollen sich $m_1^+, m_2^+, \ldots, m_\varrho^+$; $m_1^-, m_2^-, \ldots, m_\sigma^-$ auf die nach dem Prozeß übriggebliebenen bzw. neu erzeugten Teilchen beziehen. Dann muß aus statistisch-thermodynamischen Gründen die Wahrscheinlichkeit des Elementarakts proportional mit

$$\begin{aligned} &n_1^+ n_2^+ \ldots n_r^+ n_1^- n_2^- \ldots n_s^- (1 + m_1^+)(1 + m_2^+) \ldots \\ &\ldots (1 + m_\varrho^+)(1 - m_1^-)(1 - m_2^-) \ldots (1 - m_\sigma^-) \end{aligned} \tag{8}$$

angenommen werden; die des inversen Elementarakts entsprechend proportional mit

$$m_1^+ \ldots m_\varrho^+ m_1^- \ldots m_\sigma^- (1 + n_1^+) \ldots (1 + n_r^+)(1 - n_1^-) \ldots (1 - n_s^-). \tag{9}$$

* P. Jordan, ZS. f. Phys. **41**, 711, 1927. Anmerkung nach Abschluß der Arbeit: Dieselben Formeln sind kürzlich von Bothe, ZS. f. Phys. **46**, 327, 1928, erneut erörtert worden.

Bezüglich der Faktoren $(1 - m_1^+)$ usw. bei den Einsteinschen Teilchen ist nun von Dirac bei der Untersuchung der Absorption und Emission von Licht durch Atome gezeigt worden, daß ihre Gestalt unmittelbar folgt aus der Gestalt der entsprechenden Faktoren in den Formeln (5'). Entsprechend wird die Form der Glieder $(1 - m_1^-)$ usw. in (8), (9) auf (5'') zurückzuführen sein.

Was andererseits die Schwankungserscheinungen anbetrifft, so ist in A darauf hingewiesen, daß das Schwankungsquadrat der Teilchendichte in einem Volumen, welches mit einem großen Volumen bezüglich der zu einem engen Frequenzintervall $\Delta\nu$ gehörigen Wellen kommuniziert, nach den bekannten Einsteinschen Formeln einen Wert proportional zu

$$n_r(1 + n_r) \tag{10'}$$

besitzt. (Für klassische Wellen wäre es proportional n_r^2). Die von Pauli berechnete analoge Größe bei einem Fermischen Gase ist aber proportional

$$n_r(1 - n_r); \tag{10''}$$

und in (10'), (10'') zeigt sich der Unterschied vom Einsteinschen und Paulischen Gase wieder in derselben Form wie in (5'), (5'').

Von dem Einsteinschen Gas bzw. dem Boseschen Wellenfeld besitzt man auf Grund der Arbeiten, die auf S. 631 genannt wurden, bereits eine weitgehende Kenntnis. Wir beschäftigen uns im folgenden damit, in ähnlicher Weise die in A begründete Theorie des Pauligases zu vertiefen.

§ 3. Wir wiederholen hier der Deutlichkeit halber einige in A gebrachte Formeln. Die Größen b_r, $b_r^\dagger$, N_r, Θ_r sind darstellbar durch die Matrizen

$$\left.\begin{aligned} b_r &= \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}_r; \quad b_r^\dagger = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}_r; \\ N_r &= \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}_r; \quad \Theta_r = \frac{h}{4}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}_r. \end{aligned}\right\} \tag{11}$$

Hierbei ist jeweils

$$\begin{pmatrix} \alpha_{11} & \alpha_{12} \\ \alpha_{21} & \alpha_{22} \end{pmatrix}_r \tag{12}$$

eine Matrix, deren Zeilen und Spalten bezeichnet werden durch eine Reihe von Indizes, deren jeder den Wert 0 oder 1 haben kann; und zwar ist (12) eine Diagonalmatrix in bezug auf den ersten bis $(r - 1)$-ten Index und in bezug auf den $(r + 1)$-ten und die folgenden Indizes.

Neben den schon in § 1 besprochenen Gleichungen gelten die Formeln:

$$\left.\begin{aligned} b_r^\dagger b_r &= N_r; \quad b_r b_r^\dagger = 1 - N_r; \\ N_r^2 &= N_r; \quad (b_r^\dagger)^2 = (b_r)^2 = 0; \\ e^{\frac{2\pi i}{h}\Theta_r} &= i\,\frac{4}{h}\,\Theta_r. \end{aligned}\right\} \tag{13}$$

Man kann alle diese Größen ausdrücken durch drei Größen $k_1^{(r)}$, $k_2^{(r)}$, $k_3^{(r)}$, die den Multiplikationsregeln der Quaternionen folgen:

$$\left.\begin{aligned} k_1^{(r)} k_2^{(r)} &= -k_2^{(r)} k_1^{(r)} = k_3^{(r)}, \\ &\cdots\cdots\cdots\cdots \\ &\cdots\cdots\cdots\cdots \\ (k_1^{(r)})^2 &= (k_2^{(r)})^2 = (k_3^{(r)})^2 = -1, \end{aligned}\right\} \tag{14}$$

wobei durch die Punkte die beiden aus der angeschriebenen Gleichung durch zyklische Permutation der 1, 2, 3 hervorgehenden Gleichungen angedeutet sein sollen.

Nämlich:

$$\left.\begin{aligned} b_r &= -\frac{i\,k_3^{(r)} + k_2^{(r)}}{2}, \quad b_r^\dagger = -\frac{i\,k_3^{(r)} - k_3^{(r)}}{2}; \\ N_r &= -\frac{i\,k_1^{(r)} - 1}{2}; \quad \Theta_r = -\frac{h}{4}\,i\,k_2^{(r)}. \end{aligned}\right\} \tag{15}$$

Die Quaternionen $k_1^{(r)}$, $k_2^{(r)}$, $k_3^{(r)}$ werden dabei selbst dargestellt durch die Matrizen

$$k_1^{(r)} = \begin{pmatrix} -i & 0 \\ 0 & i \end{pmatrix}_r; \quad k_2^{(r)} = \begin{pmatrix} 0 & i \\ i & 0 \end{pmatrix}_r; \quad k_3^{(r)} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}_r. \tag{16}$$

§ 4. Ebenso wie in B für die Bosesche Statistik ausgeführt wurde kann man auch bei der Fermischen Statistik die Heisenberg-Diracschen Determinantenformeln für die Herleitung der antisymmetrischen Schrödingerschen Eigenfunktionen des Gesamtsystems aus denen eines Einzelatoms übertragen auf beliebige Wahrscheinlichkeitsamplituden. Eine solche Amplitude sei für ein Einzelatom gegeben durch

$$\Phi_{\alpha p}^{\beta q} = \Phi_{\alpha p}(\beta', q'). \tag{17}$$

Um die Vorzeichen der aus ihr zu bildenden Determinanten eindeutig zu machen, legen wir für die Eigenwerte β' von β in beliebiger, aber ein für allemal bestimmter Weise eine Reihenfolge fest. Wir bezeichnen die so definierte Anordnung für zwei spezielle Eigenwerte β', β'' von β durch $\beta' < \beta''$ bzw. $\beta'' < \beta'$, ohne jedoch mit dem Zeichen $<$ notwendig die Bedeutung „kleiner als" zu verbinden. Genau so verfahren wir mit den Eigenwerten q' von q und allgemein mit den Eigenwerten Q' jeder

meßbaren Größe Q beim Einzelatom. Danach kann man jeder Amplitude (17) für das Einzelatom in eindeutiger Weise eine antisymmetrische Amplitude $\Psi_{\alpha p}^{\beta q}$ für ein System von N energetisch ungekoppelten Teilchen zuordnen. Wir schreiben

$$\Psi_{\alpha p}^{\beta q} = \Psi_{\alpha p}(\beta^{(1)}, \beta^{(2)}, \ldots, \beta^{(N)};\ q^{(1)}, q^{(2)}, \ldots, q^{(N)}), \tag{18}$$

worin

$$\left.\begin{aligned} \beta^{(1)} < \beta^{(2)} < \cdots < \beta^{(N)}, \\ q^{(1)} < q^{(2)} < \cdots < q^{(N)} \end{aligned}\right\} \tag{19}$$

sein soll; und dann:

$$\Psi_{\alpha p}^{\beta q} = \frac{1}{N!} \sum_{(n)} \varepsilon_n \prod_{k=1}^{N} \Phi_{\alpha p}(\beta^{(k)}, q^{(n_k)}), \tag{20}$$

wo die Summe über alle $N!$ Permutationen $n_1, n_2, \ldots, n_N$ der Zahlen 1, 2, ..., N zu erstrecken ist, während ε_n gleich $+1$ ist für gerade Permutationen und -1 für ungerade.

Nach (20) verschwindet $\Psi_{\alpha p}^{\beta q}$, sobald zwei der Größen $\beta^{(k)}$ einander gleich werden. Das heißt physikalisch: Es kommt nicht vor, daß irgend eine nichtentartete Größe β bei zwei verschiedenen Teilchen des Systems gleichzeitig denselben Wert annimmt. Wählen wir für β insbesondere das System der Quantenzahlen, so gibt dieser Satz das Paulische Äquivalenzverbot in seiner ursprünglichen Fassung. Wir wollen im folgenden das gleichzeitige Bestehen dieses Satzes für alle Größen β als den eigentlichen Inhalt des Paulischen Äquivalenzverbots betrachten.

Wir beschäftigen uns übrigens im folgenden vorwiegend mit dem Falle, daß jede Größe β am Einzelatom nur endlich viele, sagen wir K, Eigenwerte hat. Nur gelegentlich werden wir näher hinweisen auf den Grenzübergang $K \to \infty$, der im allgemeinen keinerlei Schwierigkeit macht. Wir wollen die K Eigenwerte jeder Größe β numerieren mit $\beta'_1, \beta'_2, \ldots, \beta'_K$, und zwar so, daß die oben vorausgesetzte Anordnung der Eigenwerte gerade die Form

$$\beta'_1 < \beta'_1 < \cdots < \beta'_K \tag{21}$$

gewinnt.

§ 5. Die in solcher Weise definierten antisymmetrischen Amplituden sind nun in eindeutiger Weise darstellbar als Funktion von Argumenten

$$N'(\beta');\quad N'(q') \tag{22}$$

mit folgender Bedeutung: $N'(\beta')$ ist die Anzahl von Atomen, bei denen β den Wert β' hat; ist also β' ein diskreter Eigenwert, so ist nach dem allgemeinen Paulischen Äquivalenzverbot

$$N'(\beta') = 0 \text{ oder } 1. \tag{23}$$

Liegt dagegen β' in einem kontinuierlichen Eigenwertgebiet, so haben wir zu schreiben:

$$N'(\beta') = \sum_{k=1}^{N'} \delta(\beta' - \beta'_k), \tag{24}$$

wenn insgesamt N' Teilchen vorhanden sind; das Integral von $N'(\beta')$ über ein Teilstück des Eigenwertgebietes ist dann die Anzahl der Atome, bei denen die Werte von β in dieses Teilstück fallen.

Wir begnügen uns aber nicht mit der rein mathematischen Einführung der neuen Größen $N'(\beta')$, $N'(q')$, sondern gehen zu einer neuen physikalischen Theorie über, indem wir annehmen, das Gesamtgas sei ein System, das durch ein kanonisches System von q-Zahlgrößen

$$N(\beta'); \quad \Theta(\beta') \tag{25}$$

beschrieben werden kann, wobei die $N'(\beta')$ gerade die Eigenwerte von $N(\beta')$ darstellen. Dann sind die $N(\beta')$, $\Theta(\beta')$ in der in § 3 erläuterten Weise durch Matrizen darzustellen; den verschiedenen Eigenwerten β' entsprechen die verschiedenen Werte der in § 3 gebrauchten Indizes r, s. Insbesondere gilt für diskretes β' die Gleichung

$$N(\beta') \cdot [1 - N(\beta')] = 0, \tag{26}$$

für nicht diskretes β' kann man statt dessen schreiben:

$$N(\beta') \cdot [\delta(\beta' - \beta'') - N(\beta'')] = \begin{cases} N(\beta')\, N(\beta'') & \text{für} \quad \beta' \neq \beta'', \\ 0 & \text{für} \quad \beta' = \beta''. \end{cases} \tag{27}$$

Während nun die q-Zahlen $N(\beta')$ durch ihre physikalische Bedeutung völlig definiert sind, ist dasselbe natürlich nicht der Fall für die $\Theta(\beta')$, wenn wir von ihnen nur verlangen, daß sie kanonisch konjugiert zu den $N(\beta')$ seien. Man muß diese Nichteindeutigkeit natürlich beseitigen bzw. beschränken, wenn man eindeutige Relationen zwischen den Größen $N(\beta')$, $\Theta(\beta')$ und $N(q')$, $\Theta(q')$ erhalten will. Wir werden nun im weiteren Verlauf unserer Betrachtungen sehen: Man kann, nachdem für jede Größe β, q usw. in der oben besprochenen Weise eine Reihenfolge der Eigenwerte β' und q' usw. festgelegt ist, ein gewisses System von konjugierten Phasen $\Theta(\beta')$, $\Theta(q')$ usw. zu den $N(\beta')$, $N(q')$ usw. bestimmen derart, daß einfache und eindeutige Relationen zwischen den verschiedenen kanonischen Systemen $N(\beta')$, $\Theta(\beta')$; $N(q')$, $\Theta(q')$ usw. entstehen. Man hat dabei aber noch verschiedene Möglichkeiten für die Definition der $\Theta(\beta')$ zu den $N(\beta')$, und diese verschiedenen Möglichkeiten können eindeutig zugeordnet werden den verschiedenen konjugierten Impulsen α zu β. Wir bezeichnen

deshalb die q-Zahlgrößen, deren Theorie wir im folgenden entwickeln wollen, endgültig mit

$$N(\beta'); \quad \Theta_\alpha(\beta') \tag{28}$$

bzw.

$$N(q'); \quad \Theta_p(q') \text{ usf.} \tag{29}$$

Die gebildeten Verhältnisse besitzen offenbar die denkbar größte Analogie zu den in B erörterten Verhältnissen im Boseschen Falle, soweit man überhaupt angesichts der tiefgehenden Verschiedenheit beider Fälle eine Analogie erwarten kann.

§ 6. Die zwei K Größen

$$N(\beta'), \quad \Theta_\alpha(\beta')$$

müssen als q-Zahlen gewisse Funktionen der q-Zahlen

$$N(q'), \quad \Theta_p(q')$$

sein; dieser funktionale Zusammenhang soll jetzt besprochen werden. Im Bose-Einsteinschen Falle galt einfach

$$\left.\begin{aligned} b_\alpha(\beta') &= \sum_{q'} \Phi_{\alpha p}(\beta', q')\, b_p(q'), \\ b_\alpha^\dagger(\beta') &= \sum_{q'} b_p^\dagger(q')\, \Phi_{p\alpha}(q', \beta'); \end{aligned}\right\} \tag{30}$$

aber diese Formeln gelten nicht für das Paulische Gas. Statt dessen gelten Formeln

$$\left.\begin{aligned} a_\alpha(\beta') &= \sum_{q'} \Phi_{\alpha p}(\beta', q')\, a_p(q'), \\ a_\alpha^\dagger(\beta') &= \sum_{q'} a_p^\dagger(q')\, \Phi_{p\alpha}(q', \beta'), \end{aligned}\right\} \tag{30a}$$

wenn wir die Größen a, $a^\dagger$ durch

$$\left.\begin{aligned} a_p(q') &= v(q') \cdot b_p(q'), \\ a_p^\dagger(q') &= b_p^\dagger(q') \cdot v(q'); \end{aligned}\right\} \tag{31}$$

$$v(q') = \prod_{q'' \leqq q'} \{1 - 2\,N(q'')\} \tag{32}$$

definieren. Hier ist also $v(q')$ das Produkt der Größen $1 - 2\,N(q'')$ für $q'' = q'$ und alle vor q' kommenden q''. Es ist also $v(q')$ eine Diagonalmatrix, deren Diagonalelemente sämtlich gleich $+1$ oder -1 sind; und es wird

$$[v(q')]^2 = 1. \tag{33}$$

Der vollständige mathematische Beweis für die Richtigkeit dieser Formeln (30a) wird sich in den §§ 8 und 9 ergeben. Hier wollen wir lediglich die Multiplikationseigenschaften der Größen a, $a^\dagger$ untersuchen

und die Invarianz dieser Multiplikationseigenschaften gegen Transformationen (30a) nachweisen.

Zunächst wird

$$\left.\begin{aligned} b_p(q') \cdot v(q'') &= \begin{cases} -v(q'') \cdot b(q') & \text{für} \quad q' \leqq q'', \\ v(q'') \cdot b(q') & \text{für} \quad q' > q''; \end{cases} \\ b_p^\dagger(q') \cdot v(q'') &= \begin{cases} -v(q'') \cdot b^\dagger(q') & \text{für} \quad q' \leqq q'', \\ v(q'') \cdot b^\dagger(q') & \text{für} \quad q' > q''. \end{cases} \end{aligned}\right\} \tag{34}$$

Der Beweis ergibt sich leicht daraus, daß z. B.

$$\begin{aligned} b_p^\dagger(q') \cdot \{1 - 2N(q')\} &= N(q') e^{\frac{2\pi i}{h} \Theta_p(q')} \{1 - 2N(q')\} \\ &= -\{1 - 2N(q')\} b_p^\dagger(q') \end{aligned} \tag{35}$$

wird. Dann wird weiter

$$\left.\begin{aligned} a_p(q') a_p(q'') &= -a_p(q'') a_p(q'), \\ a_p^\dagger(q') a_p^\dagger(q'') &= -a_p^\dagger(q'') a_p^\dagger(q'). \end{aligned}\right\} \tag{36}$$

Man beweist z. B.

$$\begin{aligned} a_p(q') a_p(q'') &= v(q') b_p(q') v(q'') b_p(q'') \\ &= \begin{cases} -v(q') v(q'') b_p(q') b_p(q'') & \text{für} \quad q' \leqq q'', \\ v(q') v(q'') b_p(q') b_p(q'') & \text{für} \quad q' > q''; \end{cases} \end{aligned} \tag{37}$$

nun ist aber insbesondere für $q' = q''$:

$$b_p(q') b_p(q'') = [b_p(q')]^2 = 0, \tag{38}$$

also auch

$$[a_p(q')]^2 = 0, \tag{39}$$

womit für $q' = q''$ die Formel (36) schon bewiesen ist. Man sieht danach aus (37), daß das Produkt (36) in der Tat antisymmetrisch in q', q'' ist.

Ferner gilt

$$a_p^\dagger(q') a_p(q'') + a_p(q'') a_p^\dagger(q') = \delta(q' - q''). \tag{40}$$

Denn es wird der links stehende Ausdruck gleich

$$\left.\begin{aligned} & b_p^\dagger(q') v(q') v(q'') b_p(q'') + v(q'') b_p(q'') b_p^\dagger(q') v(q') \\ = & \begin{cases} v(q') v(q'') b_p^\dagger(q') b_p(q'') - v(q'') v(q') b_p(q'') b_p^\dagger(q') & \text{für } q' < q''; \\ v(q') v(q'') b_p^\dagger(q') b_p(q'') + v(q'') v(q') b_p(q'') b_p^\dagger(q') & \text{für } q' = q''; \\ -v(q') v(q'') b_p^\dagger(q') b_p(q'') + v(q'') v(q') b_p(q'') b_p^\dagger(q') & \text{für } q' > q''; \end{cases} \end{aligned}\right\} \tag{41}$$

also verschwindet er in der Tat für $q' \neq q''$ und wird für $q' = q''$ gleich

$$[v(q')]^2 \cdot [b_p^\dagger(q') b_p(q') + b_p(q') b_p^\dagger(q')] = 1. \tag{42}$$

Nunmehr zeigen wir, daß die Gleichungen (36) und (40) wirklich invariant sind gegen Transformationen (30a). Es ergibt sich

$$a_\alpha(\beta')\,a_\alpha(\beta'') + a_\alpha(\beta'')\,a_\alpha(\beta')$$
$$= \sum_{q'q''} \Phi_{\alpha p}(\beta', q')\,\Phi_{\alpha p}(\beta'', q'')\cdot\{a_p(q')\,a_p(q'') + a_p(q'')\,a_p(q')\} = 0; \quad (43)$$

$$a_\alpha^\dagger(\beta')\,a_\alpha(\beta'') + a_\alpha(\beta'')\,a_\alpha^\dagger(\beta')$$
$$= \sum_{q'q''} \Phi_{p\alpha}(q', \beta')\,\Phi_{\alpha p}(\beta'', q'')\cdot\{a_p^\dagger(q')\,a_p(q'') + a_p(q'')\,a_p^\dagger(q')\} = \delta(\beta' - \beta''). \quad (44)$$

Während also die Größen b, $b^\dagger$ im Paulischen Falle ebenso wie im Einsteinschen Falle die Eigenschaft haben, daß $b(\beta')$ mit $b(\beta'')$ und $b^\dagger(\beta')$ für $\beta'' \neq \beta'$ vertauschbar ist, kommt den a, $a^\dagger$ diese Eigenschaft nicht mehr zu. Trotzdem besitzen die a, $a^\dagger$ des Fermischen Gases in gewisser Hinsicht eine engere Analogie zu den b, $b^\dagger$ des Einsteingases, als die b, $b^\dagger$ des Pauligases selbst; man sieht das besonders deutlich durch die Gegenüberstellung:

Bose-Einstein	Pauli
$b_\alpha(\beta')\,b_\alpha(\beta'') - b_\alpha(\beta'')\,b_\alpha(\beta') = 0;$	$a_\alpha(\beta')\,a_\alpha(\beta'') + a_\alpha(\beta'')\,a_\alpha(\beta') = 0;$
$b_\alpha^\dagger(\beta')\,b_\alpha(\beta'') - b_\alpha(\beta'')\,b_\alpha^\dagger(\beta') = \delta(\beta' - \beta'');$	$a_\alpha^\dagger(\beta')\,a_\alpha(\beta'') + a_\alpha(\beta'')\,a_\alpha^\dagger(\beta') = \delta(\beta' - \beta'');$
$b_\alpha^\dagger(\beta')\,b_\alpha(\beta') = N(\beta');$	$a_\alpha^\dagger(\beta')\,a_\alpha(\beta') = N(\beta');$
$b_\alpha(\beta') = \sum_{q'} \Phi_{\alpha p}(\beta', q')\,b_p(q');$	$a_\alpha(\beta') = \sum_{q'} \Phi_{\alpha p}(\beta', q')\,a_p(q').$

Wir haben diese Gleichungen abgeleitet, indem wir das Paulische Äquivalenzverbot von vornherein zugrunde legten. Es zeigt sich aber, daß umgekehrt diese Multiplikationseigenschaften der a, $a^\dagger$ bereits die möglichen Eigenwerte der $N(\beta')$ bestimmen und die Vertauschbarkeit (gleichzeitige Beobachtbarkeit) von $N(\beta')$ und $N(\beta'')$ nach sich ziehen. Infolgedessen können wir sagen, daß die Existenz korpuskularer Teilchen und die Gültigkeit des Paulischen Prinzips als eine Folgerung aus den quantenmechanischen Multiplikationseigenschaften der de Broglieschen Wellenamplituden aufgefaßt werden dürfen, da in den beiden Gleichungen

$$N(\beta')\,N(\beta'') - N(\beta'')\,N(\beta') = 0, \quad (45)$$
$$N'(\beta') = 0 \text{ oder } 1 \quad (46)$$

diese Tatsachen vollständig ausgedrückt sind. Die Gleichung (45) folgt sofort. Der Beweis, daß auch (46) aus den Multiplikationsregeln der $a(\beta')$, $a^\dagger(\beta')$ folgt, ergibt sich folgendermaßen:

Auf Grund von

$$a_\alpha^\dagger(\beta')\,a_\alpha(\beta') + a_\alpha(\beta')\,a_\alpha^\dagger(\beta') = 1 \tag{47}$$

gilt wegen

$$[a_\alpha(\beta')]^2 = 0$$

die Gleichung

$$a_\alpha^\dagger(\beta')\,a_\alpha(\beta')\,a_\alpha^\dagger(\beta')\,a_\alpha(\beta')$$
$$= a_\alpha^\dagger(\beta')\,.\,[1 - a_\alpha^\dagger(\beta')\,a_\alpha(\beta')]\,.\,a_\alpha(\beta') = a_\alpha^\dagger(\beta')\,a_\alpha(\beta')\,; \tag{48}$$

also wird

$$N(\beta')\,.\,[1 - N(\beta')] = a_\alpha^\dagger(\beta')\,a_\alpha(\beta')\,.\,[1 - a_\alpha^\dagger(\beta')\,a_\alpha(\beta')] = 0. \tag{49}$$

Es sei noch betont: Da $v(\beta')$ aus den $N(\beta'')$ allein zu bilden ist (nachdem eine Reihenfolge der Eigenwerte festgelegt wurde), so kann man vermittelst

$$\left.\begin{aligned} b_\alpha(\beta') &= v(\beta')\,a_\alpha(\beta'), \\ b_\alpha^\dagger(\beta') &= a_\alpha^\dagger(\beta')\,v(\beta') \end{aligned}\right\} \tag{50}$$

die b, $b^\dagger$ eindeutig durch die a, $a^\dagger$ definieren. Man kann also in der Tat die a, $a^\dagger$ als die ursprünglichen Größen der Theorie und alle anderen Größen als Funktionen der a, $a^\dagger$ betrachten.

Endlich sei hervorgehoben, daß die Gesamtzahl N der vorhandenen Teilchen gegenüber den betrachteten Transformationen invariant bleibt:

$$N = \sum_{\beta'} N(\beta') = \sum_{q'} N(q'). \tag{51}$$

Diese Invarianz ist offenbar nur ein anderer Ausdruck dafür, daß (30a) eine unitäre Transformation ist.

Zusatz bei der Korrektur. Es zeigt sich durch eine genauere Betrachtung, die im Anhang mitgeteilt wird, daß die Multiplikationsregeln der a, $a^\dagger$ nicht nur die Eigenwerte der $N(\beta')$ schon bestimmen, sondern überhaupt die Matrizen a, $a^\dagger$ bis auf eine kanonische Transformation der Matrizendarstellung festlegen.

§ 7. Für ein eindimensionales Kontinuum mit der Schwingungsgleichung

$$\frac{\partial^2 \psi}{\partial x^2} = \frac{\partial^2 \psi}{\partial t^2}; \qquad \psi = \psi(x, t) \tag{52}$$

und der Randbedingung

$$\psi(0, t) = \psi(l, t) = 0 \tag{53}$$

war in A versuchsweise die räumliche Teilchendichte der Wellenkorpuskeln definiert durch

$$N(x) = \psi^\dagger \psi, \tag{54}$$

$$\psi = \sum_{r=1}^{\infty} b_r \sin r \frac{\pi}{l} x, \tag{55}$$

wo $N_r = b_r^\dagger b_r$ die Anzahl von Teilchen im r-ten Quantenzustand der Translation bedeutet.

Wir haben jetzt (55) zu korrigieren, indem wir die b_r durch entsprechende a_r ersetzen:

$$\psi = \sum_{r=1}^{\infty} a_r \sin r \frac{\pi}{l} x. \tag{56}$$

Die in A durchgeführte Berechnung der Dichteschwankungen kann aber sofort von (55) auf (56) übertragen werden und zeigt dann, daß (56), wie es sein muß, wirklich die richtige Formel liefert. Es wurde nämlich aus (55) erhalten, daß der fragliche quadratische Mittelwert $\overline{\Delta^2}$ proportional sei mit

$$\overline{b_r^\dagger b_r} \,.\, \overline{b_r b_r^\dagger} = \overline{N_r} \,.\, \overline{(1 - N_r)}, \tag{57}$$

wo die Querstriche die Mittelung über ein infinitesimales Frequenzgebiet im Anschluß an die Frequenz ν_r bedeuten, so daß sich die in § 2 erwähnte Formel

$$\overline{\Delta^2} = \text{const} \,.\, n_r (1 - n_r) \tag{58}$$

ergibt. Rechnet man nun entsprechend mit (56), so wird $\overline{\Delta^2}$ proportional mit

$$\overline{a_r^\dagger a_r} \,.\, \overline{a_r a_r^\dagger} = \overline{N_r} \,.\, \overline{(1 - N_r)}, \tag{59}$$

d. h. das Endergebnis bleibt ungeändert.

§ 8. Wir wollen nun den in § 6 angekündigten vollständigen Beweis für die Äquivalenz der Formeln (30 a) mit den Formeln der gewöhnlichen Darstellung im mehrdimensionalen Koordinatenraum antreten. Wir müssen uns in diesem Koordinatenraum auf solche Größen (Operatoren) beschränken, die symmetrisch in den gleichen Teilchen sind; außerdem aber beschränken wir uns in diesem Paragraphen auf Größen, die aus einer Summe bestehen, wo in jedem Summand nur ein Elektron vorkommt. Von dieser Art ist die Energie eines idealen Gases. Diese Operatoren haben also die Gestalt

$$V = V_1 + V_2 + \cdots + V_N, \tag{60}$$

wo die V_i immer dieselbe Größe darstellen, nur an verschiedenen (am 1-, 2-, ..., N-ten) Teilchen gemessen.

Unsere Wellenfunktion dagegen wird [§ 5, Gleichung (22), (23)] von den $N'(\beta'_k)$ abhängen. Dies erscheint in der Tat vom Standpunkt der Quantenmechanik als der naturgemäße Ansatz, da ja ein „maximaler Versuch" — wegen der Gleichheit der Teilchen — immer nur bestimmen kann, wie viele Teilchen im Zustand $\beta_1, \beta_2, \ldots, \beta_K$ sind, während die Frage, in welchem Zustand ein bestimmtes Elektron ist, nicht entschieden werden kann. Im Sinne des Paulischen Äquivalenzverbotes haben wir in (23) den Wertbereich der $N'(\beta'_k)$ auf 0,1 beschränkt.

Wir nehmen nun noch, wie in § 4 bereits betont, der Einfachheit halber an (was in Wahrheit niemals erfüllt ist), daß ein Elektron nur endlich viele (sagen wir K) Zustände annehmen kann, die wir also mit $\beta_1, \beta_2, \ldots, \beta_K$ bezeichnen. Dann hat die im folgenden einzuführende Wellenfunktion $\Psi(N'(\beta'_1), N'(\beta'_2), \ldots, N'(\beta'_K))$ gerade K Argumente und ist für 2^K Wertsysteme der Argumente definiert. Der Grenzübergang $K \to \infty$ scheint keine wesentlichen Schwierigkeiten zu bieten.

Die nun folgenden Betrachtungen lassen sich am einfachsten ausführen, wenn man den Zustand eines einzelnen Elektrons im mehrdimensionalen Koordinatenraum mit einer Wellenfunktion beschreibt, deren Argument β' ist. Das bedeutet, daß beim Einzelelektron die Messung eben die Bestimmung der Größe β ist, deren Wertbereich also K Zahlen $\beta'_1, \beta'_2, \ldots, \beta'_K$ umfaßt.

Haben wir nun im mehrdimensionalen Koordinatenraum eine antisymmetrische Wellenfunktion von N' Elektronen

$$\psi(\beta'^1, \beta'^2, \ldots, \beta'^{N'}), \tag{61}$$

so bestimmen wir, daß wir diesen Zustand fortan in unserem neuen N-Raume durch die Wellenfunktion $\Psi(N'(\beta'_1), N'(\beta'_2), \ldots, N'(\beta'_K))$ beschreiben wollen. Dabei sei

$$\Psi(N'(\beta'_1), \ldots, N'(\beta'_K)) = \frac{1}{\sqrt{N'!}}\psi(\beta'^1, \ldots, \beta'^{N'}). \tag{62}$$

Diese Gleichung ist so zu verstehen, daß Ψ überall 0 ist, wo nicht genau N von den $K (K > N)$ Zahlen $N'(\beta'_1), \ldots, N'(\beta'_K)$ gleich 1, die übrigen gleich 0 sind. Um den Wert an diesen Stellen zu bestimmen, setzt man rechts für die $\beta'^1, \beta'^2, \ldots, \beta'^{N'}$ jene Werte ein, für die eben $N'(\beta'_i) = 1$ ist, und zwar für β'^1 den (im Sinne der in § 4 getroffenen Anordnung) ersten, für β'^2 den folgenden, ..., für $\beta'^{N'}$ den letzten.

Wir wollen an dieser Zuordnung einer Funktion

$$\Psi(N'(\beta'_1), N'(\beta'_2), \ldots, N'(\beta'_K))$$

im neuen N-Raume zu einer Funktion

$$\psi(\beta'^1, \beta'^2, \ldots, \beta'^{N'})$$

im mehrdimensionalen Koordinatenraume im folgenden immer (also auch, wenn ψ keine Wellenfunktion ist) festhalten. Es ist zu beachten, daß es für das Vorzeichen von Ψ wichtig ist, in ψ die $\beta'_1, \beta'_2, \ldots. \beta'_K$ in der einmal festgesetzten Reihenfolge

$$(\beta'_1 < \beta'_2 < \cdots < \beta'_K)$$

einzusetzen, da nur dadurch das Vorzeichen von Ψ eindeutig geregelt wird; und nur hierdurch ist die Zuordnung einer eindeutigen Funktion Ψ zu der Funktion ψ möglich.

Umgekehrt wird aber auch ψ durch Ψ eindeutig bestimmt*: an den Stellen, für welche $\beta'^1 < \cdots < \beta'^{N'}$ gilt, durch (62), überall sonst durch die Forderung der Antisymmetrie.

Die einzelnen Teile (z. B. V_1) des Operators V [Gleichung (60)] im mehrdimensionalen Koordinatenraum sind im wesentlichen hermiteische Matrizen von K Zeilen und K Spalten. In der ν-ten Zeile und μ-ten Spalte stehe $H_{\nu\mu}$. Dann ist der ganze Operator V mit der Matrix

$$H_{\nu_1 \nu_2 \ldots \nu_{N'};\, \mu_1 \mu_2 \ldots \mu_{N'}} = H_{\nu_1 \mu_1} \delta_{\nu_2 \mu_2} \delta_{\nu_3 \mu_3} \cdots \delta_{\nu_{N'} \mu_{N'}}$$
$$+ \delta_{\nu_1 \mu_1} H_{\nu_2 \mu_2} \delta_{\nu_3 \mu_3} \cdots \delta_{\nu_{N'} \mu_{N'}} + \cdots + \delta_{\nu_1 \mu_1} \delta_{\nu_2 \mu_2} \delta_{\nu_3 \mu_3} \cdots H_{\nu_{N'} \mu_{N'}} \tag{63}$$

identisch. Wir schreiben zur Abkürzung

$$V \psi(\beta'^1 \ldots \beta'^{N'}) = \overline{\psi}(\beta'^1 \ldots \beta'^{N'}), \tag{64}$$

dann ist also

$$\overline{\psi}(\beta'_{\nu_1} \cdots \beta'_{\nu_{N'}}) = \sum_{\mu_1, \mu_2, \ldots, \mu_{N'} = 1}^{K} H_{\nu_1 \ldots \nu_{N'};\, \mu_1 \ldots \mu_{N'}}\, \psi(\beta'_{\mu_1} \cdots \beta'_{\mu_{N'}}). \tag{65}$$

Aus diesem $\overline{\psi}(\beta'^1, \beta'^2, \ldots, \beta'^{N'})$ bilden wir — genau wie in (62) — ein $\overline{\Psi}(N'(\beta'_1), \ldots, N'(\beta'_K))$ durch

$$\overline{\Psi}(N'(\beta'_1), \ldots, N'(\beta'_K)) = \overline{\psi}(\beta'^1, \ldots, \beta'^{N'}) \frac{1}{\sqrt{N'!}}. \tag{62a}$$

Wir behaupten nun, daß

$$\overline{\Psi}(N'(\beta'_1), \ldots, N'(\beta'_K)) = \Omega\, \Psi(N'(\beta'_1), \ldots, N'(\beta'_K)), \tag{66}$$

wo der Operator Ω

$$\Omega = \sum_{\varkappa, \lambda = 1}^{K} H_{\varkappa\lambda}\, a_\varkappa^\dagger a_\lambda \tag{66a}$$

ist, mit den a aus (31).

---

* Wir setzen voraus, daß Ψ wieder überall verschwindet, wo nicht

$$N'(\beta'_1) + \cdots + N'(\beta'_K) = N'$$

ist.

Gleichung (65) ist zunächst sicher richtig an allen Stellen (für alle Wertsysteme der Argumente), wo die Anzahl der 1 im Argumentsystem von $\overline{\Psi}$ nicht eben gleich N' ist. Dann verschwindet nämlich die linke Seite wegen (62), (62a), und auch auf der rechten Seite stehen lauter Nullen, da $a_\varkappa^\dagger a_\lambda$ die Anzahl der 1 nicht ändert.

An den Stellen aber, wo etwa die

$$N'(\beta'_{i_1}),\ N'(\beta'_{i_2}),\ \ldots,\ N'(\beta'_{i_{N'}})$$

(also genau N') gleich 1 sind, ist $\sqrt{N'!}\ \overline{\Psi}$ gleich

$$\overline{\psi}(\beta'_{i_1},\ \ldots,\ \beta'_{i_{N'}}),$$

also gleich

$$\sum_{\mu_1 \ldots \mu_{N'}=1}^{K} H_{i_1 \ldots i_{N'};\, \mu_1 \ldots \mu_{N'}}\, \psi(\beta'_{\mu_1},\ \ldots,\ \beta'_{\mu_{N'}})$$

$$= \sum_{\mu_1=1}^{K} H_{i_1 \mu_1}\, \psi(\beta'_{\mu_1},\ \beta'_{i_2},\ \ldots,\ \beta'_{i_{N'}}) + \sum_{\mu_2=1}^{K} H_{i_2 \mu_2}\, \psi(\beta'_{i_1},\ \beta'_{\mu_2},\ \ldots,\ \beta'_{i_{N'}}) +$$

$$\cdots + \sum_{\mu_{N'}=1}^{K} H_{i_{N'} \mu_{N'}}\, \psi(\beta'_{i_1},\ \beta'_{i_2},\ \ldots,\ \beta'_{\mu_{N'}})$$

$$= \sum_{r=1}^{N'} \sum_{\mu=1}^{K} H_{i_r \mu}\, \psi(\beta'_{i_1},\ \beta'_{i_2},\ \ldots,\ \beta'_{i_{r-1}},\ \beta'_\mu,\ \beta'_{i_{r+1}},\ \ldots,\ \beta'_{i_{N'}}), \quad (67)$$

das zweite ersieht man aus (63). Unsere Absicht ist nun, die rechte Seite von (67) durch die Ψ auszudrücken. Zu diesem Zwecke bemerken wir, daß in (67) rechts die β'_i schon in der richtigen Reihenfolge stehen, nur das jeweils auftretende β'_μ ist an der falschen Stelle. Ist etwa*

$$\beta'_{i_{s-1}} < \beta'_\mu \leqq \beta'_{i_s},$$

so können wir die Reihenfolge der β' zur richtigen machen, indem wir β'_μ über die zwischenliegenden β'_i von der Stelle zwischen $\beta'_{i_{r-1}}$ und $\beta'_{i_{r+1}}$ an die Stelle zwischen $\beta'_{i_{s-1}}$ und β'_{i_s} verschieben. Dabei multipliziert sich die antisymmetrische Funktion ψ mit $(-1)^z$, wo z die Anzahl der zwischen den beiden angegebenen Stellen stehenden β_i ist.

Wir können (67) also auch schreiben

$$\overline{\psi}(\beta'_{i_1},\ \ldots,\ \beta'_{i_{N'}}) = \sum_{r=1}^{N} \sum_{\mu=1}^{K} \pm H_{i_r \mu}\, \psi(\beta'_{i_1},\ \ldots,\ \beta'_\mu,\ \ldots,\ \beta'_{i_{N'}}), \quad (67a)$$

* Das Gleichheitszeichen kommt nicht in Frage, da dann das entsprechende ψ doch verschwindet.

wo natürlich noch das Vorzeichen $\pm$ von r und μ abhängt, aber schon

$$\beta'_{i_1} < \cdots < \beta'_\mu < \cdots < \beta'_{N'}$$

gilt. Wenn wir (62) beachten, können wir dies auch*

$$\overline{\Psi}(x_1 \ldots x_K) = \sum_{x_j=1} \sum_{x_l=0} \pm H_{jl}\, \Psi(x_1, \ldots, x_{j-1}, 0, x_{j+1}, \ldots, x_{l-1}, 1, x_l, \ldots, x_K)$$
$$+ \sum_{x_j=1} H_{jj}\, \Psi(x_1, \ldots, x_K) \tag{68}$$

schreiben, wie man leicht überlegt. In den Summanden nämlich, in (67a), wo $i_r \neq \mu$ ist [erstes Glied in (68)], sind dieselben Argumente vorhanden wie an der linken Seite, es fehlt nur β'_{i_r}, was links vorhanden war ($x_j = 1$), dagegen ist β'_μ hinzugekommen, und wir können annehmen, daß es nicht da war ($x_l = 0$), da sonst ψ doch verschwinden würde. Ist $i_r = \mu$ [zweites Glied in (68)], so sind rechts in ψ dieselben Argumente wie links.

Das Vorzeichen in (68) bestimmt sich offenbar dadurch, daß man $+$ oder $-$ setzt, je nachdem zwischen der ausgeschriebenen 0 und 1 eine gerade oder ungerade Anzahl von 1 steht. (Über ebenso viele β'_i mußte man das entsprechende β'_μ hinüberschieben.) Dies ist aber die Anzahl der 1, die links von x_l stehen, vermindert um die Anzahl der 1, die links von x_j stehen.

Obwohl jetzt die Richtigkeit der vorangehenden Formeln schon klar ist, wollen wir diese Gedanken doch zu Ende führen. Wir wissen, daß der Wertbereich der Argumente von Ψ insgesamt 2^K Stellen umfaßt, indem für jedes der $x_k = N'(\beta'_k)$ entweder $+1$ oder 0 gesetzt werden kann. Ein darauf wirkender linearer Operator ist also eine Matrix mit 2^K Zeilen und ebensoviel Spalten. Wir bezeichnen jede Zeile oder Spalte mit K Indizes (den $x_1, \ldots, x_K$ entsprechend), die jeweils 1 oder 0 sein können.

Der Operator a_λ ist entsprechend § 3 und § 6 zu definieren (wir schreiben der besseren Übersicht halber die Indizes als Argumente) durch

$$a_\lambda(x_1, x_2, \ldots, x_K;\ y_1, y_2, \ldots, y_K)$$
$$= (-1)^{x_1 + x_2 + \cdots + x_{\lambda-1}} \delta_{x_1 y_1} \delta_{x_2 y_2} \cdots \delta_{x_{\lambda-1} y_{\lambda-1}} \delta_{x_\lambda 0} \delta_{y_\lambda 1} \delta_{x_{\lambda+1} y_{\lambda+1}} \cdots \delta_{x_K y_K}, \tag{69}$$

und entsprechend ist dann $a^\dagger_\varkappa$

$$a^\dagger_\varkappa(x_1, \ldots, x_K;\ y_1, \ldots, y_K)$$
$$= (-1)^{x_1 + \cdots + x_{\varkappa-1}} \delta_{x_1 y_2} \cdots \delta_{x_{\varkappa-1} y_{\varkappa-1}} \delta_{x_\varkappa 1} \delta_{y_\varkappa 0} \cdots \delta_{x_K y_K}. \tag{69a}$$

* Wir setzen aus Bequemlichkeitsgründen x_k für $N'(\beta'_k)$.

Mit Hilfe dieser Formeln kann (68) auch so geschrieben werden:

$$\overline{\Psi}(x_1, \ldots, x_K)$$
$$\sum_{\lambda, \varkappa=1}^{K} H_{\varkappa\lambda} \sum_{\substack{y_1 \ldots y_K = 0,1 \\ z_1 \ldots z_K = 0,1}} a_\varkappa^\dagger (x_1 \ldots x_K; y_1 \ldots y_K) \cdot a_\lambda (y_1 \ldots y_K; z_1 \ldots z_K) \Psi (z_1 \ldots z_K). \quad (70)$$

Wie man sich mit einiger Mühe überlegen kann, was aber schwer hingeschrieben werden kann. Damit ist (66) gewonnen.

§ 9. Wir haben also in § 8 folgendes gesehen: Jeder antisymmetrischen Funktion, welche definiert ist in den Koordinatenräumen mit allen Anzahlen $N' < K$ von Dimensionen, entspricht durch (62) eine Funktion im neuen Raume. Es entspricht dann dem Operator (60) $V = V_1 + V_2 + \cdots + V_N$ [mit der Matrix in (64) $H_{r_1 \ldots r_{N'}; \mu_1 \ldots \mu_{N'}}$] im Koordinatenraum der Operator Ω von (68) im neuen N'-Raum. Der Operator Ω ist dabei

$$\Omega = \sum_{\varkappa, \lambda=1}^{K} H_{\varkappa\lambda} a_\varkappa^\dagger a_\lambda \quad (66\text{a})$$

mit den $a_\varkappa^\dagger$, a_λ von (69), (69a).

Es folgt hieraus, daß einer Eigenfunktion von V eine Eigenfunktion von Ω entspricht. Wenn wir noch zeigen können, daß das innere Produkt zweier Funktionen im Koordinatenraum denselben Wert hat wie dasjenige der entsprechenden Funktionen im neuen N'-Raum, so sind wir mit dem Beweis fertig. Im Koordinatenraum ist

$$(\psi \varphi) = \sum_{\beta'^1, \ldots, \beta'^{N'} = \beta_1'}^{\beta_K'} \psi (\beta'^1 \ldots \beta'^{N'}) \widetilde{\varphi} (\beta'^1 \ldots \beta'^{N'}), \quad (71)$$

was wegen der Antisymmetrie

$$(\psi \varphi) = N'! \cdot \sum_{\substack{\beta'^1, \ldots, \beta'^{N'} = \beta_1' \\ \beta'^1 < \ < \beta'^{N'}}}^{\beta_K'} \psi (\beta'^1 \ldots \beta'^{N'}) \widetilde{\varphi} (\beta'^1 \ldots \beta'^{N'}) \quad (72)$$

ergibt. Andererseits gilt im neuen Raume

$$(\Psi \Phi) = \sum_{x_1 \ldots x_K = 0,1} \Psi (x_1 \ldots x_K) \widetilde{\Phi} (x_1 \ldots x_K), \quad (73)$$

was mit Rücksicht auf (62) eben (72) ist.

Wir möchten noch bemerken, daß die q-Zahlrelationen (36), (40) natürlich sofort aus der Formel (69) hervorgehen.

§ 10. Wir müssen schließlich Operatoren betrachten, die keine Zerlegung mehr in die Gestalt (60) gestatten. Von dieser Form ist die

Energie eines nichtidealen Gases. Wir beschränken uns dabei zunächst auf solche, die in Teile zerlegbar sind, die jeweils nur zwei, stets verschiedene, Teilchen enthalten. Die Erledigung dieser Aufgabe ergibt sich durch Analogisierung der entsprechenden Formeln für die Bose-Einsteinsche Statistik*. Der Operator V läßt sich dann schreiben:

$$V = \sum_{\substack{j,\,k=1\\ j<k}}^{N'} V_{jk}, \tag{60b}$$

wobei V die Matrix

$$H_{\nu_1 \ldots \nu_{N'};\ \mu_1 \ldots \mu_{N'}}$$
$$= \sum_{\substack{j,\,k=1\\ j<k}}^{N'} H_{\nu_j \nu_k;\ \mu_j \mu_k} \delta_{\nu_1 \mu_1} \cdots \delta_{\nu_{j-1} \mu_{j-1}} \delta_{\nu_{j+1} \mu_{j+1}} \cdots \delta_{\nu_{k-1} \mu_{k-1}} \delta_{\nu_{k+1} \mu_{k+1}} \cdots \delta_{\nu_{N'} \mu_{N'}} \tag{63b}$$

entspricht. Es ist dann

$$V \psi(\beta'^1, \ldots, \beta'^{N'}) = \overline{\psi}(\beta'^1, \ldots, \beta'^{N'}) \tag{64b}$$

mit Hilfe von (63b) [ebenso wie (65)] zu berechnen. Dann ist wieder die „richtige Reihenfolge" der β auf der rechten Seite herzustellen. Dabei können die β'_i stehengelassen bleiben, β'_μ und $\beta'_{\mu'}$ müssen über eine Anzahl von β'_i hinübergeschoben werden, wobei sich wieder das Vorzeichen ändern kann.

Beachten wir wieder (62), so können wir wieder links und rechts für $\overline{\psi}$ bzw. ψ einsetzen $\overline{\Psi}$ bzw. Ψ. Unter Beachtung von (69) findet man nun, daß dem Operator (60b) nunmehr der Operator

$$\frac{1}{2!} \sum_{\substack{\lambda_1,\,\lambda_2=1\\ \varkappa_1,\,\varkappa_2=1}}^{K} H_{\varkappa_1 \varkappa_2;\ \lambda_1 \lambda_2}\, a^\dagger_{\varkappa_1} a^\dagger_{\varkappa_2} a_{\lambda_2} a_{\lambda_1} \tag{66b}$$

entspricht.

Es ist befriedigend, daß die Rückwirkung der Teilchen auf sich selbst wieder durch die nichtkommutativen Multiplikationseigenschaften der Wellenamplituden im dreidimensionalen Raume automatisch ausgeschlossen wird. Im Bose-Einsteinschen Falle wurde dieser Umstand deutlich gemacht durch Formel (40) der Arbeit von Jordan und Klein. Daß dieselbe Formel auch hier gilt, folgt aus der leicht beweisbaren Formel

$$a^\dagger_k a_k a^\dagger_l a_l - a^\dagger_k a^\dagger_l a_l a_k = \delta_{kl} a^\dagger_k a_k.$$

Wir wenden uns endlich zu dem Fall von Operatoren, die aus Summanden bestehen, welche jeweils in $n > 2$ Teilchen symmetrisch sind,

* P. Jordan und O. Klein, a. a. O.

während in der Summe alle diejenigen Summanden auszulassen sind, welche dasselbe Teilchen zweimal enthalten (was eine Wechselwirkung des Teilchens mit sich selbst bedeuten würde). Eine Verallgemeinerung der obigen Betrachtungen führt dann zu den folgenden Formeln, welche analog sind der in B angegebenen Verallgemeinerung* der Formel von Jordan und Klein:

$$\frac{1}{n!} \sum_{\lambda_1 \ldots \lambda_n = 1}^{K} \sum_{\varkappa_1 \ldots \varkappa_2 = 1}^{K} H_{\varkappa_1 \ldots \varkappa_n;\, \lambda_1 \ldots \lambda_n}\, a^\dagger_{\varkappa_1} \ldots a^\dagger_{\varkappa_n}\, a_{\lambda_n} \ldots a_{\lambda_1}. \tag{66 c}$$

§ 11. Wir können endlich die erhaltenen Ergebnisse in einer etwas anderen Form so aussprechen: Es gibt bei einem Paulischen Mehrkörperproblem eine Wahrscheinlichkeitsamplitude, welche die Wahrscheinlichkeit dafür bestimmt, daß, nachdem für die meßbaren Größen $N(\beta'_1), N(\beta'_2), \ldots, N(\beta'_K)$ ein gewisses Wertsystem $N'(\beta'_1), N'(\beta'_2), \ldots, N'(\beta'_K)$ gemessen worden ist, für die anderen, entsprechend definierten Größen $N(q'_1), N(q'_2), \ldots, N(q'_K)$ die Werte $N'(q'_1), N'(q'_2), \ldots, N'(q'_K)$ gefunden werden.

Eine solche Amplitude wird erhalten durch

$$\Phi^{N(\beta'), N(q')}_{-\theta_\alpha(\beta'), -\theta_p(\gamma')} = 0 \quad \text{für} \quad \sum_{\beta'} N'(\beta') \neq \sum_{\gamma'} N'(\gamma')$$

und

$$\Phi^{N(\beta'), N(q')}_{-\theta(\beta'), -\theta_p(\gamma')} = N'!\, \Psi^{\beta q}_{\alpha p} \quad \text{für} \quad \sum_{\beta'} N'(\beta') = \sum_{\gamma'} N'(q'), \tag{74}$$

wobei $\Phi^{\beta q}_{\alpha p}$, $\Psi^{\beta q}_{\alpha p}$ die in § 4 besprochenen Funktionen sind.

Die in A in der Form

$$\Big\{\sum_{rs} H_{rs}\, b^\dagger_r\, b_s - W\Big\}\, \Phi = 0$$

angegebene Funktionalgleichung für die zum Gesamtsystem gehörige Amplitude Φ lautet nunmehr bei unserer Vorzeichenbestimmung der in Φ enthaltenen Determinanten:

$$\Big\{\sum H_{rs}\, a^\dagger_r\, a_s - W\Big\}\, \Phi = 0;$$

diese Abänderung ist nötig, weil in A die Vorzeichenzweideutigkeit dieser Determinanten nicht ausreichend berücksichtigt wurde; die Matrizen a_r unterscheiden sich, wie wir wissen, von den b_r nur bezüglich der Vorzeichen ihrer verschiedenen Elemente. Es scheint sehr befriedigend, daß die für die Bildung des Energieausdrucks notwendige Einführung der

* Vgl. B, Formel (34).

Größen a, $a^\dagger$ gleichzeitig auch zu einfachen Multiplikationsgesetzen führte, wie wir in § 6 gesehen haben.

Es sei endlich hervorgehoben, daß die in § 6 erörterten Multiplikationsgesetze der gequantelten Amplituden $a_p(q')$ in Analogie zu den von Jordan und Pauli entwickelten relativistisch invarianten Multiplikationsregeln des ladungsfreien elektromagnetischen Feldes leicht relativistisch verallgemeinert werden können, so daß man die dem Paulischen Äquivalenzverbot entsprechende Quantelung der de Broglieschen Wellen in relativistisch invarianter Form erhält. Eine genauere Darlegung soll jedoch vorläufig zurückgestellt werden.

Zusatz bei der Korrektur: Zwischen den $2K$ Operatoren a_1, $a_2, \ldots, a_K$; $a_1^\dagger, a_2^\dagger, \ldots, a_K^\dagger$ bestehen die Relationen

$$\left.\begin{aligned} a_\varkappa a_\lambda + a_\lambda a_\varkappa &= 0 \\ a_\varkappa^\dagger a_\lambda^\dagger + a_\lambda^\dagger a_\varkappa^\dagger &= 0 \end{aligned}\right\} \tag{36}$$

und

$$a_\varkappa^\dagger a_\lambda + a_\lambda a_\varkappa^\dagger = \delta_{\varkappa\lambda}. \tag{40}$$

Wir wollen nun zeigen, daß diese ψ-Zahlrelationen die Operatoren a, $a^\dagger$ schon eindeutig bestimmen, wenn man sich auf irreduzible Matrizensysteme beschränkt und Matrizensysteme, die auseinander durch Ähnlichkeitstransformation hervorgehen, als nicht verschieden voneinander ansieht*.

Um dies einzusehen, bilden wir zunächst folgende Größen

$$\left.\begin{aligned} \alpha_\varkappa &= a_\varkappa + a_\varkappa^\dagger, \\ \alpha_{K+\varkappa} &= \frac{1}{i}(a_\varkappa - a_\varkappa^\dagger). \end{aligned}\right\} \tag{I}$$

Die $2K$ Matrizen α bestimmen umgekehrt die a eindeutig. Nun gilt für die $\alpha_\varkappa$ allgemein

$$\alpha_\varkappa \alpha_\lambda + \alpha_\lambda \alpha_\varkappa = 2\,\delta_{\varkappa\lambda}. \tag{II}$$

Man überzeugt sich z. B., wenn $\varkappa < K$, $\lambda < K$ ist, daß

$$\alpha_\varkappa \alpha_\lambda + \alpha_\lambda \alpha_\varkappa = (a_\varkappa + a_\varkappa^\dagger)\,.\,(a_\lambda + a_\lambda^\dagger) + (a_\lambda + a_\lambda^\dagger)\,.\,(a_\varkappa + a_\varkappa^\dagger) = 2\,\delta_{\varkappa\lambda}.$$

Man kann (II) auch schreiben

$$\left.\begin{aligned} \alpha_\varkappa^2 &= 1 \\ \alpha_\varkappa \alpha_\lambda &= -1\,\alpha_\lambda \alpha_\varkappa \quad \text{für} \quad \varkappa \neq \lambda, \end{aligned}\right\} \tag{IIa}$$

* Dies ist die Transformation aller Matrizen a durch dieselbe Matrix S zu $S^{-1}\,a\,S$, also nach dem Sprachgebrauch der allgemeinen Quantenmechanik die kanonische Transformation der Matrizendarstellung.

was aber mit anderen Worten so viel bedeutet, daß die $2K$ Matrizen α zusammen mit der Matrix -1 eine Gruppe aufspannen. Wenn z. B. $K = 2$ ist, so hat diese Gruppe folgende Elemente

$$\left.\begin{array}{l} 1;\ \alpha_1,\ \alpha_2,\ \alpha_3,\ \alpha_4;\ \alpha_1\alpha_2,\ \alpha_1\alpha_3,\ \alpha_1\alpha_4,\ \alpha_2\alpha_3,\ \alpha_2\alpha_4,\ \alpha_3\alpha_4; \\ -1;\ -\alpha_1,\ -\alpha_2,\ -\alpha_3,\ -\alpha_4;\ -\alpha_1\alpha_2,\ -\alpha_1\alpha_3,\ -\alpha_1\alpha_4,\ -\alpha_2\alpha_3,\ -\alpha_2\alpha_4,\ -\alpha_3\alpha_4; \\ \qquad \alpha_1\alpha_2\alpha_3,\ \alpha_1\alpha_3\alpha_4,\ \alpha_1\alpha_2\alpha_4,\ \alpha_2\alpha_3\alpha_4,\ \alpha_1\alpha_2\alpha_3\alpha_4. \\ \qquad -\alpha_1\alpha_2\alpha_3,\ -\alpha_1\alpha_3\alpha_4,\ -\alpha_1\alpha_2\alpha_4,\ -\alpha_2\alpha_3\alpha_4,\ -\alpha_1\alpha_2\alpha_3\alpha_4. \end{array}\right\} \text{(III)}$$

Das sind 32 Elemente, im allgemeinen 2^{2K+1} Elemente. Das irreduzible Matrizensystem, das (II) genügt, ist sicher eine irreduzible Darstellung dieser Gruppe (umgekehrt braucht es nicht der Fall zu sein, die Isomorphie kann ja mehrstufig sein). Wir werden nur die irreduziblen Darstellungen bestimmen.

Unsere Gruppe hat den Normalteiler $1, -1$ (das Zentrum), ihre Faktorgruppe vom Grade 2^{2K} ist abelsch. Sie hat also 2^{2K} dieser Abelschen Faktorgruppe entsprechende irreduzible Darstellungen vom Grade 1, die auch Darstellungen der ganzen Gruppe sind. Indessen kommen sie für uns nicht in Betracht, da sie den Gleichungen II nicht genügen (da sie ja kommutativ sind).

Wie viele Klassen hat unsere Gruppe? Die beiden Elemente 1 und -1 bilden je eine Klasse für sich, sonst ist aber jedes Element R mit $-1 . R$ in einer Klasse. Besteht nämlich R in (III) aus einer ungeraden Anzahl von Faktoren, so ist $\alpha R \alpha^{-1} = -1 . R$, wenn α in R nicht enthalten ist, wenn R aus einer geraden Anzahl von Faktoren besteht, ist $\alpha R \alpha^{-1} = -1 . R$, wenn α in R enthalten ist. Die Anzahl der Klassen ist also 2^{2K+1}; dies ist auch die Anzahl der voneinander verschiedenen irreduziblen Darstellungen. Da wir 2^{2K} Darstellungen schon kennen und diese nicht für uns in Betracht kommen, kann es nur eine einzige, die letzte sein, die die Gleichungen (II) befriedigt, alle anderen Lösungen von (II) gehen daraus durch Ähnlichkeitstransformation hervor.

Wir bestimmen noch die Anzahl von Zeilen und Spalten, die Dimension dieser Darstellung. (Es muß dabei natürlich 2^K herauskommen.) In der Tat ist der Grad der Gruppe 2^{2K+1} gleich der Summe der Quadrate der Dimensionen ihrer Darstellungen. Es haben 2^{2K} die Dimension 1, die letzte muß die Dimension 2^K haben, damit $(2^K)^2 + 2^{2K} . 1^2 = 2^{2K+1}$. Sie stimmt also tatsächlich mit unserem Matrizensystem (69) oder § 3 und 6 überein.

Göttingen, Institut für theoretische Physik.

Zur Erklärung einiger Eigenschaften der Spektren aus der Quantenmechanik des Drehelektrons. Zweiter Teil

J. von Neumann und E. P. Wigner

Zeitschrift für Physik *49*, 73–94 (1928)

Mit 2 Abbildungen. (Eingegangen am 2. März 1928.)

Mit Hilfe der im 1. Teil dieser Arbeit erhaltenen Ergebnisse wird das Aufbauprinzip der Serienspektren, die Auswahlregeln für die innere und magnetische Quantenzahl und der quadratische Starkeffekt behandelt.

Einleitung.

§ 1. Dem im ersten Teile dieser Arbeit gegebenen Programm entsprechend* sollen im vorliegenden Aufsatze die spektroskopischen Tatsachen, die aus dem Paulischen Modell des Drehelektrons ableitbar sind, näher untersucht werden — nachdem (im ersten Teile) die hierzu erforderlichen Grundtatsachen der Kinematik der Drehelektronen zusammengestellt wurden.

Es sei erlaubt, die dort erzielten Resultate nochmals zusammenzufassen.

Der Zustand eines Systems von n Drehelektronen ist (nach Pauli**) im Sinne der Dirac-Jordanschen Transformationstheorie durch eine „Wellenfunktion" zu beschreiben, die außer von den kartesischen Koordinaten $x_1, y_1, z_1, \ldots, x_n, y_n, z_n$ (oder auch $\mathfrak{r}_1, \mathfrak{r}_2, \ldots, \mathfrak{r}_n$) der Teilchen auch noch von ihren „Spins in der $+Z$-Richtung" $s_1, \ldots, s_n$ abhängt (die Koordinaten variieren von $-\infty$ bis $+\infty$, die Spins hingegen sind nur zweier Werte: ± 1 fähig):

$$\varphi = \varphi(x_1, \ldots, z_n; \quad s_1, \ldots, s_n). \tag{1}$$

Wenn wir $s_1, \ldots, s_n$ als Parameter behandeln, so ist φ eine Gesamtheit von 2^n gewöhnlichen Koordinatenfunktionen:

$$\varphi_{s_1, \ldots, s_n} = \varphi_{s_1, \ldots, s_n}(x_1, \ldots, z_n) = \varphi(x_1, \ldots, z_n; \; s_1, \ldots, s_n). \tag{2}$$

Demgemäß wollen wir die eigentliche Wellenfunktion als „Hyperfunktion" und die $\varphi_{s_1, \ldots, s_n}$ (die „spinfreien Wellenfunktionen") als gewöhnliche Funktionen bezeichnen.

---

* ZS. f. Phys. **47**, 203, 1928; diese Arbeit soll als I. zitiert werden. Eine genaue Kenntnis dieser Arbeit wird nicht vorausgesetzt. Dagegen dürfte die Kenntnis wenigstens der ersten Hälfte einer Arbeit des einen von uns (E. Wigner, ZS. f. Phys. **43**, 624, 1927, im folgenden als *F* zitiert) notwendig sein.

** ZS. f. Phys. **43**, 601, 1927.

Wenn derjenige Zustand des Systems, zu der die Hyperfunktion φ gehört, der räumlichen Drehung $\mathfrak{R}$:

$$\left.\begin{aligned} x' &= \alpha_{11}x + \alpha_{12}y + \alpha_{13}z, \\ y' &= \alpha_{21}x + \alpha_{22}y + \alpha_{23}z, \\ z' &= \alpha_{31}x + \alpha_{32}y + \alpha_{33}z \end{aligned}\right\} \qquad (\mathfrak{R})$$

unterworfen wird, so entsteht aus ihm ein neuer Zustand, dessen Hyperfunktion $O_{\mathfrak{R}}\varphi$ * folgendermaßen aus φ zu berechnen ist:

$$O_{\mathfrak{R}}\varphi_{s_1,\ldots,s_n}(x_1',\ldots,z_n') = \sum_{t_1,\ldots,t_n} {}^n a^{(\mathfrak{R})}_{s_1,\ldots,s_n;\,t_1,\ldots,t_n}\,\varphi_{t_1,\ldots,t_n}(x_1,\ldots,z_n), \qquad (3)$$

dabei sind die „Transformationskoeffizienten" ${}^n a^{(\mathfrak{R})}_{s_1,\ldots,s_n;\,t_1,\ldots,t_n}$ so definiert:

$${}^n a^{(\mathfrak{R})}_{s_1,\ldots,s_n;\,t_1,\ldots,t_n} = {}^1 a^{(\mathfrak{R})}_{s_1 t_1}\cdot\ldots\cdot {}^1 a^{(\mathfrak{R})}_{s_n t_n} \qquad (4)$$

und

$$\left.\begin{aligned} {}^1 a^{(\mathfrak{R})}_{-1\,-1} &= e^{-\frac{i\alpha}{2}}\cos\frac{\beta}{2}e^{-\frac{i\gamma}{2}}, & {}^1 a^{(\mathfrak{R})}_{-1\,1} &= e^{-\frac{i\alpha}{2}}\sin\frac{\beta}{2}e^{\frac{i\gamma}{2}}, \\ {}^1 a^{(\mathfrak{R})}_{1\,-1} &= -e^{\frac{i\alpha}{2}}\sin\frac{\beta}{2}e^{-\frac{i\gamma}{2}}, & {}^1 a^{(\mathfrak{R})}_{1\,1} &= e^{\frac{i\alpha}{2}}\cos\frac{\beta}{2}e^{\frac{i\gamma}{2}}, \end{aligned}\right\} \qquad (4\text{a})$$

wo α, β, γ die Eulerschen Winkel der Drehung $\mathfrak{R}$ sind.

Die ${}^n a^{(\mathfrak{R})}_{s_1,\ldots,s_n;\,t_1,\ldots,t_n}$ bilden 2^n-dimensionale orthogonale Matrizen, die eine Darstellung der dreidimensionalen Drehgruppe liefern (ein- oder zweideutig, je nachdem n gerade oder ungerade ist); deren irreduzible Bestandteile wurden bestimmt.

Außer dem Operator $O_{\mathfrak{R}}$ wurde in I. noch ein anderer betrachtet, der φ in $P_{\mathfrak{R}}\varphi$ übergeführt hat. Es war dabei

$$P_{\mathfrak{R}}\varphi_{s_1,\ldots,s_n}(x_1',\ldots,z_n') = \varphi_{s_1,\ldots,s_n}(x_1,\ldots,z_n). \qquad (5)$$

Er bedeutet also quasi die Verdrehung des Zustandes, wobei aber die Spinkoordinaten nicht mitverdreht werden.

Wir werden hier noch weitere ähnliche Transformationen gebrauchen müssen. Wenn wir nämlich unter $\mathfrak{R}$ keine Drehung, sondern eine Permutation der Zahlen $1, 2, 3, \ldots, n$ verstehen, die 1 in α_1, 2 in α_2 usw., n in α_n überführt ($\alpha_1, \alpha_2, \ldots, \alpha_n$ ist eine Permutation der Zahlen $1, 2, \ldots, n$), so verstehen wir ähnlich unter $O_{\mathfrak{R}}\varphi$ die Hyperfunktion, für die identisch gilt

$$O_{\mathfrak{R}}\varphi(\mathfrak{r}_{\alpha_1},\ldots,\mathfrak{r}_{\alpha_n};\ s_{\alpha_1},\ldots,s_{\alpha_n}) = \varphi(\mathfrak{r}_1,\ldots,\mathfrak{r}_n;\ s_1,\ldots,s_n), \qquad (3\text{a})$$

und unter $P_{\mathfrak{R}}\varphi$ die Hyperfunktion

$$P_{\mathfrak{R}}\varphi(\mathfrak{r}_{\alpha_1},\ldots,\mathfrak{r}_{\alpha_n};\ s_1,\ldots,s_n) = \varphi(\mathfrak{r}_1,\ldots,\mathfrak{r}_n;\ s_1,\ldots,s_n). \qquad (5\text{a})$$

* In der Bezeichnung F ist $O_{\mathfrak{R}}\varphi$ das $\varphi(\mathfrak{R}^{-1})$.

Wie in I. verstehen wir weiter unter $Q_{\mathfrak{R}}$ den Operator

$$Q_{\mathfrak{R}} = P_{\mathfrak{R}}^{-1} O_{\mathfrak{R}}; \quad O_{\mathfrak{R}} = P_{\mathfrak{R}} Q_{\mathfrak{R}}, \tag{6}$$

d. h. wenn $\mathfrak{R}$ eine Drehung ist, ist

$$\begin{aligned} &Q_{\mathfrak{R}} \varphi (\mathfrak{r}_1, \ldots, \mathfrak{r}_n;\ s_1, \ldots, s_n) \\ &\quad = \sum_{t_1, \ldots, t_n} {}^n a^{(\mathfrak{R})}_{s_1, \ldots, s_n;\ t_1, \ldots, t_n} \varphi (\mathfrak{r}_1, \ldots, \mathfrak{r}_n;\ t_1, \ldots, t_n), \end{aligned} \tag{7}$$

wenn $\mathfrak{R}$ die erwähnte Permutation ist, ist

$$Q_{\mathfrak{R}} \varphi (\mathfrak{r}_1, \ldots, \mathfrak{r}_n;\ s_{\alpha_1}, \ldots, s_{\alpha_n}) = \varphi (\mathfrak{r}_1, \ldots, \mathfrak{r}_n;\ s_1, \ldots, s_n). \tag{7a}$$

$Q_{\mathfrak{R}}$ bedeutet quasi das Verdrehen oder Vertauschen der Spinkoordinaten allein. Die Q sind mit den P vertauschbar. Es gilt außerdem für beliebige $\mathfrak{R}$ und $\mathfrak{S}$

$$O_{\mathfrak{R}} O_{\mathfrak{S}} = \pm O_{\mathfrak{R}\mathfrak{S}}; \quad Q_{\mathfrak{R}} Q_{\mathfrak{S}} = \pm Q_{\mathfrak{R}\mathfrak{S}}; \quad P_{\mathfrak{R}} P_{\mathfrak{S}} = P_{\mathfrak{R}\mathfrak{S}}, \quad \text{also } P_{\mathfrak{R}}^{-1} = P_{\mathfrak{R}^{-1}}. \tag{8}$$

Diese Formeln und Begriffsbildungen werden wir im Laufe der Arbeit häufig gebrauchen.

Wir werden außerdem von der Tatsache Gebrauch machen, daß $O_{\mathfrak{R}}$ ein orthogonaler Operator ist, also das innere Produkt invariant läßt. Es ist also, wenn φ und ψ beliebige Hyperfunktionen sind

$$\int O_{\mathfrak{R}} \varphi \,.\, \widetilde{O_{\mathfrak{R}} \psi} = \int \varphi \tilde{\psi}, \tag{9}$$

wie man sich leicht überzeugt. Das Integralzeichen bedeutet — wie in I. und immer in dieser Arbeit — Integration von $-\infty$ bis $+\infty$ über alle Koordinaten $x_1, \ldots, z_n$ und Summation über alle 2^n Wertsysteme der $s_1, \ldots, s_n$. Wir können dann, wie in F $O_{\mathfrak{R}} \varphi$ bzw. $O_{\mathfrak{R}} \psi$ auf andere Weise (mit Hilfe der Darstellungseigenschaften) ausdrücken und zeigen, daß (9) nur bestehen kann, wenn gewisse Relationen zwischen diesen Integralen bestehen. Wir werden dies bei der Ableitung der Auswahl und Intensitätsregeln benutzen.

Das Aufbauprinzip.

§ 2. Die Bestimmung der Terme, d. h. der Energiewerte eines Atoms wird natürlich auch bei Berücksichtigung der Elektronenmagnete auf ein Eigenwertproblem führen. In F wurde das Aufbauprinzip der Serienspektren bis zu dem Punkte abgeleitet, wo man die magnetischen Momente der Elektronen einführen muß.

Man kann den Energieoperator H in zwei Teile zerlegen, $H = H_1 + H_2$ so, daß H_1 die Energieanteile, die von der Schwerpunktsbewegung der

Elektronen und der elektrischen und magnetischen Wechselwirkung dieser Bewegung herrühren, H_2 den Rest, der also auch die magnetischen Momente der Elektronen berücksichtigt, enthält.

Das spinfreie Eigenwertproblem $H_1 y = \lambda y$, das also in F ausschließlich betrachtet wurde, enthält nur die Energieanteile, die von der Schwerpunktsbewegung der Elektronen herrühren. Ihre Eigenfunktionen waren lediglich Funktionen der Schwerpunktskoordinaten der Teilchen. Jeder Term hatte eine bestimmte Darstellungseigenschaft gegenüber Permutationen der Elektronen und eine azimutale Quantenzahl l, die bei Drehungen zur Geltung kam.

A. Unser erster Schritt wird sein, diese „spinfreien Eigenfunktionen" zu Hyperfunktionen zu ergänzen.

B. Die vorher erwähnten Darstellungseigenschaften der spinfreien Eigenfunktionen sind dann die Darstellungseigenschaften der Hyperfunktionen gegenüber den Operationen $P_{\mathfrak{R}}$. Wir werden dann die Darstellungseigenschaften gegenüber den Operationen $O_{\mathfrak{R}}$ feststellen, indem wir zunächst die Darstellungseigenschaften den $Q_{\mathfrak{R}}$ gegenüber aufsuchen und dann die Gleichung (6): $O_{\mathfrak{R}} = P_{\mathfrak{R}} Q_{\mathfrak{R}}$ benutzen.

C. Zuletzt denken wir uns dann die Energieanteile H_2, die von den magnetischen Momenten der Teilchen herrühren, langsam anwachsen. Da dabei die Invarianz des Eigenwertproblems gegenüber den Operationen $O_{\mathfrak{R}}$ erhalten bleibt, ändern sich die Darstellungseigenschaften der Hyperfunktionen den $O_{\mathfrak{R}}$ gegenüber nicht. Daß bei dem Anwachsen von H_2 die Eigenwerte von H_1 im allgemeinen weiter aufspalten, beruht natürlich darauf, daß H_1 nicht nur gegenüber $O_{\mathfrak{R}}$ ($\mathfrak{R}$ Drehung oder Permutation) invariant ist, wie $H_1 + H_2$, sondern auch die Operationen $P_{\mathfrak{R}}$ zuläßt, was $H_1 + H_2$ nicht mehr tut. Er läßt also offenbar auch $Q_{\mathfrak{R}} = P_{\mathfrak{R}}^{-1} O_{\mathfrak{R}}$ zu, was ja auch sonst klar ist.

Da in der Natur nach Heisenberg und Dirac nur solche Zustände vorkommen, deren Eigenfunktionen antisymmetrisch sind:

$$O_{\mathfrak{R}} \varphi = \pm \varphi$$

($\mathfrak{R}$ eine beliebige Permutation, es gilt $+$ oder $-$, je nachdem $\mathfrak{R}$ gerade oder ungerade ist), werden wir uns auf solche beschränken können.

Wir werden B. in zwei Schritten ausführen, indem wir zunächst solche $O_{\mathfrak{R}}$ betrachten, bei denen $\mathfrak{R}$ eine Permutation ist, sondern die antisymmetrischen Hyperfunktionen in bezug auf $O_{\mathfrak{R}}$ aus und betrachten die Wirkung der $O_{\mathfrak{R}}$, wo $\mathfrak{R}$ eine Drehung ist, nur auf diese Hyperfunktionen.

§ 3. Ist die Funktion y eine Lösung des spinfreien Eigenwertproblems $H_1 y = \lambda y$, so können wir aus y genau 2^n Hyperfunktionenlösungen dieses Eigenwertproblems bilden. Diese Hyperfunktionen werden wir mit $\psi_{\tau_1, \ldots, \tau_n}(\mathfrak{r}_1, \ldots, \mathfrak{r}_n;\ s_1, \ldots, s_n)$ bezeichnen*, wo die $\tau_1, \tau_2, \ldots, \tau_n$ gleich ± 1 sein können und zur Unterscheidung der 2^n Hyperfunktionen dienen. Es ist

$$\psi_{\tau_1, \ldots, \tau_n}(\mathfrak{r}_1, \ldots, \mathfrak{r}_n;\ s_1, \ldots, s_n) = 0, \tag{10}$$

wenn nicht $s_1 = \tau_1;\ s_2 = \tau_2; \ldots;\ s_n = \tau_n$ gilt,

$$\psi_{\tau_1, \ldots, \tau_n}(\mathfrak{r}_1, \ldots, \mathfrak{r}_n;\ \tau_1, \ldots, \tau_n) = y(\mathfrak{r}_1, \ldots, \mathfrak{r}_n). \tag{10a}$$

Es ist klar, daß diese 2^n Hyperfunktionen linear unabhängig (sogar orthogonal) sind. Auch daß sie das Eigenwertproblem $H_1 \psi = \lambda \psi$ lösen, sieht man leicht ein. Es ist nämlich klar, daß man H_1 (da es auf die Spinkoordinaten s nicht einwirkt) auf die 2^n Funktionen einer Hyperfunktion einzeln anwenden kann.

Da im allgemeinen schon mehrere, sagen wir m Funktionen y zu einem Eigenwert λ gehörten (die sich eben bei der Vertauschung der Elektronen oder Drehung unter sich transformierten), werden wir $m \,.\, 2^n$ Hyperfunktionen zu diesem Eigenwert haben.

Wir können diese $m \,.\, 2^n$ Hyperfunktionen in eine Tabelle mit m Zeilen und 2^n Spalten anordnen. In einer Zeile stehen die Hyperfunktionen, die aus demselben y_μ entstanden sind, sie unterscheiden sich durch ihre unteren Indizes. In einer Spalte stehen Hyperfunktionen, die dieselben unteren Indizes haben und aus den verschiedenen $y_1, y_2, \ldots y_m$ entstanden sind.

$$\begin{array}{lll}
\psi^{(1)}_{1,1,\ldots,1}, & \psi^{(1)}_{1,1,\ldots,-1}, \ldots & \psi^{(1)}_{-1,-1,\ldots,-1}, \\
\psi^{(2)}_{1,1,\ldots,1}, & \psi^{(2)}_{1,1,\ldots,-1}, \ldots & \psi^{(2)}_{-1,-1,\ldots,-1}, \\
\vdots & \vdots & \vdots \\
\psi^{(m)}_{1,1,\ldots,1}, & \psi^{(m)}_{1,1,\ldots,-1}, \ldots & \psi^{(m)}_{-1,-1,\ldots,-1}.
\end{array}$$

$P_{\mathfrak{R}}\, \psi^{(k)}_{\tau_1, \ldots, \tau_n}$ läßt sich durch die $\psi^{(\varkappa)}_{\tau_1, \ldots, \tau_n}$ $(\varkappa = 1, 2, \ldots, m)$ derselben Spalte allein linear ausdrücken und die Transformationseigenschaften sind offenbar die der y, dagegen läßt sich $Q_{\mathfrak{R}}\, \psi^{(k)}_{\tau_1, \ldots, \tau_n}$ durch die $\psi^{(k)}_{\sigma_1, \ldots, \sigma_n}$ derselben Zeile allein [mit Hilfe von (7) und (7a)] ausdrücken. Für $O_{\mathfrak{R}}\, \psi = P_{\mathfrak{R}}\, Q_{\mathfrak{R}}\, \psi$ braucht man zunächst die ganze Tabelle. Wie schon

* $\psi_{\tau_1, \ldots, \tau_n}$ ist also keine Funktion, sondern eine Hyperfunktion.

angekündigt, werden wir nun (§ 4) solche Linearkombinationen Φ dieser ψ einführen, bei denen möglichst wenig Φ unter sich transformieren, d. h.

$$O_{\mathfrak{R}}\,\Phi_{\eta\nu} = \sum_{\mu} {}^{\eta}b_{\mu\nu}^{(\mathfrak{R})}\,\Phi_{\eta\mu}$$

gilt, wo die $\left({}^{\eta}b_{\mu\nu}^{(\mathfrak{R})}\right)$ irreduzible Darstellungen sind. Zunächst werden wir dies für solche $\mathfrak{R}$ anstreben, die Permutationen der Elektronen darstellen.

§ 4. Wenn wir in einem $\psi_{\tau_1,\ldots,\tau_n}(\mathfrak{r}_1,\ldots,\mathfrak{r}_n;\, s_1,\ldots,s_n)$ eine Vertauschung der s vornehmen, also $Q_{\mathfrak{R}}\,\psi_{\tau_1,\ldots,\tau_n}$ bilden, so läßt sich diese Hyperfunktion durch die ursprünglichen 2^n Hyperfunktionen

$$\psi_{\sigma_1,\ldots,\sigma_n} \qquad (\sigma_1,\ldots\sigma_n = \pm 1)$$

linear ausdrücken. (Und zwar sind die Koeffizienten, wie nebenbei bemerkt sei, von allen $\psi_{\sigma_1,\ldots,\sigma_n}$ gleich Null, abgesehen von dem, dessen Indexreihe aus $\tau_1,\ldots,\tau_n$ durch die Substitution $\mathfrak{R}^{-1}$ hervorgeht.) Es liegt nun nahe, nach solchen Linearkombinationen der $\psi_{\tau_1,\ldots,\tau_n}$ zu fragen, deren Transformationsformel eine irreduzible Darstellung der symmetrischen Gruppe von n Elementen bildet.

Zunächst ist klar, daß sich von vornherein nur solche $\psi_{\tau_1,\ldots,\tau_n}$ ineinander transformieren, bei denen die Summe der unteren Indizes $\tau_1 + \cdots + \tau_n = \tau$ dieselbe ist. Jede neue Hyperfunktion wird also eine Linearkombination nur solcher ψ sein, bei denen diese Summe denselben Wert, sagen wir den Wert τ hat; solche ψ haben wir $\left(\dfrac{n}{\frac{n-|\tau|}{2}}\right)$.

Ihre Transformationsformeln sind

$$Q_{\mathfrak{R}}\,\psi_{\tau_1,\ldots,\tau_n} = \sum_{\sigma_1,\ldots,\sigma_n} c^{(\mathfrak{R})}_{\sigma_1,\ldots,\sigma_n;\,\tau_1,\ldots,\tau_n}\,\psi_{\sigma_1,\ldots,\sigma_n}, \tag{11}$$

wo $c^{(\mathfrak{R})}_{\sigma_1,\ldots,\sigma_n;\,\tau_1,\ldots,\tau_n}$ gleich 1 oder 0, je nachdem $\sigma_1,\ldots,\sigma_n$ durch $\mathfrak{R}^{-1}$ aus $\tau_1,\ldots,\tau_n$ hervorgeht oder nicht. Die $\left(c^{(\mathfrak{R})}_{\sigma_1,\ldots,\sigma_n;\,\tau_1,\ldots,\tau_n}\right)$ bilden eine (triviale) $\left(\dfrac{n}{\frac{n-|\tau|}{2}}\right)$-dimensionale Darstellung der symmetrischen Gruppe von n Elementen, deren irreduzible Bestandteile wir bestimmen müssen. Es werden nämlich diese die Transformationseigenschaften der neuen linear unabhängigen Hyperfunktionen den $Q_{\mathfrak{R}}$ gegenüber sein (siehe F, Punkt 5). Wir übergehen die einfache Rechnung und geben nur das Resultat an.

Die irreduziblen Bestandteile sind: je eine Darstellung, die folgenden Zerlegungen von n in Summanden zugeordnet sind:

$$n,\ 1+(n-1),\ 2+(n-2),\ \cdots,\qquad \frac{n-|\tau|}{2}+\frac{n+|\tau|}{2}.$$

Die Grade (Anzahl von Zeilen und Spalten) dieser Darstellungen sind der Reihe nach

$$\binom{n}{0},\ \binom{n}{1}-\binom{n}{0},\ \binom{n}{2}-\binom{n}{1},\ \cdots,\ \binom{n}{\frac{n-|\tau|}{2}}-\binom{n}{\frac{n-|\tau|}{2}-1} \tag{12}$$

Wir benennen die Darstellungen nach dem ersten Summanden in ihrer Partitio, den wir überstreichen wollen. Aus den $\binom{n}{\frac{n-|\tau|}{2}}$ Hyperfunktionen $\psi_{\tau_1,\ldots,\tau_n}$ mit $\tau+\cdots+\tau_n=\tau$ erhalten wir also neue Hyperfunktionen, die folgendermaßen zusammengefaßt werden:

eine gehört zur symmetrischen Darstellung, benannt $\overline{0}$,

$\binom{n}{1}-\binom{n}{0}$ gehören zur Darstellung, benannt $\overline{1}$,

$\binom{n}{2}-\binom{n}{1}$ „ „ „ „ $\overline{2}$,

$\binom{n}{\frac{n-|\tau|}{2}}-\binom{n}{\frac{n-|\tau|-2}{2}}$ „ „ „ „ $\overline{\frac{n-|\tau|}{2}}$.

Die neuen Hyperfunktionen [die alle aus derselben Funktion $y_k(\mathfrak{r}_2,\ldots,\mathfrak{r}_n)$ entstanden sind] bezeichnen wir also folgendermaßen: $\Phi^k_{\tau,z,\varkappa}$, wo τ die Werte $-n,\ -n+2,\ -n+4,\ldots,n$ annehmen kann, der Index $\bar{z}$ die Darstellungseigenschaft bezeichnet, von 0 bis $\frac{n-|\tau|}{2}$ und $\varkappa$ von 1 bis $\binom{n}{\frac{n-|\tau|}{2}}-\binom{n}{\frac{n-|\tau|}{2}-1}$ läuft. Sie ist eine Linearkombination der $\psi_{\tau_1,\ldots,\tau_n}$ mit $\tau_1+\cdots+\tau_n=\tau$ und es gilt

$$Q_{\mathfrak{R}}\,\Phi^k_{\tau,\bar{z},\varkappa}=\sum_{\lambda}{}^{\bar{z}}a^{(\mathfrak{R})}_{\lambda\varkappa}\,\Phi^k_{\tau,\bar{z},\lambda}, \tag{13}$$

wo also $\left({}^{\bar{z}}a^{(\mathfrak{R})}_{\lambda\varkappa}\right)$ die irreduzible Darstellung $\bar{z}$ ist.

§ 5. Andererseits wissen wir, daß

$$P_{\mathfrak{R}}\,y_k=\sum_{l}A^{(\mathfrak{R})}_{lk}\,y_l, \tag{14}$$

wo $\left(A_{kl}^{(\mathfrak{R})}\right)$ eine irreduzible Darstellung, die Darstellung der spinfreien Eigenfunktion der symmetrischen Gruppe gegenüber ist. Es ist also auch

$$\left.\begin{aligned} P_{\mathfrak{R}}\,\Phi^{k}_{\tau,\bar{z},\varkappa} &= \sum_{l} A_{lk}^{(\mathfrak{R})}\,\Phi^{l}_{\tau,\bar{z},k}, \\ P_{\mathfrak{R}}\,Q_{\mathfrak{R}}\,\Phi^{k}_{\tau,\bar{z},\varkappa} &= O_{\mathfrak{R}}\,\Phi^{k}_{\tau,\bar{z},\varkappa} = \sum_{l\lambda} A_{lk}^{(\mathfrak{R})}\,.\,{}^{\bar{z}}a_{\lambda\varkappa}^{(\mathfrak{R})}\,\Phi^{l}_{\tau,\bar{z},\lambda}. \end{aligned}\right\} \tag{15}$$

Die $O_{\mathfrak{R}}\,\Phi^{k}_{\tau,\bar{z},\varkappa}$ sind also Linearkombinationen der Φ, was ja auch selbstverständlich ist, da H_1 die Operation $O_{\mathfrak{R}}$ zuläßt. Uns interessierten aber auch die Darstellungseigenschaften der Hyperfunktionen Φ, die aus den Eigenfunktionen von $H_1 y = \lambda y$ durch Einführung der weiteren Koordinaten $s_1, s_2, \ldots, s_n$ entstehen. Wir haben diese nunmehr in (15). Unsere Aufgabe wäre nun, die $m\left[\binom{n}{z} - \binom{n}{z-1}\right]$-dimensionale Darstellung $\left(A_{lk}^{(\mathfrak{R})}\,.\,{}^{\bar{z}}a_{\lambda\varkappa}^{(\mathfrak{R})}\right)$ auszureduzieren, die irreduziblen Bestandteile zu bestimmen. Die Transformationseigenschaften dieser ändern sich nämlich auch nicht, wenn man nun H_2, d. h. den anderen Teil des Energieoperators langsam anwachsen läßt. Da uns indessen nur die antisymmetrischen Eigenfunktionen interessieren werden, genügt es nachzusehen, ob und wie oft die antisymmetrische Darstellung in $\left(A_{lk}^{(\mathfrak{R})}\,.\,{}^{\bar{z}}a_{\lambda\varkappa}^{(\mathfrak{R})}\right)$ enthalten ist. Dies ist aber leicht zu erledigen*.

Stellen wir uns nämlich vor, die Matrix $(U_{u;\,l\lambda})$ (u bezeichnet die Zeilen, läuft von 1 bis $m\left[\binom{n}{z} - \binom{n}{z-1}\right]$; k und $\varkappa$ beziehen sich auf die Spalten) hätte die Eigenschaft, $\left(A_{lk}^{(\mathfrak{R})}\,.\,{}^{\bar{z}}a_{\lambda\varkappa}^{(\mathfrak{R})}\right)$ auszureduzieren. Kommt in der ausreduzierten Darstellung $\left(B_{uv}^{(\mathfrak{R})}\right)$

$$\left(B_{uv}^{(\mathfrak{R})}\right) = (U_{u;\,l\lambda})\left(A_{lk}^{(\mathfrak{R})}\,.\,{}^{\bar{z}}a_{\lambda\varkappa}^{(\mathfrak{R})}\right)(U_{v;\,k\varkappa})^{-1} \tag{16}$$

die antisymmetrische Darstellung in die u-te Zeile, so ist jedenfalls

$$B_{uu}^{(\mathfrak{R})} = \varepsilon_{\mathfrak{R}} = \sum_{k,\varkappa,l,\lambda} C_{l\lambda,\,k\varkappa}\,A_{lk}^{(\mathfrak{R})}\,{}^{\bar{z}}a_{\lambda\varkappa}^{(\mathfrak{R})}, \tag{16a}$$

wo die $C_{l\lambda,\,k\varkappa}$ gewisse Konstanten, die $\varepsilon_{\mathfrak{R}}$ gleich $+1$ oder -1, je nachdem $\mathfrak{R}$ eine gerade oder ungerade Permutation ist.

Nun kommt der einzige wesentliche Punkt in der Überlegung. Wir beachten nämlich, daß die Matrizen

$$\left({}^{z}a_{\varkappa\lambda}^{(\mathfrak{R})}\right) = \varepsilon_{\mathfrak{R}}\left({}^{\bar{z}}a_{\varkappa\lambda}^{(\mathfrak{R})}\right) \tag{17}$$

* Siehe die schöne Diskussion bei F. Hund, ZS. f. Phys. 43, 788, 1927.

auch eine Darstellung, und zwar — da $\bar{z}$ eine solche ist — eine irreduzible Darstellung bilden. Es ist nämlich

$$\sum_{\lambda} \left({}^{z}a_{\varkappa\lambda}^{(\mathfrak{R})}\right) . \left({}^{z}a_{\lambda\mu}^{(\mathfrak{S})}\right) = \sum_{\lambda} \varepsilon_{\mathfrak{R}} \left({}^{\bar{z}}a_{\varkappa\lambda}^{(\mathfrak{R})}\right) \varepsilon_{\mathfrak{S}} \left({}^{\bar{z}}a_{\lambda\mu}^{(\mathfrak{S})}\right)$$

$$= \varepsilon_{\mathfrak{R}\mathfrak{S}} \left({}^{\bar{z}}a_{\varkappa\mu}^{(\mathfrak{R}\mathfrak{S})}\right) = \left({}^{z}a_{\varkappa\mu}^{(\mathfrak{R}\mathfrak{S})}\right), \tag{17a}$$

so daß (17) wirklich eine Darstellung gibt. Sie ist offenbar auch irreduzibel. Nun liegt es nahe, Gleichung (16a) mit $\varepsilon_{\mathfrak{R}}$ zu multiplizieren und über alle Permutationen zu summieren:

$$\sum_{\mathfrak{R}} \varepsilon_{\mathfrak{R}}^{2} = \sum_{\mathfrak{R}} 1 = \sum_{\mathfrak{R}} \sum_{k,\varkappa,l,\lambda} C_{l\lambda k\varkappa} A_{lk}^{(\mathfrak{R})} . {}^{z}a_{\lambda\varkappa}^{(\mathfrak{R})}. \tag{18}$$

Es ist sowohl $\left(A_{lk}^{(\mathfrak{R})}\right)$, wie $\left({}^{z}a_{\lambda\varkappa}^{(\mathfrak{R})}\right)$ eine irreduzible Darstellung, wir haben also, wenn sie nicht äquivalent sind,

$$\sum_{\mathfrak{R}} 1 = n! = 0, \tag{18a}$$

was offenbar ein Widerspruch ist, der nur daher rühren kann, daß wir angenommen haben, die Darstellung $\left(A_{kl}^{(\mathfrak{R})} . {}^{\bar{z}}a_{\varkappa\lambda}^{\mathfrak{R}}\right)$ enthalte die antisymmetrische, auch wenn $\left(A_{kl}^{(\mathfrak{R})}\right)$ nicht äquivalent zu $\left({}^{z}a_{\varkappa\lambda}^{\mathfrak{R}}\right)$ ist. Dies kann also höchstens der Fall sein, wenn diese äquivalent sind.

In der Tat überzeugt man sich in diesem Falle fast ebenso leicht, daß die Darstellung $\left(A_{kl}^{\mathfrak{R}} . {}^{\bar{z}}a_{\varkappa\lambda}^{\mathfrak{R}}\right)$ die antisymmetrische einmal und nur einmal enthält. (Siehe z. B. A. Speiser, Theorie der Gruppen von endlicher Ordnung. Berlin 1923. Erste Auflage. Satz 105, S. 112.)

§ 6. Nun schauen wir noch nach, welche irreduzible Darstellung $\varepsilon_{\mathfrak{R}}\left({}^{\bar{z}}a_{\varkappa\lambda}^{(\mathfrak{R})}\right)$ ist, die also $\left(A_{kl}^{(\mathfrak{R})}\right)$ äquivalent sein muß, damit überhaupt aus den $\Phi_{\tau,\bar{z},\varkappa}^{k}$ (bei festem τ und z, variablem k und $\varkappa$) eine antisymmetrische Eigenfunktion gebildet werden kann. Man findet, daß (17) zur Zerlegung von n in die Summanden

$$\underbrace{1 + 1 + \cdots 1}_{n-2z} + \underbrace{2 + \cdots 2}_{z} \tag{17b}$$

gehört*. Es sind z Zweier (daher die Bezeichnung) und $n - 2z$ Einser vorhanden. Dies muß also auch die Zerlegung von $\left(A_{kl}^{(\mathfrak{R})}\right)$ sein.

Unser Resultat ist also folgendes. Gehen wir von einem Term, der „spinfreien Differentialgleichung" $H_2 y(\mathfrak{r}_1, \ldots, \mathfrak{r}_n) = \lambda y(\mathfrak{r}_1, \ldots, \mathfrak{r}_n)$ aus, deren Darstellungseigenschaft in bezug auf die Vertauschung der Elektronen durch $\left(A_{kl}^{(\mathfrak{R})}\right)$ gegeben ist und bilden die $m . 2^n$ Hyperfunktionen $\Phi_{\tau,\bar{z},\varkappa}^{k}$, so liefert jedes Paar τ, z eine oder keine antisymmetrische Hyperfunktion, je nachdem in der „Zerlegung" von $\left(A_{kl}^{(\mathfrak{R})}\right)$ genau z Zweier und $n - 2z$

* Siehe F oder F. Hund, l. c.

Einser sind oder nicht. Diese antisymmetrischen Hyperfunktionen sind dann gewisse Linearkombinationen

$$\Psi_{\tau z} = \sum_{k\varkappa} c_{k\varkappa} \Phi^k_{\tau, \bar{z}, \varkappa},$$

wo uns indessen die Koeffizienten $c_{k\varkappa}$ nicht weiter interessieren. Die anderen Linearkombinationen der $\Phi^k_{\tau, \bar{z}, \varkappa}$ können zu Hyperfunktionen mit sicher nicht antisymmetrischen Darstellungseigenschaften zusammengefaßt werden und scheiden daher für das Folgende überhaupt aus. $\Psi_{\tau z}(\mathfrak{r}_1, \ldots, \mathfrak{r}_n; s_1, \ldots, s_n)$ ist eine antisymmetrische Hyperfunktion, die 1. an allen Stellen, wo $s_1 + \cdots + s_n \neq \tau$, verschwindet und 2. deren Funktionen alle Linearkombinationen der $y_1, \ldots, y_m$ sind. Die Darstellungseigenschaft von $y(\mathfrak{r}_1, \ldots, \mathfrak{r}_n)$ in bezug auf Vertauschung der $\mathfrak{r}$ entspricht der Zerlegung (17 b).

Ist umgekehrt die Darstellungseigenschaft (17 b) der y von vornherein vorgegeben, so kann man nur noch τ variieren und erhält alle antisymmetrischen Hyperfunktionen, wenn man es von $-(n - 2z)$ bis $n - 2z$ laufen läßt. Um den Zusammenhang mit dem Vektorzusammensetzungsmodell besser vor Augen zu haben, möchten wir dies noch durch ein Beispiel illustrieren. Es sei $n = 4$, für τ sind dann die Werte

$$1 + 1 + 1 + 1 = 4;\ 1 + 1 + 1 - 1 = 2;\ 1 + 1 - 1 - 1 = 0;$$
$$1 - 1 - 1 - 1 = -2;\ \ -1 - 1 - 1 - 1 = -4$$

möglich (§ 4).

$\bar{z}$	$\tau = -4$	-2	0	2	4
0	$\Psi_{-4\,0}$	$\Psi_{-2\,0}$	$\Psi_{0\,0}$	$\Psi_{2\,0}$	$\Psi_{4\,0}$
1		$\Psi_{-2\,1}$	$\Psi_{0\,1}$	$\Psi_{2\,1}$	
2			$\Psi_{0\,2}$		

Wenn nun z. B. y die Zerlegung $1 + 1 + 2$, also $z = 1$ hatte, so interessiert uns nur die zweite Zeile, da nur diese antisymmetrische Hyperfunktionen liefert. Man kann sie abzählen, indem man τ von $-(4 - 2 \,.\, 1) = -2$ bis $4 - 2 \,.\, 1 = 2$ laufen läßt (natürlich springt τ immer um 2).

§ 7. Damit ist gleichzeitig ein wichtiges Resultat, das in F vorausgesetzt wurde, abgeleitet, daß nämlich nur diejenigen spinfreien Eigenfunktionen in Betracht kommen, deren Zerlegung aus lauter 1 und 2 besteht. L. c. war dies damit begründet, daß jede Bahn durch Elektronen höchstens doppelt besetzt sein kann. Jetzt haben wir eine Erklärung, die aus der allgemeinen Forderung ausgeht, daß alle Eigenfunktionen antisymmetrisch sein müssen.

Nun ist es Zeit, an die Entwicklungen von 21 — 24 von F anzuschließen. Das Hundsche Aufbauprinzip ist dort bis auf die Voraussetzung abgeleitet worden, daß aus einem Term mit der azimutalen Quantenzahl l und der Zerlegung (17b) durch Berücksichtigung der Elektronenmagnete ein Singulett, Dublett, Triplett usw. entsteht, je nachdem $n - 2z + 1$ gleich 1, 2, 3 usw. ist. Wenn wir dies noch aus unseren Resultaten erschließen können, so haben wir das Aufbauprinzip vollkommen aus der Quantenmechanik abgeleitet.

Genauer gesagt, müssen wir zeigen, daß aus einem Term der spinfreien Differentialgleichung mit der Partitio (17b) und von der azimutalen Quantenzahl l genau t Terme entstehen, wo t die kleinere der beiden Zahlen $n - 2z + 1$ und $2l + 1$ ist. Die inneren Quantenzahlen j (definiert durch die Aufspaltung des Terms in $2j + 1$ Komponenten im Magnetfeld, j ist halb- oder ganzzahlig), dieser t Terme sind t aufeinanderfolgende Zahlen, die kleinste ist $\left|\frac{n-2z}{2} - l\right|$, die größte $\frac{n-2z}{2} + l$. (Siehe z. B. E. Back und A. Landé, Zeemaneffekt und Multiplettstruktur der Spektrallinien, Berlin 1925, oder F. Hund, Linienspektren und periodisches System der Elemente, Berlin 1927.)

§ 8. Wir müssen zu diesem Zwecke offenbar die innere Quantenzahl j einführen. Dies geschieht ähnlich, wie wir in F die azimutale Quantenzahl eingeführt haben. Haben wir nämlich eine (etwa antisymmetrische) Hyperfunktion, die Lösung des Eigenwertproblems

$$[H_1 + H_2]\Psi = \lambda\Psi$$

ist, so ist offenbar auch $O_{\mathfrak{R}}\Psi$ ($\mathfrak{R}$ eine beliebige Drehung) eine Lösung zum selben Eigenwert. Man schließt hieraus nach dem bekannten Schema, daß

$$O_{\mathfrak{R}}\Psi_\mu = \sum_\nu D_{\nu\mu}(\mathfrak{R})\Psi_\nu, \tag{19}$$

wo die $(D_{\nu\mu}(\mathfrak{R}))$ bis auf das Vorzeichen ($O_{\mathfrak{R}}$ ist nur bis auf das Vorzeichen bestimmt) eine Darstellung der dreidimensionalen Drehgruppe ist. Die Grade dieser Darstellungen sind (wie in I. angeführt) 1, 2, 3, 4, ... Dies ist auch die Anzahl der linear unabhängigen Eigenfunktionen, die zu diesem Term gehören, was wir mit $2j + 1$ bezeichnen. Es interessieren uns hauptsächlich diejenigen Matrizen, die den Drehungen um die Z-Achse (in dieser Richtung denken wir uns ein schwaches Magnetfeld eingeschaltet) entsprechen. Diese Matrizen können gleichzeitig auf die

Diagonalform gebracht werden und haben bei einer Drehung um den Winkel α die Gestalt (j halb oder ganz!):

$$\begin{pmatrix} e^{-ij\alpha} & 0 \cdots\cdots 0 & 0 \\ 0 & e^{-i(j-1)\alpha} \cdots\cdots 0 & 0 \\ \vdots & \vdots \qquad \ddots \qquad \vdots & \vdots \\ 0 & 0 \cdots\cdots e^{i(j-1)\alpha} & 0 \\ 0 & 0 \cdots\cdots 0 & e^{ij\alpha} \end{pmatrix} \tag{20}$$

Die $2j+1$ Eigenfunktionen, die zu einem Term mit der inneren Quantenzahl j zusammengefaßt werden, müssen also so in linear unabhängige angeordnet werden können, daß sie bei einer Drehung um die Z-Achse sich lediglich mit $e^{-ij\alpha}$, $e^{-i(j-1)\alpha}$, ..., $e^{im\alpha}$, ..., $e^{ij\alpha}$ multiplizieren. Die Hyperfunktion, die sich mit $e^{im\alpha}$ multipliziert, hat die „richtige magnetische Quantenzahl" m*.

§ 9. Aus jedem Term mit der Partitio (17 b) entstanden die antisymmetrischen Hyperfunktionen $\Psi_{\tau z}$, wo τ von $-(n-2z)$ bis $n-2z$ läuft (§ 6).

Die azimutale Quantenzahl der Eigenfunktionen $y(\mathfrak{r}_1, \ldots, \mathfrak{r}_n)$, von denen wir ausgingen, sei l. Dann sind $2l+1$ so viele vorhanden, als wir bisher betrachtet haben, da wir lediglich die Wirkung der Vertauschung der Elektronen berücksichtigten. Diese haben die „spinfreien magnetischen Quantenzahlen" $-l, -l+1, \ldots, \mu, \ldots, l-1, l$. Bezeichnen wir diese spinfreien magnetischen Quantenzahlen (abweichend von F) mit μ und fügen wir dies als oberen Index zu unserem $\Psi_{\tau z}$ zu, so ist

$$P_{\mathfrak{R}} \Psi^{\mu}_{\tau z} = e^{i\mu\alpha} \Psi^{\mu}_{\tau z}, \tag{21}$$

wenn $\mathfrak{R}$ eine Drehung um den Winkel α um die Z-Achse ist. Da wegen (7) und § 6

$$Q_{\mathfrak{R}} \Psi^{\mu}_{\tau z} = e^{i\tau\frac{\alpha}{2}} \Psi^{\mu}_{\tau z} \tag{21 a}$$

ist, ist also

$$O_{\mathfrak{R}} \Psi^{\mu}_{\tau z} = P_{\mathfrak{R}} Q_{\mathfrak{R}} \Psi^{\mu}_{\tau z} = e^{i\left(\mu+\frac{\tau}{2}\right)\alpha} \Psi^{\mu}_{\tau z}. \tag{21 b}$$

Die richtige magnetische Quantenzahl ist also $m = \mu + \frac{\tau}{2}$. Wir schreiben, um den Zusammenhang mit dem Vektorzusammensetzungsmodell besser zu sehen, diese m in eine Tabelle:

* m hat hier das entgegengesetzte Vorzeichen wie in F.

τ	$\mu = -l$	$-l+1$	$\dots$	$l-1$	l
$-n+2z$	$-\frac{n}{2}+z-l$	$-\frac{n}{2}+z-l+1$	$\cdots$	$-\frac{n}{2}+z+l-1$	$-\frac{n}{2}+z-l$
$-n+2z+2$	$-\frac{n}{2}+z-l+1$	$-\frac{n}{2}+z-l+2$	$\cdots$	$-\frac{n}{2}+z+l$	$-\frac{n}{2}+z-l+1$
$\vdots$	$\vdots$	$\vdots$		$\vdots$	$\vdots$
$n-2z-2$	$\frac{n}{2}-z-l-1$	$\frac{n}{2}-z-l$	$\cdots$	$\frac{n}{2}-z+l-2$	$\frac{n}{2}-z+l-1$
$n-2z$	$\frac{n}{2}-z-l$	$\frac{n}{2}-z-l+1$	$\cdots$	$\frac{n}{2}-z+l-1$	$\frac{n}{2}-z+l$

Jeder der $(2l+1)$ $(n-2z+1)$ Stellen dieser Tabelle entspricht eine Hyperfunktion, ihre magnetische Quantenzahl steht an der betreffenden Stelle der Tabelle. Es sind aber natürlich nicht etwa $(2l+1)$ $n-2z+1)$ Terme da, weil meistens mehrere Hyperfunktionen zu einem Term zusammengefaßt werden müssen. Wie diese Zusammenfassung zu geschehen hat, darüber gab uns § 8 Aufschluß: die $2j+1$ Hyperfunktionen eines Terms von der inneren Quantenzahl j haben die magnetischen Quantenzahlen $-j, -j+1, \dots, j$, zu solchen Komplexen müssen die Zahlen der Tabelle vereinigt werden.

Die Figuren 1 und 2 sollen diese Tabelle (abgesehen von der ersten Zeile und Spalte) für die beiden Fälle $l < \frac{n-2z}{2}$ bzw. $l > \frac{n-2z}{2}$ veranschaulichen. Die Stellen der Hyperfunktionen mit gleicher magnetischer Quantenzahl sind durch gestrichelte Linien verbunden.

Wir haben:

1 Hyperfunktion (H. F.) mit $m = \frac{n}{2} - z + l$,

2 linear unabhängige H. F. „ $m = \frac{n}{2} - z + l - 1$,

3 „ „ „ „ $m = \frac{n}{2} - z + l - 2$,

. . . . . .

3 linear unabhängige H. F. mit $m = -\frac{n}{2} + z - l + 2$,

2 „ „ „ „ $m = -\frac{n}{2} + z - l + 1$,

1 „ „ „ „ $m = -\frac{n}{2} + z - l$.

Dies ergibt einen Term mit $j = \frac{n}{2} - z + l$, einen mit $j = \frac{n}{2} - z + l - 1$, einen mit $j = \frac{n}{2} - z + l - 2$ usw. Das kleinste j hat in der Fig. 1 den Wert $\frac{n - 2z}{2} - l$, in der Fig. 2 den Wert $l - \frac{n - 2z}{2}$. Es entspricht dies genau dem in § 7 angekündigten Ergebnis.

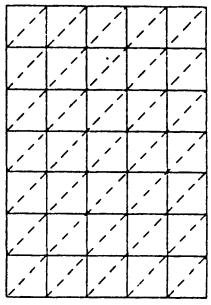


Fig. 1.

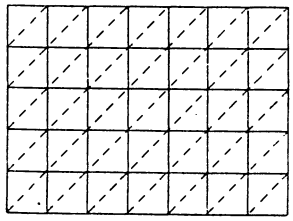


Fig. 2.

Damit ist wohl die wichtigste qualitative spektroskopische Regel abgeleitet. Man wird sich — abgesehen von der ungeheuren Leistungsfähigkeit der Quantenmechanik (in der klassischen Mechanik wäre es natürlich absolut unmöglich, eine ähnliche Rechnung durchzuführen oder z. B. die nun folgenden Intensitätsformeln streng zu berechnen) — darüber wundern, daß alles, wie man sagt, „durch die Luft" ging, d. h. ohne Bezugnahme auf die spezielle Form der Hamiltonschen Funktion, lediglich aus Symmetrieforderungen und der qualitativen Idee Paulis. Allerdings war dies bis zu einem gewisse Grade zu erwarten, daß nämlich eine dem Wesen der Dinge angepaßte Theorie diese qualitativen Erfahrungen auch ohne explizite Rechnung wird erschließen können. Dabei möchten wir bemerken, daß die verhältnismäßige Kompliziertheit unserer Überlegungen erstens daher rührt. daß das Resultat nicht ganz einfach ist und daß wir zweitens — um an Begriffsbildungen zu sparen — etwas mehr gerechnet haben, als unbedingt notwendig.

§ 10. Aus diesem Grunde auch möchten wir das Geschehene an einem Beispiel illustrieren*. Wir wählen dazu den in F angefangenen (Punkt 24) Fall. (Der Fall kommt praktisch nicht vor. Es kommt uns aber auch nur darauf an, ein möglichst einfaches Beispiel zu nehmen,

* Man kann § 10 ruhig überschlagen. Er dient nur dazu, das Vorangehende an einem Beispiel auszuführen, wobei von der Gruppentheorie kaum Gebrauch gemacht wird und an seine Stelle etwas umständliche Abzählungen treten.

das alle wesentlichen Eigenschaften beliebig komplizierter Beispiele schon aufweist.)

Wir betrachten also ein Li-Atom, mit einem Elektron in der L-Schale mit der azimutalen Quantenzahl 1 und zwei weiteren in der M-Schale, ebenfalls mit der azimutalen Quantenzahl 1.

Wenn wir von den Elektronenmagneten zuerst absehen, ist

$$L_3^3(\eta_1)P_1^{\mu_1}(\cos\vartheta_1)e^{i\mu_1\varphi_1}L_4^3(\eta_2)P_1^{\mu_2}(\cos\vartheta_2)e^{i\mu_2\varphi_2}L_4^3(\eta_3)P_1^{\mu_3}(\cos\vartheta_3)e^{i\mu_3\varphi_3} \quad (22)$$

eine Eigenfunktion dieses Zustandes, wenn man Polarkoordinaten für alle drei Elektronen einführt und η_1, η_2, η_3 proportional zu den drei Radiusvektoren, ϑ_1, ϑ_2, ϑ_3 die drei Polarwinkel und φ_1, φ_2, φ_3 die Azimute sind. Dabei kann μ_1, μ_2, μ_3 die drei Werte -1, 0, $+1$ annehmen*. Wir wollen zuerst alle linear unabhängigen Eigenfunktionen aufsuchen. Dazu müssen wir zwei Fälle unterscheiden.

I. $\mu_2 = \mu_3$, das sind die neun Möglichkeiten $\mu_1 = -1, 0, +1$; $\mu_2 = \mu_3 = -1, 0, +1$. Jede Möglichkeit gibt drei Eigenfunktionen, da wir noch η_1, ϑ_1, φ_1 mit η_2, ϑ_2, φ_2 oder mit η_3, ϑ_3, φ_3 vertauschen können. Bei einer Wechselwirkung der Elektronen erhalten wir** je einen „entarteten" und einen „symmetrischen" Term. Im ganzen erhalten wir folgende Terme:

$\mu_1 + \mu_2 + \mu_3 = \mu =$	-3	-2	-1	0	1	2	3
Antisymmetrisch	0	0	0	0	0	0	0
Entartet	1	1	2	1	2	1	1
Symmetrisch	1	1	2	1	2	1	1

wie eine einfache Abzählung ergibt (das sind 27 Eigenfunktionen, da die entarteten doppelt zu zählen sind).

II. $\mu \neq m_3$, das sind wieder neun Möglichkeiten. Jeder Möglichkeit entsprechen vier Terme: ein antisymmetrischer, zwei entartete, ein symmetrischer*. Wir haben also folgende Tabelle:

$\mu_1 + \mu_2 + \mu_3 = \mu =$	-3	-2	-1	0	1	2	3
Antisymmetrisch	0	1	2	3	2	1	0
Entartet	0	2	4	6	4	2	0
Symmetrisch	0	1	2	3	2	1	0

* Siehe z. B. E. Schrödinger, Abhandlungen zur Wellenmechanik. Leipzig 1927.

** E. Wigner, ZS. f. Phys. **40**, 492, 1926; W. Heisenberg, ebenda **41**, 239, 249—251, 1927; A. Unsöld, Ann. d. Phys. **82**, 355, 1927.

Im ganzen haben wir also folgende Terme:

$\mu =$	-3	-2	-1	0	1	2	3
Antisymmetrisch	0	1	2	3	2	1	0
Entartet	1	3	6	7	6	3	1
Symmetrisch	1	2	4	4	4	2	1

Das sind 81 linear unabhängige Eigenfunktionen, wie man auch direkt abzählen kann. Wenn wir die durch die Drehgruppe geforderten Zusammenfassungen vornehmen, so erhalten wir:

$l =$	3 (7fache Terme)	2 (5fache Terme)	1 (3fache Terme)	0 (einfache Terme)
Antisymmetrisch . . .	0	1	1	1
Entartet	1	2	3	1
Symmetrisch	1	1	2	0

Wenn wir die Wechselwirkung der Elektronen berücksichtigen, wird also der ursprüngliche Term in diese 14 Terme aufspalten. Bisher ist dies alles eine Wiederholung aus *F*.

Wenn wir jetzt die Elektronenmagnete einführen, so entstehen aus allen 81 linear unabhängigen Eigenfunktionen je 8 Hyperfunktionen. Wenn wir noch nachweisen können, daß ein (natürlich nur in den Elektronenschwerpunkten) antisymmetrischer Term der vorigen Tabelle mit $l = 0, 1, 2$ ein Quartett von *S*-, *P*-, *D*-Termen, je ein entarteter mit $l = 0, 1, 2, 3$ je ein Dublett von *S*-, *P*-, *D*-, *B*-Termen liefern, während die symmetrischen gar keine antisymmetrischen Hyperfunktionen liefern, so sind wir am Ziel.

Da es sich ohnehin nur um ein Beispiel handelt, zeigen wir dies nur an einem einzigen Term der Tabelle, an einem entarteten Term mit $l = 1$.

Wir haben, wenn wir die Elektronenmagnete einführen, folgende linear unabhängigen Eigenfunktionen (der obere Index bedeutet die spinfreie magnetische Quantenzahl, die unteren Indizes haben dieselbe Bedeutung wie in § 3):

$$\Phi^{-1}_{-1-1-1} \quad \Phi^{-1}_{-1-11} \quad \Phi^{-1}_{-11-1} \quad \Phi^{-1}_{1-1-1} \quad \Phi^{-1}_{11-1} \quad \Phi^{-1}_{1-11} \quad \Phi^{-1}_{-111} \quad \Phi^{-1}_{111}$$

$$\Phi^{0}_{-1-1-1} \quad \Phi^{0}_{-1-11} \quad \Phi^{0}_{-11-1} \quad \Phi^{0}_{1-1-1} \quad \Phi^{0}_{11-1} \quad \Phi^{0}_{1-11} \quad \Phi^{0}_{-111} \quad \Phi^{0}_{111}$$

$$\Phi^{1}_{-1-1-1} \quad \Phi^{1}_{-1-11} \quad \Phi^{1}_{-11-1} \quad \Phi^{1}_{1-1-1} \quad \Phi^{1}_{11-1} \quad \Phi^{1}_{1-11} \quad \Phi^{1}_{-111} \quad \Phi^{1}_{111}$$

Das sind 24 Eigenfunktionen. Hierbei wurde aber so getan, als ob zum ursprünglichen Terme nur eine Eigenfunktion gehörte; da er in Wahrheit zweifach entartet ist, erhöht sich die Anzahl von 24 auf 48. Aus den

acht $\Phi^{\mu}_{s_1 s_2 s_3}$ einer Reihe der Tabelle lassen sich Hyperfunktionen bilden, die bei der Vertauschung der Elektronen ($\mathfrak{r}$ und s) sich nach irreduziblen Darstellungen der symmetrischen Gruppe transformieren *.

$\Phi^{\mu}_{-1\,-1\,-1}$	ist entartet,
$\Phi^{\mu}_{-1\,-1\,1}$, $\Phi^{\mu}_{-1\,1\,-1}$, $\Phi^{\mu}_{1\,-1\,-1}$	gibt 1 symm., 2 entartete, 1 antisymm. Hyperf.,
$\Phi^{\mu}_{1\,1\,-1}$, $\Phi^{\mu}_{1\,-1\,1}$, $\Phi^{\mu}_{-1\,1\,1}$	„ 1 „ 2 „ 1 „ „
$\Phi^{\mu}_{1\,1\,1}$	ist entartet.

Wir haben im ganzen zwei antisymmetrische Hyperfunktionen aus den Φ^{μ}. Da $\mu = -1, 0, 1$ sein kann, haben wir im ganzen sechs solche Hyperfunktionen. Dabei nimmt $\mu + \frac{\Sigma\tau}{2}$ die Werte $-\frac{3}{2}, -\frac{1}{2}, -\frac{1}{2}, \frac{1}{2}, \frac{1}{2}, \frac{3}{2}$ an. Das ist ein Term mit $j = \frac{3}{2}$ und einer mit $j = \frac{1}{2}$, wir haben also ein Dublett mit den richtigen inneren Quantenzahlen.

Auswahl- und Intensitätsregeln.

§ 11. Wir wenden uns nun den Auswahl- und Intensitätsregeln zu. Dabei ist das Folgende zu beobachten: Die Paulische Theorie des rotierenden Elektrons gilt nur für den Fall, daß die Geschwindigkeit des Elektrons gegenüber der Lichtgeschwindigkeit zu vernachlässigen ist. Wie das rotierende Elektron bzw. ein System von rotierenden Elektronen relativistisch zu behandeln ist, ist zurzeit eine offene Frage**. Es hätte also für uns nicht viel Sinn, an Stelle der elektrischen Feldstärken mit den Vektorpotentialen zu rechnen, obwohl es sicher ist, daß dies der richtige Weg wäre ***.

Bevor wir die Auswahlregeln ableiten können, müssen wir die Symmetrieverhältnisse etwas genauer betrachten, als dies bisher notwendig war. Daß das Energieeigenwertproblem gegenüber einer Drehung invariant ist, ist selbstverständlich und wurde auch im vorangehenden schon benutzt. Wie steht es aber mit der Spiegelung? Wenn wir ein System, das sich im Laufe der Zeit gar nicht ändert (also eine Eigenfunktion des Energieoperators darstellt), in einem Spiegel ansehen, so muß die Zustandsfunktion des so betrachteten Zustandes auch eine Eigenfunktion des Energieoperators (und zwar zum selben Eigenwert) sein. Im Raume an und für sich ist ja rechts und links nicht bevorzugt.

* Siehe Zitate ** auf S. 87, insbesondere das zweite. (Die Arbeiten enthalten keine gruppentheoretischen Rechenmethoden.) Es ist zu beachten, daß die Φ bei der Vertauschung der $\mathfrak{r}$ allein sich wie entartete Funktionen benehmen.

** Siehe jedoch P. A. M. Dirac, Proc. Roy. Soc. (A) **117**, 610, 1928.

*** Derselbe, ebenda (A) **114**, 710, 1927.

Wie sich die Hyperfunktion eines Zustandes bei der Verdrehung desselben transformiert, haben wir in der Einleitung (und hauptsächlich in I.) gesehen. Aber auch was bei einer Spiegelung geschieht, ist leicht zu übersehen. Im Koordinatensystem, das durch Spiegelung im Ursprungspunkt aus dem alten hervorgeht, wo also $X' = -X$, $Y' = -Y$, $Z' = -Z$ ist, entspricht einem Zustand mit der Hyperfunktion

$$\Psi(x_1, y_1, z_1, \ldots, x_n, y_n, z_n; s_1, \ldots, s_n)$$

der Zustand mit der Hyperfunktion $\overline{\Psi}$

$$\overline{\Psi}(x_1, y_1, z_1, \ldots, x_n, y_n, z_n; s_1, \ldots, s_n)$$
$$= \Psi(-x_1, -y_1, -z_1, \ldots, -x_n, -y_n, -z_n; s_1, \ldots, s_n). \tag{23}$$

Dies geht daraus hervor, daß ein Nordpol in einem Spiegel zu einem Südpol wird: $|\Psi(x_1, \ldots, z_n; s_1, \ldots, s_n)|^2$ gibt die Wahrscheinlichkeit an, daß das ν-te Elektron an der Stelle x_ν, y_ν, z_ν mit dem Drehimpuls in der Z-Achse s_ν $(= \pm 1)$ gefunden wird. Dem entspricht im zweiten Koordinatensystem, daß das ν-te Elektron an der Stelle $-x_\nu$, $-y_\nu$, $-z_\nu$, aber ebenfalls mit dem Drehimpuls s_ν (in der Z'-Achse!) gefunden wird. Hieraus folgt

$$|\overline{\Psi}(x_1, \ldots, z_n; s_1, \ldots, s_n)|^2 = |\Psi(-x_1, \ldots, -z_n; s_1, \ldots, s_n)|^2, \tag{23a}$$

womit (23) bis auf die „Phasen" bewiesen ist, durch weitere Versuche beweist man sie mühelos ganz.

Das Energieeigenwertproblem läßt also folgende Operationen zu:

I. Vertauschung aller Koordinaten beliebiger Elektronen. Indessen wird dies uns hier nicht mehr interessieren, da wir uns auf antisymmetrische Lösungen beschränken.

II. Weiter ist mit Ψ auch $O_{\mathfrak{R}}\Psi$ eine Lösung des Eigenwertproblems.

III. Zuletzt ist mit $\Psi(x_1, \ldots, z_n; s_1, \ldots, s_n)$ auch

$$\Psi(-x_1, \ldots, -z_n; s_1, \ldots, s_n)$$

Lösung.

Da die Operationen von I., II., III. vertauschbar sind, kann man sie der Reihe nach einzeln betrachten. Die Darstellung, die den Operationen von I. entspricht, ist immer die antisymmetrische.

Die Darstellungen, die denen von II. entsprechen, sind natürlich die Darstellungen der dreidimensionalen Drehgruppe. Wir sahen, daß bei Elementen mit ungerader Ordnungszahl die zweideutigen, bei geraden Elementen die eindeutigen auftreten. Die Dimension der Darstellung bezeichneten wir mit $2j + 1$ und j war die innere Quantenzahl.

§ 12. Die Operation von III. war auch Substitution des Problems $H_1 y(\mathfrak{r}_1, \ldots, \mathfrak{r}_n) = \lambda y(\mathfrak{r}_1, \ldots, \mathfrak{r}_n)$. Wir nannten in F die Funktion y

normal oder gespiegelt*, je nachdem der Substitution von III. die Matrix (-1^l) oder (-1^{l+1}) entsprach. Bei unserem jetzigen Standpunkt ist es zweckmäßiger, eine andere Unterscheidung einzuführen, nämlich die Terme in positive [der Operation III. entspricht die Matrix (1)] und negative [mit der Matrix (— 1)] einzuteilen. Die richtiggestellte** Auswahlregel von F lautet dann (l ändert sich um ± 1 oder 0): positive Terme kombinieren nur mit negativen, negative nur mit positiven. Eine Hyperfunktion hat positiven oder negativen Charakter, je nachdem der Term, aus dem sie entstanden ist, diesen oder jenen Charakter hatte. Die gesperrt gedruckte Auswahlregel gilt auch bei Berücksichtigung der Elektronenmagnete streng. Es verschwindet nämlich etwa

$$\sum_{s_1, \ldots, s_n} \int x_1 \Psi(x_1, \ldots, z_n; s_1, \ldots, s_n) \widetilde{\Psi}'(x_1, \ldots, z_n; s_1, \ldots, s_n), \tag{24}$$

wenn sowohl Ψ wie Ψ' positive, oder beide negative Hyperfunktionen sind. Man sieht dies leicht ein, indem man für die x_ν, y_ν, z_ν die Variable $-x_\nu$, $-y_\nu$, $-z_\nu$ einführt.

Wir wollen jetzt die Wirkung eines schwachen magnetischen Feldes in der Z-Achse betrachten. Dieses zerstört die Symmetrie II. unserer Differentialgleichung, so daß sie nur die Achsensymmetrie in der Z-Achse behält. Da die irreduziblen Darstellungen der übrigbleibenden Gruppe sämtlich von der Dimension 1 sind, spaltet der $2j+1$-fache Eigenwert in $2j+1$ einfache Terme auf. Wir haben $2j+1$ Eigenfunktionen, die sich bei der Drehung um die Z-Achse um den Winkel α bzw. mit $e^{-ij\alpha}$, $e^{-i(j-1)\alpha}, \ldots, e^{ij\alpha}$ multiplizieren. Der Term, dessen Hyperfunktion sich mit $e^{im\alpha}$ multipliziert, hat die magnetische Quantenzahl m.

m ändert sich bei einem Übergang nicht, wenn die Strahlung in der Z-Achse polarisiert ist. Betrachten wir nämlich das Integral

$$\int z_1 \Psi_m \widetilde{\Psi}'_{m'} = J \tag{25}$$

von einem um den Winkel α verdrehten Koordinatensystem, d. h. bilden wir das Integral

$$\int O_{\mathfrak{R}} z_1 \Psi_m \cdot \widetilde{O_{\mathfrak{R}} \Psi'_{m'}} = J',$$

wo $\mathfrak{R}$ eine Drehung um die Z-Achse um den Winkel α ist, so finden wir dafür

$$\int z_1 \Psi_m e^{-im\alpha} \widetilde{\Psi}'_{m'} e^{im'\alpha} = J'. \tag{25a}$$

* Diese Einteilung ist mit der Laporte-Russellschen in ungestrichene und gestrichene identisch.

** ZS. f. Phys. 45, 601, 1927.

J muß bis auf das Vorzeichen gleich J' sein (da die Darstellung $O_{\mathfrak{R}}$ orthogonal ist), dies ist aber nur der Fall, wenn $m = m'$ ist, sonst muß $J = J' = 0$ sein.

Bei einer Strahlung, die in der XY-Ebene polarisiert ist, muß sich m um ± 1 ändern. Die Integrale

$$\int (x_1 + i y_1) \Psi_m \widetilde{\Psi}'_{m'} \quad \text{und} \quad \int (x_1 - i y_1) \Psi_m \widetilde{\Psi}'_{m'} \tag{25b}$$

verschwinden nämlich beide, wenn $|m - m'| \neq 1$. Daher verschwindet dann auch $\int x_1 \Psi_m \widetilde{\Psi}'_{m'}$ und $\int y_1 \Psi_m \widetilde{\Psi}'_{m'}$. Damit haben wir die Auswahlregeln für die magnetische Quantenzahl.

Jetzt ist es auch sehr leicht, die Regel $\Delta j = 0, \pm 1$ herzuleiten. Betrachten wir den Übergang von j zu j', wo etwa $j' \geqq j + 2$ seien. Die $2j' + 1$ Eigenfunktionen von der inneren Quantenzahl j' seien $\Psi'_{-j'}$, $\Psi'_{-j'+1}, \ldots, \Psi'_{j'}$, die $2j + 1$ von der inneren Quantenzahl j seien Ψ_{-j}, $\Psi_{-j+1}, \ldots, \Psi_j$. Da $j' \geqq j + 2$, existiert vom Zustand $\Psi'_{j'}$ kein Übergang (weder in der XY-Ebene, noch in der Z-Richtung polarisiert) zu irgend einem Ψ_μ. Die Integrale

$$\int (\alpha x_k + \beta y_k + \gamma z_k) \Psi_\mu \widetilde{\Psi}'_{j'} = 0 \tag{26}$$

verschwinden also bei beliebigem α, β, γ alle. Es gilt dies also auch für die Integrale

$$\int (\alpha' x_k + \beta' y_k + \gamma' z_k) \cdot O_{\mathfrak{R}} \Psi_\mu \cdot \widetilde{O_{\mathfrak{R}} \Psi'_{j'}} = 0 \tag{26a}$$

mit beliebigem $\mathfrak{R}$, also auch

$$\int (\alpha' x_k + \beta' y_k + \gamma' z_k) \Psi_\nu \widetilde{O_{\mathfrak{R}} \Psi'_{j'}} = 0, \tag{26b}$$

da ja die $O_{\mathfrak{R}} \Psi_\mu$ Linearkombinationen der Ψ_ν sind und umgekehrt. Nun ist

$$O_{\mathfrak{R}} \Psi'_{j'} = \sum_l D_{lj'}(\mathfrak{R}) \Psi'_l, \tag{26c}$$

was mit $\widetilde{D_{\lambda j'}}(\mathfrak{R})$ multipliziert und über $\mathfrak{R}$ integriert

$$\int_{\mathfrak{R}} \widetilde{D_{\lambda j'}}(\mathfrak{R}) O_{\mathfrak{R}} \cdot \Psi'_{j'} = c \Psi'_\lambda \tag{26d}$$

ergibt.

Multiplizieren wir also (26b) mit $D_{\lambda j'}(\mathfrak{R})$ und integrieren über $\mathfrak{R}$, so erhalten wir*

$$\int (\alpha' x_k + \beta' y_k + \gamma' z_k) \Psi_\nu \widetilde{\Psi}'_\lambda = 0. \tag{27}$$

Ein Übergang mit $|j' - j| \geqq 2$ kommt also nicht vor, j ändert sich um 0 oder ± 1. Außerdem zeigt man leicht, daß ein Übergang von $j = 0$ zu $j = 0$ nicht vorkommt.

* Da (26b) identisch in α', β', γ' gilt, können wir sie bei der Integration konstant halten.

§ 13. Hiermit haben wir also gesehen, daß die Auswahlregel für die magnetische Quantenzahl m, für die innere Quantenzahl j und die „Laportesche Regel“ auch bei Mitberücksichtigung der Elektronenmagnete gilt. Zu beachten ist natürlich, daß die Berücksichtigung der Elektronenmagnete bei ganz hohen Geschwindigkeiten nicht ganz korrekt sein dürfte, worauf wir schon in § 11 hingewiesen haben. Immerhin ist es bemerkenswert, daß uns keine einzige Durchbrechung dieser Regeln in irdischen Lichtquellen bekannt ist.

Wenn H_2 im Verhältnis zu H_1 klein ist, so gelten auch die Regeln für die Änderung von l noch gut, auch gilt dann das Verbot der Interkombination verschiedener Multiplettsysteme. Bei den normalen Serienspektren, die also aus (im Sinne von F) normalen Termen bestehen, gilt die Auswahlregel $\Delta l = \pm 1$ noch streng, da eine Durchbrechung dieser Regel gleichzeitig die Auswahlregel für j oder die Laportesche Regel verletzen würde.

Leicht zeigt man auch die einfachen Summensätze, die in F (18) und* vorkommen und die genau gelten, wenn man l durch j ersetzt. Sie regeln das Verhältnis der Intensitäten der verschiedenen Zeemankomponenten derselben Linie. Man findet diese Formeln auch bei M. Born, W. Heisenberg und P. Jordan** und bei P. A. M. Dirac***.

Dasselbe gilt vom Verschwinden des linearen Starkeffekts und den Formeln für den quadratischen, die in F Punkt 18 gegeben sind. Diese gelten also — entgegen der dortigen Annahme — streng bei beliebigem Multiplettsystem, wenigstens solange die Aufspaltungen nicht allzu groß werden. Zu beachten ist natürlich, daß auch m bei den ungeraden Elementen halbzahlig ist. Zusammenfassend kann man also sagen, daß ein linearer Starkeffekt nur im Falle der zufälligen Entartung (Wasserstoffatom) auftritt und das die quadratische Aufspaltung nach der Formel

$$\Delta E = (A + B m^2)\, \mathfrak{E}^2 \tag{28}$$

vor sich geht, wobei A und B von Term zu Term verschiedene Konstanten sind. Es ist bemerkenswert, daß diese Formel bei beliebigen Kopplungsverhältnissen in Strenge gilt (im Gegensatz z. B. zu der Landéschen g-Formel), allerdings kann man den Wert der Koeffizienten A und B hier nicht angeben.

---

* ZS. f. Phys. **45**, 601, 1927.
** Ebenda **35**, 557, 1926.
*** Proc. Roy. Soc. (A) **111**, 281, 1926.

§ 14. Wesentlich komplizierter ist die Ableitung der Summensätze, die sich auf das Verhältnis der Intensität verschiedener Linien eines Multipletts beziehen. Sie gelten nur, wenn die Multiplettaufspaltung klein ist, und man muß zu ihrer Ableitung auf die Überlegungen zurückgreifen, die wir zur Ableitung des Aufbauprinzips der Serienspektren angewendet haben. Auch erfordern sie die Kenntnis einiger mathematischer Formeln, die in der Literatur nicht vorkommen und die einfach hinzuschreiben nicht angängig erscheint. Ähnliches gilt von den Landéschen g-Formeln. Wir möchten daher die Behandlung dieser Fragen auf den „dritten Teil" verschieben.

Göttingen und Berlin, Februar 1928.

Zur Erklärung einiger Eigenschaften der Spektren aus der Quantenmechanik des Drehelektrons. Dritter Teil

J. von Neumann und E. P. Wigner

Zeitschrift für Physik *51*, 844–858 (1928)

(Eingegangen am 19. Juni 1928.)

Es wird — ausgehend von der Lösung der spinfreien (aber die spin-Freiheitsgrade der Elektronen berücksichtigenden) Differentialgleichung für ein Atom, d. h. der Schrödingergleichung, in der die magnetischen Momente der Elektronen nicht eingeführt sind — die erste Näherung für die Eigenfunktionen mit Spin abgeleitet. Hieraus ergibt sich die Landésche g-Formel, die Burger-Dorgeloschen Summensätze usw.

Dritter Teil.

§ 1. Nach dem im ersten Teile dieser Arbeit * gegebenen Programm sollen hier diejenigen Gesetzmäßigkeiten der Spektren besprochen werden, die bei Mitberücksichtigung der Elektronenmagnete nach der Quantenmechanik nicht mehr streng, sondern nur, wenn die Multiplettaufspaltung klein ist, gelten. Hierher gehört vor allem die Landésche g-Formel, von den Burger-Dorgeloschen Summensätzen diejenigen, die das Verhältnis der Intensitäten verschiedener Linien eines Multipletts regeln (während die Sätze, die das Verhältnis der Intensitäten der einzelnen Zeemankomponenten einer Linie regeln, streng gelten und in II besprochen worden sind **), und schließlich die Intervallregeln. Was die Intervallregel anbetrifft, so ist es bekanntlich bei unserem unrelativistischen Standpunkt nicht möglich, die richtige absolute Größe dieser Aufspaltung herauszubekommen; und nachdem die Berechnung mit Hilfe der Formeln des § 2 elementar ist, möchten wir auf eine nähere Ausführung verzichten.

Von bisherigen Untersuchungen über diese Frage ist zunächst die Arbeit von W. Heisenberg und P. Jordan *** zu nennen, die den Fall von Dubletts und die von P. A. M. Dirac ****, die allgemeine Multipletts behandelt. Indessen sind beide Arbeiten unter der Voraussetzung geschrieben, daß wir es mit einem Leuchtelektron zu tun haben, das sich

* ZS. f. Phys. **47**, 203, 1928; diese Arbeit soll als I zitiert werden. Der ebenfalls häufig benutzte zweite Teil (ebenda **49**, 73, 1928) als II, eine etwas ältere Arbeit des einen von uns: ZS. f. Phys. **43**, 624, 1927, als F.

** Sie sollen trotzdem von einem etwas anderen Standpunkt aus noch einmal kurz besprochen werden.

*** ZS. f. Phys. **37**, 263, 1926.

**** Proc. Roy. Soc. **111**, 281, 1926.

unter dem Einfluß eines geladenen und mit einem magnetischen Moment behafteten Kerns, des „Rumpfes", bewegt. Wir werden diese Annahme nicht machen, sondern das Atom regelrecht als Mehrelektronenproblem auffassen. Wir möchten bemerken, daß der hier eingeschlagene Weg nicht der eleganteste ist, durch den man zu den erwähnten Regeln gelangt. Wenn wir doch diesen wählen, so geschieht das, weil dieser Weg der naheliegendste ist und bei der Berechnung anderer Größen auch immer beschritten werden kann. Wir berechnen nämlich zuerst die „erste Näherung" für die Eigenfunktion und dann mit ihrer Hilfe die fraglichen Matrixelemente. Es zeigt sich, daß man dabei keine Integrale wirklich auszuwerten hat.

Um auch eine andere Methode zu zeigen, sollen mit ihrer Hilfe die „streng gültigen Summensätze" nochmals abgeleitet werden. Wir möchten empfehlen, zuerst den mathematischen Anhang etwas anzusehen*.

§ 2. Die erste Näherung für die Eigenfunktionen. In II gingen wir von einem Eigenwert der spinfreien Differentialgleichung aus, der die azimutale Quantenzahl l und die Darstellungseigenschaft gegenüber der symmetrischen Gruppe (Vertauschen der Elektronen) z hatte. Die dazu gehörigen Eigenfunktionen seien $\psi_{\mu\zeta}(\mathfrak{r}_1, \ldots, \mathfrak{r}_n)$, so daß

$$P_{\mathfrak{R}}\, \psi_{\mu\zeta}(\mathfrak{r}_1', \ldots, \mathfrak{r}_n') = \psi_{\mu\zeta}(\mathfrak{r}_1, \ldots, \mathfrak{r}_n) = \sum_{\nu} D^l_{\nu\mu}(\mathfrak{R})\, \psi_{\nu\zeta}(\mathfrak{r}_1', \ldots, \mathfrak{r}_n') \quad (1)$$

(wenn $\mathfrak{R}$ eine Drehung ist und $\mathfrak{r}_1', \ldots, \mathfrak{r}_n'$ aus $\mathfrak{r}_1, \ldots, \mathfrak{r}_n$ durch diese Drehung hervorgeht) und

$$P_{\mathfrak{S}}\, \psi_{\mu\zeta}(\mathfrak{r}_{\alpha_1}, \ldots, \mathfrak{r}_{\alpha_n}) = \psi_{\mu\zeta}(\mathfrak{r}_1, \ldots, \mathfrak{r}_n) = \sum_{\eta} \Delta^z_{\eta\zeta}(\mathfrak{S})\, \psi_{\mu\eta}(\mathfrak{r}_{\alpha_1}, \ldots, \mathfrak{r}_{\alpha_n}), \quad (1\,\mathrm{a})$$

wenn $\mathfrak{S}$ die Permutation der Zahlen $1, 2, \ldots, n$ in die Zahlen $\alpha_1, \alpha_2, \ldots, \alpha_n$ ist. Aus jedem dieser $\psi_{\mu\zeta}(\mathfrak{r}_1, \ldots, \mathfrak{r}_n)$ bildeten wir dann 2^n Hyperfunktionen, und zwar war

$$\psi_{\mu\zeta\tau_1, \ldots, \tau_n}(\mathfrak{r}_1, \ldots, \mathfrak{r}_n, s_1, \ldots, s_n) = 0, \quad (2)$$

wenn nicht $s_1 = \tau_1$, $s_2 = \tau_2, \ldots, s_n = \tau_n$ und

$$\psi_{\mu\zeta\tau_1, \ldots, \tau_n}(\mathfrak{r}_1, \ldots, \mathfrak{r}_n, \tau_1, \ldots, \tau_n) = \psi_{\mu\zeta}(\mathfrak{r}_1, \ldots, \mathfrak{r}_n). \quad (2\,\mathrm{a})$$

Dies war der Inhalt des § 1 in II.

---

* Die $2l+1$-dimensionale Darstellung der Drehgruppe bezeichnen wir mit $\left(D^l_{\varkappa\lambda}(\mathfrak{R})\right)$, wo $\varkappa$ und λ von $-l$ bis $+l$ laufen. Um bei den häufig vorkommenden Summationen die Summationsgrenzen nicht immer explizite hinschreiben zu müssen, bestimmen wir, daß wir unter $D^l_{\varkappa\lambda}$, wenn $|\varkappa| > l$ oder $|\lambda| > l$, die Zahl 0 verstehen. Dadurch ist es möglich, die Summationen immer von $-\infty$ bis ∞ auszudehnen, so daß wir nur in § 5 an einigen Stellen auf die Grenzen der Summation achten müssen.

Dann bildeten wir in II alle antisymmetrischen Linearkombinationen der $\psi_{\mu\zeta\tau_1,\ldots,\tau_n}(\mathfrak{r}_1,\ldots,\mathfrak{r}_n, s_1,\ldots,s_n)$. War insbesondere $l = 0$, so ist $\mu = 0$ und die antisymmetrischen Linearkombinationen seien

$$\begin{aligned}\Psi_m(\mathfrak{r}_1,\ldots,\mathfrak{r}_n; s_1,\ldots,s_n) &= \sum_{\zeta\tau_1\ldots\tau_n} {}^zC_{m\zeta;\tau_1,\ldots,\tau_n}\,\psi_{0\zeta\tau_1,\ldots,\tau_n}(\mathfrak{r}_1,\ldots,\mathfrak{r}_n; s_1,\ldots,s_n)\\ &= \sum_{\zeta} {}^zC_{m\zeta;s_1,\ldots,s_n}\,\psi_{0\zeta}(\mathfrak{r}_1,\ldots,\mathfrak{r}_n),\end{aligned} \tag{3}$$

wo wir zur Eindeutigkeit die Darstellungseigenschaft z der $\psi_{\mu\zeta}$ gegenüber der symmetrischen Gruppe noch als oberen Index bei ${}^zC_{m\zeta;\tau_1,\ldots,\tau_n}$ hinzufügen*. Solche antisymmetrischen Linearkombinationen hatten wir $n - 2z + 1$, m soll daher von $-(\frac{1}{2}n - z)$ bis $\frac{1}{2}n - z$ laufen. Die $n - 2z + 1$ Hyperfunktionen von (3) mußten zu einem Term mit $j = \frac{1}{2}n - z$ zusammengefaßt werden, so daß wir annehmen können, daß m die magnetische Quantenzahl ist. Aus der Orthogonalität

$$\sum_{\tau_1\ldots\tau_n}\int \Psi_m\,\widetilde{\Psi}_m = \delta_{mm'}$$

folgt

$$\sum_{\zeta\tau_1\ldots\tau_n} {}^zC_{m\zeta;\tau_1,\ldots,\tau_n}\,{}^z\widetilde{C}_{m'\zeta;\tau_1,\ldots,\tau_n} = \delta_{mm'}. \tag{4}$$

Außerdem ist nach dem Vorangehenden

$$O_{\mathfrak{R}}\,\Psi_m(\mathfrak{r}'_1,\ldots,\mathfrak{r}'_n; s_1,\ldots,s_n) = \sum_{m'} D^j_{m'm}(\mathfrak{R})\,\Psi_{m'}(\mathfrak{r}'_1,\ldots,\mathfrak{r}'_n; s_1,\ldots,s_n),$$

und mit Hilfe von (3) in II:

$$O_{\mathfrak{R}}\,\Psi_m(\mathfrak{r}'_1,\ldots,\mathfrak{r}'_n; s_1,\ldots,s_n) = \sum_{t_1\ldots t_n} a_{s_1,\ldots,s_n;t_1,\ldots,t_n}\,\Psi_m(\mathfrak{r}_1,\ldots,\mathfrak{r}_n; s_1,\ldots,s_n)$$

folgt hieraus durch (3)**:

$$\sum_{\sigma_1\ldots\sigma_n} a_{\tau_1,\ldots,\tau_n;\sigma_1,\ldots,\sigma_n}\,{}^zC_{m\zeta;\sigma_1,\ldots,\sigma_n} = \sum_{m'} D^{1/2\,n-z}_{m'm}\cdot{}^zC_{m'\zeta;\tau_1,\ldots,\tau_n}. \tag{5}$$

Nun gibt aber (3) nicht nur im Falle von $l = 0$, sondern natürlich immer alle antisymmetrischen Hyperfunktionen. Diese sind also

$$\Phi_{k\varkappa}(\mathfrak{r}_1,\ldots,\mathfrak{r}_n; \tau_1,\ldots,\tau_n) = \sum_{\zeta} {}^zC_{k\zeta;\tau_1,\ldots,\tau_n}\,\psi_{\varkappa\zeta}(\mathfrak{r}_1,\ldots,\mathfrak{r}_n), \tag{6}$$

wo für die ${}^zC_{k\zeta\tau_1,\ldots,\tau_n}$ Gleichungen (4) und (5) gelten. Aus diesen $\Phi_{k\varkappa}$ müssen wir nun Linearkombinationen bilden, die sich bei der Operation $O_{\mathfrak{R}}$

* Die ${}^zC_{m\zeta;\tau_1,\ldots,\tau_n}$ sind dann reine Zahlen, die nur von ihren Indizes abhängen. Wir bevorzugen auch sonst immer das Ausschreiben aller Indizes und Argumente, weil dies zwar die Formeln etwas verlängert, dem Leser aber viel unnötiges Nachdenken erspart.

** Die Formeln (4) und (5) werden wir im folgenden öfters verwenden.

nach irreduziblen Darstellungen der Drehgruppe transformieren. Wir bilden zunächst

$$O_{\mathfrak{R}}\,\Phi_{k\varkappa}(\mathfrak{r}_1',\ldots,\mathfrak{r}_n';\tau_1,\ldots,\tau_n)=\sum_{\sigma_1\ldots\sigma_n}a_{\tau_1,\ldots,\tau_n;\,\sigma_1,\ldots,\sigma_n}(\mathfrak{R})\cdot\Phi_{k\varkappa}(\mathfrak{r}_1,\ldots,\mathfrak{r}_n;\sigma_1,\ldots,\sigma_n)$$
$$=\sum_{\sigma_1\ldots\sigma_n}\sum_{\zeta}a_{\tau_1,\ldots,\tau_n;\,\sigma_1,\ldots,\sigma_n}\,{}^zC_{k\zeta;\,\sigma_1,\ldots,\sigma_n}\,\psi_{\varkappa\zeta}(\mathfrak{r}_1,\ldots,\mathfrak{r}_n)$$
$$=\sum_{\zeta k'}D^{1/2n-z}_{k'k}(\mathfrak{R})\,{}^zC_{k'\zeta;\,\tau_1,\ldots,\tau_n}\,\psi_{\varkappa\zeta}(\mathfrak{r}_1,\ldots,\mathfrak{r}_n)\tag{7}$$

wegen (5). Dies ergibt nunmehr mit Rücksicht auf (1):

$$O_{\mathfrak{R}}\,\Phi_{k\varkappa}(\mathfrak{r}_1',\ldots,\mathfrak{r}_n';\tau_1,\ldots,\tau_n)$$
$$=\sum_{\zeta k'}\sum_{\varkappa'}D^{1/2n-z}_{k'k}\,{}^zC_{k'\zeta;\,\tau_1,\ldots,\tau_n}\,D^{l}_{\varkappa'\varkappa}\,\psi_{\varkappa'\zeta}(\mathfrak{r}_1',\ldots,\mathfrak{r}_n')$$
$$=\sum_{k'\varkappa'}\sum_{\iota}s^{\iota}_{k'+\varkappa',k'}(\tfrac{1}{2}n-z,l)\,\tilde{s}^{\iota}_{k+\varkappa,k}(\tfrac{1}{2}n-z,l)\,D^{\iota}_{k'+\varkappa',k+\varkappa}$$
$$\cdot\,\Phi_{k'\varkappa'}(\mathfrak{r}_1',\ldots,\mathfrak{r}_n';\tau_1,\ldots,\tau_n),\tag{7a}$$

wegen (III) im Anhang und (6). Wenn wir für $\varkappa=m-k$ und für $\varkappa'=m'-k'$ setzen, erhalten wir nach Multiplikation mit s^{j}_{mk} und Summation über k unter Beachtung von (II) des Anhanges:

$$O_{\mathfrak{R}}\sum_{k}s^{j}_{mk}\,\Phi_{k,\,m-k}=\sum_{m'}D^{j}_{m'm}(\mathfrak{R})\sum_{k'}s^{j}_{m'k'}\,\Phi_{k',\,m'-k'}.\tag{8}$$

Die Funktionen

$$\Psi^{j}_{m}(\mathfrak{r}_1,\ldots,\mathfrak{r}_n;\tau_1,\ldots,\tau_n)=\sum_{k}s^{j}_{mk}(\tfrac{1}{2}n-z,l)\,\Phi_{k,\,m-k}$$
$$=\sum_{k,\,\zeta}s^{j}_{mk}(\tfrac{1}{2}n-z,l)\,{}^zC_{k\zeta;\,\tau_1,\ldots,\tau_n}\,\psi_{m-k,\,\zeta}(\mathfrak{r}_1,\ldots,\mathfrak{r}_n)\tag{9}$$

sind also antisymmetrische Hyperfunktionen, die sich nach irreduziblen Darstellungen der Drehgruppe transformieren, sie bilden die gesuchten ersten Näherungen für die Eigenfunktionen, wenn die Wirkung der magnetischen Momente der Elektronen klein ist und als Störung betrachtet werden kann.

Formel (9) gibt auch das Vektorzusammensetzungsmodell wieder, da aus ihr unmittelbar hervorgeht, daß j von $|l-\frac{1}{2}n+z|$ bis $l+\frac{1}{2}n-z$ laufen kann. Wir wollten diese Formel ableiten, da mit ihrer Hilfe Matrixelemente in der hier betrachteten Näherung im allgemeinen einfach zu berechnen sind.

§ 3. Einige streng gültige Formeln. Wir wollen zuerst die streng gültigen Summensätze auf einem etwas anderen Wege, als in F und II ableiten. Wir bezeichnen das innere Produkt, also bei Hyperfunktionen:

$$\sum_{s_1\ldots s_n}\int f(x_1,\ldots,z_n;s_1,\ldots,s_n)\,\tilde{g}(x_1,\ldots,z_n;s_1,\ldots,s_n)\,dx_1\ldots dz_n,$$

wie in II mit (f, g). Dann ist die Übergangswahrscheinlichkeit vom Zustand Ψ_m in den Zustand $\Psi'_{m'}$, wenn man die magnetischen Kräfte vernächlässigt, durch das Quadrat von

$$U^{\alpha}_{mm'} = (\Psi_m, X_\alpha \Psi'_{m'}) \tag{10}$$

gegeben, wenn wir für

$$\left.\begin{aligned} x_1 + \cdots + x_n &= X_1 \\ y_1 + \cdots + y_n &= X_2 \\ z_1 + \cdots + z_n &= X_3 \end{aligned}\right\} \tag{11}$$

setzen. Wir haben außerdem, wenn $\mathfrak{R}$ eine Drehung ist:

$$O_{\mathfrak{R}} X_\alpha \Psi'_{m'} = X''_\alpha O_{\mathfrak{R}} \Psi'_{m'}, \tag{12}$$

wo X_α aus X''_α durch die Drehung $\mathfrak{R}$ hervorgeht, also $X_\alpha = \sum_\beta \mathfrak{R}_{\alpha\beta} X''_\beta$ ist*. Bezeichnen wir $(\Psi_m, X''_\beta \Psi'_{m'})$ mit $U''^{\beta}_{mm'}$ so ist

$$\left.\begin{aligned} U^{\alpha}_{mm'} &= \sum_\beta \mathfrak{R}_{\alpha\beta} U''^{\beta}_{mm'}, \\ \sum_\alpha U^{\alpha}_{mm'} \tilde{U}^{\alpha}_{mm'} &= \sum_{\alpha\beta\beta'} \mathfrak{R}_{\alpha\beta} \mathfrak{R}_{\alpha\beta'} U''^{\beta}_{mm'} \tilde{U}''^{\beta'}_{mm'} \\ &= \sum_\beta U''^{\beta}_{mm'} \tilde{U}''^{\beta}_{mm'}. \end{aligned}\right\} \tag{13}$$

Da $O_{\mathfrak{R}}$ ein unitärer Operator ist, ist zudem

$$\begin{aligned} U^{\alpha}_{mm'} &= (O_{\mathfrak{R}} \Psi_m, O_{\mathfrak{R}} X_\alpha \Psi'_{m'}) \\ &= \sum_{\mu\mu'} D^{j}_{\mu m} D^{j'}_{\mu' m'} (\Psi_\mu, X''_\alpha \Psi'_{\mu'}) \end{aligned} \tag{14}$$

wegen (12). Erheben wir dies zum Quadrat und summieren über α von 1 bis 3 und über m', so gilt

$$\begin{aligned} \sum_{\alpha m'} |U^{\alpha}_{mm'}|^2 &= \sum_{\substack{\mu\mu' \\ \nu\nu'}} \sum_{\substack{\alpha \\ m'}} D^{j}_{\mu m} D^{j'}_{\mu' m'} \tilde{D}^{j}_{\nu m} \tilde{D}^{j'}_{\nu' m'} U''^{\alpha}_{\mu\mu'} \tilde{U}''^{\alpha}_{\nu\nu'} \\ &= \sum_{\mu\mu'} \sum_{\nu\alpha} D^{j}_{\mu m} \tilde{D}^{j}_{\nu m} U''^{\alpha}_{\mu\mu'} \tilde{U}''^{\alpha}_{\nu\mu'} \end{aligned} \tag{14a}$$

wegen der Orthogonalität von $(D^{j'}_{\mu' m'})$.

Integrieren wir dies über alle Drehungen $\mathfrak{R}$, so erhalten wir wegen der Orthogonalitätsbeziehungen der Darstellungselemente (siehe z. B. F, Formel 1 b) und (13):

$$c \sum_{\alpha m'} |U^{\alpha}_{mm'}|^2 = c' \sum_{\mu} \sum_{\alpha\mu'} |U^{\alpha}_{\mu\mu'}|^2, \tag{15}$$

* Wir bezeichnen hier die Matrix der Drehung $\mathfrak{R}$ durch $(\mathfrak{R}_{\alpha\beta})$, weiter gehöre Ψ_m zu einem Term mit der inneren Quantenzahl j, $\Psi'_{m'}$ zu einem mit der Quantenzahl j'.

d. h. die Summe der Übergangswahrscheinlichkeiten in allen Richtungen, die zum anderen (j') Term führen, ist unabhängig von m. Dies ist die ursprüngliche Gestalt der Summensätze. Hätten wir unsere Überlegungen nicht auf Hyperfunktionen, sondern auf einfache Funktionen angewandt und $P_{\mathfrak{R}}$ an Stelle von $O_{\mathfrak{R}}$ geschrieben, so hätten wir die Summenregeln von F erhalten.

Es scheint bemerkenswert, daß (15) die Auswahlregeln für die innere bzw. azimutale Quantenzahl mit enthält, wenn man diejenige für die magnetische Quantenzahl als bekannt voraussetzt. Nehmen wir nämlich $j' \leqq j - 2$ und setzen in (15) $m = j$, so verschwindet die linke Seite, da m höchstens den Wert j annehmen kann. Hieraus folgt, daß auch die rechte Seite von (15), d. h. sämtliche Übergangswahrscheinlichkeiten verschwinden.

Wenn man die magnetischen Kräfte der Lichtanregung nicht vernachlässigt, so haben die Übergangswahrscheinlichkeiten eine etwas andere Gestalt. An Stelle von $e\mathfrak{E}z_i$ steht dann $P_1 = \frac{eh\mathfrak{A}}{2\mu c}\frac{h}{2\pi i}\frac{\partial}{\partial z_i}$, wo $\mathfrak{A}$ das Vektorpotential ist und wir haben noch ein Glied, das von den magnetischen Momenten der Elektronen herrührt $P_2 = \frac{eh\mathfrak{H}}{4\pi\mu c}\tau_i$. Im allgemeinen ist P_1 viel viel größer als P_2 $\left(\text{größenordnungsmäßig } \sqrt{\frac{\mu c^2}{h\nu}}\right)$ und für P_1, d. h. für die Matrixelemente $(\Psi_m, P_1 \Psi'_{m'})$ leitet man die Summensätze genau wie für $(\Psi_m, X\Psi'_{m'})$ ab. Die Beobachtbarkeit einer Abweichung wegen P_2 ist umso weniger zu erwarten, als die Auswahlregeln für j — bei denen allein eine so feine Prüfung möglich wäre — auch für P_2 gelten.

§ 4. Die Proportionalität der Aufspaltung mit der magnetischen Quantenzahl m gilt auch bei beliebigen Kopplungsverhältnissen. Das Zusatzglied in der Hamiltonschen Funktion, das einem Magnetfeld in der Z-Achse entspricht, besteht nämlich aus zwei Teilen. Der erste Teil rührt von der Bahnbewegung der Elektronen her und lautet

$$H_1 = \frac{he\mathfrak{H}}{4\pi i\mu c}\sum_k\left(x_k\frac{\partial}{\partial y_k} - y_k\frac{\partial}{\partial x_k}\right), \tag{16}$$

wo e und μ Ladung und Masse des Elektrons sind. Der zweite Teil des Zusatzgliedes rührt von den magnetischen Momenten der Elektronen her und lautet

$$H_2 = \frac{he\mathfrak{H}}{4\pi\mu c}\sum_k \tau_k. \tag{16a}$$

Sie ist durchaus von der Größenordnung von H_1. Leicht zu berechnen sind die Matrixelemente

$$Z_1 = [\Psi_m, (H_1 + \tfrac{1}{2} H_2)\, \Psi_m]. \tag{17}$$

Verstehen wir nämlich unter $\mathfrak{R}$ eine infinitesimale Drehung um die Z-Achse um den Winkel $\Delta\alpha$, so ist

$$O_{\mathfrak{R}} \Psi_m (x_1', \ldots, z_n'; \tau_1, \ldots, \tau_n) = e^{i(\tau_1 + \cdots + \tau_n)\frac{1}{2}\Delta\alpha} . \Psi_m (x_1, \ldots, z_n; \tau_1, \ldots, \tau_n) \tag{18}$$

wegen (3), (4) und (4a) in II. Dabei gilt für

$$\left.\begin{aligned} x_k &= x_k' - \Delta\alpha\, y_k' \\ y_k &= y_k' + \Delta\alpha\, x_k' \\ z_k &= z_k', \end{aligned}\right\} \tag{18a}$$

so daß man (18) auch schreiben kann

$$\left.\begin{aligned} & O_{\mathfrak{R}} \Psi_m (x_1', \ldots, z_n'; \tau_1, \ldots, \tau_n) \\ &= \left[1 + i \sum_k \tau_k \cdot \tfrac{1}{2} \Delta\alpha\right] \left[1 + \Delta\alpha \sum_k \left(x_k' \frac{\partial}{\partial y_k'} - y_k' \frac{\partial}{\partial x_k'}\right)\right] . \Psi_m (x_1', \ldots, z_n'; \tau_1, \ldots, \tau_n) \\ &= \left\{1 + i \Delta\alpha \frac{4\pi\mu c}{h e \mathfrak{H}} \left(H_1 + \frac{1}{2} H_2\right)\right\} \Psi_m (x_1', \ldots, z_n'; \tau_1, \ldots, \tau_n). \end{aligned}\right\} \tag{19}$$

Es ist also

$$\left(H_1 + \frac{1}{2} H_2\right) \Psi_m = \frac{O_{\mathfrak{R}} - 1}{i \Delta\alpha} \frac{h e \mathfrak{H}}{4\pi\mu c} \Psi_m, \tag{19a}$$

da aber

$$O_{\mathfrak{R}} \Psi_m = e^{i m \Delta\alpha} \Psi_m = (1 + i m \Delta\alpha)\, \Psi_m$$

ist, haben wir

$$\left(H_1 + \frac{1}{2} H_2\right) \Psi_m = \frac{h e \mathfrak{H}}{4\pi\mu c} m \Psi_m, \tag{20}$$

so daß für

$$Z_1 = \left(\Psi_m, \frac{h e \mathfrak{H}}{4\pi\mu c} m \Psi_m\right) = \frac{h e \mathfrak{H}}{4\pi\mu c} m \tag{17a}$$

in der Tat die Aufspaltung des normalen Zeemaneffekts folgt.

Die Berechnung des zweiten Teiles des Zusatzgliedes

$$Z_2 = \left(\Psi_m, \frac{1}{2} H_2 \Psi_m\right) = \frac{h e \mathfrak{H}}{8\pi\mu c} \sum_k (\Psi_m, \tau_k \Psi_m) = \frac{n h e \mathfrak{H}}{8\pi\mu c} \cdot (\Psi_m, \tau_1 \Psi_m) \tag{21}$$

gestaltet sich merkwürdigerweise sehr kompliziert, wenn man streng rechnet, d. h. nicht die Näherungsformel (9) benutzt. Wir berechnen zuerst

$$O_{\mathfrak{R}} \tau_1 \Psi_m (\mathfrak{r}_1', \ldots, \mathfrak{r}_n'; \tau_1, \ldots, \tau_n) = \sum_{\sigma_1 \ldots \sigma_n} a_{\tau_1, \ldots, \tau_n; \sigma_1, \ldots, \sigma_n} \sigma_1 \Psi_m (\mathfrak{r}_1, \ldots, \mathfrak{r}_n; \sigma_1, \ldots, \sigma_n)$$

und dies ist wegen

$$\sum_{\varrho_1\ldots\varrho_n} \tilde{a}_{\varrho_1,\ldots,\varrho_n;\,\sigma_1,\ldots,\sigma_n}(\mathfrak{R})\, O_{\mathfrak{R}}\, \Psi_m(\mathfrak{r}'_1,\ldots,\mathfrak{r}'_n;\varrho_1,\ldots,\varrho_n)$$
$$= \sum_{\varrho_1\ldots\varrho_n} \tilde{a}_{\varrho_1,\ldots,\varrho_n;\,\sigma_1,\ldots,\sigma_n}(\mathfrak{R}) \sum_{m'} D_{m'm}(\mathfrak{R})\, \Psi_{m'}(\mathfrak{r}'_1,\ldots,\mathfrak{r}'_n;\varrho_1,\ldots,\varrho_n)$$
$$= \sum_{\varrho_1\ldots\varrho_n} \tilde{a}_{\varrho_1,\ldots,\varrho_n;\,\sigma_1,\ldots,\sigma_n}(\mathfrak{R}) \sum_{\nu_1\ldots\nu_n} a_{\varrho_1,\ldots,\varrho_n;\,\nu_1,\ldots,\nu_n}(\mathfrak{R})\, \Psi_m(\mathfrak{r}_1,\ldots,\mathfrak{r}_n;\nu_1,\ldots,\nu_n)$$
$$= \Psi_m(\mathfrak{r}_1,\ldots,\mathfrak{r}_n;\sigma_1,\ldots,\sigma_n)$$

gleich

$$O_{\mathfrak{R}}\,\tau_1\,\Psi_m(\mathfrak{r}'_1,\ldots,\mathfrak{r}'_n;\tau_1,\ldots,\tau_n)$$
$$= \sum_{\sigma_1\ldots\sigma_n} a_{\tau_1,\ldots,\tau_n;\sigma_1,\ldots,\sigma_n}\sigma_1 \sum_{\varrho_1\ldots\varrho_n}\sum_{m'} \tilde{a}_{\varrho_1,\ldots,\varrho_n;\sigma_1,\ldots,\sigma_n} D_{m'm}\cdot\Psi_{m'}(\mathfrak{r}'_1,\ldots,\mathfrak{r}'_n;\varrho_1,\ldots,\varrho_n),$$

worin man die Summation wegen (4) in II

$$a_{\tau_1,\ldots,\tau_n;\,\sigma_1,\ldots,\sigma_n} = a_{\tau_1\sigma_1}\cdot a_{\tau_2\sigma_2}\ldots a_{\tau_n\sigma_n}$$

über $\sigma_2, \sigma_3, \ldots, \sigma_n$ direkt ausführen kann und erhält

$$O_{\mathfrak{R}}\,\tau_1\,\Psi_m(\mathfrak{r}'_1,\ldots,\mathfrak{r}'_n;\tau_1,\ldots,\tau_n)$$
$$= \sum_{\sigma_1,\varrho_1}\sum_{m'} a_{\tau_1\sigma_1}\sigma_1\tilde{a}_{\varrho_1\sigma_1} D_{m'm}\Psi_{m'}(\mathfrak{r}'_1,\ldots,\mathfrak{r}'_n;\varrho_1,\tau_2,\ldots,\tau_n)$$
$$\sum_{m'}\sum_{\varrho} x_{\tau_1\varrho} D_{m'm}\Psi_{m'}(\mathfrak{r}'_1,\ldots,\mathfrak{r}'_n;\varrho_1,\tau_2,\ldots,\mathfrak{r}_n) \tag{22}$$

mit $x_{\tau\varrho} = \sum_{\sigma} a_{\tau\sigma}\sigma\tilde{a}_{\varrho\sigma}$

$$x_{-1\,-1} = -\cos\beta \qquad x_{-1\,1} = e^{-i\alpha}\sin\beta$$
$$x_{1\,-1} = e^{i\alpha}\sin\beta \qquad x_{1\,1} = \cos\beta.$$

Da der Summationsbuchstabe ϱ in (22) nur die beiden Werte τ_1 und $-\tau_1$ anzunehmen hat

$$O_{\mathfrak{R}}\,\tau_1\,\Psi_m(\mathfrak{r}_1,\ldots,\mathfrak{r}_n;\tau_1,\ldots,\tau_n)$$
$$= \sum_{m'} D_{m'm}(\alpha,\beta,\gamma)\,[\tau_1\cos\beta\,.\,\Psi_{m'}(\mathfrak{r}_1,\ldots,\mathfrak{r}_n;\tau_1,\tau_2,\ldots,\tau_n)$$
$$+ e^{i\tau_1\alpha}\sin\beta\,.\,\Psi_{m'}(\mathfrak{r}_1,\ldots,\mathfrak{r}_n;-\tau_1,\tau_2,\ldots,\tau_n)]. \tag{23}$$

Nun ist

$$(\Psi_m,\tau_1\Psi_m) = (O_{\mathfrak{R}}\Psi_m,\, O_{\mathfrak{R}}\tau_1\Psi_m)$$
$$= \sum_{m'm''}\sum_{\tau_1\ldots\tau_n} D_{m''m}\cdot\tilde{D}_{m'm}\cdot\Psi_{m''}(\mathfrak{r}_1,\ldots,\mathfrak{r}_n;\tau_1,\ldots,\tau_n) \tag{23a}$$
$$[\tau_1\cos\beta\,\tilde{\Psi}_{m'}(\mathfrak{r}_1,\ldots,\mathfrak{r}_n;\tau_1,\tau_2,\ldots,\tau_n) + e^{-i\tau_1\alpha}\sin\beta\,\tilde{\Psi}_{m'}(\mathfrak{r}_1,\ldots,\mathfrak{r}_n;-\tau_1,\tau_2,\ldots,\tau_n)].$$

Nun ist (F, S. 640) $D^1_{00}(\alpha,\beta,\gamma) = \cos\beta$ und $D^1_{\tau_0}(\alpha,\beta,\gamma) = -\tau\, e^{i\tau\alpha}\dfrac{\sin\beta}{\sqrt{2}}$.

Setzt man dies in (23a) ein, so kann man wegen IIIa ff. des Anhanges über alle Drehungen integrieren. Die Integration ergibt auf der rechten Seite

$$s^j_{m'm''}(j,1)\, s^j_{mm}(j,1),$$

so daß man den Faktor $s^j_{mm}(j, 1) = \frac{m}{\sqrt{j}\,\sqrt{2j+1}}$ (V des Anhanges) ausklammern kann und der Klammerausdruck unabhängig von m wird. Auf der linken Seite steht $(\Psi_m, \tau_1 \Psi_m)$ mit einer Konstanten multipliziert. Hiermit ist also bewiesen, daß auch Z_2 proportional zu m ist. Es folgt, daß die Termaufspaltung des Zeemaneffekts bei kleinen Feldstärken mit m proportional ist, d. h. die Zeemanterme bei beliebigen Kopplungsverhältnissen äquidistant sind.

§ 5. Landésche g-Formel. Für den Fall, daß die Multiplettaufspaltung klein ist, d. h. die Wirkung der Elektronenmagnete für die Eigenfunktionen in erster Näherung zu vernachlässigen ist, können wir mit Hilfe unserer Formel (9) auch Z_2 berechnen und müssen dann, wenn wir den Wert für Z_1 aus (17 a) entnehmen, für $Z_1 + Z_2$ die Landésche g-Formel erhalten.

Auch hier scheinen wir etwas weiter ausholen zu müssen. Denken wir in (5) die Abhängigkeit von den drei Eulerschen Winkeln α, β, γ eingeschrieben und differenzieren wir (5) nach γ. Wir erhalten

$$\sum_{\sigma_1 \dots \sigma_n} a_{\tau_1, \dots, \tau_n; \sigma_1, \dots, \sigma_n} \tfrac{1}{2}(\sigma_1 + \dots + \sigma_n) \,.\, {}^zC_{m\zeta; \sigma_1, \dots, \sigma_n}$$
$$= \sum_{m'} D_{m'm}\, m \,.\, {}^zC_{m'\zeta; \tau_1, \dots, \tau_n}; \tag{5 a}$$

multiplizieren wir dies — um das a von der linken Seite wegzuschaffen — mit $\tilde{a}_{\tau_1, \dots, \tau_n; \varrho_1, \dots, \varrho_n}$ und summieren wir über $\tau_1 \dots \tau_n$. Wenn wir dann noch mit ${}^z\tilde{C}_{m\zeta; \varrho_1, \dots, \varrho_n}$ multiplizieren, erhalten wir

$$\tfrac{1}{2}(\varrho_1 + \dots + \varrho_n)\; {}^zC_{m\zeta; \varrho_1, \dots, \varrho_n}\; {}^z\tilde{C}_{m\zeta; \varrho_1, \dots, \varrho_n}$$
$$= \sum_{m'} \sum_{\tau_1 \dots \tau_n} D_{m'm}\, m\, \tilde{a}_{\tau_1, \dots, \tau_n; \varrho_1, \dots, \varrho_n}\; {}^zC_{m'\zeta; \tau_1, \dots, \tau_n}\; {}^z\tilde{C}_{m\zeta; \varrho_1, \dots, \varrho_n}.$$

Summieren wir dies über $\varrho_1 \dots \varrho_n$ und benutzen rechts wiederum (5)

$$\tfrac{1}{2} \sum_{\varrho_1 \dots \varrho_n} (\varrho_1 + \dots + \varrho_n)\, |{}^zC_{m\zeta; \varrho_1, \dots, \varrho_n}|^2$$
$$\sum_{\tau_1 \dots \tau_n} \sum_{m', m''} D_{m'm}\, m\, \tilde{D}_{m''m}\; {}^z\tilde{C}_{m''\zeta; \tau_1, \dots, \tau_n}\; {}^zC_{m'\zeta; \tau_1, \dots, \tau_n},$$

was schließlich wegen (4) und über alle Drehungen integriert wegen der Orthogonalitätsrelationen der Darstellungselemente

$$\tfrac{1}{2} \sum_{\varrho_1 \dots \varrho_n} \sum_{\zeta} |{}^zC_{m\zeta; \varrho_1, \dots, \varrho_n}|^2 (\varrho_1 + \dots + \varrho_n) = m \tag{5 b}$$

ergibt.

Betrachten wir also $Z_2 = \frac{h e \mathfrak{H}}{8 \pi \mu c} \cdot (\Psi_m, (\tau_1 + \cdots + \tau_n) \Psi_m)$ und setzten für Ψ_m (9) ein, so haben wir*

$$\sum_{\tau_1 \ldots \tau_n} \int \sum_{k\zeta} s^j_{mk}(r, l)\, {}^z C_{k\zeta;\tau_1,\ldots,\tau_n} \Psi_{m-k,\zeta}$$
$$\cdot (\tau_1 + \cdots + \tau_n) \sum_{k',\zeta'} \tilde{s}^j_{mk'}(r, l)\, {}^z\tilde{C}_{k'\zeta';\tau_1,\ldots,\tau_n} \tilde{\Psi}_{m-k',\zeta'},$$

worin wir die Integration wegen der Orthogonalität der Ψ ausführen können (sie ergibt $\delta_{kk'} \cdot \delta_{\zeta\zeta'}$), und wir haben

$$Z_2 \frac{8 \pi \mu c}{h c \mathfrak{H}} = \sum_{\tau_1 \ldots \tau_n} \sum_{k\zeta} s^j_{mk}(r, l)\, \tilde{s}^j_{mk}(r, l)\, |{}^z C_{k\zeta;\tau_1,\ldots,\tau_n}|^2 (\tau_1 + \cdots + \tau_n)$$
$$= 2 \sum_k s^j_{mk}(r, l)\, \tilde{s}^j_{mk}(r, l)\, k. \tag{24}$$

Wir wissen aus dem vorangehenden § 4, daß Z_2 proportional zu m ist. Wir können also

$$\sum_k s^j_{mk}(r, l)\, \tilde{s}^j_{mk}(r, l)\, k = f(j, l, r)\, m \tag{25}$$

für $|m| \leqq j$ schreiben. Für $|m| > j$ ist dagegen natürlich

$$\sum_k s^j_{mk}(r, l)\, \tilde{s}^j_{mk}(r, l)\, k = 0. \tag{25a}$$

Bilden wir jetzt

$$\sum_{j=m}^{\infty} m f(j, l, r) = \sum_{jk} s^j_{mk}(r, l)\, \tilde{s}^j_{mk}(r, l)\, k.$$

Hierin können wir mit Hilfe von (II) des Anhanges die Summation über j ausführen. Nach ihr ist

$$\sum_{j=m}^{\infty} s^j_{mk}(r, l)\, \tilde{s}^j_{mk}(r, l) = 1, \tag{II}$$

wenn $|k| \leqq r$ und $|m - k| \leqq l$, sonst natürlich 0. Wir erhalten folglich,

$$\sum_{j=m}^{\infty} \sum_k s^j_{mk}(r, l)\, \tilde{s}^j_{mk}(r, l)\, k = \sum_{k=m-l}^{r} k = \frac{m - l + r}{2} (r - m + l + 1),$$

so daß

$$\sum_{j=m}^{\infty} f(j, l, r) = \frac{1}{m} \frac{(m - l + r)(l + r - m + 1)}{2} \tag{25b}$$

* Wir setzen der Bequemlichkeit halber r für $\frac{1}{2} n - z$.

wird. Setzen wir für m jetzt $m+1$ und ziehen die so entstandene Gleichung aus (25 b) ab, so erhalten wir

$$\sum_{j=m}^{\infty} f(j,l,r) - \sum_{j=m+1}^{\infty} f(j,l,r) = f(m,l,r)$$

$$= \frac{(m-l+r)(l+r-m+1)}{2m} - \frac{(m+1-l+r)(l+r-m)}{2(m+1)}$$

$$= \frac{m(m+1) - l(l+1) + r(r+1)}{2m(m+1)},$$

wie man leicht verifiziert. Führen wir diesen Wert für f in (24) ein, so erhalten wir für Z_2

$$Z_2 = \frac{he\mathfrak{H}}{8\pi\mu c}\, 2 \cdot \frac{j(j+1) - l(l+1) + r(r+1)}{2j(j+1)}\, m, \tag{24 a}$$

so daß wir für $Z_1 + Z_2$ mit (17 a)

$$Z_1 + Z_2 = m\frac{hc\mathfrak{H}}{4\pi\mu c}\left[1 + \frac{j(j+1) - l(l+1) + r(r+1)}{2j(j+1)}\right], \tag{26}$$

d. h. die Landésche g-Formel erhalten. (Siehe z. B. E. Back und A. Landé: Zeemaneffekt und Multiplettstruktur der Spektrallinien. Berlin 1925. S. 41, Formel (20′). Dort steht $J - \frac{1}{2}$ für unser j; $K - \frac{1}{2}$ für unser l und $R - \frac{1}{2}$ für unser $r = \frac{1}{2}n - z$).

§ 6. Wenn zur Ableitung von (26) viel Rechnung notwendig erschien, so ist es umso angenehmer zu sehen, daß man die Summensätze im wesentlichen schon sieht, wenn man die entsprechenden Ausdrücke aufgeschrieben hat.

Wir betrachten zwei Terme der spinfreien Differentialgleichung. Die Eigenfunktionen des einen seien $\psi_{\mu\zeta}(\mathfrak{r}_1, \ldots, \mathfrak{r}_n)$, er habe die azimutale Quantenzahl l und gehöre zur Partitio z, die Eigenfunktionen des anderen Terms seien $\psi'_{\mu'\zeta'}(\mathfrak{r}_1, \ldots, \mathfrak{r}_n)$, die azimutale Quantenzahl sei l', die Partitio ebenfalls z, da sonst die Übergangswahrscheinlichkeiten in der hier betrachteten Näherung überhaupt verschwinden. Wir bezeichnen das Integral

$$\int \psi_{\mu\zeta} X_\alpha \tilde{\psi}'_{\mu'\zeta'} = U^\alpha_{\mu\mu'} \delta_{\zeta\zeta'}, \tag{27}$$

wo X_α in (11) definiert ist. Der spinfreie Summensatz besagt dann, daß

$$\sum_{\alpha\mu'} U^\alpha_{\mu\mu'} \tilde{U}^\alpha_{\mu\mu'} = c, \tag{15 a}$$

d. h. unabhängig von μ ist.

Die erste Näherung (9) für die Hyperfunktion mit der inneren Quantenzahl j und der magnetischen m, die aus dem ersten Term entsteht, ist

$$\sum_{k\zeta} s^j_{mk}(r,l)\, {}^zC_{k\zeta;\tau_1,\ldots,\tau_n}\, \psi_{m-k,\zeta}(\mathfrak{r}_1,\ldots,\mathfrak{r}_n) = \Psi^j_m. \tag{9 b}$$

Für die Hyperfunktion, die aus dem zweiten Term entsteht und die Quantenzahlen j', m' hat, haben wir

$$\sum_{k'\zeta'} s^{j'}_{m'k'}(r,l')\, {}^zC_{k'\zeta';\tau_1,\ldots,\tau_n}\, \psi'_{m'-k',\zeta'}(\mathfrak{r}_1,\ldots,\mathfrak{r}_n) = \Psi'^{j'}_{m'}. \tag{9 b}$$

Die Übergangswahrscheinlichkeit ist also das Quadrat von

$$V^\alpha_{jm,j'm'} = \sum_{\tau_1\ldots\tau_n}\int \Psi^j_m X_\alpha \tilde{\Psi}'^{j'}_{m'}$$

$$= \sum_{\tau_1\ldots\tau_n}\sum_{k\zeta k'\zeta'} s^j_{mk}(r,l)\, \tilde{s}^{j'}_{m'k'}(r,l') \,.\, {}^zC_{k\zeta;\tau_1,\ldots,\tau_n}\, {}^z\tilde{C}_{k'\zeta';\tau_1,\ldots,\tau_n} \cdot \int \psi_{m-k\zeta} X_\alpha \psi_{m'-k'\zeta'},$$

oder wegen (27) und (4)

$$V^\alpha_{jm,j'm'} = \sum_k U^\alpha_{m-k,m'-k}\, s^j_{mk}(r,l)\, s^j_{m'k}(r,l'), \tag{27 a}$$

Die Summe der Übergangswahrscheinlichkeiten vom Zustand j, m in alle Zustände j', m' in allen Richtungen X, Y, Z ist

$$S_{jm} = \sum_{\alpha j' m'} V^\alpha_{jm,j'm'}\, \tilde{V}^\alpha_{jm,j'm'}$$

$$= \sum_{\alpha j' m'}\sum_{kk'} U^\alpha_{m-k,m'-k}\, \tilde{U}^\alpha_{m-k',m'-k'}\, s^j_{mk}(r,l) \,.\, \tilde{s}^j_{mk'}(r,l) \,.\, \tilde{s}^{j'}_{m'k}(r,l') \,.\, s^{j'}_{m'k'}(r,l'), \tag{28}$$

worin aber die Summation über j' wegen (II) des Anhanges leicht auszuführen ist. Wenn man für $k = m + \eta$ und $m' = m + \eta + \eta'$ setzt, erhält man

$$S_{jm} = \sum_{\alpha\eta\eta'} U^\alpha_{\eta\eta'}\, \tilde{U}^\alpha_{\eta\eta'}\, s^j_{m,m+\eta}(r,l)\, \tilde{s}^j_{m,m+\eta}(r,l), \tag{28 a}$$

was wegen (15 a)

$$S_{jm} = c\sum_\eta s^j_{m,m+\eta}(r,l)\, \tilde{s}^j_{m,m+\eta}(r,l). \tag{28 b}$$

Dies ist aber nach (IV) des Anhanges

$$S_{jm} = c\sum_\eta s^j_{m,-\eta}(l,r)\, \tilde{s}^j_{m,-\eta}(l,r) = c \tag{29}$$

wegen (II) des Anhanges. Die Summe der Übergangswahrscheinlichkeiten ist also unabhängig von j und m.

§ 7. Hiermit schließen wir unsere Rechnungen ab. Man sieht, daß man ohne besondere Mühe fast beliebig viele Zusammenhänge zwischen Übergangswahrscheinlichkeiten und dergleichen herstellen könnte. Wir glauben, daß es nützlich ist, die streng gültigen Gesetze von denen ab-

zusondern, die nur bei gewissen Kopplungsverhältnissen gelten, da namentlich die Prüfung der ersteren zur Entscheidung für oder wider die Notwendigkeit und Art von Zusatzgliedern geeignet erscheint. Sonst kam es uns hauptsächlich darauf an, die Theorie der qualitativen Struktur der Atomspektren zu einem gewissen Abschluß zu bringen.

Anhang.

Es sei $(D^l_{\varkappa\lambda}(\mathfrak{R}))$ die $2l+1$-dimensionale Darstellung der dreidimensionalen Drehgruppe, $(D^{l'}_{\varkappa\lambda}(\mathfrak{R}))$ die $2l'+1$-dimensionale. Die Matrizen $(M_{\varkappa\varkappa',\lambda\lambda'}(\mathfrak{R}))$ mit

$$M_{\varkappa\varkappa',\lambda\lambda'}(\mathfrak{R}) = D^l_{\varkappa\lambda}(\mathfrak{R}) \,.\, D^{l'}_{\varkappa'\lambda'}(\mathfrak{R})$$

bilden dann offenbar auch eine Darstellung, und zwar eine $(2l+1)(2l'+1)$-dimensionale, die aber nicht irreduzibel ist. Wir können seine irreduziblen Bestandteile aber leicht bestimmen. Die Spur von $(M_{\varkappa\varkappa',\lambda\lambda'}(\mathfrak{R}))$ ist das Produkt der Spuren von $(D^l_{\varkappa\lambda})$ und $(D^{l'}_{\varkappa'\lambda'})$, also, wenn φ der Drehwinkel ist,

$$\sum_{\varkappa=-l}^{l} e^{i\varkappa\varphi} \sum_{\varkappa'=-l'}^{l'} e^{i\varkappa'\varphi} = \sum_{j=|l-l'|}^{l+l'} \sum_{\lambda=-j}^{j} e^{i\lambda\varphi}.$$

D. h. die irreduziblen Bestandteile sind je eine Darstellung $(D^j_{\varkappa\lambda})$, wo j von $|l-l'|$ bis $l+l'$ in ganzzahligen Schritten läuft. Es muß daher eine Matrix $s_{\lambda\lambda',j\nu}$ existieren, die $(M_{\varkappa\varkappa',\lambda\lambda'})$ in die Form

$$\begin{pmatrix} D^{|l-l'|} & \cdots & 0 \\ 0 & D^{|l-l'|+1} \cdots & 0 \\ \vdots & \vdots & \vdots \\ 0 & 0 \cdots & D^{l+l'} \end{pmatrix}$$

bringt. Führen wir zur Charakterisierung der Drehung $\mathfrak{R}$ die drei Eulerschen Winkel α, β, γ ein, so können wir setzen (F, (10))

$$D^l_{\varkappa\lambda}(\alpha,\beta,\gamma) = e^{i\varkappa\alpha} d^l_{\varkappa\lambda}(\beta)\, e^{i\lambda\gamma}.$$

Für die $s_{\lambda\lambda',j\nu}$ gilt dann

$$\sum_{\lambda\lambda'} e^{i(\varkappa+\varkappa')\alpha} d^l_{\varkappa\lambda}(\beta)\, d^{l'}_{\varkappa'\lambda'}(\beta)\, e^{i(\lambda+\lambda')\gamma} s_{\lambda\lambda',j\nu} = \sum_{\nu'} s_{\varkappa\varkappa',j\nu'}\, e^{i\nu'\alpha} d^j_{\nu'\nu}(\beta)\, e^{i\nu\gamma}.$$

Wenn wir darin die Glieder mit gleicher Abhängigkeit von α und γ einander gleichsetzen, erhalten wir unter anderem

$$s_{\lambda\lambda',j\nu} = 0 \quad \text{für} \quad \lambda+\lambda' \neq \nu.$$

Wir bezeichnen also $s_{\lambda\lambda',j\nu}$ mit

$$s_{\lambda\lambda',j\nu} = s^j_{\nu\lambda}\,\delta_{\lambda+\lambda',\nu}, \tag{A}$$

was

$$\sum_{\lambda} D^l_{\varkappa\lambda} D^{l'}_{\nu-\varkappa,\mu-\lambda}\, s^j_{\mu\lambda}(l,l') = s^j_{\nu\varkappa}(l,l')\, D^j_{\nu\mu} \tag{I}$$

ergibt. Aus der Orthogonalität von $s_{\lambda\lambda', j\nu}$ folgt

$$\sum_{j\nu} s_{\lambda\lambda', j\nu}\, \tilde{s}_{\varkappa\varkappa', j\nu} = \delta_{\varkappa\lambda}\, \delta_{\varkappa'\lambda'}; \quad \sum_{\lambda\lambda'} s_{\lambda\lambda', j\nu}\, \tilde{s}_{\lambda\lambda', j'\nu'} = \delta_{jj'}\, \delta_{\nu\nu'}.$$

Dies ergibt, zusammen mit (A), wenn man $\lambda + \lambda' = \varkappa + \varkappa' = \mu$ bzw. $\nu = \nu'$ setzt:

$$\sum_j s^j_{\mu\lambda}(l, l')\, \tilde{s}^j_{\mu\varkappa}(l, l') = \delta_{\varkappa\lambda}; \quad \sum_\lambda s^j_{\nu\lambda}(l, l')\, \tilde{s}^{j'}_{\nu\lambda}(l, l') = \delta_{jj'}. \tag{II}$$

Wenn man dies mit (I) kombiniert, erhält man schließlich

$$D^l_{\varkappa\lambda}\, D^{l'}_{\nu-\varkappa,\, \mu-\lambda} = \sum_j s^j_{\nu\varkappa}(l, l')\, \tilde{s}^j_{\mu\lambda}(l, l')\, D^j_{\nu\lambda} \tag{III}$$

oder

$$D^l_{\varkappa\lambda}\, D^{l'}_{\varkappa'\lambda'} = \sum_j s^j_{\varkappa+\varkappa',\, \varkappa}(l, l')\, \tilde{s}^j_{\lambda+\lambda',\, \lambda}(l, l')\, D^j_{\varkappa+\varkappa',\, \lambda+\lambda'}. \tag{IIIa}$$

In (I), (II), (III) und (IIIa) sind die $s^j_{\varkappa\lambda}(l, l')$ reine Zahlen, die nur von ihren Indizes und Argumenten abhängen.

Wir wollen noch bemerken, daß es — da die komplexe Phase der linken Seite von (III) gleich der komplexen Phase eines jeden Gliedes der rechten Seite (nämlich $\nu\alpha + \mu\gamma$) ist — möglich ist, die $s^j_{\varkappa\lambda}(l, l')$ alle reell anzunehmen. Dann erhalten wir, wenn wir etwa in (IIIa) $\lambda = \lambda' = 0$ annehmen, mit $\tilde{D}^j_{\varkappa+\varkappa',\, 0}$ multiplizieren und über alle Drehungen integrieren,

$$\int D^l_{\varkappa,\, 0}\, D^{l'}_{\varkappa'\, 0}\, \tilde{D}^{j'}_{\varkappa+\varkappa',\, 0} = c\, s^j_{\varkappa+\varkappa',\, \varkappa}(l, l')\, s^j_{0,0}(l, l').$$

Weiter können wir $s^{j'}_{00}(l, l') = s^{j'}_{00}(l', l)$ setzen und erhalten unter anderem

$$s^j_{\varkappa+\varkappa',\, \varkappa}(l, l') = s^j_{\varkappa+\varkappa',\, \varkappa'}(l', l). \tag{IV}$$

Für den Fall $l' = 1$ wurden diese Formeln schon angegeben*. Dort findet man auch für die Koeffizienten s

$$\left.\begin{aligned}
s^{l-1}_{\varkappa-1,\, \varkappa}(l, 1) &= -\frac{\sqrt{l+\varkappa}\,\sqrt{l+\varkappa-1}}{\sqrt{2l}\,\sqrt{2l+1}}; & s^{l-1}_{\varkappa,\, \varkappa}(l, 1) &= \frac{\sqrt{l^2-\varkappa^2}}{\sqrt{l}\,\sqrt{2l+1}};\\
s^{l-1}_{\varkappa+1,\, \varkappa}(l, 1) &= -\frac{\sqrt{l-\varkappa-1}\,\sqrt{l-\varkappa}}{\sqrt{2l}\,\sqrt{2l+1}}; & s^{l}_{\varkappa-1,\, \varkappa}(l, 1) &= \frac{\sqrt{l-\varkappa+1}\,\sqrt{l+\varkappa}}{\sqrt{2l}\,\sqrt{l+1}};\\
s^{l}_{\varkappa,\, \varkappa}(l, 1) &= \frac{\varkappa}{\sqrt{l}\,\sqrt{2l+1}}; & s^{l}_{\varkappa+1,\, \varkappa}(l, 1) &= -\frac{\sqrt{l+\varkappa+1}\,\sqrt{l-\varkappa}}{\sqrt{2l}\,\sqrt{l+1}};\\
s^{l+1}_{\varkappa-1,\, \varkappa}(l, 1) &= \frac{\sqrt{l-\varkappa+1}\,\sqrt{l-\varkappa+2}}{\sqrt{2l+1}\,\sqrt{2l+2}}; & s^{l+1}_{\varkappa,\, \varkappa}(l, 1) &= \frac{\sqrt{(l+1)^2-\varkappa^2}}{\sqrt{l+1}\,\sqrt{2l+1}};\\
& & s^{l+1}_{\varkappa+1,\, \varkappa}(l, 1) &= \frac{\sqrt{l+\varkappa+1}\,\sqrt{l+\varkappa+2}}{\sqrt{2l+2}\,\sqrt{2l+1}};
\end{aligned}\right\} \tag{V}$$

* E. Wigner, ZS. f. Phys. **45**, 601, 1927.

oft sind auch die Koeffizienten für $l' = \frac{1}{2}$ wichtig*:

$$\left.\begin{aligned} s^{l-1/2}_{\varkappa-1/2,\,\varkappa}(l,\tfrac{1}{2}) &= \frac{\sqrt{l+\varkappa+\frac{1}{2}}}{\sqrt{2l+1}}; & s^{l-1/2}_{\varkappa+1/2,\,\varkappa}(l,\tfrac{1}{2}) &= \frac{\sqrt{l-\varkappa+\frac{1}{2}}}{\sqrt{2l+1}}; \\ s^{l+1/2}_{\varkappa-1/2,\,\varkappa}(l,\tfrac{1}{2}) &= -\frac{\sqrt{l-\varkappa+\frac{1}{2}}}{\sqrt{2l+1}}; & s^{l+1/2}_{\varkappa+1/2,\,\varkappa}(l,\tfrac{1}{2}) &= \frac{\sqrt{l+\varkappa+\frac{1}{2}}}{\sqrt{2l+1}}. \end{aligned}\right\} \text{(VI)}$$

Sowohl bei (V) wie bei (VI) verschwinden alle übrigen s.

Göttingen und Berlin, Juni 1928.

* Es ist $D^{1/2}_{1/2\,\tau,\,1/2\,\sigma} = a_{\tau,\,\sigma}$. Siehe I.

Fehlerberichtigung.

Im ersten Teile dieser Arbeit* kommen bedauerlicherweise mehrere Druck- und Schreibfehler vor.

S. 206 Zeile 7 von oben steht $\dot{\pm}$; richtig: ± 1,

S. 206 Anmerkung * steht $\frac{hm}{4\pi e}$; richtig: $\frac{he}{4\pi mc}$.

S. 209 Zeile 8 von oben steht $Q_{\mathfrak{R},\,\mathfrak{S}}$; richtig: $Q_{\mathfrak{R}\,\mathfrak{S}}$.

S. 210 Anmerkung * steht $l = 1, 2, \ldots$; richtig: $q = 1, 2, \ldots$

S. 211 Zeile 5 von oben steht $\left\{a^{(\mathfrak{R})}_{s_1,\,s_1}\right\}$; richtig: $\left\{a^{(\mathfrak{R})}_{s_1,\,t_1}\right\}$.

S. 214 „ 14 „ „ „ $f_{m-2,\,m-1}$; richtig: $f_{m-1,\,m-1}$.

S. 214 „ 6 „ unten „ $\sum$; richtig: $\sum\limits_{u,\,v}$.

S. 215 „ 3 „ oben „ $\cos^{2(m-2)} {}^1/_2\,\beta$; richtig: $\cos^{2(m-1)} {}^1/_2\,\beta$.

S. 216 „ 8 „ oben „ $q \cos^{q-1} {}^1/_2\,\beta$; richtig: $\sum\limits_{m=-\frac{1}{2}(q-1)}^{\frac{1}{2}q-1} e^{im\beta}$.

S. 216 „ 16 „ „ „ $e^{t(\cdots}$; richtig: $e^{i(\cdots}$.

S. 217 „ 8 „ unten „ $\beta = \frac{\pi}{2}$; richtig: $\beta = \frac{\pi}{4}$.

S. 217 „ 5 „ „ „ n_n; richtig: u_n.

S. 217 „ 3 „ „ „ $\sum t_\nu$; richtig: $\sum r_\nu$.

* ZS. f. Phys **47**, 203, 1928.

Über die Struktur der zweiatomigen Molekelspektren nach der Quantenmechanik

E. P. Wigner und E. E. Witmer*

Zeitschrift für Physik *51*, 859–886 (1928)

Mit 3 Abbildungen. (Eingegangen am 23. Juli 1928.)

Die Struktur der Molekelspektren wird mit Hilfe der gruppentheoretischen Methode behandelt. Im ersten Teile (§ 1 bis 9) werden — nach Aufstellung der ersten Näherung für die Eigenfunktionen — die Auswahl- und Intensitätsregeln und die Aufbauregel einer Rotationsbande behandelt. Im zweiten Teile (§ 10 bis 15) wird die Aufbauregel der Elektronenterme des Moleküls gegeben, indem es durch Zusammenführen von zwei Atomen entstanden gedacht wird. Es werden so diejenigen Elektronenterme des Moleküls gleichzeitig ins Auge gefaßt, die dieselbe Bandenkonvergenzstelle haben.

§ 1. Wir wollen im folgenden die Struktur der zweiatomigen Molekelspektren, wie sie sich aus der Quantenmechanik ergibt, mit Hilfe einer Methode darstellen, die schon früher zur Analyse der Atomspektren verwendet worden ist**. Diese scheint uns erstens eine einheitliche Behandlung der hier auftauchenden Fragen zu gestatten und ergibt zwangsläufig alle Regelmäßigkeiten, Auswahlverbote oder nicht ganz normale Typen von Spektren.

Von bisherigen Untersuchungen über die qualitative Struktur*** dieser Spektren sind in erster Reihe die grundlegenden Untersuchungen von F. Hund**** zu nennen, der auf halb quantenmechanischem, halb korrespondenzmäßigem Wege diese Frage zu lösen versucht hat. Besonders wichtig erscheint uns seine Idee von der adiabatischen Zusammenführung der Atome zu einem Molekül, auf die sich unser Vektorzusammensetzungsmodell gründet. Durch die Franck-Condonschen Ideen über die Be-

---

* Fellow of the National Research Council (U. S. A.) and the International Education Board.

** E. Wigner, ZS. f. Phys. **40**, 492, 883, 1927; **43**, 624, 1927, im folgenden als *F* zitiert; J. v. Neumann und E. Wigner, ebenda **47**, 203, 1928; **49**, 73, 1928; **51**, 844, 1928, als I, II, III zitiert. — Es wird hier keine Vollständigkeit angestrebt. Der Zweck ist vielmehr, nach dem Aufschreiben der ersten Näherung für die Eigenfunktionen — woraus eigentlich alles herzuleiten ist — nur einige uns besonders wichtig erscheinende Regelmäßigkeiten anzuführen.

*** Die Arbeiten von L. Mensing, ZS. f. Phys. **36**, 814, 1926; E. Fues, Ann. d. Phys. **80**, 367, 1926; **81**, 218, 1926; H. Rademacher und F. Reiche, ZS. f. Phys. **39**, 444, 1926; **41**, 453, 1927; R. de L. Kronig und I. I. Rabi, Phys. Rev. **29**, 262, 1927, beschäftigen sich in erster Reihe mit den Termen eines mehr oder weniger idealisierten Modells.

**** F. Hund, ZS. f. Phys. **36**, 657, 1926; **40**, 742, 1927; **42**, 93, 1927; **43**, 788, 1927.

deutung der Bandenkonvergenzstelle* erhält der Prozeß der Zusammenführung der Atome bzw. Trennung der Molekel in Atome einen unmittelbaren physikalischen Sinn. Weiter möchten wir die Arbeit von R. de L. Kronig** erwähnen, die einige unklare Punkte in der Hundschen Theorie aufgeklärt hat. Von W. Heisenberg*** stammt die Erklärung der ausfallenden Linien und des Intensitätswechsels in Banden mit zwei gleichen Atomen. Eine Begründung erfuhren die stillschweigenden Voraussetzungen dieser Arbeiten durch M. Born und J. R. Oppenheimer****, die unmittelbar gezeigt haben, warum man die Kernbewegung in erster Näherung vernachlässigen darf; diese Arbeit gab den unmittelbaren Anlaß zu vorliegendem Aufsatz. Vor einigen Tagen schließlich erschien eine Ankündigung einer Arbeit von J. H. van Vleck und R. S. Mulliken† über die Größe der sogenannten σ-Typ Dubletts, so daß

* J. Franck, Trans. Faraday Society (1925); E. Condon, Phys. Rev. **28**, 1182, 1926.

** R. de L. Kronig, ZS. f. Phys. **46**, 814, 1927.

*** W. Heisenberg, ebenda **41**, 239, 1927.

**** M. Born und J. R. Oppenheimer, Ann. d. Phys. **84**, 457, 1927.

† J. H. van Vleck und R. S. Mulliken, Bull. Amer. Phys. Soc. June 15, 1928, S. 13.

Wir haben diese Rechnung auch ausgeführt, indem wir zuerst die richtige nullte Näherung (7) für die Eigenfunktion in die Schrödingergleichung $H\psi = E\psi$ eingeführt haben und damit die erste Näherung $E = (\psi, H\psi)$ für den Eigenwert berechneten. Diese ergab

$$E_l = E_0 + [l(l+1) - \lambda^2] \frac{h^2}{8\pi^2} \frac{M_1 + M_2}{M_1 M_2} \int \frac{\varphi_{+,\lambda} \tilde{\varphi}_{+,\lambda}}{Z^2} dg,$$

(l Rotationsquantenzahl, M_1 und M_2 Massen der beiden Atome), wo E_0 nicht mehr von l abhängt und die Energie der Elektronenbewegung und Kernschwingung darstellt. Das reziproke Trägheitsmoment ist also

$$\frac{M_1 + M_2}{M_1 M_2} \int \frac{\varphi_{\pm,\lambda} \tilde{\varphi}_{\pm,\lambda}}{Z^2} dg.$$

Die σ-Typ-Aufspaltung tritt bei Π-Termen in der zweiten Näherung auf. Die Aufspaltung für E_l beträgt

$$\sum_n \frac{16\,|A_n|^2\, l(l+1)}{E_l - E_{n,l}}, \tag{α}$$

wo der Summationsbuchstabe n alle Σ-Terme durchläuft, $E_{n,l}$ die Energie des betreffenden Σ-Zustandes mit der Rotationsquantenzahl l ist. Die Elektronen-eigenfunktion, die zum n-ten Σ-Term gehört, sei φ_n. Dann ist

$$A_n = \frac{h^2}{8\pi^2} \frac{M_1 + M_2}{M_1 M_2} \sum_r \int \frac{1}{Z^2} \tilde{\varphi}_n \left(x_r \frac{\partial \varphi}{\partial z_r} - z_r \frac{\partial \varphi}{\partial x_r} \right) dg,$$

wo x_r, y_r, z_r die in § 4 eingeführten, die relative Lage des r-ten Elektrons kennzeichnenden Koordinaten sind, deren Gesamtheit mit Z zusammen g ergibt.

wir bezüglich dieser Frage auf diese Arbeit hinweisen zu dürfen glauben.

Wir möchten schließlich nicht unerwähnt lassen, wie wertvolle Dienste uns die Arbeiten von R. S. Mulliken* geleistet haben und Herrn Prof. F. W. Loomis für viele interessante Diskussionen und Bemerkungen auch an dieser Stelle herzlich zu danken.

§ 2. Einteilung dieser Arbeit. Im ersten Teile (§ 3 bis 8) werden wir die verschiedenen Typen von Termen und die zwischen ihnen gültigen Auswahl- und Intensitätsregeln diskutieren. Im zweiten Teile (§ 10 bis 12) werden wir ein Aufbauprinzip abzuleiten versuchen.

Bei der Ableitung der verschiedenen Typen werden wir uns zuerst die Kerne auf einer Geraden festgehalten denken (§ 3 für ungleiche, § 4 für gleiche Kerne). Dann müssen wir aber zwei verschiedene Fälle unterscheiden, je nachdem b) die Rotationsenergie größer (§ 7) oder a) kleiner als die Energie der Multiplettaufspaltung ist (§ 8 und 9). Dies entspricht dem Falle b) bzw. a) von Hund und Mulliken. Moleküle aus leichten Elementen (H_2, He_2 usw.) sind Beispiele für den ersten, Moleküle aus schweren Elementen (J_2, Hg_2 usw.) sind Beispiele für den zweiten Fall. Glücklicherweise ist das Übergangsgebiet sehr klein, da die Rotationsaufspaltung mit der Ordnungszahl fällt, die Multiplettabspaltung dagegen ansteigt. Ist die Rotationsaufspaltung groß, so berücksichtigen wir zuerst diese (§ 6) und führen danach die Elektronenmagnete, die die Multiplettaufspaltung bedingen, ein (§ 7). Ist dagegen letztere groß, so führen wir noch bei festgehaltenen Kernen die Elektronenmagnete ein und berücksichtigen erst dann die Rotation (§ 8).

Die Aufbauregel versuchen wir, wie gesagt, durch adiabatisches Zusammenführen der Kerne nach Hund zu gewinnen. Wir möchten aber bemerken, daß wir zu mit den seinigen nicht ganz übereinstimmenden Resultaten kommen. Bezüglich der Benennung der Terme weichen wir etwas von Mullikens Nomenklatur ab, da wir Namen wie azimutale Quantenzahl, S, P, D usw. Terme, auch bei Molekülen, nicht auf ganz andere Begriffe anwenden wollen, als man sie gewöhnlich bei Atomen

Die σ-Typ-Aufspaltung der Δ-Terme kommt erst in der nächsten Näherung zum Vorschein, sie sollte also im Verhältnis zur Aufspaltung bei Π-Termen klein sein.

Was die Lage der beiden entstehenden Terme anbetrifft, wurde in Tabelle 1 angenommen, daß die Summe (α) negativ ist. Dies ist nur dann sicher, wenn unter dem betrachtetem Π-Term kein Σ-Term — wenigstens desselben Multiplettsystems — mehr liegt.

* R. S. Mulliken, Phys. Rev. **28**, 481, 1202, 1926; **29**, 391, 637, 1927; **30**, 138, 785, 1927.

anwendet. Im letzten Paragraphen geben wir eine vergleichende Tabelle der beiden Nomenklaturen *.

Im folgenden denken wir uns in der Schrödingergleichung die Schwerpunktskoordinaten schon absepariert; $x_1, y_1, z_1, \ldots, x_n, y_n, z_n$ bedeuten dann die kartesischen Koordinaten der Elektronen, wofür wir auch $\mathfrak{r}_1, \mathfrak{r}_2, \ldots, \mathfrak{r}_n$ schreiben, X, Y, Z bedeuten die Koordinaten des einen Kerns, die des anderen Kerns sind dann $-\frac{M_1}{M_2}X$, $-\frac{M_1}{M_2}Y$, $-\frac{M_1}{M_2}Z$, wenn M_1 und M_2 die Massen der beiden Kerne sind und wir den Schwerpunkt in den Ursprungspunkt unseres Koordinatensystems legen.

Das Vertauschen der beiden Kerne bei Molekülen mit gleichen Kernen bedeutet das Ersetzen von X, Y, Z durch $-X, -Y, -Z$.

§ 3. Betrachten wir ein Molekül mit zwei ungleichen Kernen. Wir sehen zunächst von der Rotation des ganzen Molekül ab und denken uns die Kerne auf der Z-Achse fixiert, also $X = Y = 0$ angenommen. Das Problem hat dann die Z-Achse als Symmetrieachse, jede durch Z gehende Ebene als Symmetrieebene. Wenn wir also auch von den Entartungen, die durch die Gleichheit der Elektronen bedingt sind, für den Augenblick absehen, ist die Symmetriegruppe eine zweidimensionale Drehspiegelungsgruppe. Sie hat (siehe F) zwei irreduzible Darstellungen vom Grade 1, bei beiden entspricht einer reinen Drehung einfach die Matrix (1), d. h. die Eigenfunktion ist achsensymmetrisch. Bei der einen entspricht auch einer Spiegelung diese Matrix, bei der anderen die Matrix (-1). Alle anderen Darstellungen sind vom Grade 2. In ihnen entspricht einer Drehung um den Winkel γ die Matrix

$$\begin{pmatrix} e^{i\lambda\gamma} & 0 \\ 0 & e^{-i\lambda\gamma} \end{pmatrix} \tag{1}$$

einer Spiegelung in der YZ-Ebene die Matrix

$$\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \tag{1 a}$$

so daß jeder positiven ganzen Zahl λ eine Darstellung entspricht: λ ist der Drehimpuls der Elektronen um die Kernverbindungslinie, soweit er von ihrer Schwerpunktsbewegung herrührt. Bei Termen, die zu einem der beiden ersten Darstellungen gehören, sagen wir, daß $\lambda = 0$ ist, und

* Wie uns Herr F. Hund freundlichst mitteilt, schlägt er in einer demnächst erscheinenden Arbeit fast genau dieselbe Bezeichnungsweise vor, die wir hier verwenden.

den zweiten der beiden Darstellungen unterscheiden wir noch durch einen Strich an der 0'. Die verschiedenen Arten von Elektronentermen sind also benannt mit

$$\lambda = 0,\ 0',\ 1,\ 2,\ 3,\ 4,\ \ldots \tag{2}$$

Zieht man noch die Gleichheit der Elektronen in Betracht, so hat jeder Term noch eine Darstellungseigenschaft der symmetrischen Gruppe gegenüber, diese Darstellung bezeichnen wir mit z.

§ 4. Bei einem Molekül mit zwei gleichen Kernen kommt zur betrachteten Symmetriegruppe noch die Spiegelung in der XY-Ebene hinzu, die ganze Gruppe ist das direkte Produkt dieser Spiegelung mit der vorhin (§ 3) betrachteten Gruppe. Da in den beiden Darstellungen der einfachen Spiegelungsgruppe einer Spiegelung die Matrizen $(+1)$ und (-1) (beide Darstellungen sind vom Grade 1) entsprechen, entstehen aus jeder der vorhin beschriebenen Darstellungen zwei Darstellungen unserer vergrößerten Gruppe. Wir erhalten sie, wenn wir denjenigen Operationen, die keine Spiegelung in der XY-Ebene enthalten, dieselbe Matrix wie in § 3 zuordnen, denjenigen, die eine solche Spiegelung enthalten, entweder auch diese Matrix oder die mit -1 multiplizierte Matrix zuordnen. Im ersten Falle erhalten wir die Darstellungen, die wir mit

$$0_+,\ 0'_+,\ 1_+,\ 2_+,\ \ldots \tag{2a}$$

bezeichnen*, die des zweiten Falles bezeichnen wir mit

$$0_-,\ 0'_-,\ 1_-,\ 2_-,\ \ldots \tag{2a'}$$

§ 5. Wenn λ nicht 0 (besser gesagt nicht 0 oder 0' bzw. 0_+, $0'_+$, 0_- oder $0'_-$) ist**, können wir die beiden Eigenfunktionen, die zum selben Elektronenterm gehören, mit

$$\varphi_{+,\lambda}(Z,\ \mathfrak{r}_1,\ \ldots,\ \mathfrak{r}_n) \quad \text{und} \quad \varphi_{-,\lambda}(Z,\ \mathfrak{r}_1,\ \ldots,\ \mathfrak{r}_n) \tag{3}$$

bezeichnen. Ist λ aber 0 bzw. 0', so haben wir nur je eine Eigenfunktion

$$\varphi_0(Z,\ \mathfrak{r}_1,\ \ldots,\ \mathfrak{r}_n) \quad \text{bzw.} \quad \varphi_{0'}(Z,\ \mathfrak{r}_1,\ \ldots,\ \mathfrak{r}_n). \tag{3a}$$

Wir können in (3) und (3a) neue Koordinaten einführen, nämlich erstens die gegenseitige Konfiguration der Elektronen und Kerne, wofür wir den Buchstaben g setzen und zur Bestimmung der Lage dieser Konfiguration im Raume etwa den Winkel γ, den die $\mathfrak{r}_1 Z$-Ebene mit der YZ-Ebene

* Diese entsprechen den „in den Kernen symmetrischen" Eigenfunktionen von Hund, die 0_-, $0'_-$, 1_- usw. den „in den Kernen antisymmetrischen".

** Wenn wir von λ sprechen, oder wenn wir λ als Index von Eigenfunktionen benutzen, verstehen wir unter λ immer das ganze Symbol von (2) bzw. (2a), (2a'). Wenn dagegen λ in einer Formel als Zahl vorkommt, bedeutet es nur den Zahlenwert von λ, den man also erhält, wenn man das Komma und die $+$ und $-$ wegläßt.

einschließt. Wenn wir also $\varphi_{\pm,\lambda}(Z, \mathfrak{r}_1, \ldots, \mathfrak{r}_n) = \varphi_{\pm,\lambda}(\gamma, g)$ setzen und den Zustand um den Winkel γ' um die Z-Achse verdrehen *, erhalten wir

$$P_{\mathfrak{R}}\, \varphi_{\pm,\lambda}(\gamma + \gamma', g) = e^{\pm i\lambda\gamma'}\, \varphi_{\pm,\lambda}(\gamma + \gamma', g) = \varphi_{\pm}(\gamma, g). \tag{4}$$

Wenn wir $\gamma = 0$ und für γ' wieder γ setzen, ergibt dies

$$\varphi_{\pm,\lambda}(\gamma, g) = e^{\mp i\lambda\gamma}\, \varphi_{\pm,\lambda}(g). \tag{4a}$$

Die Eigenfunktion des in der YZ-Ebene gespiegelten Zustandes sei $\bar{\varphi}$, dann ist wegen (1a)

$$\bar{\varphi}_{+,\lambda} = \varphi_{-,\lambda} \quad \text{und} \quad \bar{\varphi}_{-,\lambda} = \varphi_{+,\lambda}, \tag{5}$$

wenn $\lambda > 0$. Sonst ist

$$\bar{\varphi}_0 = \varphi_0 \quad \text{und} \quad \bar{\varphi}_{0'} = -\varphi_{0'}. \tag{5a}$$

§ 6. Wir nehmen nun an, wir hätten den Fall b) vor uns, d. h. Rotationsenergie groß gegen Wechselwirkungsenergie der Elektronenmagnete. Dann müssen wir jetzt die Rotation des Moleküls berücksichtigen und die Kerne nicht mehr an der Z-Achse festhalten, sondern ihr Azimut α und Polabstand β in die Wellenfunktion einführen. Gleichzeitig erhöht sich die Symmetrie unseres Problems, indem wir nunmehr eine dreidimensionale Drehspiegelungsenergie haben. Diese Tatsache reicht hin (wie wir sehen werden), um die Abhängigkeit von α und β zu bestimmen. Wir haben nämlich, wenn $\mathfrak{R}$ die Drehung mit den Eulerschen Winkeln α, β, γ ist **,

$$P_{\mathfrak{R}}\, \psi_\nu^{l\lambda}(\alpha, \beta, \gamma, g) = \psi_\nu^{l\lambda}(0, 0, 0, g) = \sum_\mu D^l_{\mu\nu}(\alpha, \beta, \gamma) \cdot \psi_\mu^{l\lambda}(\alpha, \beta, \gamma, g), \tag{6}$$

woraus

$$\psi_\mu^{l\lambda}(\alpha, \beta, \gamma, g) = \sum_\nu \tilde{D}^l_{\mu\nu}(\alpha, \beta, \gamma)\, \psi_\nu^{l\lambda}(0, 0, 0, g) \tag{6a}$$

folgt. Für $\alpha = \beta = 0$ wollen wir aber wegen (3) und (4a)

$$\psi_\mu^{l\lambda}(0, 0, \gamma, g) = e^{-i\lambda\gamma}\, \varphi_{+,\lambda}(g) + c\, e^{i\lambda\gamma} \varphi_{-,\lambda}(g)$$

mit vorläufig beliebigem c haben. Es folgt hieraus

$$\psi_\mu^{l\lambda}(\alpha, \beta, \gamma, g) = \tilde{D}^l_{\mu,\lambda}(\alpha, \beta, \gamma)\, \varphi_{+,\lambda}(g) + c\, \tilde{D}^l_{\mu,-\lambda}(\alpha, \beta, \gamma)\, \varphi_{-,\lambda}(g) \tag{6b}$$

für die Eigenfunktion mit der magnetischen Quantenzahl μ, der azimutalen l (Rotationsquantenzahl) und dem Elektronendrehimpuls λ. Ist

---

* D. h. $P_{\mathfrak{R}}\, \varphi(Z, \mathfrak{r}_1, \ldots, \mathfrak{r}_n)$ bilden, mit $\mathfrak{R}$ = Drehung um die Z-Achse um den Winkel γ' (siehe I und II).

** Wir erinnern uns aus F, daß bei dreidimensionaler Symmetrie jeder Term eine azimutale Quantenzahl hat und führen hierfür den Buchstaben l ein. Dieser Term hat $2l + 1$ Eigenfunktionen, zu ihrer Unterscheidung führen wir ν ein: ν läuft von $-l$ bis $+l$.

$\lambda = 0$, so ist das zweite Glied auf der rechten Seite von (6 b) wegzulassen. α, β, γ, g stehen für $X, Y, Z, x_1 \ldots z_n$: es sind α, β, γ die Eulerschen Winkel, durch die $X, Y, Z, x_1, \ldots, z_n$ so verdreht wird, daß $Y' = X' = x_1' = 0$ wird, g ist die gegenseitige Konfiguration der Teilchen.

Zur Bestimmung der Konstante c müssen wir die Spiegelungssymmetrie unseres Problems heranzuziehen. Unter der Spiegelung im Ursprungspunkt verstehen wir das Ersetzen von $X, Y, Z, x_1, \ldots, z_n$ durch $-X, -Y, -Z, -x_1, \ldots, -z_n$. Wir können dies erreichen, wenn wir in (6 b) für $\varphi_{+,\lambda}$ und $\varphi_{-,\lambda}$ wegen (5) $\varphi_{-,\lambda}$ und $\varphi_{+,\lambda}$, für α, β, γ aber $\pi + \alpha$, $\pi - \beta$, $\pi - \gamma$ setzen. Führen wir dies in (6 b) aus, so erhalten wir *

$$\begin{aligned}\overline{\psi}_\mu^{l\lambda}(\alpha, \beta, \gamma, g) &= e^{-i\mu(\pi+\alpha)} d^l_{\mu,\lambda}(\pi - \beta) e^{-i\lambda(\pi-\gamma)} \varphi_{-,\lambda}(g) \\ &\quad + c e^{-i\mu(\pi+\alpha)} d^l_{\mu,-\lambda}(\pi - \beta) e^{i\lambda(\pi-\gamma)} \varphi_{+,\lambda}(g) \\ &= (-1)^{l+\lambda} [e^{-i\mu\alpha} d^l_{\mu,-\lambda}(\beta) e^{i\lambda\gamma} \varphi_{-,\lambda}(g) + c e^{-i\mu\alpha} d^l_{\mu,\lambda}(\beta) e^{-i\lambda\gamma} \varphi_{+,\lambda}(g)],\end{aligned}$$

das letztere wegen

$$d^l_{\mu\lambda}(\pi - \beta) = (-1)^{l+\mu} d^l_{\mu,-\lambda}(\beta).$$

Da bei positiven Termen $\overline{\psi}_\mu^{l\lambda+} = \psi_\mu^{l\lambda+}$, bei negativen $\overline{\psi}_\mu^{l\lambda-} = -\psi_\mu^{l\lambda-}$ sein muß, sind die beiden richtigen Linearkombinationen von (6 a) ** (abgesehen vom Normierungsfaktor)

$$\left.\begin{aligned}\psi_\mu^{l\lambda+} &= \widetilde{D}^l_{\mu,\lambda}(\alpha, \beta, \gamma) \varphi_{+,\lambda}(g) + (-1)^{l+\lambda} \widetilde{D}^l_{\mu,-\lambda}(\alpha, \beta, \gamma) \varphi_{-,\lambda}(g), \\ \psi_\mu^{l\lambda-} &= \widetilde{D}^l_{\mu,\lambda}(\alpha, \beta, \gamma) \varphi_{+,\lambda}(g) - (-1)^{l+\lambda} \widetilde{D}^l_{\mu,-\lambda}(\alpha, \beta, \gamma) \varphi_{-,\lambda}(g).\end{aligned}\right\} \quad (7)$$

Wenn $\lambda = 0$ oder $0'$ ist, fällt das zweite Glied in der Klammer weg, und wir erhalten bei $\lambda = 0$ für gerade l nur positive, für ungerade l nur negative Terme, für $\lambda = 0'$ umgekehrt für gerade l negative, für ungerade l positive Terme.

Haben wir es mit einem Molekül aus zwei gleichen Atomen zu tun, so müssen wir noch bestimmen, ob die Eigenfunktionen (7) in bezug auf Vertauschung der Kerne symmetrisch oder antisymmetrisch sind. Zu diesem Zwecke müssen wir X, Y, Z durch $-X, -Y, -Z$ ersetzen, was wir so tun können, daß wir zuerst φ in der Ebene senkrecht zur

* Hier bedeutet also das Überstreichen die Spiegelung im Ursprung. Dies ist eine etwas andere Transformation, als sie R. de L. Kronig (l. c.) benutzt hat. Sie hat den Vorteil, daß sie offenbar in die Symmetriegruppe des Moleküls gehört, während R. de L. Kronigs Transformation nur bei kleiner Rotation möglich ist (wir vermeiden auch seine Mittelung über α).

** Vgl. die Formel von R. de L. Kronig, l. c.

Kernverbindungslinie spiegeln, bei Termen (2a) ändert dies nichts, bei (2a') gibt dies einen Faktor — 1. Dann ersetzen wir γ durch $\pi + \gamma$, was einen Faktor $(-1)^{\lambda}$ ergibt. Diese Operationen lassen X, Y, Z ungeändert, ersetzen aber x_i, y_i, z_i durch $-x_i$, $-y_i$, $-z_i$. Wir müssen also noch das Ganze im Ursprungspunkt spiegeln, was bei positiven Termen den Faktor 1, bei negativen den Faktor — 1 gibt. Die Eigenfunktion ist also symmetrisch in den Kernen, wenn sie z. B. der Spiegelung im Ursprungspunkt gegenüber positiv ist, wir einen λ_+-Term haben

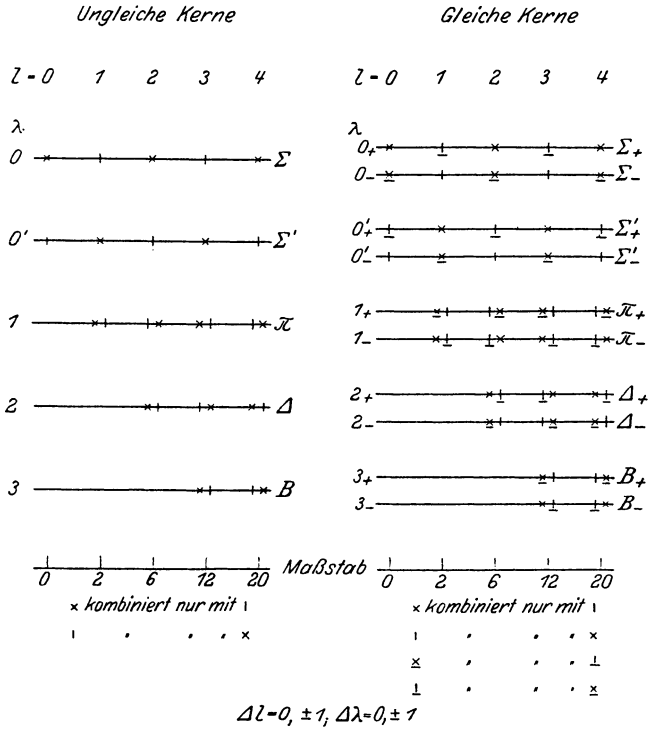


Fig. 1.

und λ gerade ist. Fehlt eine gerade Anzahl dieser Voraussetzungen, dann haben wir eine symmetrische, wenn eine ungerade Anzahl nicht zutrifft, eine antisymmetrische Eigenfunktion. Obwohl wir noch die Spinkoordinaten der Elektronen einführen müßten, wollen wir die gewonnenen Ergebnisse in der Fig. 1 zusammenstellen*. Jeder horizontale Strich repräsentiert einen Typ von Elektronenschwingungstermen, links davon die Darstellungseigenschaft des Elektronenterms § 3 und 4,

* Die Figur gilt natürlich genau für Singuletts.

l bedeutet die Rotationsquantenzahl, positive Terme sind durch $\times$, negative durch | angegeben*. In den Kernen symmetrische sind nicht weiter bezeichnet, in den Kernen antisymmetrische durch einen kleinen Strich unter dem $\underline{\times}$ oder $\underline{|}$. Die Auswahlregeln, die auch bei Hinzunahme der magnetischen Momente der Elektronen streng gelten, sind: positive Terme kombinieren nur mit negativen, negative mit positiven. In den Kernen symmetrische bzw. antisymmetrische Terme kombinieren nur unter sich. Die Regel $\Delta l = \pm 1$ oder 0 gilt nur bei kleiner Multiplettaufspaltung, was wir allerdings im Falle b) voraussetzen können. Die Regel $\Delta \lambda = \pm 1, 0$ gilt nur, solange auch die Rotationsaufspaltung klein ist, praktisch ist sie gut erfüllt**. Der horizontale Maßstab ist in der Tabelle natürlich verzerrt und entspricht nicht dem energetischen, dieser ist vielmehr $l(l+1)$ und ist unten aufgetragen.

Wir schreiben noch die Bandentypen auf, wie sie sich aus dem Vorangehenden (7) ohne weiteres ergeben und auch aus der Tabelle (abgesehen vom Verbot $0 \longleftrightarrow 0'$) abgelesen werden können.

Ungleiche Kerne.

$0 \longleftrightarrow 0$	P und R	Zweig
$0 \longleftrightarrow 0'$	Nichts	„
$0' \longleftrightarrow 0'$	P und R	„
$0, 0' \longleftrightarrow 1$	P, Q und R	„
$\lambda \longleftrightarrow \lambda$	$P, P,\ Q, Q,\ R, R,$	„
$\lambda \longleftrightarrow \lambda + 1$	$P, P,\ Q, Q,\ R, R,$	„

Gleiche Kerne.

$0_+ \longleftrightarrow 0_+$	$0_- \longleftrightarrow 0_-$	Nichts (Ausfall der reinen
$0'_+ \longleftrightarrow 0'_+$	$0'_- \longleftrightarrow 0'_-$	Rotationsschwingungsbanden)
$0_+ \longleftrightarrow 0_-$	$0'_+ \longleftrightarrow 0'_-$	P und R Zweig
$0_+ \longleftrightarrow 0'_+$	$0_- \longleftrightarrow 0'_-$	Nichts
$0_+ \longleftrightarrow 0'_-$	$0_- \longleftrightarrow 0'_+$	„
$0_+, 0'_+ \longleftrightarrow 1_+$	$0_-, 0'_- \longleftrightarrow 1_-$	P, Q, R Zweig
$0_-, 0'_- \longleftrightarrow 1_+$	$0_-, 0'_- \longleftrightarrow 1_-$	Nichts
$\lambda_+ \longleftrightarrow \lambda_+$	$\lambda_- \longleftrightarrow \lambda_-$	„
$\lambda_+ \longleftrightarrow \lambda_-$	$\lambda_- \longleftrightarrow \lambda_+$	$P, P,\ Q, Q,\ R, R,$ Zweig
$\lambda_+ \longleftrightarrow (\lambda+1)_-$	$\lambda_- \longleftrightarrow (\lambda+1)_+$	Nichts „
$\lambda_+ \longleftrightarrow (\lambda+1)_+$	$\lambda_- \longleftrightarrow (\lambda+1)_-$	$P, P,\ Q, Q,\ R, R,$ „

* Unter positiven bzw. negativen Termen verstehen wir also — ebenso wie bei Atomen — solche, deren Eigenfunktionen sich bei der Spiegelung im Ursprungspunkt mit $+1$ bzw. -1 multiplizieren. Bei Atomen wird der Spiegelungscharakter (siehe II) durch die Summe der Azimutalquantenzahlen $l_1 + \cdots + l_n$ der einzelnen Elektronen bestimmt. Ist diese gerade, so haben wir es mit einem positiven, ist sie ungerade, mit einem negativen Term zu tun.

** Die Auswahlregeln leitet man aus (7) leicht ab.

Es tritt also bei gleichen Kernen die zusätzliche Regel auf, daß λ_+-Terme mit λ_+-Termen mit $\Delta\lambda = \pm 1$ und λ_--Terme mit λ_+-Termen mit $\Delta\lambda = 0$ kombinieren. Der Wortlaut dieser Regel stimmt mit dem Wortlaut der Laporteschen Regel überein*, wenn wir l für λ sagen, so daß es naheliegend ist, alle Auswahlregeln, die hier auftreten und für die Molekelspektren charakteristisch sind, folgendermaßen zusammenfassen: λ ändert sich bei einem Übergang um ± 1 oder 0, bei gleichen Kernen ist bei $\Delta\lambda = \pm 1$ nur der Übergang von λ_+ zu $(\lambda \pm 1)_+$ oder λ_- zu $(\lambda \pm 1)_-$, bei $\Delta\lambda = 0$ von λ_+ zu λ_- möglich; $0 \leftrightarrow 0$ und $0' \leftrightarrow 0'$ gibt P- und R-Zweig, $0 \leftrightarrow 0'$ ist verboten, $0 \leftrightarrow 1$ oder $0' \leftrightarrow 1$ gibt einen P-, einen Q- und einen R-Zweig; $\lambda \leftrightarrow \lambda$ oder $\lambda \leftrightarrow \lambda + 1$ gibt zwei P-, zwei Q- und zwei R-Zweige.

Außer diesen gilt noch die Regel $\Delta l = \pm 1, 0$, solange die Multiplettaufspaltung klein ist (also praktisch immer), und die Regeln: $\times$ kombiniert nur mit $|$ bzw. $|$ nur mit $\times$, vollkommen streng.

Wenn der Kern keinen Drehimpuls hat, so fallen im Molekül mit zwei gleichen Atomen entweder die Linien $\times$ und $|$ oder die Linien $\underline{\times}$ und $\underline{|}$ gänzlich fort. Im ersten Falle haben die Kerne Fermische, im zweiten Falle Bose-Einsteinsche Statistik. Ist der Drehimpuls des Kerns $\frac{nh}{4\pi}$, so ist das Verhältnis der Intensität der Übergänge zwischen den in den Kernen symmetrischen Termen (zwischen $\times$ und $|$) zu der Intensität der Übergänge zwischen den in den Kernen symmetrischen Termen ($\underline{\times}$ und $\underline{|}$) gleich $\frac{n+2}{n}$ bzw. $\frac{n}{n+2}$, je nachdem die Kerne Bose-Einsteinsche oder Fermische Statistik haben**.

§ 7. Wir wollen noch die magnetischen Momente der Elektronen berücksichtigen. Die jetzt darzustellenden Verhältnisse des Falles b) sind bei allen Termen der leichten Elemente (etwa bis A) und bei allen Σ-Termen realisiert**. Wir können nach II sofort sagen, daß, wenn die Partitio des Terms (3) bzw. (3a) aus z-Zweiern und $n - 2z = 2r$ Einsern besteht (n Anzahl der Elektronen im Molekül), dann aus einem Term mit $l \leqq \frac{1}{2}n - z = r$ die Terme mit den $2l + 1$ inneren Quantenzahlen $j = r - l,\ r - l + 1,\ \ldots,\ r + l$, aus einem Term mit

* Trotzdem hat dies nichts mit der eigentlichen Laporteschen Regel zu tun, da diese das Verbot von $\times \leftrightarrow \times$ und $| \leftrightarrow |$ bedeutet und bei ganz beliebiger Wechselwirkung streng, während diese Regel nur bei kleiner Rotationsaufspaltung gilt. Eine ähnliche Regel hat F. Hund l. c. angegeben.

** F. Hund, l. c.

$l \geqq \frac{1}{2} n - z = r$ die $2r + 1$-Terme mit den inneren Quantenzahlen $j = l - r, l - r + 1, \ldots, l + r$ entstehen. Auch die Eigenfunktionen können wir nach (9) von III leicht hinschreiben, indem wir

$$\Psi_m^j(\mathfrak{r}_1, \ldots, \mathfrak{r}_n; \tau_1, \ldots, \tau_n) = \sum_{\varepsilon k} s_{mk}^j(r, l)^z C_{k\varepsilon; \tau_1, \ldots, \tau_n} \psi_{m-k, \varepsilon}^l(\mathfrak{r}_1, \ldots, \mathfrak{r}_n)$$

für (7) einsetzen, erhalten wir:

$$\Psi_m^{jl\lambda\pm}(\alpha, \beta, \gamma, g; \tau_1, \ldots, \tau_n)$$
$$= \sum_{k, \varepsilon} s_{mk}^j(l, r) [\widetilde{D}_{k\lambda}^l(\alpha, \beta, \gamma) \varphi_{+,\lambda}^{\varepsilon}(g) \pm (-1)^{l+\lambda} \widetilde{D}_{k,-\lambda}^l(\alpha, \beta, \gamma) \cdot \varphi_{-,\lambda}^{\varepsilon}(g)]$$
$$C_{m-k, \varepsilon; \tau_1, \ldots, \tau_n}^z. \tag{8}$$

Zur Unterscheidung der $\binom{n}{z} - \binom{n}{z-1}$ zur Partitio z gehörenden Eigenfunktionen φ haben wir den Index ε eingeführt*. Ist $\lambda = 0$, so fällt der zweite Summand in der eckigen Klammer weg.

Das Termschema wird hierdurch schon reichlich kompliziert, man muß sich z. B. bei einem Triplettsystem vorstellen, daß alle Terme der Fig 1 (abgesehen von denen mit $l = 0$) dreifach werden. Bezüglich der Intensitäten müssen wir uns vor Augen halten, daß wir es hier — ganz genau so, wie in III — mit einem spinfreien Term mit der Azimutalquantenzahl l und der Partitio z zu tun haben, alle in III abgeleiteten Regeln bleiben bestehen. Da wir indessen also z. B. bei einem Triplettsystem schon aus einer Linie sechs bzw. sieben Linien erhalten, die alle sehr nahe zueinander sind, wird die Multiplettfeinstruktur dieser Banden schwer analysierbar sein**. Theoretisch kann man sie leicht ableiten, wenn man sich nur das Gesagte vor Augen hält, daß wir nämlich einfache Multipletts aus jedem der in Fig. 1 angegebenen Terme erhalten. In dieser Weise ergeben sich auch die Intensitätsformeln.

Im Sinne von III ist nämlich die Summe der Übergangswahrscheinlichkeiten aus einem oberen Zustand (mit bestimmtem m, j und l) in alle unteren Zustände mit vorgegebenem l' (sämtliche m' und j') gleich der Summe der Übergangswahrscheinlichkeiten, die man berechnet, wenn man die Elektronenmagnete nicht einführt und die von einem m und l ausgehenden Übergänge in ein l', aber ebenfalls alle m', betrachtet.

---

* Bezüglich der Bezeichnung vgl. III.

** Glücklicherweise sind immer nur die Linien mit $\Delta l = \Delta j$ intensiv, wie man sich mit Hilfe der Intensitätsregeln bei normalen Kopplungsverhältnissen (mit denen wir es hier eigentlich zu tun haben) (Summenregeln) leicht überlegt. Eine Auswahlregel, die die Linien $\Delta l \neq \Delta j$ verbieten würde, wie sie vielfach angenommen wurde, scheint jedoch nicht zu existieren (siehe das Folgende).

Um alle Intensitätsverhältnisse zu beherrschen, müssen wir also nur diejenigen von Übergängen mit verschiedenen l und μ in verschiedene l' und μ' berechnen. Dies leisten uns im wesentlichen die Hönl-Londonschen* Formeln, die man mit Hilfe von (7) und einiger weiterer Formeln** leicht ableitet.

Hiermit haben wir den Fall b) erledigt und gehen zu a) über.

§ 8. Ist die Multiplettaufspaltung groß der Rotationsaufspaltung gegenüber, so müssen wir zuerst die magnetischen Momente der Elektronen berücksichtigen und erst dann die Rotationsfreiheitsgrade der Kerne.

Wir gehen also von (5) bzw. (5a) aus und führen die Spinkoordinaten $s_1, \ldots, s_n$ ein. Aus einem Term mit der Partition z entstehen $n - 2z + 1 = 2r + 1$ neue antisymmetrische Terme*** mit den Drehimpulsen um die Kernverbindungslinie $\iota = \pm(\lambda - r), \pm(\lambda - r + 1), \ldots, \pm(\lambda + r)$. Den Spinimpuls um die Kernverbindungslinie bezeichnen wir mit η, er kann also die Werte $-r, -r+1, \ldots, r$ annehmen, der gesamte Drehimpuls ist $\iota = \pm(\eta + \lambda)$. Wir haben für die Eigenfunktion****

$$\Phi_{\pm,\lambda,\eta}(\gamma, g, s_1, \ldots, s_n) = \sum_{\varepsilon} C^{z}_{\pm\eta,\varepsilon;s_1,\ldots,s_n} e^{\mp i\lambda\gamma} \varphi^{\varepsilon}_{\pm,\lambda}(g). \tag{9}$$

Die Größe η durchläuft dabei alle Zahlen von $-r$ bis $+r$. Zu einem Elektronenterm gehören die in (9) mit $\pm$ unterschiedenen beiden Eigenfunktionen, Eigenfunktionen mit verschiedenen η gehören zu verschiedenen Elektronentermen. Wenn wir jetzt wieder die Rotation des Moleküls berücksichtigen, müssen wir zwei neue Koordinaten α und β einführen, die Azimut und Polabstand der Kerne bestimmen. Die Symmetriegruppe des Eigenwertproblems erhöht sich indessen zu einer dreidimensionalen Drehspiegelungsgruppe, was wiederum genügen wird, um die Abhängigkeit von α und β zu bestimmen. Es gilt nämlich, wenn j die innere und m die magnetische Quantenzahl ist und $\mathfrak{R}$ die Drehung mit den drei Eulerschen Winkeln α, β, γ:

$$\left.\begin{aligned} & 0_{\mathfrak{R}} \Psi_{\nu}^{j\eta\lambda\pm}(\alpha, \beta, \gamma, g; s_1, \ldots, s_n) \\ &= \sum_{m} D^{j}_{m\nu}(\alpha, \beta, \gamma)\, \Psi_{m}^{j\eta\lambda\pm}(\alpha, \beta, \gamma, g; s_1, \ldots, s_n) \\ &= \sum_{t_1,\ldots,t_n} a_{s_1,\ldots,s_n;t_1,\ldots,t_n}(\alpha, \beta, \gamma)\, \Psi_{\nu}^{j\eta\lambda\pm}(0, 0, 0, g; t_1, \ldots, t_n), \end{aligned}\right\} \tag{10}$$

* ZS. f. Phys. **33**, 803, 1925.

** Ebenda **45**, 601, 1927.

*** Die Bezeichnungen sind wie in II und III.

**** Siehe III.

woraus

$$\Psi_m^{j\eta\lambda\pm}(\alpha, \beta, \gamma, g; s_1, \ldots, s_n) = \sum_{\nu, t_1, \ldots, t_n} \widetilde{D}^j_{m\nu}(\alpha, \beta, \gamma)$$

$$\cdot\, a_{s_1, \ldots, s_n; t_1, \ldots, t_n}(\alpha, \beta, \gamma)\, \Psi_\nu^{j\eta\lambda\pm}(0, 0, 0, g; t_1, \ldots, t_n) \tag{10a}$$

folgt. Setzen wir hierin $\alpha = \beta = 0$, so muß die Abhängigkeit von $\gamma, g, s_1, \ldots, s_n$ durch (9) gegeben sein. Wir erhalten also *

$$\Psi_m^{j\eta\lambda}(\alpha, \beta, \gamma, g; s_1, \ldots, s_n)$$

$$= \sum_{\varepsilon, t_1, \ldots, t_n} \widetilde{D}^j_{m, \lambda+\eta}(\alpha, \beta, \gamma)\, a_{s_1, \ldots, s_n; t_1, \ldots, t_n}(\alpha, \beta, \gamma) \cdot C^z_{\eta, \varepsilon; t_1, \ldots, t_n}\, \varphi^\varepsilon_{+, \lambda}(g)$$

$$+ c\, \widetilde{D}^j_{m, -\lambda-\eta}(\alpha, \beta, \gamma)\, a_{s_1, \ldots, s_n; t_1, \ldots, t_n}(\alpha, \beta, \gamma)\, C^z_{-\eta, \varepsilon; t_1, \ldots, t_n}\, \varphi^\varepsilon_{-, \lambda}(g), \tag{10b}$$

was mit Hilfe der Formel (5) von III

$$\Psi_m^{j\eta\lambda}(\alpha, \beta, \gamma, g; s_1, \ldots, s_n)$$

$$= \sum_{\varepsilon\mu} \Big\{ \widetilde{D}^j_{m, \lambda+\eta}(\alpha, \beta, \gamma)\, D^{1/2\, n-z}_{\mu, \eta}(\alpha, \beta, \gamma)\, C^z_{\mu, \varepsilon; s_1, \ldots, s_n}\, \varphi^\varepsilon_{+, \lambda}(g)$$

$$+ c\, \widetilde{D}^j_{m, -\lambda, -\eta}(\alpha, \beta, \gamma)\, D^{1/2\, n-z}_{\mu, -\eta}(\alpha, \beta, \gamma)\, C^z_{\mu, \varepsilon; s_1, \ldots, s_n}\, \varphi^\varepsilon_{-, \lambda}(g) \Big\} \tag{10c}$$

ergibt. Hierin müssen wir noch c bestimmen, was ebenso wie in § 5 mit Hilfe der Spiegelungssymmetrie unseres Problems geschieht. Wenn wir nämlich in Ψ die Koordinaten $X, Y, Z, x_1, \ldots, z_n$ durch $-X, -Y, -Z, -x_1, \ldots, -z_n$ ersetzen, müssen wir, je nachdem der Term positiv oder negativ ist, $+\Psi$ oder $-\Psi$ erhalten. Anstatt das Vorzeichen der Koordinaten zu ändern, können wir auch $\varphi^\varepsilon_{+, \lambda}$ und $\varphi^\varepsilon_{-, \lambda}$ durch $\varphi^\varepsilon_{-, \lambda}$ und $\varphi^\varepsilon_{+, \lambda}$, außerdem α, β, γ durch $\pi + \alpha, \pi - \beta, \pi - \gamma$ ersetzen. Wir erhalten (abgesehen vom Normierungsfaktor) **

$$\Psi_m^{j\eta\lambda\pm}(\alpha, \beta, \gamma, g; s_1, \ldots, s_n)$$

$$= \sum_{\varepsilon\mu} \Big\{ \widetilde{D}^j_{m, \lambda+\eta} D^r_{\mu, \eta}\, C^z_{\mu, \varepsilon; s_1, \ldots, s_n}\, \varphi^\varepsilon_{+, \lambda}(g)$$

$$\pm (-1)^{j+r+\lambda} \widetilde{D}^j_{m, -\lambda-\varepsilon} D^r_{\mu, -\eta}\, C^z_{\mu, \varepsilon; s_1, \ldots, s_n}\, \varphi^\varepsilon_{-, \lambda}(g). \tag{11}$$

Diese Formel gilt auch für $\iota = \eta + \lambda = 0$ oder $\eta = 0$. Für $\lambda = 0$ gilt sie hingegen nicht, was aber für uns nichts ausmacht, da wir es bei $\lambda = 0$ immer mit dem Falle b) zu tun haben.

§ 9. Bevor wir zur Aufbauregel übergehen, wollen wir die Auswahlregeln der Übergänge zwischen Termen vom Falle a) berechnen.

---

* $C^z_{\varepsilon, \eta; t_1, \ldots, t_n}$ ist nur für $t_1 + \cdots + t_n = \eta$ von 0 verschieden.

** $\frac{1}{2}\, n - z = r$.

Wir haben das Integral

$$\int \Psi_m^{j\eta\lambda+} \tilde{\Psi}_m'^{j'\eta'\lambda'-} X^3 \tag{12}$$

auszuwerten, wenn wir mit X^3 das Moment

$$(X^3 = z_1 + \cdots + z_n + Z\,.\,\text{Kernladungszahl})$$

bezeichnen. Für X^3 können wir auch $\sum_\nu D^1_{0\nu}(\alpha,\beta,\gamma)\, X_g^\nu$ setzen, wenn wir unter X_g^ν die Momente im Koordinatensystem verstehen, der auch zur Beschreibung von g dient. (Seine Z-Achse läuft durch den Kern, seine XZ-Ebene durch das erste Elektron, siehe § 5.) Setzen wir (11) in (12) ein, so lassen sich die auftretenden recht langen Ausdrücke auffallend leicht berechnen. Außer den bekannten Auswahlregeln ($\Delta j = \pm 1$ oder 0; $\Delta\lambda = \pm 1$ oder 0; usw.) erhalten wir noch die: $\Delta\eta = 0$, was dem Ausfallen sehr vieler Linien Rechnung trägt*.

Wir wollen noch die erhaltenen Resultate kurz zusammenfassen.

1. Ungleiche Kerne. Im Falle b), also bei leichten Elementen und allen Σ-Termen, wird ein Feinstrukturterm bestimmt durch den Elektronendrehimpuls λ, der die Werte 0, 0′, 1, 2, 3, ... durch die Azimutalquantenzahl l, die nur ganzzahlige Werte 0, 1, 2, ... annehmen, durch den positiven bzw. negativen Charakter und durch die innere Quantenzahl j, die alle Werte $|l-r|$, $|l-r|+1, \ldots, l+r$ annehmen kann. Mit $2r+1$ ist die Multiplizität des Terms** bezeichnet, so daß j bei Molekülen mit ungerader Elektronenzahl halbzahlig ist, bei gerader Elektronenzahl ganzzahlig ist. Die Auswahlregeln sind bei der Tabelle 1 für λ, l und Spiegelungscharakter beschrieben, für j gilt $\Delta j = \pm 1$, 0, wie immer. Auch die Intensitäten sind, wie im Falle normaler Serienspektren, die sie auch in Hinsicht der Multiplettstruktur sind.

Im Falle a), also bei Π-, Δ- usw. Termen schwerer Elemente, wird ein Term durch den Elektronendrehimpuls λ, der alle ganzzahligen Werte (bei Π-Termen 1, Δ-Termen 2 usw.) annehmen kann, durch das η, das von $-r$ bis $+r$ läuft, durch die innere (Rotations-) Quantenzahl j, die die Werte $|\lambda+\eta|$, $|\lambda+\eta|+1$, $|\lambda+\eta|+2, \ldots$ annehmen kann, und durch den positiven bzw. negativen Charakter des Terms bestimmt. Die Auswahlregeln sind $\Delta\lambda = \pm 1$, 0, positive Terme kombinieren nur mit negativen und umgekehrt, $\Delta\eta = 0$ und $\Delta j = \pm 1$ oder 0.

Bei den Kombinationen vom Falle a) zum Falle b) ($\Sigma - \Pi$-Übergänge bei höheren Elementen) gelten nur die Regeln $\Delta j = \pm 1$, 0;

* Diese Regel ist schon bekannt (siehe Zitate in § 1).

** Also im Singulettsystem $r = 0$, im Dublettsystem $r = {}^1/_2$ usw.

$\Delta\lambda = \pm 1, 0$, und daß positive Terme nur mit negativen und umgekehrt kombinieren. In den Diagrammen findet man gewöhnlich noch einige Linien punktiert angegeben, diese sind jedoch keine verbotenen Übergänge, auch wenn man die Rotations- und Multiplettenergie als klein annimmt, sondern nur solche, deren Intensitäten sich bei der Berechnung als verhältnismäßig klein herausstellen.

2. Moleküle mit gleichen Atomen. Für Moleküle mit zwei gleichen Atomen gelten zunächst alle Regeln, die soeben angegeben worden sind, natürlich ebenfalls. Es kommen nur einige weitere Auswahlverbote hinzu, die darauf beruhen, daß Terme, die in den Kernen symmetrisch sind, nicht mit Termen, die in den Kernen antisymmetrisch sind, kombinieren. Hierdurch werden nicht etwa Intensitätsregeln verletzt, da diese nur das Intensitätsverhältnis solcher Linien bestimmen, die entweder alle ausfallen, oder alle vorkommen.

Wir müssen also nur noch bestimmen, welche Terme von den vorhin beschriebenen symmetrisch und welche antisymmetrisch in den Kernen sind.

Im Falle b) ist dies einfach dadurch geschehen, daß wir bemerken, daß das Hinzufügen der magnetischen Momente der Elektronen an der Symmetrie bzw. Antisymmetrie des Terms in bezug auf die Kerne nichts ändert. Im Triplettsystem entstehen z. B. aus einem $\underline{\times}$-Term mit $l > 0$ drei Terme mit $j = l - 1, l, l + 1$, aber alle sind ebenfalls $\underline{\times}$-Terme. Etwas genauer müssen wir uns den Fall a) überlegen. Die Frage ist hier, ob $\Psi_m^{j\eta\lambda\pm}$ eine symmetrische oder antisymmetrische Funktion der Kerne ist, d. h. ob sie das Vorzeichen bei dem Ersetzen von X, Y, Z durch $-X$, $-Y$, $-Z$ beibehält oder verändert. Wie spiegeln zunächst wieder die Elektronen in der Symmetrieebene der Kerne, was einen Faktor $+l$ oder $-l$ gibt, je nachdem λ aus (2a) (λ_+) oder (2a') (λ_-) stammt. Dann ersetzen wir γ durch $\pi + \gamma$, was den Faktor $(-1)^\lambda$ ergibt. Jetzt haben wir $x_1, \ldots, z_n$ durch $-x_1, \ldots, -z_n$ ersetzt, während XYZ unverändert blieben. Wir ersetzen noch alle Koordinaten durch ihre negativen Werte, d. h. spiegeln das System im Ursprungspunkt, was den Faktor ± 1 ergibt.

In den Fig. 2 und 3 sind wiederum einige Diagramme für den Fall b) bzw. a) gegeben.

§ 10. Das Aufbauprinzip. Wenn man die Elektronen unter dem Einfluß von zwei sehr weit voneinander festgehaltenen Kernen betrachtet, so ist es leicht, die Eigenfunktionen und Eigenwerte dieser Zustände anzugeben. Die Eigenfunktionen sind einfach Produkte der Eigen-

funktionen von zwei Atomzuständen, der zugehörige Eigenwert ist die Summe der betreffenden Eigenwerte der Atome. Man wird also* die Eigenfunktionen und Eigenwerte in dem Falle, daß die Kerne nicht mehr sehr weit von einander sind, so zu berechnen versuchen, daß man die weit getrennten Kerne einander nähert. Wenn sie schließlich in einem Punkte zusammengetroffen sind, kennt man die Eigenwerte wieder (aus spektroskopischen Daten), da wir es dann mit einem Atom zu tun haben.

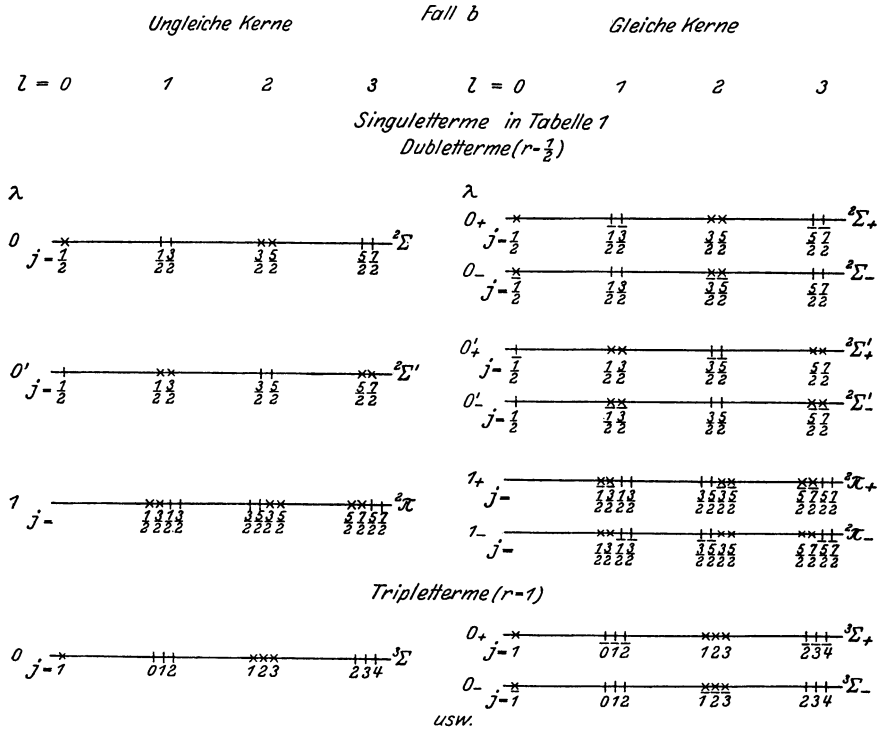


Fig. 2.

Der Vorschlag, die Elektronenterme der Moleküle auf diese Weise abzuschätzen, stammt von F. Hund*. W. Heitler und F. London** machen dagegen geltend, daß nicht alle auf diese Weise berechneten Energiewerte stationären Zuständen des Moleküls entsprechen werden, viele vielmehr im kontinuierlichen Spektrum liegen werden.

---

* F. Hund, l. c.

** W. Heitler und F. London, ZS. f. Phys. **44**, 455, 1927. Auch aus den Überlegungen von M. Born und J. R. Oppenheimer (l. c.) scheint dies hervorzugehen. Vgl. auch F. Hund, l. c. S. 761.

Diesem Umstand wollen wir Rechnung tragen, indem wir folgendermaßen verfahren. Wir denken uns zunächst die weit getrennten Atome einander genähert und berechnen, was für Elektronenterme des Moleküls aus zwei bestimmten Atomtermen entstehen. Diese Terme nennen wir

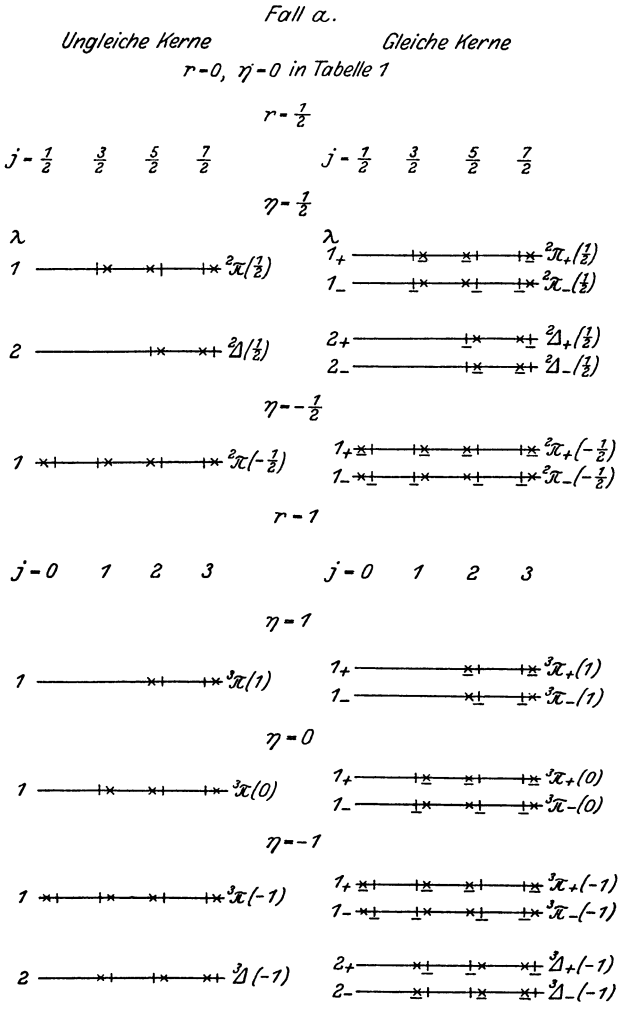


Fig. 3.

die „gruppentheoretisch möglichen Terme", da man bei ihrer Bestimmung nur von gruppentheoretischen Überlegungen ausgeht und auf energetische Fragen keinerlei Rücksicht nimmt. Zweifellos scheint, daß alle Terme, deren Energie bei großen Schwingungsquantenzahlen zu dem betreffenden Energiewerte der getrennten Atome konvergiert, auf diese Weise erfaßt

werden, wir werden sogar im allgemeinen viel mehr gruppentheoretisch mögliche Terme haben, als wirkliche existieren.

Nur zur weiteren Abschätzung der Molekelterme werden wir die Terme des Atoms, das durch Vereinigen der beiden Kerne entsteht, benutzen. Dabei muß man bei der Zuordnung streng auf die Aufrechterhaltung sämtlicher Symmetrieeigenschaften achten*. Solche „gruppentheoretisch möglichen Terme", denen auf diese Weise kein tiefliegender Atomterm des Atoms mit vereinigten Kernen zugeordnet werden kann, muß auch selber hoch — wahrscheinlich im kontinuierlichen Spektrum — liegen. Er existiert vom spektroskopischen Standpunkte aus nicht. Die Bestimmung derjenigen möglichen Terme, die zu diskreten stationären Zuständen des Moleküls führen, ist eine recht schwierige Aufgabe, die man wahrscheinlich nur — wie bei Atomen — an Hand von erfahrungsmäßig bewährten Modellen lösen können wird.

Für ungleiche Atome in S-Zuständen hat F. London** die „gruppentheoretisch möglichen Terme" bereits bestimmt. Er geht von Gleichung (14) für die Eigenfunktionen aus, die eine Darstellung der symmetrischen Gruppe von $m + m' = n$ Elementen vermitteln. Er reduziert diese Darstellung aus und bestimmt die irreduziblen Bestandteile nach einer sehr schönen Methode, womit seine Aufgabe gelöst ist. Da die Ausdehnung seiner Resultate auf P-, D- usw. Terme keinerlei Schwierigkeiten bietet, hat er die möglichen Terme bei ungleichen Atomen eigentlich bestimmt.

Etwas anders steht es bei gleichen Atomen, namentlich in dem — praktisch sehr wichtigen — Falle, wenn sie im selben Zustand sind. Hier müßte man eine wesentlich kompliziertere gruppentheoretische Aufgabe lösen***, die wir, obgleich sie durchaus lösbar ist, umgehen wollten. Deshalb haben wir eine von der seinigen wesentlich abweichende Darstellung gegeben, die auch im Falle von ungleichen Atomen etwas ein-

* Bei F. Hund ist dies nicht der Fall. Wir glauben, daß zum Teil dies die schlechte Übereinstimmung seiner Resultate mit der Erfahrung verursacht.

** ZS. f. Phys. **50**, 24, 1928. Bei der Abfassung dieser Note kannten wir bereits seine Ergebnisse und möchten ihm hierfür auch an dieser Stelle bestens danken. Vgl. auch W. Heitler, ZS. f. Phys. **47**, 835, 1928.

*** Das Bestimmen der irreduziblen Darstellungen derjenigen Gruppe, die folgende Permutationen von $2m$ Elementen enthält: 1. Diejenigen, die die ersten n Elemente unter sich und die letzten n Elemente unter sich permutieren. 2. Diejenigen, die die ersten n Elemente an die letzten n Stellen, die letzten n Elemente an die ersten n Stellen bringt. Dann müßte man wieder bestimmen, welche irreduziblen Darstellungen dieser Gruppe in einer ins Auge gefaßten Darstellung der ganzen Permutationsgruppe enthalten sind.

facher zu den mit den seinigen äquivalenten Resultaten führt, während sie bei gleichen Atomen eine vollständige Bestimmung aller Symmetrieeigenschaften gestattet.

Auch die zweite Frage bezüglich der wirklichen Existenz der „möglichen" Terme wird von London untersucht. Er stellt gewisse Formeln auf, die indessen notwendigerweise gewisse Integrale enthalten, über die wir nichts wissen. Unsere Methode hat den Vorteil, daß sie mit viel einfacheren Hilfsmitteln als die strengeren Rechnungen von London auskommt.

§ 11. Bestimmung der gruppentheoretisch möglichen Terme. Um die λ und Partitio (z) des ganzen Moleküls zu bestimmen, gehen wir zuerst von unendlich weit getrennten Atomen aus, in denen die Spin der Elektronen schon berücksichtigt sind. Durch Zusammenführen erhalten wir die vorkommenden $\iota = \lambda + \eta$ und, sofern es notwendig ist, den Spiegelungscharakter (§ 11). Wenn wir andererseits die Atome, ohne die Spin einzuführen, zusammenbringen, erhalten wir ohne Weiteres die Zahlenwerte der vorkommenden λ. Wenn auch die Partitiones dieser Terme bekannt wären, könnten wir aus den λ und z die ι berechnen (§ 7). Umgekehrt können wir, da wir die ι schon (aus der ersten Überlegung) kennen werden, aus ihnen die Partitiones z berechnen (§ 12).

A. Betrachten wir zwei unendlich weit getrennte verschiedene Atome

Der eine Kern liege im Punkte 0, 0, Z, der andere in 0, 0, Z′, wobei $\mathsf{Z}' = -\frac{M_2}{M_1}\mathsf{Z}$ ist. Die beiden Atome sind, jedes für sich, in stationären Zuständen, das erste hat m Elektronen, die innere Quantenzahl j und die Eigenfunktionen

$$\psi_\mu(x_1, y_1, z_1, \ldots, x_m, y_m, z_m; s_1, \ldots, s_m) \qquad (13)$$
$$(\mu \text{ von } -j \text{ bis } j),$$

das zweite m' Elektronen, die inneren Quantenzahlen j' und die Eigenfunktionen

$$\psi'_{\mu'}(x_{m+1}, y_{m+1}, z_{m+1}, \ldots, x_{m+m'}, y_{m+m'}, z_{m+m'}; s_{m+1}, \ldots, s_{m+m'}) \qquad (13\text{a})$$
$$(\mu' \text{ von } -j' \text{ bis } j').$$

Die Hyperfunktionen ψ_μ und $\psi'_{\mu'}$ sind antisymmetrisch. Die Eigenfunktionen des ganzen Zustandes sind die Hyperfunktionen

$$\begin{aligned}&\Psi_{\mu\mu'\alpha_1,\ldots,\alpha_m}(\mathfrak{r}_1, \ldots, \mathfrak{r}_{m+m'}, s_1, \ldots, s_{m+m'})\\ &= \psi_\mu(\mathfrak{r}_{\alpha_1} - \mathsf{Z}, \ldots, \mathfrak{r}_{\alpha_m} - \mathsf{Z}; s_{\alpha_1}, \ldots, s_{\alpha_m})\\ &\cdot \psi'_{\mu'}(\mathfrak{r}_{\alpha_{m+1}} - \mathsf{Z}', \ldots, \mathfrak{r}_{\alpha_{m+m'}} - \mathsf{Z}'; s_{\alpha_{m+1}}, \ldots, s_{\alpha_{m+m'}}),\end{aligned} \qquad (14\text{a})$$

worin Z und Z′ als Vektoren mit den Komponenten 0, 0, Z bzw. 0, 0, Z aufzufassen sind. In (14a) haben wir $(2j+1)(2j'+1)\binom{m+m'}{m}$ linear unabhängige Eigenfunktionen, da es wegen der Antisymmetrie der ψ_μ und $\psi'_{\mu'}$ weder auf die Reihenfolge der $\alpha_1, \ldots, \alpha_m$ noch der $\alpha_{m+1}, \ldots, \alpha_{m+m'}$ in der Permutation $\alpha_1, \ldots, \alpha_m, \alpha_{m+1}, \ldots, \alpha_{m+m'}$ ankommt. Bei einem bestimmten μ und μ' haben wir noch $\binom{m+m'}{m}$ Eigenfunktionen, von denen man aber nur eine antisymmetrische Linearkombination bilden kann *, nämlich

$$\Psi_{\mu\mu'} = \sum_{\mathfrak{R}} \varepsilon_{\mathfrak{R}}\, \psi_\mu(\mathfrak{r}_{\alpha_1} - \mathsf{Z}, \ldots, \mathfrak{r}_{\alpha_m} - \mathsf{Z};\, s_{\alpha_1}, \ldots, s_{\alpha_m})$$
$$\cdot\, \psi'_{\mu'}(\mathfrak{r}_{\alpha_{m+1}} - \mathsf{Z}', \ldots, \mathfrak{r}_{\alpha_{m+m'}} - \mathsf{Z}',\, s_{\alpha_{m+1}}, \ldots, s_{\alpha_{m+m'}}), \qquad (15\,\mathrm{a})$$

worin $\mathfrak{R}$ die Permutation $\alpha_1, \ldots, \alpha_m, \alpha_{m+1}, \ldots, \alpha_{m+m'}$ der Zahlen $1, 2, \ldots, m, m+1, \ldots, m+m'$ und $\varepsilon_{\mathfrak{R}} = 1$ oder -1, je nachdem diese gerade oder ungerade ist.

Aus (15a) ersehen wir, daß wir zu jedem Paare μ, μ' eine Eigenfunktion erhalten, die sich bei der Drehung um die Z-Achse mit dem Winkel γ mit $e^{i(\mu+\mu')\gamma}$ multipliziert, den Drehimpuls $\mu + \mu'$ in dieser Richtung hat.

B. Bei zwei gleichen Atomen in verschiedenen Zuständen haben wir außer (14a) noch die Eigenfunktionen

$$\Psi'_{\mu\mu'\alpha_1,\ldots,\alpha_m}(\mathfrak{r}_1, \ldots, \mathfrak{r}_{2m};\, s_1, \ldots, s_{2m})$$
$$= \psi'_\mu(\mathfrak{r}_{\alpha_1} - \mathsf{Z}, \ldots, \mathfrak{r}_{\alpha_m} - \mathsf{Z};\, s_{\alpha_1}, \ldots, s_{\alpha_m})$$
$$\cdot\, \psi_{\mu'}(\mathfrak{r}_{\alpha_{m+1}} - \mathsf{Z}', \ldots, \mathfrak{r}_{\alpha_{2m}} - \mathsf{Z}';\, s_{\alpha_{m+1}}, \ldots, s_{\alpha_{2m}}), \qquad (14\,\mathrm{b})$$

es ist diesmal $\mathsf{Z}' = -\mathsf{Z}$ und $m = m'$. Wir haben zu jedem Paare μ, μ' zwei antisymmetrische Hyperfunktionen, nämlich außer (15a) noch

$$\Psi'_{\mu\mu'} = \sum_{\mathfrak{R}} \varepsilon_{\mathfrak{R}}\, \psi'_\mu(\mathfrak{r}_{\alpha_1} - \mathsf{Z}, \ldots, \mathfrak{r}_{\alpha_m} - \mathsf{Z};\, s_{\alpha_1}, \ldots, s_{\alpha_m})$$
$$\cdot\, \psi_{\mu'}(\mathfrak{r}_{\alpha_{m+1}} + \mathsf{Z}, \ldots, \mathfrak{r}_{\alpha_{2m}} + \mathsf{Z};\, s_{\alpha_{m+1}}, \ldots, s_{\alpha_{2m}}). \qquad (15\,\mathrm{b})$$

* Man kann ja auch in dem Falle, daß alle $(m+m')!$ voneinander linear unabhängig sind, nur eine bilden.

Wir wollen ihren Spiegelungscharakter bei der Spiegelung im Ursprungspunkt bestimmen*. Je nachdem der erste bzw. der zweite betrachtete Term des Atoms positiv oder negativ war, sei w_1 bzw. w_2 gleich $+1$ oder -1. Wenn man in (15a) $-\mathfrak{r}_i$ für $\mathfrak{r}_i$ einsetzt, erhält man

$$\begin{aligned}
&\Psi_{\mu\mu'}(-\mathfrak{r}_1, \ldots, -\mathfrak{r}_m, -\mathfrak{r}_{m+1}, \ldots, -\mathfrak{r}_{2m}; s_1, \ldots, s_{2m}) \\
&= \sum_{\mathfrak{R}} \varepsilon_{\mathfrak{R}}\, \psi_\mu(-\mathfrak{r}_{\alpha_1} - \mathsf{Z}, \ldots; s_{\alpha_1}, \ldots)\, \psi'_{\mu'}(-\mathfrak{r}_{\alpha_{m+1}} + \mathsf{Z}, \ldots; s_{\alpha_{m+1}}, \ldots) \\
&= \sum_{\mathfrak{R}} \varepsilon_{\mathfrak{R}}\, w_1 w_2\, \psi_\mu(\mathfrak{r}_{\alpha_1} + \mathsf{Z}, \ldots; s_{\alpha_1}, \ldots)\, \psi'_{\mu'}(\mathfrak{r}_{\alpha_{m+1}} - \mathsf{Z}, \ldots; s_{\alpha_{m+1}}, \ldots) \\
&= w_1 w_2\, \Psi'_{\mu'\mu}(\mathfrak{r}_{\alpha_{m+1}}, \ldots, \mathfrak{r}_{\alpha_{2m}}, \mathfrak{r}_{\alpha_1}, \ldots, \mathfrak{r}_{\alpha_m}; \\
&\qquad\qquad s_{\alpha_{m+1}}, \ldots, s_{\alpha_{2m}}, s_{\alpha_1}, \ldots, s_{\alpha_m}) \\
&= (-1)^m w_1 w_2\, \Psi'_{\mu'\mu}(\mathfrak{r}_{\alpha_1}, \ldots, \mathfrak{r}_{\alpha_m}, \mathfrak{r}_{\alpha_{m+1}}, \ldots, \mathfrak{r}_{\alpha_{2m}}; s_{\alpha_1}, \ldots, s_{\alpha_{2m}}), \qquad (16)
\end{aligned}$$

das letzte wegen der Antisymmetrie von $\Psi'_{\mu'\mu}$. Es ist also

$$\Psi_{\mu\mu'} + (-1)^m w_1 w_2\, \Psi'_{\mu'\mu} \qquad (16\text{a})$$

eine positive und

$$\Psi_{\mu\mu'} - (-1)^m w_1 w_2\, \Psi'_{\mu'\mu} \qquad (16\text{b})$$

eine negative Eigenfunktion ist. Bei gleichen Atomen in verschiedenen Zuständen erhalten wir also zu jedem Paare μ, μ' eine „positive" und eine „negative" Eigenfunktion mit dem Drehimpuls $\mu + \mu'$.

C. Sind schließlich die Atome in gleichen Zuständen, so ist $\Psi'_{\mu\mu'} = \Psi_{\mu\mu'}$, und wir erhalten, wenn $\mu \neq \mu'$ ist, zu jedem Paare μ, μ', diesmal ohne Rücksicht auf die Reihenfolge, eine „positive" und eine „negative" Eigenfunktion. Wenn $\mu = \mu'$ ist, so verschwindet entweder (16a) oder (16b) identisch, und wir erhalten eine positive oder eine negative Eigenfunktion, je nachdem ($w_1 = w_2$) die Anzahl der Elektronen in einem Atom m gerade oder ungerade ist.

Nach dieser Vorbereitung können wir zu unserem eigentlichen Problem zurückkehren.

§ 12. Wir haben zwei Atome mit m bzw. m' Elektronen, ihre azimutalen Quantenzahlen sind l und l', die Spiegelungscharaktere w und w' und die Partitiones z und z'. Wir setzen r und r' für $\frac{1}{2}m - z$ bzw. $\frac{1}{2}m' - z'$ (die r bedeuten dann so etwas wie Rumpfimpuls). Unsere Aufgabe besteht nun darin, daß wir die Terme des Moleküls bei festgehaltenen Kernen (Elektronenterme) kennenlernen wollen, die aus diesen Atomen durch adiabatisches Zusammenführen der Kerne entstehen, d. h. ihre Darstellungs-

* Nur dieser ist eindeutig, wenn man die Spin eingeführt hat (vgl. I und II).

eigenschaften in bezug auf Vertauschung der Elektronenschwerpunkte und Drehspiegelung um die Z-Achse bestimmen wollen. Die spinfreien Eigenfunktionen der Atome seien $\psi_{\nu\zeta}$ und $\psi'_{\nu'\zeta'}$, worin der Index ζ zur Unterscheidung der $\binom{m}{z} - \binom{m}{z-1}$ zur Partitio z gehörenden Eigenfunktionen dient.

A. Bei ungleichen Kernen haben wir die Eigenfunktionen für das Molekül

$$\Psi_{\nu\nu'\zeta\zeta'\alpha_1\ldots\alpha_m} = \psi_{\nu\zeta}(\mathfrak{r}_{\alpha_1} - \mathsf{Z}, \ldots, \mathfrak{r}_{\alpha_m} - \mathsf{Z})\, \psi'_{\nu'\zeta'}(\mathfrak{r}_{\alpha_{m+1}}, \ldots, \mathfrak{r}_{\alpha_{m+m'}}), \quad (17)$$

worin wir die $\alpha_1, \ldots, \alpha_m$ und die $\alpha_{m+1}, \ldots, \alpha_{m+m'}$ in der natürlichen Reihenfolge denken können. Wir sehen unmittelbar, daß wir zu jedem Paare ν, ν' dieselben Partitiones für das Molekül erhalten, ja, daß dieses von den azimutalen Quantenzahlen l und l' ganz unabhängig ist.

Wir können also zur Bestimmung dieser Partitiones $l = l' = 0$ setzen, dann ist auch $\nu = \nu' = 0$. Wenn wir die R-Werte R_1, R_2 usw. für das Molekül erhalten, so müssen wir bei der Einführung der Spin die folgenden ι-Werte erhalten (§ 7):

$$\left.\begin{array}{lll} \text{wegen } R_1: & R_1, R_1 - 1, \ldots & -R_1, \\ \text{wegen } R_2: & R_2, R_2 - 1, \ldots & -R_2 \text{ usw.} \end{array}\right\} \quad (18\text{a})$$

Andererseits wissen wir, daß, wenn wir die Spin noch bei den getrennten Atomen einführen, wir bei dem ersten Atom ein $j = r$, bei dem zweiten ein $j' = r'$ erhalten. Dies ergibt die ι-Werte

$$\left.\begin{array}{llll} r + r', & r + r' - 1, & \ldots, & r - r' \\ & r + r' - 1, & \ldots, & r - r', r - r' - 1 \\ & \cdots\cdots & \cdots\cdots & \cdots\cdots \\ & & -r + r', & \ldots \quad -r - r'. \end{array}\right\} \quad (18\text{b})$$

Der Komplex (18a) muß mit dem Komplex (18b) identisch sein. Dies ergibt*

$$R_1 = r + r', \quad R_2 = r + r' - 1, \ldots \text{ bis } R = |r - r'|. \quad (19)$$

Diese Partitiones entstehen für die Molekülterme für jedes Paar ν, ν'. Andererseits ergibt das Paar ν, ν' und $-\nu$, $-\nu'$ ein $\lambda = |\nu + \nu'|$, nur wenn $\nu' = -\nu$ ist, haben wir ein $\lambda = 0$ und ein $\lambda = 0'$ und, wenn $\nu = \nu' = 0$, ein $\lambda = 0$ oder $0'$, je nachdem $(-1)^{l+l'} ww'$ gleich $+1$ oder -1 ist.

* Dies ist, da $2r$ die Anzahl der Einser in der Partitio ist, genau das Londonsche Resultat. Vgl. ZS. f. Phys. **50**, 24, 1928.

Hiermit haben wir im Falle von ungleichen Atomen das Aufbauprinzip: aus zwei Atomtermen mit l, w, r und l', w', r' (es sei $l \geqq l'$) entstehen die λ

$$\left.\begin{array}{llllll} l+l', & l+l'-1, & & \ldots & 1,0 & 0' \\ & l+l'-1, & & \ldots & 1,0 & 0' \\ & \ldots\ldots\ldots\ldots\ldots\ldots & & & & \\ & & l,\ l-1, & \ldots & 1,0 \text{ oder* } & 0' \\ & & l-1, & \ldots & 1, & \\ & & \ldots\ldots\ldots\ldots & & & \\ & & l-l', & \ldots & 1. & \end{array}\right\} \quad (20)$$

Jedes hier aufgeschriebene λ gibt je einen Term mit allen in (19) gegebenen R_1, R_2, ...

B. Im Falle von zwei gleichen Atomen in verschiedenen Zuständen haben wir außer den Eigenfunktionen

$$\Psi_{\nu\nu'\zeta\zeta'\alpha_1,\ldots,\alpha_m} = \psi_{\nu\zeta}(\mathfrak{r}_{\alpha_1} - \mathsf{Z}, \ldots)\, \psi'_{\nu'\zeta'}(\mathfrak{r}_{\alpha_{m+1}} + \mathsf{Z}, \ldots) \quad (17\,\mathrm{a})$$

noch die

$$\Psi'_{\nu\nu'\zeta\zeta'\alpha_1,\ldots,\alpha_m} = \psi'_{\nu'\zeta'}(\mathfrak{r}_{\alpha_1} - \mathsf{Z}, \ldots)\, \psi_{\nu\zeta}(\mathfrak{r}_{\alpha_{m+1}} + \mathsf{Z}, \ldots), \quad (17\,\mathrm{b})$$

und wir erhalten sowohl aus (17 a) zu jedem Paare ν, ν', wie aus (17 b) zu jedem Paare ν', ν die folgenden R für das Molekül

$$R_1 = r + r', \quad R_2 = r + r' - 1, \ldots \text{ bis } R = |r - r'|. \quad (19)$$

Das Molekül hat also alle in (20) aufgeschriebenen λ_+ und ebendiese λ_- mit allen in (19) gegebenen R.

C. Gleiche Atome in gleichen Zuständen. Die Eigenfunktionen sind

$$\Psi_{\nu\nu'\zeta\zeta'\alpha_1,\ldots,\alpha_m} = \psi_{\nu\zeta}(\mathfrak{r}_{\alpha_1} - \mathsf{Z}, \ldots)\, \psi_{\nu'\zeta'}(\mathfrak{r}_{\alpha_{m+1}} + \mathsf{Z}, \ldots). \quad (17\,\mathrm{c})$$

Zuerst können wir genau wie im Falle A verfahren und erhalten die λ-Werte

$$\left.\begin{array}{llllll} l+l', & l+l'-1, & \ldots & 1,0, & & 0' \\ & \ldots\ldots\ldots\ldots\ldots\ldots & & & & \\ & l, & & 1,0 & \text{oder* } & 0' \\ & \ldots\ldots\ldots\ldots\ldots\ldots & & & & \\ & |l-l'|, & \ldots & 1, & & \end{array}\right\} \quad (20)$$

jedes mit allen R-Werten

$$r + r', \quad r + r' - 1, \ldots, |r - r'| + 1, |r - r'|. \quad (19)$$

Ist $\nu \neq \nu'$, so haben wir für jedes Paar ν, ν' ohne Rücksicht auf die Reihenfolge einen $(\nu + \nu')_+$- und einen $(\nu + \nu')_-$-Term mit allen R von (19).

---

* Je nachdem $(-1)^{l+l'} w w'$ gleich $+1$ oder -1 ist.

Ist aber $\nu = \nu'$, so haben wir nur einen Term und müssen seinen Spiegelungscharakter bestimmen.

Diese hängt nicht von ν ab, so daß wir $\nu = 0$ setzen können. Dann ist auch $\lambda = 0$, wir haben die $2r + 1$-Werte R von (19), jedes einmal ($r = r'$)

$$2r, \quad 2r - 1, \ldots, 0. \tag{19'}$$

Um die Spiegelungscharaktere dieser $2r + 1$ Eigenfunktionen zu bestimmen, erinnern wir uns, daß wir in § 11, C die Spiegelungscharaktere der aus ihnen entstandenen Hyperfunktionen schon bestimmt haben. Wir können die dortigen Ergebnisse für $j = j' = r$ anwenden und erhalten folgende Tabelle:

Spiegelungscharaktere der Terme.

$2r$	$(-1)^m$			
$2r-1$	$(-1)^m$	$-(-1)^m$		
$2r-2$	$(-1)^m$	$-(-1)^m$	$(-1)^m$	
. . .	. . .	. . .	. . .	. . .
0	$(-1)^m$	$-(-1)^m$	$(-1)^m$	 1
$R =$	$2r$	$2r-1$	$2r-2.$	 0

Das Paar $\mu = \mu' = r$ ergibt nämlich einen Term mit dem Charakter $(-1)^m$, das Paar $\mu, \mu' = r, r-1$ zwei Terme, einen mit positivem, einen mit negativem Charakter: das Paar $\mu, \mu' = r, r-2$ ebenfalls, das Paar $\mu = \mu' = r - 1$ einen Term mit dem Charakter $(-1)^m$ usw. Die Terme, die in einer vertikalen Kolonne stehen, müssen wir zu je einem $\lambda = 0$-Term zusammenfassen. Die erste Kolonne gibt einen Term mit $R = 2r$, die zweite einen mit $R = 2r - 1$ usw.

Gegenüber der Spiegelung im Ursprung sind die Eigenfunktionen positiv oder negativ, je nachdem R gerade oder ungerade ist.

Die Terme mit $\nu = \nu'$ von (19') sind also λ_+- bzw. λ_--Terme, je nachdem $R + \lambda = R + 2\nu$ gerade oder ungerade ist.

Bei gleichen Atomen in gleichen Zuständen lautet also die Aufbauregel, wenn die Azimutalquantenzahl der Atome l, ihr Multiplettsystem r ist: Man hat je ein λ von (20')

$$\left.\begin{array}{llllll}
2l, & 2l-1, & \ldots & & 1, & 0, \quad 0' \\
 & 2l-1, & \ldots & & 1, & 0, \quad 0' \\
 & \ldots & \ldots & \ldots & \ldots & \ldots \\
 & & l, & l-1, & 1, & 0, \\
 & & & l-1, & 1, & \\
 & & & \ldots & \ldots & \ldots \\
 & & & & 1, &
\end{array}\right\} \tag{20'}$$

mit allen R von (19'). Die ungeraden λ kommen in (20') geradzahlig oft vor, es sind unter ihnen gleich viele λ_+ und λ_-. Die geradzahligen λ kommen ungeradzahlig oft vor, man hat ein λ_+ mehr oder weniger als λ_-, je nachdem R gerade oder ungerade ist. Im ersten Fall (R gerade) hat man auch lauter 0_+ und $0'_-$, im zweiten (R ungerade) 0_- und $0'_+$. Man kann dies zum Ausdruck bringen, indem man (20') schreibt:

$$\left.\begin{array}{llllll}
(2l)_\pm, & (2l-1)_\pm, & (2l-2)_\pm, \ldots & 2_\pm, & 1_\pm; & 0_\pm,\ 0'_\mp \\
 & (2l-1)_\mp, & (2l-2)_\mp, \ldots & 2_\mp, & 1_\mp; & 0_\pm,\ 0'_\mp \\
 & & (2l-2)_\pm, \ldots & 2_\pm, & 1_\pm; & 0_\pm,\ 0'_\mp \\
 & & \ldots\ldots\ldots\ldots & & & \\
 & & & 2_\mp, & 1_\mp, & \\
 & & & & 1_\pm, &
\end{array}\right\} \quad (20'a)$$

worin das obere oder untere Vorzeichen gilt, je nachdem man aus (19') ein gerades oder ungerades R wählt. Mit anderen Worten, wir haben zu jedem R von (19'): $0, 1, 2, \ldots 2r$ so viele Terme mit einem bestimmten λ, als dieses λ in (20') vorkommt. Ist diese Zahl eine gerade Zahl, so haben wir gleich viele λ_+- und λ_--Terme. Ist die Anzahl der Ziffern λ in (20') dagegen eine ungerade (wie dies z. B. für die Ziffer $\lambda = 2l$ immer zutrifft), so haben wir um eins mehr λ-Terme mit dem Index $(-1)^R$, als mit dem entgegengesetzten Index $-(-1)^R$. Alle 0-Terme haben den Index $(-1)^R$, alle 0'-Terme den Index $-(-1)^R$.

§ 13. Hiermit haben wir für die drei Kopplungsverhältnisse a) Multiplettaufspaltung groß gegen Rotationsaufspaltung, chemische Bindungsenergie gegen beide; b) Rotationsaufspaltung groß gegen Multiplettaufspaltung, chemische Bindungsenergie groß gegen beide; c) Multiplettaufspaltung groß gegen chemische Bindungsenergie, diese groß gegen die Rotationsaufspaltung, die möglichen vorkommenden Terme bestimmt.

Für die beiden ersten Fälle haben wir die vorkommenden λ- und zugehörigen r-Werte bestimmt. Der Fall c) ist in § 10 erledigt, bei ihm sind die ι-Werte so zu erhalten, daß man j und j' in bezug auf die Kernverbindungslinie quantelt. Natürlich werden sehr oft noch ganz andere Koppelungsverhältnisse vorkommen als die hier betrachteten drei Fälle, es ist aber anzunehmen, daß diese besonders häufig sein werden.

§ 14. Wir wollen noch die Resultate zusammenfassen. Wir gehen von zwei Atomen mit m' bzw. m Elektronen aus; die azimutalen Quantenzahlen sind l und l', die Multiplettsysteme r und r', die Spiegelungs-

charaktere w und w'. Bei ungleichen Atomen erhalten wir die λ von (20) für das Molekül, jedes mit jedem R-Wert, d. h. Multiplettsystem von (19). Bei gleichen Atomen in verschiedenen Zuständen haben wir doppelt soviel Terme, alle λ von (20) kommen sowohl als λ_+ wie auch als λ_- vor. Sind die Atome in gleichen Zuständen, so haben wir zu jeder Kombination eines λ aus (20′) mit einem R aus (19′) nur einen Term; welche λ_+- und welche λ_--Terme sind, ist aus (20′a) zu entnehmen. Dies sind die gruppentheoretisch möglichen Terme.

Wir müssen noch angeben, was für Molekülterme einem bestimmten Atomterm, der zu dem Atom mit den vereinigten Kernen gehört, zugeordnet werden müssen*. Ein Atomterm mit der Azimutalquantenzahl l gibt je einen Term mit $\lambda = l, l-1, \ldots, 1, 0$. Und zwar ist der letzte Term ein 0-Term, wenn l gerade und der Atomterm positiv, oder wenn l ungerade und der Atomterm negativ ist, sonst haben wir einen 0′-Term. Erfolgt die Trennung in zwei gleiche Kerne, so haben wir 0_+-, 1_--, 2_+-... usw. Terme, wenn der Atomterm positiv, dagegen 0_--, 1_+-, 2_--, ... Terme, wenn der Atomterm negativ war. Hiernach läßt sich nunmehr die im § 10 beschriebene Zuordnung der Terme der getrennten Atome zu den Termen des Moleküls und das Aufsuchen der Atomterme des Atoms mit vereinigten Kernen, die für den betreffenden Molekelterm in Frage kommen, eindeutig ausführen.

§ 15. Als Beispiel wollen wir das Spektrum von H_2 behandeln**. Als Zustände mit weit getrennten H-Atomen seien folgende betrachtet:

Atomzustände	Ionisierungsfrequenz	Gruppentheoretisch mögliche Terme	
$1s\,^2S^+ - 1s\,^2S^+$	$109\,700\,.\,\mathrm{cm}^{-1}$	$^1\Sigma_+$ (1)	$^3\Sigma_-$ (2)
$1s\,^2S^+ - 2s\,^2S^+$	27 400	$^1\Sigma_+$ (3)	$^3\Sigma_-$ (4)
		$^1\Sigma_-$ (5)	$^3\Sigma_+$ (6)
$1s\,^2S^+ - 2p\,^2P^-$	27 400	$^1\Sigma_+$ (7)	$^3\Sigma_-$ (8)
		$^1\Sigma_-$ (9)	$^3\Sigma_+$ (10)
		$^1\Pi_+$ (11)	$^3\Pi_+$ (12)
		$^1\Pi_-$ (13)	$^3\Pi_-$ (14)

Von den He-Termen müssen wir für (1) und (2) diejenigen in Betracht ziehen, deren Ionisierungsfrequenz kleiner ist als 109 700, für

* F. Hund, ZS. f. Phys. (im Erscheinen).

** Literatur siehe bei R. T. Birge, Proc. Nat. Acad. **14**, 12, 1928. Die Deutung von T. Hori, ZS. f. Phys. **44**, 834, 1927, stimmt im wesentlichen mit der unseren überein, nur Annahme einer Aufspaltung für den oberen Zustand der Lymanbanden (5) scheint uns nicht zuzutreffen.

die übrigen diejenigen, deren Ionisierungsfrequenz kleiner oder ungefähr 27 400 ist. Diese sind in fogender Tabelle aufgetragen:

He-Atomzustand	Ionisierungs-frequenz	Molekülterme
$(1s)^2\,{}^1S^+$	198 400 . cm^{-1}	${}^1\Sigma_+$ (1)
$(1s)(2s)\,{}^1S^+$	32 000	${}^1\Sigma_+$ (3)
$(1s)(2s)\,{}^3S^+$	38 500	${}^3\Sigma_+$ (6)
$(1s)(2p)\,{}^1P^-$	27 200	${}^1\Sigma_-$ (5) ${}^1\Pi_+$ (11)
$(1s)(2p)\,{}^3P^-$	29 200	${}^3\Sigma_-$ (2) ${}^3\Pi_+$ (12)

Zunächst soll also, in Übereinstimmung mit der Heitler und Londonschen Theorie, nur der eine der beiden Molekülterme, die bei hohen Schwingungsquantenzahlen unangeregte Atome geben, zum diskreten Spektrum gehören (1), der zweite (2) dagegen nicht. Von den Zuständen, die bei hohen Schwingungsquantenzahlen ein unangeregtes Atom und ein angeregtes Atom mit der Gesamtquantenzahl 2 geben, sollen zwei, (5) und (11), mit dem Grundzustand kombinieren dürfen, alle anderen Kombinationen sind durch Auswahlverbote ausgeschlossen.

Die ersteren scheinen die Lyman-, die letzteren die Wernerbanden zu sein. Nach dieser Auffassung sind also die Lymanbanden ${}^1\Sigma_- \leftrightarrow {}^1\Sigma_+$ Übergänge, und die Wernerbanden ${}^1\Pi_+ \leftrightarrow {}^1\Sigma_+$ Übergänge.

Es liegt nicht in dem Rahmen dieser Arbeit, auf weitere Spektren näher einzugehen, wir wollten nur zeigen, wie man die hier erhaltenen Resultate zur Analyse verwenden kann. Wir möchten nur noch bemerken, daß wir diese Methode mit gutem Erfolg auf die Spektren von He_2 und Na_2 angewendet haben, während das Spektrum von O_2 — namentlich die A'-Teilbande der atmosphärischen Banden — Schwierigkeiten zu bereiten scheint.

§ 16. Vergleich der Mullikenschen Bezeichnungsweise mit der unseren.

Mulliken	Hier	Mulliken	Hier
j_k	l	σ	ι
j	j	S	Σ
σ_k	λ	P	Π
σ_s	η	D	Δ

Anmerkung bei der Korrektur. In einer inzwischen erschienenen Arbeit* behandelt R. de L. Kronig die Molekülspektren

* ZS. f. Phys. 50, 347, 1928.

noch einmal und kommt zu Resultaten, die mit unseren § 1 bis 6 äquivalent sind.

In einer demnächst in dieser Zeitschrift erscheinenden Arbeit leitet F. Hund ein Aufbauprinzip ab. Seine Resultate stehen in keinem Widerspruch mit den unseren, es ist im Gegenteil zu hoffen, daß sich die beiden Methoden bei der Abschätzung von Molekeltermen ergänzen werden.

Bei Hund bewegen sich die Elektronen im ungestörten Falle nur unter der Wirkung der festgehaltenen Kerne, die Wechselwirkung der Elektronen ist entkoppelt.

Hierdurch hat F. Hund den Vorteil, daß er sich von der Kenntnis der Atomterme weitgehend unabhängig machen kann. Die von uns verwendete Methode hat dagegen den Vorteil, daß dem Auseinanderziehen der Kerne ein Vorgang entspricht, dessen Verlauf an der Bandenkonvergenz direkt abgelesen werden kann, so daß die Zuordnung der Molekülterme zu Termen der getrennten Atome durch die Spektren selbst eindeutig gegeben ist. Der Entkopplung der Wechselwirkung der Elektronen entspricht kein ähnlicher Prozeß und die Zuordnung der Terme zu den „ungestörten Termen“ kann daher vielfach willkürlich geschehen.

Göttingen, Juli 1928.

Über die elastischen Eigenschwingungen symmetrischer Systeme

E. P. Wigner

Nachrichten der Gesellschaft der Wissenschaften zu Göttingen
Mathematisch-Physikalische Klasse *1930*, 133–146

Vorgelegt von M. Born in der Sitzung am 23. Mai 1930.

1. Bekanntlich kann man in der Quantenmechanik die Symmetrieeigenschaften eines Systems zur Bestimmung seiner Bewegungstypen nutzbar machen. Durch die Angabe „die Energie[1]) ist E" ist eine endliche Mannigfaltigkeit von Zuständen in drehungsinvarianter Weise charakterisiert. So etwas ist in der klassischen Mechanik im allgemeinen nicht möglich, weil die Anfangsbedingungen eine zu große Rolle spielen.

Eine Ausnahme hiervon bilden die elastischen Schwingungen eines Punktsystems um seine Gleichgewichtskonfiguration. Die Bewegungsgleichungen sind in diesem Falle linear (d. h. die Superposition zweier möglicher Schwingungen ist wieder eine mögliche Schwingung), sodaß die Verhältnisse ganz analog zu denen der Quantenmechanik sind. Es ist daher zu erwarten, daß man mit denselben Mitteln, die man dort gebrauchen kann, auch hier zum Ziele kommt.

Die Frage der elastischen Schwingungen symmetrischer Gebilde, die wir hier untersuchen wollen, bildet auch den Gegenstand einer älteren Arbeit von Brester[2]), der sie mit ganz elementaren Hilfsmitteln vollständig gelöst hat. Wenn hier diese Aufgabe noch einmal behandelt wird, so sei das damit entschuldigt, daß die eingangs erwähnten gruppentheoretischen Mittel ihr ganz be-

1) Wenn E eine Stelle des diskreten Spektrums mit nur endlich vielen Eigenfunktionen ist.

2) C. J. Brester, Diss. Utrecht 1923, angefertigt im Göttinger Institut für theor. Physik. S. 8—90. Einen speziellen Fall behandelt auch D. M. Dennison, Astroph. Journ. 62. 87. 1925, den der Schwingungen eines CH_4-Moleküls, den wir als Beispiel gebrauchen werden.

sonders angepaßt erscheinen und daß man mit ihrer Hilfe auch zu einer besseren Übersicht der BRESTER'schen Theorie gelangt.

Wir betrachten im folgenden die Schwingungen eines Systems von n Punkten, die durch elastische Kräfte aneinandergebunden sind[3]). Der Schwerpunkt des Systems befinde sich im Nullpunkt des Koordinatensystems; in der Ruhelage seien die Koordinaten der Atome die Komponenten der Vektoren[4]) $\mathfrak{r}_1, \mathfrak{r}_2, \ldots, \mathfrak{r}_n$, während die Verrückungen aus den Gleichgewichtslagen durch die Vektoren $\mathfrak{f}_1, \mathfrak{f}_2, \ldots, \mathfrak{f}_n$ gegeben sind, die Koordinaten der Teilchen die Komponenten von $\mathfrak{r}_1 + \mathfrak{f}_1, \mathfrak{r}_2 + \mathfrak{f}_2, \ldots, \mathfrak{r}_n + \mathfrak{f}_n$ sind. Die $\mathfrak{f}_k$ hängen noch von der Zeit ab, aber sie sind immer klein gegen die $\mathfrak{r}_k$, für die Eigenschwingungen ist die Zeitabhängigkeit der $\mathfrak{f}_k$ durch einen Faktor $\sin r(t - t_0)$ gegeben. Die n Vektoren $\mathfrak{f}_1, \mathfrak{f}_2, \ldots, \mathfrak{f}_n$, einer Verrückung kann man auch zu einem $3n$ dimensionalen Vektor $\mathfrak{f}$ zusammenfassen.

Das ganze System soll eine Symmetrie haben, d. h. es soll eine Gruppe $\mathfrak{G}$ von dreidimensionalen Drehungen[5]) $(R_{\alpha\beta})$ existieren, die die Gleichgewichtskonfiguration in sich überführt. Es gilt also für jede Transformation R der Gruppe und alle k und α.

$$(1) \qquad \sum_{\beta=1}^{3} R_{\alpha\beta}\mathfrak{r}_{k\beta} = \mathfrak{r}_{l\alpha} \text{ oder einfach } R\mathfrak{r}_k = \mathfrak{r}_l,$$

wobei noch das $k$te und das $l$te Teilchen gleichartig (z. B. beide Wasserstoffatome) sein müssen[6]). Die Nummer l des Teilchens, auf dessen Platz das $k$te Teilchen durch R gerückt wird, nennen wir $R(k)$, dann lautet (1)

$$(1\text{a}) \qquad \sum_{\beta=1}^{3} R_{\alpha\beta}\mathfrak{r}_{k\beta} = \mathfrak{r}_{R(k),\alpha}; \quad R\mathfrak{r}_k = \mathfrak{r}_{R(k)}.$$

Eine Eigenschwingung des Systems kann durch die n Vektoren $\mathfrak{f}_1, \mathfrak{f}_2, \ldots, \mathfrak{f}_n$ beschrieben werden, die die Verrückungen der Atome

3) Man denke etwa an ein CH_4-Molekül, wobei die Atome durch Punkte idealisiert sind.

4) Die Vektoren $\mathfrak{r}$ sind hierdurch noch nicht eindeutig bestimmt, weil eine Drehung des ganzen Systems noch frei ist. Im folgenden denken wir sie irgendwie willkürlich festgelegt (etwa: $\mathfrak{r}_1$ liege in der Z-Achse, $\mathfrak{r}_2$ in der ZY-Ebene).

5) Die Indizes α, β beziehen sich immer auf die Koordinatenachsen $X_1 = X$, $X_2 = Y$, $X_3 = Z$ und werden zur Kennzeichnung der Komponenten eines Vektors als untere Indizes angefügt.

6) Man bezeichnet in der Kristallographie Punkte, die durch Symmetrieelemente ineinander übergeführt werden, als „gleichwertig". Gleichwertige Punkte sind immer auch gleichartig, aber nicht notwendig umgekehrt.

aus den Gleichgewichtslagen angeben, wenn die Amplitude der Schwingung eben im Maximum ist. Sind $\mathfrak{f}_1, \mathfrak{f}_2, \ldots, \mathfrak{f}_n$ Amplituden einer Eigenschwingung, und unterwirft man die durch sie gegebene Konfiguration des Systems einer Transformation R der Gruppe $\mathfrak{G}$, so erhält man wieder die Konfiguration einer Eigenschwingung derselben Frequenz, weil man ja die relative Konfiguration der Teilchen gar nicht geändert hat. Die Koordinate des $k$ten Punktes ist jetzt $R\mathfrak{r}_k + R\mathfrak{f}_k = \mathfrak{r}_{R(k)} + R\mathfrak{f}_k$, auch jetzt ist in der Nähe einer jeden Gleichgewichtslage ein Punkt.

Man kann daher eine mit dieser Konfiguration identische auch durch kleine Verrückungen der Punkte aus ihren Gleichgewichtslagen erzeugen: man muß hierzu das $l = R(k)$te Teilchen um $R\mathfrak{f}_k = R\mathfrak{f}_{R^{-1}(l)}$ verschieben. Die Verrückungen

$$R\mathfrak{f}_{R^{-1}(1)},\ R\mathfrak{f}_{R^{-1}(2)},\ \ldots,\ R\mathfrak{f}_{R^{-1}(n)}$$

$$\bar{R}\mathfrak{f}_l = R\mathfrak{f}_{R^{-1}(l)};\quad \bar{R}\mathfrak{f}_{l\alpha} = \sum_\beta R_{\alpha\beta}\mathfrak{f}_{R^{-1}(l),\beta} \tag{2}$$

bilden das Amplitudensystem $\bar{R}\mathfrak{f}$ einer Eigenschwingung, die die gleiche Frequenz, wie die Eigenschwingung $\mathfrak{f}$ hat. Die Operationen $\bar{R}$ ersetzen die „Drehungen" und „Vertauschungen von Elektonen" der Quantenmechanik.

Bezeichnet man die linear unabhängigen Eigenschwingungen einer bestimmten Frequenz mit $\mathfrak{f}^{(1)}, \mathfrak{f}^{(2)}, \ldots, \mathfrak{f}^{(f)}$, so müssen die $\bar{R}\mathfrak{f}^{(\varkappa)}$ durch diese linear ausdrückbar sein

$$\bar{R}\mathfrak{f}^{(\varkappa)} = \sum_{\lambda=1}^{f} D(R)_{\lambda\varkappa}\mathfrak{f}^{(\lambda)} \tag{3}$$

und man schließt ebenso wie in der Quantenmechanik durch Anwendung einer weiteren Operation $\bar{S}$, die einem Element S der Gruppe $\mathfrak{G}$ zugeordnet ist,

$$\bar{S}\bar{R}\mathfrak{f}^{(\varkappa)} = \sum_{\lambda=1}^{f} D(R)_{\lambda\varkappa}\bar{S}\mathfrak{f}^{(\lambda)} = \sum_{\lambda=1}^{f}\sum_{\mu=1}^{f} D(R)_{\lambda\varkappa}D(S)_{\mu\lambda}\mathfrak{f}^{(\mu)}$$
$$= \sum_{\mu=1}^{f} D(SR)_{\mu\varkappa}\mathfrak{f}^{(\mu)}$$

daß die f-dimensionalen Matrizen $(D(R)_{\lambda\varkappa})$ eine Darstellung der Gruppe $\mathfrak{G}$ bilden. Es wird auf diese Weise jeder Eigenfrequenz eine Darstellung zugeordnet. Man schließt weiter ebenso, wie in der Quantenmechanik, daß diese Darstellungen als irreduzible angenommen werden können. Es gibt also so viele Typen

von Eigenschwingungen, als[7]) $\mathfrak{G}$ irreduzible Darstellungen d. h. Klassen hat, die Anzahl der linear unabhängigen Eigenschwingungen einer Frequenz ist durch die Dimension der zugehörigen Darstellungen gegeben.

2. Es interessiert noch die Frage, zu wie vielen Frequenzen eine vorgegebene Darstellung gehört, wie viele Eigenfrequenzen eines bestimmten Typs existieren. Die analoge Frage konnte in der Quantenmechanik nicht auftreten, weil diese Zahlen dort alle unendlich waren.

Es läßt sich bekanntlich jede Bewegung des Systems, bei der die Gleichgewichtslagen nicht verändert werden, aus Eigenschwingungen zusammensetzen. Zur Erzeugung jeder Bewegung muß man noch die drei Parallelverschiebungen in Richtung der drei Koordinatenachsen und die Drehungen um die drei Achsen hinzunehmen (es wird angenommen, daß die Punkte nicht alle auf einer Geraden liegen). Diese Verrückungen können auch als Eigenschwingungen aufgefaßt werden, nur sind ihre Frequenzen Null. Die zugehörige Darstellung ist im ersten Fall $D^{(v)}(R)$, die Darstellung polarer, im zweiten Fall $D^{(v')}(R)$, die Darstellung axialer Vektoren.

Bezeichnen wir das Amplitudensystem der Verrückung, bei der nur das $k$te Teilchen seinen Platz und dieses nur in der X_α-Richtung um die Einheit verändert hat, mit $\mathfrak{e}^{(k\alpha)}$ (d. h. es sei $\mathfrak{e}_{l\beta}^{(k\alpha)} = \delta^{kl}\delta^{\alpha\beta}$). Man kann durch die $\mathfrak{e}^{(k\alpha)}$ alle Verrückungen linear ausdrücken:

$$\mathfrak{f}^{(\varkappa)} = \sum_{k=1}^{n} \sum_{\alpha=1}^{3} \mathfrak{f}_{k\alpha}^{(\varkappa)} \mathfrak{e}^{(k\alpha)}. \tag{4}$$

Umgekehrt kann man auch durch die $\mathfrak{f}$ (wenn man auch die Verrückungsvektoren der Translation und Rotation hinzunimmt) die $\mathfrak{e}$ ausdrücken (ihre Zahl ist also $3n$), die Koeffizienten seien[8])

7) Eigentlich kommen nur die in reellen irreduziblen Darstellungen in Frage und die Eigenfrequenzen zu denen eine komplexe Darstellung gehört, fallen immer mit einer Eigenfrequenz zusammen, zu der die konjugiert komplexe Darstellung gehört. Hierdurch kommen bei den Gruppen, bei denen die im reellen irreduziblen Darstellungen nicht auch im komplexen irreduzibel sind, sogenannte zufällige Entartungen zu Stande, die jedoch leicht zu übersehen und für das folgende belanglos sind.

8) Wegen der Orthogonalität der $\mathfrak{f}$ ist in Wirklichkeit, wenn man die $\mathfrak{f}$ normiert annimmt $\mathfrak{f}_{k\alpha}^{(\varkappa)} = s_{k\alpha;\varkappa}$.

$$\mathfrak{e}^{(k\alpha)} = \sum_{\substack{\text{Alle Eigen-}\\ \text{schwingungen}}}{}^{\varkappa} s_{k\alpha;\varkappa}\mathfrak{f}^{(\varkappa)}. \tag{4a}$$

Denken wir uns jetzt für einen Augenblick, daß wir die Amplitudensysteme aller $3n$ Eigenschwingungen kennen. Für die zur gleichen Frequenz gehörigen gilt (3), für das ganze System also

$$\bar{R}\mathfrak{f}^{(\varkappa)} = \sum_{\lambda} \Delta(R)_{\lambda\varkappa}\mathfrak{f}^{(\lambda)}, \tag{5}$$

wo $\Delta(R)$ eine Darstellung der Gestalt

$$\Delta(R) = \begin{pmatrix} D^{(1)}(R) & 0 & \dots \\ 0 & D^{(2)}(R) & \dots \\ \dots & \dots & \dots \end{pmatrix} \tag{5a}$$

ist. In (5a) sind die $D^{(1)}(R)$, $D^{(2)}(R)$, ... die zu den einzelnen Frequenzen im Sinne von (3) zugeordneten Darstellungen. Wenn wir also wissen wollen, wie viele Frequenzen eine gewisse Darstellung gemeinsam haben, so müssen wir bestimmen, wie oft diese Darstellung in $\Delta(R)$ vorkommt. Hierzu genügt es bekanntlich, den Charakter von $\Delta(R)$, die Summe $\sum_{\varkappa}\Delta(R)_{\varkappa\varkappa}$, für alle R zu kennen. Diesen können wir folgendermaßen bestimmen.

Wenn wir in (5) mit Hilfe von (4), (4a) an Stelle der $\mathfrak{f}$ die $\mathfrak{e}$ einführen und $R\mathfrak{e}^{(k\alpha)}$ durch die $\mathfrak{e}^{(j\beta)}$ linear ausdrücken

$$\bar{R}\mathfrak{e}^{(k\alpha)} = \sum_{j=1}^{n}\sum_{\beta=1}^{3} \bar{\Delta}(R)_{j\beta;k\alpha}\mathfrak{e}^{(j\beta)}, \tag{6}$$

so bilden die Koeffizienten wiederum eine Darstellung $\bar{\Delta}(R)$ der Gruppe $\mathfrak{G}$, die zu $\Delta(R)$ äquivalent ist und aus ihr durch Ähnlichkeitstransformation mit der $3n$-dimensionalen Matrix $s_{k\alpha;\varkappa}$ hervorgeht. Ihr Charakter $\chi(R) = \sum_{j\beta}\bar{\Delta}(R)_{j\beta;j\beta}$ ist also gleich dem von $\Delta(R)$ und es genügt den ersteren zu bestimmen.

Nun ist $\mathfrak{e}^{(k\alpha)}$ die Verrückung, die das $k$te Teilchen in der X_α-Richtung um 1 verschiebt, die anderen in der Gleichgewichtslage läßt, $\bar{R}\mathfrak{e}^{(k\alpha)}$ entsteht aus dieser Verrückung, indem man das ganze Punktsystem um R dreht. Dann ist jede Gleichgewichtslage besetzt, nur das $R(k)$te, wo das $k$te Teilchen hinkommt, nicht: dieses hat nach (2) die Verrückungskomponenten $R_{1\alpha}$, $R_{2\alpha}$, $R_{3\alpha}$. Es ist also

$$\bar{R}\mathfrak{e}^{(k\alpha)} = \sum_{\beta} R_{\alpha\beta}\mathfrak{e}^{(R^{-1}(k),\beta)} = \sum_{j\beta}\bar{\Delta}(R)_{j\beta;k\alpha}\mathfrak{e}^{(j\beta)}. \tag{7}$$

Gleichung (7) bestimmt die Darstellung $\bar{\Delta}(R)$. Um noch ihren Charakter zu berechnen, müssen wir die Summe der Dia-

gonalelemente bilden. Nun steht in der $k\alpha$ Spalte in der Hauptdiagonale sicher Null, wenn $k \neq R^{-1}(k)$ ist, in diesem Fall braucht man ja $e^{(k\alpha)}$ zum Ausdrücken von $\bar{R}e^{(k\alpha)}$ garnicht. Ist dagegen $k = R^{-1}(k)$, so steht in $\Delta(R)$ in der k Spalte in der Hauptdiagonale $R_{\alpha\alpha}$. Es ist daher

$$\bar{\Delta}(R)_{k\alpha;\,k\alpha} = \begin{matrix} 0 & \text{für } R(k) \neq k \\ R_{\alpha\alpha} & \text{für } R(k) = k \end{matrix} \tag{8}$$

und daher

$$\sum_{\alpha} \bar{\Delta}(R)_{k\alpha;\,k\alpha} = \begin{matrix} 0 & & \text{für } R(k) \neq k \\ \sum\limits_{\alpha} R_{\alpha\alpha} = \text{Spur}\,(R) & & \text{für } R(k) = k. \end{matrix} \tag{8a}$$

Bezeichnen wir noch die Anzahl der Gleichgewichtslagen, die R unverändert läßt mit u_R, so wird

$$\chi(R) = \sum_{k}\sum_{\alpha} \bar{\Delta}(R)_{k\alpha;\,k\alpha} = u_R \,\text{Spur}\,(R) = \pm\, u_R(1 + 2\cos\varphi_R), \tag{9}$$

wo Spur $(R) = \pm(1 + 2\cos\varphi_R)$ eingesetzt ist: φ_R ist der Drehwinkel von R und es gilt das obere oder das untere Vorzeichen, je nachdem R eine reine Drehung oder eine Drehspiegelung ist. Aus (9) kann man den Charakter von $\chi(R)$ berechnen, es genügt natürlich, dies für je ein Element jeder Klasse zu tun.

Schreibt man $\chi(R)$ als Linearkombination der Charaktere $\chi^{(1)}(R)$, $\chi^{(2)}(R)$, ... der verschiedenen irreduziblen Darstellungen von $\mathfrak{G}$

$$\chi(R) = a_1'\chi^{(1)}(R) + a_2'\chi^{(2)}(R) + \cdots. \tag{10}$$

so geben die Koeffizienten $a_1', a_2', \ldots$ an (sie sind nichtnegative ganze Zahlen), wie viele Eigenfrequenzen des Problems zu den verschiedenen Darstellungen $D^{(1)}(R)$, $D^{(2)}(R)$, ... gehören. Die Zahlen $a_1', a_2', \ldots$ kann man auch mit Hilfe der expliziten Formeln

$$a_p' = \frac{1}{h}\sum_{R} \chi(R)\chi^{(p)}(R) \tag{11}$$

berechnen, wo die Summation über alle h Gruppenelemente von $\mathfrak{G}$ zu erstrecken ist. Es ist indessen zu beachten, daß man auf diese Weise auch die beiden entarteten Schwingungen, die Translation und Drehungen mitbekommt. Will man diese nicht haben, so muß man aus dem nach (9) bestimmtem $\chi(R)$ die Charaktere dieser abziehen. Für die Translationen ist dies der Charakter von $D^{(v)}(R)$, d. i. $\pm(1 + 2\cos\varphi_R)$, für die Drehungen der von $D^{(v')}(R)$, d. i. $+(1 + 2\cos\varphi_R)$. Im ganzen muß man also aus dem Charakter der reinen Drehungen $2(1 + 2\cos\varphi_R)$ abziehen, den der Dreispiegelungen

unverändert lassen. Für die Anzahlen a der nichtentarteten Eigenfrequenzen der verschiedenen Typen erhält man

$$(10\text{a})\quad a_1\chi^{(1)}(R)+a_2\chi^{(2)}(R)+\cdots=\Xi(R)=\begin{cases}(u_R-2)(1+2\cos\varphi_R) & R \text{ reine Drehung}\\ -u_R(1+2\cos\varphi_R) & R \text{ Drehspiegelung}\end{cases}$$

oder die explizite Formel

$$(11\text{a})\quad a_p=\frac{1}{h}\sum_R\chi^{(p)}(R)\,\Xi(R)=\sum{}'(u_R-2)(1+2\cos\varphi_R)-\sum{}''u_R(1+2\cos\varphi_R)$$

wo die erste Summation über die reinen Drehungen, die zweite über die Drehspiegelungen von $\mathfrak{G}$ zu erstrecken ist. Dabei ist u_R die Anzahl der Teilchen, die durch R nicht verrückt werden, $\chi^{(p)}(R)$ der Charakter der irreduziblen Darstellung des betreffenden Typus. Die Anzahl der linear unabhängigen Eigenschwingungen, die zu einer Eigenschwingung eines bestimmten Typs gehören, ist die Dimension der zugehörigen Darstellung.

Die Charaktere der irreduziblen Darstellungen der meisten Symmetriegruppen hat H. Bethe[9]) bestimmt; da sie für verschiedene Probleme der Quantenmechanik wichtig sind, sind sie am Schlusse der Arbeit alle zusammengestellt.

Die Symmetriegruppe des Methans ist die Tetraedergruppe T_d, sie besteht aus der Einheit E, vier dreizähligen Drehachsen C_3 (acht Gruppenelemente), drei zweizähligen C_2 (drei Elemente), sechs Spiegelebenen σ_d (sechs Elemente) und drei vierzähligen Drehspiegelachsen S_4 (sechs Elemente). Die Charaktere der fünf irreduziblen Darstellungen sind (die Gruppe ist der symmetrischen vierten Grades holomorph) in der Tab. 1 zusammengestellt.

Tab. 1.

	$E(1)$	$C_3(8)$	$C_2(3)$	$\sigma_d(6)$	$S_4(6)$
$\chi^{(1)}$	1	1	1	1	1
$\chi^{(2)}$	2	−1	2	0	0
$\chi^{(3)}$	3	0	−1	1	−1
$\chi^{(4)}$	3	0	−1	−1	1
$\chi^{(5)}$	1	1	1	−1	−1

Um den Charakter von $\Delta(R)$ zu bestimmen, stellen wir in einer zweiten Tabelle die Anzahlen u_R der unverändert gelassenen Atome, die Drehwinkel φ_R, die Größen $\pm(1+2\cos\varphi_R)$, die Charaktere $\chi(R)$, die die entarteten Schwingungen noch mitenthalten und schließlich den Charakter $\Xi(R)$, wo diese schon abgezogen sind, zusammen.

9) Ann. d. Phys. (5) 3, 133, 1929.

Tab. 2.

	E	$C_3(8)$	$C_2(3)$	$\sigma_d(6)$	$S_4(6)$
u_R	5	2	1	3	1
φ_R	0	$2\pi/3$	π	π	$\pi/4$
$\pm(1+2\cos\varphi_R)$	3	0	-1	1	-1
$\chi(R)$	15	0	-1	3	-1
$\Xi(R)$	9	0	1	3	-1

Mit Hilfe von (10a), oder auch direkt nach (11a) ergibt sich

$$\Xi(R) = \chi^{(1)}(R) + \chi^{(2)}(R) + 2\chi^{(3)}(R).$$

Die Eigenschwingungen des Methans haben (außer den Nullfrequenzen) vier verschiedene Frequenzen, zu einer gehört eine, zu einer anderen zwei und zu den beiden letzten je drei Eigenschwingungen.

Den letzten Teil der Rechnung kann man sich ziemlich vereinfachen, indem man die Tabelle 2 nicht für das ganze System, auf einmal, sondern für alle gleichwertige Punkte separat aufgestellt (also separat für das C Atom und dann separat für die vier H Atome). Dann muß man in allen Tabellen bei den $\chi(R)$ stehen bleiben und nur in einer die Charaktere der Translation und Drehung abziehen, d. h. zu den $\Xi(R)$ übergehen, diese Tabelle muß allerdings, damit diese Abziehung möglich sei, Punkte enthalten, die nicht alle auf einer Geraden liegen, während bisher dies nur von der Gesamtheit aller Punkte verlangt wurde. Die Vereinfachung liegt darin, daß man die einzelnen $\chi(R)$ bzw. das $\Xi(R)$ einfacher in eine Summe irreduzibler Charaktere zerlegen kann, als die in der Tabelle 2 auftretende Summe dieser. Das gesamte a_p erhält man in diesem Fall durch Addition der aus den einzelnen Tabellen gewonnenen a_p.

3. Es ist noch leicht, die „aktiven" Eigenschwingungen zu bestimmen, deren Anregung mit einer Dipolstrahlung verbunden ist. Bezeichnen wir den Gesamtpolarisationsvektor mit $\mathfrak{S}$

$$\mathfrak{S}_\alpha^{(\varkappa)} = e_1 \mathfrak{f}_{1\alpha}^{(\varkappa)} + e_2 \mathfrak{f}_{2\alpha}^{(\varkappa)} + \cdots + e_n \mathfrak{f}_{n\alpha}^{(\varkappa)}, \tag{12}$$

so berechnet sich aus (3)

$$\bar{R}\mathfrak{S}_\alpha^{(\varkappa)} = \sum_{\lambda=1}^{f} D(R)_{\lambda\varkappa}(e_1 \mathfrak{f}_{1\alpha}^{(\lambda)} + e_2 \mathfrak{f}_{2\alpha}^{(\lambda)} + \cdots + e_n \mathfrak{f}_{n\alpha}^{(\lambda)}) = \sum_{\lambda=1}^{f} D(R)_{\lambda\varkappa} \mathfrak{S}_\alpha^{(\lambda)}. \tag{13}$$

Andererseits transformiert sich $\mathfrak{S}$ offenbar wie ein Vektor

$$\bar{R}\mathfrak{S}_\alpha^{(\varkappa)} = \sum_{\beta=1}^{3} R_{\alpha\beta} \mathfrak{S}_\beta^{(\varkappa)}. \tag{13a}$$

Es ergibt sich

$$(14)\quad \sum_{\beta=1}^{3} R_{\alpha\beta}\mathfrak{S}_\beta^{(\varkappa)} = \sum_{\lambda=1}^{f} D(R)_{\lambda\varkappa}\mathfrak{S}_\alpha^{(\lambda)};\quad \mathfrak{S}_\beta^{(\lambda)} = \sum_\lambda \sum_\alpha R_{\alpha\beta} D(R)_{\lambda\varkappa}\mathfrak{S}_\alpha^{(\lambda)}.$$

Summiert man die rechte Seite über alle Gruppenelemente R von $\mathfrak{G}$, so ergibt sich Null, wenn die Matrizen $(R_{\alpha\beta})$ als Darstellung von $\mathfrak{G}$ die irreduzible Darstellung $D(R)$ der betreffenden Eigenfrequenz nicht enthält[10]). Die Bedingung hierfür ist, daß

$$(15)\quad \sum_{\substack{\text{reine}\\ \text{Drehungen}}}{}' (1+2\cos\varphi_R)\chi^{(p)}(R) - \sum_{\substack{\text{Dreh-}\\ \text{spiegelungen}}}{}'' (1+2\cos\varphi_R)\chi^{(p)}(R)$$

verschwinde. In diesem Fall ist die Schwingung inaktiv, sonst aktiv[11]). Im vorangehenden Beispiel sind nur die beiden dreifachen Eigenschwinkungen aktiv.

In ähnlicher Weise kann man auch über das Vorhandensein oder Verschwinden höherer Momente Rechenschaft geben.

Zur expliziten Bestimmung der Amplitudensysteme der einzelnen Eigenschwingungen geht man am besten aus einem Vergleich der Formeln (2) und (3) aus. Diese ergibt

$$(16)\quad \bar{R}\mathfrak{f}_{l\alpha}^{(\varkappa)} = \sum_\beta R_{\alpha\beta}\mathfrak{f}_{R^{-1}(l),\beta}^{(\varkappa)} = \sum_\lambda D(R)_{\lambda\varkappa}\mathfrak{f}_{l\alpha}^{(\lambda)},$$

woraus man

$$(17)\quad \mathfrak{f}_{R^{-1}(l),\beta}^{(\varkappa)} = \sum_{\beta=1}^{3}\sum_{\lambda=1}^{f} R_{\alpha\beta} D(R)_{\lambda\varkappa}\mathfrak{f}_{l\alpha}^{(\lambda)}$$

erhält. Dies ist eine Gleichung zwischen den Verrückungen lauter gleichwertiger Punkte, die in vielen Fällen die Eigenschwingungen schon ganz zu bestimmen gestattet. Sonst muß man noch die Bewegungsgleichungen zu Hilfe nehmen.

---

10) Die Höchstzahl der aktiven Eigenschwingungstypen ist also drei, diese müssen dann alle „unentartet" sein, d. h. zu einer Eigenfrequenz dieser Typen kann nur je eine Eigenschwingung gehören. Es können auch zwei Typen aktiv sein, zu den Eigenfrequenzen einer dieser Typen gehören dann je zwei Eigenschwingungen. Schließlich ist es möglich, daß nur ein dreifach entarteter Eigenschwingungstyp aktiv ist (wie im Falle der Tetraedersymmetrie). Dies folgt daraus, daß $(R_{\alpha\beta})$ entweder aus drei eindimensionalen, oder aus einer ein und einer zweidimensionalen oder schließlich aus einer dreidimensionalen irreduziblen Darstellung von $\mathfrak{G}$ bestehen kann.

11) Es ist zu beachten, daß die Tabellen 1 und 2 für je eine Klasse nur eine Spalte enthalten, deshalb muß man jedes Produkt so oft nehmen, als die betreffende Klasse Elemente hat. Z. B. ist $\sum\chi^{(4)}(R)(1+2\cos\varphi_R) = 1.3.3 + 8.0.0 + 3.-1.-1.+6.-1.1+6.1.-1 = 0$.

Um hierin an Stelle der Darstellungskoeffizienten $D(R)_{\lambda\varkappa}$ die Charaktere einzuführen, multipliziert man (16) mit dem Charakter $\chi^{(p')}(R)$ einer irreduziblen Darstellung von $\mathfrak{G}$ und summiert über alle Gruppenelemente. Man erhält mit Hilfe der Orthogonalitätsrelationen der Darstellungskoeffizienten

$$\sum_R \sum_\beta \chi^{(p')}(R)\, R_{\alpha\beta}\, \xi^{(\varkappa)}_{R^{-1}(l),\beta} \tag{18}$$

$$= \sum_R \sum_\lambda \chi^{(p')}(R)\, D^{(p)}(R)_{\lambda\varkappa}\, \xi^{(\lambda)}_{l\alpha} = \frac{h}{f_p}\, \delta_{p'p}\, \xi^{(\varkappa)}_{l\alpha},$$

also insbesondere Null, wenn $\chi^{(p')}(R)$ der Charakter einer von $D^{(p)}(R)$ verschiedenen Darstellung ist.

4. Die Eigenfrequenzen sind die Quadratwurzeln der Eigenwerte einer $3n$-dimensionalen symmetrischen Matrix $H_{k\beta;\,la}$, die zugehörigen Eigenvektoren bilden das Amplitudensystem der entsprechenden Eigenschwingungen. Bezeichnen wir die zu der Darstellung $D^{(p)}(R)$ gehörigen Eigenfrequenzen mit $\nu_{p1}, \nu_{p2}, \ldots, \nu_{pa}, \ldots$, die zu ν_{pa} gehörigen f_p Eigenschwingungen mit $\xi^{(pa1)}, \ldots, \xi^{(paf)}$, so gilt

$$\sum_{l\alpha} H_{k\beta;\,la}\, \xi^{(pa\varkappa)}_{l\alpha} = \nu^2_{pa}\, \xi^{(pa\varkappa)}_{k\beta} \tag{19}$$

und wenn man die $\xi^{(pa\varkappa)}$ normiert annimmt $\sum_{l\alpha} (\xi^{(pa\varkappa)}_{l\alpha})^2 = 1$

$$\sum_p \sum_a \sum_\varkappa \nu^2_{pa}\, \xi^{(pa\varkappa)}_{k\beta}\, \xi^{(pa\varkappa)}_{l\alpha} = H_{k\beta;\,la}. \tag{19a}$$

Wenn man hierin $R^{-1}(l)$ für k einsetzt, mit $R_{\alpha\beta}$ und dem Charakter $\chi^{(p')}(R)$ einer irreduziblen Darstellung multipliziert, über β und über alle Gruppenelemente summiert, erhält man

$$\sum_p \sum_a \sum_\varkappa \nu^2_{pa}\, \xi^{(pa\varkappa)}_{l\alpha} \sum_\beta \sum_R R_{\alpha\beta}\, \chi^{(p')}(R)\, \xi^{(pa\varkappa)}_{R^{-1}(l),\beta} \tag{20}$$

$$= \sum_\beta \sum_R R_{\alpha\beta}\, \chi^{(p')}(R)\, H_{R^{-1}(l)\beta;\,l\alpha}.$$

Dies ist nach (18)

$$\sum_p \sum_a \sum_\varkappa \nu^2_{pa}\, \xi^{(pa\varkappa)}_{l\alpha} \frac{h}{f_p}\, \delta_{pp'}\, \xi^{(pa\varkappa)}_{l\alpha} = \sum_\beta \sum_R R_{\alpha\beta}\, \chi^{(p')}(R)\, H_{R^{-1}(l),\beta;\,l\alpha}, \tag{20a}$$

was über l und α summiert wegen der Normierung der Eigenvektoren

$$\frac{h}{f_{p'}} \sum_a \nu^2_{p'a} = \sum_{\alpha\beta} \sum_R \sum_l R_{\alpha\beta}\, \chi^{(p')}(R)\, H_{R^{-1}(l)\beta;\,l\alpha} \tag{21}$$

für die Summe der Quadrate aller zur Darstellung $D^{(p')}(R)$ ge-

hörigen Eigenfrequenzen ergibt. Die Summe rechts in (21) läßt sich, wenn die Bewegungsgleichungen des Systems, also die Matrix $H_{k\beta;\,l\alpha}$ bekannt ist, mühelos berechnen, (21) entspricht genau der HEITLER'schen Mittelwertformel [12]). Sie genügt zur Berechnung der Frequenzen derjenigen Eigenschwingungen, zu deren Darstellung keine andere Eigenschwingung mehr gehört.

Ist dies nicht der Fall, so muß man dem HEISENBERG'schen Verfahren [13]) entsprechend das Quadrat, die dritte Potenz usw. von $H_{k\beta;\,l\alpha}$ bilden, dessen Eigenwerte die vierten, sechsten usw. Potenzen der Eigenfrequenzen sind, während die zugehörigen Eigenvektoren die $\mathfrak{f}^{(pax)}$ bleiben. Wendet man auf diese das vorher für $H_{k\beta;\,l\alpha}$ geschilderte Verfahren an, so erhält man Formeln für die Summe der vierten, sechsten usw. Potenzen der zu derselben Darstellung gehörigen Eigenfrequenzen:

$$\frac{h}{f_p}\sum_a \nu_{pa}^{2t} = \sum_{\alpha\beta}\sum_{\varkappa}\sum_{l} R_{\alpha\beta}\chi^{(p)}(R)\, H^{(t)}_{R(l^{-1}),\beta;\,l\alpha}, \tag{21a}$$

wo die $H^{(t)}_{k\beta;\,l\alpha}$ Matrixelemente der $t$ten Potenzmatrix von $H_{k\beta;\,l\alpha}$ sind.

Aus diesen Gleichungen (sie sind der sogenannten irreduziblen Säkulargleichung der Quantenmechanik äquivalent) kann man die Eigenfrequenzen verhältnismäßig leicht berechnen.

In den folgenden Tabellen sind die irreduziblen Darstellungen der 32 Kristallklassen zusammengestellt [14]). Die Charaktere der Kristallklassen C, die das direkte Produkt einer anderen Kristallklasse C_0 mit der Spiegelungsgruppe $C_i = (E, i)$ (oder $C_s = (E, \sigma_h)$) sind, $C = C_0 \times C_i$ (bezw. $C = C_0 \times C_s$), sind nicht separat angegeben. Man erhält aus jeder Darstellung von C_0 je zwei von C in der Weise, daß man den Elementen A von C, die auch in C_0 vorkommen, dieselbe Matrix wie in der Darstellung von C_0 zuordnet. Die Elemente, die in C nicht vorkommen, lassen sich immer Ai (bezw. $A\sigma_h$) schreiben, wo A in C_0 vorkommt. Man ordnet ihnen dieselbe Matrix, die man A zugeordnet hat, zu, entweder durchweg mit positivem, oder durchweg mit negativem Vorzeichen. So erhält man alle Darstellungen von C, genau doppelt so viele wie C_0 hat.

Die Kristallklassen, die sich als direkte Produkte schreiben lassen, sind bekanntlich die folgenden:

12) W. HEITLER, Zs. f. Phys. 46. 47. 1927.

13) W. HEISENBERG, Zs. f. Phys. 49. 619. 1928.

14) Die Bezeichnungen sind dieselben, die auch in P. P. EWALD's Artikel in Geiger-Scheels Handbuch der Physik, Band XXIV, verwendet sind.

$C_{2h} = C_2 \times C_s = C_2 \times C_i$; $C_{3h} = C_3 \times C_s$; $C_{4h} = C_4 \times C_s = C_4 \times C_i$;
$C_{2v} = C_s \times C_s$; $C_{6h} = C_6 \times C_s = C_6 \times C_i$
$D_{3h} = D_3 \times C_s$; $D_{4h} \times C_s = D_4 \times C_i$; $D_{6h} = D_6 \times C_s = D_6 \times C_i$
$C_{3i} = C_3 \times C_i$; $V_h = V \times C_s = V \times C_i$; $D_{3d} = C_{3v} \times C_i$
$T_h = T \times C_i$; $O_h = O \times C_i$.

Die Darstellungen der zueinander holomorphen Gruppen sind in den folgenden Tabellen zusammengefaßt. In jeder Tabelle steht zuerst der Name der abstrakten Gruppe, zu der die in den folgenden Zeilen in der ersten Spalte angeführten Kristallklassen holomorph sind. In der Zeile einer jeden Kristallklasse stehen die Zeichen ihrer einzelnen Klassen mit der Anzahl der in ihnen enthaltenen Elemente (z. B. bedeutet $3\,C_2$ eine Klasse, die aus drei zweizähligen Drehungen besteht). Einander entsprechende Klassen stehen untereinander. Unter dem Strich kommen die Charaktere der irreduziblen Darstellungen, jeder Zeile entspricht eine Darstellung. Die konjugiert komplexen Darstellungen sind durch Klammern zusammengefaßt.

Spiegelungsgruppe

C_s	E	σ_h
C_2	E	C_2
C_i	E	i
	1	1
	1	-1

Vierergruppe

C_{2h}	E	C_2	σ_h	i
C_{2v}	E	C_2	σ_v	σ_v
V	E	C_2	C_2	C_2
	1	1	1	1
	1	-1	-1	1
	1	1	-1	-1
	1	-1	1	-1

3. zyklische ($\omega = -\frac{1}{2} + i\frac{1}{2}\sqrt{3}$)

C_3	E	C_3	C_3^2
	1	1	1
	$\begin{pmatrix}1\\1\end{pmatrix}$	$\begin{pmatrix}\omega\\\omega^2\end{pmatrix}$	$\begin{pmatrix}\omega^2\\\omega\end{pmatrix}$

3. Diedergruppe

C_{3v}	E	$2\,C_3$	$3\,\sigma_v$
D_3	E	$2\,C_3$	$3\,C_2$
	1	1	1
	1	1	-1
	2	-1	0

4. zyklische

C_4	E	C_2	C_4	C_4^3
S_4	E	C_2	S_4	S_4^3
	1	1	1	1
	1	1	-1	-1
	$\begin{pmatrix}1\\1\end{pmatrix}$	$\begin{pmatrix}-1\\-1\end{pmatrix}$	$\begin{pmatrix}i\\-i\end{pmatrix}$	$\begin{pmatrix}-i\\i\end{pmatrix}$

4. Diedergruppe

C_{4v}	E	C_2	$2C_4$	$2\sigma_d$	$2\sigma_v$
D_4	E	C_2	$2C_4$	$2C_2$	$2C_2$
V_d	E	C_2	$2S_4$	$2\sigma_d$	$2C_2$
	1	1	1	1	1
	1	1	1	-1	-1
	1	1	-1	1	-1
	1	1	-1	-1	1
	2	-2	0	0	0

6. zyklische $\omega = \frac{1}{2} + i\frac{1}{2}\sqrt{3}$

C_6	E	C_6	C_3	C_2	C_3^2	C_6^5
	1	1	1	1	1	1
	1	-1	1	-1	1	-1
	$\begin{pmatrix}1\\1\end{pmatrix}$	$\begin{matrix}\omega^2\\\omega^4\end{matrix}$	$\begin{matrix}\omega^4\\\omega^2\end{matrix}$	$\begin{matrix}1\\1\end{matrix}$	$\begin{matrix}\omega^2\\\omega^4\end{matrix}$	$\begin{matrix}\omega^4\\\omega^2\end{matrix}$
	$\begin{pmatrix}1\\1\end{pmatrix}$	$\begin{matrix}\omega\\-\omega^2\end{matrix}$	$\begin{matrix}\omega^2\\-\omega\end{matrix}$	$\begin{matrix}-1\\-1\end{matrix}$	$\begin{matrix}-\omega\\\omega^2\end{matrix}$	$\begin{matrix}-\omega^2\\\omega\end{matrix}$

6. Diedergruppe

C_{6v}	E	C_2	$2C_3$	$2C_6$	$3\sigma_v$	$3\sigma_d$
D_6	E	C_2	$2C_3$	$2C_6$	$3C_2$	$3C_2$
	1	1	1	1	1	1
	1	1	1	1	-1	-1
	1	-1	1	-1	1	-1
	1	-1	1	-1	-1	1
	2	2	-1	-1	0	0
	2	-2	-1	1	0	0

„Tetraedergruppe“ $\omega = -\frac{1}{2} + i\frac{1}{2}\sqrt{3}$

T	E	$3C_2$	$4C_3$	$4C_3$
	1	1	1	1
	$\begin{pmatrix}1\\1\end{pmatrix}$	$\begin{matrix}1\\1\end{matrix}$	$\begin{matrix}\omega\\\omega^2\end{matrix}$	$\begin{matrix}\omega^2\\\omega\end{matrix}$
	3	-1	0	0

146 E. WIGNER, Über die elastischen Eigenschwingungen symmetrischer Systeme.

„Oktaedergruppe“

O	E	$8C_3$	$3C_2$	$6C_2$	$6C_4$
T_d	E	$8C_3$	$3C_2$	$6\sigma_d$	$6S_4$
	1	1	1	1	1
	1	1	1	-1	-1
	2	-1	2	0	0
	3	0	-1	1	-1
	3	0	-1	-1	1

Über eine Verschärfung des Summensatzes

E. P. Wigner

Physikalische Zeitschrift *32*, 450–453 (1931)

Zur Zeit der Entdeckung der Quantenmechanik spielte der Summensatz der Übergangswahrscheinlichkeiten[1]) eine große Rolle; er diente als Fingerzeig bei der Aufstellung der Vertauschungsrelationen. Für die wirkliche Berechnung von Übergangswahrscheinlichkeiten kann man ihn dagegen nicht sehr leicht verwenden, weil er eigentlich nur die Summe der Übergangswahrscheinlichkeiten (genauer: der f-Werte) aller Sprünge, die von einem Niveau ausgehen, zu berechnen gestattet, nicht dagegen die Summe der Wahrscheinlichkeiten für die Übergänge der einzelnen Elektronen. Dies ist vielmehr nur in einer gewissen Annäherung und auch so nur in den einfachsten Fällen möglich, nämlich nur dann, wenn

1) W. Kuhn, Zeitschr. f. Phys. **33**, 408, 1925; W. Thomas, Naturw. **13**, 627, 1925; F. Reiche und W. Thomas, Zeitschr. f. Phys. **34**, 510, 1925.

die Wirkung der anderen Elektronen auf das Leuchtelektron als die einer Ladungswolke betrachtet werden kann[1]).

Deshalb wird auch eine Verschärfung (wenigstens die hier behandelte Verschärfung) des Summensatzes nicht von großem Nutzen bei der Berechnung von Übergangswahrscheinlichkeiten sein. Ich teile sie hier doch mit, weil sie im vorher erwähnten Fall, wie sich zeigen wird, einige Übergangswahrscheinlichkeiten auch einzeln zu berechnen gestattet.

1. Bezeichnen wir in einem Atom mit n-Elektronen die Summe der n Heisenbergschen Z-Koordinatenmatrixelemente, die den stationären Zuständen P und Q entsprechen, mit

$$Z_{PQ} = \int \cdots \int \psi_P^* (z_1 + z_2 + \cdots + z_n) \psi_Q, \quad (1)$$

so ist bekanntlich nach den Vertauschungsrelationen[2])

$$\sum_P (E_Q - E_P) |Z_{PQ}|^2 = -\frac{h^2 n}{8\pi^2 m}, \quad (2c)$$

wo E_P bzw. E_Q die Energien der beiden Zustände P bzw. Q sind, und die Summation über alle Zustände (über alle orthogonalen Schrödingerschen Eigenfunktionen) zu erstrecken ist, m ist die Elektronenmasse. Dieselbe Gleichung gilt für die Summe der X- und Y-Koordinatenmatrizen:

$$\sum_P (E_Q - E_P) |X_{PQ}|^2 = -\frac{h^2 n}{8\pi^2 m}, \quad (2a)$$

$$\sum_P (E_Q - E_P) |Y_{PQ}|^2 = -\frac{h^2 n}{8\pi^2 m}. \quad (2b)$$

Nun ist die Übergangswahrscheinlichkeit A_{PQ} vom Zustand Q in den Zustand P, also die prozentische Anzahl der Atome im Zustand Q, die pro Sekunde durch spontane Ausstrahlung in den Zustand P übergehen:

$$A_{PQ} = \frac{64\pi^4 e^2 (E_Q - E_P)^3}{3 h^4 c^3} \cdot (|X_{PQ}|^2 + |Y_{PQ}|^2 + |Z_{PQ}|^2), \quad (3)$$

wo e die Ladung des Elektrons ist. Es gilt daher für die Summe der Übergangswahrscheinlichkeiten der Summensatz

$$\sum_P \frac{A_{PQ}}{(E_P - E_Q)^2} = -\frac{8\pi^2 e^2 n}{h^2 m c^3} \quad (4)$$

aus dem die Vertauschungsrelationen eigentlich entstanden sind. Durch das Vorangehende ist A_{PQ} nur für $E_P < E_Q$ definiert, ist $E_P > E_Q$, so soll nach (3) $A_{PQ} = -A_{QP}$ die Übergangswahrscheinlichkeit von P nach Q mit dem negativen Vorzeichen sein.

Wir wissen andererseits, daß die Übergangswahrscheinlichkeiten A_{PQ} nur dann von Null verschieden sind, wenn die Azimutalquantenzahl des Zustandes P entweder um 1 größer, oder um 1 kleiner oder gleich ist der Azimutalquantenzahl l des Zustandes Q. Man kann daher die Summe (4) in drei Teile zerlegen

$$\sum_P^{+} \frac{A_{QP}}{(E_Q - E_P)^2} + \sum_P^{0} \frac{A_{QP}}{(E_Q - E_P)^2} + \sum_P^{-} \frac{A_{QP}}{(E_Q - E_P)^2} = \frac{8\pi^2 e^2 n}{h^2 m c^3}, \quad (4a)$$

wo die erste Summe über jene Zustände P zu erstrecken ist, deren Azimutalquantenzahl $l + 1$ ist, die zweite und dritte Summe über jene Zustände, deren Azimutalquantenzahl l bzw. $l - 1$ ist. Es fragt sich nun, ob man die drei Summen in (4 a), die wir mit

$$\left.\begin{aligned} \sum_P^{+} \frac{A_{QP}}{(E_Q - E_P)^2} &= \frac{8\pi^2 e^2 f_+^Q}{h^2 m c^3} \\ \sum_P^{0} \frac{A_{QP}}{(E_Q - E_P)^2} &= \frac{8\pi^2 e^2 f_0^Q}{h^2 m c^3} \\ \sum_P^{-} \frac{A_{QP}}{(E_Q - E_P)^2} &= \frac{8\pi^2 e^2 f_-^Q}{h^2 m c^3} \end{aligned}\right\} \quad (4b)$$

bezeichnen wollen, nicht auch gesondert bestimmen kann?

Dies ist nun zwar im allgemeinen nicht möglich, wie man schon daraus ersieht, daß bei $n = 1$, bei dem Einelektronenproblem, die Übergänge ohne Änderung der Azimutalquantenzahl verboten sind, so daß $f_0^Q = 0$ wird, während sonst

1) Nach Y. Sugiura (Phil. Mag. 4, 495, 1927) gilt dies für die inneren Elektronen des *Na* in bezug auf das Leuchtelektron, die Summe der f-Werte für die Übergänge des Leuchtelektrons allein ist 1,05. Dagegen kann man in den Röntgenspektren nach R. de L. Kronig und A. H. Kramers (Zeitschr. f. Phys. **48**, 174, 1928) die äußeren Elektronen bei der Betrachtung der Übergänge der K-Elektronen nicht mehr mit einer Ladungswolke beschreiben, weil sie — abgesehen von ihrer abschirmenden Wirkung — die Übergänge der K-Elektronen auch dadurch beeinflussen, daß wegen des Pauliprinzipes Übergänge in die von ihnen besetzten Bahnen unmöglich sind. Bei Sugiura dagegen sind die durch das Pauliprinzip verbotenen Übergänge ohnehin sehr schwach und betragen nur etwa 5 Proz. der gesamten f-Werte.

2) Da die Impulsmatrixelemente jeweils die $2\pi i m (E_P - E_Q)/h$-fachen der entsprechenden Koordinantenmatrixelemente sind.

im allgemeinen f_0^Q nicht Null ist. Man kann aber noch sehr wohl außer

$$f_+^Q + f_0^Q + f_-^Q = n \tag{5}$$

eine weitere Gleichung zwischen f_+^Q, f_0^Q, f_-^Q angeben, die dann bei $n = 1$ mit $f_0^Q = 0$ zusammen f_+^Q und f_-^Q auch gesondert zu bestimmen gestattet.

2. Führen wir an Stelle von P dreifache Indizes N, l', μ' für die Atomzustände ein, wo l' und μ' die azimutale und magnetische Quantenzahl bedeuten, und N zur Unterscheidung der verschiedenen Zustände mit derselben azimutalen und magnetischen Quantenzahl dient. Es läßt sich, wenn Q ein Zustand mit der azimutalen und magnetischen Quantenzahl l und μ ist, (2 c) ähnlich zu (4 a)

$$\left.\begin{aligned} &\sum_N (E_{Nl+1} - E_Q)\,|Z_{Nl+1\mu,Q}|^2 \\ &+ \sum_N (E_{Nl} - E_Q)\,|Z_{Nl\mu,Q}|^2 \\ &+ \sum_N (E_{Nl-1} - E_Q)\,|Z_{Nl-1\mu,Q}|^2 \\ &= \frac{h^2 n}{8\pi^2 m} \end{aligned}\right\} \tag{6}$$

schreiben, wo man über μ' nicht zu summieren hat, weil die Z-Matrixelemente alle verschwinden, wenn $\mu' \neq \mu$ ist. Setzen wir noch $\mu = l$, so ist nach den bekannten Intensitätsformeln für den Zeemaneffekt bei einem $l \to l$ Übergang, der $l/(l+1)$-te Teil der Strahlung des Zustandes Q in der Z-Richtiung polarisiert, so daß

$$|Z_{Nll,Q}|^2 = \frac{l}{l+1} \sum_{\mu'} (|X_{Nl\mu',Q}|^2 + |Y_{Nl\mu',Q}|^2 + |Z_{Nl\mu',Q}|^2) \tag{7a}$$

wird. Ähnlich ist für $\mu = l$ die Strahlung des Zustandes Q bei einem $l+1 \to l$-Übergang zum $(2l+1)/(2l+3)\,(l+1)$-ten Teil in der Z-Richtung polarisiert, so daß

$$|Z_{Nl+1l,Q}|^2 = \frac{2l+1}{(2l+3)(l+1)} \sum_{\mu'} (|X_{Nl+1\mu',Q}|^2 + |Y_{Nl+1\mu',Q}|^2 + |Z_{Nl+1\mu',Q}|^2) \tag{7b}$$

ist, während bei einem $l - 1 \to l$-Übergang die ganze Strahlung des Zustandes Q senkrecht zu Z polarisiert ist, so daß

$$Z_{Nl-1l,Q} = 0 \tag{7c}$$

gesetzt werden muß. Wenn man (7) in (6) einsetzt, folgt mit Hilfe von (3)

$$\begin{aligned} &\frac{2l+1}{(2l+3)(l+1)} \sum_{N\mu'} \frac{3h^4c^3 A_{Nl+1\mu',Q}}{64\pi^4 e^2 (E_Q - E_{Nl+1})^2} \\ &+ \frac{l}{l+1} \sum_{N\mu'} \frac{3h^4c^3 A_{Nl\mu',Q}}{64\pi^4 e^2 (E_Q - E_{Nl})^2} = -\frac{h^2 n}{8\pi^2 m} \end{aligned} \tag{8}$$

oder mit Hilfe der f ausgedrückt

$$\frac{2l+1}{(2l+3)(l+1)} f_+^Q + \frac{l}{l+1} f_0^Q = \frac{1}{3} n. \tag{9}$$

Diese Gleichungen gelten zunächst nur für einen Zustand Q, dessen magnetische Quantenzahl gleich seiner azimutalen Quantenzahl ist. Da aber nach den erwähnten Ornstein-Burgerschen Intensitätsregeln die Summe der Übergangswahrscheinlichkeiten

$$\sum_{\mu'} A_{Nl'\mu',Q}$$

von der magnetischen Quantenzahl von Q unabhängig ist, gilt (9) auch allgemein.

3. Betrachten wir zuerst ein Einelektronenproblem, dann ist $f_0^Q = 0$ und es ergibt sich aus (5) und (9)

$$3f_+^Q = \frac{(2l+3)(l+1)}{2l+1}, \quad 3f_-^Q = -\frac{(2l-1)l}{2l+1}. \tag{10}$$

Es zeigt sich also, daß f_-^Q für $l \geqq 1$ negativ ist, was soviel bedeutet, daß unter den Übergängen mit $\Delta l = 1$ die Emission stärker vertreten ist, als die Absorption. Es folgt hieraus, daß wenigstens ein Zustand mit der Azimutalquantenzahl $l - 1$ tiefer liegen muß, als selbst der tiefste Zustand mit der Azimutalquantenzahl l. Daher ist auch der Grundzustand bei dem Einelektronenproblem immer ein S-Zustand.

Bezeichnen wir das tiefste Niveau mit der Azimutalquantenzahl l mit E_l und setzen wir voraus, daß unter E_l nur ein Niveau E_{l-1} mit der Azimutalquantenzahl $l - 1$ liegt, wie das etwa im Spektrum des Wasserstoffatoms für alle l der Fall ist. Ist dann P_l ein Zustand mit der Energie E_l, so sind in der Summe $3f_-^{P_l}$ nur jene Glieder negativ, die Übergängen in die Zustände des Niveaus E_{l-1} entsprechen. Die Summe $3f_{E_{l-1}P_l}$ jener Glieder in $3f_-^{P_l}$, die diesen Übergängen entsprechen (sie ist für alle Zustände des Niveaus E_l gleich) muß also dem Betrag nach größer als $l(2l-1)/(2l+1) = l - 1 + 1/(2l+1)$ sein. Andererseits enthält die Summe $3f_+^{P_l - 1} = (2l+1)\,l/(2l-1)$ lauter positive Glieder,

und daher muß die Summe $3f_{E_l P_{l-1}}$ jener Glieder, die Übergängen in die Zustände des Niveaus E_l entsprechen, kleiner als $3f_-^{P_{l-1}} = (2l+1)\,l/(2l-1)$ sein. Da weiter nach den Summenregeln $-(2l+1)f_{E_{l-1}P_l} = (2l-1)f_{E_l P_{l-1}}$ ist, ergibt sich

$$l < 3f_{E_l P_{l-1}} < l + 1 + \frac{1}{2l-1} \qquad (11\,\mathrm{a})$$

oder

$$l - 1 + \frac{1}{2l+1} < -3f_{E_{l-1}P_l} < l. \qquad (11\,\mathrm{b})$$

Man sieht, daß bei hohem l die Übergangswahrscheinlichkeit von P_l nach E_{l-1} (und nach E_{l+1}) cca. wie l zunimmt, und je größer l ist, umso genauer aus (11) entnommen werden kann. Es folgt auch, daß je größer l ist, die Interkombinationen von P_l mit den anderen Zuständen immer mehr in den Hintergrund treten: da

$$3f_+^{P_l} = l + 2 + \frac{1}{2l+1}$$

ist und davon auf den Übergang nach E_{l+1} allein wenigstens $l+1$ entfällt, bleibt für die anderen Übergänge von P_l mit $\Delta l = 1$ einschließlich des kontinuierlichen Spektrums höchstens $1 + 1/(2l+1)$ übrig. Ähnlich ist die Summe der Glieder in $3f_-^{P_l}$ die von Absorptionsübergängen (also nicht von Übergängen nach E_{l-1}) herrühren, sicher kleiner als $1 - 1/(2l+1)$.

All dies gilt natürlich nur für das Einkörperproblem oder für Systeme, die man als Einkörperprobleme behandeln darf. Die letzten Regeln setzen zudem voraus, daß unter P_l nur ein einziger Term mit der Azimutalquantenzahl $l-1$ liegt.

Das Grenzgesetz $f_{E_l P_{l-1}} \sim l/3$ ergibt sich auch aus dem Korrespondenzprinzip unabhängig vom Kraftgesetz, wenn man beachtet, daß die Bahnen des Niveaus E_l und E_{l-1} als Kreisbahnen anzusehen sind.

4. Es sei hier noch für den Fall, daß auch Übergänge ohne Änderung der Azimutalquantenzahl vorkommen, eine Abschätzung angegeben. Wir sahen, daß der Grundzustand ein S-Zustand sein muß, wenn diese Übergänge verboten sind. Umgekehrt müssen die Übergänge mit $\Delta l = 0$ eine wesentliche Rolle spielen, wenn die Azimutalquantenzahl des Grundzustandes $l \neq 0$ ist. In der Tat folgt aus (5) und (9), wenn man aus ihnen f_+^Q eliminiert

$$\frac{l+1}{l} f_0^Q - \frac{2l+1}{l(2l-1)} f_-^Q = \frac{n}{3}. \qquad (12)$$

Da für den Grundzustand f_-^Q positiv sein muß, ist

$$f_0^Q > \frac{l}{l+1}\,\frac{n}{3}$$

und ähnlich folgt aus (9), da f_+^Q positiv sein muß

$$f_0^Q < \frac{l+1}{l}\,\frac{n}{3},$$

d. h. daß f_0^Q für den Grundzustand, wenn seine Azimutalquantenzahl groß ist, $\sim n/3$ ist.

Die Formeln (5) und (9) gelten auch, wenn man in ihnen die Azimutalquantenzahl durch die Gesamtquantenzahl (innere Quantenzahl) ersetzt. Im Falle normaler Termordnung folgen indessen die so zu erhaltenden Formeln aus (5) und (9) mit Hilfe der bekannten Intensitätsformeln für Multipletts.

(Eingegangen 31. März 1931.)

Über die Operation der Zeitumkehr in der Quantenmechanik

E. P. Wigner

Nachrichten der Gesellschaft der Wissenschaften zu Göttingen
Mathematisch-Physikalische Klasse *1932*, 546–559

Vorgelegt von M. Born in der Sitzung am 25. November 1932.

1. H. A. Kramers[1]) hat bei der Untersuchung der magnetischen Drehung der Polarisationsebene wichtige allgemeine Regelmäßigkeiten der Lösungen der quantenmechanischen Eigenwertgleichung gefunden. Er hat u. a. nachgewiesen, daß solange die äußeren, auf ein System wirkenden Kräfte rein elektrischer Natur sind, bei ungerader Teilchenzahl immer eine wenigstens zweifache Entartung der Energieniveaus auftritt. Kramers legte seinen Betrachtungen in erster Reihe die Eigenwertgleichung zugrunde, die G. Breit[2]) aufgestellt hat, und die noch in zweiter Ordnung relativistisch invariant ist. Er zeigte dann durch direkte Rechnung, daß man aus einer Lösung der Breitschen Gleichungen eine zweite finden kann, die bei ungerader Elektronenzahl von der ersten verschieden sein muß.

Im folgenden soll es versucht werden, die Kramersschen Resultate auf eine etwas allgemeinere Grundlage zu stellen. Es wird sich insbesondere zeigen, daß die von ihm benutzte Operation die Operation der Zeitumkehr ist. Außerdem soll diese Operation in das normale Schema der Gruppentheorie der Terme eingebaut werden, was, wie wir sehen werden, wegen ihres nicht linearen Charakters nicht gänzlich trivial ist.

Bekanntlich beruhen viele allgemeine Eigenschaften quantenmechanischer Systeme, insbesondere die für die Spektroskopie wichtigen, auf Symmetrieüberlegungen. Für die Anwendung dieser

1) H A. Kramers, Proc. Amsterdam, XXXIII, 959, 1930. § 2. Neuerdings hat J. H. van Vleck die Kramersschen Resultate mit sehr großem Erfolg auch auf Fragen des Paramagnetismus angewendet. (Electric and magnetic susceptibilities, Oxford 1932.)

2) G. Breit, Phys. Rev. 34. 553. 1929.

Überlegungen ist es wichtig, daß die betrachtete Energiestufe nur eine endliche Entartung aufweise. Dies ist für im Raume frei bewegliche Systeme (z. B. Atome) wegen des kontinuierlichen Spektrums, das mit der Bewegung des Systems als Ganzes verbunden ist, von vornherein nicht der Fall. Man kann sich da auf zwei Arten helfen: entweder rechnet man die Kerne nicht zum System und betrachtet diese als an feste Stellen des Raumes gebunden, oder aber man führt eine Zusatzbedingung ein, die die freie Beweglichkeit des Systems im Raume ebenfalls aufhebt, indem man außer der Energie noch vorschreibt, daß der Gesamtimpuls des Systems Null sei[3]). Bei Atomen führen beide Betrachtungen zu denselben Resultaten, bei Molekülen verwendet man bekanntlich die erstere, wenn man sich für die Elektronenterme interessiert, die zweite dagegen, wenn man auch die Bewegung der Kerne mit erfassen will. Da die Kerne in Wirklichkeit nicht an feste Punkte im Raume gebunden sind, führt die erste Betrachtungsweise nur zu näherungsweise richtigen Resultaten[4]).

Die folgenden Betrachtungen, deren Zweck die Zurückführung der KRAMERSschen Regelmäßigkeiten auf eine bisher nicht betrachtete Symmetrieart ist, können in beiden Fällen, die wir als „erste" (festgehaltene Kerne) und „zweite" (Gesamtimpuls Null) unterscheiden wollen, angewendet werden.

2. Wenn man seinen Betrachtungen ein frei bewegliches System zugrunde legen würde, würde im Falle, daß keine äußeren Felder vorhanden sind, die Symmetriegruppe des Systems die ganze inhomogene LORENTZgruppe bzw. im nicht-relativistischen Falle die entsprechende GALILEIgruppe umfassen, die außer den gewöhnlichen LORENTZ- bzw. GALILEI-Transformationen noch Verschiebungen in Raum und Zeit enthalten ($x_i' = x_i + a_i$). Dadurch, daß man nicht ein frei bewegliches System betrachtet, erfährt die Symmetrie des Problems natürlich Einschränkungen.

In der „ersten" Betrachtungsweise wird durch die festen Anziehungszentra vor allem eine absolute Ruhe definiert, und dies schließt die Transformationen, bei denen die beiden Koordinatensysteme gegeneinander bewegt sind, aus. Auch die meisten rein räumlichen Transformationen fallen aus dem gleichen Grunde fort, es bleibt von ihnen nur die Symmetriegruppe der Anziehungs-

3) Die Forderung, daß sowohl die Energie, wie auch der Gesamtimpuls feste Werte haben sollen (daß beide „scharf" sein sollen) ist bekanntlich zulässig, wenn keine äußeren Felder da sind.

4) Über die Berechtigung dieser Betrachtungsweise vgl. M. BORN und I. R. OPPENHEIMER, Ann. d. Phys. 84. 457. 1927.

zentren übrig, die bisher auch immer in Betracht gezogen wurde. Außer diesen existieren aber noch die rein zeitlichen Transformationen $t' = t + t_0$ und $t' = -t$. Durch die Angabe der Energie ist die Darstellung $\left(e^{-iEt_0/\hbar}\right)$ schon gegeben, zu der der Zustand in Bezug auf die ersteren Transformationen gehört; diese Symmetrie hat u. a. die Folge, daß Zustände mit scharfer Energie zeitlich unveränderlich sind. Die Transformationen $t' = t + t_0$ geben also nichts Neues mehr. Es bleibt demnach nur die Transformation $t' = -t$ zu untersuchen, die die Reversibilität der Zeit ausdrückt. Sie wird, wie wir sehen werden, tatsächlich die KRAMERSschen Regeln ergeben. Die Reversibilität der Zeit gilt natürlich nur so lange die Anziehungszentra auch durch ihre innere Bewegung keine Zeitrichtung auszeichnen, also insbesondere keine magnetischen Felder erzeugen[5]). Die Felder, die durch den Kernspin verursacht sind, sind allerdings so klein, daß sie in der ersten, ohnehin nur näherungsweise gültigen Betrachtung außer Acht gelassen werden können.

Ähnlich verhält es sich bei der „zweiten" Betrachtungsweise. Bei ihr ist der Gesamtimpuls des Systems in einem Koordinatensystem gleich Null gesetzt, dieses Koordinatensystem ist ausgezeichnet und kann als ruhend betrachtet werden. Von den rein räumlichen Transformationen brauchen wir die Translationen nicht zu berücksichtigen, weil das System, dessen Gesamtimpuls Null ist, in Bezug auf diese zur identischen Darstellung gehört, d. h. durch eine Translation überhaupt nicht geändert wird. Mit den übrigen rein räumlichen und mit den rein zeitlichen Transformationen verhält es sich ebenso wie vorher, mit dem einzigen Unterschied, daß die rein räumlichen gewöhnlich (wenn keine äußeren Felder vorhanden sind) die ganze dreidimensionale Drehsiegelungsgruppe umfassen.

Wir sehen also, daß sowohl das erste, wie auch das zweite quantenmechanische Problem außer den Symmetrien, die man bisher immer betrachtet hat (rein räumliche Transformationen), noch das Symmetrieelement der Zeitumkehr enthält. Es sei im folgenden die Wirkung dieses Symmetrieelements untersucht, und zwar nur im nicht relativistischen Falle.

3. Zunächst müssen wir untersuchen, welche Operation K eine Wellenfunktion φ in die Wellenfunktion $K\varphi$ desjenigen Zustandes

5) Über die Frage, unter welchen Bedingungen die Zeit „reversibel" ist, was wir fortan ausdrücklich annehmen wollen, vgl. die Diskussion von ONSAGER, Phys. Rev. 38. 2265. 1931.

überführt, dessen Zeit umgekehrt der von φ läuft, dessen Zukunft also identisch mit der Vergangenheit und dessen Vergangenheit identisch mit der Zukunft von φ ist. Nehmen wir zunächst die einfache SCHRÖDINGERsche Theorie, die den Spin unberücksichtigt läßt. Dann ist $\varphi = \varphi(x_1, x_2, \ldots, x_n)$ eine Funktion der Cartesischen Koordinaten der Teilchen allein.

In der einfachen SCHRÖDINGERschen Theorie ist bekanntlich $K\varphi = \varphi^*$, doch wollen wir die Bezeichnung K beibehalten und nur die folgenden drei Rechenregeln für die Operation K benützen:

$$\text{(I)} \qquad K^2\varphi = \varphi, \quad K^2 = 1,$$

$$\text{(II)} \qquad K(a\varphi + b\psi) = a^* K\varphi + b^* K\psi,$$

$$\text{(III)} \qquad (\varphi, \psi) = (\psi, \varphi)^* = (K\psi, K\varphi)$$

oder

$$\text{(III a)} \qquad (K\varphi, \psi) = (\varphi, K\psi)^* = (K\psi, \varphi).$$

[Hier bedeutet (u, v) das hermiteische skalare Produkt der Funktionen u und v, also das über den ganzen Konfigurationsraum erstreckte Integral von $u^* v$.] Es zeigt sich nämlich, daß die folgenden Betrachtungen in dieser Weise auch für die Theorie, die den Spin berücksichtigt im Falle gerader Elektronenzahl wörtlich gelten, mit dem einzigen Unterschied, daß K nicht mehr einfach den Übergang zum konjugiert Komplexen bedeutet, sondern eine etwas andere Operation ist, die den Regeln (I), (II), (III) ebenfalls genügt. Gleichung (II) zeigt, was für das folgende entscheidend ist, daß K kein linearer Operator ist. Alle anderen vorkommenden Operatoren werden jedoch linear sein, sodaß dies nicht immer besonders angegeben wird.

Reell ist eine Funktion, für die $K\varphi = \varphi$ gilt, imaginär eine, für die $K\varphi = -\varphi$ gilt. Von einem hermiteischen linearen Operator A sagen wir, daß er reell ist, wenn er eine reelle Funktion in eine reelle überführt, er führt dann konjugiert komplexe in konjugiert komplexe über, so daß

$$\text{(1)} \qquad AK\varphi = KA\varphi, \quad AK = KA, \quad A = KAK$$

gilt. Ein imaginärer Operator B führt reelle Funktionen in imaginäre über, es gilt für ihn

$$\text{(2)} \qquad B = -BKB.$$

Aus der Umkehrbarkeit der Zeit folgt nach 1, daß mit φ auch $K\varphi$ eine Eigenfunktion ist. Man kann φ und $K\varphi$ durch die reellen Funktionen $\varphi + K\varphi$ und $i(\varphi - K\varphi)$ ersetzen. Es können also alle Eigenfunktionen reell angenommen werden. Daher muß auch der

Energieoperator reell sein, wie wir das vom SCHRÖDINGERschen Operator auch wissen.

Dasselbe gilt auch für alle rein räumlichen Symmetrieoperatoren. Man erhält nämlich denselben Zustand, gleichgültig, ob man zuerst eine räumliche Drehung Q und dann die Zeitumkehr vornimmt, oder zuerst die Zeit umkehrt und dann im Raume dreht. Es gilt daher für alle φ

$$KQ\varphi = c_\varphi Q K\varphi, \tag{3}$$

wo c_φ eine Konstante vom Absolutbetrag 1 ist, die zunächst für verschiedene Wellenfunktionen verschieden sein könnte. Wäre aber c_ψ verschieden von c_φ, so folgt aus der Linearität von Q doch

$$\begin{aligned} KQ(\varphi+\psi) &= KQ\varphi + KQ\psi = c_\varphi Q K\varphi + c_\psi Q K\psi \\ &= c_{\varphi+\psi} Q K(\varphi+\psi) = c_{\varphi+\psi} Q K\varphi + c_{\varphi+\psi} Q K\psi \end{aligned}$$

$c_\varphi = c_{\varphi+\psi} = c_\psi$. Ersetzt man noch Q durch $O = \sqrt{c_\varphi}\, Q$, so wird

$$KO\varphi = OK\varphi \tag{3a}$$

O tatsächlich reell. Bei der einfachen SCHRÖDINGERschen Theorie ist dies auch natürlich, da die Symmetrieoperatoren reelle Transformationen der Argumente der Wellenfunktionen sind.

Der reelle Charakter von Wellenfunktionen und Operatoren hat bei der Betrachtung des Symmetrieelements $t' = -t$ entscheidende Bedeutung. Allgemein sind Operatoren, die die Zeit nicht, oder nur in gerader Potenz enthalten, reell, weil ihr Erwartungswert für φ und $K\varphi$ gleich ist. Reelle Operatoren sind demnach die Koordinaten und ihre Funktionen, das Quadrat der Geschwindigkeit, die kinetische Energie, Gesamtenergie usw. Imaginär sind dagegen die Operatoren, die die Zeit in ungerader Potenz enthalten, wie etwa die Geschwindigkeit, der Drehimpuls usw. weil ihr Erwartungswert für φ und $K\varphi$ entgegengesetzt gleich ist. Im ersten Falle ist z. B.

$$(A\varphi, \varphi) = (\varphi, A\varphi) = (K\varphi, AK\varphi) = (KAK\varphi, \varphi),$$

sodaß tatsächlich $A = KAK$ gelten muß, weil die quadratischen Formen beider Operatoren gleich sind. Im zweiten Falle erhält man aus $(\varphi, B\varphi) = -(K\varphi, BK\varphi)$ ebenso $B = -KBK$.

4. Wir sahen eben, daß es möglich ist, alle Eigenfunktionen reell anzunehmen. Daraus folgt unmittelbar, daß der Mittelwert eines rein imaginären hermiteischen Operators für alle Zustände derselben Energie

$$(\psi_1, B\psi_1) + (\psi_2, B\psi_2) + (\psi_3, B\psi_3) + \cdots = 0 \tag{4}$$

verschwindet, weil in (4) schon alle Glieder einzeln Null sind, wenn die Eigenfunktionen in rein reeller Form angenommen sind:

$$(\psi_i, B\psi_i) = (B\psi_i, \psi_i) = (K\psi_i, KB\psi_i) = -(\psi_i, BK\psi_i) = -(\psi_i, B\psi_i).$$

Für den Fall, daß keine Entartungen vorliegen, wie z. B. wenn keine räumliche Symmetrie vorhanden ist, gilt dies auch für die einzelnen Eigenfunktionen. So z. B. ist der Erwartungswert der Geschwindigheit für jeden stationären Zustand Null. Dasselbe gilt von der dritten und jeder ungeraden Potenz der Geschwindigkeitskomponenten und daher hat jede Geschwindigkeit dieselbe Wahrscheinlichkeit, wie die entgegengesetzte Geschwindigkeit. Ebenso verhält es sich mit dem Drehimpuls.

5. Gehen wir nun zu dem Fall über, daß auch eine räumliche Symmetrie vorhanden ist, und betrachten wir Eigenfunktionen $\psi_1, \ldots, \psi_l$ die in Bezug auf diese zur irreduziblen Darstellung D gehören:

$$O_R \psi_\varkappa = \sum_{\lambda=1}^{l} D(R)_{\lambda\varkappa} \psi_\lambda. \tag{5}$$

Die Zeitumkehr ist mit allen räumlichen Symmetrieelementen vertauschbar, so daß die gesamte Symmetriegruppe abstrakt das direkte Produkt der räumlichen Symmetriegruppe und der Zeitumkehr ist. Man kann aber die Darstellungstheorie nicht in der Form, wie das sonst üblich ist, anwenden, weil die Operation der Zeitumkehr nicht linear ist. Man muß vielmehr eine etwas detailliertere Überlegung anwenden, indem man die drei Fälle[6]) unterscheidet:

a) Die Darstellung D kann in eine rein reelle Form transformiert werden.

b) Die Darstellungen D und D^* sind nicht äquivalent.

c) Die Darstellungen D und D^* sind zwar äquivalent, aber sie können in keine rein reelle Form transformiert werden.

a) Im ersten Falle (wie z. B. bei allen eindeutigen Darstellungen der dreidimensionalen Drehgruppe und Drehspiegelungsgruppe, der zweidimensionalen Drehspiegelungsgruppe und der symmetrischen Gruppen) kann man D von vornherein in reeller Form annehmen. Dann folgt aus (5), wenn man zum konjugiert Komplexen übergeht

$$KO_R \psi_\varkappa = O_R K\psi_\varkappa = \sum_{\lambda=1}^{l} D(R)_{\lambda\varkappa} K\psi_\lambda. \tag{6}$$

6) Vgl. G. Frobenius und I. Schur, Berl. Ber. 186. 1906.

Dieselbe Gleichung gilt für die reellen Funktionen $\psi_\varkappa + K\psi_\varkappa = u_\varkappa$ und $i(\psi_\varkappa - K\psi_\varkappa) = v_\varkappa$, so daß sich die $u_1, \ldots, u_l$ sowohl bei räumlichen Transformationen, wie auch bei der Zeitumkehr nur unter sich transformieren. Dasselbe gilt von den $v_1, \ldots, v_l$. Es liegt daher kein Grund vor, daß diese, wenn sie von den u linear unabhängig sind, zu dem Eigenwert der u gehören sollen, es sei denn, daß eine zufällige Entartung vorliegt.

Im Falle a) bedingt also die Zeitumkehr keine zusätzliche Entartung, sie bewirkt nur, daß die Eigenfunktionen, die zur reellen Form der Darstellung gehören, durch Multiplikation mit einer Konstante alle reell gemacht werden können.

Ist A ein reeller hermiteischer Operator, und gehören ψ_i und ψ_j beide zu Eigenwerten der Klasse a), so folgt für das Matrixelement von A, das sie verbindet,

$$(\psi_i, A\psi_j)^* = (K\psi_i, KA\psi_j) = (\psi_i, A\psi_j), \tag{7}$$

daß es reell ist. Hat A noch die räumliche Symmetrie des Problems, so kann man dies noch mit (7) kombinieren und erhält einiges weitere für seine Matrixelemente. Es sei hierauf jedoch nicht weiter eingegangen, nur als Beispiel erwähnt, daß der Erwartungswert eines imaginären hermiteischen Operators, der die räumliche Symmetrie des Problems besitzt (wie z. B. von

$$\sum_n x_n p_{nx} + p_{nx} x_n + y_n p_{ny} + p_{ny} y_n + z_n p_{nz} + p_{nz} z \quad \text{oder} \quad x M_x + y M_y + z M_z)$$

für jeden stationären Zustand $a_1\psi_1 + \cdots + a_l\psi_l$ verschwindet, wenn keine zufällige Entartung vorliegt:

$$\left(\sum_\varkappa a_\varkappa \psi_\varkappa,\ B \sum_\lambda a_\lambda \psi_\lambda\right) = \sum_{\varkappa\lambda} a_\varkappa^* a_\lambda (\psi_\varkappa, B\psi_\lambda);$$

$(\psi_\varkappa, B\psi_\lambda)$ verschwindet für $\varkappa = \lambda$, weil $\psi_\varkappa$ reell ist; für $\varkappa \neq \lambda$, weil $\psi_\varkappa$ und ψ_λ zu verschiedenen Zeilen von D gehören.

b) und c) Wenn die Eigenfunktionen alle reell sind, ist die zugehörige Darstellung auch reell. Dies folgt aus (5):

$$(\psi_\lambda, O_R\psi_\varkappa) = D(R)_{\lambda\varkappa},$$

weil links sowohl ψ_λ, wie auch $O_R\psi_\varkappa$ reell sind. Kann daher eine Darstellung nicht rein reell gemacht werden, so können die zugehörigen Eigenfunktionen auch nicht reell sein. Trotzdem gehören auch $K\psi_1, \ldots, K\psi_l$ zu dem Eigenwert von $\psi_1, \ldots, \psi_l$; aber sie können durch die $\psi_1, \ldots, \psi_l$ nicht mehr linear ausgedrückt werden, wie das vorher der Fall war. Aus (5) und (3a) folgt

$$KO_R\psi_\varkappa = O_R K\psi_\varkappa = \sum_{\lambda=1}^{l} D(R)^*_{\lambda\varkappa} K\psi_\lambda. \tag{8}$$

Im Falle b) sind D und D^* nicht äquivalent, und die $K\psi_\varkappa$ die zu D^* gehören, sind demnach orthogonal auf den $\psi_\varkappa$. Die Existenz des Symmetrieelements K verdoppelt die Entartung, es gehören zwei Terme mit verschiedenen (zueinander konjugiert komplexen) Darstellungen zu demselben Eigenwert. An Stelle von (7) gilt diesmal:

$$(\psi_i, A\psi_j)^* = (K\psi_i, KA\psi_j) = (K\psi_i, AK\psi_j), \tag{9}$$

daß die Matrixelemente in ihre konjugiert komplexen Werte übergehen, wenn man die Eigenfunktionen durch die konjugiert komplexen Eigenfunktionen ersetzt, wenn dei Operator reell hermiteisch ist. Auch im Falle b) kann man noch auf weitere Eigenschaften eines reellen oder imaginären Operators schließen, wenn er auch die Symmetrie des Problems besitzt.

Auch im Falle c) bewirkt das neu hinzutretende Symmetrieelement der Zeitumkehr eine zusätzliche Entartung, es fallen in diesem Falle zwei Eigenwerte mit der gleichen Darstellung zusammen. Die Diskussion des Falles c) kann hier umsomehr unterbleiben, als die dreidimensionale Drehgruppe keine Untergruppe hat, die auch nur eine eindeutige Darstellung dieser Art hätte. Die Resultate sind ähnlich wie bei Mitberücksichtigung des Spins bei ungerader Elektronenzahl im Falle a).

6. Wir gehen jetzt zur Theorie über, die auch den Spin berücksichtigt[7]). Zuerst bestimmen wir wieder die Operation K, die „die Zeit umkehrt". Die Lagenstatistik darf K nicht ändern, dagegen wird die Wahrscheinlichkeit entgegengesetzter Geschwindigkeiten und Spin-Orientierungen für φ und $K\varphi$ gleich sein.

Die Operation K kann entweder ein unitärer Operator sein, oder ein unitärer Operator verbunden mit dem Übergang zum konjugiert Komplexen; das letztere ist der Fall[8]).

Setzen wir daher $K = UK_0$, wo U linear-unitär ist, und K_0 einfach den Übergang zum konjugiert Komplexen bedeutet. Dann müssen nach dem Vorangehenden x, y, z vertauschbar, p_x, p_y, p_z; s_x, s_y, s_z schief vertauschbar mit K sein:

7) W. Pauli, Zs. f. Phys. 43, 601, 1927.

8) E. Wigner, Gruppentheorie usw. Braunschweig 1931, S. 251—254. Dieselbe Argumentation, die dort den Übergang zum konjugiert komplexen verbietet, fordert hier diesen Übergang, da die Zeitrichtung eben umgekehrt werden soll.

$$x\varphi = KxK\varphi = UK_0xK_0U^{-1}\varphi = UxU^{-1}\varphi;$$
$$y\varphi = UyU^{-1}\varphi; \quad z\varphi = UzU^{-1}\varphi;$$
$$\frac{h}{2\pi i}\frac{\partial}{\partial x}\varphi = -UK_0\frac{h}{2\pi i}\frac{\partial}{\partial x}K_0U^{-1}\varphi = +U\frac{h}{2\pi i}\frac{\partial}{\partial x}U^{-1}\varphi;$$
$$\frac{h}{2\pi i}\frac{\partial}{\partial y}\varphi = U\frac{h}{2\pi i}\frac{\partial}{\partial y}U^{-1}\varphi; \quad \frac{h}{2\pi i}\frac{\partial}{\partial z}\varphi = U\frac{h}{2\pi i}\frac{\partial}{\partial z}U^{-1}\varphi;$$
$$s_x\varphi = -UK_0s_xK_0U^{-1}\varphi = -Us_x^*U^{-1}\varphi;$$
$$s_y\varphi = -Us_y^*U^{-1}\varphi; \quad s_z\varphi = -Us_z^*U^{-1}\varphi.$$

Die ersten beiden Gleichungen zeigen, daß U mit x, y, z, p_x, p_y, p_z vertauschbar ist und daher auf die Cartesischen Koordinaten gar nicht einwirkt. Die dritte Gleichung zeigt, daß es mit den reellen PAULIschen Spinmatrizen (s_y und s_z) schief vertauschbar, mit der imaginären (s_x) vertauschbar ist. Es kann daher bis auf eine Konstante, die man jedoch weglassen kann, nur s_x sein. Im Falle mehrerer Elektronen ergibt sich entsprechend, daß U gleich dem Produkt der s_x Operatoren aller Elektronen ist, also

$$\begin{aligned}(10)\quad & K\varphi(x_1, y_1, z_1, \ldots, x_n, y_n, z_n; \sigma_1, \ldots, \sigma_n)\\ &= s_{1x}s_{2x}\ldots s_{nx}\varphi(x_1, y_1, z_1, \ldots, x_n, y_n, z_n; \sigma_1, \ldots, \sigma_n)\\ &= (-i)^n.\sigma_1\ldots\sigma_n\varphi(x_1, y_1, z_1, \ldots, x_n, y_n, z_n; -\sigma_1, \ldots, -\sigma_n)^*.\end{aligned}$$

Dies ist genau die KRAMERSsche Transformation.

An Stelle der Gleichungen (I), (II), (III) erhalten wir somit

$$(\mathrm{I\,a})\quad K^2\varphi = UK_0UK_0\varphi = UU^*\varphi = s_{1x}\ldots s_{nx}s_{1x}^*\ldots s_{nx}^*\varphi = (-1)^n\varphi,$$
$$(\mathrm{II\,a})\quad K(a\varphi+b\psi) = UK_0(a\varphi+b\psi) = a^*UK_0\varphi+b^*UK_0\psi,$$
$$(\mathrm{III\,a})\quad (\varphi, \psi) = (K_0\psi, K_0\varphi) = (UK_0\psi, UK_0\varphi) = (K\psi, K\varphi).$$

Für gerade n sind dies genau dieselben Gleichungen, die den vorangehenden Betrachtungen zugrunde lagen, so daß diese auch in der PAULIschen Theorie wörtlich gelten, wenn die Zahl der Elektronen gerade ist. Bei ungerader Elektronenzahl dagegen gilt an Stelle von (I)

$$(\mathrm{I\,b})\qquad K^2\varphi = -\varphi; \quad K^2 = -1.$$

Mit diesem Fall wollen wir uns im nächsten Abschnitt beschäftigen.

Sind außer den Elektronen auch andere Teilchen im System enthalten, so gilt (I) oder (I b), je nachdem die Gesamtzahl der Teilchen gerade oder ungerade ist. Dabei müssen jedoch Teilchen mit s-fachem Elektronenspin s-fach gezählt werden.

7. Schon im Falle gerader Elektronenzahl ist das $K\varphi$, das wir das konjugiert Komplexe von φ nennen, nicht das konjugiert Komplexe im gewöhnlichen Sinn, sondern $UK_0\varphi = U\varphi^*$. Bei ungerader Elektronenzahl kann man aber gar nicht mehr von reellen Funktionen sprechen, da $K\varphi$ nicht mehr gleich φ sein kann, vielmehr auf φ orthogonal steht

$$(11) \qquad (K\varphi, \varphi) = (K\varphi, KK\varphi) = -(K\varphi, \varphi) = 0.$$

Es soll $K\varphi$ die zu φ konjugierte Funktion genannt werden, zu $K\varphi$ ist dann $-\varphi$ konjugiert, usw.

Reelle und imaginäre Operatoren sind auch etwas anders definiert, wie vorher. Für reelle gilt nämlich anstatt $A = K^{-1}AK = KAK$ jetzt $A = K^{-1}AK = -KAK$, für imaginäre

$$B = -K^{-1}BK = KBK.$$

Wichtige reelle Operatoren außer den in 3. erwähnten sind die Transformationen P_k der Cartesischen Koordinaten allein, die Transformationen Q_k der Spinkoordinaten, wichtige imaginäre Operatoren sind die Spingrößen s_{kx}, s_{ky}, s_{kz}.

Matrixelemente reeller hermiteischer Operatoren, die konjugierte Funktionen verbinden, verschwinden

$$(12) \quad (K\varphi, A\varphi) = (KA\varphi, K^2\varphi) = -(AK\varphi, \varphi) = -(K\varphi, A\varphi).$$

Der Erwartungswert eines reellen hermiteischen Operators ist für zwei konjugierte Funktionen gleich, der eines imaginären Operators B entgegengesetzt gleich

$$(13) \quad (\varphi, B\varphi) = (KB\varphi, K\varphi) = -(BK\varphi, K\varphi) = -(K\varphi, BK\varphi).$$

Da man die Eigenfunktionen orthogonal so wählen kann, daß sie Paare von konjugierten Funktionen bilden, verschwindet in Folge von (13) der Mittelwert eines imaginären Operators für die Gesamtheit aller Zustände gleicher Energie.

8. Infolge der Gleichung $K^2 = -1$ sind die Verhältnisse bei ungerader Elektronenzahl auch in Bezug auf Entartung wesentlich anders als bei gerader Elektronenzahl. Vor allem sieht man sofort, daß, wenn auch keine räumliche Symmetrie vorhanden ist, immer eine doppelte Entartung auftritt, die zwei orthogonalen Funktionen φ und $K\varphi$ gehören immer zu demselben Eigenwert. Die Matrixelemente reeller Operatoren sind für konjugierte Funktionen konjugiert komplex

$$(14) \qquad (\psi_i, A\psi_j) = (KA\psi_j, K\psi_i) = (K\psi_i, AK\psi_j)^*.$$

Auch bei Vorhandensein räumlicher Symmetrie verändern sich die Verhältnisse weitgehend, und wir müssen die drei Fälle von 5. wieder gesondert betrachten.

a) Ist die Darstellung des betrachteten Terms K reell $D = D^*$, so folgt aus (5)

$$KO_R\psi_\varkappa = O_R K\psi_\varkappa = \sum_{\lambda=1}^{l} D(R)_{\lambda\varkappa} K\psi_\lambda, \tag{15}$$

daß auch $K\psi_\lambda$ zur λ-Zeile von D gehört. Mithin bewirkt das Symmetrieelement K eine zusätzliche doppelte Entartung, reelle Darstellungen treten immer doppelt auf.

Für Operatoren, die die räumliche Symmetrie des Problems besitzen, läßt sich aus den üblichen Formeln der Darstellungstheorie für Matrixelemente durch Kombination mit (13) bis (15) noch einiges gewinnen. Insbesondere gilt für einen „symmetrischen" reellen hermiteischen Operator, daß die Matrixelemente alle verschwinden, die verschiedene Eigenfunktionen derselben Energie verbinden; die Erwartungswerte für die einzelnen Eigenfunktionen sind dagegen gleich. Innerhalb eines Eigenwerts verhält sich daher die reduzible Darstellung so, wie wenn sie irreduzibel wäre.

b) Die Verhältnisse liegen am einfachsten, für Eigenwerte, deren Darstellung der konjugiert komplexen Darstellung nicht äquivalent ist. Hier gelten die Betrachtungen von 5. wörtlich.

c) Wenn D und D^* äquivalent sind, so lassen sie sich durch ein unitäres S ineinander transformieren. Es sei

$$SD(R)S^{-1} = D(R)^*, \quad S^*D(R)^*S^{*-1} = D(R). \tag{16}$$

Es folgt, daß

$$S^*SD(R)S^{-1}S^{*-1} = S^*D(R)^*S^{*-1} = D(R), \tag{17}$$

S^*S mit allen $D(R)$ vertauschbar, also ein Vielfaches der Einheitsmatrix ist. Somit ist $S = cS'$ und hieraus erhält man auch $S' = cS$, so daß $c = \pm 1$ ist. In unserem Falle ist $c = -1$, da für $c = +1$ die Darstellung rein reell gemacht werden kann. Es gilt somit

$$S' = -S. \tag{18}$$

Daraus folgt auch, daß die Darstellungen der Art c) alle gerade Dimensionszahl haben. Transformiert man noch die Darstellung $D(R)$ mit einer unitären Matrix U, so geht die Matrix S, die sie in die konjugiert komplexe Form überführt in $U'SU$ über. Durch zweckmäßige Wahl der Form von $D(R)$ kann man es also immer erreichen, daß das S in (18) die Form

$$S = \begin{pmatrix} & \vdots & \vdots & \vdots & \vdots & \cdots \\ \cdots & 0 & 0 & 0 & i & \cdots \\ \cdots & 0 & 0 & -i & 0 & \cdots \\ \cdots & 0 & i & 0 & 0 & \cdots \\ \cdots & -i & 0 & 0 & 0 & \cdots \\ \cdots & \vdots & \vdots & \vdots & \vdots & \end{pmatrix} \tag{18a}$$

habe[9]). Es ist zweckmäßig, dies von vornherein anzunehmen, die halbzähligen Darstellungen der dreidimensionalen Drehgruppe (die in diese Kategorie gehören) werden gewöhnlich von vornherein in einer Form angenommen, für die S die Form (18a) hat. Weiterhin empfiehlt es sich, in Analogie zu diesen halbzähligen Darstellungen, die Zeilen und Spalten von D und S (und demgemäß auch die Eigenfunktionen) anstatt mit $1, 2, \ldots$ mit $-j, -j+1, \ldots, j-1, j$ zu numerieren, wobei j halbzählig und $2j+1$ die Dimension der Darstellung ist. Aus (16) wird dann[10])

$$D(R)^*_{\mu\nu} = i^{2\mu-2\nu} D(R)_{-\mu-\nu} = (-1)^{\mu-\nu} D(R)_{-\mu-\nu}. \tag{16a}$$

Durch Anwendung von K auf die Transformationsformel (5) der ψ_μ sieht man mit Hilfe von (18a) leicht, daß $i^{2\mu} K\psi_\mu$ zur $-\mu$-Zeile von $D(R)$ gehört

$$O_R \psi_\mu = \sum_\nu D(R)_{\nu\mu} \psi_\nu, \qquad O_R i^{2\mu} K\psi_\mu = \sum_\nu D(R)_{-\nu-\mu} i^{2\nu} K\psi_\nu. \tag{19}$$

Betrachten wir daher die Funktionen u bzw. v

$$u_\mu = \psi_\mu + i^{-2\mu} K\psi_{-\mu} \quad \text{bzw.} \quad v_\mu = i(\psi_\mu - i^{-2\mu} K\psi_{-\mu}), \tag{20}$$

so sehen wir, daß sie sich sowohl bei räumlichen Transformationen R, wie auch bei der Zeitumkehr K unter sich transformieren, so daß kein Grund für das Zusammenfallen des Eigenwertes der u mit dem Eigenwert der v vorliegt. Daher bedingt für Eigenwerte mit einer Darstellung der Art c) bei ungerader Elektronenzahl keine zusätzliche Entartung. Für die Eigenfunktionen kann von vornherein die sowohl für die u wie für die v geltende Gleichung

$$K\psi_\mu + i^{2\mu} \psi_{-\mu} = 0 \tag{21}$$

angenommen werden.

Betrachten wir, um eine Anwendung dieser Gleichungen zu sehen, die Matrixelemente eines imaginären hermiteischen Operators,

9) Vgl. I. Schur, l. c.

10) Auch die ganzzahligen Darstellungen der dreidimensionalen Drehgruppe, obwohl sie in die Kategorie a) gehören, werden gewöhnlich in solcher Form angenommen, daß für sie (16a) gilt.

die die Eigenfunktionen eines Eigenwerts der Klasse c) verbinden. Es ist

$$(22) \quad (\psi_\mu, B\psi_\nu) = -(BK\psi_\nu, K\psi_\mu) = -(-1)^{\mu-\nu}(\psi_{-\mu}, B\psi_{-\nu})^*.$$

Hat B noch die volle Symmetrie des Problems, so verschwinden alle Matrixelemente (22): für $\mu \neq \nu$ weil dann ψ_μ und ψ_ν zu verschiedenen Zeilen von D gehören, für $\mu = \nu$ dagegen müßten sie alle gleich sein und (22) zeigt, daß sie demzufolge alle rein imaginär sein müßten. Das ist aber nicht möglich, weil sie wegen des hermiteischen Charakters von B reell sein müssen. Es folgt, daß der Erwartungswert eines imaginären hermiteischen Operators, der die volle Symmetrie des Problems besitzt, für alle stationären Zustände verschwindet.

Es ist wohl nicht zu verkennen, daß bei gerader Elektronenzahl der Fall a) dieselbe Rolle spielt wie der Fall c) bei ungerader Elektronenzahl, und umgekehrt. Dies steht wohl damit in Zusammenhang, daß die Darstellung, die Q_R bildet, bei gerader Elektronenzahl zur Klasse a), bei ungerader zur Klasse c) gehört. Wenn auch andere Teilchen als Elektronen im System enthalten sind, sollte dabei eigentlich nicht „gerade Elektronenzahl", sondern „gerade Zahl der Teilchen, deren Spin halbzahlig ist" stehen.

9. Es soll nochmals auf die Verschiedenheit in der Behandlung der „Zeitumkehr" und der gewöhnlichen räumlichen Symmetrien hingewiesen werden. Sie ist dadurch bedingt, daß die Operation, die die Zeit in einer Wellenfunktion umkehrt, nicht linear ist. — Die Voraussetzung für die vorangehenden Überlegungen ist die, daß das ganze Problem invariant sei gegenüber der Transformation $t' = -t$.

Wenn das Problem auch gegenüber der Transformation $t' = -t$ nicht invariant ist, so existieren doch öfters Transformationen, die die Zeitumkehr enthalten. So z. B. besteht bei einem homogenen Magnetfeld das Symmetrieelement: Spiegelung an einer Ebene durch das Magnetfeld mit gleichzeiger Zeitumkehr. Sie bedingt, daß die Eigenfunktionen, wenn man den Faktor $e^{im\varphi}$ abspaltet, reell werden.

Die Wirkung der Zeitumkehr ist bei den beiden Betrachtungsarten des Abschnittes 1 sehr verschieden. Bei der zweiten Betrachtungsart (Gesamtimpuls Null) besagt sie z. B. bei gerader Teilchenzahl, daß Terme ohne Entartung existieren, die sich besonders einfach verhalten. So z. B. ist ihr mittleres magnetisches Moment in jeder Richtung exakt Null. Dies folgt natürlich nicht aus dem Verschwinden des Gesamtdrehimpulses, da man ja nicht

nur eine Art Teilchen hat. Bei ungerader Teilchenzahl existieren dagegen keine unentarteten Zustände.

Bei der ersten Betrachtungsart (festgehaltene Kerne) ergeben sich je nach der Symmetrie verschiedene Resultate, so z. B. ist der mittlere Drehimpuls der Elektronen bei gänzlich unsymmetrischer Konfiguration der Kerne bei gerader Elektronenzahl immer Null. Dies ist nach VAN VLECK für den Diamagnetismus der meisten Stoffe verantwortlich, weil das von der Kernbewegung herrührende magnetische Moment sehr klein ist.

Über die experimentelle Prüfung des Spinerhaltungssatzes

R. Döpel, K. Gailer und E. P. Wigner

Physikalische Zeitschrift *35*, 336–337 (1934)

1. Die Diskussion über eine vor kurzem erschienene Arbeit mit diesem Titel[1]) hat einige Ergebnisse gehabt, die wir für erwähnenswert halten und daher hier mitteilen wollen.

Wir möchten vor allem eine mehr theoretische Bemerkung einschalten. Die Erhaltungssätze der Quantenmechanik zerfallen in zwei Gruppen. In die erste Gruppe gehört der Energiesatz, der lineare Impulssatz und der Drehimpulssatz (sowie ein Satz über die Erhaltung des Spiegelungscharakters, der jedoch bei dem Stoß nicht geprüft werden kann). Eine Gültigkeitsgrenze für diese Erhaltungssätze ergibt sich aus der (gegenwärtigen) Quantenmechanik nicht, sie sollen ebenso wie in der klassischen Theorie unter allen Umständen gelten. In die zweite Gruppe der Erhaltungssätze fällt der Erhaltungssatz für

1) R. Döpel u. K. Gailer, Physik. Zeitschr. 34, 827, 1933.

den gesamten Elektronenspin (auch ein Erhaltungssatz für den gesamten Protonenspin usw.). Diese sollten in der Quantenmechanik nur insofern gelten, als man „die Wechselwirkung der Drehmomente der Elektronen usw. vernachlässigen kann“ [1]).

Der Spinerhaltungssatz wurde bisher bei zwei verschiedenen Arten von Zusammenstößen geprüft. Erstens bei solchen, bei denen die Energie der Umwandlung von vornherein als Anregungsenergie in den Atomen bzw. Molekülen vorhanden war, und zweitens in Fällen, wo sie durch die kinetische Energie der Partner geliefert wurde.

2. In die erste Gruppe kann die bekannte Stabilität des Parawasserstoffes, die Versuche von R. W. Wood und F. W. Loomis über die Jodfluoreszenz [2]), das Studium der Reaktion ${}^3Kr + {}^1Hg = {}^1Kr + {}^zHg$ (die links angefügten Zahlen deuten das Multiplettsystem des betreffenden Atomzustandes an) durch H. Beutler und W. Eisenschimmel [3]) gerechnet werden. Alle drei Versuche bestätigen den Spinerhaltungssatz in den bezüglichen Fällen – die schönen Versuche von Beutler und Eisenschimmel allerdings nur in statistischem Sinne. Dies wurde aber bei den schweren Elementen *Kr* und *Hg*, wo die Spin-Bahnwechselwirkung nicht mehr klein ist, auch nicht anders erwartet [4]).

3. Der Fall, daß die kinetische Energie der Stoßpartner die Umwandlungsenergie liefert, unterscheidet sich scheinbar theoretisch nicht von dem vorangehenden Fall. Trotzdem ist es natürlich von vornherein möglich, daß der Spinerhaltungssatz hier seine Gültigkeit verliert.

K. F. Bonhoeffer und P. Harteck [5]) haben Parawasserstoff im Entladungsrohr zum Leuchten gebracht und gefunden, daß er durch den anregenden Elektronenstoß nicht in Orthowasserstoff umgewandelt wird. Der Spinerhaltungssatz scheint also in diesem Fall bestätigt zu sein.

In der obengenannten Arbeit [6]) wurde das Leuchten von *Hg* untersucht, das von neutralen *He*-Kanalstrahlatomen angeregt wurde. Die Untersuchung war zunächst auf den Nachweis eines Triplettzustandes eingestellt, und ein solcher ($2\,{}^3S_1$) wurde durch das Vorhandensein mehrerer Linien nachgewiesen. (Das Vorhandensein noch eines Triplettzustandes $3\,{}^3D$ ist sehr wahrscheinlich.) Da gleichzeitig überhaupt keine *He*-Linien beobachtet wurden, wurde angenommen, daß die bombardierenden *He*-Atome im Grundzustand bleiben. Es liegt daher eine Stoßreaktion vor:

$$ {}^1He + {}^1Hg = {}^1He + {}^3Hg. \qquad (1) $$

Da die ganze linke Seite von (1) im Singulettzustand, die rechte dagegen, als Ganzes betrachtet, im Triplettzustand ist, liegt hier scheinbar eine Verletzung des Spinerhaltungssatzes vor.

Nun geht es aus den vorangehenden theoretischen Überlegungen hervor, daß eine strenge Gültigkeit des Spinerhaltungssatzes bei *Hg* gar nicht zu erwarten ist, weil die Kräfte, die von den magnetischen Momenten der Elektronen ausgehen, in diesem Fall bekanntlich nicht mehr klein sind. Die sogenannten Singuletterme sind keine wirklichen Singuletterme, sondern nur Terme, bei denen die Spin der beiden Valenzelektronen mit größerer Wahrscheinlichkeit antiparallel sind als parallel, und ähnlich verhält es sich mit den Triplettermen. Deshalb kann das Auftreten des Triplettzustandes $2\,{}^3S_1$ auch so interpretiert werden, daß bei diesem Term die individuellen, für die Anregung mit *He*-Atomen günstigen Eigenschaften des betreffenden Triplettzustandes seinen Triplettcharakter überkompensieren. Ein Schluß über die Gültigkeit oder Ungültigkeit des Erhaltungssatzes kann bei *Hg* nur aus einem statistischen Material gewonnen werden, aus dem Vergleich der Anregungen mehrerer Singulett- und Triplettzustände. (Vgl. das über die Arbeit von Beutler und Eisenschimmel Gesagte.)

Hierzu kommt noch ein Einwand, auf den uns Herr R. Ladenburg freundlichst aufmerksam gemacht hat. Es läßt sich die Möglichkeit nicht von der Hand weisen, daß die Atome im Zustand $2\,{}^3S_1$ zum Teil nicht direkt durch Stoß entstanden sind, sondern durch Strahlung von anderen, höher gelegenen Zuständen hervorgingen. Diese könnten entweder selber Singulettzustände sein oder aber hoch angeregte Zustände, bei denen der Singulett- bzw. Triplettcharakter noch weniger eindeutig ist als bei den niedrigeren Termen.

Aus diesen Gründen wurde beschlossen, den Spinerhaltungssatz in einem anderen Falle zu untersuchen, in dem diese Einwände leichter ausgeschaltet werden können. Als dieser wurde der Stoß von neutralem *He* auf *Mg* gewählt, und die diesbezüglichen Versuche wurden von R. Döpel und K. Gailer bereits in Angriff genommen.

(Eingegangen 19. März 1934.)

1) E. Wigner, Nachr. d. Ges. d. Wiss. zu Göttingen, S. 379, 1927.

2) Phys. Rev. **31**, 705, 1928.

3) Zeitschr. f. phys. Chem. (B) **10**, 89, 1930.

4) Wir möchten an dieser Stelle bemerken, daß auch die Versuche von J. H. Lees [Proc. Roy. Soc. London (A) **137**, 173, 1932], die J. H. Lees und H. W. B. Skinner (ebenda, S. 186) als einen Fall der Verletzung des Spinerhaltungssatzes interpretiert haben, neuerdings L. R. Maxwell (Journ. of the Franklin Inst. **214**, 533, 1932) anders gedeutet hat.

5) Zeitschr. f. phys. Chem. (B) **4**, 113, 1929.

6) R. Döpel und K. Gailer, l. c.

Separation of Rotational Coordinates from the Schrödinger Equation for N Particles

J. O. Hirschfelder and E. P. Wigner

Proceedings of the National Academy of Sciences (USA) *21*, 113–119 (1935)

Communicated November 9, 1934

1. *Introduction.*—In quantum mechanics, as well as in classical mechanics, the number of variables involved in the equations of motion may be reduced by utilizing the constants of motion of the system. For a set of N particles moving in a field-free space there are six integrals if N is greater than two. These correspond to the laws of conservation of linear and of angular momentum. It is the purpose of this paper to elimi-

nate the coördinates of rotation and of translation of the total system from the Schrödinger equation and to obtain a new set of differential equations involving only $3N$-6 variables.[1]

The separation of the translation coördinates is simple. The total energy may be decomposed into the sum of the energy due to the motion of the center of mass and to the motion relative to the center of mass. The energy of the motion about the center of mass is obtained by subtracting the operator corresponding to the energy of translation of the center of mass from the total energy operator

$$H\psi = -\frac{h^2}{2}\sum_{n=1}^{N}\sum_{k=1}^{3}\frac{1}{m_n}\frac{\partial^2\psi}{\partial x_{nk^2}} + \frac{h^2}{2M}\sum_{k=1}^{3}\left(\sum_{n=1}^{N}\frac{\partial}{\partial x_{nk}}\right)^2\psi + V\psi = E\psi, \quad (1)$$

where M is the total mass of the system.[1] It may be shown that if ψ is an eigenfunction of H and U is a function of the coördinates of the center of mass, the product $U\psi$ is also an eigenfunction. Therefore equation (1) is really independent of the translational motion of the system. For our purposes, it will be convenient to consider (1) as the Eulerian equation of the variational problem:[2]

$$\delta\int\left\{\frac{h^2}{2}\sum_n\sum_k\left|\frac{\partial\psi}{\partial x_{nk}}\right|^2 - \frac{h^2}{2M}\sum_k\left|\sum_n\frac{\partial\psi}{\partial x_{nk}}\right|^2 + V\,|\,\psi\,|^2\right\}d\tau = 0, \quad (2)$$

and the condition $\int |\psi|^2 d\tau = 1$.

The dependence of the eigenfunctions of H on the rotation of the total system is well known from group theory.[3] The eigenfunctions possess two quantum numbers l and μ, characteristic respectively of the total angular momentum lh and the Z component of angular momentum μh, and may be labeled $\psi^l_\mu(x_1, \ldots, z_N)$. If one defines for every configuration $x_1, \ldots, z_N$ a "standard position" $x_1', \ldots, z_N'$ which is the same for all configurations in which the relative position of the particles is the same, one can write

$$\psi^l_\mu(x_1 \ldots z_N) = \sum_{r=-l}^{l} D^{(l)}(R)^*_{\mu r}\, X_r\,(x_1, \ldots, z_N). \quad (3)$$

Here R is the rotation which brings the standard position $x_1', \ldots, z_N'$ into $x_1, \ldots, z_N$, the $D^{(l)}(R)_{\mu r}$ are the coefficients of the irreducible representations of the three dimensional rotation group. The

$$X_r(x_1, \ldots, z_N) = \psi_r(x_1', \ldots, z_N') \quad (3a)$$

are such functions of the $x_1, \ldots, z_N$ which depend on the relative positions of the particles only, since $x_1', \ldots, z_N'$ is the same set of numbers for all possible orientations of a relative configuration.

Once the standard position has been selected, (3) may be stubstituted into (2) and the integration performed over the rotational coördinates.

2. *Modified Schrödinger Equation in General Form.*—In order to substitute (3) into (2), it is necessary to evaluate the derivatives of the representation coefficients. Letting R and RS^{nk} be the rotation which brings the standard position into $x_{11}, \ldots, x_{N3}$ and $x_{11}, \ldots, x_{nk} + \epsilon, \ldots, x_{N3}$ (x_{nk} is the k-th coördinate of the n-th particle)

$$\frac{\partial D^{(l)}(R)_{\mu r}}{\partial x_{nk}} = \lim_{\epsilon=0} \frac{1}{\epsilon} [D^{(l)}(RS^{nk})_{\mu r} - D^{(l)}(R)_{\mu r}]. \tag{4}$$

Evidently S^{nk} is an infinitesimal rotation and can be decomposed into three infinitesimal rotations about the coördinate axes. The corresponding angles of rotation about the X, Y, Z axes may be denoted respectively by ϵS^{nk}_{32}, ϵS^{nk}_{13}, ϵS^{nk}_{21}. The S^{nk}_{ij} are functions of the coördinates, their forms depending on the choice of the standard position. If the standard position is independent of rotation and of translation of the total system, they satisfy the identities:

$$\sum_n S^{nk}_{ij} = 0 \qquad (i \neq j) \tag{5a}$$

$$\sum_n x_{na} S^{nb}_{ji} - x_{nb} S^{na}_{ji} = R_{ck} \qquad \begin{array}{l} a, b, c \text{ as well as } i, j, k \text{ form} \\ \text{cyclic permutations of } 1, 2, 3. \end{array} \tag{5b}$$

[(5b) will not be used further on.] Using (4) and the explicit form of the representation coefficients we obtain

$$\begin{aligned} \frac{\partial D^{(l)}(R)_{\mu r}}{\partial x_{nk}} &= irS^{nk}_{21} D^{(l)}(R)_{\mu r} - \\ &\frac{1}{2}(S^{nk}_{32} - i S^{nk}_{13}) \sqrt{(l+r)(l-r+1)}\, D^{(l)}(R)_{\mu r-1} + \\ &\frac{1}{2}(S^{nk}_{32} + i S^{nk}_{13}) \sqrt{(l+r+1)(l-r)}\, D^{(l)}(R)_{\mu r+1} \end{aligned} \tag{6}$$

And therefore

$$\frac{\partial \psi^l_\mu}{\partial x_{nk}} = \sum_{-l}^{l} D^{(l)}(R)^*_{\mu r} \zeta^r_{nk} \tag{7a}$$

$$\begin{aligned} \zeta^r_{nk} = \frac{\partial X_r}{\partial x_{nk}} &- ir S^{nk}_{21} X_r - \frac{1}{2}(S^{nk}_{32} + iS^{nk}_{13}) \sqrt{(l+r+1)(l-r)}\, X_{r+1} \\ &+ \frac{1}{2}(S^{nk}_{32} - iS^{nk}_{13}) \sqrt{(l+r)(l-r+1)} X_{r-1}. \end{aligned} \tag{7b}$$

These derivatives may be substituted into the variational principle, (2),

$$\delta \int dR \int d\tau \sum_{rr'} D^{(l)}(R)^*_{\mu r} D^{(l)}(R)_{\mu r'} \left[\frac{h^2}{2} \sum_{\substack{nn' \\ k}} \left(\frac{\delta_{nn'}}{m_n} - \frac{1}{M} \right) \zeta^r_{nk} \zeta^{r'*}_{n'k} + V X_r X^*_{r'} \right] = 0 \tag{8a}$$

Using the orthogonality relations between the representation coefficients,

the integration may be performed over the rotational coördinates, R. One obtains:

$$\delta \int d\tau \sum_r \left[\frac{h^2}{2} \sum_{\substack{nn' \\ k}} \left(\frac{\delta_{nn'}}{m_n} - \frac{1}{M} \right) \zeta^r_{nk} \zeta^{r*}_{n'k} + V | X_r |^2 \right] = 0 \qquad (8b)$$

with the boundary condition

$$\int \sum_r | X_r |^2 \, d\tau = \text{Const.}$$

3. *The Standard Position.*—The next problem is that of defining the standard position for any configuration of the system. The best standard position should make the final equations as simple as possible without disturbing the symmetry between equivalent particles. We investigated several possibilities for the standard position and the most simple results were obtained if the center of gravity was in the origin of the coördinate system and the principal axes of inertia parallel to the coördinate axes. Therefore, we shall define R as the orthogonal matrix which brings:

$$T_{ik} = \sum_n m_n \left(x_{ni} - \sum_{n'} \frac{m_{n'} x_{n'i}}{M} \right) \left(x_{nk} - \sum_{n'} \frac{m_{n'} x_{n'k}}{M} \right) \qquad (9)$$

into the diagonal form

$$\sum_k T_{ik} R_{kp} = t_p R_{ip}. \qquad (10)$$

The R_{ip} and the t_p are functions of the x-coördinates. It will be convenient to define the coördinates, y_{np}, of the particles referred to the principal axes of inertia

$$y_{np} = \sum_k R_{kp} \left(x_{nk} - \sum_{n'} \frac{m_{n'} x_{n'k}}{M} \right) \qquad (11)$$

They satisfy the identities:

$$\sum_n m_n y_{np} y_{np'} = 0 \quad p \neq p' \qquad (12a)$$

$$\sum_n m_n y_{np} = 0. \qquad (12b)$$

For the sake of definiteness we shall take: $t_1 \leq t_2 \leq t_3$.

The x_{nk} can be completely replaced by the three coördinates of the center of mass, the R_{ip} (three independent variables) and the y_{np} ($3N$-6 independent variables). Of course, the wave function is defined only for such y_{np} which satisfy the six identities (12). However, in order to preserve the symmetry, it is advantageous to define the X_r also for such y_{np} which do not satisfy the identities by assuming that they depend only on the relative configuration in the y-space. These generalized wave

functions might still depend on the R's and the coördinates of the center of gravity, but they do not, since they are functions only of the relative configuration in x-space.

It is easy to calculate the $S^{nk}_{pp'}$ matrix by the Rayleigh-Schrödinger perturbation method. The R_{i1}, R_{i2}, R_{i3} are the three characteristic vectors of the T_{ij} for the original configuration $x_{11} \ldots x_{N3}$. If x_{nk} is changed by the small quantity ϵ, the $S^{nk}_{pp'}\epsilon$ are the coefficients of the original R_{ip} in the p'-th characteristic vector of the "perturbed" T_{ik}. The formulas of the Rayleigh-Schrödinger theory give

$$S^{nk}_{ij} = \frac{m_n}{t_j - t_i}(R_{kj}\, y_{ni} + R_{ki}\, y_{nj}) \tag{14a}$$

and

$$\frac{\partial t_p}{\partial x_{nk}} = 2m_n\, R_{kp}\, y_{np}. \tag{14b}$$

From (14) it follows that

$$\frac{\partial X_r}{\partial x_{nk}} = \sum_{n'p} R_{kp}\left(\delta_{nn'} - \frac{m_n}{M}\right)\frac{\partial X_r}{\partial y_{n'p}} - \sum_p S^{nk}_{p+2\,p+1}\, M_p X_r \tag{15a}$$

Where

$$M_c = \sum_n \left(y_{na}\frac{\partial}{\partial y_{nb}} - y_{nb}\frac{\partial}{\partial y_{na}}\right) \quad \begin{array}{l} a,\ b,\ c \text{ are cyclical per-} \\ \text{mutations of } 1,\ 2,\ 3. \end{array} \tag{15b}$$

Inserting this into the variational principle, (8), we obtain an explicit form:

$$\delta \sum_r \int \left\{ \sum_{\substack{n'n \\ p}} \frac{h^2}{2}\left(\frac{\delta_{nn'}}{m_n} - \frac{1}{M}\right)\frac{\partial X_r}{\partial y_{np}}\frac{\partial X_r^*}{\partial y_{n'p}} + V|X_r|^2 + \right.$$
$$\left. \frac{h^2}{2}\sum_p \frac{t_{p+1} + t_{p+2}}{(t_{p+1} - t_{p+2})^2}|A_p|^2 - \frac{h^2}{2}\sum_p (A_p^* L_p X_r + A_p L_p X_r^*)\right\} d\tau. \tag{16}$$

Here

$$L_c = \frac{1}{t_a - t_b}\sum_n \left(y_{na}\frac{\partial}{\partial y_{nb}} + y_{nb}\frac{\partial}{\partial y_{na}}\right) \tag{16a}$$

$$A_1 = M_1 X_r + \frac{1}{2}\sqrt{(l + r + 1)(l - r)}X_{r+1} - \frac{1}{2}\sqrt{(l + r)(l - r + 1)}X_{r-1}$$

$$A_2 = M_2 X_r + \frac{i}{2}\sqrt{(l + r + 1)(l - r)}X_{r+1} + \frac{i}{2}\sqrt{(l + r)(l - r + 1)}X_{r-1}$$

$$A_3 = M_3 X_r + irX_r. \tag{16b}$$

The Eulerian equations for the vanishing of the first variation of (16) are

$$EX_r = H_0 X_r + G^r_{-2} X_{r-2} + G^r_{-1} X_{r-1} + G^r_0 X_r + G^r_1 X_{r+1} + G^r_2 X_{r+2} + F^r_{-1} X_{r-1} + F^r_0 X_r + F^r_1 X_{r+1} \tag{17}$$

where

$$H_0 = -\sum_{n=1}^{N} \sum_{p=1}^{3} \frac{h^2}{2m_n} \frac{\partial^2}{\partial y^2_{np}} + V. \tag{17a}$$

$$\frac{2}{h^2} G^r_{-2} = -\frac{1}{4} \sqrt{[l(l+1) - (r-1)^2]^2 - (r-1)^2} \left(\frac{t_2 + t_3}{(t_2 - t_3)^2} - \frac{t_1 + t_3}{(t_1 - t_3)^2} \right)$$

$$\frac{2}{h^2} G^r_2 = -\frac{1}{4} \sqrt{[l(l+1) - (r+1)^2] - (r+1)^2} \left(\frac{t_2 + t_3}{(t_2 - t_3)^2} - \frac{t_1 + t_3}{(t_1 - t_3)^2} \right)$$

$$\frac{2}{h^2} G^r_{-1} = -\sqrt{l(l+1) - r(r-1)}\,(L_1 - iL_2) \tag{17b}$$

$$\frac{2}{h^2} G^r_1 = \sqrt{l(l+1) - r(r+1)}\,(L_1 + iL_2)$$

$$\frac{2}{h^2} G^r_0 = 2irL_3 + \frac{t_1 + t_2}{(t_1 - t_2)^2} r^2 + \frac{1}{2}(l(l+1) - r^2) \left(\frac{t_2 + t_3}{(t_2 - t_3)^2} + \frac{t_1 + t_3}{(t_1 - t_3)^2} \right)$$

and

$$\frac{2}{h^2} F^r_{-1} = \sqrt{l(l+1) - r(r-1)} \left[\frac{t_2 + t_3}{(t_2 - t_3)^2} M_1 - i \frac{t_1 + t_3}{(t_1 - t_3)^2} M_2 \right]$$

$$\frac{2}{h^2} F^r_1 = -\sqrt{l(l+1) - r(r+1)} \left[\frac{t_2 + t_3}{(t_2 - t_3)^2} M_1 + i \frac{t_1 + t_3}{(t_1 - t_3)^2} M_2 \right] \tag{17c}$$

$$\frac{2}{h^2} F^r_0 = -2ir \frac{t_1 + t_2}{(t_1 - t_2)^2} M_3 - \sum_p \frac{t_{p+1} + t_{p+2}}{(t_{p+1} - t_{p+2})^2} M_p^2 + 2 \sum_p L_p M_p + \frac{1}{M} \sum_{nn'} \sum_p \frac{\partial^2}{\partial y_{np}\, \partial y_{n'p}}.$$

This is the Schrödinger equation after the elimination of the six integrals of motion. The unknown functions, X_r, in (16a) may be interpreted in two ways. First, if they are considered as functions of such variables, y_{np}, which satisfy the identities (12), it can be shown that the X_r need to be defined only in the quasi-hyperboloidal regions satisfying the identities.[4] Second, they may be regarded as functions of unrestricted y's, but they must be of such a form that they are invariant under all translations and rotations. In the second case, however, the t_p must be considered as the characteristic values of the T_{ik} formed from the y's (not from the x's) and this would make the actual calculations impractical since the characteristic values of a three rowed matrix are not easily

obtainable. The terms with the F in (17) can be omitted for both interpretations, if the X_r are expressed explicitly as such functions of the y, which depend only on the relative configuration.

It will be noticed that whenever any two of the t_p become equal, the principal axes of inertia as well as the standard position become indetermined. This manifests itself in the equations by producing singularities of the first and second order for these configurations. As a result, the wave functions will not be continuous as t_p and $t_{p'}$ approach equality from the two sides $t_p > t_{p'}$ and $t_{p'} > t_p$. However, a detailed discussion of this behavior will not be given in this paper.

[1] For application to the He-atom see G. Breit, *Phys. Rev.*, **35**, 569 (1930). In a paper which appeared a short time ago (*Phys. Rev.*, **46**, 383 (1934)) C. Eckart developed the theory of rotation and inner motion of a system of particles in classical dynamics. His results are such, however, that one can deduce from his eq. (26) a quantum mechanical equation, which must be equivalent to our (17), except for the use of different coördinates. It does not seem to be quite easy, however, to specify Eckart's coördinates in such a way as to conserve the symmetry between identical particles. The elimination of the translational coördinates and applications of the resulting equation were given already by D. S. Hughes and C. Eckart, *Phys. Rev.*, **36**, 694 (1930).

[2] The system must be considered as being in a finite box in order to make the integrals finite. However, the box may be made sufficiently large so as to have no effect on the system.

[3] See any book on the applications of group theory to quantum mechanics, e.g., H. Weyl, *Theory of Groups and Quantum Mechanics*, or E. Wigner, *Gruppentheorie*, etc. Formulas (3) and the expressions for the $D^{(l)}(R)_{\mu r}$ may be found in the latter one in chapters XIX and XV, respectively.

[4] This can be proved directly by substituting $f(I_1, \ldots, I_6)X_r$ for the X_r in (17), where $I_1, \ldots, I_6$ are the expressions on the left side of the identities. The new X's will still satisfy (17) for all points $I_1 = \ldots = I_6 = 0$.

Bibliography of Papers on Applied Group Theory 1926–1935

1926

Über nicht kombinierende Terme in der neueren Quantentheorie. Erster Teil. Zeitschrift für Physik *40*, 492–500 (1926)

1927

Über nicht kombinierende Terme in der neueren Quantentheorie. Zweiter Teil. Zeitschrift für Physik *40*, 883–892 (1927)

Einige Folgerungen aus der Schrödingerschen Theorie für die Termstrukturen. Zeitschrift für Physik *43*, 624–652 (1927)

Berichtigung zu der Arbeit: Einige Folgerungen aus der Schrödingerschen Theorie für die Termstrukturen. Zeitschrift für Physik *45*, 601–602 (1927)

Translated into Hungarian: A Schrödinger-Féle Elméletből A Termszerkezetekre Vonatkozóan Adódó Néhány Következtetés. Magyar Fizikai Folyoirat *12*, 85–111 (1964)

Translated into English: Some Consequences of the Schrödinger Theory for Eigenvalues. In: Quantum Theory of Angular Momentum, L.C. Biedenharn and H. van Dam (Eds.). Academic Press, New York 1965, pp. 37–67

Über die Erhaltungssätze in der Quantenmechanik. Nachrichten der Gesellschaft der Wissenschaften zu Göttingen Mathematisch-Physikalische Klasse 375–381 (1927)

1928

(With J. von Neumann) Zur Erklärung einiger Eigenschaften der Spektren aus der Quantenmechanik des Drehelektrons. Zeitschrift für Physik *47*, 203–220 (1928)

Translated into Hungarian: A Színképek Néhány Tulajdonságának Értelmezéséröe a Pörgő Elektron Kvantummechanikája Alapján. Magyar Fizikai Folyóirat *18*, 137–152 (1970)

(With P. Jordan) Über das Paulische Äquivalenzverbot. Zeitschrift für Physik *47*, 631–651 (1928)

(With J. von Neumann) Zur Erklärung einiger Eigenschaften der Spektren aus der Quantenmechanik des Drehelektrons. Zweiter Teil. Zeitschrift für Physik *49*, 73–94 (1928)

(With J. von Neumann) Zur Erklärung einiger Eigenschaften der Spektren aus der Quantenmechanik des Drehelektrons. Dritter Teil. Zeitschrift für Physik *51*, 844–858 (1928)

(With E. E. Witmer) Über die Struktur der zweiatomigen Molekülspektren nach der Quantenmechanik. Zeitschrift für Physik *51*, 859–886 (1928)

1930

Über die elastischen Eigenschwingungen symmetrischer Systeme. Nachrichten der Gesellschaft der Wissenschaften zu Göttingen Mathematisch-Physikalische Klasse 133–146 (1930)

Translated into English: Elastic Eigenvibrations of Symmetric Systems. In: Group Theory and Solid State Physics. A Selection of Papers, Vol. I, P. H. Meijer (ed.). Gordon and Breach, New York 1964, pp. 247–263

1931

Über eine Verschärfung des Summensatzes. Physikalische Zeitschrift *32*, 450–453 (1931)

1932

Über die Operation der Zeitumkehr in der Quantenmechanik. Nachrichten der Gesellschaft der Wissenschaften zu Göttingen Mathematisch-Physikalische Klasse 546–559 (1932)

Translated into Hungarian: Az Időtükrözés a Kvantummechanikában. Magyar Fizikai Folyóirat *12*, 213–223 (1964)

Translated into English: The Operation of Time Reversal in Quantum Mechanics. In: Group Theory and Solid State Physics. A Selection of Papers, Vol. I, P. H. Meijer (ed.). Gordon and Breach, New York 1964, pp. 265–278

1934

(With R. Döpel and K. Gailer) Über die experimentelle Prüfung des Spinerhaltungssatzes. Physikalische Zeitschrift *35*, 336–337 (1934)

1935

(With J. O. Hirschfelder) Separation of Rotational Coordinates from the Schrödinger Equation for N Particles. Proceedings of the National Academy of Sciences (USA) *21*, 113–119 (1935)

Related but not reprinted in the Collected Works

Gruppentheorie und ihre Anwendung auf die Quantenmechanik der Atomspektren. Vieweg, Braunschweig 1931. Reprinted by Edwards Brothers, Ann Arbor 1944

Translated into English by James J. Griffiths: Group Theory and its Application to the Quantum Mechanics of Atomic Spectra. Academic Press, New York 1959

PART III

The Mathematical Papers

The Mathematical Papers

Annotation by George W. Mackey

Eugene Wigner is above all a theoretical physicist. However he was one of the two men (Hermann Weyl was the other) who introduced a powerful new mathematical tool into quantum mechanics in its earliest years. This is the theory of group representations, invented by Frobenius in 1896, and apparently not applied outside of pure group theory until E. Artin's startling application to number theory in 1923. Wigner's first application of this theory to quantum mechanics was published only four years later in 1927. Weyl's contribution was of a completely different character and was made a few months after Wigner's.

Wigner's involvement with this new mathematical theory led him to participate in its development and thus to work which can be considered as contributions to mathematics as well as to physics – some of it indeed of considerable mathematical importance. Moreover Wigner was by no means immune to the lure of mathematics for its own sake. When his work suggested interesting generalizations or questions, he sometimes pursued these even when their importance for physics was hard to see.

Although it is certainly the best known, Wigner's work on group representations accounts for only about half of his papers with serious mathematical content. As a physicist, the single largest component of his voluminous output is his work on nuclear physics. He of course applied group representations there as well as to other branches of physics; above all to nuclear structure. However, his work on nuclear reactions led him in quite a different mathematical direction – specifically to a series of papers on a class of meromorphic functions which he called R-functions. These are meromorphic functions (analytic functions of one complex variable whose only singularities are poles) which map the upper and lower half planes into themselves and have (necessarily) all their singularities on the real axis. Wigner concerned himself chiefly with those R-functions for which there are an infinite number of poles and in which both the pole locations and their residues have certain statistical properties – e.g., well-defined densities and averages. The aspect of his work which led to the study of R-functions also led to several papers on probability theory.

Together the papers on group representations, other group theory, R-functions and probability account for about three-quarters of Wigner's contributions to mathematics. The remainder are of a miscellaneous character with only one or two papers in each category; abstract operator theory, non-associative algebra, perturbation theory, etc. All of these contributions will be described in greater detail in the body of this essay.

Of course, the distinction between a mathematical contribution and a contribution to physics is far from clear cut. An ostensible contribution to physics may be found, on close examination, to rest on a not very clearly formulated mathematical theorem of some novelty and of independent mathematical interest. Contrariwise an interesting mathematical analysis may have been suggested by a real physical problem but be interesting only under conditions not realized in the world; e.g., scattering theory with unrealistic potentials. Finally a very common type of paper in physics involves a well-defined mathematical model and a well-defined question about it which is much too difficult to settle by available mathematical techniques. Suppose that the author makes a number of approximations and qualitative assumptions of uncertain validity and by a non-air-tight chain of reasoning is led to a definite conclusion about the answer. Most mathematicians would not count this work as a contribution to mathematics. In some cases, however, a close examination may reveal that the author, somewhere in the course of his reasoning, has suggested a proof of a new and interesting mathematical theorem which he, however, leaves unformulated – perhaps even in his own mind. Should one consider this a contribution to mathematics?

Because of such difficulties it has not been easy to decide which of Wigner's papers to discuss in this essay. Despairing of finding a clear and defensible criterion I have done more or less what I felt like doing and I am sure that some readers will object to my selections. In the same spirit I have not hesitated to comment on papers that are part of the task of one of my half dozen or so collaborators on this project.

A. Wigner's Papers on Group Representations

The basic mathematical theorem underlying Wigner's discovery that the theory of group representations can be useful in quantum mechanics can be understood as a simple (but not obvious) generalization of an elementary fact about eigenvalues of commuting linear operators. Let U be a linear operator and let $\varphi_1, \varphi_2, \ldots$ be a basis for the underlying vector space consisting entirely of eigenvectors of U, $U(\varphi_j) = \lambda_j \varphi_j$. Let H be any linear operator in the same space which commutes with U $(HU = UH)$. Then $U(H(\varphi_j)) = H(U(\varphi_j)) = H(\lambda_j \varphi_j) = \lambda_j H(\varphi_j)$. In other words, H applied to any φ_j is again an eigenvector of U and one having the same eigenvalue λ_j. Let M_{λ_j} denote the linear span of all the φ_k such that $\lambda_k = \lambda_j$. Then M_{λ_j} contains all λ_j eigenvectors of U and the above argument tells us that H takes every element in M_{λ_j} back into an element in M_{λ_j} and so defines a linear operator H_{λ_j} in the subspace M_{λ_j}. Evidently in order to find a basis of eigenvectors for H, it suffices to find such a basis for each H_{λ_j}. In other words, a diagonalization of an operator U which commutes with an operator H provides a partial diagonalization of H.

Suppose now that we have several operators $U_1, U_2, \ldots$, all of which commute with H and each of which has a basis of eigenvectors. If the U_j commute among themselves, the argument of the preceding paragraph can be used to

find a basis $\varphi_1, \varphi_2, \ldots$ whose elements are simultaneous eigenvectors of all of the U_j, $U_j(\varphi_k) = \lambda_k^j \varphi_k$ and each sequence $\lambda_k^1, \lambda_k^2 \ldots$ can be considered a joint eigenvalue of $U_1, U_2, \ldots$. Let $\mathcal{M}_{\lambda_k^1, \lambda_k^2, \ldots}$ be the subspace of all φ such that $U_j(\varphi) = \lambda_k^j(\varphi)$ for all j. Then H carries each element of each $\mathcal{M}_{\lambda_k^1, \lambda_k^2, \ldots}$ into itself and, just as above, the joint eigenspaces of $U_1, U_2, \ldots$ define a partial diagonalization of H – a diagonalization that becomes less partial and more complete the more U_j one can find.

The theory of group representations enters in when one tries to extend the above results to operators U_j which commute with H but do *not* necessarily commute among themselves. One no longer need have joint eigenspaces, but it is hard to believe that some synthesis of the partial diagonalizations provided by the individual U_j does not exist. To find it one specializes to the case in which all the U_j are invertible and considers the set $\widetilde{G}$ of all products of the U_j and their inverses. It is evident that all members of $\widetilde{G}$ also commute with H, so that we might as well assume at the outset that the set of all U_j is closed with respect to the formation of products and inverses and hence is a *group of operators*. Let G be the corresponding abstract group and let U, $x \to U_x$ denote an isomorphism of G with $\widetilde{G}$. Then this mapping U of G into the invertible linear operators in our vector space is a *representation* of G in the sense of that $U_{xy} = U_x U_y$ and $U_{x^{-1}} = (U_x)^{-1}$ for all x and y in G. This particular representation has the additional property of being "faithful". This means that $U_x = U_y$ implies that $x = y$. However, since the analysis of faithful representations leads to the study of those that are not faithful it is convenient to study both kinds together. Indeed one is rather seldom interested in whether a given representation is faithful or not.

Our problem has now been replaced by the following. We have a group G, a representation U of G in a vector space $\mathcal{H}(U)$ and a linear operator H such that $U_x H = H U_x$ for all x in G. We wish to find a partial diagonalization of H generalizing that described above when the U_x commute with one another. We shall find it, almost made to order, in the general theory of the equivalence and reducibility of group representations.

This theory was first developed in the special case in which G is a finite group and the representation space $\mathcal{H}(U)$ is a finite-dimensional vector space with the complex numbers as scalars. We shall explain it first in that case and introduce generalizations later. If U is such a representation of a finite group G, it is said to be *reducible* if the vector space $\mathcal{H}(U)$ admits a proper subspace $\mathcal{M}$ (i.e., $\neq 0$ and $\mathcal{H}(U)$) which is *invariant* in the sense that $U_x(\varphi)$ is in $\mathcal{M}$ whenever φ is in $\mathcal{M}$ and x is in G. Every invariant subspace defines a new representation whose space is $\mathcal{M}$ by simply restricting all U_x to $\mathcal{M}$ and ignoring what they do outside of $\mathcal{M}$. This new representation whose space is $\mathcal{M}$ is called the *subrepresentation* defined by $\mathcal{M}$. A representation U is said to be *irreducible* if there are no proper invariant subspaces and hence no proper subrepresentation.

Now let $\mathcal{M}_1$ and $\mathcal{M}_2$ be two invariant subspaces of $\mathcal{H}(U)$. These subspaces are said to be complementary if they have only zero in common and every

vector in $\mathcal{H}(U)$ is the sum of a vector in $\mathcal{M}_1$ and a vector in $\mathcal{M}_2$. If $\mathcal{M}_1$ and $\mathcal{M}_2$ are complementary, then every element of $\mathcal{H}(U)$ is *uniquely* a sum of an element in $\mathcal{M}_1$ and an element in $\mathcal{M}_2$ and one sees that U itself is completely described by giving the subrepresentations $U^{\mathcal{M}_1}$ and $U^{\mathcal{M}_2}$ defined by $\mathcal{M}_1$ and $\mathcal{M}_2$. One says that U is the *direct sum* of the representations $U^{\mathcal{M}_1}$ and $U^{\mathcal{M}_2}$. It turns out to be possible to prove that *every* proper invariant subspace $\mathcal{M}$ has a complement so that every representation is either irreducible or the direct sum of two proper subrepresentations. Of course, one can look for subrepresentations of component representations and in a finite number of steps obtain a decomposition of every U as a direct sum of a finite number of *irreducible* subrepresentations. Finding such a deccomposition is an exact analogue of finding a basis of eigenvectors for a single operator. In neither case is the decomposition unique. However, in the operator case the eigenvalues and their multiplicities of occurrence are uniquely determined. Moreover the linear span of these basis vectors with a common eigenvalue is just the total eigenspace for that eigenvalue and is uniquely determined. The decomposition as a direct sum of eigenspaces is unique. The non-uniqueness is due to the fact that an *arbitrary* basis in each eigenspace is a basis of eigenvectors.

This partial uniqueness has an analogue in the decomposition of group representations. It involves finding an analogue for equality of eigenvalues. What are we to understand by "equality" for two group representations? We have defined group representation in a rather abstract way but if we introduce a basis $\varphi_1, \ldots, \varphi_n$ in $\mathcal{H}(U)$ every operator U_x has a matrix with respect to $\varphi_1, \ldots, \varphi_n$ and our abstract representation becomes a representation by matrices. It is obvious what it means to say that two matrix representations are equal and natural to define two abstract representations to be *equivalent* if bases can be *chosen* so that the corresponding matrix representations are equal. In more abstract terms one defines U^1 and U^2 to be *equivalent* representations of the group G if there exists a one-to-one linear transformation T of $\mathcal{H}(U)$ onto $\mathcal{H}(U)$ such that $TU_x^1T^{-1} = U_x^2$ for all x in G.

Now let U be a representation of G and let $\mathcal{H}(U)$ be a direct sum of invariant subspaces $\mathcal{M}_1 \oplus \mathcal{M}_2 + \ldots \oplus \mathcal{M}_l$ which define irreducible subrepresentations. For each irreducible representation L of G, which is equivalent to one of these subrepresentations, let $\mathcal{M}_L$ be the direct sum of all the $\mathcal{M}_j$ for which L is equivalent to the subrepresentation defined by $\mathcal{M}_j$. It is easy to prove that the subspace $\mathcal{M}_L$ contains every invariant subspace of $\mathcal{H}(U)$ whose corresponding subrepresentation is equivalent to L and no other irreducible subspaces. Hence it is uniquely determined by the equivalence class of L. It is the invariant subspaces $\mathcal{M}_L$ which are the analogues of the eigenspaces of a single operator. $\mathcal{H}(U)$ is a direct sum of the $\mathcal{M}_L$ and this decomposition is unique. The further decomposition of the $\mathcal{M}_L$ into irreducible invariant subspaces which define mutually equivalent irreducible sub-representations is not unique. The decomposition of U as a direct sum of the $U^{\mathcal{M}_L}$ is called the *canonical decomposition into primary representations*, where we define a representation to be primary if all its irreducible subrepresentations are equivalent.

We may now state the fundamental theorem which underlies Wigner's applications of the theory of group representations to the quantum mechanics of the atom in 1927 and 1928.

Theorem 1. *Let U be a representation of group G which can be decomposed as a direct sum of irreducible finite dimensional subrepresentations and let $\mathcal{H}(U) = \mathcal{M}_{L^1} \oplus \mathcal{M}_{L^2} + \ldots$ define the canonical decomposition of U into primary subrepresentations. Then any linear operator H which commutes with all U_x must map each of the $\mathcal{M}_{L^j}$ into itself.*

Corollary. *To find a basis for $\mathcal{H}(U)$ consisting of eigenvectors of H it suffices to find such a basis for the restriction H_{L^j} of H to each $\mathcal{M}_{L^j}$.*

This, however, is not the end of the story. The operators H_{L^j} which remain to be diagonalized also commute with certain representations of G. Clearly H_{L^j} commutes with the primary component of U defined by $\mathcal{M}_{L^j}$ and primary representations, unlike constant operators, are far from trivial. Theorem 1 tells us nothing but perhaps there is some other way of exploiting the commutativity of H_{L^j} with the primary L^j component of U. There is and it depends upon two lemmas whose proofs are trivial.

Lemma 1. *Let $U = L \oplus L + \ldots \oplus L$ be a direct sum of replicas of the same representation L. Then every operator H in $\mathcal{H}(U)$ is completely described by a matrix $\|H_{ij}\|$ of operators in $\mathcal{H}(L)$ and H commutes with all U_x if and only if H_{ij} commutes with L_x for all x and all pairs i, j.*

Lemma 2. *If U is irreducible and finite dimensional then H commutes with all U_x if and only if $H = \lambda I$ for some complex number λ.*

Proof. Since $\mathcal{H}(U)$ is finite dimensional H must have a nonzero eigenvector φ. $H(\varphi) = \lambda\varphi$ for some λ. Let $\mathcal{M}_\lambda$ be the vector space of all λ eigenvectors. If $\psi \in \mathcal{M}_\lambda$, $H(U_x(\psi)) = U_x(H(\psi)) = U_x(\lambda\psi) = \lambda U_x(\psi)$ for all x. Hence $U_x(\psi) \in \mathcal{M}_\lambda$ and $\mathcal{M}_\lambda$ must be a U invariant subspace of $\mathcal{H}(U)$. Since U is irreducible $\mathcal{M}_\lambda = 0$ or $\mathcal{H}(U)$ and $\mathcal{M}_\lambda = 0$ is impossible because φ is in $\mathcal{M}_\lambda$. Thus $H = \lambda I$. □

Corollary (of Proof). *If U is arbitrary and H commutes with all U_x then every eigenspace $\mathcal{M}_\lambda$ of H is a U invariant subspace of $\mathcal{H}(U)$.*

This corollary may be regarded as a sort of dual of Theorem 1 although its proof requires nothing but the definition of a group representation and no decomposition theory.

Let us return now to Theorem 1 and consider the operators in H_j in $\mathcal{M}_{L^j}$. Since U restricted to $\mathcal{M}_{L^j}$ is equivalent to the direct sum of m_j replicas of L^j it follows from Lemma 1 that H_j may be described (in the obvious way) by an $m_j \times m_j$ matrix of operators in $\mathcal{H}(L^j)$ and that each of these is a con-

stant operator by Lemma 2. It is thus clear that to diagonalize H_j it suffices to diagonalize this $m_j \times m_j$ matrix of complex numbers and that each of its eigenvalues appears d_j times in the diagonalization of H_j where d_j is the dimension of the representation L^j. In short, the fact that H_j commutes with a primary representation of dimension $d_j m_j$, where m_j is its multiplicity, tells us that we may replace diagonalizing a matrix with $d_j m_j$ rows and columns by one with m_j rows and columns. In particular, if $m_j = 1$, then H_j is already a constant operator and there is no more work to do.

Theorem 1, and the refinements of it implied by Lemmas 1 and 2, are not only useful in determining the exact eigenvalues of an operator H which commutes with a group representation U. They are also useful in obtaining a first approximation to the eigenvalues of an operator $H + V$ when H is a self adjoint operator whose eigenvalues and eigenvectors are known and V is in some sense "small". In the special case in which H has no multiple eigenvalues the so-called first order perturbation theory provides $\lambda_j + (V(\varphi_j) \cdot \varphi_j)$ as a first approximation to the eigenvalue of $H + V$ which the eigenvalue λ_j of H becomes when H is continuously deformed into it. When H has multiple eigenvalues, however, the perturbation V may cause a fixed multiple eigenvalue λ to split into as many distinct eigenvalues as the dimension of the eigenspace $\mathcal{M}_\lambda$. To determine these one must project V on $\mathcal{M}_\lambda$ and diagonalize the resulting operator V^λ on the space $\mathcal{M}_\lambda$. The eigenvalues into which λ splits will then be $\lambda + \mu_1, \lambda + \mu_2 \ldots \lambda + \mu_k$, where $\mu_1, \mu_2, \ldots, \mu_k$ are the eigenvalues of V^λ. If U is a group representation which commutes with both H and V, $\mathcal{M}_\lambda$ will be a U invariant subspace and the subrepresentation defined by $\mathcal{M}_\lambda$ will commute with V^λ. Thus Theorem 1 and its refinement may be applied to the diagonalization of the V^λ and hence to the determination of the splitting of the eigenvalues of H under the change from H to $H + V$.

The fact that H has known eigenvalues usually arises in quantum physics from the fact that H commutes with a representation U of some group G. It follows from Theorem 1 and its refinement that whenever U has irreducible constituents which are more than one dimensional, then H necessarily has multiple eigenvalues. V will usually not commute with U but will often commute with the restriction of U to some non trivial subgroup G_0 of G. When this is the case one may apply the considerations of the preceding paragraph. When the restricted representation has irreducible constituents which are more than one dimensional V^λ will have multiple eigenvalues itself and the splitting of λ will be incomplete. Of course an irreducible representation of a group G need no longer be irreducible when restricted to a subgroup. Thus in applying the above theory it becomes important not only to know the irreducible representations of the groups which occur but also to know how they decompose when restricted to certain subgroups.

Wigner was led to discover and apply the above principles by his study of the following problem. Let H be a self adjoint operator in a separable Hilbert space $\mathcal{H}$ and let $\varphi_1, \varphi_2, \ldots$ be an orthonormal basis in $\mathcal{H}$ whose elements are eigenvectors of H ($H(\varphi_j) = \lambda_j \varphi_j$). Then the tensor product $\mathcal{H} \otimes \mathcal{H} \ldots \otimes \mathcal{H}$ of

$\mathcal{H}$ with itself n times, which we will denote by $\mathcal{H}^n$, has an orthonormal basis consisting of all possible products $\varphi_{j_1} \times \varphi_{j_2} \ldots \times \varphi_{j_n}$ where $j_1, j_2, \ldots j_n$ are positive integers. Moreover if we define $\widetilde{H}^n$ to be the sum $(H \times I \ldots \times I) + (I \times H \times I \ldots \times I) + \ldots (I \times I \ldots \times H)$ one sees at once that each basis element $\varphi_{j_1} \times \varphi_{j_2} \ldots \times \varphi_{j_n}$ is an eigenvector of $\widetilde{H}^n$ with eigenvalue $\lambda_{j_1} + \lambda_{j_2} \ldots \lambda_{j_n}$. If π is any permutation of the integers $1, 2 \ldots n$ then setting $U_\pi(\varphi_{j_1} \times \varphi_{j_2} \ldots \times \varphi_{j_n}) = \varphi_{j_{\pi(1)}} \ldots \times \varphi_{j_{\pi(n)}}$ defines a permutation of the basis elements of $\mathcal{H}^n$ which, when extended by linearity, defines a unitary operator $\widetilde{U}_\pi$ in $\mathcal{H}^n$. It is clear that $\widetilde{H}^n$ commutes with all the $\widetilde{U}_\pi$. Let V be any self adjoint operator in $\mathcal{H}^n$ which also commutes with all $\widetilde{U}_\pi$ and is small enough so that $\widetilde{H}^n + V$ may be regarded as a "perturbation" of $\widetilde{H}^n$. Wigner's problem was to study the effect of the addition of V to $\widetilde{H}^n$ on the known eigenvalues $\lambda_{j_1} + \lambda_{j_2} \ldots + \lambda_{j_n}$ of $\widetilde{H}^n$. This problem arises in quantum mechanics when n replicas of the same system interact with one another, the most important special case being that of the n electrons in an atom.

Wigner's work on the problem was published in two parts in Volumes 40 (1926) and 43 (1927) of the *Zeitschrift für Physik* [1,2] and was stimulated by a slightly earlier paper of Heisenberg [3] published in Volume 38 of the same journal. When $n = 2$ the situation is rather transparent and one can derive the conclusions of the theory sketched above by simple elementary arguments. Heisenberg did this and then speculated about what one might hope to do along the same lines for larger n. In Wigner's Volume 40 paper he was able to handle the case $n = 3$ quite completely by elementary methods but had to use rather more elaborate algebraic calculations than Heisenberg. He promised to deal with $n > 3$ in Part II. At the beginning of Part II, however, he remarks that the algebra becomes progressively more difficult as n increases but that the work can be vastly simplified by applying the theory of group representations. He credits von Neumann with calling his attention to the existence of this theory and its relevance to his problem. He also credits von Neumann with correctly guessing some of his results. He then goes on to analyze the general case using the principles described in detail above.

As applied to an atom, the group representation which Wigner exploited arises from the identity of all electrons so that permuting them is a "symmetry" of the system. It did not take Wigner long to realize that an atom has other symmetries – notably rotations about the nucleus which lead to exploitable group representations in the same way. Indeed one can regard the "Hamiltonian" self adjoint operator H of the atom as a perturbation $H_0 + V$ of the operator H_0, where H_0 is what the Hamiltonian would be if the electrons did not repel one another but were subject only to the attraction of the nucleus. H_0 turns out to commute with a unitary representation U of the direct product of n replicas of the group R of all rotations and reflections about an origin in 3 space and H itself to commute with the restriction of U to the "diagonal" subgroup of $R \times R \ldots \times R$, i.e., to the subgroup of all $x_1, x_2, \ldots x_n$ with $x_1 = x_2 \ldots = x_n$. Combining the rotational and permutational symmetries one obtains a unitary representation of the natural "semi direct product" of $R \times R \ldots \times R = R^n$ and

the permutation group S^n. (In this semi direct product one defines the product of $x_1, x_2 \ldots x_n, \pi_1 \;\; y_1, y_2 \ldots y_2, \pi_2$ to be $x_1 y_{\pi_2(1)}, x_2 y_{\pi_2(2)} \ldots x_n y_{\pi_2(n)}, \pi_1 \pi_2$ where x_j and y_j are elements of R and π_1 and π_2 are elements of S_n.)

In a long paper [4] published a few months later in Volume 43 of the *Zeitschrift für Physik*, Wigner gave a systematic account of the theory of group representations, followed by a detailed treatment of the extent to which the theory of the energy levels of an atom may be understood by applying this theory in the manner indicated above – bringing both R and S_n into the picture. He warns the reader that his conclusions cannot be expected to agree with experiment in all cases because he has assumed a simplified model of the atom in which the recently discovered "spin" of the electron is neglected. This deficiency was removed the following year (1928) in a three part article by Wigner and von Neumann published in Volumes 45, 47, and 51 of the same journal [5–7].

In the course of the work, certain basic questions about the von Neumann Hilbert space formulation of quantum mechanics are elucidated. In particular it is argued that every symmetry of a quantum mechanical system is implemented, either by a unitary operator or by an anti-unitary operator, in the Hilbert space of states, and pointed out that an operator and any product of it with a complex number of modulus one define the same symmetry. Because of this last fact there is an ambiguity in defining a group of symmetries by a group of operators which forces one to consider so-called "projective" or "ray" representations along with ordinary group representations. This was noted slightly earlier by Weyl in a different context. We shall give a precise definition below in a discussion of Weyl's work. In the present context it plays a key role because the discussion of the quantum mechanics of electron spin demands a consideration of the two-dimensional projective representations of the rotation group. While anti-unitary operators are ruled out in most important cases, e.g., whenever the group in question is such that every element has a "square root" they do occur. Perhaps the most important case is that of "time reversal symmetry". Wigner's work on this subject is discussed in detail in the annotation of Arthur Wightman. It is convenient to distinguish between two rather different ways of applying the theory of group representations to quantum mechanics by the names of the men who discovered them using the terms Wigner's program and Weyl's program. In the end each man contributed to the other's program and the Volume 45 paper is certainly a contribution to Weyl's program.

Wigner's book "Gruppentheorie und ihre Anwendung auf die Quantenmechanik der Atomspektren" [8] was published three years later in 1931. This book is mainly a detailed reworking of the material just described. It is more than that in two important respects. First of all it introduces and studies the concept of "tensor operator" and a theorem about such operators now known as the Wigner-Eckart theorem. Secondly it takes advantage of Slater's important observation that quite substantial simplifications can be made by taking account of the spin of the electron before rather than after considering the consequences of the Pauli exclusion principle. Slater made his observation in an extremely influential paper [9] published in 1929.

Actually Slater's paper went considerably further than this and for two decades convinced most physicists that they could be spared the trouble of learning such an "unphysical" branch of mathematics as group theory. However, before explaining how Slater accomplished this or the nature of tensor operators it will be useful and convenient to go back to 1927 and explain a rather different application of the theory of group representations to quantum physics published by Hermann Weyl in Volume 46 of the *Zeitschrift für Physik* [10].

Weyl's fundamental paper "Quantenmechanik und Gruppentheorie" is not nearly so well known as his comprehensive book "Gruppentheorie und Quantenmechanik" which appeared in 1928 three years before Wigner's. The book contains expositions of the ideas of both Wigner and Weyl and a large part of it is devoted to systematic presentations of background material on linear algebra, group theory, group representations, and quantum mechanics. However, it is the paper which I want to describe here. Weyl's idea differed from that of Wigner in that he wanted to apply group representations to get a better understanding of the foundations of quantum mechanics in general and not so much to gain insight into particular problems. In the introduction to his paper Weyl began by distinguishing two questions in (the foundations of) quantum mechanics.

(1) How does one arrive at the self adjoint operators which correspond to various concrete physical observables?

(2) What is the physical significance of these operators, i.e., how does one deduce physical statements?

He asserts that (1) is the deeper question and has not yet been satisfactorily treated while (2) has been more or less satisfactorily answered by the (now well known) von Neumann formulation of quantum mechanics (which had been published earlier in 1927). He goes on to say that he proposes to treat question (1) with the help of group theory "Hier glaube ich mit Hilfe der Gruppentheorie zu einer tieferen Einsicht in den wahren Sachverhalt gelangt zu sein". He cites Wigner's work of earlier in the year but declares that this connection with group theory lies in quite a different direction.

In more concrete terms Weyl wanted to find a deeper reason for assuming the Heisenberg commutation relations $P_jQ_k - Q_kP_j = \hbar/i\delta_j^k$ than that provided by analogy with the classical Poisson brackets. His first step was to globalize the problem by replacing each of the self adjoint operators P_j, Q_k by the one-parameter group which it generates. If A is a self adjoint operator in a finite dimensional Hilbert space, then for each real t, e^{iAt} may be defined by the infinite series $1 + iAt + \frac{(iAt)^2}{2} + \ldots$ and it is easy to prove the following:

(i) $t \to e^{iAt}$ is a continuous unitary representation of the real line, i.e., of the group formed by all real numbers under addition.

(ii) Every finite dimensional continuous unitary representation of the real line is of the form e^{iAt} for a uniquely determined self adjoint operator A.

(iii) iA is equal to $\frac{d}{dt}(e^{iAt})\big]_{t=0}$ and so may be appropriately referred to as the infinitesimal generator of the group representation $t \to e^{iAt}$.

More or less exact analogues of these theorems are true when the Hilbert space is infinite dimensional but are much more difficult to prove at least when A has a spectrum which is in part continuous. These proofs had not been found when Weyl wrote but he took it for granted that they soon would be and made the simple calculations showing that (on a formal level) the commutation relation $P_jQ_k - Q_kP_j = \frac{\hbar}{i}\delta_j^k$ holds if and only if the two one-parameter subgroups $t \to e^{itP_j}$, $s \to e^{isQ_k}$ obey the commutation relation

$$e^{isP_k}e^{itQ_j} = e^{ist\hbar\delta_j^k}e^{itQ_j}e^{isP_k}$$

for all s and t.

Weyl's next step was to reinterpret his globalized form of the Heisenberg commutation relation by introducing a unitary operator valued function of $2n$ real variables by the definition

$$W_{t_1,t_2\ldots t_n,s_1,s_2,\ldots s_n} = e^{it_1P_1}e^{it_2P_2}\ldots e^{it_nP_n}e^{is_1Q_1}\ldots e^{is_nQ_n}.$$

One readily checks that the P_j and Q_k satisfy the globalized forms of the Heisenberg commutation relations if and only if the function $t_1\ldots s_n \to W_{t_1,t_2\ldots s_1\ldots s_n}$ satisfies the following identity

$$W_{t+t',s+s'} = e^{-i\hbar t'\cdot s}W_{t,s}W_{t',s'}$$

where t, s, t', s' are vectors with n real components and $t'\cdot s$ is the scalar product $t_1's_1 + t_2's_2\ldots t_n's_n$. This identity almost says that the mapping $t, s \to W_{t,s}$ is a unitary representation of the commutative group R^{2n} (under vector addition) of all $2n$-tuples of real numbers. It fails to do so only because the numerical factor $e^{-i\hbar(t_1's_1\ldots t_n's_n)}$ interferes. However, it says something almost equally interesting. It says that $t, s \to W_{t,s}$ *is a projective or ray representation* of the commutative $2n$-dimensional group R^{2n}.

Quite generally if G is a group and $x \to L_x$ is a linear operator valued function on G one says that L is a projective (or ray) representation if $L_{xy} = \sigma(x,y)L_xL_y$ for all x and y where $\sigma(x,y)$ is a complex number depending on x and y. When $\sigma(x,y) \equiv 1$ the definition reduces to that of group representation as given earlier. For finite groups projective representations were studied in some depth by I. Schur in papers published in 1904 and 1907. Weyl points out that projective representations are especially relevant in quantum mechanics because whenever V^1 and V^2 are unitary operators in a Hilbert space $\mathcal{H}$ and $V^2 = e^{i\theta}V^1$ for some real number θ then V^1 and V^2 define *exactly the same* permutations of the pure states of the system. It follows that in dealing with groups of symmetries, each unitary operator defining a symmetry is determined only up to a "phase factor" $e^{i\theta}$ and so the identity $L_{xy} = L_xL_y$ must be relaxed to read $L_{xy} = \sigma(x,y)L_xL_y$ where $\sigma(x,y)$ is a complex number of modulus one.

Weyl regarded it as of fundamental significance that the Heisenberg commutation relations for $P_1\ldots P_n$, $Q_1\ldots Q_n$ are formally equivalent to the assertion that W is a projective unitary representation of the group R^{2n}. This view derives much of its strength from the following two facts:

(a) Any two irreducible projective unitary representations of R^{2n} with

$$\sigma(t_1, t_2 \dots t_n, s_1 \dots s_n, t'_1 \dots t'_n, s'_1 \dots s'_n) = e^{-i\hbar(t'_1 s_1 + \dots t'_n s_n)}$$

are equivalent – in fact unitarily equivalent.

(b) Every projective multiple σ' for R^{2n} which is "non-degenerate" is equivalent to the σ of (a) in the sense that if W is any σ' representation then one can convert W into a σ representation by multiplying each operator $W_{t_1 \dots t_n\, s_1 \dots s_n}$ by a suitable complex number of modulus one.

The above considerations constitute Weyl's contribution to the first of the two problems stated in the introduction to his paper. While more suggestive than persuasive or logically compelling, they are of importance as perhaps the first step in the program of deriving fundamental relationship in quantum mechanics from group-theoretical symmetry principles.

It is interesting to compare the uniqueness theorem stated in (a) above with a theorem added in proof to a paper of Wigner and Jordan [11] also published in 1928. One says that a set of n operators $A_1, A_2, \dots, A_n$ in a Hilbert space $\mathcal{H}$ satisfies the canonical anti-commutation relations if

$$A_i A_j + A_j A_i = 0 \quad \text{and} \quad A_i A_j^* + A_j^* A_i = \delta_i^j$$

for all i and j. The theorem of Wigner and Jordan states that these relations have an irreducible solution which is unique to within equivalence. If one makes the change of variable $\alpha_j = A_j + A_j^*$, $\alpha_{n+j} = \frac{1}{i}(A_j - A_j^*)$ one obtains $2n$ operators which are self adjoint and satisfy $\alpha_j^2 = I$ and $\alpha_j \alpha_k = -\alpha_k \alpha_j$ when $j \neq k$. These operators and $\pm I$ generate a group of 2^{2n+1} elements and Jordan and Wigner show that their theorem is equivalent to showing that (to within equivalence) this group has a unique irreducible representation which is not the identity on the two element center. Although they do not say so, their theorem is also equivalent to a complete analogue of Weyl's theorem in which the group R^{2n} is replaced by the direct product of $2n$ replicas of a group of order 2. This group also has an essentially unique non-degenerate projective multiplier σ and (to within equivalence) a unique irreducible σ representation. The case $n = 2$ is given in Weyl's paper. There is a generalization which includes both theorems as special cases but it was not formulated for two decades and discussion of it will be postponed to later in this essay.

The uniqueness theorem of Wigner and Jordan was much easier to prove than that stated by Weyl because it only involves finite dimensional representations of finite groups. Indeed Weyl offered only a heuristic proof of this theorem. Moreover Weyl did not attempt to prove the (by no means trivial) infinite dimension extension of the theorem connecting self adjoint operators with unitary representations of the real line which played such a key role in his paper of 1927. He contented himself with the conjecture that this would now be proved using the techniques of recent work of von Neumann and others on spectral theory for unbounded self adjoint operators. These gaps in Weyl's paper were soon filled by work of M. H. Stone and von Neumann published in 1930 and 1931 [12, 13] and directly inspired by Weyl's work.

Wigner and Weyl not only introduced group representations into quantum mechanics in quite different ways with different goals but they reached this interaction between physics and mathematics from opposite directions. While Wigner was above all a theoretical physicist, Weyl was a pure mathematician. Moreover while Wigner had to have the existence of the theory of group representations called to his attention by von Neumann, Weyl had been actively working in the field since 1924. In particular, he and I. Schur had been responsible for the recent extension of the theory from finite groups to compact Lie groups. In 1924 Schur noticed that a great deal of the theory that had been developed by Frobenius, Burnside and himself in the preceding quarter of a century could be carried over to the orthogonal group in n dimensions by using a certain invariant integration process which had been discovered in 1897 by Hurwitz. Specifically one simply had to replace the frequently occurring summations over the set of all group elements in the finite case by integrating over the group manifold. Weyl, on learning of Schur's results, soon realized that one could extend the theory to all compact semi-simple Lie groups. He published a fully detailed account in 1925 and 1926 in a celebrated three part paper. In particular, he succeeded in determining all equivalence classes of irreducible representations for the groups in question.

With the description of Weyl's 1927 paper to refer to, let us return to Slater's paper of 1929 and explain how Slater was able (for the time being at least) to eliminate the "group pest" from the theory of atomic spectra. The idea (although Slater hardly looked at things in this way) was to do just the reverse of what Weyl had done in replacing Q and P by the one-parameter unitary groups $s \to e^{isQ}$, $t \to e^{itP}$. Let $\alpha \to U_\alpha$ be a continuous unitary representation of the group $G = SO(3)$ of all rotations about some point 0 in three dimensional Euclidean space. When this representation is restricted to the subgroup G_A of all rotations about an axis A going through 0 (and one has chosen a unit in which to measure angles) this restriction defines a unitary representation of the real line and, as indicated above, this representation may be written uniquely in the form $t \to e^{itM_A}$ where M_A is a self adjoint operator. Let M_x, M_y, M_z denote M_A when A is chosen to be the x, y, and z axes in some rectangular coordinate system. One can show that for every A, M_A is of the form $aM_x + bM_y + cM_z$ for suitable real numbers a, b and c so that the three operators M_x, M_y and M_z determine all M_A. Moreover one can show that, when the unit of angular measure is suitably chosen, M_x, M_y and M_z obey the following simple commutation relations

$$(*) \qquad \begin{aligned} &\frac{M_xM_y - M_yM_x}{i} = M_z \qquad \frac{M_yM_z - M_zM_y}{i} = M_x \\ &\frac{M_zM_x - M_xM_z}{i} = M_y\,. \end{aligned}$$

Conversely, at least when $\mathcal{H}(U)$ is finite-dimensional, it is easy to show that U is uniquely determined by the three operators M_x, M_y and M_z and moreover that (modulo one simple qualification) *every* triple of self adjoint operators which obeys the relations $*$ arises from some unitary representation of $G = SO(3)$.

This qualification is that one may have to be content with projective representations or equivalently with representations of the "double covering" $SU(2)$ of $SO(3)$. When $\mathcal{H}(U)$ is infinite dimensional the formulation and proof of this proposition is more complex and difficult but essentially the same result holds. In short, one may replace considerations involving the group representation U by considerations involving the operator triple M_x, M_y, M_z satisfying $*$.

The great advantage of such a replacement, from the point of view of the physicists of the time, was that groups and group representations were to them difficult and unphysically abstract new concepts while linear operators and matrices had become familiar to them from their appearance in the very foundations of quantum mechanics. Moreover operators M_x, M_y and M_z satisfying $*$ occur explicitly throughout quantum mechanics as the operator counterparts of the x, y and z components of angular momentum. At any rate when Slater showed how Wigner's results could be obtained by computations with "angular momentum" the physicists were delighted and by and large postponed learning group theory for a quarter of a century. The standard treatise on atomic spectroscopy for over a decade was the book, "The Theory of Atomic Spectra" [14], published by E. U. Condon and G. H. Shortley in 1935. An exposition of Slater's method, which is clearly the core of the book occupies most of their Chapters VI and VII. Toward the end of their introductory Chapter I, one finds the following statement: "We wish finally to make a few remarks concerning the place of the theory of groups in the study of the quantum mechanics of atomic spectra. The reader will have heard that this mathematical discipline is of great importance for the subject. We manage to get along without it." On the other hand, Chapter III entitled "Angular Momentum" is little more than a very complete account of the representation theory of $SU(2)$ (and hence of its quotient $SO(3)$) carried out from the point of view of a systematic study of triples of operators satisfying the commutation relations $*$.

Of course this correspondence between unitary representations of $SO(3)$ or $U(2)$ and angular momentum triples is just a very special case of the fact that every connected Lie group G admits an "infinitesimal version" which is called its "Lie algebra" and that this Lie algebra has a representation theory which largely parallels that of the group. Once one understands this, it seems almost bizarre that the physicists should have made such a big thing of the difference between working with angular momentum operators and working with representations of $SO(3)$. However, at the time most physicists did not know what a group was, much less a Lie group or its Lie algebra. The reader will find support for this contention on page 555 of "Inward bound" by A. Pais who, after calling attention to the knowledge of Lie groups possessed by graduate students in 1986 says "It was not so in, say 1960. Whatever the average well-educated theorist then knew about groups concerned rotations, Lorentz transformations and various discrete symmetries". One might also read the last eight lines of page 163 and all of page 164 of J. Bernstein's "The Life it Brings" where we learn that Murray Gell Mann, in the fall of 1960, had to be told by a mathematician that the "cute" objects he was playing with in the early development of current algebra were well known to mathematicians as Lie groups and had

an elaborate theory. The fact that Bernstein referred to these objects as Lie groups when he was really talking about Lie algebras was not exactly a mistake on his part. In those days at least, many physicists used these two terms interchangeably.

Let us return to Wigner's book and explain the concept of *tensor operator* which is introduced therein. Just a a real-valued observable in quantum mechanics is associated with a self adjoint operator, a vector-valued observable, such as a velocity or a momentum, is associated with a triple of self adjoint operators. However, as in classical mechanics, the particular triple one uses depends upon the coordinate system. A little reflection shows that this dependence must take the following form. Let U be the unitary representation of the rotation group defined by the rotational invariance of the whole system and let A_1, A_2, A_3 be the self adjoint operators describing our vector valued observable in some coordinate system. Then if we change coordinates by a rotation α the three components A_1, A_2, A_3 must be replaced by $U_\alpha A_1 U_\alpha^{-1}$, $U_\alpha A_2 U_\alpha^{-1}$, $U_\alpha A_3 U_\alpha^{-1}$. So much would be true for any triple of observables. The fact that they are the components of a vector implies further that for each $j = 1, 2, 3$ $U_\alpha A_j U_\alpha^{-1}$ is a linear combination of A_1, A_2, and A_3, say, $b_{j1}A_1 + b_{j2}A_2 + b_{j3}A_3$ and that the mapping

$$\alpha \longrightarrow \begin{pmatrix} b_{11}(\alpha^{-1}) & b_{12}(\alpha^{-1}) & b_{13}(\alpha^{-1}) \\ b_{21}(\alpha^{-1}) & b_{22}(\alpha^{-1}) & b_{23}(\alpha^{-1}) \\ b_{31}(\alpha^{-1}) & b_{32}(\alpha^{-1}) & b_{33}(\alpha^{-1}) \end{pmatrix}$$

of the rotation group is equivalent to the unique three dimensional irreducible unitary representation D^1 of this group. This implies, in particular, that each of the three components A_1, A_2, A_3 has the property that the continuum many operators $U_\alpha A_j U_\alpha^{-1}$ have a three-dimensional vector space as their linear span and that this three-dimensional vector space is independent of j and invariant under all U_α. Although this does not seem to be customary in physics, one could take the three-dimensional vector space spanned by A_1, A_2 and A_3 as a coordinate independent quantum representation of the vector observable in question.

Quite generally, Wigner defined a triple A_1, A_2, A_3 of self adjoint operators in the Hilbert space of a rotationally invariant quantum mechanical system to be a *vector operator* if it has the above stated properties with respect to the representations U and D^1 of the rotation group. He then generalized this notion by replacing D^1 by an arbitrary irreducible representation D^j of the rotation group. Since D^j is $2j + 1$-dimensional, the triple A_1, A_2, A_3 must be replaced by a $2j + 1$ triple $A_1, A_2, \ldots, A_{2j+1}$. Repeating the definition of vector operator with these changes one obtains Wigner's definition of an *irreducible tensor operator.*

Wigner introduced this notion in Chapter XXI, Section 5 of his book in order to formulate and prove a generalization of a very useful theorem about the matrix elements of an operator S which commutes with all U_α. To say that an operator S commutes with all U_α is the same as to say that it is an irreducible tensor operator in the special case in which $j = 0$. Thus to be a

component of an irreducible tensor operator for some j not necessarily equal to zero is a natural generalization of actually commuting with all U_α.

The theorem which Wigner generalized is a theorem about the matrix elements of an operator S which commutes with all U_α with respect to a so-called "angular momentum basis". This latter notion can be defined in somewhat greater generality than in the case considered by Wigner and it will be more convenient to give the more general definition. Decompose the Hilbert space as a direct sum $\mathcal{H}(U) = \mathcal{H}_1 \oplus \mathcal{H}_2 + \ldots$ of invariant subspaces each of which is such that the restriction U^k of U to $\mathcal{H}_k$ is multiplicity free for all k. This means that each U^k is *uniquely* a direct sum of irreducible representations D^l where no l occurs more than once. Let $\mathcal{H}_l$ denote the $2l+1$-dimensional subspace on which U^n is equivalent to D^l (whenever U^k actually contains D^l). Finally, choose once and for all a particular axis and consider the restriction of U^n to rotations about this axis. Under this restriction $H_{n,l}$ splits into $2l+1$ invariant subspaces which we label with an integer m with $-l \leq m \leq l$. Choosing a unit vector $\varphi_{n,l,m}$, in each of these one-dimensional subspaces one obtains an *angular momentum basis* for $\mathcal{H}(U)$.

The theorem which Wigner generalized, and which is proved in an earlier part of the book, asserts the following. If $\{\varphi_{n,l,m}\}$ is an angular momentum basis for U and S is an operator which commutes with all U_α, then the matrix element $(S(\varphi_{n,l,m}) \cdot \varphi_{n',l',m'})$ is zero unless $l = l'$ and $m = m'$ and in that case its value is independent of m. In the generalization S is replaced by the ρ component of an irreducible tensor operator for D^ω ($\rho =$ one of the integers: $-\omega, -\omega+1, \ldots \omega$). The theorem asserts that the matrix element is zero unless l' lies in the interval $|l-\omega| \leq l' \leq l+\omega$ and $m' = m+\rho$. When these conditions hold it is a product of two factors one of which is independent of the particular tensor operator chosen and the other of which depends only on n, n', l, and l' and not on m or m'. The former factor can be computed from the general properties of tensor products of the irreducible representations of the rotation group once ω, ρ, l, l', and m are known.

In the special case of vector operators this theorem appeared in a survey article by C. Eckart [15]. The generalization to other irreducible tensor operators was implicit in Eckart's paper. However, Wigner was the first to give a formal definition of irreducible tensor operator and work out the general theorem explicitly. It is now known as the Wigner-Eckart theorem.

The notion of tensor operator may be defined in Lie algebra terms and in this form played an important role in the work of Racah a decade later.

A few months before finishing work on his book, Wigner began applying the theory of group representations to problems in physics other than those of atomic spectroscopy. In [16], published in 1930, he exploited the fact that a classical dynamical system with a potential energy which is a quadratic form in the coordinates (such as in the small vibrations about an equilibrium point) has a phase space which can, in a natural way, be given the structure of a finite dimensional Hilbert space. Any finite group G of symmetries of the system gives rise to a unitary representation of G and the self adjoint operator H whose eigenvalues give the normal modes lies in the commuting algebra of this

group representation. The general principles explained earlier in the essay can then be applied to obtain information about the normal modes. Wigner refers to an earlier paper of Brester where the same results are obtained by elementary methods but asserts that using group representations produces better insight into the problem.

Two years later Wigner published a paper [17] showing that some recent work of Kramers could be explained by a slight extension of the theory of unitary group representations and proceeded to study this extension systematically. This extension consists in allowing the representing operators to be *anti-unitary* as well as unitary. A norm-preserving real linear operator U is said to be anti-unitary if the condition $U(\lambda x) = \lambda U(x)$ for complex λ is replaced by $U(\lambda x) = \overline{\lambda} U(x)$ where $\overline{\lambda}$ is the complex conjugate of λ. If U is such an extension, one checks easily that the subset of all x for which U_x is actually unitary is a closed normal subgroup of index 2. In other words, U can always be obtained from a unitary representation by adjoining a single anti-unitary operator and its products with the unitary U_x. The point is that there can be symmetries of a quantum system which are described by anti-unitary operators, the simplest example being $\psi \to \overline{\psi}$ for spinless collections of interacting particles. This will be a symmetry whenever the Hamiltonian is real and is called "time reversal". For a discussion from a somewhat different of point of view see the annotation of Wightman in Volume III.

Two other applications of group representations which Wigner found in the 1930's were intimately related to his own activities as a physicist – activities in which his main contributions were *not* group theoretical. In 1932 Wigner gave a course at Princeton on solid state physics, for which lecture notes were made available. In the ensuing six years he published nine papers on the subject. In the same period he published over a dozen papers on nuclear physics. Nuclear physics changed radically (and may even be said to have begun) with Chadwick's discovery of the neutron in 1932. Wigner became interested almost immediately. Indeed he was one of the pioneers in studying the nature of the nuclear interaction, the structure of nuclei, etc. As far as solid state physics is concerned he was also a pioneer in that in 1934 [18] he introduced and made the first calculation of the "correlation energy" in the problem of calculating the cohesive energy of metals. It was of course natural for Wigner to attempt to apply the theory of group representations to these new interests. The attempt succeeded in each case and led to two much cited papers [19, 20] published in 1936 and 1937.

We shall discuss the 1937 paper, "On the consequence of the symmetry of the nuclear Hamiltonian on the spectroscopy of the nuclei" first because it is more closely related to Wigner's work on atomic spectroscopy, and because the 1936 paper can best be discussed in connection with later work; especially the celebrated 1939 paper [21] "On unitary representations of the inhomogeneous Lorentz group".

It will be useful to begin our discussion of the 1937 paper by recalling the mathematics underlying the celebrated Pauli exclusion principle. Let $\mathcal{H}_1$ and $\mathcal{H}_2$ be the Hilbert spaces of states for two particles in quantum mechanics.

Then the Hilbert space for the system obtained by putting the two particles in interaction is the tensor product $\mathcal{H}_1 \otimes \mathcal{H}_2$ of the two Hilbert spaces – at least when the two particles are "distinguishable". More generally, if one has n particles with Hilbert spaces $\mathcal{H}_1 \ldots \mathcal{H}_n$ (and they are "distinguishable") the Hilbert space will be the n-fold tensor product $\mathcal{H}_1 \otimes \mathcal{H}_2 \ldots \otimes \mathcal{H}_n$. That this simple principle did not necessarily hold for "indistinguishable" particles became clear in the earliest applications of quantum mechanics to atoms with more than one electron. The failure of many expected spectral lines to appear was soon shown to be explainable by the hypotheses that the Hilbert space $\mathcal{H} \otimes \mathcal{H} \ldots \otimes \mathcal{H}$ (n factors), where $\mathcal{H}$ is the one electron Hilbert space, must be replaced by the anti-symmetric subspace $\mathcal{H} Ⓐ \mathcal{H} Ⓐ \ldots \mathcal{H}$. To define this one recalls the natural unitary representation U of the symmetric group S_n on $\mathcal{H} \otimes \mathcal{H} \otimes \ldots \otimes \mathcal{H}$, defined earlier and chooses the subspace on which U_π is minus the identity for all odd permutations π. The Pauli exclusion principle as first formulated by Pauli is an immediate consequence.

Later it was found that whenever one has several particles of the same type in interaction, the product of the corresponding identical Hilbert spaces must be replaced either by $\mathcal{H} Ⓐ \mathcal{H} Ⓐ \ldots \mathcal{H}$ – the anti-symmetrized tensor power – or by the subspace $\mathcal{H} Ⓢ \mathcal{H} Ⓢ \ldots \mathcal{H}$ defined by replacing $-I$ by I in the anti-symmetric definition. Accordingly one speaks of the particles in question as being fermions or bosons. Electrons, protons and neutrons turn out to be fermions and photons to be bosons.

A nucleus is formed by an interaction between n_1 protons and n_2 neutrons and the appropriate Hilbert space is

$$((\mathcal{H}_P Ⓐ \mathcal{H}_P Ⓐ \ldots \mathcal{H}_P) \otimes (\mathcal{H}_N Ⓐ \mathcal{H}_N Ⓐ \ldots \mathcal{H}_n)$$

where $\mathcal{H}_P$ is the proton Hilbert space, $\mathcal{H}_N$ is the neutron Hilbert space, and the two parentheses contain n_1 factors and n_2 factors respectively. This statement exemplifies a general principle whose precise formulation we leave to the reader.

The content of Wigner's 1937 paper is perhaps most easily explained in terms of a simple and general mathematical theorem concerning the connection between direct sums of Hilbert spaces and anti-symmetrized tensor powers. Let $\mathcal{H}_1$ and $\mathcal{H}_2$ be two Hilbert spaces and consider $(\mathcal{H}_1 \oplus \mathcal{H}_2) Ⓐ (\mathcal{H}_1 \oplus \mathcal{H}_2) Ⓐ \ldots (\mathcal{H}_1 \oplus \mathcal{H}_2)$ where there are n factors, i.e., the anti-symmetrized n-fold tensor power of the direct sum $\mathcal{H}_1 \oplus \mathcal{H}_2$. The theorem asserts that there is a *natural* isomorphism between this Hilbert space and the following direct sum of products of anti-symmetrized tensor powers

$$\overset{A}{\mathcal{H}}{}_1^n \oplus \left(\overset{A}{\mathcal{H}}{}_1^{n-1} \otimes \mathcal{H}_2\right) \oplus \left(\overset{A}{\mathcal{H}}{}_1^{n-2} \otimes (\mathcal{H}_2 Ⓐ \mathcal{H}_2)\right) + \ldots + \overset{A}{\mathcal{H}}{}_2^n .$$

(Here we have introduced the notation $\overset{A}{\mathcal{H}}{}^m$ for $\mathcal{H} Ⓐ \mathcal{H} Ⓐ \ldots \mathcal{H}$ with m factors.)

Because of this theorem one can *simplify* the problem of studying the interaction between n_1 protons and n_2 neutrons by considering simultaneously the interaction between $n_1 + k$ protons and $n_2 - k$ neutrons for $k = n_2, n_2 - 1, n_2 - 2 \ldots - n_1$. The direct sum of all of these $n_1 + n_2 + 1$ Hilbert spaces

may be identified with a single anti-symmetrized power $(\mathcal{H}_P \overset{A}{+} \mathcal{H}_N)^{n_1+n_2}$ and our problem is more or less reduced to that of considering $n_1 + n_2$ identical "particles" the Hilbert space of each of which is the direct sum $\mathcal{H}_N \oplus \mathcal{H}_P$ of the proton and neutron Hilbert spaces. A key point here is that one can look at the Hilbert space $\mathcal{H}_N \oplus \mathcal{H}_P$ as the Hilbert space of a particle which is neither a proton or a neutron but which can be either. The idea of looking at the neutron and proton as two different states of a single particle (sometimes called a nucleon) was introduced by Heisenberg in 1932 in the same year that the neutron was discovered. He used it in formulating a hypothesis about the nature of the force between two nucleons. The theorem about anti-symmetrized tensor powers cited above plays an essential role in establishing the consistency of Heisenberg's point of view.

In Wigner's 1937 paper the above described way of viewing the interaction between protons and neutrons is exploited to set up an analogy between atomic structure and nuclear structure. One replaces the n identical electrons in an atom by the n identical nucleons in a nucleus. The role of the atomic nucleus is played by the center of gravity of the nucleons and one finds that much the same group theoretical methods may be applied. The analogy would be even closer than it is "because of the fact that the electron has spin $\frac{1}{2}$" if it were not for the fact that the proton and neutron also have spin $\frac{1}{2}$. At the same time this circumstance makes it necessary to consider the representation theory of somewhat higher dimensional compact Lie groups than the rotation group and its simply connected covering group $SU(2)$ – in particular $SU(4)$. (In general $SU(n)$ is the group of all $n \times n$ unitary matrices of determinant one.)

To clarify these, perhaps somewhat cryptic, remarks let us recall the Hilbert space formulation of the concept of the "spin" of a particle. For a particle without spin the Hilbert space can always be realized as $\mathcal{L}^2(S, \mu)$ where S is physical space and μ is the volume measure in S. However, certain anomalies in atomic spectra could be explained very satisfactorily by assuming that for electrons $\mathcal{L}^2(S, \mu)$ is replaced by a slightly different Hilbert space which may be equivalently defined as $\mathcal{L}^2(S, \mu) \otimes \mathcal{H}_2$ where $\mathcal{H}_2$ is a two-dimensional Hilbert space or as $\mathcal{L}^2(S, \mu, \mathcal{H}_2)$ the space of all square summable functions from S, μ to this same two-dimensional Hilbert space. This change leaves the operators defining the coordinate and linear momentum observables essentially unchanged. More precisely it replaces them (in the $\mathcal{L}^2(S, \mu) \times \mathcal{H}_2$ form) by their tensor products $Q_j \times I$ and $P_j \times I$ where Q_j and P_j are the old operators in $\mathcal{L}^2(S, \mu)$. Each *angular* momentum operator Ω in $\mathcal{L}^2(S, \mu)$ is, however, replaced by a sum of the form $\Omega \times I + I \times \sigma$ where σ is a self adjoint operator in $\mathcal{H}_2$. The "extra" angular momentum term $I \times S$ is thought of as due to a rotation or "spin" of the electron about an axis through its center. Hence the term "spin". The mathematics of course is independent of such imaginative interpretations. The operator σ depends upon which axis of rotation one considers but in all cases has the same two eigenvalues and trace zero. Once one has chosen an axis (the Z axis, say) the space $\mathcal{H}_2$ may be written as a direct sum of two one-dimensional subspaces

in each of which σ_t has a definite value and correspondingly $\mathcal{L}^2(S,\mu)\times\mathcal{H}_2$ may be written as a direct sum of two replicas of $\mathcal{L}^2(S,\mu)$.

Notice the analogy between its direct sum decomposition into subspaces of positive and negative spin with that for nucleons into proton and neutron subspaces. It is important to remember, however, that the analogy is not complete because the proton neutron splitting does not depend upon choosing a direction in space. Moreover since neutrons and protons also have non-zero spin and in fact a spin also described by a two-dimensional Hilbert space one has to deal with both effects at the same time. The Hilbert space of a nucleon is of the form $\mathcal{L}^2(S,\mu)\otimes\mathcal{H}_2\oplus\mathcal{L}^2(S,\mu)\otimes\mathcal{H}_2$ or equivalently $\mathcal{L}^2(S,\mu)\otimes(\mathcal{H}_2\oplus\mathcal{H}_2)$ where of course $\mathcal{H}_2\oplus\mathcal{H}_2$ is a four-dimensional Hilbert space.

In the end the crucial analogy is that between the anti-symmetrized n-th power of the Hilbert space $\mathcal{L}^2(S,\mu)\otimes\mathcal{H}_2$ which underlies the quantum mechanics of the n electrons in an atom and the anti-symmetrized n-th power of the Hilbert $\mathcal{L}^2(S,\mu)\otimes(\mathcal{H}_2\oplus\mathcal{H}_2)$ or (equivalently) $\mathcal{L}^2(S,\mu)\otimes\mathcal{H}_4$ which underlies the quantum mechanics of a nucleus with n nucleons. The analogy is incomplete, because rotational symmetry is implemented rather differently in the Hilbert spaces $\mathcal{H}_2$ and $\mathcal{H}_4$ – the relevant representations of the rotation group in these two spaces being irreducible in the first case and a direct sum of two equivalent irreducible representations in the second. It is incomplete in another and more profound way which we shall describe in the next paragraph.

So far everything we have said would still be valid if the neutron and the proton were not only of different charge but also of different mass and attracted one another by forces which varied with distance and spin in quite different ways. However, a new symmetry group (more precisely a new approximate symmetry group) is available because the neutron and proton are almost identical *except for charge*. The ratio of the mass of the neutron to that of the proton is approximately 1.0014. Moreover in the very first sentence of his 1937 paper, Wigner cites two experimental papers, published in 1936 which "appear to show that the forces between all pairs of constituents of the nucleus are approximately equal". Of course, this statement must be interpreted to imply the qualification "after allowing for the electrostatic repulsion between two protons". This suggests that a good approximation to a model for a nucleus will be invariant under an action of the group of all unitary operators in $\mathcal{L}^2\mathbb{S}\otimes\mathcal{H}_2\otimes\mathcal{H}_2$ of the form $I\times I\times U$ where U is a unitary operator in $\mathcal{H}_2$ and that the action of this group on the n nucleon Hilbert space will be analogous to that of the rotation group. To be more explicit than this would require a more detailed discussion than seems appropriate. Of course, this symmetry and the rotational symmetry must be combined and one has finally to deal with an analogue of the problem of atomic spectra on which the underlying group theory is somewhat more involved.

The extra degree of group theoretical difficulty increases considerably when one considers with Wigner a less accurate but more symmetrical approximation in which the full group of all unitary operators in $\mathcal{H}_4=\mathcal{H}_2\otimes\mathcal{H}_2$ appears as a symmetry group. Whereas the direct product of the rotation group and the group $U(2)$ of all unitary operators in $\mathcal{H}_2$ has a representation theory

easily deducible from that of the rotation group (which was in turn familiar to physicists from the theory of angular momentum), that of the unitary group in four dimensions was new and unfamiliar. Indeed a substantial part of the paper is devoted to recalling the classical theory of Cartan, Schur, and Weyl in its application to the four dimensional unitary group and indicating how the theory of rotational symmetry changes when $U(4)$ replaces the rotation group.

Actually Wigner discusses two degrees of approximate symmetry in addition to that of the preceding paragraph. In one of these he allows forces which depend on the spin of the particle but not on whether it is a proton or neutron and in the other he does the opposite. As is clear from the foregoing, the mathematics of the reversal of spin and that of passing from the neutron state to the protein state are very similar and involve groups which have isomorphic Lie algebras. To emphasize this Wigner introduced the term *isotopic spin* to describe a concept which was destined to play a key role in later developments in elementary particle physics.

Concerning the significance of the work for physics, it is useful to quote a few lines from David Brink's book "Nuclear forces" first published in 1965. "They did not, however, work out more general consequences of the theory or show how it could be used for more complex nuclei. This was done by Wigner (1937) in a very important paper which also included the first work on a general method of classifying nuclear states called supermultiplet theory. This theory was much more involved than isobaric (isotopic) spin theory, and Wigner's paper was too difficult for most experimental and many theoretical physicists, and the simple consequences of charge independence were not widely recognized until many years later."

Before discussing the Wigner papers of 1936 and 1939 mentioned above it will be useful to explain a general theorem about group representations, which was known to Frobenius for finite groups, and generalizations of which to certain infinite groups play a role in Wigner's two papers. In the interests of later generalizations and unifications we shall formulate this theorem rather differently than Frobenius did. In particular we shall emphasize representations rather than their characters.

Let G be an arbitrary group (not necessarily finite) and let N be a normal subgroup of G; that is a subgroup such that xnx^{-1} is in N for all n in N and all x in G. One says that G is an extension of N and that "the extension splits" if there exists a second subgroup H such that $N \cap H$ consists only of the identity element e ($N \cap H = \{e\}$) and every element x in G is a product nh of an element n in N and an element h in H. One proves easily that the decomposition $x = nh$ is unique and that the subgroup H, when it exists is unique up to isomorphism. On the other hand, it is often possible to find more than one subgroup H' which effects the splitting when one exists.

Starting at the other end, so to speak, let N and H be two arbitrary groups and suppose that we are given a mapping $h \to \alpha_h$ taking each element h of H into an "automorphism" of N (i.e., an isomorphism of N with itself). Suppose further that $h \to \alpha_h$ is a homomorphism of H into the group of all automorphisms of N (i.e., preserves group multiplication). Consider the set G

of all pairs (n, h) where $n \in N$ and $h \in H$ and define a multiplication in G by the following rule $(n_1, h_1)(n_2, h_2) = (n_1\alpha_{h_1}(n_2), h_1h_2)$. It is straightforward to prove that under this rule G is a group, that the set of all (n, e) where e is the identity of H and $n \in N$ is a normal subgroup isomorphic to N and that the set of all (e, h) with $h \in H$ is isomorphic to H and makes G into a split extension of N. Every split extension of a group G by a normal subgroup can be constructed in this way. G is said to be a *semi-direct product* of N and H and to be the semi direct product *defined by the homomorphism* $h \to \alpha_h$. It is easy to see that a semi-direct product $N \otimes H$ reduces to the ordinary direct product whenever α_h is the identity for all h.

The theorem of Frobenius, which we are about to explain, is a theorem about the classification of all irreducible representations of a semi-direct product $N \otimes H$, when both N and H are finite and N is commutative. It reduces the problem to that of classifying the irreducible representations of H and certain of its subgroups. Both finite and infinite semi direct products occur in important places in physics. In the 1936 paper of Wigner cited above N is the group of translations in the so-called "space group" of a crystal lattice and H is a finite group. In Wigner's 1939 paper N is the four-dimensional translation group in space time and H is the Lorentz group.

Before actually stating Frobenius' theorem we shall need the notion of "induced representation". Let K be an arbitrary subgroup of a finite group G and let L be an arbitrary representations of K. We propose to use L to define a representation of the whole group G. Let $\mathcal{H}(L)$ be the vector space in which the operators L_k $(k \in K)$ act. Define $\mathcal{F}^L$ to be the vector space of all functions f from the group G to the vector space $\mathcal{H}(L)$ which satisfy the identity

$$(*) \qquad f(kx) = L_k(f(x))$$

for all $k \in K$ and all $x \in G$. Evidently the right translate of any member of $\mathcal{F}^L$ by any member of G is again a member of $\mathcal{F}^L$. Hence we obtain a representation U^L of G whose space is $\mathcal{F}^L$ by defining $U^L_x f(y) = f(yx)$. By definition U^L is the representation of G *induced* by the representation L of K. It is easy to check that the dimension of $\mathcal{F}^L$ is the dimension of $\mathcal{H}(L)$ multiplied by the number of right K cosets in G; that is, by $o(G)/o(K)$.

Now let $G = N \otimes H$ be a semi-direct product of the finite commutative group N and some other finite group H. Let $\widehat{N}$ denote the so-called character group of N; that is the set of all functions χ from N to the complex numbers of modulus one such that $\chi(n_1n_2) = \chi(n_1)\chi(n_2)$ for all n_1 and n_2 in N. These functions clearly form a group under multiplication. The members of $\widehat{N}$ are precisely the one dimensional representations of N and, since N is commutative, precisely the irreducible representations. Let α_h be the automorphism of N defined by h in H. Then for each h and χ, $n \to \chi(\alpha_h(n))$ is also a member of $\widehat{N}$ which we denote by $(\chi)\alpha_h^*$. It is easy to see that $\chi \to (\chi)\alpha_h^*$ is an automorphism of $\widehat{N}$ and that for each $\chi \in \widehat{N}$ the set of all h in H with $(\mathcal{X})\alpha_h^* = \chi$ is a subgroup of H. We denote this subgroup by H_χ.

Theorem (Frobenius). *Choose any $\chi \in \widehat{N}$ and form the subgroup G_χ of G consisting of all nh with n in N and h in H_χ. Choose any irreducible representation L of H_χ; then*

(a) *$nh \to \chi(n)L_h$ is an irreducible representation of G_χ. Denote it by χL.*
(b) *The induced representation $U^{\mathcal{X}L}$ of G is irreducible.*
(c) *Every irreducible representation of G is equivalent to $U^{\mathcal{X}L}$ for some choice of $\mathcal{X}$ and L.*
(d) *$U^{\mathcal{X}L}$ and $U^{\mathcal{X}M}$ are equivalent if and only if L and M are equivalent representations of H_χ.*
(e) *If χ_1 and χ_2 are in the same H "orbit", that is, if $\mathcal{X}_2 = (\mathcal{X}_1)\alpha_h^*$ for some H then every irreducible representation of G of the form $U^{\mathcal{X}_2 L}$ is equivalent to $U^{\mathcal{X}_1 M}$ for some M.*

Corollary. *Let C be a "cross section" for the H orbits in $\widehat{N}$, that is, a subset of $\widehat{N}$ with the property that for each $\chi \in \widehat{N}$, there is one and only one member of C which lies in the same orbit as χ. Then, as χ varies over C and L varies over the inequivalent irreducible representations of H_χ, the irreducible representation $U^{\chi L}$ of G passes through each equivalence class of representations of G once and only once.*

It follows from this theorem and its corollary that the classification of the irreducible representation of $G = N \otimes H$ proceeds in two steps. First, study the action of H on $\widehat{N}$ defined by the α_h^* and pick a cross section C for the orbits. Then, for each χ in C classify the irreducible representations of H_χ. This method does not necessarily give a complete classification. It merely reduces the classification problem to several presumably easier ones. However, in many important cases the easier problems can be solved by other methods, or even by a second application of Frobenius' theorem. Moreover, it usually happens that many of the H_χ for χ in *different* orbits are isomorphic groups, so that the number of essentially different $H\chi$'s to analyze is much less than the number of orbits. To give more insight into the situation we remark that the orbit associated with a given irreducible representation of $N \otimes H$ is precisely the set of all members of $\widehat{N}$ which occur when one regards this representation as a representation of N and reduces it into its (one-dimensional) irreducible constituents.

Wigner's 1936 paper [20] was written jointly with L. P. Bouckaert and R. Smoluchowski and is entitled: "Theory of Brillouin zones and symmetry properties of wave functions in crystals". A solid crystal is a collection of an enormous number of atoms arranged in space in a regular pattern capable of indefinite extension. For the purposes of mathematical analysis it is convenient to pass to the limit in which the extension is continued to infinity. The "regularity" of the pattern is then defined by the existence of a large discrete group of isometries of space which leave the pattern unaltered. In particular the subgroup of translations is such that every atom is a translate of a member of a finite subset of the atoms. Thus the whole pattern is determined by specifying

the nature and position of a finite number of the atoms. The group of all such isometries is called the *space group* of the crystal. In the most important cases, it is a semi-direct product of the translation subgroup and a finite group. One speaks of the space group as being *symmorphic* in these cases.

In studying the quantum mechanics of crystal structure, one starts with an approximation in which the outer electrons in each atom move freely and independently in the field generated by the ion cores held in fixed positions in a crystal lattice. In a further refinement one modifies this ion core field by adding to it the average of the fields due to the moving electrons. In either case, in the limit of an infinite crystal, one considers the motion of a single electron moving in space subject to an effective potential function v defined in space, which is not only triply periodic but invariant under the action on space of the relevant space group. As is usual in quantum mechanics, one is interested in the spectrum of the self adjoint operator which defines the motion. This operator is

$$-\frac{\hbar^2}{2m}\left(\frac{\partial^2}{\partial x^2}+\frac{\partial^2}{\partial y^2}+\frac{\partial^2}{\partial z^2}\right)+v=H_v$$

and evidently commutes with the unitary representation W of the space group Γ, defined by its natural action on Euclidean three space E_3. If it were possible to decompose W as a discrete direct sum of irreducible representations, this decomposition would imply a partial diagonalization of H_v, as explained early in this essay. However, Γ is *not* a compact group and W decomposes "continuously" rather than discretely. Nevertheless an analogue of the theorem for compact groups continuous to hold and H_v can be "continuously" decomposed into self adjoint operators parametrized by irreducible unitary representations of Γ. Moreover, each member of this continuous family of self adjoint operators turns out to have a discrete spectrum. These discrete eigenvalues vary continuously with the representation of Γ which parametrizes them and the representation theory of Γ becomes an indispensable tool. The 1936 paper of Wigner el al. (BSW), which we are now proposing to discuss, studies the representation theory of Γ with this application in view.

This paper only considers the case in which the relevant space group is symmorphic so that one can hope to deduce its representation theory from a suitable generalization of the Frobenius theorem cited above. This hope is justified; moreover, the generalization turns out to be a rather easy one because the only induced representations that occur are those in which the relevant subgroup has finite index in the whole group. In this case the definition of induced representation given above may be used without change. However, one is only interested here in unitary representations so we add the remark that U^L can be made unitary in a natural way whenever L is unitary. This said, one can extend the results of Frobenius immediately to all semi direct products $N\otimes H$ where H is finite and N is an arbitrary locally compact commutative group, provided that one changes the statement by putting "unitary" in front of representation, wherever it occurs. Of course, one must assume that the automorphisms α_h of N respect the topology of N, but this is automatically true in the space group case because N is discrete. The classification of the

unitary representations of $N \otimes H$, described by the generalization of Frobenius theorem is assumed in the BSW paper and is summarized in Section II. The reader is referred for details to a paper of Seitz, published a year earlier in "Annals of Mathematics". Seitz was a recent Ph. D. student of Wigner but seems to have attacked this problem independently.

The main concern in BSW is with the way in which the subgroups H_χ of H vary as χ varies over $\widehat{N}$ and how this influences the spectrum of the self adjoint operator H_v. They discuss the general situation in Sections III and IV and analyze three special cases in Sections V, VI, and VII. A key fact is the following. Whenever H' is a subgroup of H which is not just the identity, then the set of all χ in $\widehat{N}$ for which $H_\chi = H'$ is either empty or is a closed lower dimensional closed subset of the three dimensional torus $\widehat{N}$. Since there are only a finite number of subgroups of H, the subset $\widehat{N}^o$ of all χ for which $H_\chi = \{e\}$ is an open everywhere dense subset of $\widehat{N}$, i.e., any χ sufficiently close to a member of $\widehat{N}^o$ is in $\widehat{N}^o$ and every member of $\widehat{N}$ either in $\widehat{N}^o$ or a limit of a sequence of members of $\widehat{N}^o$. Evidently having $H_\chi = \{e\}$ is the "generic" case. In that case the set of all $\alpha_h^*(\chi)$ (which BSW call the "star" of λ) contains $o(H)$ elements and is as large as possible. Leaving out the representations associated with non generic members of $\widehat{N}$, one has a perfect one-to-one correspondence between irreducible representations and stars. The authors also study the more complex behavior of the non generic representations.

In order to study the connection between the unitary representations of Γ and the spectrum of the self-adjoint operator H_v one does not need to study the theory of continuous decomposition of Hilbert spaces (which in any case did not exist at the time). One can be led to the component self-adjoint operators in question by following F. Bloch and others in considering the eigenfunctions of the operator H_v which are not square summable but have the following quasi periodicity property instead.

$$(*) \qquad f(x+n_1, y+n_2, z+n_3) = \chi(n_1, n_2, n_3) f(x, y, z)$$

for all real triples x, y, z and all real triples n_1, n_2, n_3 contained in N. Here χ is some member of $\widehat{N}$ and one gets a different class of functions for each choice of χ. Now every χ in $\widehat{N}$ may be written (non uniquely) in the form $\chi(n_1, n_2, n_3) = e^{i(\lambda_1 n_1 + \lambda_2 n_2 + \lambda_3 n_2)}$ and it is trivial to show that every "Bloch function", that is, every function that satisfies $(*)$ above may be written (uniquely once $\lambda_1, \lambda_2, \lambda_3$ have been chosen) in the form

$$f(x, y, z) = e^{i(\lambda_1 x + \lambda_2 y + \lambda_3 z)} g(x, y, z)$$

where g is periodic with respect to translations by members n_1, n_2, n_3 of N. An easy computation then shows that f is a μ eigenfunction for H_v if and only if g is a μ eigenfunction for the operator

$$-\nabla^2 - 2i\left(\lambda_1 \frac{\partial}{\partial x} + \lambda_2 \frac{\partial}{\partial y} + \lambda_3 \frac{\partial}{\partial z}\right) + \lambda_1^2 + \lambda_2^2 + \lambda_3^2 + v\,.$$

But, this operator may be regarded as an operator defined on the compact group obtained from the additive group of all triples x, y, z of real numbers by identifying two triples when they differ by a member n_1, n_2, n_3 of N. Hence under very mild restrictions on v it is an operator with a discrete spectrum. In its dependence on $\lambda_1, \lambda_2, \lambda_3$ it depends only upon the member χ of $\widehat{N}$ which it defines by the formula above. We denote it by $H_{v,\chi}$. These operators are precisely those that H_v decomposes into when the unitary representation of N in $\mathcal{L}^2(E^3)$ is decomposed into its irreducible constituents which of course are the characters χ of N. Since one can show that each eigenvalue of each $H_{v,\chi}$ varies continuously with χ, the whole spectrum is a union of countably many closed intervals on the real line. This union does not need to be connected, and when it is not, there will be open intervals separating the connected components. The spectrum has a so-called "band structure".

We now turn our attention to one of Wigner's most celebrated papers [21]. This paper, published in 1939 in the *Annals of Mathematics*, is entitled *On unitary representations of the inhomogeneous Lorentz group*. Its basic result is an analysis of the irreducible unitary representations of the group of the title, including a complete classification of those that are "physically relevant". The analysis includes the so-called "ray" or "projective" representations (see above discussion of 1927 paper of Weyl), as well as a proof that the only possible projective multipliers are rather special. The physical motivation for the study is the fact that the Hilbert space of states of any relativistically invariant quantum mechanical system is the space of a naturally associated projective or ordinary representation of the group in question. This kind of application of the theory of group representations to quantum mechanics is much more in the spirit of Weyl's 1927 paper in the *Zeitschrift für Physik* than that of most of Wigner's work up to this point. In the terminology introduced earlier in this essay, Wigner's 1939 paper was a further contribution to "Weyl's program". At the same time it constituted a key step in the purely mathematical program of developing a general theory of unitary group representations. Except for the trivial representation in which every group element is taken into the identity, *all* of the irreducible unitary representations of the inhomogeneous Lorentz group are infinite dimensional. This was the first time that the irreducible unitary representations of such a group had been even partially classified in a more or less rigorous measure – at least explicitly. Actually many years later it was realized that the work of Stone and von Neumann on the uniqueness of the irreducible solutions of the Heisenberg commutation relations (see above) could be regarded as being such a theorem and trivially implying another. In addition to this the completion of Wigner's classification by dealing with the non-physically relevant representations was one of the stimuli eight years later for the work of Bargmann and of Gelfand and Neimark on the unitary representation theory of the groups $SL(2, R)$ and $SL(2, C)$. This, in turn, stimulated the vast and still incomplete program of classifying the irreducible unitary representations of the non compact semi simple Lie groups.

It is easy to describe the precise nature of Wigner's results by remarking that they follow from a general theorem whose statement (but not its proof)

can be immediately guessed from the theorem about the representations of *finite* semi direct products $N \otimes H$ in which we have formulated the results of Frobenius.

Indeed let $G = N \otimes H$ where N and H are arbitrary locally compact groups, N is commutative and the α_h act continuously on N. Suppose that the notion of induced representation has been generalized so that U^L is a well defined unitary representation of G, whenever L is a finite or infinite dimensional unitary representation of an arbitrary closed subgroup K of G (and not just for those in which K has finite index in G). Then it is obvious how to formulate an exact analogue of the Frobenius theorem, and the only question is whether it is true or not when N is the four dimensional group of all space-time translations and H is the Lorentz group acting on N in the natural way. Wigner's result asserts that it is, so that the problem of classifying the ordinary unitary irreducible representations of the group is reduced to classifying the ordinary unitary irreducible representation of H_χ for each χ in $\widehat{N}$. While there are a continuous infinity of characters χ in $\widehat{N}$, the H_χ are largely mutually isomorphic groups. Indeed there are only three different isomorphism classes in addition to H itself.

These are (a) the Lorentz group in three space time dimensions, (b) the two-dimensional Euclidean group, and (c) the rotation group in three dimensions. Both H itself and the Lorentz group in three space time dimensions lead to representations which, it could be plausibly argued, could not occur in physical problems. The representations of the rotation group in three dimensions were well known to Wigner because of his work in atomic spectroscopy and only the two-dimensional Euclidean group presented a problem. Note, however, that this group is also a semi-direct product – the semi-direct product of the translations and rotations in the plane. Moreover the group of rotations in the plane is commutative so that none of the H_χ in this group present problems. Thus did Wigner classify the "physical" irreducible unitary representations of the inhomogeneous Lorentz group and reduce the determination of the rest to the solution of the corresponding problem for the homogeneous Lorentz group, in three and four space time dimensions – or, essentially equivalently, the groups $SL(2, R)$ and $SL(2, C)$.

What about the projective representations? We shall content ourselves with the remark that there is a straightforward generalization of the inducing process and of the Frobenius theorem, although the latter is less straightforward when the projective multiple is non trivial on N. In Wigner's case all projective multipliers are trivial on N.

A much more serious difficulty is that of extending the proof of the Frobenius theorem to the case of semi direct products $N \otimes H$ in which H is not only infinite but continuous. The problem is that, in general, this forces the irreducible unitary representations of $N \otimes H$ to have restrictions to N which are not discretely decomposable, and direct integral theory in some form must be used. More precisely the irreducible unitary representations associated with the character χ_0 of N reduce, when restricted to N, to a multiple of a sum (or integral) of all members of $\widehat{N}$ of the form $[\chi_0]\alpha_h$, where h varies over H.

Whenever there are a continuum of right H_{χ_0} cosets an integral will be called for. The problem of dealing with direct integrals of Hilbert spaces already occurs in "diagonalizing" self-adjoint operators with continuous spectra, and, as far as bounded operators are concerned, was completely solved by Hilbert with his spectral theorem in the first decade of the twentieth century. The ideas and techniques of the spectral theorem are just what is needed for dealing rigorously with the problems at hand. These ideas and techniques have however always been considered "too abstract" by the great majority of physicists; they prefer to make do with (unrigorous) use of Dirac δ functions. Accordingly, Wigner's paper is incomplete at this crucial point. He is quite explicit about this and gives a sketch of the argument with a promise that full details will appear in a forthcoming paper of von Neumann. An incomplete manuscript of this promised paper was found among von Neumann's manuscripts, after his death in 1956, and published as it stood as part of his "Nachlaß" in his collected works. In the meantime, it had been rendered superfluous by more general work of others. More details will be found below.

It should be mentioned that the formulation in terms of induced representations is not given in Wigner's paper, nor in that of Seitz. They simply apply Frobenius' method of analysis which physicists today sometimes call the "little group" method. The H_χ are the little groups. Using the language of induced representations makes the nature and unity of the results more apparent.

In 1940 Wigner published a short note, in collaboration with von Neumann [22], which is not closely related to any of his other papers, nor to any physical applications. Its chief content is the presentation of three examples of infinite discrete groups which have different properties as far as the existence of almost periodic functions is concerned. It is an example of his interest in occasional forays into mathematics for its own sake.

A year later in 1941 he published another short note [23] of a similarly pure and (apparently) isolated character. However, the main result is surprising – almost bizarre – and the paper introduces a theme to which Wigner was to return three times beginning in 1968. The paper is about finite groups which have the following two properties:

(a) Every group element is conjugate to its inverse.

(b) If L_1 and L_2 are any two irreducible representations then their tensor (Kronecker) product is a direct sum of mutually inequivalent irreducible representations – in later terminology is "multiplicity free".

Wigner called such groups *simply reducible* (abbreviated S.R.). The main result, Theorem 2, characterizes simply reducible groups in terms which do not invoke representation theory. For each group element R, let $\rho(R)$ denote the number of square roots of R; that is, the number of group elements X such that $X^2 = R$. Let $v(R)$ denote the number of elements X such that $XR = RX$. Theorem 2 asserts that the finite group G is simply reducible if and only if $\sum_R \rho(R)^3 = \sum_R v(R)^2$ and that, in any case, the left hand side of the equation is not greater than the right hand side.

A second paper on S.R. groups [24], believed to have been written around the same time, circulated as a mimeographed typescript but did not appear in print until 1965 when Biedenharn and van Dam included a slightly revised version in their anthology "Quantum theory of angular momentum". This paper is rather longer than its predecessor and is entitled "On the matrices which reduce the Kronecker products of S.R. groups". It begins with a definition of simply reducible group, but now applied to compact groups, as well as finite groups. The main point seems to be that some of the most important properties of the representation theory of the compact groups $SO(3)$ and $SU(2)$ which play a role in atomic spectroscopy, follow from the fact that these groups are S.R. groups and may be proved for all S.R. groups. Taking advantage of this Wigner reorganized the theory and presented it in abstract form before specializing to the groups $SO(3)$ and $SU(2)$ of interest to physics. At the same time he carried the abstract theory further than the concrete theory had been carried earlier and in so doing anticipated a part of the very significant work done by Giulio Racah in the later 1940's. Specifically, he defined and established some of the properties of what are now known as *Racah coefficients*.

To be a bit more specific, let G be any simply reducible group, let L^1 and L^2 be irreducible representations of G, and let L^3 be any irreducible representation of G which is equivalent to a constituent of the decomposition of the Kronecker product $L^1 \otimes L^2$ of L^1 and L^2. This representation takes place in the tensor product $\mathcal{H}(L^1) \otimes \mathcal{H}(L^2)$ of the Hilbert spaces of L^1 and L^2, and because of the hypothesis of simple reduciblity, the constituent equivalent to L^3 takes place in a uniquely determined subspace $\mathcal{H}_{L_3}$ of $\mathcal{H}(L^1) \otimes \mathcal{H}(L^2)$. Thus, if φ_1, φ_2, and φ_3 are arbitrary vectors in $\mathcal{H}(L^1)$, $\mathcal{H}(L^2)$, and $\mathcal{H}_{L_3}$ then one can ask for the scalar product of φ_3 with $\varphi_1 \times \varphi_2$. This scalar product is a complex number which is a function of the six variables $L^1, L^2, L^3 \cdot \varphi_1, \varphi_2, \varphi_3$. When normalized by dividing by $\sqrt{d(L^3)}$, where $d(L^3)$ is the dimension of $\mathcal{H}(L^3)$, it is what is known variously in the literature as a Wigner coefficient, a Clebsch-Gordan coefficient or a $3j$ symbol. The latter term is used especially for the denotation $\begin{pmatrix} L^1, L^2, L^3 \\ \varphi_1, \varphi_2, \varphi_3 \end{pmatrix}$ for the number in question and refers to the fact that a common notation for the $2j+1$-dimensional representation of $SO(3)$ in $SU(2)$ is $\mathcal{D}^j$. We remark that the unitary equivalence between L^3, given externally, and the appropriate subspace of $\mathcal{H}(L^1) \otimes \mathcal{H}(L^2)$ is unique only up to multiplication by an arbitrary complex number of modulus one. Thus if φ_3 is chosen from $\mathcal{H}(L^3)$ rather than from $\mathcal{H}_{L_3} \subseteq \mathcal{H}(L^1) \otimes \mathcal{H}(L^2)$ the coefficients $\begin{pmatrix} L^1, L^2, L^3 \\ \varphi_1, \varphi_2, \varphi_3 \end{pmatrix}$ are ambiguous up to a (universal) multiplicative constant of modulus one. In practice one chooses canonical bases in $\mathcal{H}(L^1)$, $\mathcal{H}(L^2)$, and $\mathcal{H}(L^3)$ and chooses φ_1, φ_2, and φ_3 from the basis elements.

When one is making actual calculations of atomic spectra by perturbation theory, these coefficients enter in an essential way and one needs efficient methods for their explicit computation. This can be done by recursion formulae and by exploiting various symmetries under permutations of the variables. The rather elaborate theory that has developed was begun in Wigner's

1931 book and repeated, in Lie algebra language, in Condon and Shortley's 1935 treatise "The Theory of Atomic Spectra". The further developments in Wigner's abstract theory result from applying similar considerations to triple Kronecker products, such as $(L^1 \otimes L^2) \otimes L^3$, and comparing $(L^1 \otimes L^2) \otimes L^3$ with $L^1 \otimes (L^2 \otimes L^3)$. The analysis of this situation leads to a complex valued function of *six* irreducible representations whose value at any particular sextuple $L^1, L^2, L^3, L^4, L^5, L^6$ Wigner denotes by $\left\{ \begin{matrix} L^1, L^2, L^3 \\ L^4, L^5, L^6 \end{matrix} \right\}$ (using curly brackets) and calls a $6-j$ symbol. The $6-j$ symbols have recursion relations and symmetries just as the $3-j$ symbols do. While the $6-j$ symbols may be expressed in terms of the $3-j$ symbols, this is not the best way of dealing with them.

To set this paper of Wigner in its proper context it seems wise to digress briefly and say a few words about the work of Guilio Racah who was inspired by the 1929 paper mentioned earlier, in which Slater showed how to do almost everything in atomic spectroscopy which Wigner had done, without explicit use of groups or their representations. It turned out that Slater's methods, while valid, led to prohibitively long and tedious calculations whenever more than a very few electrons were involved. Racah set out to find short cuts, and succeeded to an astonishing degree, by applying a number of ingenious arguments and methods. The results were published in a series of four papers [25], [26], [27], and [28] with the running title "Theory of complex spectra X" where $X = I$, II, III, and IV. These papers appeared in the *Physical Review* in 1942, 1942, 1943, and 1949, respectively. They are much too complicated for brief summary. The interested reader is referred to a lengthy article by the author of this essay for further details. This article appears as an introduction to the two volume contribution of L. C. Biedenharn and J. D. Louck to the *Encyclopedia of Mathematics and its Applications*. For editorial reasons it is published at the beginning of the second volume (Volume 9 of the Encyclopedia). A description of Racah's four papers from the point of view of the theory of group representations occupies Sections 20, 21, 22, and 23. For our present purposes it will suffice to say that in the second of his four papers Racah introduces a slight variant of the $6-j$ symbols (now called the Racah coefficients) and uses them to make a considerable simplification in one of his formulae. They find a rather different use in a later paper. It is interesting that Wigner found the $6-j$ symbols while theorizing about S.R. groups while Racah found essentially the same theory while working on practical problems in atomic spectroscopy. We remark here that, slightly later, Racah found that his work could be applied to nuclear structure for much the same reasons that made it possible for Wigner to adapt his approach to atomic spectroscopy to nuclear structure in his 1937 paper on that subject [20].

It is interesting that although Racah began with the anti-group theoretical approach initiated by Slater, and developed and made standard by the book of Condon and Shortley, he finally capitulated in Part IV and brought in group theory explicitly.

In December 1941 the United States entered World War II and, in April 1942, Wigner went to work in Chicago on theoretical work connected with the

design of an atomic bomb. His work on groups and group representations was interrupted and no more papers by him on the subject appeared until 1948.

In 1948 and 1949 Wigner returned to the theme of his 1939 paper on the unitary representations of the inhomogeneous Lorentz group by publishing closely related joint papers with V. Bargmann and T. D. Newton respectively [29] and [30]. The first of these is entitled "Group theoretical discussion of relativistic wave equations" and a footnote at the very beginning informs us that "All the essential results of the present paper were obtained by the two authors independently, but they decided to publish them jointly". The idea is to exploit the fact that every relativistically invariant quantum mechanical system is intrinsically associated with a unitary representation (or projective representation) of the inhomogeneous Lorentz group in the simplest possible special case – that in which the system consists of a single particle moving freely in space. In this case one can argue that the representation must be irreducible and hence that the classification of such representations, given by Wigner in 1939, provides a classification of particles. Given any particular class, one can deduce a great deal about the quantum mechanics of particles in this class directly from the properties of the associated representation. In particular, having chosen a coordinate system, one can deduce the Hamiltonian operator H by simply restricting the representation to the one parameter subgroup of time translations and computing the infinitesimal generator of this one parameter unitary group. Given H one can write down the relativistic Schrödinger equation, referred to here as the "wave equation". Replacing the time translation by the translations in some fixed direction in space, the same procedure leads to the operators associated with the linear momentum in that direction. The authors do this in detail on a case by case basis. The most important case is that in which the "little group" is $SO(3)$. In that case a real number describing the orbit of χ is the rest mass of the particle and the irreducible representation (or projective representation) of $SO(3)$, which completes the description of the representation in question determines, its spin. Notice how much of the physics of a free relativistic particle is deducible from symmetry principles combined with group representations (Weyl's program!!).

Wigner and Bargmann have nothing to say about the operators corresponding to the position observables. This is a somewhat tricky question and is the subject of the cited 1949 paper of Wigner and T. D. Newton entitled "Localized States for Elementary Systems". Among other things, it turns out that not all of the physically acceptable irreducible unitary representations of the inhomogeneous Lorentz group can define particles which possess any such thing as a position observable. As discovered some four or five years later by Arthur Wightman, the considerations of Wigner and Newton can be made mathematically rigorous and easier to understand by introducing a certain concept from the theory of unitary group representations. This concept was first defined by the author of this essay in a paper [31], submitted to the editors in July of 1949, the same month in which the paper of Wigner and Newton appeared in print. This concept is that of a continuous system of imprimitivity for a group representation. The main point of [31] was to state and outline the proof of a

general theorem known as "the imprimitivity theorem" which has as corollaries many theorems about group representations which are of interest in physics. These include the main results in Wigner's 1939 paper [21] and the results of Stone and von Neumann on the uniqueness of the irreducible solutions of the Heisenberg commutation relations. Because of its multiple connections with Wigner's work and its use in clarifying the Wigner-Newton paper, it seems appropriate to digress briefly to define the concept of system of imprimitivity and state the imprimitivity theorem.

The notion of *discrete* system of imprimitivity for unitary representations of finite groups was known to Frobenius and is defined as follows. Let U be a unitary representation of the finite group G in the Hilbert space $\mathcal{H}(U)$ and let $\mathcal{H}(U) = \mathcal{H}_1 \oplus \mathcal{H}_2 \oplus \ldots \oplus \mathcal{H}_n$ be a direct sum decomposition of $\mathcal{H}(U)$ into mutually orthogonal subspaces $\mathcal{H}_1, \mathcal{H}_2 \ldots \mathcal{H}_n$. If U_x maps each $\mathcal{H}_j$ into itself for all x, we say that U itself is a direct sum of subrepresentations defined by the $\mathcal{H}_j$. One says that the $\mathcal{H}_j$ define a *discrete system of imprimitivity* for U, if the following weaker condition holds. For each j and each x in G, $U_x(\mathcal{H}_j) = \mathcal{H}_k$ for some k depending on j and x, but not necessarily equal to j. In other words, while the U_x do not necessarily leave the $\mathcal{H}_j$ fixed, they preserve their individuality and merely permute them amongst themselves. One says that a discrete system of imprimitivity $\mathcal{H}_1, \mathcal{H}_2 \ldots \mathcal{H}_n$ for U is *transitive*, if for each i and j there exists an $x \in G$ with $U_x(\mathcal{H}_i) \equiv \mathcal{H}_j$. Evidently, every discrete system of imprimitivity may be broken up uniquely into subsets on each of which the U_x act transitively. Correspondingly, U breaks up as a direct sum of subrepresentations each of which has some subset of the $\mathcal{H}_j$ as a *transitive* discrete system of imprimitivity. It follows that one need only study *transitive* discrete systems of imprimitivity to know them all. Suppose, then, that $\mathcal{H}_1, \mathcal{H}_2, \ldots \mathcal{H}_n$ defines a transitive discrete system of imprimitivity for U. Let K denote the set of all $x \in G$ for which $U_x(\mathcal{H}_1) = \mathcal{H}_1$. Then K is evidently a subgroup of G and the restriction of U to K defines a representation L of K in $\mathcal{H}_1$. It is easy to check that if we know only the representation L of K and they way in which G permutes the H_j we may reconstruct U and verify that it is equivalent to the representation of G induced by the representation L of K. Conversely, if K is *any* subgroup of the finite group G, and L is any unitary representation of K, then one can construct a transitive discrete system of imprimitivity for the induced representation U^L of G, as follows. Recall that the space $\mathcal{H}(U^L)$ of U^L is the set of all functions f from G to $\mathcal{H}(L)$ such that $f(kx) = L_k(f(x))$ for all $x \in G$ and all $k \in K$. Recall also that one defines a right K-coset of G to be the subset Kx of all kx for fixed x and arbitrary k in K, and that two right K-cosets Kx and Ky are either disjoint or identical. Given f in $\mathcal{H}(U^L)$, we may write it uniquely as a sum $f_1 + \ldots + f_n$ where each f_j is zero except on a single right K-coset. It is an immediate consequence of the definitions that each f_j is also in $\mathcal{H}(U^L)$ so that $\mathcal{H}(U^L)$ is a direct sum of subspaces one for each right K-coset, and consisting of all members of $\mathcal{H}(U^L)$ which vanish outside of this coset. The proof that this decomposition is a discrete system of imprimitivity for U^L is a straightforward verification. We shall call it the canonical system of imprimitivity for U^L. The so-called "imprimitivity theorem" (in the

special case of finite groups) asserts the existence of the canonical system of imprimitivity just described together with the key fact that *every* transitive system of imprimitivity arises in this way from an induced representation. The above remarks are an outline of a proof which it is easy to fill in. One may regard the theorem as an abstract characterization of induced representations. However, it is more than just an elegant theorem. Systems of imprimitivity arise naturally in various parts of mathematics and physics, and the fact that this implies that the representation is induced by a lower dimensional representation of a subgroup is quite useful. In particular Frobenius' analysis of the irreducible representations of semi-direct products of finite groups is a corollary of the imprimitivity theorem for finite groups.

Returning now to the 1949 paper [31] whose title is "Imprimitivity for representations of locally compact groups I" we now indicate what is involved in generalizing the discrete imprimitivity theorem, whose (quite elementary) proof we have just sketched, so as to apply to general separable locally compact groups. Actually, the only serious difficulty in making the generalization occurs (just as in generalizing the Frobenius semi-direct product theorem) in dealing with induced representations in which the set of right K-cosets is not discrete. When this happens, a system of imprimitivity can no longer be a direct sum decomposition of the representation space but must be a direct integral or continuous direct sum decomposition. As far as the basic definitions (but not the proofs) are concerned the difficulty can be met by introducing the concept of a "projection valued measure". This notion is not familiar to most physicists but is simple and easy to explain. Let $\mathcal{H} = \mathcal{H}_1 \oplus \mathcal{H}_2 \oplus \ldots$ be an ordinary discrete sum decomposition of a Hilbert space $\mathcal{H}$. For each subset E of the real line, let P_E be the projection operator in $\mathcal{H}$ whose range is the direct sum of all $\mathcal{H}_j$ for which j is a member of E. If $E = \emptyset$ (the empty set) let $P_E = 0$. Then $E \to P_E$ is a function defined on all subsets of the real line, having projection operators in the Hilbert space $\mathcal{H}$ as values. It is easy to check that this function has the following properties:

(a) For all E_1 and E_2, $P_{E_1}P_{E_2} = P_{E_2}P_{E_1} = P_{E_1 \cap E_2}$.
(b) Whenever the sets $E_1, E_2 \ldots$ are pairwise disjoint then $P_{E_1 \cup E_2 \cup \ldots} = P_{E_1} + P_{E_2} + \ldots$ where the sequence may be infinite.
(c) $P_R = I$ and $P_\emptyset = 0$ where R is the whole real line.

(a), (b), and (c) are the defining properties of a projection valued measure. In general it is not defined on *all* subsets but only on the measurable or Borel subsets of some measure space or Borel space which need not be the real line. In this example there is a countable subset, namely the positive integers Z^+ such that $P_{Z^+} = I$. Whenever this happens, one has a direct sum decomposition of the Hilbert space from which P can be reconstructed as in the example above. However, such examples are exceptional and the non exceptional examples may be used as substitutes for continuous direct sum decompositions. Consider the Hilbert space $\mathcal{H}$ of all square summable complex valued functions on the unit interval $0 \leq x \leq 1$. For each measurable subset E of $0 \leq x \leq 1$ let P_E denote the linear operator which takes f in $\mathcal{H}$ into $\varphi_E f$, where $\varphi_E(x) = 1$ if x is in E

and $\varphi_E(x) = 0$ if x is not in E. It is easy to see that each P_E is a projection operator in $\mathcal{H}$ and that the function taking E into P_E is a projection valued measure, in the sense that it satisfies (a), (b), and (c) above. On the other hand, for every E which has only countably many elements, $P_E = 0$, so no discrete direct sum decomposition is associated with it.

Having the notion of projection valued measure defined, it is now easy to generalize the notion of system of imprimitivity in the appropriate way. Let V be a unitary representation of the separable locally compact group G. Let $\mathcal{M}$ be any set on which G acts as a group of transformations and let $[m]x$ denote the transform of m in $\mathcal{M}$ by x in G. Then by definition of transformation group $[m]x$ is in $\mathcal{M}$ for all x and m, $[[m]x]y = [m]xy$ and $[m]e = m$ for all x, y in G and all m in $\mathcal{M}$. As usual e denotes the identity element of G. Suppose that $\mathcal{M}$ has part of the structure of a measure space, in the sense that a family of "measurable" sets has been defined which is closed under complementation and countable union, and that the transformations $m \to [m]x$ preserve this structure. One says that a projection valued measure defined on the measurable subsets of $\mathcal{M}$, and whose values are projections in $\mathcal{H}(V)$, is a system of imprimitivity for V, provided that the V_x and the P_E satisfy the following simple commutation relation

$$(*) \qquad V_x P_E = P_{[E]x^{-1}} V_x$$

where $[E]x^{-1}$ denotes the transform of the set E by x^{-1} (that is, the set of all $[m]x^{-1}$ for $m \in E$).

To see the connection of this definition with that of the very special discrete case considered above, suppose that $\mathcal{M}$ is the set Z^+ of all positive integers and every subset of $\mathcal{M}$ is measurable. Let E_j be the set whose only element is j and let $\mathcal{H}_j$ be the range of P_{E_j}. One verifies at once that $\mathcal{H}(V)$ is the direct sum of the $\mathcal{H}_j$ and that $(*)$ is equivalent to the assertion that each V_x permutes the $\mathcal{H}_j$ among themselves.

The reader should now have no trouble in formulating the general imprimitivity theorem which is the principal result of [31]. The proof is not so easy and this is not the place to give or even outline it. The chief difficulty is constructing the representation L of K which induces V. The method used in the discrete case fails because the anlogue of $\mathcal{H}_1$ is an "infinitesimal" slice of $\mathcal{H}(V)$ and is only defined "almost everywhere". Once the proof has been carried out, however, one has done all the difficult analysis necessary to rigorize the arguments which Wigner sketched in his paper [21] and said that von Neumann would fill in later. Indeed, there is an easy corollary of the imprimitivity theorem of [31] which generalizes the Frobenius theorem, stated earlier, to any separable locally compact semi direct product $N \otimes H$, in which N is commutative and in which the action of H on $\widehat{N}$ defined by the α_n^* divides $\widehat{N}$ into orbits in a manner which is not too wildly chaotic. Specifically it suffices that there should exist a cross section for the orbits which is a Borel set. When this regularity condition does not hold, the Frobenius classification still applies to a large subset of the irreducible unitary representation, but there are other representations which

can only be partially classified and ergodic theory enters the picture in a way which it would take us too far afield to describe.

The imprimitivity theorem of [31] has another application which is directly relevant to this essay and is historically interesting because it is this application which inspired the writer to discover the theorem as well as rediscover and generalize the notion of induced representation. It turns out that the commutation relation (*) may be looked upon as a generalization of the Heisenberg commutation relations in the global form, given them by Weyl in his 1927 paper [10], introducing group representations into physics and inaugurating what we have called Weyl's program. The uniqueness theorem for the irreducible solutions of these relations, formulated by Weyl and proved rigorously by Stone and von Neumann (see above), may in fact, as already mentioned above, be obtained as a corollary of the imprimitivity theorem. The same is true of the theorem of Jordan and Wigner [11] on the uniqueness of the solution of the canonical anti-commutation relations. However, in this case the finite discrete version of the imprimitivity theorem suffices.

Wightman's observation that the paper [30] of Wigner and Newton could be clarified and made more rigorous by using the notions of [31], although made in the early 1950's, was not published until 1962 when Wightman included it in a long paper [32] in "Reviews of Modern Physics" dedicated to Wigner on the latter's sixtieth birthday. (It is amusing that the Wigner-Newton paper was published in the same journal and dedicated to Einstein on *his* seventieth birthday.) Wightman's paper also includes a detailed exposition of the results of [31], in the special case of the group $\mathcal{E}^3$ generated by the translations and rotations in Euclidean 3 space. This group is a semi direct product of its translation and rotation subgroups and is a subgroup of the inhomogeneous Lorentz group. To get the operators corresponding to position observables, when they exist, Wightman restricts the irreducible representation of the inhomogeneous Lorentz group to $\mathcal{E}^3$ and finds that it has a system of imprimitivity $E \to P_E$ based on the action of the Euclidean group on physical space. The projection operator P_E then corresponds to the observable which is one, when the particle is observed to be in E, and zero when it is not.

Long before Wightman's paper was published, probably in the very early 1950's, the author of [31] (and this essay) received a postcard from our mutual friend, Irving Segal, informing him in rather vague terms that Wightman had found an application of [31] to [30]. The author of [31] found [30] incomprehensible but was engaged at the time in trying to achieve a better understanding of quantum mechanics. Eventually he was led to ideas which he assumed must be close to Wightman's, and this was verified in 1955 when he and Wightman met for the first time and discussed the matter. His approach was rather different from Wightman's and started with an axiomatic definition of the notion of a quantum mechanical particle. Further details will be found in #10 of the published version [33] of the author's 1961 American Mathematical Society Colloquium lectures and in several later publications. Suffice it to say that the commutation relation (*) can be justified *a priori* as following from applying invariance principles to position observables and that the imprimitivity theo-

rem then allows one to *deduce* the usual realizations of position, momentum and angular momentum operators. Spin appears automatically when one analyzes the angular momentum operators. (This completes our digression and we return to the main theme.)

Three years later Wigner began a brief collaboration with E. Inönü which resulted in three closely related papers [34], [35], and [36]. In the first of these the authors studied the unitary representation theory of the so-called inhomogeneous Galilei group. This is the group of symmetries of space time as it was understood before the discovery of special relativity. The authors wished to explore the connections between its representation theory and the ordinary non-relativistic Schrödinger equation for a free particle, in analogy with the work of Wigner and Bargmann, and of Wigner and Newton. The Galilei group is a semi-direct product to which the theory described above may be applied and the authors found no difficulty in determining its irreducible unitary representation. However, they found the (at first rather surprising) result that none of them were associated with position operators or with the usual non-relativistic free particle Schrödinger equation.

The second paper is entitled "On the contraction of groups and their representations" and connects with the first through the final Part III. In this part, the fact that the Galilei group is a contraction of the inhomogeneous Lorentz group is exploited to throw light on the surprising results of the first paper. Actually, the fundamental source of the mystery is that in physical applications projective representations must be considered as on a par with ordinary representations and that they play a much bigger role in the Galilei group than in the inhomogeneous Lorentz group. Specifically, up to equivalence, there is only one non-trivial projective multiplier for the inhomogeneous Lorentz group and it is trivial on the translation subgroup. By contrast, the Galilei group has continuum many inequivalent projective multipliers which are non trivial when restricted to the translation subgroup. Moreover all the particles which have mass and position operators correspond to projective representations. Indeed different masses come from representations with inequivalent projective multipliers. A full account of this aspect of things was published by V. Bargmann a year later [37]. What Wigner and Inönü point out in Part III of their paper is that in the passage to the limit described in Part II, ordinary representations can pick up a non-trivial projective multiplier and do so in the case at hand.

Parts I and II are of a more general and purely mathematical character. The contraction notion is defined in Part I and the effect of contraction on representations is discussed in Part II. The basic idea may be described as follows. Let G be a connected Lie group and let S be a closed Lie subgroup so that the space G/S of right S cosets has the structure of a differentiable manifold and right multiplication gives us an action of G on G/S ($[Sy]x = S(yx)$). Let the right coset $Se = S$ be chosen as an "origin" and let T denote the tangent space to the manifold at this origin. Then since each element x of S leaves Se fixed, it carries T onto T and defines a linear transformation in the vector space T. The semi-direct product of the additive group of T and the subgroup S defined by the just described action of S on T is what Wigner and Inönü call

the *contraction* of G by S. The Galilei group, for example, is the contraction of the inhomogeneous Lorentz group by the group generated by spatial rotations and time displacements, and the Euclidean group in two dimensions is the contraction of the rotation group in three dimensions by any one parameter subgroup. One thinks of the contraction as a sort of limiting form of the original group G in much the same sense that Galilian relativity is a limiting form of special relativity or that ordinary Euclidean space is a limiting form of a three dimensional curved Riemannian space. In the latter connection note that the coset space of $SL(2,C)$ by the subgroup $SU(2)$ of all unitary matrices is a three dimensional Riemannian manifold in a natural way and that the contraction of $SL(2,C)$ by $SU(2)$ is a double covering of the Euclidean group $\mathcal{E}^3$. Because of this interpretation, one can hope to regard unitary representations of contractions as limiting forms of unitary representations of the contracted group and this possibility is explored in Part II. Actually, the authors find it easier to discuss all of these issues in terms of the Lie algebras of the groups where convergence questions can be put into more concrete form. They call attention to a paper of I. E. Segal [38], published a year earlier, in which quite similar ideas were discussed. This paper of Segal also contains ideas close to some of those of the paper [31] on the imprimitivity theorem and was cited in the immediate predecessor of [31]. The contraction idea turned out to be of some importance and was further developed later. See, for example, [39] and the papers cited therein.

The third paper of Wigner and Inönü is mainly in the nature of a technical correction to an argument in the second paper.

These three papers were published in 1952, 1953, and 1954. In 1954 Wigner also published a one-page note [40] presenting a new way of looking at the Racah coefficients. In 1955 Wigner gave a course of lectures at Princeton. Thirteen years later these lectures were written up and published by J. D. Talman [41] under the title "Special functions – A group theoretic approach". The book begins with an elementary introduction to Lie groups, Lie algebras and their representations. In the later chapters there are some examples illustrating the known fact that many of the classical "special functions" appear as matrix coefficients of the irreducible representations of familiar Lie groups and that this fact can be used to elucidate some of the properties of these functions.

After 1955 Wigner put aside the theory of groups and group representations for over a decade, except for three isolated and rather minor contributions in 1957, 1964, and 1967. Then in 1968 he returned to the subject long enough to write three papers related to his work on S.R. groups and two survey articles. We conclude this part of our essay on Wigner's mathematical papers with brief summaries of these eight contributions.

The 1957 paper [42], written jointly with E. Hewitt, is in the spirit of generalizing theorems about the additive group of the real line to arbitrary locally compact commutative groups. W. Magnus published a paper in 1955, generalizing the Plancherel formula and the Fourier inversion theorem on the line from complex valued functions to matrix valued functions. In [42] Wigner and

Hewitt simplified the inversion formula and showed that the theorem is true in the generality indicated above.

The 1964 paper [43] is a conference report which is mainly a review of Wigner's work on the inhomogeneous Lorentz group but includes an analysis of what happens when spatial reflections are included.

The 1967 paper [44] was written jointly with A. W. Saenz. It is mainly devoted to generalizing the well known theorem according to which the Lorentz group is isomorphic to the quotient of the group $SL(2, C)$ by its center. In this generalization the Lorentz group in 3-spacelike dimensions is replaced by the Lorentz group in $n \geq 2$ spacelike dimensions. The results differ slightly depending upon whether n is even or odd.

The tensor product $L \otimes M$ of two irreducible representations of a finite group G is the restriction of the irreducible representation $L \times M$ of $G \times G$ to the G isomorphic subgroup $\widetilde{G}$ of $G \times G$ consisting of x, y with $x = y$. This subgroup is often called the diagonal subgroup. Since every irreducible subrepresentation of $G \times G$ is of the form $L \times M$, to say that every tensor product $L \otimes M$ of irreducible representations of G is multiplicity free is equivalent to saying that every irreducible representation of $G \times G$ is multiplicity free when restricted to the diagonal subgroup $\widetilde{G}$. Now this condition is the main component of the definition of simply reducible groups. Thus in view of the results of Wigner's 1941 paper [23] characterizing SR groups, it is not unreasonable to seek some characterization of finite group subgroup pairs $H \subset G$ with the property that every irreducible representation of G is multiplicity free when restricted to H. Wigner sought and found such a characterization and published it in 1968 [45]. The result is simple and elegant enough to be quoted. Let us generalize the familiar conjugacy class notion by defining two elements x and y of G to be in the same H class if $x = hyh^{-1}$ for some h in H. Wigner's necessary and sufficient condition is, then, that the set theoretic products of the H classes should commute with one another. It should be mentioned that Wigner's motivation was not the purely mathematical one indicated here. In applying a combination of perturbation theory and symmetry considerations to physical problems one often is confronted with restricting a representation of an approximate symmetry group to a more exact symmetry subgroup. To have this restricted representation turn out to be multiplicity free is very useful. Indeed, Wigner's original reason for being interested in S.R. groups is of this character.

Three years later, in 1971, Wigner published a paper [46] putting the result of the 1968 paper in a larger context. Specifically he studied the representations obtained by restricting the irreducible representations of a finite group G to a subgroup H in a systematic way which led to a general relationship connecting properties of the H classes with the multiplicities occuring in the restrictions of irreducible representations of G. The main result of the 1968 paper comes out as a corollary.

A year later in 1972 Wigner, in collaboration with F. E. Goldrich, published a third and final paper [47] on the subject. In it the authors prove a version of the theorem of the 1968 paper which applies to compact groups. The for-

mulation of the theorem has to be changed because Wigner's definition of the product of two H classes in the 1968 paper involves attaching multiplicities of occurrence to the elements in the product. The authors apply their result to obtain a new proof of the known fact that irreducible unitary representations of $U(n)$ are multiplicity free when restricted to $U(n-1)$. In the final few lines they give a further condition for the multiplicity freeness of restricted representations which is sufficient but not necessary.

In the interests of unity, it seems appropriate to mention that in the early fifties the writer of this essay found an approach to Wigner's 1941 characterization of S.R. groups through the theory of induced representations and published it in [48]. It turns out that the two sides of the mysterious equation $\sum_{x\in G}\rho(x)^3 = \sum_{x\in G} v(x)^2$ have a simple meaning in terms of the $\widetilde{G}_3 : \widetilde{G}_3$ double cosets in the triple product $G \times G \times G$. Here $\widetilde{G}_3$ is the subgroup of all x, y, z with $x = y = z$. One can show that $\frac{1}{o(G)}\sum_{x\in G}\rho(x)^3$ is equal to the number of $\widetilde{G}_3 : \widetilde{G}_3$ double cosets in $G \times G \times G$ and that $\frac{1}{o(G)}\sum_{x\in G} v(x)^2$ is equal to the number of self inverse double cosets. Thus Wigner's criterion may be restated as: every $\widetilde{G}_3 : \widetilde{G}_3$ double coset is self inverse. In this form Wigner's theorem makes sense and is true for compact groups as well as finite ones. See also [49] for more work in this direction.

In 1968 Wigner gave the annual J. Willard Gibbs lecture before the American Mathematical Society and published it in the same year [50]. Its six sections have the titles: Introduction and summary, The evolution of the physical sciences, Crystal symmetry, The role of the group of rotations in three space, Decompositon of the tensor products of representations, Symmetry problems of particle physics. Of these the next to the last is by far the longest and most detailed (six pages). It is entirely about simply reducible finite groups and the $3-j$ and $6-j$ symbols. The section on crystal symmetry is a history going back to the work of Hensel in 1830. The final section is not quite two pages long and is mainly devoted to a rather sketchy description of the eight fold way of Gell-Mann and Ne'eman with emphasis on the Gell-Mann Okubo mass formula. The important work of Wigner, Bargmann, and Newton is treated rather briefly in the section on quantum mechanics and not mentioned at all in the final section.

We conclude this account of Wigner's work on group representations with a description of a second review lecture published five years later in 1973 [51]. It is entitled "Reduction of direct products and restriction of representations to subgroups: the everyday tasks of the quantum theorists". It begins with a three page section entitled "The difference between the roles which symmetry plays in classical and quantum mechanics". This is followed by ten pages (divided into two sections), giving an even fuller treatment of the material in the fifth section of Wigner's 1968 lecture and including some identities which Wigner suggests might be new. The final section is basically an account of Wigner's 1971 paper on restrictions to subgroups.

B. Wigner's Papers on Linear Operators and Other Questions of Linear Algebra

There are ten of these of a very heterogeneous character. All but two of them are short notes of from one to five pages and they are spread out irregularly in time from 1929 to 1970 with none in the 1940's and only one in the 1950's. We shall describe their contents in chronological order.

The first two, [52] and [53], appeared in 1929 and were written jointly with von Neumann. The point of the first is to present two examples of self adjoint operators of the form

$$-\frac{h^2}{8\pi^2 m}\left(\frac{\partial^2}{\partial x^2}+\frac{\partial^2}{\partial y^2}+\frac{\partial^2}{\partial z^2}\right)+V\left(\sqrt{x^2+y^2+z^2}\right)$$

in which the discrete spectrum overlaps the continuous spectrum so as to present an apparent paradox, from the point of view of the quantum mechanics of a particle moving under the influence of the potential V. The examples are chosen so as to illustrate two distinct ways in which the situation can arise. The second is concerned with how the eigenvalues of a self adjoint operator H, which depend upon one or more real parameters, vary when the real parameters vary continuously. For simplicity only operators in finite dimensional spaces are considered and special emphasis is placed on the question of how many parameters are necessary in order to be sure of causing two variable eigenvalues to blend into one. "In general" the authors find that there must be at least three.

The first of the two long papers (36 pages) [54] appeared in 1934 and was a joint effort of Wigner, von Neumann, and P. Jordan. From one point of view it is a development of certain ideas of Jordan published in 1932 and 1933 and summarized in Part A of the present paper. From another it is a study of a class of non associative finite dimensional commutative linear algebras culminating in a complete structure theorem analogous to the celebrated Wedderburn theorem for finite dimensional associative algebras. The motivating idea is that in quantum mechanics, in the von Neumann formulation, the observables are represented by self adjoint operators and bounded self adjoint operators constitute a commutative non associate algebra under ordinary operator addition, and with $H_1 \circ H_2 = H_1H_2 + H_2H_1$ as the product. The authors express the hope that their analysis will suggest a modification of the standard von Neumann formulation which will help overcome some of the difficulties arising in reconciling quantum mechanics with special relativity. They realize of course that they will have to extend their theory to the infinite dimensional case. Algebras with closely related properties have since been studied from the point of view of pure mathematicians and are called Jordan algebras. The style of the paper strongly suggests that von Neumann did most of the writing.

A year later Wigner published a paper [55] containing his contribution to a method in perturbation theory. It is known as the Brillouin-Wigner method and involves finding an equation of the form $E - E_0 = P(E - E_0)$ where P is an infinite series and determining the energy shift by solving this equation. The

point is that this series tends to converge more rapidly and in more cases than the explicit series of the older Rayleigh-Schrödinger method. Wigner explains the relationship of his work to that of Brillouin as follows:

"The infinite series was first found by L. Brillouin[1] who obtained it by an intuitive consideration of the usual scheme. He has already pointed out in his important paper that his series converges much more rapidly than the power series of Schrödinger. He has not investigated, however, the successive approximations and their relations to the actual problem." The footnote reads: Brillouin's paper was published in 1933 in *Journal de physique.*

Wigner's next contribution in this category [56] was published a full nineteen years later in 1954. It, too, is concerned with the effect of perturbations on the eigenvalues of self adjoint operators but now with the older Rayleigh-Schrödinger theory of such perturbations. Its aim is to present a family of examples in which the first approximation to the energy shift produced by the perturbation is zero, and the second approximation is finite and non zero for only one value of a certain parameter. For all other values the second approximation is either zero or infinity.

In 1960 Wigner published two short notes [57] and [58] on anti-unitary operators. In the first of these he states and proves the analogue for anti-unitary operators of the well known theorem describing the most general unitary operator with a discrete spectrum in terms of its eigenvectors and eigenvalues. In the other he describes a way in which unitary and anti-unitary operators differ, and which can be used to distinguish them in quantum mechanical applications. The motivation of this work stems from the fact that anti-unitary operators define automorphisms of the space of states in quantum mechanics.

Let φ_1 and φ_2 be two orthogonal unit vectors in the Hilbert space $\mathcal{H}$ and let T denote either a unitary or an anti unitary operator. Let φ_1 and φ_2 both be eigenvectors of T and for all real α let $\psi_\alpha = (\varphi_1 + e^{i\alpha}\varphi_2)/\sqrt{2}$. In the unitary case $|(\psi_\alpha \cdot T(\psi_\alpha))|^2$ is independent of α while in the anti unitary case $|(\psi_\alpha \cdot T(\psi_\alpha))|^2 = \frac{1}{2}(1 + \cos 2\alpha)$.

In a paper [59] published in 1963 Wigner generalized the notion of positive definite matrix by defining a matrix to be *weakly positive* whenever it is similar to a diagonal matrix all of whose elements are positive real numbers. He then gave the following elegant and somewhat surprising characterization of such matrices.

Theorem 1. *A matrix is weakly positive if and only if it can be written as a product of two positive definite matrices.*

In addition he proved two further theorems.

Theorem 2. *If the product of a positive definite and of a weakly positive matrix is Hermitian it is also positive definite. The same holds of the product $P_1P_2P_3$ of three positive definite matrices P_1, P_2, and P_3 if this product is Hermitian.*

Theorem 3. *Theorems* 1 *and* 2 *hold also for bounded linear operators on Hilbert space.*

In 1968 Wigner published a paper [60] entitled "On a generalization of Euler's angles". Let F be the matrix $\begin{pmatrix} I_n & 0 \\ 0 & -I_m \end{pmatrix}$ where I_k denotes the k-dimensional identity matrix and $n \geq m$. The two main theorems of the paper assert

(1) If U is any $(n+m) \times (n+m)$ matrix which satisfies $U^*FU = F$ then

$$U = \begin{pmatrix} u_1 & 0 \\ 0 & u_2 \end{pmatrix} \begin{pmatrix} C & S \\ \widetilde{S} & C^1 \end{pmatrix} \begin{pmatrix} v_1 & 0 \\ 0 & v_2 \end{pmatrix},$$

where u_1 and v_1 are n-dimensional unitary matrices, u_2 and v_2 are m-dimensional unitary matrices C and C' are real diagonal matrices, $\widetilde{S}$ is the transpose of S, both are real and $C^2 - S\widetilde{S} = I_n$, $C'^2 - \widetilde{S}S = I_m$.

(2) If R is any $(n+m) \times (n+m)$ real matrix such that $\widetilde{R}FR = F$ then it can be factored in the same way except that now u_1, v_1, u_2, v_2 are real orthogonal matrices.

The degree of non uniquencess of these factorizations is specified. The reference to Euler's angles in the title is to the α, β, and γ in the following theorem of Euler. Any real orthogonal 3×3 matrix may be factored as

$$\begin{pmatrix} \cos\alpha & \sin\alpha & 0 \\ -\sin\alpha & \cos\alpha & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \cos\beta & 0 & -\sin\beta \\ 0 & 1 & 0 \\ \sin\beta & 0 & \cos\beta \end{pmatrix} \begin{pmatrix} \cos\gamma & \sin\gamma & 0 \\ -\sin\gamma & \cos\gamma & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Wigner indicates at the beginning how he was led to these considerations in his well known R matrix theory of nuclear reactions (see Feshbach's essay on Wigner's work in nuclear physics); specifically from an attempt to extend R matrix theory so as to account for peripheral reactions more generally.

Wigner's final paper in this category is a short (two page) note [61] published in 1970 with H. C. Schweiner as co-author. Principal result is an alternative to the well known Schmidt procedure for orthogonalizing linearly independent sets of vectors. This alternative gives a result independent of the order in which the original vectors are put. The result is shown to maximize a certain quartic form on the original vector.

C. Wigner's Papers on Analytic Functions and Probability

Five of the eight papers in this category are concerned with a special class of analytic functions introduced by Wigner and called R functions. All five were published between 1951 and 1954. The remaining three deal with the eigenvalue distributions of random matrices and were published in 1955, 1957, and 1958. The 1957 and 1958 papers are brief supplements to the 1955 paper. All eight

are included in the same category to avoid one extremely small category and because they are connected in two ways. First, two of the papers on R functions have a strong probabilistic flavor. Secondly, the two sets of papers were inspired by two closely related aspects of Wigner's work on nuclear physics.

An R function is, by definition, a meromorphic function R defined on the whole complex plane and having the property that $R(z)$ is in the lower half plane, the real axis or the upper half plane according as z is in the lower half plane the real axis or the upper half plane. It follows easily from this definition that all the poles of an R-function are on the real axis, are simple, and have residues which are negative real numbers. For each R function Wigner defines $(dR/dz)/(1+R^2)$ to be its *invariant derivative* and denotes it by f' where the function f is consequently determined up to an additive constant. Integrating $(dR/(1+R^2) = f'(z)dt$ one concludes that $\arctan R(z) = f(z) + \varphi$ where φ is a constant and hence that $R(z) = \tan(f(z) + \varphi)$. Evidently the poles of f' which are not on the real axis occur at those z for which $R(Z) = \pm i$ and are symmetrically placed about the real axis.

These definitions and remarks are made in Section II of Wigner's first paper on R functions [62] which was published in the *Annals of Mathematics* in 1951 and entitled "On a class of analytic functions from the quantum theory of collisions". Section II contains several other theorems about R functions the most important of which are two so-called "normal form" theorems. The first of these states that every R function has an absolutely convergent Mittag-Leffler series expansion of the special form

$$R(z) = \alpha z + \beta + \sum_{\mu} \left(\frac{\gamma_\mu^2}{Z_\mu - z} - \frac{\gamma_\mu^2}{Z_\mu} \right)$$

where $Z_1, Z_2, \ldots$ are the positions of the poles of R, $-\gamma_1^2, -\gamma_2^2 \ldots$ are the residues of these poles, $\alpha > 0$ and β is real. Wigner presents a proof which he attributes to M. M. Schiffer and V. Bargmann. The other normal form theorem expresses R in terms of the poles above and below the real axis of the invariant derivative f'. Let the poles of f' which are not on the real axis be $x_1 \pm i\kappa_1, x_2 \pm i\kappa_2, x_3 \pm i\kappa_3, \ldots$ where $x_1, x_2 \ldots$ are real and κ_1, κ_2 are real and positive and let n_j be the order of the zero of $R - i$ at $x_j + i\kappa_j$. Then f' can be shown to be of the form

$$f'(z) = \eta z + \sum_{\mu} \frac{n_\mu \kappa_\mu}{(z - x_\mu)^2 + \kappa_\mu^2}$$

where $\eta > 0$, from which it follows that

$$R(z) = \tan f(z) = \tan \left\{ \eta z + \theta + \sum_{\mu} n_\mu (\arctan \frac{z - x_\mu}{\kappa_\mu} + \arctan \frac{x_\mu}{\kappa_\mu} \right\}.$$

Wigner presents a proof for which he gives credit to von Neumann. The fact that R can be so expressed implies that, up to the choice of the constants η

and θ, R is completely determined by the zeros of $R - i$ together with their multiplicities.

For the purposes of quantum mechanical collision theory Wigner was interested only in R functions with rather special properties. These are defined and discussed in Section III, Uniform R functions, and Section IV, Statistical R functions. Before reviewing these sections we shall give a brief explanation of the relevance of R functions to physics. Consider a spinless particle moving in a force field described by a potential function V defined in space and spherically symmetrical about some origin so that $V(x, y, z) = v(\sqrt{x^2 + y^2 + z^2})$. When v vanishes sufficiently rapidly at ∞, there will be motions of the particle such that, in the distant past, as well as the distant future, it will move like a free particle. However, as a result of the distortion of the free particle motion by the force field the free particle motion in the distant future will be different from that in the distant past. Under mild restrictions on v one can prove that this transformation from one free particle motion to another is completely described by a certain unitary operator S in the underlying Hilbert space $\mathcal{H}$. This operator is called the S operator or the scattering operator. Because of the assumed rotational invariance, the unitary operator S commutes with the natural representation of the rotation group $SO(3)$ in $\mathcal{H}$, and from this it follows that S breaks down as a direct sum of unitary operators, S^l, one for each equivalence class of irreducible unitary representations of $SO(3)$. These are parameterized by the non negative integers $0, 1, 2, \ldots$ and each unitary operator S^l turns out to be describable by a function $k \to s^l(k)$ from the positive real numbers to the complex numbers of modulus one. This is so because S^l is a direct integral of the one dimensional unitary operators defined by $s^l(k)$. It turns out that the computation of the functions $s^l(k)$ can be reduced to a study of the solutions of the differential equation

$$(*) \qquad -\frac{d^2\psi}{dr^2} + \left(v(r) + \frac{l(l+1)}{r^2}\right)y = k^2 y$$

for positive real k. Indeed in the case in which $v(r) = 0$ for $r > a$ one can show that

$$s^l(k) = e^{-2ika}\left(\frac{1 + ik\psi(a)/\psi'(a)}{1 - ik\psi(a)/\psi'(a)}\right)$$

where ψ is any solution of $*$ which is asymptotic at ∞ to e^{ikr}. Of course $\psi(a)/\psi'(a)$ depends upon k^2 as well as a. Let us denote it by $R_a(k^2)$. Thus we have

$$(**) \qquad s^l(k) = e^{-2ika}\left(\frac{1 + ikR_a(k^2)}{1 - ikR_a(k^2)}\right)$$

so that $s^l(k)$ and hence S^l is completely determined by knowing the function R_a for any a for which $v(r)$ vanishes for $r > a$. The key fact is that under quite general hypotheses the functions R_a are R functions. One immediate consequence is that the functions $s^l(k)$ have extensions into the complex plane

which have a singularity structure which can be deduced from (**) and the properties of R functions.

The preceding description of the connection between R functions and the scattering operator is considerably more detailed than that given by Wigner in Section I of the paper under discussion. On the other hand, Wigner goes on to discuss the more general case in which $v(r)$ need not vanish for $r > a$ but is equal to some simple well understood function there. As a matter of fact Wigner's point of view is also somewhat different and leads him to describe the scattering by a function s such that

$$s(k^2) = \frac{F'(k^2)R(k^2) - F(k^2)}{-G'(k^2)R(k^2) + G(k^2)}$$

where F', F, G, and G are real valued functions which vary with k^2 much more slowly than R does, and can be approximated by constants. Since $F'G - G'F = 1$, $s(k^2)$ to the indicated approximation will also be an R function. Our point of view will be useful in our description of a later paper.

In Section III Wigner begins be defining what he means by the "pole density" and the "strength" of an R function and defines an R function to be *uniform* if its "pole density" and "strength" actually exist and if in addition the linear term αz in the first normal form is zero. The "strength" is the "average value" of the residues of the poles and the formal definitions of "pole density" and "average value", when they exist, are more or less what one would expect. It is clear that $(a_1R + b_1)/(a_2R + b_2)$ is an R function whenever R is an R function and a_1, b_1, a_2, b_2 are real numbers such that $a_1b_2 - b_1a_2 > 0$ and a short proof shows that if R has a pole density then $(a_1R+b_1)/(a_2R+b_2)$ has the same pole density. Another useful fact about R functions with a strength s is that this strength can be determined from the asymptotic properties of R in the upper half plane whenever the linear term αz in the first normal form is zero and in particular when R is uniform. Indeed the paper contains a proof that the imaginary part of $R(x + iy)$ converges to πs as y tends to ∞ and that this convergence is uniform in x. It also contains a proof of a converse; namely that the uniform existence of the limit in question and its independence of x together imply that R has a strength and no αz term in its first normal form. The proof of this converse is rather long and so is the proof of the next theorem. The latter asserts that a hypothesis (called the subsidiary condition) on the rapidity of convergence to the limit defining the strength of an R function implies correspondingly rapid convergence of the imaginary part of $R(x + iy)$ to πs as y increases to ∞. Other results in III have shorter proofs. The final theorem asserts that for any uniform R function, the residues of its poles are bounded above in absolute value and that there is a lower limit to the distances between successive poles.

If one knows the pole density for a uniform R function one cannot conclude very much about the statistical properties of the differences between consecutive poles. On the other hand, one might hope that, if one knows both the pole density and the strength of a uniform R function one might be able to deduce

something about the gross statistical properties of the poles of $f' = R'/(1+R^2)$. This hope turns out to be vain, and the final Section IV entitled "Statistical R functions" reports a preliminary investigation of what one can do in this direction using finer statistical properties of the sequences of poles and residues of R and f. By definition a *statistical R function* is a uniform R function in which certain specific finer statistical measures exist. The main result is the existence of a class of statistical R functions which is invariant under linear fractional transformations of the special form $R \to (\sin\varphi + R\cos\varphi)/(\cos\varphi - R\sin\varphi)$ and which are in a certain sense normalized. Connections between the statistical properties of R and f' of the sort indicated above are made plausible but are *not* proved. The details are both complicated and lengthy and are best omitted in an essay of this nature.

One year later, in 1952, Wigner published no less than three more articles about R functions. One of these is a direct continuation of Section IV of the introductory 1951 article [62]. The other two are of a quite different character. We shall deal first with the direct continuation [63]. Its contents, for our purposes, are perhaps sufficiently indicated by the following quotation from the introduction to the paper itself:

"Similarly, it is plausible that conversely, the distribution functions for κ_μ and x_μ differences, together with the statistical correlations between these quantities, determine the distribution functions for the g_μ ($g_\mu = 1/\gamma_{\mu^2}$) and the Z_μ differences and their statistical correlations.

The present article will deal with an example for the latter problem: it will give, explicitly, the distribution of the g_μ and of the differences of subsequent Z_μ for the case that the κ_μ have an arbitrary distribution, the corresponding x_μ are distributed entirely at random and $\eta > 0$."

Another of the papers on R functions [64] published in 1952 concerns the relationship discussed above in connection with Section I of the 1951 paper and labelled (**)

$$s^l(k) = e^{-2ika}\left(\frac{1 + ikR_a(k^2)}{1 - ikR_a(k^2)}\right)$$

where l and a are held fixed, R_a is an R function and s^l the complex valued function of modulus one describing the l-th component of the scattering operators. Rewriting (**) as $S(k) = e^{-2ika}((1 + ikR(k^2))/(1 - ikR(k^2)))$ where S and R are abstract functions defined on the positive real axis, Wigner studies the properties of S which imply that R must be extensible to an R function and conversely. Evidently R is uniquely determined by S. If R is indeed extensible to be an R function it is easy to check that the corresponding extension of S has the following properties:

(a) $S(k)S(-k) \equiv 1$
(b) $S(k)(S(k^*))^* \equiv 1$ (where $*$ indicates the complex conjugate)
(c) $S(k)e^{2ik\rho}$ is uniformly bounded for $\rho > a$ and $\operatorname{Im} k > 0$.

In a paper [65] published in 1951 by Schutzer and Tiomno, these properties for s^l were shown to be a consequence of "causality" on purely physical

grounds and without reference to the relationship (**). Wigner's basic question is whether (a), (b), and (c) imply conversely that R is an R function and if not by what they might be supplemented. In Part II he gives an example showing that (a), (b), and (c) do not suffice and in Part III finds a new property of S which is not implied by (a), (b), and (c) but is implied by R's being an R function. He asserts, however, that (a), (b), and (c) together with the new property still do not suffice to prove that R is an R function.

It seems appropriate to mention that a trend which almost dominated elementary particle physics for more than a decade began with a paper [66] published in 1954. It was concerned above all with exploiting the hypothesis that the scattering operator may be continued into the complex domain.

The third paper in this category [67], which appeared in 1952, was published in the American Mathematical Monthly. Its purpose was to present a new way of deriving the main properties of the elementary trigonometric and exponential functions for complex values of the argument without using their properties for real values of the argument. It does this by exploiting the fact that the tangent function is in a sense the simplest R function and the expansion theorems involved in the two normal forms for general R functions. The author takes the opportunity to acknowledge that some of the more elementary parts of his 1951 paper on R functions were published some decades earlier by G. Herglotz.

In 1954 Wigner published a joint paper with von Neumann [68] which concluded his work on the mathematical aspects of the theory of R functions. This paper is entitled "Significance of Loewner's theorem in the quantum theory of collisions". Let R be an R function and let $z_1, z_2, \ldots, z_n$ be positive real numbers with $z_1 < z_2 \ldots < z_n$ and not including any singularities of R. Form the $n \times n$ matrix $\|\kappa_{ij}\|$ where $\kappa_{ii} = R'(z_i)$ and for $i \neq j$, $\kappa_{ij} = (R(z_i) - R(z_j))/(i-j)$. It is not difficult to prove that for all $n = 1, 2, \ldots$ and all n-tuples $z_1 < z_2 \ldots < z_n$ te matrix $\|\kappa_{ij}\|$ is a positive semi definite. The main concern of the paper is the extent to which this theorem has a converse. Specifically let R be a real valued differentiable function defined on the positive real axis, except for a discrete sequence of points Z_1, Z_2 called its singularities. Suppose that the matrices $\|\kappa_{ij}\|$ defined above are all positive semi definite. To what extent can we conclude that R must be the restriction to the positive real axis of an R function in which some of the Z_j may *not* be singularities. The answer is that R can be extended over the whole complex plane with the negative real axis removed and there has all the defining properties of an R function. The proof is far from trivial and makes essential use of ideas which appear in a paper of Loewner [69] published in 1933. Indeed the authors assert that although Loewner's problem is rather different from theirs, his arguments come very near to constituting a complete proof of their result. They admit, however, that they have simplified Loewner's argument in some respects.

We have described the connection between R functions and scattering theory in only the simple special case of spinless potential scattering. It should be pointed out, however, that Wigner's interest in these functions stems from much more difficult problems in the theory of the interactions between nuclei.

We refer the reader to H. Feshbach's annotation for a description of Wigner's R matrix theory of nuclear reactions and its importance.

One can also consult Feshbach's annotation for a description of Wigner's work on the "giant resonance model" in the theory of the interaction of a single nucleon with a complex nucleus. This work also led to some purely mathematical investigations. These were published in three papers in *Annals of Mathematics* in 1955, 1957, and 1958, respectively [70, 71 72]. We conclude our essay with a brief description of the contents of these papers.

Let $\mathcal{H}$ be a Hilbert space and let $\ldots \varphi_{-2}, \varphi_{-1}, \varphi_0, \varphi_1, \varphi_2, \ldots$ be an orthonormal basis for $\mathcal{H}$ indexed by the positive and negative integers. Let H be a self-adjoint operator whose matrix elements with respect to this basis have the following properties: (a) $(H(\varphi_n) \cdot \varphi_m) = 0$ for $[m - n] > N$, where N is some fixed positive integer. (b) $(H(\varphi_n) \cdot \varphi_n) = n$. (c) $(H(\varphi_n) \cdot \varphi_m) = \pm v$ when $0 < |n - m| < N$ where v is some fixed positive number. Given v and N the matrix elements of H are completely determined up to a choice of sign at each pair $n > m$ with $n - m \leq N$. Supposing these chosen independently with equal probabilities of being ± 1, one has a well defined probability distribution on the class of possible matrices. The problem is to study the statistical properties of the eigenvalues of these matrices. The results of [70], which are admittedly in part only heuristically established, are meant to apply in the limit of very large N and v with v^2/N remaining bounded. The analysis begins with a detailed study of the eigenvalue structure in the special case in which $N = 1$. In another preliminary section, a study is made of the eigenvalue statistics for finite real symmetric matrices of very large odd dimension in which the diagonal elements are all zero and the off diagonal elements all have the same absolute value.

In the 1957 paper [71] the analysis given in [70] is improved as follows. An integral equation for one of the main results of [70] is "transformed into a much simpler equation from which the approximate expressions for ρ (for large and small values of q) which were obtained before can be obtained much more easily".

The 1958 paper [71] is only two pages long and consists mainly of an announcement that the preliminary results about finite matrices contained in the section of [70] entitled "Random sign symmetric matrices" can be considerably generalized. The final sentence reads "The heuristic proof given for the special case considered before applies equally under the more general conditions here specified".

References

[1] E. Wigner: Über nicht kombinierende Terme in der neueren Quantentheorie. Z. Phys. **40**, 492–500 (1926); volume I, part II

[2] E. Wigner: Über nicht kombinierende Terme in der neueren Quantentheorie II. Teil. Z. Phys. **43**, 883–892 (1927); volume I, part II

[3] W. Heisenberg: Mehrkörperproblem und Resonanz in der Quantenmechanik. Z. Phys. **38**, 411–426 (1926)

[4] E. Wigner: Einige Folgerungen aus der Schrödingerschen Theorie für die Termstrukturen. Z. Phys. **43**, 624–652; volume I, part II
[5] J. von Neumann und E. Wigner: Zur Erklärung einiger Eigenschaften der Spektren aus der Quantenmechanik des Drehelektrons. Z. Phys. **47**, 203–220 (1928); volume I, part II
[6] J. von Neumann und E. Wigner: Zur Erklärung einiger Eigenschaften der Spektren aus der Quantenmechanik des Drehelektrons II. Z. Phys. **49**, 73–94 (1928); volume I, part II
[7] J. von Neumann und E. Wigner: Zur Erklärung einiger Eigenschaften der Spektren aus der Quantenmechanik des Drehelektrons III. Z. Phys. **51**, 844–858 (1928); volume I, part II
[8] E. Wigner: Gruppentheorie und ihre Anwendung auf die Quantenmechanik der Atomspektren. Vieweg, Braunschweig 1931
[9] J. C. Slater: The theory of complex spectra. Phys. Rev. **34**, 1293–1322 (1929)
[10] H. Weyl: Quantenmechanik und Gruppentheorie. Z. Phys. **46**, 1–46 (1927)
[11] P. Jordan and E. Wigner: Über das Paulische Äquivalenzverbot. Z. Phys. **47**, 631–651 (1928); volume I, part II
[12] M. H. Stone: Linear transformations in Hilbert space III. Operational methods and group theory. Proc. Nat. Acad. Sci. USA **16**, 172–175 (1930)
[13] J. von Neumann: Die Eindeutigkeit der Schrödingerschen Operatoren. Math. Ann. **104**, 570–578 (1931)
[14] E. U. Condon and G. H. Shortley: The theory of atomic spectra. Cambridge University Press, Cambridge 1935
[15] C. Eckart: The application of group theory to the quantum dynamics of monatomic systems. Rev. Mod. Phys. **2**, 305–380 (1930)
[16] E. Wigner: Über die elastischen Eigenschwingungen symmetrischer Systeme. Nachr. Ges. Wiss. Gött., Math-Phys. Klasse (1930) 133–146; volume I, part II
[17] E. Wigner: Über die Operation der Zeitumkehr in der Quantenmechanik. Nachr. Ges. Wiss. Gött., Math.-Phys. Klasse (1932) 546–559; volume I, part II
[18] E. Wigner: On the interaction of electrons in metals. Phys. Rev. **46**, 1002–1011 (1934); volume IV, part II
[19] L. P. Bouckaert, R. Smoluchowski and E. Wigner: Theory of Brillouin zones and symmetry properties of wave functions in crystals. Phys. Rev. **50**, 58–67 (1936); volume IV, part II
[20] E. Wigner: On the consequences of the symmetry of the nuclear Hamiltonian on the spectroscopy of nuclei. Phys. Rev. **51**, 106–119 (1937); volume II
[21] E. Wigner: On unitary representations of the inhomogeneous Lorentz group. Ann. Math. **40**, 149–204 (1939); volume I, part III
[22] J. von Neumann and E. Wigner: Minimally almost periodic groups. Ann. Math. **41**, 746–750 (1940); volume I, part III
[23] E. Wigner: On representations of certain finite groups. Am. J. Math. **63**, 57–63 (1941); volume I, part III
[24] E. Wigner: On the matrices which reduce the Kronecker products of S.R. groups. Unpublished manuscript of circa 1940, printed in: Quantum theory of angular momentum, eds. L. C. Biedenharn and H. van Dam. Academic Press, New York 1965; volume I, part III
[25] G. Racah: Theory of complex spectra I. Phys. Rev. **61**, 186–197 (1942)
[26] G. Racah: Theory of complex spectra II. Phys. Rev. **62**, 438–462 (1942)
[27] G. Racah: Theory of complex spectra III. Phys. Rev. **63**, 367–382 (1943)
[28] G. Racah: Theory of complex spectra IV. Phys. Rev. **76**, 1352–1365 (1949)
[29] V. Bargmann and E. Wigner: Group theoretical discussion of relativistic wave equations. Proc. Nat. Acad. Sci. **34**, 211–223 (1948); volume III, part I
[30] T. D. Newton and E. Wigner: Localized states for elementary systems. Rev. Mod. Phys. **21**, 400–406 (1949); volume I, part III

[31] G. Mackey: Imprimitivity for representations of locally compact groups I. Proc. Nat. Acad. Sci. **35**, 537–544 (1949)

[32] A. Wightman: On the localizability of quantum mechanical systems. Rev. Mod. Phys. **34**, 845–872 (1962)

[33] G. Mackey: Infinite-dimensional group representations. Bull. Amer. Math. Soc. **69**, 628–686 (1963)

[34] E. Inönü and E. Wigner: Representations of the Galilei group. Il Nuovo Cimento **IX**, 705–718 (1952); volume I, part III

[35] E. Inönü and E. Wigner: On the contraction of groups and their representations. Proc. Nat. Acad. Sci. **39**, 510–524 (1953); volume I, part III

[36] E. Inönü and E. Wigner: On a particular type of convergence to a singular matrix. Proc. Nat. Acad. Sci. **40**, 119–121 (1954); volume I, part III

[37] V. Bargmann: On unitary ray representations of continuous groups. Ann. Math. **59**, 1–46 (1954)

[38] I. Segal: A class of operator algebras which are determined by groups. Duke Math. J. **18**, 221–265 (1951)

[39] G. Mackey: On the analogy between semi-simple Lie groups and certain related semi direct groups. Lie groups and their representations. Proc. Summer School on Group Representations of the Bolyai Janos Math. Soc., Budapest (1971) 339–363. Halsted, New York 1975

[40] E. Wigner: The interpretation of Racah's coefficients. Symposium on New Research Techniques in Physics, July 15–29, 1952. Published at Rio de Janeiro 1954; volume I, part III

[41] J. Talman: Special functions – a group theoretic approach. Based on 1955 lectures by E. Wigner. Benjamin, New York 1968

[42] E. Hewitt and E. Wigner: Note on a theorem of Magnus. Proc. Amer. Math. Soc. **8**, 740–744 (1957); volume I, part III

[43] E. Wigner: Unitary representations of the inhomogeneous Lorentz group including reflections. In: Group theoretical concepts and methods in elementary particle physics, Feza Gursey (ed.). Gordon and Breach, New York 1964, pp. 37–80; volume I, part III

[44] A. Saenz and E. Wigner: Unimodular matrices homomorphic to Lorentz transformations in $n \geq 2$ spacelike dimensions. Z. Naturforsch. **22**, 1293–1299 (1967); volume I, part III

[45] E. Wigner: Condition that the irreducible representations of a group, considered as representations of a subgroup, do not contain any representation of the subgroup more than once. In: Spectroscopic and group theoretic methods in physics, F. Bloch et al. (ed.). North-Holland, Amsterdam 1968, pp. 131–136; volume I, part III

[46] E. Wigner: Restriction of irreducible representations of groups to a subgroup. Proc. Royal Soc. **322**, 181–189 (1971); volume I, part III

[47] F. Goldrich and E. Wigner: Condition that all irreducible representations of a compact Lie group, if restricted to a subgroup, contain no representation more than once. Can. J. Math. **24**, 432–438 (1972); volume I, part III

[48] G. Mackey: Symmetric and anti symmetric Kronecker squares and intertwining numbers of induced representations of finite groups. Amer. J. Math. **75**, 387–405 (1953)

[49] G. Mackey: Multiplicity free representations of finite groups. Pacific J. Math. **8**, 503–510 (1958)

[50] E. Wigner: Symmetry principles in old and new physics. Bull. Amer. Math. Soc. **74**, 793–815 (1968); volume III, part I

[51] E. Wigner: Reduction of direct products and restrictions of representations to subgroups: the everyday tasks of the quantum theorists. Siam J. Appl. Math. **25**, 169–185 (1973); volume I, part III

[52] J. von Neumann and E. Wigner: Über merkwürdige diskrete Eigenwerte. Phys. Z. **30**, 465–467 (1929); volume I, part III. This paper contains an error which is discussed in Belazs' annotation, volume IV, part I

[53] J. von Neumann and E. Wigner: Über das Verhalten von Eigenwerten bei adiabatischen Prozessen. Phys. Z. **30**, 467–470 (1929); volume I, part III

[54] P. Jordan, J. von Neumann and E. Wigner: On an algebraic generalization of the quantum mechanical formalism. Ann. Math. **35**, 29–64 (1934); volume I, part III

[55] E. Wigner: On a modification of the Rayleigh-Schrödinger perturbation theory. Magy. Tud. Akad. Mat. Természettudományi Ért. **53**, 475–482 (1935); volume IV, part I

[56] E. Wigner: Application of the Rayleigh-Schrödinger perturbation theory to the hydrogen atom. Phys. Rev. **94**, 77–78 (1954); volume I, part III

[57] E. Wigner: Normal form of antiunitary operators. J. Math. Phys. **1**, 409–413 (1960); volume I, part III

[58] E. Wigner: Phenomenological distinction between unitary and antiunitary symmetry operators. J. Math. Phys. **1**, 414–416 (1960); volume I, part III

[59] E. Wigner: On weakly positive matrices. Can. J. Math. **15**, 313–317 (1963); volume I, part III

[60] E. Wigner: On a generalization of Euler's angles. In: Group theory and its Applications. Academic Press, New York 1968, pp. 119–129; volume I, part III

[61] H. Schweinler and E. Wigner: Orthogonalization methods. J. Math. Phys. **11**, 1693–1694 (1970); volume I, part III

[62] E. Wigner: On a class of analytic functions from the quantum theory of collisions. Ann. Math. **53**, 36–67 (1951); volume I, part III

[63] E. Wigner: On the connection between the distribution of poles and residues for an R-function and its invariant derivative. Ann. Math. **55**, 7–18 (1952); volume I, part III

[64] E. Wigner: Derivative matrix and scattering matrix. Revista Mexicana de Fisica **1**, 91–101 (1952); volume I, part III

[65] W. Schützer and J. Tiomno: On the connection of the scattering and derivative matrices with causality. Phys. Rev. **83**, 249–251 (1951)

[66] M. Gell-Mann, M. Goldberger and W. Thirring: Use of causality conditions in quantum theory. Phys. Rev. **95**, 1612–1627 (1954)

[67] E. Wigner: Simplified derivation of the properties of elementary transcendentals. Amer. Math. Monthley **59**, 669–683 (1952); volume I, part III

[68] J. von Neumann and E. Wigner: Significance of Loewner's theorem in the quantum theory of collisions. Ann. Math. **59**, 418–433 (1954); volume I, part III

[69] K. Loewner: Über monotone Matrixfunktionen. Math. Z. **38**, 177–216 (1933)

[70] E. Wigner: Characteristic vectors of bordered matrices with infinite dimensions. Ann. Math. **62**, 548–564 (1955); volume I, part III

[71] E. Wigner: Characteristic vectors of bordered matrices with infinite dimensions, II. Ann. Math. **65**, 203–207 (1957); volume I, part III

[72] E. Wigner: On the distribution of the roots of certain symmetric matrices. Ann. Math. **67**, 325–327 (1958); volume I, part III

Über merkwürdige diskrete Eigenwerte

J. von Neumann and E. P. Wigner

Physikalische Zeitschrift *30*, 465–467 (1929)

1. Es sind in der quantentheoretischen Literatur mehrfach anschauliche Schlüsse von der Art gemacht worden, daß z. B. aus der Tatsache, daß ein Elektron genügend kinetische Energie hat, um sich aus einem atomaren System (klassisch gerechnet) ins Unendliche zu entfernen, geschlossen wurde, daß der betreffende Energiewert zum kontinuierlichen Spektrum des genannten Systems gehört. Im folgenden soll gezeigt werden, daß derartige Überlegungen mit äußerster Vorsicht zu handhaben sind, denn es kommt häufig ein entgegengesetztes Verhalten vor. Dieser Umstand, daß ein Elektron auf einer stationären Bahn verharrt (Punkteigenwert!), obwohl es Energie genug hätte, um sich aus dem Anziehungsbereich des ihn umgebenden Systems zu befreien, ist nur scheinbar paradox. Wir werden uns an zwei verschiedenen Beispielen klar machen, daß dieses Phänomen zwei ganz verschiedene Ursachen haben kann — aber in beiden Fällen bis zu einem gewissen Grade anschaulich deutbar ist.

Wir werden stets ein Elektron im Bereiche eines geeigneten kugelsymmetrischen Potentials betrachten, also die Wellengleichung

$$-\frac{h^2}{8\pi^2 m}\Delta\psi + (V(r) - E)\psi = 0$$

$$\left(\Delta = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2} = \frac{\partial^2}{\partial r^2} + \frac{2}{r}\frac{\partial}{\partial r}\right.,$$ da wir uns auf kugelsymmetrische ψ beschränken und die Ableitungen nach den Winkeln θ und φ weglassen können). Damit $E = 0$ ein Punkteigenwert mit der kugelsymmetrischen Eigenfunktion $\psi = \psi(r)$ sei, muß also $V = \frac{h^2}{8\pi^2 m}\left(\frac{\psi''}{\psi} + \frac{2\psi'}{r\psi}\right)$ sein. Wir werden nun durch geeignete Wahl von ψ dem V die oben erwähnten, scheinbar paradoxen, Formen erteilen.

Sei zuerst $\psi = \frac{\sin(r^3)}{r^2}$, dann ergibt sich[1]) (unter Weglassung des unwesentlichen Faktors $\left.\frac{h^2}{8\pi^2 m}\right)$ $V = 2r^{-2} - 9r^4$. Ein ebener Schnitt durch den Potentialverlauf sieht also wie auf der Figur angedeutet aus. Nach allem Erwarten müßte das Elektron den Potentialabhang hinabstürzen und daher nur ein Streckenspektrum besitzen — dennoch ist eine stationäre Bahn und der Punkteigenwert $E = 0$ da!

Jedoch ist dieses Verhalten sogar auf Grund klassisch-mechanischer Analogien zu verstehen[2]).

1) Man setze versuchsweise $\psi(r) = r^a \sin(r^b)$. Damit das Quadratintegral von ψ_1 d.h. $\int_0^\infty 4\pi r^2 |(\psi(r)|^2 dr = \int_0^\infty 4\pi r^{a+2}\sin^2(r^b)dr$ endlich sei, muß mit Rücksicht auf das Verhalten bei $r = 0$: $2a + b + 2 > -1$ und auf das Verhalten bei $r = \infty$: $2a + 2 < -1$ sein. Die Regularität von $V(r)$ für $r \neq 0$ erfordert $2a + b + 1 = 0$. Beides erfüllt $a = -2$, $b = 3$.

2) Diese Bemerkung rührt von Herrn L. Szilard her.

466 Neumann u. Wigner, Über merkwürdige diskrete Eigenwerte. Physik.Zeitschr.XXX, 1929.

Auf dem Potentialabhang $V = 2r^{-2} - 9r^4$ wird ein Punkt (falls er sich längs einer durch die den Nullpunkt gehenden Geraden bewegt) die Bahn

$$\frac{m}{2}\left(\frac{dr}{dt}\right)^2 + V(r) = \text{Konst.}$$

$$\frac{dr}{dt} = \sqrt{\frac{2}{m}\,\text{Konst.} - \frac{4}{m}\frac{1}{r^2} + \frac{18}{m}r^4}$$

beschreiben.

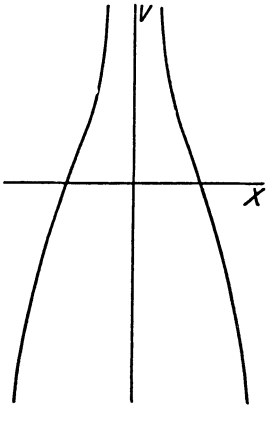
Man erkennt leicht, daß schließlich $r \to \infty$, $\frac{dr}{dt} \to \infty$, und dann ist asymptotisch

$$\frac{dr}{dt} = \frac{3\sqrt{2}}{\sqrt{m}}r^2;\ \frac{1}{r} = \frac{3\sqrt{2}}{\sqrt{m}}(t_0 - t)$$

$$r = \frac{\sqrt{m}}{3\sqrt{2}\,(t_0 - t)}$$

Der Punkt geht also noch in der endlichen Zeit $t = t_0$ ins Unendliche.

In diesem Moment hören nun die Aussagen der klassischen Mechanik auf. Es scheint aber, als ob die Quantenmechanik sich durch diese



Flucht des zu untersuchenden Punktes ins Unendliche nicht davon abhalten ließe, sein ferneres Schicksal zu verfolgen. Die stationäre Lösung $\psi(r)$ deutet an, daß er aus dem Unendlichen wiederkehrt, daß er dort sozusagen reflektiert wird — um dann das Pendeln wieder zu beginnen. Daß $\psi(r)$ für $r = \infty$ gegen o strebt, d. h. daß sich der Punkt vorwiegend bei mäßig großen r aufhält, rührt daher, daß er bei großen Werten enorme Geschwindigkeiten hat (er ist ja bis dahin einen großen Potentialberg hinuntergestürzt), also nur kurze Verweilzeiten. Da der ganze Pendelvorgang (von $r_{\min}$ nach ∞ und zurück) periodisch ist, ist es sehr vernünftig, daß das entsprechende quantenmechanische Problem eine stationäre Lösung besitzt.

Ein weiterer Beleg dafür, daß die Stationarität dieser Bahn wirklich nur durch die klassische Analogie einer Reflexion im Unendlichen verständlich ist, ergibt sich, wenn man im Sinne von Anm. [1]) der vorigen Seite allgemeiner $\psi = \frac{\sin(r^b)}{r^{1/2(b+1)}}$ $(b > 2)$ setzt, dann wird $V = (b^2 - 1)(4r^2)^{-1} - b^2 r^{2b-2}$. Man erkennt mühelos, daß bei klassisch-mechanischer Bewegung die fürs Erreichen des Unendlichen in endlicher Zeit charakteristische Bedingung $b > 2$ ist — genau das, was zur Endlichkeit des Quadratintegrals von ψ, d. h. für die Stationarität dieser Bahn, notwendig und hinreichend ist.

2. Wir gehen nun zu einem zweiten Beispiel über, das auf Grund solcher Erwägungen nicht mehr zu deuten ist.

Zuerst setzen wir $\psi = \frac{\sin r}{r}$, dann wird $V = -1$. Es ist die wohlbekannte Kugelwellenlösung, einem sich ins Unendliche entfernenden Elektron entsprechend, die Energie $E = 0$ ist dabei größer als das Potential $V = -1$, so daß ein positiver Betrag für die klassische kinetische Energie übrigbleibt. Wir sind dabei im kontinuierlichen Spektrum, da $\int_0^\infty 4\pi r^2 |\psi(r)|^2 dr = 4\pi \int_0^\infty \sin^2 r\, dr$ unendlich ist. Nun werden wir aber versuchen, ψ so abzuändern, daß sein Quadratintegral endlich wird und V doch in der Nähe von -1 variiert.

Hierzu machen wir den Ansatz:

$$\psi(r) = \frac{\sin r}{r} f(r).$$

Wenn $f(r)$ stetig ist und für $r \to \infty$ einem r^a mit $a < -\frac{1}{2}$ asymptotisch gleich ist, so ist das Quadratintegral von ψ in der Tat endlich. V dagegen bestimmt sich zu

$$V = -1 + 2\,\mathrm{tg}\, r \frac{f'(r)}{f(r)} - \frac{f''(r)}{f(r)}.$$

Damit dieses V ständig in der Nähe von -1 bleibt, und (was wir zusätzlich verlangen wollen) für $r \to \infty$ gegen -1 konvergiert, müssen $\mathrm{tg}\, r \frac{f'(r)}{f(r)}, \frac{f''(r)}{f(r)}$ stets klein sein, und für $r \to \infty$ gegen o streben.

In erster Linie muß also $\frac{f'(r)}{f(r)}$ für $\mathrm{tg}\, r = \infty$, d. i. $r = \frac{\pi}{2}, \frac{3\pi}{2}, \frac{5\pi}{2}, \ldots$ verschwinden, d. h. für $\cos r = 0$. Das ist gewiß der Fall, wenn $f(r)$ dem

$\int_0^r \cos^2 r\, dr = \frac{r}{2} + \frac{1}{4} \sin 2r$ gleich ist, oder irgendeiner Funktion davon. Wir setzen (da $f(r)$ für $r \to \infty$ asymptotisch r^a, $a < -\frac{1}{2}$ sein soll) $f(r) = [A^2 + (2r + \sin 2r)^2]^{-1}$, wo über A noch verfügt werden wird. Dann ist:

$$\frac{f'(r)}{f(r)} = -\frac{8(2r + \sin 2r)\cos^2 r}{A^2 + (2r + \sin 2r)^2}$$

$$\frac{f''(r)}{f(r)} = \frac{128(2r + \sin 2r)^2 \cos^4 r}{[A^2 + (2r + \sin 2r)^2]^2} - \frac{32\cos^4 r - 16(2r + \sin 2r)\cos r \sin r}{A^2 + (2r + \sin 2r)^2}$$

Hieraus folgt (es heben sich zwei Glieder weg, was aber unwesentlich ist)

$$V = -1 - \frac{32\cos^4 r}{A^2 + (2r + \sin 2r)^2} + \frac{128(2r + \sin 2r)^2 \cos^4 r}{[A^2 + (2r + \sin 2r)^2]^2}$$

$$= -1 - 32\cos^4 r \frac{A^2 - 3(2r + \sin 2r)^2}{[A^2 + (2r + \sin 2r)^2]^2}$$

Diesem Ausdruck sieht man es aber sofort an, daß erstens zu jedem $\varepsilon > 0$ ein A gewählt werden kann, daß er stets zwischen $-1 - \varepsilon$ und $-1 + \varepsilon$ liegt, und daß er zweitens für $r \to \infty$ gegen -1 strebt[1]).

1) Ebenso liegen alle seine Derivierten beliebig nahe bei 0.

Ein

$$\psi = \frac{\sin r}{r(A^2 + (2r + \sin 2r)^2)}$$

mit genügend großem A leistet also alles Gewünschte.

Eine Deutung wie im vorigen Punkt kommt hier nicht in Frage: V schmiegt sich ja beliebig gut an -1 an! Das einzige, was sich in Anlehnung an allgemein geläufige anschauliche Betrachtungsweisen vorbringen ließe, ist dieses: Das obige Potential V weist gewisse Wellungen mit der Periode π auf ($\cos^2 r$ und $\sin 2r$ gehen in die Formel ein), welche mit $r \to 0$ gegen 0 abfallen. Diese scheinen nun eine interferenzartige Auslöschung des Wellenpaketes ψ in den entfernteren Raumgegenden zu veranlassen[2]).

So wie wir hier in der Nähe des Potentials $V = -1$ eines mit stationären Bahnen fanden, lassen sich auch in der Nähe anderer Potentiale, z. B. $V = x$ solche angeben. Dies ist darum erwähnenswert, weil es zeigt, daß eine Diskussion der Frage, ob das durch ein homogenes elektrisches Feld gestörte Wasserstoffatom (Starkeffekt) Punkteigenwerte besitzt oder nicht, mit gewissen Schwierigkeiten verbunden ist[1]). Insbesondere muß dabei die „Glattheit" des elektrischen Störungsfeldes eine wesentliche Rolle spielen.

1) Vgl. J. R. Oppenheimer, Phys. Rev. **31**, 80, 1928, wo die Annahme des Fehlens von Punkteigenwerten gemacht wird.

2) Das Potential $V(r) = 2r^{-2} - 9r^4$ war nicht so, es war vollkommen „glatt".

(Eingegangen 25. Mai 1929.)

Über das Verhalten von Eigenwerten bei adiabatischen Prozessen

J. von Neumann und E. P. Wigner

Physikalische Zeitschrift *30*, 467–470 (1929)

1. In vielen Fragen der Quantenmechanik ist es wichtig, die Veränderung der Eigenwerte und Eigenfunktionen bei stetiger Änderung eines oder mehrerer Parameter zu untersuchen. Namentlich interessiert oft der Fall, in dem man für zwei spezielle Werte der Parameter Eigenwerte und Eigenfunktionen kennt und sich für das Zwischengebiet interessiert. Man fragt gewöhnlich, ob im Zwischengebiet Überschneidungen der Eigenwerte vorkommen, in welchen Eigenwert ein bestimmter Eigenwert übergeht, wenn man von dem einen Wertsystem der Parameter kontinuierlich in das andere Wertsystem übergeht usw. Fragen ähnlicher Art hat F. Hund aufgeworfen[1]) und insbesondere die letzte Frage für den Fall eines Parameters — auf Grund von Beispielen — dahin beantwortet, daß Überschneidungen im allgemeinen — wenn dafür kein spezieller Grund vorhanden ist — nicht vorkommen[2]). Wir wollen hier dies allgemein begründen, unsere Methode erlaubt dabei gleichzeitig die Untersuchung von Systemen mit mehreren veränderlichen Parametern.

Die Energiewerte eines Systems sind bekanntlich die Eigenwerte einer Hermiteschen[3]) Matrix $(H_{\nu\mu})$, die wir der Einfachheit halber endlichvieldimensional, etwa n-dimensional, annehmen wollen. Dabei müssen wir uns vorstellen, daß alle n^2 komplexen Größen $H_{\nu\mu}$ noch von einigen reellen Parametern $\varkappa_1, \varkappa_2, \ldots$ abhängen, und wir fragen, durch Änderung wie vieler Parameter man im allgemeinen ein Zusammenfallen zweier Eigenwerte erreichen kann. Wir werden sehen, daß dazu im allgemeinen drei reelle Parameterwerte ent-

1) F. Hund, Zeitschr. f. Phys. **40**, 742, 1927.

2) Siehe insbesondere S. 752.

3) Daß die Matrix Hermiteisch ist, ist entscheidend für das Folgende.

sprechend bestimmt werden müssen („im allgemeinen" bedeutet, daß zwischen den Zahlen $H_{\nu\mu}$ keine Beziehungen bestehen, die nicht aus dem Hermiteschen Charakter folgen)[1]). Wenn man also nur einen oder zwei Parameter verändern kann, ist es i. a. nicht möglich, ein Überschneiden der Eigenwerte zu erreichen.

Um dies zu zeigen, zählen wir die freien reellen Parameter einer n-dimensionalen Hermiteschen Matrix mit und ohne doppelte Eigenwerte ab. Die Differenz dieser Zahlen wird uns die Anzahl der zu verändernden Parameter $\varkappa_1, \varkappa_2, \ldots$ angeben, durch deren Veränderung man ein Zusammenfallen von Eigenwerten bewirken kann.

Man kann bekanntlich jede Hermitesche Matrix $(H_{\nu\mu})$ durch eine unitäre Matrix $(U_{\varrho\nu})$ auf die Diagonalform bringen

$$H_{\nu\mu} = \sum_{\varrho} E_\varrho U_{\varrho\nu} \bar{U}_{\varrho\mu} \tag{1}$$

wo $E_1, E_2, \ldots$ die Eigenwerte sind. Zur Bestimmung der $H_{\nu\mu}$ muß man also die n reellen Zahlen E_ϱ und die unitäre Matrix $(U_{\varrho\nu})$ kennen, diese letztere allerdings nur bis auf eine unitäre Matrix, die mit der Diagonalmatrix

$$\begin{vmatrix} E_1 & 0 & 0 & \ldots & 0 \\ 0 & E_2 & 0 & \ldots & 0 \\ 0 & 0 & E_3 & \ldots & 0 \\ \cdot & \cdot & \cdot & \cdots & \cdot \\ \cdot & \cdot & \cdot & \cdots & \cdot \\ \cdot & \cdot & \cdot & \cdots & \cdot \\ \cdot & \cdot & \cdot & \cdots & \cdot \\ 0 & 0 & 0 & \ldots & E_n \end{vmatrix} \tag{2}$$

vertauschbar ist. Mit einer solchen darf man nämlich $(U_{\varrho\nu})$ von hinten multiplizieren, ohne daß (1) sich ändern würde. Da eine unitäre Matrix n^2 reelle Parameter hat, ist die Anzahl der freien Parameter einer Hermiteschen Matrix $n^2 + f - v$ wo f die Anzahl der voneinander verschiedenen Zahlen unter $E_1, E_2, \ldots, E_n$ und v die Anzahl der Parameter einer mit (2) vertauschbaren unitären Matrix ist. Sind alle $E_1, E_2, \ldots, E_n$ verschieden, so ist mit (2) nur eine Diagonalmatrix vertauschbar, kann man dagegen die E_ϱ in Gruppen einteilen, etwa so, daß in den umrahmten Quadraten lauter gleiche E vorkommen

$$\begin{array}{|cc|cccc|c|c} \hline E_1 & 0 & 0 & 0 & 0 & \ldots & & 0 \\ 0 & E_1 & 0 & 0 & 0 & \ldots & & 0 \\ \hline 0 & 0 & E_2 & 0 & 0 & & \ldots & 0 \\ 0 & 0 & 0 & E_2 & 0 & & \ldots & 0 \\ 0 & 0 & 0 & 0 & E_2 & & \ldots & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & 0 & 0 & 0 & \cdots & & E_f \\ \hline \end{array} \tag{3}$$

so ist mit (3) jede Matrix vertauschbar, die nur an den den umrahmten Stellen entsprechenden Stellen von 0 verschiedene Koeffizienten hat. Soll die Matrix unitär sein, so müssen in den umrahmten Quadraten unitäre Matrizen stehen.

Kann man also die Eigenwerte in Gruppen mit $g_1, g_2, \ldots, g_f$ $(g_1 + g_2 + \ldots + g_f = n)$ Eigenwerten einteilen, so daß alle Eigenwerte derselben Gruppe gleich sind, so ist die Anzahl der freien Parameter gleich

$$n^2 + f - g_1^2 - g_2^2 - \ldots - g_f^2.$$

In der allgemeinen Hermiteschen Matrix ist $g_1 = g_2 = \ldots = g_n = 1$, die Anzahl der freien Parameter ist $n^2 + n - 1^2 - 1^2 - \ldots - 1^2 = n^2$, wie man auch auf anderem Wege leicht abzählt[1]). Sollen aber zwei Eigenwerte zusammenfallen, so ist $g_1 = 2$, $g_2 = g_3 = \ldots = g_{n-1} = 1$, die Anzahl der freien Parameter ist $n^2 + (n-1) - 2^2 - 1^2 - 1^2 - \ldots - 1^2 = n^2 + (n-1) - 4 - (n-2) = n^2 - 3$. Man sieht auch etwa für $n = 2$, daß nur ein reeller Parameter p frei ist, da nur die Matrix $\begin{pmatrix} p & 0 \\ 0 & p \end{pmatrix}$ einen doppelten Eigenwert hat. Man kann also im allgemeinen nur durch Veränderung von drei Parametern ein Zusammenfallen von zwei Eigenwerten erreichen.

Haben wir es mit einer reellen Hermiteschen Matrix zu tun, so ändert sich an dem Vorangehenden nur das, daß man überall reelle orthogonale Matrizen an die Stelle der unitären setzen muß. Die Anzahl der freien Parameter einer reellen orthogonalen Matrix von n Dimensionen ist $\frac{1}{2} n (n-1) = \binom{n}{2}$, so daß die Anzahl der freien Parameter im reellen Fall $\binom{n}{2} + f - \binom{g_1}{2} - \binom{g_2}{2} - \ldots - \binom{g_f}{2}$ beträgt. Für die allgemeine reelle symmetrische Matrix ergibt dies $\frac{1}{2} n (n+1)$; soll ein doppelter Eigenwert da sein, so haben wir $\frac{1}{2} n (n+1) - 2$: in diesem Fall genügt es schon, zwei reelle Parameter zu verändern dürfen, um zwei Eigenwerte zum Zusammenfallen zu bringen.

Bei der Analyse der Struktur der Terme atomarer Systeme konnten die Eigenwerte in verschiedene Gruppen eingeteilt werden, jede Gruppe war etwa durch eine azimutale Quantenzahl, den Spiegelungscharakter, und das Multiplettsystem gekennzeichnet und solange die bezügliche Sym-

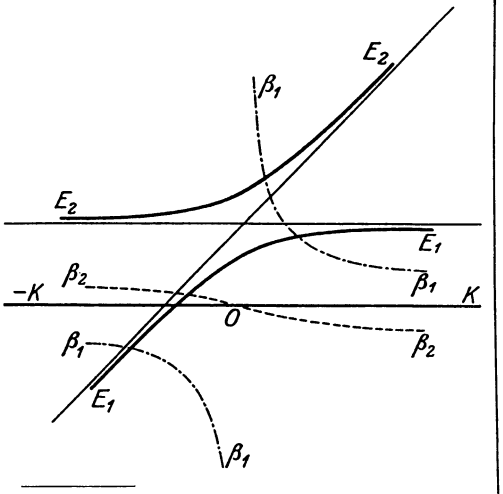
1) Man sollte meinen, daß das Zusammenfallen von zwei (natürlich reellen) Eigenwerten nur eine reelle Bedingung ergibt. Es soll aber gerade gezeigt werden, daß ein singulärer Fall vorliegt und sich die Zahl der Bedingungen so erhöht.

1) Es sind n reelle Diagonalelemente und $\frac{1}{2}(n-1)n$ komplexe Elemente oberhalb der Hauptdiagonale.

metrie nicht gestört wird, hat kein Term der einen Gruppe irgendeine „Kenntnis" von den Termen der anderen Gruppe[1]). Terme verschiedener Gruppen (d. h. verschiedener Transformationseigenschaft) haben sich beliebig überkreuzt. Außerdem waren sämtliche Terme der meisten Gruppen mehrfach entartet. Dies steht jedoch in keinem Widerspruch zu dem Vorangehenden, da hier das „im allgemeinen" nicht zutrifft, da eine genügende Anzahl von Matrixelementen immer identisch verschwindet. Ebenso wie in der Gruppentheorie der Terme angenommen werden konnte, daß keine „zufälligen Entartungen" vorkommen, können wir hier annehmen, daß keine Beziehungen existieren, die nicht aus der Symmetrie des Systems auf gruppentheoretischem Wege folgen würden, also, daß innerhalb einer Termgruppe immer der „allgemeine Fall" vorliegt.

2. Für den Fall von zwei nahe aneinander vorbeigehenden Eigenwerten wollen wir die Verhältnisse noch rechnerisch verfolgen. Dazu benutzen wir das Schrödingersche Störungsverfahren[2]), indem wir die Veränderung der Eigenwerte und Eigenfunktionen des Operators $H + \varkappa V$ bei wachsendem $\varkappa$ betrachten, wobei wir jedoch annehmen, daß die Entfernung von zwei Eigenwerten — etwa E_1 und E_2 — in der Größenordnung von $\varkappa V_{11}$ oder $\varkappa V_{22}$ liegt.

In der Figur sind die Eigenwerte als Funktion von[3]) $\varkappa$ (die ausgezogenen Kurven) aufgetragen. Es ist dabei das ursprüngliche $\varkappa$ so durch eine neue Variable $\varkappa + c$ ersetzt, daß für $\varkappa = 0$ die beiden Kurven parallel verlaufen.



1) Siehe z. B.: E. Wigner, Zeitschr. f. Phys. **40**, 883, 1927 (S. 892) und **43**, 624, 1927.

2) Sie ist hier mit dem Born, Heisenberg, Jordanschen identisch.

3) In der Figur steht K und $-K$ an Stelle von $\varkappa$ und $-\varkappa$.

Für $\varkappa = 0$ seien die Eigenfunktionen $\varphi_1(0) = \psi$ und $\varphi_2(0) = \psi'$ die Eigenwerte $E_1(0) = E - \varepsilon$ und $E_2(0) = E + \varepsilon$. Dann setzen wir an

$$\begin{aligned} \varphi_1(\varkappa) &= a_1(\varkappa)\psi + a_1'(\varkappa)\psi' + \sum_\nu a_{1\nu}(\varkappa)\psi_\nu \\ \varphi_2(\varkappa) &= a_2(\varkappa)\psi + a_2'(\varkappa)(\psi' + \sum_\nu a_{2\nu}(\varkappa)\psi_\nu \end{aligned} \tag{4}$$

und erhalten, indem wir dies in

$$\begin{aligned} (H + \varkappa V)\varphi_1(\varkappa) &= E_2(\varkappa)\varphi_1(\varkappa) \\ (H + \varkappa V)\varphi_2(\varkappa) &= E_2(\varkappa)\varphi_2(\varkappa) \end{aligned} \tag{5}$$

einführen und

$$\begin{aligned} &(\psi, V\psi) = v\,;\ (\psi', V\psi') = v'' \\ &(\psi', V\psi) = v'\,;\ (\psi, V\psi') = (V\psi, \psi') = \tilde{v}' \end{aligned} \tag{6}$$

setzen, in bekannter Weise (durch Koeffizientenvergleich) für $E_1(\varkappa)$

$$\begin{aligned} &(E - \varepsilon + \varkappa v - E_1(\varkappa))\,a_1(\varkappa) + \varkappa\tilde{v}'\,a_1'(\varkappa) = 0 \\ &\varkappa v'\,a_1(\varkappa) + (E + \varepsilon + \varkappa v'' - E_1(\varkappa))a_1'(\varkappa) = 0 \end{aligned} \tag{7}$$

Die Determinante dieses Gleichungssystems Null gesetzt ergibt für $E_1(\varkappa)$ (unter Beachtung von $v = v''$, was aus der Parallelität der $E_1(\varkappa)$ und $E_2(\varkappa)$ Kurven für $\varkappa = 0$ folgt)

$$E_1(\varkappa) = E + \varkappa v - \sqrt{\varepsilon^2 + \varkappa^2|v'|^2} \tag{8}$$

und entsprechend für $E_2(\varkappa)$

$$E_2(\varkappa) = E + \varkappa v + \sqrt{\varepsilon^2 + \varkappa^2|v'|^2}\,. \tag{8a}$$

Die beiden Kurven $E_1(\varkappa)$ und $E_2(\varkappa)$ bilden also die beiden Äste einer Hyperbel. Die Steigung der beiden Asymptoten ist $v - |v'|$ bzw. $v + |v'|$, die minimale Entfernung der beiden Äste $2\,\varepsilon$, die Entfernung nimmt auf das Doppelte zu, wenn $\varkappa = \frac{\sqrt{3}\,\varepsilon}{|v'|}$ wird, die Dauer des Vorbeigehens ist $\Delta\varkappa \sim \frac{2\sqrt{3}\,\varepsilon}{|v'|}$.

Die Eigenfunktionen $\varphi_1(\varkappa)$ und $\varphi_2(\varkappa)$ sind — wenn man sich auf die erste Näherung beschränkt — nach (4) Linearkombinationen von ψ und ψ', zu ihrer Kennzeichnung genügt es

$$\beta_1(\varkappa) = \frac{a_1(\varkappa)}{a_1'(\varkappa)} \text{ und } \beta_2(\varkappa) = \frac{a_2(\varkappa)}{a_2'(\varkappa)} \tag{9}$$

zu berechnen. Diese ergeben sich zu

$$\beta_1(\varkappa) = \frac{\varkappa\tilde{v}'}{\varepsilon - \sqrt{\varepsilon^2 + \varkappa^2|v'|^2}} = \frac{\varepsilon + \sqrt{\varepsilon^2 + \varkappa^2|v'|^2}}{\varkappa v'} \tag{10}$$

$$\beta_2(\varkappa) = \frac{\varkappa\tilde{v}'}{\varepsilon + \sqrt{\varepsilon^2 + \varkappa^2|v'|^2}} = \frac{\varepsilon - \sqrt{\varepsilon^2 + \varkappa^2|v'|^2}}{\varkappa v'} \tag{10a}$$

und sind in der Figur gestrichelt aufgetragen, wobei natürlich v' reell angenommen werden mußte.

Man sieht, daß für sehr kleine $\varkappa$

$$\beta_1(-\infty) = \frac{\tilde{v}'}{|v'|};\ \beta_2(-\infty) = -\frac{\tilde{v}'}{|v'|}, \tag{11}$$

während für sehr große $\varkappa$

$$\beta_1(\infty) = -\frac{\tilde{v}'}{|v'|};\ \beta_2(\infty) = \frac{\tilde{v}'}{|v'|}, \tag{11a}$$

gilt, so daß β_1 in β_2 und β_2 in β_1 übergeht. Für $\varkappa$ für die $|\varkappa v'| \gg 1$ ist, verhalten sich Eigenwerte und Eigenfunktionen so, als ob sie sich überkreuzen würden[1]).

Die v, v' sind die Matrixelemente von V berechnet mit denjenigen Eigenfunktionen ψ, ψ', die die richtigen Linearkombinationen für $\varkappa = 0$ sind. Es ist oft nützlich, den kleinsten Abstand 2ε und $\Delta\varkappa$, die „Dauer des Vorbeigehens" mit den Matrixelementen V_{11}, $V_{12} = \tilde{V}_{21}$, V_{22} auszudrücken, die mit Hilfe der Eigenfunktionen bei einem beliebigen $\varkappa$ gebildet werden können.

Man erhält nach etwas umständlicher Rechnung für die kleinste Entfernung der Eigenwerte

$$(E_2 - E_1)_{\min} = 2\varepsilon = \sqrt{\frac{|V_{12}|^2 (E_2 - E_1)^2}{\frac{1}{4}(V_{11} - V_{22})^2 + |V_{12}|^2}} \tag{12}$$

für $\Delta\varkappa$, die Dauer des Vorbeigehens

$$\Delta\varkappa = \frac{2\sqrt{3}\,\varepsilon}{|v'|} = \frac{\sqrt{3}(E_2 - E_1)\,|V_{12}|}{\frac{1}{4}(V_{11} - V_{22})^2 + |V_{12}|^2}. \tag{13}$$

Der Wert $(E_2 - E_1)_{\min}$ wird für

$$\varkappa = \frac{1}{4}\,\frac{(V_{11} - V_{22})(E_2 - E_1)}{\frac{1}{4}(V_{11} - {}_{22})^2 + |V_{12}|^2} \tag{14}$$

angenommen. Ist $V_{12} = 0$ wie immer, wenn E_1 und E_2 in verschiedene Termgruppen gehören, so wird sowohl (12) als auch (13) gleich 0; das sind dann ausnahmsweise doch die Verhältnisse der Überkreuzung.

Wird der Parameter $\varkappa$ nicht nur gedanklich aufgefaßt, sondern ändert man ihn wirklich am mechanischen System selber, so erfolgt die Änderung adiabatisch (d. h. $\varphi_1(-\infty)$ geht in $\varphi_1(\infty)$ und $\varphi_2(-\infty)$ — in $\varphi_2(\infty)$ über), wenn

$$\Delta\varkappa \gg \frac{h}{2\varepsilon}\frac{d\varkappa}{dt} \tag{15}$$

ist; ist dagegen

$$\Delta\varkappa \ll \frac{h}{2\varepsilon}\frac{d\varkappa}{dt}, \tag{16}$$

so geht $\varphi_1(-\infty)$ in $\varphi_2(\infty)$, $\varphi_2(-\infty)$ in $\varphi_1(\infty)$ über. In diesem Fall hat die Wellenfunktion keine Zeit sich zu ändern.

3. Eine Anwendung kann das Vorangehende vor allem auf Zuordnungsfragen finden. U. a. folgt die Zuordnungsregel für den Übergang von schwachen in starke Magnetfelder[1]). Terme mit gleichem m überkreuzen sich nicht. Sogar bei gleichzeitiger Änderung eines magnetischen und in derselben Richtung liegenden elektrischen Feldes wird man keine Überkreuzung erreichen können, da die Differentialgleichung komplex ist und zur Erreichung einer Überkreuzung drei Parameter notwendig wären. Ändert man auch den Winkel der beiden Felder, so läßt sich eine Überkreuzung erzwingen.

Bei der Anwendung auf Fragen der Zuordnung von Molekeltermen zu Termen getrennter Atome sind jedoch die Bemerkungen von F. Hund[2]) zu berücksichtigen, unsere Formeln (12), (13) gestatten prinzipiell eine Übersicht der Verhältnisse.

Auch für den Adiabatensatz dürfte das Vorangehende nicht ganz ohne Belang sein, da die Verhältnisse bei „Überkreuzungen" wohl niemals ganz klar waren.

Die wichtigste Anwendung scheint sich aber bei der Londonschen Theorie der chemischen Reaktion zu ergeben[3]). Wir wollen darauf hier nicht näher eingehen, nur auf folgenden Umstand hinweisen: Im l. c. betrachteten Falle dreier Atome, deren eines in einer durch die beiden anderen gelegten Ebene beweglich ist, haben wir zwei freie Parameter, etwa die X und V Koordinate des beweglichen Atoms. Da die Differentialgleichung reell ist, ist ein Zusammenfallen zweier Eigenwerte zwar möglich, aber nur in einzelnen Punkten. In der Tat kommt es nach F. London in erster Näherung nur in einem einzelnen Punkt vor. Es gilt das nun nach dem Vorangehenden streng, d. h. in beliebig hoher Näherung. Man kann also eine untere und eine obere Energiefläche unterscheiden, die nur einzelne Punkte gemeinsam haben. Diese Betrachtung läßt sich noch verallgemeinern und gestattet dann einen Einblick in den Mechanismus mehratomiger Systeme.

1) Auf diesen Punkt hat bereits F. Hund wiederholt hingewiesen.

1) A. Sommerfeld, Zeitschr. f. Phys. **8**, 257, 1922; A. Landé, Zeitschr. f. Phys. **19**, 112, 1923.

2) Zeitschr. f. Phys. **52**, 601, 1928. Bei streng adiabatischer Auseinanderführung der Atome würden sich nach dem Vorangehenden — wegen der Spinwechselwirkung — auch Terme verschiedener Multiplettsysteme nicht kreuzen. In diesem Fall ist jedoch die Exzentrizität der Hyperbel (wegen der Kleinheit von V_{12}, der Spinwechselwirkung) so klein, daß die Hyperbel praktisch in zwei sich schneidende Gerade ausartet. Dasselbe muß nach F. London für zwei Terme gelten, deren einer zwei aneinandergebrachten Atomen (K und F) und deren anderer zwei aneinandergebrachten Ionen (K^+ und F^-) entspricht, wenn die Überschneidung in großer Entfernung erfolgt (Zeitschr. f. Phys. **46**, 455, 1928, S. 475).

1) F. London, Sommerfeld Festschrift. S. 104 S. Hirzel. 1929.

(Eingegangen 25. Mai 1929.)

On an Algebraic Generalization of the Quantum Mechanical Formalism

P. Jordan, J. von Neumann, and E. P. Wigner

Annals of Mathematics *35*, 29–64 (1934)

(Received November 10, 1933)

Introduction

One of us has shown that the statistical properties of the measurements of a quantum mechanical system assume their simplest form when expressed in terms of a certain hypercomplex algebra which is commutative but not associative.[1] This algebra differs from the non-commutative but associative matrix algebra usually considered in that one is concerned with the commutative expression $\frac{1}{2}(A \times B + B \times A)$ instead of the associative product $A \times B$ of two matrices. It was conjectured that the laws of this commutative algebra would form a suitable starting point for a generalization of the present quantum mechanical theory. The need of such a generalization arises from the (probably) fundamental difficulties resulting when one attempts to apply quantum mechanics to questions in relativistic and nuclear phenomena.

The mathematical foundations underlying this generalization have already been investigated completely[2] and the results obtained are summarized in Part I of this paper except for the following alterations: 1) We restrict ourselves entirely to commutative systems while in the investigation mentioned the commutative law was partly excluded. Further, it is of basic importance that we assume the "field of coefficients" of the algebra considered to be always real. (In this way we meet an old objection to the quantum mechanical formalism, namely, that it does not confine itself to the real domain. Complex numbers and the like are used here only as technical expedients.) 2) We give the unpublished proof of the deducibility of the law $(a^2b)a = a(ba^2)$ from the law $a^\mu a^\nu = a^{\mu+\nu}$. 3) We give determinative conditions for the further investigation in Part III of the structure of the "r-number systems"; these conditions appear in (23), (24), and (25) below.

Part II contains a summary discussion of the concepts "trace" and "interchangeability"; the investigation of these concepts has already been begun.[2]

The essential achievement of this paper is contained in Part III. There the structure of all (formally real, finite, and commutative) r-number systems is fully determined. It is shown that, with but one exception, all such systems are merely matrix algebras. This is noteworthy in itself from the mathematical point of view. But of more importance is the fact that the above mentioned

[1] P. Jordan, Zschr. f. Phys., vol. 80, 1933, p. 285.

[2] P. Jordan, Göttinger Nachr., 1932, p. 569; 1933, p. 209.

physical considerations demand an algebra essentially more general than matrix algebras. Our results therefore point to the necessity of further generalizations. Although we found *one* r-number system which is not a matrix algebra ($\mathfrak{M}_3^8$, cf. paragraph 22), this system is too narrow for the generalization needed. It may nevertheless be usable as a first step in this direction.

Apart from the possibility of dropping the requirement of commutativity, one first thinks of omitting the condition of finiteness, all the more so since the algebra appearing in quantum mechanics is, as well known, infinite (that is, it has no finite linear basis). It may well happen that new types of algebra will arise with the removal of some of the present restrictions. This possibility is further suggested by the fact that in the ordinary quantum mechanics many important features first appeared in terms of an infinite algebra.[3] But it is too early to make conjectures concerning these difficult questions.

However, our procedure can be regarded as a phenomenological basis for the matrix calculus of the quantum theory.

We wish to express here our appreciation to Dr. C. C. Torrance for his valuable assistance in connection with the preparation of this manuscript.

Foundations of the theory

1. We consider an algebra A with elements $a, b, c, \cdots$ such that from the elements a and b we can form the sum $a + b$, the difference $a - b$, and the product ab; also from the element a we can form the element λa, where λ is a real number. (In what follows, greek letters will always denote real numbers.) In our algebra all the usual rules of calculation are to hold (including the commutative law of multiplication) with the single exception of the associative law of multiplication.[4] Our algebra will have a finite linear base, that is, there will be a finite number of linearly independent elements $a_1, a_2, \cdots, a_N$ such that each element of the algebra is linearly dependent upon these a's. Further, our algebra will be "formally real" in the sense that

(I) If $a^2 + b^2 + c^2 + \cdots = 0$, then[5] $a = b = c = \cdots = 0$.

We call this algebra an r-number algebra if it has the properties

(II) $$a^\mu a^\nu = a^{\mu+\nu}, \qquad (\mu, \nu = 1, 2, \cdots)$$

(III) $$(a^2b)a = a^2(ba),$$

[3] For example, the possibility of satisfying the commutation relations $pq - qp = \frac{h}{2\pi i} 1$, where h is a real number different from zero.

[4] That is, it is not necessary that $(ab)c = a(bc)$, though we desire that $(\lambda\mu)a = \lambda(\mu a)$ and that $(\lambda a)b = \lambda(ab)$.

[5] An algebra fulfilling this condition which, in a more general form, is better adapted to non-commutative algebras, has been called by Jordan[2] *semi-simple*. The present name, i.e., real, (used by E. Artin and I. Schreier, Hamb. Abh., vol. 5, 1926, p. 83) seems to be more suitable.

where a^μ is defined by

$$a^1 = a, \qquad a^{\mu+1} = a\,a^\mu \qquad (\mu = 1, 2, \cdots).$$

FUNDAMENTAL THEOREM 1. *According to the assumptions made above, including* (I), III *is a consequence of* II *and* II *is a consequence of* III.

The importance of this theorem lies in that (II) must be regarded as a necessary axiom because of the physical meaning of the theory,[6] whereupon the theorem shows that (III) is also a necessary law of any meaningful generalization of quantum mechanics.

The derivation of (III) from (II) is given at the end of paragraph 4. The derivation of (II) from (III) comes about in the following way:

We use the notation

$$[a, b, c] = (ab)c - a(bc). \tag{1}$$

From the commutative law it follows immediately that

$$[a, b, c] + [c, b, a] = 0, \tag{2}$$

$$[a, b, c] + [b, c, a] + [c, a, b] = 0. \tag{3}$$

If we replace b and a in (III) respectively by u and $\lambda a + \mu b + \nu c$, then an identity results in λ, μ, ν such that the terms proportional to $\lambda\mu\nu$ give

$$[ab, u, c] + [bc, u, a] + [ca, u, b] = 0. \tag{4}$$

Now suppose that it has already been shown that (II) follows from (III) for all μ and ν such that $\mu + \nu \leqq N$, where a^μ is determined by the definition given in (III). By (2) and (4) it follows that for $1 \leqq \nu \leqq N$

$$[a^{\nu+1}, b, a^{N-\nu-1}] = [a^\nu, b, a^{N-\nu}] + [a, b, a^{N-1}].$$

Hence

$$[a^\rho, b, a^{N-\rho}] = \rho[a, b, a^{N-1}] \text{ for } \rho = 1, 2, \cdots, N-1,$$

and therefore we have in particular that

$$-[a, b, a^{N-1}] = [a^{N-1}, b, a] = (N-1)\,[a, b, a^{N-1}],$$

$$[a, b, a^{N-1}] = 0, \qquad [a^\rho, b, a^{N-\rho}] = 0.$$

This gives the relation $a^{\rho+1}a^{N-\rho} = a^\rho a^{N-\rho+1} = a^1 a^N = a^{N+1}$ in the case where $b = a$. Hence (II) also holds for $\mu + \nu = N + 1$. Hence (II) holds in general and we have incidentally shown that $[a^\mu, b, a^\nu] = 0$.

In what follows we use (II) as a hypothesis; the further development will show that (III) is a consequence.

2. Before we undertake the actual investigation of A we point out that the mathematical foundation (equivalent to the physical basis, cf. footnote 1) may

[6] Cf. footnote 1.

also be based on $a \pm b$, λa, and a^μ instead of $a \pm b$, λa, and ab. For then we could define ab as $\frac{1}{2}((a+b)^2 - a^2 - b^2)$ from which it follows automatically that ab is commutative. Accordingly the postulation assumes the following form in this case:

In the algebra A with elements $a, b, c, \cdots$, one can form from the elements a and b the elements $a \pm b$, λa (λ a real number), and a^μ ($\mu = 1, 2, \cdots, a^1 = a$); all the usual rules of calculation hold for the operations $a \pm b$ and λa; and A has a finite linear basis. a^μ is a continuous function of a (that is, if we represent all the elements as linear aggregates of the basis elements, then the coefficients of a^μ depend continuously on those of a). A is "formally real", that is,

(I′) If $a^2 + b^2 + c^2 + \cdots = 0$, then $a = b = c = \cdots = 0$.

((I′) is the same as (I).) It is always the case that

(II′) $$(a^\mu + a^\nu)^2 = a^{2\mu} + a^{2\nu} + 2a^{\mu+\nu}.$$

((II′) is equivalent to (II), but it has the definition $a^1 = a$, $a^{\mu+1} = a\,a^\mu$ as a consequence.) Finally,

(*) $$(a+b+c)^2 + a^2 + b^2 + c^2 = (a+b)^2 + (a+c)^2 + (b+c)^2.$$

((*) is equivalent to $(a+b)c = ac + bc$, but because of the continuity of a^2 (and therefore also of ab), it has the condition $(\lambda a)b = \lambda(ab)$ as a consequence. This completes the list of the desired rules of calculation.—(*) could be replaced by the simpler rule $(a+b)^2 + (a-b)^2 = 2a^2 + 2b^2$.

All of these postulates, except for (*), necessarily arise from the assumed physical conditions.

Part I. Structure of the r-number algebras.

3. We now begin the investigation of the properties of A.

THEOREM 1. *If $a \neq 0$, then $a^\mu \neq 0$.*

Proof: If $a^\nu \neq 0$, then, by (I), $(a^\nu)^2 \neq 0$. Therefore $a^{\nu+1} \neq 0$ since we have $a^{\nu+1}a^{\nu-1} = a^{2\nu} = (a^\nu)^2$. (If $\nu = 1$, the factor $a^{\nu-1}$ is to be omitted.) Hence if $a \neq 0$, then $a^\mu \neq 0$.

We call an element $e = e^2 \neq 0$ a unit element (or idempotent element). If two distinct unit elements e and e' are such that $ee' = 0$, we say that they are mutually orthogonal. If $ee' = 0$, then $e + e'$ is also a unit element (because $(e+e')^2 = e^2 + e'^2 + 2ee' = e + e'$ and $e(e+e') = e^2 + ee' = e \neq 0$ so that $e + e' \neq 0$).

THEOREM 2. *Corresponding to each $a \neq 0$ there exists a unit element $e = e(a)$ (where e is a polynomial in a, and $e = e^2 \neq 0$) such that, for $\mu \geqq 2$, $ea^\mu = a^\mu$.*

Proof: Since the basis of A is finite, only a finite number of the elements $a, a^2, \cdots$ are linearly independent. Suppose that a^μ depends linearly on $a, \cdots, a^{\mu-1}$ so that

$$a^\mu + \lambda_{\mu-1}a^{\mu-1} + \cdots + \lambda_p a^p = 0.$$

Since $a \neq 0$ and $a^\mu \neq 0$, it follows that $\lambda_\rho \neq 0$ and $\rho < \mu$ if ρ is given its largest possible value. The condition $\rho > 2$ would imply that

$$a^\mu + \lambda_{\mu-1}a^{\mu-1} + \cdots + \lambda_\rho a^\rho = (a^{\mu-2} + \lambda_{\mu-1}a^{\mu-3} + \cdots + \lambda_\rho a^{\rho-2})a^2 = 0,$$

$$(a^{\mu-2} + \lambda_{\mu-1}a^{\mu-3} + \cdots + \lambda_\rho a^{\rho-2})^2 a^2 = 0,$$

so that, by Theorem 1,

$$(a^{\mu-2} + \lambda_{\mu-1}a^{\mu-3} + \cdots + \lambda_\rho a^{\rho-2})a = a^{\mu-1} + \lambda_{\mu-1}a^{\mu-2} + \cdots + \lambda_\rho a^{\rho-1} = 0.$$

Therefore if μ is given its smallest possible value it follows that $\rho = 1$ or 2. Now put

$$e = e(a) = \frac{(a^{\mu-\rho} + \lambda_{\mu-1}a^{\mu-\rho-1} + \ldots + \lambda_{\rho+1}a)^\rho}{(-\lambda_\rho)^\rho}. \tag{5}$$

Then

$$\frac{a^{\mu-\rho} + \lambda_{\mu-1}a^{\mu-\rho-1} + \ldots + \lambda_{\rho+1}a}{-\lambda_\rho}a^\rho = a^\rho.$$

Hence $ea^\rho = a^\rho$, and $ea^{\mu'} = a^{\mu'}$ if $\mu' \geqq \rho$, and therefore *a fortiori* if $\mu' \geqq 2$. Hence $e^2 = e\,e(a) = e$, and since $ea^2 = a^2 \neq 0$, it follows that $e \neq 0$, i.e., e is a unit element. (Note that, having dealt in this proof only with powers and polynomials involving one element a, we could use the associative law because of (II).)

The simplest special case of (II) is $a^2a^2 = a^4$, that is, $a^2a^2 = (a^2a)a$. If we now replace a by $\lambda a + \mu b$ and compare the terms proportional to $\lambda^3\mu$ and $\lambda^2\mu^2$ respectively, then

$$4(ab)a^2 = a^3b + a(a^2b) + 2a(a(ab)), \tag{6}$$

$$4(ab)^2 + 2a^2b^2 = a(ab^2) + b(ba^2) + 2a(b(ab)) + 2b(a(ab)). \tag{7}$$

If in (6) we replace a by a unit element e and replace b by x, then

$$2e(e(ex)) - 3e(ex) + ex = 0. \tag{8}$$

THEOREM 3. *If e is a unit element and if $N_\lambda = N_\lambda(e)$ is the set of x's such that $ex = \lambda x$, then N_λ is obviously a linear manifold, but if $\lambda \neq 0, \frac{1}{2}$, and 1, then N_λ consists of zero alone.*

Each x can be represented in the form

$$x = x_0 + x_{\frac{1}{2}} + x_1, \qquad x_0 \in N_0, \qquad x_{\frac{1}{2}} \in N_{\frac{1}{2}}, \qquad x_1 \in N_1, \tag{9}$$

in one and only one way.

Proof: If $x \in N_\lambda$, then, by (8), $(2\lambda^3 - 3\lambda^2 + \lambda)x = 0$, and if $\lambda \neq 0, \frac{1}{2}$, and 1, then $x = 0$. If $x \in N_\lambda$, then $2e(ex) - 3ex + x = (2\lambda^2 - 3\lambda + 1)x$. For $\lambda = 0$ this expression reduces to x and for $\lambda = \frac{1}{2}$ or 1 it reduces to 0. Hence, by (9),

$$\begin{cases} x_0 = 2e(ex) - 3ex + x, \\ x_{\frac{1}{2}} = -4e(ex) + 4ex, \\ x_1 = 2e(ex) - ex, \end{cases} \tag{10}$$

where the latter two formulas are obtained in a similar way. Thus equation (9) can have no solution other than (10). But (10) is a solution, for the equations $ex_0 = 0$, $ex_{\frac{1}{2}} - \frac{1}{2}x_{\frac{1}{2}} = 0$, and $ex_1 - x_1 = 0$ are equivalent to (8), and it is evident that $x = x_0 + x_{\frac{1}{2}} + x_1$.

THEOREM 4. *If $a \in N_0$ and $b \in N_0$, then $ab \in N_0$; if $a \in N_1$ and $b \in N_1$, then $ab \in N_1$; if $a \in N_0$ and $b \in N_1$, then $ab = 0$. Thus N_0 and N_1 are mutually orthogonal sub-algebras of A. If $a \in N_{\frac{1}{2}}$ and $b \in N_{\frac{1}{2}}$, then $ab \in (N_0 + N_1)$; if $a \in N_{\frac{1}{2}}$ and $b \in N_0$, then[7] $ab \in (N_0 + N_{\frac{1}{2}})$; if $a \in N_{\frac{1}{2}}$ and $b \in N_1$, then[7] $ab \in (N_{\frac{1}{2}} + N_1)$.*

Proof: in (7) we set $a = e$ and replace b by $\lambda a + \mu b$. If we consider the terms proportional to $\lambda\mu$, then

$$8(ea)(eb) + 4e(ab) - 2e(e(ab)) - (ea)b - (eb)a$$
$$- 2e(a(eb)) - 2e(b(ea)) - 2a(e(eb)) - 2b(e(ea)) = 0,$$

and if $a \in N_\rho$ and $b \in N_\sigma$, then

$$- 2e(e(ab)) + (4 - 2\rho - 2\sigma)e(ab) + (8\rho\sigma - 2\rho^2 - 2\sigma^2 - \rho - \sigma)ab = 0.$$

If we resolve ab according to Theorem 3, then

$$(8\rho\sigma - 2\rho^2 - 2\sigma^2 - \rho - \sigma)\,(ab)_0 + (8\rho\sigma - 2\rho^2 - 2\sigma^2 - 2\rho - 2\sigma + \tfrac{3}{2})\,(ab)_{\frac{1}{2}}$$
$$+ (8\rho\sigma - 2\rho^2 - 2\sigma^2 - 3\rho - 3\sigma + 2)\,(ab)_1 = 0.$$

Since this may be regarded as a resolution of zero in the sense of Theorem 3, each of the three terms separately vanish. Hence if one of the three conditions

$$8\rho\sigma - 2\rho^2 - 2\sigma^2 - \rho - \sigma \neq 0, \qquad 8\rho\sigma - 2\rho^2 - 2\sigma^2 - 2\rho - 2\sigma + \tfrac{3}{2} \neq 0,$$
$$8\rho\sigma - 2\rho^2 - 2\sigma^2 - 3\rho - 3\sigma + 2 \neq 0$$

is satisfied, the corresponding equation $(ab)_0 = 0$, $(ab)_{\frac{1}{2}} = 0$, $(ab)_1 = 0$ must be true. Substitution of ρ, $\sigma = 0, \frac{1}{2}$, and 1 leads to the desired results.

COROLLARY. *If $a \in N_{\frac{1}{2}}$, then, by Theorem 4, $a^2 \in (N_0 + N_1)$; yet were it the case that $a^2 \in N_0$ or $a^2 \in N_1$, then a would be zero.*

Proof: It would follow that $ea = \frac{1}{2}\,a$ and $ea^2 = \lambda a^2$ with $\lambda = 0$ or 1. Hence if $b = e$, then (6) becomes

$$2a^3 = ea^3 + \lambda a^3 + a^3,$$

so that $ea^3 = (1 - \lambda)a^3$ and $a^3 \in N_{1-\lambda}$. By Theorem 4, it follows that $a^5 = a^3a^2 = 0$ ($a^2 \in N_\lambda$, $a^3 \in N_{1-\lambda}$), so that $a = 0$.

THEOREM 5. *There exists one modulus 1, that is, one unit element $e = 1$ such that $1a = a$ for every a.*

Proof: We select a unit element e so that the resulting dimension δ of the algebra $N_0(e)$ is as small as possible. Suppose $N_0(e)$ contained an element $a \neq 0$. Then, by Theorem 2, it would contain a unit element e'. Since $e' \in N_0(e)$, it

[7] We shall later (Theorem 7) restrict this to $ab \in N_{\frac{1}{2}}$.

follows that $ee' = 0$. Hence $e + e'$ is also a unit element. Since $(e + e')e = e$, e belongs to $N_1(e + e')$, and if $a \in N_0(e + e')$ then, by Theorem 4, $ea = 0$ and $a \in N_0(e)$. Therefore $N_0(e + e') \subseteq N_0(e)$. $e' \in N_0(e)$, but e' does not belong to $N_0(e + e')$ since $(e + e')e' = e' \neq 0$. Therefore $N_0(e + e') \neq N_0(e)$. Hence the dimension of $N_0(e + e')$ is less than δ, contrary to assumption. Consequently $N_0(e)$ consists of 0 alone.

If $a \in N_{\frac{1}{2}}(e)$, then, by Theorem 4, $a^2 \in [N_0(e) + N_1(e)]$. By the above argument $a^2 \in N_1(e)$, and by the Corollary, $a = 0$. Hence $N_{\frac{1}{2}}(e)$ consists of 0 alone.

In the notation of (9) it follows that x always equals x_1, hence that $x \in N_1(e)$, and hence that $ex = x$. Therefore e is the modulus.

4. Because of the existence of the modulus we may restate in more complete form the considerations which led up to Theorem 2. Among the elements 1, $a, a^2, \cdots$ linear dependence must occur so that there exists a relation of the form

$$f(a) = a^\mu + \lambda_{\mu-1}a^{\mu-1} + \cdots + \lambda_1 a + \lambda_0 1 = 0. \tag{11}$$

If this relation is such that μ has the smallest possible value, then we call (11) the *characteristic equation* and we call μ the *order* of a. These notions are uniquely determined, for if there existed two equations $f(a) = 0$ and $g(a) = 0$ of this form, then $f(a) - g(a) = 0$ would be an equation of lower degree unless the polynomials f and g were identical.

Let ξ be a variable whose values are real numbers. Suppose the equation $f(\xi) = 0$ had a complex root $\xi = \rho + i\sigma$ with $\sigma \neq 0$. Then it would also have the root $\xi = \rho - i\sigma$. Hence $f(\xi)$ would be divisible by the polynomial

$$(\xi - (\rho + i\sigma))\,(\xi - (\rho - i\sigma)) = (\xi - \rho)^2 + \sigma^2.$$

Again, if $f(\xi) = 0$ had a multiple real root $\xi = \rho$, then $f(\xi)$ would again be divisible by $(\xi - \rho)^2 + \sigma^2$, where $\sigma = 0$. In either case $f(\xi)$ is of the form

$$f(\xi) = ((\xi - \rho)^2 + \sigma^2)g(\xi).$$

If $f(a) = 0$, then $f(a)g(a) = 0$, and hence $((a - \rho 1)g(a))^2 + (\sigma g(a))^2 = 0$. It follows by (I) that $(a - \rho 1)g(a) = 0$. But this is impossible since $(\xi - \rho)g(\xi)$ is a polynomial of the same form as $f(\xi)$ but of lower degree.

It follows that $f(\xi) = 0$ has μ distinct simple real roots $\xi_1, \xi_2, \cdots, \xi_\mu$ which we call the *proper values* of a.

If two polynomials $g(\xi)$ and $h(\xi)$ have the same value at each of the points $\xi = \xi_1, \xi_2, \cdots, \xi_\mu$, then $g(\xi) - h(\xi)$ is divisible by $f(\xi) = (\xi - \xi_1) \cdots (\xi - \xi_\mu)$ so that $g(\xi) - h(\xi) = f(\xi)k(\xi)$. Hence

$$g(a) - h(a) = f(a)k(a) = 0 \quad \text{and} \quad g(a) = h(a).$$

THEOREM 6. *Each element a may be represented as a linear combination of a set of pairwise orthogonal unit elements whose sum is the modulus; the coefficients in this representation are the proper values of a.*

Proof: Let $p_\nu(\xi) = \dfrac{f(\xi)}{(\xi - \xi_\nu)f'(\xi_\nu)}$ $(\nu = 1, 2, \cdots, \mu)$ be that polynomial whose value is zero when ξ has any of the values $\xi_1, \xi_2, \cdots, \xi_\mu$ except ξ_ν, in which case $p_\nu(\xi)$ has the value unity. Since $p_\nu(\xi)$ is not identically zero and since the degree of $p_\nu(\xi)$ is less than μ, it follows that $p_\nu(a) \neq 0$. Let

$$e_\nu = p_\nu(a) \qquad (\nu = 1, 2, \cdots, \mu). \tag{12}$$

For each $\xi = \xi_1, \xi_2, \cdots, \xi_\mu$, $p_\nu(\xi)^2 = p_\nu(\xi)$ and $p_\nu(\xi)p_{\nu'}(\xi) = 0$, where $\nu \neq \nu'$; also $p_1(\xi) + \cdots + p_\mu(\xi) = 1$ and $\xi_1 p_1(\xi) + \cdots + \xi_\mu p_\mu(\xi) = \xi$. It follows from the last remark before Theorem 6 that

$$\begin{cases} e_\nu^2 = e_\nu \neq 0,\ e_\nu e_{\nu'} = 0 \text{ for } \nu \neq \nu', \\ e_1 + \cdots + e_\mu = 1, \\ \xi_1 e_1 + \cdots + \xi_\mu e_\mu = a. \end{cases} \tag{13}$$

This completes the proof of all parts of the theorem.

Since the relation $g(\xi_1)p_1(\xi) + \cdots + g(\xi_\mu)p_\mu(\xi) = g(\xi)$ holds for any polynomial $g(\xi)$ for each $\xi = \xi_1, \xi_2, \cdots, \xi_\mu$, then

$$g(\xi_1)e_1 + \cdots + g(\xi_\mu)e_\mu = g(a). \tag{14}$$

If $\xi_\nu \neq 0$, we can find a polynomial $h(\xi)$ which is zero for $\xi = 0, \xi_1, \xi_2, \cdots$, and ξ_μ except ξ_ν, and which has the value unity for $\xi = \xi_\nu$. Then by (14), $h(a) = e_\nu$ and $h(\xi)$ contains no constant term. Hence $h(a)$ contains no term 1 but only terms $a, a^2, \cdots$. This proves the following

COROLLARY. *If a belongs to a sub-algebra B of A, then each of its e_ν whose $\xi_\nu \neq 0$ belongs to B.*

THEOREM 7. *Theorem 4 may be modified as follows: if $a \in N_{\frac{1}{2}}$ and if either $b \in N_0$ or $b \in N_1$, then $ab \in N_{\frac{1}{2}}$.*

Proof: By Theorem 6 and its corollary it is sufficient, since N_0 and N_1 are algebras, to consider only unit elements $b = e'$. $1 - e$ is a unit element, for $(1 - e)^2 = 1 - e$, and if $1 - e$ were zero, then it would be the case that $e = 1$ and $N_{\frac{1}{2}} = 0$, so that no proof would be necessary. Since the substitution of $(1 - e)$ for e carries N_0, $N_{\frac{1}{2}}$, and N_1 over into N_1, $N_{\frac{1}{2}}$, and N_0 respectively, it is sufficient to consider only those elements $b \in N_0$; for these elements, $ee' = 0$. (In this argument N_λ has really signified $N_\lambda(e)$.)

Since e and e' are orthogonal it follows that $e + e'$ is a unit element and according to (9) we are able to resolve a with respect to $e + e'$ as follows: $a = a_0 + a_{\frac{1}{2}} + a_1$. Since $a \in N_{\frac{1}{2}}(e)$ and $ea = \frac{1}{2}a$, we have

$$ea_0 + ea_{\frac{1}{2}} + ea_1 = \tfrac{1}{2}a_0 + \tfrac{1}{2}a_{\frac{1}{2}} + \tfrac{1}{2}a_1.$$

By application of Theorem 4 to $e + e'$ it follows that

$$ea_0 = 0, \qquad ea_{\frac{1}{2}} \in [N_0(e + e') + N_{\frac{1}{2}}(e + e')], \qquad ea_1 \in N_1(e + e')$$

since $(e + e')e = e$ and $e \in N_1(e + e')$. From the right hand side of the above equation we have that

$$(\tfrac{1}{2}a_0 + \tfrac{1}{2}a_{\frac{1}{2}}) \in [N_0(e + e') + N_{\frac{1}{2}}(e + e')], \qquad \tfrac{1}{2}a_1 \in N_1(e + e').$$

Because of the uniqueness of the resolution (9) it results that

$$ea_{\frac{1}{2}} = \tfrac{1}{2}a_0 + \tfrac{1}{2}a_{\frac{1}{2}}, \qquad ea_1 = \tfrac{1}{2}a_1.$$

By definition,

$$(e + e')a_0 = 0, \qquad (e + e')a_{\frac{1}{2}} = \tfrac{1}{2}a_{\frac{1}{2}}, \qquad (e + e')a_1 = a_1,$$

so that

$$e'a_0 = 0, \qquad e'a_{\frac{1}{2}} = -\tfrac{1}{2}a_0, \qquad e'a_1 = \tfrac{1}{2}a_1.$$

Hence $e'(e'a_{\frac{1}{2}}) = 0$. Resolving $a_{\frac{1}{2}}$ by (9) with respect to e' shows that this implies $a_{\frac{1}{2}} \in N_0(e')$. Therefore $-\tfrac{1}{2}a_0 = e'a_{\frac{1}{2}} = 0$ and $a_0 = 0$. Consequently

$$e'a = e'a_0 + e'a_{\frac{1}{2}} + e'a_1 = 0 + 0 + \tfrac{1}{2}a_1 = \tfrac{1}{2}a_1.$$

Since $ea_1 = \tfrac{1}{2}a_1$, $a_1 \in N_{\frac{1}{2}}(e)$, and $e'a \in N_{\frac{1}{2}}(e)$, we have finally that $ab \in N_{\frac{1}{2}}(e)$.

Theorem 6 and Formula (14) give us the means to prove (III), that is, that $[a^2, b, a] = 0$. If we resolve a by Theorem 6 then, by (14), it is sufficient to prove the relation $[e_\nu, b, e_{\nu'}] = 0$ for all ν and ν'. If $\nu = \nu'$ this relation follows from (2), but if $\nu \neq \nu'$, then $e_\nu e_{\nu'} = 0$ and it remains to show that

$$e(e'b) = e'(eb) \tag{15}$$

if e and e' are orthogonal unit elements. By Theorem 3 it suffices to consider the cases $b \in N_0(e)$, $b \in N_{\frac{1}{2}}(e)$, and $b \in N_1(e)$. $ee' = 0$ implies $e' \in N_0(e)$, and so theorems 4 and 7 prove the statement.

5. In accordance with the usual general terminology, the algebra A is called *irreducible* if it contains no invariant subalgebra B, that is, if no subalgebra B (except A itself and 0) is such that if $a \in A$ and $b \in B$, then $ab \in B$. A is called a *direct sum* if it is the sum of two orthogonal subalgebras B and C, that is, if every $a \in A$ can be written in the form $a = b + c$, where $b \in B$ and $c \in C$ and where B and C are such that if $b \in B$ and $c \in C$, then $bc = 0$ (neither B nor C being A or 0).[8] Every r-number algebra is completely reducible, that is, the properties of being reducible and of being a direct sum are equivalent. For it is clear that B is an invariant subalgebra if A is the direct sum of B and C. Conversely, if B is an invariant subalgebra then it has a modulus 1^*, and from its invariance it follows immediately that $N_1(1^*) = B$, $N_{\frac{1}{2}}(1^*) \subset B$, so that $N_{\frac{1}{2}}(1^*) = 0$. Consequently A is the direct sum of B and C if $N_0(1^*) = C$.

[8] If b belongs to B and C simultaneously, then $b^2 = 0$ and $b = 0$. Hence the resolution $a = b + c$ is unique, where $b \in B$ and $c \in C$.

We can represent every r-number algebra A as the direct sum of finitely many irreducible subalgebras $B_1, \cdots, B_k$ (all $\neq 0$), that is, each $a \in A$ is

$$a = b_1 + \cdots + b_k, \qquad b_1 \in B_1, \cdots, b_k \in B_k, \tag{16}$$

where if $b_i \in B_i$ and $b_j \in B_j$ for $i \neq j$, then $b_i b_j = 0$.[9] Moreover, $B_1, \cdots, B_k$ are uniquely determined except for their order. For let $C_1, \cdots, C_h$ be a second resolution of this kind. Since B_μ and C_ν are invariant subalgebras of A, so also is their intersection which is likewise an invariant subalgebra of B_μ and C_ν. Since both B_μ and C_ν are irreducible, their intersection must be 0 or else it must equal B_μ and C_ν. Hence if $B_\mu = C_\nu$, the intersection of B_μ and C_ν is zero. The sum of all the C_ν is A which in turn contains B_μ. Thus not all of these intersections can be zero. Hence for each μ there is a ν such that $B_\mu = C_\nu$. Because the C_ν are pairwise orthogonal there is only one such ν. Similarly for each ν there is one and only one μ such that $B_\mu = C_\nu$. Hence $k = h$ and the B_μ go over by a permutation into the C_ν.

The $B_1, \cdots, B_k$ are the *irreducible components* of A. It is apparent from the above discussion that it is sufficient to investigate the irreducible algebras A since the other algebras may be constructed from these in the simple manner just indicated. Consequently we shall confine our interest merely to the irreducible algebras, although we shall word our theorems for all algebras wherever the proofs do not make use of irreducibility.

6. We now examine systems of pairwise orthogonal unit elements $e_1, \cdots, e_\gamma$ ($e_\rho^2 = e_\rho \neq 0, e_\rho e_\sigma = 0$ for $\rho \neq \sigma$).

THEOREM 8. *Let $e_1, \cdots, e_\gamma$ be a system of pairwise orthogonal unit elements with the sum* 1. *Let the set of all x such that*[10]

$$e_\tau x = \tfrac{1}{2}(\delta_{\rho\tau} + \delta_{\sigma\tau})x \qquad (\tau = 1, \cdots, \gamma) \tag{17}$$

be denoted by $M^{\rho\sigma}$, where $\rho, \sigma = 1, \cdots, \gamma$.[11] *$M^{\rho\sigma}$ is obviously a linear manifold.*

Every x can be represented in the form

$$x = \sum_{\rho \leqq \sigma} x_{\rho\sigma}, \qquad x_{\rho\sigma} \in M^{\rho\sigma}, \tag{18}$$

in one and only one way.

Proof: It follows from (17) and (18) that

$$\begin{cases} 4e_\rho(e_\sigma x) = 4e_\sigma(e_\rho x) = x_{\rho\sigma} & (\rho \neq \sigma), \\ 2e_\rho(e_\rho x) - e_\rho x \qquad = x_{\rho\rho}. \end{cases} \tag{19}$$

[9] Consequently the resolution (16) is unique. Cf. footnote 8.

[10] $\delta_{\mu\nu}$ is the Kronecker symbol, that is, $\delta_{\mu\nu} = \begin{cases} 1 \text{ for } \mu = \nu, \\ 0 \text{ for } \mu \neq \nu. \end{cases}$

[11] It is obvious that $M^{\rho\sigma} = M^{\sigma\rho}$, so that for the most part we shall be able to limit ourselves to the consideration of $M^{\rho\sigma}$ for $\rho \leqq \sigma$.

Therefore (18) can have no solution other than (19) and, as we shall show, (19) is a solution of (18). ((15) shows that (19) is self-consistent.) Comparison of (19) with (10) shows that $x_{\rho\rho}$ belongs to $N_1(e_\rho)$. For $\tau \neq \rho$, $e_\tau e_\rho = 0$ so that these e_τ belong to $N_0(e_\rho)$. Hence for $\tau \neq \rho$, $e_\rho x_{\rho\rho} = x_{\rho\rho}$ and $e_\tau x_{\rho\rho} = 0$. Hence by (17), $x_{\rho\rho} \in M^{\rho\rho}$. If in (4) we set $a = b = e_\rho$, $c = x$, and $u = e_\sigma$, then for $\rho \neq \sigma$ (since $e_\rho^2 = e_\rho$ and $e_\rho e_\sigma = 0$)

$$[e_\rho, e_\sigma, x] + 2[e_\rho x, e_\sigma, e_\rho] = 0,$$

$$-e_\rho(e_\sigma x) + 2e_\rho(e_\sigma(e_\rho x)) = 0,$$

$$e_\rho(e_\sigma(e_\rho x)) = \tfrac{1}{2}e_\rho(e_\sigma x), \qquad e_\rho x_{\rho\sigma} = \tfrac{1}{2}x_{\rho\sigma}.$$

By symmetry it follows that $e_\sigma x_{\rho\sigma} = \frac{1}{2}x_{\rho\sigma}$. Hence $(e_\rho + e_\sigma)x_{\rho\sigma} = x_{\rho\sigma}$ and $x_{\rho\sigma} \in N_1(e_\rho + e_\sigma)$. If $\tau \neq \rho$ and σ, then $(e_\rho + e_\sigma)e_\tau = 0$ and $e_\tau \in N_0(e_\rho + e_\sigma)$ so that $e_\tau x_{\rho\sigma} = 0$. By (17), $x_{\rho\sigma} \in M^{\rho\sigma}$. Finally, $\left(\text{since } \sum_\rho e_\rho = 1\right)$

$$\sum_{\rho \leqq \sigma} x_{\rho\sigma} = \sum_\rho (2e_\rho(e_\rho x) - e_\rho x) + \sum_{\rho < \sigma} (2e_\rho(e_\sigma x) + 2e_\sigma(e_\rho x))$$

$$= 2\sum_{\rho,\sigma} e_\rho(e_\sigma x) - \sum_\rho e_\rho x = 2x - x = x.$$

This completes the proof of all parts of the theorem.

THEOREM 9. *Suppose $a \in M^{\rho\sigma}$ and $b \in M^{\pi\tau}$. Then: if both ρ and σ are distinct from both π and τ, then $ab = 0$; if ρ, σ and π, τ have one common index, say $\rho = \pi$ but $\sigma \neq \tau$, then $ab \in M^{\sigma\tau}$; if ρ, σ are the same pair as π, τ (except possibly for order; cf. footnote* 11), *say $\rho = \pi$ and $\sigma = \tau$, then $ab \in (M^{\rho\rho} + M^{\sigma\sigma})$ for $\rho \neq \sigma$ and $ab \in M^{\rho\rho}$ for $\rho = \sigma$.*

Proof: We set

$$e = \begin{cases} e_\rho + e_\sigma & \text{for } \rho \neq \sigma, \\ e_\rho & \text{for } \rho = \sigma, \end{cases} \qquad e' = \begin{cases} e_\pi + e_\tau & \text{for } \pi \neq \tau, \\ e_\pi & \text{for } \pi = \tau. \end{cases}$$

By (17) it follows that $ea = a$ and $e'b = b$ so that $a \in N_1(e)$ and $b \in N_1(e')$. If both ρ and σ are distinct from both π and τ, then it is obvious that $ee' = 0$ and $e' \in N_0(e)$. Hence, by Theorem 4, $e'a = 0$, $a \in N_0(e')$, and $ab = 0$.

Suppose that σ, τ, and $\rho = \pi$ are three distinct numbers. By Theorem 4, $B = N_1(e_\rho + e_\sigma + e_\tau)$ is an algebra and its modulus is $e_\rho + e_\sigma + e_\tau$. If we resolve the condition $(e_\rho + e_\sigma + e_\tau)x = x$ according to (17) and (18), the result shows that x can contain only those addends $x_{\mu\nu}$ in which $\mu, \nu = \rho, \sigma, \tau$. Hence $B = N_1(e_\rho + e_\sigma + e_\tau) = \sum_{\mu,\nu=\rho,\sigma,\tau} M^{\mu\nu}$. In a similar way it is apparent that the relations $N_1(e_\rho + e_\sigma) = \sum_{\mu,\nu=\rho,\sigma} M^{\mu\nu}$, $N_{\frac{1}{2}}(e_\rho + e_\sigma) = \sum_{\mu=\rho,\sigma} M^{\mu\tau}$, and $N_0(e_\rho + e_\sigma) = M^{\tau\tau}$ hold in B. So $a \in N_1$ and $b \in N_{\frac{1}{2}}$ and, by Theorem 7, $ab \in N_{\frac{1}{2}}$, where $N_{\frac{1}{2}} = M^{\rho\tau} + M^{\sigma\tau}$. By symmetry it follows that $ab \in (M^{\rho\sigma} + M^{\sigma\tau})$. Because of the uniqueness of the resolution (18), $ab \in M^{\sigma\tau}$.

If $\rho = \pi = \sigma \neq \tau$, we reach the same result by considering $B = N_1(e_\rho + e_\tau)$ and $N_1(e_\rho)$, $N_{\frac{1}{2}}(e_\rho)$, and $N_0(e_\rho)$. By symmetry the same result holds for $\rho = \pi = \tau \neq \sigma$. This accounts for all the cases where $\rho = \pi$ and $\sigma \neq \tau$.

If $\rho = \pi$ and $\sigma = \tau$, then $e = e'$. It follows that $a \in N_1(e)$ and $b \in N_1(e)$ so that $ab \in N_1(e)$. If $\rho = \sigma$, then $ab \in N_1(e_\rho)$, where $N_1(e_\rho) = M^{\rho\rho}$. If $\rho \neq \sigma$, then $ab \in N_1(e_\rho + e_\sigma)$, where $N_1(e_\rho + e_\sigma) = \sum_{\mu,\nu=\rho,\sigma} M^{\mu\nu} = M^{\rho\rho} + M^{\sigma\sigma} + M^{\rho\sigma}$. The relations $N_1(e_\rho) = M^{\rho\rho}$, $N_{\frac{1}{2}}(e_\rho) = M^{\rho\sigma}$, and $N_0(e_\rho) = M^{\sigma\sigma}$ (note (17) and (18)) hold in $B = N_1(e_\rho + e_\sigma)$. Hence $a \in N_{\frac{1}{2}}$, $b \in N_{\frac{1}{2}}$, and, by Theorem 4, $ab \in (N_0 + N_1)$, where $N_0 + N_1 = M^{\rho\rho} + M^{\sigma\sigma}$. This completes the proof of all parts of the theorem.

A unit element e is called *unresolvable* if it can not be represented as the sum of two orthogonal unit elements.

THEOREM 10. *Every unit element can be represented as the sum of pairwise orthogonal unresolvable unit elements.*

Proof: There exist representations of e as a sum of pairwise orthogonal unit elements:

$$e_1 + \cdots + e_\gamma = e.$$

For example, $e_1 = e$, $\gamma = 1$. Since $e_1, \cdots, e_\gamma$ are linearly independent (if $\lambda_1 e_1 + \cdots + \lambda_\gamma e_\gamma = 0$, then $e_\mu(\lambda_1 e_1 + \cdots + \lambda_\gamma e_\gamma) = \lambda_\mu e_\mu = 0$ and $\lambda_\mu = 0$ for all $\mu = 1, \cdots, \gamma$), then $\gamma \leqq N$, where N is the dimension of A. Let γ be as large as possible. Suppose an e_μ ($\mu = 1, \cdots, \gamma$) were resolvable, $e_\mu = e'_\mu + e''_\mu$. Then we would have

$$e_1 + \cdots + e'_\mu + e''_\mu + \cdots + e_\gamma = e.$$

Since $e_\mu e'_\mu = (e'_\mu + e''_\mu)e'_\mu = e'_\mu$, therefore $e'_\mu \in N_1(e_\mu)$. Likewise $e''_\mu \in N_1(e_\mu)$. For $\mu \neq \nu$, $e_\mu e_\nu = 0$ and $e_\nu \in N_0(e_\mu)$ so that $e'_\mu e_\nu = e''_\mu e_\nu = 0$. Thus the $\gamma + 1$ elements $e_\mu, \cdots, e'_\mu, e''_\mu, \cdots, e_\gamma$ are pairwise orthogonal. This contradicts our original assumption and every e_μ must be unresolvable.

THEOREM 11. *A unit element e is unresolvable when and only when the dimension of $N_1(e)$ is equal to unity.*

Proof: If e is resolvable, $e = e' + e''$, then $ee' = (e' + e'')e' = e'$ and $e' \in N_1(e)$. Likewise $e'' \in N_1(e)$, and e' and e'' are linearly independent. Hence the dimension is greater than unity.

Conversely, suppose that e is unresolvable. If $a \in N_1(e)$ then, in the resolution of a according to (13), each e_ν (for which $\xi_\nu \neq 0$) belongs to $N_1(e)$ by the corollary of Theorem 6. Hence $ee_\nu = e_\nu$. Since $e_\nu + (e - e_\nu) = e$, $(e - e_\nu)^2 = e - e_\nu$, and $e_\nu(e - e_\nu) = 0$, it follows that if $(e - e_\nu)$ were different from zero then e would be resolvable. Hence $e_\nu = e$. Hence there is at most one $\xi_\nu \neq 0$ so that $a = 0$ or $a = \xi_\nu e$. Therefore $N_1(e)$ contains only the elements λe and the dimension is unity.

On the basis of Theorem 10 we now assume a resolution of 1 in pairwise orthogonal unresolvable unit elements:

$$e_1 + \cdots + e_\gamma = 1,$$

and we form the corresponding $M^{\rho\sigma}$ in accordance with Theorem 8. Since $N_1(e_\rho) = M^{\rho\rho}$, then, by Theorem 11, each $M^{\rho\rho}$ has the dimension unity.

THEOREM 12. *Under the above assumptions* $a^2 = \alpha(e_\rho + e_\sigma)$, *where* a *runs through* $M^{\rho\sigma}$ *and* $\rho \neq \sigma$.

Proof: By Theorem 9 it follows that $a^2 \in (M^{\rho\rho} + M^{\sigma\sigma})$ and, by Theorem 11, $M^{\rho\rho}$ consists merely of the elements αe_ρ. Likewise $M^{\sigma\sigma}$ consists of the elements βe_σ, so that $a^2 = \alpha e_\rho + \beta e_\sigma$. By (III) and (17),

$$(a^2 e_\rho)a = a^2(e_\rho a), \qquad \alpha e_\rho \cdot a = (\alpha e_\rho + \beta e_\sigma)\cdot(e_\rho a),$$

$$\alpha \cdot \tfrac{1}{2}a = (\alpha + \beta)\cdot\tfrac{1}{2}\cdot\tfrac{1}{2}a, \qquad \tfrac{1}{4}(\alpha - \beta)\cdot a = 0,$$

so that if $a \neq 0$ then $\frac{1}{4}(\alpha - \beta) = 0$ and $\alpha = \beta$. If $a = 0$ then it follows immediately that $\alpha = \beta = 0$.

THEOREM 13. *Under the above assumptions there exists a linear basis* $s_1, \cdots, s_\chi$ *(or if it is desired to show the dependency of the basis on* ρ *and* σ, $s_1^{\rho\sigma}, \cdots, s_{\chi_{\rho\sigma}}^{\rho\sigma}$*) of* $M^{\rho\sigma}$ *such that the following multiplication rules are valid (see footnote* 10*):*

$$s_\mu s_\nu = \delta_{\mu\nu}(e_\mu + e_\nu) \qquad (\mu, \nu = 1, \cdots, \chi). \tag{20}$$

This basis is uniquely determined to within a (real) orthogonal transformation

$$s'_\mu = \sum_{\nu=1}^{\chi} \omega_{\mu\nu}\, s_\nu \qquad (\mu = 1, \cdots, \chi), \tag{21}$$

where the matrix $\|\omega_{\mu\nu}\|$ *is orthogonal.*

Proof: Since $ab = \frac{1}{2}((a + b)^2 - a^2 - b^2)$ it follows from Theorem 12 that ab likewise has the form $\alpha(e_\rho + e_\sigma)$. Therefore $\alpha = \alpha(a, b)$ is obviously linear in a and in b, is symmetric in a and b, and is consequently a symmetric bilinear form in a and b (that is, in the coefficients appearing in the expansion of ab with respect to $s_1, \cdots, s_\chi$). If an $\alpha(a, a)$ were not greater than zero, then it would be the case that

$$a^2 + (\sqrt{-\alpha(a, a)}\,(e_\rho + e_\sigma))^2 = \alpha(a, a)\,(e_\rho + e_\sigma) - \alpha(a, a)\,(e_\rho + e_\sigma) = 0$$

so that, by (I), $a = 0$. Hence if $a \neq 0$ then $\alpha(a, a) > 0$. Therefore the quadratic form $\alpha(a, a)$ is positive definite. Consequently $\alpha(a, b)$ can be transformed into the unit form by a linear transformation. Hence it is possible to select the basis of $M^{\rho\sigma}$ so that $\alpha(a, b)$ is the unit form of the expansion coefficients of a and b. $\alpha(s_\mu, s_\nu) = \delta_{\mu\nu}$ and (20) is proved.

It is clear that (20) is invariant under (21).

Formulas (17) and (20) together with the first and second multiplication rules of Theorem 9 (where we interchange ρ and σ) lead to the following multiplication table, providing that e_ρ and $s_\mu^{\rho\sigma}$ are selected in accordance with Theorems 12 and 13:

THEOREM 14. *Under the above assumptions it follows that* e_ρ $(\rho = 1, \cdots, \gamma)$ *and* $s_\mu^{\rho\sigma}$ $(\rho, \sigma = 1, \cdots, \gamma; \mu = 1, \cdots, \chi_{\rho\sigma}$, *where* $\rho \neq \sigma$ *and the pairs* ρ, σ *and* σ, ρ *are to be identified; see footnote* 11*) is a linear basis of* A *for which the following multiplication rules hold:*

$$
(22)\quad \begin{cases}
\text{(a)} & e_\rho\, e_\sigma = \begin{cases} e_\rho \text{ for } \rho = \sigma, \\ 0 \text{ for } \rho \neq \sigma, \end{cases} \\
\text{(b)} & e_\rho\, s_\mu^{\sigma\tau} = \begin{cases} \frac{1}{2}\, s_\mu^{\sigma\tau} \text{ for } \rho = \sigma \text{ or } \tau, \\ 0 \text{ for } \rho \neq \sigma \text{ and } \tau, \end{cases} \\
\text{(c)} & s_\mu^{\rho\sigma}\, s_\nu^{\rho\sigma} = \begin{cases} e_\rho + e_\sigma \text{ for } \mu = \nu, \\ 0 \text{ for } \mu \neq \nu, \end{cases} \\
\text{(d)} & s_\mu^{\rho\sigma}\, s_\nu^{\sigma\tau} = \sum_{\lambda=1}^{\chi_{\rho\tau}} A_{\nu\lambda\mu}^{\rho\sigma\tau}\, s_\lambda^{\rho\tau} \text{ for } \rho \neq \tau, \\
\text{(e)} & s_\mu^{\rho\sigma}\, s_\nu^{\tau\upsilon} = 0 \text{ for } \rho, \sigma \neq \tau, \upsilon.
\end{cases}
$$

The proof was given above.

7. The algebra A is completely described by (22) so that the only data we need are the integers γ and $\chi_{\rho\sigma}$ and the real numbers $A_{\nu\lambda\mu}^{\rho\sigma\tau}$. Our next problem is to deduce from the r-number character of A some conditions which will limit these numbers more closely.

THEOREM 15. *If ρ, σ, τ, and υ are four distinct numbers then, under the above assumptions,*

$$
(23)\qquad A_{\nu\lambda\mu}^{\rho\sigma\tau} = A_{\lambda\mu\nu}^{\sigma\tau\rho} = A_{\mu\nu\lambda}^{\tau\rho\sigma} = A_{\lambda\nu\mu}^{\sigma\rho\tau} = A_{\nu\mu\lambda}^{\rho\tau\sigma} = A_{\mu\lambda\nu}^{\tau\sigma\rho},
$$

$$
(24)\qquad \sum_{\nu=1}^{\chi_{\sigma\tau}} (A_{\nu\lambda\mu}^{\rho\sigma\tau}\, A_{\nu\lambda'\mu'}^{\rho\sigma\tau} + A_{\nu\lambda\mu'}^{\rho\sigma\tau}\, A_{\nu\lambda'\mu}^{\rho\sigma\tau}) = \tfrac{1}{2}\, \delta_{\lambda\lambda'}\, \delta_{\mu\mu'} \quad \text{(see footnote 10)},
$$

$$
(25)\qquad \sum_{\mu=1}^{\chi_{\sigma\upsilon}} A_{\mu\theta_\iota}^{\rho\sigma\upsilon}\, A_{\kappa\mu\lambda}^{\sigma\tau\upsilon} = \sum_{\nu=1}^{\chi_{\rho\tau}} A_{\kappa\theta\nu}^{\rho\tau\upsilon}\, A_{\lambda\nu_\iota}^{\rho\sigma\tau}.
$$

Proof: We substitute for a, b, c, and u in (4) the following combinations of the elements of the basis in Theorem 14: $s_\mu^{\rho\sigma}$, $s_\nu^{\sigma\tau}$, e_τ, $s_\lambda^{\rho\tau}$; $s_\mu^{\rho\sigma}$, $s_{\mu'}^{\rho\sigma}$, $s_\lambda^{\rho\tau}$, e_τ; $s_\iota^{\rho\sigma}$, $s_\lambda^{\sigma\tau}$, $s_\kappa^{\tau\upsilon}$, e_ρ. The first substitution gives $A_{\nu\lambda\mu}^{\rho\sigma\tau} = A_{\mu\nu\lambda}^{\tau\rho\sigma}$ and (23) follows by cyclic permutation of the indices. (24) follows from the second substitution with the aid of (23). The third substitution leads to (25).

Although the relations (23), (24), and (25) are merely necessary and not sufficient, yet they will enable us in what follows, particularly in Part III, to determine completely all r-number algebras.

We now draw a few conclusions.

THEOREM 16. *If $\chi_{\rho\tau} \neq 0$, then under the above assumptions $\chi_{\rho\sigma} = \chi_{\rho\tau}$.*

Proof: Since $\chi_{\rho\tau} \geqq 1$, we can set $\lambda = \lambda' = 1$ in (24) and it follows that $2\sum_{\nu=1}^{\chi_{\sigma\tau}} A_{\nu 1\mu}^{\rho\sigma\tau}\, A_{\nu 1\mu'}^{\rho\sigma\tau} = \frac{1}{2}\delta_{\mu\mu'}$. Therefore the vectors $2A_{11\mu}^{\rho\sigma\tau}, \cdots, 2A_{\chi_{\sigma\tau}1\mu}^{\rho\sigma\tau}$, where $\mu = 1, \cdots, \chi_{\rho\sigma}$, form a normalized orthogonal system. Since their dimension

is $\chi_{\sigma\tau}$ and since there are $\chi_{\rho\sigma}$ of these vectors, $\chi_{\rho\sigma} \leqq \chi_{\sigma\tau}$. Interchange of ρ and τ gives $\chi_{\sigma\tau} \leqq \chi_{\rho\sigma}$, whence $\chi_{\rho\sigma} = \chi_{\sigma\tau}$.

THEOREM 17. *If A is irreducible then, under the above assumptions, all the $\chi_{\rho\sigma}$ are equal to each other and distinct from zero.*

Proof: Let N be the set of all ρ for which a chain $\rho_0 = 1, \rho_1, \cdots, \rho_{\epsilon-1}, \rho_\epsilon = \rho$ $(\epsilon = 0, 1, 2, \cdots)$ exists such that every $\chi_{\rho_{\eta-1}\rho_\eta} \neq 0$ for $\eta = 1, \cdots, \epsilon$. Let M be the set of all other ρ. N is not empty since $1 \in N$. If $\rho \in N$ and $\sigma \in M$, then $\chi_{\rho\sigma} = 0$. Hence $M^{\rho\sigma}$ contains only the element 0. $e = \sum_{\rho \in N} e_\rho$ is a unit element. By (17) and (18) it follows that the condition $ex = \frac{1}{2}x$ is equivalent to the condition that, in the resolution (18) of x, only those addends $x_{\rho\sigma}$ occur for which $\rho \in N$ and $\sigma \in M$. Hence $x_{\rho\sigma} \in M^{\rho\sigma}$ and $x_{\rho\sigma} = 0$ so that $x = 0$. Therefore $N_{\frac{1}{2}}(e) = 0$ and A is the sum of the two orthogonal subalgebras $N_0(e)$ and $N_1(e)$. Since A is irreducible and $N_1(e) \neq 0$, then $N_0(e) = 0$. But every e_ρ, $\rho \in M$, is distinct from zero and belongs to $N_0(e)$ (since $ee_\rho = 0$). Hence M is empty and every ρ, $\rho = 1, \cdots, \gamma$, belongs to N. Thus we can set up the chain $\rho_0 = 1, \rho_1, \cdots, \rho_{\epsilon-1}, \rho_\epsilon = \rho$ for every ρ.

From Theorem 16 it follows that $\chi_{\rho_{\eta-1}\sigma} = \chi_{\rho_\eta\sigma}$ for every σ since $\chi_{\rho_{\eta-1}\rho_\eta} \neq 0$. Therefore $\chi_{\rho_0\sigma} = \chi_{\rho_\epsilon\sigma}$ and $\chi_{\rho\sigma} = \chi_{1\sigma}$. But we can write $\chi_{1\sigma}$ as $\chi_{\sigma 1}$. By the preceding argument it follows that $\chi_{1\sigma} = \chi_{11}$ so that $\chi_{\rho\sigma} = \chi_{11}$.

If we denote the common value of the $\chi_{\rho\sigma}$ by χ, then Theorem 14 provides the following condition on the dimension N of A:

$$N = \gamma + \frac{\gamma(\gamma - 1)}{2}\chi, \qquad \chi > 0. \tag{26}$$

It should be noted that, while N is uniquely determined by A, this by no means needs to be the case for γ and χ as introduced in Theorems 12, 13, 14 and 17. We shall see later (Theorem 23 in Part II) that γ and χ are actually so determined.

Part II. Trace, Commutativity

8. The investigation of the concepts mentioned in the title of this part is not absolutely necessary for the logical connection of the entire paper since in Part III we shall determine directly from Part I all r-number algebras. And in the concrete instances taken up in Part III these concepts can be set up directly without reference to the general theory. But it is our desire to develop in general form for the sake of eventual generalizations our present concepts which are in themselves not uninteresting. Moreover we arrive at the proof of the determinateness of γ and χ. Finally we touch on the application of these concepts to the forming of quantum mechanical concepts.

All the following considerations arise from the setting up of the operation $O_a x = ax$, where $x \in A$. If A has a linear basis $a_1, \cdots, a_N$ then a matrix $\|T_{\alpha\beta}(a)\|$, $\alpha, \beta = 1, \cdots, N$, corresponds to O_a:

$$a\, a_\alpha = \sum_{\beta=1}^{N} T_{\alpha\beta}(a)\, a_\beta \qquad (\beta = 1, \ldots, N). \tag{27}$$

If the basis is altered then $\|T_{\alpha\beta}(a)\|$ is merely transformed as follows:

$$\begin{cases} a'_\alpha = \sum_{\beta=1}^{N} \psi_{\alpha\beta}\, a_\beta, & \beta = 1, \ldots, N, \\ \|T'_{\alpha\beta}(a)\| = \|\psi_{\alpha\beta}\| \cdot \|T_{\alpha\beta}(a)\| \cdot \|\psi_{\alpha\beta}\|^{-1}. \end{cases} \tag{28}$$

Therefore the concepts which are invariant under transformations, such as trace $\|T_{\alpha\beta}(a)\| = \sum_{\alpha=1}^{N} T_{\alpha\alpha}(a)$ and the commutativity of $\|T_{\alpha\beta}(a)\|$ and $\|T_{\alpha\beta}(b)\|$ (that is, of O_a and O_b), are independent of the selection of the basis.

9. Because of the non-associativity of multiplication in A it need not be the case that O_aO_b is either O_{ab} or O_{ba}, for otherwise it would be necessary that the relation $a(bx) = (ab)x$, that is, $[a, b, x] = 0$, hold for all x. From $ab = ba$ it does not follow that $O_aO_b = O_bO_a$ (that is, the commutativity of the corresponding matrices), for this would imply the relations $a(bx) = b(ax)$ and $a(xb) = (ax)b$, that is, $[a, x, b] = 0$ for all x. Consequently it is impossible to define even the general commutativity of multiplication as a concept of interchangeability in A.

Accordingly we call a and b *interchangeable* if the operations O_a and O_b are interchangeable, that is, if the matrices $\|T_{\alpha\beta}(a)\|$ and $\|T_{\alpha\beta}(b)\|$ are interchangeable, and hence if

$$[a, x, b] = 0 \tag{29}$$

holds for all x. The *centrum* of A consists of those elements a with which all elements b are interchangeable and hence which are such that

$$[a, x, y] = 0 \tag{30}$$

holds for all x and y. In this event it follows by (2) and (3) that

$$[x, a, y] = 0 \text{ and } [x, y, a] = 0. \tag{31}$$

With regard to the preceding definition of interchangeability it must be kept in mind that it is essential that the basis of A be finite. In important instances of quantum mechanics analogous to the r-number algebras the situation is quite otherwise. There the elements $a, b, c, \cdots$ are infinite Hermitian matrices $A, B, C, \cdots$ and our product

$$AB = \frac{A \times B + B \times A}{2},$$

where $A \times B$ is the usual associative non-commutative matrix multiplication. (See footnote 1 as well as the discussion in the introduction.) A simple calculation shows that

$$[A, B, C] = \frac{B \times (A \times C - C \times A) - (A \times C - C \times A) \times B}{4},$$

so that interchangeability occurs in our sense only when $A \times C - C \times A = \lambda 1$. It follows from this that $A \times C - C \times A = 0$ only for finite matrices but not for infinite matrices.[12] In generalizations to infinite r-number algebras it is necessary to exercise caution with regard to interchangeability. It is clear that the concept of centrum must remain unaltered.

THEOREM 18. *The elements b which are interchangeable with a form an algebra $C(a)$ which contains* 1. *If one represents a according to Theorem* 6 *and formula* (13), *then the algebras $N_1(e_1), \cdots, N_1(e_\mu)$ are pairwise orthogonal and $C(a)$ is their (direct) sum.*

Proof: For the resolution (18) of a with respect to the elements $e_1, \cdots, e_\mu$ it follows that $N_1(e_\nu) = M^{\nu\nu}$ and $N_1(e_{\nu'}) = M^{\nu'\nu'}$. By Theorem 9 these two algebras (cf. Theorem 4) are orthogonal for $\nu \neq \nu'$. Consequently $N_1(e_1) + \cdots + N_1(e_\mu)$ is an algebra which contains $1 = e_1 + \cdots + e_\mu$.

To show that a is interchangeable with each element of $N_1(e_1) + \cdots + N_1(e_\mu)$ it is sufficient to show that it is interchangeable with each element of every $N_1(e_{\nu'})$. Since $a = \xi_1 e_1 + \cdots + \xi_\mu e_\mu$ it is sufficient to show that each $x \in N_1(e_{\nu'})$ is interchangeable with each e_ν. By the above argument it follows that $x \in N_1(e_\nu)$ for $\nu = \nu'$ and that $x \in N_0(e_\nu)$ for $\nu \neq \nu'$. We now have to show that $[x, y, e_\nu] = 0$ for all y, but by Theorem 3 it suffices to consider the cases $y \in N_0(e_\nu)$, $y \in N_{\frac{1}{2}}(e_\nu)$, and $y \in N_1(e_\nu)$. In other words it remains to show that $(xy)e = x(ye)$ for all six combinations of $x \in N_0(e)$ and $x \in N_1(e)$ with $y \in N_0(e)$, $y \in N_{\frac{1}{2}}(e)$, and $y \in N_1(e)$. But this follows immediately from Theorems 4 and 7.

Conversely, suppose that b is interchangeable with a. We set $x = e_{\nu'}$ in $[a, x, b] = 0$ and resolve b according to (18) with respect to the $e_1, \cdots, e_\mu$ corresponding to a. By (17) one concludes immediately that $[e_\nu, e_{\nu'}, b^{\rho\sigma}]$, i.e., $(e_\nu e_{\nu'})b^{\rho\sigma} - e_\nu(e_{\nu'}b^{\rho\sigma})$ differs from zero only if $\rho \neq \sigma$, $\nu = \rho$, and $\nu' = \sigma$ (or, what amounts to the same thing, $\nu = \sigma$ and $\nu' = \rho$) in which case it equals $-\frac{1}{4}b^{\rho\sigma}$. Consequently

$$[a, e_{\nu'}, b] = -\frac{1}{4}\sum_{\nu \neq \nu'} \xi_\nu b^{\nu\nu'}.$$

Therefore $b^{\nu\nu'} = 0$ for $\nu \neq \nu'$ if ξ_ν is not zero. So $b^{\nu\nu'} \neq 0$, $\nu \neq \nu'$ implies $\xi_\nu = 0$. By symmetry, it implies $\xi_{\nu'} = 0$. But this contradicts the fact that all the ξ_ν are distinct (note the discussion preceding Theorem 6). Hence

$$b = \sum_{\nu=1}^{\mu} b^{\nu\nu} \text{ and } b \in (N_1(e_1) + \ldots + N_1(e_\mu)).$$

[12] For example, A and C may be the matrices arising from the well-known quantum mechanical operators x and $i\frac{d}{dx}$.

THEOREM 19. *The centrum Z is an algebra which contains* 1. *Its degree is the number k of the irreducible components of A.* (*Note the end of paragraph* 5.)

Proof: By Theorem 18, the first assertion follows from the fact that Z is the common part of all $C(a)$. It can also be deduced directly from the identity[13]

$$[ab, x, y] = [a, bx, y] - [a, b, xy] + a[b, x, y] + [a, b, x]y.$$

If $A = B_1 + \cdots + B_k$ is the representation of A as the direct sum of irreducible algebras, then let $\bar{e}_1, \cdots, \bar{e}_k$ be the moduli of $B_1, \cdots, B_k$. $\bar{e}^2 = \bar{e}_i \neq 0$ and $\bar{e}_i\bar{e}_j = 0$ for $i \neq j$ (B_i and B_j orthogonal); if $x_j \in B_j$, then $\bar{e}_i x_j = x_j$ for $i = j$ and $\bar{e}_i x_j = 0$ for $i \neq j$ (also because of orthogonality). Thus $(\bar{e}_1 + \cdots \bar{e}_k) x = x$ for $x = x_j$ and accordingly for every x. Hence $\bar{e}_1 + \cdots + \bar{e}_k = 1$. According to the discussion at the end of paragraph 5 it follows that $N_1(\bar{e}_i) = 0$ so that $A = N_0(\bar{e}_i) + N_1(\bar{e}_i) = N_1(1 - \bar{e}_i) + N_1(\bar{e}_i)$. Thus the resolution of $\bar{e}_i$ according to (13) is $1 \cdot \bar{e}_i + 0 \cdot (1 - \bar{e}_i)$ so that, by Theorem 18, $C(\bar{e}_i) = A$. Hence $\bar{e}_i$ is interchangeable with every b and $\bar{e}_i \in Z$. But $\bar{e}_1, \cdots, \bar{e}_k$ are linearly independent and the dimension of Z is greater than or equal to k.

Conversely, suppose that a belongs to Z. Then every polynomial in a belongs to Z and, in particular (because of (12)), the elements $e_1, \cdots, e_\mu$ of the resolution (13). We shall show that $e_1, \cdots, e_\mu$ are linear aggregates of $\bar{e}_1, \cdots, \bar{e}_k$ and hence that a is such an aggregate. It follows from this that the dimension of Z is k.

Let $e \in Z$ be a unit element. Its resolution according to (13) is $1 \cdot e + 0 \cdot (1 - e)$. Thus $C(e) = N_0(e) + N_1(e)$. Since $e \in Z$, $C(e) = A$, i.e., $N_0(e) + N_1(e) = A$. N_0 and N_1 are of course orthogonal. If we resolve the algebras N_0 and N_1 into their irreducible components: $N_0 = B_1' + \cdots + B_\alpha'$, $N_1 = B_1'' + \cdots + B_\beta''$, then we have the resolution $A = B_1' + \cdots + B_\alpha' + B_1'' + \cdots + B_\beta''$ which must be a permutation of $A = B_1 + \cdots + B_k$. But the modulus e of N_1 is the sum of the moduli of $B_1'', \cdots, B_\beta''$ (cf. the corresponding situation relative to A in the preceding theorem). Hence e is the sum of the moduli of some of the B_i, that is, of some of the $\bar{e}_i$. Hence e is a linear aggregate of $\bar{e}_1, \cdots, \bar{e}_k$ as asserted above.

By Theorem 19, A is irreducible (that is, $k = 1$) when and only when the dimension of Z is unity, that is, when A consists of the elements $\lambda 1$ alone.

10. The trace of a is a (real) number defined by the equation

$$\text{trace } a = \Delta \text{ trace } \|T_{\alpha\beta}(a)\| = \Delta \sum_{\alpha=1}^{N} T_{\alpha\alpha}(a).$$

As we know, the trace is independent of the selection of the basis of A (cf. the end of paragraph 8). The normalization factor Δ is some positive number which we shall select later.

[13] Cf. M. Zorn, Hamb. Abh., vol. 8, 1930, p. 123. This follows from the distributive law of multiplication alone.

It is clear that the trace is linear in a. The relation (which we do not prove)

$$O_{[a,\,b,\,c]} = O_b(O_aO_c - O_cO_a) - (O_aO_c - O_cO_a)O_b$$

implies the same relation for the corresponding matrices. From this it follows that trace $[a, b, c] = 0$ (since the trace of every matrix commutator vanishes). This is noteworthy because of its analogy with certain formulas of the quantum mechanical matrix-theory, but it will not be used in the sequel. We now consider other properties of the trace which are essential for the development of the quantum mechanical calculus.

THEOREM 20. *If e and e' are two unresolvable unit elements, there exists a resolution of* 1 *in pairwise orthogonal unresolvable unit elements:* $e_1 + \cdots + e_\gamma = 1$, *where* $e_1 = e$ *and* $\gamma \geqq 2$, *so that*

$$(32) \qquad \begin{cases} e' = \alpha e_1 + (1 - \alpha)e_2 + x & (0 \leqq \alpha \leqq 1,\ x \in M^{12}), \\ x^2 = \alpha(1 - \alpha)\,(e_1 + e_2) & \text{(Cf. Theorem 12).} \end{cases}$$

Proof: We make the resolution $e' = x_0 + x_{\frac{1}{2}} + x_1$, where $x_0 \in N_0(e)$, $x_{\frac{1}{2}} \in N_{\frac{1}{2}}(e)$, and $x_1 \in N_1(e)$. Since e is unresolvable, it follows by Theorem 11 that $x_1 = \alpha e$. We resolve x_0 in the algebra $N_0(e)$ according to (13): $x_0 = \xi_\nu e_2 + \cdots + \xi_\gamma e_\gamma$. (We begin with the index 2 for formal reasons.) Since we can represent each e_ν as a sum of pairwise orthogonal unit elements (this sum being an element of $N(e_\nu)$ so that the unit elements for distinct ν are mutually orthogonal), we can arrive at the same representation with unresolvable unit elements. So we may assume that the preceding e_ν are unresolvable, but then the ξ_ν need no longer be distinct. The elements $e_2, \cdots, e_\gamma$ are all orthogonal to e_1 and their sum is the modulus $1 - e$ of $N_0(e) = N_1(1 - e)$. Hence if $e_1 = e$ then $e_1 + \cdots + e_\gamma = 1$ is a resolution of 1 of the desired sort. Since $N_{\frac{1}{2}}(e) = N_{\frac{1}{2}}(e_1) = M^{12} + \cdots + M^{1\gamma}$ (cf. the analogous argument in the proof of Theorem 9) we have

$$e' = \alpha e_1 + \xi_2 e_2 + \cdots + \xi_\gamma e_\gamma + x_2 + \cdots + x_\gamma,$$

where $x_2 \in M^{12}, \cdots, x_\gamma \in M^{1\gamma}$. By Theorem 12 we have that $x_\nu^2 = \eta_\nu(e_1 + e_\nu)$ and therefore that

$$e'^2 = (\alpha^2 + \eta_2 + \ldots + \eta_\gamma)\, e_1 + (\xi_2^2 + \eta_2)\, e_2 + \ldots + (\xi_\gamma + \eta_\gamma)\, e_\gamma + (\alpha + \xi_2)\, x_2 + \ldots + (\alpha + \xi_\gamma)\, x_\gamma + \sum_{\nu<\nu'} x_\nu x_{\nu'} \qquad (\nu, \nu' = 2, \ldots, \gamma).$$

As was the case with e', the terms of the summation belong to $M^{\nu\nu'}$ $(\nu, \nu' = 2, \cdots, \gamma;\ \nu < \nu')$, while the other terms in this equation belong respectively to $M^{11}, M^{22}, \cdots, M^{\gamma\gamma}, M^{12}, \cdots, M^{1\gamma}$. Since $e'^2 = e'$, therefore $x_\nu x_{\nu'} = 0$. With the help of formulas (22 *c*, *d*), (23) and (24) one may readily calculate that $x_\nu(x_\nu x_{\nu'}) = \frac{1}{4}\eta_\nu x_{\nu'}$ (represent x_ν and $x_{\nu'}$ as linear aggregates of the bases $s_i^{1\nu}$ and $s_j^{1\nu'}$). Hence either $x_{\nu'} = 0$ or $\eta_\nu = 0$ so that $x_\nu^2 = 0$ and $x_\nu = 0$. Hence at most one x_ν is different from zero and, by a permutation of the elements $e_2, \cdots, e_\gamma$, we

may take it to be x_2. Consequently $\eta_3 = \cdots = \eta_\gamma = 0$, $x_3 = \cdots = x_\gamma = 0$, and, since $e'^2 = e'$,

$$(33) \qquad \begin{cases} \alpha^2 + \eta_2 = \alpha, \xi_2^2 + \eta_2 = \xi_2, \xi_3^2 = \xi_3, \ldots, \xi_\gamma^2 = \xi_\gamma, \\ \alpha + \xi_2 = 1 \text{ for } x_2 \neq 0. \end{cases}$$

By (33) every ξ_ν is either zero or unity for $\nu = 3, \cdots, \gamma$. A solution remains a solution if a ξ_ν with value unity is given the value zero instead. If $\xi_\nu = 1$, then e_ν and $e' - e_\nu$ are two orthogonal unit elements provided that $e' - e_\nu \neq 0$. Since e' is unresolvable, it must be the case that $e' = e_\nu$. By a permutation of the elements $e_2, e_3, \cdots, e_\gamma$ it can be arranged that $e' = e_2$. This is in agreement with (32) when $\alpha = 0$ and $x = 0$. If, on the contrary, every ξ_ν is zero for $\nu = 3, \cdots, \gamma$, then we may draw the following conclusion from (33): If $x = 0$, then e' has one of the values $e_1, e_2, e_1 + e_2$, or 0. But it is obvious that 0 cannot be a value of e'. Neither can $e_1 + e_2$ because of our hypothesis of unresolvability. Hence e is either e_1 or e_2. This agrees with (32) in the case where α is 1 or 0 and $x = 0$. But if $x \neq 0$, then $\xi_2 = 1 - \alpha$, $\eta_2 = \alpha(1 - \alpha) > 0$ so that $0 < \alpha < 1$. This agrees with (32) in the case where α is neither 0 nor 1.

THEOREM 21. *If e and e' are non-orthogonal unit elements,* trace $(ee') > 0$.

Proof: By Theorem 10 we can represent e and e' as sums of unresolvable unit elements. Consequently it is sufficient to consider the case where e and e' are unresolvable. By (32) it follows that

$$ee' = \alpha e_1 + \tfrac{1}{2} x, \text{ where } x \epsilon M^{12}, x^2 = \alpha (1 - \alpha).$$

If we now select the basis of A defined in Theorem 14, then it follows by (22) and the definition of trace that

$$\text{trace } e_\rho = \Delta \left(1 + \tfrac{1}{2} \sum_{\rho \neq \sigma} \chi_{\rho\sigma}\right) > 0, \quad \text{trace } s_\mu^{\rho\sigma} = 0 \quad (\rho, \sigma = 1, \ldots, \gamma).$$

Hence if $\alpha > 0$, then trace $(ee') = \alpha/2$ trace $(e_1) > 0$. But since $0 \leqq \alpha \leqq 1$, therefore the only alternative is $\alpha = 0$, which in turn implies $x^2 = 0$, $x = 0$, and therefore $ee' = 0$.

THEOREM 22. *If A is irreducible then Δ can be given such a positive value that* trace $(e) = 1$ *for every unresolvable unit element e.*

Proof: For every resolution $e_1 + \cdots + e_\gamma$ of 1 in pairwise orthogonal unresolvable unit elements it follows by the trace formulas in the proof of Theorem 21 and the remarks at the end of paragraph 7 that

$$\text{trace } e_\rho = \Delta \left(1 + \frac{(\gamma - 1)\chi}{2}\right) = \frac{\Delta N}{\gamma}, \quad \text{trace } s_\mu^{\rho\sigma} = 0.$$

Thus γ, χ, and all trace e_ρ have the same values for every such resolution having the same $e_1 = e$. For any other unresolvable unit element e' it follows by (32) that trace e' = trace e_1 = trace e since trace $x = 0$ as all trace $s_\mu^{12} = 0$. There-

fore every unresolvable unit element e has the same trace. Consequently γ and χ always have the same value. We may therefore set

$$\Delta = \frac{1}{1 + \frac{(\gamma - 1)\chi}{2}} = \frac{\gamma}{N} > 0 \tag{34}$$

and arrive at the result[14] that trace $e = 1$.

We shall always select Δ in this way if A is irreducible. We now formulate two immediate consequences.

THEOREM 23. *The numbers γ and χ defined at the end of paragraph 7 for an irreducible A are uniquely determined.*

Proof: Note the remark in the proof of Theorem 22.

THEOREM 24. *If A is irreducible a unit element e may be represented (in the sense of Theorem* 10) *as a sum of unresolvable unit elements. The number of addends is then always equal to* trace e.

Proof: The theorem follows immediately from Theorem 22.

The physical meaning of an r-number algebra can be determined from Theorems 6, 21, and 22. The elements $a, b, \cdots$ correspond to the measurable quantities of a quantum system. The possible values of an element a arise from its resolution (13): $a = \xi_1 e_1 + \cdots + \xi_\mu e_\mu$ and are the proper values $\xi_1, \cdots, \xi_\mu$. If the value of a is measured and found to be ξ_ν then the *state* of the quantum system is described by the unit element e_ν. In the state e the expectation of an element b is $\frac{\text{trace } (eb)}{\text{trace } e}$. If $\eta_1 f_1 + \ldots + \eta_j f_j$ is the resolution of b by (13), then the probability that the measured value of b be η_i is $\frac{\text{trace } (ef_i)}{\text{trace } e}$. The case where e is unresolvable corresponds to a measurement which leads to a "reiner Fall."

Part III. Determination of all r-Number Algebras

11. We shall now seek all the solutions of the equations of condition (23), (24), and (25). These solutions will determine all r-number algebras which we shall assume throughout to be irreducible. We shall proceed stepwise in that we shall successively consider the cases $\gamma = 1, 2, 3$, and $\gamma \geqq 4$.

The analysis starts with the linear basis of A described in Theorem 14 together with the multiplication rules (22). But before we actually begin, we remark that if the resolution theorem 14 is applicable to an r-number algebra then the condition $\chi_{\rho\sigma} = \chi > 0$ for irreducibility is not only necessary (as we already know) but also sufficient. Under this assumption every invariant subalgebra of A is either 0 or A. This means that if any $x \neq 0$ of A is given, all elements of A may be obtained from it by forming suitable linear aggregates of expressions

[14] This conclusion is valid because of the previous unique determination of γ and χ.

of the form $(((xa)b)\cdots)k$, where $a, b, \cdots, k$ belong to A. For the formulas (18) and (19) show that in any case the elements $x_{\rho\sigma}$ can be so obtained. If some $x_{\rho\rho} \neq 0$ then we get an e_ρ, and if some $x_{\rho\sigma} \neq 0$ for $\rho \neq \sigma$ then we get an $x_{\rho\sigma}^2 \neq 0$. Therefore we can get an element $e_\rho + e_\sigma$ and from it we get $(e_\rho + e_\sigma)e_\rho = e_\rho$. Thus an e_ρ can be obtained in any case and from it we can get every element $e_\rho a_{\rho\tau} = \frac{1}{2} a_{\rho\tau}$ for $\rho \neq \tau$. Since we may select $a_{\rho\tau} \neq 0$ it follows that $a_{\rho\tau}^2 \neq 0$; this leads to $e_\rho + e_\tau$ and to $(e_\rho + e_\tau)e_\tau = e_\tau$. Hence we have all e_ρ and with them all $e_\rho \cdot 2s_\mu^{\rho\sigma} = s_\mu^{\rho\sigma}$ for $\rho \neq \sigma$. This completes the proof.

12. $\gamma = 1$. In this case $A = M^{11}$ and $e_1 = 1$ so that A consists merely of the elements $\lambda 1$. Since $\lambda 1 \pm \mu 1 = (\lambda \pm \mu)1$, $\lambda(\mu 1) = (\lambda\mu)1$, and $(\lambda 1)(\mu 1) = (\lambda\mu)1$, it follows that A is isomorphic with the algebra of real numbers.

13. $\gamma = 2$.[15] In the basis of Theorem 14 we replace e_1 and e_2 by $e_1 + e_2 = 1$ and $e_1 - e_2 = s_{\chi+1}$, and we write $s_1, \cdots, s_\chi$ in place of $s_1^{12}, \cdots, s_\chi^{12}$. We observe that $N = \chi + 2$. This leads to the basis $1, s_1, \cdots, s_{N-1}$ of A and the multiplication rules

$$s_\mu^2 = 1,\ s_\mu s_\nu = 0 \text{ for } \mu \neq \nu.$$

Since $\chi > 0$ it follows that $N \geqq 3$.

Conversely, A is an r-number algebra for each $N \geqq 3$. This may be seen by verifying (I) and (III). In fact, if $a = \alpha 1 + \alpha_1 s_1 + \cdots + \alpha_{N-1} s_{N-1}$, then $a^2 = (\alpha^2 + \alpha_1^2 + \cdots + \alpha_{N-1}^2)1 + 2\alpha\alpha_1 s_1 + \cdots + 2\alpha\alpha_{N-1} s_{N-1}$. If $a \neq 0$ the coefficient of 1 is greater than zero, so that if $a^2 + b^2 + c^2 + \cdots = 0$ then $a = b = c = \cdots = 0$. But a^2 has the form $\beta 1 + \gamma a$ so that from the definition of 1 and the commutativity of multiplication it follows that $(a^2 b)a = a^2(ba)$.

14. $\gamma = 3$. It is now necessary to solve the equations (23) and (24). In this case (24) assumes the form

$$\sum_{\nu=1}^{\chi} (A_{\nu\lambda\mu}^{123} A_{\nu\lambda'\mu'}^{123} + A_{\nu\lambda\mu'}^{123} A_{\nu\lambda'\mu}^{123}) = \tfrac{1}{2}\, \delta_{\lambda\lambda'}\, \delta_{\mu\mu'}.$$

If we define χ matrices $B_1, \cdots, B_\chi$ by the equations

$$B_\lambda = \|B_{\lambda,\mu\nu}\|,\ \mu, \nu = 1, \ldots, \chi; \qquad B_{\lambda,\mu\nu} = 2A_{\nu\lambda\mu}^{123}, \tag{35}$$

then[16]

$$B_\lambda^* B_{\lambda'} + B_{\lambda'}^* B_\lambda = 2\delta_{\lambda\lambda'}\, 1. \tag{36}$$

[15] This case was first discussed by M. Zorn to whom the authors are indebted also for valuable assistance.

[16] We denote the transposed matrix of B by B^*: $B = \|b_{\alpha\beta}\|$ and $B^* = \|b_{\beta\alpha}\|$. 1 is the unit matrix $1 = \|\delta_{\alpha\beta}\|$.

One can effect changes in the $A^{123}_{\nu\lambda\mu}$ (that is, the B_μ) by altering the bases of M^{22}, M^{13}, and M^{23}. These latter changes are unessential and, by Theorem 13 and (21), the transformations are

$$(37)\qquad (s^{12}_\mu)' = \sum_{\nu=1}^{\chi} \omega_{\mu\nu} s^{12}_\nu, \qquad (s^{13}_\mu)' = \sum_{\nu=1}^{\chi} \varphi_{\mu\nu} s^{13}_\nu, \qquad (s^{23}_\mu)' = \sum_{\nu=1}^{\chi} \psi_{\mu\nu} s^{23}_\nu,$$

where the matrices $\|\omega_{\mu\nu}\|$, $\|\varphi_{\mu\nu}\|$, and $\|\psi_{\mu\nu}\|$ are orthogonal. By (22*d*) the elements $A^{123}_{\nu\lambda\mu}$ transform according to the relation

$$(38)\qquad (A^{123}_{\nu\lambda\mu})' = \sum_{\nu',\lambda',\mu'=1}^{\chi} \omega_{\mu\mu'}\, \varphi_{\lambda\lambda'}\, \psi_{\nu\nu'}\, A^{123}_{\nu'\lambda'\mu'},$$

or, in terms of the B_μ,

$$(39)\qquad B'_\lambda = \sum_{\lambda'=1}^{\chi} \varphi_{\lambda\lambda'}\, U^* B_{\lambda'} W,$$

where $U = \|\omega_{\mu\nu}\|$, $V = \|\varphi_{\mu\nu}\|$, and $W = \|\psi_{\mu\nu}\|$ are unitary matrices. One sees immediately that for every solution of (36) there is another resulting from (39). We shall regard such solutions as being not essentially distinct.

The solution of (36) was obtained in a well-known paper[17] by A. Hurwitz who showed in particular that χ must be 1, 2, 4, or 8. Since a part of Hurwitz's investigation (unimportant for his main purpose but important for our discussion in paragraph 15) contains an essential error,[18] it is convenient to carry out the entire argument independently.

If in (36) $\lambda = \lambda'$, then $B^*_\lambda B_\lambda = 1$ so that all B_λ are orthogonal. If we set $C_\lambda = B^*_\chi B_\lambda$, $\lambda = 1, \cdots, \chi - 1$, then the C_λ are also all orthogonal. If $\lambda \neq \lambda'$ then (36) may be written in the form $B^*_\lambda B_{\lambda'} = -B^*_{\lambda'} B_\lambda$ or $B^*_{\lambda'} B_\lambda\, B^*_{\lambda'} B_\lambda = -1$. Hence if $\lambda' = \chi$ ($\lambda \neq \lambda' = \chi$) then $C^2_\lambda = -1$ and if $\lambda, \lambda' \neq \chi$ ($\lambda \neq \lambda'$) then $(C^*_{\lambda'} C_\lambda)^2 = -1$ since $C^*_{\lambda'} C_\lambda = B^*_{\lambda'} B_\lambda$. Or again, $C^*_{\lambda'} \cdot C_\lambda C^*_{\lambda'} \cdot C_\lambda = -1$ while $C^*_{\lambda'} \cdot C^*_{\lambda'} C_\lambda \cdot C_\lambda = C^{*2}_{\lambda'} \cdot C^2_\lambda = 1$ so that $C_\lambda C^*_{\lambda'} = -C^*_{\lambda'} C_\lambda$, or (after multiplication on both sides by $C_{\lambda'}$) $C_{\lambda'} C_\lambda = -C_\lambda C_{\lambda'}$. Hence the C_λ's are orthogonal,

$$(40)\qquad C^2_\lambda = -1,\ C_\lambda C_{\lambda'} = -C_{\lambda'} C_\lambda \text{ for } \lambda \neq \lambda',$$

where $\lambda, \lambda' = 1, \cdots, \chi - 1$.

If B is any orthogonal matrix, then $B_1 = BC_1, \cdots, B_{\chi-1} = BC_{\chi-1}$,

[17] "Über die Composition der quadratischen Formen von beliebig vielen Variablen," Goett. Nachr., 1898, p. 309.

[18] Namely, the assertion that for a given χ all solutions B'_λ may be obtained from one particular B_λ by means of the relation $B'_\lambda = U^* B_\lambda W$ (that is, without the use of φ in (39)), this relation assuming the form $C'_\lambda = W^* C_\lambda W$ in the case where $C_\lambda = B^*_\chi B_\lambda$. (Cf. below.) One sees that this is incorrect in the following way: in case χ is equal to 4 or 8, then $C_1 \cdot \cdots \cdot C_{\chi-1} = \pm 1$. This equation is invariant under the above transformation. Yet the left side has the opposite sign in the solution $B'_1 = -B$, $B'_2 = B_2, \cdots, B_\chi = B_\chi$. (Cf. footnote 17, pp. 313–315; our B, C, U, W are there denoted by A, B, P, Q.)

$B_\chi = B$ is the general solution of (36) in case $C_1, \cdots, C_{\chi-1}$ is the general solution of (40).

15. Let G_χ be that group whose modulus may be denoted by $\bar{1}$ and which possesses χ generators $-\bar{1}, \bar{\Gamma}_1, \cdots, \bar{\Gamma}_{\chi-1}$ which satisfy the following relations:

$$(41) \qquad \begin{cases} (-\bar{1})^2 = \bar{1}, & (-\bar{1})\bar{\Gamma}_\lambda = \bar{\Gamma}_\lambda(-\bar{1}), \\ \bar{\Gamma}_\lambda^2 = -1, & \bar{\Gamma}_\lambda\bar{\Gamma}_{\lambda'} = (-\bar{1})\bar{\Gamma}_{\lambda'}\bar{\Gamma}_\lambda \text{ for } \lambda \neq \lambda'. \end{cases}$$

It is obvious that all the elements of this group may be written as $\bar{\Gamma}_{\lambda_1} \cdots \bar{\Gamma}_{\lambda_r}$ or as $(-\bar{1})\bar{\Gamma}_{\lambda_1} \cdots \bar{\Gamma}_{\lambda_r}$, $r = 0, \cdots, \chi - 1$, $1 \leqq \lambda_1 < \cdots < \lambda_r \leqq \chi - 1$. Hence G_χ is finite and has 2^χ elements. The problem of solving (40) reduces to that of finding[19] all representations of G_χ by χ-dimensional orthogonal matrices in which $-\bar{1}$ is represented by the matrix -1. $\bar{\Gamma}_\lambda$ is then represented by C_λ, $\lambda = 1, \cdots, \chi - 1$.

We next determine the "irreducible" representations of G_χ (cf. footnote 19 for all concepts and theorems in the theory of representations), and, of course, merely the complex-unitary ones. Since no C_λ exist for $\chi = 1$ (the solution is simply that B_1 must be orthogonal) we assume that $\chi \geqq 2$.

The order of G_χ is 2^χ. The commutators of G_χ are $\bar{1}$ and $-\bar{1}$ and they form an invariant subgroup of order 2 so that the index of G_χ is $2^{\chi-1}$. An element of the group can be transformed only into itself and its product with -1, the latter being impossible for $\bar{1}$ and $-\bar{1}$ as well as for $\bar{\Gamma}_1 \cdots \bar{\Gamma}_{\chi-1}$, $(-\bar{1})\bar{\Gamma}_1 \cdots \bar{\Gamma}_{\chi-1}$ if χ is even. Hence 2 or 4 equivalence classes each contain one element according as χ is odd or even, while all others contain two elements. Their number is therefore $2^{\chi-1} + 1$ or $2^{\chi-1} + 2$. (The reader may complete the simple discussion of G_χ.) Hence G_χ has $2^{\chi-1} + 1$ or $2^{\chi-1} + 2$ inequivalent irreducible representations of which $2^{\chi-1}$ are of order 1. These are the commutative representations. But since every commutator is $\bar{1}$ or $-\bar{1}$, and is therefore representable by the matrices 1 or -1, these commutative representations are exactly those in which $-\bar{1}$ is represented by 1. The representations which are of interest to us are therefore those of order $\neq 1$. From the above discussion it follows that there are one or two such representations.

The orders of all irreducible representations are divisors of 2^χ and the sum of the squares of these orders is 2^χ, of which the representations of order 1 contribute $2^{\chi-1} \cdot 1^2 = 2^{\chi-1}$. The sum of the squares of the orders of the remaining representations is $2^\chi - 2^{\chi-1} = 2^{\chi-1}$. If χ is odd there is only one such representation and its order is $2^{\frac{\chi-1}{2}}$. If χ is even there are two and their orders are divisors of 2^χ. Hence their orders are powers of 2 and the sum of the squares of their orders can be $2^{\chi-1}$ only if each of them is $2^{\frac{\chi-2}{2}}$.

[19] The G. Frobenius—I. Schur "theory of representations," which is prominent in what follows, may be found in I. Schur's treatise "Neue Begrundung der Theorie der Gruppencharaktere," Berliner Ber., 1905, p. 406.

The (real) representation we are seeking must be made up of these irreducible (complex) representations and its order χ must be some multiple of the common degrees, i.e., $2^{\frac{\chi-1}{2}}$ or $2^{\frac{\chi-2}{2}}$. It follows immediately that $\chi = 2, 4$, or 8. All these values are even, $\chi/2^{\frac{\chi-2}{2}} = 2, 2$, or 1, and our representation consists of 2, 2, or 1 irreducible parts. But since we wish to form a real representation from complex representations, we must keep the reality conditions in mind.[20]

A complex irreducible representation can have in this regard one of three properties: 1) it may be equivalent to a real representation, 2) it may be equivalent to no real representation but instead be equivalent to its own complex-conjugate representation, or 3) it may be inequivalent to its complex-conjugate representation. In order to distinguish these cases one must form the sum of the traces (characteristics) of the matrices representing the squares of all the elements of the group. According as this sum is greater than, equal to, or less than 0 the situation is described by 1), 3), or 2).

In G_χ the square of every element is $\bar{1}$ or $-\bar{1}$. Suppose α of the elements have $\bar{1}$ for their square and suppose β of the elements have $-\bar{1}$ for their square. Then $\alpha + \beta = 2^\chi$. The matrix representations of these squared elements are 1 or -1 and their trace is $2^{\frac{\chi-2}{2}}$ or $-2^{\frac{\chi-2}{2}}$. The sum in question is therefore $2^{\frac{\chi-2}{2}}(\alpha - \beta)$ and our criterion reduces merely to the questions whether α is greater than, equal to, or less than β. Enumeration shows that the relations $=$, $<$, and $>$ hold for the values 2, 4, and 8 respectively of χ. Hence the cases 3), 2), and 1) exist. For $\chi = 2$ each of our two irreducible representations is inequivalent to its own complex-conjugate. Hence each is equivalent to the other so that we may describe the two irreducible representations as complex-conjugate to one another. The real representation we are seeking must therefore contain each of them exactly once. For $\chi = 4$ each irreducible representation is in case 2) and the real representation sought must contain each of them an even number of times. Hence it contains one of them twice but does not contain the other. For $\chi = 8$ each irreducible representation can be described as real so that both may be used as solutions. We may summarize our discussion by saying that for $\chi = 2$ there is one solution, and for $\chi = 4$ or 8 there are two inequivalent solutions.

Since $C_1 \cdots C_{\chi-1}$ is interchangeable with all C_λ (since χ is even), it must be of the form $\xi 1$ in each irreducible representation, and since its square is 1, it must be ± 1. Since the solution for $\chi = 4$ consists of two equivalent irreducible matrices and since the solution for $\chi = 8$ is irreducible, we have for $\chi = 4$ or 8 that

$$C_1 \cdots C_{\chi-1} = \pm 1. \tag{42}$$

[20] Cf. G. Frobenius and I. Schur, Berliner Ber., 1906, p. 186.

Since an equivalence transformation does not alter this relation, while the transformation $C_1' = -C_1$, $C_2' = C_2$, $\cdots$, $C_{\chi-1}' = C_{\chi-1}$ interchanges the $+$ and $-$ signs in it, it follows that this transformation carries one solution over into the other. (Cf. footnote 18.)

If, for $\chi = 2$, we mean by s_λ simply $s_\lambda = 1$, and for $\chi = 4$ or 8 we mean the same or $s_1 = -1$, $s_2 = \cdots = s_\chi = 1$, then the most general solution C_λ' arises from a solution C_λ by the transformation $C_\lambda' = s_\lambda W^* C_\lambda W$, where W is orthogonal. Hence $B_\lambda' = s_\lambda B W^* C_\lambda W$ (the relation also holds for $\lambda = \chi$ if we set $C_\chi = 1$: $B_\chi' = s_\chi B W^* W = B$). With the help of (39) it follows that $U = WB^*$ and $\varphi_{\mu\nu} = s_\mu \delta_{\mu\nu}$.

To sum up: Solutions exist for $\chi = 1, 2, 4$, and 8, but for each of these values there is essentially only one solution in the sense of (37), (38), and (39). In particular we can set U and V equal to 1 for $\chi = 1$, we can set $V = 1$ for $\chi = 2$, and we can set $V = 1$ or $V = \|\bar{s}_\mu \delta_{\mu\nu}\|$ ($\bar{s}_1 = -1$, $\bar{s}_2 = \cdots = \bar{s}_\chi = 1$) for $\chi = 4$ or 8.

16. Having shown that there is only one solution, we may better find it in another way. (36) may also be reformulated (and this was the original formulation by A. Hurwitz). In fact, (36) is equivalent to

$$(43)\qquad \begin{cases} (x_1^2 + \ldots + x_\chi^2)(y_1^2 + \ldots + y_\chi^2) = z_1^2 + \ldots + z_\chi^2, \\ z_\lambda = \sum_{\mu,\nu=1}^{\chi} B_{\lambda,\mu\nu} x_\mu y_\nu \qquad (\lambda = 1, \ldots, \chi) \end{cases}$$

($x_1, \cdots, x_\chi$ and $y_1, \cdots, y_\chi$ being two sets of independent variables). The solution may be found by the well-known substitution

$$(44)\qquad \begin{cases} X = x_1 1 + x_2 i_2 + \cdots + x_\chi i_\chi,\ Y = y_1 1 + y_2 i_2 + \cdots + y_\chi i_\chi, \\ Z = z_1 1 + z_2 i_2 + \cdots z_\chi i_\chi, \\ Xy = Z \end{cases}$$

in case $i_1 = 1, i_2, \cdots, i_\chi$ constitute the basis of the following hypercomplex number systems: for $\chi = 1$, the real numbers; for $\chi = 2$, the ordinary complex numbers; for $\chi = 4$, the quaternions;[21] for $\chi = 8$, the so-called Cayley numbers or quasi-quaternions.[22]

---

[21] Quaternions satisfy the relations $i_2^2 = i_3^2 = i_4^2 = -1$; $i_2 i_3 = -i_3 i_2 = i_4$; $i_3 i_4 - i_4 i_3 = i_2$; $i_4 i_2 = -i_2 i_4 = i_3$.

[22] These were introduced by Cayley, Corresp. Math. Phys., vol. 1, 1843, p. 452. (Published by P. H. Fuss). A summary, and at the same time historically exhaustive, description of them has been given by L. E. Dickson, Annals. of Math., vol. 20, 1918, p. 155 and particularly pp. 157–159. A detailed discussion of their algebraic properties has been given by M. Zorn, cf. footnote 13. Quasi-quaternions are defined by $i_2^2 = \ldots = i_8^2 = -1$ and each of the systems i_2, i_3, i_5; i_3, i_4, i_6; i_4, i_5, i_7; i_5, i_6, i_8; i_6, i_7, i_2; i_7, i_8, i_3; i_8, i_2, i_4 (together with $i_1 = 1$) has the formal properties of quaternions. They form a non-associative system of quantities.

If we denote the complex-conjugate of such a hypercomplex number $X = x_1 1 + x_2 i_2 + \cdots + x_\chi i_\chi$ by $\bar{X} = x_1 1 - x_2 i_2 - \cdots - x_\chi i_\chi$ and its real part by $R(X) = x_1$, then $R(X\bar{\imath}_\mu) = x_\mu$, $\mu = 1, \cdots, \chi$. Consequently, if we note that $R(XY) = R(YX)$,

$$2A_{\nu\lambda\mu}^{123} = B_{\lambda,\mu\nu} = R\,(i_\mu i_\nu \bar{\imath}_\lambda) = R\,(i_\nu \bar{\imath}_\lambda i_\mu). \tag{45}$$

By (45) it is possible to calculate these $A_{\nu\lambda\mu}^{123}$ with the help of the multiplication tables of the above hypercomplex number systems, but we shall not go into this. The general solution $(A_{\nu\lambda\mu}^{123})'$ arises from (45) by (38) where we can keep the restriction mentioned at the end of paragraph 15.

It follows from the rules (22) together with (43) or (45) that the algebra A is isomorphic with a system of matrices $A, B, C, \cdots$ of order 3, wherein the operations $a \pm b$, λa, and ab correspond to the matrix operations $A \pm B$, λA, and $\dfrac{A \times B + B \times A}{2}$ ($A \times B$ is the usual matrix multiplication). Again, e_ρ corresponds to that matrix which has the $\rho\rho$-element 1 but otherwise consists of zeros, and $s_\mu^{\rho\sigma}$ $(\rho < \sigma)$ corresponds to that matrix which has i_μ for its $\rho\sigma$-element and $\bar{\imath}_\mu$ for its $\sigma\rho$-element but otherwise consists of zeros. Consequently A is the algebra of all third order Hermitian matrices, $\|X_{\rho\sigma}\|$, where $\rho, \sigma = 1, 2, 3$ and $X_{\rho\sigma} = \bar{X}_{\sigma\rho}$ (the diagonal elements $X_{\rho\rho}$ being real numbers), with the above addition and multiplication rules. Hence the matrix elements for $\chi = 1, 2, 4$, or 8 respectively are real numbers, complex numbers, quaternions, and quasi-quaternions and $N = \gamma + \dfrac{\gamma(\gamma - 1)}{2}\chi = 3 + 3\chi = 6, 9, 15, 27$.

17. Conversely, if A is one of the above algebras, we shall ascertain whether or not it is an r-number algebra. Since we can form the system of Hermitian matrices described in the last paragraph for every order (and not merely of order three as was done there), and since we shall need the more general matrices later on, we generalize the problem as follows: A consists of all Hermitian matrices of order γ subject to the above addition and multiplication rules and having real numbers, complex numbers, quaternions, and quasi-quaternions respectively as elements. We provisionally assume merely that $\gamma \geqq 3$ (since $\gamma = 1$ or 2 leads merely to special cases of the results obtained in paragraphs 12 and 13).

It is obvious that it is necessary to verify only (I) and either (II) or (III). (I) follows from the fact that if $A = \|X_{\rho\sigma}\|$, then $\operatorname{trace}(A^2) = \sum_{\rho,\sigma=1}^{\gamma} |X_{\rho\sigma}|^2$ and hence is real and, for $A \neq 0$, positive. (II) is evident for the first three cases ($\chi = 1, 2$, and 4) since the matrix elements belong to associative number systems so that $A^2, A^3, \cdots$ result from our multiplication AB just as with the usual matrix multiplication $A \times B$ (cf. footnote 1). It remains to consider only the fourth case ($\chi = 8$). We shall soon see that there is no solution for $\gamma \geqq 4$ so

that only the case $\gamma = 3$ is in question. Instead of proving (III) we could prove its generalization (4) (which leads to (III) if one replaces a, b, c, u by a, a, a, b) which has the advantage of being linear in all four variables a, b, c, u. It suffices for this purpose to substitute all combinations of the 27 elements of the linear basis. This would lead actually to $27^4 = 531{,}441$ combinations on the basis of the multiplication rules and footnote 22. Yet this number may be reduced by a few devices so that the investigation is possible without too many possibilities. It results that (4) holds in all cases. We omit the proof which may be found in a subsequent paper of A. A. Albert in this journal. Thus for $\gamma = 3$ the fourth case also leads to a solution.

18. $\gamma \geqq 4$. Since $N_1(e_1 + e_2 + e_3) = \sum_{\rho, \sigma = 1}^{3} M^{\rho\sigma}$ is a subalgebra of A which has the same χ as A but for which $\gamma = 3$, we again have $\chi = 1, 2, 4$, or 8. We shall examine these cases successively.

$\gamma \geqq 4$, $\chi = 8$. If in (25) we set $\rho, \sigma, \tau, \upsilon$ respectively equal to 1, 2, 3, and 4, and introduce the following matrices (as in (35)):

$$\tilde{B}_\lambda = \|\tilde{B}_{\lambda,\mu\nu}\|, \tilde{B}_{\lambda,\mu\nu} = 2A^{123}_{\mu\nu\lambda}; \tilde{C}_\lambda = \|\tilde{C}_{\lambda,\mu\nu}\|, \tilde{C}_{\lambda,\mu\nu} = 2A^{124}_{\nu\mu\lambda};$$

$$\tilde{D}_\lambda = \|\tilde{D}_{\lambda,\mu\nu}\|, \tilde{D}_{\lambda,\mu\nu} = 2A^{134}_{\lambda\mu\nu}; \tilde{E}_\lambda = \|\tilde{E}_{\lambda,\mu\nu}\|, \tilde{E}_{\lambda,\mu\nu} = 2A^{234}_{\lambda\nu\mu}.$$

then it follows from (25) that

$$\tilde{C}_\iota \tilde{E}^*_\kappa = \tilde{D}_\kappa \tilde{B}^*_\iota, \tag{46}$$

and all of these matrices are orthogonal. (The orthogonality follows from (24), with either $\lambda = \lambda'$ or $\mu = \mu'$, together with (23).) If now we replace ι by ι' in (46) and form the transposed matrices, we have

$$\tilde{E}_\kappa \tilde{C}^*_{\iota'} = \tilde{B}_{\iota'} \tilde{D}^*_\kappa.$$

Multiplication of (46) on the right by this last equation gives

$$\tilde{C}_\iota \tilde{C}^*_{\iota'} = \tilde{D}_\kappa \tilde{B}^*_\iota \tilde{B}_{\iota'} \tilde{D}^*_\kappa.$$

The left side of this equation is independent of κ so that

$$\tilde{D}_\kappa \tilde{B}^*_\iota \tilde{B}_{\iota'} \tilde{D}^*_\kappa = \tilde{D}_{\kappa'} \tilde{B}^*_\iota \tilde{B}_{\iota'} \tilde{D}^*_{\kappa'}, \qquad \tilde{D}^*_{\kappa'} \tilde{D}_\kappa \cdot \tilde{B}^*_\iota \tilde{B}_{\iota'} = \tilde{B}^*_\iota \tilde{B}_{\iota'} \cdot \tilde{D}^*_{\kappa'} \tilde{D}_\kappa.$$

Hence every $\tilde{B}^*_\iota \tilde{B}_{\iota'}$ commutes with every $\tilde{D}^*_{\kappa'} \tilde{D}_\kappa$.

We now have, with regard to (23),

$$\tilde{D}_{\lambda,\mu\nu} = 2A^{134}_{\lambda\mu\nu} = 2A^{413}_{\nu\lambda\mu}.$$

Hence in the subalgebra $N_1(e_1 + e_3 + e_4) = \sum_{\rho, \sigma = 1, 3, 4} M^{\rho\sigma}$ of A the matrix $\tilde{D}_\lambda = \|\tilde{D}_{\lambda,\mu\nu}\|$ plays the same rôle as did the matrix $B_\lambda = \|B_{\lambda,\mu\nu}\|$ in paragraphs 14 to 17 in the r-number algebra with $1 = e_1 + e_2 + e_3$. (The indices 1, 2, 3 are

replaced by 4, 1, 3.) The matrices $\tilde{D}_\chi^* \tilde{D}_\kappa$, $\kappa = 1, \cdots, \chi - 1$, therefore play the rôle of the C_κ and therefore form, for our $\chi = 8$, an irreducible system (cf. paragraph 15). Since every $\tilde{B}_\iota^* \tilde{B}_{\iota'}$ is interchangeable with every $\tilde{D}_\chi^* \tilde{D}_\kappa$, they all have the form $\xi_{\iota\iota'} 1$. But with regard to (23),

$$\tilde{B}_{\lambda,\mu\nu} = 2A^{123}_{\mu\nu\lambda} = 2A^{231}_{\nu\lambda\mu}.$$

Hence in the subalgebra $N_1(e_1 + e_2 + e_3) = \sum_{\rho,\sigma=1}^{3} M^{\rho\sigma}$ of A the matrix $\tilde{B}_\lambda = \|\tilde{B}_{\lambda,\mu\nu}\|$ plays the same rôle as did the matrix $B_\lambda = \|B_{\lambda,\mu\nu}\|$ in paragraphs 14 to 17 in the r-number algebra with $1 = e_1 + e_2 + e_3$. (The indices 1, 2, 3 are replaced by 2, 3, 1.) The matrices $\tilde{B}_\chi^* \tilde{B}_\kappa$, $\kappa = 1, \cdots, \chi - 1$, play the rôle of the C_κ and form, for our $\chi = 8$, an irreducible system. Therefore the form $\xi_{\chi\kappa} 1$ with the order $\chi = 8 > 1$ is impossible.

Thus in the above case there exists no solution.

19. $\gamma \geqq 4$, $\chi = 4$. In (45) we obtained a solution for $2A^{123}_{\nu\lambda\mu}$ in the form $R(i_\nu \bar{\imath}_\lambda\, i_\mu)$, where $i_1 = 1, i_2, i_3, i_4$ compose the basis of the system of quaternions. (Cf. footnote 21.) By (38) the most general solution results from it if we replace

$$i_\nu \text{ by } \sum_{\nu'=1}^{4} \psi_{\nu\nu'}\, i_{\nu'}, \qquad i_\mu \text{ by } \sum_{\mu'=1}^{4} \varphi_{\mu\mu'}\, i_{\mu'}, \qquad i_\lambda \text{ by } \sum_{\lambda'=1}^{4} \omega_{\lambda\lambda'}\, i_{\lambda'}$$

(we make no use of the restriction proved at the end of paragraph 15). But every orthogonal transformation $\sum_{\kappa'=1}^{4} \sigma_{\kappa\kappa'}\, i_{\kappa'}$ of the basis of the quaternion system may also be represented in the form $X' i_\kappa{}^\bullet X''$, where X' and X'' represent certain quaternions of absolute value unity and $\bullet$ denotes nothing at all or else the sign for the complex conjugate.[23] The most general expression for $2A^{123}_{\nu\lambda\mu}$ is therefore (since $R(XY) = R(YX)$)

$$2A^{123}_{\nu\lambda\mu} = R(Y^{(123)}\, i_\nu^{(3\overset{1}{\bullet}2)}\, Y^{(231)}\, i_\lambda^{(1\overset{2}{\bullet}3)}\, Y^{(312)}\, i_\mu^{(2\overset{3}{\bullet}1)})$$

where the Y's are certain quaternions of absolute value unity and where the $\bullet$'s are independent of one another. Since $R(XY) = R(YX)$ it follows by (23) that this equation also holds for $2A^{231}_{\nu\lambda\mu}$ and $2A^{312}_{\nu\lambda\mu}$ if the indices 1, 2, and 3 are correspondingly permuted on the right side. Hence we may set

$$2A^{\rho\sigma\tau}_{\nu\lambda\mu} = R(Y^{(\rho\sigma\tau)}\, i_\nu^{(\tau\overset{\rho}{\bullet}\sigma)}\, Y^{(\sigma\tau\rho)}\, i_\lambda^{(\rho\overset{\sigma}{\bullet}\tau)}\, Y^{(\tau\rho\sigma)}\, i_\mu^{(\sigma\overset{\tau}{\bullet}\rho)}) \tag{47}$$

where the Y's are certain quaternions of absolute value unity and where the $\bullet$'s are independent of one another. Consequently we shall limit ourselves to the

[23] We will write $\bullet = 0$ if $\bullet$ denotes nothing, and $\bullet = -$ if it denotes the complex conjugate. The representation of rotations by means of quaternions can be verified easily.

consideration of such values of ρ, σ, and τ which may be transformed by a cyclic permutation into a monotonically increasing sequence, and we shall call them allowable sequences. (We have treated the permutations of 1, 2, and 3 in the same manner. Since ρ, σ, τ or τ, σ, ρ is always allowed in this sense, it follows that the allowable $A^{\rho\sigma\tau}_{\lambda\mu\nu}$ are sufficient for the purpose of the multiplication rules in (22d).)

On the other hand the generality of (47) can be somewhat lessened in that we can subject every $M^{\rho\sigma}$ to an arbitrary transformation in the sense of Theorem 13, (21), and (37). This is exhibited in the formulas (47) for $2A^{\rho\sigma\tau}_{\nu\lambda\mu}$, $2A^{\sigma\tau\rho}_{\nu\lambda\mu}$, $2A^{\tau\rho\sigma}_{\nu\lambda\mu}$, or $2A^{\sigma\rho\tau}_{\nu\lambda\mu}$, $2A^{\rho\tau\sigma}_{\nu\lambda\mu}$, $2A^{\tau\sigma\rho}_{\nu\lambda\mu}$ as a transformation of the corresponding i_μ, i_λ, i_ν and hence as that of the i_κ with ${}^{(\sigma\tau\rho)}_{\bullet}$ or ${}^{(\rho\tau\sigma)}_{\bullet}$. If we transform all the $M^{\rho\sigma}$ simultaneously so that in $M^{\rho\sigma}$ i_κ goes over into $X_1^{(\rho\sigma)}\, i_{\kappa\bullet}^{(\rho\sigma)}\, X_2^{(\rho\sigma)}$, and if for the sake of simplicity we identify $\rho\sigma$ and $\sigma\rho$, then we have (since $\overline{XY} = \overline{Y}\,\overline{X}$)

$$(48)\quad \begin{cases} \text{(a)} & {}^{(\rho\sigma\tau)}_{\bullet} \text{ is replaced by } {}^{(\rho\sigma\tau)}_{\bullet}\,{}^{(\rho\tau)}_{\bullet}, \\ \text{(b)} & Y^{(\rho\sigma\tau)} \text{ is replaced by } X_{\mathrm{I}}^{(\rho\sigma)}\, Y^{(\rho\sigma\tau)}\, X_{\mathrm{II}}^{(\sigma\tau)}, \text{ where} \\ & X_{\mathrm{I}}^{(\rho\sigma)} = X_2^{(\rho\sigma)} \text{ or } \bar{X}_{\mathrm{I}}^{(\rho\sigma)} \text{ when } {}^{(\sigma\tau\rho)}_{\bullet} = 0 \text{ or } \overline{}, \\ & X_{\mathrm{II}}^{(\sigma\tau)} = X_{\mathrm{I}}^{(\sigma\tau)} \text{ or } \bar{X}_2^{(\sigma\tau)} \text{ when } {}^{(\tau\rho\sigma)}_{\bullet} = 0 \text{ or } \overline{}. \end{cases}$$

We now consider (25). (Since ρ, σ, τ; ρ, σ, v; ρ, τ, v; and σ, τ, v are to be allowable, it must be the case that every ρ, σ, τ, v goes over into a monotonically increasing sequence by a cyclic permutation and hence must be allowable.) If we substitute

$$R(XY) = R(YX), \sum_{\kappa=1}^{4} R(X\, i_\kappa^{(1)}_{\bullet})\, R(Y\, i_\kappa^{(2)}_{\bullet}) = R(X\bar{Y}^{(1)}_{\bullet}{}^{(2)}_{\bullet})$$

in (47), we get

$$\begin{aligned} &R([Y^{(\sigma v\rho)}\, i_\theta^{(\rho\sigma v)}_{\bullet}\, Y^{(v\rho\sigma)}\, i_\iota^{(\sigma v\rho)}_{\bullet}\, Y^{(\rho\sigma v)}] \cdot [Y^{(v\sigma\tau)}\, i_\lambda^{(\tau v\sigma)}_{\bullet}\, Y^{(\sigma\tau v)}\, i_\kappa^{(v\sigma\tau)}_{\bullet}\, Y^{(\tau v\sigma)}]^{\overline{(v\rho\sigma)}\,\overline{(\sigma\tau v)}}_{\bullet}) \\ &= R([Y^{(\rho\tau v)}\, i_\kappa^{(v\rho\tau)}_{\bullet}\, Y^{(\tau v\rho)}\, i_\theta^{(\rho\tau v)}_{\bullet}\, Y^{(v\rho\tau)}] \cdot [Y^{(\tau\rho\sigma)}\, i_\iota^{(\sigma\tau\rho)}_{\bullet}\, Y^{(\rho\sigma\tau)}\, i_\lambda^{(\tau\rho\sigma)}_{\bullet}\, Y^{(\sigma\tau\rho)}]^{(\tau v\rho)}_{\bullet}\,{}^{(\rho\sigma\tau)}_{\bullet}). \end{aligned}$$

Since we are forming linear aggregates we can set a free quaternion variable Z_θ in place of i_θ and likewise free quaternion variables Z_ι, Z_λ, and Z_κ in place of i_ι, i_λ, i_κ. Then we can multiply out the parentheses on both sides of the equation using the law $\overline{(XY)} = \overline{Y}\,\overline{X}$. It can accordingly be seen that, because of the non-commutativity of quaternion multiplication, equality can exist only if variables which are in juxtaposition on one side are also in juxtaposition on the other side. Since Z_θ and Z_ι, as well as Z_λ and Z_κ, are in every case adjacent on the left side of the equation, they are again adjacent on the right side. Therefore $\overline{{}^{(\tau v\rho)}_{\bullet}\,{}^{(\rho\sigma\tau)}_{\bullet}}$ can not be equal to $\overline{}$, and hence ${}^{(\tau v\rho)}_{\bullet} \neq {}^{(\rho\sigma\tau)}_{\bullet}$. Similarly the invariable juxtaposition of Z_κ and Z_θ on the right side of the equation, as well as of Z_ι and Z_λ,

leads to the condition ${}^{(\upsilon\rho\sigma)}_{\bullet} \neq {}^{(\sigma\tau\upsilon)}_{\bullet}$. Now our equation assumes the form (since $R(XY) = R(YX)$)

$$R(Y^{(\tau\upsilon\sigma)}\, Y^{(\sigma\upsilon\rho)}\, i_\theta^{\,(\rho\sigma\upsilon)}_{\bullet}\, Y^{(\upsilon\rho\sigma)}\, i_\iota^{\,(\sigma\upsilon\rho)}_{\bullet}\, Y^{(\rho\sigma\upsilon)}\, Y^{(\upsilon\sigma\tau)}\, i_\lambda^{\,(\tau\upsilon\sigma)}_{\bullet}\, Y^{(\sigma\tau\upsilon)}\, i_\kappa^{\,(\upsilon\sigma\tau)}_{\bullet})$$
$$= R(Y^{(\tau\upsilon\rho)}\, i_\theta^{\,(\rho\tau\upsilon)}_{\bullet}\, Y^{(\upsilon\rho\tau)}\, Y^{(\tau\rho\sigma)}\, i_\iota^{\,(\sigma\tau\rho}_{\bullet}\, Y^{(\rho\sigma\tau)}\, i_\lambda^{\,(\tau\rho\sigma)}_{\bullet}\, Y^{(\sigma\tau\rho)}\, Y^{(\rho\tau\upsilon)}\, i_\kappa^{\,(\upsilon\rho\tau)}_{\bullet}).$$

Again we can replace $i_\theta, i_\iota, i_\lambda, i_\kappa$ by $Z_\theta, Z_\iota, Z_\lambda, Z_\kappa$ and it results that

$${}^{(\rho\sigma\upsilon)}_{\bullet} = {}^{(\rho\tau\upsilon)}_{\bullet},\ {}^{(\sigma\upsilon\rho)}_{\bullet} = {}^{(\sigma\tau\rho)}_{\bullet},\ {}^{(\tau\upsilon\sigma)}_{\bullet} = {}^{(\tau\rho\sigma)}_{\bullet},\ {}^{(\upsilon\sigma\tau)}_{\bullet} = {}^{(\upsilon\rho\tau)}_{\bullet}$$

in addition to the conditions

$${}^{(\tau\upsilon\rho)}_{\bullet} \neq {}^{(\rho\sigma\tau)}_{\bullet},\ {}^{(\upsilon\rho\sigma)}_{\bullet} \neq {}^{(\sigma\tau\upsilon)}_{\bullet}$$

already obtained, so that

$$\text{(Y)} \qquad \begin{cases} Y^{(\tau\upsilon\sigma)}Y^{(\sigma\upsilon\rho)} = Y^{(\tau\upsilon\rho)}, & Y^{(\upsilon\rho\sigma)} = Y^{(\upsilon\rho\tau)}Y^{(\tau\rho\sigma)}, \\ Y^{(\rho\sigma\upsilon)}Y^{(\upsilon\sigma\tau)} = Y^{(\rho\sigma\tau)}, & Y^{(\sigma\tau\upsilon)} = Y^{(\sigma\tau\rho)}Y^{(\rho\tau\upsilon)}. \end{cases}$$

It is to be noted that we have required of the sequences ρ, σ, τ; ρ, σ, υ; ρ, τ, υ; and σ, τ, υ that each of them be transformable into a monotonically increasing sequence by a cyclic permutation, and of course the same must hold also of $\rho, \sigma, \tau, \upsilon$.

We now see that the four $\bullet$–equations simply comprise all possible subcases of ${}^{\alpha\epsilon\beta}_{\bullet} = {}^{\alpha\eta\beta}_{\bullet}$ for which α, ϵ, β and α, η, β are allowable sequences. The two next following $\bullet$-inequalities express together that ${}^{\alpha\epsilon\beta}_{\bullet} \neq {}^{\beta\eta\alpha}_{\bullet}$, where α, ϵ, β and β, η, α are arbitrary allowable sequences. By the operation of (48 *a*) we can, for every pair $\alpha > \beta$, make a ${}^{\alpha\epsilon\beta}_{\bullet}$ vanish or make a ${}^{\beta\eta\alpha}_{\bullet}$ equal to $\overline{}$.[24] From our $\bullet$-equations and $\bullet$-inequalities it follows that in any case every ${}^{\alpha\epsilon\beta}_{\bullet}$ will vanish and every ${}^{\beta\eta\alpha}_{\bullet}$ will be equal to $\overline{}$. Hence, for $\rho < \sigma < \tau$, ${}^{\rho\sigma\tau}_{\bullet}$ equals $\overline{}$, while ${}^{\sigma\tau\rho}_{\bullet}$ and ${}^{\tau\rho\sigma}_{\bullet}$ vanish.

The Y-equations all arise by cyclic permutations of the first which may also be written

$$Y^{(\tau\upsilon\sigma)} = Y^{(\tau\upsilon\rho)}\,\overline{Y^{(\sigma\upsilon\rho)}}.$$

If now $\bar{\upsilon}$ is the immediate successor of υ (that is, $\upsilon + 1$ for $\upsilon \neq \gamma$, 1 for $\upsilon = \gamma$), then we have

$$Y^{(\tau\upsilon\sigma)} = T^{(\tau\upsilon)}\,\overline{T^{(\sigma\upsilon)}},$$

where we define $T^{(\omega\upsilon)} = Y^{(\omega\upsilon\bar{\upsilon})}$ $(\omega \neq \upsilon)$ so long as the right side has meaning, and if it is meaningless $(\omega = \bar{\upsilon})$ then $T^{(\omega\upsilon)} = 1$. In this equation it is obvious that $\tau \neq \upsilon$; for $\tau \neq \bar{\upsilon}$ the equation follows from the preceding one with $\rho = \bar{\upsilon}$, and for $\tau = \bar{\upsilon}$ it is an identity. (Note that $T^{(\rho\sigma)}$ and $T^{(\sigma\rho)}$ are not identical.)

Among the operations of (48 *b*) we select

$$X_1^{(\alpha\beta)} = T^{(\beta\alpha)},\ X_2^{(\alpha\beta)} = \overline{T^{(\alpha\beta)}} \qquad (\alpha < \beta).$$

[24] We must leave both alternatives open since, for example, $\alpha\epsilon\beta$ is never an allowable sequence for $\alpha = \gamma$ and $\beta = 1$, and the same holds for β, η, α when $\alpha = \beta + 1$.

(The left sides are symmetric in α and β while the right sides are not. Hence, for $\alpha > \beta$, we obtain valid formulas by interchanging α and β on the right sides.) If $\rho < \sigma < \tau$, then, by (48 *b*), $Y^{(\rho\sigma\tau)}$, $Y^{(\sigma\tau\rho)}$, $Y^{(\tau\rho\sigma)}$ all go over into 1 since $\overset{(\rho\sigma\tau)}{\bullet} = \overline{}$ and $\overset{(\sigma\tau\rho)}{\bullet} = \overset{(\tau\rho\sigma)}{\bullet} = 0$. We have therefore made all the Y's simultaneously equal to 1, and if $\rho < \sigma < \tau$, then, by (47),

$$2A^{\rho\sigma\tau}_{\nu\lambda\mu} = R(i_\nu \bar{i}_\lambda i_\mu). \tag{49}$$

This result is exactly similar to that for $2A^{123}_{\nu\lambda\mu}$ in (45).

The multiplication rules (22) are accordingly uniquely determined by (49) and (23). Therefore in the preceding case there is at most one solution. But in any event a solution exists, namely, the Hermitian quaternion matrices of order γ (cf. paragraph 17). Hence this is the only solution.

20. $\gamma \geqq 4$, $\chi = 2$. In (45) we have obtained a solution $2A^{123}_{\nu\lambda\mu}$ in the form $R(i_\nu \bar{i}_\lambda i_\mu)$, where $i_1 = 1$ and $i_2 = i$ constitute the basis of the system of ordinary complex numbers. By (38) the most general solution arises if we replace

$$i_\nu \text{ by } \sum_{\nu'=1}^{2} \psi_{\nu\nu'} i_{\nu'}, \qquad i_\lambda \text{ by } \sum_{\lambda'=1}^{2} \varphi_{\lambda\lambda'} i_{\lambda'}, \qquad i_\mu \text{ by } \sum_{\mu'=1}^{2} \omega_{\mu\mu'} i_{\mu'}.$$

(We make no use of the restrictions proved at the end of paragraph 15). But every orthogonal transformation $\sum_{\kappa'=1}^{2} \sigma_{\kappa\kappa'} i_{\kappa'}$ of the basis of the complex number system may also be represented in the form $X i_\kappa \bullet$, where X is a complex number of absolute value unity and $\bullet$ represents nothing at all or else the sign for the complex conjugate. The most general expression for $2A^{123}_{\nu\lambda\mu}$ is therefore (note the commutativity of complex numbers)

$$2A^{123}_{\nu\lambda\mu} = R\,(Y^{(123)}\, i_\nu \overset{(312)}{\bullet}\, i_\lambda \overset{(123)}{\bullet}\, i_\mu \overset{(231)}{\bullet}),$$

where $Y^{(123)}$ is a certain complex number of absolute value unity and where the $\bullet$'s are independent of one another. It follows by (23) that this equation also holds for $2A^{231}_{\nu\lambda\mu}$ and $2A^{312}_{\nu\lambda\mu}$ if the indices 1, 2, and 3 are correspondingly permuted on the right side, provided that we define $Y^{(123)} = Y^{(231)} = Y^{(312)}$. Therefore we may set

$$2A^{\rho\sigma\tau}_{\nu\lambda\mu} = R\,(Y^{(\rho\sigma\tau)}\, i_\nu \overset{(\tau\rho\sigma)}{\bullet}\, i_\lambda \overset{(\rho\sigma\tau)}{\bullet}\, i_\mu \overset{(\sigma\tau\rho)}{\bullet}),\ Y^{(\rho\sigma\tau)} = Y^{(\sigma\tau\rho)} = Y^{(\tau\rho\sigma)}, \tag{50}$$

where the Y's are certain complex numbers of absolute value unity and where the $\bullet$'s are independent of one another. As in paragraph 19 we limit ourselves to such values of ρ, σ, τ which may be transformed into a monotonically increasing sequence by a cyclic permutation.

Again we must determine the influence of the basis transformations of the $M^{\rho\sigma}$ (in the sense of Theorem 13, (21) and (37)) on (50). This is exhibited in the formulas (47) for $2A^{\rho\sigma\tau}_{\nu\lambda\mu}$, $2A^{\sigma\tau\rho}_{\nu\lambda\mu}$, $2A^{\tau\rho\sigma}_{\nu\lambda\mu}$, or $2A^{\sigma\rho\tau}_{\nu\lambda\mu}$, $2A^{\rho\tau\sigma}_{\nu\lambda\mu}$, $2A^{\tau\sigma\rho}_{\nu\lambda\mu}$ as a transformation of the corresponding i_μ, i_λ, i_ν and hence as that of the i_κ with $\overset{(\sigma\tau\rho)}{\bullet}$ or $\overset{(\rho\tau\sigma)}{\bullet}$.

If we transform all the $M^{\rho\sigma}$ simultaneously so that i_κ in $M^{\rho\sigma}$ goes over into $X^{(\rho\sigma)} i_\kappa^{(\rho\sigma)}$, and if for the sake of simplicity we identify $\rho\sigma$ and $\sigma\rho$, then we have (note the commutativity of complex numbers)

$$(51)\quad \begin{cases} \text{(a)} & {}^{(\rho\sigma\tau)}_{\bullet} \text{ is replaced by } {}^{(\rho\sigma\tau)}_{\bullet}\,{}^{(\rho\tau)}_{\bullet}, \\ \text{(b)} & Y^{(\rho\sigma\tau)} \text{ is replaced by } Y^{(\rho\sigma\tau)}\, X^{(\sigma\tau)}\,{}^{(\tau\rho\sigma)}_{\bullet}\, X^{(\tau\rho)}\,{}^{(\rho\sigma\tau)}_{\bullet}\, X^{(\rho\sigma)}\,{}^{(\sigma\tau\rho)}_{\bullet}. \end{cases}$$

Further, since $R(X) = R(\bar{X})$ and $\overline{XY} = \bar{X}\bar{Y}$, we can immediately

$$(51)\quad \begin{cases} (c) \text{ alter all } {}^{(\rho\sigma\tau)}_{\bullet},\ {}^{(\sigma\tau\rho)}_{\bullet}, \text{ and } {}^{(\tau\rho\sigma)}_{\bullet} \text{ simultaneously,} \\ \quad \text{and replace } Y^{(\rho\sigma\tau)} \text{ by } \overline{Y^{(\rho\sigma\tau)}} \text{ in (50).} \end{cases}$$

By application of (51 c) we can simultaneously obtain ${}^{(\rho\sigma\tau)}_{\bullet} = \overline{}$ for all ρ, σ, τ such that $\rho < \sigma < \tau$.

We now consider (25). We assume $\rho < \sigma < \tau < \upsilon$. Then it follows that ${}^{(\rho\sigma\tau)}_{\bullet} = {}^{(\rho\sigma\upsilon)}_{\bullet} = {}^{(\rho\tau\upsilon)}_{\bullet} = {}^{(\sigma\tau\upsilon)}_{\bullet} = \overline{}$. If we substitute

$$\sum_{\kappa=1}^{2} R(X i_\kappa^{(1)}_{\bullet})\, R(Y i_\kappa^{(2)}_{\bullet}) = R(X\bar{Y}^{(1)}_{\bullet}\,{}^{(2)}_{\bullet})$$

and the particular choice of some $\bullet$'s just now formulated into (50), we get

$$R(Y^{(\rho\sigma\upsilon)}\, Y^{(\sigma\tau\upsilon)}\,{}^{(\upsilon\rho\sigma)}_{\bullet}\, \bar{i}_\theta\, i_\iota^{(\sigma\upsilon\rho)}_{\bullet}\, i_\lambda^{(\tau\upsilon\sigma)}\,{}^{(\upsilon\rho\sigma)}_{\bullet}\, i_\kappa^{(\upsilon\sigma\tau)}\,{}^{(\upsilon\rho\sigma)}_{\bullet})$$
$$= R(Y^{(\rho\tau\upsilon)}\, Y^{(\rho\sigma\tau)}\,{}^{(\tau\upsilon\rho)}_{\bullet}\, i_\kappa^{(\upsilon\rho\tau)}_{\bullet}\, \bar{i}_\theta\, i_\iota^{(\sigma\tau\rho)}\,{}^{(\tau\upsilon\rho)}_{\bullet}\, i_\lambda^{(\tau\rho\sigma)}\,{}^{(\tau\upsilon\rho)}_{\bullet}).$$

Since we are forming linear aggregates we can set free complex variables Z_θ, Z_ι, Z_λ, Z_κ in place of i_θ, i_ι, i_λ, i_κ. The above equality can hold only if the two parentheses are equal or if one is equal to the conjugate of the other. But since $\overline{Z_\theta}$ occurs on both sides the latter possibility is excluded. Consequently it must be the case that

$${}^{(\sigma\upsilon\rho)}_{\bullet} = {}^{(\sigma\tau\rho)}_{\bullet}\,{}^{(\tau\upsilon\rho)}_{\bullet},\quad {}^{(\tau\upsilon\rho)}_{\bullet}\,{}^{(\upsilon\rho\sigma)}_{\bullet} = {}^{(\tau\rho\sigma)}_{\bullet}\,{}^{(\tau\upsilon\rho)}_{\bullet},\quad {}^{(\upsilon\sigma\tau)}_{\bullet}\,{}^{(\upsilon\rho\sigma)}_{\bullet} = {}^{(\upsilon\rho\tau)}_{\bullet}.$$

The first equation can also be written ${}^{(\tau\upsilon\rho)}_{\bullet} = {}^{(\sigma\tau\rho)}_{\bullet}\,{}^{(\sigma\upsilon\rho)}_{\bullet}$ so that if $\bar{\rho}$ is the immediate successor of ρ (that is, $\rho + 1$ inasmuch as $\rho \neq \gamma$ because of the condition $\rho < \sigma < \tau < \upsilon$)

$${}^{(\tau\upsilon\rho)}_{\bullet} = {}^{((\tau\rho))}_{\bullet}\,{}^{((\upsilon\rho))}_{\bullet},$$

where we define ${}^{((\omega\rho))}_{\bullet} = {}^{(\bar{\omega}\rho)}_{\bullet}$ $(\omega \neq \rho)$ so long as the right side has meaning, and if it is meaningless $(\omega = \bar{\rho})$ then ${}^{((\omega\rho))}_{\bullet} = 0$.

For $\rho < \sigma < \tau$ it follows that ${}^{(\rho\sigma\tau)}_{\bullet} = \overline{}$ and ${}^{(\sigma\tau\rho)}_{\bullet} = {}^{((\sigma\rho))}_{\bullet}\,{}^{((\tau\rho))}_{\bullet}$, where ${}^{((\tau\rho))}_{\bullet}$ is arbitrary. If for ${}^{((\tau\rho))}_{\bullet} = \overline{}$ we apply the operation (51 c) and for ${}^{((\tau\rho))}_{\bullet} = 0$ we do nothing, then it results that ${}^{(\rho\sigma\tau)}_{\bullet} = \overline{{}^{((\tau\rho))}_{\bullet}}$ and ${}^{(\sigma\tau\rho)}_{\bullet} = {}^{((\sigma\rho))}_{\bullet}$, where ${}^{(\tau\rho\sigma)}_{\bullet}$ is arbitrary. If to this result we apply the operation (51 a) with ${}^{(\alpha\beta)}_{\bullet} = {}^{((\alpha\beta))}_{\bullet}$ $(\alpha > \beta)$, then it results that ${}^{(\rho\sigma\tau)}_{\bullet} = \overline{}$ and ${}^{(\sigma\tau\rho)}_{\bullet} = 0$, where ${}^{(\tau\rho\sigma)}_{\bullet}$ is arbitrary. From these equations we have

$${}^{(\upsilon\rho\sigma)}_{\bullet} = {}^{(\tau\rho\sigma)}_{\bullet},\quad {}^{(\upsilon\sigma\tau)}\,{}^{(\upsilon\rho\sigma)}_{\bullet} = {}^{(\upsilon\rho\tau)}_{\bullet}.$$

The first states that ${}^{(\omega\rho\sigma)}_{\bullet}$ (naturally $\rho < \sigma < \omega$) is independent of ω so that ${}^{(\omega\rho\sigma)}_{\bullet} = {}^{(((\rho\sigma)))}_{\bullet}$ and the second states that ${}^{(((\sigma\tau)))}_{\bullet}\ {}^{(((\rho\sigma)))}_{\bullet} = {}^{(((\rho\tau)))}_{\bullet}$. Therefore ${}^{(((\rho\sigma)))}_{\bullet} = {}^{[\rho]}_{\bullet}{}^{[\sigma]}_{\bullet}$ and ${}^{(\omega\rho\sigma)}_{\bullet} = {}^{[\rho]}_{\bullet}{}^{[\sigma]}_{\bullet}$. Therefore ${}^{(\rho\sigma\tau)}_{\bullet} = \overline{}$, ${}^{(\sigma\tau\rho)}_{\bullet} = 0$, and ${}^{(\tau\rho\sigma)}_{\bullet} = {}^{[\rho]}_{\bullet}{}^{[\sigma]}_{\bullet}$. If for ${}^{[\rho]}_{\bullet} = \overline{}$ we here apply the operation (51 c) and for ${}^{[\rho]}_{\bullet} = 0$ we do nothing, then it results that ${}^{(\rho\sigma\tau)}_{\bullet} = \overline{{}^{[\rho]}_{\bullet}}$, ${}^{(\sigma\tau\rho)}_{\bullet} = {}^{[\rho]}_{\bullet}$, and ${}^{(\tau\rho\sigma)}_{\bullet} = {}^{[\sigma]}_{\bullet}$. If we apply the operation (51 a) with ${}^{(\alpha\beta)}_{\bullet} = {}^{[\beta]}_{\bullet}$ ($\alpha > \beta$), then we have ${}^{(\rho\sigma\tau)}_{\bullet} = \overline{}$ and ${}^{(\sigma\tau\rho)}_{\bullet} = {}^{(\tau\rho\sigma)}_{\bullet} = 0$.

In the second equation following (51 c) the coefficients in the parentheses must be equal to one another so that for $\rho < \sigma < \tau < \upsilon$ it follows that $Y^{(\rho\sigma\upsilon)}\ Y^{(\sigma\tau\upsilon)}\ {}^{(\upsilon\rho\sigma)}_{\bullet} = Y^{(\rho\tau\upsilon)}\ Y^{(\rho\sigma\tau)}\ {}^{(\tau\upsilon\rho)}_{\bullet}$ so that with the $\bullet$ just obtained, $Y^{(\rho\sigma\upsilon)}\ Y^{(\sigma\tau\upsilon)} = Y^{(\rho\tau\upsilon)}\ Y^{(\rho\sigma\tau)}$ and hence $Y^{(\rho\sigma\tau)} = Y^{(\sigma\tau\upsilon)}\ \overline{Y^{(\rho\tau\upsilon)}}\ Y^{(\rho\sigma\upsilon)}$. Consequently

$$Y^{(\rho\sigma\tau)} = T^{(\sigma\tau)}\ \overline{T^{(\rho\tau)}}\ T^{(\rho\sigma)},$$

where we set $T^{(\alpha\beta)} = Y^{(\alpha\beta\gamma)}$ ($\alpha < \beta$) so long as the right side has meaning, and if it is meaningless ($\beta = \gamma$) then $T^{(\alpha\beta)} = 1$. We now select from the operations of (51 b)

$$X^{(\alpha\beta)} = \overline{T^{(\alpha\beta)}} \qquad (\alpha < \beta).$$

(The left side is symmetric in α and β while the right side is not so that we obtain the formulas valid for $\alpha > \beta$ by interchange of α and β on the right side.) Then all $Y^{(\rho\sigma\tau)}$ go over into 1.

By (50) it follows that, for $\rho < \sigma < \tau$,

$$2A^{\rho\sigma\tau}_{\nu\lambda\mu} = R(i_\nu\ \overline{i_\lambda}\ i_\mu) \tag{52}$$

which is similar to (49) and to $2A^{1\,2\,3}_{\nu\lambda\mu}$ in (45).

Hence the multiplication rules (22) are uniquely determined by (52) and (23). There is therefore at most one solution in the preceding case. But a solution exists in any event, for example, the Hermitian complex matrices of order γ (cf. paragraph 17). Hence it is the only solution.

21. $\gamma \geqq 4$, $\chi = 1$. In this case it is not worth while to carry out the general theory of paragraphs 14 to 17. Since $\lambda = \mu = \nu = 1$ in $A^{\rho\sigma\tau}_{\nu\lambda\mu}$, therefore $2A^{\rho\sigma\tau}_{\nu\lambda\mu} = Y^{\rho\sigma\tau}$ and (24) leads to the result $Y^{\rho\sigma\tau} = \pm 1$. The basis transformations of $\mathrm{M}^{\rho\sigma}$ in the sense of Theorem 13, (21), and (37) are $(s_1^{\rho\sigma})' = X^{\rho\sigma}\ s_1^{\rho\sigma}$ and $X^{\rho\sigma} = \pm 1$, where for the sake of simplicity we identify $\rho\sigma$ and $\sigma\rho$. Hence

$$Y^{(\rho\sigma\tau)} \text{ is replaced by } Y^{(\rho\sigma\tau)}\ X^{(\sigma\tau)}\ X^{(\tau\rho)}\ X^{(\rho\sigma)}. \tag{53}$$

In the preceding case it follows from (25) that

$$Y^{(\rho\sigma\upsilon)}\ Y^{(\sigma\tau\upsilon)} = Y^{(\rho\tau\upsilon)}\ Y^{(\rho\sigma\tau)}. \tag{54}$$

The relation

$$Y^{(\rho\sigma\tau)} = T^{(\sigma\tau)}\ T^{(\rho\tau)}\ T^{(\rho\sigma)}$$

follows from (54) in exactly the same way as did the analogous relation for the complex Y at the end of paragraph 20. We similarly find that every T is either a Y or 1 and hence is ± 1. Therefore we select from the operations of (53)

$$X^{(\alpha\beta)} = T^{(\alpha\beta)} \qquad (\alpha < \beta).$$

(With reference to the symmetry conditions in α and β, cf. the remark at the end of paragraph 20.) Then all $Y^{(\rho\sigma\tau)}$ go over into 1. Hence we have in general that

$$2A^{\rho\sigma\tau}_{\nu\lambda\mu} = 1. \tag{54}$$

The multiplication rules (22) are uniquely determined by (54) and (23). Hence in the preceding case there is at most one solution. But in any event a solution exists, for example, the Hermitian (real) matrices of order γ (cf. paragraph 17). Hence it is the only solution.

22. We now summarize our results.

In the course of the preceding investigations we have learned the properties of the following algebras:

α. $\mathfrak{R}$, the algebra of the real numbers in which $a \pm b$, λa, and ab are defined in the usual way.

β. $\mathfrak{S}_N$, $N = 1, 2, \cdots$, the algebra with the linear basis $1, s_1, \cdots, s_{N-1}$, in which $a \pm b$ and λa are defined in the usual way but in which ab is defined by

$$11 = 1,\ 1\, s_\mu = s_\mu, \qquad s_\mu s_\nu = \delta_{\mu\nu} 1.$$

γ. $\mathfrak{M}^\chi_\gamma$, $\chi = 1, 2, 4, 8$; $\gamma = 1, 2, \cdots$, the algebra of the Hermitian matrices of order γ in which the elements are real numbers, ordinary complex numbers, quanternions, and quasi-quaternions respectively for $\chi = 1, 2, 4$, and 8. $a \pm b$ and λa are defined in the usual way, but ab is defined by $\dfrac{a \times b + b \times a}{2}$, where $a \times b$ represents the usual matrix multiplication.

We can summarize the results of paragraphs 12 to 21 as follows:

FUNDAMENTAL THEOREM 2. *The only irreducible r-number algebras are the following:*

For $\gamma = 1$: $\mathfrak{R}$.

For $\gamma = 2$: *All* $\mathfrak{S}_N$, $N = 3, 4, \cdots$. ($\mathfrak{S}_1$ *is* $\mathfrak{R}$ *and* $\mathfrak{S}_2$ *is reducible.*)

For $\gamma = 3$: *All* $\mathfrak{M}^\chi_3$, $\chi = 1, 2, 4, 8$.

For $\gamma \geqq 4$: *All* $\mathfrak{M}^\chi_\gamma$, $\chi = 1, 2, 4$. (*All* $\mathfrak{M}^\chi_1$ *are* $\mathfrak{R}$; *the* $\mathfrak{M}^\chi_2$ *are respectively the* $\mathfrak{S}_{\chi+2}$; *the* $\mathfrak{M}^\chi_3$ *are already classified under* $\gamma = 3$; *the* $\mathfrak{M}^8_\gamma$, $\gamma \geqq 4$, *are not r-number algebras.*)

In connection with this result we mention the following circumstance: let A' be an associative, but not necessarily commutative algebra having the real

numbers as coefficients. Then $a \pm b$, λa, and $a \times b$ are the fundamental operations and all the usual rules of calculation are valid except for $a \times b = b \times a$. We now define $ab = \dfrac{a \times b + b \times a}{2}$ and we consider a subsystem A of A' which is closed for $a \pm b$, λa, and ab (but not necessarily for $a \times b$) and which is "formally real" ($a^2 = a \times a = aa$; cf. (I) in paragraph 1 and footnote 5). Since (II) is obviously satisfied (a^n is the same for ab and $a \times b$), A is an r-number algebra. By the theorem of Wedderburn[25] we could select a matrix algebra for A' without loss of generality. We say (in agreement with the terminology of footnote 1) that an r-number algebra which is isomorphic with such an A *results from quasi-multiplication.* As one of us has shown elsewhere (cf. footnote 1), essentially new results, in contrast with the present content of quantum mechanics, are to be expected only in such r-number algebras which do *not* result from quasi-multiplication.

Now all the r-number algebras of our list, with the exception of $\mathfrak{M}_3^8$, result from quasi-multiplication. This is obvious in the case of $\mathfrak{R}$ and the $\mathfrak{M}_\gamma^\chi$ ($\chi = 1, 2, 4$) so that only the $\mathfrak{S}_N$ need examination. But in any event the $\mathfrak{S}_N$ arise by the operation $ab = \dfrac{a \times b + b \times a}{2}$ from a system in which 1 is the modulus, $s_\mu^2 = 1$, and $s_\mu s_\nu = -s_\nu s_\mu (\mu \neq \nu, \mu, \nu = 1, \cdots, N-1)$. We now take the quaternion-like, non-commutative, but associative system $1, i, j, k$ with $i^2 = j^2 = k^2 = 1$, $ij = -ji = k$, $jk = -kj = i$, $ki = -ik = j$ and form $N-1$ sets $1^{(1)}, i^{(1)}, j^{(1)}, k^{(1)}; \cdots; 1^{(N-1)}, i^{(N-1)}, j^{(N-1)}, k^{(N-1)}$ of this sort. The direct products $1 = 1^{(1)} \cdot 1^{(2)} \cdot \cdots \cdot 1^{(N-1)}$, $s_1 = j^{(1)} \cdot 1^{(2)} \cdot \cdots \cdot 1^{(N-1)}$, $s_2 = i^{(1)} \cdot j^{(2)} \cdot \cdots \cdot 1^{(N-1)}, \cdots, s_{N-1} = i^{(1)} \cdot i^{(2)} \cdot \cdots \cdot j^{(N-1)}$ perform the desired function.

On the other hand it can be shown that $\mathfrak{M}_3^8$ can not result from quasi-multiplication. The proof of this remarkable result was given by A. Albert, to whom we proposed the problem, and can be found in his next paper in this journal. Hence there is one single (irreducible) r-number algebra which does not arise from quasi-multiplication: $\mathfrak{M}_3^8$.

ROSTOCK, GERMANY, AND PRINCETON, N. J.

[25] Proc. London Math. Soc., 1907, Vol. (2) 6, p. 99.

On Unitary Representations of the Inhomogeneous Lorentz Group*

E. P. Wigner

Annals of Mathematics *40*, 149–204 (1939)

(Received December 22, 1937)

1. Origin and Characterization of the Problem

It is perhaps the most fundamental principle of Quantum Mechanics that the system of states forms a *linear manifold*,[1] in which a unitary *scalar product* is defined.[2] The states are generally represented by wave functions[3] in such a way that φ and constant multiples of φ represent the same physical state. It is possible, therefore, to normalize the wave function, i.e., to multiply it by a constant factor such that its scalar product with itself becomes 1. Then, only a constant factor of modulus 1, the so-called phase, will be left undetermined in the wave function. The linear character of the wave function is called the superposition principle. The square of the modulus of the unitary scalar product (ψ, φ) of two normalized wave functions ψ and φ is called the transition probability from the state ψ into φ, or conversely. This is supposed to give the probability that an experiment performed on a system in the state φ, to see whether or not the state is ψ, gives the result that it is ψ. If there are two or more different experiments to decide this (e.g., essentially the same experiment,

---

* Parts of the present paper were presented at the Pittsburgh Symposium on Group Theory and Quantum Mechanics. Cf. Bull. Amer. Math. Soc., 41, p. 306, 1935.

[1] The possibility of a future non linear character of the quantum mechanics must be admitted, of course. An indication in this direction is given by the theory of the positron, as developed by P. A. M. Dirac (Proc. Camb. Phil. Soc. *30*, 150, 1934, cf. also W. Heisenberg, Zeits. f. Phys. *90*, 209, 1934; *92*, 623, 1934; W. Heisenberg and H. Euler, ibid. *98*, 714, 1936 and R. Serber, Phys. Rev. *48*, 49, 1935; *49*, 545, 1936) which does not use wave functions and is a non linear theory.

[2] Cf. P. A. M. Dirac, *The Principles of Quantum Mechanics*, Oxford 1935, Chapters I and II; J. v. Neumann, *Mathematische Grundlagen der Quantenmechanik*, Berlin 1932, pages 19–24.

[3] The wave functions represent throughout this paper states in the sense of the "Heisenberg picture," i.e. a single wave function represents the state for all past and future. On the other hand, the operator which refers to a measurement at a certain time t contains this t as a parameter. (Cf. e.g. Dirac, l.c. ref. 2, pages 115-123). One obtains the wave function $\varphi_s(t)$ of the Schrödinger picture from the wave function φ_H of the Heisenberg picture by $\varphi_s(t) = \exp(-iHt/\hbar)\varphi_H$. The operator of the Heisenberg picture is $Q(t) = \exp(iHt/\hbar)\, Q \exp(-iHt/\hbar)$, where Q is the operator in the Schrödinger picture which does not depend on time. Cf. also E. Schrödinger, Sitz. d. Kön. Preuss. Akad. p. 418, 1930.

The wave functions are complex quantities and the undetermined factors in them are complex also. Recently attempts have been made toward a theory with real wave functions. Cf. E. Majorana, Nuovo Cim. *14*, 171, 1937 and P. A. M. Dirac, in print.

performed at different times) they are all supposed to give the same result, i.e., the transition probability has an invariant physical sense.

The wave functions form a description of the physical state, not an invariant however, since the same state will be described in different coördinate systems by different wave functions. In order to put this into evidence, we shall affix an index to our wave functions, denoting the Lorentz frame of reference for which the wave function is given. Thus φ_l and $\varphi_{l'}$ represent the same state, but they are different functions. The first is the wave function of the state in the coordinate system l, the second in the coördinate system l'. If $\varphi_l = \psi_{l'}$ the state φ behaves in the coördinate system l exactly as ψ behaves in the coördinate system l'. If φ_l is given, all $\varphi_{l'}$ are determined up to a constant factor. Because of the invariance of the transition probability we have

$$(1) \qquad |(\varphi_l, \psi_l)|^2 = |(\varphi_{l'}, \psi_{l'})|^2$$

and it can be shown[4] that the aforementioned constants in the $\varphi_{l'}$ can be chosen in such a way that the $\varphi_{l'}$ are obtained from the φ_l by a linear unitary operation, depending, of course, on l and l'

$$(2) \qquad \varphi_{l'} = D(l', l)\varphi_l .$$

The unitary operators D are determined by the physical content of the theory up to a constant factor again, which can depend on l and l'. Apart from this constant however, the operations $D(l', l)$ and $D(l_1', l_1)$ must be identical if l' arises from l by the same Lorentz transformation, by which l_1' arises from l_1. If this were not true, there would be a real difference between the frames of reference l and l_1. Thus the unitary operator $D(l', l) = D(L)$ is in every Lorentz invariant quantum mechanical theory (apart from the constant factor which has no physical significance) completely determined by the Lorentz transformation L which carries l into $l' = Ll$. One can write, instead of (2)

$$(2a) \qquad \varphi_{Ll} = D(L)\varphi_l .$$

By going over from a first system of reference l to a second $l' = L_1 l$ and then to a third $l'' = L_2 L_1 l$ or directly to the third $l'' = (L_2 L_1)l$, one must obtain—apart from the above mentioned constant—the same set of wave functions. Hence from

$$\varphi_{l''} = D(l'', l')D(l', l)\varphi_l$$

$$\varphi_{l''} = D(l'', l)\varphi_l$$

it follows

$$(3) \qquad D(l'', l')D(l', l) = \omega D(l'', l)$$

[4] E. Wigner, *Gruppentheorie und ihre Anwendungen auf die Quantenmechanik der Atomspektren.* Braunschweig 1931, pages 251–254.

or

$$D(L_2)D(L_1) = \omega D(L_2L_1), \tag{3a}$$

where ω is a number of modulus 1 and can depend on L_2 and L_1. Thus the $D(L)$ form, up to a factor, a representation of the inhomogeneous Lorentz group by linear, unitary operators.

We see thus[5] that there corresponds to every invariant quantum mechanical system of equations such a representation of the inhomogeneous Lorentz group. This representation, on the other hand, though not sufficient to replace the quantum mechanical equations entirely, can replace them to a large extent. If we knew, e.g., the operator K corresponding to the measurement of a physical quantity at the time $t = 0$, we could follow up the change of this quantity throughout time. In order to obtain its value for the time $t = t_1$, we could transform the original wave function φ_l by $D(l', l)$ to a coördinate system l' the time scale of which begins a time t_1 later. The measurement of the quantity in question in this coördinate system for the time 0 is given—as in the original one—by the operator K. This measurement is indentical, however, with the measurement of the quantity at time t_1 in the original system. One can say that the representation can replace the equation of motion, it cannot replace, however, connections holding between operators at one instant of time.

It may be mentioned, finally, that these developments apply not only in quantum mechanics, but also to all linear theories, e.g., the Maxwell equations in empty space. The only difference is that there is no arbitrary factor in the description and the ω can be omitted in (3a) and one is led to real representations instead of representations up to a factor. On the other hand, the unitary character of the representation is not a consequence of the basic assumptions.

The increase in generality, obtained by the present calculus, as compared with the usual tensor theory, consists in that no assumptions regarding the field nature of the underlying equations are necessary. Thus more general equations, as far as they exist (e.g., in which the coördinate is quantized, etc.) are also included in the present treatment. It must be realized, however, that some assumptions concerning the continuity of space have been made by assuming Lorentz frames of reference in the classical sense. We should like to mention, on the other hand, that the previous remarks concerning the time-parameter in the observables, have only an explanatory character, and we do not make assumptions of the kind that measurements can be performed instantaneously.

We shall endeavor, in the ensuing sections, to determine all the continuous[6] unitary representations up to a factor of the inhomogeneous Lorentz group, i.e., all continuous systems of linear, unitary operators satisfying (3a).

---

[5] E. Wigner, l.c. Chapter XX.

[6] The exact definition of the continuous character of a representation up to a factor will be given in Section 5A. The definition of the inhomogeneous Lorentz group is contained in Section 4A.

2. Comparison With Previous Treatments and Some Immediate Simplifications

A. Previous treatments

The representations of the Lorentz group have been investigated repeatedly. The first investigation is due to Majorana,[7] who in fact found all representations of the class to be dealt with in the present work excepting two sets of representations. Dirac[8] and Proca[8] gave more elegant derivations of Majorana's results and brought them into a form which can be handled more easily. Klein's work[9] does not endeavor to derive irreducible representations and seems to be in a less close connection with the present work.

The difference between the present paper and that of Majorana and Dirac lies—apart from the finding of new representations—mainly in its greater mathematical rigor. Majorana and Dirac freely use the notion of infinitesimal operators and a set of functions to all members of which every infinitesimal operator can be applied. This procedure cannot be mathematically justified at present, and no such assumption will be used in the present paper. Also the conditions of reducibility and irreducibility could be, in general, somewhat more complicated than assumed by Majorana and Dirac. Finally, the previous treatments assume from the outset that the space and time coördinates will be continuous variables of the wave function in the usual way. This will not be done, of course, in the present work.

B. Some immediate simplifications

Two representations are *physically equivalent* if there is a one to one correspondence between the states of both which is 1. invariant under Lorentz transformations and 2. of such a character that the transition probabilities between corresponding states are the same.

It follows from the second condition[5] that there either exists a unitary operator S by which the wave functions $\Phi^{(2)}$ of the second representation can be obtained from the corresponding wave functions $\Phi^{(1)}$ of the first representation

$$\Phi^{(2)} = S\Phi^{(1)} \tag{4}$$

or that this is true for the conjugate imaginary of $\Phi^{(2)}$. Although, in the latter case, the two representations are still equivalent physically, we shall, in keeping with the mathematical convention, not call them equivalent.

The first condition now means that if the states $\Phi^{(1)}$, $\Phi^{(2)} = S\Phi^{(1)}$ correspond to each other in one coördinate system, the states $D^{(1)}(L)\Phi^{(1)}$ and $D^{(2)}(L)\Phi^{(2)}$ correspond to each other also. We have then

$$D^{(2)}(L)\Phi^{(2)} = SD^{(1)}(L)\Phi^{(1)} = SD^{(1)}(L)S^{-1}\Phi^{(2)}. \tag{4a}$$

[7] E. Majorana, Nuovo Cim. *9*, 335, 1932.

[8] P. A. M. Dirac, Proc. Roy. Soc. A. *155*, 447, 1936; Al. Proca, J. de Phys. Rad. *7*, 347, 1936.

[9] Klein, Arkiv f. Matem. Astr. och Fysik, *25A*, No. 15, 1936. I am indebted to Mr. Darling for an interesting conversation on this paper.

As this shall hold for every $\Phi^{(2)}$, the existence of a unitary S which transforms $D^{(1)}$ into $D^{(2)}$ is the condition for the equivalence of these two representations. Equivalent representations are not considered to be really different and it will be sufficient to find one sample from every infinite class of equivalent representations.

If there is a closed linear manifold of states which is invariant under all Lorentz transformations, i.e. which contains $D(L)\psi$ if it contains ψ, the linear manifold perpendicular to this one will be invariant also. In fact, if φ belongs to the second manifold, $D(L)\varphi$ will be, on account of the unitary character of $D(L)$, perpendicular to $D(L)\psi'$ if ψ' belongs to the first manifold. However, $D(L^{-1})\psi$ belongs to the first manifold if ψ does and thus $D(L)\varphi$ will be orthogonal to $D(L)D(L^{-1})\psi = \omega\psi$ i.e. to all members of the first manifold and belong itself to the second manifold also. The original representation then "decomposes" into two representations, corresponding to the two linear manifolds. It is clear that, conversely, one can form a representation, by simply "adding" several other representations together, i.e. by considering as states linear combinations of the states of several representations and assume that the states which originate from different representations are perpendicular to each other.

Representations which are equivalent to sums of already known representations are not really new and, in order to master all representations, it will be sufficient to determine those, out of which all others can be obtained by "adding" a finite or infinite number of them together.

Two simple theorems shall be mentioned here which will be proved later (Sections 7A and 8C respectively). The first one refers to unitary representations of any closed group, the second to irreducible unitary representations of any (closed or open) group.

The representations of a closed group by unitary *operators* can be transformed into the sum of unitary representations with matrices of finite dimensions.

Given two non equivalent irreducible unitary representations of an arbitrary group. If the scalar product between the wave functions is invariant under the operations of the group, the wave functions belonging[28] to the first representation are orthogonal to all wave functions belonging to the second representation.

C. Classification of unitary representations according to von Neumann and Murray[10]

Given the operators $D(L)$ of a unitary representations, or a representation up to a factor, one can consider the algebra of these operators, i.e. all linear combinations

$$a_1D(L_1) + a_2D(L_2) + a_3D(L_3) + \cdots$$

of the $D(L)$ and all limits of such linear combinations which are bounded operators. According to the properties of this representation algebra, three classes of unitary representations can be distinguished.

[10] F. J. Murray and J. v. Neumann, Ann. of Math. *37*, 116, 1936; J. v. Neumann, to be published soon.

The first class of *irreducible* representations has a representation algebra which contains all bounded operators, i.e. if ψ and φ are two arbitrary states, there is an operator A of the representation algebra for which $A\psi = \varphi$ and $A\psi' = 0$ if ψ' is orthogonal to ψ. It is clear that the center of the algebra contains only the unit operator and multiply thereof. In fact, if C is in the center one can decompose $C\psi = \alpha\psi + \psi'$ so that ψ' shall be orthogonal to ψ. However, ψ' must vanish since otherwise C would not commute with the operator which leaves ψ invariant and transforms every function orthogonal to it into 0. For similar reasons, α must be the same for all ψ. For irreducible representations there is no closed linear manifold of states, (excepting the manifold of all states) which is invariant under all Lorentz transformations. In fact, according to the above definition, a φ' arbitrarily close to any φ can be represented by a finite linear combination

$$a_1D(L_1)\psi + a_2D(L_2)\psi + \cdots + a_nD(L_n)\psi.$$

Hence, a closed linear invariant manifold contains every state if it contains one. This is, in fact, the more customary definition for irreducible representations and the one which will be used subsequently. It is well known that all finite dimensional representations are sums of irreducible representations. This is not true,[10] in general, in an infinite number of dimensions.

The second class of representations will be called *factorial*. For these, the center of the representation algebra still contains only multiples of the unit operator. Clearly, the irreducible representations are all factorial, but not conversely. For finite dimensions, the factorial representations may contain one irreducible representation several times. This is also possible in an infinite number of dimensions, but in addition to this, there are the "continuous" representations of Murray and von Neumann.[10] These are not irreducible as there are invariant linear manifolds of states. On the other hand, it is impossible to carry the decomposition so far as to obtain as parts only irreducible representations. In all the examples known so far, the representations into which these continuous representations can be decomposed, are equivalent to the original representation.

The third class contains all possible unitary representations. In a finite number of dimensions, these can be decomposed first into factorial representations, and these, in turn, in irreducible ones. Von Neumann[10] has shown that the first step still is possible in infinite dimensions. We can assume, therefore, from the outset that we are dealing with factorial representations.

In the theory of representations of finite dimensions, it is sufficient to determine only the irreducible ones, all others are equivalent to sums of these. Here, it will be necessary to determine all factorial representations. Having done that, we shall know from the above theorem of von Neumann, that all representations are equivalent to finite or infinite sums of factorial representations.

It will be one of the results of the detailed investigation that the inhomogeneous Lorentz group has no "continuous" representations, all representations

can be decomposed into irreducible ones. Thus the work of Majorana and Dirac appears to be justified from this point of view a posteriori.

D. Classification of unitary representations from the point of view of infinitesimal operators

The existence of an infinitesimal operator of a continuous one parametric (cyclic, abelian) unitary group has been shown by Stone.[11] He proved that the operators of such a group can be written as $\exp(iHt)$ where H is a (bounded or unbounded) hermitean operator and t is the group parameter. However, the Lorentz group has many one parametric subgroups, and the corresponding infinitesimal operators $H_1, H_2, \cdots$ are all unbounded. For every H_i an everywhere dense set of functions φ can be found such that $H_i\varphi$ can be defined. It is not clear, however, that an everywhere dense set can be found, to all members of which every H can be applied. In fact, it is not clear that one such φ can be found.

Indeed, it may be interesting to remark that for an irreducible representation the existence of one function φ to which all infinitesimal operators can be applied, entails the existence of an everywhere dense set of such functions. This again has the consequence that one can operate with infinitesimal operators to a large extent in the usual way.

PROOF: Let $Q(t)$ be a one parametric subgroup such that $Q(t)Q(t') = Q(t + t')$. If the infinitesimal operator of all subgroups can be applied to φ, the

$$\lim_{t=0} t^{-1}(Q(t) - 1)\varphi \tag{5}$$

exists. It follows, then, that the infinitesimal operators can be applied to $R\varphi$ also where R is an arbitrary operator of the representation: Since $R^{-1}Q(t)\,R$ is also a one parametric subgroup

$$\lim_{t=0} t^{-1}(R^{-1}Q(t)R - 1)\varphi = \lim_{t=0} R^{-1}\cdot t^{-1}(Q(t) - 1)R\varphi$$

also exists and hence also (R is unitary)

$$\lim_{t=0} t^{-1}(Q(t) - 1)R\varphi.$$

Every infinitesimal operator can be applied to $R\varphi$ if they all can be applied to φ, and the same holds for sums of the kind

$$a_1R_1\varphi + a_2R_2\varphi + \cdots + a_nR_n\varphi. \tag{6}$$

These form, however, an everywhere dense set of functions if the representation is irreducible.

If the representation is not irreducible, one can consider the set N_0 of such wave functions to which every infinitesimal operator can be applied. This set is

[11] M. H. Stone, Proc. Nat. Acad. *16*, 173, 1930, Ann. of Math. *33*, 643, 1932, also J. v. Neumann, ibid, *33*, 567, 1932.

clearly linear and, according to the previous paragraph, invariant under the operations of the group (i.e. contains every $R\varphi$ if it contains φ). The same holds for the closed set N generated by N_0 and also of the set P of functions which are perpendicular to all functions of N. In fact, if φ_p is perpendicular to all φ_n of N, it is perpendicular also to all $R^{-1}\varphi_n$ and, for the unitary character of R, the $R\varphi_p$ is perpendicular to all φ_n, i.e. is also contained in the set P.

We can decompose thus, by a unitary transformation, every unitary representation into a "normal" and a "pathological" part. For the former, there is an everywhere dense set of functions, to which all infinitesimal operators can be applied. There is no single wave functions to which all infinitesimal operators of a "pathological" representation could be applied.

According to Murray and von Neumann, if the original representation was factorial, all representations into which it can be decomposed will be factorial also. Thus every representation is equivalent to a sum of factorial representations, part of which is "normal," the other part "pathological."

It will turn out again that the inhomogeneous Lorentz group has no pathological representations. Thus this assumption of Majorana and Dirac also will be justified a posteriori. Every unitary representation of the inhomogenous Lorentz group can be decomposed into normal irreducible representations. It should be stated, however, that the representations in which the unit operator corresponds to every translation have not been determined to date (cf. also section 3, end). Hence, the above statements are not proved for these representations, which are, however, more truly representations of the homogeneous Lorentz group, than of the inhomogeneous group.

While all these points may be of interest to the mathematician only, the new representation of the Lorentz group which will be described in section 7 may interest the physicist also. It describes a particle with a continuous spin.

Acknowledgement. The subject of this paper was suggested to me as early as 1928 by P. A. M. Dirac who realised even at that date the connection of representations with quantum mechanical equations. I am greatly indebted to him also for many fruitful conversations about this subject, especially during the years 1934/35, the outgrowth of which the present paper is.

I am indebted also to J. v. Neumann for his help and friendly advice.

3. Summary of Ensuing Sections

Section 4 will be devoted to the definition of the inhomogeneous Lorentz group and the theory of characteristic values and characteristic vectors of a homogeneous (ordinary) Lorentz transformation. The discussion will follow very closely the corresponding, well-known theory of the group of motions in ordinary space and the theory of characteristic values of orthogonal transformations.[12] It will contain only a straightforward generalization of the methods usually applied in those discussions.

[12] Cf. e.g. E. Wigner, l.c. Chapter III. O. Veblen and J. W. Young, *Projective Geometry*, Boston 1917. Vol. 2, especially Chapter VII.

In Section 5, it will be proved that one can determine the physically meaningless constants in the $D(L)$ in such a way that instead of (3a) the more special equation

$$D(L_1)D(L_2) = \pm D(L_1L_2) \tag{7}$$

will be valid. This means that instead of a representation up to a factor, we can consider representations up to the sign. For the case that either L_1 or L_2 is a pure translation, Dirac[13] has given a proof of (7) using infinitesimal operators. A consideration very similar to his can be carried out, however, also using only finite transformations.

For representations with a finite number of dimensions (corresponding to an only finite number of linearly independent states), (7) could be proved also if both L_1 and L_2 are homogeneous Lorentz transformations, by a straightforward application of the method of Weyl and Schreier.[14] However, the Lorentz group has no finite dimensional representation (apart from the trivial one in which the unit operation corresponds to every L). Thus the method of Weyl and Schreier cannot be applied. Its first step is to normalize the indeterminate constants in every matrix $D(L)$ in such a way that the determinant of $D(L)$ becomes 1. No determinant can be defined for general unitary operators.

The method to be employed here will be to decompose every L into a product of two involutions $L = MN$ with $M^2 = N^2 = 1$. Then $D(M)$ and $D(N)$ will be normalized so that their squares become unity and $D(L) = D(M)D(N)$ set. It will be possible, then, to prove (7) without going back to the topology of the group.

Sections 6, 7, and 8 will contain the determination of the representations. The pure translations form an invariant subgroup of the whole inhomogeneous Lorentz group and Frobenius' method[15] will be applied in Section 6 to build up the representations of the whole group out of representations of the subgroup, by means of a "little group." In Section 6, it will be shown on the basis of an as yet unpublished work[24] of J. v. Neumann that there is a characteristic (invariant) set of "momentum vectors" for every irreducible representation. The irreducible representations of the Lorentz group will be divided into four classes. The momentum vectors of the

1st class are time-like,

2nd class are null-vectors, but not all their components will be zero,

3rd class vanish (i.e., all their components will be zero),

4th class are space-like.

Only the first two cases will be considered in Section 7, although the last case

[13] P. A. M. Dirac, mimeographed notes of lectures delivered at Princeton University, 1934/35, page 5a.

[14] H. Weyl, Mathem. Zeits. *23*, 271; *24*, 328, 377, 789, 1925; O. Schreier, Abhandl. Mathem. Seminar Hamburg, *4*, 15, 1926; *5*, 233, 1927.

[15] G. Frobenius, Sitz. d. Kön. Preuss. Akad. p. 501, 1898, I. Schur, ibid, p. 164, 1906; F. Seitz, Ann. of Math. *37*, 17, 1936.

may be the most interesting from the mathematical point of view. I hope to return to it in another paper. I did not succeed so far in giving a complete discussion of the 3rd class. (All these restrictions appear in the previous treatments also.)

In Section 7, we shall find again all known representations of the inhomogeneous Lorentz group (i.e., all known Lorentz invariant equations) and two new sets.

Sections 5, 6, 7 will deal with the "restricted Lorentz group" only, i.e. Lorentz transformations with determinant 1 which do not reverse the direction of the time axis. In section 8, the representations of the extended Lorentz group will be considered, the transformations of which are not subject to these conditions.

4. Description of the Inhomogeneous Lorentz Group

A.

An inhomogeneous Lorentz transformation $L = (a, \Lambda)$ is the product of a translation by a real vector a

$$x_i' = x_i + a_i \qquad (i = 1, 2, 3, 4) \tag{8}$$

and a homogeneous Lorentz transformation Λ with real coefficients

$$x_i' = \sum_{k=1}^{4} \Lambda_{ik} x_k . \tag{9}$$

The translation shall be performed after the homogeneous transformation. The coefficients of the homogeneous transformation satisfy three conditions: (1) They are real and Λ leaves the indefinite quadratic form $-x_1^2 - x_2^2 - x_3^2 + x_4^2$ invariant:

$$\Lambda F \Lambda' = F \tag{10}$$

where the prime denotes the interchange of rows and columns and F is the diagonal matrix with the diagonal elements $-1, -1, -1, +1$.—(2) The determinant $|\Lambda_{ik}| = 1$ and—(3) $\Lambda_{44} > 0$.

We shall denote the Lorentz-hermitean product of two vectors x and y by

$$\{x, y\} = -x_1^* y_1 - x_2^* y_2 - x_3^* y_3 + x_4^* y_4 . \tag{11}$$

(The star denotes the conjugate imaginary.) If $\{x, x\} < 0$ the vector x is called space-like, if $\{x, x\} = 0$, it is a null vector, if $\{x, x\} > 0$, it is called time-like. A real time-like vector lies in the positive light cone if $x_4 > 0$; it lies in the negative light cone if $x_4 < 0$. Two vectors x and y are called orthogonal if $\{x, y\} = 0$.

On account of its linear character a homogeneous Lorentz transformation is completely defined if Λv is given for four linearly independent vectors $v^{(1)}$, $v^{(2)}$, $v^{(3)}$, $v^{(4)}$.

From (11) and (10) it follows that $\{v, w\} = \{\Lambda v, \Lambda w\}$ for every pair of vectors v, w. This will be satisfied for every pair if it is satisfied for all pairs $v^{(i)}$, $v^{(k)}$

of four linearly independent vectors. The reality condition is satisfied if $(\Lambda v^{(i)})^* = \Lambda(v^{(i)*})$ holds for four such vectors.

The scalar product of two vectors x and y is positive if both lie in the positive light cone or both in the negative light cone. It is negative if one lies in the positive, the other in the negative light cone. Since both x and y are time-like $|x_4|^2 > |x_1|^2 + |x_2|^2 + |x_3|^2$; $|y_4|^2 > |y_1|^2 + |y_2|^2 + |y_3|^2$. Hence, by Schwarz's inequality $|x_4^* y_4| > |x_1^* y_1 + x_2^* y_2 + x_3^* y_3|$ and the sign of the scalar product of two real time-like vectors is determined by the product of their time components.

A time-like vector is transformed by a Lorentz transformation into a time-like vector. Furthermore, on account of the condition $\Lambda_{44} > 0$, the vector $v^{(0)}$ with the components 0, 0, 0, 1 remains in the positive light cone, since the fourth component of $\Lambda v^{(0)}$ is Λ_{44}. If $v^{(1)}$ is another vector[16] in the positive light cone $\{v^{(1)}, v^{(0)}\} > 0$ and hence also $\{\Lambda v^{(1)}, \Lambda v^{(0)}\} > 0$ and $\Lambda v^{(1)}$ is in the positive light cone also. The third condition for a Lorentz transformation can be formulated also as the requirement that every vector in (or on) the positive light cone shall remain in (or, respectively, on) the positive light cone.

This formulation of the third condition shows that the third condition holds for the product of two homogeneous Lorentz transformations if it holds for both factors. The same is evident for the first two conditions.

From $\Lambda F \Lambda' = F$ one obtains by multiplying with Λ^{-1} from the left and $\Lambda'^{-1} = (\Lambda^{-1})'$ from the right $F = \Lambda^{-1} F (\Lambda^{-1})'$ so that the reciprocal of a homogeneous Lorentz transformation is again such a transformation. The homogeneous Lorentz transformations form a group, therefore.

One easily calculates that the product of two inhomogeneous Lorentz transformations (b, M) and (c, N) is again an inhomogeneous Lorentz transformation (a, Λ)

$$(b, M)(c, N) = (a, \Lambda) \tag{12}$$

where

$$\Lambda_{ik} = \sum_j M_{ij} N_{jk}; \qquad a_i = b_i + \sum_j M_{ij} c_j, \tag{12a}$$

or, somewhat shorter

$$\Lambda = MN; \qquad a = b + Mc. \tag{12b}$$

B. Theory of characteristic values and characteristic vectors of a homogeneous Lorentz transformation

Linear homogeneous transformations are most simply described by their characteristic values and vectors. Before doing this for the homogeneous Lorentz group, however, we shall need two rules about orthogonal vectors.

[16] Wherever a confusion between vectors and vector components appears to be possible, upper indices will be used for distinguishing different vectors and lower indices for denoting the components of a vector.

[1] *If* $\{v, w\} = 0$ *and* $\{v, v\} > 0$, *then* $\{w, w\} < 0$; *if* $\{v, w\} = 0$, $\{v, v\} = 0$, *then* w *is either space-like, or parallel to* v (*either* $\{w, w\} < 0$, *or* $w = cv$).

PROOF:

$$v_4^* w_4 = v_1^* w_1 + v_2^* w_2 + v_3^* w_3 . \tag{13}$$

By Schwarz's inequality, then

$$| v_4 |^2 | w_4 |^2 \leqq (| v_1 |^2 + | v_2 |^2 + | v_3 |^2)(| w_1 |^2 + | w_2 |^2 + | w_3 |^2). \tag{14}$$

For $| v_4 |^2 > | v_1 |^2 + | v_2 |^2 + | v_3 |^2$ it follows that $| w_4 |^2 < | w_1 |^2 + | w_2 |^2 + | w_3 |^2$. If $| v_4 |^2 = | v_1 |^2 + | v_2 |^2 + | v_3 |^2$ the second inequality still follows if the inequality sign holds in (14). The equality sign can hold only, however, if the first three components of the vectors v and w are proportional. Then, on account of (13) and both being null vectors, the fourth components are in the same ratio also.

[2] *If four vectors* $v^{(1)}$, $v^{(2)}$, $v^{(3)}$, $v^{(4)}$ *are mutually orthogonal and linearly independent, one of them is time-like, three are space-like.*

PROOF: It follows from the previous paragraph that only one of four mutually orthogonal, linearly independent vectors can be time-like or a null vector. It remains to be shown therefore only that one of them *is* time-like. Since they are linearly independent, it is possible to express by them any time-like vector

$$v^{(t)} = \sum_{k=1}^{4} \alpha_k v^{(k)}.$$

The scalar product of the left side of this equation with itself is positive and therefore

$$\left\{ \sum_k \alpha_k v^{(k)}, \quad \sum_k \alpha_k v^{(k)} \right\} > 0$$

or

$$\sum_k | \alpha_k |^2 \{v^{(k)}, v^{(k)}\} > 0 \tag{15}$$

and one $\{v^{(k)}, v^{(k)}\}$ must be positive. Four mutually orthogonal vectors are not necessarily linearly independent, because a null vector is perpendicular to itself. The linear independence follows, however, if none of the four is a null vector.

We go over now to the characteristic values λ of Λ. These make the determinant $| \Lambda - \lambda 1 |$ of the matrix $\Lambda - \lambda 1$ vanish.

[3] *If* λ *is a characteristic value,* λ^*, λ^{-1} *and* λ^{*-1} *are characteristic values also.*

PROOF: For λ^* this follows from the fact that Λ is real. Furthermore, from $| \Lambda - \lambda 1 | = 0$ also $| \Lambda' - \lambda 1 | = 0$ follows, and this multiplied by the determinants of ΛF and F^{-1} gives

$$| \Lambda F | \cdot | \Lambda' - \lambda 1 | \cdot | F |^{-1} = | \Lambda F \Lambda' F^{-1} - \lambda \Lambda | = | 1 - \lambda \Lambda | = 0,$$

so that λ^{-1} is a characteristic value also.

[4] *The characteristic vectors* v_1 *and* v_2 *belonging to two characteristic values* λ_1 *and* λ_2 *are orthogonal if* $\lambda_1^* \lambda_2 \neq 1$.

PROOF:

$$\{v_1, v_2\} = \{\Lambda v_1, \Lambda v_2\} = \{\lambda_1 v_1, \lambda_2 v_2\} = \lambda_1^* \lambda_2 \{v_1, v_2\}.$$

Thus if $\{v_1, v_2\} \neq 0, \lambda_1^* \lambda_2 = 1$.

[5] *If the modulus of a characteristic value* λ *is* $|\lambda| \neq 1$, *the corresponding characteristic vector* v *is a null vector and* λ *itself real and positive.*

From $\{v, v\} = \{\Lambda v, \Lambda v\} = |\lambda|^2\{v, v\}$ the $\{v, v\} = 0$ follows immediately for $|\lambda| \neq 1$. If λ were complex, λ^* would be a characteristic value also. The characteristic vectors of λ and λ^* would be two different null vectors and, because of [4], orthogonal to each other. This is impossible on account of [1]. Thus λ is real and v a real null vector. Then, on account of the third condition for a homogeneous Lorentz transformation, λ must be positive.

[6] *The characteristic value* λ *of a characteristic vector* v *of length null is real and positive.*

If λ were not real, λ^* would be a characteristic value also. The corresponding characteristic vector v^* would be different from v, a null vector also, and perpendicular to v on account of [4]. This is impossible because of [1].

[7] *The characteristic vector* v *of a complex characteristic value* λ (*the modulus of which is* 1 *on account of* [5]) *is space-like*: $\{v, v\} < 0$.

PROOF: λ^* is a characteristic value also, the corresponding characteristic vector is v^*. Since $(\lambda^*)^*\lambda = \lambda^2 \neq 1$, $\{v^*, v\} = 0$. Since they are different, at least one is space-like. On account of $\{v, v\} = \{v^*, v^*\}$ both are space-like. If all four characteristic values were complex and the corresponding characteristic vectors linearly independent (which is true except if Λ has elementary divisors) we should have four space-like, mutually orthogonal vectors. This is impossible, on account of [2]. Hence

[8] *There is not more than one pair of conjugate complex characteristic values, if* Λ *has no elementary divisors. Similarly, under the same condition, there is not more than one pair* λ, λ^{-1} *of characteristic values whose modulus is different from* 1. Otherwise their characteristic vectors would be orthogonal, which they cannot be, being null vectors.

For homogeneous Lorentz transformations which do not have elementary divisors, the following possibilities remain:

(a) There is a pair of complex characteristic values, their modulus is 1, on account of [5]

$$\lambda_1 = \lambda_2^* = \lambda_2^{-1}; \qquad |\lambda_1| = |\lambda_2| = 1, \tag{16}$$

and also a pair of characteristic values λ_3, λ_4, the modulus of which is not 1. These must be real and positive:

$$\lambda_4 = \lambda_3^{-1}; \qquad \lambda_3 = \lambda_3^* > 0. \tag{16a}$$

The characteristic vectors of the conjugate complex characteristic values are conjugate complex, perpendicular to each other and space-like so that they can be normalized to -1

$$\begin{gathered} v_1 = v_2^*; \qquad \{v_1, v_2\} = \{v_1, v_1^*\} = 0 \\ \{v_1, v_1\} = \{v_2, v_2\} = -1 \end{gathered} \tag{17}$$

those of the real characteristic values are real null vectors, their scalar product can be normalized to 1

$$v_3 = v_3^* \qquad v_4 = v_4^* \qquad \{v_3, v_4\} = 1$$
$$\{v_3, v_3\} = \{v_4, v_4\} = 0. \tag{17a}$$

Finally, the former pair of characteristic vectors is perpendicular to the latter kind

$$\{v_1, v_3\} = \{v_1, v_4\} = \{v_2, v_3\} = \{v_2, v_4\} = 0. \tag{17b}$$

It will turn out that all the other cases in which Λ has no elementary divisor are special cases of (a).

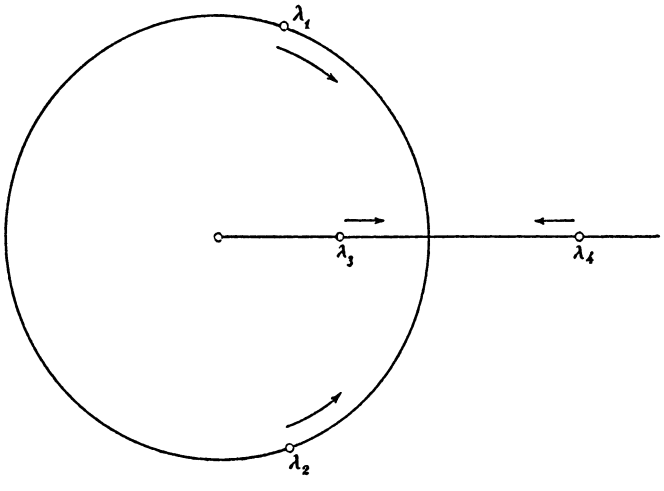


FIG. 1. Position of the characteristic values for the general case a) in the complex plane. In case b), λ_3 and λ_4 coincide and are equal 1; in case c), λ_1 and λ_2 coincide and are either $+1$ or -1. In case d) both pairs $\lambda_3 = \lambda_4 = 1$ and $\lambda_1 = \lambda_2 = \pm 1$ coincide.

(b) There is a pair of complex characteristic values λ_1, $\lambda_2 = \lambda_1^{-1} = \lambda_1^*$, $\lambda_1 \neq \lambda_1^*$, $|\lambda_1| = |\lambda_2| = 1$. No pair with $|\lambda_3| \neq 1$, however. Then on account of [8], still $\lambda_3 = \lambda_3^*$ which gives with $|\lambda_3| = 1$, $\lambda_3 = \pm 1$. Since the product $\lambda_1\lambda_2\lambda_3\lambda_4 = 1$, on account of the second condition for homogeneous Lorentz transformations, also $\lambda_4 = \lambda_3 = \pm 1$. The double characteristic value ± 1 has two linearly independent characteristic vectors v_3 and v_4 which can be assumed to be perpendicular to each other, $\{v_3, v_4\} = 0$. According to [2], one of the four characteristic vectors must be time-like and since those of λ_1 and λ_2 are space-like, the time-like one must belong to ± 1. This must be positive, therefore $\lambda_3 = \lambda_4 = 1$. Out of the time-like and space-like vectors $\{v_3, v_3\} = -1$ and $\{v_4, v_4\} = 1$, one can build two null vectors $v_4 + v_3$ and $v_4 - v_3$. Doing this, case (b) becomes the special case of (a) in which the real positive characteristic values become equal $\lambda_3 = \lambda_4^{-1} = 1$.

(c) All characteristic values are real; there is however one pair $\lambda_3 = \lambda_3^*$,

$\lambda_4 = \lambda_3^{-1}$, the modulus of which is not unity. Then $\{v_3, v_3\} = \{v_4, v_4\} = 0$ and $\lambda_3 > 0$ and one can conclude for λ_1 and λ_2, as before for λ_3 and λ_4 that $\lambda_1 = \lambda_2 = \pm 1$. This again is a special case of (a); here the two characteristic values of modulus 1 become equal.

(d) All characteristic values are real and of modulus 1. If all of them are $+1$, we have the unit matrix which clearly can be considered as a special case of (a). The other case is $\lambda_1 = \lambda_2 = -1$, $\lambda_3 = \lambda_4 = +1$. The characteristic vectors of λ_1 and λ_2 must be space-like, on account of the third condition for a homogeneous Lorentz transformation; they can be assumed to be orthogonal and normalized to -1. This is then a special case of (b) and hence of (a) also. The cases (a), (b), (c), (d) are illustrated in Fig. 1.

The cases remain to be considered in which Λ has an elementary divisor. We set therefore

$$\Lambda_e v_e = \lambda_e v_e\,; \qquad \Lambda_e w_e = \lambda_e w_e + v_e\,. \tag{18}$$

It follows from [5] that either $|\lambda_e| = 1$, or $\{v_e, v_e\} = 0$. We have $\{v_e, w_e\} = \{\Lambda_e v_e, \Lambda_e w_e\} = |\lambda_e|^2\{v_e, w_e\} + \{v_e, v_e\}$. From this equation

$$\{v_e, v_e\} = 0 \tag{19}$$

follows for $|\lambda_e| = 1$, so that (19) holds in any case. It follows then from [6] that λ_e is real, positive and v_e, w_e can be assumed to be real also. The last equation now becomes $\{v_e, w_e\} = \lambda_e^2\{v_e, w_e\}$ so that either $\lambda_e = 1$ or $\{v_e, w_e\} = 0$. Finally, we have

$$\{w_e, w_e\} = \{\Lambda_e w_e, \Lambda_e w_e\} = \lambda_e^2\{w_e, w_e\} + 2\lambda_e\{w_e, v_e\} + \{v_e, v_e\}.$$

This equation now shows that

$$\{w_e, v_e\} = 0 \tag{19a}$$

even if $\lambda_e = 1$. From (19), (19a) it follows that w_e is space-like and can be normalized to

$$\{w_e, w_e\} = -1. \tag{19b}$$

Inserting (19a) into the preceding equation we finally obtain

$$\lambda_e = 1. \tag{19c}$$

[9] *If Λ_e has an elementary divisor, all its characteristic roots are* 1.

From (19c) we see that the root of the elementary divisor is 1 and this is at least a double root. If Λ had a pair of characteristic values $\lambda_1 \neq 1$, $\lambda_2 = \lambda_1^{-1}$, the corresponding characteristic vectors v_1 and v_2 would be orthogonal to v_e and therefore space-like. On account of [5], then $|\lambda_1| = |\lambda_2| = 1$ and $\{v_1, v_2\} = 0$. Furthermore, from $\{w_e, v_1\} = \{\Lambda_e w_e, \Lambda_e v_1\} = \lambda_1\{w_e, v_1\} + \lambda_1\{v_e, v_1\}$ and from $\{v_e, v_1\} = 0$ also $\{w_e, v_1\} = 0$ follows. Thus all the four vectors v_1, v_2, v_e, w_e would be mutually orthogonal. This is excluded by [2] and (19).

Two cases are conceivable now. Either the fourfold characteristic root has only one characteristic vector, or there is in addition to v_e (at least) another characteristic vector v_1. In the former case four linearly independent vectors v_e, w_e, z_e, x_e could be found such that

$$\Lambda_e v_e = v_e \qquad \Lambda_e w_e = w_e + v_e$$
$$\Lambda_e z_e = z_e + w_e \qquad \Lambda_e x_e = x_e + z_e.$$

However $\{v_e, x_e\} = \{\Lambda_e v_e, \Lambda_e x_e\} = \{v_e, x_e\} + \{v_e, z_e\}$ from which $\{v_e, z_e\} = 0$ follows. On the other hand

$$\{w_e, z_e\} = \{\Lambda_e w_e, \Lambda_e z_e\} = \{w_e, z_e\} + \{w_e, w_e\} + \{v_e, z_e\} + \{v_e, w_e\}.$$

This gives with (19a) and (19b) $\{v_e, z_e\} = 1$ so that this case must be excluded.

(e) There is thus a vector v_1 so that in addition to (18)

$$\Lambda_e v_1 = v_1 \tag{18a}$$

holds. From $\{w_e, v_1\} = \{\Lambda_e w_e, \Lambda_e v_1\} = \{w_e, v_1\} + \{v_e, v_1\}$ follows

$$\{v_e, v_1\} = 0. \tag{19d}$$

The equations (18), (18a) will remain unchanged if we add to w_e and v_1 a multiple of v_e. We can achieve in this way that the fourth components of both w_e and v_1 vanish. Furthermore, v_1 can be normalized to -1 and added to w_e also with an arbitrary coefficient, to make it orthogonal to v_1. Hence, we can assume that

$$v_{14} = w_{e4} = 0; \qquad \{v_1, v_1\} = -1; \qquad \{w_e, v_1\} = 0. \tag{19e}$$

We can finally define the null vector z_e to be orthogonal to w_e and v_1 and have a scalar product 1 with v_e

$$\{z_e, z_e\} = \{z_e, w_e\} = \{z_e, v_1\} = 0; \qquad \{z_e, v_e\} = 1. \tag{19f}$$

Then the null vectors v_e and z_e represent the momenta of two light beams in opposite directions. If we set $\Lambda_e z_e = av_e + bw_e + cz_e + dv_1$ the conditions $\{z_e, v\} = \{\Lambda_e z_e, \Lambda_e v\}$ give, if we set for v the vectors v_e, w_e, z_e, v_1 the conditions $c = 1; b = c; 2ac - b^2 - d^2 = 0; d = 0$. Hence

$$\begin{aligned} \Lambda_e v_e &= v_e & \Lambda_e w_e &= w_e + v_e \\ \Lambda_e v_1 &= v_1 & \Lambda_e z_e &= z_e + w_e + \tfrac{1}{2} v_e. \end{aligned} \tag{20}$$

A Lorentz transformation with an elementary divisor can be best characterized by the null vector v_e which is invariant under it and the space part of which forms with the two other vectors w_e and v_1 three mutually orthogonal vectors in ordinary space. The two vectors w_e and v_1 are normalized, v_1 is invariant under Λ_e while the vector v_e is added to w_e upon application of Λ_e. The result of the application of Λ_e to a vector which is linearly independent of v_e, w_e and v_1 is, as we saw, already determined by the expressions for $\Lambda_e v_e$, $\Lambda_e w_e$ and $\Lambda_e v_1$.

The $\Lambda_e(\gamma)$ which have the invariant null vector v_e and also w_e (and hence also

v_1) in common and differ only by adding to w_e different multiples γv_e of v_e, form a cyclic group with $\gamma = 0$, the unit transformation as unity:

$$\Lambda_e(\gamma)\Lambda_e(\gamma') = \Lambda_e(\gamma + \gamma').$$

The Lorentz transformation $M(\alpha)$ which leaves v_1 and w_e invariant but replaces v_e by αv_e (and z_e by $\alpha^{-1}z_e$) has the property of transforming $\Lambda_e(\gamma)$ into

$$M(\alpha)\Lambda_e(\gamma)M(\alpha)^{-1} = \Lambda_e(\alpha\gamma). \qquad (+)$$

An example of $\Lambda_e(\gamma)$ and $M(\alpha)$ is

$$\Lambda_e(\gamma) = \begin{Vmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & \gamma & \gamma \\ 0 & -\gamma & 1 - \frac{1}{2}\gamma^2 & -\frac{1}{2}\gamma^2 \\ 0 & \gamma & \frac{1}{2}\gamma^2 & 1 + \frac{1}{2}\gamma^2 \end{Vmatrix};$$

$$M(\alpha) = \begin{Vmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & \frac{1}{2}(\alpha + \alpha^{-1}) & \frac{1}{2}(\alpha - \alpha^{-1}) \\ 0 & 0 & \frac{1}{2}(\alpha - \alpha^{-1}) & \frac{1}{2}(\alpha + \alpha^{-1}) \end{Vmatrix}.$$

These Lorentz transformations play an important rôle in the representations with space like momentum vectors.

A behavior like (+) is impossible for finite unitary matrices because the characteristic values of $M(\alpha)^{-1}\Lambda_e(\gamma)M(\alpha)$ and $\Lambda_e(\gamma)$ are the same—those of $\Lambda_e(\gamma\alpha) = \Lambda_e(\gamma)^\alpha$ the αth powers of those of $\Lambda_e(\gamma)$. This shows very simply that the Lorentz group has no true unitary representation in a finite number of dimensions.

C. Decomposition of a homogeneous Lorentz transformation into rotations and an acceleration in a given direction

The homogeneous Lorentz transformation is, from the point of view of the physicist, a transformation to a uniformly moving coördinate system, the origin of which coincided at $t = 0$ with the origin of the first coördinate system. One can, therefore, first perform a rotation which brings the direction of motion of the second system into a given direction—say the direction of the third axis—and impart it a velocity in this direction, which will bring it to rest. After this, the two coördinate systems can differ only in a rotation. This means that every homogeneous Lorentz transformation can be decomposed in the following way[17]

$$\Lambda = RZS \qquad (21)$$

[17] Cf. e.g. L. Silberstein, *The Theory of Relativity*, London 1924, p. 142.

where R and S are pure rotations, (i.e. $R_{i4} = R_{4i} = S_{i4} = S_{4i} = 0$ for $i \neq 4$ and $R_{44} = S_{44} = 1$, also $R' = R^{-1}$, $S' = S^{-1}$) and Z is an acceleration in the direction of the third axis, i.e.

$$Z = \begin{Vmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & a & b \\ 0 & 0 & b & a \end{Vmatrix}$$

with $a^2 - b^2 = 1$, $a > b > 0$. The decomposition (21) is clearly not unique. It will be shown, however, that Z is uniquely determined, i.e. the same in every decomposition of the form (21).

In order to prove this mathematically, we chose R so that in $R^{-1}\Lambda = I$ the first two components in the fourth column $I_{14} = I_{24} = 0$ become zero: R^{-1} shall bring the vector with the components Λ_{14}, Λ_{24}, Λ_{34} into the third axis. Then we take $I_{34} = (\Lambda_{14}^2 + \Lambda_{24}^2 + \Lambda_{34}^2)^{\frac{1}{2}}$ and $I_{44} = \Lambda_{44}$ for b and a to form Z; they satisfy the equation $I_{44}^2 - I_{34}^2 = 1$. Hence, the first three components of the fourth column of $J = Z^{-1}I = Z^{-1}R^{-1}\Lambda$ will become zero and $J_{44} = 1$, because of $J_{44}^2 - J_{14}^2 - J_{24}^2 - J_{34}^2 = 1$. Furthermore, the first three components of the fourth row of J will vanish also, on account of $J_{44}^2 - J_{41}^2 - J_{42}^2 - J_{43}^2 = 1$, i.e. $J = S = Z^{-1}R^{-1}\Lambda$ is a pure rotation. This proves the possibility of the decomposition (21).

The trace of $\Lambda\Lambda' = RZ^2R^{-1}$ is equal to the trace of Z^2, i.e. equal to $2a^2 + 2b^2 + 2 = 4a^2 = 4b^2 + 4$ which shows that the a and b of Z are uniquely determined. In particular $a = 1$, $b = 0$ and Z the unit matrix if $\Lambda\Lambda' = 1$, i.e. Λ a pure rotation.

It is easy to show now that the group space of the homogeneous Lorentz transformations is only doubly connected. If a continuous series $\Lambda(t)$ of homogeneous Lorentz transformations is given, which is unity both for $t = 0$ and for $t = 1$, we can decompose it according to (21)

$$\Lambda(t) = R(t)Z(t)S(t). \tag{21a}$$

It is also clear from the foregoing, that $R(t)$ can be assumed to be continuous in t, except for values of t, for which $\Lambda_{14} = \Lambda_{24} = \Lambda_{34} = 0$, i.e. for which Λ is a pure rotation. Similarly, $Z(t)$ will be continuous in t and this will hold even where $\Lambda(t)$ is a pure rotation. Finally, $S = Z^{-1}R^{-1}\Lambda$ will be continuous also, except where $\Lambda(t)$ is a pure rotation.

Let us consider now the series of Lorentz transformations

$$\Lambda_s(t) = R(t)Z(t)^sS(t) \tag{21b}$$

where the b of $Z(t)^s$ is s times the b of $Z(t)$. By decreasing s from 1 to 0 we continuously deform the set $\Lambda_1(t) = \Lambda(t)$ of Lorentz transformations into a set of rotations $\Lambda_0(t) = R(t)S(t)$. Both the beginning $\Lambda_0(0) = 1$ and the end $\Lambda_s(1) = 1$ of the set remain the unit matrix and the sets $\Lambda_s(t)$ remain continuous in t for

all values of s. This last fact is evident for such t for which $\Lambda(t)$ is not a rotation: for such t all factors of (21b) are continuous. But it is true also for t_0 for which $\Lambda(t_0)$ is a rotation, and for which, hence $Z(t_0) = 1$ and $\Lambda_s(t_0) = \Lambda_1(t_0) = \Lambda(t_0)$. As $Z(t)$ is everywhere continuous, there will be a neighborhood of t_0 in which $Z(t)$ and hence also $Z(t)^s$ is arbitrarily close to the unit matrix. In this neighborhood $\Lambda_s(t) = \Lambda(t)$. $S(t)^{-1}Z(t)^{-1}Z(t)^s\ S(t)$ is arbitrarily close to $\Lambda(t)$; and, if the neighborhood is small enough, this is arbitrarily close to $\Lambda(t_0) = \Lambda_s(t_0)$.

Thus (21b) replaces the continuous set $\Lambda(t)$ of Lorentz transformations by a continuous set of rotations. Since these form an only doubly connected manifold, the manifold of Lorentz transformations can not be more than doubly connected. The existence of a two valued representation[18] shows that it is actually doubly and not simply connected.

We can form a new group[14] from the Lorentz group, the elements of which are the elements of the Lorentz group, together with a way $\Lambda(t)$, connecting $\Lambda(1) = \Lambda$ with the unity $\Lambda(0) = E$. However, two ways which can be continuously deformed into each other are not considered different. The product of the element "Λ with the way $\Lambda(t)$" with the element "I with the way $I(t)$" is the element ΛI with the way which goes from E along $\Lambda(t)$ to Λ and hence along $\Lambda I(t)$ to ΛI. Clearly, the Lorentz group is isomorphic with this group and two elements (corresponding to the two essentially different ways to Λ) of this group correspond to one element of the Lorentz group. It is well known,[18] that this group is holomorphic with the group of unimodular complex two dimensional transformations.

Every continuous representation of the Lorentz group "up to the sign" is a singlevalued, continuous representation of this group. The transformation which corresponds to "Λ with the way $\Lambda(t)$" is that $d(\Lambda)$ which is obtained by going over from $d(E) = d(\Lambda(0)) = 1$ continuously along $d(\Lambda(t))$ to $d(\Lambda(1)) = d(\Lambda)$.

D. The homogeneous Lorentz group is simple

It will be shown, first, that an invariant subgroup of the homogeneous Lorentz group contains a rotation (i.e. a transformation which leaves x_4 invariant).— We can write an arbitrary element of the invariant subgroup in the form RZS of (21). From its presence in the invariant subgroup follows that of $S \cdot RZS \cdot S^{-1} = SRZ = TZ$. If X_π is the rotation by π about the first axis, $X_\pi Z X_\pi = Z^{-1}$ and $X_\pi TZX_\pi^{-1} = X_\pi TX_\pi X_\pi ZX_\pi = X_\pi TX_\pi Z^{-1}$ is contained in the invariant subgroup also and thus the transform of this with Z, i.e. $Z^{-1}X_\pi TX_\pi$ also. The product of this with TZ is $TX_\pi TX_\pi$ which leaves x_4 invariant. If $TX_\pi TX_\pi = 1$ we can take $TY_\pi TY_\pi$. If this is the unity also, $TX_\pi TX_\pi = TY_\pi TY_\pi$ and T commutes with $X_\pi Y_\pi$, i.e. is a rotation about the third axis. In this case the

[18] Cf. H. Weyl, *Gruppentheorie und Quantenmechanik*, 1st. ed. Leipzig 1928, pages 110–114, 2nd ed. Leipzig 1931, pages 130–133. It may be interesting to remark that essentially the same isomorphism has been recognized already by L. Silberstein, l.c. pages 148–157.

space like (complex) characteristic vectors of TZ lie in the plane of the first two coördinate axes. Transforming TZ by an acceleration in the direction of the first coördinate axis we obtain a new element of the invariant subgroup for which the space like characteristic vector will have a not vanishing fourth component. Taking this for RZS we can transform it with S again to obtain a new $SRZ = TZ$. However, since S leaves x_4 invariant, the fourth component of the space like characteristic vectors of this TZ will not vanish and we can obtain from it by the procedure just described a rotation which must be contained in the invariant subgroup.

It remains to be shown that an invariant subgroup which contains a rotation, contains the whole homogeneous Lorentz group. Since the three-dimensional rotation group is simple, all rotations must be contained in the invariant subgroup. Thus the rotation by π around the first axis X_π and also its transform with Z and also

$$ZX_\pi Z^{-1} \cdot X_\pi = Z \cdot X_\pi Z^{-1} X_\pi = Z^2$$

is contained in the invariant subgroup. However, the general acceleration in the direction of the third axis can be written in this form. As all rotations are contained in the invariant subgroup also, (21) shows that this holds for all elements of the homogeneous Lorentz group.

It follows from this that the homogeneous Lorentz group has apart from the representation with unit matrices only true representations. It follows then from the remark at the end of part B, that these have all infinite dimensions. This holds even for the two-valued representations to which we shall be led in Section 5 equ. (52D), as the group elements to which the positive or negative unit matrix corresponds must form an invariant subgroup also, and because the argument at the end of part B holds for two-valued representations also. One easily sees furthermore from the equations (52B), (52C) that it holds for the inhomogeneous Lorentz group equally well.

5. Reduction of Representations Up to a Factor to Two-Valued Representations

The reduction will be effected by giving each unitary transformation, which is defined by the physical content of the theory and the consideration of reference only up to a factor of modulus unity, a "phase," which will leave only the sign of the representation operators undetermined. The unitary operator corresponding to the translation a will be denoted by $T(a)$, that to the homogeneous Lorentz transformation Λ by $d(\Lambda)$. To the general inhomogeneous Lorentz transformation then $D(a, \Lambda) = T(a)d(\Lambda)$ will correspond. Instead of the relations (12), we shall use the following ones.

$$T(a)T(b) = \omega(a, b)T(a + b) \tag{22B}$$

$$d(\Lambda)T(a) = \omega(\Lambda, a)T(\Lambda a)d(\Lambda) \tag{22C}$$

$$d(\Lambda)d(I) = \omega(\Lambda, I)d(\Lambda I). \tag{22D}$$

The ω are numbers of modulus 1. They enter because the multiplication rules (12) hold for the representatives only up to a factor. Otherwise, the relations (22) are consequences of (12) and can in their return replace (12). We shall replace the $T(a)$, $d(\Lambda)$ by $\Omega(a)T(a)$ and $\Omega(\Lambda)d(\Lambda)$ respectively, for which equations similar to (22) hold, however with

$$\omega(a, b) = 1; \qquad \omega(\Lambda, a) = 1; \qquad \omega(\Lambda, I) = \pm 1. \tag{22'}$$

A.

It is necessary, first, to show that the undetermined factors in the representation $D(L)$ can be assumed in such a way that the $\omega(a, b)$, $\omega(\Lambda, a)$, $\omega(\Lambda, I)$ become —apart from regions of lower dimensionality—continuous functions of their arguments. This is a consequence of the continuous character of the representation and shall be discussed first.

(a) From the point of view of the physicist, the natural definition of the continuity of a representation up to a factor is as follows. The neighborhood δ of a Lorentz transformation $L_0 = (b, I)$ shall contain all the transformations $L = (a, \Lambda)$ for which $| a_\kappa - b_\kappa | < \delta$ and $| \Lambda_{ik} - I_{ik} | < \delta$. The representation up to a factor $D(L)$ is continuous if there is to every positive number ϵ, every normalized wave function φ and every Lorentz transformation L_0 such a neighborhood δ of L_0 that for every L of this neighborhood one can find an Ω of modulus 1 (the Ω depending on L and φ) such that $(u_\varphi, u_\varphi) < \epsilon$ where

$$u_\varphi = (D(L_0) - \Omega D(L))\varphi. \tag{23}$$

Let us now take a point L_0 in the group space and find a normalized wave function φ for which $| (\varphi, D(L_0)\varphi) | > 1/6$. There always exists a φ with this property, if $|(\varphi, D(L_0)\varphi)| < 1/6$ then $\psi = \alpha\varphi + \beta D(L_0)\varphi$ with suitably chosen α and β will be normalized and $| (\psi, D(L_0)\psi) | > 1/6$. We consider then such a neighborhood $\mathfrak{N}$ of L_0 for all L of which $| (\varphi, D(L)\varphi) | > 1/12$. It is well known[19] that the whole group space can be covered with such neighborhoods. We want to show now that the $D(L)\varphi$ can be multiplied with such phase factors (depending on L) of modulus unity that it becomes strongly continuous in the region $\mathfrak{N}$.

We shall chose that phase factor so that $(\varphi, D(L)\varphi)$ becomes real and positive. Denoting then

$$(D(L_1) - D(L))\varphi = U_\varphi, \tag{23'}$$

the (U_φ, U_φ) can be made arbitrarily small by letting L approach sufficiently near to L_1, if L_1 is in $\mathfrak{N}$. Indeed, on account of the continuity, as defined above, there is an $\Omega = e^{i\kappa}$ such that $(u, u) < \epsilon$ if L is sufficiently near to L_1 where

$$u = (D(L_1) - e^{i\kappa}D(L))\varphi.$$

[19] This condition is the "separability" of the group. Cf. e.g. A. Haar, Ann. of Math., 34, 147, 1933.

Taking the absolute value of the scalar product of u with φ one obtains

$$|(\varphi, D(L_1)\varphi) - \cos\kappa(\varphi, D(L)\varphi) - i\sin\kappa(\varphi, D(L)\varphi)| = |(\varphi, u)| \leqq \sqrt{\epsilon},$$

because of Schwartz's inequality. If only $\sqrt{\epsilon} < 1/12$, the κ must be smaller than $\pi/2$ because the absolute value is certainly greater than the real part, and both $(\varphi, D(L_1)\varphi)$ and $(\varphi, D(L)\varphi)$ are real and greater than $1/12$.

As the absolute value is also greater than the imaginary part, we

$$\sin\kappa < 12\sqrt{\epsilon}.$$

On the other hand,

$$U_\varphi = u + (e^{i\kappa} - 1)D(L)\varphi,$$

and thus

$$(U_\varphi, U_\varphi)^{\frac{1}{2}} \leqq (u, u)^{\frac{1}{2}} + |e^{i\kappa} - 1| \leqq \sqrt{\epsilon} + 2\sin\kappa/2$$

$$(U_\varphi, U_\varphi) \leqq 625\,\epsilon.$$

(b) It shall be shown next that if $D(L)\varphi$ is strongly continuous in a region and $D(L)$ is continuous in the sense defined at the beginning of this section, then $D(L)\psi$ with an arbitrary ψ is (strongly) continuous in that region also. We shall see, hence, that the $D(L)$, with any normalization which makes a $D(L)\varphi$ strongly continuous, is continuous in the ordinary sense: There is to every L_1, ϵ and *every* ψ a δ so that $(U_\psi, U_\psi) < \epsilon$ where

$$U_\psi = (D(L_1) - D(L))\psi$$

if L is in the neighborhood δ of L_1.

It is sufficient to show the continuity of $D(L)\psi$ where ψ is orthogonal to φ. Indeed, every ψ' can be decomposed into two terms, $\psi' = \alpha\varphi + \beta\psi$ the one of which is parallel, the other perpendicular to φ. Since $D(L)\varphi$ is continuous, according to supposition, $D(L)\psi' = \alpha D(L)\varphi + \beta D(L)\psi$ will be continuous also if $D(L)\psi$ is continuous.

The continuity of the representation up to a factor requires that it is possible to achieve that $(u_\psi, u_\psi) < \epsilon$ and $(u_{\psi+\varphi}, u_{\psi+\varphi}) < \epsilon$ where

$$u_\psi = (D(L_1) - \Omega_\psi D(L))\psi, \tag{23a}$$

$$u_{\psi+\varphi} = (D(L_1) - \Omega_{\psi+\varphi}D(L))(\psi + \varphi), \tag{23b}$$

with suitably chosen Ω's. According to the foregoing, it also is possible to choose L and L_1 so close that $(U_\varphi, U_\varphi) < \epsilon$.

Subtracting (23′) and (23a) from (23b) and applying $D(L)^{-1}$ on both sides gives

$$(\Omega_\psi - \Omega_{\psi+\varphi})\psi + (1 - \Omega_{\psi+\varphi})\varphi = D(L)^{-1}(u_{\psi+\varphi} - u_\psi - U_\varphi)$$

The scalar product of the right side with itself is less than 9ϵ. Hence both $|\Omega_\psi - \Omega_{\psi+\varphi}| < 3\epsilon^{\frac{1}{2}}$ and $|1 - \Omega_{\psi+\varphi}| < 3\epsilon^{\frac{1}{2}}$ or $|1 - \Omega_\psi| < 6\epsilon^{\frac{1}{2}}$. Because of $U_\psi = u_\psi - (1 - \Omega_\psi)D(L)\psi$, the $(U_\psi, U_\psi)^{\frac{1}{2}} < (u_\psi, u_\psi)^{\frac{1}{2}} + |1 - \Omega_\psi|$ and thus $(U_\psi, U_\psi) < 49\epsilon$.

This completes the proof of the theorem stated under (b). It also shows that not only the continuity of $D(L)\varphi$ has been achieved in the neighborhood of L_0 by the normalization used in (a) but also that of $D(L)\psi$ with every ψ, i.e., the continuity of $D(L)$.

It is clear also that every finite part of the group space can be covered by a finite number of neighborhoods in which $D(L)$ can be made continuous. It is easy to see that the ω of (22) will be also continuous in these neighborhoods so that is is possible to make them continuous, apart from regions of lower dimensionality than their variables have. In the following only the fact will be used that they can be made continuous in the neighborhood of any a, b and Λ.

B.

(a) We want to show next that all $T(a)$ commute. From (22B) we have

$$T(a)T(b)T(a)^{-1} = c(a, b)T(b) \tag{24}$$

where $c(a, b) = \omega(a, b)/\omega(b, a)$ and hence

$$c(a, b) = c(b, a)^{-1}. \tag{24a}$$

Transforming (24) with $T(a')$ one obtains

$$T(a')T(a)T(b)T(a)^{-1}T(a')^{-1} = c(a, b)T(a')T(b)T(a')^{-1}$$

or $$\omega(a', a)T(a' + a)T(b)\omega(a', a)^{-1}T(a' + a)^{-1} = c(a, b)c(a', b)T(b)$$

or

$$c(a, b)c(a', b) = c(a + a', b). \tag{25}$$

It follows[20] from (25) and the partial continuity of $c(a, b)$ that

$$c(a, b) = \exp\left(2\pi i \sum_{\kappa=1}^{4} a_\kappa f_\kappa(b)\right) \tag{26}$$

and, since this is equal to $c(b, a)^{-1} = \exp(-2\pi i \sum b_\kappa f_\kappa(a))$

$$\sum_{\kappa=1}^{4} (a_\kappa f_\kappa(b) + b_\kappa f_\kappa(a)) = n(a, b) \tag{27}$$

where $n(a, b)$ is an integer. Setting in (27) for b the vector $e^{(\lambda)}$ the λ component of which is 1, all the others zero and for $f_\kappa(e^{(\lambda)}) = -f_{\kappa\lambda}$

$$f_\lambda(a) = n(a, e^{(\lambda)}) + \sum_\kappa a_\kappa f_{\kappa\lambda},$$

and putting this back into (27) we obtain

$$\sum_{\kappa,\lambda=1}^{4} f_{\kappa\lambda}(a_\lambda b_\kappa + b_\lambda a_\kappa) + \sum_{\kappa=1}^{4} a_\kappa n(b, e^{(\kappa)}) + b_\kappa n(a, e^{(\kappa)}) = n(a, b). \tag{28}$$

[20] G. Hamel, Math. Ann. 60, 460, 1905, quoted from H. Hahn, *Theorie der reellen Funktionen*. Berlin 1921, pages 581–583.

Assuming for the components of a and b such values which are transcendental both with respect to each other and the $f_{\kappa\lambda}$ (which are fixed numbers), one sees that (28) cannot hold except if the coefficient of every one vanishes

$$f_{\kappa\lambda} + f_{\lambda\kappa} = 0; \qquad n(b, e^{(\kappa)}) = 0, \tag{29}$$

so that (26) becomes

$$c(a, b) = \exp\left(2\pi i \sum_{\kappa,\lambda=1}^{4} f_{\kappa\lambda} a_\lambda b_\kappa\right). \tag{30}$$

It is necessary now to consider the existence of an operator $d(\Lambda)$ satisfying (22C). Transforming this equation with the similar equation containing b instead of a

$$d(\Lambda)T(b)d(\Lambda)^{-1}d(\Lambda)T(a)d(\Lambda)^{-1}d(\Lambda)T(b)^{-1}d(\Lambda)^{-1}$$
$$= \omega(\Lambda, b)T(\Lambda b)\omega(\Lambda, a)T(\Lambda a)\omega(\Lambda, b)^{-1}T(\Lambda b)^{-1} = \omega(\Lambda, a)c(\Lambda b, \Lambda a)T(\Lambda a),$$

while the first line is clearly $d(\Lambda)c(b, a)T(a)d(\Lambda)^{-1} = \omega(\Lambda, a)c(b, a)T(\Lambda a)$ whence

$$c(b, a) = c(\Lambda b, \Lambda a) \tag{31}$$

holds for every Lorentz transformation Λ. Combined with (30) this gives

$$\sum_{\kappa\lambda}\left(f_{\kappa\lambda} a_\kappa b_\lambda - \sum_{\nu\mu} f_{\nu\mu}\Lambda_{\nu\kappa}\Lambda_{\mu\lambda} a_\kappa b_\lambda\right) = n'(a, b),$$

where $n'(a, b)$ is again an integer. As this equation holds for every a, b

$$f_{\kappa\lambda} = \sum_{\nu\mu} f_{\nu\mu}\Lambda_{\nu\kappa}\Lambda_{\mu\lambda}; \qquad f = \Lambda' f \Lambda$$

must hold also, for every Lorentz transformation. However, the only form invariant under all Lorentz transformations are multiples of the F of (10). Actually, because of (29), f must vanish and $c(a, b) = 1$, all the operators corresponding to translations commute

$$T(a)T(b) = T(b)T(a). \tag{32}$$

It is well to remember that it was necessary for obtaining this result to use the existence of $d(\Lambda)$ satisfying (22C).

(b) Equation (32) is clearly independent of the normalization of the $T(a)$. If we could fix the translation operators in four linearly independent directions $e^{(1)}$, $e^{(2)}$, $e^{(3)}$, $e^{(4)}$ so that for each of these directions

$$T(ae^{(\kappa)})T(be^{(\kappa)}) = T((a + b)e^{(\kappa)}) \tag{33}$$

be valid for every pair of numbers a, b, then the normalization

$$T(a_1e^{(1)} + a_2e^{(2)} + a_3e^{(3)} + a_4e^{(4)}) = T(a_1e^{(1)})T(a_2e^{(2)})T(a_3e^{(3)})T(a_4e^{(4)}) \tag{33a}$$

and (32) would ensure the general validity of

$$T(a)T(b) = T(a + b). \tag{34}$$

As the four linearly independent directions $e^{(1)}, \cdots, e^{(4)}$ we shall take four null vectors. If e is a null vector, there is, according to section 3, a homogeneous Lorentz transformation[21] Λ_e such that $\Lambda_e e = 2e$.

We normalize $T(e)$ so that

$$d(\Lambda_e)T(e)d(\Lambda_e)^{-1} = T(e)^2. \tag{35}$$

This is clearly independent of the normalization of $d(\Lambda_e)$. We further normalize for all (positive and negative) integers n

$$d(\Lambda_e)^n T(e) d(\Lambda_e)^{-n} = T(2^n e). \tag{35a}$$

It follows from this equation also that

$$T(2^n e)^2 = d(\Lambda_e)^n T(e)^2 d(\Lambda_e)^{-n} = d(\Lambda_e)^n d(\Lambda_e) T(e) d(\Lambda_e)^{-1} d(\Lambda_e)^{-n} = T(2^{n+1}e). \tag{36}$$

This allows us to normalize for every positive integer k

$$T(k \cdot 2^{-n} e) = T(2^{-n} e)^k \tag{35b}$$

in such a way that the normalization remains the same if we replace k by $2^m k$ and n by $n + m$. This ensures, together with (36), the validity of

$$\begin{aligned} T(\nu e)T(\mu e) &= T((\nu + \mu)e) \\ d(\Lambda_e)T(\nu e)d(\Lambda_e)^{-1} &= T(2\nu e) \end{aligned} \tag{36a}$$

for all dyadic fractions ν and μ.

It must be shown that if $\nu_1, \nu_2, \nu_3, \cdots$ is a sequence of dyadic fractions, converging to 0, $\lim T(\nu_i e) = 1$. From $T(a) \cdot T(0) = \omega(a, 0)T(a)$ it follows that $T(0)$ is a constant. According to the theorem of part (A)(b), the $T(\nu e)$, if multiplied by proper constants Ω_ν will converge to 1, i.e., by choosing an arbitrary φ, it is possible to make both $(1 - \Omega_\nu T(\nu e))\varphi = u$ and $(1 - \Omega_\nu T(\nu e)) \cdot d(\Lambda_e)^{-1}\varphi = u'$ arbitrarily small, by making ν small. Applying $d(\Lambda_e)$ to the second expression, one obtains, for (36a), that $(1 - \Omega_\nu T(2\nu e))\varphi = d(\Lambda_e)u'$ is also small. On the other hand, applying $T(\nu e)$ to the first expression one sees that $(T(\nu e) - \Omega_\nu T(2\nu e))\varphi = T(\nu e)u$ approaches zero also. Hence, the difference of these two quantities $(1 - T(\nu e))\varphi$ goes to zero, i.e. $T(\nu_i e)\varphi$ converges to φ if $\nu_1, \nu_2, \nu_3, \cdots$ is a sequence of dyadic fractions approaching 0.

Now $\nu_1, \nu_2, \nu_3, \cdots$ be a sequence of dyadic fractions converging to an arbitrary number a. It will be shown then that $T(\nu_i e)$ converges to a multiple of $T(ae)$ and this multiple of $T(ae)$ will be the normalized $T(ae)$. Again, it follows from the continuity that there are such Ω_i that $\Omega_i T(\nu_i e)\varphi$ converges to $T(ae)\varphi$. The $\Omega_j^{-1} T(\nu_j e)^{-1} \Omega_i T(\nu_i e)\varphi$ will converge to φ, therefore, as both i and j tend to infinity. However, according to the previous paragraph, $T((\nu_i - \nu_j)e)\varphi$ tends to φ and thus $\Omega_j^{-1}\Omega_i$ tends to 1. It follows that Ω_i^{-1} converges to a definite number Ω. Hence $\Omega_i^{-1} \cdot \Omega_i T(\nu_i e)\varphi$ converges to $\Omega T(ae)\varphi$ which will be denoted, henceforth, by $T(ae)$. For the $T(ae)$, normalized in this way, (33) will hold,

[21] The index e denotes here the vector e for which $\Lambda_e e = 2e$; this Λ_e has no elementary divisor.

since if $\mu_1, \mu_2, \mu_3, \cdots$ are dyadic fractions converging to b, we obtain, with the help (36a)

$$T(ae)T(be)\varphi = \lim_{i,j=\infty} T((\nu_i + \mu_j)e)\varphi = T((a+b)e)\varphi.$$

This argument not only shows that it is possible to normalize the $T(ae^{(k)})$ and hence by (33a) the $T(a)$ so that (34) holds for them but, in addition to this, that these $T(a)$ will be continuous in the ordinary sense.

C.

It is clear that (34) will remain valid if one replaces $T(a)$ by $\exp(2\pi i\{a, c\})T(a)$ where c is an arbitrary vector. This remaining freedom in the normalization of $T(a)$ will be used to eliminate the $\omega(\Lambda, a)$ from (22C).

Transforming (22C) $d(\Lambda)T(a)d(\Lambda)^{-1} = \omega(\Lambda, a)T(\Lambda a)$ with $d(M)$ one obtains on the left side $\omega(M, \Lambda)d(M\Lambda)T(a)\omega(M, \Lambda)^{-1}\, d(M\Lambda)^{-1} = \omega(M\Lambda, a)T(M\Lambda a)$ while the right side becomes $\omega(\Lambda, a)\omega(M, \Lambda a)T(M\Lambda a)$. Hence

$$\omega(M\Lambda, a) = \omega(M, \Lambda a)\omega(\Lambda, a). \tag{37}$$

On the other hand, the product of two equations (22C) with the same Λ but with a and b respectively, instead of a yields with the help of (34)

$$\omega(\Lambda, a)\omega(\Lambda, b) = \omega(\Lambda(a+b)).$$

Hence

$$\omega(\Lambda, a) = \exp(2\pi i\{a, f(\Lambda)\}),$$

where $f(\Lambda)$ is a vector which can depend on Λ. Inserting this back into (37) one obtains

$$\{a, f(M\Lambda)\} = \{\Lambda a, f(M)\} + \{a, f(\Lambda)\} + n,$$

$$\{a, f(M\Lambda) - \Lambda^{-1}f(M) - f(\Lambda)\} = n,$$

where n is an integer which must vanish since it is a linear function of a. Hence

$$f(M\Lambda) = \Lambda^{-1}f(M) + f(\Lambda). \tag{38}$$

If we can show that the most general solution of the equation is

$$f(\Lambda) = (\Lambda^{-1} - 1)v_0, \tag{39}$$

where v_0 is a vector independent of Λ, the $\omega(\Lambda, a)$ will become $\omega(\Lambda, a) = \exp(2\pi i\{(\Lambda - 1)a, v_0\})$. Then $\omega(\Lambda, a)$ in (22C) will disappear if we replace $T(a)$ by $\exp(2\pi i\{a, v_0\})T(a)$.

The proof that (39) is a consequence of (38) is somewhat laborious. One can first consider the following homogeneous Lorentz transformations

$$
(40)\qquad
X(\alpha_1, \gamma_1) = \begin{Vmatrix} C_1 & 0 & 0 & S_1 \\ 0 & c_1 & s_1 & 0 \\ 0 & -s_1 & c_1 & 0 \\ S_1 & 0 & 0 & C_1 \end{Vmatrix}; \qquad
Y(\alpha_2, \gamma_2) = \begin{Vmatrix} c_2 & 0 & -s_2 & 0 \\ 0 & C_2 & 0 & S_2 \\ s_2 & 0 & c_2 & 0 \\ 0 & S_2 & 0 & C_2 \end{Vmatrix}
$$

$$
Z(\alpha_3, \gamma_3) = \begin{Vmatrix} c_3 & s_3 & 0 & 0 \\ -s_3 & c_3 & 0 & 0 \\ 0 & 0 & C_3 & S_3 \\ 0 & 0 & S_3 & C_3 \end{Vmatrix}
$$

where $c_i = \cos \alpha_i$; $s_i = \sin \alpha_i$; $C_i = \mathrm{Ch}\gamma_i$; $S_i = \mathrm{Sh}\gamma_i$. All the $X(\alpha, \gamma)$ commute. Let us choose, therefore, two angles α_1, γ_1 for which $1 - X(\alpha_1, \gamma_1)^{-1}$ has a reciprocal. It follows then from (38)

$$
X(\alpha, \gamma)^{-1} f(X(\alpha_1, \gamma_1)) + f(X(\alpha, \gamma)) = X(\alpha_1, \gamma_1)^{-1} f(X(\alpha, \gamma)) + f(X(\alpha_1, \gamma_1))
$$

$$
(41)\quad \text{or}\quad f(X(\alpha, \gamma)) = [1 - X(\alpha_1, \gamma_1)^{-1}]^{-1}[1 - X(\alpha, \gamma)^{-1}] f(X(\alpha_1, \gamma_1))
$$

$$
f(X(\alpha, \gamma)) = (1 - X(\alpha, \gamma)^{-1}) v_X ,
$$

where v_X is independent of α, γ. Similar equations hold for the $f(Y(\alpha, \gamma))$ and $f(Z(\alpha, \gamma))$. Let us denote now $X(\pi, 0) = X$; $Y(\pi, 0) = Y$; $Z(\pi, 0) = Z$. These anticommute in the following sense with the transformations (40):

$$
(42)\qquad YX(\alpha, \gamma)Y = ZX(\alpha, \gamma)Z = X(\alpha, \gamma)^{-1}.
$$

From (38) one easily calculates

$$
f(YX(\alpha, \gamma)Y) = (YX(\alpha, \gamma)^{-1} + 1) f(Y) + Y f(X(\alpha, \gamma)),
$$

or, because of (41) and (42), after some trivial transformations

$$
(43)\qquad (1 - X(\alpha, \gamma))(1 - Y)(v_X - v_Y) = 0.
$$

As α, γ can be taken arbitrarily, the first factor can be dropped. This leaves $(1 - Y)(v_X - v_Y) = 0$, or that the first and third components of v_X and v_Y are equal. One similarly concludes, however, that $(1 - X)(v_Y - v_X) = 0$ and thus that the first three components of v_X, v_Y and also of v_Z are equal.

For $\gamma_1 = \gamma_2 = \gamma_3 = 0$ the transformations (40) are the generators of all rotations, i.e. all Lorentz transformations R not affecting the fourth coördinate. As the 4-4 matrix element of these transformations is 1, the expression $(1 - R^{-1})v$ is independent of the fourth component of v and $(1 - R^{-1})v_X = (1 - R^{-1})v_Y = (1 - R^{-1})v_Z$. It follows from (38) that if $f(R) = (1 - R^{-1})v_X$ and $f(S) = (1 - S^{-1})v_X$, then $f(SR) = (1 - R^{-1}S^{-1})v_X$. Thus $f(R) = (1 - R^{-1})v_X$ is valid with the same v_X for all rotations.

Now

$$
f(X(\alpha, \gamma)R) = R^{-1}(1 - X(\alpha, \gamma)^{-1})v_X + (1 - R^{-1})v_X = (1 - (X(\alpha, \gamma)R)^{-1})v_X .
$$

One easily concludes from (38) that the $f(E)$ corresponding to the unit operation vanishes and $f(\Lambda^{-1}) = -\Lambda f(\Lambda)$. Hence $f(R^{-1}X(\alpha, \gamma)^{-1}) = (1 - X(\alpha, \gamma)R)v_X$; and one concludes further that for all Lorentz transformations $\Lambda = RX(\alpha, \gamma)S$, (39) holds with $v_0 = -v_X$ if R and S are rotations. However, every homogeneous Lorentz transformation can be brought into this form (Section 4C). This completes the proof of (39) and thus of $\omega(\Lambda, a) = 1$.

D.

The quantities $\omega(a, b)$ and $\omega(\Lambda, a)$ for which it has just been shown that they can be assumed to be 1, are independent from the normalization of $d(\Lambda)$. We can affix therefore an arbitrary factor of modulus 1 to all the $d(\Lambda)$, without interfering with the normalizations so far accomplished. In consequence hereof, the ensuing discussion will be simply a discussion of the normalization

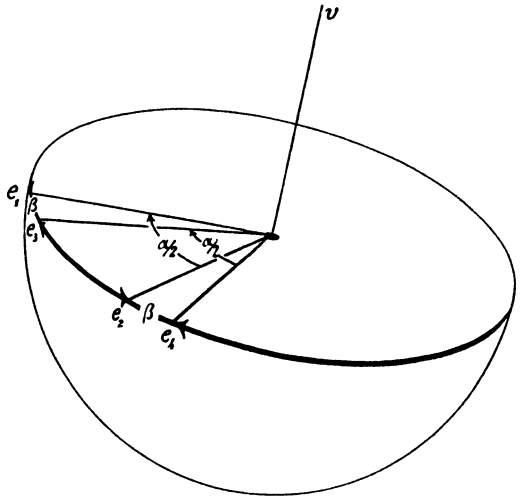


FIG. 2

of the operators for the homogeneous Lorentz group and the result to be obtained will be valid for that group also.

Partly because the representations up to a factor of the three dimensional rotation group may be interesting in themselves, but more particularly because the procedure to be followed for the Lorentz group can be especially simply demonstrated for this group, the three dimensional rotation group shall be taken up first.

It is well known that the normalization cannot be carried so far that $\omega(\Lambda, I) = 1$ in (22D) and there are well known representations for which $\omega(\Lambda, I) = \pm 1$. We shall allow this ambiguity therefore from the outset.

One can observe, first, that the operator corresponding to the unity of the group is a constant. This follows simply from $d(\Lambda)d(E) = \omega(\Lambda, E)d(\Lambda)$. The square of an operator corresponding to an involution is a constant, therefore.

The operator corresponding to the rotation about the axis e by the angle π, normalized so that its square be actually 1, will be denoted by $\tilde{e}$; $\tilde{e}^2 = 1$. The $\tilde{e}$ are—apart from the sign—uniquely defined.

A rotation R about v by the angle α is the product of two rotations by π about e_1 and e_2 where e_1 and e_2 are perpendicular to v and e_2 arises from e_1 by rotation about v with $\alpha/2$. Choosing for every v an arbitrary e_1 perpendicular to v, we can normalize, therefore

$$d(R) = \pm\tilde{e}_1\tilde{e}_2 . \tag{44}$$

Now $d(R)$ commutes with every $d(S)$ if S is also a rotation about v. This is proved in equations (24)–(30). The f_{11} in (30) must vanish on account of (29).

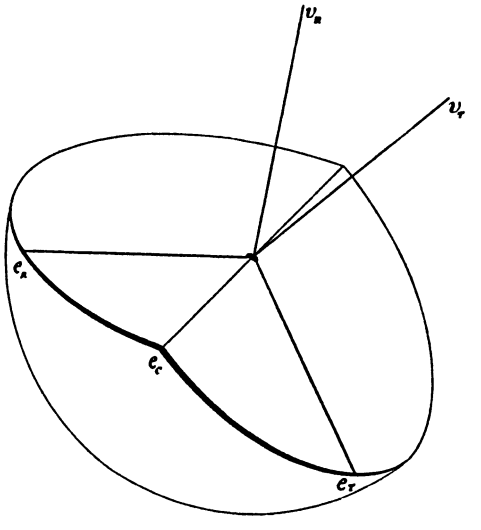


FIG. 3

(Also, both R and S can be arbitrarily accurately represented as powers of a very small rotation about v). Hence, transforming (44) by $d(S)$ one obtains

$$d(R) = \pm d(S)\tilde{e}_1 d(S)^{-1}\cdot d(S)\tilde{e}_2 d(S)^{-1}. \tag{44a}$$

Now $d(S)\tilde{e}_1 d(S)^{-1}$ corresponds to a rotation by π about an axis, perpendicular to v and enclosing an angle β with e_1, where β is the angle of rotation of S. Since the square of $d(S)\tilde{e}_1 d(S)^{-1}$ is also 1, (44a) is simply another way of writing $d(R) = \tilde{e}_3\tilde{e}_4$ as a product of two $\tilde{e}$ and we see that the normalization (44) is independent of the choice of the axis e_1 (Cf. Fig. 2).

For computing $d(R)d(T)$ we can draw the planes perpendicular to the axes of rotation of R and T and use for $d(R) = \tilde{e}_R\tilde{e}_C$ such a development that the axis e_C of the second involution coincide with the intersection line of the above-mentioned planes, while for $d(T) = \tilde{e}_C\tilde{e}_T$ we choose the first involution to be a rotation about this intersection line (Fig. 3). Then, the product

$$d(R)d(T) = \pm\tilde{e}_R\tilde{e}_C\tilde{e}_C\tilde{e}_T = \pm\tilde{e}_R\tilde{e}_T \tag{45}$$

will automatically have the normalization corresponding to (44). This shows that the operators normalized in (44) give a representation up to the sign.

For the Lorentz group, the proof can be performed along the same line, only the underlying geometrical facts are less obvious. Let Λ be a Lorentz transformation without elementary divisors with the characteristic values $e^{2i\gamma}$, $e^{-2i\gamma}$, $e^{2\chi}$, $e^{-2\chi}$ and the characteristic vectors v_1, $v_2 = v_1^*$, v_3, v_4, as described in section 4B.

We want to make $\Lambda = MN$ with $M^2 = N^2 = 1$. For $\Lambda N = M$, we have $\Lambda N \Lambda N = 1$ and thus $\Lambda N \Lambda = N$. Setting $Nv_i = \sum_k \alpha_{ik} v_k$, we obtain $\Lambda N \Lambda v_i = \sum \lambda_k \alpha_{ik} \lambda_i v_i = \sum \alpha_{ik} v_k$. Because of the linear independence of the v_k this amounts to $\lambda_i \lambda_k \alpha_{ik} = \alpha_{ik}$: all α_{ik} are zero, except those for which $\lambda_i \lambda_k = 1$. As in none of the cases (a), (b), (c), (d) of section 4B is λ_1 or λ_2 reciprocal to one of the last two λ, the vectors v_1 and v_2 will be transformed by N into a linear combination of v_1 and v_2 again, and the same holds for v_3 and v_4. This means that N can be considered as the product of two transformations $N = N_s N_t$, the first in the $v_1 v_2$ plane, the second in the $v_3 v_4$ plane. (Instead of $v_1 v_2$ plane one really should say $v_1 + v_2$, $iv_1 - iv_2$ plane, as v_1 and v_2 are complex themselves. This will be meant always by $v_1 v_2$ plane, etc.). The same holds for M also.

Both N_s and N_t must satisfy the first and third condition for Lorentz transformations (cf. 4A) and both determinants must be either 1, or -1. Furthermore, the square of both of them must be unity.

If both determinants were $+1$, the N_t had to be unity itself, while N_s could be the unity or a rotation by π in the $v_1 v_2$ plane. Thus v_1, v_2, v_3, v_4 would be characteristic vectors of N itself.

If both determinants are -1 (this will turn out to be the case), N_s is a reflection on a line in the $v_1 v_2$ plane and N_t a reflection in the $v_3 v_4$ plane, interchanging v_3 and v_4. In this case v_1, v_2, v_3, v_4 would not all be characteristic vectors of N.

If v_1, v_2, v_3, v_4 are characteristic vectors of N, they are characteristic vectors of $M = \Lambda N$ also. Then both M and N would be either unity, or a rotation by π in the $v_1 v_2$ plane. If both of them were rotations in the $v_1 v_2$ plane, their product Λ would be the unity which we want to exclude for the present. We can exclude the remaining cases in which the determinants of N_s and N_t are $+1$ by further stipulating that neither M nor N shall be the unity in the decomposition $\Lambda = MN$.

Hence N is the product of a reflection in the $v_1 v_2$ plane

$$Ns'_\nu = s'_\nu\,; \qquad Ns_\nu = -s_\nu \tag{46a}$$

where s_ν and s'_ν are two perpendicular real vectors in the $v_1 v_2$ plane

$$s'_\nu = e^{i\nu} v_1 + e^{-i\nu} v_2\,; \qquad s_\nu = i(e^{i\nu} v_1 - e^{-i\nu} v_2), \tag{46b}$$

and of a reflection in the $v_3 v_4$ plane

$$Nt'_\mu = t'_\mu\,; \qquad Nt_\mu = -t_\mu\,, \tag{46c}$$

where again t_μ, t'_μ are real vectors in the v_3v_4 plane, perpendicular to each other, t_μ being space-like, t'_μ time-like:

$$t'_\mu = e^\mu v_3 + e^{-\mu} v_4 \; ; \qquad t_\mu = e^\mu v_3 - e^{-\mu} v_4 \, . \tag{46d}$$

Thus N becomes a rotation by π in the purely space like $s_\nu t_\mu$ plane. The M can be calculated from $M = \Lambda N$

$$\begin{aligned}
Ms'_\nu &= \Lambda N s'_\nu = \Lambda s'_\nu = e^{i\nu+2i\gamma} v_1 + e^{-i\nu-2i\gamma} v_2 \\
&= \tfrac{1}{2} e^{2i\gamma}(s'_\nu - is_\nu) + \tfrac{1}{2} e^{-2i\gamma}(s'_\nu + is_\nu) = \cos 2\gamma \cdot s'_\nu + \sin 2\gamma \cdot s_\nu \\
Ms_\nu &= \sin 2\gamma \cdot s'_\nu - \cos 2\gamma \cdot s_\nu \\
Mt'_\mu &= \Lambda N t'_\mu = \Lambda t'_\mu = e^{\mu+2\chi} v_3 + e^{-\mu-2\chi} v_4 \\
&= \tfrac{1}{2} e^{2\chi}(t'_\mu + t_\mu) + \tfrac{1}{2} e^{-2\chi}(t'_\mu - t_\mu) = \mathrm{Ch}\, 2\chi \cdot t'_\mu + \mathrm{Sh}\, 2\chi \cdot t_\mu \\
Mt_\mu &= -\mathrm{Sh}\, 2\chi \cdot t'_\mu - \mathrm{Ch}\, 2\chi \cdot t_\mu \, .
\end{aligned} \tag{46e}$$

Thus M also becomes a product of two reflections, one in the $v_1v_2 = s'_\nu s_\nu$ the other in the $v_3v_4 = t'_\mu t_\mu$ plane. This completes the decomposition of Λ into two involutions. One of the involutions can be taken to be a rotation by π in an arbitrary space like plane, intersecting both the v_1v_2 and the v_3v_4 planes, as the freedom in choosing ν and μ allows us to fix the lines s_ν and t_μ arbitrarily in those planes. The involution characterized by (46) will be called $N_{\nu\mu}$ henceforth. The other involution M is then a similar rotation, in a plane, however, which is completely determined once the $s_\nu t_\mu$ plane is fixed. It will be denoted by $M_{\nu\mu}$ (it is, in fact $M_{\nu\mu} = N_{\nu+\gamma\ \mu+\chi}$). One sees the complete analogy to the three dimensional case if one remembers that γ and χ are the half angles of rotation.

The $d(M)$ and $d(N)$ so normalized that their squares be 1 shall be denoted by $d_1(M_{\nu\mu})$ and $d_1(N_{\nu\mu})$. We must show that the normalization for

$$d(\Lambda) = \pm d_1(M_{\nu\mu}) d_1(N_{\nu\mu}) \tag{47}$$

is independent of ν and μ. For this purpose, we transform

$$d(\Lambda) = \pm d_1(M_{00}) d_1(N_{00}) \tag{47a}$$

with $d(\Lambda_1)$ where Λ_1 has the same characteristic vectors as Λ but different characteristic values, namely $e^{i\nu}$, $e^{-i\nu}$, e^μ and $e^{-\mu}$. Since $\Lambda_1 M_{00} \Lambda_1^{-1} = M_{\nu\mu}$ and $\Lambda_1 N_{00} \Lambda_1^{-1} = N_{\nu\mu}$ we have $d(\Lambda_1) d_1(M_{00}) d(\Lambda_1)^{-1} = \omega d_1(M_{\nu\mu})$ where $\omega = \pm 1$, as the squares of both sides are 1. Hence, (47a) becomes if transformed with $d(\Lambda_1)$ just

$$d(\Lambda_1) d(\Lambda) d(\Lambda_1)^{-1} = \pm d_1(M_{\nu\mu}) d_1(N_{\nu\mu}). \tag{47b}$$

The normalization (47) would be clearly independent of ν and μ if $d(\Lambda_1)$ commuted with $d(\Lambda)$.

Again, the argument contained in equations (24)-(30) can be applied and shows that

$$d(\Lambda_1) d(\Lambda) d(\Lambda_1)^{-1} = \exp\,(2\pi i f(2\gamma\mu - 2\chi\nu)) d(\Lambda) \tag{48}$$

holds for every γ, χ, ν, μ. However, the exponential in (48) must be 1 if $\gamma = 0$; $\nu = 2\pi/n$; $\chi = \frac{1}{2}n\mu$ since in this case $\Lambda = \Lambda_1^n$. Thus $\exp(-4\pi^2 if\mu) = 1$ for every μ and $f = 0$ and the left side of (47b) can be replaced by $d(\Lambda)$; the normalization in (47) is independent of ν and μ.

In order to have the analogue of (45), we must show that, having two Lorentz transformations $\Lambda = M_{\nu\mu}N_{\nu\mu}$ and $I = P_{\alpha\beta}Q_{\alpha\beta}$ we can chose ν, μ and α, β so that $N_{\nu\mu} = P_{\alpha\beta}$ i.e. that the plane of rotation $s_\nu t_\mu$ of $N_{\nu\mu}$ coincide with the plane of rotation of $P_{\alpha\beta}$. As the latter plane can be made to an arbitrary space-like plane intersecting both the w_1w_2 and the w_3w_4 planes (where w_1, w_2, w_3, w_4 are the characteristic vectors of I), we must show the existence of a space like plane, intersecting all four planes v_1v_2, v_3v_4, w_1w_2, w_3w_4. Both the first and the second pair of planes are orthogonal.

One can show[22] that if Λ and I have no common null vector as characteristic

[22] We first suppose the existence of a real plane p intersecting all four planes v_1v_2, v_3v_4, w_1w_2, w_3w_4. If p intersects v_1v_2 the plane q perpendicular to p will intersect the plane v_3v_4 perpendicular to v_1v_2. Indeed, the line which is perpendicular to both p and v_1v_2 (there is such a line as p and v_1v_2 intersect) is contained in both q and v_3v_4. This shows that if there is a plane intersecting all four planes, the plane perpendicular to this will have this property also.

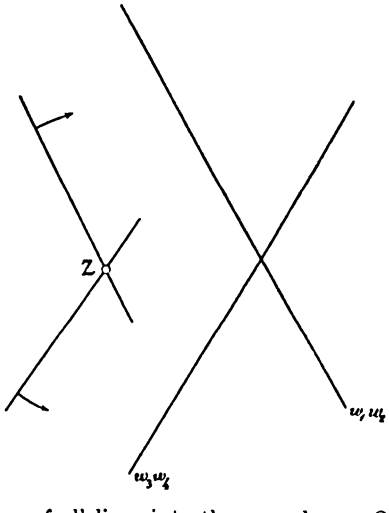


Fig. 4 gives a projection of all lines into the x_1x_2 plane. One sees that there are, in general, two intersecting planes, only in exceptional cases is there only one.

If the plane p—the existence of which we suppose for the time being—contains a time-like vector, q will be space-like (Section 4B, [1]). Both in this case and if p contains only space-like vectors, the theorem in the text is valid. There is a last possibility, that p is tangent to the light cone, i.e. contains only space like vectors and a null vector v. The space-like vectors of p are all orthogonal to v, otherwise p would contain time-like vectors also. In this case the plane q, perpendicular to p will contain v also. The line in which v_1v_2 intersects p is space-like and orthogonal to the vector in which v_3v_4 intersects p. The latter intersection must coincide with v, therefore, as no other vector of p is orthogonal to

vector, there are always two planes, perpendicular to each other which intersect four such planes. One of these is always space like. It is possible to assume, therefore, that both $N_{\nu\mu}$ and $P_{\alpha\beta}$ are the rotation by π in this plane. Thus

$$\begin{aligned} d(\Lambda)d(I) &= \pm d_1(M_{\nu\mu})d_1(N_{\nu\mu})d_1(P_{\alpha\beta})d_1(Q_{\alpha\beta}) \\ &= \pm d_1(M_{\nu\mu})d_1(Q_{\alpha\beta}), \end{aligned} \tag{49}$$

and $d(\Lambda)d(I)$ has the normalization corresponding to the product of two involutions, neither of which is unity. This is, however, also the normalization adopted for $d(\Lambda I)$. Hence

$$d(\Lambda)d(I) = \pm d(\Lambda I) \tag{49a}$$

holds if Λ, I and ΛI are Lorentz transformations corresponding to one of the cases (a), (b), (c) or (d) of section 4B and if Λ and I have no common characteristic null vector. In addition to this (49a) holds also, assuming $d(E) = \pm 1$, if any of the transformations Λ, I, ΛI is unity, or if both characteristic null vectors of Λ and I are equal, as in this case the planes v_3v_4 and w_3w_4 and also v_1v_2 and w_1w_2 coincide and there are many space like planes intersecting all.

If Λ and I have one common characteristic null vector, $v_3 = w_3$, the others, v_4 and w_4 respectively, being different, one can use an aritifice to prove (49a) which will be used in later parts of this section extensively. One can find a Lorentz transformation J so that none of the pairs $I - J$; $\Lambda - IJ$; $\Lambda IJ - J^{-1}$ has a common characteristic null vector. This will be true, e.g. if the characteristic null vectors of J are v_4 and another null vector, different from v_3, w_4 and the characteristic vectors of ΛI. Then (49a) will hold for all the above pairs and

$$\begin{aligned} d(\Lambda)d(I) &= \pm d(\Lambda)d(I)d(J)d(J^{-1}) = \pm d(\Lambda)d(IJ)d(J^{-1}) \\ &= \pm d(\Lambda IJ)d(J^{-1}) = \pm d(\Lambda I). \end{aligned}$$

---

any space-like vector in it. Hence, v is the intersection of p and v_3v_4 and is either v_3 or v_4. One can conclude in the same way that v coincides with either w_3 or w_4 also and we see that if p is tangent to the light cone the two transformations Λ and I have a common null vector as characteristic vector. Thus the theorem in the text is correct if we can show the existence of an arbitrary real plane p intersecting all four planes v_1v_2, v_3v_4, w_1w_2, w_3w_4.

Let us draw a coördinate system in our four dimensional space, the x_1x_2 plane of which is the v_1v_2 plane, the x_3 and x_4 axes having the directions of the vectors $v_3 - v_4$ and $v_3 + v_4$, respectively. The three dimensional manifold M characterized by $x_4 = 1$ intersects all planes in a line, the v_1v_2 plane in the line at infinity of the x_1x_2 plane, the v_3v_4 plane in the x_3 axis. The intersection of M with the w_1w_2 and w_3w_4 planes will be lines in M with directions perpendicular to each other. They will have a common normal through the origin of M, intersecting it at reciprocal distances. This follows from their orthogonality in the four dimensional space.

A plane intersecting v_1v_2 and v_3v_4 will be a line parallel to x_1x_2 through the x_3 axis. If we draw such lines through all points of the line corresponding to w_1w_2, the direction of this line will turn by π if we go from one end of this line to the other. Similarly, the lines going through the line corresponding to w_3w_4 will turn by π in the *opposite* direction. Thus the first set of lines will have at least one line in common with the second set and this line will correspond to a real plane intersecting all four planes v_1v_2, v_3v_4, w_1w_2, w_3w_4. This completes the proof of the theorem referred to in the text.

This completes the proof of (49a) for all cases in which Λ, I and ΛI have no elementary divisors. It is evident also that we can replace in the normalization (47) the d_1 by d. One also concludes easily that $d(M)^2$ is in the same representation either $+1$ for all involutions M, or -1 for every involution. The former ones will give real representations, the latter ones representations up to the sign.

If Λ has an elementary divisor, it can be expressed in the v_e, w_e, z_e, v_1 scheme as the matrix (Cf. equ. (20))

$$\Lambda_e = \begin{Vmatrix} 1 & 1 & \frac{1}{2} & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{Vmatrix}$$

and can be written, in the same scheme, as the product of two Lorentz transformations with the square 1

$$\Lambda_e = M_0 N_0 = \begin{Vmatrix} 1 & -1 & \frac{1}{2} & 0 \\ 0 & -1 & 1 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{Vmatrix} \cdot \begin{Vmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{Vmatrix}.$$

We can normalize therefore $d(\Lambda_e) = \pm d(M_0)d(N_0)$. If Λ can be written as the product of two other involutions also $\Lambda_e = M_1 N_1$ the corresponding normalization will be identical with the original one. In order to prove this, let us consider a Lorentz transformation J such that neither of the Lorentz transformations J, $N_0 J$, $N_1 J$, $\Lambda_e J = M_0 N_0 J = M_1 N_1 J$ have an elementary divisor. Since the number of free parameters is only 4 in case (e), while 6 for case (a), this is always possible. Then, for (45a)

$$\begin{aligned} d(M_0)d(N_0)d(J) &= \pm d(M_0)d(N_0 J) = \pm d(M_0 N_0 J) \\ &= \pm d(M_1 N_1 J) = \pm d(M_1)d(N_1 J) = \pm d(M_1)d(N_1)d(J) \end{aligned}$$

and thus $d(M_0)d(N_0) = \pm d(M_1)d(N_1)$. This shows also that even if ΛI is in case (e), $\omega(\Lambda, I) = \pm 1$, since (49) leads to the correct normalization.

If $\Lambda = MN$ has an elementary divisor, I not, $d(\Lambda)d(I)$ still will have the normalization corresponding to the product of two involutions. One can find again a J such that neither of the transformations J, J^{-1}, IJ, NIJ, $MNIJ$, have an elementary divisor. Then

$$\begin{aligned} d(M)d(N)d(I) &= \pm d(M)d(N)d(I)d(J)d(J)^{-1} \\ &= \pm d(M)d(N)d(IJ)d(J)^{-1} = \pm d(M)d(NIJ)d(J)^{-1} \\ &= d(\Lambda IJ)d(J^{-1}). \end{aligned}$$

The last product has, however, the normalization corresponding to two involutions, as was shown in (49a), since neither ΛIJ, nor J^{-1} is in case (e).

Lastly, we must consider the case when both Λ and I may have an elementary divisor. In this case, we need a J such that neither of J, J^{-1}, IJ have one. Then, because of the generalization of (49a) just proved, in which the first factor is in case (c)

$$\begin{aligned} d(\Lambda)d(I) &= \pm d(\Lambda)d(I)d(J)d(J^{-1}) = \pm d(\Lambda)d(IJ)d(J^{-1}) \\ &= \pm d(\Lambda IJ)d(J^{-1}) \end{aligned}$$

which has the right normalization.

This completes the proof of

$$\omega(\Lambda, I) = \pm 1 \tag{50}$$

for all possible cases, and the normalization of all $D(L)$ of a representation of the inhomogeneous Lorentz group up to a factor, is carried out in such a way that the normalized operators give a representation up to the sign. It is even carried so far that in the first two of equations (22) $\omega = 1$ can be set. We shall consider henceforth systems of operators satisfying (7), or, more specifically, (22B) and (22C) with $\omega(a, b) = \omega(\Lambda, a) = 1$ and (22D) with $\omega(\Lambda, I) = \pm 1$.

E.

Lastly, it shall be shown that the renormalization not only did not spoil the partly continuous character of the representation, attained at the first normalization in part (A) of this section, but that the same holds now *everywhere*, in the ordinary sense for $T(a)$ and, apart from the ambiguity of sign, also for $d(\Lambda)$. For $T(a)$ this was proved in part $(B)(b)$ of this section, for $d(\Lambda)$ it means that to every Λ_1, ϵ and φ there is such a δ that *one* of the two quantities

$$((d(\Lambda_1) \mp d(\Lambda))\varphi, (d(\Lambda_1) \mp d(\Lambda))\varphi) < \epsilon \tag{51}$$

if Λ is in the neighborhood δ of Λ_1. The inequality (51) is equivalent to

$$((1 \mp d(\Lambda_0))\varphi, (1 \mp d(\Lambda_0))\varphi) < \epsilon, \tag{51a}$$

where $\Lambda_0 = \Lambda_1^{-1}\Lambda$ now can be assumed to be in the neighborhood of the unity. Thus, the continuity of $d(\Lambda)$ at $\Lambda = E$ entails the continuity everywhere.[23] In fact, it would be sufficient to show that the $d(X)$, $d(Y)$ and $d(Z)$ corresponding to the transformations (40) converge to ± 1, as α, γ approach 0, since one can write every transformation in the neighborhood of the unit element as a product $\Lambda = Z(0, \gamma_3)Y(0, \gamma_2)X(0, \gamma_1)X(\alpha_1, 0)Y(\alpha_2, 0)Z(\alpha_3, 0)$ and the parameters $\alpha_1, \cdots, \gamma_3$ will converge to 0 as Λ converges to 1. However, we shall carry out the proof for an arbitrary Λ without an elementary divisor.

For $d(\Lambda)$, equations (46) show that as Λ approaches E (i.e., as γ and χ approach zero) both M_{00} and N_{00} approach the same involution, which we shall call K. Let us now consider a wave function $\psi = \varphi + d_1(K)\varphi$ or, if this vanishes $\psi = \varphi - d_1(K)\varphi$. We have $d_1(K)\psi = \pm\psi$. If Λ is sufficiently near to unity,

[23] J. von Neumann, Sitz. d. kön. Preuss. Akad. p. 76, 1927.

$d_1(N_{00})\psi$ will be sufficiently near to $\Omega d_1(K)\psi = \pm\Omega\psi$ and all we have to show is that Ω approaches ± 1. The same thing will hold for $d_1(M_{00})$. Indeed from $d_1(N_{00})\psi - \Omega\psi = u$ it follows by applying $d_1(N_{00})$ on both sides $\psi - \Omega^2\psi = (d_1(N_{00}) + \Omega)u$. As (u, u) goes to zero, Ω must go to ± 1, and consequently, also $d_1(N_{00})\psi$ goes to ψ or to $-\psi$. Applying $d_1(M_{00})$ to this, one sees that $d_1(M_{00})d_1(N_{00})\psi = d(\Lambda)\psi$ goes to $\pm\psi$ as Λ goes to unity. The argument given in (A)(b) shows that this holds not only for ψ but for every other function also, i.e. $d(\Lambda)$ converges to $\pm 1 = d(E)$ as Λ approaches E. Thus $d(\Lambda)$ is continuous in the neighborhood of E and hence everywhere.

According to the last remark in part 4, the operators $\pm d(\Lambda)$ form a single valued representation of the group of complex unimodular two dimensional matrices C. Let us denote the homogeneous Lorentz transformation which corresponds in the isomorphism to C by $\tilde{C}$. Our task of solving the equs. (22) has been reduced to finding all single valued unitary representation of the group with the elements $[a, C] = [a, 1]\,[0, C]$, the multiplication rule of which is $[a, C_1]\,[b, C_2] = [a + \tilde{C}_1 b, C_1C_2]$. For the representations of this group $D[a, C] = T(a)d[C]$ we had

$$\begin{aligned} T(a)T(b) &= T(a + b) \\ d[C]T(a) &= T(\tilde{C}a)\,d[C] \\ d[C_1]\,d[C_2] &= d[C_1C_2]. \end{aligned} \tag{52a}$$

It would be more natural, perhaps, from the mathematical point of view, to use henceforth this new notation for the representations and let the d depend on the C rather than on the $\tilde{C}$ or Λ. However, in order to be reminded on the geometrical significance of the group elements, it appeared to me to be better to keep the old notation. Instead of the equations (22B), (22C), (22D) we have, then

$$T(a)T(b) = T(a + b) \tag{52B}$$

$$d(\Lambda)T(a) = T(\Lambda a)\,d(\Lambda) \tag{52C}$$

$$d(\Lambda)\,d(I) = \pm d(\Lambda I). \tag{52D}$$

6. Reduction of the Representations of the Inhomogeneous Lorentz Group to Representations of a "Little Group"

This section, unlike the other ones, will often make use of methods, which though commonly accepted in physics, must be further justified from a rigorous mathematical point of view. This has been done, in the meanwhile, by J. von Neumann in an as yet unpublished article and I am much indebted to him for his coöperation in this respect and for his readiness in communicating his results to me. A reference to his paper[24] will be made whenever his work is necessary for making inexact considerations of this section rigorous.

[24] J. von Neumann, Ann. of Math. to appear shortly.

A.

Since the translation operators all commute, it is possible[24] to introduce such a coördinate system in Hilbert space that the wave functions $\varphi(p, \zeta)$ contain momentum variables p_1, p_2, p_3, p_4 and a discrete variable ζ so that

$$T(a)\varphi(p, \zeta) = e^{i\{p,a\}}\varphi(p, \zeta). \tag{53}$$

p will stand for the four variables p_1, p_2, p_3, p_4.

Of course, the fact that the Lorentzian scalar product enters in the exponent, rather than the ordinary, is entirely arbitrary and could be changed by changing the signs of p_1, p_2, p_3.

The unitary scalar product of two wave functions is not yet completely defined by the requirements so far made on the coördinate system. It can be a summation over ζ and an arbitrary Stieltjes integral over the components of p:

$$(\psi, \varphi) = \sum_{\zeta} \int \psi(p, \zeta)^* \varphi(p, \zeta)\, df(p, \zeta). \tag{54}$$

The importance of introducing a weight factor, depending on p, for the scalar product lies not so much in the possibility of giving finite but different weights to different regions in p space. Such a weight distribution $g(p, \zeta)$ always could be absorbed into the wave functions, replacing all $\varphi(p, \zeta)$ by $\sqrt{g(p, \zeta)} \cdot \varphi(p, \zeta)$. The necessity of introducing the $f(p, \zeta)$ lies rather in the possibility of some regions of p having zero weight while, on the other hand, at other places points may have finite weights. On account of the definite metric in Hilbert space, the integral $\int df(p, \zeta)$ over any region r, for any ζ, is either positive, or zero, since it is the scalar product of that function with itself, which is 1 in the region r of p and the value ζ of the discrete variable, zero otherwise.

Let us now define the operators

$$P(\Lambda)\varphi(p, \zeta) = \varphi(\Lambda^{-1}p, \zeta). \tag{55}$$

This equation defines the function $P(\Lambda)\varphi$, which is, at the point p, ζ, as great as the function φ at the point $\Lambda^{-1}p, \zeta$. The operator $P(\Lambda)$ is not necessarily unitary, on account of the weight factor in (54). We can easily calculate

$$P(\Lambda)T(a)\varphi(p, \zeta) = T(a)\varphi(\Lambda^{-1}p, \zeta) = e^{i\{\Lambda^{-1}p,a\}}\varphi(\Lambda^{-1}p, \zeta),$$

$$T(\Lambda a)P(\Lambda)\varphi(p, \zeta) = e^{i\{p,\Lambda a\}}P(\Lambda)\varphi(p, \zeta) = e^{i\{p,\Lambda a\}}\varphi(\Lambda^{-1}p, \zeta),$$

so that, for $\{\Lambda^{-1}p, a\} = \{p, \Lambda a\}$, we have

$$P(\Lambda)T(a) = T(\Lambda a)P(\Lambda). \tag{56}$$

This, together with (52C), shows that $d(\Lambda)P(\Lambda)^{-1} = Q(\Lambda)$ commutes with all $T(a)$ and, therefore, with the multiplication with every function of p, since the exponentials form a complete set of functions of p_1, p_2, p_3, p_4. Thus

$$d(\Lambda) = Q(\Lambda)P(\Lambda), \tag{57}$$

where $Q(\Lambda)$ is an operator in the space of the ζ alone[24] which can depend, however, on the particular value of p in the underlying space:

$$Q(\Lambda)\varphi(p, \zeta) = \sum_{\eta} Q(p, \Lambda)_{\zeta\eta}\varphi(p, \eta). \tag{57a}$$

Here, $Q(p, \Lambda)_{\zeta\eta}$ are the components of an ordinary (finite or infinite) matrix, depending on p and Λ. From (57), we obtain

$$\begin{aligned} d(\Lambda)\varphi(p, \zeta) &= \sum_{\eta} Q(p, \Lambda)_{\zeta\eta} P(\Lambda)\varphi(p, \eta) \\ &= \sum_{\eta} Q(p, \Lambda)_{\zeta\eta}\varphi(\Lambda^{-1}p, \eta). \end{aligned} \tag{57b}$$

As the exponentials form a complete set of functions, we can approximate the operation of multiplication with any function of p_1, p_2, p_3, p_4 by a linear combination

$$f(p)\varphi = \sum_{n} c_n T(a_n)\varphi. \tag{58}$$

If we choose $f(p)$ to be such a function that

$$f(p) = f(\Lambda p) \tag{58a}$$

the operation of multiplication with $f(p)$ will commute with all operations of the group. It commutes evidently with the $T(a)$ and the $Q(p, \Lambda)$, and on account of (56) and (58), (58a) also with $P(\Lambda)$. Thus the operation of (58) belongs to the centrum of the algebra of our representation. Since, however, we assume that the representation is factorial (cf. 2), the centrum contains only multiples of the unity and

$$f(p)\varphi(p, \zeta) = c\varphi(p, \zeta). \tag{58b}$$

This can be true only if φ is different from zero only for such momenta p which can be obtained from each other by homogeneous Lorentz transformations, because $f(p)$ needs to be equal to $f(p')$ only if there is a Λ which brings them into each other.

It will be sufficient, henceforth, to consider only such representations, the wave functions of which vanish except for such momenta which can be obtained from one by homogeneous Lorentz transformations. One can restrict, then, the definition domain of the φ to these momenta.

These representations can now naturally be divided into the four classes enumerated in section 3, and two classes contain two subclasses. There will be representations, the wave functions of which are defined for such p that

(1) $\{p, p\} = P > 0$ (3) $p = 0$

(2) $\{p, p\} = P = 0; p \neq 0$ (4) $\{p, p\} = P < 0.$

The classes 1 and 2 contain two sub-classes each. In the positive subclasses P_+ and 0_+ the time components of all momenta are $p_4 > 0$, in the negative

subclasses P_- and 0_- the fourth components of the momenta are negative. Class 3 will be denoted by 0_0. If P is negative, it has no index.

From the condition that $d(\Lambda)$ shall be a unitary operator, it is possible to infer[24] that one can introduce a coördinate system in Hilbert space in such a way that

$$\int_r df(p, \zeta) = \int_{\Lambda r} df(p, \eta) \tag{59}$$

if $Q(p, \Lambda)_{\zeta\eta} \neq 0$ for the p of the domain r. Otherwise, r is an arbitrary domain in the space of p_1, p_2, p_3, p_4 and Λr is the domain which contains Λp if r contains p. Equation (59) holds for all ζ, η, except for such pairs for which $Q(p, \Lambda)_{\zeta\eta} = 0$. It is possible, hence, to decompose the original representation in such a way that (59) holds within every reduced part. Neither $T(a)$ nor $d(\Lambda)$ can have matrix elements between such η and ζ for which (59) does not hold.

In the third class of representations, the variable p can be dropped entirely, and $T(a)\varphi(\zeta) = \varphi(\zeta)$, i.e., all wave functions are invariant under the operations of the invariant subgroup, formed by the translations. The equation $T(a)\varphi(\zeta) = \varphi(\zeta)$ is an invariant characterization of the representations of the third class, i.e., a characterization which is not affected by a similarity transformation. Hence, the reduced parts of a representation of class 3 also belong to this class.

Since no wave function of the other classes can remain invariant under all translations, no representation of the third class can be contained in any representation of one of the other classes. In the other classes, the variability domain of p remains three dimensional. It is possible, therefore, to introduce instead of p_1, p_2, p_3, p_4 three independent variables. In the cases 1 and 2 with which we shall be concerned most, p_1, p_2, p_3 can be kept for these three variables. On account of (59), the Stieltjes integral can be replaced by an ordinary integral[24] over these variables, the weight factor being $|p_4|^{-1} = (P + p_1^2 + p_2^2 + p_3^2)^{-\frac{1}{2}}$

$$\{\psi, \varphi\} = \sum_\zeta \int\int\int_{-\infty}^{\infty} \psi(p, \zeta)^* \varphi(p, \zeta)\, |p_4|^{-1} dp_1\, dp_2\, dp_3. \tag{59a}$$

In fact, with the weight factor $|p_4|^{-1}$ the weight of the domain r i.e., $W_r = \int\int\int_r |p_4|^{-1} dp_1\, dp_2\, dp_3$ is equal to the weight of the domain $W_{\Lambda r}$ as required[25] by (59). Having the scalar product fixed in this way, $P(\Lambda)$ becomes a unitary operator and, hence, $Q(\Lambda)$ will be unitary also.

We want to give next a characterization of the representations with a given P, which is independent of the coördinate system in Hilbert space. It follows from

[25] The invariance of integrals of the character of (59a) is frequently made use of in relativity theory. One can prove it by calculating the Jacobian of the transformation

$$p'_i = \Lambda_{i1}p_1 + \Lambda_{i2}p_2 + \Lambda_{i3}p_3 + (P + p_1^2 + p_2^2 + p_3^2)^{\frac{1}{2}} \qquad (i = 1, 2, 3)$$

which comes out to be $(P + p_1^2 + p_2^2 + p_3^2)^{\frac{1}{2}}(P + p_1'^2 + p_2'^2 + p_3'^2)^{-\frac{1}{2}}$. Equ. (59a) will not be used in later parts of this paper.

(53), that in a representation with a given P the wave functions $\psi_1, \psi_2, \cdots$ which are different from zero only in a finite domain of p, form an everywhere dense set, to all elements of which the infinitesimal operators of translation can be applied arbitrarily often

$$\begin{aligned} \lim_{h=0} h^{-n}(T(he) - 1)^n\psi &= \lim_{h=0} h^{-n}(e^{ih\{p,e\}} - 1)^n\psi \\ &= i^n\{p, e\}^n\psi, \end{aligned} \tag{60}$$

where e will be a unit vector in the direction of a coördinate axis or oppositely directed to it. Hence for all members ψ of this everywhere dense set

$$\lim_{h=0} \sum_k \pm h^{-2}(T(2he_k) - 2T(he_k) + 1)\psi = (p_1^2 + p_2^2 + p_3^2 - p_4^2)\psi = -P\psi, \tag{61}$$

where e_k is a unit vector in (or opposite) the kth coördinate axis and the $\pm$ is $+$ for $k = 4$, and $-$ for $k = 1, 2, 3$.

On the other hand, there is no φ for which

$$\lim_{h=0} \sum_k \pm h^{-2}(T(2he_k) - 2T(he_k) + 1)\varphi \tag{61a}$$

if it exists, would be different from $-P\varphi$. Suppose the limit in (61a) exists and is $-P\varphi + \varphi'$. Let us choose then a normalized ψ, from the above set, such that $(\psi, \varphi') = \delta$ with $\delta > 0$ and an h so that the expression after the lim sign in (61a) assumes the value $-P\varphi + \varphi' + u$ with $(u, u) < \delta/3$ and also the expression after the lim sign in (61), with oppositely directed e_k becomes $-P\psi + u'$ with $(u', u') < \delta/3$. Then, on account of the unitary character of $T(a)$ and because of $T(-a) = T(a)^{-1}$

$$\begin{aligned} &\Big(\psi, \sum_k \pm h^{-2}(T(2he_k) - 2T(he_k) + 1)\varphi\Big), \\ &\qquad = \Big(\sum_k \pm h^{-2}(T(-2he_k) - 2T(-he_k) + 1)\psi, \varphi\Big), \end{aligned}$$

or

$$-P(\psi, \varphi) + (\psi, \varphi') + (\psi, u) = -P(\psi, \varphi) + (u', \varphi),$$

which is clearly impossible.

Thus if the lim in (61a) exists, it is $-P\varphi$ and this constitutes a characterization of the representation which is independent of similarity transformations. Since, according to the foregoing, it is always possible to find wave functions for a representation, to which (61a) can be applied, every reduced part of a representation with a given P must have this same P and no representation with one P can be contained in a representation with an other P. The same argument can be applied evidently to the positive and negative sub-classes of class 1 and 2.

B.

Every automorphism $L \to L^\circ$ of the group allows us to construct from one representation $D(L)$ another representation

$$D^\circ(L) = D(L^\circ). \tag{62}$$

This principle will allow us to restrict ourselves, for representations with finite, positive or negative P, to one value of P which can be taken respectively, to be $+1_+$ and -1. It will also allow in cases 1 and 2 to construct the representations of the negative sub-classes out of representations of the positive sub-classes.

The first automorphism is $a^\circ = \alpha a, \Lambda^\circ = \Lambda$. Evidently Equs. (12) are invariant under this transformation. If we set, however,

$$T^\circ(a)\varphi = T(\alpha a)\varphi; \qquad d^\circ(\Lambda)\varphi = d(\Lambda)\varphi,$$

then the occurring p

$$T^\circ(a)\varphi = T(\alpha a)\varphi = e^{i\{p,\alpha a\}}\varphi = e^{i\{\alpha p,a\}}\varphi,$$

will be the p occurring for the unprimed representation, multiplied by α. This allows, with a real positive α, to construct all representations with all possible numerical values of P, from all representation with one numerical value of P. If we take α negative, the representations of the negative sub-classes are obtained from the representations of the positive sub-class.

In case $P = 0_0$ evidently all representations go over into themselves by the transformation (62). In case $P = 0_+$ and $P = 0_-$ it will turn out that for positive α, (62) carries every representation into an equivalent one.

C.

On account of (53) and (56), (57), the equs. (52B) and (52C) are automatically satisfied and the $Q(p, \Lambda)_{\zeta\eta}$ must be determined by (52D). This gives

$$\sum_{\eta\vartheta} Q(p, \Lambda)_{\zeta\eta} Q(\Lambda^{-1}p, I)_{\eta\vartheta}\, \varphi(I^{-1}\Lambda^{-1}p, \vartheta) = \pm \sum_{\vartheta} Q(p, \Lambda I)_{\zeta\vartheta}\, \varphi(I^{-1}\Lambda^{-1}p, \vartheta). \tag{63}$$

Since this must hold for every φ, one would conclude

$$\sum_{\eta} Q(p, \Lambda)_{\zeta\eta} Q(\Lambda^{-1}p, I)_{\eta\vartheta} = \pm\, Q(p, \Lambda I)_{\zeta\vartheta}. \tag{63a}$$

Actually, this conclusion is not justified, since two wave functions must be considered to be equal even if they are different on a set of measure zero. Thus one cannot conclude, without further consideration, that the two sides of (63a) are equal at every point p. On the other hand,[24] the value of $Q(p, \Lambda)_{\zeta\eta}$ can be changed on a set of measure zero and one can make it continuous in the neighborhood of every point, if the representation is continuous. This allows then, to justify (63a). It follows from (63a) that $Q(p, 1)_{\zeta\eta} = \delta_{\zeta\eta}$.

Let us choose[15] now a basic p_0 arbitrarily. We can consider then the subgroup of all homogeneous Lorentz transformations which leave this p_0 unchanged. For all elements λ, ι of this "little group," we have

$$\begin{aligned} \sum_\eta Q(p_0, \lambda)_{\zeta\eta} Q(p_0, \iota)_{\eta\vartheta} &= \pm Q(p_0, \lambda\iota)_{\zeta\vartheta} \\ q(\lambda)q(\iota) &= \pm q(\lambda\iota), \end{aligned} \tag{64}$$

where $g(\lambda)$ is the matrix $q(\lambda)_{\zeta\eta} = Q(p_0, \lambda)_{\zeta\eta}$. Because of the unitary character of $Q(\Lambda)$, the $Q(p_0, \Lambda)_{\zeta\eta}$ is unitary matrix and $q(\lambda)$ is unitary also.

If we consider, according to the last paragraph of Section 5, the group formed out of the translations and unimodular two-dimensional matrices, rather than Lorentz transformations, the $\pm$ sign in (64) can be replaced by a $+$ sign. In this case, λ and ι are unimodular two-dimensional matrices and the little group is formed by those matrices, the corresponding Lorentz transformations $\bar{\lambda}$, $\bar{\iota}$ to which leave p_0 unchanged $\bar{\lambda}p_0 = \bar{\iota}p_0 = p_0$.

Adopting this interpretation of (64), one can also see, conversely, that the representation $q(\lambda)$ of the little group, together with the class and P of the representation of the whole group, determines the latter representation, apart from a similarity transformation. In order to prove this, let us define for every p a two-dimensional unimodular matrix $\alpha(p)$ in such a way that the corresponding Lorentz transformation

$$\bar{\alpha}(p)p_0 = p \tag{65}$$

brings p_0 into p. The $\alpha(p)$ can be quite arbitrary except of being an almost everywhere continuous function of p, especially continuous for $p = p_0$ and $\alpha(p_0) = 1$. Then, we can set

$$\begin{aligned} d(\alpha(p)^{-1})\varphi(p_0, \zeta) &= \varphi(p, \zeta), \\ d(\alpha(p))\varphi(p, \zeta) &= \varphi(p_0, \zeta). \end{aligned} \tag{66}$$

This is equivalent to setting in (58)

$$Q(p, \alpha(p)) = 1 \tag{66a}$$

and can be achieved by a similarity transformation which replaces $\varphi(p, \zeta)$ by $\sum_\eta Q(p_0, \alpha(p)^{-1})_{\zeta\eta}\varphi(p, \eta)$. As the matrix $Q(p_0, \alpha(p)^{-1})$ is unitary, this is a unitary transformation. It does not affect, furthermore, (53) since it contains p only as a parameter.

Assuming this transformation to be carried out, (66) will be valid and will define, together with the $d(\lambda)$, all the remaining $Q(p, \Lambda)$ uniquely. In fact, calculating $d(\Lambda)\varphi(p, \zeta)$, we can decompose Λ into three factors

$$\Lambda = \alpha(p). \quad \alpha(p)^{-1}\Lambda\alpha(\bar{\Lambda}^{-1}p). \quad \alpha(\bar{\Lambda}^{-1}p)^{-1} \tag{67}$$

The second factor $\beta = \alpha(p)^{-1}\Lambda\alpha(\tilde{\Lambda}^{-1}p)$ belongs into the little group: $\tilde{\alpha}(p)^{-1}\tilde{\Lambda}\tilde{\alpha}(\tilde{\Lambda}^{-1}p)p_0 = \tilde{\alpha}(p)^{-1}\tilde{\Lambda}\cdot\tilde{\Lambda}^{-1}p = \tilde{\alpha}(p)^{-1}p = p_0$. We can write, therefore ($\tilde{\Lambda}^{-1}p = p'$)

$$\begin{aligned} d(\Lambda)\varphi(p, \zeta) &= d(\alpha(p))d(\beta)d(\alpha(p'))^{-1}\varphi(p, \zeta) \\ &= d(\beta)d(\alpha(p'))^{-1}\varphi(p_0, \zeta) \qquad (67a) \\ &= \sum_\eta q(\beta)_{\zeta\eta}d(\alpha(p')^{-1})\varphi(p_0, \eta) = \sum_\eta q(\beta)_{\zeta\eta}\varphi(p', \eta). \end{aligned}$$

This shows that all representations of the whole inhomogeneous Lorentz group are equivalent which have the same P and the same representation of the little group. Further than this, the same holds even if the representations of the little group are not the same for the two representations but only equivalent to each other. Let us assume $q_1(\Lambda) = sq_2(\Lambda)s^{-1}$. Then by replacing $\varphi(p, \zeta)$ by $\sum_\eta s(\zeta, \eta)\varphi(p, \eta)$ we obtain a new form of the representation for which (53) still holds but $q_2(\beta)$ for the little group is replaced by $q_1(\beta)$. Then, by the transformation just described (Eq. (66)), we can bring $d(\Lambda)$ for both into the form (67a). The equivalence of two representations of the little group must be defined as the existence of a *unitary* transformation which transforms them into each other. (Only unitary transformations are used for the whole group, also).

On the other hand, if the representations of the whole group are equivalent, the representations of the little group are equivalent also: the representation of the whole group determines the representation of the little group up to a similarity transformation uniquely.

The representation of the little group was defined as the set of matrices $Q(p_0, \lambda)_{\zeta\eta}$ if the representation is so transformed that (53) and (66a) hold. Having two equivalent representations D and $SDS^{-1} = D^\circ$ for both of which (53) and (66a) holds, the unitary transformation S bringing the first into the second must leave all displacement operators invariant. Hence, it must have the form (57a), i.e., operate on the ζ only and depend on p only as on a parameter.

$$S\varphi(p, \zeta) = \sum_\eta S(p)_{\zeta\eta}\varphi(p, \eta). \tag{68}$$

Denoting the matrix Q for the two representations by Q and Q°, the condition $SD(\Lambda) = D^\circ(\Lambda)S$ gives that

$$\sum_\eta S(p)_{\zeta\eta}Q(p, \Lambda)_{\eta\vartheta} = \sum_\eta Q^\circ(p, \Lambda)_{\zeta\eta}S(\Lambda^{-1}p)_{\eta\vartheta} \tag{68a}$$

holds, for every Λ, for almost every p. Setting $\Lambda = \alpha(p_1)$ we can let p approach p_1 in such a way that (68a) remains valid. Since Q is a continuous function of p both $Q(p, \Lambda)$ and $Q^\circ(p, \Lambda)$ will approach their limiting value 1. It follows that there is no domain in which

$$S(p_1) = S(\alpha(p_1)^{-1}p_1) = S(p_0) \tag{69}$$

would not hold, i.e., that (69) holds for almost every p_1. Since all our equations must hold only for almost every p, the $S(p)_{\zeta\eta}$ can be assumed to be independent

of p and (68a) then to hold for every p also. It then follows that the representations of the little group in D and D° are transformed into each other by $S_{\xi\eta}$.

The definition of the little group involved an arbitrarily chosen momentum vector p_0. It is clear, however, that the little groups corresponding to two different momentum vectors p_0 and p are holomorphic. In fact they can be transformed into each other by $\alpha(p)$: If Λ is an element of the little group leaving p invariant then $\alpha(p)^{-1}\Lambda\alpha(p) = \beta$ is an element of the little group which leaves p_0 invariant. We can see furthermore from (67a) that if Λ is in the little group corresponding to p, i.e. $\Lambda p = p$ then the representation matrix $q(\beta)$ of the little group of p_0, corresponding to β, is identical with the representation matrix of the little group of p, corresponding to $\Lambda = \alpha(p)\beta\alpha(p)^{-1}$. Thus when characterizing a representation of the whole inhomogeneous Lorentz group by P and the representation of the little group, it is not necessary to say which p_0 is left invariant by the little group.

D.

Lastly we shall determine the constitution of the little group in the different cases.

1_+. In case 1_+ we can take for p_0 the vector with the components 0, 0, 0, 1. The little group which leaves this invariant obviously contains all rotations in the space of the first three coördinates. This holds for the little group of all representations of the first class.

0_0. In case 0_0, the little group is the whole homogeneous Lorentz group.

-1. In case $P = -1$ the p_0 can be assumed to have the components 1, 0, 0, 0. The little group then contains all transformations which leave the form $-x_2^2 - x_3^2 + x_4^2$ invariant, i.e., is the $2 + 1$ dimensional homogeneous Lorentz group. The same holds for all representations with $P < 0$.

0_+. The determination of the little group for $P = 0_+$ is somewhat more complicated. It can be done, however, rather simply, for the group of unimodular two dimensional matrices. The Lorentz transformation corresponding to the matrix $\begin{Vmatrix} a & b \\ c & d \end{Vmatrix}$ with $ad - bc = 1$ brings the vector with the components x_1, x_2, x_3, x_4, into the vector with the components x_1', x_2', x_3', x_4', where[18]

$$\begin{Vmatrix} a & b \\ c & d \end{Vmatrix} \begin{Vmatrix} x_4 + x_3 & x_1 + ix_2 \\ x_1 - ix_2 & x_4 - x_3 \end{Vmatrix} \begin{Vmatrix} a^* & c^* \\ b^* & d^* \end{Vmatrix} = \begin{Vmatrix} x_4' + x_3' & x_1' + ix_2' \\ x_1' - ix_2' & x_4' - x_3' \end{Vmatrix}. \tag{70}$$

The condition that a null-vector p_0, say with the components 0, 0, 1, 1 be invariant is easily found to be $|a|^2 = 1$, $c = 0$. Hence the most general element of the little group can be written

$$\begin{Vmatrix} e^{-i\beta/2} & (x + iy)e^{i\beta/2} \\ 0 & e^{i\beta/2} \end{Vmatrix}, \tag{71}$$

with real x, y, β and $0 \leqq \beta < 4\pi$. The general element (71) can be written as $t(x, y)\ \delta(\beta)$ where

$$t(x, y) = \begin{Vmatrix} 1 & x + iy \\ 0 & 1 \end{Vmatrix}; \qquad \delta(\beta) = \begin{Vmatrix} e^{-i\beta/2} & 0 \\ 0 & e^{i\beta/2} \end{Vmatrix}. \tag{71a}$$

The multiplication rules for these are

$$t(x, y)t(x', y') = t(x + x', y + y'), \tag{71b}$$

$$\delta(\beta)t(x, y) = t(x \cos \beta + y \sin \beta, - x \sin \beta + y \cos \beta)\delta(\beta), \tag{71c}$$

$$\delta(\beta)\delta(\beta') = \delta(\beta + \beta'). \tag{71d}$$

One could restrict the variability domain of β in $\delta(\beta)$ from 0 to 2π. As $\delta(2\pi)$ commutes with all elements of the little group, it will be a constant and from $\delta(2\pi)^2 = \delta(4\pi) = 1$ it can be $\delta(2\pi) = \pm 1$. Hence $\delta(\beta + 2\pi) = \pm\delta(\beta)$ and inserting a $\pm$ into equation (71d) one could restrict β to $0 \leqq \beta < 2\pi$.

These equations are analogous to the equations (52)–(52D) and show that the little group is, in this case, isomorphic with the inhomogeneous rotation group of two dimensions, i.e. the two dimensional Euclidean group.

It may be mentioned that the Lorentz transformations corresponding to $t(x, y)$ have elementary divisors, and constitute all transformations of class e) in 4B, for which $v_e = p_0$. The transformations $\delta(\beta)$ can be considered to be rotations in the ordinary three dimensional space, about the direction of the space part of the vector p_0. It is possible, then, to prove equations (71) also directly.

7. The Representations of the Little Groups

A. Representations of the three dimensional rotation group by unitary transformations.

The representations of the three dimensional rotation group in a space with a finite member of dimensions are well known. There is one irreducible representation with the dimensions 1, 2, 3, 4, $\cdots$ each, the representations with an odd number of dimensions are single valued, those with an even number of dimensions are two-valued. These representations will be denoted by $D^{(j)}(R)$ where the dimension is $2j + 1$. Thus for single valued representations j is an integer, for double valued representations a half integer. Every finite dimensional representation can be decomposed into these irreducible representations. Consequently those representations of the Lorentz group with positive P in which the representation of the little group—as defined by (64)—has a finite number of dimensions, can be decomposed into such representations in which the representation of the little group is one of the well known irreducible representations of the rotation group. This result will hold for all representations of the inhomogeneous Lorentz group with positive P, since we shall show that even the infinite dimensional representations of the rotation group can be decomposed into the same, finite, irreducible representations.

In the following, it is more appropriate to consider the subgroup of the two dimensional unimodular group which corresponds to rotations, than the rotation group itself, as we can restrict ourselves to single valued representations in this case (cf. equations (52)). From (70), one easily sees[18] that the condition for $\begin{Vmatrix} a & b \\ c & d \end{Vmatrix}$ to leave the vector with the components 0, 0, 0, 1 invariant is that it shall be unitary. It is, therefore, the two dimensional unimodular unitary group the representations of which we shall consider, instead of the representations of the rotation group.

Let us introduce a discrete coördinate system in the representation space and denote the coefficients of the unitary representation by $q(R)_{\kappa\lambda}$ where R is a two dimensional unitary transformation. The condition for the unitary character of the representation $q(R)$ gives

$$\sum_{\kappa} q(R)^*_{\kappa\lambda} q(R)_{\kappa\mu} = \delta_{\lambda\mu}; \qquad \sum_{\lambda} q(R)^*_{\kappa\lambda} q(R)_{\nu\lambda} = \delta_{\kappa\nu}, \tag{72}$$

$$\sum_{\kappa} | q(R)_{\kappa\lambda} |^2 = 1; \qquad \sum_{\lambda} | q(R)_{\kappa\lambda} |^2 = 1. \tag{72a}$$

This show also that $| q(R)_{\kappa\lambda} | \leqq 1$ and the $q(R)_{\kappa\lambda}$ are therefore, as functions of R, square integrable:

$$\int | q(R)_{\kappa\lambda} |^2 \, dR$$

exists if $\int \cdots dR$ is the well known invariant integral in group space. Since this is finite for the rotation group (or the unimodular unitary group), it can be normalized to 1. We then have

$$\sum_{\kappa} \int | q(R)_{\kappa\lambda} |^2 dR = \sum_{\lambda} \int | q(R)_{\kappa\lambda} |^2 dR = 1. \tag{73}$$

The $(2j + 1)^{\frac{1}{2}} D^{(j)}(R)_{kl}$ form,[26] a complete set of normalized orthogonal functions for R. We set

$$q(R)_{\kappa\lambda} = \sum_{jkl} C^{\kappa\lambda}_{jkl} D^{(j)}(R)_{kl}. \tag{74}$$

We shall calculate now the integral over group space of the product of $D^{(j)}(R)^*_{kl}$ and

$$q(RS)_{\kappa\mu} = \sum_{\lambda} q(R)_{\kappa\lambda} q(S)_{\lambda\mu}. \tag{75}$$

The sum on the right converges uniformly, as for (72a)

$$\sum_{\lambda=N}^{\infty} | q(R)_{\kappa\lambda} q(S)_{\lambda\mu} | \leqq \left(\sum_{\lambda=N}^{\infty} | q(R)_{\kappa\lambda} |^2 \sum_{\lambda=N}^{\infty} | q(S)_{\lambda\mu} |^2 \right)^{\frac{1}{2}} \leqq \left(\sum_{\lambda=N}^{\infty} | q(S)_{\lambda\mu} |^2 \right)^{\frac{1}{2}}$$

can be made arbitrarily small by choosing an N, independent of R, making the last expression small. Hence, (75) can be integrated term by term and gives

$$\int D^{(j)}(R)^*_{kl} q(RS)_{\kappa\mu} dR = \sum_{\lambda} \int D^{(j)}(R)^*_{kl} q(R)_{\kappa\lambda} q(S)_{\lambda\mu} dR. \tag{76}$$

[26] H. Weyl and F. Peter, Math. Annal. *97*, 737, 1927.

Substituting $\sum_m D^{(j)}(RS)_{km} D^{(j)}(S^{-1})_{ml}$ for $D^{(j)}(R)_{kl}$ one obtains

$$(77) \quad \sum_m D^{(j)}(S^{-1})^*_{ml} \int D^{(j)}(RS)^*_{km} q(RS)_{\kappa\mu} dR = \sum_\lambda q(S)_{\lambda\mu} \int D^{(j)}(R)^*_{kl} q(R)_{\kappa\lambda} dR.$$

In the invariant integral on the left of (77), R can be substituted for RS and we obtain, for (74) and the unitary character

$$(78) \qquad \sum_m D^{(j)}(S)_{lm} C^{\kappa\mu}_{jkm} = \sum_\lambda q(S)_{\lambda\mu} C^{\kappa\lambda}_{jkl}.$$

Multiplying (78) by $D^{(h)}(S)^*_{in}$, the integration on the right side can be carried out term by term again, since the sum over λ converges uniformly

$$\sum_{\lambda=N}^{\infty} | C^{\kappa\lambda}_{jkl} q(S)_{\lambda\mu} | \leqq \left(\sum_{\lambda=N}^{\infty} | C^{\kappa\lambda}_{jkl} |^2 \sum_{\lambda=N}^{\infty} | q(S)_{\lambda\mu} |^2 \right)^{\frac{1}{2}} \leqq \left(\sum_{\lambda=N}^{\infty} | C^{\kappa\lambda}_{jkl} |^2 \right)^{\frac{1}{2}}.$$

This can be made arbitrarily small, as even $\sum_\lambda \sum_{jkl} (2j+1)^{-1} | C^{\kappa\lambda}_{jkl} |^2$ converges, for (74) and (72a). The integration of (78) yields thus

$$(79) \qquad \sum_\lambda C^{\kappa\lambda}_{jkl} C^{\lambda\mu}_{hin} = \delta_{jh} \delta_{li} C^{\kappa\mu}_{jkn}.$$

From $q(R)q(E) = q(R)$ follows $q(E) = 1$ and then $q(R^{-1}) = q(R)^{-1} = q(R)^\dagger$. This, with the similar equation for $D^{(j)}(R)$ gives

$$(80) \qquad \begin{aligned} \sum_{jkl} C^{\kappa\lambda}_{jkl} D^{(j)}(R^{-1})_{kl} &= q(R^{-1})_{\kappa\lambda} = q(R)^*_{\lambda\kappa} \\ &= \sum_{jkl} C^{\lambda\kappa *}_{jlk} D^{(j)}(R)^*_{lk} = \sum_{jkl} C^{\lambda\kappa *}_{jlk} D^{(j)}(R^{-1})_{kl}, \end{aligned}$$

or

$$(81) \qquad C^{\kappa\lambda}_{jkl} = C^{\lambda\kappa *}_{jlk}.$$

On the other hand $q(E)_{\kappa\lambda} = \delta_{\kappa\lambda}$ yields

$$(82) \qquad \sum_{jk} C^{\kappa\lambda}_{jkk} = \delta_{\kappa\lambda}.$$

These formulas suffice for the reduction of $q(R)$. Let us choose for every finite irreducible representation $D^{(j)}$ an index k, say $k = 0$. We define then, in the original space of the representation $q(R)$ vectors $v^{(\kappa jl)}$ with the components

$$C^{\kappa 1}_{jkl}, C^{\kappa 2}_{jkl}, C^{\kappa 3}_{jkl}, \cdots.$$

The vectors $v^{(\kappa jl)}$ for different j or l are orthogonal, the scalar product of those with the same j and l is independent of l. This follows from (79) and (81)

$$(83) \qquad (v^{(\mu j' l')}, v^{(\kappa jl)}) = \sum_\lambda C^{\mu\lambda *}_{j'kl'} C^{\kappa\lambda}_{jkl} = \sum_\lambda C^{\kappa\lambda}_{jkl} C^{\lambda\mu}_{j'l'k} = \delta_{jj'} \delta_{ll'} C^{\kappa\mu}_{jkk}.$$

The $v^{(\kappa jl)}$ for all κ, j, l, form a complete set of vectors. In order to show this, it is sufficient to form, for every ν, a linear combination from them, the ν component of which is 1, all other components 0. This linear combination is

$$(84) \qquad \sum_{\kappa jl} C^{\nu\kappa}_{jlk} v^{(\kappa jl)}.$$

In fact, the λ component of (84) is, on account of (79) and (82)

$$\sum_{\kappa j l} C^{\nu\kappa}_{jlk} C^{\kappa\lambda}_{jkl} = \sum_{jl} C^{\nu\lambda}_{jll} = \delta_{\nu\lambda}. \tag{85}$$

However, two v with the same j and l but different first indices κ are not orthogonal. We can choose for every j an l, say $l = 0$ and go through the vectors $v^{(1j0)}, v^{(2j0)}, \cdots$ and, following Schmidt's method, orthogonalize and normalize them. The vectors obtained in this way shall be denoted by

$$w^{(nj0)} = \sum_{\lambda} \alpha^{j}_{n\lambda} v^{(\lambda j0)}. \tag{86}$$

Then, since according to (83) the scalar products $(v^{(\kappa jl)}, v^{(\lambda jl)})$ do not depend on l, the vectors

$$w^{(njl)} = \sum_{\lambda} \alpha^{j}_{n\lambda} v^{(\lambda jl)} \tag{86a}$$

will be mutually orthogonal and normalized also and the vectors $w^{(njl)}$ for all n, j, l will form a complete set of orthonormal vectors. The same holds for the set of the conjugate complex vectors $w^{(njl)*}$. Using these vectors as coördinate axes for the original representation $q(R)$, we shall find that $q(R)$ is completely reduced. The ν component of the vector $q(R)v^{(\kappa jl)*}$ obtained by applying $q(R)$ on $v^{(\kappa jl)*}$ is

$$\sum_{\mu} q(R)_{\nu\mu}(v^{(\kappa jl)*})_{\mu} = \sum_{\mu} q(R)_{\nu\mu} C^{\mu\kappa}_{jlk}. \tag{87}$$

The right side is uniformly convergent. Hence, its product with $(2h + 1)$ $D^{(h)}(R)^{*}_{in}$ can be integrated term by term giving

$$\sum_{\mu} \int (2h + 1) D^{(h)}(R)^{*}_{in} q(R)_{\nu\mu} C^{\mu\kappa}_{jlk}\, dR = \sum_{\mu} C^{\nu\mu}_{hin} C^{\mu\kappa}_{jlk} = \delta_{hj}\delta_{ln} C^{\nu\kappa}_{jik}. \tag{88}$$

Thus we have for almost all R

$$\sum_{\mu} q(R)_{\nu\mu}(v^{(\kappa jl)*})_{\mu} = \sum_{i} C^{\nu\kappa}_{jik} D^{(j)}(R)_{il} = \sum_{i} D^{(j)}(R)_{il}(v^{(\kappa ji)*})_{\nu}, \tag{88a}$$

or

$$q(R)v^{(\kappa jl)*} = \sum_{i} D^{(j)}(R)_{il} v^{(\kappa ji)*}. \tag{88b}$$

Since both sides are supposed to be strongly continuous functions of R, (88b) holds for every R. In (86a), for every n, the summation must be carried out only over a finite number of λ. We can write therefore immediately

$$q(R)w^{(njl)*} = \sum_{i} D^{(j)}(R)_{il} w^{(nji)*}. \tag{89}$$

This proves that the original representation decomposes in the coördinate system of the w into well known finite irreducible representations $D^{(j)}(R)$. Since the w form a complete orthonormal set of vectors, the transition corresponds to a unitary transformation.

This completes the proof of the complete reducibility of all (finite and infinite dimensional) representations of the rotation group or unimodular unitary group. It is clear also that the same consideration applies for all closed groups, i.e., whenever the invariant integral $\int dR$ converges.

The result for the inhomogeneous Lorentz group is: For every positive numerical value of P, the representations of the little group can be, in an irreducible representation, only the $D^{(0)}, D^{(\frac{1}{2})}, D^{(1)}, \cdots$, both for P_+ and for P_-. All these representations have been found already by Majorana and by Dirac and for positive P there are none in addition to these.

B. Representations of the two dimensional Euclidean group

This group, as pointed out in Section 6, has a great similarity with the inhomogeneous Lorentz group. It is possible, again[24], to introduce "momenta", i.e. variables ξ, η and υ instead of the ζ in such a way that

$$t(x, y)\varphi(p_0, \xi, \eta, \upsilon) = e^{i(x\xi+y\eta)}\varphi(p_0, \xi, \eta, \upsilon). \tag{90}$$

Similarly, one can define again operators $R(\beta)$

$$R(\beta)\varphi(p_0, \xi, \eta, \upsilon) = \varphi(p_0, \xi', \eta', \upsilon), \tag{91}$$

where

$$\begin{aligned} \xi' &= \xi \cos\beta - \eta \sin\beta, \\ \eta' &= \xi \sin\beta + \eta \cos\beta. \end{aligned} \tag{91a}$$

Then $\delta(\beta)R(\beta)^{-1} = S(\beta)$ will commute, on account of (71c), with $t(x, y)$ and again contain ξ, η as parameter only. The equation corresponding to (57a) is

$$\delta(\beta)\varphi(p_0, \xi, \eta, \upsilon) = \sum_{\omega} S(\beta)_{\upsilon\omega}\varphi(p_0, \xi', \eta', \omega). \tag{92}$$

One can infer from (90) and (92) again that the variability domain of ξ, η can be restricted in such a way that all pairs ξ, η arise from one pair ξ_0, η_0 by a rotation, according (91a). We have, therefore two essentially different cases:

a.) $\xi^2 + \eta^2 = \Xi \neq 0$

b.) $\xi^2 + \eta^2 = \Xi = 0$, i.e. $\xi = \eta = 0$.

The positive definite metric in the ξ, η space excludes the other possibilities of section 6 which were made possible by the Lorentzian metric for the momenta, necessitated by (55).

Case b) can be settled very easily. The "little group" is, in this case, the group of rotations in a plane and we are interested in one and two valued irreducible representations. These are all one dimensional ($e^{is\beta}$)

$$S(\beta) = e^{is\beta} \tag{93}$$

where s is integer or half integer. These representations were also all found by Majorana and by Dirac. For $s = 0$ we have simply the equation $\Box\varphi = 0$,

for $s = \pm\frac{1}{2}$ Dirac's electron equation without mass, for $s = \pm 1$ Maxwell's electromagnetic equations, etc

In case a) the little group consists only of the unit matrix and the matrix $\begin{Vmatrix} -1 & 0 \\ 0 & -1 \end{Vmatrix}$ of the two dimensional unimodular group. This group has two irreducible representations, as (1) and (-1) can correspond to the above two dimensional matrix of the little group. This gives two new representations of the whole inhomogeneous Lorentz group, corresponding to every numerical value of Ξ. Both these sets belong to class 0_+ and two similar new sets belong to class 0_-.

The final result is thus as follows: The representations P_{+j} of the first subclass P_+ can be characterized by the two numbers P and j. From these P is positive, otherwise arbitrary, while j is an integer or a half integer, positive, or zero. The same holds for the subclass P_-. There are three kinds of representations of the subclass 0_+. Those of the first kind 0_{+s} can be characterized by a number s, which can be either an integer or a half integer, positive, negative or zero. Those of the second kind $0_+(\Xi)$ are single valued and can be characterized by an arbitrary positive number Ξ, those of the third kind $0'_+(\Xi)$ are double-valued and also can be characterized by a positive Ξ. The same holds for the subclass 0_-. The representations of the other classes (0_0 and P with $P < 0$) have not been determined.

8. Representations of the Extended Lorentz Group

A.

As most wave equations are invariant under a wider group than the one investigated in the previous sections, and as it is very probable that the laws of physics are all invariant under this wider group, it seems appropriate to investigate now how the results of the previous sections will be modified if we go over from the "restricted Lorentz group" defined in section 4A, to the extended Lorentz group. This extended Lorentz group contains in addition to the translations all the homogeneous transformations X satisfying (10)

$$XFX' = F \tag{10'}$$

while the homogeneous transformations of section 4A were restricted by two more conditions. From (10′) it follows that the determinant of X can be $+1$ or -1 only. If its -1, the determinant of $X_1 = XF$ is $+1$. If the four-four element of X_1 is negative, that of $X_2 = -X_1$ is positive. It is clear, therefore, that if X is a matrix of the extended Lorentz group, one of the matrices X, XF, $-X$, $-XF$ is in the restricted Lorentz group. For $F^2 = 1$, conversely, all homogeneous transformations of the extended Lorentz group can be obtained from the homogeneous transformations of the restricted group by multiplication with one of the matrices

$$1, F, -1, -F. \tag{94}$$

The group elements corresponding to these transformations will be denoted by E, F, I, IF. The restricted group contains those elements of the extended group which can be reached continuously from the unity. It follows that the transformation of an element L of the restricted group by F, I, or IF gives again an element of the restricted group. This is, therefore, an invariant subgroup of the extended Lorentz group. In order to find the representations of the extended Lorentz group, we shall use again Frobenius' method.[15]

We shall denote the operators corresponding in a representation to the homogeneous transformations (94) by $d(E) = 1, d(F), d(I), d(IF)$. For deriving the equations (52) it was necessary only to assume the existence of the transformations of the restricted group, it was not necessary to assume that these are the only transformations. These equations will hold, therefore, for elements of the restricted group, in representations of the extended group also. We normalize the indeterminate factors in $d(F)$ and $d(I)$ so that their squares become unity. Then we have $d(F)d(I) = \omega d(I)d(F)$ or $d(I) = \omega d(F)d(I)d(F)$. Squaring this, one obtains $\omega^2 = \pm 1$. We can set, therefore

$$\begin{aligned} d(IF) &= d(I)d(F) = \pm d(F)d(I) \\ d(F)^2 &= d(I)^2 = 1; \qquad d(IF)^2 = \pm 1. \end{aligned} \tag{95}$$

Finally, from

$$d(F)D(L_1)d(F) = \omega(L_1)D(FL_1F) \tag{96}$$

we obtain, multiplying this with the similar equation for L_2

$$\omega(L_1)\omega(L_2) = \omega(L_1L_2)$$

which, gives $\omega(L) = 1$ as the inhomogeneous Lorentz group (or the group used in (52B)-(52D)) has the only one dimensional representation by the unity (1). In this way, we obtain

$$d(F)D(L)d(F) = D(FLF), \tag{96a}$$

$$d(I)D(L)d(I) = D(ILI), \tag{96b}$$

$$d(IF)D(L)d(IF)^{-1} = D(IFLFI). \tag{96c}$$

B.

Given a representation of the extended Lorentz group, one can perform the transformations described in section 6A, by considering the elements of the restricted group only. We shall consider here only such representations of the extended group, for which, after having introduced the momenta, all representations of the restricted group are either in class 1 or 2, i.e. $P \geqq 0$ but not 0_0. Following then the procedure of section 6, one can find a set of wave functions for which the operators $D(L)$ of the restricted group have one of the forms, given in section 6 as irreducible representations. We shall proceed, next to find the operator $d(F)$. For the wave functions belonging to an irreducible $D(L)$ of the

restricted group, we can introduce a complete set of orthonormal functions $\psi_1(p, \zeta)$, $\psi_2(p, \zeta)$, $\cdots$. We then have

$$D(L)\psi_\kappa(p, \zeta) = \sum_\mu D(L)_{\mu\kappa}\psi_\mu(p, \zeta). \tag{97}$$

The infinite matrices $D(L)_{\mu\kappa}$ defined in (97) are unitary and form a representation which is equivalent to the representation by the operators $D(L)$. The $D(L)$, $d(F)$ are, of course, operators, but the $D(L)_{\mu\kappa}$ are components of a matrix, i.e. numbers. We can now form the wave functions $d(F)\psi_1$, $d(F)\psi_2$, $d(F)\psi_3$, $\cdots$ and apply $D(L)$ to these. For (96a) and (97) we have

$$\begin{aligned} D(L)d(F)\psi_\kappa &= d(F)D(FLF)\psi_\kappa \\ &= d(F)\sum_\mu D(FLF)_{\mu\kappa}\psi_\mu \\ &= \sum_\mu D(FLF)_{\mu\kappa}d(F)\psi_\mu . \end{aligned} \tag{97a}$$

The matrices $D^\circ(L)_{\mu\kappa} = D(FLF)_{\mu\kappa}$ give a representation of the restricted group (FLF is an element of the restricted group, we have a new representation by an automorphism, as discussed in section 6B). We shall find out whether $D^\circ(L)$ is equivalent $D(L)$ or not. The translation operation in D° is

$$T^\circ(a) = d(F)T(a)d(F) = T(Fa) \tag{98}$$

which, together with (53) shows that D° has the same P as $D(L)$ itself. In fact, writing

$$U_1\varphi(p, \zeta) = \varphi(Fp, \zeta) \tag{99}$$

one has $U_1^{-1} = U_1$ and one easily calculates $U_1T^\circ(a)U_1 = T(a)$. Similarly for $U_1d^\circ(\Lambda)U_1$ one has

$$\begin{aligned} U_1d^\circ(\Lambda)U_1\varphi(p, \zeta) &= U_1d(F\Lambda F)U_1\varphi(p, \zeta) \\ &= d(F\Lambda F)U_1\varphi(Fp, \zeta) = \sum_\eta Q(Fp, F\Lambda F)_{\zeta\eta}U_1\varphi(F\Lambda^{-1}p, \eta) \\ &= \sum_\eta Q(Fp, F\Lambda F)_{\zeta\eta}\varphi(\Lambda^{-1}p, \eta). \end{aligned} \tag{99a}$$

This means that the similarity transformation with U_1 brings $T^\circ(a)$ into $T(a)$ and $d^\circ(\Lambda)$ into $Q(Fp, F\Lambda F)P(\Lambda)$. Thus the representation of the "little group" in $U_1d^\circ(\Lambda)U_1$ is

$$q^\circ(\lambda) = Q(Fp_0, F\lambda F).$$

For this latter matrix, one obtains from (67a)

$$\begin{aligned} q^\circ(\lambda) &= Q(Fp_0, F\lambda F) = q(\alpha(Fp_0)^{-1}F\lambda F\alpha(Fp_0)) \\ &= q(\lambda^\circ) \end{aligned} \tag{100}$$

where λ° is obtained from λ by transforming it with $F\alpha(Fp_0)$.

The representations $D^\circ(L)$ and $D(L)$ are equivalent if the representation

$q(\lambda)$ is equivalent to the representation which coördinates $q(\lambda^\circ)$ to λ. The $\alpha(Fp_0)$ is a transformation of the restricted group which brings p_0 into $\alpha(Fp_0)p_0 = Fp_0$. (Cf. (65).) This transformation is, of course, not uniquely determined but if $\alpha(Fp_0)$ is one, the most general can be written as $\alpha(Fp_0)\iota$, where $\iota p_0 = p_0$ is in the little group. For $q(\iota^{-1}\alpha(Fp_0)^{-1} \Lambda\, \alpha(Fp_0)\iota) = q(\iota)^{-1}q(\alpha(Fp_0)^{-1} \Lambda\, \alpha(Fp_0))\, q(\iota)$, the freedom in the choice of $\alpha(Fp_0)$ only amounts to a similarity transformation of $q^\circ(\lambda)$ and naturally does not change the equivalence or non equivalence of $q^\circ(\lambda)$ with $q(\lambda)$.

For the case P_+, we can choose p_0 in the direction of the fourth axis, with components 0, 0, 0, 1. Then $Fp_0 = p_0$ and $\alpha(Fp_0) = 1$. The little group is the group of rotations in ordinary space and $F\lambda F = \lambda$. Hence $q^\circ(\lambda) = q(\lambda)$ and $D^\circ(\Lambda)$ is equivalent to $D(\Lambda)$ in this case. The same holds for the representations of class P_-.

For 0_+ we can assume that p_0 has the components 0, 0, 1, 1. Then the components of Fp_0 are 0, 0, -1, 1. For $\alpha(Fp_0)$ we can take a rotation by π about the second axis and $F\alpha(Fp_0)$ will be a diagonal matrix with diagonal elements 1, -1, 1, 1, i.e., a reflection of the second axis. Thus if λ is the transformation in (70), $\lambda^0 = \alpha(Fp_0)^{-1}F\lambda F\alpha(Fp_0)$ is the transformation for which

$$\lambda^0\begin{pmatrix} x_4 + x_3 & x_1 - ix_2 \\ x_1 + ix_2 & x_4 - x_3 \end{pmatrix}\lambda^{0\dagger} = \begin{pmatrix} x_4' + x_3' & x_1' - ix_2' \\ x_1' + ix_2' & x_4' - x_3' \end{pmatrix}. \tag{101}$$

This is, however, clearly $\lambda^0 = \lambda^*$. Thus the operators of $q^0(\lambda)$ are obtained from the operators $q(\lambda)$ by (cf. (71a))

$$\begin{aligned} t^0(x, y) &= t(x, -y) \\ \delta^0(\beta) &= \delta(-\beta). \end{aligned} \tag{101a}$$

For the representations 0_{+s} with discrete s, the $q^0(\lambda)$ and $q(\lambda)$ are clearly inequivalent as $\delta^0(\beta) = (e^{-is\beta})$ and $\delta(\beta) = (e^{is\beta})$, except for $s = 0$, when they are equivalent. For the representations $0_+(\Xi)$, $0_+'(\Xi)$, the $q^0(\lambda)$ and $q(\lambda)$ are equivalent, both in the single valued and the double valued case, as the substitution $\eta \to -\eta$ transforms them into each other. The same holds for representations of the class 0_-. If $D^0(L)$ and $D(L)$ are equivalent

$$U^{-1}D^0(L)U = D(L), \tag{102}$$

the square of U commutes with all $D(L)$. As a consequence of this, U^2 must be a constant matrix. Otherwise, one could form, in well known manner,[27] an idempotent which is a function of U^2 and thus commutes with $D(L)$ also. Such an idempotent would lead to a reduction of the representation $D(L)$ of the restricted group. As a constant is free in U, we can set

$$U^2 = 1 \tag{102a}$$

[27] J. von Neumann, Ann. of Math. *32*. 191. 1931; ref. 2. p. 89.

C.

Returning now to equation (97a), if $D^0(L) = D(FLF)$ and $D(L)$ are equivalent ($P > 0$ or 0_+, 0_- with continuous Ξ or $s = 0$) there is a unitary matrix $U_{\mu\nu}$, corresponding to U, such that

$$\text{(102b)} \qquad \begin{aligned} \sum_\mu D(FLF)_{\kappa\mu} U_{\mu\nu} &= \sum_\mu U_{\kappa\mu} D(L)_{\mu\nu} \\ \sum_\mu U_{\kappa\mu} U_{\mu\nu} &= \delta_{\kappa\nu} . \end{aligned}$$

Let us now consider the functions

$$\text{(103)} \qquad \varphi_\nu = \psi_\nu + \sum_\mu U_{\mu\nu} d(F) \psi_\mu .$$

Applying $D(L)$ to these

$$\text{(103a)} \qquad \begin{aligned} D(L)\varphi_\nu &= D(L)\psi_\nu + \sum_\mu U_{\mu\nu} D(L) d(F) \psi_\mu \\ &= D(L)\psi_\nu + \sum_\mu U_{\mu\nu} d(F) D(FLF) \psi_\mu \\ &= \sum_\mu D(L)_{\mu\nu} \psi_\mu + \sum_{\mu\kappa} U_{\mu\nu} d(F) D(FLF)_{\kappa\mu} \psi_\kappa \\ &= \sum_\mu D(L)_{\mu\nu} \left(\psi_\mu + \sum_\kappa U_{\kappa\mu} d(F) \psi_\kappa \right) = \sum_\mu D(L)_{\mu\nu} \varphi_\mu . \end{aligned}$$

Similarly

$$\text{(103b)} \qquad \begin{aligned} d(F)\varphi_\nu &= d(F)\psi_\nu + \sum_\mu U_{\mu\nu} \psi_\mu \\ &= \sum_\mu U_{\mu\nu} \left(\psi_\mu + \sum_\kappa U_{\kappa\mu} d(F) \psi_\kappa \right) = \sum_\mu U_{\mu\nu} \varphi_\mu . \end{aligned}$$

Thus the wave functions φ transform according to the representation in which $D(L)_{\mu\nu}$ corresponds to L and $U_{\mu\nu}$ to $d(F)$. The same holds for the wave functions

$$\text{(104)} \qquad \varphi'_\nu = \psi_\nu - \sum_\mu U_{\mu\nu} d(F) \psi_\mu ,$$

except that in this case $(-U_{\mu\nu})$ corresponds to $d(F)$. The ψ_ν and $d(F)\psi_\nu$ can be expressed by the φ and φ'. If the ψ and $d(F)\psi$ were linearly independent, the φ and φ' will be linearly independent also. If the $d(F)\psi$ were linear combinations of the ψ, either the φ or the φ' will vanish.

If we imagine a unitary representation of the group formed by the L and FL in the form in which it is completely reduced out as a representation of the group of restricted transformations L, the above procedure will lead to a reduction of that part of the representation of the group of the L and FL, for which $D(L)$ and $D(FLF)$ are equivalent.

If $D(L)_{\mu\nu}$ and $D^\circ(L)_{\mu\nu}$ are inequivalent, the ψ_κ and $d(F)\psi_\nu = \psi'_\nu$ are orthogonal. This is again a generalization of the similar rule for finite unitary

representations.[28] One can see this in the following way: Denoting $M_{\kappa\nu} = (\psi_\kappa, \psi'_\nu)$ one has

$$M_{\kappa\nu} = (\psi_\kappa, \psi'_\nu) = (D(L)\psi_\kappa, D(L)\psi'_\nu)$$
$$= \sum_{\mu\lambda} D(L)^*_{\mu\kappa} D^\circ(L)_{\lambda\nu} M_{\mu\lambda};$$
$$M = D(L)^\dagger M D^\circ(L).$$

Hence

$$D(L)M = MD^\circ(L); \qquad M^\dagger D(L) = D^\circ(L)M^\dagger. \tag{105}$$

From these, one easily infers that $MM^\dagger$ commutes with $D(L)$, and $M^\dagger M$ commutes with $D^\circ(L)$. Hence both are constant matrices, and if neither of them is zero, M and $M^\dagger$ are, apart from a constant, unitary. Thus $D(L)$ would be equivalent $D^\circ(L)$ which is contrary to supposition. Hence $MM^\dagger = 0$, $M = 0$ and the ψ are orthogonal to the $d(F)\psi = \psi'$. Together, they give a representation of the group formed by the restricted Lorentz group and F. If they do not form a complete set, the reduction can be continued as before.

One sees, thus, that introducing the operation F "doubles" the number of dimensions of the irreducible representations in which the little group was the two dimensional rotation group, while it does not increase the underlying linear manifold in the other cases. This is analogous to what happens, if one adjoins the reflection operation to the rotation groups themselves.[29]

D.

The operations $d(I)$ can be determined in the same manner as the $d(F)$ were found. A complete set of orthonormal functions corresponding to an irreducible representation of the group formed by the L and FL shall be denoted by ψ_1, $\psi_2, \cdots$. For this, we shall assume (97) again, although the $D(L)$ contained therein is now not necessarily irreducible for the restricted group alone but contains, in case of 0_{+s} or 0_{-s} and finite s, both s and $-s$. We shall set, furthermore

$$d(F)\psi_\kappa = \sum_\mu d(F)_{\mu\kappa}\psi_\mu. \tag{106}$$

We can form then the functions $d(I)\psi_1$, $d(I)\psi_2, \cdots$. The consideration, contained in (97a) shows that these transform according to $D(ILI)_{\mu\kappa}$ for the transformation L of the restricted group:

$$D(L)d(I)\psi_\kappa = \sum_\mu D(ILI)_{\mu\kappa} d(I)\psi_\mu. \tag{106a}$$

Choosing for L a pure translation, a consideration analogous to that performed in (98) shows that the set of momenta in the representation $L \to D\,(ILI)$ has the opposite sign to the set of momenta in the representation $D(L)$. If the latter

[28] Cf. e.g. E. Wigner, ref. 4, Chapter XII.

[29] I. Schur, Sitz. d. kön. Preuss. Akad. pages 189, 297, 1924.

belongs to a positive subclass, the former belongs to the corresponding negative subclass and conversely. Thus the adjunction of the transformation I always leads to a "doubling" of the number of states, the states of "negative energy" are attached to the system of possible states. One can describe all states ψ_1, $\psi_2, \cdots, d(I)\psi_1, d(I)\psi_2, \cdots$ by introducing momenta p_1, p_2, p_3, p_4 and restricting the variability domain of p by the condition $\{p, p\} = P$ alone without stipulating a definite sign for p_4.

As we saw before, the $d(I)\psi_1$, $d(I)\psi_2$, are orthogonal to the original set of wave functions $\psi_1, \psi_2, \cdots$. The result of the application of the operations $D(L)$ and $d(F)$ to the $\psi_1, \psi_2, \cdots$ (i.e., the representation of the group formed by the L, FL) was given in part C. The $D(L)d(I)\psi_\kappa$ are given in (106a). On account of the normalization of $d(I)$ we can set

$$d(I)d(I)\psi_\kappa = \psi_\kappa . \tag{106b}$$

For $d(F)d(I)\psi_\kappa$ we have two possibilities, according to the two possibilities in (95). We can either set

$$d(F)d(I)\psi_\kappa = d(I)d(F)\psi_\kappa = \sum_\mu d(F)_{\mu\kappa} d(I)\psi_\mu , \tag{107}$$

or

$$d(F)\cdot d(I)\psi_\kappa = -d(I)d(F)\psi_\kappa = -\sum_\mu d(F)_{\mu\kappa} d(I)\psi_\mu . \tag{107a}$$

Strictly speaking, we thus obtain two different representations. The system of states satisfying (107) could be distinguished from the system of states for which (107a) is valid, however, only if we could really perform the transition to a new coördinate system by the transformation I. As this is, in reality, impossible, the representations distinguished by (107) and (107a) are not different in the same sense as the previously described representations are different.

I am much indebted to the Wisconsin Alumni Research Foundation for their aid enabling me to complete this research.

MADISON, WIS.

Minimally Almost Periodic Groups

J. von Neumann and E. P. Wigner

Annals of Mathematics *41*, 746–750 (1940)

(Received November 7, 1939)

1. Given a group $\mathcal{G}$ it is of some interest to decide which elements of $\mathcal{G}$ can be "told apart" by almost periodic functions of $\mathcal{G}$ or, which is the same thing (cf. below) by finite dimensional bounded linear representations of $\mathcal{G}$. That is: *For two $a, b \in \mathcal{G}$ we define $a \sim b$ by either of these two properties*:

(I) *For every almost periodic function $f(x)$ in $\mathcal{G}$ $f(a) = f(b)$.*

(II) *For every finite dimensional linear unitary representation $D(x)$ of $\mathcal{G}$ $D(a) = D(b)$.*

The general theory of almost periodic functions in groups[1] deals with these questions. The equivalence of (I) and (II) is shown in Ap, p. 480, Theorem 33. It is also known there, Ap, p. 481, Theorem 34, that the above notion determines an invariant subgroup $\mathcal{G}_0$ of $\mathcal{G}$, such that $a \sim b$ holds if and only if a and b belong to the same coset of $\mathcal{G}_0$ in $\mathcal{G}$.

The smaller $\mathcal{G}_0$, the more almost-periodic functions—and finite-dimensional bounded linear representations—$\mathcal{G}$ possesses. Hence we termed $\mathcal{G}$ *maximally almost periodic if* $\mathcal{G}_0 = (1)$ (i.e., $a \sim b$ for $a = b$ only), and *minimally almost periodic* if $\mathcal{G}_0 = \mathcal{G}$ (i.e. always $a \sim b$), cf. Ap, p. 482, Def. 17.

Maximally almost-periodic groups $\mathcal{G}$ are easily constructed; obviously every finite dimensional, linear, unitary group is such. For more general examples cf. Ap, pp. 482–483, Theorem 36, A and B.

Minimally almost-periodic groups $\mathcal{G}$ are somewhat pathological; they are characterized by the fact that their only almost-periodic functions are the constants: or equivalently, their only finite dimensional, linear, unitary representation is $D(x) \equiv 1$. The existence of such groups was shown in Ap, p. 483, using an interesting theorem of B. L. van der Waerden, combined with earlier results of J. v. Neumann and the representation theory of the affine group. (Cf. loc. cit.)

Thus these examples are really obtained in a rather complicated way.

In this note we will construct examples of various kinds of groups, including minimally almost-periodic ones, by simpler direct methods. All these examples will be enumerably infinite (discrete) groups, whereas those of Ap (cf. above) are continuous (Lie) groups. We will also gain some general viewpoints concerning the nature of maximal and minimal almost periodicity.

2. If $\mathcal{G}'$ is an invariant subgroup of $\mathcal{G}$ with finite index, then its factor group $\mathcal{G}/\mathcal{G}'$ is finite. Hence the regular representation of $\mathcal{G}/\mathcal{G}'$ is finite dimensional,

[1] J. v. Neumann, *Almost periodic functions in a group*, Amer. Math. Soc. Trans. 36 (1934), pp. 445–492. To be quoted as "Ap".

linear, unitary. It is the same for $\mathcal{G}$. It represents any two different elements of $\mathcal{G}/\mathcal{G}'$—hence any two elements of which belong to different cosets of $\mathcal{G}'$ in $\mathcal{G}$—differently. Hence $\mathcal{G}_0 \subseteq \mathcal{G}'$.

We see therefore: *If $\mathcal{G}_1$ is the intersection of all invariant subgroups $\mathcal{G}'$ of $\mathcal{G}$ with finite index, then*

$$\mathcal{G}_0 \subseteq \mathcal{G}_1 .$$

We do not have in general, however, $\mathcal{G}_0 = \mathcal{G}_1$. Let $\mathcal{G}$ be the group of all proper rotations in 3 dimensions. Since $\mathcal{G}$ is itself a finite dimensional, linear, unitary group, so it is maximally almost periodic. Hence $\mathcal{G}_0 = (1)$. On the other hand one verifies easily that the only invariant subgroup $\mathcal{G}'$ of $\mathcal{G}$ with finite index is $\mathcal{G}' = \mathcal{G}$, therefore $\mathcal{G}_1 = \mathcal{G}$.

3. Abelian groups—if they are separable and locally compact, which is the case for the enumerably infinite (discrete) ones—are maximally almost periodic, cf. Ap, pp. 482–483, Theorem 36, B. Hence one might be led to expect that the minimally almost-periodic groups—especially the enumerably infinite ones—will be found at the other extreme: among the free groups. This is not so, however. The free group of $h(= 2, 3, \cdots)$ generators, $\mathcal{G} = \mathcal{F}^{(h)}$, is maximally almost periodic. Indeed, the reader who is familiar with the general theory of the $\mathcal{F}^{(h)}$ will find no difficulty in constructing for any $a \in \mathcal{F}^{(h)}$, $a \neq 1$, an invariant aubgroup $\mathcal{G}'$ of $\mathcal{F}^{(h)}$ with finite index, such that $a \notin \mathcal{G}'$. Hence we have $\mathcal{G}_1 = (1)$ and *a fortiori* $\mathcal{G}_0 = (1)$ for $\mathcal{G} = \mathcal{F}^{(h)}$, cf. §2 above. (For another example of a maximally almost-periodic group which is closely related to free groups, cf. $\mathcal{G} = \mathcal{G}^{(\beta)}$ in §5 below.)

Thus both extremes—Abelian groups as well as free groups—are maximally almost periodic. The minimally almost periodic groups must be somewhere in between.

4. Our examples are based on the following lemma:

LEMMA 1. *Let U be a finite dimensional unitary matrix. Assume that there exists for every s $(= 1, 2, \cdots)$ a $t = t(s)$ $(= 1, 2, \cdots)$ which is an integral multiple of s such that for a suitable pair of reciprocal matrices V_t, V_t^{-1} there is $U^t = V_t U V_t^{-1}$. Then $U = 1$.*

PROOF: Let n $(= 1, 2, \cdots)$ be the dimensionality (order) of the matrices under consideration. Let $\alpha_1, \cdots, \alpha_n$ be the characteristic values of U,—in an arbitrary order, but with the right multiplicities.

Consider a $t = t(s)$. U^t has the characteristic values $\alpha_1^t, \cdots, \alpha_n^t$ while $V_t U V_t^{-1}$ has again the characteristic values $\alpha_1, \cdots, \alpha_n$. Therefore the $\alpha_1, \cdots, \alpha_n$ must be a permutation of the $\alpha_1^t, \cdots, \alpha_n^t$.

Since there are infinitely many different $t = t(s)$ (as $t(s) \geqq s$ so $s \to \infty$ implies $t(s) \to \infty$), and only a finite number of $\alpha_1, \cdots, \alpha_n$, so for every $i = 1, \cdots, n$ two different t_i', t_i'' with $\alpha_i^{t_i'} = \alpha_i^{t_i''}$ exist. So $\alpha_i^{t_i'-t_i''} = 1$, that is: All α_i are roots of unity. Now choose s, and *a fortiori* $t = t(s)$, as a common multiple of the

root of unity exponents of $\alpha_1, \cdots, \alpha_n$. Then $\alpha_1^t = \cdots = \alpha_n^t = 1$, and as $\alpha_1, \cdots, \alpha_n$ are a permutation of $\alpha_1^t, \cdots, \alpha_n^t$, therefore $\alpha_1 = \cdots = \alpha_n = 1$.

So all characteristic values of U are equal to 1, and consequently $U = 1$.

From this we conclude immediately:

LEMMA 2. *Let $\mathcal{G}$ be a group, and $\mathcal{G}_0$ as described in §1 above. Consider a fixed $a \in \mathcal{G}$. Assume that there exists for every s $(= i, 2, \cdots)$ a $t = t(s)$ $(= i, 2, \cdots)$ which is an integral multiple of s, such that for a suitable $b_t \in \mathcal{G}$ there is $a^t = b_t a b_t^{-1}$. Then $a \in \mathcal{G}_0$.*

PROOF: Apply Lemma 1 to $U = D(a)$ and $V_t = D(b_t)$, where $D(x)$ is a finite dimensional, linear, unitary representation of $\mathcal{G}$. Then $D(a) = 1$, hence $a \in \mathcal{G}_0$ by (II) in §1.

5. We will now discuss the following groups as to their types of almost periodicity:

(α) $\mathcal{G} = \mathcal{G}^{(\alpha)}$, the group of all linear transformations

$$a'(u, v): \qquad x \to x' \equiv ux + v,$$

where $u \neq 0$, u, v rational.

(β) $\mathcal{G} = \mathcal{G}^{(\beta)}$, the group of all rational transformations

$$a''\begin{pmatrix} u, & v \\ w, & z \end{pmatrix}: \qquad x \to x'' \equiv \frac{ux + v}{wx + z}$$

where $\operatorname{Det}\begin{pmatrix} u, & v \\ w, & z \end{pmatrix} = uz - vw = 1$,

u, v, w, z rational integers.

(γ) $\mathcal{G} = \mathcal{G}^{(\gamma)}$, same as above, but u, v, w, z must be rational.

Discussion of (α): Since

$$a'(s, 0)a'(1, v)a'(s, 0)^{-1} = a'(1, sv) = (a'(1, v))^s$$

$(s = 1, 2, \cdots)$, so Lemma 2 applies to $a'(1, v)$ (with $t = t(s) = s$). Hence $a'(1, v) \in \mathcal{G}_0$.

On the other hand $D_\lambda(a'(u, v)) = (\exp(2\pi i\lambda \ln |u|))$ is a finite (one!) dimensional, linear, unitary representation of $\mathcal{G}^{(\alpha)}$ for every real λ. If $u \neq 1$, then $D_\lambda(a'(u, v)) \neq 1$ for some suitable real λ, hence $a'(u, v) \notin \mathcal{G}_0$ in this case.

Thus $\mathcal{G}_0$ is the set of all $a'(1, v)$, and consequently $\mathcal{G}$ is neither maximally nor minimally almost periodic.

Discussion of (β): This is the *modular group*; we observe (without giving the—well known—proofs, since these facts are not essential in our deductions): $\mathcal{G}^{(\beta)}$ is generated by the two elements

$$a_1 = a''\begin{pmatrix} 0, & -1 \\ 1, & 0 \end{pmatrix}: \qquad x \to x'' \equiv -\frac{1}{x},$$

$$a_2 = a''\begin{pmatrix} 0, & -1 \\ 1, & 1 \end{pmatrix}: \qquad x \to x'' \equiv -\frac{1}{x+1}.$$

We have $a_1^2 = a_2^3 = 1$, and these are the only relations which hold for a_1, a_2. So $\mathcal{G}^{(\beta)}$ is the free group of 2 generators of order 2 and 3.

The almost periodic character of $\mathcal{G}^{(\beta)}$ is determined as follows:

For any $p = 2, 3, \cdots$ the $a''\begin{pmatrix} u, & v \\ w, & z \end{pmatrix}$ with $\begin{pmatrix} u, & v \\ w, & z \end{pmatrix} \equiv \begin{pmatrix} 1, & 0 \\ 0, & 1 \end{pmatrix} \cdots \bmod p$ form an invariant subgroup $\mathcal{G}'_p$ of $\mathcal{G}$. Two $a''\begin{pmatrix} u, & v \\ w, & z \end{pmatrix}$ and $a''\begin{pmatrix} \bar{u}, & \bar{v} \\ \bar{w}, & \bar{z} \end{pmatrix}$ belong to the same coset of $\mathcal{G}'_p$ in $\mathcal{G}$ if and only if $\begin{pmatrix} u, & v \\ w, & z \end{pmatrix} \equiv \begin{pmatrix} \bar{u}, & \bar{v} \\ \bar{w}, & \bar{z} \end{pmatrix} \cdots \bmod p$. Hence the number of these cosets is finite, i.e.: the index of $\mathcal{G}'_p$ in $\mathcal{G}$ is finite.

Therefore if $a = a''\begin{pmatrix} u, & v \\ w, & z \end{pmatrix} \in \mathcal{G}_1$ (cf. §2 above), then $a''\begin{pmatrix} u, & v \\ w, & z \end{pmatrix} \in \mathcal{G}'_p$ for all $p = 2, 3, \cdots$, i.e.: $\begin{pmatrix} u, & v \\ w, & z \end{pmatrix} \equiv \begin{pmatrix} 1, & 0 \\ 0, & 1 \end{pmatrix} \cdots \bmod p$ for all $p = 2, 3, \cdots$. Consequently $\begin{pmatrix} u, & v \\ w, & z \end{pmatrix} = \begin{pmatrix} 1, & 0 \\ 0, & 1 \end{pmatrix}$, $a = 1$. Hence $\mathcal{G}_1 = (1)$, and so *a fortiori* $\mathcal{G}_0 = (1)$ (cf. §2 above). Thus $\mathcal{G}^{(\beta)}$ is maximally almost periodic.

Discussion of (γ): Since

$$a''\begin{pmatrix} s, & 0 \\ 0, & \frac{1}{s} \end{pmatrix} a''\begin{pmatrix} 1, & v \\ 0, & 1 \end{pmatrix} a''\begin{pmatrix} s, & 0 \\ 0, & \frac{1}{s} \end{pmatrix}^{-1} = a''\begin{pmatrix} 1, & s^2 v \\ 0, & 1 \end{pmatrix} = \left(a''\begin{pmatrix} 1, & v \\ 0, & 1 \end{pmatrix}\right)^{(s^2)},$$

$(s = i, 2, \cdots)$, so Lemma 2 applies to $a''\begin{pmatrix} 1, & v \\ 0, & 1 \end{pmatrix}$ (with $t = t(s) = s^2$). Hence $a''\begin{pmatrix} 1, & v \\ 0, & 1 \end{pmatrix} \in \mathcal{G}_0$.

Since $\mathcal{G}_0$ is an invariant subgroup of $\mathcal{G}$, we can now infer successively that the following further elements belong to $\mathcal{G}_0$:

$$a''\begin{pmatrix} 0, & -1 \\ 1, & 0 \end{pmatrix} a''\begin{pmatrix} 1, & -v \\ 0, & 1 \end{pmatrix} a''\begin{pmatrix} 0, & -1 \\ 1, & 0 \end{pmatrix}^{-1} = a''\begin{pmatrix} 1, & 0 \\ v, & 1 \end{pmatrix}$$

$$a''\begin{pmatrix} 1, & v_1 \\ 0, & 1 \end{pmatrix} a''\begin{pmatrix} 1, & 0 \\ v_2, & 1 \end{pmatrix} a''\begin{pmatrix} 1, & v_3 \\ 0, & 1 \end{pmatrix} = a''\begin{pmatrix} 1 + v_1 v_2, & v_1 + v_3 + v_1 v_2 v_3 \\ v_2 & 1 + v_2 v_3 \end{pmatrix}.$$

The general $a''\begin{pmatrix} u, & v \\ w, & z \end{pmatrix}$ with $uz - vw = 1$; u, v, w, z rational, has the above form if $w \neq 0$: For this purpose we must determine v_1, v_2, v_3 with

$$v_2 = w, \qquad 1 + v_1 v_2 = u, \qquad 1 + v_2 v_3 = z, \qquad v_1 + v_3 + v_1 v_2 v_3 = v.$$

The three first equations can be satified directly by choosing v_2, v_1, v_3 respectively. The last one then holds automatically, since $uz - vw = 1$. So $a''\begin{pmatrix} u, & v \\ w, & z \end{pmatrix} \in \mathcal{G}_0$ when $w \neq 0$. In particular $a''\begin{pmatrix} 0, & -1 \\ 1, & 0 \end{pmatrix} \in \mathcal{G}_0$. Since

$$a''\begin{pmatrix} -v, & w \\ -z, & -w \end{pmatrix} a''\begin{pmatrix} 0, & -1 \\ 1, & 0 \end{pmatrix} = a''\begin{pmatrix} u, & v \\ w, & z \end{pmatrix},$$

the same can be asserted when $z \neq 0$. And as $uz - vw = 1$ excludes $w = z = 0$, we have thus $a''\begin{pmatrix} u, & v \\ w, & z \end{pmatrix} \epsilon \mathfrak{G}_0$ unrestrictedly.

Hence $\mathfrak{G}_0 = \mathfrak{G}$, and consequently $\mathfrak{G}^{(\gamma)}$ is minimally almost-periodic.

It seems to be worth emphasizing that the fundamental difference between the groups $\mathfrak{G}^{(\beta)}$ and $\mathfrak{G}^{(\gamma)}$ is brought about by requiring u, v, w, z to be rational only, instead of requiring them to be rational integers.

As $\mathfrak{G}^{(\beta)}$ is a subgroup of $\mathfrak{G}^{(\gamma)}$, the interesting situation arises that there are representations of $\mathfrak{G}^{(\beta)}$ which are not contained in any representation of $\mathfrak{G}^{(\gamma)}$. It is well known that this could not occur for finite or compact groups.

We mention finally that the minimal almost-periodicity of the common Lorentz group can be readily inferred from earlier work.[2]

INSTITUTE FOR ADVANCED STUDY AND
PRINCETON UNIVERSITY

[2] E. P. Wigner, *On unitary representations of the inhomogeneous Lorentz group*, Annals of Math., 40 (1939), pp. 149–204. Cf. in particular pp. 164–168.

On Representations of Certain Finite Groups*

E. P. Wigner

American Journal of Mathematics *63*, 57–63 (1941)

1. The purpose of this paper is the derivation of a classification of the representations of finite groups with special reference to groups which satisfy the following two conditions:

a. Every element is equivalent to its reciprocal, i. e., all classes are ambivalent.

b. The Kronecker (or "direct") product of any two irreducible representations of the group contains no representation more than once.

Groups of this character will be called S. R. groups (simply reducible). The symmetric permutation groups of the third and fourth degree, the quaternion group, the three dimensional rotation group, the two dimensional unimodular unitary group are S. R. groups. . The significance of condition *a* is that every representation is equivalent to the conjugate imaginary representation. One sees this most easily by assuming the representation to be unitary. Then, the traces of reciprocal elements are conjugate complex. They are, on the other hand, equal, since they belong to the same class. Thus all traces are real and conjugate complex representations are equivalent.

The groups of most eigen-value problems occurring in quantum theory are S. R. This is important for the following reason.

Let us assume that we have two eigen-value problems $H_1\psi_1 = \lambda_1\psi_1$ and $H_2\psi_2 = \lambda_2\psi_2$ which allow the same group; ψ_1 and ψ_2 shall be defined in different spaces. One often considers then [1] the "united system" the wave functions Ψ of which are defined in the product space of the spaces of ψ_1 and ψ_2. The "unperturbed" eigen-value equation is $(H_1 + H_2)\Psi = \Lambda\Psi$. The multiplicity of the eigen-value $\Lambda = \lambda_1 + \lambda_2$ is the product of the multiplicities of λ_1 and λ_2. The eigen-value Λ splits up if one introduces a small perturbation term into the last equation. If this perturbation allows the same group as the original two problems and if this group satisfies the above condition *b*, the characteristic functions of the eigen-values into which Λ splits can be determined in "first approximation" by the invariance of the eigen-value problem under the group. The properties of S. R. groups to be derived here

* Received May 1, 1940.

[1] For a more complete discussion cf. e. g. E. Wigner: *Gruppentheorie*, etc., Braunschweig, 1931.

give a basis for a suitable normalization of (and numerous relations between) these eigen-functions which will be dealt with elsewhere.

We shall denote the different irreducible representations of a group by letters j, k, l, etc. The identical representation (in which the matrix (1) corresponds to every element) by 0. The elements of the group will be P, Q, R, S, T, etc. The rows and columns of the representations will be designated by small Greek letters κ, λ, μ, ν, etc. The $\kappa\lambda$ element of the matrix which corresponds in the j-th representation to the element R will be denoted by $\begin{bmatrix} jR \\ \kappa\lambda \end{bmatrix}$ so that one has

$$\sum_{\lambda} \begin{bmatrix} jR \\ \kappa\lambda \end{bmatrix} \begin{bmatrix} jS \\ \lambda\mu \end{bmatrix} = \begin{bmatrix} j\,RS \\ \kappa\ \mu \end{bmatrix} \tag{1}$$

The character will be abbreviated to

$$\sum_{\kappa} \begin{bmatrix} jR \\ \kappa\kappa \end{bmatrix} = [j\,;R]. \tag{1a}$$

The summation over the indices κ, λ etc. referring to the rows or columns of the representations will always run over all values. The unit element of the group will be E, the degree of the representations j is

$$[j\,;E] = [j]. \tag{1b}$$

A star will denote the conjugate complex. The Kronecker product of the representations j and k coordinates to the group element R the matrix $M_{\kappa\mu\,;\,\lambda\nu}$ the rows and columns of which are denoted by double indices $\kappa\mu$ and $\lambda\nu$ respectively. We set

$$M_{\kappa\mu\,;\,\lambda\nu} = \begin{bmatrix} jR \\ \kappa\lambda \end{bmatrix} \begin{bmatrix} kR \\ \mu\nu \end{bmatrix}. \tag{2}$$

The character corresponding to the element R is

$$\sum_{\kappa\mu} M_{\kappa\mu;\kappa\mu} = \sum_{\kappa\mu} \begin{bmatrix} jR \\ \kappa\kappa \end{bmatrix} \begin{bmatrix} kR \\ \mu\mu \end{bmatrix} = [j\,;R][k\,;R]. \tag{2a}$$

The significance of condition b for the groups under consideration becomes evident if one reduces the Kronecker product of two representations, i. e. brings it into the form in which it appears as the sum of irreducible representations. The matrix by which this transformation can be carried out is—apart from some phase factors—uniquely determined.

It may be useful to write down the well known orthogonality and completeness relations for irreducible representations. These are

$$\sum_{R} \begin{bmatrix} jR \\ \kappa\lambda \end{bmatrix}^{*} \begin{bmatrix} kR \\ \mu\nu \end{bmatrix} = \frac{h}{[j]}\, \delta_{jk}\delta_{\kappa\mu}\delta_{\lambda\nu} \tag{3}$$

$$\sum_R [j;R]^*[k;R] = \sum_C n_c[j;C][k;C] = h\delta_{jk}. \tag{3a}$$

The summation is to be extended in this and all similar formulas over all group elements;

$$h = \sum_R 1 \tag{4}$$

is the order of the group. The summation over C in the second part of (3a) is to be extended over all different classes, n_c is the number of elements of the class C. The completeness relations yield

$$\sum_{j\kappa\lambda} \frac{[j]}{h} \begin{bmatrix} jR \\ \kappa\lambda \end{bmatrix}^* \begin{bmatrix} jS \\ \kappa\lambda \end{bmatrix} = \delta_{R,S} \tag{5}$$

$$\sum_j [j;R]^*[j;S] = h\,\Delta_{R,S}/n_R = \sum_Q \delta_{RQ,QS}. \tag{5a}$$

The summation over j is to be extended over all different irreducible representations, $\delta_{R,S}$ is 1 for $R = S$, zero otherwise, $\Delta_{R,S}$ is 1 if R and S are in the same class, zero otherwise, n_R is the number of the elements of the class of R. All representations are assumed to be in the unitary form, i. e.

$$\sum_\lambda \begin{bmatrix} jR \\ \kappa\lambda \end{bmatrix}^* \begin{bmatrix} jR \\ \mu\lambda \end{bmatrix} = \delta_{\kappa\mu};\ \sum_\kappa \begin{bmatrix} jR \\ \kappa\lambda \end{bmatrix}^* \begin{bmatrix} jR \\ \kappa\nu \end{bmatrix} = \delta_{\lambda\nu}. \tag{6}$$

The irreducible representations can be classified,[2] in general, into three groups: those which can be transformed into a real form, those which cannot but are equivalent to the conjugate complex representation, and those which are not equivalent to the conjugate complex representation. In analogy to the notation customarily used for the two dimensional unimodular unitary group, we shall call the representations of the first kind *integer* representations. Correspondingly $c_j = 1$ will hold for representations j which can be transformed into a real form, $c_j = -1$ will hold for *half integer* representations j which cannot be transformed into a real form but are equivalent to the conjugate complex representation. Finally $c_j = 0$ if the representation j is not equivalent to the conjugate complex of j. According to G. Frobenius and I. Schur [2]

$$\sum_R [j;R^2] = c_j h. \tag{7}$$

2. The number of square roots of an element R will be denoted by $\zeta(R)$

$$\zeta(R) = \sum_S \delta_{R,S^2}. \tag{8}$$

We have

$$\sum_R \zeta(R)^2 = \sum_{R,S} \zeta(R)\delta_{R,S^2} = \sum_S \zeta(S^2) = \sum_{S,T} \delta_{S^2,T^2}$$

[2] G. Frobenius and I. Schur, Berl. Ber. 1906, p. 186.

One can replace S by TR in the last summation and obtain

$$\sum_R \zeta(R)^2 = \sum_{R,T} \delta_{TRTR,T^2} = \sum_{R,T} \delta_{R,TR^{-1}T^{-1}}$$

as $TRTR = T^2$ if and only if $R = TR^{-1}T^{-1}$. For a given R, there will be a T such that $R = TR^{-1}T^{-1}$ only if R and R^{-1} are in the same class, i. e. if R is in an ambivalent class. In this case, the number of T satisfying $R = TR^{-1}T^{-1}$ is equal to h/n_R, since each of the n_R members of the class of R is obtained h/n_R times when T runs over all h elements of the group. Hence

$$\sum_R \zeta(R)^2 = \sum_R{}' h/n_R = h. \text{ (number of ambivalent classes).} \tag{9}$$

The second summation is to be extended only over the elements of the ambivalent classes. The result thus obtained [3] holds for every finite group:

THEOREM 1. *The sum of the squares of the numbers of square roots of all elements of a finite group is equal to the order of the group, multiplied by the number of ambivalent classes.*

All classes are ambivalent in the S. R. groups. Hence

$$\sum_R \zeta(R)^2 = hn \tag{9a}$$

holds for these, where n is the number of all classes.

The number of times the representation i is contained in the Kronecker product of the representations j and k is given by the equation

$$(i, j, k) = \sum_C [i; C]^* [j; C] [k; C] n_c/h \tag{10}$$

where the summation has to be extended, as in (3a) over all classes. Multiplying (7) by $[j; S]$ and summing over j gives, for (5a)

$$\begin{aligned} \sum_j c_j h[j; S] &= \sum_{R,j} [j; R^2][j; S] \\ &= \sum_R \Delta_{R^2,S}\, h/n_S = h\zeta(S). \end{aligned} \tag{11}$$

The $R^2 = S$ equation is satisfied for $\zeta(S)$ group elements R but R^2 is in the class of S for $n_S\zeta(S)$ group elements.

LEMMA 1. *The Kronecker product of two integer representations or of two half integer representations of a S. R. group contains only integer representations; the Kronecker product of an integer and a half integer represen-*

[3] This must have been known to the authors of Reference 2 since it follows immediately from a comparison of the last sentence of § 4 with the sentence in italics on page 201.

tation contains only half integer representations. The unitary matrix which transforms an integer representation into the conjugate complex form is symmetric, that which transforms a half integer representation into the conjugate complex form is skew symmetric.[2] Hence the unitary matrix S which transforms the Kronecker product M of two integer or two half integer representations into the conjugate complex form is symmetric:

$$SM = M^*S; \qquad S = S'. \tag{12}$$

If the unitary U brings M into the reduced form $UMU^{-1} = M_r$

$$S_rM_r = M^*_rS_r; \qquad S_r = U^*SU^{-1} \tag{12a}$$

and S_r is again symmetric. Since the corresponding parts of M_r and M^*_r are equivalent and since M_r does not contain any irreducible representation more than once, S_r is a step matrix just as M_r is and every submatrix of S_r is symmetric on account of the symmetry of S_r. Hence, every submatrix of M_r, i. e. all irreducible parts of M, can be transformed into the conjugate complex form by a symmetric matrix and are integer representations.

If M is the product of an integer and a half integer representation, S will be skew symmetric and the same will hold for S_r and its submatrices. Consequently, all the irreducible parts of M will be half integer representations.

For a S. R. group $c_ic_jc_k = 1$ if (ijk) is different from zero.

Since the (ijk) are positive integers or zero,

$$(ijk)^2 \geq c_ic_jc_k(ijk). \tag{13}$$

The equality sign can hold only if either $(ijk) = 0$, or $(ijk) = 1$ and $c_ic_jc_k = 1$. Hence

$$\sum_{ijk} (ijk)^2 \geq \sum_{ijk} c_ic_jc_k(ijk) \tag{13a}$$

and the equality sign can hold only if for all i, j, k either $(ijk) = 0$ or $(ijk) = 1$ and $c_ic_jc_k = 1$. This is the case, according to the definition of S. R. groups and Lemma 1, for S. R. groups and conversely, if the equality sign holds in (13a), the group must be a S. R. group.

Because of (11), we have

$$\begin{aligned}\sum_{ijk} c_ic_jc_k(i, j, k) &= \sum_{ijk} c_ic_jc_k[i; C]^*[j; C][k; C]n_c/h \\ &= \sum_{ijk} c_ic_jc_k[i; R]^*[j; R][k; R]/h = \sum_R \zeta(R)^3/h\end{aligned}$$

For the left side of (13a) we have, because of (5a)

$$
\begin{aligned}
&\sum_{ijk} (i, j, k)^2 \\
&\quad = \sum_{ijk} \sum_{CC'} [i; C]^*[j; C][k; C][i; C'][j; C']^*[k; C']^* n_c n_{c'}/h^2 \\
(14a) \qquad &\quad = \sum_{ik} \sum_{C} [i; C]^2 [k; C]^2 n_c/h = \sum_{C} h/n_c \\
&\quad = \sum_{R} h/n_R^2 = \sum_{R} v_R^2/h
\end{aligned}
$$

where $v_R = h/n_R$ is the number of elements which commute with R. Hence (13a) is equivalent to

THEOREM 2. *The inequality*

$$\sum_{R} \zeta(R)^3 \leq \sum_{R} v_R^2 \tag{15}$$

holds for every finite group. The equality sign in (15) *holds for all finite S. R. groups and only these.*

3. The Kronecker product of a representation with itself can be decomposed into a symmetric part

$$B_{\kappa\mu;\lambda\nu} = \frac{1}{2}\begin{bmatrix} jR \\ \kappa\lambda \end{bmatrix}\begin{bmatrix} jR \\ \mu\nu \end{bmatrix} + \frac{1}{2}\begin{bmatrix} jR \\ \kappa\nu \end{bmatrix}\begin{bmatrix} jR \\ \mu\lambda \end{bmatrix} \tag{16}$$

and an antisymmetric part

$$A_{\kappa\mu;\lambda\nu} = \frac{1}{2}\begin{bmatrix} jR \\ \kappa\lambda \end{bmatrix}\begin{bmatrix} jR \\ \mu\nu \end{bmatrix} - \frac{1}{2}\begin{bmatrix} jR \\ \kappa\nu \end{bmatrix}\begin{bmatrix} jR \\ \mu\lambda \end{bmatrix} \tag{16a}$$

It is easy to see that both B and A form a representation of the group. The irreducible parts of both B and A are integer representations in case of *S. R.* groups. The irreducible parts of the B for integer j and the irreducible parts of the A for half integer j will be called even representations. Conversely, the irreducible parts of the B for integer j will be called odd representations. This notation is taken again from the theory of representations of the two dimensional unitary group.

THEOREM 3. *In S. R. groups no representation can be both even and odd.*

The trace of the symmetric part of the square of the representation j is

$$X_{js}(R) = \tfrac{1}{2}[j; R]^2 + \tfrac{1}{2}[j; R^2] \tag{17}$$

and the trace of the antisymmetric part is

$$X_{ja}(R) = \tfrac{1}{2}[j; R]^2 - \tfrac{1}{2}[j; R^2] \tag{17a}$$

The condition that two representations have no common part is that the sum of the products of their characters vanish. Theorem 3 is equivalent therefore with the validity of

$$(18) \qquad \tfrac{1}{4}\sum_R ([j;R]^2 + c_j[j;R^2])([k;R]^2 - c_k[k;R^2]) = 0$$

for all j and k. Since the left side of (18) by its nature cannot be negative (it is $\Sigma n_i m_i$ where n_i and m_i are the numbers of times the representation i is contained in the first and second representation of (18)), the validity of (18) for all j and k is equivalent with the vanishing of

$$\begin{aligned}(18a) \qquad \sum_R \sum_{jk} ([j;R]^2 + c_j[j;R^2])([k;R]^2 - c_k[k:R^2]) \\ = \sum_R (h/n_R + \zeta(R^2))(h/n_R - \zeta(R^2)) \\ = \sum_R v_R^2 - \sum_R \zeta(R^2)^2\end{aligned}$$

where (5a) and (11) have been utilized. Now evidently

$$\sum_S \zeta(S)^3 = \sum_{S,R} \zeta(S)^2 \delta_{S,R^2} = \sum_R \zeta(R^2)^2$$

so that (18a) vanishes on account of Theorem 2. Hence Theorem 3 is valid.

Of course, the Kronecker product of an even and an odd representation e. g., contains, in general, both even and odd representations. It has not been shown, either, that every integer representation is either even, or odd. In fact, one can easily find a group which has an integer representation which does not occur in the square of any representation. A group of this character is formed by the elements $1, -1, x, -x, y, -y, z, -z$, with the multiplication rules $x^2 = y^2 = 1$, $z^2 = -1$, $xy = -yx = z$, $xz = -zx = y$, $zy = -yz = x$.

PRINCETON UNIVERSITY.

Localized States for Elementary Systems

T. D. Newton and E. P. Wigner

Reviews of Modern Physics *21*, 400–406 (1949)

It is attempted to formulate the properties of localized states on the basis of natural invariance requirements. Chief of these is that a state, localized at a certain point, becomes, after a translation, orthogonal to all the undisplaced states localized at that point. It is found that the required properties uniquely define the set of localized states for elementary systems of non-zero mass and arbitrary spin. The localized functions belong to a continuous spectrum of an operator which it is natural to call the position operator. This operator has automatically the property of preserving the positive energy character of the wave function to which it is applied (and it should be applied only to such wave functions). It is believed that the development here presented may have applications in the theory of elementary particles and of the collision matrix.

INTRODUCTION

It is well known that invariance arguments suffice to obtain an enumeration of the relativistic equations for elementary systems.[1] The concept of an "elementary system" is, however, not quite identical with the intuitive concept of an elementary particle. Intuitively, we consider a particle "elementary" if it does not appear to be useful to attribute a structure to it. The definition under which the aforementioned enumeration can be made is a more explicit one: it requires that all states of the system be obtainable from the relativistic transforms of *any* state by superpositions. In other words, there must be no relativistically invariant distinction between the various states of the system which would allow for the principle of superposition. This condition is often referred to as irreducibility condition. Relativistic transform is meant to include in the above connection not only the customary Lorentz transformations but also rotations and displacements in space and time.

The role of elementary systems as initial and final states of collision phenomena, and hence their connection with the theory of the collision matrix, will be discussed at the end of this article. We wish to turn now to the connection of elementary systems with elementary particles.

Two conditions seem to play the most important role in the concept of an elementary particle. The first one is that its states shall form an elementary system in the sense given above. This condition is quite unambiguous. The second condition is less clear cut: it is that it should not be useful to consider the particle as a union of other particles. In the case of an electron or a proton both conditions are fulfilled and there is no question as to the elementary nature of these particles. Only the first condition is fulfilled for a hydrogen atom in its normal state and we do not consider it to be an elementary particle.

The situation is more ambiguous, for instance, in the case of a π-meson. Qualitatively, a π-meson differs in no way from a very sharp resonance state, formed by the collision of a μ-meson and a neutrino. Strictly speaking, the states of a π-meson do not form an elementary system because, after a sufficiently long time, it can be either in the dissociated or in the undissociated state and the distinction between these is surely invariant relativistically. Nevertheless, the life time of the π-meson is very long as compared with any relevant unit of time (such as h/mc^2) and within this time interval its states do form an elementary system. On the other hand, the properties of a π-meson are very different from what one would expect from a compound consisting of a μ-meson and a neutrino. Thus the second condition for an elementary particle is fulfilled. It is this condition which has no counterpart in the definition of an elementary system. As a result of this circumstance the concept of an elementary system is much broader than that of an elementary particle; as was mentioned above, a hydrogen atom in its normal state forms an elementary system.

Every system, even one consisting of an arbitrary number of particles, can be decomposed into elementary systems. These elementary systems can be specified in a relativistically invariant manner, as containing only certain states. Thus, the restriction to the normal state of the hydrogen atom selected an elementary system from all the states of the hydrogen atom, which, together, do not form an elementary system. The usefulness of the decomposition into elementary systems depends on how often one has to deal with linear combinations containing several elementary systems.

The great drawback of using the elementary systems as a basis of the theory is that their existence follows by a rather abstract argument from the principles of quantum mechanics. As a result, the expressions for some of the most important operators "get lost" in the process. The only physical quantities for which the theory directly provides expressions are the basic quantities of the components of the momentum-energy vector and the six components of the relativistic angular momentum tensor. The subject of the present article is an attempt to find general, invariant theoretic principles

[1] E. P. Wigner, Ann. of Math. 40, 149 (1939). The concept of an elementary system, which will be explained below, is a description of a set of states which forms, in mathematical language, an irreducible representation space for the inhomogeneous Lorentz group.

on the basis of which operators for the position coordinates can be found.

If we restrict ourselves to an elementary system, the physical interpretation of the operators to be found is unique: they will correspond to the position of the particle if we deal with an elementary particle. Otherwise they may correspond to the center of mass of the system. If the system is not elementary, the interpretation will not be unique and neither will our postulates lead us to a uniquely determined set of operators.

Before proceeding with our argument, we wish to refer to other investigations with somewhat similar objectives. The problem of the center of mass in relativity theory has been treated particularly by Eddington[2] and by Fokker[3] on the basis of non-quantum mechanics. Their work was evaluated and a quantum mechanical generalization thereto given by Pryce.[4] We shall have frequent occasion to refer to his results. Ideas related to Pryce's work have been first put forward by Schrödinger[5] and, more recently, by Finkelstein[6] and also by Møller.[7]

The present paper arose from a reinvestigation of the irreducible representations[8] of de Sitter space which was undertaken by one of us.[9] These representations are in a one to one correspondence with relativistically invariant wave equations for elementary systems in de Sitter space. At the conclusion of the investigation it appeared that the physical content of the equations which were obtained could be understood much more readily if position operators could be defined on an invariant theoretic basis. As an introduction to this, a similar investigation was undertaken in flat space with the results given in the following sections.

POSTULATES FOR LOCALIZED STATES AND POSITION OPERATORS

The position operator could easily be written down if the wave function of the state (or the states) were known for which the three space coordinates are zero at $t=0$. If ψ is such a function and $T(a)$ the operator of displacement by a_x, a_y, a_z, a_t, the wave function $T(a)^{-1}\psi$ represents a state for which the space coordinates are a_x, a_y, a_z at time a_t. Thus the knowledge of the wave functions corresponding to the state $x=y=z=0$ at $t=0$ (and the knowledge of the displacement operators) entails the knowledge of all localized states, i.e., of all characteristic functions of the position operators. From these, the position operators are easily obtained. For this reason we concentrated on obtaining the wave functions of those states which are, at time $t=0$, localized at the origin of the coordinate system.

We postulate that the states which represent a system localized at time $t=0$ at $x=y=z=0$: (a) form a linear set S_0, i.e., that the superposition of two such localized states be again localized in the same manner; (b) that the set S_0 be invariant under rotations about the origin and reflections both of the spatial and of the time coordinate; (c) that if a state ψ is localized as above, a spatial displacement of ψ shall make it orthogonal to all states of S_0; (d) certain regularity conditions, amounting essentially to the requirement that all the infinitesimal operators of the Lorentz group be applicable to the localized states, will be introduced later.

It is to be expected that the states localized at a certain point have the same properties as characteristic functions of a continuous spectrum, i.e., they will not be square integrable but the limits of square integrable functions. It seems to us that the above postulates are a reasonable expression for the localization of the system to the extent that one would naturally call a system unlocalizable if it should prove to be impossible to satisfy these requirements.

We shall carry out our calculations in the realization of the elementary systems which was described by Bargmann and Wigner[10] and will proceed with the calculation.

Particle with no spin (Klein-Gordon particle)

The determination of the localized state is particularly simple in this case. It will be carried out in some detail in spite of this, because the same steps occur in the consideration of systems with spin.

The wave functions are defined, in this case, on the positive shell of a hyperboloid $p_0^2=p_1^2+p_2^2+p_3^2+\mu^2$ and we shall use p^1, p^2, p^3 as independent variables. In any formula, p_0 is an abbreviation for $(p_1^2+p_2^2+p_3^2+\mu^2)^{\frac{1}{2}}$. The invariant scalar product is

$$(\psi,\phi)=\int\int\int\psi(p_1p_2p_3)^*\phi(p_1p_2p_3)dp_1dp_2dp_3/p_0. \quad (1)$$

The wave function Φ in coordinate space becomes

$$\Phi(x^1,x^2,x^3,x^0)=(2\pi)^{-\frac{3}{2}}\int\phi(p_1,p_2,p_3)\times\exp(-i\{x,p\})dp_1dp_2dp_3/p_0, \quad (2)$$

where

$$\{x,p\}=x^0p^0-x^1p^1-x^2p^2-x^3p^3=x_0p_0-x_1p_1-x_2p_2-x_3p_3, \quad (3)$$

is the Lorentz invariant scalar product. Throughout this paper, the covariant and contravariant components are equal for the time (0) coordinate, oppositely equal for the space (1, 2, 3) coordinates. This governs the

[2] A. S. Eddington, *Fundamental Theory* (Cambridge University Press, London, 1946).
[3] A. D. Fokker, *Relativitatstheorie* (Groningen, Noordhoff, 1929).
[4] M. H. L. Pryce, Proc. Roy. Soc. 195A, 62 (1948).
[5] E. Schrödinger, Berl. Ber. 418 (1930); 63 (1931).
[6] R. J. Finkelstein, Phys. Rev. 74, 1563A (1948).
[7] Chr. Møller, Comm. Dublin Inst. for Adv. Studies A, No. 5 (1949); also A. Papapetrou, Acad. Athens 14, 540 (1939).
[8] L. H. Thomas, Ann. of Math. 42, 113 (1941).
[9] T. D. Newton, Princeton Dissertation (1949).
[10] V. Bargmann and E. P. Wigner, Proc. Nat. Acad. Sci. 34, 211 (1948).

raising and lowering of all indices. Occasionally, we shall use for the scalar product of two space-like vectors the notation (x, p) so that, e.g., $\{x, p\} = x^0p^0 - (x\cdot p)$.

The linear manifolds which are invariant with respect to rotations about $p_1 = p_2 = p_3 = 0$ are, for any integer j, the $2j+1$ functions

$$P_m{}^j(\vartheta, \phi)f(p) \quad (m = -j, -j+1, \cdots j-1, j), \quad (4)$$

where p, ϑ and ϕ are polar coordinates for p_1, p_2, p_3 and f is an arbitrary function. The $P_m{}^j$ are the well-known spherical harmonics. The sets (4) are also invariant with respect to inversion, i.e., replacement of p_1, p_2, p_3 by $-p_1$, $-p_2$, $-p_3$. Naturally, not only a single set (4) has these properties of invariance but the sum of an arbitrary number of such sets as long as one includes with one function (4) all $2j+1$ functions and their linear combinations. The $f(p)$ could be different for different j.

Under time reversal $\psi(p_1p_2p_3)$ goes over into[11]

$$\Theta\psi(p_1, p_2, p_3) = \psi(-p_1, -p_2, -p_3)^*. \quad (5)$$

We understand by time reversal the operation which makes out of a wave function ψ the wave function $\theta\psi$ on which every experiment, if carried out at $-t$, yields same results as the same experiment carried out on ψ at time t. Because of (5) and our postulate (*b*), if the $P_m{}^j(\theta, \phi)f(p)$ are localized at the origin, the set $P_{-m}{}^j(\theta, \phi)f(p)^*$, i.e., the $P_{+m}{}^j(\theta, \phi)f(p)^*$ are also localized. The same is then true for the sum and difference of the corresponding pairs of functions which shows that the $f(p)$ can be assumed to be real without loss of generality.

The displacement operator in momentum space is simply multiplication with $\exp(-i\{a, p\})$;

$$T(a)\psi = \exp(-i\{a, p\})\psi. \quad (6)$$

We shall have to consider purely space-like displacements, i.e., assume that $a^0 = 0$. It then follows from our postulate (*c*) that, in particular, $\exp(i(a, p))\psi$ is orthogonal to ψ if ψ is localized, or that

$$\int\int\int |\psi(p_1p_2p_3)|^2 \exp i(a_1p_1 + a_2p_2 + a_3p_3)dp_1dp_2dp_3/p_0 = 0, \quad (6a)$$

for any non-vanishing vector a. This shows that only the zero wave number part occurs in the expansion of $|\psi|^2/p_0$ into a Fourier integral. Hence $|\psi|^2/p_0$ is a constant, $|\psi|$ proportional to $p_0{}^{\frac{1}{2}}$. Comparing this with (4), we see that only the set $j=0$ can be chosen. Since, furthermore, we saw that $f(p)$ can be assumed to be real, we have

$$\psi^2 = (2\pi)^{-\frac{3}{2}}p_0. \quad (7)$$

As was anticipated, (ψ, ψ) is infinite, the localized function is part of a continuous spectrum.

[11] E. P. Wigner, Gottinger Nachrichten 546 (1932).

As far as postulates (*a*), (*b*), (*c*) are concerned, ψ could be a discontinuous function, being $+p_0{}^{\frac{1}{2}} = (p^2+\mu^2)^{\frac{1}{4}}$ for some p, and $-(p^2+\mu^2)^{\frac{1}{4}}$ for the remaining p. However, no matter how ψ is chosen, consistently with (7), there is, in this case, only one state localized at the origin because if there were two, say ψ_1, and ψ_2, the ψ_1 would have to be orthogonal not only to $\psi_1 \exp(-i(a, p))$ but also to $\psi_2 \exp(-i(a, p))$ from which not only $|\psi|^2 \sim p_0$ but also $\psi_1{}^*\psi_2 \sim p_0$ and hence the proportionality of ψ_1 to ψ_2 follows.

In order to eliminate the discontinuous ψ as localized state, we introduce the further regularity condition that

$$(M^{0k}\psi_n, M^{0k}\psi_n)/(\psi_n, \psi_n) \quad (8)$$

shall remain finite as the normalizable wave functions ψ_n approach ψ. The M_{0k} is the infinitesimal operator of a proper Lorentz transformation in the x^0x^k plane, its operator is[10]

$$M^{0k} = ip^0\partial/\partial p_k. \quad (8a)$$

This further postulate eliminates all discontinuous ψ and we obtain for the wave function of the only state which is localized at the origin

$$\psi = (2\pi)^{-\frac{1}{2}}p_0{}^{\frac{1}{2}}. \quad (9)$$

The regularity requirement (*d*) actually asks for the finiteness of (8) for all M^{kl}. However, if one substitutes M^{23}, M^{31} or M^{12} for M^{0k} in (8), the resulting expression is automatically bounded—in fact their sum is $j(j+1)$. Hence requiring the applicability of the M^{23}, M^{31}, M^{12} to ψ does not introduce a new condition.

The localized wave function in coordinate space is obtained by (2). It is, apart from a constant[12]

$$\Psi(r) = (\mu/r)^{5/4}H_{5/4}{}^{(1)}(i\mu r). \quad (9a)$$

It goes to zero at $r=\infty$ as $e^{-\mu r}$, at $r=0$ it becomes infinite as $r^{-5/2}$. It is, of course, not square integrable since it is part of a continuous spectrum.

Applying the operator of displacement to (9) we obtain for the wave function of the state which is localized at x^1, x^2, x^3 at time $t=0$

$$T(-x)\psi = (2\pi)^{\frac{1}{2}}p_0{}^{\frac{1}{2}} \exp - i(p^1x^1 + p^2x^2 + p^3x^3) = (2\pi)^{-\frac{1}{2}}p_0{}^{\frac{1}{2}}e^{-i(p\cdot x)}.$$

This must be a characteristic function of the operator q^k for the k-coordinate with the characteristic value x^k. The operator q^k is therefore defined by

$$q^k\phi(p) = (2\pi)^{-3}\int p_0{}^{\frac{1}{2}}e^{-i(p\cdot x)}x^k(p_0')^{\frac{1}{2}} \times e^{i(p'\cdot x)}\phi(p')dxdp'/p_0'; \quad (10)$$

dx and dp' stand for $dx^1dx^2dx^3$ and $dp_1'dp_2'dp_3'$. One

[12] G. A. Campbell and R. M. Foster, "Fourier Integrals for Practical Applications," American Telephone and Telegraph Company (1931).

can transform (10) in a well-known fashion

$$q^k\phi(p)=\left(i\frac{\partial}{\partial p^k}-\frac{i}{2}\frac{p^k}{p_0^2}\right)(2\pi)^{-3}\times\int p_0^{\frac{1}{2}}e^{i(p'-p)\cdot x}(p_0')^{-\frac{1}{2}}\phi(p')dxdp'$$

$$=-i\left(\frac{\partial}{\partial p_k}+\frac{p^k}{2p_0^2}\right)\phi(p). \tag{11}$$

These expressions are valid for finite as well as for vanishing rest mass. It is remarkable that the operator q^k can be transformed into coordinate space and retains a relatively simple form

$$q^k\Phi(x)=x^k\Phi(x)+\frac{1}{8\pi}\int\frac{\exp(-\mu|x-y|)}{|x-y|}\frac{\partial\Phi(y)}{\partial y_k}dy. \tag{12}$$

x and y stand for the spatial part of the four vectors x^μ and y^μ and dy indicates integration over y^1, y^2, y^3. The customary q^k operator contains only the first term of (11).

It may be well to remember at this point that the position operators to which our postulates lead necessarily commute with each other so that only Pryce's case (*e*) can be used for comparison. In fact, our q^k is identical with his $\bar{q}^k$. It may be pointed out, second, that a state which is localized at the origin in one coordinate system, is not localized in a moving coordinate system, even if the origins coincide at $t=0$. Hence our operators q^k have no simple covariant meaning under relativistic transformations. This is not the case for the customary operators q^k either. Furthermore, even though it appears that $\Phi(x)=\delta(x)$ is invariant under relativistic transformations which leave the origin unchanged, this is not much more than a mathematical quirk. One sees this best by transforming the δ-function to momentum space through the inversion of (2). The result, p_0, seems to have a simple covariant meaning. However, it does not represent a square integrable function and if one approximates it by one, say by $\psi_\alpha=p_0\exp(-\alpha^2p_0^2)$, the Lorentz transform of ψ_α will not approach ψ_α with decreasing α. In fact, as soon as $\alpha\mu\ll 1$, the scalar product of ψ_α and its transform will be independent of α and smaller than the norm of ψ_α.

Particles with spin and finite mass

We again use the description given in reference 10, i.e., define wave functions on the positive hyperboloid $p_0^2=p_1^2+p_2^2+p_3^2+\mu^2$ and use in addition to p_1, p_2, p_3 the $2s$ spin variables $\xi_1, \xi_2, \cdots\xi_{2s}$ all of them four-valued. The wave functions which describe the possible states of the system will be symmetric functions of the ξ and satisfy the $2s$ equations

$$\sum_\kappa \gamma_\alpha^\kappa p_\kappa\psi=\mu\psi \quad \alpha=1, 2, \cdots 2s. \tag{13}$$

The consistency of these was shown before.[10] The γ_α apply to ξ_α, two γ with different first indices commute, with the same first index they satisfy the well-known relations

$$\gamma_\alpha^\kappa\gamma_\alpha^\lambda+\gamma_\alpha^\lambda\gamma_\alpha^\kappa=2g^{\kappa\lambda}. \tag{13a}$$

The great difference between the present case and that of zero spin consists in the limitation (13) of the permissible wave functions, in addition to the limitation to the positive hyperboloid. This latter limitation can be taken care of by using only the p_1, p_2, p_3 as independent variables, the former limitation cannot be taken care of in an equally simple fashion. We shall make extensive use, however, of a device, most successfully employed by Schrödinger[5] and define operators

$$E_\alpha=\tfrac{1}{2}(p_0)^{-1}(\textstyle\sum_\kappa \gamma_\alpha^\kappa p_\kappa+\mu)\gamma_\alpha^0. \tag{14}$$

This is a projection operator: $E_\alpha^2=E_\alpha$ and $E_\alpha\psi$ automatically satisfies the corresponding Eq. (13). Denoting the product of all the E_α by E

$$E=E_1E_2\cdots E_{2s} \tag{14a}$$

any $E\psi$ is a permissible wave function, satisfying all Eqs. (13).

For the scalar product, we shall use the expression

$$(\psi,\phi)=\int p_0^{-2s-1}\textstyle\sum_\xi \psi^*\phi dp. \tag{15}$$

It follows from this at once, because of our postulate (*c*), and since (6) is valid in this case also, that every wave function which is localized at the origin satisfies the analogon of (7):

$$\textstyle\sum_\xi|\psi|^2=(2\pi)^{-3}p_0^{2s+1}. \tag{16}$$

The operator for time reversal is

$$\Theta\psi(p_1, p_2, p_3)=C\psi(-p_1, -p_2, -p_3)^*, \tag{17}$$

where C is a matrix which operates on the ξ coordinates and satisfies the equations

$$\begin{aligned} C\gamma_\alpha^{0*}&=\gamma_\alpha^0C; \quad (\alpha=1, 2, \cdots, 2s)\\ C\gamma_\alpha^{k*}&=-\gamma_\alpha^kC. \quad (\alpha=1, 2, \cdots, 2s; k=1, 2, 3). \end{aligned} \tag{17a}$$

If γ^0, γ^2, γ^3 are real, γ^1 imaginary

$$C=\prod_{\alpha=1}^{2s}\gamma_\alpha^2\gamma_\alpha^3; \quad C^2=(-)^{2s}. \tag{17b}$$

Since C, as defined above, is a real matrix we also have $\Theta^2=(-)^{2s}$ which is true independently of the choice of the γ-matrices. The operator for the inversion of the space coordinates is

$$I\psi(p_1, p_2, p_3)=\gamma_1^0\gamma_2^0\cdots\gamma_{2s}^0\psi(-p_1, -p_2, -p_3); \tag{18}$$

it commutes with the E_α of (14).

In order to determine the sets of wave functions which are invariant under rotations, we first define the analogue of the pure spin function for the relativistic

Eqs. (13). For this purpose we define auxiliary functions v_m which are independent of p_1, p_2, p_3 and functions of the ξ only. They satisfy the equations

$$\gamma_\alpha{}^0 v_m = v_m \quad (\alpha=1, 2, \cdots, 2s) \tag{19}$$

and

$$\tfrac{1}{2}i\sum_\alpha \gamma_\alpha{}^1\gamma_\alpha{}^2 v_m = m v_m \quad (m=-s, -s+1, \cdots, s-1, s). \tag{19a}$$

Since the γ^0 and the $i\gamma^1\gamma^2$ commute, it is possible to assume temporarily that they are all diagonal. Equation (19) then demands that they all belong to the characteristic value $+1$ of each γ^0; there are 2^{2s} such functions. However, we are interested only in symmetric functions of the ξ and there are only $2s+1$ of these. They are distinguished by the index m: the v_m has non-zero components only for those ξ for which $s+m$ of the $i\gamma^1\gamma^2$ are $+1$, the remaining $s-m$ are -1. For these ξ the value of v_m is $((s+m)!(s-m)!/(2s)!)^{\frac{1}{2}}2^{-s}$ so that v_m is normalized in the sense

$$\sum_\xi |v_m|^2 = \sum_\xi v_m{}^2 = 1. \tag{19b}$$

Physically, m corresponds to the spin angular momentum about the x^3 axis, the parity of v_m is even because of (19) and (18).

The v_m are not permissible wave functions because they do not satisfy the wave Eqs. (13). We therefore define as spin functions

$$V_m(p_1, p_2, p_3, \xi_1, \cdots, \xi_{2s}) = E v_m \quad (m=-s, \cdots, s). \tag{20}$$

They are permissible wave functions of even parity and V_m represents a state of angular momentum $m\hbar$ about the third coordinate axis. Their normalization is, instead of (19b)

$$\sum_\xi |V_m|^2 = ((p_0+\mu)/2p_0)^{2s}. \tag{20a}$$

The most general solution of (13) is a linear combination of the V_m multiplied with arbitrary functions of the p_1, p_2, p_3. A set of wave functions which is invariant under rotations and reflections contains wave functions of the form

$$\psi_{jm} = \sum_{l,m'} S(l, s)_{j,m-m',m'} P^l{}_{m-m'}(\vartheta, \phi) f_l(p) V_{m'}. \tag{21}$$

The p, ϑ, ϕ are again polar coordinates for p^1, p^2, p^3; the f_l are arbitrary unknown functions of the length of p. However, if one function of the form (21) occurs in the set, all others with different m but the same f_l also occur. The summation over l is to be extended over all even values between $|j-s|$ and $j+s$ if the parity of the ψ_j is to be even, over all odd values of l if the parity of ψ_j is odd. The $S(l, s)$ are the customary coefficients[13] giving a total j from wave functions with given "orbital momentum" l and "spin momentum" s.

Since the polar angles ϑ, ϕ are indeterminate for $p=0$, the $f_l(p)$ must vanish for $p=0$ unless $l=0$. Otherwise, the ψ_{jm} would become singular at $p=0$ and the M^{0k} could not be applied to them in the sense of the boundedness of (8). (Actually it is necessary to postulate this equation for the square of M^{0k} instead only for M^{0k}.) It follows that ψ_{jm} vanishes at $p=0$ unless the series (21) contains a term with $l=0$. However, (16) shows that ψ cannot vanish at this point if the rest mass is finite and that, hence, every localizable wave function must have an $l=0$ term in its expansion. This happens only if $j=s$ and the wave function has even parity. If the parity of ψ_{jm} is odd, only f_1, f_3, etc., enter (21) and these still vanish at $p=0$. It follows that the wave functions which are localized at the origin all have angular momentum $j=s$ and the form

$$\psi_m = \sum_{l=0}^{2s} \sum_{m'} S(l, s)_{s,m-m',m'} P^l{}_{m-m'}(\vartheta, \phi) f_l V_{m'}. \tag{21a}$$

We now skip the part of the calculation which deals with the determination of the f_l and give only the result: The wave functions localized at the origin are the $2s+1$ functions

$$\psi_m = (2\pi)^{-3/2} 2^s p_0{}^{2s+\frac{1}{2}} (p_0+\mu)^{-s} \times V_m(p_1 p_2 p_3; \xi_1, \cdots \xi_{2s}) \tag{21b}$$

(i.e., the $l=0$ term alone remains from (21a)). Actually, this result is far from being surprising.[14]

The operator for the position coordinate can be calculated in exactly the same way as this was done in the case of zero spin, and gives

$$q^k = E \prod_{\alpha=1}^{2s} (1+\gamma_\alpha{}^0) \frac{p_0{}^{2s+\frac{1}{2}}}{(p_0+\mu)^s} \left(-i\frac{\partial}{\partial p_k}\right) \frac{p_0{}^{-\frac{1}{2}}}{(p_0+\mu)^s} E. \tag{22}$$

For $s=\frac{1}{2}$ this again agrees with Pryce's result[4] for his case (e), i.e., for his operator $\bar{q}$.

The significance of the projection operators E in (22) is only to annihilate any negative energy part of the wave function to which it is applied and to produce a purely positive energy wave function. Since q^k is the position operator only for wave functions which are

[13] See e.g., E. P. Wigner, *Gruppentheorie, etc.* (V. Vieweg & Sohn, Braunschweig, 1931). The composition of the V and the spherical harmonics P^l to the ψ_j is the same operation as the composition of the spin functions with a definite S and the space coordinate functions with a definite L, to functions of both, with definite J. This composition is explained in Chapter XXII. The coefficients of the composition, i.e., our $S(l, s)$ are calculated p. 202 ff (they are denoted by $s^{(LS)}$).

[14] The proof runs as follows. One first shows, by considering $\psi_m \pm \Theta\psi_{-m}$ that the f_l of (21a) can all be assumed to be real. One then subdivides ψ_m into two parts: the $l=0$ part of the sum (21a) will be denoted by ψ^0, the rest ψ^r. As we have seen, ψ^0 is finite at $p=0$, while ψ^r vanishes at that point. The proof then consists in showing that there can be no region in which ψ^r is finite but very much smaller than ψ^0. It then follows from the continuity of both ψ^0 and ψ^r that the latter vanishes everywhere. Inserting $\psi^0+\psi^r$ for ψ in (16), one can neglect in the aforementioned region the square of ψ^r as compared with the other terms. The right side, as well as the term from the square of ψ^0, are independent of ϑ and φ. This must be true, therefore, also for the term arising from the cross product of ψ^0 and ψ^r. This term is, however, a sum of expressions $P_m{}^l(\vartheta,\varphi) f_0 f_l$ which cannot be independent of ϑ, φ except if all f_l with $l>0$ vanish (f_0 is finite by assumption). It then follows that the f_l vanish everywhere and (21a) reduces to a single term. This can be obtained from (16) by taking the square root on both sides.

defined on the positive hyperboloid alone, the E on the right could logically be omitted. Both E can be omitted[5] if one calculates a matrix element between two purely positive energy wave functions. The factors involving p_0 are necessary in order to make $i\partial/\partial p_j$ hermitian: because of the factor p_0^{-2s-1} in the volume element (15), an operator is hermitian if it looks hermitian after multiplication with $p_0^{s+\frac{1}{2}}$ on the right and division with the same factor on the left. The operator $\prod_{\alpha=1}^{2s} \frac{1}{2}(1+\gamma_\alpha^0)$ is a projection operator, i.e., it is identical with its square and could therefore be inserted into (22) once more before the second E, thus making (22) somewhat more symmetric. The position operator (11) for the Klein-Gordon particle is a special case of (22) and can be obtained from (22) by setting $s=0$.

If one displaces a state by a and measures its x^k coordinate afterwards, the result will be greater by a^k than the x^k coordinate measured on the undisplaced state. This leads to the relation

$$T(-a)q^kT(a)=q^k+a^k \tag{23}$$

for $a^0=0$. Inserting the expression (6) for $T(a)$ and going to the limit of very small a^k, one obtains

$$(q^kp^l-p^lq^k)\phi=-i\delta_{kl}\phi \tag{23a}$$

where ϕ is any permissible wave function. Actually one obtains by direct calculation, using in particular the identity

$$E_\alpha(1+\gamma_\alpha^0)E_\alpha=p_0^{-1}(p_0+\mu)E_\alpha \tag{24}$$

the commutation relation

$$q^kp^l-p^lq^k=-i\delta_{kl}E. \tag{25}$$

The commutation relations of the q^k with p_0 are naturally also the usual ones as p_0 is a function of the p^k alone. Since the q^k are the components of a vector operator in three-dimensional space, their commutation relations with the spatial components of M^{kl} are also the usual ones.

We wish to remark, finally, that a consideration, similar to the above, has been carried out also for the equations with zero mass. In the case of spin 0 and $\frac{1}{2}$, we were led back to the expressions for localized systems which were given in (9) and (21b). However, for higher but finite s, beginning with $s=1$ (i.e., Maxwell's equations), we found that no localized states in the above sense exist. This is an unsatisfactory, if not unexpected, feature of our work. The situation is not entirely satisfactory for infinite spin either.

DISCUSSION

One might wonder, first, what the reason is that our localized states are not the δ-functions in coordinate space which are usually considered to represent localized states. The reason is, naturally, that all our wave functions represent pure positive energy states. This is not true of the δ-function. Similarly, our operator (22) transforms positive energy functions into positive energy functions.

It is often stated that a measurement of the position of a particle, such as an electron, if carried out with a greater precision than the Compton wave-length, would lead to pair production and that it is, therefore, natural that the position operators do not preserve the positive energy nature of a wave function. Since a position measurement on a particle should result in a particle at a definite position, and not in a particle and some pairs, this consideration really denies the possibility of the measurement of the position of the particle. If this is accepted it still remains strange that pair creation renders the position measurement impossible to the same degree in such widely different systems as an electron, a neutron and even a neutriono. The calculations given above prove, at any rate, that there is nothing absurd in assuming the measurability of the position, and the existence of localized states, of elementary systems of non-zero mass. Moreover, the postulates (*a*), (*b*), (*c*) and (*d*), which are based on considerations of invariance, define the localized states and position operators uniquely for all non-zero mass elementary systems.

No similarly unique definition of localized states is possible for composite systems. Although it remains easy to show that definite total angular momenta j can be attributed to localized states, one soon runs into difficulties with the rest of the argument. In particular, the summation in (16) must be extended not only over the spin coordinates ξ but also over all states with different total rest mass and different intrinsic spin. As a result, one can, e.g., find states which can coexist as localized states in the sense of our axioms even though their j values are different. This is also what one would expect on ordinary reasoning since, if the system contains several particles, the states in which any one of them is localized at the origin satisfy our postulates. This holds also for the states in which another one of the particles is so localized or for states in which an arbitrary linear combination of the coordinates is zero. As a result, not only is the number of localized states greatly augmented but, further, one must expect to find many such large sets for which our postulates hold, although no two sets can be considered to be localized simultaneously. In other words, each set of localized states is not only much larger for composite systems but one also has to make a choice between many sets all of which satisfy our postulates by themselves. It does not appear that one can proceed much further in the definition of localized states for composite systems without making much more specific assumptions. Naturally, one can define as localized states those, which, in any of the elementary parts of the composite system, appear localizable. It appears reasonable to assume that this definition corresponds to the center of mass of the whole system.

One may wonder, even in the case of elementary particles, whether the determination of the localized

states and position operators has much significance. Such doubts might arise particularly strongly if one is inclined to consider the collision matrix as the future form of the theory. One must not forget, however, that the customary exposition of this theory refers only to questions about cross sections. There is another interesting set of questions referring to the position of the scattered particles: how much further back (i.e., closer to the scattering center) are they than if they had gone straight to the scattering center and then continued in the new direction without any delay.[15] In order to answer such questions in the relativistic region, one will need some definition of localized states for elementary systems. From this point of view it is satisfactory that the localized states could be defined without ambiguity just for these systems.

[15] L. Eisenbud, Princeton Dissertation (1948).

On a Class of Analytic Functions from the Quantum Theory of Collisions

E. P. Wigner

Annals of Mathematics *53*, 36–67 (1951)

(Received April 27, 1950)

I. Origin of the Problem

The present investigation has its origin in the quantum theory of collisions. It appears appropriate, therefore, to give here a very short review of the subject even though some of the material to be covered in this section may be found in standard publications, such as the Theory of Atomic Collisions by N. F. Mott and H. S. W. Massey (Oxford University Press 1933).

Although the underlying problem can be formulated much more generally, it will be sufficient for the purpose of the present paper to consider a single particle in the field of some scattering center. The wave function of the particle will be denoted by φ and it will be sufficient to restrict ourselves to spherically symmetric wave functions, i.e., to the case that φ depends only on the distance r of the particle from the scattering center. Hence, the Schrodinger equation for φ can be written in the form

$$L\varphi_E \equiv -\frac{d^2\varphi_E}{dr^2} + V(r)\varphi_E = E\varphi_E\,. \tag{1}$$

E is the energy of the particle, $V(r)$ the potential due to the scattering center; φ must vanish for $r = 0$.

It will be assumed that the variability domain of r can be divided into two parts. The region $r > a$, where a is a fixed distance, will be called the region of the long range interaction. $V(r)$ is supposed to be known in this region, and to go, for $r = \infty$, sufficiently fast to zero so that the two independent solutions $F(r)$ and $G(r)$ of (1) have, at very large r, the asymptotic forms $k^{-\frac{1}{2}} \sin kr$ and $k^{-\frac{1}{2}} \cos kr$ with $k^2 = E$. The purpose of the factors $k^{-\frac{1}{2}}$ is to make the F and G obey the equation

$$F'G - G'F = 1. \tag{2}$$

The region $r < a$ will be called, on the other hand, the region of short range interaction. Very little is known about $V(r)$ in this region and even less about the permissible solution of (1), i.e., the solution which vanishes at $r = 0$. On the other hand, one is usually not much interested in the behavior of the wave function at $r < a$. It is reasonable, therefore, to express the effect of the short range interaction by giving the ratio of φ_E to its derivative $\varphi_E' = d\varphi_E/dr$ for $r = a$:

$$R(E) = \varphi_E(a)/\varphi_E'(a). \tag{3}$$

If $R(E)$, the "derivative function", as well as the solutions $F(r)$ and $G(r)$ for $r > a$ are known, the wave function is, apart from a constant factor, uniquely

determined for $r > a$ by the requirement of the continuity of φ_E and that of φ_E' at $r = a$:

$$\varphi_E(r) = F(r) + SG(r) \qquad (r \geqq a) \tag{4}$$

wherein

$$S(E) = \frac{F'R - F}{-G'R + G}. \tag{4a}$$

The F, G, F', G' are the values and radial derivatives of the functions $F(r)$, $G(r)$ at $r = a$.

By multiplication with a constant (i.e., r independent) factor $\varphi_E(r)$ goes over into

$$\varphi_E(r) = F(r) + \frac{S}{i - S}(F(r) + iG(r)). \tag{5}$$

The interpretation of this equation is as follows: The asymptotic behavior of the first term is the same ($k^{-\frac{1}{2}} \sin kr$) as would obtain if $V(r)$ vanished throughout. Hence the second term represents the addition to the wave function which is caused by the presence of the potential $V(r)$. It has, at large r, the behavior of an outgoing wave $k^{-\frac{1}{2}}e^{ikr}$ and is what an observer at large r would interpret as the scattered wave. Hence the absolute square of its coefficient is proportional to the scattering cross section σ. A somewhat more detailed consideration shows that

$$\sigma = \frac{4\pi}{E} \frac{S^2}{1 + S^2}. \tag{6}$$

This equation, together with (4a), gives the scattering cross section in terms of the "short range interaction", characterized by R, and the "long range interaction" which finds expression in the quantities F, G, F', G'. The former depends only on the properties of $V(r)$ for $r < a$, the latter only on the properties of $V(r)$ for $r > a$.

It would first appear that nothing has been gained by the preceding development because the $R(E)$ of (3) remains as unknown as the $V(r)$ was for $r < a$. It can be shown, however, that R permits a convergent expansion into an, in general, infinite series

$$R(E) = \sum_{\mu} \frac{\gamma_\mu^2}{E_\mu - E}. \tag{7}$$

It follows that R is an analytic, in fact a meromorphic function, all the poles of which are on the real axis and all have negative residua. The proof of (7) can be obtained,[1] in the simplest case which was outlined above, by a modification of the expansion of the Green's function of (1), similar to that given in L. Bieberbach's *Theorie der Differentialgleichungen* (Julius Springer, Berlin 1923).

[1] For further details cf. e.g. E. P. Wigner and L. Eisenbud, Phys. Rev. 72, 29, (1947).

However, the generality of (7) is actually much greater than would appear herefrom. The L in (1) need not be a differential operator but can be an arbitrary self adjoint operator with the property that it acts, on any function which vanishes for $r < a$, as a differential operator of the form (1). Even this condition can be weakened to the requirement that the probability of the particle being inside a can be expressed as a hermitian form of $\varphi(r)$ and that a current at a, which is also a hermitian form of $\varphi(r)$, can be defined so that these quantities satisfy the usual continuity equation.[2] Furthermore, the concept of R can be extended[1] to functions with several variables $\varphi_E(x_1, x_2, \cdots, x_m)$ in which case $R(E)$ becomes an infinite matrix function of E. Because of the appearance of functions with the same essential properties under such a variety of conditions, it appears desirable to characterize these functions[3] in a more abstract fashion and to determine some of their common properties. This will be done for scalar (non matrix) R functions in Section II.

The direct incentive to the present investigation was furnished by the following problem. In many cases of considerable physical interest the variation of the F, G, F', G' with energy is so much smaller than that of R that it is reasonable to consider them to be independent of E. We shall see in the next section that if one does this, i.e., considers the F, G and their derivatives at a as constants, the $S(E)$ of (4a) permits an expansion similar to (7)

$$S = \text{const} + \sum_{\mu} \frac{r_\mu^2}{H_\mu - E}. \tag{7a}$$

Inserting this into (6) one finds that $\sigma/(4\pi/E)$, i.e., the cross section measured in terms of the maximum possible cross section $4\pi/E$, has maxima where S is infinite and the width of each maximum is equal to the negative residue of the corresponding pole, i.e., equal to r_μ^2. It is reasonable now that the width of these levels depend on the long range interaction and indeed the r_μ of (7a) contain what is known as the penetration factor. It seems reasonable to assume, on the other hand, that the density of the resonances H_μ, the distribution function for the distances $H_{\mu+1} - H_\mu$ of successive resonances, as well as the distribution function of the widths r_μ^2 apart from the absolute scale, depend only on the short range interaction R. Expressed mathematically, one could expect that the density of the H_μ and the above distribution functions be invariant under a linear fractional transformation: that the number of the $H_{\mu+1} - H_\mu$

[2] Cf. G. Goertzel, Phys. Rev. 73, 1463 (1948) and E. P. Wigner, ibid. 70, 15 (1946) especially page 18.

[3] They were first considered, probably, in connection with problems of radio signal transmission, by the group of workers at Bell Telephone Laboratories. Cf. G. A. Campbell, Bell System Tech. Journ. 1, No. 2, p. 1 (1922), O. J. Zobel ibid. 2, 1, (1923), R. M. Foster, ibid. 3, 259, (1924). The great generality of the occurrence of expressions of the form (3) in problems of collision theory finds particularly striking expression in G. Goertzel's treatment of the collision of particles obeying Dirac's equation (Phys. Rev. 73, 1463, (1948)). Cf. also the author's review article Nuclear Reactions and Level Widths, Am. Journ. of Physics, 17, 99, (1949).

which exceed a certain number I be the same as the number of $E_{\mu+1} - E_\mu$ which exceed the same number I, and that the frequency of the deviation of γ_μ^2 from the average by a certain factor be the same as the frequency of the deviation of r_μ^2 from the average r_μ^2 by the same factor. This we will indeed find corroborated under rather general conditions. This will be the subject of Sections III and IV.

II. Properties of R-Functions

The R function will be defined as a meromorphic function, the imaginary part of which is non negative on the upper half plane and non positive on the lower half plane. However, a real constant will not be considered an R function.

The variable E of R will be denoted by z henceforth and a number of auxiliary theorems for R established. It is clear, first of all, that:

II.1. *A linear combination of R functions with positive* (positive shall mean, throughout this article, real positive) *coefficients is again an R function.*

II.2. *An R function is real on the real axis and assumes real values only on the real axis.* This follows from the fact that the imaginary part of an analytic function, unless it is a constant, cannot assume its minimum or its maximum inside a closed domain. It follows that the terms "non negative" and "non positive" in the definition of an R function can be replaced by "positive" and "negative", respectively.

It also follows that all the derivatives of R are real at any regular point of the real axis and that all the coefficients of the power series development of R at such points are real. Hence, in the neighborhood of such points, R assumes conjugate complex values for conjugate complex values of its argument

$$R(z^*) = (R(z))^* \tag{8}$$

and this equation will hold throughout the complex plane. This permits one to restrict many of the considerations to the upper half plane.

II.3. *The derivative of an R function is positive at every regular point of the real axis.* In fact, let us denote the first two terms of the power series expansion of R around the real point z_0 by $R_0 + R_n(z - z_0)^n + \cdots$ where R_0 is real and $R_n \neq 0$. Since there is a neighborhood of z_0 in which these two terms dominate, $R_0 + R_n(z - z_0)^n$ must have a positive imaginary part in the upper, negative imaginary part in the lower half plane if $z - z_0$ is sufficiently small. This can be the case only if $n = 1$ and $R_n = R_1 \geqq 0$. Since, according to assumption, $R_n \neq 0$, we have $R_1 > 0$.

II.4. *All poles of an R function must be on the real axis.* In the neighborhood of a pole $a(z-z_0)^{-n}$ the R function's imaginary part assumes both positive and negative values. This can be the case only if z_0 is real.

II.5. *All poles of an R function are simple and their residues negative.* In a sufficiently small neighborhood of a pole z_0 the value of R is dominated by the highest singularity $a(z - z_0)^{-n}$ or $ar^{-n}e^{-in\varphi}$ if we introduce polar coordinates for $z - z_0 = re^{i\varphi}$. This must be real for $\varphi = 0$ whence a is real. Since its derivative with respect to r must be positive for $\varphi = 0$, the a must be negative. Unless

$n = 1$, the imaginary part of $r^{-n}e^{-in\varphi}$ will change sign in the upper half plane. The diagram of a typical R function is given in Fig. 1.

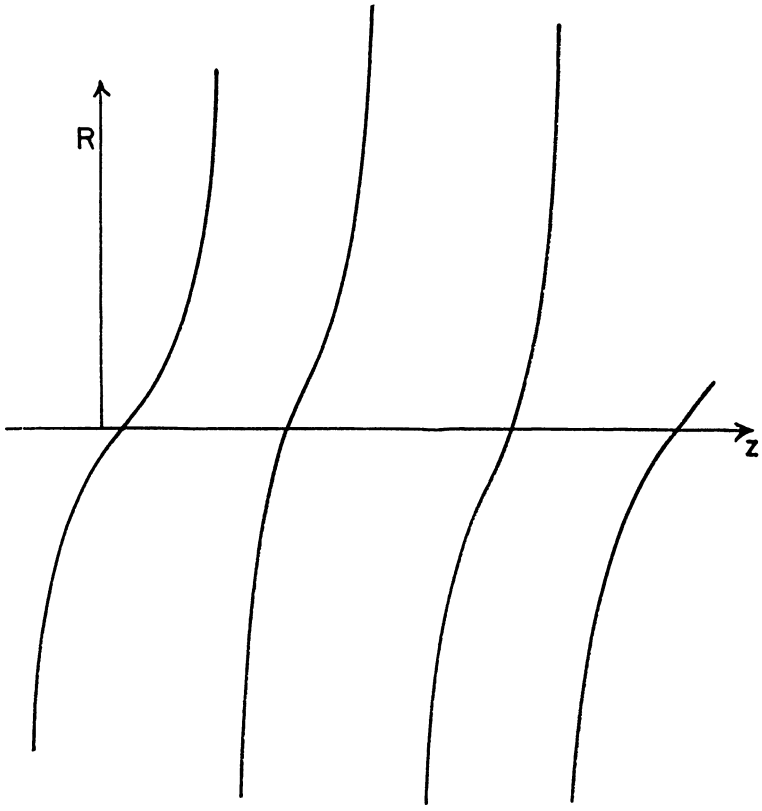


FIG. 1. Behavior of a typical R function for real values of its argument

II.6. The functions

$$\frac{az + b}{a'z + b'} \quad \text{with} \quad ab' - a'b > 0 \tag{9}$$

and

$$a \operatorname{tg}(b(z - z_0)) \text{ with } ab > 0 \tag{9a}$$

(all a, b, a', b' real) are R functions.

II.7. *The only entire R functions are the linear functions $R = az + b$ with positive a and real b.*

It follows from II.2 that R can assume real values only on the real axis. If R has no singularities on this axis it is, because of II.3, a monotonously increasing function on this axis. It has been observed by Mr. D. W. Blackett, to whom the first proof of this theorem is due, that, as a consequence of the above, R will not assume outside of a circle about $z = 0$ any of the real values which it

assumes inside the circle so that, by Picard's theorem, it can have no essential singularity anywhere. However, one can prove the theorem also without reference to Picard's theorem, e.g., by considering the complex phase of $(R(z) - R(0))/z$. This phase is clearly always between $-\pi$ and π so that the above expression is a constant. Hence $R(E)$ must be a linear function of z.

II.8. *Every R function can be expanded into an absolutely convergent Mittag-Leffler series:*

$$R = \alpha z + \beta + \sum_{\mu} \left(\frac{\gamma_\mu^2}{Z_\mu - z} - \frac{\gamma_\mu^2}{Z_\mu} \right) \tag{10}$$

with α positive, β and γ real. The proof of this theorem, which the writer had used for the considerations given in Sections III and IV, has been given by Drs.

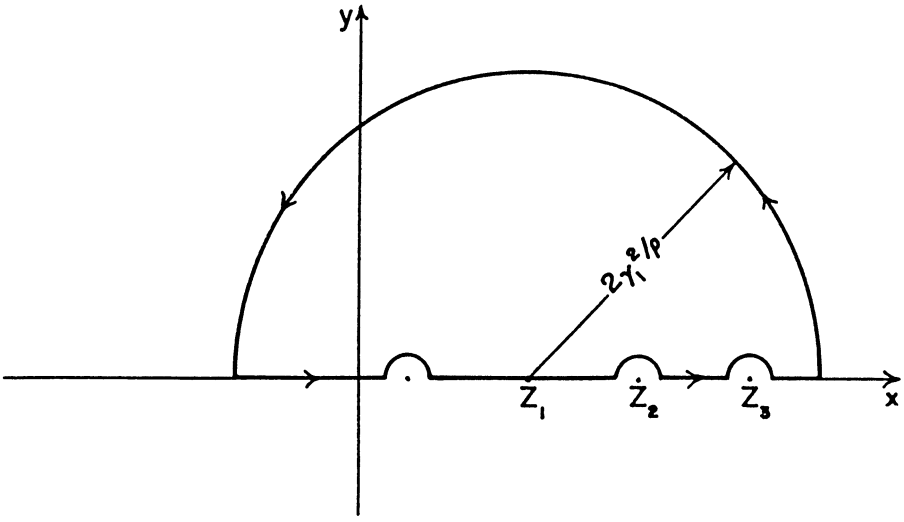


FIG. 2. Path along which the imaginary part of $R - \gamma_1^2(Z_1 - z)^{-1}$ is everywhere larger than $-\frac{1}{2}\rho$.

M. Schiffer and V. Bargmann and the writer is much indebted to them for permission to reproduce their proof. It starts from the following Lemma.

II.8a. *If Z_1 is a pole of R with the meromorphic part $\gamma_1^2/(Z_1 - z)$ then*

$$R_1(z) = R(z) - \gamma_1^2/(Z_1 - z) \tag{11}$$

is also an R function. Clearly R_1 is a meromorphic function, it needs to be shown only that the sign of its imaginary part is the same as the sign of the imaginary part of z. Suppose this is not the case and $-iR_1$ has a negative real part for some point z' of the upper half plane which will be denoted by $-\rho$ where $\rho > 0$. Let us now consider $-iR_1$ along a closed path as indicated in Fig. 2. This runs along the real axis, where $-iR_1$ is imaginary, but avoids the poles, except Z_1, which is a regular point of R_1. It then follows a semicircle about Z_1 back to

the real axis. The radius of this semicircle is larger than $2\gamma_1^2/\rho$ and large enough to include z'. Along this semicircle, the real part of $-iR_1$ is larger than $-\rho/2$. The poles along the real axis can be avoided so closely that the real part of $-iR_1$ be positive. As a result, the real part of $-iR_1$ is nowhere smaller than $-\rho/2$, it cannot assume the value $-\rho$ inside the path. This proves 8a.

Let us now subtract the singularities of $R(z)$ one after the other, i.e. consider the function.

$$R_n(z) = R(z) - \sum_{\mu=1}^{n} \frac{\gamma_\mu^2}{Z_\mu - z}. \tag{12}$$

This is, according to 8a, still an R function. If z_0 is a real regular point, the derivative of R_n must be positive by II.3. Hence

$$\sum_{\mu=1}^{n} \frac{\gamma_\mu^2}{(Z_\mu - z_0)^2} < R'(z_0) \tag{12a}$$

and this holds for every n. Hence the infinite series $\sum \gamma_\mu^2/(Z_\mu - z_0)^2$ converges absolutely and this holds also for the series

$$\sum \left(\frac{\gamma_\mu^2}{Z_\mu - z_0} - \frac{\gamma_\mu^2}{Z_\mu} \right) = \sum \frac{\gamma_\mu^2 z_0}{(Z_\mu - z_0) Z_\mu} \leqq (|z_0| + z_0^2/|Z_s|) \sum \frac{\gamma_\mu^2}{(Z_\mu - z_0)^2} \tag{13}$$

where Z_s is the pole with the smallest absolute value. It is easy to see that the series (13) converges also for non real z_0 and one can consider, therefore, the function

$$R(z) - \sum \left(\frac{\gamma_\mu^2}{Z_\mu - z} - \frac{\gamma_\mu^2}{Z_\mu} \right).$$

This is a limit of R functions and is, therefore, according to II.2, itself an R function or a constant. It has no poles anywhere on the complex plane and is, by II.7, a linear function of z with coefficients as described under II.8. This completes the proof of II.8.

II.9. *A linear fractional function of an R function with positive determinant and real coefficients is an R function*

$$S = \frac{a_1 R + b_1}{a_2 R + b_2} \qquad (a_1 b_2 - a_2 b_1 > 0). \tag{14}$$

S is clearly meromorphic: its poles are where $R = -b_2/a_2$. Furthermore, if one considers S as a function of R, it follows from II.6 that the imaginary part of (14) has the same sign as the imaginary part of R. This, however, has the same sign as the imaginary part of z.

The simplest function of the form (14) is $-b/R$ with $b > 0$. It follows now that $R_1 - b/R_2$ is an R function. Substituting for R_2 a similar expression, one finds that if the b are all positive, the continued fraction

$$R = R_1 - b_2/(R_2 - b_3/(R_3 - \cdots - b_n/R_n) \cdots)) \tag{15}$$

is also an R function if $R_1, R_2, R_3, \cdots$ are all R functions.

It follows from (14) and II.8 that, if the F, G, F', G' are considered as constants, the S of (4a) can be written in the form (7a). The absence of the linear term of (10) follows from the particular property of the R of (3) that they go to a finite complex number (in fact to a purely imaginary number) if the imaginary part of z becomes infinite and that this number is independent of the real part of z. This then follows also for S which entails the absence of the αz term of (10). The details of this last argument will be worked out below.

II.10. The S of (14) is, because of II.6, an R function of an R function. The sign of the imaginary part of such a function always follows the sign of the imaginary part of the argument. Analytic functions with this property will be called "following" functions. If such a "following" function is many valued, the imaginary part of every value must have the same sign as the imaginary part of the argument. Every R function is a following function but a following function is an R function only if it is meromorphic. The "following" functions form a group, except for the ambiguities which may occur in the definition of an analytic function of an analytic function, because the inverse of a following function, and a following function of a following function, are also following functions. This is, of course, not true for R functions, except if the second one has only a finite number of poles and it can have only one pole if its inverse is also to be an R function.

Next, a new normal form will be given to R functions in which they will be represented as the tg of a following function:

$$R = \operatorname{tg} f; \qquad f' = R'/(1 + R^2). \tag{16}$$

(This f' will later be called also the invariant derivative of R.) The f' has singularities (except at ∞) only where $R = \pm i$. These points $z_\mu = x_\mu \pm i\kappa_\mu (\kappa_\mu > 0$, cf. (8)) can have no condensation point except at infinity: a condensation point would be an essential singularity of $(R - i)^{-1}$ and hence also of R. If one expands R at the points z_μ into a power series $R(z) = i + a_n(z - z_\mu)^n + \cdots$ one finds for f'

$$f' = \frac{na_n(z - z_\mu)^{n-1} + \cdots}{2ia_n(z - z_\mu)^n + \cdots} \sim \frac{n}{2i(z - z_\mu)}. \tag{16a}$$

II.11. *The poles of f' lie symmetrically to the real axis, the residues have the form $n/2i$ where n is a positive integer for the poles above the real axis; the residues of the poles which lie symmetrically with respect to the real axis are oppositely equal.* Hence the meromorphic part of a pair of such poles has the form

$$\frac{n\kappa_\mu}{(z - x_\mu)^2 + \kappa_\mu^2} \tag{17}$$

where x_μ is real, κ_μ positive.

It is suggestive, therefore, to consider for f' a sum of expressions (17)

$$f' = \sum \frac{\kappa_\mu}{(z - x_\mu)^2 + \kappa_\mu^2}. \tag{*}$$

The multiplicity n of the roots z_μ, which occur in (17), is taken care of in (*) by including, for $n > 1$, the same term several times. Hence

$$(**) \qquad f = \sum \left(\operatorname{arctg} \frac{z - x}{\kappa_\mu} + \operatorname{arctg} \frac{x_\mu}{\kappa_\mu} \right).$$

Since the f of (**) is a following function and since the tg of it has no essential singularities in the finite plane, the tg of the f of (**) is an R function. If the sum in (*) is finite, it is in fact a rational R function. One can, furthermore, add a linear R function to the f of (**) without changing its character and thus obtain the expression

$$(18) \qquad R = \operatorname{tg} f = \operatorname{tg} \left\{ \eta z + \vartheta + \sum \left(\operatorname{arctg} \frac{z - x_\mu}{\kappa_\mu} + \operatorname{arctg} \frac{x_\mu}{\kappa_\mu} \right) \right\}$$

with positive η and κ_μ and real ϑ and x_μ which is clearly an R function if the series in it converges. Conversely, we have the theorem

II.12. *Every R function can be given the form* (18) *with a convergent series as the argument of the* tg. The proof of this theorem, which the writer had used for the considerations to be given in Section IV without having a rigorous proof for it, has been given by Dr. v. Neumann and the writer is much indebted to him for permission to reproduce his proof. It is patterned on the Schiffer-Bargmann proof given under II.8.

It follows from II.11 that f' has no other singularity in the finite plane but those at the $x_\mu \pm i\kappa_\mu$ with meromorphic parts given in (17). Hence e.g.

$$(19) \qquad f_1 = f - n \operatorname{arctg} (z - x_1)/\kappa_1$$

has no singularity at $x_1 \pm i\kappa_1$. One then has the Lemma which is the analogue of 8a:

II.12a. *The f_1 of* (19) *is a following function.* In order to prove this, let us consider the imaginary part of f_1. In the neighborhood of a singularity $z_\mu = x_\mu + i\kappa_\mu$ of the upper half plane, (16a) shows that f has the form $-\frac{1}{2}ni \ln (z - z_\mu)$ and hence one can draw a small circle around z_μ such that the imaginary part of f_1 be positive on that circle. Furthermore, f and also f_1 are regular and real on the real axis, so that the imaginary part of f_1 vanishes there. At a large distance from z_1, the imaginary part could become negative in the upper half plane but, by taking the distance large enough, it can be accomplished that it do not sink below $-\varepsilon$. This is true because the imaginary part of f is positive in the upper half plane, and because the imaginary part of the second term of (19) goes to zero for large $| z - x_1 |$ as $n/| z - x_1 |$. Hence, the imaginary part of f_1 is nowhere smaller than $-\varepsilon$ on the boundary of the domain which is bordered by the real axis and a sufficiently large semicircle, and from which small circles around the z_μ are excluded. Since, furthermore, the imaginary part of f_1 is a single valued function of z, it cannot sink below $-\varepsilon$ anywhere in the inside of this domain. As this holds for every ε, Lemma 12a follows.

Let us now consider the functions

$$f_N = f - \sum_{\mu=1}^{N} (\operatorname{arctg}(z - x_\mu)/\kappa_\mu + \operatorname{arctg}(x_\mu/\kappa_\mu)). \tag{20}$$

They are still following functions and, hence, the imaginary part of the sum is smaller than the imaginary part of f. It follows, as under II.8, that the imaginary part of the sum in (20), even if extended over all μ, converges at every regular point z_0 of f.

For $z_0 = is$ this means that

$$\tfrac{1}{2} \sum \ln \frac{1 + 2\kappa_\mu s/(x_\mu^2 + \kappa_\mu^2 + s^2)}{1 - 2\kappa_\mu s/(x_\mu^2 + \kappa_\mu^2 + s^2)}$$

converges. Since $\kappa_\mu s/(x_\mu^2 + \kappa_\mu^2 + s^2) < 1$ and since for $\xi < 1$ the $\frac{1}{2} \ln (1 + \xi)/(1 - \xi) > \xi$, the infinite sum

$$\sum \frac{\kappa_\mu}{x_\mu^2 + \kappa_\mu^2 + s^2}$$

must also converge for $s > 0$. Let us assume, for the sake of definiteness, that the terms in this sum are so arranged that the $x_\mu^2 + \kappa_\mu^2$ form an increasing, or at least a non decreasing, series. Since the $z_\mu = x_\mu + i\kappa_\mu$ have no condensation point, the numbers $x_\mu^2 + \kappa_\mu^2$ must tend to infinity. It then follows that

$$\sum \frac{\kappa_\mu}{x_\mu^2 + \kappa_\mu^2} = \sum \frac{\kappa_\mu}{|z_\mu|^2} \tag{21}$$

converges since from a definite μ on, its terms are surely less than twice the terms of the preceding sum. Since the denominator of (*) is larger than $|z|^2 - 2|z|\,|z_\mu| + |z_\mu|^2$ it follows in a similar way that (*) also converges wherever all terms of it are regular, and, indeed, that it is bounded for real z in every finite interval. As a result, the integral (**) of (*) converges also so that one can consider the function f_∞ obtained from (20) by letting N go to infinity. This is now a following function which has no singularity in the finite plane. It is, therefore, an R function and thus by II.7 a linear function as described there. This completes the demonstration of II.12 and of the possibility of representing every R function in the form (18). We also have verified (*) with the surmised non negative additional term η

$$f'(z) = \frac{R'}{1 + R^2} = \eta + \sum_\mu \frac{\kappa_\mu}{\kappa_\mu^2 + (z - x_\mu)^2}. \tag{18a}$$

This series converges for all values of z except if a term in (18a) becomes infinite by itself.

Actually, one has obtained in this way an infinite number of representations for R. One of these is given by (18), the others are obtained by representing

$a^{-1}R - b$ in that form (a positive, b real). Having done that, one can obtain for R an expression

$$R = b + a \operatorname{tg}\left\{\eta z + \vartheta + \sum_{\mu}\left(\operatorname{arctg}\frac{z - x_\mu}{\kappa_\mu} + \operatorname{arctg}\frac{x_\mu}{\kappa_\mu}\right)\right\} \tag{22}$$

with every real b and positive a. In the expansion of the same R, the η, ϑ, x_μ and κ_μ will depend on the choice of a and b.

III. Uniform R Functions

In the present and particularly the following Section, we finally come to the question which prompted the present investigation. We shall define, first, "uniform" R functions, in which the "density" of poles, suitably defined, and the "average" residue of poles, is constant, i.e. independent of the position of the interval on the real axis. This concept will then be refined in the following section by the stipulation that not only the average of the distances between poles, but also the statistical distribution of these distances, be independent of the position of the interval on the real axis. The same will be postulated also for the residues: not only their average but also their distribution over all possible values shall be constant. Naturally, it will be necessary to define these concepts more accurately.

An R function will be said to have a pole density if for every ε, there is an $L(\varepsilon)$ such that the number of poles in any interval of length $L > L(\varepsilon)$ on the real axis is between $(1 - \varepsilon)\rho L$ and $(1 + \varepsilon)\rho L$. The ρ in these expressions will be called the density of poles.

An R function will be said to have a definite strength if for every ε there is an $L(\varepsilon)$ such that the sum of the residues of all the poles within any interval of length $L > L(\varepsilon)$ on the real axis is between $-(1 - \varepsilon)sL$ and $-(1 + \varepsilon)sL$ with $s > 0$. The s of these expressions is then the strength of the R function.

An R function will be called uniform if it has both a pole density and a definite strength, and if there is no linear term αz in its expansion by (10). In some cases, it will be necessary to strengthen the above conditions by the subsidiary stipulation that there be positive numbers A and ν such that $L(\varepsilon) < A/\varepsilon^\nu$ if $\varepsilon < 1$. This will be referred to as the subsidiary condition for the rapidity of the convergence of the average of the γ_μ^2 (cf. (10)) to its asymptotic value.

A trivial example of a uniform R function is (s/π) tg $(\pi\rho z)$. Less trivial examples will be given below. It should be noted that the above requirements are very strong in the statistical sense inasmuch as a common $L(\varepsilon)$ must exist for the whole infinite real axis and that the above definition does not allow even rare exceptional regions in which the number of poles, or the sum of their residues, could be outside the above limits. We shall next establish a number of simple theorems for the density of poles, and for the strength, of R functions.

III.1. *The pole density of an R function is invariant under fractional linear transformations.* The S of (14) has poles where $R = -b_2/a_2$, i.e. has one pole between each pair of successive poles of R (cf. II.3 and Fig. 1). Hence, the num-

ber of poles of S, in any interval in which R has N poles, is between $N - 1$ and $N + 1$. Hence the $L(2\varepsilon)$ for S can be chosen in an arbitrary fashion provided that it is greater than the $L(\varepsilon)$ for R and also greater than $1/(\rho\varepsilon)$. If this is the case, the number of poles of S in any interval of length $L > L(2\varepsilon)$ is greater than

$$(1 - \varepsilon)\rho L - 1 > (1 - 2\varepsilon)\rho L.$$

Similarly, the number of poles within any similar interval[4] L will be smaller than $(1 + \varepsilon)\rho L + 1 < (1 + 2\varepsilon)\rho L$. It is clear from the above that

III.1′. *If R satisfies the subsidiary condition concerning the limitation $L(\varepsilon) < A/\varepsilon^{\nu}$, its transform will satisfy a similar condition.*

III.2. *If an R function has a definite strength s, and if there is no linear term αz in its expansion* (10), *the imaginary part of R will converge uniformly* (*in x*) *to πs as the imaginary part y of $z = x + iy$ goes to infinity.*

In order to show this convergence we note that since β, γ_μ, Z_μ are real, the imaginary part of the R of (10) is

$$\frac{1}{2i}(R(x + iy) - R(x - iy)) = \sum_\mu \frac{\gamma_\mu^2 y}{(x - Z_\mu)^2 + y^2}. \tag{23}$$

In order to find the y_ε (which is independent of x) such that for $y > y_\varepsilon$ (23) differ by less than $3\pi s\varepsilon$ from πs, we first find an L such that the sum of all γ_μ^2 in any interval of length L be between $sL(1 - \varepsilon)$ and $sL(1 + \varepsilon)$. Then we find the interval v which permits one to approximate the infinite integral

$$\int d\xi/(\xi^2 + 1) = \pi$$

better than by the factor $(1 + \varepsilon)$

$$2\sum_{n=0}^{\infty} \frac{v}{n^2 v^2 + 1} < \pi(1 + \varepsilon) \tag{24a}$$

$$2\sum_{n=0}^{\infty} \frac{v}{(n + 1)^2 v^2 + 1} > \pi(1 - \varepsilon). \tag{24b}$$

Actually, the first sum is clearly larger than π because it gives the area of the circumscribed rectangles of Fig. 3, the second sum is smaller than π because it gives the area of the inscribed rectangles. Since the two expressions differ by $2v$, (24a) and (24b) will be both satisfied if $2v < \pi\varepsilon$. We then take $y_\varepsilon = L/v = 2L/\pi\varepsilon$.

We now can replace the sum (23) by partial sums, the nth of which (n going from $-\infty$ to ∞) contains those terms of (23) for which $nL \leqq Z_\mu - x < (n + 1)L$. If we replace in each of these sums the $Z_\mu - x$ in the denominator by nL for

[4] We shall use here and subsequently capital letters to denote intervals on the real axis. The symbol of an interval, if it occurs in a formula, always denotes the length of the interval.

$n > 0$ and by $(n - 1)L$ for $n < 0$, we have increased the total sum. Hence the sum (23) is greater than $\sum_n (S_n + S_{-n-1})y\,(n^2L^2 + y^2)^{-1}$ where S_μ is the sum of the γ_μ^2 which correspond to the Z_μ of the interval between νL and $(\nu + 1)L$. By assumption, all S_ν are smaller than $Ls(1 + \varepsilon)$ and hence, the total sum

$$(25)\quad \frac{1}{2i}(R(x + iy) - R(x - iy)) < \sum_n \frac{2Ls(1 + \varepsilon)y}{n^2 L^2 + y^2} = \sum s(1 + \varepsilon)\frac{2L/y}{n^2L^2/y^2 + 1}.$$

Now for $L/y < v$, the last sum is, by (24a), smaller than $\pi(1 + \varepsilon)$ so that the whole expression (23) is smaller than $\pi s(1 + \varepsilon)^2$. A similar argument leads to the result that the same sum is larger than $\pi s(1 - \varepsilon)^2$ which then proves the proposition.

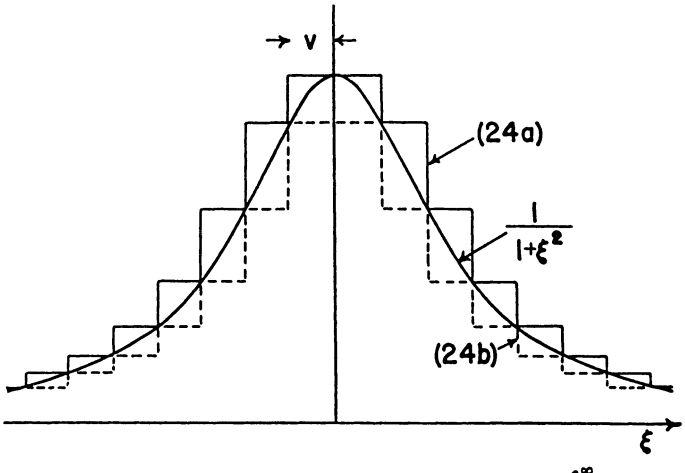


FIG. 3. Relation of the sums (24a), (24b) to the integral $\int_{-\infty}^{\infty} d\xi/(1 + \xi^2)$

One can again follow the above argument with the estimate for $L(\varepsilon)$ given in the subsidiary condition. One finds that the imaginary part is between $\pi s(1 \pm \varepsilon)$ if $y > A/(\pi 2^\nu \varepsilon^{\nu+1})$. Hence

III.2′. *If the subsidiary condition for the strength of R holds, there are two numbers A' and ν' so that it is sufficient to choose $y > A'/\varepsilon^{\nu'}$ in order to bring the imaginary part of R between the limits $\pi s(1 \pm \varepsilon)$.* The calculation which has been sketched above shows, in fact, that $\nu' = \nu + 1$ whence it follows that the imaginary part of $R(x + iy)$ approaches its asymptotic value at least as fast as Const $y^{-1/(\nu+1)}$

III.3. Conversely, *if the imaginary part of $R(x + iy)$ converges with increasing y, uniformly in x, and for all x, to the same number πs, then R has a definite strength s and has no linear term in its expansion* (10). The following proof owes an important simplification[5] to Dr. M. Schiffer to whom I am indebted on this score also. We first have to show the following Lemmas.

[5] The original proof compared the sum of the residua of an interval on the real axis with the integral, extended over an interval of similar length and on a line parallel to the real

III.3a. *If the imaginary part of* $R(x + iy)$ *is* πs, *the sum of the* γ_μ^2 *with* Z_μ *between* $x - y$ *and* $x + y$ *is smaller than* $2\pi sy$. This follows directly from (23) by omitting on the right side all terms the Z_μ of which is not in the interval $x - y$, $x + y$ and replacing in the remaining terms $(x - Z_\mu)^2$ by the larger y^2.

The derivative of the R of (10) is

$$R' = \alpha + \sum \gamma_\mu^2/(Z_\mu - z)^2.$$

Hence the absolute value of $R' - \alpha$ is, at $z = x + iy$,

$$|R'(x + iy) - \alpha| < \sum \frac{\gamma_\mu^2}{(Z_\mu - x)^2 + y^2}.$$

Because of (12a), the expression on the right side converges absolutely, and uniformly in y (except if x is equal to one of the Z_μ). Since all terms of the sum converge to zero with increasing y, the sum also converges to zero. Hence

III.3b. $R'(x + iy)$ *converges to* α *with increasing* y. If the imaginary part of $R(x + iy)$ is to converge, with increasing y, to a definite value πs, the real part of $R'(x + iy)$ cannot converge to a finite number α. Hence it follows from the premise of III.3, that the α of the expansion (10) vanishes.

Let us now proceed with the proof. The sum of the γ_μ^2 in an interval M can be expressed as the integral

$$\sum_{Z_\mu > x}^{Z_\mu < x+M} \gamma_\mu^2 = \frac{1}{2\pi i} \int_Q R\, dz = \frac{1}{2\pi i} \left(\int_{2,4} R\, dz + \int_{1,3} R\, dz \right) \tag{26}$$

where Q is a quadrangle, as indicated in Fig. 4. Its two horizontal sides 2 and 4 are obtained by shifting the interval M by iy and $-iy$, the vertical sides 1 and 3 go through x and through $x + M$, (neither of which is a pole), their length is $2y$. We shall find that the contribution of the horizontal sides of Q to the integral is closely enough Ms, and that the contribution of the vertical sides can be neglected if M is sufficiently larger than y. Lemma 3a will be needed for this estimate.

The values of R are conjugate to each other on opposite points of the horizontal sides. Hence only the imaginary part of R contributes to this part of the integral. If y is so large that the imaginary part is between the limits $\pi Ms(1 \pm \delta)$, the contribution of these sides will be between the limits

$$Ms(1 - \delta) < \frac{1}{2\pi i} \int_{2,4} R\, dz < Ms(1 + \delta). \tag{27}$$

In order to estimate the contribution of the vertical sides, we make use of the

axis, but removed from it to large imaginary values. As a result, the estimate of the integrals over the sections 1 and 3 of Q (cf. Fig. 4) was not necessary but the contribution of the residua near the ends of the interval on the real axis had to be estimated. This made the proof more lengthy than the present proof.

expansion (10). The contribution from $\alpha z + \beta$ naturally vanishes, (since $\alpha = 0$), the contribution from the μth term is

$$\frac{1}{2\pi i}\int_{1,3}\left(\frac{\gamma_\mu^2}{Z_\mu - z} - \frac{\gamma_\mu^2}{Z_\mu}\right)dz = \frac{1}{\pi}\cdot\gamma_\mu^2(\varphi_3 - \varphi_1). \tag{28}$$

In this $2\varphi_3$ and $2\varphi_1$ are the angles under which the sides 3 and 1 appear from Z_μ. These angles are positive if Z_μ is to the left of the corresponding side, negative if Z_μ is to the right therefrom.

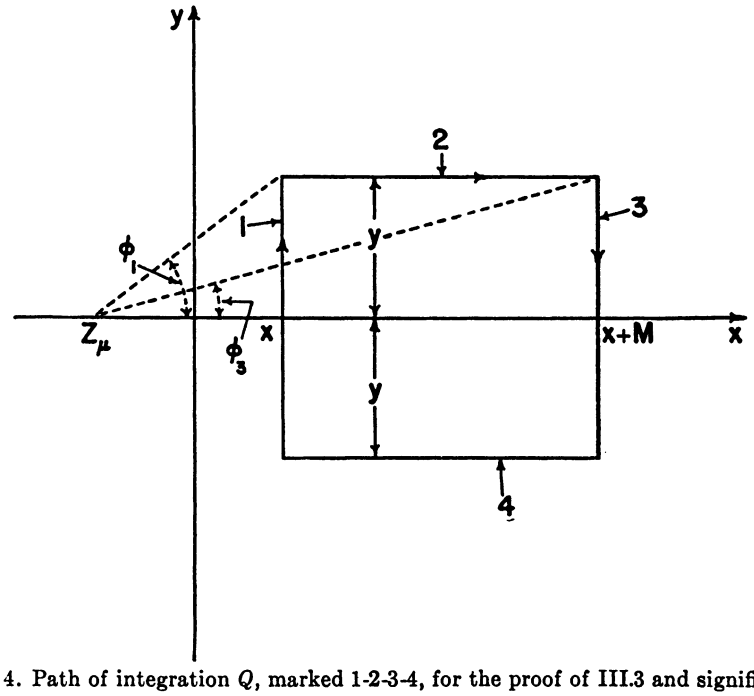


FIG. 4. Path of integration Q, marked 1-2-3-4, for the proof of III.3 and significance of the angles φ_3 and φ_1.

In order to estimate (28) we shall subdivide the real axis, beginning from x, into sections of length $2y$. For sake of simplicity, we shall assume that M is an integral multiple of this, say $M = 2my$.

For a Z_μ which is to the left of both sides, one can easily see that $\varphi_3 - \varphi_1 < y/(x - Z_\mu) - y/(x + M - Z_\mu)$ and also smaller than π. We use the latter estimate for the interval between $x - 2y$ and x, the former for the other intervals. The sum of the γ_μ^2 will be estimated by Lemma 3a. This way, we obtain as an upper bound for the contribution of the Z_μ to the left of x

$$2s(1+\delta)y\left(\pi + \sum\left(\frac{1}{2n} - \frac{1}{2n+2m}\right)\right) < 2sy(1+\delta)\left(\pi + \frac{1}{2}\int_1^\infty \frac{2m\,d\xi}{\xi(\xi+2m)}\right)$$
$$= sy(1+\delta)[2\pi + \ln(1+2m)].$$

The last part is obtained by setting $\xi = 2n$ and noting that the integrand is convex from below. The contribution of the Z_μ between x and $x + M$ is negative, that of the $Z_\mu > x + M$ can be estimated as above. Hence the contribution of the vertical sides is surely smaller than

$$\frac{1}{2\pi i}\int_{1,3} R\,dz < 2sy(1 + \delta)(2\pi + \ln(2m + 1)). \tag{29}$$

A lower limit for the contribution of the vertical sides to (26) can be obtained in the same way, except that in this case, the Z_μ between x and $x + M$ have to be considered. For these Z_μ, the contributions of both φ_3 and φ_1 are negative. They can be easily estimated to be each smaller in absolute value than

$$2s(1 + \delta)y\left(\pi + \frac{1}{2} + \frac{1}{4} + \frac{1}{2(m - 1)}\right) < sy(1 + \delta)(2\pi + \ln(2m - 1)).$$

Hence (29) holds not only for the contribution of the vertical sides of Q but also for the absolute value of this contribution.

From (26), (27) and (29) one infers that the sum of the γ_μ^2 over the interval M differs from Ms by a factor less than $1 \pm \varepsilon$ if

$$(1 + \delta)(1 + m^{-1}(2\pi + \ln(2m + 1))) < 1 + \varepsilon. \tag{30}$$

This, however, can be always accomplished by first choosing y so large that $\delta < \frac{1}{3}\varepsilon$ and then choosing an m so large that the second bracket of the left side also be smaller than $1 + \frac{1}{3}\varepsilon$. The interval M thus becomes $2my$. Conversely, for any interval M of this or larger size, an upper limit $M\pi S(1 + \varepsilon)$ for the sum γ_μ^2 can be obtained by means of the above argument, using a path of integration Q, the vertical sides of which are equal to the $2y$ defined above.

If, in the sense of III.2′, the above y can be assumed to have the form $A'/\delta^{\nu'}$, m can be assumed to be proportional to ε^{-2} and we see that the converse of 2′ also holds:

III.3′. *If there are positive numbers A' and ν' such that for $y > A'/\varepsilon^{\nu'}$ the imaginary part of $R(x + iy)$ differs from a constant πs by less than a factor* $(1 \pm \varepsilon)$ *then not only does R have a definite strength but the subsidiary condition also holds for the convergence of the average of the γ_μ^2.* One can easily calculate that if ν is the lowest exponent in the subsidiary condition, the imaginary part of $R(x + iy)$ cannot approach its asymptotic value uniformly faster than $y^{-1/(\nu-1)}$.

While the subsidiary condition concerning the rapidity with which the average of the γ_μ^2 approaches its asymptotic value has played only a subordinate role in the theorems 1, 2, 3, the following theorem cannot be established on the basis of mere uniform convergence of the sums of γ_μ^2. It states that

III.4. *If an R function has a definite strength and if the convergence of the average of γ_μ^2 to its asymptotic value is as rapid as described in the subsidiary condition, and provided further that the expansion of R into the form* (10) *has no linear term αz, the real part of R approaches the same value r for all x if the imaginary part y of $z = x + iy$ grows beyond all limits. Furthermore, this approach is so rapid that*

there are positive constants A'' and ν'' such that the absolute value of the difference between r and the real part R is smaller than ε if $y > A''/\varepsilon^{\nu''}$ where A'' and ν'' are independent of x.

Once III.4 is established, it will also follow from III.3′ and III.4 that the convergence of the imaginary part of R, in the sense that

$$\left| \frac{1}{2i} (R(x + iy) - R(x - iy)) - \pi s \right| < \varepsilon \quad \text{if} \quad y > A'/\varepsilon^{\nu'},$$

entails the convergence of the real part $\frac{1}{2}(R(x + iy) + R(x - iy))$ in a similar sense though possibly with different constants instead of the A' and ν'.

The proof of III.4 will be carried out in two steps: the first step will demonstrate the convergence of the real part of $R(x + iy)$ with increasing y to a constant, only the second step will be concerned with the rapidity of that convergence.

The first part of the proof will be based on the expansion (10) with $\alpha = 0$ and on the subsidiary assumption that the sum of the γ_μ^2 for the Z_μ of any interval of length L is between the limits $\pi s(1 \pm \varepsilon)$ if $\varepsilon < 1$ and $L > A/\varepsilon^\nu$. Under this assumption, it will be shown that the sum of each of the two terms under the summation sign of (10) converges separately in a certain sense and that the real part of the first sum, which is the only one which depends on $z = x + iy$, approaches zero as y goes to infinity.

The sum appearing in (10) converges absolutely and can be, therefore, arranged in an arbitrary fashion. Let us denote with $\sum_n$ the sum over those μ the Z_μ of which is either between l_n and l_{n+1}, or between $-l_{n+1}$ and $-l_n$ where

$$l_n = \tfrac{1}{2}(A/\delta^\nu)n^2. \tag{31}$$

The δ in (31) is quite arbitrary and the following calculations are valid, therefore, for every $\delta \leqq 1$. In fact, the final result could be obtained also by setting $\delta = 1$ in (31). Because of (31), $\sum_n$ contains terms from two intervals of length $(A/\delta^\nu)(n + \frac{1}{2}) > A/(n^{-1/\nu}\delta)^\nu$ and the sum of the γ_μ^2 in each of these intervals is, therefore, between the limits $s(l_{n+1} - l_n)(1 \pm n^{-1/\nu}\delta)$.

We can rewrite (10) with this notation for $\alpha = 0$ as

$$R = \beta + \sum_{n=0}^{\infty} \sum\nolimits_n \left(\frac{\gamma_\mu^2}{Z_\mu - z} - \frac{\gamma_\mu^2}{Z_\mu} \right). \tag{32}$$

Let us now consider the series

$$R_0(z) = \sum_{n=0}^{\infty} \left(\sum\nolimits_n \frac{\gamma_\mu^2}{Z_\mu - z} \right). \tag{33}$$

If it should turn out that the series in (33) converges for all z (excepting $z = Z_\mu$), it would be possible to write

$$R(z) = R_0(z) + \beta - R_0(0). \tag{33a}$$

The reason for the convergence of (33) is, of course, that the $\sum_n$ will contain about equally many positive and negative terms.

The imaginary part of (33) is the whole imaginary part of R and hence the series of the imaginary parts of (33) converges absolutely. Thus, we can restrict ourselves to the consideration of the real part of (33) and show that it converges. We shall see, at the same time, that it approaches zero as y becomes very large. The real part of (33) is

$$\tfrac{1}{2}(R_0(x + iy) + R_0(x - iy)) = \sum_n (t_n + t_{n'}) \tag{34}$$

where

$$t_n = \sum_n \frac{\gamma_\mu^2 Z_\mu}{(Z_\mu - x)^2 + y^2}; \qquad t_{n'} = \sum_n \frac{-\gamma_\mu^2 x}{(Z_\mu - x)^2 + y^2}. \tag{34a}$$

The series $\sum t_{n'}$ surely converges because it is simply $-x/y$ times the series for the imaginary part of R. In particular, for large y, the $\sum t_{n'}$ becomes $-x\pi s/y$ and hence goes to zero as $1/y$. We can restrict ourselves, therefore, to the consideration of the series $\sum t_n$.

When estimating t_n from the point of view of the convergence of the series $\sum t_n$, we can further restrict ourselves to such n that $l_n \geqq |x|$. The contribution of the smaller n to this series consists of a finite number of terms and cannot influence its convergence. Furthermore, each of these terms goes to zero as y^{-2} as y approaches ∞.

If $l_n \geqq |x|$, the terms $l_n \leqq Z_\mu < l_{n+1}$ give a positive contribution to t_n, the terms $-l_{n+1} \leqq Z_\mu < -l_n$ a negative contribution. Hence we shall obtain an upper limit for t_n by replacing in its denominator Z_μ by l_n in the former and by $-l_{n+1}$ in the latter terms. Similarly, we can replace Z_μ in the numerator by l_{n+1} and $-l_n$ in the two types of terms. We obtain this way the upper limit

$$t_n < \frac{l_{n+1}\sum_{n+}\gamma_\mu^2}{(l_n - x)^2 + y^2} - \frac{l_n \sum_{n-}\gamma_\mu^2}{(l_{n+1} - x)^2 + y^2}. \tag{35}$$

The $\sum_{n+}$ contains the γ_μ^2 for which $l_n < Z_\mu < l_{n+1}$, the $\sum_{n-}$ those for which $-l_{n+1} < Z_\mu < -l_n$. We will further increase (35) by inserting for the $\sum_{n+}\gamma_\mu^2$ its maximum and for $\sum_{n-}\gamma_\mu^2$ its minimum value. Thus t_n is smaller than

$$\begin{aligned} t_n &< s(l_{n+1} - l_n)\left[\frac{l_{n+1}(1 + n^{-1/\nu}\delta)}{(l_n - x)^2 + y^2} - \frac{l_n(1 - n^{-1/\nu}\delta)}{(l_{n+1} - x)^2 + y^2}\right] \\ &= \frac{(l_{n+1} - l_n)^2}{(l_{n+1} - x)^2 + y^2} + \frac{(l_{n+1} - l_n)^2 l_{n+1}(l_{n+1} + l_n - 2x)}{[(l_{n+1} - x)^2 + y^2][(l_n - x)^2 + y^2]} \\ &\quad + \frac{(l_{n+1} - l_n)\delta}{n^{1/\nu}}\left[\frac{l_{n+1}}{(l_n - x)^2 + y^2} + \frac{l_n}{(l_{n+1} - x)^2 + y^2}\right]. \end{aligned} \tag{35a}$$

One can clearly obtain an entirely similar estimate for the lower limit of t_n and hence also for its absolute value. One notices, furthermore, that the first two terms of the right side of (35a) go to zero as n^{-2}, the last one as $n^{-1-1/\nu}$. Hence the infinite series of the t_n converges absolutely.

The series (33) was decomposed above into four parts. The first of these was

its imaginary part which converges absolutely and which approached, for very large y, the value $i\pi s$. The second part consisted of the sum of the $t_{n'}$—its value is $-x/y$ times the imaginary part of R. The third part was a finite number of terms t_n, each of which approaches zero as y increases. The fourth part was the sum of the t_n for which $l_n \geqq |x|$, the absolute value of these terms is estimated in (35a). These t_n hence form an absolutely convergent series which is majorized by the series of terms given by the right side of (35a). However, all terms of this latter series monotonously decrease to zero as y increases. Consequently, the limit of this majorizing series goes to zero as y increases. The same then holds of the fourth part of the sum (33). It follows that the sum (33) converges absolutely and that the function defined by its approaches $i\pi s$ as y approaches infinity. It also follows that R approaches, for very large y, the value $i\pi s + \beta - R_0(0)$ — i.e. a value which is independent of x.

We now wish to show that this approach is as given in III.4. In order to do this, we first show that for $x = 0$, the $R(x + iy)$ is closer than ε to its limiting value if $y > A'/\varepsilon^{\nu'}$ where an upper limit for A' and ν' can be given solely in terms of A and ν. We note, for this purpose, that for $x = 0$, the $R_0(z)$ has only two parts in the above sense because the second and third parts vanish in this case and (35a) holds for all n. Furthermore, the approach of the first part, i.e. of the imaginary part of R, has been shown already in III.2′ to be sufficiently rapid, the corresponding A' and ν' depending only on A and ν (as noted there, $A' = A/\pi 2^{\nu}$, $\nu' = \nu + 1$). The fourth and only real part is majorized by the series formed of the right side of (35a) which is for $x = 0$

$$t_n < \frac{4(n+\frac{1}{2})^2}{(n+1)^4+\mu^2}\left(1 + \frac{(n+1)^2(2n^2+2n+1)}{n^4+\mu^2}\right) + \frac{(2n+1)\delta}{n^{1/\nu}}\left(\frac{(n+1)^2}{n^4+\mu^2} + \frac{n^2}{(n+1)^4+\mu^2}\right) \tag{35b}$$

wherein $\mu = 2\delta^{\nu}y/A$. The sum of these expressions can, however, be easily estimated: the first line's sum can be shown for $\mu > 1$ to be smaller than $B\mu^{-\frac{1}{2}}$ where B is a numerical constant, the sum of the second line can be similarly estimated to be smaller than $\varepsilon C_\nu \mu^{-\frac{1}{2}\nu}$ where C_ν is again a number depending only on ν. Hence the sum of the t_n, i.e. the real part of $R_0(iy)$ can be shown to be smaller than

$$\tfrac{1}{2}(R_0(iy) + R_0(-iy)) < \frac{BA^{\frac{1}{2}}}{\delta^{\nu/2}(2y)^{\frac{1}{2}}} + \frac{\delta^{\frac{1}{2}}C_\nu A^{\frac{1}{2}\nu}}{(2y)^{\frac{1}{2}}}. \tag{36}$$

This sum will be smaller than ε if both terms are smaller than $\frac{1}{2}\varepsilon$. In order to obtain a low estimate for ν', we may write for δ, which was left free at the beginning of the development $\delta = \varepsilon^{1-1/\nu}$. It then follows from (36) that the real part of $R_0(iy)$ is smaller than ε if y is larger than either of the expressions

$$y > \frac{2B^2A}{\varepsilon^{\nu+1}} \qquad y > \frac{2^{2\nu-1}C_\nu^{2\nu}A}{\varepsilon^{\nu+1}}. \tag{36a}$$

Alternately, we can say that the real part of $R(y)$ is closer than ε to its limiting value, for y which satisfy (36a).

The important thing about this last statement is that the limit of y, given in (36a), depends, apart from ε, only on A and ν: the B and C_ν are completely determined by these quantities. We can apply, therefore, the above result to the R function $R(x_0 + z)$ which is also a uniform R function and is characterized, in terms of our subsidiary condition, by the same A and ν. Hence the real part of $R(x_0 + iy)$ also is closer than ε to its limit for $y = \infty$ if (36a) holds. We know already that this limiting value is the same as the limiting value of $R(iy)$ so that III.4 is now proved in all detail. In fact, we can set $\nu' = \nu + 1$ therein, while A' is a more complicated expression in A and ν.

III.5. *If $R(z)$ is a uniform R function with a rapidity of convergence as given in the subsidiary condition, the same will hold true* (though possibly with different constants A and ν) *for any fractional linear expression S of R*, as given in (14).

It follows from III.1 that S has the same pole density as R and from III.1′ that the density also satisfies the subsidiary condition. It follows from III.2 and 4 that $R(x + iy)$ converges to $r + i\pi s$ if $y > \infty$ where s is the strength of R and r a real constant. It also follows from 2′ and 4 that y increases only with a power of y (in fact with the power $\varepsilon^{-\nu-1}$). This then is true also for S which converges in the same way to

$$\frac{a_1(r + i\pi s) + b_1}{a_2(r + i\pi s) + b_2} = \frac{(a_1 r + b_1)(a_2 r + b_2) + \pi^2 a_1 a_2 s^2 + i(a_1 b_2 - b_1 a_2)\pi s}{(a_2 r + b_2)^2 + \pi^2 a_2^2 s^2} \tag{37}$$

Hence, by III.3, S also has a definite strength, this being given by the imaginary part of (37), divided by π, and because of III.3′, the rapidity of the convergence of the average of γ_μ^2 is in accordance with the subsidiary condition.

This completes the set of theorems on uniform R functions. It may be worth while to add only the remark that:

III.6. *For every uniform R function, the distance of two closest poles never can exceed a definite limit and that there is an upper limit for the residue of any pole.* The former limit is, of course, equal to the $L(1)$ of the definition of the density of R functions; the latter is $2sL(1)$ where, however, $L(1)$ is the L occurring in the definition of the strength.

Let us now return briefly to the questions of Section I. It is reasonable to assume that in many cases the R, as defined by (2), satisfies some condition very similar to the uniformity condition given at the beginning of this Section, together with the subsidiary condition, perhaps with $\nu = 2$. Furthermore, although the reason herefore is not entirely clear theoretically, the R may have the form (3) or, more accurately (33), i.e. that it converges to a purely imaginary value for very large values of the imaginary part of the argument. It then follows from the development of the present Section, that the S of (4a) are also uniform R functions which all have the same density. On the other hand, the strength s_S of S is, because of (33), (4a) and (5), given in terms of the strength s of R by

$$s_S = \frac{s}{G^2 + \pi^2 s^2 G'^2}. \tag{38}$$

We can conclude from the above that the density of resonances is independent of the long range interaction, their average width depends by (38).

IV. Statistical R Functions

It would be natural, next, to investigate the distribution of the positions of the poles $z_\mu = x_\mu + i\kappa_\mu$ of f' as function of the distribution of the poles Z_μ of R and of the corresponding residues γ_μ^2. It turns out, however, that this is not an entirely simple problem. The difficulty is, fundamentally, that the distribution of the poles of f' does not in itself determine the distribution of the pole distances and residues of R. Rather, the latter distributions depend also on the relation of the poles of f' to each other, whether these are placed at random in the upper half plane, or occur frequently in closely spaced pairs, etc. The same holds also conversely, for the determination of the distribution of the poles of f' by the frequency of the γ_μ^2 and $Z_{\mu+1} - Z_\mu$. Hence it seems unlikely that one can obtain far reaching relations as long as one considers only probability functions which depend only on the location and character of single poles. As a consequence, the results of the present Section will be quite limited and will refer mainly to the influence of fractional linear transformations on the distribution of γ_μ^2 and of distances of poles of R.

One can consider the linear fractional transformation (14), as applied to a uniform R function which satisfies the subsidiary condition, as the succession of three transformations. The first of these

$$R_0 = R/\pi s - r \tag{39}$$

reduces the R function to a function of the type of R_0 of (33): it converges for very large imaginary part of the argument to a purely imaginary number. Furthermore, since the strength of the R_0 of (39) is $1/\pi$, this imaginary number is, by III.2, just i. An R function of this nature, i.e. one for which $R(i\infty) = i$ will be called normalized. The transformation (39) does not change the position of the poles and multiplies all residues by the same number. Hence (39) does not affect the distribution of the distances of successive poles nor the statistics of the ratios of the residues.

The second transformation is an "orthogonal" linear fractional transformation

$$S_0 = \frac{aR_0 + b}{-bR_0 + a} \qquad (a^2 + b^2 = 1) \tag{40}$$

which, as we know from III.1, III.5, and in particular (37), does not affect the density of the poles, and transforms normalized R functions into normalized R functions. It may influence, however, the frequency with which a certain distance of poles occurs and also influence the statistics of the residues, without affecting, however, the average value of either quantity. The quantities a, b in (40) are real, since the sign of both of them can be changed without changing the transformation, it would be possible to assume that $b \geqq 0$.

The last transformation is again an ordinary linear transformation

$$S = cS_0 + c' \tag{41}$$

to which the same remarks apply as were made to (39). The succession of (39), (40), (41) gives indeed (14) if

$$a = \frac{b_2'}{(a_2'^2 + b_2'^2)^{\frac{1}{2}}} \qquad b = \frac{-a_2'}{(a_2'^2 + b_2'^2)^{\frac{1}{2}}} \tag{40a}$$

$$c = \frac{a_1' b_2' - a_2' b_1'}{a_2'^2 + b_2'^2} \qquad c' = \frac{a_1' a_2' + b_1' b_2'}{a_2'^2 + b_2'^2} \tag{41a}$$

where

$$\begin{aligned} a_1' &= \pi s a_1 , & b_1' &= b_1 + \pi s r a_1 \\ a_2' &= \pi s a_2 , & b_2' &= b_2 + \pi s r a_2 . \end{aligned} \tag{40b}$$

Since the transformations (39) and (41) affect the statistics of pole distances and of residues only in a trivial fashion, most of the rest of the present Section will be devoted to orthogonal fractional transformations (40). The subscript 0 will be omitted henceforth from the R_0 and its transform S_0. The effect of orthogonal fractional transformations on an R function can be surveyed relatively easily by the f', defined in (16).

With $\operatorname{tg} \varphi = b/a$, the transformation (40) can also be written as

$$S = \frac{R \cos \varphi + \sin \varphi}{-R \sin \varphi + \cos \varphi} = \frac{R + \operatorname{tg} \varphi}{1 - R \operatorname{tg} \varphi} \tag{42}$$

where the φ is determined, by the transformation, only up to an integer multiple of π. If one writes, in the sense of (16), $R = \operatorname{tg} f$, (42) yields $S = \operatorname{tg}(f + \varphi)$. Hence

IV.1. *The transformation* (42) *replaces f by $f + \varphi$ and leaves f' invariant.* Hence $f' = R'/(1 + R^2)$ will be called the invariant derivative of the normalized uniform R function. The R functions (42), which have a common invariant derivative, will be called a family of normalized R functions.

IV.2. *There is one, and only one, transform of R which has a pole at a given real point x.* The corresponding φ is given by the equation $\operatorname{ctg} \varphi = R(x)$. As φ increases the poles of R move continuously to the left; at $\varphi = \pi$ every pole has arrived at the original position of the next smaller pole. As a result of the present remark, in order to prove a theorem for all members of a family of R functions, i.e. for all S of (42) with $0 \leqq \varphi < \pi$, one can prove it for all positions of a single pole situated between two adjoining poles of one member of the family.

IV.3. *If an orthogonal fractional transform S of R has a pole at x, the residue of this pole is $-1/f'(x)$.* This follows from the fact that $f' = R'/(1 + R^2)$ is also equal to $S'/(1 + S^2)$. Inserting into this expression, for S, its expansion (10) gives the desired result.

The above remarks enable one to survey the family (42) of normalized R in the following way. One first plots $f'(z)$ for real z. This is common for all S of (42) with any φ. The poles of the S, which has a pole at x, can be obtained by marking off, from x, such intervals on the real axis that the areas bordered by these in-

tervals, by f', and by lines perpendicular to the real axis, be integer multiples of π. The length of the side of such an area which joins the real axis with the f' curve is the reciprocal of the negative residue of the pole which is at its foot. The procedure is illustrated in Fig. 5.

IV.4. *If the* η *of* (18) *is positive (not zero) it is a lower bound of* $f'(z)$ *for real* z *and its reciprocal an upper bound for the absolute values of the residues of* R *and* S. *If* $f'(z)$ *has also an upper bound* $f'_{\max}$ *for real* z *the reciprocal of this is a lower bound for the* $\kappa_\mu > 1/f'_{\max}$ *and also the absolute values of the residues of* R *and* S. Furthermore, *the absolute value of* f'' *is also bounded on the real axis in this case, the upper limit being given by* (43).

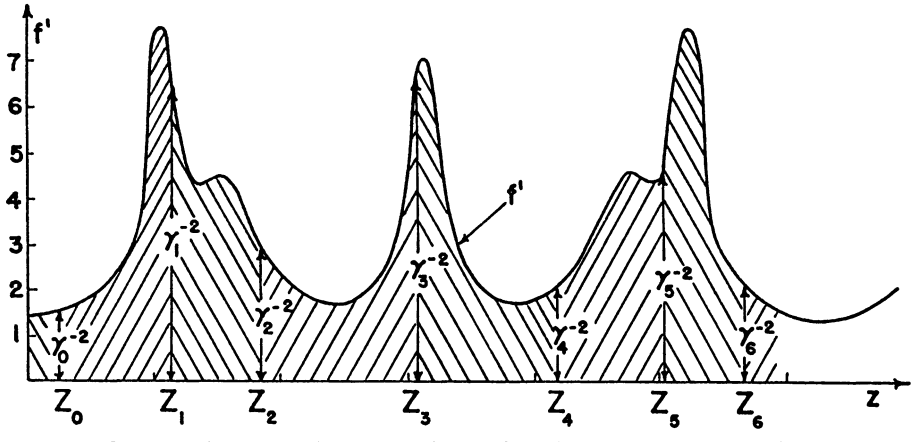


FIG. 5. Construction to obtain the position of the poles Z_1, Z_2, Z_3, $\cdots$ and the reciprocals of the corresponding residua γ_1^2, γ_2^2, γ_3^2, $\cdots$ of an R function from its invariant derivative $f' = R'/(1 + R^2)$. The area of the shaded regions is π, the ordinates which border them give the positions of the poles and their length is equal to the reciprocal of the corresponding residues (with opposite sign). The f' curve is common to all R functions of a family (40). If the construction is started from a different Z_0, the poles and residues of another member of the family will be obtained.

Only the last part of this theorem needs to be verified, the others immediately follow from (18a) and IV.3. We have for real z

$$|f''(z)| = \left|\sum_\mu \frac{2\kappa_\mu(z - x_\mu)}{(\kappa_\mu^2 + (z - x_\mu)^2)^2}\right| \leqq \sum_\mu \frac{\kappa_\mu}{\kappa_\mu^2 + (z - x_\mu)^2}\left|\frac{2\kappa_\mu(z - x_\mu)}{\kappa_\mu^2 + (z - x_\mu)^2}\right| f'_{\max}.$$

Every term has been multiplied by $f'_{\max}\kappa_\mu$ which is larger than 1. The sum on the right side will be further augmented if we replace the second factor by 1. Hence we obtain for real x

$$|f''(x)| \leqq (f'(x) - \eta)f'_{\max} \leqq (f'_{\max} - \eta)f'_{\max}. \tag{43}$$

We next prove the following theorem.

IV.5. *Let* R *be an* R *function in the expansion* (18) *of which* $\eta > 0$ *and the invariant derivative of which is also bounded from above* $f'(x) < f'_{\max}$ *on the real axis.*

Consider then the R function R_ε with $\varepsilon < 1/f'_{\max}$ in the expansion (18) of which η, ϑ, κ_μ are the same as in the expansion of R, the x_μ are, however, replaced by $x_\mu + \xi_\mu$ with $-\varepsilon < \xi_\mu < \varepsilon$. The poles of R shall be denoted by Z_μ, those of R_ε by $Z_\mu + \zeta_\mu$, the residues of R be denoted by $-\gamma_\mu^2$, those of R_ε by $-(\gamma_\mu^2 + \delta_\mu)$. *Then* (45a) *gives an upper limit for* ζ_μ, (45b) *for* δ_μ.

The value of f at the pole Z_μ must be equal to $\pi(\mu + \frac{1}{2})$

$$\eta Z_\mu + \vartheta + \sum_\nu \left(\operatorname{arctg} \frac{Z_\mu - x_\nu}{\kappa_\nu} + \operatorname{arctg} \frac{x_\nu}{\kappa_\nu}\right) = \pi(\mu + \tfrac{1}{2}). \tag{44a}$$

The argument of the tg in the expansion of R_ε by (18) shall be denoted by f_ε, hence the invariant derivative of R_ε by f'_ε. The value of f_ε at $z = Z_\mu + \zeta_\mu$ is also $\pi(\mu + \frac{1}{2})$. We shall substitute for this the value of f_ε at Z_μ plus the integral of f'_ε between Z_μ and $Z_\mu + \zeta_\mu$

$$\begin{aligned} \eta Z_\mu + \vartheta + \sum_\nu \left(\operatorname{arctg} \frac{Z_\mu - x_\nu - \xi_\nu}{\kappa_\nu} + \operatorname{arctg} \frac{x_\nu + \xi_\nu}{\kappa_\nu}\right) \\ + \int_{Z_\mu}^{Z_\mu + \zeta_\mu} f'_\varepsilon(x)\, dx = \pi(\mu + \tfrac{1}{2}). \end{aligned} \tag{44b}$$

Comparing the two equations we obtain

$$\begin{aligned} \sum_\nu \left\{\operatorname{arctg} \frac{\xi_\nu \kappa_\nu}{\kappa_\nu^2 + (Z_\mu - x_\nu)(Z_\mu - x_\nu - \xi_\nu)} + \operatorname{arctg} \frac{\xi_\nu \kappa_\nu}{\kappa_\nu^2 + x_\nu(x_\nu + \xi_\nu)}\right\} \\ = \int_{Z_\mu}^{Z_\mu + \zeta_\mu} f'_\varepsilon(x)\, dx. \end{aligned} \tag{44}$$

The value of all arctg is between $-\pi/2$ and $\pi/2$ since they vanish for $\xi = 0$ and their argument cannot become infinite since $|\xi_\nu| < \varepsilon < 1/f'_{\max} < \kappa_\nu$. Hence the absolute value of the left side will be increased by substituting for the arctg the absolute value of its argument and by replacing the ξ_ν in the numerator by ε. The absolute value of the right side is surely larger than $\eta\,|\zeta_\mu|$. It follows that

$$\eta\,|\zeta_\mu| < \sum_\nu \left\{\frac{\varepsilon\kappa_\nu}{\kappa_\nu^2 + (Z_\mu - x_\nu)(Z_\mu - x_\nu - \xi_\nu)} + \frac{\epsilon\kappa_\nu}{\kappa_\nu^2 + x_\mu(x_\nu + \xi_\nu)}\right\}. \tag{45}$$

Again, because of $|\xi_\nu| < \kappa_\nu$, omitting the ξ_ν from the denominator of the right side does not increase any of the terms by more than a factor 2. The resulting expression is then $\varepsilon(f'(Z_\mu) + f'(0) - 2\eta)$. Thus we have

$$|\zeta_\mu| < \frac{2\varepsilon}{\eta}\,(f'(Z_\mu) + f'(0) - 2\eta) < \frac{4\varepsilon}{\eta}\,(f'_{\max} - \eta). \tag{45a}$$

Our construction further shows that

$$\frac{1}{\gamma_\mu^2 + \delta_\mu} - \frac{1}{\gamma_\mu^2} = f'_\varepsilon(Z_\mu + \zeta_\mu) - f'(Z_\mu).$$

Now $f_\varepsilon'(z) - f'(z)$ can be easily estimated for real z

$$\sum_\mu \frac{\kappa_\mu}{\kappa_\mu^2 + (z - x_\mu)^2} \frac{2(z - x_\mu - \xi_\mu)\xi_\mu + \xi_\mu^2}{\kappa_\mu^2 + (z - x_\mu - \xi_\mu)^2}$$

$$< \sum_\mu \frac{\kappa_\mu}{\kappa_\mu^2 + (z - x_\mu)^2} \left(\frac{|\xi_\mu|}{\kappa_\mu} + \frac{\xi_\mu^2}{\kappa_\mu^2} \right) < 2\varepsilon(f'(z) - \eta) f'_{\max}$$

since $|\xi_\mu| < \varepsilon$ and $1/\kappa_\mu < f'_{\max}$. Making use of the fact that both γ_μ^2 and $\gamma_\mu^2 + \delta_\mu$ are larger than $1/\eta$ we obtain

$$\eta^2 |\delta_\mu| < 2\varepsilon(f'_{\max} - \eta) f'_{\max} + f'(Z_\mu + \zeta_\mu) - f'(Z_\mu).$$

The last difference is, by the mean value theorem and (43), smaller than $\zeta_\mu(f'_{\max} - \eta) f'_{\max}$ and we have therefore, by (45a) and IV.4

$$|\delta_\mu| < 4\varepsilon(f'_{\max} - \eta) f'^2_{\max}/\eta^3. \tag{45b}$$

We now come to the definition of statistical R functions. An R function will be called statistical with distributions for spacing $D(Z)$ and strength $s(r)$ if R is uniform with density $\rho = D(0) = s(0)$ and strength $s = \int s(r)\, dr$ and if, for any positive $\varepsilon < 1$ there is an $L(\varepsilon)$ such that for any interval[4] $L > L(\varepsilon)$ on the real axis

$$\int_0^\infty |D(Z) - D_L(Z)/L|\, dZ < \varepsilon \tag{46}$$

$$\int_0^\infty |s(r) - s_L(r)/L|\, dr < \varepsilon \tag{46a}$$

where $D_L(Z)$ is the number of distances between closest poles, contained in L, which are larger than Z. Similarly $s_L(r)$ is the number of residues which are smaller than $-r$ and which belong to the poles of the interval L. Obviously $D_L(Z)$ and $s_L(r)$ are monotonously decreasing functions of their argument and this holds for D and s also. Because of III.6, both vanish beyond a maximum value of their argument.

The above definition considers, at best, statistical properties of pairs of poles. As was mentioned before, it would be reasonable to consider also more intricate relations between properties of poles, such as for instance the number of adjoining triplets of poles with distances exceeding Z and Z' between them, or the frequency of adjoining pairs with a distance in excess of Z and residues exceeding r and r'. This would necessitate introducing a series of distribution functions with an increasing number of variables (similar to the Fock space of quantum theory) which will not be undertaken.

Let us now consider the functions

$$R = \operatorname{tg}\left[\eta z + \vartheta + \sum_{k=1}^{K} \operatorname{arctg}\left[\operatorname{Th}(\lambda_k \kappa_k) \operatorname{tg}(\lambda_k(z - x_k))\right] \right] \tag{47}$$

in which the sum over k is a finite one; all greek symbols excepting ϑ are positive. These will be proved to be statistical R functions the distributions of which are invariant under orthogonal fractional transformations.

The R of (47) is a following function of following functions and has no essential singularity in the finite complex plane. It is, therefore, an R function. For $z \to i\infty$ it assumes the value i and is, hence, a normalized uniform R function. It can, of course, also be written as a rational function of tg functions but that would offer no obvious advantage. The invariant derivative of (47) is

$$\begin{aligned} f' &= \eta + \sum_k \frac{\lambda_k \operatorname{Th}(\lambda_k \kappa_k)(1 + \operatorname{tg}^2 \lambda_k(z - x_k))}{1 + \operatorname{Th}^2(\lambda_k \kappa_k) \operatorname{tg}^2(\lambda_k(z - x_k))} \\ &= \eta + \sum_k \frac{\lambda_k \operatorname{Th}(\lambda_k \kappa_k)}{\cos^2(\lambda_k(z - x_k)) + \operatorname{Th}^2(\lambda_k \kappa_k) \sin^2(\lambda_k(z - x_k))}. \end{aligned} \tag{47a}$$

By integrating this equation term by term on the right side, one can bring (47) into the form (18). The poles of f' are located where any of the denominators of the right side of (47a) vanish:

$$x_k + \pi(n + \tfrac{1}{2})/\lambda_k \pm i\kappa_k \tag{48}$$

with every integer n. This means that the poles of f' form equidistant points on lines which are parallel to the real axis and at distances κ_k from it.

It follows from (47a) easily that f' is bounded for real z

$$f'_{\max} = \eta + \sum_k \lambda_k \operatorname{Cth}(\lambda_k \kappa_k). \tag{48a}$$

Hence the R of (47) fulfills the conditions of theorems IV.4 and IV.5.

The last preparatory theorem which we need is

IV.6. *Given any set* $\lambda_1/\eta, \lambda_2/\eta, \cdots, \lambda_K/\eta$ *of real numbers, there is an integer* $N(\varepsilon)$ *so large that among any* N *consecutive integers* $n + 1, n + 2, \cdots, n + N$ *there is at least one, say* n', *for which the fractional part of every* $n'\lambda_k/\eta$ *(with* $k = 1, 2, \cdots, K$*) is smaller than* ε. The fractional part of a number is its excess over the nearest integer. This theorem follows almost immediately from Kronecker's theorem.[6]

We now prove the theorem

IV.7. *If* R *is of the form* (47) *then there is, for every* ε *a length* $L(\varepsilon)$ *such that, for any two intervals* L, L' *of length* $L(\varepsilon)$ *on the real axis we have*

$$\int_0^\infty |D_L(Z) - D_{L'}(Z)|\, dZ < \varepsilon L(\varepsilon) \tag{49a}$$

$$\int_0^\infty |s_L(r) - s_{L'}(r)|\, dr < \varepsilon L(\varepsilon) \tag{49b}$$

[6] Cf. Theorem 5 of Chapter VII in Diophantische Approximationen by J. F. Koksma (Julius Springer, Berlin 1936) p. 83. A more direct proof, based only on Theorem 1 of Chapter I of Koksma's book (which follows immediately from Dirichlet's principle) is given in the Appendix to this paper.

where $D_L(Z)$ is the number of those distances between adjoining poles contained in L which are greater than Z and $s_L(r)$ is the number of those poles of L the residues of which are smaller than $-r$. The $D_{L'}(Z)$ and $s_{L'}(r)$ are the same numbers for the interval L'. It is clearly sufficient to prove this theorem for the case that the lower boundary of L is at 0.

The function $s_L(r) - s_{L'}(r)$ is a step function to which every pole in L contributes a rectangle of height 1 extending from zero to the absolute value of its residue, every pole of L' a rectangle of height -1 with a similar extension. When calculating the integral (49*b*) we will pair off most of the poles of L with a pole in L'. If the residues of these poles are r and r', their joint contribution to $s_L(r) - s_{L'}(r)$ will be a rectangle of height ± 1 and width $|r - r'|$. This way we can consider $s_L(r) - s_{L'}(r)$ as a sum of contributions of the various paired-off poles and of the several not paired-off poles. Since the integral of the absolute value is not larger than the sum of the integrals of the absolute values of the parts, the left side (49b) will be smaller than the sum of the absolute values of the residues of the not paired-off poles, plus the sum of the absolute values of the differences of the residues of the pairs of poles.

We are now ready to choose $L(\varepsilon) = \pi m N/\eta$. The integer N herein is, first of all, large enough so that the sum of the negative residues in any interval of length $\pi N/\eta$ be smaller than $2\pi s N/\eta$ and the total number of poles less than $2\pi N\rho/\eta$ (these are very mild requirements) and second, N shall be so large that there shall be, among any set of numbers $n + 1, \cdots, n + N$, at least one which, if multiplied by any of the λ_k/η, shall have a fractional part smaller than $\lambda_m\varepsilon/(4\pi A\rho)$ where A is the factor of ε in (45b) and λ_m the smallest of the λ. Finally the integer m is chosen so large that $4s/m < \frac{1}{2}\varepsilon$. (The strength s of the R of (47) is $1/\pi$. However, it seems easier to follow the ensuing argument if this fact is not made use of.)

The proof will be carried by finding, close to the lower boundary of L', a quasiperiod of R. The behavior of R in L' beyond that quasiperiod will be very similar to the behavior of R in L. The two intervals were chosen so long that the part of L' which is below the quasiperiod become negligible and the same hold for the part of L to which no part of L' corresponds.

Let us assume that the lower boundary of L' is at b while that of L is at 0. Let us then find the smallest $\pi n'/\eta > b$ with integer n' such that the fractional parts of $n'\lambda_k/\eta$ be all smaller than $\lambda_m\varepsilon/4\pi A\rho$. By assumption then $0 < \pi n'/\eta - b < \pi N/\eta$. The poles of L' which are between b and $\pi n'/\eta$ will not be paired off with poles in L. However, the total sum of their negative residues is, by assumption, less than $2\pi s N/\eta$ since they are all in an interval which is shorter than $\pi N/\eta$. Hence they will contribute to (49b), if this is calculated as outlined above, less than $2\pi s N/\eta$. Similarly, we will consider the poles of L which are between $\pi m N/\eta - \pi n'/\eta$ and $\pi m N/\eta$ as not paired off, their contribution to (49b) is also less than $2\pi s N/\eta$. The rest of the poles will be paired off. The total contribution of the not paired-off poles is, therefore, less than $4\pi s N/\eta = 4sL(\varepsilon)/m < \varepsilon L(\varepsilon)/2$.

In order to pair off the rest of the poles, let us consider the R function $R(z - \pi n'/\eta)$. The poles of this R function, in the interval between 0 and

$(\pi/\eta)(mN - n')$, will have the same residues as the poles of R which we wish to pair off. The poles in L with which we wish to pair them off, are also in the interval 0 to $(\pi/\eta)(mN - n')$. Let us consider, therefore, the x_μ which occur in the expansion of $R(z - \pi n'/\eta)$ into the form (18): they will be, by (48), of the form

$$x_k + \pi(n + \tfrac{1}{2})/\lambda_k + i\kappa_k + \pi n'/\eta = x_k + \pi(n + \tfrac{1}{2} + n'\lambda_k/\eta)/\lambda_k + i\kappa_k \tag{50}$$

for all k and n. Since, however, the fractional parts of the $n'\lambda_k/\eta$ are smaller than $\lambda_m\varepsilon/4\pi A\rho$, the poles (50) of the invariant derivative of $R(z - \pi n'/\eta)$ differ from the poles (48) of the invariant derivative of R by real numbers which are smaller than $\varepsilon/4A\rho$. Hence by IV.5, and in particular (45b), the corresponding residues of $R(z)$ and of $R(z - \pi n'/\eta)$ will differ by less than $\varepsilon/4\rho$. Hence, if we pair off corresponding residues of $R(z)$ and $R(z - \pi n'/\eta)$, each pair will contribute to (49b) less than $\varepsilon/4\rho$. The total number of poles in the interval 0 to $(\pi/\eta)(mN - n')$ is, however, surely smaller than $2\pi mN\rho/\eta$ so that the total contribution of the pairs of poles to (49b) is less than $\pi\epsilon mN/2\eta = \varepsilon L(\varepsilon)/2$. Hence $L(\varepsilon) = \pi mN/\eta$ satisfies (49b) if m and N are as given above. An entirely similar argument, which makes use of (45a) instead of (45b), shows that the same $L(\varepsilon)$ also satisfies (49a)(the coefficient of ε in (45a) is smaller than in (45b)). This then proves IV.7.

The $D_L(Z)$ is, if the interval L is the sum of other intervals L′, L'', the sum of the D for these intervals: $D_L(Z) = D_{L'}(Z) + D_{L''}(Z)$. The same holds of $s_L(Z)$. Let us take, hence, two intervals J and K, both larger than $m'L(\frac{1}{2}\varepsilon)$ where $L(\varepsilon)$ is the function defined in the preceding proof and m' an integer $m' > 8/\varepsilon$. Let us consider J as the sum of j intervals of length $L(\frac{1}{2}\varepsilon)$ and one interval J' which is shorter than $L(\frac{1}{2}\varepsilon)$. Then $j \geqq m'$ and the length of J is $J = jL(\frac{1}{2}\varepsilon) + J'$. Let us decompose K in a similar way into k intervals of length $L(\frac{1}{2}\varepsilon)$ and one interval K' shorter than $L(\frac{1}{2}\varepsilon)$ so that the length of K is $K = kL(\frac{1}{2}\varepsilon) + L'$. We can then majorize the integral

$$\begin{aligned}\int_0^\infty &| KD_J(Z) - JD_K(Z) | \, dZ \\ &= \int_0^\infty | (kL(\tfrac{1}{2}\varepsilon) + K')D_J(Z) - (jL(\tfrac{1}{2}\varepsilon) + J')D_K(Z) | \, dZ\end{aligned} \tag{51}$$

by 1.) jk integrals over $L(\frac{1}{2}\varepsilon) \mid D_L(Z) - D_{L'}(Z) \mid$, where both L and L' are intervals of length $L(\frac{1}{2}\varepsilon)$, plus 2.) j integrals over $\mid K'D_L(Z) \mid$, plus 3.) k integrals over $\mid J'D_L(Z) \mid$ and 4.) two integrals over $\mid KD_{J'}(Z) \mid$ and over $\mid JD_{K'}(Z) \mid$, respectively. The first jk integrals are smaller than $\frac{1}{2}\epsilon L(\frac{1}{2}\varepsilon)^2$ each, so that their total contribution to (51) is less than $\frac{1}{2}jk\epsilon L(\frac{1}{2}\varepsilon)^2$. The second integral is $K'L(\frac{1}{2}\varepsilon)$ since the integral

$$\int_0^\infty D_L(Z)\, dZ = L$$

is the sum of all pole distances in L, i.e. is L itself. As a result, the total contribution of the second type of integrals is $jK'L(\frac{1}{2}\varepsilon)$. Similarly, the contribution of the third type of integrals is $kJ'L(\frac{1}{2}\varepsilon)$. The sum of the last two integrals is $KJ' + JK'$. We have hence, by dividing (51) with JK and because of $JK > jkL(\frac{1}{2}\varepsilon)^2$ and $J', K' < L(\frac{1}{2}\varepsilon)$

$$\int_0^\infty \left| \frac{1}{J} D_J(Z) - \frac{1}{K} D_K(Z) \right| dZ < \tfrac{1}{2}\varepsilon + \frac{1}{k} + \frac{1}{j} + \frac{1}{k} + \frac{1}{j} < \varepsilon \tag{52}$$

since $k, j \geqq m' > 8/\varepsilon$. This proves that there is a length $M(\varepsilon) = m'L(\frac{1}{2}\varepsilon)$ such that if both J and K are intervals longer than $M(\varepsilon)$, (52) is valid. One can conclude, hence, in a well known manner that $D_M(Z)/M$ converges, with increasing length of the interval M, strongly to a limiting function $D(Z)$ and that (46) is valid for this $D(Z)$ and a $D_L(Z)$ with a sufficiently long interval L. The same thing applies to (46a) whence it follows that

IV.8. *The R functions of the form* (47) *with $\eta > 0$ are statistical R functions.* Since the orthogonal fractional transforms (42) of R have, by IV.1, the same form, this applies also to the orthogonal fractional transforms of R. We can, finally, easily prove now

IV.9. *The distributions for spacing and strength of R functions of the form* (47) *are invariant under orthogonal fractional transformation if the λ_k are incommensurable with η.*

The proof will make use of the preceding theorem, i.e. that both R and its orthogonal fractional transform

$$S = \operatorname{tg}\left[\eta z + \vartheta + \varphi + \sum_{k=1}^{K} \operatorname{arctg}\left(\operatorname{Th}\left(\lambda_k \kappa_k\left(\operatorname{tg} \lambda_k(z - x_k)\right)\right)\right)\right] \tag{53}$$

are statistical R functions. If the distribution functions of the two were different it would not be possible to find, for every ε, two intervals L and M of arbitrarily long but equal extension such that

$$\frac{1}{L}\int_0^\infty | D_L(Z) - E_M(Z) | \, dZ < \varepsilon \tag{53a}$$

$$\frac{1}{L}\int_0^\infty | s_L(r) - t_M(r) | \, dr < \varepsilon \tag{53b}$$

where D_L and s_L have the same significance as in (49a), (49b) and E and t denote the same quantities but refer to S rather than to R. Contrariwise, if we find for every ε arbitrarily long intervals L and M such that these equations are satisfied, we will have proved the identity of the distribution functions of R and S.

The proof now proceeds along the principle used for IV.7 but utilizing Kronecker's theorem directly, rather than in the modified form IV.6. According to that theorem, if the $\lambda_1/\eta, \lambda_2/\eta, \cdots, \lambda_K/\eta$ are all irrational and incommensurable, then there is, for every δ, an integer N such that the fractional parts

of the $\lambda_k N/\eta$ differ by less than δ from given fractions which we shall choose as the fractional parts of $\lambda_k\varphi/\pi\eta$. In other words, for this N

$$\left|\frac{\lambda_k}{\eta}N - \frac{\lambda_k\varphi}{\pi\eta} - n_k\right| < \delta \qquad (k = 1, 2, \cdots, K) \tag{54}$$

holds with integer n_k. The above assumes not only that the λ_k/η are irrational but also that they are not commensurable. Actually, this assumption is unnecessary: if λ_1 and λ_2 are integer multiples of a common measure $\lambda_1 = m_1\lambda$, $\lambda_2 = m_2\lambda$ we can stipulate (54) for λ, instead of λ_1 and λ_2, but replace δ by the smaller of δ/m_1 and δ/m_2. Hence we can assume the existence of the integer N with the property (54). We now choose the interval M as the displacement, by $(\pi N + \varphi)/\eta$, of the interval L. The poles and residues of S in this interval are the same as the poles and residues of $S(z - \pi N/\eta - \varphi/\eta)$ in the interval L. Hence we can compare the functions R and $S(z - \pi N/\eta - \varphi/\eta)$ in the interval L.

Both of these are R functions with the same η and ϑ in the normal form (18). The poles of the invariant derivative of R are given by (48), those of $S(z - \pi N/\eta - \varphi/\eta)$ are

$$\begin{aligned} &x_k + \pi N/\eta + \varphi/\eta + \pi(n + \tfrac{1}{2})/\lambda_k + i\kappa_k \\ &\qquad = x_k + \pi(n + \tfrac{1}{2} + \lambda_k N/\eta + \lambda_k\varphi/\pi\eta)/\lambda_k + i\kappa_k \end{aligned} \tag{55}$$

for all $k = 1, 2, \cdots, K$ and all integers n. Again, the poles (55) differ from the poles (48) by real numbers which are, because of (54), all smaller than $\pi\delta/\lambda_m$ where λ_m is the smallest of the λ. Hence, by IV.5, the poles and residues of R and of $S(z - \pi N/\eta - \varphi/\eta)$ can be brought arbitrarily close to each other by choosing a very small δ for (54). Hence, the distribution functions for the residues and pole spacings must be very nearly the same for these functions in every reasonably long interval L. The same then holds for $R(z)$ and $S(z)$ if the intervals L and M, which one compares, differ by a displacement $(\pi N + \varphi)/\eta$. This then proves (53a), (53b) and hence IV.8.

The above theorem shows the existence of a wide class of normalized statistical R functions the distribution functions of which are invariant under orthogonal fractional transformations. Nevertheless, this is a very partial result. From the construction illustrated in Figure 5 one would expect that the distribution functions of almost all normalized statistical R functions are invariant under orthogonal fractional transformations, i.e. that such invariance obtains for all normalized statistical R functions except for a set of measure 0 if such a measure could reasonably be defined in the space of these functions. However, this must remain a conjecture as far as the present paper is concerned. The fact that one has to exclude exceptional R functions when stating the conjecture is evident both from the construction illustrated in Figure 5 and from examples which one can give explicitly. The simplest one known to the writer has the form (47) but with a single λ which is commensurable with η (for

IV.8, it would be sufficient to assume that one of the λ is not commensurable with η):

$$R = \operatorname{tg}(z + \operatorname{arctg}(\vartheta \operatorname{tg} nz)) = \frac{\operatorname{tg} z + \vartheta \operatorname{tg} nz}{1 - \vartheta \operatorname{tg} z \operatorname{tg} nz} \tag{56}$$

with arbitrary positive $\vartheta \neq 1$ but integer n. This R, as well as its invariant derivative, have a period π and the construction of Fig. 5 shows that (56) is not invariant under the transformation (42).

From the point of view of the remarks made at the beginning of this Section, as well as after the definition of statistical R functions, it is interesting to note that the statistical nature of the R functions (47), as well as their invariance, was proved by showing that these, and their invariant derivatives, have quasi-periods. As a result, what we have shown above is actually more than implied in the definition of the statistical nature of R functions: it applies not only to the distribution of single residues and to the spacings between pairs of poles, but also to statistical connections between several residues and spacings of several successive poles.

Returning to the original set of questions, the above calculations prove that the density of resonances is independent of the long range interaction and that the dependence of the average value of their widths depends on the long range interaction in the manner given by (37). They make it very probable that the distribution of the spacings of the resonances and that the distribution of their widths, in terms of the average width, are also independent of the long range interaction.

The writer wishes to reiterate his expression of gratitude to all who have helped him while he was working on the present paper, but most particularly to Professors Schiffer, Bargmann, von Neumann, Tukey, and Bochner. They have helped him not only with the proofs where their names are mentioned in the text, but also in other ways, in particular with references to the existing literature.

Appendix

It has been suggested that a more direct proof would be desirable for Theorem IV.6. Instead, the following, somewhat more general theorem will be proved.

Let us consider the unit cube in K dimensional space, i.e. the points λ, all the coordinates λ_k of which are between 0 and 1

$$0 \leqq \lambda_k < 1. \qquad (k = 1, \cdots, K) \tag{57}$$

Two points λ and μ will be said to be in the neighborhood ε of each other if all the K numbers $\lambda_k - \mu_k$ are closer than ε either to 0 or to 1. Geometrically, the points of the neighborhood ε of a point λ form a cube of edge length 2ε at the center of which λ is located. In order to use this language one must consider the underlying space periodic, with a period 1 in the direction of every coordinate axis, so that a "cube" may consist of unconnected regions in the

unit cube: one part of the "cube" may adjoin one of the boundaries of the unit cube, the other the opposite boundary.

The product of a point λ with a number is defined in a similar fashion: the k coordinate of $n\lambda$ is $n\lambda_k$ if $n\lambda_k < 1$. Otherwise, one has to subtract from $n\lambda_k$ an integer so that the result be smaller than 1 (i.e. the k coordinate of $n\lambda$ is the fractional part of $n\lambda_k$). We shall consider the infinite set of points $0, \lambda, 2\lambda, 3\lambda, \cdots$, which we shall call the set $n\lambda$, and prove for these the following theorem.

Given any point λ and any positive number ε, there is an integer $N(\lambda, \varepsilon)$ such that no point of the infinite set $n\lambda$ is further than ε from any N consecutive points $(n_0 + 1)\lambda, (n_0 + 2)\lambda, \cdots, (n_0 + N)\lambda$ whatever the value of the integer n_0 is.

It is easy to see, first of all, that all points of the set are condensation points (Kronecker). In order to prove then the above theorem, let us cover the above unit cube with a finite number of such domains that a) all points of a domain are in the neighborhood $\varepsilon/3$ of each other and b) that every point of the unit cube be inner point of at least one domain. This can be done, e. g. by choosing an integer $l > 3/\varepsilon$ and considering, as the above domains, the points which are in the neighborhood $\varepsilon/6$, of a point the coordinates of which are integer multiples of l^{-1}. This way the unit cube is covered with l^K neighborhoods. Next, we select those domains which contain in their interior at least one point of the set $n\lambda$ and call these domains selected domains.

We now choose an N_1 so large that there be, among the $0, \lambda, 2\lambda, \cdots, N_1\lambda$ at least one point in each selected domain. This is possible since there is only a finite number of selected domains and since each of these has a condensation point of the $n\lambda$ set. Consider then a number $N_2 > N_1$ for which $N_2\lambda$ is in the neighborhood $\varepsilon/6$ of 0. This is again possible since 0 is either a point $\nu\lambda$ of the set $n\lambda$, in which case $\nu\lambda, 2\nu\lambda, \cdots$ are all 0, or if no $n\lambda$ is 0, there is a finite neighborhood of 0 so that none of the first N_1 points $n\lambda$ falls into that neighborhood. Since, however, there are points $n\lambda$ in every neighborhood of 0, there must be a point $N_2\lambda$ with $N_2 > N_1$ which is closer to 0 than any of the first N_1 of the $n\lambda$. One of these, however, is in the neighborhood $\varepsilon/6$ of 0 so that $N_2\lambda$ is in an even closer neighborhood of it.

We can now choose $N = N_1 + N_2$. In fact, there surely is among the points $n_0\lambda, (n_0 + 1)\lambda, \cdots, (n_0 + N_2)\lambda$ one which is in the neighborhood $\varepsilon/2$ of 0, no matter what n_0 is: since $n_0\lambda$ is in a selected domain, there is an n_1 with $n_1 \leqq N_1$ which is in the neighborhood $\varepsilon/3$ of it. Then, since $N_2\lambda$ is in the neighborhood $\varepsilon/6$ of 0, the point $(N_2 - n_1)\lambda + n_0\lambda = (n_0\lambda - n_1\lambda) + N_2\lambda$ will be in the neighborhood $\varepsilon/2$ of 0. Furthermore, since for every point of the set $n\lambda$, there is one, say $n_3\lambda$ with $n_3 \leqq N_1$ which is in the neighborhood $\varepsilon/3$ of it, n_3 further steps will bring one from the $\varepsilon/2$ neighborhood of 0 to the $5\ \varepsilon/6$ neighborhood of that point. This then proves the theorem.

PRINCETON UNIVERSITY

On the Connection Between the Distribution of Poles and Residues for an R Function and Its Invariant Derivative

E. P. Wigner

Annals of Mathematics 55, 7–18 (1952)

(Received March 5, 1951)

I. Introduction

The present paper is the continuation of an earlier article.[1] It was proposed, in that article, to denote as R functions such meromorphic functions the imaginary part of which has the same sign as the imaginary part of the variable z. It has been shown then (II.8) that these functions permit an expansion (the first normal form of R)

$$R(z) = \alpha z + \beta + \sum_\mu \left(\frac{\gamma_\mu^2}{Z_\mu - z} - \frac{\gamma_\mu^2}{Z_\mu} \right) \tag{1}$$

which converges absolutely (except if z coincides with one of the poles Z_μ). The constants $\alpha, \beta, \gamma_\mu, Z_\mu$, are all real in (1), $\alpha \geqq 0$. For a special class of R functions, the uniform R functions, the linear term αz vanishes and the rest of the series can be written as (equations (33), (33a))

$$R(z) = \beta' + \sum_\mu \frac{1}{g_\mu (Z_\mu - z)}; \qquad g_\mu = \gamma_\mu^{-2}. \tag{1a}$$

The β' of (1a) is not necessarily equal to the β of (1); the g_μ are real and positive. The series (1a) is to be arranged according to increasing values of the absolute value $| Z_\mu |$ and will converge then. Furthermore, it is equal to the series which one obtains by arranging the terms according to increasing values of $| Z_\mu - a |$ where a is any real number.

Every R function can also be written as the tangent of a real analytic function $f(z)$ (real for real z)

$$R(z) = \operatorname{tg} f(z) \tag{2a}$$

where the derivative of f

$$f' = \frac{df(z)}{dz} = \frac{R'}{1 + R^2} \tag{2b}$$

also permits an expansion into partial fractions

$$f'(z) = \eta + \sum_\mu \frac{\kappa_\mu}{(z - x_\mu)^2 + \kappa_\mu^2} \tag{2}$$

with real η, κ_μ, x_μ and $\eta \geqq 0$. The formulae (2), (2a), (2b) constitute the second normal form (II 12) of R; the function f' is called its invariant derivative. One

[1] E. P. Wigner, *On a class of analytic functions from the quantum theory of collisions.* Ann. of Math. **53** (1951), pp. 36–67.

sees that the poles of R are all on the real axis and have negative residues. The poles $x_\mu \pm i\kappa_\mu$ of f' are complex: those in the upper half plane have residues $-i/2$ (or $-in/2$ with integer n) those in the lower half plane $i/2$.

Finally, the aforementioned paper defines "statistical" R functions, i.e., such R functions for which the distance of subsequent poles Z_μ and the magnitude of the residues $-g_\mu^{-1}$ shows a statistical distribution. (The pole subsequent to Z_μ is the one which is nearest to the right of it on the real axis.) The statement about the distribution of the residues reads, in detail, as follows. The number of g_μ^{-1} which exceed a positive number r and which belong to poles of an interval L of length L was denoted by $s_L(r)$. This quantity depends, of course, on the position of the interval on the real axis. The existence of a distribution for the residues is then equivalent with the existence of a function $s(r)$ such that

$$\int_0^\infty | s(r) - s_L(r)/L | \, dr \tag{3}$$

approach zero with increasing L and that this approach be uniform with respect to the position of the interval L. A similar statement applies also to the distance Z of subsequent poles from each other which leads to the definition of a similar function $D(Z)$ for statistical R functions.

Statistical R functions (for the complete definition cf reference 1) are also uniform and permit the expansion (1a); they were called "normalized" if the average sum of the (negative) residues of the poles in a unit interval on the real axis was $1/\pi$ and if $\beta' = 0$ in (1a). They approach i as the imaginary part of z goes to infinity. This and the ensuing uniformity will hold also for the R functions which form the subject of the present investigation. However, it will turn out that the definition of the statistical nature would have to be somewhat weakened to apply to them.

One can easily verify by direct calculation that the orthogonal fractional transform

$$R_\phi = \frac{\sin\phi + R\cos\phi}{\cos\phi - R\sin\phi} \qquad (\phi \text{ real}) \tag{4}$$

of a uniform normalized R function is again an R function of this nature and convince oneself by direct calculation that the invariant derivative (2b) of R_ϕ is identical with the invariant derivative of R. It was not proved but made plausible in the earlier article that, "apart from rare exceptions", the statistical distributions, for level spacing and for the magnitude of residues, are the same for a normalized statistical R function and all its orthogonal transforms, i.e., that the functions $s(r)$ and $D(Z)$ are the same for R and its orthogonal fractional transforms. It is plausible, finally, that if these distributions are given, and suitable correlation functions for the statistics of two and more subsequent pole distances and residues defined, these determine, in their turn, the statistical distribution of the κ_μ, the distances of subsequent x_μ, and the statistical correlations of these quantities. Similarly, it is plausible that, conversely, the distribution functions

for κ_μ and for the x_μ differences, together with the statistical correlations between these quantities, determine the distribution functions for the g_μ and the Z_μ differences and their statistical correlations.

The present article will deal with an example for the latter problem: it will give, explicitly, the distribution of the g_μ and of the differences of subsequent Z_μ for the case that the κ_μ have an arbitrary distribution, the corresponding x_μ are distributed entirely at random and $\eta > 0$. The probability for the occurrence of an x_μ, κ_μ pair with $x < x_\mu < x + \delta x$, $\kappa < \kappa_\mu < \kappa + \delta\kappa$ will be denoted by $K(\kappa)\,\delta x\,\delta\kappa$. It will be assumed that this probability is independent of the presence or absence of further poles in the neighborhood of $x + i\kappa$, i.e., that there are no statistical correlations between the positions of the poles of f'. The average number of poles of f' in the upper half plane and within a strip of unit width parallel to the imaginary axis

$$\int_0^\infty K(\kappa)\,d\kappa = \sigma \tag{5}$$

will be denoted by σ.

As was mentioned before, it will be explicitly assumed that the distribution function $s(r)$ for the residues of the poles of R is invariant under orthogonal fractional transformations. Actually, we shall use, instead of $s(r)$ the quantity $G(g)$. This is so defined that $G(g)\,\delta g$ is the average number of the g_μ which belong to the poles Z_μ within unit interval on the real axis and which are in the interval $g < g_\mu < g + \delta g$.

Clearly

$$\int_0^\infty G(g)\,dg = \rho \tag{6}$$

where ρ is the density of the poles of R. The connection between the functions s and G is given by

$$s(r) = \int_0^{r^{-1}} G(g)\,dg. \tag{7}$$

The reason for using the parameters $g = \gamma^{-2}$ in (1a) is that most of the subsequent formulae are simpler when expressed in terms of the g than in terms of the γ. The strength s, which is in our case $1/\pi$, is

$$\int_0^\infty G(g)\,dg/g = s. \tag{8}$$

The corresponding equation in terms of $s(r)$ read $\int s(r)\,dr = s$.

The number of distances, between the successive poles of a very long interval, which exceed Z, divided by the length of that interval, will again be denoted by $D\,(Z)$. Clearly $D\,(\infty) = 0$,

$$D(0) = \rho \tag{9}$$

and the probability that the distance of an arbitrary pole from its right hand neighbor (or from its left hand neighbor) be between Z and $Z + \delta Z$ is $-D'(Z)\delta Z/\rho$. The function D is monotonically decreasing.

Before we go over to the actual calculations, we wish to add three remarks to the considerations of the earlier paper.

I.1. *The invariant derivative f' of a not normalized ($s \neq \pi^{-1}$) uniform R function goes (uniformly in x) to zero if the imaginary part of its argument $z = x + iy$ approaches infinity.* It follows from III.2 and III.4 that the denominator $1 + R^2$ of f' converges uniformly to $1 - \pi^2 s^2$ which is different from zero. On the other hand, the absolute value of its numerator R'

$$(10) \qquad |R'| = \left| \sum \frac{1}{g_\mu (Z_\mu - z)^2} \right| < \sum \frac{1}{g_\mu} \frac{1}{(Z_\mu - x)^2 + y^2}.$$

This last expression, however, is just y^{-1} times the imaginary part of R. Since the latter converges uniformly to $i\pi s$, the R', and hence also f', converges uniformly to zero. It follows that a uniform R function is surely normalized if its invariant derivative $f'(x + iy)$ remains finite as $y \to \infty$.

An immediate consequence of I.1 is I.2a. *If R is uniform but not normalized, the κ_μ in the expansion* (2) *of its invariant derivative have a finite upper bound.* This follows from the fact that f' has a pole at $x_\mu + i\,\kappa_\mu$. It further follows from this lemma that the x_μ have no finite condensation point since, f' being meromorphic, its poles have none.

Let us consider now the sum in (2) for a uniform not normalized R. The general term of this has, for $z = iy$, the form

$$(11) \qquad \frac{\kappa_\mu(\kappa_\mu^2 + x_\mu^2 - y^2 + 2iyx_\mu)}{\kappa_\mu^4 + 2\kappa_\mu^2(x_\mu^2 - y^2) + (x_\mu^2 + y^2)^2}.$$

If we choose y so large that $y > 3\kappa_\mu$ for every μ, the last term of the denominator will be more than nine times greater than $\kappa_\mu^2 y^2$. Hence adding $4\kappa_\mu^2 y^2$ to the denominator will not decrease (11) by more than a factor of 2. The absolute value of the real part of the resulting expression is, therefore, less than

$$(11a) \qquad \frac{\kappa_\mu(\kappa_\mu^2 + x_\mu^2 + y^2)}{\kappa_\mu^4 + 2\kappa_\mu^2(x_\mu^2 + y^2) + (x_\mu^2 + y^2)^2} = \frac{\kappa_\mu}{\kappa_\mu^2 + x_\mu^2 + y^2}.$$

Since, however, the sum over μ of the expressions (11a) converges for $y = 0$ (it is the sum in (2) for $z = 0$), it goes to zero as y tends to infinity. The same is then true of the real part of (11) and a similar argument shows that it is true also of the imaginary part of (11). The same proof applies if z has the form $x + iy$, except that $x_\mu - x$ occurs wherever x_μ stood in the preceding expressions. One thus has

I.2. *The infinite sum in the expansion of* (2) *of the invariant derivative of a uniform but not normalized R function tends to zero (uniformly in x) as the imaginary part y of its argument $z = x + iy$ tends to infinity.*

It now follows from I.1 that

I.3. *The constant η in the expansion* (2) *of the invariant derivative of a uniform but not normalized R function vanishes; if $\eta > 0$, the corresponding R function is normalized.*

It appears from the above that the "invariant derivative" of a not normalized R function has simpler properties than the invariant derivative of a normalized R function. Even though this may be correct, one must remember that only if R is normalized can its essential properties be invariant under orthogonal fractional transformations (4): otherwise, not even the strength of R is invariant.

It is a little surprising that the equation $R(x + i\infty) = i$ follows from the inequality $\eta > 0$ while the inequality $R(x + i\infty) \neq i$ entails the equality $\eta = 0$. The "explanation" of this unusual behavior is that the poles of f' spread further and further away from the real axis as R becomes more nearly normalized. The η in (2) is, then, a limiting case of η/π poles with infinite imaginary parts per unit abscissa. It may be worth while to remark also that $\eta > 0$ is only a sufficient but not a necessary condition for the normalized nature of R. It is not difficult to give examples of uniform normalized R, the invariant derivative of which has no constant term η.

II. Calculation of the Distribution Function G for the g_μ

The theorem which we wish to show in this section follows intuitively from the construction illustrated in Figure 5 of the previous paper.[1] The probability that a pole Z_i of R shall fall into an interval between z and $z + \Delta$ is proportional to the ordinate $f'(z)$ at z. If it falls there, there will be a pole in that interval the g of which is $f'(z')$ with $z < z' < z + \Delta$. Hence the probability $G(g)\,\delta g$ that g_μ be between g and $g + \delta g$ is proportional to the total length of the domains in which $f'(z)$ is between g and $g + \delta g$, multiplied by the ordinate $f' = g$ in these domains.

Hence, if we denote with $LP(g)\,\delta g$ the total measure of the domains which are within a very long interval L and in which $g < f'(z) < g + \delta g$ we can expect this distribution function to be related to the distribution function $G(g)$ for the g_μ of the corresponding R function by

$$G(g) = (g/\pi)P(g). \tag{12}$$

The proportionality constant in (12) follows from (8) with $s = 1/\pi$ and the obvious equation

$$\int_0^\infty P(g)\,dg = 1. \tag{12a}$$

An exact formulation of this theorem is as follows:

II.1. *If all the fractional orthogonal transforms R_ϕ of R by* (4) *have the same distribution function $G(g)$ for the $g_\mu = \gamma_\mu^{-2}$ in their expansion by* (1) *or* (1a) *and if all the poles of their invariant derivative f' by* (2) *are further than $\kappa_{\min}$ from the real axis, i.e. if all $\kappa_\mu > \kappa_{\min}$, then* (12) *holds between the distribution functions G and P.*

In order to prove this theorem, we first note that for real z

$$(13)\qquad |f''(z)| < \sum_{\mu} \left| \frac{2\kappa_\mu(x_\mu - z)}{((z - x_\mu)^2 + \kappa_\mu^2)^2} \right| < \sum \frac{1}{(z - x_\mu)^2 + \kappa_\mu^2} < \frac{f'(z)}{\kappa_{\min}}$$

and

$$(13a)\qquad |f'''(z)| < \sum_{\mu} \left| \frac{6\kappa_\mu(z - x_\mu)^2 - 2\kappa_\mu^3}{((z - x_\mu)^2 + \kappa_\mu^2)^3} \right| < \sum \frac{6\kappa_\mu}{((z - x_\mu)^2 + \kappa_\mu^2)^2} < \frac{6f'(z)}{\kappa_{\min}^2}.$$

From these inequalities we can obtain an estimate for the minimum length of the intervals which constitute the domain D of z in which $g_1 < f'(z) < g_2$. We

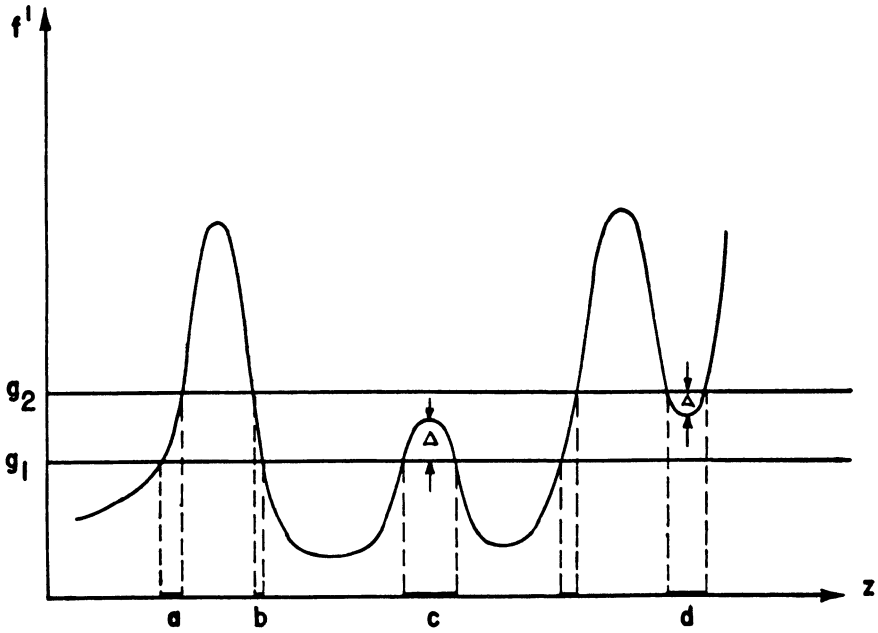


FIG. 1

note that $f'(z)$ can pass through the (g_1, g_2) interval in the four different ways which are illustrated in Figure 1. If f' enters and leaves the (g_1, g_2) interval at different sides (cases a and b) its sojourn in the interval has the length

$$(14)\qquad l = \frac{g_2 - g_1}{f''(z')} > \frac{(g_2 - g_1)\kappa_{\min}}{g_2}$$

where z' is a point in D. If f' enters and leaves the (g_1, g_2) interval at the same side (cases c and d of Figure 1), the length of sojourn of f' depends on how deeply it dips into the interval. If f' dips from above into the interval and leaves it on the same side, the maximum of $g_2 - f'$ shall be denoted by Δ, otherwise Δ shall be the maximum of $f' - g_1$. In both cases, the length of sojourn l is clearly greater than $(8\Delta/f''')^{\frac{1}{2}}$ where f''' is the highest value the second derivative of f' assumes in D. Hence, by (13a), we have

$$(14a)\qquad l > (8\Delta/f''')^{\frac{1}{2}} > (\Delta/g_2)^{\frac{1}{2}}\kappa_{\min}.$$

Let us now denote the primitive function of P by P_1. Then, if L is a sufficiently long interval the measure of the domain $D \times L$ (which is the part of L where where $g_1 < f'(z) < g_2$) will be between the two members $L(P_1(g_2) - P_1(g_1))$ $(1 \pm \varepsilon)$. Let us find, further, a $\Delta < \varepsilon(g_2 - g_1)/2$ such that the measure of the domain in a sufficiently long L, in which either $g_1 - \Delta < f' < g_1 + \Delta$ or $g_2 - \Delta < f' < g_2 + \Delta$, is smaller than $\varepsilon L(P_1(g_2) - P_1(g_1))$.Let us then calculate the right side of (14a) with this Δ and call l henceforth the smaller of the two quantities (14), (14a). Let us finally choose an integer n such that $\pi/n < (g_1 - \Delta)l(1 - \varepsilon)\varepsilon$. Starting with an arbitrary point z_1, we now find, successively, the points z_2, z_3, such that the area under the f' curve which is between z_i and z_{i+1} be always π/n.

We denote by D_2 the domain which contains those intervals of z for which $g_1 - \Delta < f'(z) < g_2 + \Delta$, except if f' does not dip at least Δ into the $(g_1 - \Delta, g_2 + \Delta)$ interval. Each interval of D_2 contains at least one point of the domain

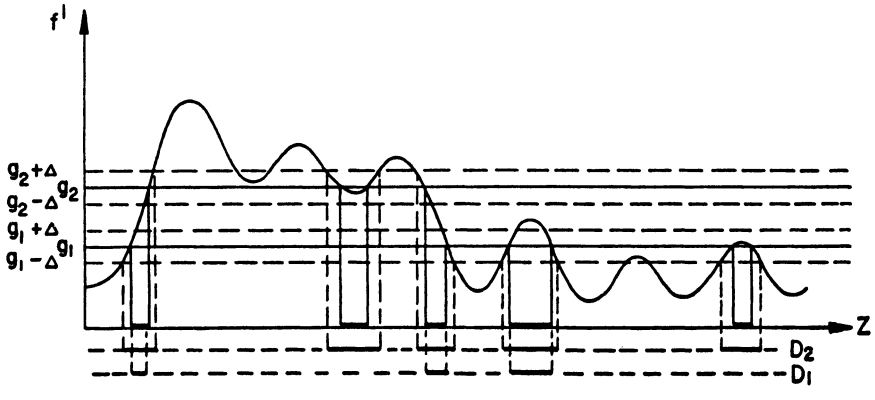


FIG. 2

D in which $g_1 < f'(z) < g_2$, and each interval of D_2 has at least the length $l(1 - \varepsilon)$. Hence, the area under f' and over each of these intervals is at least $l(1 - \varepsilon)(g_1 - \Delta)$ and contains, therefore, more $l(1 - \varepsilon)(g_1 - \Delta)/(\pi/n) - 1 > (1/\varepsilon) - 1$ successive points z_i. In Figure 2, the intervals of D are heavily marked on the z axis, those of D_2 (and of D_1) are indicated below it.

This second domain, D_1, contains all the intervals of the domain D except those in which $f'(z)$ does not reach into the $(g_1 + \Delta, g_2 - \Delta)$ interval. Again, each interval of D_1 is longer than $l(1 - \varepsilon)$ and contains, therefore, more than $(1/\varepsilon) - 1$ successive points z_i.

The domain D_2 is contained in the domain of all the z for which $g_1 - \Delta < f'(z) < g_2 + \Delta$. Hence the measure of its part $D_2 \times L$ contained in L is smaller than $L(P_1(g_2) - P_1(g_1))$ $(1 + \varepsilon + \varepsilon)$. On the other hand, D_2 contains D. On the contrary, D_1 is contained in D but contains the domain of z for which $g_1 + \Delta < f'(z) < g_2 - \Delta$. The measure of the intersection $D_1 \times L$ of D_1 and L is, therefore, larger than $L(P_1(g_2) - P_1(g_1))(1 - \varepsilon - \varepsilon)$.

The areas under f' between any pair of points z_i and z_{i+n} will be just π. As

a result, $f'(z_1)$, $f'(z_{n+1})$, $f'(z_{2n+1})$, $\cdots$, will be successive g_μ values for the R_ϕ which has a pole at z_1. Similarly, $f'(z_i)$, $f'(z_{i+n})$, $f'(z_{i+2n})$, $\cdots$, are the successive g_μ values of the R_ϕ which has a pole at z_i(cf. IV.2). The R_ϕ which have poles at $z_1, z_2, z_3, \cdots, z_n$, respectively, shall be called, for brevity, $R_1, R_2, \cdots, R_n$.

We now find an interval Λ which is, firstly, longer than L so that the preceding considerations apply to it. Second, the interval Λ shall be long enough so that the number of g_μ of the poles of this interval, which are between g_1 and g_2, shall approximate by a factor between $1 \pm \varepsilon$, the expression $\Lambda(G_1(g_2) - G_1(g_2))$ where G_1 is the primitive function of G. It will be possible to accomplish this simultaneously for all the n functions $R_1, R_2, \cdots, R_n$. This completes the preparations for the proof.

Let us denote the number of z_i which are in a domain with the starred symbol of that domain. Then $(D \times \Lambda)^*$ will be the number of those poles of the functions $R_1, R_2, \cdots, R_n$ which are in Λ and the g_μ values of which are between g_1 and g_2. Hence, according to the preceding paragraph,

$$n\Lambda(G_1(g_2) - G_1(g_1))(1 - \varepsilon) < (D \times \Lambda)^* < n\Lambda(G_1(g_2) - G_1(g_1))(1 + \varepsilon). \tag{15}$$

We shall now estimate the area under f' which is over the domain $D_2 \times \Lambda$. Since, as has been pointed out before, the measure of the domain D_2 is smaller than $\Lambda(P_1(g_2) - P_1(g_1))(1 + 2\varepsilon)$, the area in question is smaller than

$$\Lambda(P_1(g_2) - P_1(g_1))(1 + 2\varepsilon)(g_2 + \Delta) < \Lambda(P_1(g_2) - P_1(g_1))g_2(1 + \varepsilon)^3.$$

On the other hand, since D_2 consists of intervals each with at least $(1/\varepsilon) - 1$ successive points z_i and since the area associated with each such point is π/n, the area in question is surely larger than $(D_2 \times \Lambda)^*(\pi/n)((1/\varepsilon) - 2)/((1/\varepsilon) - 1)$ and, since D_2 contains D, it is also larger than $(D \times \Lambda)^*(\pi/n)(1 + \varepsilon)^{-1}$. Hence,

$$(D \times \Lambda)^* < \Lambda(P_1(g_2) - P_1(g_1))g_2(n/\pi)(1 + \varepsilon)^4. \tag{16a}$$

A similar consideration involving D_1 instead of D_2 yields

$$(D \times \Lambda)^* > \Lambda(P_1(g_2) - P_1(g_1))g_1(n/\pi)(1 - \varepsilon)^4. \tag{16b}$$

The combination of (16a), (16b) with (15) gives

$$\begin{aligned} \frac{P_1(g_2) - P_1(g_1)}{\pi} g_1 \frac{(1 - \varepsilon)^4}{1 + \varepsilon} &< G_1(g_2) - G_1(g_1) \\ &< \frac{P_1(g_2) - P_1(g_1)}{\pi} g_2 \frac{(1 + \varepsilon)^4}{1 - \varepsilon}. \end{aligned} \tag{17}$$

Since this inequality is valid for any ε it follows that:

$$(g_1/\pi) \int_{g_1}^{g_2} P(g)\, dg \leqq \int_{g_1}^{g_2} G(g)\, dg \leqq (g_2/\pi) \int_{g_1}^{g_2} P(g)\, dg \tag{18}$$

(the P_1 and G_1 are the primitive functions of P and G) and since (18) is valid for any g_1, g_2 we finally have obtained (12). It may be worth while to remark that the proof did not presuppose the uniformity of the convergence of an in-

tegral of the type (3) with respect to the location of the interval L. It did presuppose, however, the existence of an invariant distribution function G for the g_μ. It appears natural to assume this for the R functions to be considered in the next section while the condition about the uniformity of the convergence of (3) cannot be expected to apply to them.

Before going over to the calculation of $P(g)$ in the case of a random distribution of the poles f' we note that the average of f' can be calculated from the series (2) if one assumes that the number of poles with very large κ is not too large. One obtains for

$$\bar{f}' = \int_0^\infty gP(g)\,dg = \eta + \pi\sigma \tag{19}$$

where σ is the average number of the pairs of poles of f' per unit interval on the real axis. It then follows from (12) and (6) that

$$\rho = \sigma + \eta/\pi. \tag{19a}$$

This equation again shows that, in a sense, the term η in (2) corresponds to η/π pairs of poles of f' per unit interval on the real axis.

III. Calculation of the Distribution Function G in Case of Random Distribution of the Poles f'

As outlined in the Introduction, we mean by random distribution of the poles that the probabilities for the positions of the poles in the upper half plane be statistically independent and that their real parts be entirely at random. (If f' has a pole at $x + i\kappa$ it also has a pole at $x - i\kappa$). The probability that there be a pole $x_\mu + i\kappa_\mu$ with $x < x_\mu < x + \delta$; $0 < \kappa < \kappa_\mu < \kappa + \delta$ was denoted by $K(\kappa)\,\delta x\,\delta\kappa$ and these probabilities are independent. As a result, (A4) of the Appendix can be applied directly to calculate the probability distribution of the sum in (2). If this sum is between $g - \eta$ and $g - \eta + \delta$, the quantity f' will be between g and $g + \delta$. Hence we have:

$$P(g) = (2\pi)^{-1}\int_{-\infty}^{\infty} d\nu \exp\left\{i\nu(g-\eta) + \int_0^\infty d\kappa \int_{-\infty}^{\infty} dx K(\kappa)(e^{-i\nu\kappa/(\kappa^2+x^2)} - 1)\right\}. \tag{20}$$

The point τ of the manifold M of the Appendix is characterized by the variables κ, x; the function $F(\tau)$ is $\kappa/(\kappa^2 + x^2)$. The last integral in (20) can be evaluated to give (20) the form

$$\begin{aligned} P(g) &= (2\pi)^{-1} \\ &\cdot \int_{-\infty}^{\infty} d\nu \exp\left\{i\nu(g-\eta) - i\pi\nu \int_0^\infty K(\kappa)e^{-i\nu/2\kappa}\left(J_0\left(\frac{\nu}{2\kappa}\right) + iJ_1\left(\frac{\nu}{2\kappa}\right)\right)d\kappa\right\}. \end{aligned} \tag{20a}$$

Hence (12) gives for the fraction of the g_μ which are in unit interval at g

$$\begin{aligned} G(g) &= \frac{g}{2\pi^2} \\ &\cdot \int_{-\infty}^{\infty} d\nu \exp\left\{i\nu(g-\eta) - \tfrac{1}{2}i\pi\nu^2 \int_0^\infty K(\nu/2k)e^{-ik}(J_0(k) + iJ_1(k))\,dk/k^2\right\}. \end{aligned} \tag{21}$$

One verifies easily that (21) vanishes for $g < \eta$ and that the integral of $G(g)/g$, which is π^{-1} times the integrand for $\nu = 0$, is equal to $1/\pi$. This confirms (8) with $s = 1/\pi$. In order to verify (6), one can calculate the integral of $G(g)$ from $-\infty$ to ∞. Since negative g do not contribute to this integral it is equal to the integral (6) from 0 to ∞. Carrying out the calculation one finds, on the right side by means of (5), the right side of (19a) instead of ρ.

IV. Calculation of the Distribution Function D

An argument, entirely analogous to that given in Section II, leads to IV.1. *Denote with d_Z the domain of x for which*

$$\int_x^{x+Z} f'(x)\, dx < \pi. \tag{22a}$$

Then, under the conditions of II.1 *for f' and R, the integral*

$$(\pi L)^{-1} \int_{d_Z \times L} f'(x)\, dx \to D(Z) \tag{22}$$

extended over the intersection of d_Z and an interval of length L, converges to $D(Z)$ as the length L of L grows beyond all limits. The formulation II.1 which corresponds to the above formulation of IV.1 would replace (12) by the equation

$$G_1(g) \equiv \int_g^\infty G(g)\, dg = (\pi L)^{-1} \int_{D_g \times L} f'(x)\, dx \tag{23}$$

in which D_g is the domain of x in which

$$f'(x) > g. \tag{23a}$$

The proof of IV.1 will not be given because of its great similarity with that of II.1 and because of its lengthiness.

It is again possible to calculate $D(Z)$ explicitly in case of random distribution of the poles of f', as defined in the preceding section. For this purpose, we first consider the simultaneous probability $P(g, I)\, dg\, dI$ that

$$\begin{gathered} g < f'(x) < g + dg \\ I < \int_x^{x+Z} f'(x)\, dx < I + dI. \end{gathered} \tag{24a}$$

This last expression is, in terms of η, the x_μ and κ_μ

$$\int_x^{x+Z} f'(x)\, dx = \eta Z + \sum_\mu \operatorname{arctg} (Z - x_\mu)/\kappa_\mu + \operatorname{arctg} x_\mu/\kappa_\mu \tag{25}$$

Z is to be considered, for the time being, as a fixed parameter. According to (A5), $P(g, \mathrm{I})$ is

$$P(g, I) = \frac{1}{(2\pi)^2} \iint_{-\infty}^{\infty} d\nu\, d\lambda \exp \{i\nu(g - \eta) + i\lambda(I - \eta Z) + \int_{-\infty}^{\infty} dx \int_0^{\infty} d\kappa K(\kappa)\, [e^{-i\nu\kappa/(\kappa^2+x^2) - i\lambda \operatorname{arctg}(Z-x)/\kappa - i\lambda \operatorname{arctg} x/\kappa} - 1]\}. \tag{24}$$

Naturally, (24) gives $P(g)$ if we integrate it over all values of I.

We can express the integral in (22) as an integral over $f'(x) = g$. The domain of x in which $f'(x)$ is between g and $g + \delta g$ and which contributes to (22) is the fraction of L in which (22a) is less than π and $g < f'(x) < g + \delta g$. This fraction is just the integral of (24), extended from 0 to π over I, and extended from g to $g + \delta g$ over g. Hence

$$D(Z) = \frac{1}{\pi} \int_0^{\infty} g\, dg \int_0^{\pi} P(g, I)\, dI. \tag{26}$$

The integration of (24) with respect to I can be carried out directly. The result is g times a function of g which has the form of a Fourier integral. The integral of this product is, therefore, $2\pi i$ times the derivative with respect to ν, of the Fourier transform of that function, at $\nu = 0$. Hence we have

$$D(Z) = \frac{1}{2\pi^2} \int_{-\infty}^{\infty} d\lambda \frac{e^{i\pi\lambda} - 1}{i\lambda} \left\{\eta + \int_{-\infty}^{\infty} dx \int_0^{\infty} d\kappa K(\kappa) \frac{\kappa}{\kappa^2 + x^2} \cdot \exp\left(i\lambda \operatorname{arctg} \frac{x - Z}{\kappa} - i\lambda \operatorname{arctg} \frac{x}{\kappa}\right)\right\} \times \exp\left\{-\lambda\eta Z + \int_{-\infty}^{\infty} dx \int_0^{\infty} d\kappa K(\kappa)[e^{i\lambda \operatorname{arctg}(x-Z)/\kappa - i\lambda \operatorname{arctg}(x/\kappa)} - 1]\right\}. \tag{27}$$

In this case, the writer has not succeeded in carrying out the integrations with respect to x.

As a check on (27), one can verify (9). When doing this, one must realize that the right side of (27) is discontinuous at $Z = 0$ and that $D(0)$ is the value of $D(Z)$ for very small Z. Carrying out the calculation, one obtains again the right side of (19a), instead of ρ, for $D(0)$.

Appendix

Let us consider a one or more dimensional manifold M (in the text, this manifold is the upper half of the complex plane; i.e., two dimensional) in which a volume is defined. The points of the manifold shall be denoted by τ, the volume element by $d\tau$. There are points distributed in the manifold according to a statistical law: the probability of a point being at τ within $d\tau$ is denoted by $T(\tau)\, d\tau$ and these probabilities shall be independent of each other. Let us consider then the sum

$$F = F(\tau_1) + F(\tau_2) + F(\tau_3) \cdots \tag{A1}$$

where F is a continuous function of τ and $\tau_1, \tau_2, \cdots$ are a possible arrangement

of the points in the manifold, their number finite or infinite. The question which we wish to answer is: if $T(\tau_1)T(\tau_2)T(\tau_3) \cdots d\tau_1, d\tau_2, d\tau_3, \cdots$ is the probability that the configuration $\tau_1, \tau_2, \tau_3, \cdots$ be realized within the accuracy $d\tau_i$ for τ_i, what is the probability $P(F)\,dF$ that the sum (A1) assumes a value between F and $F + dF$?

We can rearrange the sum (A1) by considering the contributions of the different volume elements $d\tau$ rather than the different points 1, 2, 3, $\cdots$. The volume element $d\tau$ at τ gives with the probability $1 - T(\tau)\,d\tau$ no contribution, with the probability $T(\tau)\,d\tau$ the contribution $F(\tau)$. The distribution function $P(F)$ for the whole F is the convolution of all these distribution functions, one for each volume element. Its Fourier transform is the infinite product of the Fourier transforms of the distribution functions for all the infinitesimal domains into which the whole manifold can be decomposed. Introducing the variable ν for the domain $d\tau$ at τ, the Fourier transform $g_\tau(\nu)$ in question is:

$$g_\tau(\nu) = 1 + T(\tau)\,d\tau(e^{-i\nu F(\tau)} - 1). \tag{A2}$$

The product of all these $g(\nu)$ is

$$g(\nu) = \exp\left\{\int_M T(\tau)(e^{-i\nu F(\tau)} - 1)\,d\tau\right\} \tag{A3}$$

where the integral in the exponent has to be extended over the whole manifold M. The distribution function $P(F)$ is, therefore:

$$\begin{aligned} P(F) &= (1/2\pi)\int_{-\infty}^{\infty} d\nu e^{i\nu F} \exp\left\{\int_M T(\tau)(e^{-i\nu F(\tau)} - 1)\,d\tau\right\} \\ &= (1/2\pi)\int_{-\infty}^{\infty} d\nu \exp\left\{i\nu F + \int_M T(\tau)(e^{-i\nu F(\tau)} - 1)\,d\tau\right\}. \end{aligned} \tag{A4}$$

It is clear that one can calculate, in similar fashion, also the simultaneous probability $P(F, G)\,dF\,dG$ for two sums of the nature (A1) to assume specified values. One obtains

$$\begin{aligned} P(F, G) = (2\pi)^{-2} &\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} d\nu\,d\lambda \\ &\cdot \exp\left\{i\nu F + i\lambda G + \int_M T(\tau)(e^{-i\nu F(\tau) - i\lambda G(\tau)} - 1)\,d\tau\right\} \end{aligned} \tag{A5}$$

and that this can be extended also to more than two variables.

The above is a somewhat simplified version of the arguments which are, perhaps, best known from the derivation[2] of the "Holtsmark distribution". It is not essentially new. It is for this reason that no attempt was made to make the derivation rigorous.

[2] J. Holtsmark, Ann. d. Physik **58**, 577 (1919), S. Chandrasekhar, Rev. Mod. Phys. **15**, 1 (1943), p. 70 ff.

Derivative Matrix and Scattering Matrix

Eugene P. Wigner

Revista Mexicana de Fisica *1*, 91–101 (1952)

I. Introduction

R. Kronig [1] was the first to point out, on the basis of his and Kramers' [2] work, that the principle of causality must entail some properties of the collision and scattering matrices. Starting from Kronig's suggestion, Schutzer and Tiomno [3] gave, for non relativistic particles, a derivation of the well known theorem [4] that the poles of the scattering function $S(k)$ lie either in the lower half plane or on the imaginary axis of k. Schutzer and Tiomno's work has been extended, since, by Toll and by Van Kampen [5] to the case of relativistic particles with zero rest mass which formed also the subject of Kronig's and Kramers' early considerations [2]. Results similar to these were obtained in the course of the last years also in communication engineering, following the pioneering work of Campbell, Zobel and Foster [6], and of Cauer [7], by Brune [8], by Fränz [9] and, particularly, by Richards [10]. It was possible to derive from the principle of causality also that $S(k)$ satisfies the further well known equations [4]

$$S(k)\,S(-k) = 1 \qquad S(k)\,(S(k^*))^* = 1 \tag{1}$$

and that if ρ is larger than the range $a_{\min}$ of the scattering forces

$$S(k)\,e^{2ik\rho} < \infty \quad (\text{for } \rho > a_{\min},\ \operatorname{Im} k > 0) \tag{2}$$

is uniformly bounded in the upper half plane of k.

The above theorems about the function $S(k)$ have been derived by Schutzer and Tiomno also from the properties of the $R(E)$ function. In the case of simple scattering with definite angular momentum, which is the only one with which Schutzer and Tiomno's paper as well as the present one deal, $R(E)$ is the ratio of the wave function to its radial derivative, at a point $r = a > a_{\min}$ outside the range of the scattering forces. Hence $R(E)$ depends not only on E but also on a. However, this quantity will be considered to be a constant in all that follows. If the wave function outside a has the form $A \sin kr + B \cos kr$, the connection between R and the scattering function becomes

$$S(k) = e^{-2ika}\,\frac{1 + ik\,R(k^2)}{1 - ik\,R(k^2)}\,. \tag{3}$$

$R(E)$ as function of $E = k^2$ is a single valued real function [11] of E

$$R(E^*) = (R(E))^* \tag{4}$$

which has, for finite E, no other singularities but poles [12] and its imaginary part is positive in the upper, negative in the lower half plane. It follows from this property that the poles of R are all at real E and have negative residues. A function $R(E)$ which has these properties will be called a permissible R. One can then formulate Schutzer and Tiomno's secondary result also by stating that an $S(k)$, derived by (3) from a permissible R, has all the properties given in the first paragraph.

The present note will first inquire whether, conversely, the properties of $S(k)$ enumerated above suffice to establish the properties of R. This can easily be seen as far as the single valued and real nature (4) of R are concerned but will be answered in the negative concerning the theorems about the poles and residues of R. Hence we shall inquire for necessary conditions which $S(k)$ must fulfill in order to be derivable, by (3), from a permissible R. This will lead to an alternation theorem for the poles of S which are on the imaginary axis.

II. $S(k)$ Which Do Not Lead to Permissible $R(E)$

A rather simple general expression which always satisfies the requirements (1) and (2) for scattering functions is

$$S(k) = \frac{f(ik)}{f(-ik)} e^{-2ibk} \tag{5}$$

in which

$$f(x^*) = f(x)^* \tag{5a}$$

is a real function, $b < a_{\min}$ and

$$\frac{f(-k_1 + ik_2)}{f(k_1 - ik_2)} < \infty \qquad \text{for } k_1 > 0 \tag{5b}$$

is bounded if $k_1 + ik_2$ is in the right half plane. These conditions are clearly all satisfied if $f(x)$ is a real polynomial of x, i.e., $f(ik)$ a real polynomial of ik. Furthermore, the poles of S will have the proper location if the roots of $f(x)$ are all either real or lie in the left half plane. Scattering functions of this form play an important role in communication engineering [6], [7]. Our procedure will be to calculate the R from (5) by means of the inversion of (3)

$$R = \frac{i}{k} \frac{1 - Se^{2iak}}{1 + Se^{2iak}} \qquad (a > a_{\min}) \tag{6}$$

and ascertain whether the R obtained in this way

$$R = \frac{k^{-1} \operatorname{tg} \alpha k + F(k)}{1 - F(k) k \operatorname{tg} \alpha k}; \tag{7}$$

where

$$\alpha = a - b > 0 \tag{7a}$$

$$F(k) = \frac{f(ik) - f(-ik)}{ik[f(ik) + f(-ik)]} \tag{7b}$$

give a permissible R function in the sense of the second paragraph of the introduction.

Clearly, $k^{-1}\operatorname{tg}\alpha k$, $k\operatorname{tg}\alpha k$ as well as $F(k)$ are real given functions of k. It follows that the R of (7) is a single valued real function of $k^2 = E$ which can have, furthermore, no essential singularities for finite E. This point was noted already by Schutzer and Tiomno [3]. The only further conditions which remain to be verified concern, therefore, the position of the poles of R and their residues. It easily follows from the general theorems on R functions [11] that these conditions would also be satisfied if F as well as EF were themselves permissible R functions. This, however, is not the case in general.

In order to find a simple case in which (7) is not permissible we note that this is a continuous function of α at $\alpha = 0$ in every finite region of k except possible where $F(k)$ has a singularity. However, if $F(k)$ should have a negative imaginary part for a k for which k^2 has a positive imaginary part, this will be true also for a sufficiently small finite α and (7) will represent, for such an α, a not permissible R. Hence (7) can represent a permissible R for all permissible S only if this is true also for the expression which one obtains by setting $\alpha = 0$ in (7), i. e. if it is true for the F of (7b). One sees that the permissible nature of F is a necessary condition for the permissible nature of R at all admissible a and b while the permissible nature of F and EF would be a sufficient condition.

The two simplest f which give an S satisfying all conditions but may give a not permissible F and, hence, a not permissible R, are

$$f(-ik) = -k^2 - 2ibk + c \tag{8a}$$

with either $b > 0$ or $b^2 > c$ and

$$f(-ik) = (-k^2 - 2ib_1k + c_1)(-k^2 - 2ib_2k + c_2) \tag{8b}$$

in which either $b_1 > 0$ or $b_1^2 > c_1$, and either $b_2 > 0$ or $b_2^2 > c_2$. If the b is positive, the sum of the roots of $f(x) = x^2 + 2bx + c$ is negative and they must lie in the left half plane if they are complex. If $b^2 > c$, both roots are real. The same considerations apply to (8b). It follows that the S obtain from (8a) or (8b) by means of (5) will satisfy all conditions of the first paragraph. Nevertheless, the

$$F(k) = \frac{2b}{c - k^2} = \frac{2b}{c - E} \tag{9a}$$

will have a negative residue at its pole $E = c$ only if b is positive. In the other case, in fact, F is "antipermissible", i. e. it has a negative imaginary part in the upper, a positive one in the lower half plane. Similarly, the F obtained from (8b) by means of (7b)

$$F = \frac{-2(b_1 + b_2)E + 2b_1c_2 + 2b_2c_1}{E^2 - (c_1 + c_2 + 4b_1b_2)E + c_1c_2} \tag{9b}$$

will have complex poles if $4c_1c_2 > (c_1 + c_2 + 4b_1b_2)^2$ which is not incompatible with the conditions enumerated above. One can put, for instance $c_1 = c_2 = -2$, $b_1 = b_2 = 1$. In this case, the imaginary part of F will assume arbitrarily large positive values at some points of the lower half plane and arbitrarily large negative values at the opposite points of the upper half plane. We can conclude from these examples that (1) and (2), together with the conditions for the location of the poles of S, do not guarantee the permissible nature of the R obtained from the S.

III. Alternation Theorem

Since the conditions on S which were enumerated in the first paragraph of the introduction do not necessarily lead to a permissible R, it is natural to ask whether the condition to yield a permissible R leads to simple additional properties of S. This will be found to be the case.

All the poles Z_ν of R are real and the zeros x_ν also lie on the real axis; there is one zero between any two successive poles [13]. Let $Z_0, Z_{-1}, Z_{-2}, \ldots$ be the negative poles in decreasing order, $Z_1, Z_2, Z_3, \ldots$ the positive poles in increasing order, the zero between Z_ν and $Z_{\nu+1}$ shall be denoted by x_ν

$$Z_\nu < x_\nu < Z_{\nu+1} \tag{10}$$

$$Z_0 \leq 0 < Z_1 \,. \tag{10a}$$

x_0 can be positive or negative. Then, every permissible R allows a product expansion [14]

$$R(E) = C\frac{E - x_0}{E - Z_0}\prod\frac{1 - E/x_\nu}{1 - E/Z_\nu}\prod\frac{1 - E/x_{-\nu}}{1 - E/Z_{-\nu}} \tag{11}$$

with positive real C. The two products $\prod$ can be finite or infinite.

It follows from (3) that the poles of S satisfy the equation

$$1 - ik\,R(k^2) = 1 + \kappa\,R(-\kappa^2) = 0 \tag{12}$$

where $\kappa = -ik$. With (11), this gives

$$\begin{aligned}(\kappa^2 + Z_0)\prod(1 + \kappa^2/Z_\nu)\prod(1 + \kappa^2/Z_{-\nu}) &= \\ = -C\kappa(\kappa^2 + x_0)\prod(1 + \kappa^2/x_\nu)\prod(1 + \kappa^2/x_{-\nu})\,. &\end{aligned} \tag{13}$$

The left side of this equation is shown as function of the real κ, schematically, in the upper part of Figure 1; the right side of (13) in the two lower diagrams. The first product $\prod$ is positive in both cases for all real κ. The zeros of the left side lie at $\kappa + \pm\sqrt{(-Z_{-\nu})} = \pm\zeta_\nu$ with $\nu = 0, 1, 2, \ldots$ and the whole left side

changes at every ζ_ν and $-\zeta_\nu$; it is negative at $\kappa = 0$. The zeros of the right side lie at $\kappa = 0$ and $\kappa = \pm\sqrt{(-x_{-\nu})} = \pm\xi_\nu$ with $\nu = 0, 1, 2$ if $x_0 < 0$ and $\nu = 1, 2, 3, \ldots$ if $x_0 > 0$. In the former case the right side is positive between 0 and ξ_0; in the latter case it is negative between 0 and ξ_1.

Let us take up the second case ($x_0 > 0$) first. It follows from (10) that

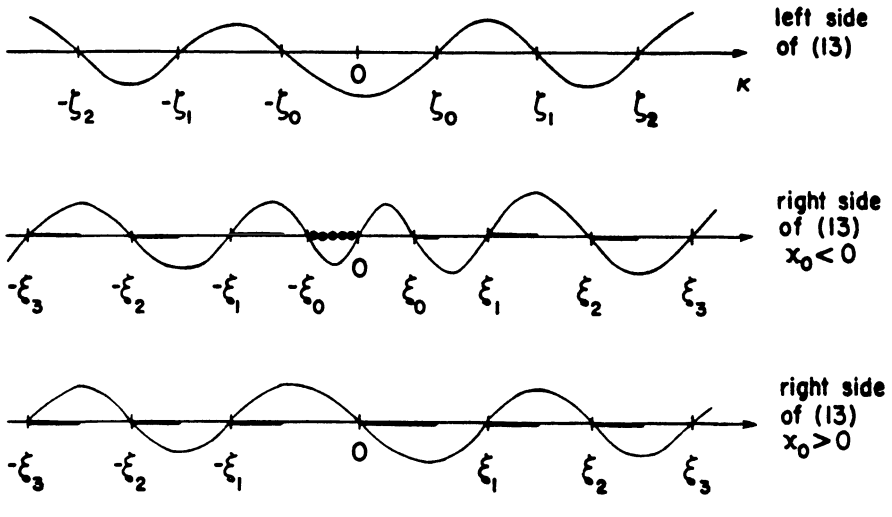
$$0 < \zeta_0 < \xi_1 < \zeta_1 < \xi_2 < \zeta_2 \ldots \tag{14}$$

Hence (13) will have a solution κ_0 between 0 and ζ_0 because the left side of (13) increases in this interval from a negative value to zero; the right side drops from zero to a negative value. On the other hand, (13) can have no root between ζ_0 and ξ_1 because the two sides of (13) have opposite signs in this interval. Again, there will be a root $\xi_1 < \kappa_1 < \zeta_1$ and in general in every interval (ξ_ν, ζ_ν). For negative κ, the situation will be opposite: the roots $-\kappa'_\nu$ will be so situated that $\zeta_{\nu-1} < \kappa'_\nu < \xi_\nu$. One can summarize this by stating that the absolute values of the positive imaginary poles $k = i\kappa_0, i\kappa_1, i\kappa_2, \ldots$ alternate with those of the negative imaginary poles $k = -i\kappa'_1, -i\kappa'_2, -i\kappa'_3, \ldots$

$$0 < \kappa_0 < \kappa'_1 < \kappa_1 < \kappa'_2 < \kappa_2 < \kappa'_3 \ldots \tag{15}$$

The situation is very similar if x_0 is negative, except that, in this case, either (15) holds or there can be two negative imaginary poles $k = -i\kappa'_0$ and $k = -i\kappa''_0$ the absolute values of which are both smaller than that of the smallest positive imaginary pole κ_0.

The situation as described above certainly seems the most natural one and can be expected to be the usual one. It is conceivable, however, that instead of the *a* single pole between 0 and ζ_0, for instance, one has three poles or, in fact, any odd number of poles. This can happen if the curves of Figure 1 are not as smooth as drawn but show appreciable curvature. In such a case, every heavily drawn interval in Figure 1 can contain an arbitrary odd number



of poles while the $(-\zeta_0, 0)$ interval, i.e. the first negative interval, will contain an even number of poles. The final rule which emerges is therefore as follows. If one orders the purely imaginary poles of S according to their absolute value, one first finds an even number (or zero) of negative imaginary poles, then an odd number of negative imaginary poles, and odd numbers of positive and of negative imaginary poles then continue to alternate.

The above alternating rule is by no means a sufficient condition for S to yield, by (6), a permissible R; it is only a necessary condition. If the properties of S, enumerated in the first paragraph, suffice to ensure the principle of causality [3], this principle cannot suffice to derive the properties of R, given in the second paragraph. Since these properties of R follow from the possibility to define a local probability and flux density, it would be interesting to ascertain whether they remain valid for particles, such as the light quantum, for which no such local densities can be defined.

References

[1] R. Kronig, Physica *12*, 543 (1946). This remark gave the original stimulus to the article of reference 3 and hence also to the present paper

[2] H. A. Kramers, Atti Cong. di Fisica Como 1927, p. 545, R. de. L. Kronig, Ned. T. Natuurk *9*, 402 (1942)

[3] W. Schutzer and J. Tiomno. Phys. Rev. *83*, 249 (1951)

[4] For the older literature of Chr. Moeller, K. Danske Vidensk. Selsk. *23*, no. 1 (1945); *22*, no. 19 (1946)

[5] Personal communication

[6] G. A. Campbell, Bell System Tech. Journ. *1*, no. 2 (1922), O. J. Zobel, ibid. *2*, 1 (1923), R. M. Foster, ibid. *3*, 259 (1924)

[7] W. Cauer, Arch. Elektrotechnik *17*, Chapter II (1926), Sitzungsber. Preuss. Akad. 1927, 1931, Math. Ann. *105*, 101 (1931), *106*, 369 (1932), Elektrische Nachrichten Technik *9*, 157 (1932)

[8] O. Brune, Journ. of Math. and Phys. *10*, 191 (1931). This article contains numerous errors

[9] K. Franz, Elektrische Nachrichten Technik *21*, 8 (1944)

[10] P. I. Richards, Duke Math. Journ. *14*, 777 (1947). I am much indebted to Dr. N. Greenspan for the last three references.

[11] The properties of R are discussed in some detail in the writer's article, Ann. of Math. *53*, 36 (1951) and in a forthcoming article Amer. Math. Monthly. The concept of the derivative matrix (and derivative function) was used by E. P. Wigner and L. Eisenbud, Phys. Rev. *72*, 29 (1947)

[12] S, as function of E, has a branch point at $E = 0$ but is a single valued function of k in the case of pure scattering and will therefore always be considered to be a function of k. On the other hand, R is a single valued function of both $E = k^2$ and k in the non-relativistic case but its properties are simpler if it is considered to be a function of E and will always be considered to be a function of E. In the relativistic case, or if, in addition to scattering, reactions are also possible, (i.e. if S and R become matrices), S has in general several branch points and R is single valued only as a function of E

[13] Cf. e.g. the first article of reference 11

[14] The product expansion (10) for somewhat specialized R was given in references 9 and 10. A general proof is given in the second reference of 11

Representations of the Galilei Group

E. Inönü and E. P. Wigner

Il Nuovo Cimento *9*, 705–718 (1952)

(ricevuto il 22 Maggio 1952)

Summary. — While the transition to a moving coordinate system $(x \to x + vt,\ t \to t)$ commutes in classical mechanics with displacements $(x \to x + a,\ t \to t)$, the corresponding operations in Schrödinger's nonrelativistic wave mechanics do not commute. In fact, the two operations, taken in different orders, differ by a factor $\exp[im\boldsymbol{v}\cdot\boldsymbol{a}/h]$. The present article considers the possibility of a nonrelativistic wave mechanics in which the transformations of the wave functions obey the same commutation relations as the transformations themselves. It shows that in such a mechanics position and momentum operators can exist only if it is reducible, i.e. only if the set of all states can be decomposed into subsets which are themselves invariant with respect to all permissible transformations of classical mechanics (rotations, displacements, proper Galilei transformations).

Introduction.

It is well known that every relativistically invariant wave equation defines a unitary representation of the inhomogeneous Lorentz group (1) (also called Poincaré group). It is not true, however, that, conversely, a physically meaningful theory corresponds to every unitary representation of this group. It seems, in fact, that the existence of particles with space-like momentum vectors (case 4 of reference 1, p. 186) would contradict the principle of causality and a system of wave functions which are all individually invariant with respect to displacements (case 3 of reference 1) cannot represent particles.

The purpose of the present article is to investigate the representations of the symmetry group of nonrelativistic mechanics (the Galilei group) and to

(1) E. P. WIGNER: *Ann. of Math.*, **40**, 149 (1939). Actually, as was discussed in this article, the representation need not be a true representation, only a representation up to a factor. This free factor corresponds to the indeterminate phase factor in the wave functions.

ascertain which of these representations may correspond to physically meaningful theories. The answer to this question is obvious for the representations which correspond to Schrödinger's nonrelativistic equation. These representations are again representations up to a factor: although the symmetry transformations which correspond to transition to a moving coordinate system commute with those which correspond to a displacement, the operators $\exp[im\boldsymbol{v}\cdot\boldsymbol{x}/h]$ and $\boldsymbol{x}\to\boldsymbol{x}+\boldsymbol{a}$ which correspond to these symmetry transformations do not commute. Their product, taken in the two different orders, differs by the factor $\exp[im\boldsymbol{v}\cdot\boldsymbol{a}/h]$. Furthermore, V. BARGMANN has shown ([2]) that these representations are essentially the only « up to a factor » representations. For this reason, the present work will be restricted to the investigation of the true representations of the Galilei group. The next section will be devoted to the description of this group and the determination of its representations. The ensuing section will deal with an attempt to find a physical interpretation to these equations. It will conclude that these representations do not correspond to equations describing particles. The investigation will include a search for states localized at one point, states with definite velocity components in all directions and states with definite velocity in only one direction.

1. – The Galilei group.

A. *Definition and Elementary Properties.*

The proper Galilei group contains the translations in space and time, the rotations and the transitions to a uniformly moving coordinate system (proper Galilei transformations or accelerations). The general element will be denoted by $G=(\boldsymbol{a}, b, \boldsymbol{v}, R)$ where $\boldsymbol{a}$ represents the space translations $\boldsymbol{x}'=\boldsymbol{x}+\boldsymbol{a}$, b represents the time translation $t'=t+b$, $\boldsymbol{v}$ represents the acceleration $\boldsymbol{x}'=\boldsymbol{x}+\boldsymbol{v}t$ and R represents the rotation $\boldsymbol{x}'=R\boldsymbol{x}$. The order of the transformations is from right to left. The relations between the elementary transformations are:

$$(\boldsymbol{a}_1)(\boldsymbol{a}_2) = (\boldsymbol{a}_2)(\boldsymbol{a}_1) = (\boldsymbol{a}_1+\boldsymbol{a}_2) \tag{1}$$

$$(b_1)(b_2) = (b_2)(b_1) = (b_1+b_2) \tag{2}$$

$$(\boldsymbol{v}_1)(\boldsymbol{v}_2) = (\boldsymbol{v}_2)(\boldsymbol{v}_1) = (\boldsymbol{v}_1+\boldsymbol{v}_2) \tag{3}$$

$$(R_1)(R_2) = (R_1R_2) \tag{4}$$

$$(\boldsymbol{a})(b) = (b)(\boldsymbol{a}) \tag{5}$$

([2]) Private communication.

$$(6) \qquad (\boldsymbol{a})(\boldsymbol{v}) = (\boldsymbol{v})(\boldsymbol{a})$$

$$(7) \qquad (R)(\boldsymbol{a}) = (\boldsymbol{a}')(R) \qquad \text{where } \boldsymbol{a}' = R\boldsymbol{a}$$

$$(8) \qquad (\boldsymbol{v})(b) = (b)(\boldsymbol{a})(\boldsymbol{v}) \qquad \text{where } \boldsymbol{a} = b\boldsymbol{v}$$

$$(9) \qquad (b)(R) = (R)(b)$$

$$(10) \qquad (R)(\boldsymbol{v}) = (\boldsymbol{v}')(R) \qquad \text{where } \boldsymbol{v}' = R\boldsymbol{v}\,.$$

We note the following points:

a) Space and time translations commute among themselves and with each other. The same is true for space translations and accelerations.

b) Accelerations do not commute with time translations. With respect to rotations they behave like space translations.

Using the relations (1)-(10) one easily calculates that the product of two inhomogeneous Galilei transformations

$$(11) \qquad (\boldsymbol{a}_1, b_1, \boldsymbol{v}_1, R_1)(\boldsymbol{a}_2, b_2, \boldsymbol{v}_2, R_2) = (\boldsymbol{a}, b, \boldsymbol{v}, R)\,,$$

where

$$\begin{aligned} \boldsymbol{a} &= \boldsymbol{a}_1 + b_2\boldsymbol{v}_1 + R_1\boldsymbol{a}_2\,, \\ b &= b_1 + b_2\,, \\ \boldsymbol{v} &= \boldsymbol{v}_1 + R_1\boldsymbol{v}_2\,, \\ R &= R_1R_2\,. \end{aligned}$$

The inverse is

$$(12) \qquad (\boldsymbol{a}, b, \boldsymbol{v}, R)^{-1} = \left(R^{-1}(b\boldsymbol{v} - \boldsymbol{a}),\ -b,\ -R^{-1}\boldsymbol{v},\ R^{-1}\right)\,,$$

the unit element is $(\boldsymbol{O}, O, \boldsymbol{O}, E)$

B. *Determination of the Irreducible Representations.*

Let $T(\boldsymbol{a})$, $\theta(b)$, $G(\boldsymbol{v})$, $O(R)$ be respectively the operators corresponding to the elementary transformations; then $U(G) = T(\boldsymbol{a})\theta(b)G(\boldsymbol{v})O(R)$ corresponds to the general Galilei transformation. If a state is described by a wave function ψ_g in the g frame, operating on ψ_g with any operator, say $T(\boldsymbol{a})$, results in a function $\psi_{g'}$ which describes the same state in the frame $g' = (\boldsymbol{a})g$. In other words, an experiment performed on $\psi_{g'}$ at the point $\boldsymbol{x}'$, t' will give the same result as one performed on ψ_g at the point $\boldsymbol{x}$, t, where $\boldsymbol{x}' = \boldsymbol{x} + \boldsymbol{a}$ and $t' = t$. We remark here that we shall be using throughout time independent wave functions (Heisenberg representation); these may be considered

to be the values for $t = 0$ of the time-dependent wave functions. The change of the state with time will be described by means of the operator $\theta(b)$. Thus the value of the time-dependent wave function at $t = t_0$ will be given by

$$\psi(t_0) = \theta(-t_0)\psi\,,$$

(where ψ is the time-independent function) as $\theta(-t_0)$ takes one into a frame with $t' = t - t_0$ and $t' = 0$ corresponds to $t = t_0$.

To form a true representation the operators $U(G)$ must satisfy the relation

$$U(G_1)U(G_2) = U(G_1G_2)\,, \tag{13}$$

where G_1G_2 is given in (11); or since the relations (1)-(10) are equivalent to (11), the elementary operators $T(\boldsymbol{a})$, $\theta(b)$, $G(\boldsymbol{v})$ and $O(R)$ must satisfy the relations (1)-(10) with the operators substituted for the corresponding transformations. We shall determine the representations by Frobenius' method which builds up the representations of the whole group out of the representations of an invariant subgroup by means of a «little group». This method was applied previously to the determination of the representations of the inhomogeneous Lorentz group ([1]).

For the invariant subgroup we can take either the group of the space and time translations or of the space translations and accelerations. It appears more convenient to use the second. As $T(\boldsymbol{a})$ and $G(\boldsymbol{v})$ commute with each other as well as among themselves, it is possible to introduce such a coordinate system in Hilbert space that the wave functions $\psi(\boldsymbol{p}, \boldsymbol{q}, \zeta)$ depend on two vectors $\boldsymbol{p}$, $\boldsymbol{q}$ and a discrete variable ζ and for which

$$T(\boldsymbol{a})\psi(\boldsymbol{p}, \boldsymbol{q}, \zeta) = \exp\,[i\boldsymbol{p}\cdot\boldsymbol{a}]\psi(\boldsymbol{p}, \boldsymbol{q}, \zeta)\,, \tag{14}$$

$$G(\boldsymbol{v})\psi(\boldsymbol{p}, \boldsymbol{q}, \zeta) = \exp\,[i\boldsymbol{q}\cdot\boldsymbol{v}]\psi(\boldsymbol{p}, \boldsymbol{q}, \zeta)\,. \tag{15}$$

On defines an operator $P(b)$ by

$$P(b)\psi(\boldsymbol{p}, \boldsymbol{q}, \zeta) = \psi(\boldsymbol{p}, \boldsymbol{q} - b\boldsymbol{p}, \zeta)\,. \tag{16}$$

Thus the function $\varphi = P(b)\psi$ has at the point $(\boldsymbol{p}, \boldsymbol{q}, \zeta)$ the same value as ψ at the point $(\boldsymbol{p}, \boldsymbol{q} - b\boldsymbol{p}, \zeta)$. We have then

$$G(\boldsymbol{v})P(b)\psi(\boldsymbol{p}, \boldsymbol{q}, \zeta) = \exp\,[i\boldsymbol{q}\cdot\boldsymbol{v}]P(b)\psi(\boldsymbol{p}, \boldsymbol{q}, \zeta) = \exp\,[i\boldsymbol{q}\cdot\boldsymbol{v}]\psi(\boldsymbol{p}, \boldsymbol{q}, -b\boldsymbol{p}, \zeta)\,.$$

One can calculate $P(b)T(b\boldsymbol{v})G(\boldsymbol{v})\psi(\boldsymbol{p}, \boldsymbol{q}, \zeta)$ in a similar way and finds

$$G(\boldsymbol{v})P(b) = P(b)T(b\boldsymbol{v})G(\boldsymbol{v})\,. \tag{17}$$

Together with

$$P(b)^{-1} = P(-b)\,, \qquad P(b)T(\boldsymbol{a}) = T(\boldsymbol{a})P(b)\,, \tag{18}$$

this shows that $P(b)$ satisfies the same equations which (5), (8) require $\theta(b)$ to satisfy. Hence $\theta(b)P(b)^{-1}$ commutes with all $G(\boldsymbol{v})$ and $T(\boldsymbol{a})$ and therefore with any function of $\boldsymbol{q}$ and $\boldsymbol{p}$. One can set therefore

$$\theta(b) = V(b)P(b)\,, \tag{19}$$

where $V(b)$ operates on ζ only, though it may depend on $\boldsymbol{p}$ and $\boldsymbol{q}$; or

$$\theta(b)\psi(\boldsymbol{p}, \boldsymbol{q}, \zeta) = \sum_{\eta} V(\boldsymbol{p}, \boldsymbol{q}, b)_{\zeta\eta}\psi(\boldsymbol{p}, \boldsymbol{q} - b\boldsymbol{p}, \eta)\,. \tag{20}$$

Next we define an operator $Y(R)$ by

$$Y(R)\psi(\boldsymbol{p}, \boldsymbol{q}, \zeta) = \psi(R^{-1}\boldsymbol{p}, R^{-1}\boldsymbol{q}, \zeta)\,.$$

A similar argument, using the relations (7) and (10) shows that $O(R)Y(R)^{-1}$ does not operate on $\boldsymbol{p}$ or $\boldsymbol{q}$, so that

$$O(R)\psi(\boldsymbol{p}, \boldsymbol{q}, \zeta) = \sum_{\eta} Z(\boldsymbol{p}, \boldsymbol{q}, R)_{\zeta\eta}\psi(R^{-1}\boldsymbol{p}, R^{-1}\boldsymbol{q}, \eta)\,. \tag{21}$$

The operators $T(\boldsymbol{a})$, $\theta(b)$, $G(\boldsymbol{v})$, $O(R)$ as determined in (14), (15), (20), (21) satisfy all the commutation relations (1)-(10) except (2), (4) and (9). Before determining V and Z on the basis of these relations we determine the variability domain of the vectors $\boldsymbol{p}$, $\boldsymbol{q}$. This will be different for the different irreducible represensations of the Galilei group.

If the $T(\boldsymbol{a})$, $G(\boldsymbol{v})$ constituted the whole group, it would be possible to restrict the variability domain of the pair $\boldsymbol{p}$, $\boldsymbol{q}$ to a single point $\boldsymbol{p} = \boldsymbol{p}_0$, $\boldsymbol{q} = \boldsymbol{q}_0$ and the irreducible representations would all become one dimensional. Because of the group elements (R) and (b) this is not possible: (20) and (21) show, that in the representation in which $\boldsymbol{p}_0$, $\boldsymbol{q}_0$ are permissible values of the variables, the pairs $R\boldsymbol{p}_0$, $R\boldsymbol{q}_0$ and $\boldsymbol{p}_0$, $\boldsymbol{q}_0 - b\boldsymbol{p}_0$ are likewise permissible values. This is true for all rotations R and all values of b.

On the other hand, in an irreducible representation, the variability domain of the pair $\boldsymbol{p}$, $\boldsymbol{q}$ will be restricted to those pairs which can be obtained from a single pair $\boldsymbol{p}_0$, $\boldsymbol{q}_0$ by a succession of the above transformations. Their general form is

$$\boldsymbol{p} = R\boldsymbol{p}_0\,, \qquad \boldsymbol{q} = R\boldsymbol{q}_0 - bR\boldsymbol{p}_0\,, \tag{22}$$

with R an arbitrary rotation, b an arbitrary real number. The definition domain (22) corresponds to the definition domain of the four-dimensional

vectors $\boldsymbol{p}$ in the case of the inhomogeneous Lorentz group, which contained all the vectors $\Lambda \boldsymbol{p}_0$ where Λ was a homogeneous Lorentz transformation.

Depending on the original pair of vectors $\boldsymbol{p}_0$, $\boldsymbol{q}_0$, there are four types of manifolds (22) which will give rise to four classes of representations. In class I, $\boldsymbol{p}_0$ and $\boldsymbol{q}_0$ are both finite and have different directions. In this case the manifold (22) can be defined by the equations

$$\text{(22.I)} \qquad \boldsymbol{p}\cdot\boldsymbol{p} = \boldsymbol{p}_0\cdot\boldsymbol{p}_0 = P^2\,; \qquad |\boldsymbol{p}\wedge\boldsymbol{q}| = |\boldsymbol{p}_0\wedge\boldsymbol{q}_0| = S$$

and characterized by the positive P and S. This is the general and most important case.

In the second case, $\boldsymbol{p}$ and $\boldsymbol{q}$ are collinear. These manifolds can be characterized by a positive P

$$\text{(22.II)} \qquad \boldsymbol{p}\cdot\boldsymbol{p} = P^2\,.$$

In the third case

$$\text{(22.III)} \qquad \boldsymbol{p} = 0\,, \qquad \boldsymbol{q}\cdot\boldsymbol{q} = Q^2\,,$$

the manifold can be characterized by a positive Q. In the last case

$$\text{(22.IV)} \qquad \boldsymbol{p} = 0\,, \qquad \boldsymbol{q} = 0\,.$$

The result of the operations $O(R)$ and $\theta(b)$ can be now determined in the same way as the result of $d(\Lambda)$ was determined in reference 1. Only the result of this calculation will be given; it is different for the four different classes.

I. In class I the discrete variable ζ of (14), (15) can assume only one value and can be suppressed, therefore. One has

$$\text{(23.I)} \qquad O(R)\psi(\boldsymbol{p}, \boldsymbol{q}) = \psi(R^{-1}\boldsymbol{p}, R^{-1}\boldsymbol{q})\,,$$

$$\text{(24.I)} \qquad \theta(b)\ \psi(\boldsymbol{p}, \boldsymbol{q}) = \psi(\boldsymbol{p}, \boldsymbol{q} - b\boldsymbol{p})\,.$$

One can convince oneself directly that the $O(R)$, $\theta(b)$ defined by these equations, together with the $T(\boldsymbol{a})$ and $G(\boldsymbol{v})$ given by (14), (15), form a representation of the Galilei group, i.e. that these operators satisfy the relations (1)-(10). For class I, the «little group» consists only of the identity and its representation is one dimensional. This causes ζ to assume only one value.

The representations of class I can be characterized by the two positive constants P, S and each such pair P, S defines a representation.

II. In this case the variables $\boldsymbol{q}$ can be replaced by the single variable s by $\boldsymbol{q} = s\boldsymbol{p}$ and (15) can be replaced by

$$G(\boldsymbol{v})\psi(\boldsymbol{p}, s) = \exp\,[is\boldsymbol{p}\cdot\boldsymbol{v}]\psi(\boldsymbol{p}, s)\,. \tag{15.II}$$

ζ can again assume only one value and has been suppressed. For $\theta(b)\psi$ one has

$$\theta(b)\psi(\boldsymbol{p}, s) = \psi(\boldsymbol{p}, s - b)\,, \tag{24.II}$$

which expresses the same relation as (24.I). In order to define $O(R)$ one has to select a permissible but otherwise arbitrary $\boldsymbol{p}_0$ and coordinate to every $\boldsymbol{p}$ a rotation $r(\boldsymbol{p})$ so that

$$r(\boldsymbol{p})\boldsymbol{p}_0 = \boldsymbol{p}\,. \tag{25}$$

Then

$$P_\varphi = r(\boldsymbol{p})^{-1}Rr(R^{-1}\boldsymbol{p})\,, \tag{25a}$$

leaves $\boldsymbol{p}_0 = P_\varphi\boldsymbol{p}_0$ invariant, i.e. is a rotation about $\boldsymbol{p}_0$, its angle of rotation is denoted by φ. This is a function of R and $\boldsymbol{p}$. For $O(R)\psi$ one has

$$O(R)\psi(\boldsymbol{p}, s) = \exp\,[im\varphi(R, \boldsymbol{p})]\psi(R^{-1}\boldsymbol{p}, s)\,, \tag{23.II}$$

where m is an arbitrary integer. The little group of this case is the group of rotations about $\boldsymbol{p}_0$, its representations are one dimensional (as a result of which ζ can assume only one value); they can be characterized by the integer m. The representations of class II are characterized by the positive number P, and the integer m.

III. In the remaining two cases all states are invariant with respect to translations and this makes it evident that representations of the classes III and IV cannot correspond to equations which describe particles. These representations are actually only representations of smaller groups, viz. the factor group of the invariant subgroups formed by the $T(\boldsymbol{a})$ and the $T(\boldsymbol{a})G(\boldsymbol{v})$ respectively.

Since the variables $\boldsymbol{p}$ are restricted to the value O, they can be suppressed and one has, instead of (14)

$$T(\boldsymbol{a})\psi = \psi\,. \tag{14.III}$$

(15) remains valid in case III while (23.II) is replaced by

$$O(R)\psi(\boldsymbol{q}) = \exp\,[im\varphi]\psi(R^{-1}\boldsymbol{q})\,, \tag{23.III}$$

φ is again defined by (25), (25a), except that a $\boldsymbol{q}_0$ and $\boldsymbol{q}$ replace $\boldsymbol{p}_0$ and $\boldsymbol{p}$. Instead of (24.II), one has

$$\theta(b)\psi(\boldsymbol{q}) = \exp\,[-i\varepsilon b]\psi(\boldsymbol{q})\,. \tag{24.III}$$

ε being an arbitrary real number. The representations of class III are characterized by the real ε, positive Q and integer m. The little group consists in this case of the rotations about $\boldsymbol{q}_0$ and displacements in time; its representations are one dimensional which explains the absence of the variable ζ.

IV. In this case both $\boldsymbol{p}$ and $\boldsymbol{q}$ can be suppressed and all wave functions are invariant with respect to both $T(\boldsymbol{a})$ and $G(\boldsymbol{v})$. The little group consists of all operations (R) and (b) and one has

$$O(R)\psi(\zeta) = \sum_{\eta} D^{(l)}(R)_{\zeta\eta}\psi(\eta)\,, \tag{23.IV}$$

$$\theta(b)\ \psi(\zeta) = \exp\,[-i\varepsilon b]\psi(\zeta)\,. \tag{24.IV}$$

The representations are characterized by the real variable ε and the integer l, the total number of linearly independent wave function is $2l+1$. The $D^{(l)}(R)$ are the representations of the ordinary three dimensional rotation group.

C. *The Improper Galilei Group.*

The operator I corresponding to inversion $(x' = -x,\ y' = -y,\ z' = -z)$ can be found on the basis of the commutation relations

$$T(-\boldsymbol{a})I = IT(\boldsymbol{a})\,, \qquad G(-\boldsymbol{v})I = IG(\boldsymbol{v})\ ; \tag{26}$$

its square will be 1 and it will commute with $0(R)$ and $\theta(b)$. These relations can be satisfied, in case of class I representations by

$$I\psi(\boldsymbol{p}, \boldsymbol{q}) = \pm\,\psi(-\boldsymbol{p}, -\boldsymbol{q})\,. \tag{27.I}$$

The $\pm$ defines a parity for these representations. For representations of class II, unless $m = 0$, the extension of the proper Galilei group to include the inversion unites the representations P, m and P, $-m$. If we denote the wave functions which belong to the former representation of the proper Galilei group by ψ_+, those which belong to the latter representation of the proper Galiley group by ψ_-, the most natural definition of the inversion operator is

$$\begin{cases} I\psi_+(\boldsymbol{p}, s) = \psi_-(-\boldsymbol{p}, s) \\ I\psi_-(\boldsymbol{p}, s) = \psi_+(-\boldsymbol{p}, s)\,. \end{cases} \tag{27.II}$$

In the case of $m = 0$, the indices $+$, $-$ can be omitted; the representations P, O can be extended to representations of the Galilei group which includes the operation of inversion. In this case $I\psi(\boldsymbol{p}, s)$ can be either $\psi(-\boldsymbol{p}, s)$ or $-\psi(-\boldsymbol{p}, s)$ which again defines a parity for the P, O representations. Introducing a $-$ sign in (27.II) only gives another form to the same representation: the representations P, $m \neq 0$ have no parity. The reader will recognize the similarity with the extension of the two dimensional rotation group to the two dimensional rotation-reflection group.

The same situation prevails also with the representations of class III except that $\boldsymbol{q}$ plays the role which $\boldsymbol{p}$ played above. The representations of class IV also can be extended to include the operation of inversion and the extension again leads to a distinction of two types of representations, of even and odd parity.

The operation of time inversion θ will commute with $T(\boldsymbol{a})$ and $O(R)$ and

$$\theta(-b)\theta = \theta\theta(b)\,, \qquad G(-\boldsymbol{v})\theta = \theta G(\boldsymbol{v})\,. \tag{28}$$

We further postulate that θ involve complex conjugation K. These requirements can be most naturally satisfied, for class I representations, by

$$\theta\psi(\boldsymbol{p}, \boldsymbol{q}) = K\psi(-\boldsymbol{p}, \boldsymbol{q}) \equiv \psi(-\boldsymbol{p}, \boldsymbol{q})^*\,. \tag{29.I}$$

For the representations of class II one can set

$$\theta\psi_\pm(\boldsymbol{p}, s) = K\psi_\pm(-\boldsymbol{p}, -s)\,. \tag{29.II}$$

The situation is again similar to that encountered when extending the representations of the two-dimensional rotation group to include time inversion and this applies also to class III. The representations of class IV are actually representations of the direct product of the three-dimensional rotation group with the group of time displacement. The effect af adjoining to this group the operation of time inversion has been discussed before in detail.

It should be remarked that the above discussion of the extension of the representations to include time inversion is not exhaustive. It deals only with the most simple and natural possibility. The results of the following sections are not affected by this incompleteness.

D. *Definition of Scalar Product.*

The wave functions of the first class depend on two vectors $\boldsymbol{p}$, $\boldsymbol{q}$ whose variability domain is restricted by the conditions $\boldsymbol{p}^2 = P^2, |\boldsymbol{p} \wedge \boldsymbol{q}| = S$. Hence we define the scalar product (ψ, φ) of two functions $\psi(\boldsymbol{p}, \boldsymbol{q})$, $\varphi(\boldsymbol{p}, \boldsymbol{q})$ by the

integral,

$$(\psi, \varphi) = \int \psi^*(\boldsymbol{p}, \boldsymbol{q})\varphi(\boldsymbol{p}, \boldsymbol{q})\, \delta(\boldsymbol{p}^2 - P^2)\, \delta(|\boldsymbol{p} \wedge \boldsymbol{q}| - S)\, \mathrm{d}\boldsymbol{p}\, \mathrm{d}\boldsymbol{q}\,. \tag{30.I}$$

All our operators are unitary with respect to the scalar product (30). This is evident as far as the $T(\boldsymbol{a})$ and $G(\boldsymbol{v})$ of (14), (15) are concerned but remains valid also for $O(R)$ and $\theta(b)$ defined in (23.I) and (24.I). One can show this by the substitutions $\boldsymbol{p}' = R^{-1}\boldsymbol{p}$, $\boldsymbol{q}' = R^{-1}\boldsymbol{q}$ and $\boldsymbol{p}' = \boldsymbol{p}$, $\boldsymbol{q}' = \boldsymbol{q} - b\boldsymbol{p}$ which, because of

$$R(\boldsymbol{p} \wedge \boldsymbol{q}) = R\boldsymbol{p} \wedge R\boldsymbol{q}\,,$$

leave the δ functions of (30) unchanged. Similarly, for the representations of class II the unitary scalar product becomes

$$(\psi, \varphi) = \int \psi^*(\boldsymbol{p}, s)\varphi(\boldsymbol{p}, s)\delta(\boldsymbol{p}^2 - P)\, \mathrm{d}\boldsymbol{p}\, \mathrm{d}s\,. \tag{30.II}$$

2. - Physical Interpretation of the Representations of the Galilei Group.

Of the four classes of representations only the first two can have a straightforward physical intepretation because only in these the operation of displacement has an effect on the wave functions. As the first class is the more general of the two we shall discuss this one in detail limiting ourselves to brief remarks about the second.

A. *Search for localized states.*

If a representation is to describe in some sense the motion of a particle one will expect to find among its wave functions some describing states which represent, at a definite time, $t = 0$, a particle at a definite point of space, say at $x = y = z = 0$. Knowing these functions implies knowing all characteristic functions of the position operator and consequently the position operator itself. Both the non-relativistic Schrödinger equation and the various relativistic equations for particles with finite mass and arbitrary spin or zero mass and zero or one-half spin have localized states ([3]).

We shall require that the wave functions localized at $x = y = z = 0$ at $t = 0$ satisfy the following requirements:

(a) They must form a linear set S_0; i.e. the superposition of two such states must again be localized in the same manner.

([3]) T. D. NEWTON and E. P. WIGNER: *Rev. Mod. Phys.*, **21**, 400 (1949).

(*b*) If ψ is localized, then $T(\boldsymbol{a})\psi$ must be orthogonal to all states of S_0 for $\boldsymbol{a} \neq 0$.

(*c*) The set S_0 must be invariant under rotations around the origin and reflections in space and time.

One can easily see that the postulate (*c*) is compatible with (*b*). Using the unitary nature of $O(R)$, one finds

$$(T(a)O(R)\psi, O(R)\varphi) = (O(R)^{-1}T(a)O(R)\psi, \varphi) = (T(R^{-1}a)\psi, \varphi) = 0\,,$$

so that $O(R)S_0$ satisfies postulate (*b*) if S_0 does. The same applies for space and time inversions. We shall see, however, that postulate (*b*) already leads to a contradiction.

For this purpose, let us write down (*b*) by means of (30.I) for a function ψ which belongs to S_0

$$\int \exp\,[i\boldsymbol{p}\cdot\boldsymbol{a}]\,|\psi(\boldsymbol{p}, \boldsymbol{q})|^2\,\delta(\boldsymbol{p}^2 - P^2)\,\delta(|\boldsymbol{p}\wedge\boldsymbol{q}| - S)\,\mathrm{d}\boldsymbol{p}\,\mathrm{d}\boldsymbol{q} = 0\,, \tag{31}$$

for $\boldsymbol{a} \neq 0$. We can average this equation with respect to $\boldsymbol{a}$ over any domain which does not include $\boldsymbol{a} = 0$. In particular, averaging over a sphere of radius a gives

$$\int \frac{\sin a|\boldsymbol{p}|}{a|\boldsymbol{p}|}\,|\psi(\boldsymbol{p}, \boldsymbol{q})|^2\,\delta(\boldsymbol{p}^2 - P^2)\,\delta(|\boldsymbol{p}\wedge\boldsymbol{q} - S)\,\mathrm{d}\boldsymbol{p}\,\mathrm{d}\boldsymbol{q} = \frac{\sin aP}{aP} = 0\,. \tag{31a}$$

Because of the $\delta(\boldsymbol{p}^2 - P^2)$ factor, the $|\boldsymbol{p}|$ in the integrand of (31*a*) can be replaced by P and the corresponding factor brought outside the integral. The latter then becomes the normalization integral for ψ. Hence (31*a*) shows that the condition (31) cannot be satisfied. The same argument applies also to representations of class II.

Clearly, the $\delta(\boldsymbol{p}^2 - P^2)$ factor in the definitions (30) of the scalar product is responsible for this result. A linear manifold of functions in which $|\boldsymbol{p}|$ can assume all values does contain localized states. However, such a linear manifold is not irreducible.

Our equation (31*a*) shows that not even approximate localization is possible for an irreducible manifold. In fact the integral (31*a*) immediately gives the transition probability between an arbitrary wave function ψ and the linear combination of wave functions

$$\varphi = \frac{1}{4\pi}\int T(a\boldsymbol{\Omega})\psi\,\mathrm{d}\boldsymbol{\Omega}\,.$$

$\boldsymbol{\Omega}$ is a unit vector, $\mathrm{d}\boldsymbol{\Omega}$ indicates integration over all directions.

If ψ is approximately localized at the origin, φ is a linear combination of wave functions which are approximately localized somewhere on a sphere of ra-

dius a. Furthermore, it is easy to see that the norm $(\varphi, \varphi) \leqslant 1$. Hence (31$a$) shows that the transition probability of ψ into a function φ which should be localized at a distance a from ψ is

$$|(\psi, \varphi)|^2 / (\varphi, \varphi) \geqslant \sin^2 aP/(aP)^2 .$$

It decreases for large a at best as a^{-2}. This is the best that one can accomplish in approximate localization and if one chooses ψ accordingly the transition probability of ψ into all states of $T(\boldsymbol{a})S_0$ will be proportional to a^{-2} for large a. Since the volume of a thin spherical shell is $4\pi a^2 \, \mathrm{d}a$, it appears that a particle localized at O as well as possible is, with equal probabilities, in any spherical shell of unit thickness about O.

B. *Search for States with Definite Velocity.*

Again, we shall try to find a linear manifold V_0 of wave functions which represent states with zero velocity. We shall stipulate that if φ is such a state $G(\boldsymbol{v})\varphi$ be orthogonal to all states of V_0 if $\boldsymbol{v} \neq 0$ and further that V_0 shall be invariant under translations $T(\boldsymbol{a})$. It would be natural to postulate further that V_0 be invariant also under rotations, reflections, etc., but it will be impossible to meet even the requirements with respect to $G(\boldsymbol{v})$ and $T(\boldsymbol{a})$.

According to the above postulates $T(\boldsymbol{a})\varphi$ has zero velocity if φ has this property. Hence

$$(G(\boldsymbol{v})T(\boldsymbol{a})\varphi, \varphi) = 0 , \tag{32}$$

for all a and $\boldsymbol{v} \neq 0$. Writing out (32) again by means of (30.I)

$$\int \exp\,[i\boldsymbol{p}\cdot\boldsymbol{a}] \exp\,[i\boldsymbol{q}\cdot\boldsymbol{v}] |\varphi(\boldsymbol{p}, \boldsymbol{q})|^2 \delta(\boldsymbol{p}^2 - P^2)\, \delta(|\boldsymbol{p} \wedge \boldsymbol{q}| - S)\, \mathrm{d}\boldsymbol{p}\, \mathrm{d}\boldsymbol{q} = 0 ,$$

for $\boldsymbol{v} \neq 0$. Since the $\exp\,[i\boldsymbol{p}\cdot\boldsymbol{a}]$ form a complete set of functions of $\boldsymbol{p}$

$$\int \exp\,[i\boldsymbol{q}\cdot\boldsymbol{v}] |\varphi(\boldsymbol{p}, \boldsymbol{q})|^2 \, \delta(|\boldsymbol{p} \wedge \boldsymbol{q}| - S)\, \mathrm{d}\boldsymbol{q} = 0 , \tag{32a}$$

must hold for almost all $|\boldsymbol{p}| = P$. Let $\boldsymbol{p}_0$ be such a $\boldsymbol{p}$ and let us decompose both $\boldsymbol{q} = \boldsymbol{r} + \alpha \boldsymbol{p}_0/P$ and $\boldsymbol{v} = \boldsymbol{s} + \beta \boldsymbol{p}_0/P$ perpendicular and parallel to $\boldsymbol{p}_0$. The last equation then becomes

$$\int \exp\,[i\boldsymbol{r}\cdot\boldsymbol{s} + i\alpha\beta] |\varphi(\boldsymbol{p}_0, \boldsymbol{q})|^2 \, \delta(P|\boldsymbol{r}| - S)\, \mathrm{d}\boldsymbol{q} = 0.$$

This equation can be integrated over any domain which does not include $\boldsymbol{v} = 0$. Hence we multiply it with an arbitrary function $f(\beta)$ and integrate it over the cylinder $|\boldsymbol{s}| = v_0$. The integration over the angle which gives the direction of $\boldsymbol{s}$ on the cylinder can be carried out and replaces the first factor of the integrand by $2\pi J_0(|\boldsymbol{r}|v_0)$ which can be replaced, on account of the δ function, by $2\pi J_0(Sv_0/P)$. As this does not vanish in general, we have

$$\int f(\beta) \exp[i\alpha\beta] |\varphi(\boldsymbol{p}_0, \boldsymbol{q})|^2 \delta(P|\boldsymbol{r}| - S) \, \mathrm{d}\boldsymbol{q} \, \mathrm{d}\beta = 0 ,$$

which, because of the arbitrariness of f, shows that

$$\int |\varphi(\boldsymbol{p}_0, \boldsymbol{q})|^2 \delta(|\boldsymbol{p}_0 \wedge \boldsymbol{q}| - S) \, \mathrm{d}\boldsymbol{q} = 0 , \tag{32b}$$

vanishes for almost all $\boldsymbol{q}$. This, however, entails the vanishing of the normalization integral (30.I) for φ. It follows that no states with zero velocity can be found among the states of the representations of class I. This result can be extended also to the representations of class II.

It may be of some interest to remark that states with definite velocity can be found if one does not require that the velocity be invariant with respect to displacements in space.

C. *Other Attempts at Interpretation.*

We have tried to define states for which at least one component of the velocity vanishes and found that this is the case for a φ which vanishes unless $\boldsymbol{p}$ has the direction in which $\boldsymbol{v}$ is to have zero component. From this, states for which the velocity has a prescribed component in a given direction can be easily found by (15). However, one finds that the totality of the states obtained in this way does not form a complete set of functions and no operator can be defined for the component of the velocity in a prescribed direction.

Conclusion.

We have not succedeed in giving a physical content to any of the irreducible representations of the inhomogeneous Galilei group. In particular, we were unable to describe states localized at one space point or states with a definite velocity. There exist some states with a definite velocity in only one direction, but since they do not form a complete set their field of application is very limited. The identification of the infinitesimal translation operator

of the Galilei group with the momentum operator leads to absurd results. Unless one should find a physical interpretation radically different from the one here attempted, one is forced to conclude that these representations are similar to those representations of the Lorentz group, the momentum vectors of which either vanish or are space-like. It appears surprising that merely by going over from the representations up to a factor to true representations one encounters such a radical change in interpretability.

RIASSUNTO

Mentre la transizione a un sistema di coordinate in movimento ($x \to x + vt$, $t \to t$) in meccanica classica commuta con gli spostamenti ($x \to x + a$, $t \to t$), le operazioni corrispondenti non commutano nella meccanica ondulatoria non relativistica di Schrödinger. Precisamente, le due operazioni prese in ordine differente differiscono di un fattore $\exp[im\boldsymbol{v}\cdot\boldsymbol{a}/h]$. Il presente articolo considera la possibilità di una meccanica ondulatoria non relativistica in cui le trasformazioni delle funzioni d'onda obbediscano alle stesse relazioni di commutazione delle trasformazioni stesse. Si mostra che gli operatori di posizione e di quantità di moto possono esistere in tale meccanica solo se essa è riducibile, cioè solo se l'insieme di tutti gli stati può essere decomposto in sottogruppi che sono essi stessi invarianti rispetto a tutte le trasformazioni permesse della meccanica classica (rotazioni, spostamenti, trasformazioni proprie di Galileo).

Simplified Derivation of the Properties of Elementary Transcendentals

E. P. Wigner

American Mathematical Monthly *59*, 669–683 (1952)

The purpose of the present paper is to give a derivation of the properties of the elementary transcendentals from a more unified point of view than is customary. It will be assumed that the general properties of analytic functions are established. In particular we shall use the theorem that the real (or imaginary) part of an analytic function cannot assume its maximum value in the inside of a closed domain in which it is regular. Similarly, use will be made of Weierstrass's theorem about a set of uniformly converging analytic functions, of Liouville's and Rolle's theorems.§ All of these theorems can be established without the use of the properties of the transcendental functions, using only elementary calculations and the concept of convergence. The present paper shows how the properties of the elementary transcendentals can then be derived from the aforementioned theorems by means of the methods of the theory of complex functions. From the point of view of this theory, this appears to be more appropriate than taking over the results concerning the properties of these functions, at least for real arguments, from other parts of analysis. An outline of our procedure follows.

As a first step, a special class of meromorphic functions, the so-called R functions, will be introduced and a number of properties of these functions, including two expansion theorems, established (Section 1). Some of these properties will not be used for the derivation of the properties of elementary transcendentals; the numbers of the corresponding theorems have a star attached to them.

The tangent then will be defined in Section 2 as the simplest periodic R function. The partial fraction and infinite product expansion of the tangent is a special case of the expansion theorems mentioned before while the addition theorem for this function will follow from its periodic nature. It is easy to obtain then the similar expansions for the other trigonometric and the exponential functions.

The writer wishes to admit at the outset that his definition of the exponential function is less direct than that of Caratheodory,† for instance. On the other hand, the derivation of the various expansion theorems which will be given below is less artificial and more general than the customary derivations.

§ *Cf. e.g.* A. Hurwitz and R. Courant, *Allgemeine Funktionentheorie etc.* Verlag Julius Springer, Berlin 1922, III 2, page 296; I 3, p. 63 and I 3, p. 56.

† C. Caratheodory, *Funktionentheorie*, Verlag Birkhauser, Basel 1950, Sec. 234, p. 233.

1. Properties of the R functions. R functions are defined ‡ as meromorphic functions of a complex variable z, the imaginary part of which is non-negative in the upper and non-positive in the lower half-plane. For the sake of the reader's convenience, some rather obvious properties of such functions are repeated here from an earlier article of the writer.‡ A few of these properties had been given previously by Fränz.‡

THEOREM 1.1. *If a set of R functions converges to a meromorphic function, this function itself is an R function.*

THEOREM 1.2. *An R function is real on the real axis and assumes real values only on the real axis.*

THEOREM 1.3. *The derivative of an R function is positive at every regular point on the real axis unless R is a constant.*

THEOREM 1.4. *All poles of an R function are on the real axis, they are simple and their residues negative.* The poles will be denoted henceforth by Z_μ, the corresponding residues by $-\gamma_\mu^2$; both the Z and γ are then real. A typical R function for real values of its argument is given in Figure 1.

‡ As far as the writer could determine, R functions appear in the mathematical literature first (though not under this name) in K. Loewner's article *Ueber monotone Matrixfunktionen*, Math. Zeits. 38, 177 (1933). In this article Loewner considers functions $R(m)$ of real symmetric or hermitean matrices m which satisfy the condition that $R(m_1)-R(m_2)$ be positive definite if m_1-m_2 is positive definite and if the characteristic values of m_1 and m_2 are in an interval (λ, λ'). He shows that the $R(m)$ are R functions which are regular in the (λ, λ') interval. The writer is much indebted for this quotation to S. Sherman. Certain generalizations of R functions also play an important role in the theory of moments. *Cf. e.g.* J. A. Shohat and J. D. Tamarkin, The Problem of Moments, Publication of the American Mathematical Society, New York 1943, particularly Chapters I and II. The R functions have also played a substantial if not always obvious role in electrical engineering, in particular in the theory of the reactance of circuits, ever since the pioneering investigations of O. J. Zobel, G. A. Campbell and R. M. Foster (Bell System Technical Journal, Volumes 1, 2 and 3, 1922–24). *Cf.* also W. Cauer, Preuss. Akad. Ber. 1931, p. 30. A more modern review of the theory has been given recently by H. W. Bode, Network Analysis and Feedback Amplifier Design (Van Nostrand, 1945). However, most of this work deals with rational R functions for which most of the contents of the present article are obvious. An exception is the article of K. Fränz, Elektrische Nachrichten Technik, 21, 8 (1944) who already considers transcendental R functions. The writer is much indebted for this quotation to Dr. T. Teichmann. For odd R functions, the expansion theorems 1.10 and 1.12 also follow from the work of P. I. Richards, Duke Math. Journ., 14, 777, 1947, which is based on W. Cauer's article in Bull. Amer. Math. Soc., 38, 713, 1932. This article anticipates, for odd R functions, many of the results of our Section 1. The connection between his *iPRF* function $Z(s)$ and the R functions is $Z(s)=-iR(is)$. *Cf.* also Richards' summary article, Quarterly Journ. of Math., 6, 21, 1948. The writer's interest in these functions was raised by his work on the quantum mechanical theory of reaction and scattering processes. *Cf.* in this connection E. P. Wigner and L. Eisenbud, Physical Rev. 72, 29 (1947) and the summary article Amer. Journ. of Physics, 17, 99 (1949). The first general investigation of the mathematical properties of these functions is given in Ann. of Math. 53, 36 (1951). It was not entirely possible to eliminate all repetition between this last article and the present one.

THEOREM 1.5. *A linear combination of R functions with positive coefficients is an R function.*

THEOREM 1.6. *The linear function* $\alpha z+\beta$ *is an R function if* α *is positive and* β *real.*

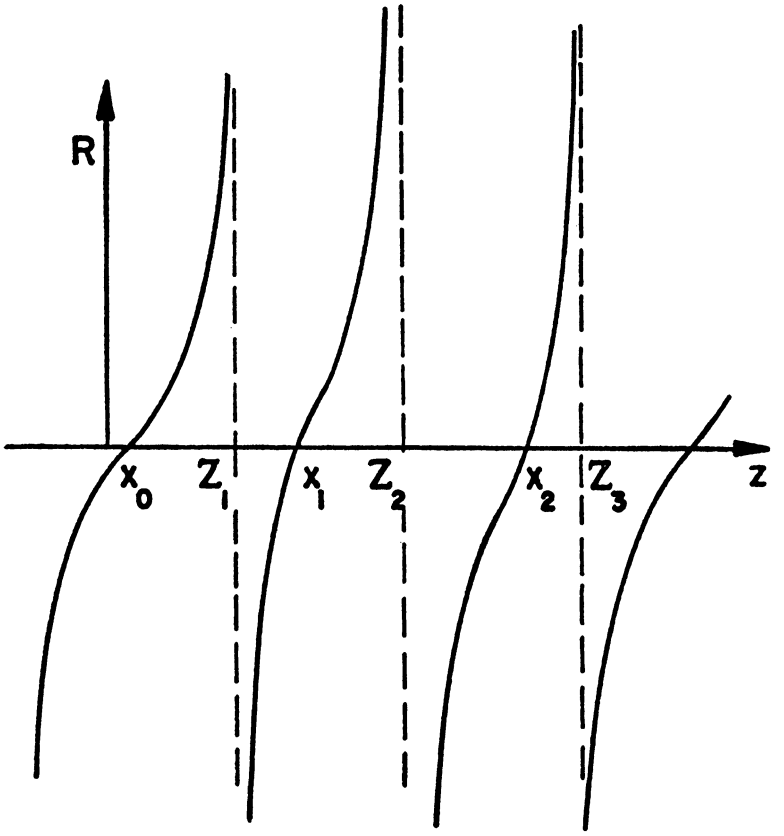


FIG. 1

THEOREM 1.7. *A linear fractional transform of an R function*

$$S = \frac{a_1R + b_1}{a_2R + b_2} \tag{1}$$

with real coefficients and positive determinant $a_1b_2-b_1a_2$ *is again an R function.*

THEOREM 1.8. *A real rational function is an R function if*

a) *the degree of its numerator does not exceed that of the denominator by more than one,*

b) *its value remains below a finite number as* $z\to-\infty$,

c) *its poles are real and their residues negative.*

The rational function in question can be written in the form

$$R(z) = \alpha z + \beta + \sum\left(\frac{\overset{2}{\gamma_\mu}}{Z_\mu - z} - \frac{\overset{2}{\gamma_\mu}}{Z_\mu}\right) \tag{2}$$

the number of terms in the sum being the degree of the denominator. Because of the second condition, $\alpha \geqq 0$, because of the last condition Z_μ and γ_μ are real; β is real also. Hence, by Theorem 1.6 and Theorem 1.7, all terms of (2) are R functions and it is an R function by Theorem 1.5.

THEOREM 1.9. *The series* (2), *finite or infinite, with positive or vanishing* α, *real* β, γ *and* Z *is an* R *function if the* Z_μ *have no finite condensation point and if the series on the right converges at least for one value of* $z \neq 0$. *In this case, it converges absolutely and uniformly in every closed domain of* z *which excludes the poles* Z.

In fact, the general term of the series in (2) has the form

$$\frac{\gamma^2}{Z - z} - \frac{\gamma^2}{Z} = \left(\frac{\gamma}{Z}\right)^2 z\left(1 + \frac{z}{Z - z}\right)$$

which shows: if the series converges for a $z \neq 0$, the series with the positive general term $(\gamma/Z)^2$ converges too, and so the series in (2) converges absolutely and uniformly in every closed domain excluding the poles Z, and represents, by Weierstrass' theorem, an analytic function. This function is clearly meromorphic and an R function by Theorems 1.8 and 1.1.

We now come to the expansion theorems for R functions.

THEOREM 1.10. *Every* R *function can be expanded in the form* (2).

In order to prove this without the use of the properties of the elementary transcendentals, we first establish the following Lemma due to M. Schiffer and V. Bargmann.§

LEMMA 1.10a. *Let* $R(z)$ *be an* R *function,* Z_1 *one of its poles and* $\gamma_1^2/(Z_1 - z)$ *its meromorphic part at* Z_1. *Then*

$$R_1(z) = R(z) - \overset{2}{\gamma_1}/(Z_1 - z) \tag{3}$$

is also an R *function.*

The proof of this theorem, due to Schiffer and Bargmann, first observes that R_1 is clearly meromorphic. If its imaginary part assumed a negative value $-\rho$ at a point z' of the upper half-plane, one could draw a circle about Z_1 with a radius larger than $2\gamma_1^2/\rho$ and large enough to contain z'. The imaginary part of R_1 is larger than $-\frac{1}{2}\rho$ along the part of this circle which lies in the upper half-

§ Cf. E. P. Wigner, Ann. of Math. 53, 36 (1951). Only the last part of the present proof, beginning with lemma 1.10b, is new.

plane because the imaginary part of R is positive and the absolute value of $\gamma_1^2/(Z_1-z)$ is smaller than $\frac{1}{2}\rho$ along this semicircle. Furthermore, R_1 is real at every point of regularity of the real axis, including Z_1. Finally, if Z_μ is a pole of R within the circle, one can draw a sufficiently small semicircle about Z_μ into the upper half-plane so that the imaginary part of R_1 becomes positive thereon. In order to demonstrate this, one observes that, by Theorem 1.4, Z_μ is a simple pole of R_1. Hence there is a neighborhood $|\zeta|<\epsilon$ of Z_μ with $\zeta=\zeta_1+i\zeta_2=Z_\mu-z$ where $R_{1\mu}=R_1-\gamma_\mu^2/\zeta$ is regular. Furthermore $R_{1\mu}$ remains real on the real axis so that, by Rolle's theorem, the imaginary part of $R_{1\mu}$ for $|\zeta|<\epsilon$, $\zeta_2>0$ is larger than $M\zeta_2$ where M is the minimum of the derivative of $R_{1\mu}$ in the direction of the imaginary axis. Hence the imaginary part of R_1 is, in the same region, larger than $(M+\gamma_\mu^2/|\zeta|^2)\zeta_2$. The bracket of this expression can however be made positive by choosing $|\zeta|^2$ small enough. Thus z' can be surrounded by a continuous path (*cf.* Figure 2) containing no singularities of R_1, on which the imaginary part of R_1 is everywhere larger than $-\frac{1}{2}\rho$. Hence the imaginary part of R_1 cannot be negative in the upper half-plane and a similar argument shows that it cannot be positive in the lower half-plane. Thus R_1 is an R function.

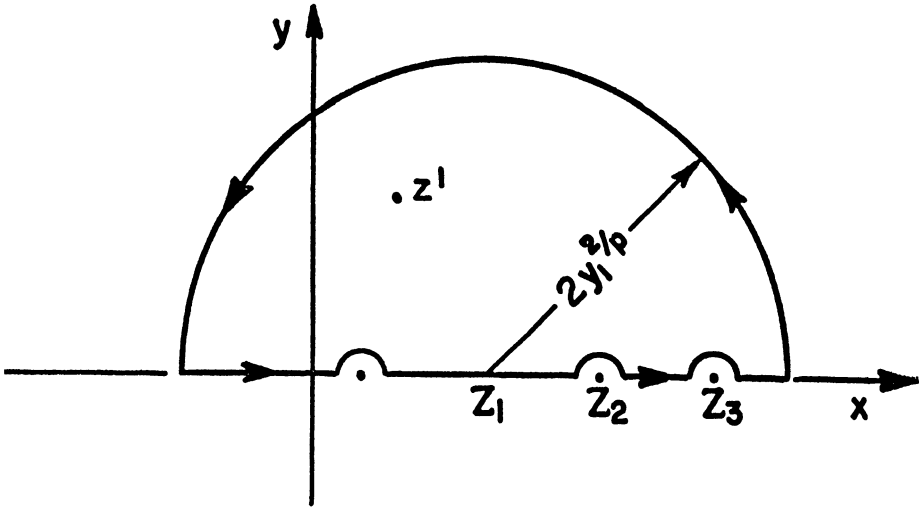


FIG. 2

We can subtract in a similar way any finite number of singularities from R, and the resulting function

$$R_n = R - \sum[\gamma_\mu^2/(Z_\mu - z) - \gamma_\mu^2/Z_\mu] \tag{4}$$

will remain an R function. Hence, its derivative at a real point of regularity z_0 will be positive by Theorem 1.3:

$$\sum\gamma_\mu^2/(Z_\mu - z_0)^2 < R'(z_0). \tag{5}$$

It follows that the whole series $\gamma_\mu^2/(Z_\mu - z_0)^2$ converges if z_0 is real. However, since $|\gamma_\mu^2/(Z_\mu - x - iy)^2| < \gamma_\mu^2/(Z_\mu - x)^2$, the series (5) converges absolutely also for complex z_0. The same is true of the series

$$\left|\frac{\gamma_\mu^2}{Z_\mu - z_0} - \frac{\gamma_\mu^2}{Z_\mu}\right| = |z_0| \sum \left|\frac{\gamma_\mu^2}{(Z_\mu - z_0)Z_\mu}\right|. \tag{6}$$

Except for a finite number of terms, its terms are smaller than $2|z_0|$ times the corresponding terms of (5). Hence, the conditions of Theorem 1.9 are fulfilled and R_n converges, as we subtract all the poles, by the second part of Theorem 1.9, to an analytic function $I(z)$. This $I(z)$ can have no finite singularity. As a limiting case of R functions, $I(z)$ is itself an R function. In order to prove Theorem 1.10, it remains to be shown only that such a function must be linear. This follows from Lemma 1.10b:

LEMMA 1.10b: *An entire R function $I(z)$ is a linear function $\alpha z + \beta$, with real β, $\alpha \geqq 0$.*

Because of Theorem 1.3, unless $I(z)$ is a constant, it is for real z a monotonically increasing function. Let us denote its value at $z=0$ with β and consider the function

$$J(z) = 1/(\beta - I(z)). \tag{7}$$

This is, by Theorem 1.7, an R function, negative for positive z, positive for negative z. It has only one singularity, at $z=0$. Let us subtract this singularity by Theorem 1.10a. The result

$$J_1(z) = J(z) + \gamma^2/z \tag{8}$$

is again an entire R function. It must be a constant unless it is monotonically increasing. Since $J(z)$ is negative for positive z, and conversely, $J_1(\infty) \leqq 0$ and $J_1(-\infty) \geqq 0$. Hence $J_1(z) = 0$ and $I(z) = \beta + z/\gamma^2$ which proves Theorem 1.10b and hence, Theorem 1.10.

THEOREM 1.11. *Given a set of real numbers x_μ, Z_μ without finite condensation points for which*

$$\cdots < Z_{-2} < x_{-2} < Z_{-1} < x_{-1} < Z_0 < x_0 < Z_1 < x_1 < Z_2 < x_2 < \cdots \tag{9}$$

and $Z_0 \leqq 0 < Z_1$, $c > 0$, the products (finite or infinite) in

$$R(z) = c\frac{x_0 - z}{Z_0 - z} \prod_\nu \frac{1 - z/x_\nu}{1 - z/Z_\nu} \prod_\nu \frac{1 - z/x_{-\nu}}{1 - z/Z_{-\nu}} \tag{9a}$$

converge absolutely and uniformly in every finite domain to an R function.

The ν in both products of (9a), as in all formulae below, runs over all positive integers. If the set (9) breaks off at a finite Z_l, the last factor in the first $\prod$ is $(1 - z/Z_l)^{-1}$; if the set starts with a finite x_s the last factor in the second

$\prod$ is $z-x_s$. The last two rules can be replaced by the convention of adding an $x_l=\infty$ to the set of x in the former and a $Z_s=-\infty$ to the set of Z in the latter case. Hence (9) always begins with a Z and ends with an x if it has a finite beginning or end.

To show Theorem 1.11 it will be necessary to establish the well-known theorem about the convergence of infinite products without the use of the exponential function.

LEMMA 1.11a. *The absolute convergence of* $\prod(1+a_\nu)$ *follows from the convergence of* $\sum|a_\nu|$.

Indeed, it can be shown that any partial product $\prod'(1+a_\nu)$ is closer to 1 than the corresponding sum $\sum|2a_\nu|$ as long as this sum does not exceed 1. We shall prove

$$|\prod'(1+a_\nu)-1|<2\sum'|a_\nu| \tag{10}$$

by induction. It is correct if the partial product contains only one factor. Let us assume that (10) is correct if the partial product contains only $n-1$ factors. It then follows that

$$\begin{aligned}|\prod'(1+a_\nu)(1+a_n)-1| &\leqq |\prod'(1+a_\nu)-1|+|a_n||\prod'(1+a_\nu)| \\ &< 2\sum'|a_\nu|+|a_n|(1+2\sum'|a_\nu|) \\ &\leqq 2(|a_n|+\sum'|a_\nu|).\end{aligned} \tag{11}$$

The last part holds because $\sum'|a_\nu|\leqq\frac{1}{2}$ by assumption. This proves the inequality (10) for n factors and hence in general. The lemma 1.11a is an immediate consequence of (10).

Given now any closed domain, there is only a finite number of Z_μ the absolute value of which is smaller than the maximum value of $2|z|$ in the domain. Apart from the corresponding factors in the products in (9a), $|1-z/Z_\mu|>\frac{1}{2}$ holds for all z of the domain. The product of these factors will converge absolutely and uniformly by Lemma 1.11a if the series

$$\sum|a_\mu|=\sum\left|\frac{1-z/x_\mu}{1-z/Z_\mu}-1\right|=|z|\sum\left|\frac{Z_\mu^{-1}-x_\mu^{-1}}{1-z/Z_\mu}\right|\leqq 2|z|\sum|1/Z_\mu-1/x_\mu| \tag{12}$$

converges. The last series is, however, because of (9), smaller than $2|z/Z_1|$ for the first product and smaller than $2|z/x_{-1}|$ for the second product in (9a). Thus, the product of the factors considered represents, by Weierstrass' theorem, an analytic function in the domain considered and the right side of (9a) is a meromorphic function.

In order to prove that (9a) is an R function, it is, because of Theorem 1.1, necessary only to prove that every finite partial product of (9a) is an R function. These partial products are real rational functions and the degree of their numerators is equal to that of the denominator, unless the series (9) has a beginning

or end, in which case the degrees differ by ± 1. Second, all factors of (9a) tend to 1 as z goes to $\pm\infty$ save the exceptional factors $(1-z/Z_l)^{-1}$ and $z-x_s$, if such are present. At any rate $R(z)$ is bounded from above as $z\to-\infty$. Finally for $\mu\neq 0$, every factor is positive at $z=Z_\mu$, except

$$(1-z/x_\mu)/(1-z/Z_\mu)=Z_\mu(1-z/x_\mu)/(Z_\mu-z)$$

and the numerator of this expression is also positive at $z=Z_\mu$ both for $\mu<0$ and for $\mu>0$. Hence the residues of all poles Z_μ with $\mu\neq 0$ are negative and one can easily see that this holds also for the pole Z_0. It thus follows from Theorem 1.8 that the finite partial products of (9a) are R functions and from Theorems 1.11a and 1.1 that (9a) is also an R function.

THEOREM 1.12. *Every R function $R(z)$ can be expanded§ into a product of the form* (9a).

Because of Theorem 1.3, the poles Z_μ and zeros x_μ satisfy the conditions (9) whence (9a) represents an R function with the same poles and zeros as $R(z)$. Hence Theorem 1.12 will be proved if it can be shown that the ratio $r(z)$ of two R functions with the same poles and zeros is a constant. Since both zeros and poles of an R function must be, by Theorems 1.3 and 1.4, simple, $r(z)$ is an entire function. It cannot assume a negative real value on the real axis since both R functions are negative in the same (Z_μ, x_μ) intervals, positive elsewhere. It cannot assume a negative real value on the upper half-plane either since the imaginary parts of both R functions are positive there. For the same reason r cannot be negative on the lower half-plane. However, *an entire function $r(z)$ which does not assume any value of a finite interval (r_1, r_2) is necessarily a constant.* One sees this most easily by considering $\rho=((r-r_1)/(r-r_2))^{1/2}$. Since the radicand of this expression assumes negative real values only if r is in the interval (r_1, r_2), and since $r(z)$ does not assume values in this interval, ρ is a single-valued function which does not assume any purely imaginary value. Its real part is either positive for every z, or negative for every z. Application of Liouville's theorem to $1/(\rho+i)$ in the former and $1/(\rho-i)$ in the latter case shows that it is a constant. This then proves Theorem 1.12.

This result about the relation of two R functions which have the same poles and zeros can be generalized to the case when two other real values, a and $b>a$, are assumed by both R functions at the same two sets of values of z. By applying the above result to $(R-a)/(b-R)$ one finds:

THEOREM 1.13.* *If two R functions assume the real value a at the same values of z, and if this is true for another real number b, the two R functions can be obtained from each other by a fractional linear transformation which leaves a and b invariant.*

§ Under certain more restrictive assumptions, this theorem has been shown already by K. Fränz, *l.c.* footnote p. 670.

We now come to a couple of further theorems which are not necessary for the development of Section 2. They give the connection between differential equations and R functions. Let us consider the function $\phi(\xi, z)$ of the complex variable z and real variable ξ which obeys the differential equation

$$\frac{d^2\phi}{d\xi^2} = U(\xi)\frac{d\phi}{d\xi} + V(\xi)\phi - z\phi \tag{13}$$

in which U, V are real continuous functions of ξ in the interval $(0, \xi_0)$. Given z-independent initial conditions at $\xi=0$, it is well known that† ϕ and $d\phi/d\xi$, as functions of z, are entire functions.

THEOREM 1.14.* *The solution $R(\xi, z)$ of Riccati's equation*

$$\frac{dR(\xi, z)}{d\xi} = R^2 + UR - V + z \tag{14}$$

with the initial condition $R(0, z) = R_0$ and real R_0 is, for every value ξ of the interval $(0, \xi_0)$, an R function of z.

One obtains (14) from (13) by the substitution $R = -(d\phi/d\xi)/\phi$. As quotient of two entire functions, R is meromorphic. Let us separate $R = R_r + iR_s$ into real and imaginary parts. The equation for R_s is

$$dR_s/d\xi = 2R_rR_s + UR_s + y \tag{14a}$$

where y is the imaginary part of $z = x + iy$. It is necessary to show now that $R_s(\xi)$ has the same sign as y.

Because of $R_s(0, z) = 0$, this is evident for sufficiently small values of ξ. Furthermore, if y is positive, R_s has a positive slope at a point where $R_s = 0$. Hence, as ξ increases, R_s cannot become negative as long as R obeys the differential equation (14). Under the same condition, R_s cannot become positive if y is negative.

Naturally, R can have singularities and will not obey (14) at these. In order to prove that R_s remains positive for positive y as ξ passes through a singularity ξ_0, we note that $Q = -1/R$ vanishes at, and is regular in the neighborhood of, ξ_0. Hence, in the neighborhood of ξ_0, the imaginary part Q_s of $Q = Q_r + iQ_s$ vanishes more strongly than its derivative. It follows that we can set $Q_s = 0$ when calculating $dQ_s/d\xi$ in the neighborhood of ξ_0. The differential equation for Q_s then shows that $dQ_s/d\xi$ has, in the neighborhood of ξ_0, the same sign as yQ_r^2, *i.e.* as y. According to the preceding paragraph, this is true also for $Q_s = R_s|Q|^2$ as long as $\xi < \xi_0$ so that Q_s, and hence also R_s, retain their sign as ξ passes

† *Cf. e.g.* E. L. Ince, Ordinary Differential Equations. Dover Publications, New York, 1944. Section 3.31, p. 72.

through ξ_0. This then proves‡ Theorem 1.14.

Theorem 1.14 permits the definition of a wide variety of R functions and, of course, the R functions derive their significance from Theorems 1.14 and 1.15. However, the latter will not be proved here.§

In order to formulate Theorem 1.15, we consider an n dimensional real euclidean space E_n with rectangular coordinates $\xi_1, \xi_2, \cdots, \xi_n$. There are two regular closed surfaces S_1 and S in E_n, the former being entirely inside S. Given further a self-adjoint real operator§§ H defined for functions within S and for which

$$Hf = -\Delta f = -(\partial^2/\partial\xi_1^2 + \partial^2/\partial\xi_2^2 + \cdots + \partial^2/\partial\xi_n^2)f \tag{15}$$

holds if f vanishes within S_1 and belongs to the definition domain of the Laplace operator Δ. This operator is defined to be self-adjoint also and applicable to all twice continuously differentiable functions in S which have zero normal derivatives thereon. Let us consider then a solution of the equation

$$H\phi(\xi, z) = z\phi(\xi, z) \tag{16}$$

such that ϕ is, on the surface S, real and independent of z

$$\phi(\xi, z) = \psi(\xi) \qquad \text{on } S. \tag{16a}$$

THEOREM 1.15.* *Under the conditions of the preceding paragraph the surface integral*

$$-\int \psi(\xi)\, \mathrm{grad}_n\, \phi(\xi, z)dS = R(z) \tag{17}$$

extended over the surface S is an R function of z; the grad_n *denotes the component of* grad ϕ *that is normal to the surface S.*

For the sake of completeness, three further and rather obvious criteria will be given here which suffice, in conjunction with Theorem 1.14, to prove that the logarithmic derivative of the Bessel functions and of a few other functions, including the Γ function, are R functions and hence permit partial fraction and product representations of the form (2) and (9a).

THEOREM 1.16.* *$aR(bz+c)$ is an R function of z if a, b, c are real and ab positive, and if $R(z)$ is an R function.*

‡ This theorem, although in a different form, has been known for a long time. *Cf. e.g.* L. Bieberbach, *Theorie der Differentialgleichungen*, Julius Springer, Berlin 1923, Section 2, Chapter III, p. 161.

§ For a plausibility argument cf. E. P. Wigner and L. Eisenbud, reference 1.

§§ For the definition of self-adjoint operators *cf. e.g.* B.v.Sz. Nagy, *Spektraldarstellung Linearer Transformationen des Hilbertschen Raumes*. Vol. 5 of *Ergebnisse der Mathematik und Grenzgebiete*. Springer Verlag, Berlin 1942 Section V4, p. 34.

THEOREM 1.17.* *Let $R_+(z)$ be an R function all the poles of which are positive, then*

$$R_0(z) = \frac{1}{z}\left[R_+(z^2) - R_+(0)\right] \tag{18}$$

is an odd R function. If $R_+(z) > 0$ for all negative z

$$R_0(z) = z\,R_+(z^2) \tag{18a}$$

is also an odd R function.

The first part follows from the expansion (2) for R_+ in which Z_μ can be replaced by $\zeta_\mu^2 = Z_\mu$ with positive ζ_μ. The identity

$$\frac{1}{z}\left(\frac{1}{\zeta_\mu^2 - z^2} - \frac{1}{\zeta_\mu^2}\right) = \frac{1}{2\zeta_\mu^2}\left(\frac{1}{\zeta_\mu - z} + \frac{1}{-\zeta_\mu - z}\right)$$

together with Theorem 1.9 then directly gives (18). In the second case, $-1/R_+(z)$ satisfies the conditions of the first case for R_+. Hence $-(1/z)(1/R_+(z^2) - 1/R_+(0))$ is an R function and remains one if $-R_+(0)/z$ is added to it. Hence (18a) is also an R function.

THEOREM 1.18.* *Conversely, if $R_0(z)$ is an odd R function which is regular at $z=0$, then $\sqrt{z}R_0(\sqrt{z})$ is also an R function of z.*

2. The elementary transcendentals. Let us consider the class t of R functions which (a) are periodic with period 1, (b) have only one singularity for $-\frac{1}{2} < z \leqq \frac{1}{2}$, (c) converge to i as z goes to $i\infty$. The existence of R functions satisfying these conditions will be apparent when the expansion (2) will be found for them. We now proceed to find this expansion.

Let us denote the position of the singularity in the $-\frac{1}{2} < z \leqq \frac{1}{2}$ interval by $\frac{1}{2} - x_0$, its residue by $-c < 0$. Because of the periodicity condition, there will be similar singularities at all points $n - \frac{1}{2} - x_0$, the n being any integer, and only at these points. Hence, if R functions of the class t exist, their expansion by (2) must be of the form

$$t(z) = \alpha z + \beta + c\sum\left(\frac{1}{n - \frac{1}{2} - x_0 - z} - \frac{1}{n - \frac{1}{2} - x_0}\right), \tag{19}$$

where n assumes every integral value from $-\infty$ to ∞.

It is worth while to remark that because of the absolute convergence of the series (19) one can unite the terms $\nu = n$ and $\nu = 1 - n$; ν then will run over all positive integers. Then

$$
\begin{aligned}
t(z) &= \alpha z + \beta + \sum\left(\frac{c}{\nu - \frac{1}{2} - x_0 - z} - \frac{c}{\nu - \frac{1}{2} + x_0 + z}\right) \\
&\quad - \sum\left(\frac{c}{\nu - \frac{1}{2} - x_0} - \frac{c}{\nu - \frac{1}{2} + x_0}\right) \qquad \text{(19a)} \\
&= \alpha z + \beta' + \sum\left(\frac{c}{\nu - \frac{1}{2} - x_0 - z} - \frac{c}{\nu - \frac{1}{2} + x_0 + z}\right),
\end{aligned}
$$

since the sums of both brackets of the first equation converge absolutely and the second of these is independent of z.

In order to impose the boundary condition $t(i\infty)=i$, we calculate the sum in (19a) for $z=x+iy$

$$
\begin{aligned}
&\sum \frac{c}{\nu + \frac{1}{2} - x_0 - z} - \frac{c}{\nu + \frac{1}{2} + x_0 + z} \\
&\qquad = \sum \frac{2c(x_0 + x + iy)}{(\nu + \frac{1}{2})^2 - (x_0 + x + iy)^2} \qquad \text{(19b)} \\
&\qquad = \frac{2ic}{y} \sum \frac{1 - i(x_0 + x)/y}{((\nu + \frac{1}{2})/y)^2 + (1 - i(x_0 + x)/y)^2}.
\end{aligned}
$$

The factor by which the νth term of (19b) differs from $[1+((\nu+\frac{1}{2})/y)^2]^{-1}$ converges to 1 as y goes to infinity. Hence the sum (19b) converges to the sum of these expressions and, by the definition of the integral, (19b) converges to

$$
\frac{2ic}{y} \sum \frac{1}{1 + (\nu + \frac{1}{2})^2/y^2} \to 2ic \int_0^\infty \frac{d\xi}{1 + \xi^2} = i\pi c \qquad \text{(19c)}
$$

if we denote the last integral by $\frac{1}{2}\pi$.

Since t converges to a fixed value i as $z \to i\infty$ and since this is true also of the last term in (19a), $\alpha z+\beta'$ must also converge to a fixed number. This is possible only if $\alpha=0$ and since β' is real, it also must vanish. It then follows that $\pi c=1$.

THEOREM 2.1. *The R functions of the class t have the form $t(x_0+z)$ with real x_0 where*

$$
t(z) = \frac{1}{\pi} \sum\left(\frac{1}{\nu - \frac{1}{2} - z} - \frac{1}{\nu - \frac{1}{2} + z}\right) = \frac{1}{\pi} \sum\left(\frac{1}{n - \frac{1}{2} - z} - \frac{1}{n - \frac{1}{2}}\right) \qquad \text{(20)}
$$

where ν runs over all positive integers, n over all integers.

It is easy to convince oneself that the t given by (20) fulfills the three conditions of the definition of the class. This is evident for (b) and the above calculation verifies (c). To verify (a) let N be so large that the first N terms of the series (20) represent both $t(z)$ and $t(z+1)$ with a smaller error than ϵ, then (20) shows that $t(z)-t(z+1)$ is within 2ϵ equal to

$$\frac{1}{N-\frac{1}{2}-z}+\frac{1}{N+3/2+z]}$$

and this goes to 0 as N increases.

The function $t(z)$ has singularities at $\pm\frac{1}{2}$ and assumes in the interval $(-\frac{1}{2},\frac{1}{2})$ every real value. Let us consider therefore the R functions

$$\frac{\alpha t(z)+\beta}{-\beta t(z)+\alpha} \tag{21}$$

with real α, β (*cf.* Theorem 1.7). The function (21) is clearly periodic, has only one singularity for $-\frac{1}{2}<z\leqq\frac{1}{2}$, namely at the point where $t(z)=\alpha/\beta$, and one can verify directly that its value is i for $z=i\infty$. It therefore is in the class t and is equal to $t(x_0+z)$ with a definite x_0. In order to determine this x_0, one may insert $z=0$ in (21), because of $t(0)=0$, its value then is $\beta/\alpha=t(x_0)$. Introducing this into (21) gives the addition theorem

$$t(x_0+z)=\frac{t(z)+t(x_0)}{1-t(x_0)t(z)}. \tag{21a}$$

The above derivation proves (21a) only for $-\frac{1}{2}<x_0\leqq\frac{1}{2}$. However, since both sides of (21a) are analytic functions of x_0, it must hold also in general.

The function $t(z)$ is odd, so vanishes at $z=0$. In order to calculate the value of its derivative t' at $z=0$, we note from (21a) with $x_0=1/2$ that

$$t(\tfrac{1}{2}+z)=-1/t(z). \tag{22}$$

The residue of the left side of (22) at $z=0$ is $-1/\pi$, hence the derivative of $t(z)$ at $z=0$ is

$$t'(0)=\pi. \tag{22a}$$

From the odd and real nature of t it follows that it is imaginary on the imaginary axis and, because of (22), it is also imaginary if z has the form $\frac{1}{2}+iy$.

The poles of $-1/t(z)$ are all the integers. From this and its odd nature follows the expansion

$$-\frac{1}{t(z)}=-\frac{1}{\pi z}+\frac{1}{\pi}\sum\left(\frac{1}{\nu-z}-\frac{1}{\nu+z}\right). \tag{23}$$

Let us calculate, finally, the addition formula for an arbitrary fractional linear transform of t

$$e(z)=\frac{a_1t(z)+b_1}{a_2t(z)+b_2}.$$

For $e(x_0+z)$ one obtains a fractional linear function of $e(z)$ with the matrix

$$\begin{Vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{Vmatrix} \begin{Vmatrix} 1 & t(x_0) \\ -t(x_0) & 1 \end{Vmatrix} \begin{Vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{Vmatrix}^{-1}.$$

This will be a diagonal matrix if $a_1=-a_2=i$; $b_1=b_2=1$. Hence the addition theorem for

$$e(z) = \frac{it(z)+1}{-it(z)+1} \tag{24}$$

has the form $e(x_0+z)=f(x_0)e(z)$. Writing $z=0$ herein one finds $f(x_0)=e(x_0)$ and

$$e(z+z') = e(z)e(z'). \tag{24a}$$

Evidently, $e(z)$ is periodic with period 1. From (24) one finds $e(0)=1$, $e(i\infty)=0$, $e(-i\infty)=\infty$, that $|e(z)|=1$ for real z and from the remark after (22a) that it is real for imaginary z: it assumes every positive value as z drops from $i\infty$ to $-i\infty$.

The value of the derivative e' of e at $z=0$ can be obtained from (24) and (22a): $e'(0)=2it'(0)=2\pi i$. The differential equation

$$e'(z) = 2\pi i e(z) \tag{25}$$

then easily follows from (24a). It shows that $e(z)$ can have no poles and is an entire function. The various formulae for the trigonometric and hyperbolic functions follow either from (21a) or (24a) by means of

$$\begin{aligned} t(z) &= \tan \pi z, \\ e(z) &= e^{2\pi i z}. \end{aligned} \tag{6}$$

The infinite product representation for the tangent is a direct consequence of Theorem 1.12:

$$\begin{aligned} t(z) &= \frac{\pi}{2}\,\frac{z}{z+\frac{1}{2}} \prod \frac{1-z/\nu}{1-z/(\nu-\frac{1}{2})} \prod \frac{1+z/\nu}{1+z/(\nu+\frac{1}{2})} \\ &= \frac{\cdots(1-z/2)(1-z)z(1+z)(1+z/2)\cdots}{\cdots(1-2z/3)(1-2z)(1+2z)(1+2z/3)(1+2z/5)\cdots}. \end{aligned} \tag{27}$$

Calculating $t'(z)/t(z)$, one obtains the partial fraction representation of $s(2z)$ and, hence, of $s(z)$ and $c(z)$. The most obvious way to obtain the infinite product representation for $c(z)=\cos \pi z$ is to show that it satisfies the differential equation $c'(z)/c(z)=-\pi t(z)$. The infinite product for the sine can be deduced either in the same fashion or by multiplying the infinite products for $c(z)$ and $t(z)$.

It is worth while yet to point out that every complex number $x+iy$ can be expressed in the form $re(\phi)$ with real r and ϕ. Set $r=(x^2+y^2)^{1/2}$ and $t(\phi)=(r-x)/y=y/(r+x)$; there is one and only one ϕ between $-\frac{1}{2}$ and $\frac{1}{2}$ which satisfies this last equation. This then shows that $\pi\phi$ is the angle in the Gauss plane. The remark after (24a) shows, furthermore, that there is an imaginary

number $i\rho$ such that $r=e(i\rho)$ if r is positive. Hence, unless $x+iy=0$, there is a complex $\phi+i\rho$ such that $e(\phi+i\rho)=x+iy$. This defines the logarithm; its properties, including its derivative, follow from the properties of the exponential function.

One can also easily see that the π defined in (19c) is the ratio of the circumference of a circle to its diameter. For this purpose one extends the integral from $-\infty$ to ∞ and deforms the path of integration to a small circle about i; the integrand $dz/(z-i)(z+i)$ is easily seen to approach ds/r as the radius of the circle contracts.

On the Contraction of Groups and Their Representations

E. Inönü and E. P. Wigner

Proceedings of the National Academy of Sciences (USA) *39*, 510–524 (1953)

Communicated April 21, 1953

Introduction.—Classical mechanics is a limiting case of relativistic mechanics. Hence the group of the former, the Galilei group, must be in some sense a limiting case of the relativistic mechanics' group, the representations of the former must be limiting cases of the latter's representations. There are other examples for similar relations between groups. Thus, the inhomogeneous Lorentz group must be, in the same sense, a limiting case of the de Sitter groups. The purpose of the present note is to investigate, in some generality, in which sense groups can be limiting cases of other groups (Section I), and how their representations can be obtained from the representations of the groups of which they appear as limits (Section II). Section III deals briefly with the transition from inhomogeneous Lorentz group to Galilei group. It shows in which way the representation up to a factor of the Galilei group, embodied in

the Schrödinger equation, appears as a limit of a representation of the inhomogeneous Lorentz group and also gives the reason why no physical interpretation is possible for the real representations of that group.

I. CONTRACTION OF GROUPS

Let us consider an arbitrary Lie group with n parameters a^i and infinitesimal operators I_i. These shall be given, as usual by

$$I_i = \lim_{h \to 0} \frac{g(he_i) - g(0)}{h} \tag{1}$$

where 0 are the parameters of the unit element and e_i differs from 0 by a unit increase of a_i. The I_j are skew hermitean matrices if the group consists of unitary matrices; instead of them one often uses the hermitean quantities iI_j. However, all our equations remain somewhat simpler if expressed in terms of the I_j. The structure constants C are defined by

$$[I_i, I_j] = \sum_k C_{ij}^k I_k \tag{2}$$

If we subject the I_i to a linear homogeneous non-singular transformation, the C will be replaced by other constants. These are obtained from the C by contragradient transformations of its upper and lower indices. However, such a transformation has, naturally, no effect on the structure of the group. Let us denote the transformation in question by

$$J_\nu = \sum_i I_i U^i{}_\nu. \tag{3}$$

It corresponds to the transformation

$$a^i = \sum_\kappa U^i{}_\kappa b^\kappa \tag{3a}$$

according to which the J are obtained by the same equation (1) as the I, except that the e_i have to be replaced in it by a similar quantity, defined with respect to the b.

The above transformation may lead to a new group only if the matrix U of (3) is singular. We shall call the operation of obtaining a new group by a singular transformation of the infinitesimal elements of the old group a *contraction* of the latter. The reason for this term will become clear below. The singular matrix will be a limiting case of a non-singular matrix. The latter will depend linearly on a parameter ϵ which will tend to zero:

$$U^i{}_\nu = u^i{}_\nu + \epsilon w^i{}_\nu. \tag{4}$$

For $0 < \epsilon < \epsilon_0$ the determinant of (4) is different from zero, it vanishes for $\epsilon = 0$.

We shall transform (4) into a normal form by a non-singular and ϵ

independent transformation of both the J_ν and I_i. If the matrices of the corresponding transformations are denoted by α and β, respectively, u and w will be replaced by $\beta u \alpha^{-1}$ and $\beta w \alpha^{-1}$. It is possible, by such a transformation, to give u and w the form

$$u = \begin{Vmatrix} 1 & 0 \\ 0 & 0 \end{Vmatrix} \qquad w = \begin{Vmatrix} v & 0 \\ 0 & 1 \end{Vmatrix}. \tag{5}$$

The number of rows and columns in the unit matrix in u, and in v, is equal to the rank r of u. It is advantageous to label the transformed J and I with a pair of indices, the first referring to the subdivision of u given in (5), the second specifying the various J and I within that subdivision. Hence, (3) assumes the form

$$\begin{aligned} J_{1\nu} &= I_{1\nu} + \epsilon \sum_{\mu=1}^{r} v_{\mu\nu} I_{1\mu} \qquad && (\nu = 1, 2, \ldots, r) \\ J_{2\nu} &= \epsilon I_{2\nu}. && (\nu = 1, 2, \ldots, n - r) \end{aligned} \tag{6}$$

The corresponding transformation of the group parameters is

$$\begin{aligned} a_{1\nu} &= b_{1\nu} + \epsilon \sum_{\mu=1}^{r} v_{\nu\mu} b_{1\mu} \qquad && (\nu = 1, 2, \ldots, r) \\ a_{2\nu} &= \epsilon b_{2\nu}. && (\nu = 1, 2, \ldots, n - r) \end{aligned} \tag{6a}$$

It is well to remember that the parameters a lead to the infinitesimal elements I, the parameters b to the J. The last equation shows that a given set of parameters b correspond, with decreasing ϵ, to smaller and smaller values of the parameters a_2. In the limit $\epsilon = 0$ (if such a limit exists), one will have contracted the whole group to an infinitesimally small neighborhood of the group defined by the $a_{1\nu}$ alone. This justifies the name given to the process considered.

The transformation of the infinitesimal elements which we carried out also changes the structure constants and we shall write for (2)

$$[I_{\alpha\nu}, I_{\beta\mu}] = \sum_{\kappa=1}^{r} C_{\alpha\nu,\,\beta\mu}{}^{1\kappa} I_{1\kappa} + \sum_{\kappa=1}^{n-r} C_{\alpha\nu,\,\beta\mu}{}^{2\kappa} I_{2\kappa} \tag{7}$$

wherein α and β can assume the values 1 and 2. This gives for

$$\begin{aligned} [J_{1\nu}, J_{1\mu}] &= [I_{1\nu}, I_{1\mu}] + \epsilon \sum (v_{\nu\nu'}\delta_{\mu\mu'} + \delta_{\nu\nu'} v_{\mu\mu'} + \epsilon v_{\nu\nu'} v_{\mu\mu'}) [I_{1\nu'}, I_{1\mu'}] \\ &= \sum_{\kappa} C_{1\nu,\,1\mu}{}^{1\kappa} J_{1\kappa} + \frac{1}{\epsilon} \sum_{\kappa} C_{1\nu,\,1\mu}{}^{2\kappa} J_{2\kappa} + 0(1). \end{aligned}$$

Hence, if the commutator of $J_{1\nu}$ and $J_{1\mu}$ is to converge, as $\epsilon \to 0$, to a linear combination of the J, the structure constants

$$C_{1\nu,\,1\mu}{}^{2\kappa} = 0 \tag{8}$$

i.e., the $I_{1\nu}$ must span a subgroup. On the other hand, if this happens to be the case, the structure constants will converge to definite values $c_{\alpha\nu,\,\beta\mu}{}^{\gamma\kappa}$ as $\epsilon \to 0$

$$\begin{aligned} c_{1\nu,\,1\mu}{}^{1\kappa} &= C_{1\nu,\,1\mu}{}^{1\kappa} & c_{1\nu,\,1\mu}{}^{2\kappa} &= C_{1\nu,\,1\mu}{}^{2\kappa} = 0 \\ c_{1\nu,\,2\mu}{}^{1\kappa} &= 0 & c_{1\nu,\,2\mu}{}^{2\kappa} &= C_{1\nu,\,2\mu}{}^{2\kappa} \\ & & c_{2\nu,\,2\mu}{}^{1\kappa} &= c_{2\nu,\,2\mu}{}^{2\kappa} = 0. \end{aligned} \tag{9}$$

These structure constants satisfy Jacobi's identities since the structure constants for the J do this for non-vanishing ϵ. We shall say that the above operation is a contraction of the group with respect to the infinitesimal elements $I_{1\nu}$ or that the infinitesimal elements $I_{2\mu}$ are contracted. We then have, from (9).

THEOREM 1. *Every Lie group can be contracted with respect to any of its continuous subgroups and only with respect to these.* The subgroup with respect to which the contraction is undertaken will be called S. *The contracted infinitesimal elements form an abelian invariant subgroup of the contracted group. The subgroup S with respect to which the contraction was undertaken is isomorphic with the factor group of this invariant subgroup. Conversely, the existence of an abelian invariant subgroup and the possibility to choose from each of its cosets an element so that these form a subgroup S, is a necessary condition for the possibility to obtain the group from another group by contraction.*

It is easy to visualize now the effect of the contraction on the whole group. The subgroup S with respect to which the contraction is undertaken remains unchanged and it is advantageous to choose the group parameters in such a way that $a_{2\nu} = 0$ throughout S. Then (6a) can be replaced by

$$a_{1\nu} = b_{1\nu} \qquad a_{2\nu} = \epsilon b_{2\nu} \tag{6b}$$

and this can be assumed to be valid throughout the whole group, not only in the neighborhood of the unit element. As ϵ decreases, a fixed range of the parameter b will describe an increasingly small surrounding of S. As ϵ tends to 0, the range of the $b_{2\nu}$ will become infinite and describe only those group elements which differ infinitesimally from the elements of S. The elements which are in the neighborhood of the unit element of the original group but have finite parameters $b_{2\nu}$ will commute and form the aforementioned commutative invariant subgroup. Naturally, the elements of this invariant subgroup will not commute, in general, with the elements of the subgroup S: the change of the parameters $a_{2\nu} = \epsilon b_{2\nu}$ will be, upon transformation by finite elements of S, of the same order of magnitude as these parameters themselves. Naturally, the convergence

of the original group toward the contracted group is a typically non-uniform convergence.

Every Lie group can be contracted with respect to any of its one parametric subgroups. If the three dimensional rotation group is contracted in this way, one obtains the Euclidean group for two dimensions. Contraction of the homogeneous Lorentz group with respect to the subgroup which leaves the time coordinate invariant yields the homogeneous Galilei group, contraction of the inhomogeneous Lorentz group with respect to the group generated by spatial rotations and time displacements yields the full Galilei group. Contraction of the de Sitter groups yields the inhomogeneous Lorentz group. It should be remarked, finally, that if a group, obtained by contraction of another group with respect to the subgroup S, is contracted again with respect to S, the second contraction remains without effect.

The above considerations show a certain similarity with those of I. E. Segal.[1] However, Segal's considerations are more general than ours as he considers a sequence of Lie groups the structure constants of which converge toward the structure constants of a non-isomorphic group. In the above, we have considered only one Lie group but have introduced a sequence of coordinate systems therein and investigated the limiting case of these coordinate systems becoming singular. As a result of our problem being more restricted, we could arrive at more specific results.

II. Contraction of Representations

If one applies the transformation (6) to the infinitesimal elements of a representation of the group to be contracted, and lets ϵ tend to zero, the $J_{2\nu}$ will also tend to zero. The representation will become isomorphic to the representation of the subgroup S, i.e., will be a representation of the factor group of the invariant subgroup. In order to obtain a faithful representation, one must *either subject the $J_{2\nu}$ to an ϵ dependent transformation, or consider the $J_{2\nu}$ which correspond to different representations*, e.g., go to higher and higher dimensional representations as ϵ decreases. We shall give examples for both procedures.

(*a*) *Representations of the Contracted Group by Means of ϵ Dependent Transformations.*—The first procedure is applicable only if the infinitesimal elements are not bounded operators, i.e., as far as irreducible representations are concerned, only if the group is not compact. The simplest non-compact non-commutative group is that of the linear transformations $x' = e^{\alpha}x + \beta$. The general group element is $O_{\alpha,\beta} = T(\beta)R_{\alpha}$ with the group relations $T(\beta)T(\beta') = T(\beta + \beta')$; $R_{\alpha}R_{\alpha'} = R_{\alpha+\alpha'}$ and $R_{\alpha}T(\beta) = T(e^{\alpha}\beta)R_{\alpha}$. The only faithful irreducible unitary representation of this group can be given in the Hilbert space of square integrable functions of $0 < x < \infty$

$$R_\alpha\psi(x) = e^{(1/2)\alpha}\psi(e^\alpha x) \qquad T_\beta\psi(x) = e^{i\beta x}\psi(x). \tag{10}$$

The infinitesimal operators are

$$I_\alpha = I_1 = {}^1/_2 + x\frac{d}{dx} \qquad I_\beta = I_2 = ix \tag{10a}$$

with the structure relation

$$[I_1, I_2] = I_2. \tag{10b}$$

Contraction with respect to the group of transformations $x' = e^{\alpha x}$ leaves the group unchanged: the only non-vanishing structure constant is C_{12}^2 and (9) shows that this does not change. We can try, therefore, to transform I_1 and ϵI_2 with an ϵ dependent unitary matrix so that I_1 remain unchanged, the transformed ϵI_2 converge to I_2

$$S_\epsilon^{-1}I_1S_\epsilon = I_1 \qquad \epsilon S_\epsilon^{-1}I_2S_\epsilon \rightarrow I_2. \tag{11}$$

This is indeed possible: one has to choose $S_\epsilon = R_{\ln\epsilon}$. This commutes with I_1. It follows from the group relation that $S_\epsilon^{-1}T(\beta)S_\epsilon = R_{-\ln\epsilon}T(\beta)R_{\ln\epsilon} = T(\beta/\epsilon)$ and one has, hence, as $h \rightarrow 0$

$$\begin{aligned}\epsilon S_\epsilon^{-1}I_2S_\epsilon &= \epsilon S_\epsilon^{-1}\lim h^{-1}(T(h) - 1)S_\epsilon \\ &= \lim \epsilon h^{-1}(T(h/\epsilon) - 1) = I_2\end{aligned} \tag{11a}$$

Hence $\epsilon S_\epsilon^{-1}I_2S_\epsilon$ not only converges to $J_2 = I_2$ but remains equal to it for all ϵ.

It is more surprising, perhaps, to see that the same device is possible also if one contracts the group with respect to the subgroup of the $T(\beta)$. The contracted group is, in this case, the two parametric abelian group. We demand, in this case, that S_ϵ commute with I_β and hence, by (10a), that it be multiplication with a function of x. Because of S' unitary nature, we can give it the form $\exp(if(x, \epsilon))$. Transformation of the ϵI_α of (10a) with this gives

$$\begin{aligned}\epsilon S_\epsilon^{-1}I_\alpha S_\epsilon &= \epsilon e^{-if(x,\epsilon)}({}^1/_2 + x\,d/dx)e^{if(x,\epsilon)} \\ &= \epsilon({}^1/_2 + x\,d/dx) + \epsilon ix\,d/dx\,f(x, \epsilon).\end{aligned} \tag{12}$$

The first part of this converges to 0 as it should since (12) should converge to an operator which commutes with ix. The second part converges to $J_2 = ixf'(x) = ig(x)$ if one sets

$$f(x, \epsilon) = \epsilon^{-1}f(x). \tag{12a}$$

Hence, the transformations of the contracted group corresponding to the parameters α, β is multiplication with

$$e^{i\alpha g(x) + i\beta x} \tag{13}$$

which is again a faithful (reducible) representation of the contracted group.

The operators corresponding to finite group elements could have been obtained directly by transforming $O_{\alpha,\epsilon\beta} = T(\epsilon\beta)R_\alpha$ with $R_{\ln\epsilon}$ in the case of contraction with respect to R_α. Similarly, (13) could have been obtained directly by transforming $O_{\epsilon\alpha,\beta} = T(\beta)R_{\epsilon\alpha}$ with $\exp(\epsilon^{-1}if(x))$.

It is not clear how generally one can obtain a faithful representation of the contracted group as a limit of an ϵ dependent transform of a faithful representation of the original group and the substitution (6b) of its parameters. Certainly, the procedure is not applicable to irreducible representations of compact groups or, more generally, if the infinitesimal operators are bounded.

(*b*) *Representations of the Contracted Group from a Sequence of Representations.*—We shall now give a few examples for the second procedure, i.e., obtaining a representation of the contracted group by choosing a sequence of unitary representations $D^{(1)}, D^{(2)}, \ldots, D^{(l)}, \ldots$ so that each of the operators

$$I^{(l)}_{1\nu},\ \epsilon I^{(l)}_{2\mu} \qquad (\nu = 1, 2, \ldots, r;\ \mu = 1, 2, \ldots, n - r) \tag{15}$$

converge to a finite operator as $\epsilon \to 0$ and $l \to \infty$. Alternately, we can ask that the transformation corresponding to finite group elements

$$D^{(l)}(b_{1\nu}, \epsilon b_{2\mu}) \tag{15a}$$

converge to a unitary representation of the contracted group as $\epsilon \to 0$, $l \to \infty$. The b are the parameters of the contracted group; the operator (15) corresponds, in the representation l, to the group element of the original group the parameters a of which are given by (6b). The formulation making use of the finite group elements (15a) is unambiguous because it deals with the convergence of unitary operators; the first one is usually easier to attack directly.

The convergence of the sequences of (15) and (15a) will depend not only on the values which ϵ assumes and on the corresponding representations $D^{(l)}$, it will also depend in which form that representation is assumed. Hence, method (*a*) can be considered as a special case of the present method in which all $D^{(l)}$ are unitary equivalent.

The contracted group is always an open group because the variability domain of the b is infinite. Hence its representations are, as a rule, infinite dimensional. If the $D^{(l)}$ are finite dimensional, they should be considered to affect only a finite number of the coordinates of Hilbert space.

The simplest non-commutative compact group is the three-dimensional rotation group. All its subgroups are one parametric, we shall contract it with respect to the rotations about the z axis of a rectangular coordinate

system of ordinary space. The contracted group is the Euclidean group of the plane, i.e., the inhomogeneous two-dimensional rotation group. We shall choose for $D^{(l)}$ the representation which is usually denoted[2] by $D^{(l)}$: it is $2l + 1$ dimensional (l being any integer) and is usually described in a space the coordinate axes of which are labeled with $m = -l, -l + 1, \ldots, l - 1, l$. Hence, we label the coordinate axes of Hilbert space with all integers m from $-\infty$ to ∞. In keeping with our previous notation, we call the infinitesimal element which corresponds to rotations about the z axis $M_z = I_1$. One can then write for $|m| \leq l$, $|m'| \leq l$

$$(I_1^{(l)})_{mm'} = (M_z^{(l)})_{mm'} = im\delta_{mm'} \tag{16}$$

$$(I_{2x}^{(l)})_{mm'} = (M_x^{(l)})_{mm'} = -\frac{1}{2}\sqrt{(l-m)(l+m')}\,\delta_{m'm+1} + \frac{1}{2}\sqrt{(l-m')(l+m)}\,\delta_{m'm-1}$$

$$(I_{2y}^{(l)})_{mm'} = (M_y^{(l)})_{mm'} = \frac{i}{2}\sqrt{(l-m)(l+m')}\,\delta_{m'm+1} + \frac{i}{2}\sqrt{(l-m')(l+m)}\,\delta_{m'm-1}.$$

All matrix elements vanish if either $|m|$ or $|m'|$ is larger than l. As $l \to \infty$, the I_1 converge to a definite operator

$$(J_1)_{mm'} = im\,\delta_{mm'}. \qquad (-\infty < m, m' < \infty) \tag{17a}$$

In fact, the convergence is strong in the sense that $I_1\varphi$ converges strongly to $J_1\varphi$ if φ is in the definition domain of J_1. If ϵ and l converge to zero and infinity, respectively, in such a way that $l\epsilon \to \Xi$, the other two infinitesimal elements will also converge to J_{2x}, J_{2y}, respectively, where

$$(J_{2x})_{mm'} = \frac{1}{2}\Xi(\delta_{m'm-1} - \delta_{m'm+1}) \tag{17b}$$

$$(J_{2y})_{mm'} = \frac{i}{2}\Xi\,(\delta_{m'm-1} + \delta_{m'm+1}).$$

It follows from the fact that there are only a finite number (two) nonvanishing matrix elements in both I_{2x} and I_{2y} that the operators satisfy the commutation relations of the contracted group

$$[J_{2x}, J_{2y}] = 0 \qquad [J_1, J_{2x}] = J_{2y} \qquad [J_1, J_{2y}] = -J_{2x}. \tag{18}$$

It further follows from

$$(I_1^{(l)})^2 + (I_{2x}^{(l)})^2 + (I_{2y}^{(l)})^2 = -l(l+1)$$

by multiplication with ϵ^2 and going to the limit in the above way that

$$J_{2x}^2 + J_{2y}^2 = -\Xi^2. \tag{18a}$$

It was important, for obtaining convergent sequences of $I_1^{(l)}$ and $\epsilon I_2^{(l)}$, to have assumed the $D^{(l)}$ in the form given by (16). This form was reduced out with respect to the subgroup S and this caused, in this case, the convergence of the sequence $I_1^{(l)}$. The convergence of the $\epsilon I_2^{(l)}$ does not actually follow from the convergence of the $I_1^{(l)}$, and hence from the reduced out form of the representations $D^{(l)}$, but was made at least possible by this circumstance.

Before going over to the investigation of the Lorentz groups, it may be worth while to make a final remark about the above contraction, even though it has little to do with our subject. We shall determine, first, the matrices which correspond to finite group elements. For this purpose, it is useful to consider that form of the above representation in which the Hilbert space consists of functions of x and y and the infinitesimal operators have the natural form for infinitesimal operators of the Euclidean group (α and r are polar coordinates $x = r \cos \alpha$, $y = r \sin \alpha$)

$$J_1' = -\partial/\partial\alpha = y\partial/\partial x - x\partial/\partial y \tag{19}$$

$$J_{2_x}' = \partial/\partial y \qquad J_{2_y}' = -\partial/\partial x. \tag{19a}$$

One should keep in mind that J_{2_x}' arose from I_{2_x} which is the infinitesimal rotation about the x axis and corresponds to a displacement in the $-y$ direction. Similarly J_{2_y}' corresponds to a displacement in the x direction.

The function $\varphi(\alpha, r)$ corresponds in the new Hilbert space to the vector which has the components φ_m in the Hilbert space of (17a), (17b). It further follows from (18a) that

$$\left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2}\right) \varphi = -\Xi^2 \varphi \tag{20}$$

whence one can write

$$\begin{aligned} \varphi(x, y) &= \int e^{-i\Xi(x \cos \alpha' + y \sin \alpha')} g(\alpha')d\alpha' \\ &= \int e^{-i\Xi r \cos (\alpha - \alpha')} g(\alpha')d\alpha' \\ &= \int e^{-i\Xi r \cos \alpha'} g(\alpha - \alpha')d\alpha'. \end{aligned} \tag{20a}$$

All integrations are from 0 to 2π. Expanding $g(\alpha - \alpha')$ into a Fourier series of $\alpha - \alpha'$, one finds

$$\varphi(x, y) = \int e^{-i\Xi r \cos \alpha'} \sum g_m e^{-im(\alpha - \alpha')} d\alpha' \tag{20b}$$

m runs from $-\infty$ to ∞. The last form makes it easy to calculate J_1' by (19), the first form permits one to calculate J_{2_x}' and J_{2_y}' easily. Comparison of the expressions obtained in this way with (17a) and (17b) shows that $g_m = \varphi_m$. In order to remain in keeping with the usual notation, we define[3]

$$2\pi J_m(z) = i^m \int e^{-iz\cos\alpha'} e^{im\alpha'} d\alpha'.$$

J_m is then the ordinary Bessel function of order m. This permits one to write for (20b)

$$\varphi(\alpha, r) = 2\pi \sum \varphi_m i^{-m} e^{-im\alpha} J_m(\Xi r). \tag{21}$$

Because of (19a), one can write down at once the finite group operations. In particular, for the displacement $T(\xi, \eta)$ by ξ and η in the x and y directions one has

$$T(\xi, \eta)\varphi(x, y) = \varphi(x - \xi, y - \eta). \tag{22}$$

Hence, denoting the matrix for the same operation in the original Hilbert space by $T(\xi, \eta)_{mm'}$, one has

$$\varphi(x - \xi, y - \eta) = 2\pi \sum_m \sum_m T(\xi, \eta)_{mm'} \varphi_{m'} i^{-m} e^{-im\alpha} J_m(\Xi r). \tag{22a}$$

This permits an explicit determination of the $T(\xi, \eta)_{mm}'$. We shall not carry this out completely but set only $r = 0$ in (22a). Since all $J_m(0) = 0$ except $J_0(0) = 1$, the summation over m disappears on the right side. The left side becomes, at the same time by (21)

$$\varphi(-\xi, -\eta) = 2\pi \sum \varphi_m i^{-m} e^{-im(\beta + \pi)} J_m(\Xi\rho) \tag{22b}$$

where β, ρ are the polar coordinates for ξ, η. Comparing (22a) and (22b) one finds, with $\xi = 0$, $\beta = {}^1/_2\pi$, $\eta = \rho$

$$J_m(\Xi\rho) = T(0, \rho)_{0m}. \tag{23}$$

The group relations and the form of the infinitesimal operators (19a) gives at once the most important relations for Bessel functions, such as the addition theorem, differential equation (cf. (20)), etc. Up to this point the argument is not new but merely a repetition, for the two dimensional Euclidean group, of a similar reasoning given before[4] for the rotation group. This led to the equation[4]

$$D^{(l)}(0, \beta 0)_{0m} = \bar{P}^m(\cos\beta) = \left(\frac{(l - m)!}{(l + m)!}\right)^{1/2} P_l^m(\cos\beta) \tag{24}$$

in which P_l^m is Legendre's associated function of the first kind, $\bar{P}_l^m$ is normalized to the same value as $P_l^0 = P_l$. Furthermore, 0, β, 0 is the rotation about the x axis so that, by the definition of $T(0, \rho)$

$$T(0, \rho)_{m'm} = (-)^{m'} \lim D^{(l)} \quad (0, \Xi\rho/l, 0)_{m'm}. \tag{25}$$

This, together with (23), gives the asymptotic expression for the associated Legendre functions[5]

$$\lim_{l \to \infty} \bar{P}_l^m(\cos(\rho/l)) = J_m(\rho). \tag{25a}$$

III. Contraction of Lorentz Groups

Let us consider, first, the inhomogeneous Lorentz group with one space-like, one time-like dimension. It is given by the transformations

$$\begin{aligned} x' &= x \cosh \lambda + t \sinh \lambda + a_x \\ t' &= x \sinh \lambda + t \cosh \lambda + a_t. \end{aligned} \tag{26}$$

We wish to contract it with respect to the subgroup of time displacements $t' = t + a_t$. The infinitesimal elements of (26) are: time displacement I_1, space displacement I_{2x} and "rotation" in space-time $I_{2\lambda}$. Their commutation relations read

$$[I_1, I_{2x}] = 0 \qquad [I_1, I_{2\lambda}] = -I_{2x} \qquad [I_{2x}, I_{2\lambda}] = -I_1. \tag{27}$$

Hence, by (9), the commutation relations of the contracted group are

$$[J_1, J_{2x}] = 0 \qquad [J_1, J_{2\lambda}] = -J_{2x} \qquad [J_{2x}, J_{2\lambda}] = 0. \tag{27a}$$

The "rotations" in space-time, together with the displacements in space, form a commutative invariant subgroup.

The matrices

$$\begin{Vmatrix} \cosh \lambda & \sinh \lambda & a_x \\ \sinh \lambda & \cosh \lambda & a_t \\ 0 & 0 & 1 \end{Vmatrix} \tag{26a}$$

form a natural, though not unitary, representation of the group of transformations (26). We can carry out the contraction by setting $a_t = b_t$, $\lambda = \epsilon v$, $a_x = \epsilon b_x$ or $\lambda = v/c$, $a_x = b_x/c$ and letting ϵ converge to 0, or c converge to infinity. If we do this directly in (26a), the representation will not remain faithful for the contracted group. We shall transform therefore (26a) with a suitable ϵ (or c) dependent matrix: multiply the first row with c, the first column with $1/c$. If c goes to infinity in the matrix obtained in this way, one obtains the transformations of the contracted group

$$\begin{aligned} x' &= x + vt + b_x \\ t' &= \qquad t + b_t. \end{aligned} \tag{27a}$$

It is the inhomogeneous Galilei group with one spatial dimension. The transformations $x' = x + vt + b_x$, $t' = t$ form the commutative invariant subgroup.

The same contraction can be carried out for an inhomogeneous Lorentz group with an arbitrary number of spatial dimensions. The only difference is that the subgroup S, with respect to which the contraction is carried out, contains not only the displacements in time as in the above example, but also all purely spatial rotations, i.e., all homogeneous transformations

which leave t invariant. The invariant subgroup of the contracted group consists of all spatial displacements and Galilei transformations: $x_i' = x_i' + v_i t + b_i, \quad t' = t$.

Contraction of the Unitary Representations of the Lorentz Groups.—We shall be principally concerned here with the group of the special theory of relativity, i.e., the inhomogeneous Lorentz group with three space-like and one time-like dimension. The subgroup S with respect to which we shall contract it contains the displacements in time, the spatial rotations, and the products of these operations. The contracted group is the ordinary Galilei group, i.e., the group of classical mechanics.

We shall denote the displacement operators in the direction of the three space-like axes by I_k ($k = 1, 2, 3$), the displacement in the direction of the time axis by I_0. The rotations in the kl plane will be denoted by I_{kl}, the acceleration in the direction of the k axis by I_{k0}. I_0 and I_{kl} span the subgroup S.

The quantity

$$I_1^2 + I_2^2 + I_3^2 - I_0^2 = P \tag{28}$$

is a constant in every irreducible representation and the irreducible representations can be divided into three classes according to the value of this constant.[6] In the first class, $P < 0$ and the momenta (which are $-i$ times the infinitesimal operators) are space-like. In the second class $P = 0$ and the momenta form a null vector, $P > 0$ in the third class.

It is generally admitted that the representations of the first class have no physical significance because the momenta of all observed particles are time-like or null vectors. Hence we shall investigate only the simplest one of these representations. Its operators are most easily given in the Hilbert space of functions of three variables p_1, p_2, p_3 which are restricted to the outside of a sphere of radius $\sqrt{-P}$. The expressions for the infinitesimal operators are

$$I_k = ip_k \qquad I_0 = -i(p_1^2 + p_2^2 + p_3^2 + P)^{1/2}. \tag{28a}$$

This last equation also shows the reason for the variables p_k to be restricted by $p_1^2 + p_2^2 + p_3^2 > -P$: if this inequality is not fulfilled, I_0 ceases to be skew hermitean and, hence, the representation is not unitary. Further

$$I_{kl} = p_l \partial/\partial p_k - p_k \partial/\partial p_l \tag{28b}$$

and

$$I_{k0} = -(p_1^2 + p_2^2 + p_3^2 + P)^{1/2} \partial/\partial p_k. \tag{28c}$$

It is useful to introduce new variables instead of the p_1 both in order to simplify the definition domain of the variables and also to bring the operations of the subgroup S into a form which is independent of P. This can

be done, most simply, by introducing as new variables

$$p_0 = (p_1^2 + p_2^2 + p_3^2 + P)^{1/2} \qquad \Omega_k = p_k/(p_0^2 - P)^{1/2}. \tag{29}$$

The Ω_k are restricted to a unit sphere, p_0 changes from 0 to ∞. In these variables, the scalar product between φ and ψ reduces to

$$(\varphi, \psi) = \int d\Omega \int dp_0 (p_0^2 - P)^{1/2} \bar{\varphi}\psi \tag{29a}$$

$d\Omega$ being the surface element of the unit sphere over which the integration with respect to the Ω is to be extended. In terms of the new variables, the infinitesimal operators assume the form

$$I_{kl} = \Omega_l \, \partial/\partial\Omega_k - \Omega_k \, \partial/\partial\Omega_l \qquad I_0 = -ip_0 \tag{30}$$

and

$$I_k = i\Omega_k (p_0^2 - P)^{1/2} \tag{30a}$$

$$I_{k0} = -(p_0^2 - P)^{1/2}\Omega_k \, \partial/\partial p_0 - \frac{p_0}{(p_0^2 - P)^{1/2}} \sum_l \Omega_l (\Omega_l \, \partial/\partial\Omega_k - \Omega_k \, \partial/\partial\Omega_l).$$

If we now set $J_k = \epsilon I_k$, the J_k will converge to zero unless $-P$ becomes inversely proportional to ϵ^2, i.e., unless $-\epsilon^2 P$ converges to a definite limit P. If this is assumed, the second term of ϵI_{k0} will converge to zero and the infinitesimal elements of the representation of the contracted group become

$$J_{kl} = \Omega_l \, \partial/\partial\Omega_k - \Omega_k \, \partial/\partial\Omega_l \qquad J_0 = -ip_0 \tag{31}$$

$$J_{k0} = -\mathrm{P}\Omega_k \, \partial/\partial p_0 \qquad J_k = i\,\mathrm{P}\Omega_k. \tag{31a}$$

These operators indeed span a unitary representation of the Galilei group: they correspond to case II with $m = 0$ of a recent determination of these representations.[7] The p_k of this article correspond to our $\mathrm{P}\Omega_k$, the variable s is given by $i\,\mathrm{P}\,\partial/\partial p_0$. One also understands now why it was impossible to find a physical interpretation to this representation: it is the limiting case of a representation of the relativistic group with imaginary mass $P^{1/2}$. The same is probably true of the other true representations[7] of the Galilei group.

We go over now to the investigation of the simplest representation with positive P, i.e., the representation of the Klein-Gordon equation. The infinitesimal operators are again given by (28a), (28b), (28c). However, since P is positive, the variability domain of the p extends over the whole three dimensional space. The scalar product of two functions φ and ψ is now given by

$$(\varphi, \psi) = \int\int\int dp_1 \, dp_2 \, dp_3 \, p_0^{-1} \, \bar{\varphi}\psi \tag{32}$$

in which p_0 is still given by (29).

If we set, in the sense of (6),

$$J_k = \epsilon I_k = \epsilon i p_k \tag{33}$$

the operation of time displacement becomes

$$iJ_0 = iI_0 = (-I_1^2 - I_2^2 - I_3^2 + P)^{1/2} = \epsilon^{-1}(-J_1^2 - J_2^2 - J_3^2 + \epsilon^2 P)^{1/2} \tag{33a}$$

Since the J_k are skew hermitean, their squares are positive definite hermitean operators. Since $\epsilon^2 P$ is also positive the expectation value $(\varphi, iJ_0\varphi)$ of iJ_0 is, for any state φ, greater than ϵ^{-1} times the expectation value of iJ_1. It follows that if $J_1\varphi$ converges to a vector in Hilbert space as $\epsilon \to 0$, the vector $J_0\varphi$ must grow beyond all limits. The same is true, of course, for the other J_k. It follows that the representations considered cannot be contracted in the sense discussed in the previous sections and the same is true of all representations of the classes $P \geq 0$.

It is possible, however, to contract these representations to representations up to a factor of the Galilei group. The commutation relations of the infinitesimal elements of representations up to a factor differ from the commutation relations of real representations by the appearance of a constant in the structural relations. Hence

$$[J_\alpha, J_\beta] = \sum c_{\alpha\beta}^\gamma J_\gamma + b_{\alpha\beta}1 \tag{34}$$

where $c_{\alpha\beta}^\gamma$ are the structure constants of the group to be represented (in our case, the inhomogeneous Galilei group) and the $b_{\alpha\beta}1$ are multiples of the unit operator. One will, therefore, obtain infinitesimal elements of representations up to a factor if one sets, instead of (6)

$$J_{1\nu} = I_{1\nu} - a_{1\nu}1 \qquad J_{2\nu} = \epsilon I_{2\nu} - a_{2\nu}1 \tag{34a}$$

in which all a may depend on ϵ. Since the additional terms in (34a) commute with all other operators, these additional terms will not affect the left side of (34). Hence, they must be compensated also on the right side and this is done by the additional terms $b_{\alpha\beta}1$. The point of introducing the terms $a1$ in (34a), which then necessitates the introduction of the b in (34), is that the right sides of (34a) may converge to finite non-vanishing operators even if the $I_{1\nu}$, $\epsilon I_{2\nu}$ cannot be made to converge.

The above generalization of the concept of contraction indeed allows a contraction of the representations given by (28a), (28b), (28c) also for $P > 0$. As we let P go to infinity, I_0 will also tend to infinity ($(I_0\varphi, I_0\varphi)$ converges to infinity for all φ). However, subtracting $-iP^{1/2}1$ from I_0, it will converge to

$$\begin{aligned} J_0 &= \lim -i(-I_1^2 - I_2^2 - I_3^2 + P)^{1/2} + iP^{1/2} \\ &= (i/2P^{1/2})(I_1^2 + I_2^2 + I_3^2) = (i/2P^{1/2}\epsilon^2)(J_1^2 + J_2^2 + J_3^2). \end{aligned} \tag{35}$$

This shows that J_0 will converge to a finite operator if the $J_k = \epsilon I_k$ do and if $P^{1/2}\epsilon^2$ converges to a finite constant m as $P \to \infty$, $\epsilon \to 0$. Both can be accomplished by assuming the representation in such a form that, instead of (28)

$$I_k = ip_k/\epsilon \qquad I_0 = -i(P + (p_1^2 + p_2^2 + p_3^2)/\epsilon^2)^{1/2}. \tag{36}$$

This is indeed possible because the variability domain of the p_k is unrestricted and the above form of the infinitesimal elements can be obtained by unitary transformation of the operators given in (28a). Such a transformation leaves the I_{kl} of (28b) unchanged but transforms the I_{k0} of (28c) into

$$I_{k0} = -\epsilon(P + (p_1^2 + p_2^2 + p_3^2)/\epsilon^2)^{1/2}\partial/\partial p_k. \tag{36b}$$

Hence we shall have

$$\begin{aligned} J_0 &= \lim -i(P + (p_1^2 + p_2^2 + p_3^2)/\epsilon^2)^{1/2} + iP^{1/2} = \\ &\qquad -(i/2m)(p_1^2 + p_2^2 + p_3^2) \\ J_{kl} &= p_l\partial/\partial p_k - p_k\partial/\partial p_l \end{aligned} \tag{37.1}$$

$$\begin{aligned} J_k &= \lim \epsilon(ip_k/\epsilon) = ip_k \\ J_{k0} &= \lim -\epsilon(P + (p_1^2 + p_2^2 + p_3^2)/\epsilon^2)^{1/2}\epsilon\partial/\partial p_k \\ &= \lim -(P\epsilon^4 + \epsilon^2(p_1^2 + p_2^2 + p_3^2))^{1/2}\partial/\partial p_k = -m\partial/\partial p_k. \end{aligned} \tag{37.2}$$

The reader familiar with the transition from the Klein-Gordon to the Schrödinger equation will recognize the increase of the rest mass with increasing c and the elimination of this rest mass by the subtraction of $-iP^{1/2}1$ from the infinitesimal operator of the time-displacement operator. The infinitesimal operators (37.1), (37.2) for the contracted group are in fact those of Schrödinger's theory. It is likely that a similar contraction is possible also for the other representations with positive rest mass (i.e., $P > 0$) but this and the behavior of the representations with $P = 0$ will not be further discussed here.

[1] Segal, I. E., *Duke Math. J.*, **18**, 221 (1951).

[2] Cf. e.g. Wigner, E., *Gruppentheorie und ihre Auwendiengen* etc., Friedr. Vieweg, Braunschweig (1931) and Edwards Brothers, Ann Arbor (1944). Chapter XV.

[3] Cf. Jahnke, E., and Emde, F., *Tables of Functions*, Dover Publications, 1943, p. 149; or Watson, G. N., *Treatise on Bessel Functions*, Cambridge Univ. Press, 1922, p. 19ff.

[4] Reference 2, Chapter XIX, particularly p. 230, 232. Cf. also Wigner, E. P., *J. Franklin Inst.*, **250**, 477 (1950), and Godement, R., *Trans. Amer. Math. Soc.* **73**, 496 (1952).

[5] For $m = 0$ this is given on p. 65, of Watson's *Bessel Functions* (ref. 3).

[6] Cf. e.g. Bargmann, V., and Wigner, E. P., these PROCEEDINGS, **34**, 211 (1948) and further literature quoted there.

[7] Inonu, E., and Wigner, E. P., *Nuovo cimento*, **9**, 705 (1952).

On a Particular Type of Convergence to a Singular Matrix

E. Inönü and E. P. Wigner

Proceedings of the National Academy of Sciences 40, 119–121 (1954)

Communicated December 14, 1953

It was attempted, in an earlier investigation,[1] to consider limiting cases of continuous groups G which are themselves groups. For this purpose, linear combinations of the infinitesimal elements I of G were formed

$$J_i = \sum(u_{i\nu} + \epsilon w_{i\nu})I_\nu \tag{1}$$

in which u was singular but the determinant $|u + \epsilon w|$ did not vanish identically in ϵ. It was attempted then to find the conditions for the J to remain infinitesimal operators of a group even in the limiting case $\epsilon \to 0$. The group generated by the J in the limiting case was called a contraction of G.

It was stated, in the course of the derivation of the conditions for the contraction, that it is always possible to find two non-singular matrices β and α such that

$$\beta u \alpha^{-1} = \begin{Vmatrix} 1 & 0 \\ 0 & 0 \end{Vmatrix} \qquad \beta w \alpha^{-1} = \begin{Vmatrix} v & 0 \\ 0 & 1 \end{Vmatrix}. \tag{2}$$

In (2), the 1, 0, v are submatrices: the rank of the 1 in $\beta u \alpha^{-1}$ is equal to that of u. This statement is, in general, incorrect[2] as can be seen already on the example of the two dimensional matrices

$$u = \begin{Vmatrix} 1 & 0 \\ 0 & 0 \end{Vmatrix} \qquad w = \begin{Vmatrix} 0 & 1 \\ 1 & 0 \end{Vmatrix}.$$

When checking the calculation which led to the form (2), it was found that it contained a continuity assumption which appears reasonable but which was not stated explicitly. Since the continuity assumption in question may be of interest also apart from the problem of the contraction of groups, we shall formulate it below.

Let us consider two n dimensional vector spaces I and J and mappings $U(\epsilon)$ of I on J. For $0 < \epsilon < \epsilon_0$, the mapping shall be one-to-one; $U(0)$ shall map I on a subspace J_u of J. We shall call the mapping $U(\epsilon)$ *reversibly continuous* at $\epsilon = 0$ if $U(0)U(\epsilon)^{-1}$, applied to the vectors of J_u, converges to the identical mapping of J_u on itself as $\epsilon \to 0$. An equivalent formulation hereto is that the mapping $U(0)$ $U(\epsilon)^{-1}U(0)$ of I unto J shall converge, as $\epsilon \to 0$, to $U(0)$. In fact, let i be a vector in I and write $j = U(0)i$. The convergence of $U(0)\ U(\epsilon)^{-1}j \to j$ is synonymous with the convergence of $U(0)\ U(\epsilon)^{-1}U(0)i$ to $U(0)i$.

If the convergence of $u + \epsilon w$ to u is reversibly continuous, the existence of the normal forms (2) does indeed follow. *Thus the addition of this assumption will guarantee the validity of the aforementioned considerations.*[1] One can see this perhaps most simply as follows. Let us denote by I_0 the subspace of I for which $uI_0 = 0$. Similarly, let us denote the subspaces uI and wI_0 by J_u and J_{0w}. We can first prove then that J_u and J_{0w} together span the whole space J. In fact, if the vector i_0 is in I_0, $wi_0 = 0$ would entail $(u + \epsilon w)i_0 = 0$ for all ϵ which is contrary to assumption. Furthermore, wi_0 cannot be contained in J_u because if it were

$$u(u + \epsilon w)^{-1}wi_0 = (1/\epsilon)u(u + \epsilon w)^{-1}(u + \epsilon w)i_0 = (1/\epsilon)ui_0 = 0$$

would not converge to wi_0, contrary to the postulate of reversibly continuous convergence. Hence, wI_0 contains as many linearly independent vectors as I_0 and these are also linearly independent of the vectors in uI. It follows that $uI = J_u$ together with wI_0 span the whole space J. If we choose, therefore, a coordinate system in I so that some of its axes i_{0k} span I_0, the others i_{1k} the rest of I, we shall have

$$ui_{0k} = 0 \qquad wi_{0k} = j_{0k} \qquad ui_{1k} = j_{uk}. \tag{3}$$

The last two of these are the definitions of j_{0k} and j_{uk} which, together, span J. We can set, therefore

$$wi_{1k} = \sum_\lambda v_{\lambda k}j_{u\lambda} + \sum_\lambda z_{\lambda k}j_{0\lambda}. \tag{3a}$$

Hence if we write

$$i_{uk} = i_{1k} - \sum z_{\lambda k}i_{0\lambda} \tag{3b}$$

the i_{uk} and i_{0k} will span I and we have, in addition to (3) which remain valid

$$ui_{uk} = j_{uk} \qquad wi_{uk} = \sum_\lambda v_{\lambda k}j_{u\lambda}. \tag{3c}$$

Hence u and w will in fact have the form (2) if we choose in I the coordinate axes i_{uk}, i_{0k} and in J the axes j_{uk}, j_{0k}.

In matrix language, one can formulate our theorem as follows. Be u and w n-dimensional square matrices and the rank r of u be smaller than n while $u + \epsilon w$ be non-singular except for isolated ϵ. Then the necessary and sufficient condition for

$$u(u + \epsilon w)^{-1}u \to u \qquad \text{as } \epsilon \to 0 \tag{4}$$

is that the coefficient of ϵ^{n-r} in the determinant $|u + \epsilon w|$ shall not vanish. In order to prove this, we first note that (4) remains invariant if one substitutes $\beta u \alpha^{-1}$ for u and $\beta w \alpha^{-1}$ for w where α and β are non-singular and independent of ϵ. Ac-

cording to our Theorem, if (4) is fulfilled, β and α can be so chosen that (2) be fulfilled. This form shows, however, that the lowest term in the determinant $|\beta u\alpha^{-1} + \beta w\alpha^{-1}|$, if considered as a polynomial in ϵ, is ϵ^{n-r}. Hence, the lowest term in the determinant $|u + \epsilon w|$ is $|\beta|^{-1}|\alpha|\epsilon^{n-r}$, i.e., the coefficient of ϵ^{n-r} in the determinant $|u + \epsilon w|$ does not vanish.

Conversely, let us assume that u and w are of dimension n, u of rank r, and the coefficient of ϵ^{n-r} in $|u + \epsilon w|$ does not vanish. The same holds then for $u' = \beta u\alpha^{-1}$ and $w' = \beta w\alpha^{-1}$ and we can choose β and α so that $u' = \beta u\alpha^{-1}$ have the form given in (2). Furthermore, the algebraic complement of any element in the first r rows of $|u' + \epsilon w'|$ goes to zero at least as ϵ^{n-r}. Hence, the ratio of such a complement to the determinant $|u' + \epsilon w'|$, i.e., any element in the first r rows of $(u' + \epsilon w')^{-1}$, converges to a finite number (or to 0) as $\epsilon \to 0$. It follows that the elements in the first r rows of $(u' + \epsilon w')^{-1}u'$ converge to the same numbers to which the corresponding elements of $(u' + \epsilon w')^{-1}(u' + \epsilon w')$ converge. Hence, as far as the first r rows are concerned

$$u'(u' + \epsilon w')^{-1}u' \to u'(u' + \epsilon w')^{-1}(u' + \epsilon w') = u'. \tag{4a}$$

However (4a) is clearly correct also as far as the last $n - r$ rows are concerned because these vanish on both sides because of the form (2) of $u' = \beta u\alpha^{-1}$. Hence, (4a) follows from the premises and since α and β are independent of ϵ, this holds also for (4).

[1] Inonu, E., and Wigner, E. P., these PROCEEDINGS, **39**, 510 (1953).

[2] We are much indebted to Dr. V. Bargmann for his friendly comments concerning this point.

Application of the Rayleigh-Schrödinger Perturbation Theory to the Hydrogen Atom

E. P. Wigner

The Physical Review 94, 77–78 (1954)

(Received November 27, 1953)

The motion of a single electron in the electrostatic field of a nucleus is treated by the Rayleigh-Schrödinger perturbation method, the whole electrostatic potential being considered as the "perturbation." The contribution of the first approximation to the energy vanishes. The second approximation gives a finite ionization energy which is, however, incorrect numerically. The first approximation also vanishes for potentials $\sim r^{-n}$ with $0<n<3$ but the second approximation is finite only for $n=1$: it vanishes for $n>1$, and is infinite for $n<1$.

THERE are several instances in which simple approximation methods yield correct results even though the conditions for the applicability of those methods are not fulfilled. The best known example is the calculation of the Rutherford scattering by Born's approximation.[1] The calculation of the interaction of a large number of particles[2] by the Rayleigh-Schrödinger method is another example. In all these cases, the wave function given by the approximation method is quite meaningless. Thus, in the case of the interaction of many particles, the wave function obtained by the Rayleigh-Schrödinger perturbation calculation's second approximation represents a state in which, at most, two particles are in excited states while, actually, the average number of particles in excited states is, if the number of particles is large, a constant fraction of the number of particles. Nevertheless, the energy given by the second approximation of the Rayleigh-Schrödinger theory appears to be meaningful. One may gain the impression, from these examples, that the second approximation is always correct if it gives a finite result and that this applies particularly to the electrostatic interaction.

It may be of some interest to note, in this connection, that the aforementioned perturbation method yields, in second approximation, a finite result if applied to the problem of an electron moving in the attractive field of a nucleus. In fact, among all the potentials cr^{-n}, this applies only in the case of $n=1$. The starting point, in every case, is a free electron moving in a very large "box." Nevertheless, the numerical result of the perturbation calculation gives a grossly inaccurate result for the Coulomb potential. The wave function is not only inaccurate but virtually meaningless, but this fact does not entail an inaccurate energy value in the cases discussed in the first paragraph.

As the starting point we choose a spherical "box" with a radius R. The unperturbed spherically symmetric wave functions are, then,

$$\psi_n=(\sin\alpha_n r)/C_n r, \tag{1}$$

with the normalization constant C_n being equal to

$$C_n^2=(2\pi/\alpha_n)(\alpha_n R-\sin\alpha_n R\cos\alpha_n R). \tag{1a}$$

The α_n are determined by the boundary condition to be used at the surface of the box. The first approximation to the energy, because of an e^2/r potential, goes to zero with increasing R because the average distance of the particle from the origin is of the order R in the state ψ_1. The second approximation becomes

$$-E=\sum_2^\infty \frac{\left(\int \psi_1(e^2/r)\psi_n 4\pi r^2 dr\right)^2}{(\hbar^2/2m)(\alpha_n^2-\alpha_1^2)}. \tag{2}$$

The denominator contains the energy difference between the states ψ_n and ψ_1. Inserting (1) into (2),

$$\begin{aligned}-E&=\sum_2^\infty \frac{32\pi^2e^4m}{C_1^2C_2^2(\alpha_n^2-\alpha_1^2)\hbar^2}\left(\int_0^R \sin\alpha_1 r\sin\alpha_n r dr/r\right)^2\\ &=\sum_2^\infty \frac{8\pi^2e^4m}{C_1^2C_2^2(\alpha_n^2-\alpha_1^2)\hbar^2}\left(\int_0^R [\cos(\alpha_n-\alpha_1)r\right.\\ &\qquad\left.-\cos(\alpha_n+\alpha_1)r]dr/r\right)^2.\end{aligned} \tag{2a}$$

The integrals can be transformed by substituting $x=(\alpha_n-\alpha_1)r$ and $x=(\alpha_n+\alpha_1)r$, respectively.

$$\begin{aligned}-E=\sum_2^\infty &\frac{8\pi^2e^4m}{C_1^2C_2^2(\alpha_n^2-\alpha_1^2)\hbar^2}\\ &\times\left(\int_{(\alpha_n-\alpha_1)R}^{(\alpha_n+\alpha_1)R}(\cos x-1)\frac{dx}{x}\right)^2.\end{aligned} \tag{2b}$$

In order to define the unperturbed system completely, one has to give the value of the logarithmic derivative of the unperturbed wave functions (1) for $r=R$. We set $\rho\psi'(R)=\psi(R)$, which gives

$$\cot\alpha_n R=\frac{1}{\alpha_n\rho}+\frac{1}{\alpha_n R}. \tag{3}$$

This equation determines the values of α_n. We have two cases to distinguish: ρ is either infinite or finite. In the latter case, we assume that ρ remains the same as we increase the size of the "box," i.e., increase R to

[1] See G. Wentzel, *Geiger-Scheel's Handbuch der Physik* (Springer, Berlin, 1933), Vol. 24, Part 1, p. 703.
[2] S. Watanabe, Z. Physik 112, 159 (1939); 113, 482 (1939); W. Macke, Z. Naturforsch. A5, 192 (1950).

infinity. If this is the case, $\alpha_n R \to n\pi$ as $R \to \infty$, no matter what the value of ρ is. In other words, the value of the expression (2b) converges to the same value as $R \to \infty$, no matter what the value of ρ is as long as ρ remains finite as $R \to \infty$. In this case, then, $C^2 = 2\pi R$, and

$$-E = \sum_2^\infty \frac{2e^4 m}{R^2(\alpha_n^2 - \alpha_1^2)\hbar^2}\left(\int_{(n-1)\pi}^{(n+1)\pi} (1-\cos x)\frac{dx}{x}\right)^2$$

$$= \sum_2^\infty \frac{2e^4 m}{\pi^2\hbar^2(n^2-1)}(\mathrm{Ci}(n+1)\pi - \mathrm{Ci}(n-1)\pi)^2, \tag{4}$$

where Ci is the cosine integral

$$\mathrm{Ci}x = \int_0^x \frac{\cos x - 1}{x} dx. \tag{4a}$$

As we anticipated, the second approximation (2) gives a finite value in this case. The sum (4) was evaluated numerically and gave

$$-E = 0.1093 e^4 m/\hbar^2, \tag{5}$$

as contrasted with the correct value of 0.5 $e^4 m/\hbar^2$. As a matter of fact, it is not surprising that the value (5) is so far from being correct. Although the first approximation correction to the nth characteristic value of the problem,

$$-\int_0^R \psi_n^2 (e^2/r) 4\pi r^2 dr = \frac{e^2}{R}\mathrm{Ci}(2n\pi), \tag{6}$$

vanishes for very large R, it vanishes only as $1/R$. Furthermore, it increases with n roughly like $\gamma + \ln 2\pi n$. The unperturbed energy differences, on the other hand, go to zero as R^{-2}. Hence, in first approximation, the energy of ψ_2 is actually lower than that of ψ_1 if R is large enough, and the same holds for a large number of other ψ_n. This reversal of the order of the first-order energy values shows again that the perturbation calculation cannot be given a direct physical interpretation in this case.

It remains to be seen whether the other boundary condition, $\psi'(R) = 0$, gives more satisfactory results. In this case ψ_1 is a constant,

$$\psi_1 = (3/4\pi R^3)^{\frac{1}{2}}, \tag{7}$$

while the other ψ are still given by (1). The equation defining α_n becomes

$$\alpha_n R = \tan \alpha_n R; \quad \alpha_1 = 0. \tag{8}$$

Denoting the solutions of this equation with $q_n = \alpha_n R = \tan q_n$, the integral

$$\int_0^R \psi_1 (e^2/r) \psi_n 4\pi r^2 dr = 6^{\frac{1}{2}} \frac{e^2}{R} \frac{(1+q_n^2)^{\frac{1}{2}} + (-)^n}{q_n^2} \tag{9}$$

can be expressed in closed form. For the ionization energy $-E$ one obtains

$$-E = \frac{12 e^4 m}{\hbar^2} \sum_2^\infty \frac{[(1+q_n^2)^{\frac{1}{2}} + (-)^n]^2}{q_n^6} \sim 0.05 \frac{e^4 m}{\hbar^2}, \tag{10}$$

i.e., a substantially smaller value than (5).

The reason it was at all possible to expect that the Rayleigh-Schrödinger perturbation method yield a correct expression for the energy is that this approximation method gives a power series expansion for the energy in terms of the perturbation parameter. A power series expansion in terms of the perturbation parameter $\lambda = e^2$ in fact does exist for the energy values of a particle in a Coulomb field: the coefficients of all powers of λ are zero, excepting that of the second power which is $m/2\hbar^2$ for the normal state. The reason that the perturbation method nevertheless gives an incorrect value in second approximation is that the energy is a function of the size R of the "box" which underlies the perturbation calculation. This function could well be of a form similar to

$$(e^4 m/2\hbar^2)[1 - c\exp(-me^2 R/\hbar^2)].$$

This expression gives the correct energy value for $R = \infty$ and has a convergent power series expansion in terms of e^2. However, this power series expansion does not converge term by term to the power series expansion of the energy for $R = \infty$ and, in particular, the coefficient of the e^4 term is too small by a factor $1 - c$.

The fact that the Rayleigh-Schrödinger second approximation gives a definite value for the potential λ/r, but for no other potential of the form λr^{-n}, can be seen already by means of a dimensional argument. The dimension of λ is [energy][length]^n. Since the dimension of the matrix elements of the perturbation energy have the dimension of energy, they are proportional to R^{-n}. The square of the matrix element, divided by a typical difference of the unperturbed energy $\hbar^2/2mR^2$, is therefore proportional to R^{2-2n}. This gives a definite value only for $n = 1$ but tends to infinity as R increases for $n < 1$. It tends to zero if $n > 1$.

The square integral of the first approximation to the wave function tends to infinity with increasing R for all $n < 2$. In order to calculate it we have to divide the square of the matrix element with the square of a quantity of the order $\hbar^2/2mR^2$. Hence, the square integral will be proportional to R^{4-2n} and increase beyond all bounds also in the case of the Coulomb potential $n = 1$. In fact, the first order wave function in our case is closely approximated by ψ_2 multiplied with a very large factor and is almost as poor an approximation to the real wave function as the unperturbed wave function ψ_1. The second approximation to the energy can be finite because it is not the expectation value of the first order wave function.

Significance of Loewner's Theorem in the Quantum Theory of Collisions

E. P. Wigner and J. von Neumann

Annals of Mathematics *59*, 418–433 (1954)

(Received January 6, 1954)

I. Introduction

It can be shown[1] that under rather general conditions one can define "components" of the quantum mechanical wave function of a collision system between which relations of the form

$$v_\nu = \sum_\mu R_{\nu\mu} d_\mu \tag{1}$$

obtain. In (1), the v and d are "components" of the wave function, i.e. scalar products of the state function φ with suitably defined other functions v_ν and d_ν

$$v_\nu = (v_\nu, \varphi) \qquad d_\nu = (d_\nu, \varphi). \tag{2}$$

R is often called the derivative matrix because in the most common application[2] of the above formalism the v_ν contain a delta function and the d_ν a δ' function. Hence the v_ν are integrals over values of φ and the d_ν integrals over the derivatives of φ, and R expresses the former in terms of the latter. $R_{\nu\mu} = R_{\nu\mu}(E)$ depends on the energy of the system and the same is true, of course, of any bilinear form of R with a constant, i.e. energy independent real vector ξ:

$$R(E) = \sum_{\nu\mu} R_{\nu\mu}(E)\xi_\nu \xi_\mu . \tag{3}$$

One can show,[1] furthermore, that if it is possible to define a probability density and a probability current in the region in which the particles are not interacting, the matrices

$$\left\| \begin{matrix} R'(E_1) & \dfrac{R(E_1) - R(E_2)}{E_1 - E_2} & \cdots & \dfrac{R(E_1) - R(E_n)}{E_1 - E_n} \\ \dfrac{R(E_2) - R(E_1)}{E_2 - E_1} & R'(E_2) & \cdots & \dfrac{R(E_2) - R(E_n)}{E_2 - E_n} \\ \vdots & \vdots & & \vdots \\ \dfrac{R(E_n) - R(E_1)}{E_n - E_1} & \dfrac{R(E_n) - R(E_2)}{E_n - E_2} & \cdots & R'(E_n) \end{matrix} \right\| \tag{4}$$

are positive definite, no matter what the direction of the vector ξ in (3) is as long as this converges and no matter what the value of n and of the

$$E_1, E_2, \cdots, E_n$$

[1] To be published elsewhere. Cf. also Notes on Lectures on Collision Theory. E. P. Wigner, notes by E. Vogt.

[2] E. P. Wigner and L. Eisenbud, Physical Rev. **72**, 29 (1947).

are as long as they are real and positive. The R' in (4) are the derivatives of R; it follows from the aforementioned considerations that these exist except, perhaps, at isolated points where R may become singular.

It was reasonable to expect that the above conditions on the function R entail its "R function" nature, i.e. that R can be extended over the whole complex plane and becomes a meromorphic function of the variable E, the imaginary part of which is positive in the upper, negative in the lower half plane. This indeed would be the case if one could assume that the E_k in (4) can assume negative as well as positive values. Since there is no apparent reason to extend the postulate embodied in (4) to negative E, it will be possible to conclude only that R can be defined over the whole complex plane excepting the negative real axis, that its imaginary parts are positive and negative in the upper and the lower half planes, respectively (which excludes singularities from anywhere but the real axis), and that its singularities on the positive real axis are only poles. It then follows easily that these poles are of first order and their residues are negative.[3] However, our postulates do not suffice to exclude essential singularities on the negative real axis.

The objective of the aforementioned[1] and of the present paper is very similar to a recent article by Van Kampen.[4] Van Kampen investigates the analytic nature of the collision matrix and of the derivative matrix in a one dimensional problem of great importance. The physical assumptions which led to the postulate of the positive definite nature of (4) are very closely akin to those used by Van Kampen and used even before by Schutzer and Tiomno.[5] Our result, however, differs from that of Van Kampen inasmuch as he seems to arrive, in his particular case, at an everywhere meromorphic R.

Mathematically, our formulation is very similar to K. Loewner's work[6] on monotonic matrix functions. Loewner investigates functions f of real symmetric matrices with characteristic values in an interval I such that if m_1 and m_2 are two such matrices and $m_1 - m_2$ is positive definite, $f(m_1) - f(m_2)$ be also positive definite. Loewner shows that an f which satisfies these conditions for matrices of any dimensionality can be extended to become an analytic function in the whole complex plane with imaginary parts positive and negative in the upper and lower half planes, respectively. Loewner's f is regular in the interval I.

Loewner's problem might appear to have very little to do with ours. However, one of Loewner's intermediate formulae is identical with our postulate and our use of continued fractions was also stimulated by his procedure. It is true that our condition (4) is different from Loewner's intermediate result inasmuch as

[3] Cf. a discussion of a class of functions which include these of J. A. Shohat and J. D. Tamarkin. The Problem of Moments. Publication of the American Mathematical Society, New York 1943.

[4] N. G. Van Kampen, Physical Rev. **89**, 1072 (1953), **91**, 1267 (1953).

[5] W. Schutzer and J. Tiomno, Physical Rev. **83**, 249 (1951).

[6] K. Loewner, Ueber monotone Matrix funktionen. Math. Zeits. **38**, 177 (1933).

our (4) must remain positive definite even if the E_k are not in the same domain of regularity, while Loewner's intermediate formula contains only values of the variable which are all in the same domain of regularity. It is this difference which necessitates the considerations of Section II. Apart from this point, however, Loewner's very deep considerations constitute a full proof of our premise. It is only the desire for a simpler and more concise proof of these results, starting directly from (4), which prompted the present investigation.

II. The singularities of R

In keeping with mathematical custom, we shall denote the independent variable by z rather than E. The postulates on the functions R follow.

(1) R is defined for all real positive z excepting isolated points $Z_1, Z_2, Z_3, \cdots$ which will be called its singularities. It is a real valued differentiable function in its domain of definition.

It will be noted that nothing is required of R at the singularities Z_ν. Hence, given an $R(z)$, any number of isolated points at which R is entirely regular, can be excluded from its definition domain and it still will continue to satisfy our postulates. Naturally, such points will be restored to R's definition domain in the course of the development.

(2) For positive real $z_1, \cdots, z_n$, neither of which is coincident with any of the singularities Z_ν, the matrices $\| \kappa_{ik} \|$ are positive semidefinite if

$$\kappa_{ii} = R'(z_i) \tag{5}$$

$$\kappa_{ik} = \frac{R(z_i) - R(z_k)}{z_i - z_k}. \qquad (i \neq k) \tag{5a}$$

This shall hold for any n. The R' is the derivative of R.

In order to prove the first two theorems, it is necessary to use postulate (2) only for $n = 1$ and $n = 2$. All z in this section shall be in the definition domain of R, excepting z_0 which will occasionally coincide with a singularity Z. From (2) we have for $n = 1$

$$R'(z) \geqq 0 \tag{6}$$

$n = 2$ adds to (6)

$$R'(z_1)R'(z_2) \geqq \left(\frac{R(z_1) - R(z_2)}{z_1 - z_2}\right)^2. \tag{6a}$$

It follows from (6a) at once that the $>$ sign holds in (6) everywhere unless R is a constant. Excluding this case, we infer that R is monotonically increasing in every interval which belongs to its domain of definition.

We shall first show that the singularities Z_ν fall into two groups, the first of which will be eliminated in the course of this section.

THEOREM 1. *R is either continuous and non decreasing at a singularity Z or tends to $+\infty$ and $-\infty$ as z approaches Z from the left and the right.*

$R(z)$ increases as z tends toward Z from below. Hence it either has a limiting value $R(Z-)$ as $z \to Z$ or goes to ∞. Similarly, R has a limiting value $R(Z+)$ as z approaches Z from above or tends to $-\infty$. In either case, it is possible to find two intervals (z_1, Z), (Z, z_4) contained in R's definition domain such that $R(z_2) \neq R(z_3)$ if $z_2 \neq Z$ is in the first, $z_3 \neq Z$ in the second interval. If $R(Z-) \leqq R(Z+)$, the only condition is that there be no further singularity between z_1 and z_4. If $R(Z-) > R(Z+)$ (including the case when one of these is infinite) it is necessary, in addition, to choose a number between $R(Z-)$ and $R(Z+)$ and choose $R(z_1)$ larger, $R(z_4)$ smaller than this number.

We can then integrate (6a) in the form

$$\frac{R'(z)R'(z')}{(R(z) - R(z'))^2} \geqq \frac{1}{(z - z')^2}$$

with respect to z' from z_3 to z_4 if z is in the interval (z_1, Z) and z_3 in the interval (Z, z_4)

$$R'(z)\left(\frac{1}{R(z) - R(z_4)} - \frac{1}{R(z) - R(z_3)}\right) > \frac{1}{z - z_4} - \frac{1}{z - z_3}. \tag{7}$$

This can be integrated with respect to z from z_1 to z_2 if z_2 is in the interval (z_1, Z):

$$\ln\left|\frac{R(z_2) - R(z_4)}{R(z_1) - R(z_4)}\frac{R(z_1) - R(z_3)}{R(z_2) - R(z_3)}\right| > \ln\left|\frac{z_2 - z_4}{z_1 - z_4}\frac{z_1 - z_3}{z_2 - z_3}\right|. \tag{7a}$$

If z_2 and z_3 both tend to Z (from opposite directions) the right side tends to ∞. The left side can tend to infinity if $R(z_2)$-$R(z_3) \to 0$—in this case R is continuous and at least non decreasing at Z—or if *both* $|R(z_2)|$ and $|R(z_3)|$ tend to ∞. This proves our theorem.

We shall call the first type of singularities apparent, the second type real. Clearly, $R'(z) \to \infty$ as z approaches a real singularity from either side. We shall denote with $R(Z)$ the value to which $R(z)$ converges when z approaches an apparent singularity.

THEOREM 2. *$R'(z)$ is a continuous function of z in R's domain of definition, and if one sets $R(Z)$ equal to its limiting value obtained above at an apparent singularity Z, it is continuously differentiable also at the apparent singularity.*

As a result of this theorem, it will be possible to define $R(Z)$ for an apparent singularity and to show that $R'(Z)$ also exists at these points. In order to prove Theorem 2 we shall have to establish a number of lemmas.

LEMMA 2a. *If the interval $(z - a, z + a)$ is in the definition domain of R*

$$\begin{aligned} &\tfrac{1}{2}(R'(z + a) + R'(z - a)) \geqq \sqrt{R'(z - a)R'(z + a)} \\ &\geqq (1/2a)(R(z + a) - R(z - a)) \geqq (1/2a')(R(z + a') - R(z - a')) \geqq R'(z) \end{aligned} \tag{8}$$

where $0 < a' < a$.

The first inequality is obvious, the second a direct consequence of (6a). In order to establish the last two, we form the derivative

$$\frac{d}{da}\frac{R(z+a)-R(z-a)}{2a} = \frac{R'(z+a)+R'(z-a)}{2a} - \frac{R(z+a)-R(z-a)}{2a^2}. \tag{8a}$$

This is positive because the first term on the right side is positive by (6) and larger than the geometric average of $R'(z+a)$ and $R'(z-a)$ divided by a. This, on its turn, by virtue of the second inequality, is larger than the second term on the right side of (8a). Hence, under the conditions stated,

$$\frac{R(z+a)-R(z-a)}{2a} \tag{8b}$$

is a monotonically increasing function of a. Since its limit at $a = 0$ is the last expression in (8), Lemma 2a is established. The function $R'(z)$ is, because of (8), a convex function of z in the sense of Jensen[7] in any interval which belongs to the definition domain of R. It therefore follows from Jensen's work that $R'(z)$ is continuous in every such interval. This establishes Theorem 2 except as far as the existence and behavior of $R'(z)$ at apparent singularities is concerned. However, Theorem 2 will be used particularly at apparent singularities and it does not seem possible to prove it at such singularities without proving it also at regular points. For this reason, we shall not make explicit use of Jensen's classic article even though the rest of this section is very similar to his work.

We note that except for the last part of (8), which becomes meaningless in this case, (8) remains true if $z = Z$ is a singularity of R. This is important for

LEMMA 2b. *If z_0 is an apparent singularity or a regular point of R, the derivative $R'(z)$ remains bounded as z approaches z_0.*

The third and fourth expressions of (8) are, by Theorem 1, positive if $z = z_0$ is an apparent singularity. Because of the remark preceding this Lemma

$$\frac{R(z_0+a')-R(z_0-a')}{2a'}$$

is bounded in a' if z_0 is an apparent singularity and there is no other singularity between $z_0 - a$ and $z_0 + a$. Since $R(z_0 - a') < R(z_0) < R(z_0 + a')$ in this case,

$$\frac{R(z_0+a')-R(z_0)}{a'} \tag{8c}$$

is also bounded. Since the interval $(z_0, z_0 + a)$ contains no singularity, the last part of (8) shows that (8c) is larger than $R'(z_0 + \frac{1}{2}a')$ and this is also bounded. The same is true of $R'(z_0 - \frac{1}{2}a')$ and this establishes Lemma 2b if z_0 is an apparent singularity. The argument is even simpler if z_0 is a regular point.

[7] J. L. W. V. Jensen, Sur les fonctions convexes et les inégalités entre les valeurs moyennes. Acta Mathematica, **30**, 175 (1906). This article shows also that a convex function, in addition to being continuous, has a right derivative, and a left derivative, at every point of convexity. The present article will use the positive definite nature of (4) to demonstrate the existence of the higher derivatives of R.

It follows that $R'(z)$ tends to infinity only in the neighborhood of real singularities.

LEMMA 2c. *If z approaches a regular point or an apparent singularity z_0 from above, $R'(z)$ converges to a definite value. The same is true if the approach is from below.*

Let us assume that d is the smallest condensation point of any sequence $R'(z_\nu)$ in which z_ν converges to z_0 from above. Because of (6), $d \geqq 0$. We have to show then that d is the only condensation point of such sequences.

If $R'(z_\nu) \leqq d + \varepsilon$ and if x_1 and $x_2 > x_1$ are both in the interval (z_0, z_ν)

$$\frac{R(x_2) - R(x_1)}{x_2 - x_1} \leqq d + \varepsilon. \tag{9}$$

In order to see this, we choose a z_μ such that the interval (z_μ, z_ν) include both x_1 and x_2 and $R'(z_\mu) - d \leqq \varepsilon$. Then, by (8), this will be true also at the midpoint $(\frac{1}{2})$ $(z_\mu + z_\nu)$ of z_μ and z_ν. For the same reason, it will be true at the midpoint of this point and both z_ν and z_μ. This procedure can be continued and one concludes that there is an everywhere dense set of points in the interval z_μ, z_ν such that $R'(z_\alpha) \leqq d + \varepsilon$ for any z_α of this set. If z_α converges toward x_1 from below, z_β to x_2 from above,

$$\frac{R(z_\beta) - R(z_\alpha)}{z_\beta - z_\alpha} > \frac{R(x_2) - R(x_1)}{z_\beta - z_\alpha} \tag{9a}$$

because of the monotonic nature of R. However, the left side of (9a) is, by (6a) smaller than the average of $R'(z_\beta)$ and $R'(z_\alpha)$ and hence smaller than or equal to $d + \varepsilon$. This then holds also for the right side of (9a). As z_β approaches x_2 and z_α approaches x_1, the right side of (9a) approaches the left side of (9) and this establishes the inequality (9). It follows that all differential quotients to the left of z_ν are smaller than or equal to $d + \varepsilon$. Since d is the smallest condensation point of any set $R'(z_\nu)$ in which z_ν converges to z_0 from above, it follows that $R'(z)$ converges to d no matter how z approaches to z_0 from above. Similarly, $R'(z)$ converges to a definite value d' as z approaches z_0 from below.

We shall show next, that the right derivative of R at a regular point or an apparent singularity z_0 is equal to the limiting value d of $R'(z)$ as z approaches z_0 from above. Indeed, if $h > 0$, then by Rolle's theorem

$$\frac{R(z_0 + h) - R(z_0)}{h} = R'(z_0 + \vartheta h) \tag{10}$$

for some ϑ with $0 < \vartheta < 1$. As $h \to 0$ also $\vartheta h \to 0$, hence the right hand side of (10) converges to d. This shows that the right derivative exists also at an apparent singularity and is at any rate continuous from the right. The same holds if left is substituted for right.

If differentiability in postulate (1) means the equality of right and left derivative, as was tacitly assumed so far, the above already shows the continuity of $R'(z)$ in the original definition domain of R. However, it is easy to see that (6a) excludes the possibility that right and left derivative, d and d', be different at an apparent singularity. One can write (6a) in the form

$$\left(\frac{(R(z_0 + a) - R(z_0)) + (R(z_0) - R(z_0 - a))}{2a}\right)^2 \leqq R'(z_0 + a)R'(z_0 - a). \tag{10a}$$

Now let $a \to 0$. Then the left hand side of (10a) converges to $\frac{1}{2}(d + d')^2$, while the right hand side converges to dd'. Hence

$$\left(\frac{d + d'}{2}\right)^2 \leqq dd',$$

which implies $d = d'$. Thus right and left derivative are equal at every point of the original definition domain and at every apparent singularity and Theorem 2 is established.

Theorem 2 shows that the inclusion of the apparent singularities into the definition domain of R preserves the validity of postulate (1). It preserves, however, also postulate (2). The quadratic form of a matrix of (5), with a z_k equal to an apparent singularity, is the limiting value (because of Theorems 1 and 2), of quadratic forms with all z_k in the original definition domain of R. As a limiting value of non negative numbers it is itself non negative. It follows that *the definition domain of R can be extended without impairing the defining postulate* (1) *and* (2) *so as to include all singularities at which R does not increase beyond all limits when the singularity is approached from either side.*

III. The Schiffer-Bargmann lemma

Some of the elementary properties of R functions are equally valid for those treated in the present paper. Thus Theorems 1.2, 1.3, 1.5, 1.6 given in a recent article[8] are immediate consequences of our postulates (1) and (2). We shall need here only the equivalent of Theorem 1.7.

LEMMA 3a. *If R is a function of the class described by* (1) *and* (2) *with zeros at the values x_ν of the positive real variable z and real singularities at Z_ν, and with no apparent singularities, the function $S = -1/R$ also belongs to this class and has real singularities at x_ν and apparent ones at Z_ν.*

It is clear from $S' = R'/R^2$ that S satisfies postulate (1) and that its singularities are as given in the lemma. In particular S is continuous at the points Z_ν and will assume the value 0 at these points when its definition domain will be extended to include these points. Furthermore, the matrix $\| \lambda_{ik} \|$ with

$$\lambda_{ii} = S'(z_i) = R'(z_i)/(R(z_i))^2 \tag{11}$$

$$\lambda_{ik} = \frac{S(z_i) - S(z_k)}{z_i - z_k} = \frac{1}{R(z_i)} \frac{R(z_i) - R(z_k)}{z_i - z_k} \frac{1}{R(z_k)}. \qquad (i \neq k) \tag{11a}$$

is obtained from the matrix $\| \kappa_{ik} \|$ of (5), (5a) by multiplying it from left and right with the diagonal matrix

$$D_{ik} = \delta_{ik}/R(z_i).$$

Hence, $\| \lambda_{ik} \|$ is positive definite if this holds for $\| \kappa_{ik} \|$. This proves Lemma 3a.

[8] E. P. Wigner, Amer. Math. Monthly, **59**, 669 (1952).

Let us now extend the definition domain of S to include the points Z_ν. Because of the continuity of S' at these points, and since S vanishes at them, there is a neighborhood of every Z_ν in which with $0 < \vartheta < 1$

$$S(Z_\nu + a) = aS'(Z_\nu + \vartheta a) \tag{12}$$

and S' is a continuous function of its argument. It follows from this that R is, in the neighborhood of its real singularity Z,

$$R(Z_\nu + a) = -1/S(Z_\nu + a) = -1/S'(Z_\nu + \vartheta a)a. \tag{12a}$$

Since, by the remark after (6a), $S'(Z_\nu)$ cannot be zero unless S and hence R be a constant, nor can it become infinite on account of Theorem 2, we have

LEMMA 3b. *In the neighborhood of a real singularity Z_ν, the expression $a\,R(Z_\nu + a)$ tends to a negative number as a approaches* 0.

We shall now prove the analogue of the Schiffer-Bargmann lemma for R functions (Lemma 1, 10a, reference 8).

THEOREM 3. *If R satisfies the postulates* (1), (2) *and Z_ν is a real singularity in the sense of Theorem* 1, $S = -1/R$ *then*

$$R_\nu(z) = R(z) - 1/S'(Z_\nu)(Z_\nu - z) \tag{13}$$

also satisfies these postulates and Z_ν is only an apparent singularity of R_ν.

Clearly, R_ν satisfies the differentiability requirements of (1). In order to prove (2), we shall show that the positive semidefinite nature of the matrices of the form (5), (5a) for R_ν with n dimensions follows from the positive semidefinite nature of the similar matrices for R with $n + 1$ dimensions. The corresponding $n + 1$ values of the variable z are $z_1, z_2, \cdots, z_n, Z_\nu + a$. The corresponding quadratic form F will consist of the quadratic form of the n dimensional $\| \kappa_{ik} \|$, given by (5), (5a) proper, plus terms which also involve R or R' at the point $Z_\nu + a$

$$F = \sum_{i,k=1}^{n} \kappa_{ik}\xi_i\xi_k + 2\sum_{k=1}^{n} \frac{R(Z_\nu + a) - R(z_k)}{Z_\nu + a - z_k}\xi_k\xi_{n+1} + R'(Z_\nu + a)\xi_{n+1}^2 \geqq 0. \tag{14}$$

We choose a small enough so that (12a) shall hold, write for

$$\xi_{n+1} = -\sum_{i=1}^{n} \frac{\xi_i a}{Z_\nu - z_i} \tag{14a}$$

and obtain

$$\begin{aligned} F = {} & \sum_{i,k=1}^{n} \kappa_{ik}\xi_i\xi_k - 2\sum_{i,k=1}^{n}\left[\frac{1}{S'(Z_\nu + \vartheta a)a} + R(z_k)\right] \\ & \frac{\xi_k\xi_i a}{(Z_\nu + a - z_k)(Z_\nu - z_i)} + \sum_{i,k=1}^{n} \frac{S'(Z_\nu + a)}{a^2(S'(Z_\nu + \vartheta a))^2}\frac{\xi_i\xi_k a^2}{(Z_\nu - z_i)(Z_\nu - z_k)} \geqq 0. \end{aligned} \tag{14b}$$

In the last term, $R' = S'/S^2$ was substituted and (12a) used for S. As a approaches 0, all the S' in (14b) will approach $S'(Z_\nu)$ and one obtains

$$F = \sum_{i,k=1}^{n}\left[\kappa_{ik} - \frac{1}{S'(Z_\nu)}\frac{1}{(Z_\nu - z_k)(Z_\nu - z_i)}\right]\xi_i\xi_k \geqq 0 \tag{14c}$$

assuming that all z_k are regular points of R. This is, however, just the quadratic form of the matrix, the positive definite nature of which guarantees (2) for the function $R_\nu(z)$ since

$$\frac{R_\nu(z_i) - R_\nu(z_k)}{z_i - z_k} = \frac{1}{z_i - z_k}\left[R(z_i) - R(z_k) - \frac{1}{S'(Z_\nu)}\left(\frac{1}{Z_\nu - z_i} - \frac{1}{Z_\nu - z_k}\right)\right]$$

$$= \kappa_{ik} - \frac{1}{S'(Z_\nu)}\,\frac{1}{(Z_\nu - Z_k)(Z_\nu - z_i)}$$

for $i \neq k$ and a similar equation holds for $R_\nu'(z_k)$. Hence (2) is also fulfilled for R_ν and this also belongs to the class of functions defined by (1) and (2).

In order to see the last part of Theorem 3 we have to show that Z_ν is not a real singularity of R_ν. We have, in fact

$$R_\nu(Z_\nu + a) = \left[-\frac{1}{S'(Z_\nu + \vartheta a)} + \frac{1}{S'(Z_\nu)}\right]\frac{1}{a}.$$

Since S' is continuous $aR_\nu(Z_\nu + a)$ tends to 0 as a approaches 0 and Z_ν must be, by Lemma 3b, an apparent singularity of R_ν. This proves Theorem 3.

IV. Continued fraction expansion of R

The Schiffer-Bargmann lemma suggests the following construction. Take an arbitrary but fixed point ζ from R's definition domain and denote $R(\zeta) = b$ and form the function $R_0(z) = R(z) - b$. This satisfies the postulates (1) and (2) and vanishes at $z = \zeta$. Let us denote its derivative at this point by r_0. It now follows from Lemma 3a that $-1/R_0$ also satisfies postulates (1) and (2) and that it has a singularity at ζ. From the Schiffer-Bargmann lemma it now follows that $-1/R_0$ can be written in the form

$$-\frac{1}{R_0} = \frac{1}{r_0(\zeta - z)} + b_0 + R_1 \tag{15}$$

where R_1 again satisfies our postulates and b_0 is so chosen that $R_1(\zeta) = 0$. One can proceed with R_1 in the same way and arrive at a sequence of functions satisfying our postulates

$$-\frac{1}{R_n} = q_{n+1} + R_{n+1} \qquad (n = 0, 1, 2, \cdots) \tag{15a}$$

where $R_{n+1}(\zeta) = 0$ and q_{n+1} has the form

$$q_{n+1}(z) = \frac{1}{r_n(\zeta - z)} + b_n \qquad (n = 0, 1, 2, \cdots) \tag{15b}$$

with $r_n > 0$. The sequence can break off only if an R_n vanishes—in this case the original R was a rational function of z.

The above equations give R in terms of continued fraction

$$R = b - \cfrac{1}{q_1 - \cfrac{1}{q_2 - \ddots - \cfrac{1}{q_{n+1} + R_{n+1}}}} \tag{16}$$

and define, unless R is rational, an infinite continued fraction representation therefor. The successive approximants[9] $S_n = A_n/B_n$ can be obtained by the well known recursive formulae

$$\begin{aligned} A_{n+1} &= q_{n+1} A_n - A_{n-1} \\ B_{n+1} &= q_{n+1} B_n - B_{n-1} \end{aligned} \tag{17}$$

$A_0 = q_0 = b$, $B_0 = 1$, $A_1 = q_0 q_1 - 1$, $B_1 = q_1$. Because of (15b), A_n and B_n are polynomials of $1/(\zeta - z)$ of degree n. Their leading terms are

$$A_n \approx \frac{b}{r_0 r_1 \cdots r_{n-1}(\zeta - z)^n}; \qquad B_n \approx \frac{1}{r_0 r_1 \cdots r_{n-1}(\zeta - z)^n}. \tag{18}$$

It follows from the relation[10]

$$B_{n+1} A_n - A_{n+1} B_n = 1 \tag{19}$$

that A_n and B_n cannot vanish at the same z so that the approximants never become indeterminate expressions. R can be expressed in terms of the A, B by substituting $q_{n+1} + R_{n+1}$ for q_{n+1} when calculating the $(n + 1)^{\text{th}}$ approximant

$$R = \frac{(q_{n+1} + R_{n+1})A_n - A_{n-1}}{(q_{n+1} + R_{n+1})B_n - B_{n-1}}. \tag{20}$$

This corresponds to the expression (16). By means of (19), this can be given also the form

$$R - \frac{A_n}{B_n} = \frac{1}{B_n[B_n(q_{n+1} + R_{n+1}) - B_{n-1}]}. \tag{20a}$$

Let us assume that all apparent poles of R have been eliminated. If the continued fraction expansion (16) breaks off at any $R_{n+1} = 0$, the R is a rational R function as one easily sees by induction: All q_n are R functions by (15b) and so is, by (15a), R_n. However, if R_m is an R function, so is $q_m + R_m$ and hence also $-1/(q_m + R_m) = R_{m-1}$. Finally, R is an R function. In this case, surely, the definition domain of R can be extended over the whole complex plane with the exclusion of isolated points on the real axis and the imaginary part of $R(z)$

[9] Cf. e.g. H. S. Wall, Analytic Theory of Continued Fractions (Van Nostrand 1948), p. 15. The notation of this book will be used as far as possible.

[10] Page 16, reference 9.

has the same sign as the imaginary part of z. We can restrict ourselves, therefore, to the case in which no R_n becomes identically zero and prove

THEOREM 4. *There is a circle in the complex plane about every regular point ζ of R in which the infinite continued fraction indicated in* (16) *converges uniformly. It represents within that circle, by Weierstrass' theorem, an analytic function.*

Let us first choose a positive number $a < \zeta$ in such a way that R, and hence R_0, have no singularity between $\zeta - a$ and $\zeta + a$. Then R_0 will be negative in the $(\zeta - a, \zeta)$, positive in the $(\zeta, \zeta + a)$ interval. Hence the only singularity of $-1/R_0$ between $\zeta - a$ and $\zeta + a$ will be at ζ and, by (15), R_1 will have again no singularity in this interval. One can conclude, on the basis of (15a), by induction in the same way.

LEMMA 4a. *All R_n are regular and increase monotonically in the $(\zeta - a, \zeta + a)$ interval, assuming the value* 0 *at ζ.*

Lemma 4a will be used to obtain a lower limit for the $| q_{n+1}(z) |$ and $| q_{n+1}(z) + R_{n+1}(z) |$. With (15a), Lemma 4a gives

$$\begin{aligned} q_{n+1}(\zeta - a) + R_{n+1}(\zeta - a) &> 0 \\ q_{n+1}(\zeta + a) + R_{n+1}(\zeta + a) &< 0. \end{aligned} \tag{21}$$

Since the sign of R_{n+1} alone would be just opposite, the same inequalities hold for q_{n+1} fortiori. This gives on account of (15b)

$$| b_n | < \frac{1}{r_n a}. \tag{21a}$$

Hence, in a circle with radius $a' = a/(\alpha + 1)$, about ζ

$$| q_{n+1}(z) | = \left| \frac{1}{r_n(\zeta - z)} + b_n \right| \geqq \frac{\alpha + 1}{r_n a} - | b_n | > \frac{\alpha}{r_n a}. \tag{21b}$$

We note here another inequality which is not necessary for proving Theorem 4 but will be used later. $R_{n+1}(z) < R_{n+1}(\zeta + a)$ for any z between ζ and $\zeta + a$. Furthermore, by (21), $R_{n+1}(\zeta + a)$ is smaller than $| q_{n+1}(\zeta + a) |$. Hence, for $0 < \vartheta < 1$

$$\begin{aligned} | q_{n+1}(\zeta + \vartheta a) + R_{n+1}(\zeta + \vartheta a) | &> | q_{n+1}(\zeta + \vartheta a) - q_{n+1}(\zeta + a) | \\ &= \frac{1 - \vartheta}{\vartheta r_n a}. \end{aligned} \tag{21c}$$

Naturally, the estimate (21c) for $q + R$ applies only on the real axis while (21b) is valid within the circle of radius $a/(\alpha + 1)$.

We shall obtain, finally, an estimate for r_n from (21). The difference of the two inequalities gives, with (15b), directly

$$\frac{R_{n+1}(\zeta + a) - R_{n+1}(\zeta - a)}{2a} < \frac{1}{r_n a^2}. \tag{22}$$

Because of Lemma 2a, the fraction in (22) is larger than $R_{n+1}'(\zeta) = r_{n+1}$. Hence

$$r_n r_{n+1} < 1/a^2. \tag{22a}$$

Thus, (21b) becomes

$$| q_{n+1}(z) q_n(z) | > \frac{\alpha^2}{r_n r_{n-1} a^2} > \alpha^2 \tag{22b}$$

valid for all z within the circle of radius $a' = a/(\alpha + 1)$ about ζ. Similarly, (21c) yields for $0 < \vartheta < 1$

$$\begin{aligned} | q_{n+1}(\zeta + \vartheta a) + R_{n+1}(\zeta + \vartheta a) | \, | q_n(\zeta + \vartheta a) + R_n(\zeta + \vartheta a) | \\ > (1 - \vartheta)^2/\vartheta^2. \end{aligned} \tag{22c}$$

We shall conclude from (22b)

LEMMA 4c. *Given any $\alpha > 2$, there is a circle about ζ in the complex plane such that*

$$\left| \frac{B_n}{B_{n-2}} \right| > \tfrac{1}{2}\alpha^2 - 1. \tag{23}$$

This will again be shown by induction. One sees from (18) that there is a circle in which the lemma holds for $n = 2$ and we choose our circle to lie entirely inside this circle and have a smaller radius than $a/(\alpha + 1)$. The product of two successive equations (17) $B_{n+1} + B_{n-1} = q_{n+1} B_n$ and $B_n + B_{n-2} = q_n B_{n-1}$ gives

$$\frac{B_{n+1}}{B_{n-1}} = \frac{q_n q_{n+1} B_n}{B_n + B_{n-2}} - 1. \tag{23a}$$

Hence, by supposition $| B_{n-2}/B_n | < 2/(\alpha^2 - 2)$ and

$$\left| \frac{B_{n+1}}{B_{n-1}} \right| \geqq \frac{| q_n q_{n+1} |}{1 + | B_{n-2}/B_n |} - 1 \geqq \frac{| q_n q_{n+1} | \, (\alpha^2 - 2)}{\alpha^2} - 1$$

and because of (22b)

$$\left| \frac{B_{n+1}}{B_{n-1}} \right| > \alpha^2 - 2 - 1 > \tfrac{1}{2}\alpha^2 - 1 \tag{23b}$$

since $\alpha > 2$. It now follows that, within the circle of Lemma 4c, the $| B_n |$ increase at least as fast as the geometrical series with the quotient $\frac{1}{2}\alpha^2 - 1$. Since the difference of successive approximants is, by (19)

$$\frac{A_n}{B_n} - \frac{A_{n+1}}{B_{n+1}} = \frac{1}{B_n B_{n+1}} \tag{24}$$

the continued fraction expansion converges, within the circle of Lemma 4c, at least as well as a geometrical series with a quotient $2/(\alpha^2 - 2)$.

THEOREM 5. *The infinite continued fraction indicated in* (16) *converges uniformly in every closed domain which excludes the real axis.*

Actually, this theorem immediately follows from Theorem 4 and Hamburger's theorem.[11] The proof which can be patterned on Jensen's proof of Van Vleck's theorem,[12] is particularly simple in our case.

Let us denote

$$B_{n+1}\bar{B}_n = \tau_n + i\sigma_n \tag{25}$$

and multiply the second equation (17) by the conjugate complex of B_n

$$B_{n+1}\bar{B}_n = q_{n+1}\,|\,B_n\,|^2 - \bar{B}_n B_{n-1}\,.$$

The imaginary part hereof becomes with (15b)

$$\sigma_n = \frac{\eta}{r_n}\,|\,B_n\,|^2 + \sigma_{n-1} \tag{26}$$

where η is the imaginary part of $1/(\zeta - z)$. This will be positive if our domain is above the real axis which we shall assume for the sake of simplicity. Let us denote the minimum of η in our closed domain by η_0, then $0 < \eta_0 \leqq \eta$. Since $\sigma_0 = \eta/r_0$ is already positive, the σ's form an increasing series of positive numbers.

Let us multiply two successive equations (26)

$$(\sigma_n - \sigma_{n-1})(\sigma_{n+1} - \sigma_n) = \frac{\eta^2}{r_n r_{n+1}}\,|\,B_n B_{n+1}\,|^2 > \eta_0^2 a^2\,|\,B_{n+1}\bar{B}_n\,|^2. \tag{26a}$$

The last part follows from (22a). Because of the increasing nature of the σ, the left side is smaller than $(\sigma_{n+1} - \sigma_{n-1})^2$, the right side is, because of (25), larger than σ_n^2. Hence

$$\sigma_{n+1} - \sigma_{n-1} > \eta_0 a\sigma_n > 0. \tag{26b}$$

One now easily concludes by induction that σ increases at least as fast as a geometrical series

$$\sigma_n \geqq \sigma^*(1 + a\eta_0)^{\frac{1}{2}n}. \tag{27}$$

Assuming (27) for σ_n and σ_{n-1} (26b) gives

$$\begin{aligned}\sigma_{n+1} &> \sigma^*[(1 + a\eta_0)^{\frac{1}{2}n}a\eta_0 + (1 + a\eta_0)^{\frac{1}{2}(n-1)}] \\ &> \sigma^*(1 + a\eta_0)^{\frac{1}{2}(n-1)}(a\eta_0 + 1) = \sigma^*(1 + a\eta_0)^{\frac{1}{2}(n+1)}.\end{aligned} \tag{27a}$$

Theorem 5 now follows from (24) just as Theorem 4 did. The same argument applies if the closed domain is below the real axis.

It follows from Weierstrass' theorem and Theorems 4 and 5 that the infinite continued fraction of (16) represents an analytic function over the whole com-

[11] Theorem 27.2, page 114, reference 9.

[12] Cf. O. Perron, Die Lehre von den Kettenbruchen, Teubner 1913, page 264 ff.

plex plane, excluding the parts of the real axis to the left of $\zeta - a/3$ and to the right of $\zeta + a/3$ where a is the distance from ζ of the closest real singularity of R, or ζ itself, whichever is smaller.

We have to prove finally,

THEOREM 6. *The infinite continued fraction indicated in* (16) *is equal to R on the real points of the circle of Theorem* 4.

This follows from (20a), the fact that the $|B_n|$ increase at least as fast as a geometrical series (cf. (23b)) and the fact that among two successive.

$$|B_n(q_{n+1} + R_{n+1}) - B_{n-1}| \tag{28}$$

at least one is larger than 1 if n is large enough. In fact, from

$$\begin{aligned} |q_{n+1} + R_{n+1}| &< |B_{n-1}/B_n| + 1/|B_n| \\ |q_n + R_n| &< |B_{n-2}/B_{n-1}| + 1/|B_{n-1}| \end{aligned} \tag{28a}$$

it would follow by multiplication

$$|q_{n+1} + R_{n+1}|\,|q_n + R_n| < \left(\left|\frac{B_{n-2}}{B_n}\right| + \frac{1}{|B_n|}\right)\left(1 + \frac{1}{|B_{n-1}|}\right). \tag{28b}$$

For $n \to \infty$, the superior limit of the right hand side is equal to that of $|B_{n-2}/B_n|$, which is, by (23), $\leqq 2/(\alpha^2 - 2) \leqq 1$. The left side is, by (22c) and since in the circle in question ϑ is smaller than $\frac{1}{3}$, larger than 4. Hence, from a certain n on.

$$\left|R - \frac{A_n}{B_n}\right| < \frac{1}{|B_n|} \quad \text{or} \quad \left|R - \frac{A_{n+1}}{B_{n+1}}\right| < \frac{1}{|B_{n+1}|}. \tag{29}$$

Actually, because of (24), the fact that one of the expressions (29) tends to zero, shows that both tend to zero. This establishes Theorem 6.

We finally come to

THEOREM 7. *A function R which satisfies the postulates* (1) *and* (2) *can be extended over the complex plane, excepting the negative real axis, as an analytic function. Its singularities are all on the real axis and, unless it is a constant, its imaginary part is positive in the upper, negative in the lower half plane.*

In order to prove the first part of the theorem, we have to show that the analytic continuation of a function which represents R in the $(\zeta - a/3, \zeta + a/3)$ interval, will represent R over the whole real axis. This can be seen, most simply perhaps, by covering the real axis with overlapping intervals in each of which either R or $-1/R$ is analytic. This is indeed possible because ζ was an arbitrary regular point. The fact that R has no singularity at complex z follows most simply from Theorem 6 and Weierstrass' theorem. Finally, all approximants of the continued fraction expansion of R have the property of mapping the upper and lower half planes on themselves. Hence, the imaginary part of the function to which they converge must be non negative in the upper, non positive in the lower half plane. Since the imaginary part of an analytic function, which is not a constant, cannot assume its minimum (or maximum) in the interior of a closed

domain, the imaginary part of R cannot vanish except on the real axis. This completes the proof of Theorem 7.

The above development gives R as a limit of meromorphic functions which map the upper and lower half planes on themselves (called R-functions by the writer)[8]. It is worthwhile to point out therefore

THEOREM 8. *All R functions satisfy the postulates* (1) *and* (2). *The same applies to limits of such functions which have only isolated singularities on the positive real axis.*

It is clear that the functions so defined meet the differentiability criteria (1). If R is itself an R function, it permits an expansion of the form[8]

$$R = \alpha z + \beta + \sum \left(\frac{\gamma_\mu^2}{Z_\mu - z} - \frac{\gamma_\mu^2}{Z_\mu} \right) \tag{30}$$

with non negative α, γ_μ and real β, Z_μ. In order to show that (30) satisfies postulate (2) also, one can show that all its terms do so, i.e., give positive semi-definite matrices $\| \kappa_{ik} \|$ when substituted for R in (5), (5a). This is evident: $\alpha z + \beta$ gives a matrix, all elements of which are $\alpha \geqq 0$, and a typical term under the summation sign gives the matrix $\| \kappa_{ik} \|$

$$\kappa_{ii} = \frac{\gamma_\mu^2}{(Z_\mu - z_i)^2} \tag{30a}$$

$$\kappa_{ik} = \frac{\gamma_\mu^2}{z_i - z_k} \left(\frac{1}{Z_\mu - z_i} - \frac{1}{Z_\mu - z_k} \right) = \frac{\gamma_\mu^2}{(Z_\mu - z_i)(Z_\mu - z_k)} \tag{30b}$$

which is also positive semidefinite. It is also clear that the limits of such functions also satisfy postulate (2). Hence, e.g.,

$$\int_{-2}^{-1} \frac{dZ}{Z - z} = \ln \frac{z + 2}{z + 1} \tag{31}$$

although clearly not an R function, satisfies postulates (1) and (2).

The above considerations can be carried out also if R is defined not on the positive real axis, but any connected part of the real axis. The partial fraction expansion of R functions, and of functions which are limits of R functions (i.e., the functions of reference[3]), has been given already by Markoff.[13] The above considerations, as well as Loewner's, use the partial fraction expansion, conversely, to show that the functions satisfying our postulates are functions of this nature.

It may be well to reiterate that the preceding considerations are less far-reaching, and that they solve a different problem, than Loewner's original work. As to assumptions: we explicitly assume differentiability for all positive z, except isolated points. Although, in order to carry out the above considerations, it is necessary only to assume differentiability in an arbitrarily small interval, this is nevertheless more than Loewner has assumed. As to results: we consider only the case that the $\| \kappa_{ik} \|$ of (5) is a definite matrix for all dimensions, while Loew-

[13] Reference 12, page 385. We are much indebted to Dr. H. S. Wall for this quotation.

ner also considers, and solves, the problem of the matrices $\| \kappa_{ik} \|$ being definite up to a certain n only. Similarly, our problem is different from Loewner's. He is interested in monotonic matrix functions which lead to our postulate (2). The purpose of the above considerations is to obtain the criteria for functions satisfying postulate (2) which was obtained in a different way, and the validity of which extends across the singularities of R. This last condition is responsible for the necessity of much of Section II and of the last part of the present Section.

PRINCETON UNIVERSITY
THE INSTITUTE FOR ADVANCED STUDY

Characteristic Vectors of Bordered Matrices with Infinite Dimensions I

E. P. Wigner

Annals of Mathematics *62*, 548–564 (1955)

(Received April 18, 1955)

Introduction

The statistical properties of the characteristic values of a matrix the elements of which show a normal (Gaussian) distribution are well known (cf. [6] Chapter XI) and have been derived, rather recently, in a particularly elegant fashion.[1] The present problem arose from the consideration of the properties of the wave functions of quantum mechanical systems which are assumed to be so complicated that statistical considerations can be applied to them. Since the physical problem has been given rather recently in some detail in another journal [3], it will not be reviewed here. Actually, the model which underlies the present calculations shows only a limited similarity to the model which is believed to be correct. Nevertheless, the calculation which follows may have some independent interest; it certainly provided the encouragement for a detailed investigation of the model which may reproduce some features of the actual behavior of atomic nuclei.

All the remaining work will deal with real symmetric matrices of very high dimensionality. The first and last problems concern infinite bordered matrices; the second one a finite matrix the consideration of which served as an intermediate step toward the solution of the last one. We mean by a bordered matrix the sum of a diagonal matrix $\mathbf{k}$ and a border matrix $\mathbf{v}$. The diagonal elements of $\mathbf{k}$ are all the integers $\cdots, -2, -1, 0, 1, 2, \cdots$. The border matrix $\mathbf{v}$ has non vanishing elements only up to a distance N from the diagonal, the absolute value of all the non vanishing elements is the same

$$(1)\qquad \begin{aligned} |v_{mn}| &= v && \text{for } |m-n| \leqq N, (-\infty < m, n < \infty) \\ &= 0 && \text{for } |m-n| > N. \end{aligned}$$

Since the matrix $H = \mathbf{k} + \mathbf{v}$ is symmetric, $v_{mn} = v_{nm}$. Subject to this condition, however, the signs of the v_{ij} are random, i.e. we consider ensembles of matrices with all possible signs of v_{mn} subject to the conditions of symmetry. In the first of the problems considered $N = 1$, in the third one both N and v are very large in such a way, however, that $v^2/N = q$ remains limited. The first problem will be solved completely, i.e. the characteristic values and vectors given explicitly.

In order to formulate the third problem, we denote the characteristic values and vectors of an H by λ and $\psi^{(\lambda)}$

$$(2)\qquad H\psi^{(\lambda)} = \lambda\psi^{(\lambda)} \quad \text{or} \quad \sum_n H_{mn}\psi_n^{(\lambda)} = \lambda\psi_m^{(\lambda)}.$$

[1] Personal communication of Professor V. Bargmann.

We shall inquire then for the expectation value of those

$$(\psi_0^{(\lambda)})^2 \tag{3}$$

the λ of which is in unit interval at x which will be considered to be a continuous variable. One can imagine this to be calculated by solving the characteristic value problem (2) for each of the permissible sign combinations of the v_{mn}, choosing those which have a characteristic value λ between x and $x + \delta$, calculating (3) for these H and λ, adding the expressions obtained in this way and dividing the result by δ times the number of all H considered. The expectation value obtained in this way will be denoted by $\sigma(x)$, it is the quantity which we inquired for. An integral equation will be obtained for σ and some of its properties, but not σ itself, will be obtained explicitly. However, the mathematical analysis which leads to the integral equation for σ will lack rigor; in particular the convergence of the procedure which defines σ will not be proved. Many important statistical properties of the characteristic functions and characteristic values of H can easily be obtained from σ. The fact that the density of the characteristic values of σ is 1 follows easily from the invariance properties of the set of permissible H.

The second problem concerns a finite symmetric matrix H_0 albeit of very large dimension. Its diagonal elements are zero, the off-diagonal elements $|v_{mn}| = v$ are real but have random signs as before, m and n assume only the values $-N, -N+1, \cdots, N-1, N$ so that this H is $2N+1$ dimensional. The problem is again to calculate $\sigma(x)$ as defined above for the third problem. The averages have to be taken in this case only over $2^{N(2N+1)}$ matrices and the fact that $\sigma(x/(N)^{1/2}\, v)$, properly normalized, converges to a limiting function can be proved in this case. This limiting function will be explicitly determined. It also gives the density of the characteristic values of H_0.

Singly bordered symmetric matrix

We shall calculate first the characteristic vectors of

$$H_1 = \begin{Vmatrix} & \cdot & \cdot & \cdot & \cdot & \cdot & \\ & \cdot & \cdot & \cdot & \cdot & \cdot & \\ & \cdot & \cdot & \cdot & \cdot & \cdot & \\ \cdots & -2 & v_{-1-2} & 0 & 0 & 0 & \cdots \\ \cdots & v_{-1-2} & -1 & v_{-10} & 0 & 0 & \cdots \\ \cdots & 0 & v_{-10} & 0 & v_{01} & 0 & \cdots \\ \cdots & 0 & 0 & v_{01} & 1 & v_{12} & \cdots \\ \cdots & 0 & 0 & 0 & v_{12} & 2 & \cdots \\ & \cdot & \cdot & \cdot & \cdot & \cdot & \\ & \cdot & \cdot & \cdot & \cdot & \cdot & \\ & \cdot & \cdot & \cdot & \cdot & \cdot & \end{Vmatrix} \tag{4}$$

in which the absolute values of all v are the same. Neither the characteristic values nor the squares of the components of the characteristic vectors will be changed if H_1 is transformed by a diagonal matrix s all the diagonal elements of which are $s_k = \pm 1$. Such a transformation will leave the diagonal elements of (4) unchanged but will replace v_{kk+1} by $s_k s_{k+1} v_{kk+1}$. By choosing $s_0 = 1$, s_k to have the sign of $-v_{k-1,k}$ for $k > 0$ and the sign of $-v_{k,k+1}$ for $k < 0$, all non diagonal elements will become negative, their value will be denoted by $-v$.

The resulting matrix $\hat{H}_1$ can be transformed into $\hat{H}_1 + 1$ by renumbering the rows and columns. Hence, the set of characteristic values λ_k will be unchanged if one replaces each λ_k by $\lambda_k + 1$. Furthermore, $\hat{H}_1$ can be transformed into $-\hat{H}_1$ by interchanging the kth and $-k^{\text{th}}$ rows and columns and transforming it by an s with $s_k = 1$ for even k and $s_k = -1$ for odd k. Hence the set of characteristic values contains, with every λ_k, also $-\lambda_k$. Since the λ_k will change continuously as v is increased and since for $v = 0$ the set of λ_k consists of all integers, this will be true for all v. Without loss of generality, one can set

$$\lambda_k = k. \tag{4a}$$

The characteristic vector of k shall be denoted by $\psi^{(k)}$. It then follows from the remarks about transforming $\hat{H}_1$ into $\hat{H}_1 + 1$ that $\psi_l^{(k)} = \psi_{l+1}^{(k+1)}$ i.e. the lth component of $\psi^{(k)}$ depends only on the difference $l - k$

$$\psi_l^{(k)} = \psi_{l-k}^{(0)} = \psi_{l-k} \tag{4b}$$

where, for simplicity $\psi^{(0)}$ has been replaced by ψ. It finally follows from the remark about the transformation of H_1 into $-H_1$ that

$$\psi_{-l}^{(0)} = c(-)^l \psi_l^{(0)}.$$

Clearly $c = \pm 1$ and again by continuity from the $v = 0$ case, $c = 1$. This gives

$$\psi_{-l} = (-)^l \psi_l\,. \tag{4c}$$

The equation $\hat{H}_1 \psi = 0$ reads explicitly

$$-v\psi_{l-1} + l\psi_l - v\psi_{l+1} = 0. \tag{5}$$

This reminds one of the recursive formula for Bessel functions, ([5], 17.21, p. 359)

$$-J_{l-1}(z) + (2l/z)J_l(z) - J_{l+1}(z) = 0. \tag{5a}$$

Hence it follows that one can set

$$\psi_l = J_l(2v) = \psi_{l+k}^{(k)}\,. \tag{5b}$$

This satisfies all conditions since for large l

$$\psi_l \approx v^l/l! \approx (ev)^l l^{-l}, \tag{5c}$$

ψ_k goes to zero fast enough for $\sum \psi_l^2 < \infty$. We note that the orthogonality of $\psi^{(0)}$ and $\psi^{(k)}$ gives

$$\textstyle\sum_l J_l(2v) J_{l-k}(2v) = 0 \qquad k \neq 0. \tag{6}$$

The summation runs over all integers from $-\infty$ to ∞. The essential result of the calculation is (5c) which shows how ψ_l decreases as $l \to \infty$. Because of (4c), the decrease is essentially the same as $l \to -\infty$.

It should be noted that one can obtain a formal solution of the equation $H_1\varphi = \lambda\varphi$ for any λ. It is necessary for this purpose only to set

$$\varphi = Z_{l-\lambda}(2v)$$

where the $Z_l = \alpha J_l + \beta N_l$ are all the same linear combinations of regular and irregular Bessel functions. However, the φ obtained in this way cannot be normalized.

Preliminary remarks on the strength function

We shall consider, below, very large sets of real symmetric matrices. Each of these will have certain characteristic values λ; the corresponding real normalized characteristic vector will be denoted by $\psi^{(\lambda)}$, its components by $\psi_k^{(\lambda)}$. The orthogonality relations of the $\psi_k^{(\lambda)}$ are

$$\sum_\lambda \psi_k^{(\lambda)}\psi_l^{(\lambda)} = \delta_{kl}\,; \qquad \sum_k \psi_k^{(\lambda)}\psi_k^{(\mu)} = \delta_{\lambda\mu}\,. \tag{7}$$

We next define a function $S(x)$ as follows. Let us consider all the characteristic values λ of all the matrices of the set which are below x. Their number shall be denoted by $\mathfrak{N}\Lambda(x)$ where $\mathfrak{N}$ is the number of matrices in the set. Let us consider then the sum of the squares of a particular component $k = 0$ of all the characteristic vectors $\psi^{(\lambda)}$ for which $\lambda < x$. Then

$$\mathfrak{N}S(x) = \sum_{\text{set}} \sum_{\lambda<x} (\psi_0^{(\lambda)})^2, \tag{8}$$

the summation is extended over all matrices of the set and all the characteristic vectors of these which satisfy the condition $\lambda < x$. The right side is clearly independent of the sign and other arbitrarinesses in the definition of the characteristic vectors should multiple characteristic values occur. Because of the normalization relations (7), if $x \to \infty$, the sum over λ gives 1 for each matrix separately and one finds

$$S(\infty) = 1. \tag{8a}$$

We shall calculate now the moments

$$M_\nu = \int_{-\infty}^{\infty} x^\nu \, dS(x) = \mathfrak{N}^{-1} \sum_{\text{set}} \sum_\lambda \lambda^\nu (\psi_0^{(\lambda)})^2 \tag{9}$$

the summation is to be extended over all characteristic vectors of all matrices of the set. If a typical matrix of the set is denoted by $H = (H_{mn})$ one can write

$$\sum_n H_{mn}\psi_n^{(\lambda)} = \lambda\psi_m^{(\lambda)}$$

and repeated application of H to both sides gives

$$\sum_n (H^\nu)_{mn}\psi_n^{(\lambda)} = \lambda^\nu\psi_m^{(\lambda)}.$$

Multiplication with $\psi_k^{(\lambda)}$ and summation over λ yields by means of (7) the well known equation

$$(H^\nu)_{mk} = \sum_\lambda \lambda^\nu \psi_m^{(\lambda)} \psi_k^{(\lambda)}. \tag{9a}$$

Setting $m = k = 0$ herein and summing over all matrices of the set gives

$$M_\nu = \mathfrak{N}^{-1} \sum_{\text{set}} (H^\nu)_{00} = \text{Av}(H^\nu)_{00} . \tag{9b}$$

Av will denote the average of the succeeding expression over all matrices of the set.

The M_ν will be calculated in the following section for a certain set of matrices in the limiting case that the dimension $2N + 1$ of these matrices becomes infinite. It will be shown, then, that $S(x)$, which is a step function for every finite N, becomes a differentiable function and its derivative $S'(x) = \sigma(x)$ will be called the strength function. In the last section, infinite sets of infinite matrices will be considered. However, all powers of these matrices will be defined and $(H^\nu)_{00}$ involves, for every ν, only a finite part of the matrix. It will be seen that the definition of the average of this quantity for the infinite set of H does not involve any difficulty. However, a similar transition to a limiting case $N \to \infty$ will be carried out with this set as with the aforementioned set and this transition will not be carried through in a rigorous manner in either case.

The expression "strength function" originates from the fact that the absorption of an energy level depends, under certain conditions, only on the square of a definite component of the corresponding characteristic vector. This component was taken, in (8), to be the 0 component. Hence $S(x_1) - S(x_2)$ is the average strength of absorption by all energy levels in the (x_1 , x_2) interval.

Random sign symmetric matrix

The matrices to be considered are $2N + 1$ dimensional real symmetric matrices; N is a very large number. The diagonal elements of these matrices are zero, the non diagonal elements $v_{ik} = v_{ki} = \pm v$ have all the same absolute value but random signs. There are $\mathfrak{N} = 2^{N(2N+1)}$ such matrices. We shall calculate, after an introductory remark, the averages of $(H^\nu)_{00}$ and hence the strength function $S'(x) = \sigma(x)$. This has, in the present case, a second interpretation: it also gives the density of the characteristic values of these matrices. This will be shown first.

Let us consider one of the above matrices and choose a characteristic value λ with characteristic vector $\psi^{(\lambda)}$. Clearly, λ will be a characteristic value also of all those matrices which are obtained from the chosen one by renumbering rows and columns. However, the components $\psi_k^{(\lambda)}$ of the corresponding characteristic vectors will be all possible permutations of the components of the original matrix' characteristic vector. It follows that if we average $(\psi_k^{(\lambda)})^2$ over the aforementioned matrices, the result will be independent of k. Because of the normalization condition (7), it will be equal to $1/(2N + 1)$.

Let us denote now the average number of characteristic values of the matrices

of our set which are smaller than x by $\Lambda(x)$. More accurately, this is the number of all characteristic values of all matrices of our set which are smaller than x, divided by the number of matrices in our set. Clearly $\Lambda(-\infty) = 0$, $\Lambda(\infty) = 2N + 1$. Since every characteristic value λ which is common to n_λ of the matrices contributes $n_\lambda/\mathfrak{N}$ to $\Lambda(x)$ if $\lambda < x$ and since it contributes $\mathfrak{N}^{-1}n_\lambda/(2N + 1)$ to $S(x)$, we have

$$\Lambda(x) = (2N + 1)S(x). \tag{10}$$

This is true for any value of N, as $N \to \infty$, the step functions $\Lambda(x)$ and $S(x)$ become differentiable and the density of characteristic values per unit interval becomes

$$\Lambda'(x) = (2N + 1)\sigma(x). \tag{10a}$$

Calculation of the moments

We now proceed with the calculation of the average of $(H^\nu)_{kk}$ or $(H^\nu)_{00}$ since the former quantity is independent of k. We first note that the diagonal elements of H^2 are

$$\sum_k v_{0k}v_{k0} = \sum v_{0k}^2 = 2Nv^2. \tag{11}$$

Hence the average of this quantity is also

$$M_2 = 2Nv^2. \tag{11a}$$

Let us consider now the expression for the νth moment

$$M_\nu = \mathfrak{N}^{-1} \sum_{\text{set}} \sum_{i_1 \cdots i_{\nu-1}} v_{0i_1}v_{i_1i_2} \cdots v_{i_{\nu-1}0}\,. \tag{12}$$

Since the diagonal elements of our matrix vanish, the summation in (12) can be restricted to those sequences $0, i_1, i_2, \cdots, i_{\nu-1}, 0$ in which no two succeeding members are equal. A pair of unequal numbers jl, each of which characterizes a row or column of our matrices (i.e. for which $-N \leqq j, l \leqq N$) will be called a "step"; the step lj will be called the reverse of jl. The sequence $0, i_1, i_2, \cdots, i_{\nu-1}, 0$ contains ν steps: $0i_1, i_1i_2, i_2i_3, \cdots, i_{\nu-1}0$. Each matrix of our set attributes a sign to every step but the signs of a step and of its reverse are always the same.

Set summation and the summation over the i can be interchanged in (12). The set summation will give 0 for each sequence $0, i_1, i_2, \cdots, i_{\nu-1}, 0$ unless it contains every step jl and its reverse an even number of times. If it does and if its first and last members are the same, it will be called a valid sequence. Thus, $0, i, j, i, l, m, n, m, n, m, l, i, 0$ is a valid sequence for $\nu = 12$ if the i, j, l, m, n, are all different. For a valid sequence, the set summation of (12) gives $\mathfrak{N}v^\nu$. Hence M_ν is equal to v^ν times the number of valid sequences of length $\nu + 1$.

Clearly, there is no valid sequence if ν is odd. Hence, $M_{2\nu+1} = 0$ and we can restrict ourselves to the calculation of the $M_{2\nu}$, i.e. the number of valid sequences of length $2\nu + 1$. A typical valid sequence of length $2\nu + 1$ is $0, i_1, i_2, \cdots,$

$i_{\nu-1}, i_\nu, i_{\nu-1}, \cdots, i_2, i_1, 0$, no matter what the i are as long as no two successive members of the sequence are equal. There are $(2N)^\nu$ valid sequences of this type and it will be shown that the leading term in the expression for the total number of all valid sequences of length $2\nu + 1$ is a numerical multiple of this, which will be denoted by $t_\nu(2N)^\nu$. The accurate expression for the number of valid sequences contains terms with lower powers of N but we shall assume that N is so large that we can restrict ourselves to the leading term. Hence

$$M_{2\nu} = t_\nu(2N)^\nu v^{2\nu}. \tag{13}$$

Because of (11a), $t_1 = 1$ and we proceed to calculate the general t_ν.

Excluding the first member, there cannot be more than ν different members in a valid sequence of length $2\nu + 1$. Let us call the steps "free" which end in a member which did not occur in the sequence before and let us call the steps "repetitive" which end in a member which did occur before. The sequence can be valid only if the number of repetitive steps is at least equal to the number of free steps. Since the number of free steps is equal to the number of different members of the sequence, the latter number cannot be more than one half of the number 2ν of all steps. This proves the assertion.

The number of valid sequences which contain less than ν different members is proportional to $N^{\nu-1}$ and their number is, therefore, negligible. The total number of sequences with only $\nu - 1$ different members is less than $(2\nu)!\,(2N + 1)^{\nu-1}$ and the number of valid sequences is smaller than this.

It follows from the preceding observation that it will be sufficient to count those valid sequences which have ν different members, i.e. contain ν free steps. It is permissible, therefore, to call only those sequences valid which contain ν different members. Since the number of repetitive steps in these sequences is also ν and since every free step must be repeated (in the same or the opposite direction) if the sequence is to be valid, every repetitive step will in fact be equal to a free step or to the reverse thereof. This justifies the name "repetitive" for the not free steps of a valid sequence.

A valid sequence in the above restricted sense defines a type sequence. The type sequence has 2ν members, its μth member is the number of the free steps minus the number of the repetitive steps among the first μ steps of the sequence to which it is coordinated. Thus, for instance, the sequence 0, 5, 3, 5, 0 and 0, 3, 0, -2, 0 are both valid sequences for $2\nu = 4$ (if $N \geqq 5$). The type sequence of the former is 1, 2, 1, 0, that of the latter is 1, 0, 1, 0. *All type sequences start with* 1 *and end with* 0, *successive members of it differ by* ± 1.

There are exactly $2N(2N - 1) \cdots (2N - \nu + 1) = (2N)!/(2N - \nu)! \sim (2N)^\nu$ *valid sequences for every type sequence.* If the μth member of the type sequence is larger than the preceding one, the corresponding step in the valid sequence is free and may lead to any number $j(-N \leqq j \leqq N)$ which has not been used before. If, on the other hand, the μth member of the type sequence is smaller than the preceding member, the μth step of the valid sequence is a repetitive one and must be the reverse of the step which originally led to the element from which this step starts. Hence, in the latter case, the μth step of the valid

sequence is completely determined. (It follows that in a valid sequence in the restricted sense the steps are not actually repeated but balanced by the reverse step. This fact is not material for the remainder of the calculation.) *It follows that t_ν in* (13) *is the number of type sequences of length* 2ν.

Before obtaining a recursive formula for t_ν, we note that the only two type sequences for $2\nu = 4$ are the ones given above: 1, 2, 1, 0 and 1, 0, 1, 0 the corresponding valid sequences have the form 0, i, j, i, 0 and 0, i, 0, j, 0 with $i \neq j$. Hence $t_2 = 2$. For $2\nu = 6$ the type sequences are 1, 2, 3, 2, 1, 0; 1, 2, 1, 2, 1, 0; 1, 2, 1, 0, 1, 0; 1, 0, 1, 2, 1, 0; 1, 0, 1, 0, 1, 0. Hence, $t_3 = 5$. The valid sequences of the second type are 0, i, j, i, l, i, 0 with i, j, l all different.

The number of type sequences which contain no 0 before the last member will be denoted by t_ν'. From these sequences, one can obtain a type sequence of length $2\nu - 2$ by omitting the first and last member and subtracting 1 from each remaining member. Hence

$$t_\nu' = t_{\nu-1}\,. \qquad (t_1' = t_0 = 1). \tag{14}$$

If the first 0 member of the type sequence is at the position 2κ (it must be at an even position) the first 2κ members of it form a 0 free type of length 2κ, the remainder an arbitrary type sequence of length $2\nu - 2\kappa$. Hence

$$t_\nu = \sum_{\kappa=1}^{\nu} t_\kappa' t_{\nu-\kappa} = \sum_{\kappa=1}^{\nu} t_{\kappa-1} t_{\nu-\kappa}\,. \qquad (\nu = 1, 2, 3, \cdots). \tag{14a}$$

These recursive equations permit the successive calculation of the t_ν. One can obtain a closed formula for them by writing

$$t(x) = \sum_{\nu=0}^{\infty} t_\nu x^\nu. \tag{15}$$

The recursive formula (14a) then gives

$$t(x) = 1 + xt(x)^2. \tag{15a}$$

The 1 on the right side is necessary because (14a) is not valid for $\nu = 0$. It follows that

$$t(x) = (2x)^{-1}(1 \pm (1 - 4x)^{\frac{1}{2}}). \tag{15b}$$

Actually, the lower sign has to be taken. It gives

$$t_\nu = \frac{1}{2}\binom{\frac{1}{2}}{\nu + 1}(-4)^{\nu+1} = \frac{(2\nu)!}{\nu!(\nu + 1)!}. \tag{15c}$$

This can also be proved by induction by means of (14a). Finally, by (13)

$$M_{2\nu} = \frac{(2\nu)!}{\nu!(\nu + 1)!}(2N)^\nu v^{2\nu}. \tag{16}$$

Calculation of the strength function

It follows from (9) and (15) that

$$\int_{-\infty}^{\infty} x^{2\nu}\, dS(x) = \int_{-\infty}^{\infty} x^{2\nu}\sigma(x)\, dx = \frac{2\nu!}{\nu!(\nu + 1)!}(2N)^\nu v^{2\nu}. \tag{17}$$

The fact that $S(x)$ is differentiable and that, therefore, the first integral can be replaced by the second one ($\sigma(x) = S'(x)$) will be a consequence of (17). From the vanishing of the odd moments $M_{2\nu+1}$ it follows that $\sigma(x)$ is an even function of x. The form of the even moments suggests the introduction of new variables

$$\rho(\xi) = N\sigma(N\xi) \tag{18}$$

in terms of which (17) reads

$$\int \xi^{2\nu}\rho(\xi)\, d\xi = \frac{(2\nu)!\, 2^{\nu}}{\nu!(\nu+1)!}\frac{v^{2\nu}}{N^{\nu}} = \frac{(2\nu)!(2q)^{\nu}}{\nu!(\nu+1)!} \tag{18a}$$

where $q = v^2/N$.

Professor W. Feller kindly pointed out to the writer that the analysis of the original manuscript leading to (20) can be simplified by calculating directly (cf. also [2], Chapters 14 and 16),

$$\begin{aligned} \int e^{ik\xi}\rho(\xi)\, d\xi &= \sum_{\nu=0}^{\infty} \int \frac{(ik\xi)^{2\nu}}{(2\nu)!}\rho(\xi)\, d\xi \\ &= \sum_{\nu} \frac{(-)^{\nu}(2qk^2)^{\nu}}{\nu!(\nu+1)!} = \frac{2J_1(q'k)}{q'k} \end{aligned} \tag{19}$$

where $q' = (8q)^{\frac{1}{2}}$. The second member of (19) follows from the vanishing of the odd moments, the third member is the well known series (see [5], 17.1, p. 355) for the Bessel function of order 1. In order to obtain ρ itself, the Bessel function must be represented as a Fourier integral. Such a representation is provided by the expression (see [5], Example 1, p. 366),

$$\begin{aligned} J_1(z) &= \frac{\frac{1}{2}z}{\Gamma(3/2)\Gamma(1/2)} \int_0^{\pi} e^{iz\cos\varphi} \sin^2\varphi\, d\varphi \\ &= \frac{z}{\pi}\int_{-1}^{1} e^{izw}(1-w^2)^{\frac{1}{2}}\, dw. \end{aligned} \tag{19a}$$

The last part was obtained by substituting w for $\cos\varphi$. Substitution of $q'k$ for z and ξ for $q'w$ then gives (remembering that $q'^2 = 8q$)

$$\frac{2J_1(q'k)}{q'k} = \frac{1}{4\pi q}\int_{-q'}^{q'} e^{ik\xi}(8q-\xi^2)^{\frac{1}{2}}\, d\xi. \tag{19b}$$

Comparison of (19b) with (19) yields

$$\begin{aligned} \rho(\xi) &= (4\pi q)^{-1}(8q-\xi^2)^{\frac{1}{2}} && \text{for } \xi^2 < 8q \\ &= 0 && \text{elsewhere.} \end{aligned} \tag{20}$$

The original analysis did not make use of the properties of Bessel functions. It showed, on the basis of (18a), that the moments of $(8q/\xi)\, ds/d\xi$ and of $\xi\, ds/d\xi - 3s$ are equal where $ds/d\xi = \xi\rho$. This led to the differential equation

$$(8q/\xi)\, ds/d\xi = \xi\, ds/d\xi - 3s \tag{19c}$$

and hence to (20). The integration constant was determined from the normalization of ρ, i.e. from (18a) with $\nu = 0$. It was necessary then to refer to Corollary 1.1 in Shohat and Tamarkin's book [4], p. 11, or to Carleman's general theorem [1], p. 115, or [4], p. 19, to infer from the asymptotic form of the right side of (18a) that (20) is the only solution of (18a). This also justified (17). In any case, it is easy to verify a posteriori that the ρ of (20) satisfies (17).

For $\sigma(x)$, (18) gives

$$\sigma(x) = \frac{(8Nv^2 - x^2)^{\frac{1}{2}}}{4\pi Nv^2} \text{ (for } -(8N)^{\frac{1}{2}}v < x < (8N)^{\frac{1}{2}}v). \tag{20a}$$

This gives both the distribution of the characteristic values of the random sign symmetrical real matrices defined at the beginning of this section and also their strength function. Since, on the average, all the components of the characteristic vectors have the same absolute value, the two distributions are naturally identical. The reader will notice that even though some of the matrices of the set considered have characteristic values as high as $2Nv$, the characteristic values in excess of $(8N)^{\frac{1}{2}}v$ become increasingly rare as $N \to \infty$ and their total number constitutes, in the limit, a negligible fraction of all characteristic values of the matrices of the set.

Infinite symmetric matrices with wide random borders

The set of matrices $H = \mathbf{k} + \mathbf{v}$ underlying the following calculation has been described in some detail in the introduction. Of the two, $\mathbf{k}$ is unbounded but its characteristic vectors are clearly the unit vectors parallel to the coordinate axes; $\mathbf{v}$ is bounded for all N, its bounds being $\pm 2Nv$. Hence the characteristic value problem of H is defined.

Some of the remarks which apply for the singly bordered matrix also apply to the present *set* of matrices. In particular, $\mathbf{k}$ can be transformed into $-\mathbf{k}$ by renumbering rows and columns and the set of matrices $\mathbf{v}$ is also invariant under this transformation. It follows that the average number of characteristic values λ per unit interval at x is an even function of x. It also follows that if $\psi_k^{(\lambda)}$ is the characteristic vector of one matrix of the set, the set contains another matrix with characteristic vector $\psi_{-k}^{(-\lambda)}$. Hence, if $\mathfrak{N}$ is the number of matrices of the set which we consider and if this set of matrices includes either both or neither of the aforementioned matrices

$$\mathfrak{N}S(x) = \sum_{\text{set}} \sum_{\lambda<x} (\psi_0^{(\lambda)})^2 = \sum_{\text{set}} \sum_{\lambda>-x} (\psi_0^{(\lambda)})^2. \tag{21}$$

Because of (7), the sum of the expression in (21), extended over all λ, is just $\mathfrak{N}$. Hence, the last sum is also equal to

$$\mathfrak{N}S(x) = \mathfrak{N} - \sum_{\text{set}} \sum_{\lambda<-x} (\psi_0^{(\lambda)})^2 = \mathfrak{N} - \mathfrak{N}S(-x) \tag{21a}$$

so that

$$S(x) + S(-x) = 1 \tag{21b}$$

and the strength function $\sigma(x) = S'(x)$ is again even.

Second, it is again possible to transform $\mathbf{k}$ into $\mathbf{k} + 1$ by renumbering the rows and columns and the set of matrices $\mathbf{v}$ is also invariant under this transformation. Hence, the average number of characteristic values λ per unit interval at x is a periodic function of x with period 1.

It may be well to repeat here that we shall be interested, eventually, in a very wide border, i.e. in the case that N in (1) is very large. At the same time v shall be very large also, in such a way, however, that $v^2/N = q$ remains constant. The problem originates from the consideration of an "unperturbed problem" in which the spacing of the characteristic values is $1/N$, the matrix of the unperturbed problem being $N^{-1}\mathbf{k}$. The "perturbation" $N^{-1}\mathbf{v}$ has matrix elements connecting characteristic values of the unperturbed problem which differ up to 1 (this quantity being chosen as the unit of energy). The diagonal elements of the square of the perturbation $N^{-2}v^2$ are $2NN^{-2}v^2$ and this quantity is denoted by $2q$.

Calculation of the moments

Because of the even nature of the strength function, it will suffice to calculate the even moments thereof. Their calculation will be based again on (9b), i.e. on the calculation of the set average of $(H^{2\nu})_{00}$. Before carrying out the calculation, it should be noted that

$$(H^{2\nu})_{00} = \sum_n (H^\nu)_{0n}(H^\nu)_{n0} = \sum_n (H^\nu)^2_{0n} . \tag{22}$$

It will be shown that $(H^\nu)_{0n}$ is different from zero only for $-\nu N \leqq n \leqq \nu N$ and that only those v_{mn} influence it for which $-\nu N \leqq m, n \leqq \nu N$. These statements are evident for $\nu = 1$. Since

$$(H^\nu)_{0n} = \sum_m (H^{\nu-1})_{0m}(m\delta_{mn} + v_{mn}) \tag{22a}$$

and since $v_{mn} = 0$ for $| m - n | > N$, the statements follow by induction. They show that, when calculating the set average of $(H^{2\nu})_{00}$, the average has to be taken only over a finite set of matrices, namely those in which the v_{mn} for $-2\nu N \leqq m, n \leqq 2\nu N$ have all possible signs consistent with the condition of symmetry.

The calculation of $M_{2\nu} = \mathrm{Av}(H^{2\nu})_{00}$ will closely resemble the calculation of the preceding section. It will turn out again that the term in $M_{2\nu}$ which is proportional to the highest power of N has a factor $(v^2N)^\nu \approx N^{2\nu}$ and only the terms proportional to $N^{2\nu}$ will be calculated. The proportionality constant of $N^{2\nu}$ will depend on q and we shall write

$$M_{2\nu} = \mathrm{Av}(H^{2\nu})_{00} = T_\nu(q)N^{2\nu}. \tag{23}$$

We use again an expression similar to (12) for

$$M_{2\nu} = \mathrm{Av} \sum_{i_1 i_2 \cdots i_{2\nu-1}} H_{0i_1}H_{i_1 i_2}\cdots H_{i_{2\nu-1}0} \tag{24}$$

and define the steps $0i_1$, i_1i_2, i_2i_3, $\cdots$, $i_{2\nu-1}0$ as before. However, there are, in this case, three kinds of steps: (1) the free steps which lead to an i which has

not occurred before in the sequence $0i_1i_2i_3 \cdots i_{2\nu-1}0$, (2) the repetitive steps which occurred before in the same or in the reverse direction, and (3) the waiting steps which correspond to diagonal elements of the matrix. It is possible again to define type sequences each member of which differs from the preceding one by ± 1 or 0. If the μth member of the type sequence is larger than the preceding one, the corresponding step is a free step: if it is smaller, it is a repetitive step. Finally, if $\mu - 1^{\text{th}}$ and μ^{th} members are equal, the corresponding step is a waiting step. It is true again that every free step must be repeated in the same or in the opposite direction. Furthermore, since the diagonal element in the 0 row vanishes, neither the first, nor the last step, can be a waiting step; the first one must be a free step, the last one a repetitive step.

We shall single out again that part of the sum (24) which belongs to types which have no 0 members, excepting the last. In the corresponding terms of (24), none of the i is 0. The sum of these terms will be denoted by $T'_\nu(q)N^{2\nu}$. The first and last factors in a typical sum of this type are v_{oi_1} and v_{i_1o}, their product v^2. The rest of the terms give all the terms of

$$\sum_{i_2i_3\cdots i_{2\nu-2}} H_{i_1i_2}H_{i_2i_3}\cdots H_{i_{2\nu-2}i_1} = (H^{2\nu-2})_{i_1i_1}\,. \tag{25}$$

However, the set of matrices (H_{i_1+m,i_1+n}) is identical with the set of matrices $(H_{mn} + i_11)$ so that the average of (25) is equal to the average of $[(H + i_11)^{2\nu-2}]_{00}$. This gives for the contribution of the zero-free types to (24)

$$\begin{aligned} T'_\nu(q)N^{2\nu} &= \sum_k v^2\mathrm{Av}[(H + k1)^{2\nu-2}]_{00} \\ &= \sum_k v^2\mathrm{Av}\sum_\lambda \binom{2\nu-2}{\lambda} k^{2\nu-2-\lambda}(H^\lambda)_{00}\,. \end{aligned} \tag{25a}$$

The summation over k can be replaced by integration from $-N$ to N. It gives 0 for odd λ, for even λ it gives $2N^{2\nu-1-\lambda}/(2\nu - 1 - \lambda)$. It simplifies the notation to write 2λ for λ and let λ run over the integers. At the same time, κ will be written for ν and the expression (23) substituted for the Av

$$\begin{aligned} T'_\kappa(q)N^{2\kappa} &= 2v^2\sum_{\lambda=0}^{\kappa-1}\binom{2\kappa-2}{2\lambda} T_\lambda(q)N^{2\lambda}\frac{N^{2\kappa-2\lambda-1}}{2\kappa-2\lambda-1} \\ &= 2qN^{2\kappa}\sum_{\lambda=0}^{\kappa-1}\binom{2\kappa-2}{2\lambda}\frac{T_\lambda(q)}{2\kappa-2\lambda-1} \qquad (\text{for } \kappa > 0). \end{aligned} \tag{25b}$$

This is the analogue of the very much simpler equation (14) of the preceding section. The highest T_λ which occurs on the right side is $T_{\kappa-1}$.

The various terms which enter the sum (24) can be distinguished according to the position 2κ at which the first 0 occurs in their type sequence. In the terms in which 0 occurs at the position 2κ, the index $i_{2\kappa} = 0$ so that they contribute altogether to the sum (24)

$$T'_\kappa(q)N^{2\kappa}\mathrm{Av}H_{0i_{2\kappa+1}}H_{i_{2\kappa+1}i_{2\kappa+2}}\cdots H_{i_{2\nu-1}0} = T'_\kappa(q)N^{2\kappa}\mathrm{Av}(H^{2\nu-2\kappa})_{00}\,. \tag{26}$$

The whole sum (24) will contain terms which correspond to any κ from 1 to ν; it will be a sum of expressions (26) with all these κ. Hence

$$\begin{aligned} T_\nu(q)N^{2\nu} &= \textstyle\sum_{\kappa=1}^{\nu} T'_\kappa(q)N^{2\kappa}\mathrm{Av}(H^{2\nu-2\kappa})_{00} \\ &= \textstyle\sum_{\kappa=1}^{\nu} T'_\kappa(q)T_{\nu-\kappa}(q)N^{2\nu} \qquad (\text{for } \nu > 0) \end{aligned} \tag{26a}$$

where (23) was used to express the Av. This last expression is the analogue of the second member of (14a). Substitution of (25b) into (26a) gives a recursive formula for the T_ν alone

$$\begin{aligned} T_\nu &= 2q \textstyle\sum_{\kappa=1}^{\nu} \sum_{\lambda=0}^{\kappa-1} \binom{2\kappa-2}{2\lambda} \dfrac{T_\lambda T_{\nu-\kappa}}{2\kappa-2\lambda-1} + \delta_{\nu o} \\ &= 2q \textstyle\sum_{\kappa=0}^{\nu-1} \sum_{\lambda=0}^{\kappa} \binom{2\kappa+1}{2\lambda} \dfrac{T_\lambda T_{\nu-\kappa-1}}{2\kappa+1} + \delta_{\nu o}. \end{aligned} \tag{26b}$$

The last term was added to make (26b) valid also for $\nu = 0$. The highest T that occurs on the right side has index $\nu - 1$ and the equations (26b) determine successively all T_ν. One obtains, either by direct calculation or from (26b)

$$T_0 = 1 \qquad T_1 = 2q \qquad T_2 = 8q^2 + 2q/3. \tag{26c}$$

The equations of the preceding section for the M_ν are contained in (26b) as a limiting case for $q = \infty$. Hence the coefficient of q^ν in T_ν is $2^\nu(2\nu)!/(\nu!(\nu+1)!)$. The coefficient of $q^{\nu-1}$ is

$$2^{3\nu-4}\nu/3 - 2^{\nu-2}(2\nu-1)!((\nu-1)!)^{-2}/3.$$

The coefficients of q and q^2 are $2/(2\nu-1)$ and $4(2\nu-2)/(2\nu-3)$, respectively. However, no closed expression could be obtained for T_ν.

The following estimate of T_ν, although very crude, will show by Carleman's theorem [1], p. 115 or [4] p. 19, that the moments (26b) uniquely determine the distribution function. The explicit expressions (26c) show that

$$T_\nu < 2(q+1)^\nu \nu^{2\nu} \tag{27}$$

is valid for $\nu = 0, 1$. It will be proved, in general, by induction. If (27) is valid for $\lambda = 0, 1, 2, \cdots, \nu - 1$, we have, for $\kappa \leqq \nu - 1$

$$\begin{aligned} \textstyle\sum_{\lambda=0}^{\kappa} \binom{2\kappa+1}{2\lambda} T_\lambda &< 2 \textstyle\sum_{\lambda=0}^{\kappa} \binom{2\kappa+1}{2\lambda} (q+1)^\lambda \lambda^{2\lambda} < 2 \textstyle\sum_0^\kappa \binom{2\kappa+1}{2\lambda} \\ &\cdot (q+1)^\kappa \kappa^{2\lambda} < (q+1)^\kappa[(1+\kappa)^{2\kappa+1} + (1-\kappa)^{2\kappa+1}]. \end{aligned} \tag{27a}$$

The last term on the right side can be omitted for $\kappa \geqq 1$. For $\kappa = 0$, omission of the last term makes (27a) an equality. Hence, we have, by (26b) and (27)

$$\begin{aligned} T_\nu &\leqq 2q \textstyle\sum_{\kappa=0}^{\nu-1} (q+1)^\kappa (1+\kappa)^{2\kappa+1} 2(q+1)^{\nu-\kappa-1}(\nu-\kappa-1)^{2(\nu-\kappa-1)}/(2\kappa+1) \\ &< 4(q+1)^\nu \left\{ \frac{1(\nu-1)^{2\nu-2}}{1} + \frac{2^3(\nu-2)^{2\nu-4}}{3} + \cdots + \frac{(\nu-1)^{2\nu-3}1^2}{2\nu-3} + \frac{\nu^{2\nu-1}}{2\nu-1} \right\}. \end{aligned}$$

There are ν terms in the bracket and the last one is largest among them. Hence

$$T_\nu < 4(q+1)\nu^{2\nu}/(2\nu-1) < 2(q+1)^\nu \nu^{2\nu} \tag{27b}$$

for $\nu = 2, 3, \cdots$. This proves (27).

It follows that

$$(T_\nu)^{-\frac{1}{2}\nu} \to (q+1)^{-\frac{1}{2}}\nu^{-1} \tag{27c}$$

so that the series with the general term (27c) diverges and the Perron-Carleman criterion for the uniqueness of the moment problem (28b) is satisfied.

Equations for the strength function

We shall write the basic equation for the strength function at once in terms of σ

$$\int_{-\infty}^{\infty} x^{2\nu}\sigma(x)\,dx = M_{2\nu} = T_\nu(q)N^{2\nu}. \tag{28}$$

By introducing a proper scale for the variable of the strength function

$$\rho(\xi) = N\sigma(N\xi) \tag{28a}$$

(28) transforms into

$$\int_{-\infty}^{\infty} \xi^{2\nu}\rho(\xi)\,d\xi = T_\nu(q). \tag{28b}$$

All odd moments of σ and ρ vanish. The existence of all moments implies that ρ goes to zero at $\xi \to \pm\infty$ faster than any power of ξ. Substitution of (28b) into (26b) gives for $\nu \neq 0$

$$\int d\xi\xi^{2\nu}\rho(\xi) = 2q\sum_{\kappa=0}^{\nu-1}\sum_{\lambda=0}^{\kappa}\binom{2\kappa+1}{2\lambda}\int d\xi\int d\zeta\rho(\xi)\rho(\zeta)\frac{\xi^{2\lambda}\zeta^{2\nu-2\kappa-2}}{2\kappa+1}. \tag{29}$$

All integrals in which no limit is given are to be extended from $-\infty$ to ∞. Together with $\int\rho(\xi)\,d\xi = 1$, (29) completely determines the even function ρ. The summation over λ can be carried out in (29) by the binomial theorem and gives

$$\int d\xi\xi^{2\nu}\rho(\xi) = q\sum_{\kappa=0}^{\nu-1}\int d\xi\int d\zeta\rho(\xi)\rho(\zeta)\zeta^{2\nu-2\kappa-2}\frac{(1+\xi)^{2\kappa+1}+(1-\xi)^{2\kappa+1}}{2\kappa+1}. \tag{29a}$$

Integration of the identity

$$\sum_{\kappa=0}^{\nu-1} x^{2\kappa} = \frac{x^{2\nu}-1}{x^2-1}$$

gives

$$\sum_{\kappa=0}^{\nu-1}\frac{u^{2\kappa+1}}{2\kappa+1} = \frac{1}{2}\int_{-u}^{u} dx\,\frac{x^{2\nu}-1}{x^2-1}. \tag{29b}$$

Hence (29a) can be written also as

$$\text{(29c)} \quad \int d\xi \xi^{2\nu} \rho(\xi) = \tfrac{1}{2} q \int d\xi \int d\zeta \rho(\xi) \rho(\zeta) \zeta^{2\nu-1} \left(\int_{-\alpha}^{\alpha} dx + \int_{-\beta}^{\beta} dx \right) \frac{x^{2\nu} - 1}{x^2 - 1}$$

where $\alpha = (1 + \xi)/\zeta$ and $\beta = (1 - \xi)/\zeta$. Introducing $z = \zeta x$ as new variable instead of x gives

$$\text{(29d)} \quad \int d\zeta \zeta^{2\nu} \rho(\zeta) = \tfrac{1}{2} q \int d\xi \int d\zeta \rho(\xi) \rho(\zeta) \left(\int_{1-\xi}^{1+\xi} dz + \int_{-1+\xi}^{1-\xi} dz \right) \frac{z^{2\nu} - \zeta^{2\nu}}{z^2 - \zeta^2}.$$

The purpose of the following transformations is to bring the right side into a form in which ν appears only as the exponent of ζ^2. This cannot be done by simply interchanging the variables z and ζ in the term which contains $z^{2\nu}$ because the singularity which would then result at $z = \zeta$. In addition, it appears worthwhile to simplify the last equation somewhat. One can replace the two integrals with respect to z by two similar integrals extending from $\xi - 1$ to $\xi + 1$ and from $-\xi - 1$ to $-\xi + 1$. Since the remainder of the integrand is an even function of ξ, the two integrals are equal and can be replaced by twice one of them. Finally, one can interchange the integrations with respect to z and ξ to obtain, still only for $\nu \neq 0$,

$$\text{(30)} \quad \begin{aligned} \int d\zeta \zeta^{2\nu} \rho(\zeta) &= q \int dz \int d\zeta \int_{z-1}^{z+1} d\xi \rho(\xi) \rho(\zeta) \frac{z^{2\nu} - \zeta^{2\nu}}{z^2 - \zeta^2} \\ &= q \int dz \int d\zeta R_1(z) \rho(\zeta) \frac{z^{2\nu} - \zeta^{2\nu}}{z^2 - \zeta^2} \end{aligned}$$

in which

$$\text{(30a)} \quad R_1(z) = \int_{z-1}^{z+1} d\xi \rho(\xi)$$

is again an even function which drops to zero at $z = \pm\infty$ faster than any power of z. Since the integrand on the right side of (30) has no singularity, it is permissible to exclude two narrow strips $|z \pm \zeta| < \varepsilon$ from the integration. Then, the integration variables z and ζ can be interchanged in the term with $z^{2\nu}$ to give, as $\varepsilon \to 0$

$$\text{(30b)} \quad \int d\zeta \zeta^{2\nu} \rho(\zeta) = \iint_{|z \pm \zeta| > \varepsilon} q \, dz \, d\zeta [R_1(\zeta) \rho(z) + \rho(\zeta) R_1(z)] \frac{\zeta^{2\nu}}{\zeta^2 - z^2}.$$

Even though both integrands are even functions of ζ, the application of the theorem of moments to equate the integrands is not permissible because (30b) is not valid for $\nu = 0$.

In order to render (30b) valid also for $\nu = 0$, we calculate first

$$\text{(31)} \quad \begin{aligned} \int_{|z \pm \zeta| > \varepsilon} \frac{dz}{z^2 - \zeta^2} &= \frac{1}{\zeta} \ln \frac{2\zeta + \varepsilon}{\varepsilon} && \text{for } 0 < \zeta < \varepsilon \\ &= \frac{1}{\zeta} \ln \frac{2\zeta + \rho}{2\zeta - \varepsilon} && \text{for } \varepsilon < \zeta. \end{aligned}$$

These expressions are proportional to $1/\varepsilon$ as long as ζ is of the order ε. They drop very fast as ζ increases: the integral is 1 for $\zeta \sim \varepsilon^{\frac{1}{2}}$ and of the order of $\varepsilon^{\frac{1}{2}}$ for $\zeta \sim \varepsilon^{\frac{1}{2}}$. Hence, if $f(\zeta)$ is an even function with two continuous derivatives which is small enough at $\zeta = \pm\infty$,

$$\lim_{\varepsilon=0} \iint_{|z\pm\zeta|>\varepsilon} dz\, d\zeta\, \frac{f(\zeta)}{z^2 - \zeta^2} = cf(0); \qquad c = \pi^2/2 \tag{31a}$$

c is the integral of (31) from $\zeta = -\infty$ to ∞ or twice the integral from 0 to ∞. The reader will recognize that the double integral of $(\zeta^2 - z^2)^{-1}$ over the whole ζz plane excluding the $|\zeta \pm z| < \varepsilon$ strips is not absolutely convergent: it gives a negative value if first integrated over z then ζ, the opposite value if the integration is carried out in the reverse order. It is clear from (31a) that the subtraction of a term $f(\zeta)$ in the square bracket of (30b) will not render this equation invalid for $\nu \geqq 1$ as long as the even function $f(\zeta)$ drops faster at $\pm\infty$ than any power of ζ because $\zeta^{2\nu}f(\zeta)$ vanishes at $\zeta = 0$. On the other hand, the value of the integral for $\nu = 0$, which is zero as the integral now stands, will become 1 if

$$f(0) = 2/\pi^2 q. \tag{32}$$

Hence

$$\int d\zeta\zeta^{2\nu}\rho(\zeta) = \lim_{\varepsilon=0} \int d\zeta\zeta^{2\nu} \int_{|z-\zeta|<\varepsilon} q\, dz\, \frac{R_1(\zeta)\rho(z) + \rho(\zeta)R_1(z) - f(\zeta)}{\zeta^2 - z^2} \tag{32a}$$

is valid now for all ν. It can be written in a somewhat more transparent form

$$\begin{aligned} &\int d\zeta\zeta^{2\nu}\rho(\zeta) - \int d\zeta\zeta^{2\nu}q\left[R_1(\zeta)\int dz\,\frac{\rho(z) - \rho(\zeta)}{\zeta^2 - z^2} + \rho(\zeta)\int dz\,\frac{R_1(z) - R_1(\zeta)}{\zeta^2 - z^2}\right] \\ &\qquad = \lim_{\varepsilon=0} \int d\zeta\zeta^{2\nu} \int_{|z\pm\zeta|>\varepsilon} q\, dz\, \frac{2R_1(\zeta)\rho(\zeta) - f(\zeta)}{\zeta^2 - z^2}. \end{aligned} \tag{32b}$$

In the square bracket, the strips $|z \pm \zeta| < \varepsilon$ need not be excluded from the integration since the integrand is regular everywhere. Because of (31a), the right side can be replaced by

$$\delta_{\nu 0}cq[f(0) - 2R_1(0)\rho(0)] = \tfrac{1}{2}\pi^2 q\delta_{\nu 0}[2/\pi^2 q - 2R_1(0)\rho(0)]. \tag{32c}$$

Hence, the left side of (32b) contains the moments of an even function of and the form (32c) of the right side shows that these all vanish except the zeroth moment. It follows that the function vanishes and so does its zeroth moment

$$\rho(\zeta) = qR_1(\zeta)\int \frac{\rho(z) - \rho(\zeta)}{\zeta^2 - z^2}\, dz + q\rho(\zeta)\int \frac{R_1(z) - R_1(\zeta)}{\zeta^2 - z^2}\, dz \tag{33}$$

$$R_1(0)\rho(0) = 1/\pi^2 q. \tag{33a}$$

The last equation means, in terms of the strength function σ (cf. (28a))

$$N\sigma(0) \int_{-N}^{N} \sigma(x)\, dx = 1/\pi^2 q. \tag{33b}$$

Discussion of the equations (33) for the strength function

As was mentioned before, the problem of the preceding section represents the limiting case $q = \infty$ of the present problem. Hence, (20) is a solution of (33), (33a) for very large q. In the opposite limiting case of small q, the total width of ρ will be much smaller than 1 and R_1 can be replaced by 1 over the significant region. Hence, in this case

$$\rho(\zeta) \sim q \int \frac{\rho(z) - \rho(\zeta)}{\zeta^2 - z^2}\, dz. \tag{34}$$

This equation is solved by

$$\rho(\zeta) \sim \frac{\text{const}}{q^2\pi^2 + \zeta^2} = \frac{q}{\pi^2 q^2 + \zeta^2} \tag{34a}$$

the constant being determined by (33a). However (34a) is valid only as long as $\zeta \ll 1$ since only in this region is $R_1 \sim 1$.

Neither (20) nor (34a) give the asymptotic behavior of $\rho(\zeta)$ correctly. In particular already the second moment of (34a) is divergent. The asymptotic behavior can be calculated on the basis of the observation that ρ and R_1 surely decrease to 0 at $\zeta \to \infty$ faster than any power of ζ and that, hence, the largest contribution to the integrals in (33) comes from the region at very small z. Hence, the $\zeta^2 - z^2$ in the denominators can be replaced by ζ^2. Since the integrals of ρ and R_1 are 1 and 2, respectively, we have asymptotically

$$\rho(\zeta) \approx qR_1(\zeta)/\zeta^2 + 2q\rho(\zeta)/\zeta^2. \tag{35}$$

The last term on the right side is much smaller than the left side and can be omitted. The resulting equation has the asymptotic solution

$$\rho(\zeta) \approx \text{const.}\ (2q\zeta^2 \ln \zeta/e^2)^{-\zeta}. \tag{35a}$$

The calculations reported in the present paper were carried out during the summer of 1954. The writer is much indebted to the Wisconsin Alumni Research Foundation for supporting this work.

PRINCETON UNIVERSITY

REFERENCES

[1] T. CARLEMAN, Les Fonctions Quasi-analytic, Gauthier-Villars, Paris, 1926.
[2] W. FELLER, An Introduction to Probability Theory and its Applications, John Wiley, 1950.
[3] A. LANE, R. G. THOMAS and E. P. WIGNER, *Giant Resonance Interpretation of the Nucleon-Nucleus Interaction*, Physical Review, 98 (1955), pp. 693–701.
[4] J. A. SHOHAT and J. D. TAMARKIN, The Problem of Moments, Amer. Math. Soc., New York, 1943.
[5] E. T. WHITTAKER and G. N. WATSON, A Course of Modern Analysis, Cambridge University Press, 1927.
[6] S. WILKS, Mathematical Statistics, Princeton University Press, 1943.

Characteristic Vectors of Bordered Matrices with Infinite Dimensions II

E. P. Wigner

Annals of Mathematics *65*, 203–207 (1957)

(Received June 19, 1956)

The matrices which form the subject of an earlier study [1] consist of a diagonal matrix **k**

$$k_{mn} = m\delta_{mn} \qquad (-\infty < m, n < \infty) \tag{1}$$

and of the "border". This is a real symmetric matrix **v**, all the elements of which vanish beyond a certain distance from the diagonal

$$v_{mn} = 0 \quad \text{for} \quad |m - n| > N \qquad (-\infty < m, n < \infty) \tag{1a}$$

while those closer to the diagonal have all the same absolute value

$$|v_{mn}| = v \quad \text{for} \quad |m - n| \leqq N. \tag{1b}$$

Since **v** is symmetric $v_{mn} = v_{nm}$. Subject to this condition, however, the signs of the v_{mn} are random; i.e., ensembles of matrices are considered with all possible signs of the v_{mn} subject to the conditions of symmetry.

The problem is connected with the real normed characteristic vectors $\psi^{(\lambda)}$ of $H = \mathbf{k} + \mathbf{v}$:

$$\sum_n H_{mn}\psi_n^{(\lambda)} = \lambda\psi_m^{(\lambda)}. \tag{2}$$

The average value of $(\psi_0^{(\lambda)})^2$ over all matrices of the ensemble was denoted by $\sigma(\lambda)$, the characteristic value λ being considered as a continuous variable. A more nearly exact definition of $\sigma(\lambda)$ is given in the aforementioned article. A complete solution could be given for $N = 1$; the more interesting case of a very large N led, in general, to an integral equation. In order to formulate this integral equation, a more proper scale for the variable λ was introduced

$$\rho(\lambda/N) = N\sigma(\lambda). \tag{3}$$

It was further assumed that, together with N, also v tends to infinity in such a way, however, that

$$v^2/N = q \tag{3a}$$

remains finite. Then, the following integral equation (33) was obtained for ρ, partly by heuristic arguments

$$\rho(\zeta) = qR_1(\zeta)\int \frac{\rho(z) - \rho(\zeta)}{\zeta^2 - z^2}\,dz + q\rho(\zeta)\int \frac{R_1(z) - R_1(\zeta)}{\zeta^2 - z^2}\,dz. \tag{4}$$

All integrations on which no limits are given are to be extended from $-\infty$ to ∞;

$$R_1(\zeta) = \int_{\zeta-1}^{\zeta+1} \rho(z)dz. \tag{4a}$$

It is easily seen that $\rho(\zeta)$ is an even function, it tends to zero for large ζ so rapidly that all its moments exist. Its zeroth moment (integral) is 1, its second moment $2q$.

The integral equation (4) could not be solved in closed form and no such solution will be presented here. However, it will be transformed into a much simpler equation from which the approximate expressions for ρ (for large and small values of q) which were obtained before can be obtained much more easily. For this purpose we consider the function $r(z)$ which is analytic in the upper half plane, tends to zero in that plane as $|z| \to \infty$, and the real part of which is $\rho(z)$ for real z. This function is well known to be given in the upper half plane by the integral

$$r(z) = \frac{1}{i\pi} \int \frac{\rho(\zeta)}{\zeta - z} d\zeta \qquad \text{(for Im } z > 0). \tag{5}$$

It is easy to verify that the real part of $r(z)$ tends to $\rho(\zeta)$ as z approaches the real ζ from above

$$\operatorname{Re} r(z) = \rho(z) \qquad \text{(for real } z). \tag{5a}$$

The imaginary part of $r(z)$ on the real axis is obtained in the same way as the principal value integral

$$\operatorname{Im} r(z) = \frac{1}{\pi} P \int \frac{\rho(\zeta)}{z - \zeta} d\zeta \qquad \text{(for real } z). \tag{5b}$$

Naturally, the integral representation (5) is not valid in the lower half plane; in fact, the real part of the integral in (5) approaches $-\rho(\zeta)$ as z approaches the real ζ from below. In the lower half plane $r(z)$ must be obtained by analytic extension from the upper half plane, not by (5). As a matter of fact, it will not be necessary to make such an extension, the complex variable z can be restricted for the purposes of the present article to the upper half plane including the real axis.

Let us consider now the analytic function which is in the same relation to R_1 as r is to ρ. The function

$$r_1(z) = \int_{-1}^{1} r(z + \zeta) d\zeta \tag{5c}$$

satisfies all the conditions which this function must satisfy and is therefore this function. It is also even in z.

At very large $|z|$, because of the rapid drop of $\rho(\zeta)$ at $\zeta \to \pm\infty$, and since $\int \rho(\zeta)\, d\zeta = 1$, one obtains from (5)

$$r(z) \to i/\pi z \qquad \text{(for Im } z > 0). \tag{6}$$

Let us now calculate the imaginary parts of $r(z)$ and $r_1(z)$ on the real axis.

Since ρ is an even function

$$\text{(7)}\qquad \begin{aligned} \operatorname{Im} r(z) &= \frac{1}{\pi} P \int \frac{\rho(\zeta)}{z-\zeta}\, d\zeta = \frac{1}{2\pi} P \int \left(\frac{\rho(\zeta)}{z-\zeta} + \frac{\rho(-\zeta)}{z+\zeta} \right) d\zeta \\ &= \frac{1}{2\pi} P \int \rho(\zeta) \frac{2z}{z^2-\zeta^2}\, d\zeta. \end{aligned}$$

An easy calculation gives, furthermore,

$$\begin{aligned} P \int_{-\infty}^{\infty} \frac{2z\, d\zeta}{z^2-\zeta^2} &= 2P \int_0^{\infty} \left(\frac{1}{\zeta+z} - \frac{1}{\zeta-z} \right) d\zeta \\ &= 2 \lim_{N\to\infty} \lim_{\epsilon\to 0} \left(\int_0^{z-\epsilon} d\zeta + \int_{z+\epsilon}^{N} d\zeta \right) \left(\frac{1}{\zeta+z} - \frac{1}{\zeta-z} \right) \\ &= 2 \lim_{N\to\infty} \lim_{\epsilon\to 0} \left\{ \left[\ln \frac{\zeta+z}{z-\zeta} \right]_0^{z-\epsilon} + \left[\ln \frac{\zeta+z}{\zeta-z} \right]_{z+\epsilon}^{N} \right\} \\ &= 2 \lim_{N\to\infty} \lim_{\epsilon\to 0} \left\{ \ln \frac{2z-\epsilon}{\epsilon} + \ln \frac{N+z}{N-z} - \ln \frac{2z+\epsilon}{\epsilon} \right\} = 0. \end{aligned}$$

It now follows by means of (7) that for real z

$$\text{(7a)}\qquad \operatorname{Im} r(z) = \frac{z}{\pi} P \int \frac{\rho(\zeta)}{z^2-\zeta^2}\, d\zeta = \frac{z}{\pi} P \int \frac{\rho(\zeta)-\rho(z)}{z^2-\zeta^2}\, d\zeta$$

since the last equation shows that the term added on the right side vanishes. The principal value integral can now be replaced by the proper integral since the integrand has no singularity any more.

As (4a) shows, the even character of ρ entails the even character of R_1. Hence, the same calculation gives

$$\text{(7b)}\qquad \operatorname{Im} r_1(z) = \frac{z}{\pi} \int \frac{R_1(\zeta)-R_1(z)}{z^2-\zeta^2}\, d\zeta$$

and (4), multiplied by z/π, can be given the form

$$\text{(8)}\qquad (\operatorname{Re} zr/\pi) = q(\operatorname{Re} r_1)(\operatorname{Im} r) + q(\operatorname{Re} r)(\operatorname{Im} r_1) \qquad \text{(for real } z\text{)}.$$

It follows that the real part of the analytic function $w = zr/\pi + iqrr_1$ vanishes on the real axis. Since r and r_1 are both regular in the upper half plane, including the real axis, this holds also of w. Finally, since both r and r_1 tend to zero as $|z| \to \infty$, and because of (6), $w \to i/\pi^2$ for very large z. It follows that $w - i/\pi^2$ is regular in the upper half plane, tends to zero for $|z| \to \infty$ and its real part vanishes on the real axis. Hence $w - i/\pi^2 = 0$

$$\text{(9)}\qquad zr/\pi + iqrr_1 = i/\pi^2.$$

This is the simplified form of (4) to which we were referring. It can be written in an even slightly simpler form if one introduces $p(z) = \pi r(z)$ and $p_1(z) =$

$\pi r_1(z)$

$$zp + iqpp_1 = i \tag{9a}$$

while the relation between p and p_1 remains

$$p_1(z) = \int_{z-1}^{z+1} p(\zeta)\, d\zeta. \tag{9b}$$

The consequences of (4), obtained before, will now be derived from (9) or the equivalent (9a). It will be noted, first, that $\operatorname{Im} r(z)$ is an odd function of the real variable z. This follows most easily from (7a). Similarly, (7b) shows that $\operatorname{Im} r_1$ is also odd. Hence, both r and r_1 are real functions of iz

$$\overline{r(iz)} = r(\overline{iz}); \qquad \overline{r_1(iz)} = r_1(\overline{iz}). \tag{10}$$

In particular, r and r_1 are real for $z = 0$ and equal to $\rho(0)$ and $R_1(0)$, respectively. Hence, insertion of $z = 0$ into (9) gives

$$q\rho(0)R_1(0) = 1/\pi^2. \tag{11}$$

This is equation (33a) of Reference 1. It follows from the development given there that (11) should be a consequence of (4). The preceding argument verifies this directly.

Since $2q$ is the second moment of ρ, this will change very little in a unit interval if q is large. Hence, by (5c), $r_1 \approx 2r$ for large q. Thus (9) becomes in this case

$$2iqr^2 + (z/\pi)r - i/\pi^2 = 0$$

or

$$r(z) = +\frac{iz}{4\pi q} + \left[\frac{1}{2\pi^2 q} - \frac{z^2}{16\pi^2 q^2}\right]^{\frac{1}{2}}.$$

In this case for real z

$$\begin{aligned} \rho(z) = \operatorname{Re} r(z) &= (4\pi q)^{-1}(8q - z^2)^{\frac{1}{2}} && \text{for } z^2 < 8q \\ &= 0 && \text{for } z^2 > 8q. \end{aligned}$$

This is equation (20) of Reference 1.

If q is very small, the width of ρ will be very small and the integral of (5c) will be closely approximated by the similar integral from $-\infty$ to ∞ as long as $z < 1$. For real z, the value of this integral is 1 because this is the contribution of the real part of r and because the imaginary part of r, being odd, does not contribute at all. Hence (9) becomes in this case

$$((z/\pi) + iq)r = i/\pi^2.$$

The real part of r then is for real z

$$\rho(z) = \operatorname{Re} r(z) = q/(z^2 + \pi^2 q^2) \qquad (\text{for } -1 < z < 1)$$

which is (34a) of Reference 1.

PRINCETON UNIVERSITY

REFERENCE

1. E. P. WIGNER, *Characteristic Vectors of Bordered Matrices with Infinite Dimensions*, Ann. of Math., 62 (1955), pp. 548–564. The following errors may be disturbing: The coefficient of q^2 in T_ν given around the middle of page 560 is incorrect (the expression given is neither derived nor used subsequently). In (27c), replace $(T_\nu)^{-\frac{1}{2}\nu}$ by $(T_\nu)^{-1/2\nu}$. In (31), replace ρ by ϵ.

On a Theorem of Magnus

E. Hewitt and E. P. Wigner[1]

Proceedings of the American Mathematical Society *8*, 740–744 (1957)

1. In a recent paper [2],[2] W. Magnus has shown that analogues of the Fourier inversion and Plancherel theorems hold for matrix-valued functions on the real line R. We propose to show that these theorems actually hold for an arbitrary locally compact Abelian group, and that Magnus's inversion integral (l. c. (1.4)) can be simplified. For all group- and integral-theoretic notation, terms, and facts used here without explanation, see [1].

2. Let G be a locally compact Abelian group, written additively, with character group X. Elements of G will be denoted "s", "t", and elements of X by "χ", with or without subscripts. The differential of Haar measure on G $[X]$ will be denoted $dt[d\chi]$ and these measures are to be so chosen that equality obtains in the Fourier inversion theorem [1, p. 143] and Plancherel's theorem [1, p. 145].

2.1. Let U be a continuous n-dimensional unitary representation of G, so that: $U(s+t)=U(s)U(t)$ for all s, $t\in G$; $U(0)=I$; and all coefficients u_{jk} of U are continuous functions on G. Then the reduction theorem states that there exist a unitary matrix V and characters $\chi_1, \cdots, \chi_n \in X$ such that

Received by the editors February 16, 1956 and, in revised form, October 1, 1956.

[1] The first-named author is a fellow of the John Simon Guggenheim Memorial Foundation.

[2] Numbers in brackets refer to the bibliography.

2.1.1 $$U(t) = V^{-1}\begin{pmatrix} \chi_1(t) & 0 & \cdots & 0 \\ 0 & \chi_2(t) & \cdots & 0 \\ \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & \cdots & \chi_n(t) \end{pmatrix} V$$

for all $t \in G$.

The set $\{\chi_1, \cdots, \chi_n\}$ of characters is completely determined by U, although obviously their order is not.

2.2. We now select a symmetric compact neighborhood A of the identity in X having (finite) positive measure. Let $E = E(\chi_1, \cdots, \chi_n)$ be the function on X^n that is equal to 1 if $\chi_j\chi_k^{-1} \in A$ for all j, k and is equal to 0 otherwise. It is obvious that

2.2.1 $$E(\chi_1, \cdots, \chi_n) = E(\chi_1\chi, \cdots, \chi_n\chi)$$

for all $\chi \in X$.

3. Let $f \in L_1(G)$. We say that Fourier inversion holds for f if $\hat{f} \in L_1(X)$ and $\int_X \hat{f}(\chi)\overline{\chi(s)}d\chi = f(s)$ for all $s \in G$. (Recall that $\hat{f}(\chi) = \int_G f(t)\chi(t)dt$.) It is known [1, p. 143] that Fourier inversion holds if f is in $L_1(G)$ and is also a linear combination of continuous positive definite functions. We can now establish our generalization of Magnus's theorem.

3.1. THEOREM. *Let $F = F(t)$ be an $n \times n$ complex matrix function on G such that $F_{jk} \in L_1(G)$ and Fourier inversion holds for F_{jk}, for all coefficients F_{jk} of F. For every representation as in* 2.1, *let*

3.1.1 $$F^{\frown}(U) = \int_G F(t)U(t)dt.$$

Then for all $s \in G$, the equality

3.1.2 $$\kappa \cdot F(s) = \int_{X^n} F^{\frown}(U)U(-s)E(\chi_1, \cdots, \chi_n)d\chi_1 \cdots d\chi_n$$

holds, where κ depends only on n and A and $0 < \kappa < \infty$. The integral in 3.1.2 *is an n-fold group integral over X^n, in which the integrand vanishes unless $\chi_j\chi_k^{-1} \in A$ for all j and k. The matrix $F^{\frown}(U)U(-s)$ depends symmetrically on the characters $\chi_1, \cdots, \chi_n$ associated with U by* 2.1.1, *and hence the integrand in* 3.1.2 *is well defined.*

PROOF. Let j and k be arbitrary integers, $1 \leqq j \leqq n$ and $1 \leqq k \leqq n$. Using 3.1.1 and 2.1.1, we can write the (j, k)th entry of the integral in 3.1.2 as

$$3.1.3 \qquad \sum_{\alpha,\beta=1}^{n} \int_{X^n} \int_G F_{j\alpha}(t)\chi_\beta(t-s)\bar{v}_{\beta\alpha}v_{\beta k}E(\chi_1, \cdots, \chi_n)dtd\chi_1 \cdots d\chi_n.$$

(Here and also below we make free use of Fubini's theorem.) For each α and β, the integral in 3.1.3 can be written as

$$3.1.4 \qquad \bar{v}_{\beta\alpha}v_{\beta k}\int_{X^n} \hat{F}_{j\alpha}(\chi_\beta)\chi_\beta(-s)E(\chi_1, \cdots, \chi_n)d\chi_1 \cdots d\chi_n.$$

To evaluate 3.1.4, suppose for simplicity that $\beta=1$. The integral in 3.1.4, taken over $\chi_2, \cdots, \chi_n$, will be independent of the value of χ_1. To see this, we substitute 2.2.1 in 3.1.4 and obtain

$$\begin{aligned} &\int_{X^n} E(\chi_1, \chi_2, \cdots, \chi_n)d\chi_2 \cdots d\chi_n \\ 3.1.5 \qquad &= \int_{X^n} E(\chi_1\chi, \chi_2\chi, \cdots, \chi_n\chi)d\chi_2 \cdots d\chi_n \\ &= \int_{X^n} E(\chi_1\chi, \chi_2, \cdots, \chi_n)d\chi_2 \cdots d\chi_n. \end{aligned}$$

The last step follows because the Haar integral over X is invariant. Since χ is arbitrary in X, our assertion about 3.1.4 follows. By hypothesis, Fourier inversion holds for $F_{j\alpha}$. Therefore, denoting the value of 3.1.5 by κ (it is easy to see that $0<\kappa<\infty$), we can write 3.1.4 as

$$3.1.6 \qquad \bar{v}_{\beta\alpha}v_{\beta k}\kappa F_{j\alpha}(s).$$

Hence 3.1.3 is equal to $\sum_{\alpha,\beta=1}^{n} \bar{v}_{\beta\alpha}v_{\beta k}\kappa F_{j\alpha}(s) = \sum_{\alpha=1}^{n} \delta_{\alpha k}\kappa F_{j\alpha}(s) = \kappa F_{jk}(s)$. This proves the theorem.

3.2. Remark. The foregoing proof shows that the matrices V and V^{-1} in 2.1.1 play a quite inessential rôle for Fourier inversion. They do not affect the value of the integral 3.1.2, although $F^{\frown}(U)U(-s)$ does depend upon the choice of V. This fact makes it unnecessary to integrate over the unitary group in order to invert the mapping $F \to F^{\frown}$. Note that Magnus's inversion formula involves integrating over the unitary group.

4. The Plancherel theorem is very easy to establish in the present context. For a matrix M, let M^* be the conjugate transpose of M and let Tr M be the trace of M.

4.1. Theorem. *Let $F=F(t)$ be an $n\times n$ complex matrix-valued function on G such that $F_{jk}\in L_1(G)\cap L_2(G)$ for all coefficients F_{jk} of F. Let $F^{\frown}$ (U) be defined as in* 3.1.1. *Then*

$$\kappa \operatorname{Tr} \int_G F(t)F^*(t)dt$$

4.1.1

$$= \operatorname{Tr} \int_{X^n} F^\frown(U)F^{\frown *}(U)E(\chi_1, \cdots, \chi_n)d\chi_1 \cdots d\chi_n.$$

PROOF. Let $D(\chi_1, \cdots, \chi_n)$ denote the diagonal matrix with entries $\chi_1, \cdots, \chi_n$. For a fixed representation U, write $U = V^{-1}D(\chi_1, \cdots, \chi_n)V$. Let $P(t) = VF(t)V^{-1}$, and let $Q(\chi_1, \cdots, \chi_n)$ be the matrix $\{\hat{P}_{jk}(\chi_k)\}_{j,k=1}^n$. Then we have

$$F^\frown(U) = \int_G F(t)U(t)dt = V^{-1} \cdot \int_G P(t) \cdot D(\chi_1(t), \cdots, \chi_n(t))dt \cdot V$$

4.1.2

$$= V^{-1}Q(\chi_1, \cdots, \chi_n)V.$$

It is plain from 4.1.2 that

$$\operatorname{Tr}(QQ^*) = \operatorname{Tr}(F^\frown F^{\frown *}).$$

Hence we find

$$\operatorname{Tr} = \int_{X^n} F^\frown F^{\frown *}(U)E(\chi_1, \cdots, \chi_n)d\chi_1 \cdots d\chi_n$$

4.1.3

$$= \operatorname{Tr} \int_{X^n} Q(\chi_1, \cdots \chi_n)Q^*(\chi_1, \cdots \chi_n)E(\chi_1, \cdots, \chi_n)d\chi_1 \cdots d\chi_n$$

$$= \sum_{j,k=1}^n \int_{X^n} |\hat{P}_{jk}(\chi_k)|^2 E(\chi_1, \cdots, \chi_n)d\chi_1 \cdots d\chi_n.$$

By the usual Plancherel theorem and the argument used in 3.1, we can rewrite the last term in 4.1.3 as

4.1.4

$$\kappa \cdot \sum_{j,k=1}^n \int_G |P_{jk}(t)|^2 dt.$$

Now using again elementary facts about the trace, we can rewrite 4.1.4 as

$$\kappa \int_G \operatorname{Tr}[P(t)P^*(t)]dt = \kappa \int_G \operatorname{Tr}[VF(t)F^*(t)V^{-1}]dt$$

4.1.5

$$= \kappa \int_G \operatorname{Tr}[F(t)F^*(t)]dt = \kappa \operatorname{Tr} \int_G F(t)F^*(t)dt.$$

Combining 4.1.2–4.1.5, we have the present theorem.

4.2. REMARK. If some coefficient F_{jk} of F is in $L_2(G)$ but not in

$L_1(G)$, then the integral 3.1.1 does not exist as an absolutely convergent integral. Nevertheless, a theorem analogous to 4.1 follows with equal ease.

5. We note finally the obvious fact that an "isomorphic" theory can be constructed using the integral

$$\hat{}F(U) = \int_G U(t)F(t)dt$$

and the inversion integral

$$\int_{X^n} U(-s)\hat{}F(U)E(\chi_1, \cdots, \chi_n)d\chi_1 \cdots d\chi_n.$$

Bibliography

1. Lynn H. Loomis, *An introduction to abstract harmonic analysis*, New York, Van Nostrand, 1953.

2. Wilhelm Magnus, *A Fourier theorem for matrices*, Proc. Amer. Math. Soc. vol. 6 (1955) pp. 880–890. (Also appeared as Division of Electromagnetic Research, Institute of Mathematical Sciences, New York University, Research Report No. BR-8 (1954)).

Institute for Advanced Study,
University of Washington and
Princeton University

Normal Form of Antiunitary Operators

E. P. Wigner

Journal of Mathematical Physics *1*, 409–413 (1960)

(Received April 25, 1960)

Antiunitary operators are characterized in a manner similar to the characterization of unitary operators by their characteristic vectors and characteristic values. It is shown that a complete orthonormal set of vectors can be defined, some of which are invariant under the antiunitary operator. The rest of the vectors, which are always even in number, form pairs in such a way that the antiunitary operator transforms each member of a pair into a multiple of the other member of the same pair [Eq. (11)]. The extent to which the vectors of the orthonormal set are determined by the antiunitary operator is ascertained and the number of free parameters in the various cases of degeneracy found.

1.

ANTIUNITARY operators[1] play a significant role in the theory of the invariance of quantum mechanical equations. The symmetry operators which involve the operation of time-inversion are antiunitary. The antiunitary operators are antilinear, i.e., if φ and ψ are two vectors of the complex Hilbert space in which the antiunitary operator A is defined and if a and b are two complex numbers,

$$A(a\varphi+b\psi)=a^*A\varphi+b^*A\psi. \tag{1}$$

The asterisk denotes the conjugate complex. Furthermore, A changes the scalar product into its conjugate complex

$$(A\varphi,A\psi)=(\varphi,\psi)^*=(\psi,\varphi). \tag{2}$$

Actually, (1) follows from (2) so that the latter equation can serve as the definition of the antiunitary nature of A. However, unless the Hilbert space has only a finite number of dimensions, it is also necessary to specify that A has an inverse. This is also antiunitary.

If A is antiunitary, A^2 defined by

$$A^2\psi=A(A\psi) \tag{3}$$

is unitary. This follows directly from the defining Eqs. (1) and (2), and it is also clear that A^2 has an inverse if A does.

If Av_k is given for all the members of a complete orthonormal set of vectors $v_1,v_2,\cdots$, its antilinear property defines it for all vectors $v=\Sigma a_kv_k$:

$$Av=A(\sum a_kv_k)=\sum a_k^*Av_k. \tag{4}$$

Hence, the normal form of A will be obtained by specifying a complete set of orthonormal vectors v_k for which Av_k has a particularly simple form. These vectors are the analogs of the characteristic vectors for unitary operators and will be, indeed, characteristic vectors of A^2. However, this property does not define them completely.

If $v_1,v_2,\cdots$ form a complete orthonormal set, $Av_1,Av_2,\cdots$ also form such a set. The orthonormal nature of the latter set follows directly from (2), the completeness from the existence of the inverse of A. If w were orthogonal to all Av_k, then $A^{-1}w$ would be orthogonal to all v_k.

We mention further for the sake of completeness, that if K is the operation of complex conjugation so that, in a particular coordinate system,

$$K\psi=\psi^*;\quad K^2=1, \tag{5}$$

AK is unitary and it follows that every antiunitary operator can be written in the form

$$A=UK, \tag{6}$$

where U is unitary. It follows from (6) that

$$A^2=UKUK=UU^*K^2=UU^*, \tag{7}$$

where U^* is the conjugate complex of U in the coordinate system in which (5) is valid. Since UU^* is equivalent to its conjugate complex

$$UU^*=UU^*UU^{-1}=U(UU^*)^*U^{-1}, \tag{8}$$

its characteristic values are either real or pairwise conjugate complex. It follows that the square of an antiunitary operator is equivalent to a rotation. The last four equations will not be used explicitly.

2.

It will be assumed that the spectrum of $A^2=\Lambda$ is discrete. The complications which arise if Λ has a continuous spectrum are not serious, but their elimination is cumbersome. Let us consider then a characteristic vector of Λ:

$$\Lambda v=A^2v=\omega v. \tag{9}$$

Since Λ is unitary, $|\omega|=1$. It then follows that Av is also a characteristic vector of Λ,

$$\Lambda Av=A^2Av=AA^2v=A\omega v=\omega^*Av, \tag{10}$$

[1] Some of the results of the present article can be obtained on the basis of theorems derived by E. Cartan in his *Leçons sur la Géométrie Projective Complexe* (Gauthier-Villars, Paris, 1931). I am much indebted to Professor S. Bochner for drawing my attention to the very profound investigations contained in this treatise, which deals with general linear and antilinear transformations. However, the direct derivations, given in the text of the present paper, are hardly longer than the reinterpretation and amplification of Cartan's results (see particularly pp. 124–137) would have been.

and belongs to the characteristic value ω^*. Unless $\omega=1$ or $\omega=-1$, $\omega\neq\omega^*$ and Av is orthogonal to v. Hence, if we choose an arbitrary orthonormal base, $v_{\omega 1}, v_{\omega 2}, \cdots$ among the characteristic vectors of ω, we can define, if ω is complex,

$$v_{\omega^*,k}=\omega^{\frac{1}{2}}Av_{\omega,k} \quad \text{or} \quad Av_{\omega,k}=(\omega^{\frac{1}{2}})^*v_{\omega^*k} \tag{11}$$

and the $v_{\omega^*,k}$ will form a full base of orthonormal characteristic vectors to ω^*. The sign of the square root in (11) is best fixed in such a way that the imaginary part of $\omega^{\frac{1}{2}}$ shall have the same sign as the imaginary part of ω. Then $(\omega^*)^{\frac{1}{2}}=(\omega^{\frac{1}{2}})^*$. The purpose of the $\omega^{\frac{1}{2}}$ factor will become evident at once.

Application of A to both sides of (11) gives

$$Av_{\omega^*,k}=(\omega^{\frac{1}{2}})^*A^2v_{\omega,k}=(\omega^{\frac{1}{2}})^*\omega v_{\omega,k}=\omega^{\frac{1}{2}}v_{\omega,k}, \tag{12}$$

so that the choice of the characteristic vectors to ω^* made in (11) renders this equation valid also if ω is replaced by ω^*. The $v_{\omega,k}$ may be called characteristic vectors of A also.[2] However, in contrast to the unitary case, the characteristic vectors of A to ω also define the characteristic vectors of A to ω^* if we want (11) to hold. If one recalls that Λ is equivalent to a rotation it is not surprising that a certain amount of simplification results if a relation exists between the characteristic vectors of ω and of ω^*. In the case of a rotation one would set $v_{\omega^*,k}=v_{\omega,k}{}^*$.

Let us consider now a characteristic vector v to the characteristic value 1:

$$\Lambda v=A^2v=v. \tag{13}$$

It then follows from (10) that Av is also a characteristic vector to the characteristic value 1 and so is, unless it vanishes, $v_{11}=c(v+Av)$; c is a real normalization constant. It follows from (13) that

$$Av_{11}=Ac(v+Av)=c(Av+v)=v_{11}, \tag{14}$$

so that v_{11} is invariant under A. If $v=-Av$ we choose $v_{11}=iv$ and have again

$$Av_{11}=Aiv=-iAv=iv=v_{11}. \tag{15}$$

Next we consider another characteristic vector $v'=\Lambda v'$ which is orthogonal to v_{11}:

$$(v_{11},v')=0. \tag{16}$$

Because of (2) and (14),

$$(v_{11},Av')=(A^2v',Av_{11})=(\Lambda v',v_{11})=(v',v_{11})=0, \tag{17}$$

Av' will also be orthogonal to v_{11}. We can write therefore $v_{12}=c(v'+Av')$ or, if this vanishes, $v_{12}=iv'$, and this will still be orthogonal to v_{11} and also invariant under A. Proceeding in the same way, a full orthonormal base $v_{11}, v_{12}, \cdots$ of characteristic vectors of Λ to the characteristic value 1 can be found which are invariant under A,

$$Av_{1k}=v_{1k}. \tag{18}$$

The vectors which satisfy (18) can be called the invariant vectors of A. The procedure just used to ensure (18) is similar to the separation of real and imaginary parts of a number.

Let us finally consider a characteristic vector of Λ to -1:

$$\Lambda v_{-11}=A^2v_{-11}=-v_{-11}. \tag{19}$$

In this case again, because of (10), Av_{-11} is also a characteristic value to -1. Furthermore, Av_{-11} is orthogonal to v_{-11} because of (2) and (19):

$$(v_{-11},Av_{-11})=(A^2v_{-11},Av_{-11})=-(v_{-11},Av_{-11}). \tag{20}$$

Hence we can write

$$v_{-1^*1}=iAv_{-11} \quad v_{-11}=i^*Av_{-1^*1}=-iAv_{-1^*1}. \tag{21}$$

If Λ has further linearly independent characteristic vectors to -1, a normalized v_{-12} can be found which is orthogonal to both v_{-11} and v_{-1^*1}. Furthermore, the same will be true of $v_{-1^*2}=iAv_{-12}$. Thus, for instance,

$$\begin{aligned}(v_{-1^*2},v_{-11})&=(iAv_{-12},v_{-11})=-i(Av_{-11},A^2v_{-12})\\&=-i(-iv_{-1^*1},-v_{-12})=0.\end{aligned} \tag{22}$$

Hence, proceeding in the same way, one can find a full orthonormal base of characteristic vectors of Λ to -1,

$$\Lambda v_{-1k}=-v_{-1k} \quad \Lambda v_{-1^*k}=-v_{-1^*k}, \tag{23}$$

for which

$$v_{-1^*k}=iAv_{-1k} \quad v_{-1k}=-iAv_{-1^*k}=i^*A_{-1^*k} \tag{24}$$

holds. These equations are formally identical with the Eqs. (11) for complex characteristic values if one considers -1 to be two conjugate complex characteristic values -1 and -1^* of Λ, which happen to coincide. The v_{-1k} belong to the characteristic value -1, the v_{-1^*k} to the characteristic value -1^*. Equation (24) becomes a special case of (11) if one sets $(-1)^{\frac{1}{2}}=i$; $(-1^*)^{\frac{1}{2}}=i^*=-i$.

3.

On summarizing the preceding results, we can characterize an antiunitary operator by two sets of vectors, which jointly form a complete orthonormal set, together with the characteristic values $\omega_1, \omega_1^*, \omega_2, \omega_2^*, \cdots$ belonging to the second set. These characteristic values are pairwise conjugate complex, of modulus 1, but are not equal to 1. The first set of vectors are invariant under the antiunitary operator, i.e., (18) applies to them; (11) is valid for the members of the second set. The ω may also be equal to -1, but this characteristic value always occurs in pairs and one member of the pair is denoted by -1, the other by -1^*.

It will be shown now that any two sets of vectors v_{1k} and $v_{\omega k}$ which jointly form a complete orthonormal

[2] The two vectors $v_{\omega k}$ and $v_{\omega^* k}$ form a plane in our Hilbert space. The line which corresponds to this plane in Cartan's projective space is the invariant line of the passage cited in footnote reference 1.

set, together with the corresponding ω, give an antiunitary operator by means of (4), (11), and (18). In other words, the sets v_{1k} and $v_{\omega k}$ are not subject to any further conditions except that there are just as many vectors bearing the index ω as there are with the index ω^*. The number of vectors in the first set is arbitrary and so are the values of ω except that $\omega \neq 1$, $|\omega|=1$ and they occur in conjugate complex pairs.

In order to prove the preceding assertion we consider two vectors φ and ψ and expand them in terms of the orthonormal set

$$\begin{aligned} \varphi &= \sum_k a_k v_{1k} + \sum_{\omega k} b_{\omega k} v_{\omega k} \\ \psi &= \sum_k c_k v_{1k} + \sum_{\omega k} d_{\omega k} v_{\omega k}. \end{aligned} \tag{25}$$

$A\varphi$ and $A\psi$ are then given by

$$\begin{aligned} A\varphi &= \sum_k a_k^* v_{1k} + \sum_{\omega k} b_{\omega k}^* (\omega^{\frac{1}{2}})^* v_{\omega^* k} \\ A\psi &= \sum_k c_k^* v_{1k} + \sum_{\omega k} d_{\omega k}^* (\omega^{\frac{1}{2}})^* v_{\omega^* k}. \end{aligned} \tag{26}$$

Both conditions (1) and (2) of the antiunitary nature of A can be verified to be consequences of (26) and the orthonormality of the v_{1k}, $v_{\omega k}$, provided that

$$(\omega^*)^{\frac{1}{2}} = (\omega^{\frac{1}{2}})^* \quad |\omega| = 1. \tag{27}$$

For $\omega = -1$, this last condition is spelled out explicitly in (23). As was mentioned before, (27) can most simply be assured for complex ω by using that sign for $\omega^{\frac{1}{2}}$ for which the signs of the imaginary parts of ω and of $\omega^{\frac{1}{2}}$ are the same.

4.

Evidently, the two sets v_{1k}, $v_{\omega k}$ and the corresponding ω completely determine A. Conversely, A determines the number of vectors contained in the set v_{1k}—this is the multiplicity of the characteristic value 1 of A^2—and the value of the ω and their multiplicities. However, the vectors v are not completely determined by A and the present section will be devoted to the determination of the freedom that remains in the choice of these vectors.

Let us denote two other orthonormal sets which characterize the same antilinear operator by w_{1k} and $w_{\omega k}$. Since the w_{1k} form a base for the characteristic functions to the characteristic value 1 of $\Lambda = A^2$, they are connected with the v_{1k} by a nonsingular transformation

$$w_{1k} = \sum r_{kl} v_{1l}. \tag{28}$$

In fact, it follows from the orthonormality of the v_{1k} and of the w_{1k} that r is unitary. This is, however, not the only condition on r: If the vectors w_{1k} are to be invariant under A, i.e., if they satisfy (18),

$$\Lambda w_{1k} = \sum r_{kl}^* A v_{1l} = \sum r_{kl}^* v_{1l} = w_{1k}, \tag{29}$$

the r_{kl} must be real. Hence, two different invariant sets of vectors of the same antiunitary operator are related to each other by a rotation

$$r = r^* \quad rr' = rr^\dagger = 1. \tag{30}$$

The prime denotes the transpose, the dagger the Hermitian adjoint.

For complex ω, the sets $w_{\omega k}$ and $v_{\omega k}$ span the same linear manifold. Hence, we have

$$w_{\omega k} = \sum u_{kl}^{(\omega)} v_{\omega l}, \tag{31}$$

and it again follows from the orthonormality of the w_ω and v_ω that $u^{(\omega)}$ is unitary. By calculating $Aw_{\omega k}$ again, we find

$$Aw_{\omega k} = \sum u_{kl}^{(\omega)*} A v_{\omega l} = \sum u_{kl}^{(\omega)*} (\omega^*)^{\frac{1}{2}} v_{\omega^* l}, \tag{32}$$

so that if we want $Aw_{\omega k} = (\omega^*)^{\frac{1}{2}} w_{\omega^* k}$ to remain valid, we must have

$$u^{(\omega^*)} = u^{(\omega)*}, \tag{33}$$

i.e., the unitary transformations which belong to conjugate complex characteristic value are conjugate complex.

The preceding argument does not apply if $\omega = -1$. It is indeed clear that in this case the w_{-1k} may be linear combinations of the v_{-1k} and of the v_{-1^*k} because all these belong to the characteristic value -1 of Λ. Hence we set

$$w_{-1k} = \sum s_{kl} v_{-1l} + \sum l_{kl} v_{-1^*l}. \tag{34}$$

The condition (24) that $w_{-1^*k} = iAw_{-1k}$ now reads

$$\begin{aligned} w_{-1^*k} &= iA\left(\sum s_{kl} v_{-1l} + \sum l_{kl} v_{-1^*l}\right) \\ &= \sum s_{kl}^* iAv_{-1l} + \sum l_{kl}^* iAv_{-1^*l} \\ &= \sum -l_{kl}^* v_{-1l} + \sum s_{kl}^* v_{-1^*l}. \end{aligned} \tag{35}$$

Hence, the sets of vectors w_{-1} and w_{-1^*} are obtained from the sets v_{-1}, v_{-1^*} by the transformation

$$S = \begin{Vmatrix} s & l \\ -l^* & s^* \end{Vmatrix}. \tag{36}$$

This will guarantee that (24) is valid for the w_{-1}, w_{-1^*} if it is valid for the v_{-1}, v_{-1^*} because the second set of Eqs. (24) can be obtained from the first set by applying A to these. However, in order to make the w_{-1}, w_{-1^*} an orthonormal set, the S of (36) must be unitary. The conditions for this are obtained by setting $SS^\dagger = 1$ or, in terms of the submatrices s and l,

$$ss^\dagger + ll^\dagger = 1 \quad sl' = ls'. \tag{37}$$

It is easy to see that if the conditions (37) are satisfied, S becomes a simplectic matrix, i.e., it leaves the form

$$F = \begin{Vmatrix} 0 & 1 \\ -1 & 0 \end{Vmatrix} \tag{38}$$

invariant in the sense that

$$SFS'=F. \tag{39}$$

It follows that the sets w_{-1k}, w_{-1^*k} are obtained from the sets v_{-1k}, v_{-1^*k} by a unitary simplectic transformation.

The calculation of the last paragraph shows that the role of vectors $v_{\omega k}$, $v_{\omega^* k}$ for $\omega=-1$ is quite different from the role of the vectors $v_{\omega k}$, $v_{\omega^* k}$ for complex ω. The fact that the same Eq. (24) holds for $\omega=-1$ and for complex ω is somewhat accidental.

It may be well to note at this point that the equation

$$Aw=\nu w \tag{40}$$

with complex ν does not imply that ν is one of the ω. In fact, (40) holds with $w=(\nu^{\frac{1}{2}})^* v_{1k}$ and an arbitrary ν.

5.

Lastly, we shall determine the number of free parameters in an antiunitary transformation which can be characterized by l invariant vectors; $2m$ vectors with the characteristic value -1; 2ρ different complex characteristic values with positive imaginary parts and their complex conjugates with multiplicities $c_1, c_2, \cdots, c_\rho$. These are then also the multiplicities of the corresponding conjugate complex characteristic values. Hence,

$$l+2m+2c_1+2c_2+\cdots+2c_\rho=n, \tag{41}$$

where n is the number of dimensions of the underlying Hilbert space which will be assumed to be finite dimensional in the present section.

The number of free parameters will be calculated by adding the free parameters necessary to characterize the complete orthonormal set v_{1k}, $v_{\omega k}$ and the ω, and subtracting the number of parameters contained in the transformations which alter the v but leave A unchanged. These were determined in the preceding section.

A complete orthonormal set in n dimensions can be characterized by $2n-1+(2n-3)+\cdots+3+1=n^2$ parameters. The number of free parameters in the ω is just ρ. Hence, $n^2+\rho$ parameters are necessary to characterize the v and the ω.

A rotation in the l-dimensional space of the v_{1k} does not change A. The number of parameters of such a rotation is $\frac{1}{2}l(l-1)$. Similarly, a $2m$-dimensional unitary simplectic transformation remains free for the vectors $v_{-11}, \cdots, v_{-1m}, v_{-1^*1}, \cdots, v_{-1^*m}$. The number of parameters of such a transformation is $m(2m+1)$. Finally, an arbitrary unitary transformation of the vectors $v_{\omega 1}, v_{\omega 2}, \cdots$ leaves A also unchanged if the conjugate complex transformation is applied to the vectors $v_{\omega^* 1}, v_{\omega^* 2}, \cdots$. The number of parameters in such a transformation is just the square of the corresponding c. Hence, the total number of free parameters in the antiunitary transformation is

$$p=n^2+\rho-\tfrac{1}{2}l(l-1)-m(2m+1)-\sum_1^\rho c_r^2$$

$$=n^2-\tfrac{1}{2}l(l-1)-m(2m+1)-\sum_1^\rho (c_r^2-1). \tag{42}$$

For even n, the number of parameters is just n^2 if all the characteristic values are complex and simple. Two invariant vectors decrease the number of parameters by 1, two characteristic values -1 by 3, if a complex characteristic value is doubly degenerate (the same then holds for the conjugate complex characteristic value) the number of parameters is also decreased by three.

The number of free parameters is also n^2 if n is odd and there are $n-1$ simple complex characteristic values and one invariant vector. Multiplicities among the complex characteristic values and the presence of a characteristic value -1 (which is always at least double) reduce the number of free parameters as in the case of even n.

The fact that the number of parameters is n^2 in the general case could have been inferred from the possibility of representing an antiunitary transformation in the form (6), i.e., as the product of a unitary transformation and complex conjugation. The number of free parameters in an n-dimensional unitary transformation is just n^2. The decrease in the number of free parameters (by 3) caused by the presence of a single pair of characteristic values -1 is remarkable.

6.

The preceding results will now be formulated in the language of projection operators and thus extended to the case in which there is a continuous spectrum. However, the proofs, which are rather obvious, will be omitted.

Consider again the unitary operator $\Lambda=A^2$. If 1 and -1 belong to the point spectrum of Λ, denote the corresponding projection operators by E_1 and E_{-1}. The projection operator which belongs to an interval J of the unit circle in the complex plane will be denoted by E_J. All these projection operators are self-adjoint, commute with Λ and with each other; the product of two of them is equal to the projection operator which corresponds to the intersection of the domains to which the two factors correspond. Furthermore,

$$\Lambda E_1=E_1 \quad \Lambda E_{-1}=-E_{-1} \quad \lim \Lambda E_J=\omega E_J, \tag{43}$$

where the lim in the last equation indicates that J is an infinitely narrow interval around ω. We define the antiunitary operators

$$A_1=AE_1 \quad A_{-1}=AE_{-1} \quad A_J=AE_J, \tag{44}$$

then

$$A=A_1+A_{-1}+\lim \sum A_J. \tag{45}$$

The lim again indicates that the intervals J are infinitely narrow; they cover all the unit circle with the exceptions of the points 1 and -1. The intervals J will be assumed to lie either entirely in the upper half-plane, or entirely in the lower half-plane. The interval J^* will be the conjugate complex of the interval J.

It is good to recall, for the rest of this discussion, that A^{-1} is also an antiunitary operator and is, in fact, given by

$$A^{-1}=\Lambda^{-1}A=A\Lambda^{-1}. \tag{46}$$

A transforms every projection operator into the projection operator which corresponds to the conjugate complex domain

$$AE_1A^{-1}=E_1 \quad AE_{-1}A^{-1}=E_{-1} \quad AE_JA^{-1}=E_{J^*}. \tag{47}$$

These equations can be given a variety of forms by combining them with (43) and (44). The most interesting of these forms gives the projection operators in terms of the A_J. Thus

$$A_1^2=AE_1AE_1=E_1A^2E_1=E_1\Lambda E_1=E_1^2=E_1. \tag{48}$$

Similarly,

$$A_{-1}^2=-E_{-1} \quad \lim A_{J^*}A_J=\omega E_J. \tag{49}$$

Whereas, if J and L do not overlap,

$$A_{L^*}A_J=0. \tag{50}$$

These equations form a substitute for the equations involving the characteristic vectors v of A. As an example, we show that v_{-1k} and Av_{-1k} are orthogonal or, in the present language, that $E_{-1}\varphi$ and $AE_{-1}\varphi$ are orthogonal for any φ

$$\begin{aligned}(E_{-1}\varphi,AE_{-1}\varphi)&=(A^2E_{-1}\varphi,AE_{-1}\varphi)\\&=(-E_{-1}\varphi,AE_{-1}\varphi)=0.\end{aligned} \tag{51}$$

The second form follows from the antiunitary nature of A, the third from (43).

Phenomenological Distinction Between Unitary and Antiunitary Symmetry Operators

E. P. Wigner

Journal of Mathematical Physics *1*, 414–416 (1960)

(Received April 25, 1960)

It is well known that one always can find as many orthogonal states (i.e., states between which the transition probability is zero) as the Hilbert space has dimensions which are invariant under a given unitary transformation. The corresponding vectors are characteristic vectors of the unitary operator. In contrast, most antiunitary operators leave not more than one state invariant. However, if there are two orthogonal invariant states, a consideration of the states for which the transition probability is $\frac{1}{2}$ into both invariant states surely provides a distinction. In the antiunitary case, one of these states is also invariant, another one is transformed into an orthogonal state, the rest are in between. In the unitary case, the transition probability between original state and transformed state is the same for all states for which the transition probability is $\frac{1}{2}$ into two orthogonal states. This provides a "directly observable" distinction between unitary and antiunitary transformations.

1.

THE invariance transformations of quantum mechanics are transformations in a complex Hilbert space which leave the absolute value of the scalar product of any two vectors φ, ψ invariant:

$$|(\varphi,\psi)| = |(T\varphi,T\psi)|. \tag{1}$$

The reason for the invariance of (1) is that this absolute value (or, rather, its square) is, according to the usual physical interpretation of quantum mechanics, operationally meaningful: It is the transition probability between the two states characterized by the two vectors φ and ψ. It is well known that the transformations T which satisfy (1) fall into one of two categories: They can be unitary or antiunitary.

The last statement should be formulated more precisely. The physical state does not determine the state vector φ completely, all multiples of φ (the whole "ray of φ") describe the same state. It is customary and useful to normalize the state vectors, i.e., to choose a vector from the ray of φ which is normalized,

$$(\varphi,\varphi)=1. \tag{2}$$

Even then, a phase factor (i.e., a factor of modulus 1) remains free in φ. The same applies to the transformed state, the ray of $T\varphi$. Note that (1) remains valid if φ, ψ, $T\varphi$, and $T\psi$ are multiplied by arbitrary phase factors. The precise formulation of the statement at the end of the preceding paragraph stipulates the possibility of a certain choice of the state vector from the ray of the transformed state, which corresponds to any choice of a state vector φ from the ray of the original state. The choice referred to therefore replaces the physically given ray correspondence by a vector correspondence. The theorem states that the choice can be made either in such a way that for any two state vectors φ, ψ and any two numbers a and b,

$$T(a\varphi+b\psi)=aT\varphi+bT\psi \quad\text{and}\quad (\varphi,\psi)=(T\varphi,T\psi) \tag{3a}$$

(unitary case), or in such a way that

$$T(a\varphi+b\psi)=a^*T\varphi+b^*T\psi \quad\text{and}\quad (\varphi,\psi)=(T\psi,T\varphi). \tag{3b}$$

In the latter case T (and also the ray correspondence from which it derives) is called antiunitary.

It is easy to show, by abstract reasoning, that no transformation can fall into both classes, i.e., that if the choice of $T\varphi$ from its ray can be made in such a way that (3a) becomes valid, it cannot be made in such a way that (3b) becomes valid, and conversely.

It follows that it must be possible to ascertain the unitary or antiunitary nature of a transformation by considering only transition probabilities, i.e., absolute values of scalar products. However, it may be of some interest to spell this out in detail, and it is the aim of the present article to do this. Use will be made, for this purpose, of the normal form of antiunitary operators obtained in the preceding article.[1] Similarly, the unitary operators will be described by their characteristic values and the corresponding characteristic vectors.

We note for further reference that if φ and ψ are normalized state vectors, they represent the same state if $\varphi=\omega\psi$ with an arbitrary ω of modulus 1. In this case $|(\varphi,\psi)|=1$. On the other hand, if

$$(\varphi,\psi)=0, \tag{4}$$

φ and ψ represent states which have, at least in some respect, opposite properties. If (4) holds for two states, we call them orthogonal: The transition probability between them is zero. Orthogonality is, therefore, an operationally verifiable relation between two states. The number of the mutually orthogonal states is equal to the dimension of the underlying Hilbert space. We shall assume henceforth that this is larger than 1.

[1] E. P. Wigner, J. Math. Phys. 1, 409 (1960), preceding article.

2.

Let us consider, first, the states which are left unchanged by the transformation. For these,

$$T\varphi=\omega\varphi, \tag{5}$$

and we shall call them invariant states. If T is unitary, φ is one of its characteristic vectors. It follows that there is at least one set of mutually orthogonal states which contains as many members as *any* set of mutually orthogonal states contains. In particular, unless there are at least two orthogonal invariant states, the symmetry operator cannot be unitary and must be, therefore, antiunitary.

If the spectrum of the unitary operator is simple, the invariant states are isolated, but if it has a characteristic value of multiplicity l, the corresponding invariant states form a continuous manifold with $2l-2$ parameters. The characteristic functions have l complex or $2l$ real parameters but the normalization condition subjects these to one real equation and one real parameter, characterizing the phase factor, is physically meaningless.

Let us consider now an antiunitary transformation. We decompose the state vector into the invariant and characteristic vectors of the operator T

$$\varphi=\sum_k a_k v_{1k}+\sum_{\omega k} b_{\omega k} v_{\omega k}. \tag{6}$$

If $T\varphi=\omega'\varphi$ (i.e., φ is an invariant state), we can consider $\varphi'=\omega'^{\frac{1}{2}}\varphi$ and have

$$T\varphi'=(\omega'^{\frac{1}{2}})^*T\varphi=(\omega'^{\frac{1}{2}})^*\omega\varphi=(\omega'^{\frac{1}{2}})^*\varphi=\varphi'. \tag{7}$$

Since φ and $\omega'^{\frac{1}{2}}\varphi$ represent the same state, it suffices, in the antiunitary case, to find those state vectors for which

$$T\varphi=\varphi. \tag{8}$$

Since, for the φ of (6)

$$\begin{aligned} T\varphi&=\sum a_k^* v_{1k}+\sum b_{\omega k}^*(\omega^{\frac{1}{2}})^* v_{\omega^* k} \\ &=\sum a_k^* v_{1k}+\sum b_{\omega^* k}^*\omega^{\frac{1}{2}} v_{\omega k}, \end{aligned} \tag{9}$$

φ will satisfy (8) if

$$a_k=a_k^* \tag{10}$$

and

$$\omega^{\frac{1}{2}} b_{\omega^* k}^*=b_{\omega k}. \tag{11}$$

Since (11) must hold for all ω and all k, ω can be replaced therein by ω^* to give

$$(\omega^*)^{\frac{1}{2}} b_{\omega k}^*=b_{\omega^* k} \quad \text{or} \quad b_{\omega^* k}^*=\omega^{\frac{1}{2}} b_{\omega k}. \tag{12}$$

Insertion of the latter expression into (11) gives

$$\omega b_{\omega k}=b_{\omega k} \quad \text{or} \quad b_{\omega k}=0 \tag{13}$$

since $\omega\neq 1$. It follows that if the symmetry operator is antiunitary, the invariant states can be described by invariant vectors, i.e., by real linear conbinations of the v_{1k}.

It follows that the structure of the set of invariant states is, in general, very different for unitary and antiunitary operators. An antiunitary operator may have no invariant state, or it may have only one. However, no matter how many invariant states it has, they form a continuous manifold each member of which can be changed continuously into any other. If the antiunitary operator has n orthogonal invariant vectors, the manifold of invariant states is an $n-1$ dimensional continuous manifold. Its state vector can be described by n real coefficients the sum of the squares of which is 1. There is no arbitrary phase factor in this case because (8) already determines the phase factor.

3.

The number and topological properties of the invariant states actually permit a phenomenological distinction between unitary and antiunitary transformations. The following distinction is, however, more direct.

Consider two orthogonal invariant states, φ_1 and φ_2. If there are no such, the transformation is surely antiunitary. Next, consider the states for which the transition probability is $\frac{1}{2}$ into both φ_1 and φ_2. The state vectors of these states are

$$\psi_\alpha=2^{-\frac{1}{2}}(\varphi_1+e^{i\alpha}\varphi_2). \tag{14}$$

Consider finally the transition probability of these states into the states $T\psi_\alpha$,

$$P(\alpha)=|(\psi_\alpha,T\psi_\alpha)|^2. \tag{15}$$

If T is unitary,

$$T\psi_\alpha=2^{-\frac{1}{2}}(\omega_1\varphi_1+e^{i\alpha}\omega_2\varphi_2), \tag{16}$$

and the transition probability becomes

$$\begin{aligned} P(\alpha)&=\tfrac{1}{4}|(\varphi_1+e^{i\alpha}\varphi_2,\ \omega_1\varphi_1+e^{i\alpha}\omega_2\varphi_2)|^2=\tfrac{1}{4}|\omega_1+\omega_2|^2 \\ &=\tfrac{1}{2}(1+\mathrm{Re}\,\omega_1\omega_2^*). \end{aligned} \tag{17}$$

It is independent of α, i.e., the same for all the states ψ_α with transition probabilities $\frac{1}{2}$ into φ_1 and φ_2.

Let us assume next that T is antiunitary. In this case

$$T\psi_\alpha=2^{-\frac{1}{2}}(\varphi_1+e^{-i\alpha}\varphi_2), \tag{18}$$

and the transition probability becomes

$$\begin{aligned} P(\alpha)&=\tfrac{1}{4}|(\varphi_1+e^{i\alpha}\varphi_2,\ \varphi_1+e^{-i\alpha}\varphi_2)|^2=\tfrac{1}{4}|1+e^{-2i\alpha}|^2 \\ &=\tfrac{1}{2}(1+\cos 2\alpha). \end{aligned} \tag{19}$$

It varies, for the states in question, between 0 and 1.

The preceding argument can yet be greatly generalized. In the unitary case, if the transition probabilities into n orthogonal invariant states $\varphi_1, \varphi_2, \cdots, \varphi_n$ are prescribed to be $r_1^2, r_2^2, \cdots, r_n^2$ so

that

$$\sum r_k^2 = 1, \tag{20}$$

the general form of the state vector is

$$\psi_{\alpha_1 \cdots \alpha_n} = \sum r_k e^{i\alpha_k} \varphi_k, \tag{21}$$

with arbitrary $\alpha_1, \cdots, \alpha_n$. Nevertheless, the transition probability from any of these $\psi_{\alpha_1 \cdots \alpha_n}$ into the corresponding $T\psi_{\alpha_1 \cdots \alpha_n}$ is the same, namely,

$$\begin{aligned} &|(\psi_{\alpha_1 \cdots \alpha_n}, T\psi_{\alpha_1 \cdots \alpha_n})|^2 \\ &\quad = |(\sum r_k e^{i\alpha_k} \varphi_k, \sum r_k e^{i\alpha_k} \omega_k \varphi_k)|^2 \\ &\quad = |\sum r_k^2 \omega_k|^2. \end{aligned} \tag{22}$$

In the antiunitary case, on the other hand,

$$T\psi_{\alpha_1 \cdots \alpha_n} = \sum r_k e^{-i\alpha_k} \varphi_k, \tag{23}$$

and the transition probability becomes

$$|(\psi_{\alpha_1 \cdots \alpha_n}, T\psi_{\alpha_1 \cdots \alpha_n})|^2 = |\sum r_k^2 e^{-2i\alpha_k}|^2. \tag{24}$$

Unless one of the r_k^2 is larger than $\frac{1}{2}$, this still will assume every value between 0 and 1 for suitably chosen α.

The striking difference in the relation of the original and transformed states to each other shows particularly clearly how definite the relations in question are in either case.

On Weakly Positive Matrices

E. P. Wigner

Canadian Journal of Mathematics *15*, 313–317 (1963)

A matrix is said to be positive definite if it is hermitian and if all of its characteristic values are positive. It is well known,[1] and easy to prove, that the necessary and sufficient condition for a matrix P to be positive definite is that its hermitian quadratic form

$$(v, Pv) > 0 \tag{1}$$

with any vector $v \neq 0$ be positive. (This will imply, in the present article, that it is real.) It is easy to see from (1) that if P_1 and P_2 are positive definite, the same holds of $a_1P_1 + a_2P_2$ if a_1 and a_2 are positive numbers. Similarly, if P is positive definite, so is $X^\dagger PX$ for any non-singular X, where $X^\dagger$ is the hermitian adjoint of X. Indeed, the hermitian quadratic form of $X^\dagger PX$ with the vector v,

$$(v, X^\dagger PXv) = (Xv, PXv), \tag{1a}$$

is equal to the hermitian quadratic form of P with the vector Xv and is, therefore, positive.

We shall call a matrix W weakly positive if all of its characteristic values are positive and the characteristic vectors form a complete set. An equivalent definition is that W have a diagonal form with (real) positive diagonal elements

$$W = XDX^{-1}, \tag{2}$$

where X is a non-singular matrix and D a diagonal matrix with positive diagonal elements.

Clearly, a matrix is positive definite if it is weakly positive and hermitian. In this case, the X of (2) can be assumed to be unitary. Conversely, a matrix which is positive definite is also weakly positive. It is also clear that a weakly positive matrix is necessarily non-singular and that its reciprocal is also weakly positive. Similarly, the reciprocal of a positive definite matrix is also positive definite.

Let us note, finally, that any XPX^{-1} is weakly positive if X is non-singular and P positive definite because a positive definite P can be written as UDU^{-1} with a unitary U so that XPX^{-1} becomes $(XU)D(XU)^{-1}$.

THEOREM 1. *A matrix is weakly positive if and only if it can be written as the product of two positive definite matrices*

$$W = P_1P_2. \tag{3}$$

Received April 16, 1962.

[1]See, for instance, P. R. Halmos, *Finite dimensional vector spaces* (Princeton University Press, 1942), § 64).

To prove this, one notes that the W of (2) is equal to P_1P_2, where

$$P_1 = XX^\dagger, \qquad P_2 = (X^\dagger)^{-1}DX^{-1} \tag{4}$$

or

$$P_1 = XDX^\dagger, \qquad P_2 = (XX^\dagger)^{-1}, \tag{5}$$

where the dagger denotes the hermitian adjoint. All P_1, P_2 in (4) and (5) are positive definite because they have the form (1*a*) with $P = 1$ or $P = D$ and various non-singular X.

The converse, that P_1P_2 is weakly positive if P_1 and P_2 are positive definite can be demonstrated in two steps. First, if λ is a characteristic value of P_1P_2,

$$P_1P_2v = \lambda v, \tag{6}$$

one can form the scalar product of (6) with P_2v:

$$(P_2v, P_1P_2v) = \lambda(P_2v, v) = \lambda(v, P_2v). \tag{6a}$$

The left side is the hermitian quadratic form of P_1 with the non-vanishing vector P_2v, and hence positive. Since P_2 is positive definite, (v, P_2v) is also positive and λ becomes the ratio of two positive numbers.

If P_1P_2 had a deficiency of characteristic vectors, there would be a characteristic value λ for which

$$P_1P_2v = \lambda v \qquad \text{and} \qquad P_1P_2w = \lambda w + v \tag{7}$$

would hold[2] with a suitable w. The scalar product of the first equation with P_2w is

$$(P_2w, P_1P_2v) = \lambda(P_2w, v) \qquad \text{or} \qquad (P_2w, P_1P_2v) = \lambda(w, P_2v). \tag{7a}$$

Similarly, the scalar product of the second equation (7) with P_2v is

$$(P_1P_2w, P_2v) = \lambda(w, P_2v) + (v, P_2v). \tag{7b}$$

Since P_1 is hermitian, the left sides of (7*b*) and of (7*a*) are equal. It then follows that $(v, P_2v) = 0$—a contradiction since P_2 is positive definite. Hence, (7) cannot hold and the characteristic vectors of P_1P_2 form a complete set.

The second part of Theorem 1 also follows from the fact that, if P_2 is positive definite, there is a square root q of $P_2 = q^2$ which is also positive definite.[3] However,

$$P_1P_2 = P_1q^2 = q^{-1}\cdot qP_1q\cdot q \tag{8}$$

so that, if v' is a characteristic vector of the positive matrix qP_1q,

$$qP_1qv' = \lambda v'. \tag{8a}$$

$q^{-1}v' = v$ is a characteristic vector of P_1P_2:

[2]See any description of the so-called Jordan canonical form, for instance, Halmos, Appendix I, or N. Dunford and J. T. Schwartz, *Linear operators*, Part I (Interscience Publishers, New York, 1958), p. 563, example 17.

[3]See, for instance, Halmos, § 66, or Dunford and Schwartz, p. 569, Theorem 11.

$$P_1P_2v = q^{-1}\cdot qP_1q\cdot v' = q^{-1}\lambda v' = \lambda v. \tag{8b}$$

Hence, the characteristic values of P_1P_2 and of qP_1q are the same. Further, since every vector can be expressed linearly in terms of the v', the expression for the vector $q^{-1}x$ in terms of the v' gives the expression for x in terms of the v so that these indeed form a complete set.

THEOREM 2. *If the product of a positive definite and of a weakly positive matrix is hermitian, it is also positive definite. The same holds of the product $P_1P_2P_3$ of three positive definite matrices P_1, P_2, P_3 if it is hermitian.*

Since a weakly positive matrix is the product of two positive definite matrices, it suffices to prove the second part of the theorem. This is a generalization, to three factors, of the well-known theorem[4] that, if the product of *two* positive definite matrices is hermitian, it is also positive definite.

In order to prove Theorem 2, consider the matrices

$$H(\lambda) = [(1-\lambda)P_3 + \lambda P_1]P_2P_3 \tag{9}$$

with positive real $\lambda \leqslant 1$. As a linear combination of two hermitian matrices $P_3P_2P_3$ and $P_1P_2P_3$ with positive coefficients, $H(\lambda)$ is itself hermitian, its characteristic values real. For $\lambda = 0$ they are all positive since P_2 and hence also $P_3P_2P_3$ is positive definite (see (1*a*)). As λ increases, the characteristic values change continuously and they will remain positive for all $\lambda \leqslant 1$ unless there is a 0 characteristic value for some intermediate λ. This is not possible, however, because, if v is non-zero, the same holds for P_3v, and hence for P_2P_3v, and finally for $H(\lambda)v$. The last conclusion follows from the positive definite nature of $(1-\lambda)P_3 + \lambda P_1$ which is (for $0 \leqslant \lambda \leqslant 1$) a linear combination of positive definite matrices with positive coefficients.

Remarks. It may be worth noting that $a_1W_1 + a_2W_2$ need not be weakly positive for weakly positive W_1, W_2 and positive a_1, a_2. Thus, if $a_1 = a_2 = 1$,

$$W_1 = \begin{Vmatrix} 10 & 6 \\ 1 & 1 \end{Vmatrix}, \qquad W_2 = \begin{Vmatrix} 10 & 1 \\ 6 & 1 \end{Vmatrix}, \tag{10}$$

the determinant of $W_1 + W_2$ is negative and hence $W_1 + W_2$ is not weakly positive. It is also easy to find two weakly positive matrices, or four positive definite matrices, the product of which is not positive even though it is hermitian. Two such W, with their product, are

$$\begin{Vmatrix} 1 & 1+\epsilon \\ -\epsilon & -\epsilon \end{Vmatrix} \cdot \begin{Vmatrix} 1 & \epsilon+\epsilon^2 \\ -1 & -\epsilon \end{Vmatrix} = \begin{Vmatrix} -\epsilon & 0 \\ 0 & -\epsilon^3 \end{Vmatrix}$$

with a sufficiently small but positive ϵ.

Theorem 3 will extend Theorems 1 and 2 to bounded linear operators in Hilbert space. For this purpose, the notions of positive definiteness and weak

[4]See, for instance, Halmos, §§ 56, 72.

positivity have to be formulated for such operators. A bounded linear operator will be called positive definite if it is self-adjoint and if it has a positive lower bound so that

$$(v, Pv) > b(v, v), \qquad b > 0. \tag{11}$$

Such an operator has a reciprocal which is positive definite and also a square root which is also positive definite.[5]

A bounded linear operator will be called weakly positive if it can be expressed in the form

$$W = XDX^{-1}, \tag{12}$$

where D is positive definite and both X and its reciprocal X^{-1} are bounded. Since X can be written[6] as $X = HU$ with a self-adjoint, in fact positive definite, H, and a unitary U, and since UDU^{-1} is also a positive definite operator, it can be assumed that all three factors in (12) are positive definite.

Let E_λ be the self-adjoint projection operators in the spectral decomposition of D,

$$D = \int \lambda dE_\lambda. \tag{13}$$

W will also have a spectral decomposition

$$W = \int \lambda dF_\lambda, \tag{14}$$

where, however, the projection operators

$$F_\lambda = XE_\lambda X^{-1} \tag{14a}$$

are not, in general, self-adjoint. They have the multiplication properties of the E_λ

$$F_\lambda F_\mu = F_\mu F_\lambda = F_\lambda \qquad \text{for } \lambda \leqslant \mu. \tag{14b}$$

Note that the integrals in (13) and (14) have to be extended only between the positive lower and upper bounds of D, i.e., over a finite positive interval.

THEOREM 3. *Theorems* 1 *and* 2 *hold also for bounded linear operators in Hilbert space.*

The first part of the extension of Theorem 1 follows from the fact that the P_1 and P_2 given by (4) or (5) satisfy (3) identically. As was mentioned before, it can be assumed that X is self-adjoint so that

[5]See, for instance, F. Riesz and B. Sz-Nagy, *Leçons d'analyse fonctionelle* (Akadémiai Kiadó, Budapest, 1953), Section 120, or F. J. Murray, *Introduction to linear transformations in Hilbert space* (Princeton University Press, 1941), Chap. VIII, § 3.

[6]See Riesz and Nagy, Section 110. Since X has a bounded reciprocal, the same holds for $H = (X^\dagger X)^{\frac{1}{2}}$ and Hv covers the whole Hilbert space. Hence, according to Riesz and Nagy, U is isometric rather than merely quasi-isometric. Further, if there were a vector W orthogonal to all Uv, then $H^{-1}w$ would be orthogonal to all $HUv = Xv$. This is not possible if X has a reciprocal. Hence U is unitary in the case considered here.

$$W = X^2 \cdot X^{-1}DX^{-1} \tag{15}$$

is a possible decomposition of W into the product of two positive definite operators. Indeed, both X^2 and $X^{-1}DX^{-1}$, as well as the reciprocals of these operators, are bounded. Thus, if b and B are the lower and upper bounds of D,

$$b(X^{-1}v, X^{-1}v) < (X^{-1}v, DX^{-1}v) = (v, X^{-1}DX^{-1}v) < B(X^{-1}v, X^{-1}v) \tag{15a}$$

and if ξ and Ξ are lower and upper bounds of X,

$$\Xi^{-2}(v, v) < (X^{-1}v, X^{-1}v) < \xi^{-2}(v, v) \tag{15b}$$

so that

$$b\Xi^{-2}(v, v) < (v, X^{-1}DX^{-1}v) < \xi^{-2}B(v, v), \tag{15c}$$

i.e., $X^{-1}DX^{-1}$ has positive upper and lower bounds. The same can be demonstrated for the other P which occur in (4) or (5).

The second part of the extension of Theorem 1 follows by the argument contained in equations (8). The existence of a positive definite square root of a positive definite operator was mentioned before and bounds for $D = qP_1q$ can be established in the same way as they were established for $X^{-1}DX^{-1}$.

As to the extension of Theorem 2, one can note that a lower bound of $H(\lambda)^2$ is $b^2b_2^2b_3^2$, where b_2 and b_3 are the lower bounds for P_2 and P_3 and b is the lower bound of P_1 or P_3, whichever is smaller. Hence, b is also a lower bound for $P_\lambda = (1 - \lambda)P_3 + \lambda P_1$ for every λ between 0 and 1. It then follows that b^2 is a lower bound for P_λ^2, so that

$$\begin{aligned}(v, H(\lambda)^2v) &= (H(\lambda)v, H(\lambda)v) = (P_\lambda P_2P_3v, P_\lambda P_2P_3v)\\ &= (P_2P_3v, P_\lambda^2 P_2P_3v) > b^2(P_2P_3v, P_2P_3v)\\ &= b^2(P_3v, P_2^2 P_3v) > b^2b_2^2 (P_3v, P_3v) > b^2 b_2^2 b_3^2(v, v).\end{aligned}$$

It then follows that the spectrum of $H(\lambda)$ does not extend over the interval $(-bb_2b_3, bb_2b_3)$ as long as $0 \leqslant \lambda \leqslant 1$. However, the lower bound of $H(\lambda)$ is a continuous function of λ and since it is larger than bb_2b_3 for $\lambda = 0$ and cannot enter the $(-bb_2b_3, bb_2b_3)$ interval, it always stays above this interval. Hence, $H(\lambda)$ and, in particular, $H(1) = P_1P_2P_3$ are positive definite.

Princeton University

Unitary Representations of the Inhomogeneous Lorentz Group Including Reflections

E. P. Wigner

In: Group Theoretical Concepts in Elementary Particle Physics (Feza Gursey, ed.).
Gordan and Breach, New York 1964, pp. 37–80

I. Introduction

The present notes do not relate to a subject which is now at the center of interest. To be sure, symmetry considerations play a great and perhaps still increasing role in theoretical physics. However, we shall be concerned only with what may be called kinematical symmetry properties, i.e., symmetry properties of the space-time continuum in which the physical events occur. In contrast to the dynamical laws of symmetry, which relate to symmetries of interaction and of properties of particles, the kinematical laws of symmetry are supposed to be well known. The purpose of the lectures which are recorded in these notes was to make this statement more nearly true.

Let me introduce the discussion by admitting that the existence of symmetries and invariances, and their significance even many years ago, but particularly their increasing significance in the past years, appears to me to border on the mysterious. If we look at the body of our knowledge in physics, we must note that it divides the phenomena into two groups, in fact into two sharply distinguished groups. Into the first group falls our knowledge of what may be called initial conditions, that we are on the Earth, that there is a Sun around, and so on. We maintain that these initial conditions are accidental, that there is no precise theory which would explain them, that there *can be no such theory*. Into the second category falls our knowledge of the "laws of motion" which tell us how to foresee the future state of the World, assuming that we know the present state. These laws, we are convinced, are exact, of almost unbelievable precision, and of a

beauty and simplicity which is much greater and deeper than any that we could have invented.

This sharp division of our knowledge and even of the knowable is strange enough. It is stranger yet to assume that the laws of motion have a structure themselves, that they conform with some general principles which, though they do not determine the laws of motion, do impose certain restrictions on them. Further, these superprinciples are so much easier to divine than the laws of motion themselves that we in fact know sufficiently many of them to reject the vast majority of possible laws of motion as unsatisfactory even if we can quote no single phenomenon which would directly contradict them. The superprinciples I am speaking about are, of course, the laws of invariance.

As was mentioned before, we shall be concerned only with the geometrical or kinematical principles of invariance. These are supposed to be firmly established as consequences of the invariance of the laws of motion under the inhomogeneous Lorentz group. They do not have the timely and almost fashionable significance of the other invariances which refer more directly and more exclusively to the laws of motion; these will and should dominate our conference. However, I firmly believe that the geometrical laws of invariance continue to be of fundamental importance. Neither is it true that they form a closed book. First, it is not certain that it is the inhomogeneous Lorentz group which plays the fundamental role, and I am very much looking forward to the remarks which Professor Gürsey will make on the de Sitter groups. Mr. T. Philips and I have also devoted lately a good deal of attention to the possibility that one of these groups is the fundamental one and have come to some rather surprising conclusions. Second, the interpretation of the reflection operators, to which most of this discussion will be devoted, has raised a number of difficult questions on the role of measurability. Lastly, I should like to mention the possibility of using a number system for the Hilbert space which is different from the ordinary complex numbers. Dr. Michel's lectures will implicitly deal with this problem also. This is, as far as I know, virgin territory.

Much of the material which follows was taken from a rather old but unpublished manuscript by V. Bargmann, A. S. Wightman and myself. A short outline of the subsequent sections follows.

(2) The restricted Poincaré group. This is also called inhomogeneous proper Lorentz group; it contains no reflections.

(3) The quantum mechanical restricted Poincaré group. The quantum mechanical operators which correspond to the symmetry operations of the group of section 2 form a group which is homomorphic but not necessarily isomorphic to that group. The group of the operators is called the quantum mechanical restricted Poincaré group. This will be described and analyzed.

(4) The relevant representations of the quantum mechanical restricted Poincaré group. The representations which describe systems with non-zero rest-mass and positive energy are determined. Only unitary representations are considered.

(5) The irreducible representations of the restricted quantum mechanical Poincaré group. The irreducible ones of the representations of the preceding section are specified and some remarks made on the reduction of representations of non-compact groups.

(6) The extended Poincaré group. The Poincaré group which contains space, time, and space-time reflections, is described and the physical meaning of these reflections discussed.

(7) The full quantum mechanical Poincaré groups. The quantum mechanical operators which correspond to the operations of the preceding section always form a group which is, similar to the case of the restricted group, always homomorphic but not necessarily isomorphic to the extended Poincaré group. Depending on the nature of the system considered, the operators form one of four groups. These are determined and analyzed.

(8) Inclusion of space reflection symmetry. This section determines the representations of the group of quantum mechanical operators which correspond to proper Poincaré transformations and to transformations involving space reflections. Again, only unitary representations are considered and those which describe systems with finite rest-mass and positive energy.

(9) Adjoining the time reflection operators. The considerations of the preceding section are extended to the full Poincaré group. The antiunitary operators which correspond to time inversions and to space-time inversions are determined.

(10) Attempt at physical interpretation. The representations which correspond to light quanta and to particles which are different from their anti-particles are determined. It is assumed in this section that all states form a Hilbert space and hence that any two state vectors, which are not constant multiples of each other, are distinguishable.

(11) Critique of the interpretation. The effect of the probable existence of superselection rules is traced on the results of the preceding section.

Sections 2 to 5 review material that is "well known" and are somewhat sketchy.

2. The Restricted Poincaré Group

The restricted Poincaré group's operations can be obtained as products of the operations of the displacement group in four dimensions and of the proper Lorentz transformatives Λ. The latter are defined as the real four dimensional matrices which leave the form G

$$G = \begin{Vmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{Vmatrix} \tag{2.1}$$

invariant, i.e. for which

$$\Lambda G \Lambda^T = G. \tag{2.2}$$

It is further postulated that the determinant of Λ is 1, and the element in the upper left corner, $\Lambda_{tt} > 0$. It is then, in fact, $\Lambda_{tt} \geq 1$. The product of the displacement by the four vector a and the Lorentz transformation Λ will be denoted by (a, Λ); the group equation then reads

$$(a_1, \Lambda_1)(a_2, \Lambda_2) = (a_1 + \Lambda_1 a_2, \quad \Lambda_1 \Lambda_2) \tag{2.3}$$

The unit element is (0.1).

The four dimensional vector a can be uniquely characterized also by the hermitean matrix $h(a)$

$$h(a) = \begin{Vmatrix} a_t - a_z & a_x + ia_y \\ a_x - ia_y & a_t + a_z \end{Vmatrix} \tag{2.4}$$

The determinant of h is then just the Lorentz length of the vector a.

Let us consider next the transformation

$$Bh(a)B^\dagger = h(a'). \tag{2.5}$$

The dagger denotes hermitean adjoint. By (2.5), B transforms the vector a linearly into a vector a' and if the determinant of B is 1, the

transformation will leave the Lorentz length of a unchanged. It can easily be seen that the transformation $a \to a'$ defined by (5) is a Lorentz transformation. It can also be shown rather easily that every proper Lorentz transformation can be represented under the form (5) and that conversely, the Lorentz transformation defines B except for its sign: evidently B and $-B$ transform a into the same a'. The group equation (3) reads in terms of h and B

$$(h_1, B_1)(h_2, B_2) = (h_1 + B_1 h_2 B_1{}^\dagger, B_1 B_2). \tag{2.6}$$

It is useful to replace (2.6) by a simpler set of equations, even though less general ones. One writes

$$(h, B) = [h](B) \tag{2.7}$$

hence

$$[h_1][h_2] = [h_1 + h_2] \tag{2.7a}$$

$$(B_1)(B_2) = (B_1 B_2) \tag{2.7b}$$

$$(B)[h] = [BhB^\dagger](B). \tag{2.7c}$$

There is one further equation which will be needed

$$B(\Lambda^T) = B(\Lambda)^\dagger \tag{2.8}$$

which will not be proved here.

3. The Quantum Mechanical Restricted Poincaré Group

It can be shown, and I understand Dr. Emch has a simplified proof for this, that every symmetry transformation can be replaced by a unitary or antiunitary transformation. This means that given an active symmetry transformation, such as (a, Λ), there is a unitary or antiunitary operator $O(a, \Lambda)$ which transforms the quantum mechanical description of any state ψ into the description $\psi' = O(a, \Lambda)\psi$ of the state into which the active symmetry transformation (a, Λ) transforms ψ. This can be deduced from the fact that the transition probability $|(\psi, \phi)|^2$ between any pair of ψ, ϕ states, and the transformed states ψ', ϕ', is the same. As a matter of fact, the unitary (or antiunitary) operator is determined only up to a factor: if ω is a number of absolute value 1, the operators $O(a, \Lambda)$ and $\omega O(a, \Lambda)$ cannot be distinguished. On the other hand, such a "phase factor" is the only

freedom left in O. Throughout these notes, ω will denote a number of modulus 1.

Consider the succession of two symmetry operations s_1 and s_2, to which the quantum mechanical operators $O(s_1)$ and $O(s_2)$ (or $\omega_1 O(s_1)$ and $\omega_2 O(s_2)$) correspond. The operator which corresponds to the succession of s_1 and s_2 (s_2 being carried out first) can differ from $O(s_1)O(s_2)$ only in a numerical factor of modulus 1. Hence

$$O(s_1)O(s_2) = \omega(s_1, s_2)O(s_1 s_2) \tag{3.1}$$

the quantum mechanical operators form, except for a factor of modulus 1, a representation of the symmetry group. This representation is unitary if all the $O(s)$ are unitary; it is called a corepresentation if some of the O are antiunitary. We shall see presently that as far as the restricted Poincaré group is concerned (or any connected group), the representation must be unitary. This follows simply from the fact that every element s which is close enough to the unit element of the group can be written as the square $s = s_1^2$ of an element. It then follows from (1) that $O(s)$ is unitary, no matter whether $O(s_1)$ is unitary or antiunitary. Once it is established that to all elements in the neighborhood of unity a *unitary* operator corresponds, the same will be true of all elements of the group which can be obtained as products of these elements. These are in our case all the elements of the group which we have called "restricted" or "proper".

There is a second theorem which is based on the fact that $O(s)$ can be replaced by $\omega(s)O(s)$. It can be shown that these $\omega(s)$ can be chosen in the case of the restricted Poincaré group in such a way that the $\omega(s_1, s_2)$ in (1) can be replaced by the $\pm$ sign

$$O(a_1, \Lambda_1)O(a_2, \Lambda_2) = \pm O(a_1 + \Lambda_1 a_2, \Lambda_1\Lambda_2). \tag{3.2}$$

Actually, this theorem cannot be proved in the same way as if the underlying group were compact. In this case one can decompose $O(s)$ into a direct sum of finite dimensional matrices and one then can operate with the determinants of these matrices. In our case, unless the O are trivial, (i.e. all $O(s) = 1$), they are infinite dimensional and the determinant is not a meaningful concept. Nevertheless, the choice of the $\omega(s) = \omega(a, \Lambda)$ can be made so that (2) holds.

Because of (2), it is reasonable to define a group which has "twice more" elements than the restricted Poincaré group and to

consider the operators $O(a, \Lambda)$ as a representation of this larger group. It can be shown that this larger group is isomorphic with the group (h, B) which we encountered in the preceding section. Hence, the group of these quantities, with the multiplication law (2.6), will be called the quantum mechanical restricted Poincaré group and we arrived at the theorem that the quantum mechanical symmetry operators which correspond to the elements of the restricted Poincaré group form a true unitary representation of the group of the $(h, B) = [h](B)$, i.e. of the quantum mechanical restricted Poincaré group. It simplifies the notation to introduce particular symbols for the operators which correspond to h and to B: they will be T_h and L_B (the T stands for translation, L for Lorentz transformation). The fact that these generate a true representation of the group of the (h, B) entails the transcription of the equations (2.7):

$$T_{h_1} T_{h_2} = T_{h_1 + h_2}. \tag{3.3a}$$

$$L_{B_1} L_{B_2} = L_{B_1 B_2} \tag{3.3b}$$

$$L_B T_h = T_{BhB^\dagger} L_B \tag{3.3c}$$

4. The Relevant Representations of the Quantum Mechanical Restricted Poincaré Group

This section continues the review of results which are, actually, quite old, but uses a different and simpler language. The reason that the standard (Frobenius–Burnside–Schur) theory of group representations cannot be applied in the present case is that this theory applies only to compact groups whereas the Lorentz or Poincaré groups are not compact but "infinite". In view of the thoughts provoked by the theory of Regge poles, it may be useful to point out that every unitary representation of a compact group, whether finite or infinite dimensional, can be transformed into a direct sum of finite dimensional unitary irreducible representations. This applies in particular also to the rotation group. Hence the infinite dimensional irreducible representations of such a group cannot be unitary and the investigations of M. Froissart show that they are, at least in the case of the rotation group, highly pathological. On the other hand, a locally compact group which has infinite volume (is not globally

compact), such as the Lorentz group, has no unitary faithful representation in finite dimensions. Hence the Lorentz or Poincaré groups, all the factor groups of which are non-compact, have only infinite dimensional unitary representations. The only exception is the trivial representation by the unit matrix: this belongs to the trivial factor group consisting of a single element which is generated by the whole group as its own invariant subgroup.

In order to make full use of the simplification which the characterization (h, B) of the elements of the quantum mechanical restricted Poincaré group by means of a hermitean and a unimodular matrix, h and B, affords, it is useful to replace the variables p_t, p_x, p_y, p_z which were used in earlier work, by a matrix

$$\mathbf{p} = \begin{Vmatrix} p_t + p_z & -p_x - ip_y \\ -p_x + ip_y & p_t - p_z \end{Vmatrix} \tag{4.1}$$

Similar to h, this is a hermitean matrix but the sign conventions are different from (2.4). In effect, $\mathbf{p}$ is the h which corresponds to the contragradient vector Gp (cf. (2.1)):

$$\mathbf{p} = h\,(Gp), \tag{4.2}$$

Hence, the wave functions, or state vectors, will depend on $\mathbf{p}$ plus a discrete variable, rather than on the four vector p and a discrete variable.

The mathematical meaning of p or $\mathbf{p}$ is connected with the operators T_h which correspond to pure displacements. These form an abelian invariant subgroup and can be assumed, therefore, to be "diagonal". Physically, the four vector p is the energy-momentum vector. The rigorous establishment of the formula

$$T_{h(a)}\psi(p) = \exp[i(a_t p_t - a_x p_x - a_y p_y - a_z p_z)]\psi(p) \tag{4.3}$$

requires some careful attention but this will not be given here. Omitting these details one can say that one considers such Hilbert vectors which transform, under the operations of the abelian group of displacements $[h]$, according to an irreducible, and hence one dimensional, representation. We denote the state vectors which transform according to such an irreducible representation by $|p, \zeta\rangle$ and will have, instead of (3)

$$T_{h(a)}|p, \zeta\rangle = \exp[i(a_t p_t - a_x p_x - a_y p_y - a_z p_z)]|p, \zeta\rangle. \tag{4.4}$$

We shall have to remember, however, that the actual states may be not only linear combinations of a discrete set of $|p, \zeta\rangle$ but also integrals over a continuous manifold of p values. Actually, only the latter situation will prevail and the $|p, \zeta\rangle$ themselves will be idealized "members of the continuous spectrum", or "distributions". As a result, the operation with the $|p, \zeta\rangle$ is not rigorous and should be further justified. This is indeed possible, on the basis of the work of Mautner and von Neumann, but will not be carried out. Clearly, the minus signs in the exponents of (4.3) or (4.4) are arbitrary and are chosen for future convenience; they are permissible, however, and mean only that the particular state which transforms according to the representation $\exp[i(a_t p_t - a_x p_x - a_y p_y - a_z p_z)]$ is denoted by $|p\rangle$ and not the one which transforms according to $\exp[i(a_t p_t + a_x p_x + a_y p_y + a_z p_z)]$. At any rate, because of the unitary character of the representation, the exponents must be imaginary so that the p are real. The variable ζ is necessary because there may be several (in fact there may be an infinite number of) state vectors with the same transformation property (4) with respect to displacements. In the most important case in which the system considered is a single elementary particle, the index ζ will assume $2S + 1$ values if the spin is S.

In order to transcribe (4) into our notation, we introduce the symbol

$$\langle h\mathbf{p}\rangle = \tfrac{1}{2}\ \mathrm{Trace}\ h\mathbf{p} \tag{4.5}$$

For $h = h(a)$ this is just the Lorentz product of the vectors a and $\mathbf{p}$

$$\langle h(a)\mathbf{p}\rangle = \tfrac{1}{2}\ \mathrm{Trace}\ [h(a)\mathbf{p}] = p_t a_t - p_x a_x - p_y a_y - p_t a_t \tag{4.5a}$$

so that (4) becomes

$$T_h|\mathbf{p},\zeta\rangle = \exp(i\langle h\mathbf{p}\rangle)|\mathbf{p},\zeta\rangle. \tag{4.6}$$

Let us consider now the vectors $L_B|\mathbf{p}, \zeta\rangle$. This is the Lorentz transform of $|\mathbf{p}, \zeta\rangle$ so that it is reasonable to expect that it is some $|\mathbf{p}', \eta\rangle$ or a linear combination of such vectors with different η. We therefore apply T_h to $L_B|\mathbf{p},\zeta\rangle$ to see how it transforms

$$T_h(L_B|\mathbf{p}, \zeta\rangle) = T_h L_B|\mathbf{p}, \zeta\rangle = L_B T_{B^{-1}hB^{\dagger -1}}|\mathbf{p}, \zeta\rangle$$
$$= L_B \exp(i\langle B^{-1}h\, B^{\dagger -1}\mathbf{p}\rangle)|\mathbf{p}, \zeta\rangle = \exp(i\langle B^{-1}h\, B^{\dagger -1}\mathbf{p}\rangle)L_B|\mathbf{p}, \zeta\rangle. \tag{4.7}$$

The last step is correct because the exponential is a coefficient and L_B is linear. Since the $\langle \ldots \rangle$ symbol indicates a trace, the first factor can be shifted to the end. Hence

$$T_h(L_B|\mathbf{p}, \zeta\rangle) = \exp(i\langle h\, B^{\dagger -1}\mathbf{p}\, B^{-1}\rangle)L_B|\mathbf{p}, \zeta\rangle. \tag{4.8}$$

$L_B|\mathbf{p}, \zeta\rangle$ transforms, under translations T_h, as the vectors $|B^{\dagger -1}\mathbf{p}\, B^{-1}, \zeta\rangle$ and is therefore a linear combination of these vectors:

$$L_B|\mathbf{p}, \zeta\rangle = \sum_\eta A_\eta |B^{\dagger -1}\mathbf{p}B^{-1}, \eta\rangle \tag{4.9}$$

The A_η may depend on B, or $\mathbf{p}$, and on ζ.

As a result of (6) and (9), two of the three group equations, namely (3.3a) and (3.3c) are automatically satisfied. Hence, only (3.3b) remains to determine the A and, of course, similarity transformations are still possible to make the determination unique. We first carry out such a transformation, in the space of the sets $|\mathbf{p}, \zeta\rangle$ with the same $\mathbf{p}$ but different ζ. Such a transformation will leave (6) unchanged but will be used to simplify (9) at least for some $\mathbf{p}$ and B. We consider in particular the $\mathbf{p} = m1$ vectors, where m is a positive real number, which are parallel to positive time axis and the true Lorentz transformations, i.e. symmetric $\Lambda = \Lambda^T$. For these, because of (2.8), $B = B^\dagger$ is hermitean. For such a $\mathbf{p}$ and B, (9) reads

$$L_B|m1, \zeta\rangle = \sum A_\eta |mB^{-2}, \eta\rangle. \tag{4.10}$$

We now denote the sum on the right side by the single term $|mB^{-2}, \zeta\rangle\rangle$. In other words, we set

$$|\mathbf{p}, \zeta\rangle\rangle = L_B|m1, \zeta\rangle \tag{4.11}$$

where m and B are defined by the equation

$$\mathbf{p} = mB^{-2} \tag{4.11a}$$

and m must be a positive number, B a unimodular hermitean matrix. This definition of $|\mathbf{p}, \xi\rangle\rangle$ will be permissible if there is one and only one solution of (11a) with a positive m and a unique unimodular hermitean B. We shall assume that $\mathbf{p}$ is positive definite so that the four vector p is time-like and its time component is positive. It then follows from (11a) that the determinant of $\mathbf{p}$ is m^2 so that m is the positive square root of this determinant. Hence, $m^{-1}\mathbf{p}$ is a unimodular

positive definite hermitean matrix and B^{-1} is its square root. Now every positive definite hermitean matrix has one and only one positive definite square root, and this will be unimodular if the square is unimodular. Hence, there will always be one and only one B^{-1} which satisfied (11a) if we add the further requirement that B^{-1}, and hence B, shall be positive definite.

We now note that the $|mB^{-2}, \zeta\rangle\rangle$ for the same mB^{-2} but different ζ, are orthogonal and have the same norm as the $|m1, \zeta\rangle$. This is true because L_B is unitary and because the $|m1, \zeta\rangle$ should be mutually orthogonal. Hence, replacing the $|mB^{-2}, \zeta\rangle\rangle$ by $|mB^{-2}, \zeta\rangle$ is a unitary transformation in the space of the vectors $|\mathbf{p}', \zeta\rangle$ with $p' = mB^{-2}$ and we can write, finally, for positive definite hermitean B

$$L_B|m1, \zeta\rangle = |mB^{-2}, \zeta\rangle. \tag{4.12}$$

Writing B^{-1} for B and applying L_B to both sides, this gives also

$$L_B|mB^2, \zeta\rangle = |m1, \zeta\rangle \tag{4.12a}$$

This also holds for unimodular positive definite hermitean B.

Let us next consider such B which leave $m1$ unchanged, i.e. for which

$$B^{\dagger -1}m1B^{-1} = m1 \tag{4.13}$$

so that the first arguments on the left and right sides of (9) are the same. It follows from (13) that $1 = B^{\dagger}B$, i.e. that B is unitary. This is natural since the unitary B correspond to spatial rotations and these indeed leave the vector of $m1$, which is parallel to the time axis, invariant. For unitary B, (9) reads with $\mathbf{p} = m1$

$$L_B|m1, \zeta\rangle = \sum_{\eta} \Delta_m(B)_{\eta\zeta}|m1, \eta\rangle. \tag{4.14}$$

We have replaced the A of (13) by a more elaborate symbol which exhibits all the quantities on which the A depends.

Writing now B_2 for B in (14), and applying L_{B_1}, to both sides, we find in the usual way

$$\Delta_m(B_1)\Delta_m(B_2) = \Delta_m(B_1B_2). \tag{4.15}$$

The matrices Δ_m in (14) are a representation of the group of the B for which (14) is valid, that is the unimodular unitary matrices in two dimensions, the SU_2 group. The dimension of Δ_m is equal to the

number of values ζ can assume in $|m1, \zeta\rangle$, i.e. the number of linearly independent quantities which transform under displacements according to the same representation (4) with $p_t = m$; $p_x = p_y = p_z = 0$. However, we know from (11) and (12) that this number is equal to the number of values ζ can assume in any $|\mathbf{p}, \zeta\rangle$ as long as $\mathbf{p}$ can be written as mB^2, i.e. as long as the determinant of $\mathbf{p}$ is m^2 and it is positive definite. This also shows that the number of linearly independent states, which transform under displacements according to (4) or (6) with any positive definite $\mathbf{p}$, with the determinant m^2, is the same.

If we wish, we can now carry out a unitary transformation of the $|\mathbf{p}, \zeta\rangle$, with the same $\mathbf{p}$, which depends only on the determinant of $\mathbf{p}$. Such a transformation will leave all the preceding equations unchanged but will induce a similarity transformation of each Δ_m. One can bring Δ_m by means of such a similarity transformation into the reduced form so that it becomes the direct sum of the well-known irreducible transformations of SU_2. This last step is not essential but completes, at any rate, the calculation. It remains to be shown that the equations (12), and (14), together with (3.3b), permit expressing the A of (9) in terms of the Δ_m and we shall carry out this last step now.

$L_B|\mathbf{p}, \zeta\rangle$ will be calculated by decomposing B into three factors $B = B_1B_2B_3$ so that the corresponding L_B is also decomposed in a similar way by (3.3b)

$$L_B|\mathbf{p}, \zeta\rangle = L_{B_1}L_{B_2}L_{B_3}|\mathbf{p}, \zeta\rangle \tag{4.16}$$

and each expression $L_{B_3}|\mathbf{p}, \zeta\rangle$, $L_{B_2}L_{B_3}|\mathbf{p}, \zeta\rangle$, $L_{B_1}L_{B_2}L_{B_3}|\mathbf{p}, \zeta\rangle$ can be calculated successively by one of the equations (12), (14). As to the first step, it is clear that only (12) can be used and this only if p is time-like and has a positive time component, i.e. if $\mathbf{p}$ is positive definite. In this case we can divide $\mathbf{p}$ by the square root of its determinant

$$m = (\det)\mathbf{p}^{1/2} \tag{4.17}$$

so that $\mathbf{p}/m$ is unimodular positive definite and take B_3 as the positive definite hermitean square root of it

$$B_3 = (\mathbf{p}/m)^{1/2} \tag{4.18}$$

so that $\mathbf{p} = mB_3^2$ and by (12a)

$$L_{B_3}|\mathbf{p}, \zeta\rangle = |m1, \zeta\rangle. \tag{4.19}$$

Note that in general a positive definite hermitean matrix in two dimensions has four square roots but two of these are indefinite and have negative determinants, one is negative definite and is oppositely equal to the last one which is the one meant in (18).

The next factor B_2 will be unitary so that $L_{B_2}L_{B_3}|\mathbf{p}, \zeta\rangle$ can be calculated by (14). It will be a linear combination of vectors $|m1, \eta\rangle$. Because of (9) we know that L_{B_1} will transform this into a linear combination of $|B^{\dagger -1}\mathbf{p}B^{-1}, \eta\rangle$ so that we shall need (12) to calculate the effect of L_{B_1}. We see in fact from (12) that B_1 must satisfy the equation

$$mB_1^{-2} = B^{\dagger -1}\mathbf{p}B^{-1}$$

so that

$$B_1 = (Bm\mathbf{p}^{-1}B^\dagger)^{1/2}. \tag{4.20}$$

Since $B_1B_2B_3 = B$, we obtain for B_2

$$B_2 = B_1^{-1}B\,B_3^{-1} = (B^{\dagger -1}\mathbf{p}B^{-1})^{1/2}B\mathbf{p}^{-1/2}. \tag{4.21}$$

We must verify that this is unitary so that (14) can be used to calculate $L_{B_2}(L_{B_3}|\mathbf{p}, \zeta\rangle) = L_{B_2}|m1, \zeta\rangle$.

We have in fact, since the first and last factors of (21) are hermitean

$$\begin{aligned} B_2{}^\dagger B_2 &= \mathbf{p}^{-1/2}B^\dagger(B^{\dagger -1}\mathbf{p}B^{-1})^{1/2}(B^{\dagger -1}\mathbf{p}B^{-1})^{1/2}B\mathbf{p}^{-1/2} \\ &= \mathbf{p}^{-1/2}B^\dagger B^{\dagger -1}\mathbf{p}B^{-1}B\mathbf{p}^{-1/2} = \mathbf{p}^{-1/2}\mathbf{p}\mathbf{p}^{-1/2}. \end{aligned} \tag{4.22}$$

The last expression can easily be seen to be 1 so that B_2 is indeed unitary; (21) shows that it is also unimodular. Hence we can write, by means of (14)

$$L_{B_2}L_{B_3}|\mathbf{p}, \zeta\rangle = \sum_\eta \Delta_m[(B^{\dagger -1}\mathbf{p}B^{-1})^{1/2}B\mathbf{p}^{-1/2}]_{\eta\zeta}|m1, \eta\rangle \tag{4.23}$$

and finally, by (12) and (20)

$$\begin{aligned} L_B|\mathbf{p}, \zeta\rangle &= L_{B_1}L_{B_2}L_{B_3}|\mathbf{p}, \zeta\rangle \\ &= \sum_\eta \Delta_m[(B^{\dagger -1}\mathbf{p}B^{-1})^{1/2}B\mathbf{p}^{-1/2}]_{\eta\zeta}|B^{\dagger -1}\mathbf{p}B^{-1}, \eta\rangle. \end{aligned} \tag{4.24}$$

This is then the most general unitary representation of the proper quantum mechanical Poincaré group as long as we can restrict ourselves to states with positive time-like momenta, i.e. to states which transform under displacements according to (4) with a time-like p with positive p_t. The L_B defined by (24) automatically satisfies (3.3b)—as was to be expected. It follows that (6), (12), (14), or (6)

and (24) fully determine the representation as long as the Δ_m in (14) are unitary representations of SU_2.

It was mentioned before that the $|\mathbf{p}, \zeta\rangle$ are not bona fide vectors in Hilbert space but members of a continuous spectrum. Their normalization corresponds to

$$\langle \mathbf{p}', \zeta' | \mathbf{p}, \zeta \rangle = \delta^4(\mathbf{p} - \mathbf{p}')\delta_{\zeta\zeta'}. \tag{4.25}$$

The general normalizable Hilbert vectors are integrals over the four components of $\mathbf{p}$ and sums over ζ

$$\sum_{\zeta} \int d^4\mathbf{p}\psi(\mathbf{p}, \zeta)|\mathbf{p}, \zeta\rangle \tag{4.26}$$

the norm of (26) being

$$\sum_{\zeta} \int d^4\mathbf{p}|\psi(\mathbf{p}, \zeta)|^2. \tag{4.26a}$$

It is easy to verify that the representation obtained is continuous as long as the Δ_m is continuous. This is the case for all well-known irreducible representations of SU_2 and, hence, for the sums of these representations.

It should be noted in the second place that the square roots which occur in the final formula (24) can be expressed very simply in terms of a formula discovered in Klein's papers by Gürsey

$$\mathbf{p}^{1/2} = \frac{m1 + \mathbf{p}}{\sqrt{2(m + \langle \mathbf{p} \rangle)}}. \tag{4.27}$$

This is valid for every hermitean two dimensional matrix, in particular also for $B^{\dagger-1}\mathbf{p}B^{-1}$. Similarly,

$$\mathbf{p}^{-1/2} = \frac{m1 + 2\langle \mathbf{p} \rangle 1 - \mathbf{p}}{m\sqrt{2(m + \langle \mathbf{p} \rangle)}}. \tag{4.27a}$$

It should be noted, finally, that all the Δ_m must contain either only integer, or only half-integer, representations of the SU_2 group. Otherwise it is, for instance, not true that the L_{-1} is simply multiplication by a constant ($+1$ or -1) as it must be since in this case the transformed states may differ only by the same constant from the not transformed states.

The preceding calculation is an adaptation of Frobenius' method to determine representations of a group from the representations of an invariant subgroup, in our case the subgroup of displacements.

5. The Irreducible Representations of the Restricted Quantum Mechanical Poincaré Group

Clearly, the representation obtained can be irreducible only if the Δ_m is one of the irreducible representations $D^{(0)}$, $D^{(1/2)}$, $D^{(1)}$, . . . of SU_2. Otherwise, we can assume that Δ_m is in a reduced form and the vectors which belong to any one of the irreducible parts thereof form an invariant manifold. Hence, in an irreducible representation, Δ_m must be one of the irreducible $D^{(s)}$.

It is clear, similarly, that if there are $|\mathbf{p}, \zeta\rangle$ with more than one value of det $\mathbf{p} = m^2$, the representation will be reducible because the $|\mathbf{p}, \zeta\rangle$ with a single value of det $\mathbf{p}$ form an invariant linear manifold. Hence the invariant representations are given by (6), (14), (24) in which, however, Δ_m is replaced by one of the $D^{(s)}$ and m is not an abbreviation for $(\det \mathbf{p})^{1/2}$ but a parameter, characteristic of the representation. This is the situation, at least as long as the $\mathbf{p}$ vector in (4.4) is time-like for any one of the $|\mathbf{p}, \zeta\rangle$. According to the preceding discussion, it is then time-like for all such quantities and the length of $\mathbf{p}$, i.e. the determinant of $\mathbf{p}$, has the same value for all. It is the square of the mass of the particle the transformation property of which is given by the irreducible representation, whereas S is the spin of the particle. Similarly, the sign of p_t, that is of $\langle \mathbf{p} \rangle$, is the same for all p for which $|\mathbf{p}, \zeta\rangle$ exists.

If one adopts the more rigorous point of view sketched at the end of the preceding section, i.e. considers state vectors of the form (4.26), one will have to introduce a singular ψ function, containing a factor $\delta(\det \mathbf{p} - m^2) = \delta(p_t^2 - p_x^2 - p_y^2 - p_z^2 - m^2)$. This will indeed restrict det $\mathbf{p}$ to a single value. As a result, ψ will be effectively a function of only three independent variables and one usually chooses p_x, p_y, p_z for these.

The expression $\psi(\mathbf{p}, \zeta) = \psi(p_x, p_y, p_z' \zeta)\delta(p_t^2 - p_x^2 - p_y^2 - p_z^2 - m^2)$ for ψ renders the normalization integral in (4.26a) infinite.

This has to be redefined, therefore, for irreducible representations by omitting one of the δ functions. It then becomes

$$\sum_{\zeta} \int d^4\mathbf{p}|\psi(\mathbf{p}, \zeta)|^2\delta(p_t^2 - p_x^2 - p_y^2 - p_z^2 - m^2)$$

$$= \sum_{\zeta} \int d^3\vec{p}|\psi(\vec{p}, \zeta)|^2/2p_t \tag{5.1}$$

if the integration over p_t is carried out. The p_t on the right side of (1) is not an independent variable but an abbreviation for the value given by the δ-function, namely

$$p_t = (p_x^2 + p_y^2 + p_z^2 + m^2)^{1/2}. \tag{5.1a}$$

It follows from the preceding consideration that a reducible representation of the quantum mechanical Poincaré group cannot in general be decomposed into irreducible representations, not even into an infinite number of such representations. It is in general an integral over such representations. These connections have been uncovered by Mautner and von Neumann and further exploited particularly by G. Mackey. It also follows that, rigorously speaking, the integral over $d^4\mathbf{p}$ in (4.26a) is not an ordinary integral but a Stieltjes integral the differential of which may have finite jumps at certain values of det $\mathbf{p}$, corresponding to finite weights of some values of this quantity. The irreducible representations are then those in which only one value of det $\mathbf{p}$ has a finite weight, all other values and intervals of values zero weight.

These remarks conclude our discussion of the representations of the quantum mechanical Poincaré group. The last remarks are not specific for this group but their essentials apply to all representations of non-compact (infinite) groups. Thus, even the representations of the displacement group in one dimension have the form

$$T_a\psi(p, \zeta) = \exp(ipa)\psi(p, \zeta) \tag{5.2}$$

with the norm of ψ defined as

$$\sum_{\zeta} \int |\psi(p, \zeta)|^2 d\mu(p) \tag{5.2a}$$

where the $\mu(p)$ in the integral is a nondecreasing function of p which can, however, have finite jumps. The values of p where $\mu(p)$ has a

finite jump have finite weights; in the regions in which $\mu(p)$ is differentiable and has a finite derivative, only intervals of p have finite weights. For an irreducible representation $\mu(p)$ is constant everywhere except for a single jump at a definite value of p.

6. The Extended Poincaré Group

The extended or "full" Poincaré group consists of combinations of displacements and Lorentz transformations (a, L), just like the restricted Poincaré group, with the difference, however, that L can be *any* real transformation which satisfies the equation

$$LGL^T = G. \tag{6.1}$$

Since the determinant of L^T is equal to that of L, both determinants must be ± 1. Because of the tt component of (6.1)

$$L_{tt}^2 - L_{tx}^2 - L_{ty}^2 - L_{tz}^2 = 1 \tag{6.2}$$

L_{tt} is either larger than 1, or smaller than -1. If $L_{tt} \leq -1$, the transformation L can be multiplied by $-G = \Theta$ and this element will be larger or equal to $+1$. It will still satisfy (6.1). If the determinant of the transformation obtained in this way is -1, it can be multiplied by $G = \Sigma$ and an element of the restricted Lorentz group will result. Assuming that the transformations of the restricted Lorentz group form a connected manifold, it follows conversely, that each transformation of the extended Lorentz group can be written in one of the four forms

$$\Lambda, \Lambda\ \Sigma, \Lambda\Theta, \Lambda I \tag{6.3}$$

where Λ is an element of the restricted group, $\Sigma = G$ corresponds to space inversion, $\Theta = -G$ to time inversion and $I = \Sigma\Theta = -1$ to space and time inversion. The restricted group, i.e. the transformations Λ form an invariant subgroup of the extended group and the other symbols in (6.3) stand for the three cosets thereof. These form the factor group of the restricted Lorentz group; this factor group is isomorphic to the four group.

We shall assume that the laws of physics are invariant under the extended Poincaré group. Hence, we shall look for sets of unitary and antiunitary operators which satisfy

$$O_{(a_1\ L_2)}O_{(a_2\ L_2)} = \omega(a_1, L_1, a_2, L_2)O_{a_1+L_1a_2,\ L_1L_2} \tag{6.4}$$

because the group equations (2.3) remain valid even if the Λ are replaced by transformations of the extended Lorentz group. We shall assume that $O_{(a,L)}$ is unitary if L is a member of the restricted group, or of the coset $\Lambda\Sigma$. If it is a member of one of the cosets $\Lambda\Theta$ or ΛI, it will be assumed to be antiunitary. However, this requirement will not be motivated in detail.

The physical interpretation of the operators Σ, Θ, I has undergone considerable revision in the last five years. The old interpretation of Σ as a simple space inversion operator which replaces the position x, y, z of a particle by $-x$, $-y$, $-z$, and affects its velocity in a similar way, cannot be maintained if Σ is to represent an accurate symmetry operation. Nevertheless, Σ interpreted in this way is an approximate symmetry operation and the old interpretation, therefore, remains useful. This is a point which is particularly emphasized by Wightman. However, if one wishes Σ to represent, as far as we know, a rigorously valid symmetry operation, the geometrical transformation of the old interpretation must be coupled by a replacement of particle by antiparticle. In other words, Σ must mean what ΣC used to mean. This is the interpretation which is usually in my own mind if I speak about Σ. If this is adopted, Σ represents an accurate symmetry operation but it may be useful, for many applications, to introduce another symmetry operation which corresponds to the old interpretation and which has only approximate validity. The original interpretation of the time inversion operation, aptly characterized by Lüders as inversion of the direction of motion, will be maintained; all present experiments confirm this as a valid symmetry operation. The operation I is a succession of the two operations Σ and Θ; their order as physical operations is immaterial.

With the new interpretation of Σ, which incidentally has been proposed independently by scores of people, the symmetry laws which are believed to be valid rigorously are in a one-to-one correspondence with the extended Poincaré group. This may be significant.

In order to determine the operators which satisfy the equations (6.4), we shall proceed in the same way as we proceeded when determining the operators satisfying the similar equations with L replaced by members of the restricted Lorentz group. The first step is to normalize the operators $O_{(a,L)}$ so that the phase factors are eliminated.

7. The Full Quantum Mechanical Poincaré Groups

We will now carry out the normalization of the operators $O_{a,\,L}$, considered in the preceding section. We will continue to assume that all these operators are, apart from a phase factor, completely defined, i.e. we assume that no superselection rules are operative in the Hilbert space in which these operators act.

Let us consider first those operators $O_{a,\,L} = O_{a,\,\Lambda}$ in which L is a member of the restricted Lorentz group. The phases of these operators can be adjusted in the same way as if they were the only operators present. They will be then in a one-to-one correspondence with the elements (h,B) of the restricted quantum mechanical Poincaré group and will form representations of this group. Let us denote them by $U_{h,\,B}$; this will be a reducible or irreducible representation of the group of the (h,B).

The unitary operator which corresponds to Σ will be denoted by U_Σ and its phase so normalized that

$$U_\Sigma{}^2 = 1. \tag{7.1}$$

This is possible because the square of U_Σ corresponds to the physical operation which leaves all states unchanged so that the corresponding operator must be a multiple of the unit operator. The constant factor of the unit operator can be eliminated by a proper choice of the phase factor of U_Σ. Thus (1) is the result of an arbitrary choice which appears useful.

For the same reason, the squares of the operators A_Θ and A_I, which correspond to time and to space-time inversions, must also be multiples of the unit operator. However, the introduction of a phase factor ω does not change the square of these operators because they are antiunitary; $(\omega A)^2 = \omega A \omega A = \omega\omega^* A^2 = A^2$. On the other hand, just because of the anti-unitary nature of A_Θ and A_I, their squares are necessarily either $+1$ or -1. There is no choice involved in this case. In fact, it follows from the associative law that $A_\Theta{}^2 A_\Theta = A_\Theta A_\Theta{}^2$ and since we have seen that $A_\Theta{}^2 = \epsilon_\Theta 1$, we have

$$\epsilon_\Theta 1 A_\Theta = \epsilon_\Theta A_\Theta = A_\Theta \epsilon_\Theta 1 = \epsilon_\Theta{}^* A_\Theta. \tag{7.2}$$

Hence ϵ_Θ is real and since it is of absolute value 1, we have

$$A_\Theta = \epsilon_\Theta 1 \qquad \epsilon_\Theta = \pm 1 \tag{7.3}$$

and similarly

$$A_I^2 = \epsilon_I 1 \qquad \epsilon_I = \pm 1. \tag{7.4}$$

The values of ϵ_Θ and ϵ_I cannot be changed by changing phases and their values are characteristic of particular solutions of the equations (6.4). In other words, for some solutions of these equations $A_\Theta^2 = 1$, $A_I^2 = -1$, for others the signs are different. There are four types of solutions of these equations, corresponding to the four possible choices of signs in (3) and (4). However, both A_Θ and A_I can yet be multiplied with a phase factor. Since it follows from (6.4) that $A_I = \omega U_\Sigma A_\Theta$ we choose the phase of A_I in such a way that

$$A_I = U_\Sigma A_\Theta. \tag{7.5}$$

This is again a convention. It then follows from (5) that

$$A_I^2 = U_\Sigma A_\Theta U_\Sigma A_\Theta = \epsilon_I 1.$$

Multiplication of this equation by U_Σ on the left and A_Θ on the right gives because of (1) and (3)

$$A_\Theta U_\Sigma \epsilon_\Theta = \epsilon_I U_\Sigma A_\Theta$$

or

$$A_\Theta U_\Sigma = \epsilon_\Theta \epsilon_I A_I. \tag{7.6}$$

This completes the normalization of the operators U_Σ, A_Θ and A_I. The phase factor of A_Θ, in particular, has not been specified but this will not prove to be necessary. The preceding results are summarized in Table 1.

TABLE 1

Multiplication Table of the Reflection Operators

	U_Σ	A_Θ	A_I
U_Σ	1	A_I	A_Θ
A_Θ	$\epsilon_\Theta \epsilon_I A_I$	$\epsilon_\Theta 1$	$\epsilon_I U_\Sigma$
A_I	$\epsilon_\Theta \epsilon_I A_\Theta$	$\epsilon_\Theta U_\Sigma$	$\epsilon_I 1$

Evidently, the operators U_Σ, A_Θ, A_I generate a group in every case. The nature of this group depends, however, on the values of ϵ_Θ and

ϵ_I; it is completely determined by the preceding multiplication table. The next table gives a description of these groups.

TABLE 2

The Groups Formed by the Reflection Operators

Case ϵ_Θ	ϵ_I	Elements of order 2	Elements of order 4	Name of group
1	1	U_Σ, A_Θ, A_I	None	Four group
1	-1	$-1, U_\Sigma, -U_\Sigma, A_\Theta, -A_\Theta$	$A_I, -A_I$	Dihedral group
-1	1	$-1, U_\Sigma, -U_\Sigma, A_I, -A_I$	$A_\Theta, -A_\Theta$	Dihedral group
-1	-1	$-1, U_\Sigma, -U_\Sigma$	$A_\Theta, -A_\Theta, A_I, -A_I$	$C_2 \otimes C_4$

Naturally, each group has a unit element of order 1.

We shall specify, next, the phases of the remaining operators $O_{a,L}$. We know from (6.3) that L has one of the forms $\Lambda, \Lambda\Sigma, \Lambda I$. We have already replaced $O_{a,L}$ by the two operators $U_{h,B}$ and $U_{h,-B}$ if $L = \Lambda$ is a member of the restricted group. These latter operators are completely determined and h and B are given in terms of a and Λ by (2.4) and (2.5) with $a' = \Lambda a$. Similarly, we replace $O_{a,\Lambda\Sigma}$ by $U_{h,B}U_\Sigma$ and $U_{h,-B}U_\Sigma$. Since $U_{h,B}$ is either equal to $U_{h,-B}$ or is oppositely equal thereto, $U_{h,B}U_\Sigma$ and $U_{h,-B}U_\Sigma$ are in a similar relation to each other. The same holds of the replacements of $O_{a,\Lambda\Theta}$ by $U_{h,B}A_\Theta$ and $U_{h,-B}A_\Theta$ and of the replacement of $O_{a,\Lambda I}$ by $U_{h,B}A_I$ and $U_{h,-B}A_I$. In other words, the operators $O_{a,L}$ are specified only up to a factor ± 1 if we deal with a representation in which $U_{0,-1} = L_{-1} = -1$, they are completely specified if we deal with a representation in which $U_{0,-1} = L_{-1} = 1$. As was pointed out after (4.27a), one of these alternatives must hold after our normalization in every meaningful representation of the restricted Poincaré group. It will be shown that the operators obtained in this way satisfy a simplified form of the equations (6.4) which imply

$$O_{a_1 L_1} O_{a_2 L_2} = \pm O_{a_1 + L_1 a_2, L_1 L_2}. \tag{7.7}$$

Even the ambiguity of sign can be eliminated if we coordinate the operators not to the elements (a, L) of the full Poincaré group but to

elements of one of the full quantum mechanical Poincaré groups. The fundamentally new element here is that the symmetry operators do not necessarily form a representation or corepresentation of a definite group but may form a corepresentation on one of four distinct groups. Each of these groups is the semidirect product of the restricted quantum mechanical Poincaré group, with elements (h, B), and of one of the groups of Table 2, with the multiplication laws given in Table 1. Hence, the transformation operators can be written in a normal form $U_{h,B}O_r$ where $U_{h,B} = T_h L_B$ form one of the representations of the restricted group which were determined before and the O_r are the elements of one of the groups of Table 2. The multiplication laws of the $U_{h,B}$ are given by (3.3) and can be united into

$$U_{h_1\, B_1} U_{h_2\, B_2} = U_{h_1 + B_1 h_2 B_1^\dagger\, B_1 B_2}. \tag{7.8}$$

The multiplication laws of the O_r (their number is four or eight depending on the group which is being represented) are given by Table 1. To these come

$$U_\Sigma U_{h,B} U_\Sigma^{-1} = U_\Sigma^{-1} U_{h,B} U_\Sigma = U_{\mathbf{h},\mathbf{B}} \tag{7.9a}$$

$$A_\Theta U_{h,B} A_\Theta^{-1} = A_\Theta^{-1} U_{h,B} A_\Theta = U_{-\mathbf{h},\mathbf{B}} \tag{7.9b}$$

$$A_I U_{h,B} A_I^{-1} = A_I^{-1} U_{h,B} A_I = U_{-h,B} \tag{7.9c}$$

where $\mathbf{h}$ is the matrix obtained from h by using the contragradient rather than the cogradient components, as indicated in (4.2). One can also write

$$\mathbf{h}_\pm = \sigma h^* \sigma \tag{7.10}$$

where the star denotes transition to the conjugate complex and

$$\sigma = \begin{Vmatrix} 0 & i \\ -i & 0 \end{Vmatrix} \tag{7.10a}$$

Similarly

$$\mathbf{B} = \sigma B^* \sigma. \tag{7.10b}$$

The relations (7.9) will be proved below. Actually, because of $U_\Sigma^{-1} = U_\Sigma$, $A_I^{-1} = \epsilon_I A_I$ and $A_\Theta = U_\Sigma A_I$, it will suffice to prove only (7.9a) and (7.9c).

The operators $O_{a,\,L}$ which correspond to the full Poincaré group are

$$O_{a,\Lambda} = U_{h,\,\pm B} \tag{7.11}$$

$$O_{a,\Lambda\Sigma} = U_{h,\,\pm B}U_{\Sigma} \tag{7.11a}$$

$$O_{a,\Lambda\Theta} = U_{h,\,\pm B}A_{\Theta} \text{ and } U_{h,\,\pm B}\,\epsilon_{\Theta}A_{\Theta} \tag{7.11b}$$

$$O_{a,\Lambda I} = U_{h,\,\pm B}A_{I} \text{ and } U_{h,\,\pm B}\,\epsilon_{I}A_{I} \tag{7.11c}$$

$h = h(a)$ is given by (2.4), B by (2.5) with $a' = \Lambda a$. The coordination of operators to the elements (a, L) of the Poincaré group is unique only if $U_{0,\,-1} = L_{-1} = 1$ and if the $\epsilon_{\Theta} = \epsilon_{I} = 1$ group is chosen. In all other cases, two operators correspond at least to some physical operations; if $U_{0,\,-1} = L - 1 = -1$, two operators correspond to each.

The relations (8), (9), together with the multiplication laws of Table 1 suffice to put every product of transformation operators into the normal form $U_{h,B}O_{r}$. Thus, for instance, if we deal with the $\epsilon_{\Theta} = 1$, $\epsilon_{I} = -1$ group, the product

$$\begin{aligned} U_{h_1B_1}(-A_I)\,.\,U_{h_2B_2}\,.\,A_{\Theta} &= U_{h_1B_1}\,A_I^{-1}U_{h_2B_2}A_I(-A_I)A_{\Theta} \\ &= U_{h_1B_1}U_{-h_2B_2}(-A_IA_{\Theta}) = U_{h_1-B_1h_2B_1^{\dagger},\;B_1B_2}(-U_{\Sigma}). \end{aligned}$$

The minus sign could also be pulled to the front—it could have been pulled to the front right at the start.

Equations (9) remain to be established. This will be done in two steps: It will be proved first that they are valid within a phase factor and second this phase factor will be shown to have the value 1. We proceed with the first step which essentially amounts to a calculation of the h and of the B which correspond to Σa and $\Sigma\Lambda\Sigma$, respectively, and to the other transforms of a and Λ.

It follows from the group equations that

$$U_{\Sigma}O_{a,\,\Lambda}U_{\Sigma}^{-1} = \omega(a, \Lambda)\,O_{\Sigma a,\,\Sigma\Lambda\Sigma} \tag{7.12}$$

Σa is the vector the spatial components of which are oppositely equal to those of a whereas its time component is a_t. It then follows from (4.2) and (10) (since $\Sigma = G$) that

$$h\Big(\Sigma\; a\Big) = \mathbf{h}(a) = \sigma h(a)^{*}\sigma. \tag{7.13}$$

This completes the calculation of the h which corresponds to the space-inverted a. Let us now calculate the two dimensional unimodular $\hat{B}$ which corresponds by (2.5) to $\Sigma\Lambda\Sigma$, in terms of the B which corresponds to Λ. The former is defined by

$$\hat{B}h(a)\hat{B}^\dagger = h\Big(\Sigma\,\Lambda\,\Sigma\,a\Big) \tag{7.14}$$

which is valid for every a, whereas B is defined by

$$B\,h(b)B^\dagger = h(\Lambda b) \tag{7.14a}$$

valid for all b. Replace Σa by b in (14) and use (13) to obtain

$$\hat{B}h\Big(\Sigma\,b\Big)\hat{B}^\dagger = h\Big(\Sigma\,\Lambda b\Big) = \sigma h(\Lambda b)^*\sigma. \tag{7.14b}$$

Substitution of $h(\Lambda b)$ from (14a) and the use of (13) for $h(\Sigma b)$ gives finally

$$\hat{B}\sigma h(b)^*\sigma\hat{B}^\dagger = \sigma B^* h(b)^* B^{*\dagger}\sigma \tag{7.14c}$$

or

$$\hat{B}\sigma h(b)^*(\hat{B}\sigma)^\dagger = \sigma B^* h(b)^*(\sigma B^*)^\dagger. \tag{7.15}$$

Since this holds for every b, the $h(b)^*$ in (15) is an arbitrary hermitean matrix and $\hat{B}\sigma$ can differ from σB^* only by a factor of modulus 1. Since, however, the determinants are all 1, the factor can be only ± 1. It follows that

$$\hat{B} = \pm\,\sigma B^*\sigma. \tag{7.16}$$

Actually, both (13) and (16) could have been obtained by direct computation.

If we replace $O_{a,\Lambda}$ in (12) by the corresponding $U_{h,B}$, this equation becomes by (13) and (16)

$$U_\Sigma U_{h,B} U_\Sigma^{-1} = \omega(h, B) U_{\sigma h^*,\,\sigma B^*\sigma}. \tag{7.17}$$

The $\pm$ sign of (16) is not necessary for (17) because $U_{\sigma h\sigma^*,\,-\sigma B^*\sigma}$ can differ from $U_{\sigma h^*\sigma,\,\sigma B^*\sigma}$ only in sign and this can be absorbed into the $\omega(h, B)$. Since the operators on both sides of (17) are completely specified, i.e. their phase factors are determined, $\omega(h, B)$ is an entirely definite function of h and B.

The second equation which we shall need

$$A_I U_{h,B} A_I^{-1} = \omega'(h, B) U_{-h,B} \tag{7.18}$$

is even easier to establish on the basis of

$$A_I O_{a\ \Lambda} A_I^{-1} = \omega'(a, \Lambda) O_{Ia,\, I\Lambda I} \tag{7.19}$$

if we recall that I is the matrix -1. Hence $Ia = -a$, $I\Lambda I = \Lambda$ and $h(-a) = -h(a)$. This, together with $A_\Theta = U_\Sigma A_I$, proves also (9b) up to a factor.

We now come to the second step of the demonstration of (9) which shows that the ω in (17) and (18) must be equal to 1. Since (17) is valid for every h, B, let us write down this equation for two other matrices j, C

$$U_\Sigma U_{j,\,C} U_\Sigma^{-1} = \omega(j, C) U_{\sigma j^* \sigma,\, \sigma C^* \sigma}. \tag{7.17a}$$

The product of (17) and (17a) is

$$U_\Sigma U_{h,\,B} U_{j,\,C} U_\Sigma^{-1} = \omega(h, B)\omega(j, C) U_{\sigma h^* \sigma,\ B^* \sigma} U_{\sigma j^* \sigma, \sigma C^* \sigma}. \tag{7.20}$$

The product of the last two U can be replaced by a single U by means of the group equation (8). The left side then becomes

$$U_\Sigma U_{h+BjB^\dagger, BC} U_\Sigma^{-1} = \omega(h + BjB^\dagger, BC) U_{\sigma(h+BjB^\dagger)^* \sigma,\ \sigma(BC)^* \sigma}.$$

The right side follows from (17) which is valid no matter what the two indices of U are. The operator on the right side is now equal to the product of the two operators on the right side of (20) so that these cancel. Hence

$$\omega(h + BjB^\dagger, BC) = \omega(h, B)\omega(j, C) \tag{7.21}$$

and comparison with (8) shows that the ω form a one dimensional representation of the restricted quantum mechanical Poincaré group. Since, however, this group has no invariant subgroup with an abelian factor group (except the whole group), the ω must be equal to 1. The ω' of (19) can be shown to be 1 in the same fashion. This then establishes all equations (9) since, as was mentioned before, all of these follow from any two of them by means of the multiplication laws of the reflection operators.

The multiplication laws of the invariance operators, (8), (9) with one of the alternatives of Table 1, define the four possible full quantum mechanical Poincaré groups. However, again, not all representations of these groups are physically admissible. In addition to $U_{0,\,-1} = L_{-1}$ being either 1 or -1, it is necessary that the representative of the element -1 of the reflection group be the

operator -1. It is clear from this circumstance, and the appearance of four distinct possibilities, that to a certain extent, the invariance operators had to be forced into groups. This was true, though to a lesser extent, also for the case of the restricted group. It seems nevertheless, that it is useful to normalize the invariance operators in some way, and that it is simplest to normalize them in such a way that they form a group. It is difficult to carry out calculations with "ray operators", i.e. as operators with entirely indeterminate phase factors, even though, in principle, this is surely possible since an arbitrary normalization of the phase factors surely cannot add anything to the physical content of the theory.

We now proceed to the determination of the representations, or rather corepresentations, of the four groups. The next section adjoins U_Σ to the restricted group; in this case no distinction between the four groups is yet necessary.

8. Inclusion of Space Reflection Symmetry

The present section will determine the representations of the restricted quantum mechanical Poincaré group, when augmented by space reflection symmetry. The group equations are (7.8), (7.1) and (7.9a).

Let us consider, first, the operations of the restricted group and form the vectors $|\mathbf{p}, \zeta\rangle$ for which (4.6) and (4.24) apply. It will be shown that the same equations apply also for $U_\Sigma|\sigma\mathbf{p}^*\sigma,\zeta\rangle$. This is what has to be expected since U_Σ will invert the signs of the spatial components of p but will leave the angular momenta unchanged. In fact applying T_h to $U_\Sigma|\sigma\mathbf{p}^*\sigma, \zeta\rangle$ one obtains, because of (7.9a)

$$\begin{aligned} T_h U_\Sigma|\sigma\mathbf{p}^*\sigma, \zeta\rangle &= U_\Sigma T_{\sigma h^*\sigma}|\sigma\mathbf{p}^*\sigma, \zeta\rangle \\ &= U_\Sigma \exp(i\langle\sigma h^*\sigma\sigma\mathbf{p}^*\sigma\rangle)|\sigma\mathbf{p}^*\sigma, \zeta\rangle. \end{aligned}$$

Since $\sigma^2 = 1$ and since the first factor in the bracket can be shifted to the last position, the exponent can be replaced by $i\langle h^*\mathbf{p}^*\sigma^2\rangle = i\langle h^*\mathbf{p}^*\rangle$. Since, furthermore, $\langle h\mathbf{p}\rangle$ is a real quantity, the exponent is also equal to $i\langle h\mathbf{p}\rangle$. Hence indeed

$$T_h(U_\Sigma|\sigma\mathbf{p}^*\sigma, \zeta\rangle) = \exp(i\langle h\mathbf{p}\rangle)(U_\Sigma|\sigma\mathbf{p}^*\sigma, \zeta\rangle). \tag{8.1}$$

Let us now apply L_B to $U_\Sigma|\sigma\mathbf{p}^*\sigma, \xi\rangle$. Again (7.9a) gives, together with (4.24)

$$\begin{aligned} L_B U_\Sigma|\sigma\mathbf{p}^*\sigma, \zeta\rangle &= U_\Sigma L_{\sigma B^*\sigma}|\sigma\mathbf{p}^*\sigma, \zeta\rangle \\ &= U_\Sigma \sum_\eta \Delta_m(u)_{\eta\zeta}|(\sigma B^*\sigma)^{\dagger-1}\sigma\mathbf{p}^*\sigma(\sigma B^*\sigma)^{-1}, \eta\rangle \end{aligned} \tag{8.2}$$

where the unitary argument u of Δ_m is

$$u = [(\sigma B^*\sigma)^{\dagger-1}\sigma\mathbf{p}^*\sigma(\sigma B^*\sigma)^{-1}]^{1/2}\sigma B^*\sigma(\sigma\mathbf{p}^*\sigma)^{-1/2}. \tag{8.2a}$$

Let us verify, first that the unitary matrix in the ket of (8.2) is indeed equal to $\sigma(B^{\dagger-1}\mathbf{p}B^{-1})^*\,\sigma$. This follows from $\sigma^\dagger = \sigma^{-1} = \sigma$. For the same reason u of (2a) is equal to (^{T} denotes the transpose)

$$u = (\sigma B^{T-1}\mathbf{p}^*B^{*-1}\sigma^{-1})^{1/2}\sigma B^*\sigma(\sigma\mathbf{p}^*\sigma^{-1})^{-1/2}. \tag{8.2b}$$

Since σ is unitary, both roots contain the unitary transform of a quantity. However, $(\sigma q\sigma^{-1})^n = \sigma q^n\sigma^{-1}$ for every n; hence

$$u = \sigma[(B^{T-1}\mathbf{p}^*B^{*-1})^{1/2}B^*\mathbf{p}^{*-1/2}]\sigma^{-1}. \tag{8.2c}$$

As both u and σ are unitary and unimodular, this holds also for $\sigma^{-1}u\sigma$, i.e. the expression in the bracket. One can easily verify this also directly. Further, the transform of a unimodular unitary matrix

$$\left\| \begin{matrix} \alpha & \beta \\ -\beta^* & \alpha^* \end{matrix} \right\|, \quad |\alpha|^2 + |\beta|^2 = 1$$

with σ is the conjugate complex of this matrix. This is one of the identities given also by Gürsey. Hence

$$u = (B^{\dagger-1}\mathbf{p}B^{-1})^{1/2}B\mathbf{p}^{-1/2} \tag{8.3}$$

and (2) becomes

$$\begin{aligned} &L_B(U_\Sigma|\sigma\mathbf{p}^*\sigma, \zeta\rangle) \\ &\quad = \sum_\eta \Delta_m[(B^{\dagger-1}\mathbf{p}B^{-1})^{1/2}B\mathbf{p}^{-1/2}]_{\eta\zeta}(U_\Sigma|\sigma(B^{\dagger-1}\mathbf{p}B^{-1})^*\sigma, \eta\rangle. \end{aligned} \tag{8.4}$$

Comparison of this with (4.24), and of (1) with (4.6), show that $U_\Sigma|\sigma\mathbf{p}^*\sigma_1\zeta\rangle$ transforms under the $U_{h,B}$ just as $|\mathbf{p}, \zeta\rangle$. If neither of the two quantities

$$|\mathbf{p}, \zeta\rangle + U_\Sigma|\sigma\mathbf{p}^*\sigma, \zeta\rangle = |\mathbf{p}, \zeta\rangle_1 \tag{8.5a}$$

and

$$|\mathbf{p}, \zeta\rangle - U_\Sigma|\sigma\mathbf{p}^*\sigma, \zeta\rangle = |\mathbf{p}, \zeta\rangle_2 \tag{8.5b}$$

is zero, they still obey the equations (4.6) and (4.24). Further, because of $U_\Sigma^2 = 1$, they also satisfy the equations

$$U_\Sigma|\mathbf{p}, \zeta\rangle_1 = U_\Sigma|\mathbf{p}, \zeta\rangle + |\sigma\mathbf{p}^*\sigma, \zeta\rangle = |\sigma\mathbf{p}^*\sigma, \zeta\rangle_1 \tag{8.6a}$$

and

$$U_\Sigma|\mathbf{p}, \zeta\rangle_2 = U_\Sigma|\mathbf{p}, \zeta\rangle - |\sigma\mathbf{p}^*\sigma, \zeta\rangle = -|\sigma\mathbf{p}^*\sigma, \zeta\rangle_2. \tag{8.6b}$$

Hence, if the expression in (5b) is zero, the two terms at the left of (5a) are equal and (6b) reduces to

$$U_\Sigma|\mathbf{p}, \zeta\rangle = |p, \zeta\rangle = |\sigma\mathbf{p}^*\sigma, \zeta\rangle, \tag{8.6}$$

i.e. the equation which we wanted to prove. Similarly, if the expression (5a) is zero, the two terms at the left of (5b) become equal and (6) holds for $|\mathbf{p}, \zeta\rangle_2$ if U_Σ is replaced by $-U_\Sigma$. This will not affect any of the normalizations which have been stipulated, such as $U_\Sigma^2 = 1$.

Finally, if neither of the expressions (5) vanishes, one can use these as the basic states instead of the $|\mathbf{p}, \zeta\rangle$ and the $U_\Sigma|\mathbf{p}, \zeta\rangle$. The representation is then reducible because both $U_{h,B}$ and U_Σ transform the states $|\mathbf{p}, \zeta\rangle_1$ into linear combinations of these same states. The same applies for $|\mathbf{p}, \zeta\rangle_2$. In this case (6) is not valid but the manifold of states has both states of even parity (the $|\mathbf{p}, \zeta\rangle_1$) and states of odd parity (the $|\mathbf{p}, \zeta\rangle_2$). The question whether, in view of possible superselection rules, this statement is physically meaningful, will be discussed later. At any rate, even if all the operators are physically meaningful, replacement of U_Σ by $-U_\Sigma$—which is permissible—will change the states of even parity into states of odd parity, and vice versa: only a relative parity can be defined.

The preceding discussion also shows that, as far as irreducible linear manifolds are concerned, the adjoining of U_Σ to the transformations of the restricted Poincaré group does not change the dimension of these: these are still $2S + 1$ states of a given momentum p. This will not be true always when the antiunitary operators A_Θ and A_I are adjoined. The situation reminds one of that in atomic spectroscopy: the space reflection operator does not increase the degeneracy although

it does introduce a new quantum number, the parity. In some cases the time inversion operator increases the degeneracy, in others this is not the case.

In the next section, where A_I and hence $A_\Theta = U_\Sigma A_I$ will be adjoined to the group, we shall deal with irreducible linear manifolds. The preceding consideration then shows that one can select from any such manifold states for which (4.6), (4.24), and also (6) are satisfied. The manifold may contain also states for which (6) is satisfied with a minus sign but we shall, at any rate, start our calculation with the states for which (6) is valid. In some cases, that is for some of the four full groups, all states of an irreducible manifold will satisfy (6). For the rest of the groups the manifold of the states will contain such $|\mathbf{p}, \zeta\rangle$ for which (6) is valid, and also such $|\mathbf{p}, \zeta\rangle$ for which (6) is valid with the negative sign. These two kinds of states will then be transformed into each other by the operations A_I or A_Θ.

The amount of computation that was necessary to establish (4), i.e. the fact that the $U_\Sigma|\sigma\mathbf{p}^*\sigma, \zeta\rangle = U_\Sigma|p_1\, \zeta\rangle$ transform under the operations of the restricted group just as the $|\mathbf{p}, \zeta\rangle$ do, is surprising. I do not know how this calculation could be simplified. Its result was, however, foreseeable.

9. Adjoining the Time Reflection Operators

As was mentioned before, the discussion of the present section, which will be the last one of largely mathematical nature, will be principally concerned with irreducible linear manifolds which are invariant under all operations of the full Poincaré group. The question whether any linear set of states which is invariant under the operations of the full Poincaré group can be obtained by a composition of such irreducible sets will be discussed at the end of this section. It will turn out that certain, though rather simple, modifications of the equations will be necessary for such a composition. The term composition is to be understood here in the extended sense discussed at the end of Section 4 and in Section 5 so that it may consist of the formation of direct integrals, rather than of only direct sums, of Hilbert spaces. It should be recalled, finally, that the present discussion is wholly restricted to such cases in which the standard normalization of the displacement operators, yielding (4.6), yields only

states with positive time-like momentum vectors, i.e. to positive definite $\mathbf{p}$.

Under the conditions enumerated, it is possible to select a set of states $|\mathbf{p}, \zeta\rangle$ such that the normalized $O_{a,\Lambda}$ and O_Σ, that is the $U_{h,B} = T_h L_B$ and U_Σ, have the effect given by (4.6), (4.24) and (8.6). In fact, it can be assumed that the manifold $|\mathbf{p}, \zeta\rangle$ contains only states with a definite mass, $\det \mathbf{p} = m^2$, and also that the Δ_m in (4.24) is one of the irreducible representations $D^{(S)}$ of US_2. If the states $A_I|\mathbf{p}, \zeta\rangle$ are adjoined to the states $|\mathbf{p}, \zeta\rangle$ the resulting linear manifold will be invariant under all transformations $O_{a,L}$ of the full Poincaré group. In order to show this, we shall call the linear combinations of the vectors $|\mathbf{p}, \zeta\rangle$ the first manifold, the linear combinations of the vectors $A_I|\mathbf{p}, \zeta\rangle$ the second manifold. Further, we shall call unitary subgroup the group of operators obtained from those of the restricted group by adjoining the space reflection operator U_Σ. That is, the unitary subgroup is the group the representations of which were determined in the preceding section.

The invariance of the first manifold, that is of the $|\mathbf{p}, \xi\rangle$ under the unitary group is exhibited explicitly in the equations quoted, i.e. (4.6), (4.24), (8.6). The invariance of the second manifold follows from (7.9c)

$$U_{h,B} A_I = A_I U_{-h,B}. \tag{9.1}$$

Because of (1), if one wants to apply $U_{h,B}$ to $A_I|\mathbf{p}, \zeta\rangle$ one may first apply $U_{-h,B}$ to this vector, and then A_I to the resulting vector. The first operation will give a vector in the first manifold, the A_I then transforms this into a vector of the second manifold. Hence, as was to be expected, both sets are invariant under the unitary subgroup.

The invariance of the union of the two manifolds under the remaining operations of the full Poincaré group then follows from (1) and the fact that all the remaining operations can be obtained from the unitary operations by multiplication with O_I. Hence it suffices to show that the union of the two manifolds is invariant also under the operators $U_{h,B} A_I$. Because of $A_I^2 = \pm 1$, it is clear that the vectors of the second manifold, i.e. the linear combinations of the $A_I|\mathbf{p}, \zeta\rangle$, are transformed into similar linear combinations of the first manifold.

$$
\begin{aligned}
&U_{h,B}A_I\Big(\sum_\zeta \int \psi(\mathbf{p},\zeta)A_I|\mathbf{p},\zeta\rangle d^4p\Big)\\
&= U_{h,B}\Big(\sum_\zeta \int \psi(\mathbf{p},\zeta)^* A_I{}^2|\mathbf{p},\zeta\rangle d^4p\Big)\\
&= \epsilon_I U_{h,B}\Big(\sum_\zeta \int \psi(\mathbf{p},\zeta)^*|\mathbf{p},\zeta\rangle d^4p\Big)
\end{aligned}
\tag{9.2}
$$

and this is, because of the invariance of the first manifold, itself a member of that manifold.

In order to demonstrate that $U_{h,B}A_I$, if applied to a member of the first manifold, becomes, conversely, a member of the second manifold, one has to make use again of (1):

$$
\begin{aligned}
&U_{h,B}A_I\Big(\sum_\zeta \int \psi(\mathbf{p},\zeta)|\mathbf{p},\zeta\rangle d^4p\Big)\\
&= A_I\Big(U_{-h,B}\sum_\zeta \int \psi(p,\zeta)|\mathbf{p},\zeta\rangle d^4p\Big)
\end{aligned}
\tag{9.2a}
$$

The vector on the right side is obtained by applying A_I to a vector of the first manifold, it is therefore itself a vector of the second manifold. The union of the two manifolds, i.e. the linear combinations of the $|\mathbf{p},\zeta\rangle$ and $A_I|\mathbf{p},\zeta\rangle$, are invariant under all the operations of the full Poincaré group.

If the manifold of the $A_I|\mathbf{p},\zeta\rangle$ is linearly independent of the manifold of the $|\mathbf{p},\zeta\rangle$, the introduction of the inversion operator A_I leads to a doubling of states. This need not be the case, however, and is usually not assumed to be the case. Hence, the possibility of the existence of linear relations in the manifold composed of the $|\mathbf{p},\zeta\rangle$ and $A_I|\mathbf{p},\zeta\rangle$ needs to be investigated.

Let us apply T_h to $A_I|\mathbf{p},\zeta\rangle$. The group equations then give

$$T_h(A_I|\mathbf{p},\zeta\rangle) = A_I T_{-h}|\mathbf{p},\zeta\rangle = A_I \exp(-i\langle h\mathbf{p}\rangle)|\mathbf{p},\zeta\rangle.$$

Since A_I is antiunitary

$$T_h(A_I|\mathbf{p},\zeta) = \exp(i\langle h\mathbf{p}\rangle)(A_I|\mathbf{p},\zeta\rangle). \tag{9.3}$$

Under the displacement group, the $A_I|\mathbf{p},\zeta\rangle$, transform in the same way as the $|\mathbf{p},\eta\rangle$ with the same $\mathbf{p}$, but any η. Hence, if there is a linear relation in our manifold, there is such a relation already between the $|\mathbf{p},\eta\rangle$ and $A_I|\mathbf{p},\zeta\rangle$ with a single $\mathbf{p}$.

Let us now apply L_B to an $A_I|\mathbf{p}, \zeta\rangle$. The result is

$$\begin{aligned} L_B(A_I|\mathbf{p}, \zeta\rangle) &= A_I L_B|\mathbf{p}, \zeta\rangle = A_I \sum_\eta D^{(S)}(u)_{\eta\zeta}|\mathbf{p}', \eta\rangle \\ &= \sum_\eta D^{(S)}(u)_{\eta\zeta}{}^* A_I|\mathbf{p}', \eta\rangle \end{aligned} \tag{9.4}$$

where (cf. (4.24))

$$\mathbf{p}' = B^{\dagger-1}\mathbf{p}B^{-1}, \quad u = (B^{\dagger-1}\mathbf{p}B^{-1})^{1/2}B\mathbf{p}^{-1/2}. \tag{9.4a}$$

It follows that $A_I|\mathbf{p}, \zeta\rangle$, which transforms as $|\mathbf{p}, \zeta\rangle$ under displacements, transforms also under the homogeneous operations B of the restricted group in a similar way to $|\mathbf{p}, \zeta\rangle$, the only exception being that instead of $D^{(S)}$, its conjugate complex occurs in the transformation (4) of $A_I|\mathbf{p}, \zeta\rangle$. However, $D^{(S)*}$ is equivalent to $D^{(S)}$ and can be transformed into this by the well known matrix C

$$D^{(S)}(u) = CD^{(S)}(u)^*C^{-1} \tag{9.5}$$

where

$$C_{\mu\nu}{}^{(S)} = (-)^{S+\mu}\delta_{\mu,-\nu}. \quad (-S \leq \mu, \nu \leq S). \tag{9.5a}$$

Its only finite elements are on the skew diagonal. It follows that the transformation formulae for

$$|\mathbf{p}, \zeta\rangle\rangle = \sum_\eta C_{\eta\zeta}{}^{(S)}|\mathbf{p}, \eta\rangle \tag{9.6}$$

are exactly the same as for $A_I|\mathbf{p}, \zeta\rangle$. It then follows easily—most directly by the consideration of the B for which $\mathbf{p}' = \mathbf{p}$—that the only linear relations that can hold between the $A_I|\mathbf{p}, \zeta\rangle$ and the $|\mathbf{p}, \eta\rangle$ are

$$A_I|\mathbf{p}, \zeta\rangle = \omega|\mathbf{p}, \zeta\rangle\rangle = \omega \sum_\eta C_{\eta\zeta}{}^{(S)}|\mathbf{p}, \eta\rangle \tag{9.7}$$

the ω being independent of ζ. Further, application of the general L_B on (7) shows that ω is also independent of $\mathbf{p}$. This is a special case of the general rule that if two sets of vectors belong to the same irreducible representation and one vector of the first set is obtained from the corresponding vector of the other set by multiplication with a constant factor, then each vector of the first set can be obtained by multiplication with the same constant from the corresponding vector of the second set. Since ω in (7) is the same for all vectors $A_I|\mathbf{p}, \zeta\rangle$,

it can also be eliminated by replacing A_I by $\omega^* A_I$. It follows that if there is any linear relation between the $|\mathbf{p}, \zeta\rangle$ and the $A_I|\mathbf{p}, \zeta\rangle$, then all the relations

$$A_I|\mathbf{p}, \zeta\rangle = \sum_{\eta\eta} C_{\eta\zeta}^{(S)}|\mathbf{p}, \eta\rangle \tag{9.8}$$

can be assumed to hold.

Before concluding that the relation (8) is possible, the remaining relations $U_\Sigma A_I = \epsilon_\Theta \epsilon_I A_I U_\Sigma$ and $A_I^2 = \epsilon_I$ must be checked. As for the latter, we calculate

$$A_I^2|\mathbf{p}, \zeta\rangle = A_I A_I|\mathbf{p}, \zeta\rangle = A_I \sum C_{\eta\zeta}^{(S)}|\mathbf{p}, \eta\rangle$$

$$= \sum_{\lambda\eta} C_{\lambda\eta}^{(S)} C_{\eta\zeta}^{(S)}|\mathbf{p}, \lambda\rangle.$$

The coefficient of $|\mathbf{p}, \lambda\rangle$ on the right is the $\lambda\zeta$ element of the square of $C^{(S)}$. This is, however, $(-)^{2S}$ as can easily be calculated from (5a). Hence, (8) can hold only if

$$\epsilon_I = (-)^{2S} \tag{9.9}$$

For $\epsilon_I = -(-)^{2S}$ the relation (8) cannot be valid and since we have already concluded that this is the only possible relation, no linear equation between the $|\mathbf{p}, \zeta\rangle$ and $A_I|\mathbf{p}, \zeta\rangle$ is possible and the manifold of these vectors, which is invariant under all the transformations of the full Poincaré group, is irreducible.

Let us now consider the last remaining relation, $U_\Sigma A_I = \epsilon_\Theta \epsilon_I A_I U_\Sigma$. We apply U_Σ to both sides of (8) and obtain by means of (8.6)

$$U_\Sigma A_I|\mathbf{p}, \zeta\rangle = \epsilon_\Theta \epsilon_I A_I U_\Sigma|\mathbf{p}, \zeta\rangle = \epsilon_\Theta \epsilon_I A_I|\sigma\mathbf{p}^*\sigma, \zeta\rangle. \tag{9.10}$$

On the other hand

$$U_\Sigma \sum_\eta C_{\eta\zeta}^{(S)}|\mathbf{p}, \zeta\rangle = \sum_\eta C_{\eta\zeta}^{(S)}|\sigma\mathbf{p}^*\sigma, \zeta\rangle. \tag{9.10a}$$

This last expression is, if (8) is valid, equal to the right side of (10) if $\epsilon_\theta \epsilon_I = 1$ but is oppositely equal thereto if $\epsilon_\Theta \epsilon_I = -1$. In this case, U_Σ if applied to the vectors of the linear manifold $A_I|\mathbf{p}, \zeta\rangle$, transforms them according to (8.6b) or, expressed differently, A_I changes the parity of the states $|\mathbf{p}, \zeta\rangle$. It is worth recalling that, as long as we consider U_Σ to be a valid symmetry operation, this fact of the *opposite* nature of the parties of the states of the two manifolds,

$|\mathbf{p}, \zeta\rangle$ and $A_I|\mathbf{p}, \zeta\rangle$ cannot be eliminated by attaching a phase factor to U_Σ. At any rate, if $\epsilon_\Theta\epsilon_I = -1$, equation (8) leads to a contradiction and cannot be a valid relation. Hence in this case, similar to the case $\epsilon_I = -(-)^{2S}$, we have a doubling of all states and the manifolds $|\mathbf{p}, \zeta\rangle$ and $A_I|\mathbf{p}, \zeta\rangle$ are linearly independent, in fact orthogonal.

It now follows that only if $(-)^{2S} = \epsilon_I = \epsilon_\Theta$ can (8) be valid. Hence the reflection operators will cause a doubling of the irreducible manifold $|\mathbf{p}, \zeta\rangle$, which is invariant under the restricted group, unless $\epsilon_I = \epsilon_\Theta = (-)^{2S}$. Naturally, nothing that has been said so far proves that (8) is indeed valid, and that there is no doubling of the states, even in the case of $\epsilon_\Theta = \epsilon_I = (-)^{2S}$. However, we can enforce the validity of (8) by the same method which we used to show that (8.6) is valid, possibly after changing the sign of U_Σ, in every irreducible representation. For this purpose, we again define vectors

$$A_I|\mathbf{p}, \zeta\rangle + \sum_\eta C_{\eta\zeta}^{(S)}|\mathbf{p}, \eta\rangle = |\mathbf{p}, \zeta\rangle_1 \tag{9.11a}$$

and

$$A_I|\mathbf{p}, \zeta\rangle - \sum_\eta C_{\eta\zeta}^{(S)}|\mathbf{p}, \eta\rangle = |\mathbf{p}, \zeta\rangle_2 \tag{9.11b}$$

Under the operations $U_{h,B}$ of the restricted group, the vectors $|\mathbf{p}, \zeta\rangle_1$ transform as the $A_I|\mathbf{p}, \zeta\rangle$ (i.e. according to (3) and (4)). The same is true of the $|\mathbf{p}, \zeta\rangle_2$. If A_I is applied to $|\mathbf{p}, \zeta\rangle_1$, it becomes

$$A_I|\mathbf{p}, \zeta\rangle_1 = \epsilon_I|\mathbf{p}, \zeta\rangle + \sum_\eta C_{\eta\zeta}^{(S)}A_I|\mathbf{p}, \eta\rangle. \tag{9.12}$$

Because of

$$\sum_{\zeta'} C_{\eta\zeta'}^{(S)}C_{\zeta'\zeta}^{(S)} = (-)^{2S}\delta_{\zeta\eta}$$

(12) is equal to

$$\sum_{\zeta'} C_{\zeta'\zeta}^{(S)}|\mathbf{p}, \zeta'\rangle_1$$

if $(-)^{2S} = \epsilon_I$. Hence the linear manifold $|\mathbf{p}, \zeta\rangle_1$ is invariant under A_I if $(-)^{2S} = \epsilon_I$. The same applies to the manifold of the $|\mathbf{p}, \zeta\rangle_2$. A similar consideration shows that both manifolds are invariant also under U_Σ if $\epsilon_I\epsilon_\Theta = 1$. They are therefore also invariant under the operations of the full Poincaré group and the linear set of the $|\mathbf{p}, \zeta\rangle$ and $A_I|\mathbf{p}, \zeta\rangle$ could be decomposed into two invariant sets. In an irreducible invariant set, (8) can be assumed to be valid, possibly after changing the sign of A_I. In a reducible set, however, there may be states for which (8) is valid, and other states for which the two sides of (8) are oppositely equal. Hence, in a reducible system, there may be a relative I-parity just as there may be a relative Σ-parity. One of these parities can be replaced by a Θ-parity. Since A_Θ is not believed to cross superselection lines, the definition of the relative Θ parity is less subject to the doubts which will be discussed in the last section.

The discussion of the present section is an adaptation of the general discussion of corepresentations to the mathematically relatively simple case which faced us here. The general discussion itself shows much similarity to the oft-mentioned Frobenius method. There is a great difference between the result of the present discussion and the general mathematical discussion of the corepresentations of a definite group. The latter shows that every irreducible representation of the unitary subgroup can be extended only to one irreducible corepresentation of the whole group. The reason for the difference is not that the general theory is inapplicable but that the "whole group" is not uniquely defined. It remains true, also in our case, that every irreducible representation of the unitary subgroup (i.e. the restricted Poincaré group augmented by space reflections) can be extended to only one corepresentation of every one of the full groups. There are, however, four quantum mechanical full Poincaré groups and every irreducible representation can be extended to a corepresentation of each of them. The customary assumption is that the full group is characterized by $\epsilon_\Theta = \epsilon_I = (-)^{2S}$—the case which does not lead to a doubling of states. The other possibilities will be explored in the next section.

The following table gives an abbreviated summary of the results of this section. The two rows of the matrices which occur in it refer to the first and second linear manifold discussed at the beginning of this section. A_Θ can be obtained as $U_\Sigma A_I$.

TABLE 3

Summary of Results

Characterization of full Poincaré group ϵ_θ	ϵ_I	U_Σ times $\mathbf{p} \to p = \sigma p^* \sigma$	A_I
$(-)^{2S}$	$(-)^{2S}$	1	C
$-(-)^{2S}$	$(-)^{2S}$	$\begin{Vmatrix} 1 & 0 \\ 0 & -1 \end{Vmatrix}$	$\begin{Vmatrix} 0 & C \\ C & 0 \end{Vmatrix}$
$(-)^{2S}$	$-(-)^{2S}$	$\begin{Vmatrix} 1 & 0 \\ 0 & -1 \end{Vmatrix}$	$\begin{Vmatrix} 0 & C \\ -C & 0 \end{Vmatrix}$
$-(-)^{2S}$	$-(-)^{2S}$	$\begin{Vmatrix} 1 & 0 \\ 0 & 1 \end{Vmatrix}$	$\begin{Vmatrix} 0 & C \\ -C & 0 \end{Vmatrix}$

In all cases considered, the basic vectors have been chosen in such a way that the operators of the restricted group have the form

$$\begin{Vmatrix} U_{h,B} & 0 \\ 0 & U_{h,B} \end{Vmatrix} \tag{9.13}$$

As far as the restricted group is concerned, the two linear manifolds do not differ. Hence, the vectors which correspond to the second rows of the matrices of Table 3 or of (13) are not the $A_I|\mathbf{p}, \zeta\rangle$ themselves but the vectors

$$(-)^{2S} \sum C_{\zeta\eta} A_I|\mathbf{p}, \eta\rangle = (-)^{-S+\zeta} A_I|\mathbf{p}, -\zeta\rangle.$$

Both manifolds are also invariant under U_Σ but they differ from each other if $\epsilon_\Theta \epsilon_I = -1$, that is the second and third case. In these cases, A_I (or A_Θ) transforms a positive Σ-parity state into a negative one, and vice versa. A_I is, of course, antiunitary even though this is not indicated in the table which gives only the transformation of the basic vectors. Finally, it should be noted that in all cases considered, two representations or corepresentations are considered to be

equivalent if their operators can be transformed into each other by a similarity transformation ($S \ldots S^{-1}$ where S is the same for all operators), followed by a phase renormalization, i.e., multiplication with a numerical factor of modulus 1 which can be different from operator to operator.

10. Attempt at Physical Interpretation

The present notes do not consider the zero-mass representations of the Poincaré group. However, the extension of the representations of the restricted group with zero mass to corepresentations of the full group can be carried out along the lines employed for the finite mass representations. The following table gives the number of independent states for a definite momentum vector both for the zero mass and for the finite mass representations.

TABLE 4

Number of Independent States for Definite Momentum, in Corepresentations

Group		Representation of restricted group					
		Finite mass			Zero mass		
ϵ_Θ	ϵ_I	$S = 0$	S integer	half int.	$S = 0$	S integer	half int.
1	1	1	$2S + 1$	$4S + 2$	1	2	4
1	−1	2	$4S + 2$	$4S + 2$	2	2	4
−1	1	2	$4S + 2$	$4S + 2$	2	4	2
−1	−1	2	$4S + 2$	$2S + 1$	2	4	2

The problem now is to determine the "type", i.e. the full quantum mechanical Poincaré group according to which the various particles, such as photon, electron, etc., transform. The hope is that the coordination can be made uniquely, that it explains some of the *a priori* surprising facts such as the equality of the rest masses of electron and position.

Let us try to determine, first, the "type" of the photon. This has no mass and the behavior of the classical description of the photon, that of the electromagnetic field, under restricted Lorentz transformations is such that $S = 1$ has to be attributed to it. If we apply a

time inversion to the electromagnetic field, the electric vector will not change its direction, the magnetic will. However, if this operation is carried out twice, no change remains so that $\epsilon_\Theta = 1$. If the time reversal is followed by a space inversion, the electric vector again remains unchanged, the magnetic changes sign. This follows from the new way of looking at reflections, discussed in section 6. We can decompose the operation of space inversion into a reflection in a plane parallel to the electric vector, and a rotation by π about an axis perpendicular to this plane. Both operations change the direction of the electric vector. In the case of the magnetic vector, the first operation has no effect. Hence, $I = \Sigma\Theta$ leaves the direction of the field vectors unchanged and this is then true also of I^2. Hence, we also have $\epsilon_I = 1$. The preceding consideration uses classical concepts but the use of quantum concepts can be expected to give the same result. Hence the type of the light quantum is $\epsilon_\Theta = \epsilon_I = 1$, or $++$, and the same result could be expected always if the relation between classical and quantum descriptions is as close as in the case of the light quantum. Table 4 shows that a particle of type $++$ with zero mass and $S = 1$ has two independent states of directions of polarization and this is in agreement with what we know about photons.

Let us consider now the case of a charged particle or, more generally, a particle which is different from its antiparticle. We also assume that the particle has a finite mass. If particle and antiparticle are different, it must be possible to define an operator Q with two characteristic values, $+1$ and -1, so that the particle states belong to the characteristic value $+1$, the antiparticle states to the characteristic value -1.

Since the manifold of the particle states is invariant under the operations of the restricted group, and since the same applies to the antiparticle states, these operations must commute with Q. The operations in question have the form (9.13)

$$\begin{Vmatrix} U_{h,B} & 0 \\ 0 & U_{h,B} \end{Vmatrix} \tag{10.1}$$

in the scheme used in Table 3. An operator which commutes with the operators (1) for every h and B must have the form

$$\begin{Vmatrix} Q_{11}1 & Q_{12}1 \\ Q_{21}1 & Q_{22}1 \end{Vmatrix} \tag{10.2}$$

The Q_{ik} are members, the 1 indicates that Q_{11} 1, for instance, which is to be applied to the members of the first manifold, merely multiplies these members by Q_{11}.

The manifold of particle states, and also the manifold of antiparticle states, must also be invariant under $A_\Theta = U_\Sigma A_I$. This means that Q commutes with A_Θ. If we consider, for instance, the operators of the third row, $\epsilon_\Theta = (-)^{2S}$, $\epsilon_I = -(-)^{2S}$ of Table 3, we obtain using $A_\Theta = U_\Sigma A_I$

$$\begin{Vmatrix} 0 & C \\ C & 0 \end{Vmatrix} \begin{Vmatrix} Q_{11}{}^*1 & Q_{12}{}^*1 \\ Q_{21}{}^*1 & Q_{22}{}^*1 \end{Vmatrix}$$

$$= \begin{Vmatrix} Q_{11}1 & Q_{12}1 \\ Q_{21}1 & Q_{22}1 \end{Vmatrix} \begin{Vmatrix} 0 & C \\ C & 0 \end{Vmatrix} \tag{10.3}$$

The conjugate complex signs appear on the left because A_Θ is antiunitary. Finally, Q must anticommute with U_Σ because space reflection transforms particles into antiparticles and vice versa. It is easy to see from Table 3 that this last condition cannot be fulfilled in the first and last case enumerated there, i.e. in the cases in which $\epsilon_\Theta = \epsilon_I$; it will be fulfilled in the two other cases if $Q_{11} = Q_{22} = 0$. Hence, (3) now gives

$$Q_{21}{}^*C = Q_{12}C \qquad Q_{12}{}^*C = Q_{21}C \tag{10.4}$$

or $Q_{12} = Q_{21},^*$, a condition which is fulfilled anyway because of the hermitean nature of Q. Had we considered the case $\epsilon_\Theta = -(-)^{2S}$, $\epsilon_I = (-)^{2S}$, the result would have been $Q_{12} = -Q_{21}{}^*$ which is inconsistent with the self-adjoint nature of Q. Hence, a particle-antiparticle system transforms according to the group characterized by $\epsilon_\Theta = (-)^{2S}$; $\epsilon_I = -(-)^{2S}$.

The operator Q remains to be more fully determined. Since its two characteristic values are 1 and -1, the determinant of the Q_{ik} must be -1 and this leads to

$$Q = \begin{Vmatrix} 0 & \omega 1 \\ \omega 1 & 0 \end{Vmatrix} \tag{10.5}$$

where $|\omega| = 1$. If we now transform all operators of the representation by

$$S = \begin{Vmatrix} \sqrt{\omega^*1} & 0 \\ 0 & \sqrt{\omega}1 \end{Vmatrix} \tag{10.6}$$

the operators of the unitary subgroup, (1), and the U_Σ of Table 3, will remain unchanged. The same is true, however, also of A_I because of its antiunitary nature. This is true then for all symmetry operators which are products of proper Poincaré transformations with 1, U_Σ, 0_I or $0_\theta = U_\Sigma 0_I$. On the other hand Q will go over into

$$Q = \begin{Vmatrix} 0 & 1 \\ 1 & 0 \end{Vmatrix} \tag{10.7}$$

which we may adopt as the normalized form of this operator. The particle states are then linear combinations of the kets

$$|\mathbf{p}, \zeta\rangle + (-)^{2S} \sum_\eta C_{\zeta\eta} A_I |\mathbf{p}, \eta\rangle \tag{10.8}$$

whereas the antiparticle states are linear combinations of

$$|\mathbf{p}, \zeta\rangle - (-)^{2S} \sum_\eta C_{\zeta\eta} A_I |\mathbf{p}, \eta\rangle \tag{10.8a}$$

C is the matrix defined in (9.5) or (9.5a).

The preceding consideration can be extended to other particles, perhaps even to unstable ones. There is some doubt in this regard because the unstable particles do not belong, strictly speaking, to any irreducible representation of the Poincaré groups—a time displacement does not leave them in the manifold which can be considered to be a state of the unstable particle. A good deal more can be said about this question. At any rate, the determination of the type leads, in all cases, to one of the types encountered so far, that is to $\epsilon_\Theta = \epsilon_I = (-)^{2S}$ (normal type) or to $\epsilon_\Theta = (-)^{2S} = -\epsilon_I$. The usual assumption, on the contrary, is that the transformation properties of all particles corresponds to the $\epsilon_\Theta = \epsilon_I = (-)^{2S}$ type. This is different from the conclusion here arrived at.

It might be noted that no example for the other types, corresponding to the second and fourth rows of Table 3, has been found. In fact, one can adduce empirical evidence to conclude that there are no particles in nature which transform according to one of these groups. If this is true, $\epsilon_\Theta = (-)^{2S}$ is universally valid.

It may be worth mentioning that the increased number of states in the case of all groups, except the $\epsilon_\Theta = \epsilon_I = (-)^{2S}$ group permits a wider variety of properties of the larger manifold of states. Nevertheless, the well-known argument of Lee and Yang concerning the

absence of an electric dipole moment, as a result of the time inversion symmetry, remains valid in all cases.

II. Critique of the Interpretation

The most serious doubts concerning the validity of the preceding considerations arises from the fact that, according to the interpretation of the space inversion symmetry, given in section 6, this operation may well "cross" a superselection line. This means, in more simple words, that it may transform every realizable state into such a state which, although also realizable by itself, cannot be superposed over the original state. Let us consider, for instance, an arbitrary state ψ_e of the electron. U_Σ transforms this into a definite state ψ_p of the positron. This, in itself, is also a realizable state. However, states such as $\alpha\psi_e + \beta\psi_p$ are probably not realizable and it is in fact difficult to imagine how this state could be distinguished from any state $\alpha\psi_e + \omega\beta\psi_p$, where $|\omega| = 1$. This is what is meant by the statement that a superselection line divides the states of the positron from those of the electron.

If the presence of a definite state such as $\alpha\psi_e + \beta\psi_p$ is not verifiable, it is not possible to prove that there is one, and except for a phase factor only one, unitary or antiunitary operator which transforms the vector of any state into the vector of the state obtained from it by the symmetry operation. In other words, the operator U_Σ cannot be defined as an operator in one Hilbert space but must be regarded as an operator which transforms the vectors of one Hilbert space (that of the electron states) into vectors of another Hilbert space (that of the positron states). Both of these Hilbert spaces contain, in a sense, fewer states than their direct product which is the space we used.

The preceding difficulty applies not only to U_Σ itself but to all operators which transform particles into antiparticles. It does not apply to A_Θ but does apply for instance, to A_I. It is possible, for instance, to redefine A_I by changing only its effect when applied to positron states by multiplying it with a factor. If this factor is properly chosen, the square of A_I will be 1—or, for another choice, any other member of modulus 1. This shows that if the superselection rule is valid, the quantity $A_I^2 = \epsilon_I$ is not uniquely defined. A change in the definition of A_I is not possible if a state $\alpha\psi_e + \beta\psi_p$

is distinguishable from the state $\alpha\psi_e + \omega\beta\psi_p$ with $\omega \neq 1$. Thus if $A_I\psi_e = \psi'_p$ and $A_I\psi_p = \psi'_e$, we have

$$A_I(\alpha\psi_e + \beta\psi_p) = \alpha^*\psi'_p + \beta^*\psi'_e \tag{11.1}$$

and

$$A_I^2(\alpha\psi_e + \beta\psi p) = \alpha A_I\psi'_p + \beta A_I\psi'_e \tag{11.1a}$$

so that if this is to be, within a factor, $\alpha\psi_e + \beta\psi_p$, we must have

$$A_I\psi'_p = \epsilon\psi_e, \qquad A_I\psi'_e = \epsilon\psi_p \tag{11.1b}$$

with the same factor ϵ in both equations. If we now change $A_I\psi_p$ to $\omega\psi'_e$, $A_I\psi'_p$ to $\omega\epsilon\psi_e$, leaving $A_I\psi_e = \psi'_p$ and $A_I\psi'_e = \epsilon\psi_p$, we shall have

$$\begin{aligned} A_I^2(\alpha\psi_e + \beta\psi_p) &= A_I(\alpha^*\psi'_p + \omega\beta^*\psi'_e) \\ &= \alpha A_I\psi'_p + \omega^*\beta A_I\psi'_e \\ &= \omega\alpha\epsilon\psi_e + \omega^*\beta\epsilon\psi_p \end{aligned} \tag{11.2}$$

which, if $\omega \neq \omega^*$, is distinguishable from $\alpha\psi_e + \beta\psi_p$. Hence (as we know) the change in the phase factor in *part* of the range of A_I is not possible if the two states, $\omega\alpha\epsilon\psi_e + \omega^*\beta\epsilon\psi_p$ and $\alpha\psi_e + \beta\psi_p$ are distinguishable. It is possible, however, if these two states are undistinguishable, i.e. if a superselection rule divides the electron states from the positron states.

If the point of view just sketched is adopted, only A_Θ (and its products with the transformations of the restricted group) remains as reflection symmetry to which a unitary or antiunitary (actually antiunitary) operator corresponds. The same consideration which led us, in section 9, to four quantum mechanical Poincaré groups, then leads to only two such groups. These can be characterized by the value of $\epsilon_\Theta = A_\Theta^2$, or by that of $\epsilon_\Theta(-)^{2S}$. These quantities can assume the values $+1$ or -1. In actual fact, however, $\epsilon_\Theta(-)^{2S}$ has the value $+1$ for all known particles and it was mentioned in the preceding section that there are (empirical) reasons for believing that no particles for which $\epsilon_\Theta(-)^{2S} = -1$ exist, at least as far as particles with non-zero mass are concerned.

If the operations which involve space inversions are eliminated from the group, no corepresentation leads to a doubling of states. The manifold of states which belongs to the corepresentation $\epsilon_\Theta = (-)^{2S}$

$= -\epsilon_I$ decomposes, if only invariance under the subgroup is demanded which does not contain space (or space-time) inversions, into two invariant manifolds, spanned by the vectors (10.8) and (10.8a), respectively. One of the conditions which led to the determination of these vectors was in fact time inversion invariance. This does not mean, however, that the existence of particles and antiparticles of equal restmass and spin remains unexplained by symmetry operations. *Space (or space-time) inversion still transforms particles into antiparticles* but no unique unitary or antiunitary operator can be coordinated to these symmetry operations. They play a role similar to that played by all symmetry operations in elementary classical theory; rotational invariance shows that if there is an orbit with energy E, many other orbits with the same energy can be derived therefrom by rotations. In our case of space inversion, which is a single operation, only one new state can be obtained which may or may not be identical with the original state. In the former case we say that the particle is identical with its antiparticle, in the second case particle and antiparticle are different.

It is difficult to tell which of the two points of view will prevail: the one adopted in section 9, or the one sketched here. The present one is more secure as it does not assume as much distinguishability between states as was assumed in the preceding work. On the other hand, it is not sure that particle and antiparticle states are in every case separated by a superselection rule. It is not sure, either, that the assumption of an essentially unique operator U_Σ leads to an incorrect result even if this assumption cannot be based on firmly established principles of measurability. Many concepts in physics, though not based on such principles, proved very fruitful.

This concludes the criticism of the considerations of the preceding sections, as far as their conceptual basis is concerned. It remains to be mentioned that the paucity of the results is somewhat disappointing. I had hoped that the consideration of the reflection operators would introduce some more general additive quantity, or quantities, which then could be identified with electric charge or baryon number. This has not been the case, at least not in the framework of the present quantum mechanical theory which operates with a Hilbert space with the ordinary complex members as vector components.

It is not only customary, it is almost axiomatic in present theory, to associate the direct product of the Hilbert spaces of two systems (or at

least the symmetric or antisymmetric part of this product) with the union of the two systems. This then leads to rules, such as that the parity of a composite system is the product of the parities of the constituents. The cluster theorem should help to establish such rules even more firmly in present quantum field theories. The product rule may in turn provide a definite preference for the normalization of the operators of the constituent, elementary systems. In some cases the normalization of the operators which is preferable in order to maintain the product rule may be different from the normalization which appears most natural from the point of view of the constituent system alone. Such indications have not been taken into account in the present discussion and all the normalizations were adopted solely on the basis of simplicity in the description of the elementary system considered. It is clear, however, that no arbitrary normalization can have any effect on the physical content of the theory—it can only make the description more simple or more cumbersome. We have considered only the simplicity of the description from the point of view of the systems considered, not from the point of view of the description of unions of such systems.

It would be very difficult to give a full enumeration of all who have contributed to the problems discussed in the present series of lectures. In addition to my closest colleagues, Bargmann and Wightman, I have learned a great deal from a participant at this conference, Professor L. Michel, and derived much stimulation from a colleague in Leiden, Dr. R. M. F. Houtappel, and from R. Haag. Most of what has been reported in the preceding pages was known to all of us for several years. (See A. S. Wightman's article in the collection *Les Problèmes Mathematiques de la Théorie Quantique des Champs*, Centre National de la Récherche Scientifique, Paris, 1959). There are, however, also some publications on the subject. Those which occur to me at once are due to J. C. Taylor (*Nucl. Phys.* **3**, 606(1957)), to J. A. McLennan (*Phys. Rev.* **109**, 986 (1958)) and to H. Ekstein (*Phys. Rev.* **120**, 1917 (1960)).

On the Matrices Which Reduce the Kronecker Products of Representations of S. R. Groups*

E. P. Wigner

In: Quantum Theory of Angular Momentum (L. C. Biedenharn and H. van Dam, eds.). Academic Press, New York 1965, pp. 87–133

I. Simply Reducible Groups 89
1. Definition of Simply Reducible Groups 89
2. Invariant Vectors in Representation Space (Three-j-Symbols or Clebsch Gordan Coefficients); Covariant and Contravariant Components 89
3. The Six-j-Symbols (Racah Coefficients) 96
4. The Metric in Representation Space 100
5. Further Properties of the Six-j-Symbols 101
6. Expressions for Six-j-Symbols in Terms of Group Integrals 103
II. The Three-Dimensional Rotation Group 104
7. The Representations of the Three-Dimensional Rotation Group 104
8. The Three-Dimensional Rotation Group is Simply Reducible 107
9. The Three-j-Symbols for the Three-Dimensional Rotation Group 109
10. Calculation of the Racah Coefficients 111
11. Calculation of the Three-j-Symbols 114
12. Calculation of the Representation Coefficients 117
III. Applications to Quantum Mechanics 119
13. The Hilbert Space as a Union of Representation Spaces; Irreducible Tensors 119
14. Transition to Contravariant Components as Time Inversion 126
15. Union of Two Quantum Mechanical Systems; Nine-j-Symbols 127
16. Combinations of Irreducible Tensors 131
References 132

The present article deals with simply reducible groups which are finite or compact and with their unitary representations (more precisely: with their representations which are in unitary form) in general. A simply reducible group satisfies two criteria. First, all classes are ambivalent, i.e., contain with an element X also its reciprocal, X^{-1}. Since the characters of all elements of a class are equal to each other in every representation, it follows that the characters of X and X^{-1} are equal. If the representation is in unitary form, it is at once evident that the characters of reciprocals are conjugate complex. It therefore follows from the ambivalent nature of all classes that all characters are real in every representation of a simply reducible group. As a

* *Editors' note*: The original version of Professor Wigner's paper differed in only minor ways from the later version given here. The changes are entirely of a clarifying nature, except that references to a few articles which have appeared since 1940 have been added, and the more modern terminology (Racah coefficients, multiplicity free representations, etc.) occasionally mentioned. Also, the left-handed coordinate system used in the original manuscript was changed into a right-handed one by interchanging X and Y. We feel that the usefulness of these changes outweighs that of historical accuracy.

result, every representation is equivalent to the conjugate complex representation.

The second condition for a group to be called simply reducible is that the irreducible representations of the group be, in the terminology of G. R. Mackey, multiplicity free. This means that the Kronecker (direct) product of two irreducible representations, j_1 and j_2, does not contain any irreducible representation more than once. If the direct product of j_1 and j_2 contains j_3, we set the symbol $(j_1 j_2 j_3) = 1$; if it does not contain j_3, we set $(j_1 j_2 j_3) = 0$. Since all representations to be considered are equivalent to their conjugate complex, if the product of j_1 and j_2 contains j_3, the product of j_2 and j_3 contains j_1. Hence, the symbol $(j_1 j_2 j_3)$ is symmetric in its arguments. The multiplicity free nature of the representations is the condition for the validity of what is often called the Eckart-Wigner theorem in quantum mechanics, and it was this circumstance which attracted this writer's attention to the subject.

The present article, which is published here for the first time, is the continuation of the article[6] in which simply reducible groups were defined. A summary of the relevant results of this article is given in Reference 6, and its notation is followed in the present article. Group elements are denoted by capitals, R, S, T, ... Summation over these is always extended over all elements in case of finite groups; the invariant integral is meant in the case of continuous groups. In the former case, h is the order of the group; in the latter case, h is the volume thereof. Other symbols to be used are:

$[j]$ dimension of the irreducible representation j; in the case of the three-dimensional rotation group $[j] = 2j + 1$.

$\begin{bmatrix} j & R \\ \kappa & \lambda \end{bmatrix}$ $\kappa\lambda$ matrix element of the matrix which the unitary irreducible representation j coordinates to the group element R. The customary notation is $\begin{bmatrix} j & R \\ \kappa & \lambda \end{bmatrix} = D^{(j)}(R)_{\kappa\lambda}$.

$\begin{pmatrix} j_1 & j_2 & j_3 \\ \kappa_1 & \kappa_2 & \kappa_3 \end{pmatrix}$ coefficient of the vector addition model. An equivalent quantity was calculated for the three-dimensional rotation group in the writer's book.[3] The relation between the two notations is, explicitly,

$$s^{(ll)}_{\mathscr{L}\mu\nu} = (-)^{l-l+\mu+\nu}(2\mathscr{L}+1)^{1/2} \begin{pmatrix} l & l & \mathscr{L} \\ \mu & \nu & -\mu-\nu \end{pmatrix}.$$

$\begin{Bmatrix} j_1 & j_2 & j_3 \\ l_1 & l_2 & l_3 \end{Bmatrix}$ is called, in the case of the three-dimensional rotation group, the Racah coefficient[7] $W(j_1 l_1, j_2 l_2, j_3 l_3)$.

Much more extensive tables for both coefficients than are contained in the present article have been published by H. A. Jahn[8] and L. C. Biedenharn.[9] and also others.*

The star denotes the conjugate imaginary.

* Editors' note: See Bibliography, Section E, p. 321.

S. R. GROUPS

I. Simply Reducible Groups

1. *Definition of Simply Reducible Groups*

The chief purpose of the present paper is to derive some formulas for the irreducible representations of the three-dimensional rotation or the two-dimensional unimodular unitary group. Many of the formulas and relations to be derived were given already in the books quoted in References 3 and 5. It seems appropriate, however, to give a more systematic and exhaustive treatment to this subject than has been done hitherto. The results of the first five sections will be valid not only for the above continuous groups, but for all finite S.R. (simply reducible) groups because the special properties of the above groups will not be utilized in the first five sections.

Equations derived in earlier parts of the paper will be quoted in the customary manner when used again. Some equations expressing symmetry properties of certain symbols form an exception under this rule and will be used without being quoted. The numbers of these equations are set in bold face type where they are derived.

The S.R. groups were defined[6] as groups for which (a) the direct product of two irreducible representations contains no irreducible representation more than once, and (b) every element is equivalent with (contained in the class of) its reciprocal. The notation adopted in the first part of Reference 6 will be used here also. All equations of Sections 2–5 are valid for finite S.R. groups and also for the above continuous groups if the summation over the group elements is replaced by the invariant group integral and h is replaced by the volume of the group, i.e., the integral of 1 over the group. In order to avoid the necessity of distinguishing between the two kinds of groups, the necessary results in Reference 6 (Lemma 1, Theorem 3), will be established for the above continuous groups by direct enumeration in Section 7.

2. *Invariant Vectors in Representation Space (Three-j-Symbols or Clebsch Gordan Coefficients); Covariant and Contravariant Components*

Let us denote the matrix elements of the unitary matrix which brings the Kronecker product of the representations j_1 and j_2 into the reduced form by

$$U_{j_3\kappa_3;\,\kappa_1\kappa_2} = [j_3]^{1/2}\begin{pmatrix} j_1 & j_2 & j_3 \\ \kappa_1 & \kappa_2 & \kappa_3 \end{pmatrix}. \tag{1}$$

All representations are assumed to be in unitary form. In U, j_3 and κ_3 specify the row, κ_1 and κ_2 the column into which the matrix element (1) belongs. The second factor on the right side of (1) will be called a "three-j-symbol." Because of (1), we have

$$\begin{bmatrix} j_1 & R \\ \kappa_1 & \lambda_1 \end{bmatrix}\begin{bmatrix} j_2 & R \\ \kappa_2 & \lambda_2 \end{bmatrix} = \sum_{j_3}[j_3]\begin{pmatrix} j_1 & j_2 & j_3 \\ \kappa_1 & \kappa_2 & \kappa_3 \end{pmatrix}^*\begin{bmatrix} j_3 & R \\ \kappa_3 & \lambda_3 \end{bmatrix}^*\begin{pmatrix} j_1 & j_2 & j_3 \\ \lambda_1 & \lambda_2 & \lambda_3 \end{pmatrix} \tag{2}$$

or

$$\begin{pmatrix} j_1 & j_2 & j_3 \\ \kappa_1 & \kappa_2 & \kappa_3 \end{pmatrix}\begin{bmatrix} j_1 & R \\ \kappa_1 & \lambda_1 \end{bmatrix}\begin{bmatrix} j_2 & R \\ \kappa_2 & \lambda_2 \end{bmatrix} = \begin{bmatrix} j_3 & R \\ \kappa_3 & \lambda_3 \end{bmatrix}^*\begin{pmatrix} j_1 & j_2 & j_3 \\ \lambda_1 & \lambda_2 & \lambda_3 \end{pmatrix}. \tag{2a}$$

The left side is the $j_3\ \kappa_3;\ \lambda_1\ \lambda_2$ matrix element of the product of U with the Kronecker product of the representations j_1 and j_2, the right side is the same matrix element of the product of the reduced form of the above Kronecker product with U. *The summation convention is adopted in these and all following formulas for the greek letters*, i.e., a summation must be carried out over all greek letters occurring twice in the same expression. In order to make it unnecessary to give the limits of summation, we shall set the three-j-symbols equal to zero if the last j does not occur in the Kronecker product of the first two. Then all summations can be extended over all possible values of the variables. This holds, naturally, for the greek letters also, the possible values of which serve to denote the rows and columns of the corresponding representation.

Equation (2) can be given, in addition to (2a), further particular forms. Multiplying this equation by $\begin{bmatrix} j_3 & R \\ \kappa_3 & \lambda_3'' \end{bmatrix}$ and summing over κ_3, one obtains because of the unitary nature of the representation

$$\begin{bmatrix} j_1 & R \\ \kappa_1 & \lambda_1 \end{bmatrix}\begin{bmatrix} j_2 & R \\ \kappa_2 & \lambda_2 \end{bmatrix}\begin{bmatrix} j_3 & R \\ \kappa_3 & \lambda_3 \end{bmatrix}\begin{pmatrix} j_1 & j_2 & j_3 \\ \kappa_1 & \kappa_2 & \kappa_3 \end{pmatrix} = \begin{pmatrix} j_1 & j_2 & j_3 \\ \lambda_1 & \lambda_2 & \lambda_3 \end{pmatrix}. \tag{2b}$$

This expression can be said to express the fact that the three-j-symbols are invariant under the operations of the group. The conjugate imaginary of (2b) can be written, again on account of the unitary nature of the representations

$$\begin{bmatrix} j_1 & R \\ \lambda_1 & \kappa_1 \end{bmatrix}\begin{bmatrix} j_2 & R \\ \lambda_2 & \kappa_2 \end{bmatrix}\begin{bmatrix} j_3 & R \\ \lambda_3 & \kappa_3 \end{bmatrix}\begin{pmatrix} j_1 & j_2 & j_3 \\ \kappa_1 & \kappa_2 & \kappa_3 \end{pmatrix}^* = \begin{pmatrix} j_1 & j_2 & j_3 \\ \lambda_1 & \lambda_2 & \lambda_3 \end{pmatrix}^* \tag{2c}$$

where R^{-1} has been replaced by R. This equation shows that the conjugate imaginaries of the three-j-symbols are also invariant under the group operations though the role of rows and columns is interchanged.

Another form of (2) expresses the fact that while $U^{-1} \ldots U$ transforms the reduced form of the Kronecker product into the Kronecker product $U \ldots U^{-1}$ transforms the Kronecker product into its reduced form. This equation reads

$$\begin{pmatrix} j_1 & j_2 & j_3 \\ \kappa_1 & \kappa_2 & \kappa_3 \end{pmatrix}\begin{bmatrix} j_1 & R \\ \kappa_1 & \lambda_1 \end{bmatrix}\begin{bmatrix} j_2 & R \\ \kappa_2 & \lambda_2 \end{bmatrix}\begin{pmatrix} j_1 & j_2 & j_3' \\ \lambda_1 & \lambda_2 & \lambda_3 \end{pmatrix}^* = [j_3]^{-1}\,\delta_{j_3 j_3'}\begin{bmatrix} j_3 & R \\ \kappa_3 & \lambda_3 \end{bmatrix}^*. \tag{2c}$$

A last form of (2) is given by (5).

If we again adopt the notation that $(-1)^{2j}=1$ if j is an integer representation, $(-1)^{2j}=-1$ if j is a half integer representation, the above convention allows us to write

$$\begin{pmatrix} j_1 & j_2 & j_3 \\ \lambda_1 & \lambda_2 & \lambda_3 \end{pmatrix} = (-1)^{2j_1+2j_2+2j_3} \begin{pmatrix} j_1 & j_2 & j_3 \\ \lambda_1 & \lambda_2 & \lambda_3 \end{pmatrix} \tag{3}$$

because the three-j-symbol in (3) vanishes on account of the lemma in Reference 6 unless $(-1)^{2j_1+2j_2+2j_3}=1$. We shall write, furthermore $(-1)^j=+1$, if j is an even representation; $(-1)^j=-1$ if j is an odd representation (cf. Theorem 3 of reference 6). If j is integer but neither even nor odd, the value of $(-1)^j$ is immaterial for the following and can be taken to be $+1$. If j is half integer, $(-1)^j$ can be assumed to be i or $-i$ arbitrarily, provided that its value, once adopted, is maintained throughout.

The unitary character of the matrix (1) is expressed by

$$\sum_{j_3} [j_3] \begin{pmatrix} j_1 & j_2 & j_3 \\ \kappa_1 & \kappa_2 & \kappa_3 \end{pmatrix} \begin{pmatrix} j_1 & j_2 & j_3 \\ \lambda_1 & \lambda_2 & \kappa_3 \end{pmatrix}^* = \delta_{\kappa_1\lambda_1}\delta_{\kappa_2\lambda_2} \tag{4}$$

$$\begin{pmatrix} j_1 & j_2 & j \\ \lambda_1 & \lambda_2 & \lambda \end{pmatrix} \begin{pmatrix} j_1 & j_2 & j' \\ \lambda_1 & \lambda_2 & \lambda' \end{pmatrix}^* = [j]^{-1}(j_1 \quad j_2 \quad j)\delta_{jj'}\delta_{\lambda\lambda'}. \tag{4a}$$

Here again $(j_1 \quad j_2 \quad j)=1$, if the Kronecker product of j_1 and j_2 contains j; it is zero otherwise. The normalization of the three-j-symbols adopted in (1) is such that

$$\begin{pmatrix} j_1 & j_2 & j_3 \\ \lambda_1 & \lambda_2 & \lambda_3 \end{pmatrix} \begin{pmatrix} j_1 & j_2 & j_3 \\ \lambda_1 & \lambda_2 & \lambda_3 \end{pmatrix}^* = (j_1 \quad j_2 \quad j_3). \tag{4b}$$

The three-j-symbols are not completely determined by the condition that they satisfy (2). The reduced form of the Kronecker product commutes with all diagonal matrices in which the diagonal elements corresponding to one representation are the same (and only with these). Hence, the transforming matrix (1) can be multiplied by a diagonal matrix of this character. This corresponds to multiplying the three-j-symbol of (1) by a quantity depending only on j_3 (and $j_1 j_2$), but independent of κ_1, κ_2, κ_3. If we require that (1) shall continue to be unitary, the modulus of the indeterminate factor, $\omega(j_1, j_2, j_3)$, must be 1.

One can bring, by means of the orthogonality relations, the representation coefficient from the right side of (2) to the left side. This gives

$$\sum_R \begin{bmatrix} j_1 & R \\ \kappa_1 & \lambda_1 \end{bmatrix} \begin{bmatrix} j_2 & R \\ \kappa_2 & \lambda_2 \end{bmatrix} \begin{bmatrix} j_3 & R \\ \kappa_3 & \lambda_3 \end{bmatrix} = h \begin{pmatrix} j_1 & j_2 & j_3 \\ \kappa_1 & \kappa_2 & \kappa_3 \end{pmatrix}^* \begin{pmatrix} j_1 & j_2 & j_3 \\ \lambda_1 & \lambda_2 & \lambda_3 \end{pmatrix}. \tag{5}$$

The purpose of the factor $[j]^{1/2}$ in (1) was to make this expression symmetric in the representation coefficients. It follows from (5) at once that

$$\left|\begin{pmatrix} j_1 & j_2 & j_3 \\ \kappa_1 & \kappa_2 & \kappa_3 \end{pmatrix}\right|^2 = \left|\begin{pmatrix} j_2 & j_1 & j_3 \\ \kappa_2 & \kappa_1 & \kappa_3 \end{pmatrix}\right|^2 = \left|\begin{pmatrix} j_1 & j_3 & j_2 \\ \kappa_1 & \kappa_3 & \kappa_2 \end{pmatrix}\right|^2 \text{ etc.} \tag{5a}$$

i.e., that the interchange of the columns does not affect the absolute value of the three-j-symbols. It will be shown now that it is possible to normalize the latter in such a way that

$$\begin{pmatrix} j_1 & j_2 & j_3 \\ \kappa_1 & \kappa_2 & \kappa_3 \end{pmatrix} = (-1)^{j_1+j_2+j_3} \begin{pmatrix} j_2 & j_1 & j_3 \\ \kappa_2 & \kappa_1 & \kappa_3 \end{pmatrix} = (-1)^{j_1+j_2+j_3} \begin{pmatrix} j_1 & j_3 & j_2 \\ \kappa_1 & \kappa_3 & \kappa_2 \end{pmatrix} \tag{6}$$

and hence

$$\begin{pmatrix} j_1 & j_2 & j_3 \\ \kappa_1 & \kappa_2 & \kappa_3 \end{pmatrix} = \begin{pmatrix} j_2 & j_3 & j_1 \\ \kappa_2 & \kappa_3 & \kappa_1 \end{pmatrix} = \begin{pmatrix} j_3 & j_1 & j_2 \\ \kappa_3 & \kappa_1 & \kappa_2 \end{pmatrix}, \tag{6a}$$

i.e., that the three-j-symbols remain unchanged under an even permutation of their columns and are multiplied by $(-1)^{j_1+j_2+j_3}$ for an odd permutation. According to the lemma of Reference 6 and our convention for the values of $(-1)^j$ the $(-1)^{j_1+j_2+j_3} = \pm 1$, unless the three-$j$-symbols in (6) and (6a) vanish.

In order to obtain (6) and (6a), let us first choose a triple $\kappa_{10}, \kappa_{20}, \kappa_{30}$ of κ for which the three-j-symbol does not vanish. Let us multiply then the three-j-symbols of $j_1 j_2 j_3$ by a factor of modulus 1 which makes

$$\begin{pmatrix} j_1 & j_2 & j_3 \\ \kappa_{10} & \kappa_{20} & \kappa_{30} \end{pmatrix} \tag{*}$$

real and positive. We can then multiply all the symbols

$$\begin{pmatrix} j_2 & j_1 & j_3 \\ \kappa_2 & \kappa_1 & \kappa_3 \end{pmatrix} \tag{†}$$

with a second factor of modulus 1 so that the first part of (6) shall be correct for the triple $\kappa_{10}, \kappa_{20}, \kappa_{30}$. This is evidently possible if j_2 and j_1 are different, because in this case the factors by which (*) and (†) can be multiplied are entirely independent of each other. For $j_1 = j_2$ on the other hand, (6) only expresses the fact that the representation j_3 is in the symmetric or antisymmetric part of the Kronecker product of j_1 with itself, for even and odd values of $j_3 - 2j_1$, respectively. In a similar way, all the equations (6) and (6a) can be made to hold for a single triple $\kappa_1 = \kappa_{10}, \kappa_2 = \kappa_{20}, \kappa_3 = \kappa_{30}$ for which (*) does not vanish. It follows then from (5)—its left side is clearly unaltered if one interchanges the indices 1 and 2—that (6) and (6a) hold for all $\kappa_1, \kappa_2, \kappa_3$. In the following, we shall not necessarily retain the normalization which makes (*) real but may multiply all three-j-symbols with factors of modulus unity which depend symmetrically on j_1, j_2, j_3. This will leave equations (6) and (6a) valid. In the case of the three-dimensional rotation group, it will turn

out to be possible to choose all three-j-symbols real and this will be the normalization adopted.

An important special case of (2) is obtained by taking for j_2 the identical representation $j_2 = 0$. The Kronecker product of this with j_1 is of course j_1 itself so that j_3 must be j_1 also. Since j_2 has only one row, no summation over κ_2 is necessary. We shall denote the only row of $j_2 = 0$ by $\kappa_2 = 0$. It then follows from (2) that

$$[j]^{1/2}\begin{pmatrix} j & 0 & j \\ \kappa & 0 & \mu \end{pmatrix} = \begin{pmatrix} j & \\ \kappa & \mu \end{pmatrix} \tag{7}$$

is the unitary matrix which transforms the representation j into the conjugate complex form

$$\begin{bmatrix} j & R \\ \kappa & \lambda \end{bmatrix} = \begin{pmatrix} j & \\ \kappa & \mu \end{pmatrix}^* \begin{bmatrix} j & R \\ \mu & \nu \end{bmatrix}^* \begin{pmatrix} j & \\ \lambda & \nu \end{pmatrix}. \tag{8}$$

The unitary condition (4) and (4a) becomes

$$\begin{pmatrix} j & \\ \kappa & \mu \end{pmatrix}^* \begin{pmatrix} j & \\ \kappa & \mu' \end{pmatrix} = \delta_{\mu\mu'}; \qquad \begin{pmatrix} j & \\ \kappa' & \mu \end{pmatrix}^* \begin{pmatrix} j & \\ \kappa & \mu \end{pmatrix} = \delta_{\kappa'\kappa}, \tag{9}$$

and instead of (6) we have

$$\begin{pmatrix} j & \\ \kappa & \mu \end{pmatrix} = (-1)^{2j} \begin{pmatrix} j & \\ \mu & \kappa \end{pmatrix} \tag{10}$$

a result well known from the investigations of Frobenius and Schur.[1] Since $(-1)^{4j} = 1$ for all j, it is permissible to interchange the indices on *both* "one-j-symbols" in (8).

The one-j-symbol defined in (7) can be made to play the role of a metric tensor by which to raise and lower indices. Thus we can write

$$\begin{pmatrix} j_1 & j_2 & j_3 \\ \kappa_1 & \kappa_2 & \kappa_3 \end{pmatrix} \begin{pmatrix} j_3 & \\ \kappa_3 & \lambda_3 \end{pmatrix}^* = \begin{pmatrix} j_1 & j_2 & \lambda_3 \\ \kappa_1 & \kappa_2 & j_3 \end{pmatrix} \tag{10a}$$

and

$$\begin{pmatrix} j_1 & j_2 & j_3 \\ \lambda_1 & \lambda_2 & \lambda_3 \end{pmatrix} \begin{pmatrix} j & \\ \lambda_1 & \kappa_1 \end{pmatrix}^* \begin{pmatrix} j & \\ \lambda_2 & \kappa_2 \end{pmatrix}^* \begin{pmatrix} j & \\ \lambda_3 & \kappa_3 \end{pmatrix}^* = \begin{pmatrix} \kappa_1 & \kappa_2 & \kappa_3 \\ j_1 & j_2 & j_3 \end{pmatrix}. \tag{10b}$$

If one wants to use the notation of (10a) and (10b), defining covariant (straight) and contravariant (time inverted) three-j-symbols, one has to be careful when substituting numbers for the j and κ. It is necessary then to use a different type of number for the designation of the representations $j = 0, 1, 2, \ldots$ and the rows and columns of the representations $\lambda = 0, 1, 2, \ldots$.

Because of (9), one can lower an index by the conjugate complex of the metric

$$\begin{pmatrix} j_3 & \\ \kappa_3 & \lambda_3 \end{pmatrix} \begin{pmatrix} j_1 & j_2 & \lambda_3 \\ \kappa_1 & \kappa_2 & j_3 \end{pmatrix} = \begin{pmatrix} j_1 & j_2 & j_3 \\ \kappa_1 & \kappa_2 & \kappa_3 \end{pmatrix}. \tag{10c}$$

If one wanted to adopt this notation consistently, it would be appropriate to write

$$\begin{pmatrix} j & \\ \kappa & \lambda \end{pmatrix}^* = \begin{pmatrix} \kappa & \lambda \\ j & \end{pmatrix} \tag{10d}$$

whence the raising of the index in (10a) and (10b) would follow the natural rule. It should be remembered though that the metric tensor (7) is symmetric only for integer j; for half integer j it is skew symmetric. Hence the raising of an index must be done by using the left index of the metric tensor as summation index; in the lowering process the right index must be used as summation index. Since the three-j-symbols will be all real in the case of the three-dimensional rotation group, the same will hold also for the metric or one-j-symbol. Hence raising and lowering of indices is done by the same tensor in this case; nevertheless, the order of the indices remains to be observed. If the indices refer to a half-integer j, the scalar products $a_\mu b^\mu$ and $a^\mu b_\mu$ have opposite signs.

The idea of using the contravariant and covariant notation for three-j-symbols is due to C. Herring. It will not be made full use of in the present article.

It follows from (8) and (9) that

$$\begin{bmatrix} j_1 & R \\ \kappa_1 & \lambda_1 \end{bmatrix}\begin{bmatrix} j_2 & R \\ \kappa_2 & \lambda_2 \end{bmatrix}\begin{bmatrix} j_3 & R \\ \kappa_3 & \lambda_3 \end{bmatrix}\begin{pmatrix} j_1 & \\ \lambda_1 & \nu_1 \end{pmatrix}^*\begin{pmatrix} j_2 & \\ \lambda_2 & \nu_2 \end{pmatrix}\begin{pmatrix} j_3 & \\ \lambda_3 & \nu_3 \end{pmatrix}^*$$

$$= \begin{pmatrix} j_1 & \\ \kappa_1 & \mu_1 \end{pmatrix}^*\begin{pmatrix} j_2 & \\ \kappa_2 & \mu_2 \end{pmatrix}^*\begin{pmatrix} j_3 & \\ \kappa_3 & \mu_3 \end{pmatrix}^*\begin{bmatrix} j_1 & R_1 \\ \mu_1 & \nu_1 \end{bmatrix}^*\begin{bmatrix} j_2 & R_2 \\ \mu_2 & \nu_2 \end{bmatrix}^*\begin{bmatrix} j_3 & R_3 \\ \mu_3 & \nu_3 \end{bmatrix}^*$$

and this equation gives with the aid of (5) when summed over all R

$$\begin{pmatrix} j_1 & j_2 & j_3 \\ \kappa_1 & \kappa_2 & \kappa_3 \end{pmatrix}^*\begin{pmatrix} j_1 & j_2 & j_3 \\ \lambda_1 & \lambda_2 & \lambda_3 \end{pmatrix}\begin{pmatrix} j_1 & \\ \lambda_1 & \nu_1 \end{pmatrix}^*\begin{pmatrix} j_2 & \\ \lambda_2 & \nu_2 \end{pmatrix}^*\begin{pmatrix} j_3 & \\ \lambda_3 & \nu_3 \end{pmatrix}^*$$

$$= \begin{pmatrix} j_1 & \\ \kappa_1 & \mu_1 \end{pmatrix}^*\begin{pmatrix} j_2 & \\ \kappa_2 & \mu_2 \end{pmatrix}^*\begin{pmatrix} j_3 & \\ \kappa_3 & \mu_3 \end{pmatrix}^*\begin{pmatrix} j_1 & j_2 & j_3 \\ \mu_1 & \mu_2 & \mu_3 \end{pmatrix}\begin{pmatrix} j_1 & j_2 & j_3 \\ \nu_1 & \nu_2 & \nu_3 \end{pmatrix}^*.$$

Dividing this by the first factor on the left side one sees that

$$\begin{pmatrix} j_1 & j_2 & j_3 \\ \lambda_1 & \lambda_2 & \lambda_3 \end{pmatrix}\begin{pmatrix} j_1 & \\ \lambda_1 & \nu_1 \end{pmatrix}^*\begin{pmatrix} j_2 & \\ \lambda_2 & \nu_2 \end{pmatrix}^*\begin{pmatrix} j_3 & \\ \lambda_3 & \nu_3 \end{pmatrix}^* = c\begin{pmatrix} j_1 & j_2 & j_3 \\ \nu_1 & \nu_2 & \nu_3 \end{pmatrix}^* \tag{11}$$

where c does not depend on the ν. One can bring by means of the orthogonality relations (9), the last three factors of (11) to the right side. The conjugate complex of the resulting equation is

$$(c^*)^{-1}\begin{pmatrix} j_1 & j_2 & j_3 \\ \lambda_1 & \lambda_2 & \lambda_3 \end{pmatrix}^* = \begin{pmatrix} j_1 & \\ \lambda_1 & \nu_1 \end{pmatrix}^*\begin{pmatrix} j_2 & \\ \lambda_2 & \nu_2 \end{pmatrix}^*\begin{pmatrix} j_3 & \\ \lambda_3 & \nu_3 \end{pmatrix}^*\begin{pmatrix} j_1 & j_2 & j_3 \\ \nu_1 & \nu_2 & \nu_3 \end{pmatrix}.$$

Interchanging herein the λ with the ν, one obtains (11) back again and

$$|c|^2 = (-1)^{2j_1+2j_2+2j_3} = 1. \tag{12}$$

One sees from (11) furthermore that c is a symmetric function of the j. Hence one can make $c = 1$ by multiplying the three-j-symbols by factors of modulus unity which depend only on the j and are symmetric functions of these. The only equation which could be changed by this renormalization is (7). One convinces oneself easily, however, that $c = 1$ automatically holds if $j_2 = 0$, provided that

$$\begin{pmatrix} & 0 & \\ 0 & & 0 \end{pmatrix} = 1 \tag{13}$$

i.e., if we choose the unit matrix for the matrix which transforms the unit matrix into its conjugate complex. One then has, instead of (11)

$$\begin{pmatrix} j_1 & j_2 & j_3 \\ \nu_1 & \nu_2 & \nu_3 \end{pmatrix}^* = \begin{pmatrix} j_1 & j_2 & j_3 \\ \lambda_1 & \lambda_2 & \lambda_3 \end{pmatrix} \begin{pmatrix} j_1 & \\ \lambda_1 & \nu_1 \end{pmatrix}^* \begin{pmatrix} j_2 & \\ \lambda_2 & \nu_2 \end{pmatrix}^* \begin{pmatrix} j_3 & \\ \lambda_3 & \nu_3 \end{pmatrix}^*. \tag{11a}$$

Because of (10) and (3), the order of the greek indices can be reversed in all one-j-symbols of (11a) simultaneously. It shows that *the fully covariant and fully contravariant components of a three-j-symbol are conjugate complex.* For the sake of future reference, we rewrite (2) by means of (8) and (11), (9), (10)

$$\begin{aligned}
&\begin{bmatrix} j_1 & R \\ \kappa_1 & \lambda_1 \end{bmatrix} \begin{bmatrix} j_2 & R \\ \kappa_2 & \lambda_2 \end{bmatrix} \\
&= \sum_{j_3} [j_3] \begin{pmatrix} j_1 & j_2 & j_3 \\ \mu_1 & \mu_2 & \mu_3 \end{pmatrix} \begin{pmatrix} j_1 & \\ \mu_1 & \kappa_1 \end{pmatrix}^* \begin{pmatrix} j_2 & \\ \mu_2 & \kappa_2 \end{pmatrix}^* \begin{pmatrix} j_3 & \\ \kappa_3 & \lambda_3 \end{pmatrix}^* \begin{bmatrix} j_3 & R \\ \mu_3 & \kappa_3 \end{bmatrix} \begin{pmatrix} j_1 & j_2 & j_3 \\ \lambda_1 & \lambda_2 & \lambda_3 \end{pmatrix} \\
&= \sum_{j_3} [j_3] \begin{pmatrix} \kappa_1 & \kappa_2 & j_3 \\ j_1 & j_2 & \mu_3 \end{pmatrix} \begin{bmatrix} j_3 & R \\ \mu_3 & \kappa_3 \end{bmatrix} \begin{pmatrix} j_3 & \\ \kappa_3 & \lambda_3 \end{pmatrix}^* \begin{pmatrix} j_1 & j_2 & j_3 \\ \lambda_1 & \lambda_2 & \lambda_3 \end{pmatrix}
\end{aligned} \tag{14}$$

which shows that the row-index of the representation should be a contravariant (upper) index.

The particular case of (4a) in which $j' = 0$, $j_1 = j_2$ is significant for many applications

$$\begin{pmatrix} j_1 & j & j_1 \\ \lambda_1 & \lambda & \lambda_2 \end{pmatrix} \begin{pmatrix} j_1 & \\ \lambda_1 & \lambda_2 \end{pmatrix}^* = [j_1]^{1/2} \delta_{j_0} \delta_{\lambda_0}, \tag{15}$$

Again, this can be written as

$$\begin{pmatrix} j_1 & \lambda_1 & j \\ \lambda_1 & j_1 & \lambda \end{pmatrix} = [j_1]^{1/2} \delta_{j_0} \delta_{\lambda_0}. \tag{15a}$$

It may be well to record here to what extent the indeterminate phases in our different symbols are fixed. Naturally, all the irreducible representations can be transformed by any unitary matrix, and this will affect the three-j-symbols and also (7). Even if the form of the representations is supposed to

be fixed, all one-j-symbols [except (13)] can be multiplied by a factor depending on j. The three-j-symbols will then be multiplied by the $-\frac{1}{2}$ power of the product of the corresponding three factors. If, however, all one-j-symbols (7) are supposed to be fixed, only the signs of the three-j-symbols can be changed and these only simultaneously for all three-j-symbols containing the same three j's.

3. *The Six-j-Symbols* (*Racah Coefficients*)

It is clear that some connection must exist between the three-j-symbols which corresponds to the associative law of ordinary multiplication. One can expand by (14) the product of three representation coefficients in three different ways, one of which is

$$\begin{aligned}
&\begin{bmatrix} j_1 & R \\ \kappa_1 & \lambda_1 \end{bmatrix}\begin{bmatrix} j_2 & R \\ \kappa_2 & \lambda_2 \end{bmatrix}\begin{bmatrix} j_3 & R \\ \kappa_3 & \lambda_3 \end{bmatrix} \\
&= \sum_j [j] \begin{pmatrix} j_1 & j_2 & j \\ \mu_1 & \mu_2 & \kappa \end{pmatrix}\begin{pmatrix} j_1 & j_2 & j \\ \lambda_1 & \lambda_2 & \eta \end{pmatrix}\begin{pmatrix} j_1 & \\ \mu_1 & \kappa_1 \end{pmatrix}^* \\
&\quad \cdot \begin{pmatrix} j_2 & \\ \mu_2 & \kappa_2 \end{pmatrix}^* \begin{pmatrix} j & \\ \lambda & \eta \end{pmatrix}^* \begin{bmatrix} j & R \\ \kappa & \lambda \end{bmatrix}\begin{bmatrix} j_3 & R \\ \kappa_3 & \lambda_3 \end{bmatrix} \\
&= \sum_{jj_4} [j][j_4] \begin{pmatrix} j_1 & j_2 & j \\ \mu_1 & \mu_2 & \kappa \end{pmatrix}\begin{pmatrix} j_1 & j_2 & j \\ \lambda_1 & \lambda_2 & \eta \end{pmatrix}\begin{pmatrix} j & j_3 & j_4 \\ \mu & \mu_3 & \mu_4 \end{pmatrix}\begin{pmatrix} j & j_3 & j_4 \\ \lambda & \lambda_3 & \lambda_4 \end{pmatrix} \\
&\quad \cdot \begin{pmatrix} j & \\ \lambda & \eta \end{pmatrix}^* \begin{pmatrix} j & \\ \mu & \kappa \end{pmatrix}^* \begin{pmatrix} j_1 & \\ \mu_1 & \kappa_1 \end{pmatrix}^* \begin{pmatrix} j_2 & \\ \mu_2 & \kappa_2 \end{pmatrix}^* \begin{pmatrix} j_3 & \\ \mu_3 & \kappa_3 \end{pmatrix}^* \begin{pmatrix} j_4 & \\ \eta_4 & \lambda_4 \end{pmatrix}^* \begin{bmatrix} j_4 & R \\ \mu_4 & \eta_4 \end{bmatrix}.
\end{aligned} \tag{15b}$$

The left side of (15b) is symmetric in 1 and 3 and the right side must also remain unchanged if one interchanges these two indices. This gives an equation which is valid for all R. Since the representation coefficients are orthogonal, their coefficients must be equal on the two sides of the equation. After some obvious simplifications, this assumes the form

$$\begin{aligned}
&\sum_{\substack{j \\ \kappa\mu}} [j] \begin{pmatrix} j_1 & j_2 & j \\ \mu_1 & \mu_2 & \kappa \end{pmatrix}\begin{pmatrix} j_3 & j_4 & j \\ \mu_3 & \mu_4 & \mu \end{pmatrix}\begin{pmatrix} j & \\ \mu & \kappa \end{pmatrix}^* \\
&\qquad \cdot \sum_{\eta\lambda} \begin{pmatrix} j_1 & j_2 & j \\ \lambda_1 & \lambda_2 & \eta \end{pmatrix}\begin{pmatrix} j_3 & j_4 & j \\ \lambda_3 & \lambda_4 & \lambda \end{pmatrix}\begin{pmatrix} j & \\ \lambda & \eta \end{pmatrix}^* \\
&= \sum_{j\kappa\mu} [j] \begin{pmatrix} j_3 & j_2 & j \\ \mu_3 & \mu_2 & \kappa \end{pmatrix}\begin{pmatrix} j_1 & j_4 & j \\ \mu_1 & \mu_4 & \mu \end{pmatrix}\begin{pmatrix} j & \\ \mu & \kappa \end{pmatrix}^* \\
&\qquad \cdot \sum_{\eta\lambda} \begin{pmatrix} j_3 & j_2 & j \\ \lambda_3 & \lambda_2 & \eta \end{pmatrix}\begin{pmatrix} j_1 & j_4 & j \\ \lambda_1 & \lambda_4 & \lambda \end{pmatrix}\begin{pmatrix} j & \\ \lambda & \eta \end{pmatrix}^*.
\end{aligned} \tag{16}$$

All summations occurring in (16) are so indicated. Equations (16) form a set of linear equations for the quantities appearing on the left side after the dot. One can solve them by multiplying (16) with

$$\begin{pmatrix} j_1 & j_2 & j' \\ \mu_1 & \mu_2 & \kappa' \end{pmatrix}^* \begin{pmatrix} j_3 & j_4 & j' \\ \mu_3 & \mu_4 & \mu' \end{pmatrix}^* \begin{pmatrix} j' & \\ \mu' & \kappa' \end{pmatrix}$$

and summing over $\mu_1\mu_2\mu_3\mu_4\kappa'\mu'$. The quantities after the dot on the left will appear then as linear combinations of the expressions after the dot on the right side of (16), with coefficients which depend on all j but are independent of the λ. We can, further, interchange the lower indices of the one-j-symbols by means of (10) and write,

$$\begin{pmatrix} j_1 & j_2 & j' \\ \lambda_1 & \lambda_2 & \eta \end{pmatrix} \begin{pmatrix} j_3 & j_4 & j' \\ \lambda_3 & \lambda_4 & \lambda \end{pmatrix} \begin{pmatrix} j' & \\ \eta & \lambda \end{pmatrix}^*$$
$$= \sum_j (-1)^{2j_4}[j] \begin{Bmatrix} j_1 & j_2 & j' \\ j_3 & j_4 & j \end{Bmatrix} \begin{pmatrix} j_3 & j_2 & j \\ \lambda_3 & \lambda_2 & \eta \end{pmatrix} \begin{pmatrix} j_1 & j_4 & j \\ \lambda_1 & \lambda_4 & \lambda \end{pmatrix} \begin{pmatrix} j & \\ \eta & \lambda \end{pmatrix}^*, \quad (17)$$

where the summation convention is again used for the greek indices. The coefficients given by the curly brackets will be called six-j-symbols.

It is best now to bring the three-j-symbols of the right side of (17) over to the left side. This can be done by means of the orthogonality relations (4) and (9), multiplying (17) with

$$\begin{pmatrix} j_3 & j_2 & j'' \\ \lambda_3 & \lambda_2 & \eta'' \end{pmatrix}^* \begin{pmatrix} j_1 & j_4 & j'' \\ \lambda_1 & \lambda_4 & \lambda'' \end{pmatrix}^* \begin{pmatrix} j'' & \\ \eta'' & \lambda'' \end{pmatrix}.$$

This gives, after a slight change in notation and by means of (11a)

$$\begin{Bmatrix} j_1 & j_2 & j_3 \\ l_1 & l_2 & l_3 \end{Bmatrix} = \begin{pmatrix} j_1 & j_2 & j_3 \\ \iota_1 & \iota_2 & \iota_3 \end{pmatrix} \begin{pmatrix} j_1 & l_2 & l_3 \\ \kappa_1 & \lambda_2 & \mu_3 \end{pmatrix} \begin{pmatrix} l_1 & j_2 & l_3 \\ \mu_1 & \kappa_2 & \lambda_3 \end{pmatrix} \begin{pmatrix} l_1 & l_2 & j_3 \\ \lambda_1 & \mu_2 & \kappa_3 \end{pmatrix}$$
$$\cdot \begin{pmatrix} j_1 & \\ \iota_1 & \kappa_1 \end{pmatrix}^* \begin{pmatrix} j_2 & \\ \iota_2 & \kappa_2 \end{pmatrix}^* \begin{pmatrix} j_3 & \\ \iota_3 & \kappa_3 \end{pmatrix}^* \begin{pmatrix} l_1 & \\ \lambda_1 & \mu_1 \end{pmatrix}^* \begin{pmatrix} l_2 & \\ \lambda_2 & \mu_2 \end{pmatrix}^* \begin{pmatrix} l_3 & \\ \lambda_3 & \mu_3 \end{pmatrix}^*$$
$$= \begin{pmatrix} \kappa_1 & \kappa_2 & \kappa_3 \\ j_1 & j_2 & j_3 \end{pmatrix} \begin{pmatrix} j_1 & \mu_2 & l_3 \\ \kappa_1 & l_2 & \mu_3 \end{pmatrix} \begin{pmatrix} l_1 & j_2 & \mu_3 \\ \mu_1 & \kappa_2 & l_3 \end{pmatrix} \begin{pmatrix} \mu_1 & l_2 & j_3 \\ l_1 & \mu_2 & \kappa_3 \end{pmatrix}. \quad (18)$$

It follows from (18) and (6a) that

$$\begin{Bmatrix} j_1 & j_2 & j_3 \\ l_1 & l_2 & l_3 \end{Bmatrix} = \begin{Bmatrix} j_2 & j_3 & j_1 \\ l_2 & l_3 & l_1 \end{Bmatrix} = \begin{Bmatrix} j_3 & j_1 & j_2 \\ l_3 & l_1 & l_2 \end{Bmatrix} \quad \textbf{(19)}$$

and with (6) also

$$\begin{Bmatrix} j_1 & j_2 & j_3 \\ l_1 & l_2 & l_3 \end{Bmatrix} = \begin{Bmatrix} j_2 & j_1 & j_3 \\ l_2 & l_1 & l_3 \end{Bmatrix} = \begin{Bmatrix} j_1 & j_3 & j_2 \\ l_1 & l_3 & l_2 \end{Bmatrix} \text{etc.} \quad \textbf{(19a)}$$

so that the columns of a six-j-symbol can be interchanged arbitrarily. It follows in a similar way that

$$\begin{Bmatrix} j_1 & j_2 & j_3 \\ l_1 & l_2 & l_3 \end{Bmatrix} = \begin{Bmatrix} l_1 & l_2 & j_3 \\ j_1 & j_2 & l_3 \end{Bmatrix} = \begin{Bmatrix} l_1 & j_2 & l_3 \\ j_1 & l_2 & j_3 \end{Bmatrix} = \begin{Bmatrix} j_1 & l_2 & l_3 \\ l_1 & j_2 & j_3 \end{Bmatrix}, \tag{19b}$$

i.e., that it is possible to interchange in two columns simultaneously the upper representation with the lower. The symmetry of the six-j-symbols corresponds to that of a regular tetrahedron; the representations in one column correspond to opposite edges. Those which occur in the same three-j-symbol in (18) meet at one corner. The six-j-symbol vanishes unless the Kronecker product of any two edges meeting at one corner contains the representation corresponding to the third edge of that corner.

By substituting (11a) into the conjugate complex of (18) one easily verifies that it is equal to (18) itself, i.e., the six-j-symbols are *real.*

Interchanging now the indices 1 and 3 in (17), the expressions on the right side of (17) will appear as linear combinations of the expressions on the left side. Since these are, as functions of λ_1, λ_2, λ_3, λ_4 evidently linearly independent, it is clear that the two systems of coefficients form reciprocal matrices:

$$\sum_j [j] \begin{Bmatrix} j_1 & j_2 & j' \\ j_3 & j_4 & j \end{Bmatrix} \cdot [j''] \begin{Bmatrix} j_3 & j_2 & j \\ j_1 & j_4 & j'' \end{Bmatrix} = \delta_{j'j''}. \tag{20}$$

Since the first and last column in the second six-j-symbol can be turned upside down, (20) expresses the fact that

$$R_{j'j} = ([j][j'])^{1/2} \begin{Bmatrix} j_1 & j_2 & j' \\ j_3 & j_4 & j \end{Bmatrix} \tag{20a}$$

is a real orthogonal matrix. The possible values of j' are those which are contained in both Kronecker products $j_1 \times j_2$ and $j_3 \times j_4$. The possible values of j are the common parts of $j_1 \times j_4$ and $j_3 \times j_2$. According to (20a), the number of the common irreducible parts of $j_1 \times j_2$ and $j_3 \times j_4$ must be equal with the number of common irreducible parts of $j_1 \times j_4$ and $j_2 \times j_3$—a fact which can be also directly verified.

One can interchange the first two columns of the first three-j-symbol on the right side of (17). This introduces a factor $(-1)^{j+j_2+j_3}$. After this, one can apply (17) again to the right side of the resulting equation which becomes hereupon

$$\sum_{jj''} (-1)^{j+j_2+j_3} [j][j''] \cdot \begin{Bmatrix} j_1 & j_2 & j' \\ j_3 & j_4 & j \end{Bmatrix} \begin{Bmatrix} j_2 & j_3 & j \\ j_1 & j_4 & j'' \end{Bmatrix} \begin{pmatrix} j_1 & j_3 & j'' \\ \lambda_1 & \lambda_3 & \eta \end{pmatrix} \begin{pmatrix} j_2 & j_4 & j'' \\ \lambda_2 & \lambda_4 & \lambda \end{pmatrix} \begin{pmatrix} j'' & \\ \eta & \lambda \end{pmatrix}^* .$$

On the other hand, one can first interchange the first two columns in the first three-j-symbol on the left side of (17) and then use (17). This gives

$$\sum_{j''}(-1)^{2j_4+j_1+j_2+j'}[j'']\begin{Bmatrix} j_2 & j_1 & j' \\ j_3 & j_4 & j'' \end{Bmatrix}\begin{Bmatrix} j_3 & j_1 & j'' \\ \lambda_3 & \lambda_1 & \eta \end{Bmatrix}\begin{pmatrix} j_2 & j_4 & j'' \\ \lambda_2 & \lambda_4 & \lambda \end{pmatrix}\begin{pmatrix} j'' & \\ \eta & \lambda \end{pmatrix}^*.$$

Comparing this with the preceding expression yields

$$\sum_{j}(-1)^{j+j'+j''}[j]\begin{Bmatrix} j_1 & j_2 & j' \\ j_3 & j_4 & j \end{Bmatrix}\begin{Bmatrix} j_2 & j_3 & j \\ j_1 & j_4 & j'' \end{Bmatrix} = \begin{Bmatrix} j_2 & j_1 & j' \\ j_3 & j_4 & j'' \end{Bmatrix}. \tag{21}$$

Bringing the right side over to the left by (20) gives a more symmetric form to this

$$\sum_{ll'}(-1)^{j+l+l'}[j][l][l']\begin{Bmatrix} j_1 & j_2 & j \\ j_3 & j_4 & l \end{Bmatrix}\begin{Bmatrix} j_1 & j_4 & l \\ j_2 & j_3 & l' \end{Bmatrix}\begin{Bmatrix} j_1 & j_3 & l' \\ j_4 & j_2 & j' \end{Bmatrix} = \delta_{jj'}. \tag{21a}$$

The arrangement of the j_k in (21a) is as follows: j_1 is always in the upper left corner. It would be equally possible, however, to bring any other j_k to that place. Otherwise, the arrangement of the j_k in all three six-j-symbols is different from each other but such that both six-j-symbols, in which a summation index (l or l') occurs, limit the values of this index to the same set. This is, for l, e.g., the common part $j_1 \times j_4$ and $j_2 \times j_3$. Similarly, j and j' are also limited to the same set. As special cases of the above formulas we may note the following ones. Writing 0 for l_1 in (18), i.e., substituting the unit representation therefor, gives zero, unless $l_3 = j_2$, $l_2 = j_3$. In this case the summation can be carried out by (4b), (7), (9), and (11a) and gives

$$\begin{Bmatrix} j_1 & j_2 & j_3 \\ 0 & j_3 & j_2 \end{Bmatrix} = (-1)^{j_1+j_2+j_3}([j_2][j_3])^{-1/2}(j_1 j_2 j_3). \tag{18b}$$

Writing now 0 for j'' in (20) we obtain in case of $j_1 = j_2$, $j_3 = j_4$

$$\sum_{j}(-1)^{j_1+j_3+j}[j]\begin{Bmatrix} j_1 & j_1 & j' \\ j_3 & j_3 & j \end{Bmatrix} = \delta_{j'0}([j_1][j_3])^{1/2}. \tag{20b}$$

Writing 0 for j'' in (21) gives in a similar way

$$\sum_{j}[j]\begin{Bmatrix} j_1 & j_2 & j' \\ j_1 & j_2 & j \end{Bmatrix} = (-1)^{2j_1+2j_2}. \tag{21b}$$

We finally note a formula which immediately follows from (17) and which will be useful for the calculation of the three-j-symbols. One obtains it by bringing a three-j-symbol in (17) from the right to the left side by (11a) and (4a). After changing the notation somewhat, one has

$$\begin{Bmatrix} j_1 & j_2 & j_3 \\ l_1 & l_2 & l_3 \end{Bmatrix}\begin{pmatrix} j_1 & j_2 & j_3 \\ \iota_1 & \iota_2 & \iota_3 \end{pmatrix}$$
$$= \begin{pmatrix} j_1 & l_2 & l_3 \\ \iota_1 & \lambda_2 & \eta_3 \end{pmatrix}\begin{pmatrix} l_1 & j_2 & l_3 \\ \eta_1 & \iota_2 & \lambda_3 \end{pmatrix}\begin{pmatrix} l_1 & l_2 & j_3 \\ \lambda_1 & \eta_2 & \iota_3 \end{pmatrix}\begin{pmatrix} & l_1 & \\ \lambda_1 & & \eta_1 \end{pmatrix}^*\begin{pmatrix} & l_2 & \\ \lambda_2 & & \eta_2 \end{pmatrix}^*\begin{pmatrix} & l_3 & \\ \lambda_3 & & \eta_3 \end{pmatrix}^*$$
$$= \begin{pmatrix} j_1 & \eta_2 & l_3 \\ \iota_1 & l_2 & \eta_3 \end{pmatrix}\begin{pmatrix} l_1 & j_2 & \eta_3 \\ \eta_1 & \iota_2 & l_3 \end{pmatrix}\begin{pmatrix} \eta_1 & l_2 & j_3 \\ l_1 & \eta_2 & \iota_3 \end{pmatrix}. \tag{22}$$

4. *The Metric in Representation Space*

If one subjects the representation j_1 to a unitary transformation

$$\left[\begin{bmatrix} j_1 & R \\ \mu & \nu \end{bmatrix}\right] = u^*_{\kappa\mu} \begin{bmatrix} j_1 & R \\ \kappa & \lambda \end{bmatrix} u_{\lambda\nu} \tag{23}$$

all the preceding formulae remain valid if the three-j-symbols are replaced at the same time by

$$\left(\begin{pmatrix} j_1 & j_2 & j_3 \\ \mu & \kappa_2 & \kappa_3 \end{pmatrix}\right) = \begin{pmatrix} j_1 & j_2 & j_3 \\ \kappa & \kappa_2 & \kappa_3 \end{pmatrix} u_{\kappa\mu} \qquad (j_2, j_3 \neq j_1). \tag{23a}$$

The one-j-symbols have to be replaced by

$$\left(\begin{pmatrix} j & \\ \mu & \nu \end{pmatrix}\right) = \begin{pmatrix} j & \\ \kappa & \lambda \end{pmatrix} u_{\kappa\mu} u_{\lambda\nu}. \tag{23b}$$

This shows that the index κ_1 of three-j-symbol $\begin{pmatrix} j_1 & j_2 & j_3 \\ \kappa_1 & \kappa_2 & \kappa_3 \end{pmatrix}$ has indeed the character of the index of a vector in the unitary representation space of the representation j_1. The same holds, naturally, of the other greek indices in three-j-symbols. The one-j-symbol on the other hand, is a tensor, symmetric or antisymmetric, depending on the integer or half integer character of j. The possibility to use the one-j-symbol as a metric tensor is connected with the possibility of introducing, in the spaces of the irreducible representations, the operation of the time inversion. In fact, if ψ_λ are the components of a vector in representation space then $\phi_\lambda = (\psi^\lambda)^*$ are the components of the vector obtained by time inversion where

$$\psi^\lambda = \psi_\kappa \begin{pmatrix} j & \\ \kappa & \lambda \end{pmatrix}^*. \tag{24}$$

Since the operation of time inversion is independent of the choice of coordinates in the representation space and invariant under the transformations of the underlying group, which are all spatial, the transition to the contravariant components must also have these properties. This can be verified also directly and, unless one is willing to accept the above arguments, such a verification is necessary to establish the fact that the calculus with co- and contravariant components is consistent. As far as the group transformations are concerned, this is a consequence of (8), but the verification will not be carried out in detail. If the components of the vector ψ in the second coordinate system are $\bar{\psi}_\nu$, one has

$$\bar{\psi}^\nu = \bar{\psi}_\mu \left(\begin{pmatrix} j & \\ \mu & \nu \end{pmatrix}\right)^* = \psi_{\kappa'} u_{\kappa'\mu} \begin{pmatrix} j & \\ \kappa & \lambda \end{pmatrix}^* u^*_{\kappa\mu} u^*_{\lambda\nu} = \psi_\kappa \begin{pmatrix} j & \\ \kappa & \lambda \end{pmatrix}^* u^*_{\lambda\nu} = \psi^\lambda u^*_{\lambda\nu}. \tag{24a}$$

This gives the transformation law for the contravariant components. One easily verifies now that the contraction of two vectors ψ and ϕ

$$\bar{\psi}^{\nu}\bar{\phi}_{\nu}=\psi^{\lambda}u^{*}_{\lambda\nu}\phi_{\kappa}u_{\kappa\nu}=\psi^{\kappa}\phi_{\kappa} \tag{25}$$

is independent of the coordinate system. This assures the consistency of the calculus. It should be noted, however, that for half integer j the antisymmetric nature of the one-j-symbol renders $\psi^{\kappa}\phi_{\kappa}=-\psi_{\kappa}\phi^{\kappa}$.

The question of invariants with respect to coordinate transformations in representation space naturally arises. One such invariant is clearly

$$\begin{pmatrix} j_1 & j_2 & j_3 \\ \kappa_1 & \kappa_2 & \kappa_3 \end{pmatrix}\begin{pmatrix} \kappa_1 & \kappa_2 & \kappa_3 \\ j_1 & j_2 & j_3 \end{pmatrix}.$$

However, the value of this is given by (4b) to be $(j_1j_2j_3)$. It is reasonable to try, next, combinations of the form

$$\begin{pmatrix} j_1 & \kappa_1 & j_2 \\ \kappa_1 & j_1 & \kappa_2 \end{pmatrix}\begin{pmatrix} \kappa_2 & j_3 & \kappa_3 \\ j_2 & \kappa_3 & j_3 \end{pmatrix}.$$

However, these vanish, on account of (15a), unless $j_2=0$, i.e., unless the three-j-symbol is actually a one-j-symbol. In this case, the expression is again trivial.

The simplest nontrivial invariants with respect to coordinate transformations in the representation spaces are made up of four three-j-symbols; (18) shows that the six-j-symbol is such an invariant. In view of (4a), one easily convinces oneself that it is the most general invariant consisting of only four three-j-symbols which cannot be fully contracted. This explains the significance of the six-j-symbols.

5. *Further Properties of the Six-j-Symbols*

One can derive a relation for the six-j-symbols which is similar both in its derivation and its applications to Eq. (17), valid for the three-j-symbols. For this purpose, one has to consider such combinations of three-j-symbols which can be expressed in different ways by a series of three-j-symbols. In (15b), we started from a combination of representation coefficients which we expressed in two ways as a series of representation coefficients. Comparison of coefficients of the two series gave us a relation for the three-j-symbols and will give us a relation for six-j-symbols in this section.

In

$$\begin{pmatrix} l_1 & k_2 & j_3 \\ \lambda_1 & \kappa_2 & \iota_3 \end{pmatrix}\begin{pmatrix} j_1 & l_2 & k_3 \\ \iota_1 & \lambda_2 & \kappa_3 \end{pmatrix}\begin{pmatrix} k_1 & j_2 & l_3 \\ \kappa_1 & \iota_2 & \lambda_3 \end{pmatrix}\begin{pmatrix} l_1 & l_2 & l_3 \\ \lambda_1' & \lambda_2' & \lambda_3' \end{pmatrix}\begin{pmatrix} k_1 & k_2 & k_3 \\ \kappa_1' & \kappa_2' & \kappa_3' \end{pmatrix}$$
$$\cdot\begin{pmatrix} & l_1 & \\ \lambda_1 & & \lambda_1' \end{pmatrix}^{*}\begin{pmatrix} & l_2 & \\ \lambda_2 & & \lambda_2' \end{pmatrix}^{*}\begin{pmatrix} & l_3 & \\ \lambda_3 & & \lambda_3' \end{pmatrix}^{*}\begin{pmatrix} & k_1 & \\ \kappa_1 & & \kappa_1' \end{pmatrix}^{*}\begin{pmatrix} & k_2 & \\ \kappa_2 & & \kappa_2' \end{pmatrix}^{*}\begin{pmatrix} & k_3 & \\ \kappa_3 & & \kappa_3' \end{pmatrix}^{*} \tag{26}$$

(17) can be applied to the product of the third and fourth and to the product of the second and fifth three-j-symbols (together with the corresponding one-j-symbols). One obtains in this way

$$\sum_{lk} \begin{pmatrix} l_1 & k_2 & j_3 \\ \lambda_1 & \kappa_2 & \iota_3 \end{pmatrix} (-1)^{2k_1+2k_2+2l_2}[l][k] \begin{Bmatrix} k_1 & j_2 & l_3 \\ l_1 & l_2 & l \end{Bmatrix} \begin{pmatrix} l_1 & j_2 & l \\ \lambda_1' & \iota_2 & \lambda \end{pmatrix} \begin{pmatrix} k_1 & l_2 & l \\ \kappa_1' & \lambda_2 & \lambda \end{pmatrix}^*$$
$$\cdot \begin{Bmatrix} j_1 & l_2 & k_3 \\ k_1 & k_2 & k \end{Bmatrix} \begin{pmatrix} k_1 & l_2 & k \\ \kappa_1' & \lambda_2 & \kappa \end{pmatrix} \begin{pmatrix} j_1 & k_2 & k \\ \iota_1 & \kappa_2' & \kappa' \end{pmatrix} \begin{pmatrix} & k & \\ \kappa & & \kappa' \end{pmatrix}^* \begin{pmatrix} & l_1 & \\ \lambda_1 & & \lambda_1' \end{pmatrix}^* \begin{pmatrix} & k_2 & \\ \kappa_2 & & \kappa_2' \end{pmatrix}^*$$

where (11a) also has been made use of. This expression can be contracted by means of the orthogonality relation (4a) after which only the summation over l remains. The resulting expression has two three-j-symbols in which the last upper index is l. The product of these can be transformed again by (17) which gives the expression

$$\sum_{lj} (-1)^{2k_1+2l_2}[l][j] \begin{Bmatrix} k_1 & j_2 & l_3 \\ l_1 & l_2 & l \end{Bmatrix} \begin{Bmatrix} j_1 & l_2 & k_3 \\ k_1 & k_2 & l \end{Bmatrix} \begin{Bmatrix} l_1 & j_2 & l \\ j_1 & k_2 & j \end{Bmatrix}$$
$$\cdot \begin{pmatrix} l_1 & k_2 & j_3 \\ \lambda_1 & \kappa_2 & \iota_3 \end{pmatrix} \begin{pmatrix} j_1 & j_2 & j \\ \iota_1 & \iota_2 & \iota \end{pmatrix} \begin{pmatrix} l_1 & k_2 & j \\ \lambda_1' & \kappa_2' & \iota' \end{pmatrix} \begin{pmatrix} & j & \\ \iota & & \iota' \end{pmatrix}^* \begin{pmatrix} & l_1 & \\ \lambda_1 & & \lambda_1' \end{pmatrix}^* \begin{pmatrix} & k_2 & \\ \kappa_2 & & \kappa_2' \end{pmatrix}^*.$$

This can be contracted again by (11a) and (4a) and the summation over j drops out hereupon. One obtains in this way for (26) the left side of (26a)

$$\sum_{l} (-1)^{2l}[l] \begin{Bmatrix} j_1 & l_2 & k_3 \\ k_1 & k_2 & l \end{Bmatrix} \begin{Bmatrix} k_1 & j_2 & l_3 \\ l_1 & l_2 & l \end{Bmatrix} \begin{Bmatrix} l_1 & k_2 & j_3 \\ j_1 & j_2 & l \end{Bmatrix} \begin{pmatrix} j_1 & j_2 & j_3 \\ \iota_1 & \iota_2 & \iota_3 \end{pmatrix}$$
$$= \sum_{l} (-1)^{2l}[l] \begin{Bmatrix} j_1 & l_2 & k_3 \\ l & j_2 & j_3 \end{Bmatrix} \begin{Bmatrix} k_1 & j_2 & l_3 \\ l & k_2 & k_3 \end{Bmatrix} \begin{Bmatrix} l_1 & k_2 & j_3 \\ l & l_2 & l_3 \end{Bmatrix} \begin{pmatrix} j_1 & j_2 & j_3 \\ \iota_1 & \iota_2 & \iota_3 \end{pmatrix}. \quad (26a)$$

The right side results from the remark that (26) is invariant under a cyclic permutation of the indices 1, 2, 3. The last factor in (26a) can be dropped, of course. If one brings one of the six-j-symbols over the left side to the right side by means of the orthogonality relation (20) one obtains after some rewriting

$$\begin{Bmatrix} j_1 & j_2 & j_3 \\ l_1 & l_2 & l_3 \end{Bmatrix} \begin{Bmatrix} j_1 & j_2 & j_3 \\ k_1 & k_2 & k_3 \end{Bmatrix}$$
$$= \sum_{kl} (-1)^{2j_1+2l}[k][l] \begin{Bmatrix} j_1 & l_2 & l_3 \\ k & k_2 & k_3 \end{Bmatrix} \begin{Bmatrix} l_1 & j_2 & l_3 \\ k & k_2 & l \end{Bmatrix} \begin{Bmatrix} k_1 & j_2 & k_3 \\ k & l_2 & l \end{Bmatrix} \begin{Bmatrix} k_1 & k_2 & j_3 \\ l_1 & l_2 & l \end{Bmatrix}. \quad (26b)$$

This equation is analogous to (22) and will be used for the calculation of the six-j-symbols. Note that the exponent of -1 could be replaced by $2j_3 + 2k$ since $2j_1 + 2l + 2j_3 + 2k$ is necessarily even, since this holds for $2j_1 + 2j_2 + 2j_3$, for $2l + 2j_2 + 2k$ and, of course, for $4j_2$.

6. *Expressions for Six-j-Symbols in Terms of Group Integrals*

It was mentioned before that if one subjects the representation j to a unitary transformation with the matrix u, the three-j-symbols will transform according to

$$\begin{pmatrix} j_1 & j_2 & j \\ \lambda_1 & \lambda_2 & \lambda \end{pmatrix} \to \begin{pmatrix} j_1 & j_2 & j \\ \lambda_1 & \lambda_2 & \eta \end{pmatrix} u_{\eta\lambda}.$$

On the other hand, the one-j-symbols transform according to

$$\begin{pmatrix} & j & \\ \kappa & & \lambda \end{pmatrix} \to \begin{pmatrix} & j & \\ \mu & & \nu \end{pmatrix} u_{\mu\kappa} u_{\nu\lambda}.$$

We saw that, as a result of these expressions, the six-j-symbols are invariant under a unitary transformation of the underlying representations. On the other hand, a change in the sign in one of the four three-j-symbols of (18a) (cf. the last paragraph of Section 2) will change the sign of the six-j-symbol also.

It appears reasonable, therefore, to consider such products of the six-j-symbols which remain unchanged if the sign of the three-j-symbols is changed. Such a combination is, first of all [cf. (18) and (11a)]

$$\begin{Bmatrix} j_1 & j_2 & j_3 \\ j_1' & j_2' & j_3' \end{Bmatrix}^2$$

$$= \begin{pmatrix} j_1 & j_2 & j_3 \\ \kappa_1 & \kappa_2 & \kappa_3 \end{pmatrix}^* \begin{pmatrix} j_1 & j_2' & j_3' \\ \kappa_1' & \mu_2 & \mu_3 \end{pmatrix}^* \begin{pmatrix} & j_1 & \\ \kappa_1 & & \kappa_1' \end{pmatrix} \begin{pmatrix} j_1' & j_2 & j_3' \\ \mu_1 & \kappa_2 & \mu_3 \end{pmatrix} \begin{pmatrix} j_1' & j_2' & j_3 \\ \mu_1' & \mu_2 & \kappa_3 \end{pmatrix} \begin{pmatrix} & j_1' & \\ \mu_1 & & \mu_1' \end{pmatrix}^*$$

$$\begin{pmatrix} j_1 & j_2 & j_3 \\ \iota_1 & \iota_2 & \iota_3 \end{pmatrix} \begin{pmatrix} j_1 & j_2' & j_3' \\ \iota_1' & \lambda_2 & \lambda_3 \end{pmatrix} \begin{pmatrix} & j_1 & \\ \iota_1 & & \iota_1' \end{pmatrix}^* \begin{pmatrix} j_1' & j_2 & j_3' \\ \lambda_1 & \iota_2 & \lambda_3 \end{pmatrix}^* \begin{pmatrix} j_1' & j_2' & j_3 \\ \lambda_1' & \lambda_2 & \iota_3 \end{pmatrix}^* \begin{pmatrix} & j_1' & \\ \lambda_1 & & \lambda_1' \end{pmatrix}$$

$$= \sum_{RSTU} h^{-4} \begin{bmatrix} j_1 & R \\ \kappa_1 & \iota_1 \end{bmatrix} \begin{bmatrix} j_2 & R \\ \kappa_2 & \iota^2 \end{bmatrix} \begin{bmatrix} j_3 & R \\ \kappa_3 & \iota_3 \end{bmatrix} \begin{bmatrix} j_1 & S \\ \kappa_1 & \iota_1 \end{bmatrix}^* \begin{bmatrix} j_2' & S \\ \mu_2 & \lambda_2 \end{bmatrix} \begin{bmatrix} j_3' & S \\ \mu_3 & \lambda_3 \end{bmatrix} \begin{bmatrix} j_1' & T \\ \lambda_1 & \mu_1 \end{bmatrix}$$

$$\cdot \begin{bmatrix} j_2 & T \\ \iota_2 & \kappa_2 \end{bmatrix} \begin{bmatrix} j_3' & T \\ \lambda_3 & \mu_3 \end{bmatrix} \begin{bmatrix} j_1' & U \\ \lambda_1 & \mu_1 \end{bmatrix}^* \begin{bmatrix} j_2' & U \\ \lambda_2 & \mu_2 \end{bmatrix} \begin{bmatrix} j_3 & U \\ \iota_3 & \kappa_3 \end{bmatrix}$$

$$= \sum_{RSTU} h^{-4} [j_1; RS^{-1}][j_2; RT][j_3; RU][j_1'; TU^{-1}][j_2'; SU][j_3'; ST]$$

where (5) and (8) have been used; $[j; X]$ is the character of the element X in the representation j. Changing the notation somewhat, this gives

$$\begin{Bmatrix} j_1 & j_2 & j_3 \\ j_1' & j_2' & j_3' \end{Bmatrix}^2 = h^{-3} \sum_{R_1 R_2 R_3} [j_1; R_1][j_2; R_2][j_3; R_3]$$
$$\cdot [j_1'; R_2 R_3^{-1}][j_2'; R_3 R_1^{-1}][j'; R_1 R_2^{-1}]. \quad (27)$$

The equations

$$\begin{Bmatrix} j_1 & j_2 & j' \\ j_1 & j_2 & j \end{Bmatrix} = \frac{(-1)^{2j}}{h^2} \sum_{RS} [j_1; SR][j_2; S^{-1}R][j; S][j'; R], \tag{27a}$$

$$\begin{Bmatrix} j_1 & j_2 & j' \\ j_3 & j_2 & j \end{Bmatrix} \begin{Bmatrix} j_1 & j_2 & j'' \\ j_3 & j_2 & j' \end{Bmatrix} \begin{Bmatrix} j_1 & j_2 & j \\ j_3 & j_2 & j'' \end{Bmatrix}$$
$$= \frac{(-1)^{2j_1}}{h^4} \sum_{RXYZ} [j_1; RY][j_2; RXYZR][j_3; RXZ][j; X][j'; Y][j''; Z] \tag{27b}$$

$$\begin{Bmatrix} j_1 & j_2 & j_3 \\ l_1 & l_2 & l_3 \end{Bmatrix} \begin{Bmatrix} j_1 & l_2 & l_3 \\ j_1 & j_2 & j_3 \end{Bmatrix} \begin{Bmatrix} l_1 & j_2 & l_3 \\ j_1 & j_2 & j_3 \end{Bmatrix} \begin{Bmatrix} l_1 & l_2 & j_3 \\ j_1 & j_2 & j_3 \end{Bmatrix}$$
$$= \frac{(-1)^{2l_1}}{h^5} \sum_{URVSW} [j_1; R][j_1; S][j_2; UVWRWS]$$
$$\cdot [j_3; R^{-1}UVWSV][l_1; U][l_2; V][l_3; W] \tag{27c}$$

follow in a similar way. The first of these formulas provides for finite groups a reasonably convenient way to determine the absolute value of the six-j-symbols. The remaining ones permit the determination of those signs of the six-j-symbols which cannot be chosen arbitrarily. In most cases, it will be easier to compare the six-j-symbols by the formulas of the preceding sections than by Eqs. (27), and use these for the evaluation of sums or integrals over the group. The left sides of (27b) and (27c) show a high degree of symmetry which is not manifest on the right sides. This shows that the character sums satisfy certain identities which, however, have not been explored.

II. The Three-Dimensional Rotation Group

7. *The Representations of the Three-Dimensional Rotation Group*

The following sections will be devoted to an application of the formulas of the preceding sections to the representations of the two-dimensional unimodular unitary group. For this purpose, we briefly summarize first the well known results concerning the representations of this group. We shall parametrize the group in the customary fashion, by writing the elements as a product of a diagonal, a real and another diagonal matrix. The group elements can be characterized in this way by three angles, α, β, γ, the variability domain of α extends from -2π to 2π, β varies from 0 to π, and γ from $-\pi$ to π. The group element corresponding to the angles α, β, γ is

$$Z_\alpha Y_\beta Z_\gamma = \begin{Vmatrix} e^{-i\alpha/2} & 0 \\ 0 & e^{i\alpha/2} \end{Vmatrix} \cdot \begin{Vmatrix} \cos\frac{1}{2}\beta & -\sin\frac{1}{2}\beta \\ \sin\frac{1}{2}\beta & \cos\frac{1}{2}\beta \end{Vmatrix} \cdot \begin{Vmatrix} e^{-i\gamma/2} & 0 \\ 0 & e^{i\gamma/2} \end{Vmatrix}. \tag{28}$$

The matrices in (28) form naturally a representation of their own group; it is the representation $j = \frac{1}{2}$. In conformity herewith, we denote the two

rows and columns of (28) by $-\frac{1}{2}$ and $\frac{1}{2}$. The representation j is the symmetric part of the $2j$ power of (28), defined as[2]

$$\begin{bmatrix} j & R \\ \kappa & \lambda \end{bmatrix} = \frac{[(j+\kappa)!(j-\kappa)!(j+\lambda)!(j-\lambda)!]^{1/2}}{(2j)!} \cdot \sum_{\substack{\nu_1+\cdots+\nu_{2j}=\kappa \\ \mu_1+\cdots+\mu_{2j}=\lambda}} \begin{bmatrix} \frac{1}{2} & R \\ \nu_1 & \mu_1 \end{bmatrix} \begin{bmatrix} \frac{1}{2} & R \\ \nu_2 & \mu_2 \end{bmatrix} \cdots \begin{bmatrix} \frac{1}{2} & R \\ \nu_{2j} & \mu_{2j} \end{bmatrix}. \quad (29)$$

where κ and λ can assume the values $-j, -j+1, \ldots, j-1, j$, and

$$\frac{(2j)!}{(j+\kappa)!(j-\kappa)!}$$

is the number of ways the $\nu_1, \ldots, \nu_{2j}$ can be picked so as to have $j+\kappa$ of the ν to be $+\frac{1}{2}$ and $j-\kappa$ to be $-\frac{1}{2}$, which makes their sum equal to κ. One can easily convince oneself that (29) forms a unitary representation of the group, for all values of $j = 0, \frac{1}{2}, 1, \frac{3}{2}, 2, \ldots$.

In particular, (29) gives for the representatives of the diagonal matrices

$$\begin{bmatrix} j & Z_a \\ \kappa & \lambda \end{bmatrix} = \frac{[(j+\kappa)!(j-\kappa)!(j+\lambda)!(j-\lambda)!]^{1/2}}{(2j)!} \cdot \sum_{\substack{\nu_1+\cdots+\nu_{2j}=\kappa \\ \mu_1+\cdots+\mu_{2j}=\lambda}} e^{i(\nu_1+\cdots+\nu_{2j})a} \delta_{\nu_1\mu_1} \cdots \delta_{\nu_{2j}\mu_{2j}}$$

$$= \frac{(j+\kappa)!(j-\kappa)!}{(2j)!} \sum_{\nu_1+\cdots+\nu_{2j}=\kappa} e^{i\kappa a} \delta_{\kappa\lambda} = \delta_{\kappa\lambda} e^{i\kappa a} \quad (30)$$

again diagonal matrices and the representatives of real matrices are, evidently, real matrices. These can be calculated to be[3]

$$\begin{bmatrix} j & Y_\beta \\ \kappa & \lambda \end{bmatrix} = \sum_\eta (-1)^{j+\lambda-\eta} \frac{[(j+\kappa)!(j-\kappa)!(j+\lambda)!(j-\lambda)!]^{1/2}}{(\eta-\kappa-\lambda)!\eta!(j+\kappa-\eta)!(j+\lambda-\eta)!} \cdot \cos^{2\eta-\kappa-\lambda}\tfrac{1}{2}\beta \sin^{2j+\kappa+\lambda-2\eta}\tfrac{1}{2}\beta. \quad (30a)$$

The summation over η has to be extended over those integers for which none of the factorials in the denominator becomes negative. We shall note the special case of $\kappa = \pm j$ of (30a). Evidently, the sum over the ν in (29) reduces in this case to the single term $\nu_1 = \nu_2 = \ldots = \nu_{2j} = \frac{1}{2}$ and all terms of the summation are equal. Hence

$$\begin{bmatrix} j & R \\ j & \lambda \end{bmatrix} = \left(\frac{(2j)!}{(j-\lambda)!(j+\lambda)!}\right)^{1/2} e^{ij\alpha} \cos^{j+\lambda}\tfrac{1}{2}\beta \sin^{j-\lambda}\tfrac{1}{2}\beta e^{i\lambda\gamma}. \quad (30b)$$

We note here that those unitary matrices which have the same characteristic values $e^{\frac{1}{2}i\phi}$, $e^{-\frac{1}{2}i\phi}$ form a class ϕ. Every class is represented among the

diagonal matrices Z_ϕ, in fact Z_ϕ and $Z_{-\phi}$ are in the same class. The character of the representation j can be easily obtained from (30)

$$[j;\phi]=\sum_{\kappa=-j}^{j} e^{i\kappa\phi}. \tag{31}$$

It follows from (31) easily that the Kronecker product of the representations j and j' contains as irreducible parts the representations

$$j\times j'=|j-j'|,|j-j'|+1,|j-j'|+2,\ldots,j+j'-1,j+j'. \tag{31a}$$

Evidently, the dimension of the representation j is

$$[j]=2j+1. \tag{31b}$$

The matrix which transforms the representation into the conjugate complex form must be, apart from an arbitrary factor which we shall choose to be 1, equal to

$$\begin{pmatrix} j & \\ \kappa & \lambda \end{pmatrix}=\begin{bmatrix} j & Y_{-\pi} \\ \kappa & \lambda \end{bmatrix}. \tag{32}$$

This follows from the fact that this matrix must commute with all real matrices of the representation and that it must transform those of the form (30) into their reciprocals. Both conditions are evidently satisfied by the matrix corresponding to Y_π. One sees from (28) and (29) easily that

$$\begin{pmatrix} j & \\ \kappa & \lambda \end{pmatrix}=(-1)^{j-\lambda}\delta_{\kappa,-\lambda}=(-1)^{j+\kappa}\delta_{\kappa,-\lambda}. \tag{32a}$$

The possibility to assume the form (32a) for the one-j-symbols can be easily seen without reference to the preceding developments if one assumes that the representations with integer j can be brought into a real form, those with half integer j cannot (i.e., that a representation is integer or half integer according to the integer or half integer character of its j). For integer j, in fact, the representation can be first assumed to be real in which case the unit matrix transforms it into the conjugate complex form. If we now transform the representation by a unitary matrix u^{-1}, it will be brought into the conjugate complex form by a further transformation with $u'u$. Denoting the matrix (32a) (which is symmetric for integer j) by J and the unit matrix by I, we can choose

$$u=\tfrac{1}{2}(1-i)(iI+J) \tag{32b}$$

and we have on account of $J^2=I$,

$$u'u=\tfrac{1}{4}(1-i)^2(-I+2iJ+I)=J.$$

For half integer j (32a) already has—apart from the order of rows and columns—the usual form.

From the fact that (32a) commutes with the matrices corresponding to Y_β it follows that

$$\begin{pmatrix} j & \\ \kappa & \lambda \end{pmatrix}\begin{bmatrix} j & Y_\beta \\ \lambda & -\mu \end{bmatrix} = \begin{bmatrix} j & Y_\beta \\ \kappa & \lambda \end{bmatrix}\begin{pmatrix} j & \\ \lambda & -\mu \end{pmatrix} \tag{33}$$

or

$$\begin{bmatrix} j & Y_\beta \\ -\kappa & -\mu \end{bmatrix} = (-1)^{\kappa-\mu}\begin{bmatrix} j & Y_\beta \\ \kappa & \mu \end{bmatrix}. \tag{33a}$$

The condition that Z_π transforms Y_β into its reciprocal gives, using the real orthogonal character of the matrix corresponding to Y,

$$\begin{bmatrix} j & Z_\pi \\ \kappa & \lambda \end{bmatrix}\begin{bmatrix} j & Y_\beta \\ \lambda & \mu \end{bmatrix} = \begin{bmatrix} j & Y_\beta \\ \lambda & \kappa \end{bmatrix}\begin{bmatrix} j & Z_\pi \\ \lambda & \mu \end{bmatrix}$$

or

$$\begin{bmatrix} j & Y_\beta \\ \kappa & \mu \end{bmatrix} = (-1)^{\kappa-\mu}\begin{bmatrix} j & Y_\beta \\ \mu & \kappa \end{bmatrix} \tag{33b}$$

and

$$\begin{bmatrix} j & Y_\beta \\ \kappa & \mu \end{bmatrix} = \begin{bmatrix} j & Y_\beta \\ -\mu & -\kappa \end{bmatrix}. \tag{33c}$$

One has, furthermore, for $Y_\beta = Y_\pi(Y_{\pi-\beta})^{-1}$

$$\begin{bmatrix} j & Y_\beta \\ \kappa & \lambda \end{bmatrix} = \begin{bmatrix} j & Y_\pi \\ \kappa & \mu \end{bmatrix}\begin{bmatrix} j & Y_{\pi-\beta} \\ \lambda & \mu \end{bmatrix}$$

or

$$\begin{bmatrix} j & Y_\beta \\ \kappa & \lambda \end{bmatrix} = (-1)^{j-\kappa}\begin{bmatrix} j & Y_{\pi-\beta} \\ \lambda & -\kappa \end{bmatrix} = (-1)^{j+\lambda}\begin{bmatrix} j & Y_{\pi-\beta} \\ -\kappa & \lambda \end{bmatrix} \tag{33d}$$

all these formulas are written down here only for the sake of easier references. They can be, of course, also deduced directly from (30a) and are further illustrated by the table at the end of Section 12, which gives some of the matrices corresponding to Y_β for $\beta = \frac{1}{2}\pi$.

8. *The Three-Dimensional Rotation Group is Simply Reducible*

We now wish to establish the validity of the results of the paper of Reference 6, which are mentioned in this reference, for the unimodular unitary group.

The validity of the Lemma follows from the fact that both the considerations of Frobenius and Schur which are used for the derivation of the Lemma and also the derivation itself can be carried out without assuming that the underlying group is finite. It then follows from the half integer character of the representation $j = \frac{1}{2}$ and the above-described structure of the Kronecker

products that (as is well known) the representations with half integer numerical values of j (even dimensions) are half integer and the representations with integer numerical values of j (odd dimensions) are integer.

In order to determine the even and odd character of the representations, it is necessary to show, first, that the representations $j = 2j', 2j' - 2, 2j' - 4, \ldots$ are in the symmetric part $j' \bar{\times} j'$ of the Kronecker product $j' \times j'$ of the representation j' with itself, while $j = 2j' - 1,\ 2j - 3 \ldots$ are in the antisymmetric part $j' \underline{\times} j'$ of $j' \times j'$. This statement is evidently correct for $j' = 0$. It is also correct for $j' = \frac{1}{2}$ in which case $j' \bar{\times} j'$ must be three-dimensional (and is $j = 1$) while $j' \underline{\times} j'$ must be one-dimensional (and is $j = 0$). Assuming the above structure for all $j \leq j'$, it will be shown now for $j' + \frac{1}{2}$ which will make it generally valid.

Let us consider the symmetric part of the Kronecker product of $(j \times \frac{1}{2})$ with itself. This is[4]

$$(j' \times \tfrac{1}{2}) \bar{\times} (j' \times \tfrac{1}{2}) = [(j' + \tfrac{1}{2}) \dot{+} (j' - \tfrac{1}{2})] \bar{\times} [(j' + \tfrac{1}{2}) \dot{+} (j' - \tfrac{1}{2})]. \tag{34}$$

The trace of the right side can be easily expressed in terms of the traces of the Kronecker products of the representations involved. It is, according to (17) of Reference 6, identical with the trace of

$$(j' + \tfrac{1}{2}) \bar{\times} (j' + \tfrac{1}{2}) \dot{+} (j' + \tfrac{1}{2}) \times (j' - \tfrac{1}{2}) \dot{+} (j' - \tfrac{1}{2}) \bar{\times} (j' - \tfrac{1}{2}).$$

and is, according to our surmise, equivalent to the sum of the representations

$$\begin{aligned} &(j' + \tfrac{1}{2}) \bar{\times} (j' + \tfrac{1}{2}), 2j', 2j' - 1, 2j' - 2, 2j' - 3, 2j' - 4, 2j' - 5, \ldots \\ &\qquad\qquad 2j' - 1, \qquad 2j' - 3, \qquad 2j' - 5, \ldots . \end{aligned} \tag{34a}$$

On the other hand, one easily verifies by the same equation that the trace of $(j' \times \frac{1}{2}) \bar{\times} (j' \times \frac{1}{2})$ is also equal to the trace of

$$(j' \bar{\times} j') \times (\tfrac{1}{2} \bar{\times} \tfrac{1}{2}) \dot{+} (j' \underline{\times} j') \times (\tfrac{1}{2} \underline{\times} \tfrac{1}{2})$$

which is equal, according to our surmise, to

$$\begin{aligned} &[(2j') \dot{+} (2j' - 2) \dot{+} (2j' - 4) \ldots] \times (1) \dot{+} [(2j' - 1) \dot{+} (2j' - 3) \dot{+} \\ &(2j' - 5) \ldots] \times (0) = [(2j' + 1) \dot{+} (2j') \dot{+} (2j' - 1)] \dot{+} [(2j' - 1) \dot{+} (2j' - 2) \\ &\qquad\qquad \dot{+} (2j' - 1) \\ &\dot{+} (2j' - 3)] \dot{+} [(2j' - 3) \ldots \\ &\dot{+} (2j' - 3) \ldots . \end{aligned} \tag{34b}$$

Comparing (34a) and (34b) one sees that $(j' + \frac{1}{2}) \bar{\times} (j' \times \frac{1}{2})$ contains the representations $j = (2j' + 1), (2j' - 1), (2j' - 3), \ldots$ which is the desired result.

It follows that if j' is integer the representations $2j',\ 2j' - 2,\ 2j' - 4$, are even while $2j' - 1, 2j' - 3, \ldots$ are odd representations. If j' is half integer, the first series consists of odd and the second set of even representations; an integer representation j is even or odd according to the even or odd

character of j. Thus no representation is both even and odd, in accordance with Theorem 3 which had to be verified. The notation half integer, integer, odd and even for the representations of S.R. groups was chosen so that the numerical value of j shall have the character which the representation denotes.

For the definition of the $(-1)^j$ (cf. Section 2) we shall adopt the convention

$$(-1)^j = i^{2j}. \tag{35}$$

It may be remarked, however, that the convention $(-1)^j = (-i)^{2j}$ would lead to identical formulas, since the exponent of -1 is an integer in all equations.

9. *The Three-j-Symbols for the Three-Dimensional Rotation Group*

A few simple relations immediately follow from the above normalizations and the definition (2) and (2a) of the three-j-symbols. Writing in (2a) Z_α for R, we have, for (30), (2a)

$$\begin{pmatrix} j_1 & j_2 & j_3 \\ \lambda_1 & \lambda_2 & \lambda_3 \end{pmatrix} e^{i(\lambda_1+\lambda_2)\alpha} = e^{-i\lambda_3\alpha} \begin{pmatrix} j_1 & j_2 & j_3 \\ \lambda_1 & \lambda_2 & \lambda_3 \end{pmatrix}$$

or

$$\begin{pmatrix} j_1 & j_2 & j_3 \\ \lambda_1 & \lambda_2 & \lambda_3 \end{pmatrix} = 0 \qquad \text{unless } \lambda_1 + \lambda_2 + \lambda_3 = 0. \tag{36}$$

Setting, on the other hand, $R = Y_{-\pi}$ in (2a), we have

$$\begin{pmatrix} j_1 & j_2 & j_3 \\ -\lambda_1 & -\lambda_2 & \kappa_3 \end{pmatrix} (-1)^{j_1-\lambda_1+j_2-\lambda_2} = (-1)^{j_3+\kappa_3} \begin{pmatrix} j_1 & j_2 & j_3 \\ \lambda_1 & \lambda_2 & -\kappa_3 \end{pmatrix}$$

or

$$\begin{pmatrix} j_1 & j_2 & j_3 \\ -\lambda_1 & -\lambda_2 & -\lambda_3 \end{pmatrix} = \begin{pmatrix} j_1 & j_2 & j_3 \\ \lambda_1 & \lambda_2 & \lambda_3 \end{pmatrix} (-1)^{j_1+j_2+j_3} \tag{36a}$$

since $j_3 + \kappa_3$ being an integer $(-1)^{j_3+\kappa_3} = (-1)^{-j_3-\kappa_3}$ and since the three-j-symbols vanishes anyway unless $\lambda_1 + \lambda_2 + \lambda_3 = 0$.

Finally, (11a) and (32a) show that

$$\begin{pmatrix} j_1 & j_2 & j_3 \\ \lambda_1 & \lambda_2 & \lambda_3 \end{pmatrix}^* = (-1)^{j_1+j_2+j_3+\lambda_1+\lambda_2+\lambda_3} \begin{pmatrix} j_1 & j_2 & j_3 \\ -\lambda_1 & -\lambda_2 & -\lambda_3 \end{pmatrix}$$

so that the normalization adopted makes the three-j-symbols *real.*

It follows from (30b) and the well known orthogonality relations for representation coefficients that

$$\int \left| \begin{bmatrix} j & R \\ j & \lambda \end{bmatrix} \right|^2 dR = \frac{(2j)!}{(j-\lambda)!(j+\lambda)!} \int \cos^{2j+2\lambda} \tfrac{1}{2}\beta \sin^{2j-2\lambda} \tfrac{1}{2}\beta \, dR = \frac{h}{2j+1}. \tag{37}$$

where $h = \int dR$ is the volume of the group. Hence, we have, for (5), (30b), (33a), and (33c)

$$h\begin{pmatrix} j_1 & j_2 & j_3 \\ j_1 & -j_2 & j_2-j_1 \end{pmatrix}\begin{pmatrix} j_1 & j_2 & j_3 \\ \lambda & -\lambda+j_3 & -j_3 \end{pmatrix}$$

$$=\int\begin{bmatrix} j_1 & R \\ j_1 & \lambda \end{bmatrix}\begin{bmatrix} j_2 & R \\ -j_2 & -\lambda+j_3 \end{bmatrix}\begin{bmatrix} j_3 & R \\ j_2-j_1 & -j_3 \end{bmatrix}dR$$

$$=(-1)^{j_2+j_3-\lambda}$$

$$\cdot\frac{[(2j_1)!(2j_2)!(2j_3)!]^{1/2}\int\cos^{2j_1+2\lambda}\tfrac{1}{2}\beta\sin^{2j_2+2j_3-2\lambda}\tfrac{1}{2}\beta\,dR}{[(j_1-\lambda)!(j_1+\lambda)!(j_2-j_3+\lambda)!(j_2+j_3-\lambda)!(j_3-j_2+j_1)!(j_3+j_2-j_1)!]^{1/2}}$$

$$=\frac{(-1)^{j_2+j_3-\lambda}[(2j_1)!(2j_2)!(2j_3)!]^{1/2}(j_1+\lambda)!(j_2+j_3-\lambda)!h}{[(j_1-\lambda)!(j_1+\lambda)!(j_2-j_3+\lambda)!(j_2+j_3-\lambda)!\cdot(j_3-j_2+j_1)!(j_3+j_2-j_1)!]^{1/2}(j_1+j_2+j_3+1)!.} \tag{37a}$$

For $\lambda=j_1$, this gives

$$\begin{pmatrix} j_1 & j_2 & j_3 \\ j_1 & -j_2 & j_2-j_1 \end{pmatrix}\begin{pmatrix} j_1 & j_3 & j_2 \\ j_1 & -j_3 & j_3-j_1 \end{pmatrix}$$

$$=\frac{(-1)^{2j_1}(2j_1)!}{(j_1+j_2+j_3+1)!}\left[\frac{(2j_2)!(2j_3)!}{(j_1+j_2-j_3)!(j_1-j_2+j_3)!}\right]^{1/2}. \tag{37b}$$

This equation shows that if we give to the first factor the sign of $(-1)^{2j_3}$ (which we are yet at liberty to do), the sign of the second factor will be automatically $(-1)^{2j_2}$ since $(-1)^{2j_1}=(-1)^{2j_2+2j_3}$. On account of (6) and (36a), the sign $(-1)^{2j_3}$ will hold also for the three-j-symbol obtained by interchanging j_1 and j_2 in the first factor of (37b). The equations obtained from (37b) by cyclic permutations of the j_1, j_2, j_3 then show that the above convention will hold true for all three-j-symbols obtained from those on the left side of (37b) by an arbitrary interchange of the j. This normalization of the sign of the three-j-symbols leaves, furthermore, (7) and (32a) valid and will be adopted henceforth. In order to obtain the value of the first factor of (37b) separately, one can make a cyclic permutation of the j and multiply the resulting equation with (37b). Dividing this product by the equation obtained from (37b) by the other cyclic permutation of the j, one obtains an equation for the square of one three-j-symbol of (37b) alone. This gives

$$\begin{pmatrix} j_1 & j_2 & j_3 \\ j_1 & -j_2 & j_2-j_1 \end{pmatrix}=(-1)^{2j_3}\left[\frac{(2j_1)!(2j_2)!}{(j_1+j_2+j_3+1)!(j_1+j_2-j_3)!}\right]^{1/2} \tag{37c}$$

and now (37a)

$$\begin{pmatrix} j_1 & j_2 & j_3 \\ \lambda & j_3-\lambda & -j_3 \end{pmatrix}$$

$$=\frac{(-1)^{j_2-j_3-\lambda}[(2j_3)!(j_1+j_2-j_3)!(j_1+\lambda)!(j_2+j_3-\lambda)!]^{1/2}}{[(j_1+j_2+j_3+1)!(-j_1+j_2+j_3)!(j_1-j_2+j_3)!(j_1-\lambda)!(j_2-j_3+\lambda)!]^{1/2}}. \tag{37d}$$

This equation gives immediately all three-j-symbols in which one j is $\frac{1}{2}$. We have

$$\begin{pmatrix} j & j+\frac{1}{2} & \frac{1}{2} \\ \lambda & \frac{1}{2}-\lambda & -\frac{1}{2} \end{pmatrix} = \frac{(-1)^{j-\lambda}(j+1-\lambda)^{1/2}}{[(2j+1)(2j+2)]^{1/2}} \tag{37e}$$

and from this, all others can be obtained by interchanging the first two columns, or using (36a), or both.

10. *Calculation of the Racah Coefficients*

We now go over to the calculation of the six-j-symbols. Since these do not depend on the particular form in which the representations are assumed, it should be possible to calculate them directly, i.e., without first computing the three-j-symbols. This is in fact the case and can be done, for instance, in the following way.

One can remark, first, that the orthogonality relations (20) directly give those six-j-symbols for which the orthogonal matrix (20a) has only one dimension. This is the case, e.g., if $j_2+j_4=\pm(j_1-j_3)$ or $\pm(j_2-j_4)=j_1+j_3$. If $j_3=\frac{1}{2}$, the orthogonal matrix (20a) has, if it is not actually one-dimensional, only two rows and two columns and its opposite elements are therefore equal, apart from the sign. This means, e.g.,

$$\begin{Bmatrix} j_1 & j_2 & j_4-\frac{1}{2} \\ \frac{1}{2} & j_4 & j_2-\frac{1}{2} \end{Bmatrix} (2j_2 . 2j_4)^{1/2} = \pm \begin{Bmatrix} j_1 & j_2 & j_4+\frac{1}{2} \\ \frac{1}{2} & j_4 & j_2+\frac{1}{2} \end{Bmatrix} ((2j_2+2)(2j_4+2))^{1/2}$$

which, together with (19a) allows to reduce j_2 and j_4 simultaneously by $\frac{1}{2}$. This reduction can be carried to the point at which $j_2+j_4=j_1-\frac{1}{2}$, after which the resulting six-j-symbol is given by the first remark. The other six-j-symbols with $j_3=\frac{1}{2}$ can be calculated in a similar way. This being carried out, the calculation of the remaining six-j-symbols can proceed much in the same way as will be described toward the end of this section.

The reason that the above-outlined procedure will not be followed is that it yields only the absolute values of the six-j-symbols. The determination of the signs, on the other hand, becomes inelegant. Some of them can be chosen arbitrarily, as mentioned at the beginning of Section 6. This being done, the remaining signs can be obtained by the above orthogonality relations with an occasional reference to (21a). The whole discussion becomes, however, somewhat tedious. We shall make use therefore in the following calculation of the results of the preceding section.

If the six-j-symbol

$$\begin{Bmatrix} j_1 & j_2 & j_3 \\ l_1 & l_2 & l_3 \end{Bmatrix}$$

is different from zero, the following numbers must be integers:

$$\begin{aligned} J &= j_1+j_2+j_3 & J_1 &= j_1+l_2+l_3 \\ J_2 &= l_1+j_2+l_3 & J_3 &= l_1+l_2+j_3 \end{aligned} \tag{38}$$

and since they will frequently occur in the ensuing formulas, we shall use the above abbreviations for them.

A general formula for the six-j-symbols is contained in (22). By inserting j_1 for ι_1, $-j_2$ for ι_2, and j_2-j_1 for ι_3, one obtains a reasonably simple expression for the six-j-symbols. In the case $l_1+l_2=j_3$ one can set more advantageously $\iota_3=j_3$ and $\iota_1=-j_1$ or $\iota_2=-j_2$. Then, because of the last three-j-symbol on the right side, λ_1 and η_2 can assume only the values $-l_1$ and $-l_2$, respectively, and the right side of (22) reduces to a single term. One obtains with (37c) and (37d)

$$\begin{Bmatrix} j_1 & j_2 & j_3 \\ l_1 & l_2 & l_3 \end{Bmatrix} = (-1)^{j_1+j_2+j_3} \cdot \left[\frac{(J+1)!(J-2j_1)!(J-2j_2)!(J_1-2l_2)!(J_2-2l_1)!(2l_1)!(2l_2)!}{(J_1+1)!(J_2+1)!(J-2j_3)!(J_1-2j_1)! \cdot (J_2-2j_2)!(J_1-2l_3)!(J_2-2l_3)!(2j_3+1)!} \right]^{1/2}$$

$$\text{for } l_1+l_2=j_3. \quad (39)$$

This equation immediately gives all six-j-symbols in which one representation is $j=0$ or $j=\frac{1}{2}$. For $l_1=0$, we must set $l_2=j_3$ and $l_3=j_2$; otherwise the six-j-symbol vanishes.

$$\begin{Bmatrix} j_1 & j_2 & j_3 \\ 0 & j_3 & j_2 \end{Bmatrix} = (-1)^{j_1+j_2+j_3}[(2j_2+1)(2j_3+1)]^{-1/2}. \quad (39a)$$

This is identical with (18b) and thus independent of the normalizations of Section 9.

For $l_1=\frac{1}{2}$, we have to distinguish between two cases:

$$\begin{Bmatrix} j_1 & j_2 & j_3 \\ \frac{1}{2} & j_3-\frac{1}{2} & j_2+\frac{1}{2} \end{Bmatrix} = (-1)^{j_1+j_2+j_3} \left[\frac{(J-2j_2)(J-2j_3+1)}{(2j_2+1)(2j_2+2)2j_3(2j_3+1)} \right]^{1/2}, \quad (39b)$$

$$\begin{Bmatrix} j_1 & j_2 & j_3 \\ \frac{1}{2} & j_3-\frac{1}{2} & j_2-\frac{1}{2} \end{Bmatrix} = (-1)^{j_1+j_2+j_3} \left[\frac{(J+1)(J-2j_1)}{2j_2(2j_2+1)2j_3(2j_3+1)} \right]^{1/2}. \quad (39c)$$

We may note finally

$$\begin{Bmatrix} j_1 & j_2 & j_3 \\ \frac{1}{2}J-j_1 & \frac{1}{2}J-j_2 & \frac{1}{2}J-j_3 \end{Bmatrix} = (-1)^J \left[\frac{(J+1)!(J-2j_1)!(J-2j_2)!(J-2j_3)!}{(2j_1+1)!(2j_2+1)!(2j_3+1)!} \right]^{1/2}. \quad (40)$$

One can obtain a recursive formula for the six-j-symbols by setting $k_2=\frac{1}{2}$, $k_1=j_3-\frac{1}{2}$, $k_3=j_1-\frac{1}{2}$ in (26b). Both k and l will assume two values, $l_3\pm\frac{1}{2}$ and $l_1\pm\frac{1}{2}$, respectively, and the six-j-symbol will appear as a sum of

four six-j-symbols in which both j_1 and j_3 are smaller by $\frac{1}{2}$ than in the original one:

$$\frac{(2l_1+1)(2l_3+1)[(J+1)(J-2j_2)]^{1/2}}{[(J_1-2l_2)(J_1-2l_3)(J_3-2l_1)(J_3-2l_2)]^{1/2}}\begin{Bmatrix} j_1 & j_2 & j_3 \\ l_1 & l_2 & l_3 \end{Bmatrix}$$

$$= -\left[\frac{J_1+1}{J_1-2l_3}(J_2+1)(J_2-2j_2)\frac{J_3+1}{J_3-2l_1}\right]^{1/2}\begin{Bmatrix} j_1-\frac{1}{2} & j_2 & j_3-\frac{1}{2} \\ l_1-\frac{1}{2} & l_2 & l_3-\frac{1}{2} \end{Bmatrix}$$

$$-\left[\frac{J_1-2j_1+1}{J_1-2l_2}(J_2-2l_1+1)(J_2-2l_3)\frac{J_3+1}{J_3-2l_1}\right]^{1/2}\begin{Bmatrix} j_1-\frac{1}{2} & j_2 & j_3-\frac{1}{2} \\ l_1-\frac{1}{2} & l_2 & l_3+\frac{1}{2} \end{Bmatrix}$$

$$-\left[\frac{J_1+1}{J_1-2l_3}(J_2-2l_1)(J_2-2l_3+1)\frac{J_3-2j_3+1}{J_3-2l_2}\right]^{1/2}\begin{Bmatrix} j_1-\frac{1}{2} & j_2 & j_3-\frac{1}{2} \\ l_1+\frac{1}{2} & l_2 & l_3-\frac{1}{2} \end{Bmatrix}$$

$$+\left[\frac{J_1-2j_1+1}{J_1-2l_2}(J_2+2)(J_2-2j_2+1)\frac{J_3-2j_3+1}{J_3-2l_2}\right]^{1/2}\begin{Bmatrix} j_1-\frac{1}{2} & j_2 & j_3-\frac{1}{2} \\ l_1+\frac{1}{2} & l_2 & l_3+\frac{1}{2} \end{Bmatrix}. \tag{41}$$

This formula is valid, of course, only if $J-2j_2=j_1-j_2+j_3>0$. It permits one to diminish the j occurring in a six-j-symbol until this assumes a form in which one of the equations (39a, b, or c) is applicable. The following tables were obtained in this way. The numerical ones illustrate, in particular, the orthogonality relations: every table, considered as a matrix, is orthogonal if the weight factor $((2j+1)(2l+1))^{1/2}$ is inserted. It may be mentioned that this fact also follows a successive calculation of all six-j-symbols.

$$\begin{Bmatrix} j_1 & j_2 & j_3 \\ 1 & j_3-1 & j_2-1 \end{Bmatrix} = (-1)^{j_1+j_2+j_3} \cdot\left[\frac{J(J+1)(J-2j_1-1)(J-2j_1)}{(2j_2-1)2j_2(2j_2+1)(2j_3-1)2j_3(2j_3+1)}\right]^{1/2} \tag{42a}$$

$$\begin{Bmatrix} j_1 & j_2 & j_3 \\ 1 & j_3-1 & j_2 \end{Bmatrix} = (-1)^{j_1+j_2+j_3} \cdot\left[\frac{2(J+1)(J-2j_1)(J-2j_2)(J-2j_3+1)}{2j_2(2j_2+1)(2j_2+2)(2j_3-1)2j_3(2j_3+1)}\right]^{1/2} \tag{42b}$$

$$\begin{Bmatrix} j_1 & j_2 & j_3 \\ 1 & j_3-1 & j_2+1 \end{Bmatrix} = (-1)^{j_1+j_2+j_3} \cdot\left[\frac{(J-2j_2-1)(J-2j_2)(J-2j_3+1)(J-2j_3+2)}{(2j_2+1)(2j_2+2)(2j_2+3)(2j_3-1)2j_3(2j_3+1)}\right]^{1/2} \tag{42c}$$

$$\begin{Bmatrix} j_1 & j_2 & j_3 \\ 1 & j_3 & j_2 \end{Bmatrix} = (-1)^{j_1+j_2+j_3} \cdot\frac{2(j_1^2+j_1-j_2^2-j_2-j_3^2-j_3)}{[2j_2(2j_2+1)(2j_2+2)2j_3(2j_3+1)(2j_3+2)]^{1/2}}. \tag{42d}$$

Formulas (42) give the values of all six-j-symbols in which one of the j is 1, some of them by means of the symmetry relations (19).

Table for $\begin{Bmatrix} 3/2 & 3/2 & j \\ 3/2 & 3/2 & l \end{Bmatrix}$

$l \backslash j$	0	1	2	3
0	$-1/4$	$1/4$	$-1/4$	$1/4$
1	$1/4$	$-11/60$	$1/20$	$3/20$
2	$-1/4$	$1/20$	$3/20$	$1/20$
3	$1/4$	$3/20$	$1/20$	$1/140$

Table for $\begin{Bmatrix} 3/2 & 3/2 & j \\ 2 & 2 & l \end{Bmatrix}$

$l \backslash j$	0	1	2	3
$\frac{1}{2}$	$\frac{5^{1/2}}{10}$	$-\frac{18^{1/2}}{20}$	$\frac{14^{1/2}}{20}$	$-\frac{2^{1/2}}{10}$
$\frac{3}{2}$	$-\frac{5^{1/2}}{10}$	$\frac{2^{1/2}}{10}$	0	$-\frac{2^{1/2}}{10}$
$\frac{5}{2}$	$\frac{5^{1/2}}{10}$	$-\frac{2^{1/2}}{60}$	$-\frac{14^{1/2}}{28}$	$-\frac{18^{1/2}}{70}$
$\frac{7}{2}$	$-\frac{5^{1/2}}{10}$	$-\frac{2^{1/2}}{10}$	$-\frac{14^{1/2}}{70}$	$-\frac{2^{1/2}}{140}$

Table for $\begin{Bmatrix} 3/2 & 2 & j \\ 3/2 & 2 & l \end{Bmatrix}$

$l \backslash j$	1/2	3/2	5/2	7/2
1/2	$1/20$	$-1/10$	$3/20$	$-1/5$
3/2	$-1/10$	$3/20$	$-1/10$	$-1/10$
5/2	$3/20$	$-1/10$	$-47/420$	$-1/35$
7/2	$-1/5$	$-1/10$	$-1/35$	$-1/280$

Table for $\begin{Bmatrix} 2 & 2 & l \\ 2 & 2 & j \end{Bmatrix}$

$l \backslash j$	0	1	2	3	4
0	$1/5$	$-1/5$	$1/5$	$-1/5$	$1/5$
1	$-1/5$	$1/6$	$-1/10$	0	$2/15$
2	$1/5$	$-1/10$	$-3/70$	$4/35$	$2/35$
3	$-1/5$	0	$4/35$	$1/14$	$1/70$
4	$1/5$	$2/15$	$2/35$	$1/70$	$1/630$

11. *Calculation of the Three-j-Symbols*

The three-j-symbols in which one of the j is equal to the sum of two others can be obtained rather simply from (2a). Writing therein $-\kappa_3 = j_3 = j_1 + j_2$ only the term with $\kappa_1 = j_1$, $\kappa_2 = j_2$ remains on the left side. The representation coefficients are all of the form (30b) or can be brought to this form by (33a). One obtains with (37c)

$$\begin{pmatrix} j_1 & j_2 & j_1+j_2 \\ \lambda_1 & \lambda_2 & -\lambda_1-\lambda_2 \end{pmatrix} = (-1)^{j_1-j_2+\lambda_1+\lambda_2} \cdot \left[\frac{(2j_1)!(2j_2)!(j_1+j_2+\lambda_1+\lambda_2)!(j_1+j_2-\lambda_1-\lambda_2)!}{(2j_1+2j_2+1)!(j_1+\lambda_1)!(j_1-\lambda_1)!(j_2+\lambda_2)!(j_2-\lambda_2)!}\right]^{1/2}. \tag{43}$$

This yields, together with (39d), (32a), and (22), a general formula for the three-j-symbols. For this purpose, we write $\frac{1}{2}J-j_1$ for l_1, $\frac{1}{2}J-j_2$ for l_2 and $\frac{1}{2}J-j_3$ for l_3 in (22) which gives to all three-j-symbols on the right side the form (43). The six-j-symbol is given by (40) and one obtains

$$\begin{pmatrix} j_1 & j_2 & j_3 \\ \lambda_1 & \lambda_2 & \lambda_3 \end{pmatrix} = \sum_{\kappa_1\kappa_2\kappa_3} \frac{(-1)^{\kappa_1+\kappa_2+\kappa_3-\frac{1}{2}J}[(J-2j_1)!(J-2j_2)!(J-2j_3)!]^{1/2}}{[(J+1)!]^{1/2}(\frac{1}{2}J-j_1-\kappa_1)!(\frac{1}{2}J-j_1+\kappa_1)!}$$
$$\cdot \frac{[(j_1-\lambda_1)!(j_1+\lambda_1)!(j_2-\lambda_2)!(j_2+\lambda_2)!(j_3-\lambda_3)!(j_3+\lambda_3)!]^{1/2}}{(\frac{1}{2}J-j_2-\kappa_2)!(\frac{1}{2}J-j_2+\kappa_2)!(\frac{1}{2}J-j_3-\kappa_3)!(\frac{1}{2}J-j_3+\kappa_3)!} \tag{44}$$
$$\kappa_3-\kappa_2=\lambda_1,\ \kappa_1-\kappa_3=\lambda_2,\ \kappa_2-\kappa_1=\lambda_3.$$

The summation runs in reality only over one index because the last line determines all three κ if one is given. The summation has to be extended over all values of κ_1, κ_2, κ_3, compatible with the last line of (44), for which all factorials in the denominator become non-negative integers. Equation (44) is essentially identical with Eq. (27) in Chapter XVII of Reference 3. One can show that the connection between the s and our three-j-symbols is

$$s^{(ll)}_{\mathscr{L}\mu\nu} = (-1)^{l-l+\mu+\nu}(2\mathscr{L}+1)^{1/2}\begin{pmatrix} l & l & \mathscr{L} \\ \mu & \nu & -\mu-\nu \end{pmatrix}. \tag{44a}$$

One can obtain a recursive formula for the three-j-symbols from (22), which is similar to (41) but much more easily manageable. Setting $l_1=\frac{1}{2}$, $l_2=j_3-\frac{1}{2}$, $l_3=j_2-\frac{1}{2}$ in (22) all symbols in which l_1 occurs can be expressed directly by (37e) and (39c), the one-j-symbols by (32a). Only two terms of the summation remain, corresponding to $\lambda_1=\pm\frac{1}{2}$. One obtains in this way

$$\begin{pmatrix} j_1 & j_2 & j_3 \\ \iota_1 & \iota_2 & \iota_3 \end{pmatrix} = -\left[\frac{(j_2-\iota_2)(j_3+\iota_3)}{(J+1)(J-2j_1)}\right]^{1/2}\begin{pmatrix} j_1 & j_2-\frac{1}{2} & j_3-\frac{1}{2} \\ \iota_1 & \iota_2+\frac{1}{2} & \iota_3-\frac{1}{2} \end{pmatrix}$$
$$+\left[\frac{(j_2+\iota_2)(j_3-\iota_3)}{(J+1)(J-2j_1)}\right]^{1/2}\begin{pmatrix} j_1 & j_2-\frac{1}{2} & j_3-\frac{1}{2} \\ \iota_1 & \iota_2-\frac{1}{2} & \iota_3+\frac{1}{2} \end{pmatrix}. \tag{45}$$

The three-j-symbols given in the next table[5] were calculated by this formula and (37e). Since the three-j-symbol vanishes unless the sum of the greek indices is zero, all symbols with the same value of the j can be given in one figure. The center of the figure corresponds to $\iota_1=\iota_2=\iota_3=0$—if all j are integers; ι_1 is constant in horizontal rows, ι_2 in rows of 60° to the right up, ι_3 in rows of 60° to the left up. The values of ι_1, ι_2, ι_3 corresponding to the different points of the diagrams in case of integer j is given in the first diagram, in case ι_1 and ι_2 are half integer (ι_3 therefore integer) in the second diagram. In this case, the center of the diagram is marked by $\otimes$. Each of the

remaining diagrams contains the values of the three-j-symbol given at the left of the diagram. The notation in the remaining cases is analogous.

	-220		-211		-202		
$-12-1$		-110		-101		$-1-12$	
	$01-1$		000		$0-11$		
$11-2$		$10-1$		$1-10$		$1-21$	
	$20-2$		$2-1-1$		$2-20$		

	$-\frac{3}{2}\ \frac{3}{2}\ 0$		$-\frac{3}{2}\ \frac{1}{2}\ 1$			$-\frac{3}{2}\ -\frac{1}{2}\ 2$	
$-\frac{1}{2}\ \frac{3}{2}\ -1$		$-\frac{1}{2}\ \frac{1}{2}\ 0$		$-\frac{1}{2}\ -\frac{1}{2}\ 1$			
			$\otimes$				
	$\frac{1}{2}\ \frac{1}{2}\ -1$		$\frac{1}{2}\ -\frac{1}{2}\ 0$			$\frac{1}{2}\ -\frac{3}{2}\ 1$	
$\frac{3}{2}\ \frac{1}{2}\ -2$		$\frac{3}{2}\ -\frac{1}{2}\ -1$		$\frac{3}{2}\ -\frac{3}{2}\ 0$			

$(0\ j\ j)$	$\frac{(-1)^{2j}}{(2j+1)^{1/2}}$,	$\frac{(-1)^{2j-1}}{(2j+1)^{1/2}}$,	$\cdots$,	$\frac{-1}{(2j+1)^{1/2}}$,	$\frac{1}{(2j+1)^{1/2}}$		

			$6^{-1/2}$		$-3^{-1/2}$		
$(\frac{1}{2}\ \frac{1}{2}\ 1)$				$\otimes$			
		$-3^{-1/2}$		$6^{-1/2}$			

	$-\frac{1}{2}$		$6^{-1/2}$		$-12^{-1/2}$		
$(\frac{1}{2}\ \frac{3}{2}\ 1)$				$\otimes$			
		$12^{-1/2}$		$-6^{-1/2}$		$\frac{1}{2}$	

	$20^{-1/2}$		$-10^{-1/2}$		$(3/20)^{-1/2}$		$-5^{-1/2}$
$(\frac{1}{2}\ \frac{3}{2}\ 2)$				$\otimes$			
	$-5^{-1/2}$		$(3/20)^{1/2}$		$-10^{-1/2}$		$20^{-1/2}$

	$-6^{-1/2}$		$(2/15)^{1/2}$		$-10^{-1/2}$		$15^{-1/2}$		$-30^{-1/2}$
$(\frac{1}{2}\ \frac{5}{2}\ 2)$					$\otimes$				
	$30^{-1/2}$		$-15^{-1/2}$		$10^{-1/2}$		$-(2/15)^{-1/2}$		$6^{-1/2}$

S. R. GROUPS

			$42^{-1/2}$	$-21^{-1/2}$	$14^{-1/2}$	$-(2/21)^{-1/2}$	$(5/42)^{1/2}$	$7^{-1/2}$			
$(\frac{1}{2}\ \frac{5}{2}\ 3)$						$\otimes$					
	$-7^{-1/2}$	$(5/42)^{1/2}$	$-(2/21)^{-1/2}$	$14^{-1/2}$	$-21^{-1/2}$	$42^{-1/2}$					
					$-6^{-1/2}$	$6^{-1/2}$					
$(1\ 1\ 1)$				$6^{-1/2}$	0	$-6^{-1/2}$					
					$-6^{-1/2}$	$6^{-1/2}$					
			$30^{-1/2}$		$-10^{-1/2}$		$5^{-1/2}$				
$(1\ 1\ 2)$		$-10^{-1/2}$		$(2/15)^{1/2}$		$-10^{-1/2}$					
	$5^{-1/2}$		$-10^{-1/2}$		$30^{-1/2}$						
			$-15^{-1/2}$	$10^{-1/2}$	$-10^{-1/2}$	$15^{-1/2}$					
$(1\ 2\ 2)$		$(2/15)^{1/2}$	$-30^{-1/2}$	0	$30^{-1/2}$	$-(2/15)^{-1/2}$					
			$-15^{-1/2}$	$10^{-1/2}$	$-10^{-1/2}$	$15^{-1/2}$					
			$105^{-1/2}$	$-35^{-1/2}$	$(2/35)^{1/2}$	$-(2/21)^{1/2}$	$7^{-1/2}$				
$(1\ 2\ 3)$		$-21^{-1/2}$	$(8/105)^{1/2}$	$-(3/35)^{1/2}$	$(8/105)^{1/2}$	$-21^{-1/2}$					
	$7^{-1/2}$	$-(2/21)^{1/2}$	$(2/35)^{1/2}$	$-35^{-1/2}$	$105^{-1/2}$						

12. *Calculation of the Representation Coefficients*

We go over, finally, to formulas concerning the representation coefficients themselves. On account of (30) these are essentially equivalent with formulas relating to the representatives of Y_β in $R = Z_\alpha Y_\beta Z_\gamma$.

Using (2a) with $j_3 = j_1 + j_2$, (43) gives a simple expression for the three-j-symbols occurring and we can set $\lambda_1 = j_1$ in order to have by (30b) and (33b) a simple form for the j_1 representation coefficient. Writing κ for κ_3, ν for κ_1 and $-j_1 - \lambda$ for λ_2 we obtain, using (33a), for $\lambda < j_2 - j_1$

$$\begin{bmatrix} j_1 + j_2 & Y_\beta \\ \kappa & \lambda \end{bmatrix}$$
$$= \sum_\nu \frac{(2j_1)!}{(j_1+\nu)!(j_1-\nu)!} \left[\frac{(j_1+j_2-\kappa)!(j_1+j_2+\kappa)!(j_2-j_1-\lambda)!}{(j_2-\nu-\kappa)!(j_2+\nu+\kappa)!(j_2+j_1-\lambda)!}\right]^{1/2}$$
$$\cdot \begin{bmatrix} j_2 & Y_\beta \\ \kappa+\nu & j_1+\lambda \end{bmatrix} \cos^{j_1+\nu}\tfrac{1}{2}\beta \sin^{j_1-\nu}\tfrac{1}{2}\beta. \tag{46}$$

The summation over ν must be extended, as in all similar formulas, over all values for which in the denominators only factorials of non-negative integers occur.

One can specialize this by assuming $j_1 = \frac{1}{2}$

$$\begin{bmatrix} j & Y_\beta \\ \kappa & \lambda \end{bmatrix} = \left(\frac{j+\kappa}{j-\lambda}\right)^{1/2} \begin{bmatrix} j-\frac{1}{2} & Y_\beta \\ \kappa-\frac{1}{2} & \lambda+\frac{1}{2} \end{bmatrix} \sin \tfrac{1}{2}\beta + \left(\frac{j-\kappa}{j-\lambda}\right)^{1/2} \begin{bmatrix} j-\frac{1}{2} & Y_\beta \\ \kappa+\frac{1}{2} & \lambda+\frac{1}{2} \end{bmatrix} \cos \tfrac{1}{2}\beta. \tag{46a}$$

This provides a useful formula for the calculation of the representation coefficients [except for $\lambda = +j$ which can be obtained from (30b), for example] while (46) allows one to take larger steps. For $j = 1$, (46a) gives easily

$$\begin{bmatrix} 1 & Y_\beta \\ \kappa & \lambda \end{bmatrix} = \begin{Vmatrix} \frac{1}{2}(1+\cos\beta) & -2^{-1/2}\sin\beta & \frac{1}{2}(1-\cos\beta) \\ 2^{-1/2}\sin\beta & \cos\beta & -2^{-1/2}\sin\beta \\ \frac{1}{2}(1-\cos\beta) & 2^{-1/2}\sin\beta & \frac{1}{2}(1+\cos\beta) \end{Vmatrix} \tag{46b}$$

(in all these matrices, the order of rows and columns is $-j, -j+1, \ldots, j-1, j$).

The disadvantage of Eqs. (46), (46a), and (30a) is that they express the representation coefficients by means of the powers of $\cos \frac{1}{2}\beta$ and $\sin \frac{1}{2}\beta$. The general behavior of these functions can be visualized much more readily if they are given in terms of cosines and sines of multiples of β. This is done, for the representatives of Z_α, in (30). For the Y_β one can start out from the remark that Y_β is in the class of Z_β and can be obtained from Z_β

$$Y_\beta = X Z_\beta X^{-1} \tag{47}$$

by transforming the latter with

$$X = X_{\pi/2} = \begin{Vmatrix} 2^{-1/2} & i2^{-1/2} \\ i2^{-1/2} & 2^{-1/2} \end{Vmatrix}. \tag{47a}$$

Hence

$$\begin{bmatrix} j & Y_\beta \\ \kappa & \lambda \end{bmatrix} = \begin{bmatrix} j & X_{\pi/2} \\ \kappa & \mu \end{bmatrix} \begin{bmatrix} j & Z_\beta \\ \mu & \mu' \end{bmatrix} \begin{bmatrix} j & X_{\pi/2} \\ \lambda & \mu' \end{bmatrix}^* = \begin{bmatrix} j & X_{\pi/2} \\ \kappa & \mu \end{bmatrix} \begin{bmatrix} j & X_{\pi/2} \\ \lambda & \mu \end{bmatrix}^* e^{i\mu\beta} \tag{47b}$$

where,

$$\begin{bmatrix} j & X_{\pi/2} \\ \kappa & \mu \end{bmatrix} = \begin{bmatrix} j & Z_{\pi/2} \\ \kappa & \nu \end{bmatrix} \begin{bmatrix} j & Y_{\pi/2} \\ \nu & \nu' \end{bmatrix} \begin{bmatrix} j & Z_{-\pi/2} \\ \nu' & \mu \end{bmatrix} = i^{\kappa-\mu} \begin{bmatrix} j & Y_{\pi/2} \\ \kappa & \mu \end{bmatrix}. \tag{47c}$$

Substituting this into (47b)

$$\begin{bmatrix} j & Y_\beta \\ \kappa & \lambda \end{bmatrix} = \sum_\mu \begin{bmatrix} j & Y \\ \mu & \kappa \end{bmatrix} \begin{bmatrix} j & Y \\ \mu & \lambda \end{bmatrix} i^{\lambda-\kappa} e^{i\mu\beta} \tag{48}$$

one sees that the knowledge of the representatives of $Y = Y_{\pi/2}$ suffices for obtaining the desired formula for the representatives of an arbitrary Y_β. It is possible, by means of (33d), to reduce the summation over μ to non-negative values of this quantity

$$\begin{bmatrix} j & Y_\beta \\ \kappa & \lambda \end{bmatrix} = 2\sum{}' \begin{bmatrix} j & Y \\ \mu & \kappa \end{bmatrix} \begin{bmatrix} j & Y \\ \mu & \lambda \end{bmatrix} \operatorname{tr} \mu\beta \tag{48a}$$

where the summation over μ extends over all positive values of this quantity and over the value $\mu = 0$ with a factor $\frac{1}{2}$;

$$\operatorname{tr} \mu\beta = \begin{cases} \cos \mu\beta & \text{if } \kappa - \lambda = 0 \pmod 4 \\ \sin \mu\beta & \text{if } \kappa - \lambda = 1 \pmod 4 \\ -\cos \mu\beta & \text{if } \kappa - \lambda = 2 \pmod 4 \\ -\sin \mu\beta & \text{if } \kappa - \lambda = 3 \pmod 4. \end{cases} \tag{48b}$$

The calculation of the $\begin{bmatrix} j & Y \\ \mu & \kappa \end{bmatrix}$ can be easily carried out by means of (46a) setting $\sin \frac{1}{2}\beta = \cos \frac{1}{2}\beta = 2^{-1/2}$ and one is further aided in this calculation by the relations (33). The first few of them are

$$2^{-1/2} \begin{Vmatrix} 1 & -1 \\ 1 & 1 \end{Vmatrix}; \qquad 2^{-1} \begin{Vmatrix} 1 & -2^{1/2} & 1 \\ 2^{1/2} & 0 & -2^{1/2} \\ 1 & 2^{1/2} & 1 \end{Vmatrix};$$

$$2^{-3/2} \begin{Vmatrix} 1 & -3^{1/2} & 3^{1/2} & -1 \\ 3^{1/2} & -1 & -1 & 3^{1/2} \\ 3^{1/2} & 1 & -1 & -3^{1/2} \\ 1 & 3^{1/2} & 3^{1/2} & 1 \end{Vmatrix};$$

$$2^{-2} \begin{Vmatrix} 1 & -2 & 6^{1/2} & -2 & 1 \\ 2 & -2 & 0 & 2 & -2 \\ 6^{1/2} & 0 & -2 & 0 & 6^{1/2} \\ 2 & 2 & 0 & -2 & -2 \\ 1 & 2 & 6^{1/2} & 2 & 1 \end{Vmatrix}$$

III. Applications to Quantum Mechanics

13. *The Hilbert Space as a Union of Representation Spaces; Irreducible Tensors*

The application of the notions developed in the previous sections as it occurs in quantum mechanics shall be summarized in this section. Since

it is not necessary for these applications to specify the group with which we are dealing (which is the symmetry group of the underlying problem), we shall not insert into our formulas the values of three-j-symbols etc., obtained in the previous sections for the rotation group. For the same reason, the real character of the three-j-symbols and one-j-symbols, although true in practically all applications, will not be assumed here.

In quantum mechanics, the different representation spaces are united into a single Hilbert space. In fact, the Hilbert space of quantum theory usually consists of the direct sum of an infinite number of each of the representation spaces. However, practically all the vectors in this Hilbert space which we shall have to deal with will be confined to one of the representation spaces, i.e., will have vanishing components in the direction of the axes of all but one representation space.

We introduce in each representation space l a set of orthogonal vectors ψ_ν^l. This is a slight change of notation from that used in Section 4: in that section ψ_ν was the ν component of a vector. Henceforth, ψ_ν^l will denote a unit vector, situated in the representation space l. The general vector of this representation space is

$$\phi^l = a^{l\nu}\psi_\nu^l \tag{49}$$

while the general vector of the Hilbert space is

$$\phi = \sum_l a^{l\nu}\psi_\nu^l. \tag{49a}$$

There are linear unitary operators P_R in the Hilbert space which permit one to express the invariance of this space with respect to the transformations R of the underlying group. The transformation of the unit vectors is given by

$$P_R\psi_\nu^l = \begin{bmatrix} l & R \\ \mu & \nu \end{bmatrix} \psi_\mu^l \tag{50}$$

i.e., the P_R leave the representation spaces within the Hilbert space invariant. Equation (50) conforms with the general covariant-contravariant formalism: it was noted after (14) that the first index of the representation coefficients is contravariant.

Because of its linearity, P_R applied to ϕ^l of (49) gives

$$P_R\phi_l = a^{l\nu}P_R\psi_\nu^l = a^{l\nu}\begin{bmatrix} R & l \\ \mu & \nu \end{bmatrix} \psi_\mu^l. \tag{51}$$

This equation permits P_R to be interpreted also as a transformation of the $a^{l\nu}$ to

$$a^{l\nu} \to \begin{bmatrix} l & R \\ \mu & \nu \end{bmatrix} a^{l\nu}. \tag{51a}$$

This equation expresses the fact that the contravariant components $a^{l\nu}$ of ϕ transform by the usual law of transformation for contravariant components.

If S is another group element the calculation of $P_S(P_R\psi^l_\nu)$ can be carried out using the linear nature of the P and (50)

$$\begin{aligned} P_S(P_R\psi^l_\nu) &= P_S\left(\begin{bmatrix} l & R \\ \mu & \nu \end{bmatrix}\psi^l_\mu\right) = \begin{bmatrix} l & R \\ \mu & \nu \end{bmatrix} P_S\psi^l_\mu \\ &= \begin{bmatrix} l & R \\ \mu & \nu \end{bmatrix}\begin{bmatrix} l & S \\ \eta & \mu \end{bmatrix}\psi^l_\eta = \begin{bmatrix} l & SR \\ \eta & \nu \end{bmatrix}\psi^l_\eta = P_{SR}\psi^l_\nu . \end{aligned} \tag{52}$$

A similar equation holds also for the general vector ϕ. It follows from this that

$$P_S P_R = P_{SR} \tag{52a}$$

The reader versed in quantum mechanics will recognize the Hilbert space as the space of all state vectors. The ψ^l_ν usually represent the characteristic functions of a definite eigenvalue of the energy operator; the ϕ^l of (49) is the general characteristic function of this eigenvalue. The energy operator H is assumed to be invariant under the transformations of the underlying group ($P_R H P_{R^{-1}} = H$) and a representation l of this group is associated with every eigenvalue of H. Usually, there is an infinity of eigenvalues with which the same representation l is associated and this holds for all l. This accounts for the remark made earlier in this section that the Hilbert space is, usually, a direct sum of infinitely many of each of the representation spaces.

The next important notions are "irreducible tensor operators of rank t" t_τ where τ runs from $-t$ to t. They satisfy the equations

$$P_R t_\tau P_{R^{-1}} = \begin{bmatrix} t & R \\ \sigma & \tau \end{bmatrix} t_\sigma \tag{53}$$

which is again consistent with (52a) inasmuch as

$$P_S P_R t_\tau P_{R^{-1}} P_{S^{-1}} = P_{SR} t_\tau P_{R^{-1}S^{-1}} .$$

It will be noted that, under this definition, the energy operator H is an irreducible tensor operator of rank 0, i.e., an invariant operator or a scalar, in the case which we have called above usual.

The operators of integer rank can be Hermitian. In fact, assuming the representation t to be real it is directly clear from (53) that the components $P_R t_\tau P_{R^{-1}}$ are hermitean if this holds for the original components t_τ. If one does not want to assume the representation in the real form, it is not useful to define the self adjoint character for the single operators t_τ, but only for their entity. This shall imply, at any rate, that the hermitean adjoints of the t_τ are linear combinations of the t_τ themselves

$$t^\dagger_\tau = a_{\sigma\tau} t_\sigma . \tag{53a}$$

Transforming this equation by P_R, yields, because of $P_R t_\tau^\dagger P_{R^{-1}} = (P_R t_\tau P_{R^{-1}})^\dagger$, the relation that $a_{\sigma\tau}$ transforms the representation t into its conjugate imaginary form. Hence, apart from a constant, $a_{\sigma\tau}$ must be the one-j-symbol corresponding to the representation t and we can assume the operators t_τ to be multiplied by a constant in such a way that

$$t_\tau^\dagger = \begin{pmatrix} t & \\ \sigma & \tau \end{pmatrix}^* t_\sigma = t^\tau \tag{54}$$

holds exactly. In our covariant-contravariant formalism (54) states that the contravariant components of a self-adjoint tensor operator are the hermitean adjoints of the covariant components.

It follows from (54) that tensor operators corresponding to a half integer rank cannot form a self-adjoint set, i.e., cannot satisfy (53a). This becomes evident if one takes the Hermitian adjoint of (54) and expresses on the right side the $t_\sigma^\dagger$ by the t_ρ themselves. Because the t_ρ are necessarily linearly independent, this gives (10) for the one-j-symbol of t, however, without the $(-1)^{2t}$ factor. It follows that $2t$ is even and t an integer. It should be remarked that the self-adjoint character of the set t_τ does not imply that all t_τ are Hermitian themselves. Thus, for instance, in the case of the rotation group only the t_0 are hermitean, and these only for even values of t. The t_0 for odd values of t are skew hermitean.

The simplest "tensor operator" is of rank 0, i.e., a scalar and the number 1 (or the energy operator of the above-mentioned example) is an example thereof. It is clear that linear combinations of the operators "multiplication with x, y, or z" will form the components of a tensor operator of rank 1, i.e., of a vector operator, and we shall proceed to determine these linear combinations explicitly. The following calculation will, at the same time, prove the vector operator character of (61).

Let us set

$$\begin{aligned} v_{-1} &= s_{x-1}x + s_{y-1}y + s_{z-1}z \\ v_0 &= s_{x0}x + s_{y0}y + s_{z0}z \\ v_1 &= s_{x1}x + s_{y1}y + s_{z1}z. \end{aligned} \tag{55}$$

Writing $x_1 = x$, $x_2 = y$, $x_3 = z$ one can abbreviate (55) as

$$v_\tau = \sum_a s_{a\tau} x_a. \tag{55a}$$

One easily convinces oneself that the definition

$$P_R f(x, y, z) = f(\sum R_{ax} x_a, \sum R_{ay} x_a, \sum R_{az} x_a) \tag{56}$$

or

$$P_R f(x) = f(R^{-1}x) \tag{56a}$$

satisfies (52a). It follows for the operators x_a (i.e., multiplication with x_a)

$$P_R x_a P_{R^{-1}} = \sum_b R_{ba} x_b \tag{57}$$

and for

$$P_R v_\tau P_{R^{-1}} = \sum_{ab} s_{a\tau} R_{ba} x_b. \tag{58}$$

Combined with (53), this gives

$$\sum_{ab} R_{ba} s_{a\tau} x_b = \begin{bmatrix} 1 & R \\ \sigma & \tau \end{bmatrix} v_\sigma = \sum_b \begin{bmatrix} 1 & R \\ \sigma & \tau \end{bmatrix} s_{b\sigma} x_b,$$

whence

$$Rs = s[1; R] \tag{59}$$

$[l; R]$ being the matrix which corresponds to R in the representation l. According to (59), s is the matrix which transforms R into the representation $l = 1$.

Naturally, (59) leaves a common multiplicative constant in all s free. This corresponds to the circumstances that a tensor operator does not lose its character if all its components are multiplied by the same constant. We shall choose this constant in such a way that v forms a self-adjoint set in the sense of (54); this will fix the argument of the indeterminate constant. The calculation of the s which satisfies (59) is quite straightforward. Thus if $R = Z_a$ is a rotation about Z, the representation is given by (30) while Z_a itself is

$$Z_a = \begin{Vmatrix} \cos\alpha & \sin\alpha & 0 \\ -\sin\alpha & \cos\alpha & 0 \\ 0 & 0 & 1 \end{Vmatrix}. \tag{60}$$

Hence (59) gives $s_{z-1} = s_{z1} = s_{x0} = s_{y0} = 0$; $s_{x-1} = is_{y-1}$; $s_{x1} = -is_{y1}$. The form of $[1; Y_\beta]$ is given in (46b) and the corresponding Y_β is obtained by a cyclic permutation of (60). Normalizing $s_{z0} = i$, one obtains in this way

$$\begin{aligned} v_{-1} &= 2^{-1/2} i(x - iy) \\ v_0 &= iz \\ v_1 &= -2^{-1/2} i(x + iy). \end{aligned} \tag{61}$$

It forms a self-adjoint set. Another, equally important set of vector operators is given by the infinitesimal rotations themselves. Let us define

$$M_a = \lim_{\varepsilon \to 0} \frac{1}{\varepsilon} (P_{a(\varepsilon)} - 1) \tag{62}$$

where a may be any of the three axes x, y, z and $a(\varepsilon)$ is the rotation by ε

about the axis a. Then

$$\begin{aligned} J_{-1} &= -\frac{i}{\sqrt{2}}(M_x - iM_y) \\ J_0 &= -iM_z \\ J_1 &= \frac{i}{\sqrt{2}}(M_x + iM_y) \end{aligned} \tag{62a}$$

form a set of (not self-adjoint) vector operators.

Since the effect of P_R on the wave functions is directly given by (50), the effect of those operators on the ψ_ν^l can be calculated directly. Let us calculate first $M_a\psi_\nu^l$. This is

$$M_a\psi_\nu^l = \lim \frac{1}{\varepsilon}(P_{a(\varepsilon)} - 1)\psi_\nu^l = \lim\left(\frac{1}{\varepsilon}\begin{bmatrix} l & a(\varepsilon) \\ \mu & \nu \end{bmatrix} - \frac{1}{\varepsilon}\delta_{\mu\nu}\right)\psi_\mu^l. \tag{63}$$

We need the $\begin{bmatrix} l & a(\varepsilon) \\ \mu & \nu \end{bmatrix}$ up to first powers of ε. For $a = Z$ and Y, these are directly given by (30) and (30a)

$$\lim \frac{1}{\varepsilon}\left(\begin{bmatrix} l & Z(\varepsilon) \\ \mu & \nu \end{bmatrix} - \delta_{\mu\nu}\right) = i\nu\delta_{\mu\nu} \tag{64}$$

and

$$\begin{aligned} &\lim \frac{1}{\varepsilon}\left(\begin{bmatrix} l & Y(\varepsilon) \\ \mu & \nu \end{bmatrix} - \delta_{\mu\nu}\right) \\ &\quad = \tfrac{1}{2}\sqrt{(l+\nu+1)(l-\nu)}\,\delta_{\mu\nu+1} - \tfrac{1}{2}\sqrt{(l+\nu)(l-\nu+1)}\,\delta_{\mu\nu-1}. \end{aligned} \tag{64a}$$

Both are, naturally, skew Hermitian. For M_y, one has from the analogue of (47) $Y_\beta = Z^{-1}X_\beta Z$

$$\begin{aligned} &\lim \frac{1}{\varepsilon}\left(\begin{bmatrix} l & X(\varepsilon) \\ \mu & \nu \end{bmatrix} - \delta_{\mu\nu}\right) \\ &\quad = \frac{i}{2}\sqrt{(l+\nu+1)(l-\nu)}\,\delta_{\mu\nu+1} + \frac{i}{2}\sqrt{(l+\nu)(l-\nu+1)}\,\delta_{\mu\nu-1}. \end{aligned} \tag{64b}$$

Hence, by (63)

$$\begin{aligned} M_x\psi_\nu^l &= \tfrac{1}{2}i\sqrt{(l+\nu+1)(l-\nu)}\,\psi_{\nu+1}^l + \tfrac{1}{2}i\sqrt{(l+\nu)(l-\nu+1)}\,\psi_{\nu-1}^l \\ M_y\psi_\nu^l &= \tfrac{1}{2}\sqrt{(l+\nu+1)(l-\nu)}\,\psi_{\nu+1}^l - \tfrac{1}{2}\sqrt{(l+\nu)(l-\nu+1)}\,\psi_{\nu-1} \\ M_z\psi_\nu^l &= i\nu\psi_\nu^l \end{aligned} \tag{65}$$

and

$$\begin{aligned} J_{-1}\psi^l_\nu &= 2^{-1/2}\sqrt{(l+\nu)(l-\nu+1)}\psi^l_{\nu-1} \\ J_0\psi^l_\nu &= \nu\psi^l_\nu \\ J_1\psi^l_\nu &= -2^{-1/2}\sqrt{(l+\nu+1)(l-\nu)}\psi^l_{\nu+1}. \end{aligned} \tag{65a}$$

Another, at least partial derivation of these formulas will be given at the end of this section.

For $t=2$, one obtains in a way similar to that which led to (61) that

$$\begin{aligned} q_{-2} &= -\tfrac{1}{2}(x-iy)^2 & q_2 &= -\tfrac{1}{2}(x+iy)^2 \\ q_{-1} &= -z(x-iy) & q_1 &= z(x+iy) \\ & q_0 = 6^{-1/2}(x^2+y^2-2z^2) \end{aligned} \tag{66}$$

form the components of a self-adjoint set of irreducible tensor operators of rank 2. A simpler derivation of (66) will be given below.

Let us calculate now an expression of the form $(\psi^l_\nu, t_\tau\psi^l_\nu)$. Because of the unitary character of the P_R and the invariance of the sets ψ^l, ψ^l, t as expressed in (50) and (53), we have

$$\begin{aligned} (\psi^l_\kappa, t_\tau\,\psi^{l'}_\lambda) &= (P_R\psi^l_\kappa, P_R t_\tau P_{R^{-1}} P_R\psi^{l'}_\lambda) \\ &= \begin{bmatrix} l & R \\ \kappa' & \kappa \end{bmatrix}^* \begin{bmatrix} t & R \\ \tau' & \tau \end{bmatrix} \begin{bmatrix} l' & R \\ \lambda' & \lambda \end{bmatrix} (\psi^l_{\kappa'}, t_{\tau'}\psi^{l'}_{\lambda'}) \\ &= \begin{pmatrix} l & \\ \mu' & \kappa' \end{pmatrix} \begin{bmatrix} l & R \\ \mu' & \mu \end{bmatrix} \begin{pmatrix} l & \\ \mu & \kappa \end{pmatrix}^* \begin{bmatrix} t & R \\ \tau' & \tau \end{bmatrix} \begin{bmatrix} l' & R \\ \lambda' & \lambda \end{bmatrix} (\psi^l_{\kappa'}, t_{\tau'}\psi^{l'}_{\lambda'}). \end{aligned}$$

Integration of both sides over all elements of the group yields because of (5)

$$\begin{aligned} (\psi^l_\kappa, t_\tau\psi^{l'}_\lambda) &= \begin{pmatrix} l & t & l' \\ \mu & \tau & \lambda \end{pmatrix} \begin{pmatrix} l & \\ \mu & \kappa \end{pmatrix}^* t_{ll'} \\ &= \begin{pmatrix} \kappa & t & l' \\ l & \tau & \lambda \end{pmatrix} t_{ll'} \end{aligned} \tag{67}$$

where

$$t_{ll'} = \begin{pmatrix} \kappa' & t & l' \\ l & \tau' & \lambda' \end{pmatrix}^* (\psi^l_{\kappa'}, t_{\tau'}\psi^{l'}_{\lambda'}) \tag{67a}$$

is independent of κ, τ, λ. Equation (67) gives a complete analytic expression for the consequences of the invariance of the quantum mechanical equations, as long as one deals only with one Hilbert space. It shows, in particular, that the matrix element (67) vanishes unless the direct product of the representations l and t contains l', i.e., unless l, t, l' can form a vector triangle. In this case, it permits us to compare the matrix elements of the various components τ of t with each other and also gives the ratios of these matrix elements for the

various κ, λ of the representation spaces. In the all-important case of a scalar operator, it shows that (67) vanishes unless $l = l'$ and, because of (7) and (32a), $\kappa = \lambda$. In this case, the value of (67) is independent of κ.

A simple calculation shows that if the set t_τ is self-adjoint

$$t_{ll'} = (-1)^{l-t-l'} t_{l'l}^*. \tag{67b}$$

Comparing (67) with (65a) one sees that the three-j-symbols the middle j of which is 1, are essentially equal to the matrix elements of the infinitesimal operators. This permits some simplification in the derivation of (65a).

14. *Transition to Contravariant Components as Time Inversion*

The form of (67) conforms with the covariant-contravariant formalism if we recollect that the covariant index κ on ψ_κ^l, the conjugate complex of which enters (67), is really a contravariant index. It may be worthwhile, however, to expound more fully at this point upon the connection between the covariant-contravariant formalism and the operation of time inversion than has been possible heretofore.

Let us denote the operation of time inversion by θ, an antiunitary operator. If the set of wave function ψ^j, i.e., the representation space j, is invariant under time inversion, $\theta\psi_\nu^j = \psi^{j\nu}$ (the appropriateness of this definition will become evident below) must be a linear combination of the ψ_ν^j. Since, furthermore, all operations of the underlying group commute with θ, application of θ to (50) gives for

$$P_R\psi^{j\nu} = P_R\theta\psi_\nu^j = \theta \begin{bmatrix} j & R \\ \mu & \nu \end{bmatrix} \psi_\mu^j = \begin{bmatrix} j & R \\ \mu & \nu \end{bmatrix}^* \psi^{j\mu} \tag{68}$$

(note that θ in antiunitary, i.e., $\theta a^\nu \psi_\nu = a^{\nu *}\theta\psi_\nu$). One sees that the ψ_ν^j transform by the conjugate complex representation. It follows then with the aid of (8) that

$$\begin{pmatrix} & j & \\ \nu & & \nu' \end{pmatrix} \psi^{j\nu'}$$

transforms under the P_R like ψ_ν^j. If it is not equal to ψ_ν^j, they can differ only by a factor of modulus 1 and a change of the base vectors ψ_ν^j by the square root of this factor will render

$$\psi_\nu^j = \begin{pmatrix} & j & \\ \nu & & \nu' \end{pmatrix} \psi^{j\nu'} = \begin{pmatrix} & j & \\ \nu & & \nu' \end{pmatrix} \theta\psi_\nu^j \tag{69}$$

or

$$\psi^{j\nu} = \theta\psi_\nu^j = \psi_{\nu'}^j \begin{pmatrix} & j & \\ \nu' & & \nu \end{pmatrix}^*. \tag{69a}$$

A base which satisfies (69a) may be called real and the preceding consideration, based on the existence of the operator θ, shows that it is always possible to choose the base real. This operator is simply complex conjugation in the simple Schrödinger theory. However, j is necessarily an integer in this case. If one includes the spin dependence of the wave function, the operation θ is somewhat more complicated but permits the definition of the contravariant vectors $\psi^{j\nu}$ and of a real base also if j is half integer.

If the ϕ^j of (49) is invariant with respect to time inversion, the vector with the components $a^{j\nu}$ will be called real. Since θ is an antiunitary operator, we have, for $\theta\phi$

$$\theta\phi^j = a^{j\nu *}\theta\psi^j_\nu = a^{j\nu *}\psi^j_{\nu'}\begin{pmatrix} j & \\ \nu' & \nu \end{pmatrix}^* \tag{70}$$

This will be equal to $a^{j\nu}\psi^j_\nu$ if

$$a^{j*}_\nu = \begin{pmatrix} j & \\ \nu & \nu' \end{pmatrix}^* a^{j\nu'*} = a^{j\nu} \tag{71}$$

i.e., if the covariant and the contravariant components of a^j are conjugate complex. It is easy to see that (71) can be satisfied only if j is integer, i.e., there are no real vectors in the spaces of half integer representations. This corresponds to the well known fact that no state of a system with an odd number of spinor particles can be invariant with respect to time inversion. The above consideration and Eq. (68) to (71) give a more precise meaning to the connection between the operation of time inversion in quantum mechanics and the covariant-contravariant transformations.

Whatever was said above about the reality of wave functions and their invariance with respect to time inversion applies *mutatis mutandis* also to operators t. It is more appropriate, however, to call an operator Q which satisfies the equation

$$Q\theta = \theta Q \tag{72}$$

"even" (with respect to time inversion) rather than "real" because the term "real" is used for operators synonymously with "self-adjoint." It remains true, of course, that the energy operator is even in the absence of magnetic fields.

15. *Union of Two Quantum Mechanical Systems; Nine-j-Symbols*

The states of the union of two quantum mechanical systems are best described in a Hilbert space which is the direct product of the Hilbert spaces of the two systems which are to be united. The axes of this Hilbert space can be characterized by a pair of the symbols which characterize the axes of the Hilbert spaces to be united. Thus if ψ^l and χ^k_κ are the axes of these latter

Hilbert spaces, $\psi^l_\lambda \chi^k_\kappa$ will be the axes of their direct product. We shall assume that the two Hilbert spaces undergo similar transformations when subjected to the operators P_R and that the transformation of the product space is given by

$$P_R \psi^l_\lambda \chi^k_\kappa = \begin{bmatrix} l & R \\ \lambda' & \lambda \end{bmatrix} \begin{bmatrix} k & R \\ \kappa' & \kappa \end{bmatrix} \psi^l_{\lambda'} \chi^k_{\kappa'}. \tag{73}$$

Quite apart from the aforementioned problem, one may be interested in the problem of the transformation properties of products such as those that occur in (73). We wish to form, in particular, linear combinations of these products which transform by irreducible representations. On the basis of our covariant-contravariant formalism, it seems reasonable to consider

$$\begin{aligned} \Psi^j_\iota &= [j]^{1/2} \begin{pmatrix} j & \lambda & \kappa \\ \iota & l & k \end{pmatrix} \psi^l_\lambda \chi^k_x \\ &= [j]^{1/2} \begin{pmatrix} j & \\ \iota & \iota' \end{pmatrix} \begin{pmatrix} j & l & k \\ \iota' & \lambda & \kappa \end{pmatrix}^* \psi^l_\lambda \chi^k_\kappa. \end{aligned} \tag{74}$$

The last line follows by (10b) and (10c). Applying (73) to this equation, one obtains in fact

$$P_R \Psi^j_\iota = \begin{bmatrix} j & R \\ \iota' & \iota \end{bmatrix} \Psi^j_{\iota'}. \tag{74a}$$

The factor $[j]^{1/2}$ was introduced in (74) in order to give Ψ^j_ι unit length. Conversely, one can express the $\psi^l_\lambda \chi^k_\kappa$ in terms of the Ψ

$$\psi^l_\lambda \chi^k_\kappa = \sum_j [j]^{1/2} \begin{pmatrix} j & l & k \\ \iota & \lambda & \kappa \end{pmatrix} \Psi^j_{\iota'} \begin{pmatrix} j & \\ \iota' & \iota \end{pmatrix}^*. \tag{74b}$$

The expression following the three-j-symbol is the contravariant component $\Psi^{j\iota}$.

The above calculus describes the vector addition model. The ψ^l and χ^k can be the wave functions of two particles: Ψ is then the wave function of the system consisting of both. The possible j values of this range from $|l-k|$ to $l+k$. In another application $\psi^l = \psi^L$ is the positional wave function, and $\chi^k = \chi^S$ the spin function; $\Psi^j = \Psi^J$ is then the total wave function. J ranges from $|L-S|$ to $L+S$.

Concepts very similar to the above can be introduced also for operators. Let r_ρ be the components of an irreducible tensor operator of rank r which acts in the Hilbert space of the ψ, and s_σ have a similar significance in the Hilbert space of the χ. In other words,

$$r_\rho \psi^l_\lambda \chi^k_\kappa = (r_\rho \psi^l_\lambda) \chi^k_\kappa; \qquad s_\sigma \psi^l_\lambda \chi^k_\kappa = \psi^l_\lambda (s_\sigma \chi^k_\kappa). \tag{75}$$

One can define then, in analogy to (74), the operators

$$t_\tau = [t]^{1/2} \begin{pmatrix} t & \\ \tau & \tau' \end{pmatrix} \begin{pmatrix} t & r & s \\ \tau' & \rho & \sigma \end{pmatrix}^* r_\rho s_\sigma . \tag{76}$$

One easily convinces oneself that the t form a set of irreducible tensor operators of rank t. In fact, the proof thereof does not assume that the r and s commute—as follows from (75)—the t are components of a tensor operator even if they do not, as long as the order of the r and s is the same in each term of (76). We shall return to this point later.

The commutative nature of the r and s is necessary, however, to establish the fact that if those form self-adjoint sets, the t will be self-adjoint also.

$$\begin{aligned} t_\tau^\dagger &= [t]^{1/2} \begin{pmatrix} t & \\ \tau & \tau' \end{pmatrix}^* \begin{pmatrix} t & r & s \\ \tau' & \rho & \sigma \end{pmatrix} s_\sigma^\dagger r_\rho^\dagger \\ &= [t]^{1/2} \begin{pmatrix} t & \\ \tau & \tau' \end{pmatrix}^* \begin{pmatrix} t & r & s \\ \tau' & \rho' & \sigma' \end{pmatrix} \begin{pmatrix} r & \\ \rho & \rho' \end{pmatrix}^* \begin{pmatrix} s & \\ \sigma & \sigma' \end{pmatrix}^* s_\sigma r_\rho \\ &= [t]^{1/2} \begin{pmatrix} t & r & s \\ \tau & \rho & \sigma \end{pmatrix}^* s_\sigma r_\rho = \begin{pmatrix} t & \\ \tau' & \tau \end{pmatrix}^* t_{\tau'} . \end{aligned} \tag{76a}$$

The second line follows by (54), the last one by the remark after (11a) and (76) together with the commutative nature of the r and s.

We shall now calculate expressions similar to (67), however with the Ψ^j instead of the ψ^j, and assume the form (76) for t. We shall encounter in the calculation expressions of the form

$$\begin{aligned} (\psi_\lambda^l \chi_\kappa^k, r_\rho s_\sigma \psi_{\lambda'}^{l'} \chi_{\kappa'}^{k'}) &= (\psi_\lambda^l, r_\rho \psi_{\lambda'}^{l'})(\chi_\kappa^k, s_\sigma \chi_{\kappa'}^{k'}) \\ &= \begin{pmatrix} l & r & l' \\ \lambda'' & \rho & \lambda' \end{pmatrix} \begin{pmatrix} k & s & k' \\ \kappa'' & \sigma & \kappa' \end{pmatrix} \begin{pmatrix} l & \\ \lambda'' & \lambda \end{pmatrix}^* \begin{pmatrix} k & \\ \kappa'' & \kappa \end{pmatrix}^* r_{ll'} s_{kk'} . \end{aligned} \tag{77}$$

The first line follows from (75) and the definition of the scalar product in the product of two Hilbert spaces; the second line is a consequence of (67). Using now (74) and (76), we can express the $t_{jj'}$ of

$$(\Psi_\iota^j, t_\tau \Psi_{\iota'}^{j'}) = \begin{pmatrix} j & t & j' \\ \iota'' & \tau & \iota' \end{pmatrix} \begin{pmatrix} j & \\ \iota'' & \iota \end{pmatrix}^* t_{jj'} \tag{67a}$$

in terms of $r_{ll'} s_{kk'}$

$$t_{jj'} = ([j][t][j'])^{1/2} \begin{Bmatrix} j & t & j' \\ l & r & l' \\ k & s & k' \end{Bmatrix} r_{ll'} s_{kk'} . \tag{78}$$

The expression obtained for the left side of (67a) by the direct substitution of (74) and of (76) for $\Psi^{j'}$, Ψ^j, and t, respectively, contains five three-j-symbols, neither of which is, however, the one which appears at the right side.

In order to give the expression directly obtained the form which appears on the right side, one has to apply, first, (17) to two of the three-j-symbols (it does not matter which) in order to exchange some of their columns. After this, (22) can be applied twice to give the form (67a). One obtains in this way for the nine-j-symbol of (78)

$$\begin{Bmatrix} j & t & j' \\ l & r & l' \\ k & s & k' \end{Bmatrix} = \sum_m (-1)^{2m}[m] \begin{Bmatrix} j & l & k \\ s & k' & m \end{Bmatrix} \begin{Bmatrix} t & r & s \\ l & m & l' \end{Bmatrix} \begin{Bmatrix} j' & l' & k' \\ m & j & t \end{Bmatrix}. \tag{78a}$$

There are, of course, numerous symmetry and orthogonality relations for (78a) which can be obtained either from (78a) through the relations of Section 3 or even more easily from the alternate expression for the nine-j-symbol

$$\begin{Bmatrix} j_{11} & j_{12} & j_{13} \\ j_{21} & j_{22} & j_{23} \\ j_{31} & j_{32} & j_{33} \end{Bmatrix} = \begin{pmatrix} j_{11} & j_{12} & j_{13} \\ \lambda_{11} & \lambda_{12} & \lambda_{13} \end{pmatrix} \begin{pmatrix} j_{21} & j_{22} & j_{23} \\ \lambda_{21} & \lambda_{22} & \lambda_{23} \end{pmatrix} \begin{pmatrix} j_{31} & j_{32} & j_{33} \\ \lambda_{31} & \lambda_{32} & \lambda_{33} \end{pmatrix}$$
$$\cdot \begin{pmatrix} j_{11} & j_{21} & j_{31} \\ \lambda_{11} & \lambda_{21} & \lambda_{31} \end{pmatrix}^* \begin{pmatrix} j_{12} & j_{22} & j_{32} \\ \lambda_{12} & \lambda_{22} & \lambda_{32} \end{pmatrix}^* \begin{pmatrix} j_{13} & j_{23} & j_{33} \\ \lambda_{13} & \lambda_{23} & \lambda_{33} \end{pmatrix}^*. \tag{78b}$$

The most direct derivation of (78b) starts from the equation, mentioned after (78), which expresses $(\Psi^j_\iota, t_\tau \Psi^{j}_{\iota'})$ directly in terms of (74) and (76) and which contains five six-j-symbols. One can calculate from this equation

$$\begin{pmatrix} j & & \\ & \iota'' & \iota \end{pmatrix} \begin{pmatrix} j & t & j' \\ \iota'' & \tau & \iota' \end{pmatrix}^* (\Psi^j_\iota, t_\tau \Psi^{j'}_{\iota'}) = t_{jj'}.$$

Comparison of this equation with (78) directly gives (78b). This equation shows that, in a sense, the nine-j-symbols are the most symmetric invariant expressions in the product space of the representation spaces. The fact that the invariant expression (78b) can be written in terms of the six-j-symbols, also follows from a theorem of Biedenharn.[10] We shall not discuss the symmetry and othogonality relation of the nine-j-symbols in detail.

There are a few special cases in which (78) assumes a much simpler form. In the most important of these t is a scalar operator, i.e., an invariant under the elements of the group. This is, of course, possible only if r and s have the same rank and the scalar is given in this case, on account of (76) and (7), (13)

$$S = (-)^{2r}[r]^{-1/2} \begin{pmatrix} r & \\ & \sigma \\ \rho & \end{pmatrix}^* r_\rho s_\sigma. \tag{79}$$

Inserting $t = 0$ into (78a) causes the last six-j-symbol to vanish unless $j = j'$ (which is natural, since t is a scalar) and unless $m = l'$. Hence, only one

term of (78a) remains and the last two six-j-symbols are directly given by (18b). One obtains with (7)

$$(\Psi'^{j}_{\iota}, S\Psi^{j}_{\iota'}) = \delta_{\iota\iota'}[r]^{-1/2}(-1)^{j+k'+l+r}\begin{Bmatrix} j & k & l \\ r & l' & k' \end{Bmatrix} r_{ll'}s_{kk'}. \tag{79a}$$

The prime on the first Ψ' should indicate that although this Ψ' has the same j as the second, it need not be identical with it: it is linear combination of the $\psi^l\chi^k$ while Ψ' is a linear combination of the $\psi^{l'}\chi^{k'}$.

Another case in which the nine-j-symbol reduces to a six-j-symbol is in which the operator s (or the operator r) is a scalar. In this case

$$t_\tau = r_\tau s. \tag{80}$$

For $s=0$, the first two six-j-symbols vanish unless $k=k'$, $t=r$ (which is natural), and unless $m=l$. Hence the summation in (78a) again reduces to a single term and the first two six-j-symbols are again given by (18b). We have for the nine-j-symbol

$$\begin{Bmatrix} j & t & j' \\ l & t & l' \\ k & 0 & k \end{Bmatrix} = \frac{(-1)^{j'+t+l'+k}}{([k][t])^{1/2}}\begin{Bmatrix} j & t & j' \\ l' & k & l \end{Bmatrix} \tag{80a}$$

and for

$$t_{jj'} = (-1)^{j'+t+l+k}\left(\frac{[j'][j]}{[k]}\right)^{1/2}\begin{Bmatrix} j & t & j' \\ l' & k & l \end{Bmatrix} r_{ll'}s. \tag{80b}$$

With $s=1$, this formula gives the ratio $t_{j'j}/r_{l'l}$ of the matrix elements of an operator $r_\tau = t_\tau$ between the states Ψ' of the compound system and the states ψ of one of the components. A well known problem covered by (80b) is the contribution of the orbital (or of the spin) magnetic moment to the magnetic moment of the atom. In this case the two systems which are united to a common system are the orbital motion and the spins of the particles. The former is described by a coordinate wave function ψ, its l is usually denoted by L; the latter is described by a spin wave function χ, its k is usually denoted by S. The operator of the orbital momentum is the vector operator J of (62a), its matrix elements with respect to the ψ can be obtained directly from (65a). The problem is to determine the matrix elements of this operator with respect to the wave functions Ψ' of the united system, i.e., the total wave functions of the atom. This is given by (67) with the $t_{jj'}$, given by (80b). In fact for $t=1, j=j'=J$, $k=S$, and $l=l'=L$, the six-j-symbol in (80b) is essentially Landé's g-factor.

16. *Combinations of Irreducible Tensors*

It has been mentioned before that the law of combination for irreducible tensor operators which is given in (76) has much wider applications

than given in the preceding section. The expression given by (76) will be a tensor operator of rank t not only if r and s act in separate Hilbert spaces. It will be a tensor operator also if they are arbitrary irreducible tensor operators in the same Hilbert space; they do not even have to commute. This equation is therefore eminently suited to make tensor operators of a higher rank from lower rank tensor operators. Thus, for instance, the tensor operator of (66) can be obtained in a more simple way than the one used for its derivation by substituting in (76) the v of (61) for both r and s. The number t has to be set $t=2$, corresponding to the rank 2 of (66), both $r=s=1$ because of the vector operator character of the v of (61). A similar tensor operator can be constructed from the J_τ of (62a).

The tensor operator of rank 2 obtained from J_τ in this way, and the higher order tensor operators obtained by a repeated application of (76) to J_τ, permit a more natural and more symmetric formulation of several rather common questions. Thus, for instance, it is customary to define the quadrupole moment as the expectation value of the operator $-\sqrt{6}q_0$ for the wave function ψ_j^j (i.e., ψ_ν^j for $\nu=j$). This is then, by definition, the quadrupole moment of all ψ_r^j. Instead of this, one may consider the scalar operator [cf. (79)]

$$Q=5^{-1/2}\begin{pmatrix}2\\ \tau\rho\end{pmatrix}^* q_\rho J_\tau^2 \tag{81}$$

where J^2 is the aforementioned tensor operator

$$J_\tau^2=5^{1/2}\begin{pmatrix}&2&\\ \tau&&\tau'\end{pmatrix}\begin{pmatrix}2&1&1\\ \tau'&\rho'&\sigma'\end{pmatrix}^* J_{\rho'}J_{\sigma'} \tag{81a}$$

or

$$Q=\begin{pmatrix}2&1&1\\ \rho&\rho'&\sigma'\end{pmatrix}^* q_\rho J_{\rho'}J_{\sigma'}. \tag{81b}$$

The expectation value of this operator is the same for all states ψ_ν^j and, for a given j, its magnitude is proportional to the quadrupole moment as commonly defined.

References

1. G. Frobenius and I. Schur, *Sitzber. Deut. Akad. Wiss. Berlin, Kl. Math. Phys. Tech.*, p. 186 (1906).
2. Cf. F. D. Murnaghan, "Group Representations," Chapter 3 (Baltimore, Maryland, 1938).
3. The parametrization of the group and the form (29), (30a) of the representations is identical with that adopted in E. Wigner, "Gruppentheorie und ihre Anwendungen auf etc." (Braunschweig, 1931), Chapter XV. This chapter also gives the isomorphism of the three-dimensional rotation and two-dimensional unimodular unitary groups.

4. $(j' + \frac{1}{2}) \dotplus (j' - \frac{1}{2})$ denotes the sum of the representations $j' + \frac{1}{2}$ and $j' - \frac{1}{2}$ (i.e. the reducible representation containing $j' + \frac{1}{2}$ and $j' - \frac{1}{2}$). Cf. F. D. Murnaghan, Reference 2.
5. A collection of formulas for these quantities (which are essentially identical with the three-j-symbols) is given in "The Theory of Atomic Spectra" by E. U. Condon and G. H. Shortley (Cambridge Univ. Press, London and New York, 1935), Chapter III. Relations of this character are, of course, well known in quantum mechanics and many of the six-j-symbols are implicitly contained in the formulas for the relative intensities of spectral lines and the magnitude of the Zeeman effect (Hönl, Kronig, Rubinowicz). Cf. the monographs of Reference 3 and 5, particularly the latter.
6. E. P. Wigner, *Am. J. Math.*, **63**, 57, 1941. For the present article, the following definitions and theorems are of importance. An irreducible representation which can be transformed into real form is called an integer representation even if the form in which it is used is not real. An irreducible representation which is equivalent to its conjugate complex but cannot be transformed into real form is called a half integer representation. The representations $D^{(j)}$ of the three dimensional rotation group with integer j are examples for the former, those with half odd integer j, examples for the latter. The Kronecker product of two integer representations or of two half integer representations contains only integer representations; the Kronecker product of an integer and a half integer representation contains only half integer representations. A representation which occurs in the symmetrized part of the Kronecker product of an integer representation with itself, or in the antisymmetrized part of the Kronecker product of a half integer representation with itself, is called an even representation. Representations occurring in the antisymmetrized square of an integer representation, or the symmetrized square of a half integer representation are called odd representations. Both even and odd representations are integer; examples for the even and odd are again the $D^{(j)}$ with even and odd j. Some representations may not occur in the Kronecker product of any representation with itself; these are then neither even, nor odd. However, no representation can be both even and odd and every integer representation of the three-dimensional rotation group (or of the two-dimensional unitary group) is either even or odd. The identical representation's j is 0; it is an even representation.
7. G. Racah, *Phys. Rev.* **62**, 438 (1942); **78**, 622 (1950).
8. H. A. Jahn, *Proc. Roy. Soc.* **A205**, 192 (1951).
9. L. C. Biedenharn, Oak Ridge National Laboratory Report No. 1098 (1952).
10. L. C. Biedenharn, *J. Math. Phys.* **31**, 287 (1953).

Unimodular Matrices Homomorphic to Lorentz Transformations in $n \geq 2$ Spacelike Dimensions

A. W. Sáenz and E. P. Wigner

Zeitschrift für Naturforschung *22a*, 1293–1299 (1967)

(Z. Naturforschg. **22 a**, 1293—1299 [1967]; received 30 June 1967)

Dedicated to Professor Pascual Jordan *on the occasion of his 65th birthday*

We generalize the usual homomorphism between 2×2 unimodular matrices $B(\Lambda)$ and restricted 4×4 Lorentz matrices Λ to the case of one timelike and $n \geqq 2$ spacelike dimensions. For every such n which is even (odd), this generalization associates homomorphically to each restricted (orthochronous) $(n+1)$-dimensional Lorentz matrix a set of $N \times N$-dimensional unimodular matrices, where $N = 2^{n/2}$ or $2^{(n-1)/2}$, depending on whether n is even or odd. In the case $n \geqq 2$, we prove the theorem that, if $B(\Lambda)$ and $B(\Lambda^{\mathrm{T}})$ are any two such unimodular matrices associated with Λ and its adjoint Λ^{T}, respectively, then $B(\Lambda^{\mathrm{T}}) = \omega\, B(\Lambda)^{\dagger}$, where ω is an Nth root of unity and $\dagger$ means hermitean adjoint. We also prove that for $n > 3$ one can select these two unimodular matrices so that this equation holds with $\omega = 1$, but that no such selection is possible for $n = 2, 3$.

The purpose of the present paper is to establish a homomorphic correspondence between a subgroup of the unimodular matrices and the Lorentz transformations in $(n+1)$-dimensional space with one timelike and $n \geqq 2$ spacelike dimensions. This correspondence, which generalizes the well-known correspondence for $n = 3$ between 2×2 unimodular matrices and 4×4 restricted Lorentz transformations, is established in Section 1. We have devoted Section 2 to two proofs, which we believe to be simple, of a theorem on the unimodular matrix associated with the transpose of an $(n+1)$-dimensional Lorentz matrix for $n \geqq 2$. This theorem is essentially known for the case of three spatial dimensions [1, 2], the only accessible proof being, however, that in reference [2].

1. Homomorphism Between Lorentz Transformations and Unimodular Matrices for $n \geqq 2$

Consider a fixed irreducible set τ_k $(k = 1, 2, \ldots, n)$ of $n \geqq 2$ hermitean matrices which satisfy the equation [3]

$$\tau_k \tau_l + \tau_l \tau_k = 2\, \delta_{kl}\, 1 \quad (k, l = 1, 2, \ldots, n). \quad (1)$$

[1] See, for example, p. 41, Eq. (2.8) of the article of E. P. Wigner on "Unitary Representations of the Inhomogeneous Lorentz Group Including Reflections" in Group Theoretical Concepts and Methods in Elementary Particle Physics, Gordon and Breach, New York 1964.

[2] I. M. Gel'fand, R. A. Minlos, and Z. Ya. Shapiro, Representations of the Rotation and Lorentz Groups and Their Applications, Pergamon Press, New York 1963, p. 170—171.

[3] Even if not stated explicitly, the inequality $n \geqq 2$ is always supposed to hold in this paper. Lower case Greek and Latin indices run over $0, 1, \ldots, n$ and $1, 2, \ldots, n$, respectively, and all equations involving free indices of these two types should be understood to hold for all values in these respective ranges. Summations over Greek and Latin indices also run over the full ranges of these indices.

It is well known[4] that the τ_k and their products with positive and negative signs, are homomorphic to an abstract group of order 2^{n+1}. Exploiting this fact, one infers, in the first place, that for even n there is exactly one irreducible set $\{\tau_k\}$ $(k=1, 2, \ldots, n)$ within equivalence, the dimension of these matrices being $2^{n/2}$. Secondly, one concludes that there are precisely two such inequivalent irreducible sets for the case of odd n, the corresponding τ_k having dimension $2^{(n-1)/2}$. In the latter case, $\tau_n = \lambda_n \tau_1 \tau_2 \ldots \tau_{n-1}$ in one of these irreducible sets and $\tau_n = -\lambda_n \tau_1 \tau_2 \ldots \tau_{n-1}$ in the other one, where $\lambda_n = 1\,(i)$ if n is of the form $4l+1\,(4l+3)$, l being zero or a positive integer. If we chose the τ_k to be unitary, which is always possible because they generate a representation of a finite group, then the τ_k are also hermitean, since their squares are unity. This hermiticity property will be supposed to hold in what follows.

For each $(n+1)$-vector $(x^0, x^1, \ldots, x^n)$ we define the $N \times N$ matrix $X(x)$ as the linear combination

$$X(x) = \sum_\mu x^\mu \tau_\mu \tag{2}$$

of the matrices $\tau_0, \tau_1, \ldots, \tau_n$; τ_0 being the $N \times N$ unit matrix[4].

The symbol $\Lambda = \| \Lambda^\mu{}_\nu \|$ $(\mu, \nu = 0, 1, \ldots, n)$ will denote the matrix of a transformation leaving the form $(x^0)^2 - (x^1)^2 - \ldots - (x^n)^2$ invariant, so that

$$\Lambda G \Lambda^{\mathrm{T}} = G, \tag{3}$$

where $G = \| g_{\mu\nu} \|$ is a diagonal matrix with $g_{00} = 1$, $g_{11} = \ldots = g_{nn} = -1$ and ^{T} means transpose. In particular, $(\Lambda^{\mathrm{T}})^\mu{}_\nu = \Lambda^\nu{}_\mu$. Moreover, Λ will always refer either to an orthochronous LORENTZ transformation $(\Lambda^0{}_0 > 0)$ or to a restricted one $(\Lambda^0{}_0 > 0$ and $\det \Lambda = 1)$.

The $N \times N$ matrix $B(\Lambda)$ is defined as a unimodular solution of the equation

$$X(x') = X(\Lambda x) = B(\Lambda)\, X(x)\, B(\Lambda)^\dagger, \tag{4}$$

for every Λ belonging to the groups of LORENTZ matrices, to be specified below, for which such solutions exist. Here † denotes hermitean adjoint and

$$x'^\mu = (\Lambda x)^\mu = \sum_\nu \Lambda^\mu{}_\nu x^\nu .$$

[4] P. JORDAN and E. P. WIGNER, Z. Phys. 47, 631 [1928], especially pages 650 and 651. Although only the case of even n is explicitly considered in this reference, the methods of the cited two pages are easily applied to the case of odd $n \geqq 3$.

Eq. (4) is supposed to hold identically in the x^ν. Hence the familiar argument using the linear independence of the τ_k shows that (4) can be regarded as a shorthand for the $n+1$ equations

$$\sum_\mu \Lambda^\mu{}_\nu \tau_\mu = B(\Lambda)\, \tau_\nu\, B(\Lambda)^\dagger . \tag{5}$$

The matrices Λ for which (4) or (5) admits a unimodular solution are given by the two theorems, valid for $n \geqq 2$, which we now state.

Theorem 1. If n is even (odd), then $B(\Lambda)$ exists for each real orthochronous (restricted) $(n+1)$-dimensional LORENTZ *matrix Λ and is determined by* (5) *up to a multiplicative constant equal to an arbitrary N^{th} root of unity.*

That the matrices Λ in Theorem 1 are the only ones for which such unmodular solutions $B(\Lambda)$ exist is implied by

Theorem 2. If B is an arbitrary unimodular N-dimensional matrix having the property that all the $B\,\tau_\nu\,B^\dagger$ are linear combinations

$$B\,\tau_\nu\,B^\dagger = \sum_\mu L^\mu{}_\nu \tau_\mu \tag{5'}$$

of the τ_μ, then the unique matrix $\| L^\mu{}_\nu \|$ corresponding to B in the sense of (5′) *is a real $(n+1)$-dimensional orthochronous (restricted)* LORENTZ *matrix for even (odd) n.*

The set $\mathcal{B}_n$ of all matrices associated in the sense of (5′) [or (5)] with the group of orthochronous or restricted LORENTZ matrices in the respective cases of n even or n odd forms a group, as follows easily with the aid of (5′). Theorems 1 and 2 inform us that there is a 1-to-N onto mapping of the pertinent one of these two LORENTZ groups onto $\mathcal{B}_n$. Just as in the well known case $n=3$, this onto mapping is homomorphic because, if the unimodular matrices B_1 and B_2 correspond to the LORENTZ matrices Λ_1 and Λ_2, respectively, then it is clear, for example from (4), that $B_1 B_2$ corresponds to $\Lambda_1 \Lambda_2$.

The irreducibility of the set $\tau_1, \ldots, \tau_n$ entails that of $\mathcal{B}_n$. In fact, $\det \tau_k = -1$ for $n = 2, 3$, as is well known, and $\det \tau_k = 1$ for $n > 3$, as follows from Eq. (17) of the present paper and the evenness of $N/2$ for $n > 3$. Since the τ_k are two-dimensional if $n = 2, 3$, we therefore see that the matrices $i\tau_k$ are unimodular for these two values of n. Moreover, for $n \geqq 2$ these matrices satisfy the equations

$$(i\tau_k)\, \tau_\nu (i\tau_k)^\dagger = \tau_k \tau_\nu \tau_k = \mp \tau_\nu$$

because of the hermiticity of the τ_k and (1). Hence we conclude that the obviously irreducible set $\{i\tau_k\}$

$(k=1, \ldots, n)$ is a subset of $\mathcal{B}_n$ if $n=2, 3$, and hence that $\mathcal{B}_2$ and $\mathcal{B}_3$ are irreducible. Similarly, if $n>3$, one can show that the irreducible set $\{\tau_k\}$ of n unimodular matrices τ_k is in $\mathcal{B}_n$, so that $\mathcal{B}_n$ is irreducible for such n. Conversely, the irreducibility of $\mathcal{B}_n$ entails that of $\{\tau_k\}$ for $n \geqq 2$.

Let us now prove the stated theorems.

Proof of Theorem 1.

We begin by showing that $B(\Lambda)$ exists for each restricted LORENTZ matrix Λ for $n \geqq 2$.

Every such Λ of $n+1$ dimensions can be written as a product

$$\Lambda = R_2 \Lambda_k R_1 , \tag{6}$$

where R_1 and R_2 are proper spatial rotations and Λ_k is a pure LORENTZ transformation along one of the spatial axes x^k $(k=1, 2, \ldots, n)$. It suffices to show that (5) has a unimodular solution for each LORENTZ matrix which is of one of these two types to conclude that (5) also has such a solution for any given restricted Λ, since because of (6) and a previous remark the unimodular matrix

$$B(R_2)\, B(\Lambda_k)\, B(R_1)$$

corresponds to this Λ in the sense of (5).

We first prove that to the pure LORENTZ transformation

$$\begin{aligned} x'^0 &= \cosh \chi\, x^0 + \sinh \chi\, x^k , \\ x'^k &= \sinh \chi\, x^0 + \cosh \chi\, x^k , \\ x'^l &= x^l \quad (l \neq 0, k), \end{aligned} \tag{7}$$

with an arbitrary χ, there corresponds the unimodular matrix

$$B_k = \cosh \tfrac{1}{2} \chi\, \tau_0 + \sinh \tfrac{1}{2} \chi\, \tau_k \tag{8}$$

in the sense of (5).

To establish the unimodularity of B_k, we observe that Eqs. (1) state that each τ_k has unit square and anticommutes with each τ_l with $l \neq k$. Therefore, if this inequality obtains, $\tau_l = -\tau_k \tau_l \tau_k = -\tau_k \tau_l \tau_k^{-1}$, whence

$$\operatorname{Trace} \tau_l = -\operatorname{Trace} \tau_k \tau_l \tau_k^{-1} = -\operatorname{Trace} \tau_l = 0.$$

Since $\tau_k^2 = 1$, the eigenvalues of τ_k can only be ± 1. In view of the tracelessness of τ_k, half of these eigenvalues, i. e., $N/2$ of them, are equal to $+1$ and the remaining half are -1. Consequently, because of the hermitean nature of τ_k, there exists a unitary matrix U_k such that

$$\tau_k = U_k \left\| \begin{array}{c:c} 1 & 0 \\ \hdashline 0 & -1 \end{array} \right\| U_k^\dagger , \tag{9}$$

where 0 and 1 denote the $\frac{1}{2} N$-dimensional zero and unit matrices, respectively. From (8), (9), and $\tau_0 = 1$,

$$\begin{aligned} B_k &= U_k \left\{ \cosh \tfrac{1}{2} \chi \left\| \begin{array}{c:c} 1 & 0 \\ \hdashline 0 & 1 \end{array} \right\| + \sinh \tfrac{1}{2} \chi \left\| \begin{array}{c:c} 1 & 0 \\ \hdashline 0 & -1 \end{array} \right\| \right\} U_k^\dagger \\ &= U_k \left\| \begin{array}{c:c} e^{\chi/2} \cdot 1 & 0 \\ \hdashline 0 & e^{-\chi/2} \cdot 1 \end{array} \right\| U_k^\dagger , \end{aligned} \tag{10}$$

so that B_k is unitarily equivalent to a unimodular matrix and is therefore unimodular.

We proceed to show that B_k corresponds to Λ_k. From (9), the hermiticity of B_k, and the fact that $\tau_0 = \tau_k^2 = 1$, we can obtain by direct calculation:

$$\begin{aligned} B_k \tau_0 B_k^\dagger = B_k^2 &= (\cosh \tfrac{1}{2} \chi\, \tau_0 + \sinh \tfrac{1}{2} \chi \tau_k)^2 \\ &= \cosh \chi\, \tau_0 + \sinh \chi\, \tau_k , \end{aligned} \tag{11 a}$$

and similarly

$$B_k \tau_k B_k^\dagger = \sinh \chi\, \tau_0 + \cosh \chi\, \tau_k , \tag{11 b}$$

$$B_k \tau_l B_k^\dagger = \tau_l \quad (l \neq k). \tag{11 c}$$

From (11 a) to (11 c), B_k is seen to obey the $n+1$ Eqs. (5) for the pure LORENTZ transformation (7).

Next we deal with a rotation by any angle φ, which takes place in the plane containing the spatial axes x^k and x^l $(k \neq l)$:

$$\begin{aligned} x'^k &= x^k \cos \varphi - x^l \sin \varphi , \\ x'^l &= x^k \sin \varphi + x^l \cos \varphi , \\ x'^\mu &= x^\mu \quad (\mu \neq k, l). \end{aligned} \tag{12}$$

We assert that the matrix

$$B_{kl} = \cos \tfrac{1}{2} \varphi\, \tau_0 + i \sin \tfrac{1}{2} \varphi\, i^{-1} \tau_k \tau_l \tag{13}$$

is unimodular and corresponds to this rotation.

That B_{kl} is unimodular can be shown by combining the facts that the matrix $i^{-1} \tau_k \tau_l$ $(k \neq l)$ in (13) is hermitean and that it is also, because of (1), of unit square and traceless, with procedures parallel to those just employed in proving the unimodularity of B_k. Computations analogous to those carried out in the case of B_k serve to verify that B_{kl} is indeed a solution of (5) pertaining to the spatial rotation (12).

We now show that, for even n, (5) has a unimodular solution $B(\Lambda)$ for every orthochronous Λ, and not merely for each restricted Λ. This is in sharp contrast with the situation obtaining for odd $n \geqq 3$, in which case (5) has such a unimodular solution only if Λ is a restricted LORENTZ matrix, as will be proved later on in this section.

It is well known that each improper orthochronous Λ can be written, for example, as the product

$$\Lambda = i_n \Lambda_0 \tag{14}$$

of the spatial reflection i_n:

$$\begin{aligned} x'^n &= -x^n, \\ x'^\mu &= x^\mu \ (\mu = 0, 1, \ldots, n-1), \end{aligned} \tag{15}$$

and a restricted LORENTZ matrix Λ_0. Since $B(\Lambda_0)$ exists, as is shown by our previous discussions, we only need to prove the existence of a unimodular matrix $B^{(n)}$ which corresponds to i_n for each even n to deduce that for every n of this type (5) possesses the unimodular solution $B(\Lambda) = B^{(n)} B(\Lambda_0)$ for every orthochronous Λ. Such a $B^{(n)}$ is given by

$$\begin{aligned} B^{(2)} &= i\tau_1, \\ B^{(n)} &= \tau_1 \tau_2 \ldots \tau_{n-1} \quad , n \geqq 4 \end{aligned} \tag{16}$$

for every even n [5].

To prove that the matrices $B^{(n)}$ in (16) are unimodular, one notices that

$$\det \tau_k = (-1)^{N/2}, \tag{17}$$

an equation which holds because $N/2$ of the eigenvalues of τ_k are 1 and the remaining $N/2$ are -1. Cobining (16) and (17) with the fact that $N/2$ is odd for $n = 2$, 3 and even for $n \geqq 4$, the asserted unimodularity property follows immediately.

For even n, one easily sees from (1) that the $B^{(n)}$ in (16) satisfy the equations

$$\begin{aligned} B^{(n)} \tau_\mu B^{(n)\dagger} &= \tau_\mu \quad (\mu = 0, 1, \ldots, n-1), \\ B^{(n)} \tau_n B^{(n)\dagger} &= -\tau_n \ , \end{aligned} \tag{18}$$

which are necessary and sufficient conditions for these $B^{(n)}$ to correspond to i_n.

For odd n, on the other hand, because of

$$\tau_n = \lambda_n \tau_1 \tau_2 \ldots \tau_{n-1},$$

there can be no $B^{(n)}$ to satisfy (18). The first equation (18), for $\mu = 0$, shows that $B^{(n)}$ would have to be unitary. This unitarity and the rest of the same equation shows that $B^{(n)}$ would have to commute with all $\tau_1, \ldots, \tau_{n-1}$. It cannot anticommute then with $\tau_n = \lambda_n \tau_1 \tau_2 \ldots \tau_{n-1}$.

We complete the proof of Theorem 1 by showing that, if the unimodular matrices B_1 and B_2 correspond to the same LORENTZ matrix Λ, then B_1 and B_2 are equal to within a factor of an Nth root of unity. Our method of proving this assertion is parallel to the one applied in reference [2] to the special case $n = 3$.

One easily finds from (5) and the nonsingularity of B_1 and B_2:

$$B_1^{-1} B_2 \tau_\mu = \tau_\mu B_1^\dagger B_2^{\dagger -1} \quad (\mu = 0, 1, \ldots, n). \tag{19}$$

Setting $\mu = 0$ in (19), we infer that

$$B_1^{-1} B_2 = B_1^\dagger B_2^{\dagger -1}$$

and hence that

$$B_1^{-1} B_2 \tau_k = \tau_k B_1^{-1} B_2 \quad (k = 1, 2, \ldots, n). \tag{19'}$$

However, as was proved in reference [4], the irreducible set $\{\tau_k\}$ $(k = 1, 2, \ldots, n)$ generates an irreducible representation of a certain finite abstract group. Therefore, (19′) entails that $B_1^{-1} B_2$ commutes with all the matrices of this representation and thus that it is a multiple $\omega 1$ of the $N \times N$ unit matrix, i. e., that

$$B_2 = \omega B_1 . \tag{20}$$

Taking the determinant of both sides of (20), the unimodularity of B_1 and B_2 yields that $\omega^N = 1$, so that B_1 and B_2 are related in the stated manner.

Proof of Theorem 2.

We first dispose of the easy matters of the uniqueness and reality of the matrix $L \equiv \| L^\mu{}_\nu \|$ in (5′). The first property follows immediately from the linear independence of the matrices τ_μ. The reality of L can be proved by taking the hermitean adjoint of both sides of (5′) and using the equation $\tau_\mu^\dagger = \tau_\mu$, thus obtaining

$$\sum_\mu \overline{L}^\mu{}_\nu \tau_\mu = \sum_\mu L^\mu{}_\nu \tau_\mu ,$$

where $^-$ means complex conjugation. Hence $\overline{L}^\mu{}_\nu = L^\mu{}_\nu$ because of the linear independence of the τ_μ. Incidentally, for each n, the hermiticity of the τ_μ can also be shown to be necessary to insure the reality

[5] The existence of a unimodular matrix $B^{(n)}$ corresponding to i_n for even n for a given irreducible set $\{\tau_k\}$ $(k = 1, 2, \ldots, n)$ can also be established as follows. For every n of this type, we adjoin to this set the matrix $\tau_{n+1} = \lambda_{n+1} \tau_1 \tau_2 \ldots \tau_n$ and hence obtain an irreducible set $\{\tau_k\}$ $(k = 1, 2, \ldots, n+1)$ of $n+1$ matrices of dimension $2^{n/2}$ fulfilling (1). [Plainly, (1) would not be obeyed by all these $n+1$ matrices if n had been assumed to be odd.] Now, we know that there exists for each even n a $2^{n/2}$ dimensional, unimodular matrix $B^{(n)}$ satisfying (5) for this last set of τ_k in the case of the restricted LORENTZ transformation (spatial rotation) in $n+1$ spacelike dimensions given by $x'^\mu = x^\mu$ $(\mu = 1, 2, \ldots, n-1)$, $x'^n = -x^n$, $x'^{n+1} = -x^{n+1}$. Such $B^{(n)}$ are immediately seen to obey (18) and therefore to correspond to i_n.

of the matrix L corresponding to each unimodular N-dimensional matrix B.

Lemma 1. The matrix $X(y) = \Sigma y^\mu \tau_\mu$ is positive (negative) definite if and only if the $(n+1)$-vector y is timelike and $y^0 > 0$ $(y^0 < 0)$. It is nonsingular and indefinite if and only if y is spacelike. It is semidefinite if and only if y is a null vector.

Proof. If y is timelike it can be written as $y = \Lambda' \cdot (a, 0, 0, \ldots, 0)$, where $a \neq 0$ has the same sign as y^0 and Λ' is a restricted LORENTZ matrix. Hence, by Theorem 1, there exists a unimodular $B(\Lambda')$ such that

$$\begin{aligned} X(y) &= B(\Lambda')\, X((a, 0, 0, \ldots, 0))\, B(\Lambda')^\dagger \\ &= B(\Lambda')\, (a\tau_0)\, B(\Lambda')^\dagger = a\, B(\Lambda')\, B(\Lambda')^\dagger, \end{aligned} \tag{21a}$$

so that $X(y)$ is positive (negative) definite if a, and therefore y^0, is positive (negative).

If y is spacelike, there is a restricted LORENTZ matrix Λ'' such that $y = \Lambda'' \cdot (0, b, 0, \ldots, 0)$, and therefore a unimodular matrix $B(\Lambda'')$ in terms of which

$$\begin{aligned} X(y) &= B(\Lambda'')\, X((0, b, 0, \ldots, 0))\, B(\Lambda'')^\dagger \\ &= b\, B(\Lambda'')\, \tau_1\, B(\Lambda'')^\dagger . \end{aligned} \tag{21 b}$$

Because the characteristic values of τ_1 are 1 and -1, (21 b) implies that $X(y)$ is nonsingular and indefinite for such spacelike y.

If y is a null vector, a restricted LORENTZ matrix Λ''' and a unimodular $B(\Lambda''')$ can be found such that $y = \Lambda''' \cdot (c, c, 0, \ldots, 0)$. Therefore,

$$\begin{aligned} X(y) &= B(\Lambda''')\, X((c, c, 0, \ldots, 0))\, B(\Lambda''')^\dagger \\ &= c\, B(\Lambda''')\, (\tau_0 + \tau_1)\, B(\Lambda''')^\dagger \end{aligned} \tag{21 c}$$

in this case. Now, the characteristic values of $\tau_0 + \tau_1$ are 0 and 2, so that (21 c) informs us that $X(y)$ is semidefinite for each null vector y. One can prove this property of $\tau_0 + \tau_1$ by first observing that (1) implies that $(\tau_0 + \tau_1)^2 = 2(\tau_0 + \tau_1)$, whence $\tau_0 + \tau_1$ has at least one of the numbers 0, 2 as characteristic values. Not all of the characteristic values of the matrix $\tau_0 + \tau_1$ are the same for, if they were, it would be a multiple of the unit matrix and would thus commute with all the τ_k, contrary to (1), so that both 0 and 2 are eigenvalues of $\tau_0 + \tau_1$.

Lemma 2. For each $(n+1)$-vector y,

$$\det X(y) = (y^2)^{N/2}, \tag{22}$$

where $y^2 \equiv (y^0)^2 - (y^1)^2 - (y^2)^2 - \ldots - (y^n)^2$. (23)

Proof. Suppose that y is spacelike. Then we infer from (21 b), the unimodularity of $B(\Lambda'')$ in this equation, (17), and (23), that

$$\begin{aligned} \det X(y) &= b^N \det \tau_1 = b^N (-1)^{N/2} = (-b^2)^{N/2} \\ &= (y^2)^{N/2}. \end{aligned}$$

A similar, simpler, proof yields (22) when y is timelike.

Proceeding as in the previous two cases, one obtains for each null vector y, employing (21 c) and the singular nature of $\tau_0 + \tau_1$:

$$\det X(y) = c^N \det(\tau_0 + \tau_1) = 0,$$

so that (22) also holds in this case.

Returning to Theorem 2, let B denote an arbitrary unimodular matrix for which there exists a matrix L such that (5′) is fulfilled, so that the equation

$$B\, X(y)\, B^\dagger = X(L\, y) \tag{24}$$

holds identically for this B and L. Applying Lemma 2 to (24), one concludes that

$$((L\, y)^2)^{N/2} = (y^2)^{N/2},$$

that is,

$$(L\, y)^2 = \omega(L, y)\, y^2, \tag{25}$$

where $\omega(L, y)$ is an $\frac{1}{2} N^{\text{th}}$ root of unity [Actually, $\omega(L, y) = \pm 1$ since L is a real matrix].

Eq. (25) entails trivially that

$$(L\, y)^2 = y^2 \tag{26}$$

for any null vector y. We wish to show that (26) also obtains for each timelike and each spacelike y. By virtue of (25), this is equivalent to proving that $(L\, y)^2$ is timelike or spacelike if y^2 has one of these respective properties.

Let us apply Lemma 1 to a timelike vector y having $y^0 > 0$. Then $X(y)$ is positive definite and so is $X(L\, y) = B\, X(y)\, B^\dagger$, since B is nonsingular, which implies that $L\, y$ is timelike and $(L\, y)^0 > 0$. Similarly, if y is timelike and $y^0 < 0$, we conclude that then $L\, y$ is timelike and $(L\, y)^0 < 0$. The fact that $L\, y$ is spacelike when y has this last property is proved analogously.

Hence, we have shown that L is a real homogeneous LORENTZ transformation. Moreover, $L^0{}_0 > 0$, i. e., L is orthochronous, since $(L\, y)^0$ has the same sign as y^0 for each timelike y.

It remains to be established that, for every odd $n \geqq 3$, the matrix L corresponding in the sense of (5′) to each unimodular $N \times N$ matrix B is a restricted LORENTZ matrix. This will certainly be the

case if there does not exist for any such n an $N \times N$ unimodular matrix which satisfies (5′) when L is chosen to be the improper LORENTZ matrix i_n pertaining to the spatial reflection (15). This was, however, established after (18).

2. Generalization to $n \geqq 2$ of a Theorem on the Unimodular Matrices Associated with the Transpose of a Lorentz Matrix

Whe shall present two proofs of the following theorem:

Theorem 3. One has

$$B(\Lambda^{\mathrm{T}}) = \omega\, B(\Lambda)^{\dagger} \tag{27}$$

for every orthochronous or restricted $(n+1)$-dimensional LORENTZ *matrix Λ in the respective cases of even n or of odd $n \geqq 3$, where ω is an Nth root of unity.*

In these proofs and in all of the subsequent discussions, symbols such as ω_0, ω_1, ω_2, ω', ω'' will serve to denote N^{th} roots of unity.

First Proof

Since, if the unimodular matrices $B(\Lambda_1)$ and $B(\Lambda_2)$ correspond to the respective LORENTZ matrices Λ_1, Λ_2, then $B(\Lambda_1 \Lambda_2)$ corresponds to $\Lambda_1 \Lambda_2$, one finds from Theorem 1

$$B(\Lambda_1 \Lambda_2) = \omega_1 B(\Lambda_1)\, B(\Lambda_2). \tag{28 a}$$

Similarly, $$B(\Lambda^{-1}) = \omega_2 B(\Lambda)^{-1}. \tag{28 b}$$

If theorem 3 holds for Λ_1 and Λ_2, it also holds for $\Lambda_1 \Lambda_2$. Indeed, its validity for Λ_1 and Λ_2 entails

$$\begin{aligned} B((\Lambda_1 \Lambda_2)^{\mathrm{T}}) &= B(\Lambda_2^{\mathrm{T}} \Lambda_1^{\mathrm{T}}) = \omega' B(\Lambda_2^{\mathrm{T}})\, B(\Lambda_1^{\mathrm{T}}) \\ &= \omega'' B(\Lambda_2)^{\dagger} B(\Lambda_1)^{\dagger} \\ &= \omega'' (B(\Lambda_1)\, B(\Lambda_2))^{\dagger} \\ &= \omega''' B(\Lambda_1 \Lambda_2)^{\dagger}, \end{aligned}$$

where we have used (28 a) and elementary properties of the roots of unity. Hence (27) holds for the product $\Lambda = \Lambda_1 \Lambda_2, \ldots, \Lambda_r$ of any finite number r of factors if it holds for each of these factors.

But each restricted or improper orthochronous LORENTZ matrix Λ can be expressed as a product (6) of two proper spatial rotations and a pure LORENTZ transformation Λ_k or as a product (14) of a spatial reflection matrix i_n and a restricted LORENTZ matrix, respectively. Hence it suffices to prove (27) for the case when Λ is equal to a proper spatial rotation R, or to Λ_k or i_n.

Let $\Lambda = R$. Applying (4) to the vector $x = (1, 0, 0, \ldots, 0)$, or setting $\nu = 0$ in (5), we find that $x' = R\,x = x$ and $X(x) = X(x') = \tau_0 = 1$. Therefore

$$B(R)\, B(R)^{\dagger} = 1,$$

i. e., $B(R)$ is unitary. In view of this fact, the property $R^{\mathrm{T}} = R^{-1}$, and (28 b),

$$B(R^{\mathrm{T}}) = B(R^{-1}) = \omega_0 B(R)^{-1} = \omega_0 B(R)^{\dagger},$$

so that (27) holds for such rotations.

If $\Lambda = \Lambda_k$ or $\Lambda = i_n$, then $\Lambda^{\mathrm{T}} = \Lambda$. Moreover, the unimodular matrix B_k in (8) corresponding to Λ_k and the unimodular matrix $B^{(n)}$ in (16) pertaining to i_n for even n have the properties

$$B_k^{\dagger} = B_k \quad \text{and} \quad B^{(n)\dagger} = \pm B^{(n)},$$

where the $+$ $(-)$ sign obtains if $n = 4m+2$ ($n = 2$ or $n = 4m$), for $m = 1\ 2, \ldots$. Consequently, if $\Lambda = \Lambda_k$ or $\Lambda = i_n$, we infer by employing Theorem 1 that

$$B(\Lambda^{\mathrm{T}}) = B(\Lambda) = \omega_0' B(\Lambda)^{\dagger},$$

so that (27) also obtains for these two types of Λ. Therefore (27) is valid for arbitrary orthochronous (restricted) Λ for even n (odd $n \geqq 3$).

Second Proof

This proof is based on the equation

$$X(G\,x) = x^2\, X(x)^{-1}, \tag{29}$$

where G is the diagonal matrix in (3) with diagonal elements $1, -1, \ldots, -1$ and x^2 is defined in (23). Eq. (29) holds for any $(n+1)$-vector x for which $x^2 \neq 0$.

In order to prove that such x satisfy this equation, we employ (1), (2), and the equations $g_{00} = 1$, $g_{kk} = -1$ to write

$$\begin{aligned} X(x)\, X(G\,x) &= (\sum_{\mu} x^{\mu} \tau_{\mu})\, (\sum_{\nu} g_{\nu\nu}\, x^{\nu} \tau_{\nu}) \\ &= \tfrac{1}{2} \sum_{\mu,\nu} x^{\mu} x^{\nu} (g_{\nu\nu}\, \tau_{\mu} \tau_{\nu} + g_{\mu\mu}\, \tau_{\nu} \tau_{\mu}) \\ &= \tfrac{1}{2} \sum_{k,l} x^k x^l (-\tau_k \tau_l - \tau_l \tau_k) \\ &+ \tfrac{1}{2} \sum_{k} x^k x^0 (\tau_k \tau_0 - \tau_0 \tau_k) \\ &+ \tfrac{1}{2} \sum_{l} x^0 x^l (-\tau_l \tau_0 + \tau_0 \tau_l) \\ &+ \tfrac{1}{2} (x^0)^2 (\tau_0^2 + \tau_0^2) \\ &= [-\sum_{k} (x^k)^2 + (x^0)^2]\, 1 = x^2\, 1. \end{aligned}$$

Therefore, if $x^2 \neq 0$, $X(x)^{-1}$ exists and obeys (29).

To prove Theorem 3, we begin by using (3) and the property $G^2 = 1$ to write

$$\Lambda^{\mathrm{T}} = G\,\Lambda^{-1}\,G\,. \tag{3'}$$

We also observe that, since Λ^{-1} and G are LORENTZ transformations,

$$(\Lambda^{-1}\,G\,x)^2 = (G\,x)^2 = x^2\,, \tag{30}$$

so that, in particular, $\Lambda^{-1}\,G\,x$ and $G\,x$ are not null vectors if x is not a null vector.

Let Λ pertain to one of the LORENTZ groups specified in Theorem 3 and let $x^2 \neq 0$. We then obtain, by combining (29) with (3′), (4), (28 b), and (30):

$$\begin{aligned} X(\Lambda^{\mathrm{T}}\,x) &= X(G(\Lambda^{-1}\,G\,x)) \\ &= (\Lambda^{-1}\,G\,x)^2\,X(\Lambda^{-1}\,G\,x)^{-1} \\ &= x^2[B(\Lambda^{-1})\,X(G\,x)\,B(\Lambda^{-1})^{\dagger}]^{-1} \\ &= x^2[B(\Lambda)^{-1}\cdot x^2\,X(x)^{-1}\cdot B(\Lambda)^{\dagger-1}]^{-1} \\ &= x^2\,B(\Lambda)^{\dagger}\,X(x)\,(B(\Lambda)^{\dagger})^{\dagger}\,. \end{aligned} \tag{31}$$

Now, as is clear geometrically and easy to establish algebraically, each null $(n+1)$-vector can be written as the sum of two $(n+1)$-vectors which are not null. Hence, because of the linearity of $X(x)$ in x, (31) is also valid when $x^2 = 0$, and thus holds for all $(n+1)$-vectors x[6]. We therefore conclude that $B(\Lambda)^{\dagger}$ corresponds to Λ^{T}, and therefore that it differs at most by a factor of an Nth root of unity from $B(\Lambda^{\mathrm{T}})$.

It may be of interest to point out that (27) allows one to prove very rapidly that $B(\Lambda)$ is unitary of and only if Λ is a spatial rotation, and that a necessary and sufficient condition for $B(\Lambda)$ to be hermitean or skew hermitean is that Λ be symmetric.

In conclusion, we shall prove that for $n = 2$, 3 it is impossible to select a $B(\Lambda)$ so that (27) obtains with the + sign for all relevant Λ, but that for $n > 3$ one can make such a selection.

If B is a unimodular matrix corresponding to Λ, then $B^{\dagger}$ corresponds to Λ^{T} by Theorem 3. Hence the most general unimodular matrices $B(\Lambda)$ and $B(\Lambda^{\mathrm{T}})$ corresponding to Λ and Λ^{T} are

$$B(\Lambda) = \omega'\,B\,, \tag{32 a}$$

$$B(\Lambda^{\mathrm{T}}) = \omega''\,B^{\dagger} \tag{32 b}$$

and therefore

$$B(\Lambda^{\mathrm{T}}) = \omega''\,B^{\dagger} = \omega''(\bar{\omega}'\,B(\Lambda))^{\dagger} = \omega'\,\omega''\,B(\Lambda)^{\dagger}\,.$$

[6] The fact that (31) holds for null $(n+1)$-vectors x can also be proved by a continuity argument.

If $\Lambda^{\mathrm{T}} \neq \Lambda$, ω' and ω'' can be chosen independently, and therefore (27) can be made to hold with $\omega = 1$ for each $n \geqq 1$ for such Λ.

Let us now investigate the case when $\Lambda = \Lambda^{\mathrm{T}}$. If B is an arbitrary unimodular matrix corresponding to a Λ of this type, then clearly

$$B^{\dagger} = \omega'''\,B\,. \tag{33}$$

If Λ is any LORENTZ transformation, symmetric or not, such that (5) has unimodular solutions corresponding to it then among these solutions there is one of the form

$$B = \sum_{A} c_A\,\tau_A\,, \tag{34}$$

where the summation runs over a set of linearly independent matrices τ_A which are products of the τ_k and which therefore have the property $\tau_A^{\dagger} = \pm\,\tau_A$, and all the coefficients c_A are real and are evidently not all zero. One can prove (34) by combining the fact that a restricted or improper orthochronous Λ can be written in one of the respective forms (6) or (14) with (8) and with the simple result that to each proper spatial rotation there corresponds a unimodular matrix which is the product of matrices of the type (13). From Eqs. (33) and (34),

$$\sum_{A} \pm\,c_A\,\tau_A = \omega'''\sum_{A} c_A\,\tau_A\,,$$

whence $\omega''' = 1$, so that (33) becomes

$$B^{\dagger} = \pm\,B \tag{33'}$$

for a B of the form (34). Because of (32 a) and (33′), the most general matrix $B(\Lambda)$ corresponding to a symmetric Λ obeys

$$\begin{aligned} B(\Lambda^{\mathrm{T}}) &= B(\Lambda) = \omega'\,B = \pm\,\omega'\,B^{\dagger} \\ &= \pm\,\omega'(\bar{\omega}'\,B(\Lambda))^{\dagger} = \pm\,(\omega')^2 B(\Lambda)^{\dagger}\,. \end{aligned} \tag{35}$$

Let Λ be a spatial rotation (12) in the $x^k - x^l$ plane with $\varphi = \pi$. Then $\Lambda^{\mathrm{T}} = \Lambda$ and the unimodular matrix $B = \tau_k\,\tau_l = -B^{\dagger}$ corresponds to this Λ. Hence (39) obtains with a − sign for this Λ. Suppose now that $n = 2$, 3. Then $N = 2$, so that $(\omega')^2 = 1$, and therefore (35) yields $B(\Lambda^{\mathrm{T}}) = -B(\Lambda)^{\dagger}$ for the rotation in question, no matter how $B(\Lambda)$ is chosen.

If $N \geqq 4$, N is divisible by 4, and hence i is an Nth root of unity in this case. Therefore we can always choose ω' in (35) for any n of this type so that $\pm\,(\omega')^2 = 1$, i. e., so that (27) holds with $\omega = 1$ when $\Lambda^{\mathrm{T}} = \Lambda$. Consequently, (27) can be made to hold with this value of ω for all Λ of interest when $n \geqq 4$.

On a Generalization of Euler's Angles

E. P. Wigner

In: Group Theory and Its Applications. Academic Press, New York 1968, pp. 119–129

I. Origin of the Problem 119
II. Summary of Results 122
III. Proof 124
IV. Corollary 128
References 129

I. Origin of the Problem

The considerations to be presented arose from an attempt to extend **R** matrix theory so as to account for peripheral reactions more generally.

As is well known, **R** matrix theory deals principally with nuclear reactions. It divides the configuration space into an internal and an external region. It introduces on the boundary B separating these two regions a suitable orthogonal set of functions. The **R** matrix itself provides a relation

$$v_s = \sum_t R_{st} d_t \tag{1}$$

between the expansion coefficients v_s of the value of a wave function on B and the expansion coefficients d_t of the normal derivative of that same wave function. The validity of Eq. (1) is the condition that the wave function can be continued into the internal region.

The purpose of the theory is to obtain a relation

$$e_s = \sum_t S_{st} i_t \tag{2}$$

between the amplitudes of the outgoing waves e_s, and those of the incoming waves i_t, in those parts of the configuration space in which the distance between colliding or separating particles is large. In the standard form of the theory,*

* There are several excellent reviews of this theory, including those by Lane and Thomas (*1*), by Vogt (*2*), and by Breit (*3*). For the standard theory of direct reactions, see, for instance, the article by Austern (*4*).

we assume a one-to-one correspondence between "channels" (which characterize the type, state of excitation, and relative angular momentum of the colliding or separating particles) and the aforementioned orthogonal system on B in such a way that each pair of expansion coefficients v_s, d_s depends only on the corresponding e_s, i_s

$$\begin{aligned} v_s &= \mathscr{E}_s e_s + \mathscr{I}_s i_s \\ d_s &= \mathscr{E}'_s e_s + \mathscr{I}'_s i_s. \end{aligned} \tag{3}$$

$\mathscr{E}_s$ and $\mathscr{E}'_s$ are the value and the derivative, at the boundary B, of the radial part of the outgoing wave carrying unit current in channel s. $\mathscr{I}_s$ and $\mathscr{I}'_s$ have the same significance with respect to the incoming wave. Equation (3) implies that the interaction in the external region is very simple; according to the usual assumption it can be represented by a potential acting between the colliding or separating particles as entities. It causes no change in the structure of the particles nor a transformation between them.

If interactions do cause such changes at considerable separation of these particles, and if we do not wish to extend the internal region to a correspondingly large part of the configuration space, we can take into account the interaction at larger distances as peripheral reactions by generalizing Eq. (3). This will render the usual properties of the $\mathbf{R}$ of Eq. (1) more pronounced. Hence, we shall write, instead of Eq. (3),

$$\begin{aligned} i_s &= \sum_t \alpha_{st} d_t + \beta_{st} v_t, \\ e_s &= \sum_t \gamma_{st} d_t + \delta_{st} v_t, \end{aligned} \tag{4}$$

expressing, this time, $\mathbf{i}$ and $\mathbf{e}$ in terms of $\mathbf{v}$ and $\mathbf{d}$. Actually, because of time-inversion invariance, there are some relations between the matrices that enter (4) ($\boldsymbol{\alpha} = \boldsymbol{\gamma}^*$, $\boldsymbol{\beta} = \boldsymbol{\delta}^*$), but we shall disregard these. The problem that prompted the considerations to be presented is the determination of those relations between the matrices $\boldsymbol{\alpha}$, $\boldsymbol{\beta}$, $\boldsymbol{\gamma}$, $\boldsymbol{\delta}$ that follow from the unitary nature of the theory. This can be expressed by the postulate that the probability current entering the external region from the internal one be equal to the current flowing out from it at very large distances.

The aforementioned unitarity condition gives the equation

$$\sum i(v_s d_s^* - d_s v_s^*) = \sum |e_s|^2 - |i_s|^2. \tag{5}$$

f the $\mathbf{e}$ and $\mathbf{i}$ satisfy the relations (4), this must be valid for arbitrary $\mathbf{v}$ and $\mathbf{d}$. 'he $\mathbf{v}$, $\mathbf{d}$, $\mathbf{e}$, $\mathbf{i}$ can be considered to be vectors with the scalar product

$$(\mathbf{v}, \mathbf{d}) = \sum v_s^* d_s \tag{6}$$

so that (5) can be given the form

$$i(\mathbf{d}, \mathbf{v}) - i(\mathbf{v}, \mathbf{d}) = (\mathbf{e}, \mathbf{e}) - (\mathbf{i}, \mathbf{i}). \tag{7}$$

This equation can be further simplified if we unite **d** and **v** into a vector **D** with twice the dimension of **v** or **d**, and similarly, unite **i** and **e** into a vector **I** of the same dimension as **D**. Then, (7) becomes

$$(\mathbf{D}, \mathbf{s}_y \mathbf{D}) = (\mathbf{I}, \mathbf{s}_z \mathbf{I}) \tag{8}$$

where

$$\mathbf{s}_y = \begin{pmatrix} 0 & i\mathbf{1} \\ -i\mathbf{1} & 0 \end{pmatrix}, \qquad \mathbf{s}_z = \begin{pmatrix} -\mathbf{1} & 0 \\ 0 & \mathbf{1} \end{pmatrix}, \tag{8a}$$

the matrices $\mathbf{1}$ in (8a) having the same dimension as the vectors **d**, **v**, etc. Finally, we write

$$\mathbf{D} = \mathbf{T}\mathbf{D}', \qquad \mathbf{T} = 2^{-1/2}\begin{pmatrix} \mathbf{1} & i\mathbf{1} \\ i\mathbf{1} & \mathbf{1} \end{pmatrix}, \tag{9}$$

the unitary $\mathbf{T}$ having been so chosen that (the dagger denotes Hermitian adjoint)

$$\mathbf{T}^\dagger \mathbf{s}_y \mathbf{T} = \mathbf{s}_z, \tag{9a}$$

which then gives, instead of Eq. (8)

$$(\mathbf{D}', \mathbf{s}_z \mathbf{D}') = (\mathbf{I}, \mathbf{s}_z \mathbf{I}). \tag{9b}$$

According to (4) **i** and **e** depend linearly on **v** and **d**. Hence, we can write

$$\mathbf{I} = \mathbf{A}\mathbf{D} = \mathbf{A}\mathbf{T}\mathbf{D}', \qquad \mathbf{A} = \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix}, \tag{10}$$

so that if we write $\mathbf{AT} = \mathbf{U}$, (9b) implies that $\mathbf{U}$ leaves the form $\mathbf{s}_z$ invariant: $\mathbf{U}^\dagger \mathbf{s}_z \mathbf{U} = \mathbf{s}_z$. It is also clear that in order to return from the relation

$$\mathbf{I} = \mathbf{U}\mathbf{D}' = \mathbf{U}\mathbf{T}^\dagger \mathbf{D} \tag{10a}$$

between **I** and **D** to a relation between **i**, **e**, **v**, and **d**, we will need $\mathbf{U}$ in the form decomposed into submatrices with half the dimension of $\mathbf{U}$.

The problem to be considered further is somewhat more general than the one, just outlined, to which the generalization of $\mathbf{R}$ matrix theory leads. Instead of giving a normal form of the submatrices of the $\mathbf{U}$ that satisfies the equation $\mathbf{U}^\dagger \mathbf{s}_z \mathbf{U} = \mathbf{s}_z$, we shall consider the matrices $\mathbf{U}$

$$\mathbf{U}^\dagger \mathbf{F} \mathbf{U} = \mathbf{F}, \tag{11}$$

which leave the form

$$\mathbf{F} = \begin{pmatrix} +\mathbf{1}_n & 0 \\ 0 & -\mathbf{1}_m \end{pmatrix} \tag{12}$$

invariant where $\mathbf{1}_n$ and $\mathbf{1}_m$ are unit matrices that may have different dimensions n and m, and we assume $n \geqslant m$. The dagger denotes Hermitian adjoint, the tilde ($\sim$) will denote the transpose. Similarly, we shall consider real matrices $\mathbf{R}$ that leave $\mathbf{F}$ invariant

$$\tilde{\mathbf{R}}\mathbf{F}\mathbf{R} = \mathbf{F}. \tag{13}$$

The matrices $\mathbf{U}$ and $\mathbf{R}$ that satisfies Eqs. (2) and (3) form, of course, well-known groups. The $\mathbf{R}$ here is, of course, not the $\mathbf{R}$-matrix of (1).

II. Summary of Results

The representations of $\mathbf{U}$ and $\mathbf{R}$ that will be arrived at show a great similarity to Euler's representation of rotations in three-dimensional space and the corresponding representation of two-dimensional unitary matrices

$$\mathbf{U} = \begin{pmatrix} u_1 & 0 \\ 0 & u_2 \end{pmatrix} \begin{pmatrix} c & s \\ -s & c \end{pmatrix} \begin{pmatrix} v_1 & 0 \\ 0 & v_2 \end{pmatrix} \tag{14}$$

where u_1, u_2, v_1, v_2 are numbers of modulus 1, c and s are real $c^2 + s^2 = 1$. It will be shown that there is a corresponding representation for unitary matrices of arbitrary dimensions, that is, for matrices that satisfy Eq. (11) with an $\mathbf{F}$ that is an $(n+m)$-dimensional unit matrix. The generalization of Eq. (14) reads in this case

$$\mathbf{U} = \begin{pmatrix} \mathbf{u}_1 & \mathbf{0} \\ \mathbf{0} & \mathbf{u}_2 \end{pmatrix} \begin{pmatrix} \mathbf{c} & \mathbf{s} \\ -\tilde{\mathbf{s}} & \mathbf{c}' \end{pmatrix} \begin{pmatrix} \mathbf{v}_1 & \mathbf{0} \\ \mathbf{0} & \mathbf{v}_2 \end{pmatrix} \tag{15}$$

where $\mathbf{u}_1$ and $\mathbf{v}_1$ are n-dimensional unitary matrices, $\mathbf{u}_2$ and $\mathbf{v}_2$ are m-dimensional and unitary, $\mathbf{c}$ and $\mathbf{c}'$ are real diagonal matrices, the last $n-m$ diagonal elements of $\mathbf{c}$ being 1, the first m identical with those of $\mathbf{c}'$. The $\mathbf{s}$ of Eq. (15) is a real $n \times m$ matrix. We shall call such a matrix diagonal if, for $n > m$, the last $n-m$ rows are 0, and if the matrix formed by the first m rows is diagonal. Similarly, an $m \times n$ matrix will be called diagonal if its transpose is diagonal; its last $n-m$ columns are then 0. In this sense, $\mathbf{s}$ is diagonal, as is also $\tilde{\mathbf{s}}$, and they satisfy the equation

$$\mathbf{c}^2 + \mathbf{s}\tilde{\mathbf{s}} = \mathbf{1}_n \tag{15a}$$

and hence

$$(\mathbf{c}')^2 + \tilde{\mathbf{s}}\mathbf{s} = \mathbf{1}_m. \tag{15b}$$

However, we shall be more interested in the $\mathbf{U}$ that satisfy Eq. (11) with the $\mathbf{F}$ of Eq. (12). For these, the representation to be obtained is

$$\mathbf{U} = \begin{pmatrix} \mathbf{u}_1 & \mathbf{0} \\ \mathbf{0} & \mathbf{u}_2 \end{pmatrix} \begin{pmatrix} \mathbf{C} & \mathbf{S} \\ \tilde{\mathbf{S}} & \mathbf{C}' \end{pmatrix} \begin{pmatrix} \mathbf{v}_1 & \mathbf{0} \\ \mathbf{0} & \mathbf{v}_2 \end{pmatrix}; \tag{16}$$

$\mathbf{u}_1$, $\mathbf{v}_1$, $\mathbf{u}_2$, $\mathbf{v}_2$ are again unitary with dimensions n and m, the $\mathbf{C}$, $\mathbf{C}'$ again real and diagonal, $\mathbf{C}'$ is m-dimensional, and the first m diagonal elements of $\mathbf{C}$ are identical with the corresponding elements of $\mathbf{C}'$, the rest 1. The $\mathbf{S}$ is again $n \times m$, again real and diagonal, and instead of (15a) and (15b) we now have

$$\mathbf{C}^2 - \mathbf{S}\tilde{\mathbf{S}} = \mathbf{1}_n \tag{16a}$$

$$\mathbf{C}'^2 - \tilde{\mathbf{S}}\mathbf{S} = \mathbf{1}_m. \tag{16b}$$

Evidently, the constituents of the right side of (16) are not uniquely determined by $\mathbf{U}$. This remains unchanged if the first and last factors are changed as follows:

$$\begin{aligned} \mathbf{u}_1 &\to \mathbf{u}_1\boldsymbol{\omega}, & \mathbf{v}_1 &\to \boldsymbol{\omega}^\dagger\mathbf{v}_1, \\ \mathbf{u}_2 &\to \mathbf{u}_2\boldsymbol{\omega}', & \mathbf{v}_2 &\to \boldsymbol{\omega}'^\dagger\mathbf{v}_2, \end{aligned} \tag{17}$$

where $\boldsymbol{\omega}$ and $\boldsymbol{\omega}'$ are unitary, $\boldsymbol{\omega}$ commutes with $\mathbf{C}$ (every diagonal matrix does), and $\boldsymbol{\omega}'$ is equal to $\boldsymbol{\omega}$ in those rows and columns in which the diagonal element of $\mathbf{C}$ is not 1, arbitrary otherwise. If no diagonal element of $\mathbf{S}$ is zero, only the last $n - m$ diagonal elements of $\mathbf{C}$ are 1 and $\boldsymbol{\omega}$ will split up into an $m \times m$ unitary matrix and an $(n - m) \times (n - m)$ unitary matrix, the remaining matrix elements being 0. In this case, $\boldsymbol{\omega}'$ is identical with the $m \times m$ submatrix of $\boldsymbol{\omega}$.

The representation to be obtained for the real matrices $\mathbf{R}$ is the same as (16) except that $\mathbf{u}_1$, $\mathbf{u}_2$, $\mathbf{v}_1$, $\mathbf{v}_2$ are real orthogonal matrices in this case. If the $-\mathbf{1}_m$ in (12) is replaced by $\mathbf{1}_m$ and $n = 2$, $m = 1$, we recognize Euler's original formula

$$\mathbf{R} = \begin{pmatrix} \cos\alpha & \sin\alpha & 0 \\ -\sin\alpha & \cos\alpha & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \cos\beta & 0 & -\sin\beta \\ 0 & 1 & 0 \\ \sin\beta & 0 & \cos\beta \end{pmatrix} \begin{pmatrix} \cos\gamma & \sin\gamma & 0 \\ -\sin\gamma & \cos\gamma & 0 \\ 0 & 0 & 1 \end{pmatrix}; \tag{14a}$$

$\mathbf{u}_2$ and $\mathbf{v}_2$ are, in this case, 1; $\mathbf{u}_1$ and $\mathbf{v}_1$ rotations by α and γ in the plane of $\mathbf{1}_2$; the diagonal elements of $\mathbf{C}$ are $\cos\beta$ and 1, whereas the only diagonal

element of $\mathbf{C}'$ is $\cos\beta$. The $\mathbf{S}$ is a 2×1 matrix, the second row of which is 0; $\tilde{\mathbf{S}}$ is its transpose, a 1×2 matrix.

It may be worthwhile to state explicitly the similarities, as well as the differences, between Euler's formulas (14) and (14a) and their generalizations (15) and (16). The first factors are in both cases elements of a subgroup. However, this subgroup is commutative in Euler's formulas (14) and (14a), and noncommutative in the generalizations (15) and (16). The elements of the same subgroup appear also as the last factor in both cases. The middle factors are elements of another subgroup; this is commutative in both cases.* Naturally, the generalized formulas apply to a much greater variety of groups than the original formulas.

III. Proof

Let us subdivide $\mathbf{U}$ in the same fashion in which $\mathbf{F}$ in (12) is subdivided

$$\mathbf{U} = \begin{pmatrix} \mathbf{U}_{11} & \mathbf{U}_{12} \\ \mathbf{U}_{21} & \mathbf{U}_{22} \end{pmatrix} \tag{18}$$

where $\mathbf{U}_{11}$ is an $n \times n$ matrix, $\mathbf{U}_{12}$ has n rows and m columns, etc. Equation (11) then reads

$$\mathbf{U}_{11}^{\dagger}\,\mathbf{U}_{11} = \mathbf{1}_n + \mathbf{U}_{21}^{\dagger}\,\mathbf{U}_{21} \tag{18a}$$

$$\mathbf{U}_{11}^{\dagger}\,\mathbf{U}_{12} = \mathbf{U}_{21}^{\dagger}\,\mathbf{U}_{22} \tag{18b}$$

$$\mathbf{U}_{12}^{\dagger}\,\mathbf{U}_{11} = \mathbf{U}_{22}^{\dagger}\,\mathbf{U}_{21} \tag{18c}$$

$$\mathbf{U}_{22}^{\dagger}\,\mathbf{U}_{22} = \mathbf{1}_m + \mathbf{U}_{12}^{\dagger}\,\mathbf{U}_{12}. \tag{18d}$$

Equation (18c) is just the adjoint to (18b) and need not be considered separately. Since $\mathbf{U}_{21}^{\dagger}\,\mathbf{U}_{21}$ is self-adjoint and positive semidefinite, it can be written as

$$\mathbf{U}_{21}^{\dagger}\,\mathbf{U}_{21} = \mathbf{v}_1^{\dagger}\,\mathbf{S}^2\,\mathbf{v}_1 \tag{19}$$

where $\mathbf{S}$ is a diagonal matrix with real, nonnegative diagonal elements; $\mathbf{v}_1$ is unitary. Further, it can be assumed that the diagonal elements of $\mathbf{S}$ form a nonincreasing sequence. Since the rank of $\mathbf{U}_{21}$ is at most m, at least $n - m$ of

* One can recognize a certain similarity between the decompositions considered here and Iwasawa's decomposition of an arbitrary matrix into a product of *two* matrices, one orthogonal, the other triangular. Cf., e.g., Helgason (*5*). In Iwasawa's decomposition also the matrices which can appear as the first factor are elements of a subgroup; the same applies to the matrices which can appear as the second factor.

the characteristic values of $\mathbf{U}_{21}^{\dagger}\,\mathbf{U}_{21}$ vanish. As a result, at least the last $n-m$ diagonal elements of $\mathbf{S}^2$ are zero, and as a rule only these diagonal elements will vanish. At any rate, $\mathbf{S}$ can be replaced by an $n \times m$ matrix and $\mathbf{S}\tilde{\mathbf{S}}$ substituted for $\mathbf{S}^2$ in (19). It follows from (18a) that

$$\mathbf{U}_{11}^{\dagger}\,\mathbf{U}_{11} = \mathbf{v}_1^{\dagger}(\mathbf{1}_n + \mathbf{S}\tilde{\mathbf{S}})\,\mathbf{v}_1 = \mathbf{v}_1^{\dagger}\,\mathbf{C}^2\,\mathbf{v}_1 \tag{19a}$$

where $\mathbf{C}$ is defined by (16a). Evidently, the diagonal elements of $\mathbf{C}$ also form a nondecreasing sequence, the last $n-m$ of them are 1.

Since $\mathbf{C}\mathbf{v}_1$ is nonsingular, we can write

$$\mathbf{U}_{11} = \mathbf{u}_1\,\mathbf{C}\mathbf{v}_1 \tag{20}$$

and, inserting this into (19a), we find that $\mathbf{u}_1$ is unitary. It now follows from (19) that the ranks of $\tilde{\mathbf{S}}\mathbf{v}_1$ and of $\mathbf{U}_{21}$ are the same. In fact, if $\tilde{\mathbf{S}}\mathbf{v}_1\boldsymbol{\varphi} = \mathbf{0}$, where $\boldsymbol{\varphi}$ is an n-dimensional vector, it follows from (19) that $\mathbf{U}_{21}^{\dagger}\,\mathbf{U}_{21}\boldsymbol{\varphi} = \mathbf{0}$, or

$$(\boldsymbol{\varphi}, \mathbf{U}_{21}^{\dagger}\,\mathbf{U}_{21}\,\boldsymbol{\varphi}) = (\mathbf{U}_{21}\,\boldsymbol{\varphi}, \mathbf{U}_{21}\,\boldsymbol{\varphi}) = 0$$

so that $\mathbf{U}_{21}\boldsymbol{\varphi} = 0$ also. Hence, the ranges of $\tilde{\mathbf{S}}\mathbf{v}_1$ and of $\mathbf{U}_{21}$ have equal dimensions and there is an $m \times m$ matrix $\mathbf{u}_2$ such that

$$\mathbf{U}_{21} = \mathbf{u}_2\,\tilde{\mathbf{S}}\mathbf{v}_1. \tag{20a}$$

Insertion of this into (19) shows that $\mathbf{u}_2$ is length preserving on the range of $\tilde{\mathbf{S}}\mathbf{v}_1$ and since it is arbitrary on the perpendicular complement of this range, it can be assumed to be unitary. It will then map the perpendicular complement of the range of $\tilde{\mathbf{S}}\mathbf{v}_1$ into the perpendicular complement of the range of $\mathbf{U}_{21}$.

The same argument that was just applied to (18a) can be applied to (18d). It gives, instead of (20) and (20a),

$$\mathbf{U}_{22} = \mathbf{u}_p\,\mathbf{C}'\,\mathbf{v}_2 \tag{21}$$

and

$$\mathbf{U}_{12} = \mathbf{u}_p'\,\tilde{\mathbf{S}}_p\,\mathbf{v}_2 \tag{21a}$$

where $\mathbf{u}_p$ and $\mathbf{v}_2$ are unitary, $\mathbf{C}'$ real diagonal, all three $m \times m$ matrices. $\mathbf{S}_p$ is $m \times n$, its last $n-m$ columns being 0, and it is real and diagonal in its first m columns; $\mathbf{u}_p'$ is $n \times n$ and can be assumed to be unitary. Instead of (16b) we now have

$$(\mathbf{C}')^2 - \mathbf{S}_p\,\tilde{\mathbf{S}}_p = \mathbf{1}_m. \tag{21b}$$

We assume again that the diagonal elements of $\mathbf{S}_p$ and $\mathbf{C}'$ form a nonincreasing sequence. The only difference from the preceding case is that, even if $n > m$, no diagonal element of $\mathbf{C}'$ needs to be 1.

We have so far reduced $\mathbf{U}$ to the form

$$\mathbf{U} = \begin{pmatrix} \mathbf{u}_1 \mathbf{C} \mathbf{v}_1 & \mathbf{u}_p' \tilde{\mathbf{S}}_p \mathbf{v}_2 \\ \mathbf{u}_2 \tilde{\mathbf{S}} \mathbf{v}_1 & \mathbf{u}_p \mathbf{C}' \mathbf{v}_2 \end{pmatrix}. \tag{22}$$

In order to verify (16) completely, we have yet to show that $\tilde{\mathbf{S}}_p = \mathbf{S}$, $\mathbf{u}_p = \mathbf{u}_2$, $\mathbf{u}_p' = \mathbf{u}_1$, and that the first m diagonal elements of $\mathbf{C}$ are the same as the diagonal elements of $\mathbf{C}'$. This must follow from (18b), which becones

$$\mathbf{v}_1^\dagger \mathbf{C} \mathbf{u}_1^\dagger \mathbf{u}_p' \tilde{\mathbf{S}}_p \mathbf{v}_2 = \mathbf{v}_1^\dagger \mathbf{S} \mathbf{u}_2^\dagger \mathbf{u}_p \mathbf{C}' \mathbf{v}_2 \tag{22a}$$

or

$$\mathbf{u}_1^\dagger \mathbf{u}_p' \tilde{\mathbf{S}}_p \mathbf{C}'^{-1} = \mathbf{C}^{-1} \mathbf{S} \mathbf{u}_2^\dagger \mathbf{u}_p. \tag{22b}$$

Let us denote $\mathbf{u}_1^\dagger \mathbf{u}_p' = \mathbf{w}_1$, $\mathbf{u}_2^\dagger \mathbf{u}_p = \mathbf{w}_2$; these are unitary. Similarly, we set $\tilde{\mathbf{S}}_p(\mathbf{C}')^{-1} = \mathbf{T}_1$, $\mathbf{C}^{-1}\mathbf{S} = \mathbf{T}$. These are $n \times m$ real diagonal matrices with nonnegative diagonal elements. Furthermore, if s_p is a diagonal element of $\mathbf{S}_p$, the corresponding diagonal element of $\mathbf{C}'$ is, because of (21b), $(1 + s_p^2)^{1/2}$ and hence the corresponding diagonal element to $\mathbf{T}_1$ is $s_p(1 + s_p^2)^{-1/2}$. Since the diagonal elements s_p form a nonincreasing sequence, the same is true of the diagonal elements $s_p(1 + s_p^2)^{-1/2}$ of $\mathbf{T}_1$. The same applies to $\mathbf{T}$, because of (16a) which was established by (19a), and because the diagonal elements of $\mathbf{S}$ were also assumed to form a nonincreasing sequence.

In terms of the $\mathbf{w}$ and $\mathbf{T}$, (22b) becomes

$$\mathbf{w}_1 \mathbf{T}_1 = \mathbf{T} \mathbf{w}_2 \tag{23}$$

and the adjoint of this is

$$\mathbf{T}_1^\dagger \mathbf{w}_1^\dagger = \mathbf{w}_2^\dagger \mathbf{T}^\dagger. \tag{23a}$$

The product of the last two equations is

$$\mathbf{T}_1^\dagger \mathbf{T}_1 = \mathbf{w}_2^\dagger \mathbf{T}^\dagger \mathbf{T} \mathbf{w}_2. \tag{23b}$$

The left side is an $m \times m$ real diagonal matrix; the right side is the transform, by the unitary matrix $\mathbf{w}_2$, of a similar diagonal matrix. It follows that the diagonal elements of $\mathbf{T}_1^\dagger \mathbf{T}_1$ and $\mathbf{T}^\dagger \mathbf{T}$ are pairwise equal and, since both form a nonincreasing sequence, the two diagonal matrices are equal:

$$\mathbf{T}_1^\dagger \mathbf{T}_1 = \mathbf{T}^\dagger \mathbf{T}. \tag{24}$$

Since, furthermore, the diagonal element of $\mathbf{T}_1^\dagger \mathbf{T}_1$ that corresponds to the diagonal element s_p of $\tilde{\mathbf{S}}_p$ is $s_p^2/(1+s_p^2)$, and the diagonal element of $\mathbf{T}^\dagger \mathbf{T}$ that corresponds to the diagonal element s of $\mathbf{S}$ is $s^2/(1+s^2)$, these two are equal. Since both s_p and s were assumed to be nonnegative, they are pairwise equal and we conclude

$$\tilde{\mathbf{S}}_p = \mathbf{S}. \tag{24a}$$

It then follows from (16a) and (21b) that the first m diagonal elements of $\mathbf{C}$ and $\mathbf{C}'$ (all of which are positive) are equal. The preceding is a somewhat roundabout version of the argument that, because of the positive semidefinite nature of $\mathbf{S}_p$, $\mathbf{C}'$, $\mathbf{T}_1$, and $\mathbf{T}_1^\dagger \mathbf{T}_1$, these are all single-valued functions of each other and the same applies to the quartet $\mathbf{S}$, $\mathbf{C}$, $\mathbf{T}$, and $\mathbf{T}^\dagger \mathbf{T}$. The argument was made in detail because $\mathbf{S}_p$, $\mathbf{T}_1$, $\mathbf{S}$, and $\mathbf{T}$ are not square matrices.

We have yet to prove that $\mathbf{u}_p = \mathbf{u}_2$, $\mathbf{u}_p' = \mathbf{u}_1$ or, rather, that $\mathbf{u}_p$ can be replaced by $\mathbf{u}_2$, $\mathbf{u}_p'$ by $\mathbf{u}_1$ without affecting the validity of (21) and (21a). For this purpose, we rewrite (23b) by means of (24)

$$\mathbf{w}_2 \mathbf{T}_1^\dagger \mathbf{T}_1 = \mathbf{T}_1^\dagger \mathbf{T}_1 \mathbf{w}_2; \tag{25}$$

that is, $\mathbf{w}_2$ commutes with $\mathbf{T}_1^\dagger \mathbf{T}_1$. It therefore also commutes with $\mathbf{C}'$ and $\tilde{\mathbf{S}}_p = \mathbf{S}$. Hence, (21) and (21a) will remain valid if we replace

$$\mathbf{u}_p \to \mathbf{u}_p \mathbf{w}_2^{-1}, \qquad \mathbf{v}_2 \to \mathbf{w}_2 \mathbf{v}_2, \qquad \mathbf{u}_p' \to \mathbf{u}_p' \mathbf{w}_2^{-1}. \tag{25a}$$

Since $\mathbf{w}_2 = \mathbf{u}_2^\dagger \mathbf{u}_p$, the new $\mathbf{u}_p$ will be $\mathbf{u}_p \mathbf{u}_p^\dagger \mathbf{u}_2 = \mathbf{u}_2$ and we now have

$$\mathbf{u}_p = \mathbf{u}_2. \tag{25b}$$

We next return to (22b). We note that since the last $n-m$ rows of $\mathbf{S}$ are 0, the first m rows of $\mathbf{C}$ are equal to those of $\mathbf{C}'$, all of them diagonal.

$$\mathbf{S}(\mathbf{C}')^{-1} = \mathbf{C}^{-1}\mathbf{S} = \mathbf{T}. \tag{26}$$

Hence, (22b) reads, with (24a), (25b), and (26),

$$\mathbf{u}_1^\dagger \mathbf{u}_p' \mathbf{T} = \mathbf{w}_1 \mathbf{T} = \mathbf{T}. \tag{26a}$$

We recall that $\mathbf{w}_1 = \mathbf{u}_1^\dagger \mathbf{u}_p'$. Let us assume that the first k diagonal elements of $\mathbf{S}$, and hence of $\mathbf{T}$, are different from 0. Then, (26a) means that the first k columns of $\mathbf{w}_1$ contain a 1 at the diagonal, the rest is 0. Since $\mathbf{w}_1$ is unitary, the same applies to the first k rows:

$$\mathbf{w}_1 = \begin{pmatrix} \mathbf{1}_k & \mathbf{0} \\ \mathbf{0} & \mathbf{w}_1' \end{pmatrix}. \tag{27}$$

The subdivision in (27) is k ($k \leqslant m$) and $n-k$. Since $\mathbf{S}$ is diagonal and only its first k diagonal elements are different from 0, we have

$$\mathbf{w}_1 \mathbf{S} = \mathbf{S}, \qquad \tilde{\mathbf{S}} \mathbf{w}_1^\dagger = \tilde{\mathbf{S}}, \tag{27a}$$

and multiplying the last equation by $\mathbf{w}_1$, we obtain

$$\tilde{\mathbf{S}} = \tilde{\mathbf{S}} \mathbf{w}_1. \tag{27b}$$

Since the diagonal elements of $\mathbf{C}$ are 1 wherever those of $\mathbf{S}$ are 0, with the same subdivision as in (27)

$$\mathbf{C} = \begin{pmatrix} \mathbf{C}_k & \mathbf{0} \\ \mathbf{0} & \mathbf{1}_{n-k} \end{pmatrix}. \tag{27c}$$

It follows that

$$\mathbf{w}_1 \mathbf{C} = \mathbf{C} \mathbf{w}_1. \tag{27d}$$

Hence, (20) and (20a) will remain valid if we replace

$$\begin{aligned} \mathbf{u}_1 &\to \mathbf{u}_1 \mathbf{w}_1, \\ \mathbf{v}_1 &\to \mathbf{w}_1^\dagger \mathbf{v}_1, \end{aligned} \tag{28}$$

as a result of which

$$\mathbf{u}_1 = \mathbf{u}_p'. \tag{28a}$$

As a result of (24a), (25b), (28a), the form (22) is now identical with (16).

The proofs of (15) and the modification of (16) in which $\mathbf{U}$ is real (and the $\mathbf{u}$ and $\mathbf{v}$ are also real) can be carried out in the same way. Actually, the proof carried out is very simple and the apparent complication arises only from the fact that the $\mathbf{S}$ is not a square matrix and we are less used to operating with general rectangular matrices than with square ones. For this reason also, the case $m = n$ originally considered by this writer, appears to be simpler.

IV. Corollary

As was mentioned before, the original objective of the preceding considerations concerned the theory of peripheral nuclear reactions, which will not be dealt with here. There is, however, one immediate consequence of the form (16) for complex and real matrices that leave $\mathbf{F}$ invariant that is worth mentioning. It relates to the characteristic values of $\mathbf{U}^\dagger \mathbf{U}$ (or of $\mathbf{U}\mathbf{U}^\dagger$). We obtain from (16)

$$\mathbf{U}^\dagger \mathbf{U} = \begin{pmatrix} \mathbf{v}_1^\dagger & \mathbf{0} \\ \mathbf{0} & \mathbf{v}_2^\dagger \end{pmatrix} \begin{pmatrix} \mathbf{C} & \mathbf{S} \\ \tilde{\mathbf{S}} & \mathbf{C}' \end{pmatrix}^2 \begin{pmatrix} \mathbf{v}_1 & \mathbf{0} \\ \mathbf{0} & \mathbf{v}_2 \end{pmatrix}. \tag{29}$$

Hence, $\mathbf{U}^\dagger \mathbf{U}$ is the unitary transform of

$$\begin{pmatrix} \mathbf{C} & \mathbf{S} \\ \tilde{\mathbf{S}} & \mathbf{C}' \end{pmatrix}^2 = \begin{pmatrix} \mathbf{C}^2 + \mathbf{S}\tilde{\mathbf{S}} & \mathbf{CS} + \mathbf{SC}' \\ \tilde{\mathbf{S}}\mathbf{C} + \mathbf{C}'\tilde{\mathbf{S}} & \tilde{\mathbf{S}}\mathbf{S} + (\mathbf{C}')^2 \end{pmatrix}. \tag{29a}$$

Since the first m diagonal elements of $\mathbf{C}$ and $\mathbf{C}'$ are identical, if we subdivide (29a) into sections m, $n-m$, and m, it assumes the form

$$\begin{pmatrix} (\mathbf{C}')^2 + \mathbf{S}_m^2 & \mathbf{0} & 2\mathbf{S}_m\mathbf{C}' \\ \mathbf{0} & \mathbf{1} & \mathbf{0} \\ 2\mathbf{S}_m\mathbf{C}' & \mathbf{0} & (\mathbf{C}')^2 + \mathbf{S}_m^2 \end{pmatrix} \tag{29b}$$

where $\mathbf{S}_m$ is the upper $m \times m$ section of $\mathbf{S}$. The secular equation of (29b) decomposes into m two-dimensional parts of the form

$$\begin{pmatrix} c^2 + s^2 & 2sc \\ 2sc & c^2 + s^2 \end{pmatrix} \tag{29c}$$

where s is a diagonal element of $\mathbf{S}$, $c = (1 + s^2)^{1/2}$, the corresponding diagonal element of $\mathbf{C}'$ (or of $\mathbf{C}$). Hence, $\mathbf{U}^\dagger \mathbf{U}$ has m pairs of characteristic values that are reciprocals of each other. It also has $n-m$ characteristic values 1. For $m = 0$ this is, or course, a trivial result.

In the case of real $\mathbf{U}$, we can further infer from (29), since the $\mathbf{v}$ are real, that the characteristic vectors of $\mathbf{U}^\dagger \mathbf{U}$ (or of $\mathbf{U}\mathbf{U}^\dagger$) can be assumed to be real.

REFERENCES

1. A. M. Lane and R. G. Thomas, *Rev. Mod. Phys.* **30**, 257 (1958).
2. E. Vogt, *in* "Nuclear Reactions," p. 215. North-Holland Publ., Amsterdam, 1959.
3. G. Breit, *in* "Encyclopedia of Physics," Vol. 41/1, p. 274. Springer Verlag, Berlin, 1959.
4. N. Austern, *in* "Fast Neutron Physics," Vol. IV, p. 1113. Wiley (Interscience), New York, 1963.
5. S. Helgason, "Differential Geometry and Symmetric Spaces," p. 219 ff. Academic Press, New York, 1962.

Condition That the Irreducible Representations of a Group, Considered as Representations of a Subgroup, Do Not Contain Any Representation of the Subgroup More Than Once

E. P. Wigner

In: Spectroscopic and Group Theoretic Methods in Physics (F. Bloch et al., eds.).
North-Holland, Amsterdam 1968, pp. 131–136

Origin of the problem

The interest in the condition contained in the title stems from the desirability to specify a definite form for all irreducible representations of a group, i.e. to specify a particular coordinate system, with some simple properties, in representation space. In order to do this, let us first decompose each representation space into orthogonal subspaces, each such subspace being the representation space of an irreducible representation of a subgroup. If these irreducible representations are all inequivalent, their subspaces are uniquely given and can be characterized by the corresponding irreducible representation of the subgroup. Once this is done, one can proceed in the same way with the subgroup and, if possible, continue until an abelian subgroup of the last group is reached. Hence, if there is a sequence of subgroups $G_1, G_2, \ldots$ of the original group G_0 such that G_{l+1} is a subgroup of G_l of the nature specified in the title, a coordinate axis in the space of an irreducible representation of G_0 can be specified by enumerating the irreducible representations of all the subgroups $G_1, G_2, \ldots$ to the representation spaces of which the vector belongs.

Let us consider, as an example, the symmetric group S_5 of the permutations of 5 letters. Since S_{n-1} is a subgroup of S_n of the nature given in the title, the sequence of $G_1, G_2, \ldots$ can be chosen as S_4, S_3, S_2; the last one

being abelian. Let us consider the representation $3+1+1$ of S_5. We then obtain the tree

$$\begin{array}{lccccccccc}
S_5 & & & & & 3+1+1 & & & & \\
 & & & & \swarrow & & \searrow & & & \\
S_4 & & & 2+1+1 & & & & & 3+1 & \\
 & & \swarrow & & \searrow & & & \swarrow & & \searrow \\
S_3 & 1+1+1 & & & 2+1 & & & 2+1 & & 3 \\
 & \downarrow & & \swarrow & & \searrow & \swarrow & & \searrow & \downarrow \\
S_2 & 1+1 & & 1+1 & & 2 & 1+1 & & 2 & 2
\end{array}$$

No two arrows issuing from the symbol of any representation end at identical symbols. A vector in the six dimensional space of $3+1+1$ can therefore be characterized uniquely by one of the branches of the tree, such as $2+1+1$, $2+1$, $1+1$.

Many of the classical groups have sequences of subgroups as given in the title so that coordinate systems can be introduced in the spaces of their representations in the manner given above; each axis of these coordinate systems will be labelled by irreducible representations of a sequence of subgroups, each subgroup of the sequence being contained in the preceding subgroup of the sequence.

Establishment of the condition

Denote elements of the subgroup by $p, q, r, \ldots$, elements of the full group by $P, Q, R, \ldots$. Let $\sum_p$ be the summation over all the elements of the subgroup, $\sum_P$ the summation over all the elements of the full group.

Consider vectors a in group space with components $a_E, \ldots, a_R, \ldots$, such that the matrix $\sum_R a_R\, D(R)$ commutes with all $D(s)$ of the subgroup and that this condition be satisfied for all irreducible (and hence also for all) representations D of the whole group. The vectors a form a linear manifold. It will be shown, furthermore, that if a matrix $c^{(0)}$ commutes with all $D^{(0)}(s)$ of a single irreducible representation, $c^{(0)}$ is contained among the matrices $\sum a_R\, D^{(0)}(R)$: the a_R being given by

$$a_R = \frac{l_0}{h} \sum_{ik} c_{ik}^{(0)}\, D^{(0)}\,(R^{-1})_{ki}, \tag{1}$$

l_0 is the dimension of $D^{(0)}$ and h the order of the full group. That the a of (1) is in the linear manifold means that

$$\sum_R a_R\, D^{(j)}(R) = \sum_R \frac{l_0}{h} \sum_{ik} c_{ik}^{(0)}\, D^{(0)}(R^{-1})_{ki}\, D^{(j)}(R) \tag{2}$$

commutes with all $D^{(j)}(s)$ for every j. This is evident since (2) is zero if $D^{(j)}s$ is not equivalent to $D^{(0)}$; it is $c^{(0)}$ if $D^{(j)}=D^{(0)}$. Hence, all matrices which commute with the $D^{(j)}(s)$, for every s but any single j, are contained in the set $\sum a_R\, D(R)$.

Let us determine now the linear manifold of the a. By definition

$$\sum_R a_R D(R)\, D(s) = \sum_R a_R D(s)\, D(R) \tag{3}$$

or

$$\sum a_R D(R) = \sum a_R D(sRs^{-1})$$

or

$$\sum_R a_R D(R) = \sum_R a_{s^{-1}Rs} D(R). \tag{4}$$

Since this must hold for every irreducible representation, and since these (or their matrix elements) form a complete set in group space

$$a_R = a_{s^{-1}Rs} \tag{5}$$

and this holds for every element s of the subgroup. We shall call the complex of elements which can be transformed into each other by an element of the subgroup, a subclass. Hence, the components of a in the direction of the elements which are in the same subclass are equal. A full base for the vectors a has as many members as there are subclasses in the whole group; a vector which has finite components in the direction of the subclass which contains the element Q and zero components in every other direction is

$$a_R = \sum_s \delta(sQs^{-1}R^{-1}) \tag{6}$$

where

$$\delta(P)\begin{cases} = 1 & \text{for } P = E = e, \text{ the unit element}, \\ = 0 & \text{otherwise}. \end{cases} \tag{6a}$$

Evidently, the product of two subclasses consists of full subclasses and every class of the whole group consists of full subclasses. A subclass is either contained in the subgroup or has no element therein.

Let us consider now the vectors b in group space which satisfy two conditions. First $\sum b_R\, D(R)$ commute with all $D(s)$ for every representation (this is the same condition which the vectors a satisfy), second $\sum b_R\, D(R)$ commute with all $\sum a_R\, D(R)$. The vectors b also form a linear manifold and because of the first condition (5) is valid for the b also,

$$b_R = b_{s^{-1}Rs}. \tag{7}$$

Furthermore, if there is a matrix $c^{(0)}$ which commutes with all $D^{(0)}(s)$ of a single representation, and with all matrices in the space of this representation which commute with all matrices commuting with all $D^{(0)}(s)$, it is contained among the $\sum b_R\, D^{(0)}(R)$. The corresponding b_R is given by the right side of (1) and the proof that the $\sum b_R\, D(R)$ have the necessary properties for all $D^{(j)}$ follow from eq. (2) as before. If the representation $D^{(0)}$, considered as representation of the subgroup, does not contain any irreducible representation more than once, the matrices which commute with all $D^{(0)}(s)$ also commute with all matrices which commute with all $D^{(0)}(s)$. If this is true for all irreducible representations, $D^{(j)}$, the vectors b are the same as the vectors a and conversely, if every vector a satisfies also the second condition for the b, no irreducible representation of the full group, if considered as representation of the subgroup, contains any representation of the subgroup more than once. Hence, the identity of the linear manifolds a and b is the necessary and sufficient condition for the circumstance given in the title of this note.

The sum $\sum b_T\, D(T)$ will commute with every $\sum a_R\, D(R)$ if it commutes with every member of the base (6). Hence

$$\sum_T b_T D(T) \sum_R \sum_s \delta(sQs^{-1}R^{-1})\, D(R) = \sum_R \sum_s \delta(sQs^{-1}R^{-1})\, D(R) \sum_T b_T D(T)$$

is the second condition on the b. Because of (6a), it can be written also

$$\sum_T \sum_s b_T D(TsQs^{-1}) = \sum_T \sum_s b_T\, D(sQs^{-1}T). \tag{8}$$

This second condition will be a consequence of the first, if (8) is satisfied for every vector (6) of the full base of vectors a. Hence we set

$$b_T = \sum_r \delta(rPr^{-1}T^{-1}). \tag{9}$$

Introducing this into (8), one obtains

$$\sum_r \sum_s D(rPr^{-1}\, sQs^{-1}) = \sum_r \sum_s D(sQs^{-1}\, rPr^{-1}). \tag{10}$$

The left side contains the matrices which correspond to the elements of the complex $\mathfrak{P}\mathfrak{Q}$ where $\mathfrak{P}$ is the subclass which contains P and $\mathfrak{Q}$ is the subclass which contains Q, and it contains each equally often. The right side has the same relation with respect to the complex $\mathfrak{Q}\mathfrak{P}$. Since, because of the completeness of the D, eq. (10) can be valid only if the same group elements are present equally often on left and right side, we obtain

$$\mathfrak{P}\mathfrak{Q} = \mathfrak{Q}\mathfrak{P}. \tag{11}$$

The necessary and sufficient condition that no irreducible representation, considered as representation of the subgroup, shall contain any representation of the subgroup more than once, is that the subclasses (with respect to the subgroup) commute.

REMARKS

If the subgroup is the whole group, (11) is the well-known relation of the commutative nature of the multiplication of classes. If, on the other hand, P and Q commute with all elements of the subgroup, $\mathfrak{P}=P$ and $\mathfrak{Q}=Q$ and (11) becomes the condition that the centralizer of the subgroup must be abelian. This is, therefore, a necessary condition for the validity of (11). It can be proved that this necessary condition is satisfied for all maximal subgroups, i.e., each proper subgroup which is not a proper subgroup of a proper subgroup. For, let P and Q be two elements of the centralizer of the subgroup which do not commute, $PQ \neq QP$. This can be the case only if neither is an element of the subgroup since, by definition, both P and Q commute with all elements of the subgroup. If, however, P and Q are not elements of the subgroup – which will be denoted now by S – one can form a new subgroup consisting of S, SP, $SP^2, \ldots$. This does not contain Q because all its elements commute with P whereas Q does not. Hence, the subgroup S was not maximal. Conversely, if S is maximal, (11) is satisfied for all the subclasses consisting of an element of the centralizer of the subgroup. It may not be satisfied for other subclasses and, indeed, Mackey has constructed a group[1]) which has no subgroup satisfying the condition contained in our title.

Let us observe that, on the other hand, *if all subclasses are ambivalent* (i.e., if all subclasses contain the reciprocal of every element which they contain) *then the subclasses also commute.* The subgroup will then have the property given in the title of this article.

To prove our assertion, let us observe, first, that if the subclasses are ambivalent, one can map each subclass, in a one-to-one fashion, on itself by mapping every element on its reciprocal. Furthermore, if the element R_1 is in the class C_1, and R_2 in the class of C_2, the products of R_1 and R_2 and of their maps R_1^{-1} and R_2^{-1} taken in the opposite order, will be in the same subclass: $R_2^{-1}R_1^{-1}=(R_1R_2)^{-1}$ is in the subclass of R_1R_2. Hence, the complexes C_1C_2 and C_2C_1 will have equally many elements in every subclass. Since both complexes consist of complete subclasses, i.e., contain every element of each subclass equally many times, the two complexes are identical and our assertion is demonstrated.

The group of all permutations of the first $n-1$ elements is a subgroup S_{n-1} of the group S_n of the permutations of all n elements and the subclasses of S_{n-1} are ambivalent. One can see this most easily by decomposing every permutation P into a product of disjoint cycles

$$P = (\alpha_1 \dots \alpha_\rho)(\alpha_{\rho+1} \dots \alpha_\sigma) \dots (\alpha_{\tau+1} \dots \alpha_{n-1} n). \tag{12}$$

The cycle containing n was written as last factor and n was chosen as its last member. The element p of the subgroup S_{n-1} which transforms P into its reciprocal consists of the product of transpositions

$$p = (\alpha_1 \alpha_\rho)(\alpha_2 \alpha_{\rho-1}) \dots (\alpha_{\rho+1} \alpha_\sigma)(\alpha_{\rho+2} \alpha_{\sigma-1}) \dots (\alpha_{\tau+1} \alpha_{n-1})(_{\tau+2} \alpha_{n-2}) \dots . \tag{12a}$$

If the two elements of a transposition in p should be identical (e.g., $\alpha_2 = \alpha_{\rho-1}$), the corresponding transposition can be omitted from (12a). It may be noted that it is not necessary to move the n in the last factor of (12) in order to transform it into its reciprocal.

It follows that the subclasses of the subgroup S_{n-1} of permutations of $1, 2, \dots, n-1$ are ambivalent and hence that they commute. It then follows that, if an irreducible representation of the full group S_n is considered as a representation of S_{n-1}, it contains each irreducible representation of S_{n-1} once or not at all – a fact which also follows from the standard theory of the representations of the symmetric group S_n developed by Frobenius, Schur and Young[2]).

References

1) G. W. Mackey, personal communication. Dr. Mackey also informed me that he found a shorter and more abstract proof of the principal result of the present paper.

2) See, e.g , M. Hamermesh, *Group Theory* (Addison-Wesley, Reading, Mass., 1962) eq. (7-52), p. 210.

Orthogonalization Methods

H. C. Schweinler and E. P. Wigner

Journal of Mathematical Physics *11*, 1693–1694 (1970)

(Received 20 October 1969)

An essentially unique method of forming an orthonormalized set of vectors from an independent set of n vectors is presented. The resulting set is proven to maximize a certain quartic form in the original vectors.

The best-known orthogonalization procedure, and the one most commonly used by mathematicians, is Schmidt's procedure.[1] It has the great advantage that it can be used, without further modifications, for a countably infinite set of vectors. On the other hand, the orthogonal set of vectors which it furnishes has no simple relation to the original set of vectors—the members of the set depend, in fact, on the order in which the vectors are arranged.

Physicists often use another orthogonalization method.[2] Let $\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_n$ be the vectors to be orthogonalized; they will be assumed to span a finite-dimensional complex Hilbert space. One then forms the Hermitian matrix M (Gram's matrix[1]), where

$$M_{kl} = (\mathbf{v}_k, \mathbf{v}_l). \tag{1}$$

If the vectors $\mathbf{v}$ are linearly independent, which is being assumed, M is positive definite. One forms its characteristic vectors $\mathbf{u}_\lambda$, with components $u_{k\lambda}$,

$$\sum_l M_{kl} u_{l\lambda} = p_\lambda u_{k\lambda}, \tag{2}$$

and the p are then positive. The $\mathbf{u}_\lambda$ will be assumed to be normalized so that

$$\sum_k u^*_{k\kappa} u_{k\lambda} = \delta_{\kappa\lambda}, \tag{3}$$

$$\sum_\kappa u^*_{k\kappa} u_{l\kappa} = \delta_{kl}. \tag{3'}$$

The orthogonal vectors $\mathbf{w}$ in the space of the $\mathbf{v}$ are then defined by means of the

$$\mathbf{w}^1_\kappa = \sum_k u_{k\kappa} \mathbf{v}_k. \tag{4}$$

One has

$$\begin{aligned}(\mathbf{w}^1_\kappa, \mathbf{w}^1_\lambda) &= \sum_k \sum_l u^*_{k\kappa} u_{l\lambda} (\mathbf{v}_k, \mathbf{v}_l) = \sum_{kl} u^*_{k\kappa} M_{kl} u_{l\lambda} \\ &= \sum_k u^*_{k\kappa} p_\lambda u_{k\lambda} = \delta_{\kappa\lambda} \cdot p_\lambda.\end{aligned} \tag{5}$$

Hence, if one sets

$$\mathbf{w}_\kappa = p_\kappa^{-\frac{1}{2}} \mathbf{w}^1_\kappa = \sum_k p_\kappa^{-\frac{1}{2}} u_{k\kappa} \mathbf{v}_k, \tag{6}$$

the $\mathbf{w}$ form an orthonormal set. Since the p_κ are positive, $p_\kappa^{-\frac{1}{2}}$ can also be assumed to be positive. Conversely,

$$\mathbf{v}_k = \sum_\kappa u^*_{k\kappa} p_\kappa^{\frac{1}{2}} \mathbf{w}_k. \tag{6'}$$

As matrix equations, (6) and (6′) are $\mathbf{w} = \mathbf{v} u p^{-\frac{1}{2}}$ and $\mathbf{v} = \mathbf{w} p^{\frac{1}{2}} u^\dagger$, respectively, p being a diagonal matrix with diagonal elements p_κ.

The set of $\mathbf{w}$ is, evidently, independent of the order in which the $\mathbf{v}$ are arranged, and one expects it to satisfy some extremal condition which distinguishes it among all the sets of orthonormal vectors

$$\mathbf{Z}_\kappa = \sum_\lambda S_{\lambda\kappa} \mathbf{w}_\lambda, \tag{7}$$

where S is an arbitrary unitary matrix. Of course, if S is a phase matrix $\exp(i\delta) = \operatorname{diag}(\exp(i\delta_1), \exp(i\delta_2), \cdots, \exp(i\delta_n))$, or a permutation matrix P, or a product $\exp(i\delta)P$ of such, the resulting set $\mathbf{Z}$ will be identical with the set of $\mathbf{w}$, except possibly for phase. One such extremal theorem is the subject of this paper. It states that the quantity

$$m = \sum_\kappa \left(\sum_k |(\mathbf{Z}_\kappa, \mathbf{v}_k)|^2 \right)^2 \tag{8}$$

assumes its maximal value of $\sum p_\kappa^2$ for the $\mathbf{w}$ set of orthogonal vectors, i.e., if the S in (7) is a unit matrix or a generalized permutation matrix. In one sense, (8) is the simplest expression involving the original set of vectors $\mathbf{v}$ and the orthonormal set $\mathbf{Z}$: If one does not square the sum over k in (8), the result is simply the sum of the squares of the lengths of the $\mathbf{v}$ and, hence, independent of the S in (7). Perhaps an even more natural extremal principle would be to set

$$m' = \sum_{k\kappa} |(\mathbf{Z}_\kappa, \mathbf{v}_k)|^4, \tag{8'}$$

but this does not lead to a simple set of equations such as the preceding one.

In order to express m in terms of S and $\mathbf{w}$, we first

calculate

$$(\mathbf{w}_\lambda, \mathbf{v}_k) = \left(\mathbf{w}_\lambda, \sum_\kappa p_\kappa^{\frac{1}{2}} u_{k\kappa}^* \mathbf{w}_\kappa\right) = p_\lambda^{\frac{1}{2}} u_{k\lambda}^*, \tag{9a}$$

and hence

$$(\mathbf{Z}_\kappa, \mathbf{v}_k) = \sum_\lambda S_{\lambda\kappa}^* p_\lambda^{\frac{1}{2}} u_{k\lambda}^*, \tag{9b}$$

and, because of (3),

$$\begin{aligned}\sum_k |(\mathbf{Z}_\kappa, \mathbf{v}_k)|^2 &= \sum_k \sum_{\lambda\mu} S_{\lambda\kappa} p_\lambda^{\frac{1}{2}} u_{k\lambda} S_{\mu\kappa}^* p_\mu^{\frac{1}{2}} u_{k\mu}^* \\ &= \sum_\lambda |S_{\lambda\kappa}|^2 p_\lambda, \end{aligned}\tag{9c}$$

so that

$$m = \sum_\kappa \left(\sum_\lambda |S_{\lambda\kappa}|^2 p_\lambda\right)^2. \tag{9d}$$

The p_λ are given by the $\mathbf{v}$, and we want to find the maximum of (9d) by varying the unitary matrix S. It is clear that the maximum exists, since (9d) is bounded and the domain of S is compact. We will assume for the present, further, that the p_λ are all different—the maximum of m is a continuous function of the p_λ, and its value for a set in which some of the p_λ are equal can be obtained as a limit of the maxima of m for a set of p all the p_λ of which are different.

The maximum is found, and the above statements about the generalized permutation matrix are verified, by observing that we can write m in the form

$$\begin{aligned} m &= \sum_\kappa \{(S^\dagger pS)_{\kappa\kappa}\}^2 \\ &= \sum_\kappa \sum_\lambda (S^\dagger pS)_{\kappa\lambda}(S^\dagger pS)_{\lambda\kappa} - \sum_{\kappa\neq\lambda}\sum{}' (S^\dagger pS)_{\kappa\lambda}(S^\dagger pS)_{\lambda\kappa} \\ &= \sum_\lambda p_\lambda^2 - \sum_{\kappa\neq\lambda}\sum{}' |(S^\dagger pS)_{\lambda\kappa}|^2. \end{aligned}\tag{10}$$

m is bounded above by $\sum p_\lambda^2$, and it can attain this bound if $S^\dagger pS$ is a diagonal matrix d. In this case, $pS = Sd$, or

$$p_\lambda S_{\lambda\kappa} = S_{\lambda\kappa} d_\kappa, \tag{11}$$

so the set of p coincides with the set of d, and $S_{\lambda\kappa}$ vanishes unless $p_\lambda = d_\kappa$. Since we have assumed that the p_λ are all different, only one $S_{\lambda\kappa}$ is nonzero. The unitarity of S requires that the absolute value of this nonzero term be unity, and its phase is free. Thus S is a generalized permutation matrix.

The preceding discussion assumed that all the characteristic roots p of M are different. One can conclude, if some of the p are equal, by continuity arguments that no m, for any choice of the orthogonal $\mathbf{Z}$ in (8), can be larger than the m obtained for $\mathbf{Z}_\kappa = \mathbf{w}_\kappa$. However, if some p_κ are equal, the corresponding $\mathbf{w}$ can be subjected to a unitary transformation without affecting the value of m. In such a case, the S which yields the maximum value of m can contain such a unitary transformation. This was foreseeable: If some p are equal, the corresponding $\mathbf{w}$ are determined only up to a unitary transformation between them.

* Research done in part at Oak Ridge National Laboratory and sponsored by the U.S. Atomic Energy Commission under contract with Union Carbide Corporation.

[1] R. Courant and D. Hilbert, *Methoden der mathematischen Physik* (Springer, Berlin, 1931), 2nd ed., Vol. I, p. 4; see also pp. 29–32.

[2] R. Landshoff, Z. Physik **102**, 201 (1936). Actually, Landshoff's orthonormalized functions are, in our notation,

$$\sum_l (M^{-\frac{1}{2}})_{lk} \mathbf{v}_l = \sum_\kappa u_{k\kappa}^* \mathbf{w}_\kappa.$$

Restriction of Irreducible Representations of Groups to a Subgroup

E. P. Wigner

Proceedings of the Royal Society of London A *322*, 181–189 (1971)

(*Received* 12 *November* 1970)

The number and character of the irreducible representations of a subgroup, contained in an irreducible representation of the whole group (if this representation is restricted to the subgroup) play an important role in quantum mechanics. They give the number and type of the states, generated by a symmetry breaking perturbation, from a state which has the symmetry of the whole group. Three equations are derived here for the number and character of the representations of the subgroup, resulting from the restriction of the irreducible representations of the whole group. These equations contain an earlier rule as a special case.

1. Introduction and notation

Some time ago, the present writer (Wigner 1968) derived a condition which a subgroup must satisfy if no irreducible representations of this subgroup is to be contained more than once in any irreducible representation of the full group (when this representation is restricted to the subgroup). This condition has some relevance for certain symmetry problems of physics. It appeared, however, that the condition derived must be a special case of more general relations and the purpose of the present article is to obtain those more general relations. These are given by equations (II) and (III).

We shall denote the elements of the full group by capital letters, $P, Q, R, S, \ldots$, those of the subgroup by small letters, $p, q, \ldots$. The order of the full group will be G, that of the subgroup g. The irreducible representations of the full group will be denoted by D with suitable indices, those of the subgroup by d. The indices $J, K, \ldots$, will serve to distinguish the different irreducible representations of the full group, $j, k, \ldots$, those of the subgroup. Hence, for instance, $D^J(R)$ is the matrix which corresponds to the group element R in the irreducible representation J of the full group. The symbols $J, K, \ldots$, as well as the symbols $j, k, \ldots$, if not used as indices but as factors in formulae, denote the dimension (number of rows and columns) of the corresponding representation. We shall be concerned only with representations by means of non-singular matrices.

Finally, we shall denote by (J,j) the number of times the irreducible representation D^J of the full group contains the irreducible representation d^j of the subgroup, if the former is restricted to the subgroup. This last clause will often be omitted henceforth and we shall simply speak about the number of times d^j is contained in D^J. The aforementioned article gives a criterion for all (J,j) to be 0 or 1, i.e. for $(J,j)^2 = (J,j)$ for all J and all j.

We shall begin with a very obvious relation between the quantities (J,j). This is done partly in order to assure a measure of completeness of the equations to be

derived, partly because the relation in question is often disregarded when surmises are put forward.

The present article will be concerned solely with finite groups.

2. Restriction of the regular representation to the subgroup

The elements of the regular representation D of the full group are given by

$$D(P)_{Q,R} = \delta_{Q,PR}, \tag{1}$$

where $\delta_{A,B} = 0$, unless the elements A and B are identical. The rows and columns of the regular representation are labelled in (1) by the group elements themselves.

If restricted to the subgroup, this representation contains the regular representation thereof G/g times. In order to see this, let us decompose the full group into right cosets. We give the elements of the full group, therefore, the form qT_i where q is an element of the subgroup and T_i is the same element for the whole coset i, but different for different cosets. There are, naturally, G/g cosets. For an element p of the subgroup, the elements of the regular representation assume then the form

$$D(p)_{qT_i,rT_j} = \delta_{qT_i,prT_j}. \tag{2}$$

This is zero unless qT_i and prT_j are in the same coset, i.e. unless $i = j$ and unless $q = pr$. Hence, we also have

$$D(p)_{qT_i,rT_j} = \delta_{ij}\delta_{q,pr}, \tag{2a}$$

which shows that indeed the regular representation of the full group, if restricted to the subgroup, decomposes into G/g regular representations thereof.

Since the regular representation of the full group contains the irreducible representation D^J just J times, and since the subgroup's regular representation contains d^j just j times, we have

$$\sum_j J(J,j) = (G/g)j. \tag{I}$$

As was mentioned before, this relation should be well known. It can be derived in several other ways.

3. Properties of the subclasses

The subclasses were defined in the earlier publication (Wigner 1968) as the sets of elements of the full group which can be transformed into each other by elements of the subgroup. The subclasses have a number of rather simple properties which will be given below. The first seven of these are evident and only the last one (h) will be explicitly demonstrated.

(a) If two subclasses have one element in common, they are identical.

(b) Every subclass is contained in one of the classes of the full group. In other words, all classes consist of complete subclasses.

(c) A subclass which contains an element of the subgroup is contained therein entirely. It forms a class of the subgroup.

(*d*) The subclass, as a set, commutes with every element of the subgroup, i.e. if s is a subclass, the sets ps and sp are identical.

(*e*) The elements of the subgroup which commute with an element of the full group (the subgroup's normalizer of that element) form a group. For different elements of a subclass, these normalizers are conjugate and have the same order.

(*f*) The number of the elements in each subclass is a divisor of the order of the subgroup. It is the index of the normalizer groups of (*e*) within the subgroup.

(*g*) The reciprocals of the elements of a subclass also form a subclass.

(*h*) The product of two subclasses consists of entire subclasses.

Let us denote the two subclasses by s and s'; they should contain n and n' elements, respectively, so that the set ss' contains nn' elements. We have to demonstrate that any two elements of the same subclass, P and pPp^{-1} are contained equally often in ss'.

If P were contained in ss' more often than pPp^{-1}, the sets ss and $pss'p^{-1}$ would not be identical: the second one would contain pPp^{-1} more often than the first. However, according to (*d*), the set psp^{-1} is identical with the set s, and $ps'p^{-1}$ is identical with s'. Hence, the product $psp^{-1}ps'p^{-1} = pss'p^{-1}$ is identical with ss'.

The statement (*h*) implies that the subclasses form an algebra. If the subgroup consists solely of elements of the centre, this algebra is the usual group algebra, commutative only if the full group is commutative. If the subgroup is the full group, the subclasses become the usual classes of the group and their algebra is commutative.

4. Recapitulation of the Frobenius–Schur theory (see Schur 1905, 1908)

Consider an arbitrary representation D in reduced form

$$D(P)_{J\mu\alpha;K\nu\beta} = \delta_{JK}\delta_{\mu\nu}D^J(P)_{\alpha\beta}. \tag{3}$$

The indices μ and ν take care of the fact that D^J can be contained in D several times. Consider now the set of matrices A which commute with all $D(P)$

$$\sum_{K\nu\beta} D(P)_{J\mu\alpha;K\nu\beta}A_{K\nu\beta;L\lambda\gamma} = \sum_{K\nu\beta} A_{J\mu\alpha;K\nu\beta}D(P)_{K\nu\beta;L\lambda\gamma}. \tag{4}$$

Substitution of (3) into (4) gives

$$\sum_{\beta} D^J(P)_{\alpha\beta}A_{J\mu\beta;L\lambda\gamma} = \sum_{\beta} A_{J\mu\alpha;L\lambda\beta}D^L(P)_{\beta\gamma}. \tag{5}$$

It then follows from Schur's lemma that

$$A_{J\mu\alpha;K\nu\beta} = \delta_{JK}\delta_{\alpha\beta}a^J_{\mu\nu}, \tag{6}$$

the $a^J_{\mu\nu}$ being arbitrary.

We shall call the number of linearly independent matrices of a linear set of matrices the dimension of the set. Hence, the dimension of the set of the matrices A commuting with D is

$$\sum_J (D,J)^2 = \text{dimension of } A, \tag{7}$$

where (D, J) denotes the number of times the irreducible representation D^J is contained in D.

Let us consider, next, the set of all matrices B which commute with all matrices A

$$\sum_{K\nu\beta} A_{J\mu\alpha;K\nu\beta} B_{K\nu\beta;L\lambda\gamma} = \sum_{K\nu\beta} B_{J\mu\alpha;K\nu\beta} A_{K\nu\beta;L\lambda\gamma}. \tag{8}$$

Substitution of (6) into (8) gives

$$\sum_{\nu} a^J_{\mu\nu} B_{J\nu\alpha;L\lambda\gamma} = \sum_{\nu} B_{J\mu\alpha;L\nu\gamma} a^L_{\nu\lambda}, \tag{9}$$

and since this must hold for all $a^J_{\mu\nu}$

$$B_{J\mu\alpha;K\nu\beta} = \delta_{JK}\delta_{\nu\mu} b^J_{\alpha\beta}. \tag{10}$$

This was to be expected, since the $D(P)$ all commute with all the A and—as (10) shows—essentially only the $D(P)$ and their linear combinations commute with all A.

The dimension of the set of the B is

$$\sum_J J^2 = \text{dimension of } B, \tag{11}$$

where the summation is to be extended over those representations D^J which are contained in D. Note that the b^J in (10) is independent of ν and μ.

The basic observation of the Frobenius–Schur theory (see Schur 1905, 1908) is that the dimension of sets of matrices, such as that of the A or B, which commute with the matrices of a given set, remains the same if the matrices of that set, such as the $D(P)$, are replaced by an equivalent set $XD(P)X^{-1}$ with a P independent but otherwise arbitrary X. Let us replace, therefore, the $D(P)$ of (3) by the matrices (1) of the regular representation.

The set of A' which commute with the $D(P)$ of (1) is determined by

$$\sum_R \delta_{Q,PR} A'_{R,S} = \sum_R A'_{Q,R} \delta_{R,PS}, \tag{12}$$

or

$$A'_{P^{-1}Q,S} = A'_{Q,PS}. \tag{13}$$

Setting $P = Q$ in this, replacing $A'_{E,S}$ by a'_S (E being the unit element), and substituting QS by R, we obtain

$$A'_{Q,R} = a'_{Q^{-1}R}. \tag{14}$$

Conversely, if (14) is satisfied with any a'_P, so is (13). It follows that, if D is the regular representation, the dimension of the set A is equal to the order of the group. This gives, together with (7), a well-known relation which did not have to be derived here. However, we will need the form of the A', given by (14).

Let us determine, next, the corresponding form B' of the B. We must satisfy the equations

$$\sum_R A'_{Q,R} B'_{R,S} = \sum_R B'_{Q,R} A'_{R,S}. \tag{14a}$$

Using (14) and the fact that (14a) must be valid for any choice of the G numbers $a'_{Q^{-1}R}$, one finds

$$B'_{Q,R} = b'_{QR^{-1}} \tag{15}$$

with the G numbers b'_P being again arbitrary, so that the dimension of the sets of the B' (and hence also that of the B) is also G. Comparison of (15) with (11) again leads to a well-known relation which it was not necessary to re-derive. However, again, we will need the form (15) of the B'.

5. Sets of matrices which commute also with the restriction of a representation to the subgroup

We shall consider, next, subsets of the matrices B which commute not only with the matrices A defined by (4) but also with the matrices $D(p)$ obtained by restricting the representation $D(P)$ to the subgroup. We shall denote these matrices by C. The form of the representation $D(P)$, which was assumed in (3) to be the reduced form, will be further specified: it will be assumed that each $D^J(P)$ is reduced out as a representation of the subgroup. Hence, the indices $\alpha, \beta, \ldots$, will be replaced by triple indices j, ϵ, α. The first of these refers to the representation d^j of the subgroup which appears in the restriction of D^J to the subgroup, ϵ takes care of the fact that the restriction of D^J may contain d^j several times; it can assume, therefore, (J,j) values. The indices $\alpha, \beta, \gamma, \ldots$, will henceforth refer to rows or columns of d^j. Hence, the elements of $D(p)$ of (3) will now be written in the form

$$D(p)_{J\mu j\epsilon\alpha;K\nu k\zeta\beta} = \delta_{JK}\delta_{\mu\nu}\delta_{jk}\delta_{\epsilon\zeta}d^j(p)_{\alpha\beta}. \tag{16}$$

For the subset of the matrices B which commute with the $D(p)$, we shall use the symbol c^J instead of b^J, so that (10) assumes for these the form

$$C_{J\mu j\epsilon\alpha;K\nu k\zeta\beta} = \delta_{JK}\delta_{\mu\nu}c^J_{j\epsilon\alpha;k\zeta\beta}. \tag{17}$$

The added condition for these is that

$$\sum_{K\nu}\sum_{k\zeta\beta} D(p)_{J\mu j\epsilon\alpha;K\nu k\zeta\beta}C_{K\nu k\zeta\beta;L\lambda l\eta\gamma} = \sum_{K\nu}\sum_{k\zeta\beta} C_{J\mu j\epsilon\alpha;K\nu k\zeta\beta}D(p)_{K\nu k\zeta\beta;L\lambda l\eta\gamma}. \tag{18}$$

Insertion of (16) and (17) into (18) gives again by Schur's theorem

$$c^J_{j\epsilon\alpha;k\zeta\beta} = \delta_{jk}c^{Jj}_{\epsilon\zeta}\delta_{\alpha\beta}. \tag{19}$$

The c^{Jj} can be arbitrary matrices. Since they have (J,j) rows and columns, the dimension of the set of the C is

$$\sum_{Jj}(J,j)^2 = \text{dimension of } C. \tag{20}$$

The summation over J is to be extended over all the irreducible representations which occur in D (all occur in the regular representation), each to be counted once. We recall that the $b^J_{\alpha\beta}$ of (10) was independent of μ.

Let us apply (20) to the regular representation. The dimension of the set C will then be obtained by finding the matrices B', of the form (15), which commute with the $D(p)$, of the form (1). Such a matrix will be denoted by C'.

The condition that C' commutes with the $D(p)$ of (1) now reads

$$\sum_R \delta_{Q,pR}C'_{R,S} = \sum_R C'_{Q,R}\delta_{R,pS}. \tag{21}$$

We replace $C'_{Q,R}$ by $c'_{QR^{-1}}$; as (15) shows, this will ensure that C' commutes with all A'. Hence, (21) becomes

$$c'_{p^{-1}QS^{-1}} = c'_{QS^{-1}p^{-1}}. \tag{22}$$

The necessary and sufficient condition for (22) is that c' be the same for all elements of every subclass, i.e. that it be a subclass function. The dimension of the set of the matrices C' is, therefore, the number of subclasses and (20) thus gives

$$\sum_{J,j}(J,j)^2 = \text{number of subclasses.} \tag{II}$$

If the subgroup consists of the unit element alone, a representation D^J, if restricted to the unit element, contains J identical representations of this element. The number of subclasses is equal to the number of elements and the well known formula $\Sigma J^2 = G$ results. If the subgroup is the whole group, the D^J does not decompose at all and the sum over j alone of $\Sigma(J,j)^2 = 1$. Hence, the sum on the left side of (II) becomes equal to the number of non-equivalent irreducible representations. The subclasses are classes in this case and again (II) reduces to a well-known result. The case in which we are interested is, however, neither of these extreme cases.

6. The centre of the set of the C

The centre of the set of the matrices C consists of those matrices Z which commute with all matrices A, with all $D(q)$, and also with all the matrices C. Since they form a subset of the matrices C, because of (17) and (19), they must have the form

$$Z_{J\mu j\epsilon\alpha;K\nu k\zeta\beta} = \delta_{JK}\delta_{\mu\nu}\delta_{jk}z^{Jj}_{\epsilon\zeta}\delta_{\alpha\beta}. \tag{23}$$

This will commute with the C if z^{Jj} commutes with all the c^{Jj}, and if this is true for all J and j. This will be true only if all the matrices z^{Jj} are multiples of the unit matrix:

$$z^{Jj}_{\epsilon\zeta} = z_{Jj}\delta_{\epsilon\zeta}. \tag{24}$$

The z_{Jj} are arbitrary, so that every Jj combination in D contributes one dimension to the set of Z:

$$\sum_{Jj}(J,j)^\epsilon = \text{dimension of } Z, \tag{25}$$

$(J,j)^\epsilon$ being 1 if D^J is contained in D and d^j contained in D^J, zero otherwise.

Again, we consider the case that D is the regular representation; the centre of the set of the matrices C' will consist of the matrices Z'. Since the Z' are members of the set B', we can write by (15)

$$Z'_{Q,R} = z'_{QR^{-1}}, \tag{26}$$

and since they are also members of the subset C', (22) also holds for them and the z' are equal for all elements of a subclass

$$z'_P = z'_{pPp^{-1}}. \tag{27}$$

The added condition on the Z' is that they commute with any matrix C', i.e.

$$\sum_R c'_{QR^{-1}} z'_{RS^{-1}} = \sum_R z'_{QR^{-1}} c'_{RS^{-1}} \tag{28}$$

must be valid for any c'_P which is a subclass function. Writing P^{-1} for QR^{-1} on the left side of (28) and for RS^{-1} on the right side, this assumes the form

$$\sum_P c'_{P^{-1}} z'_{PQS^{-1}} = \sum_P c'_{P^{-1}} z'_{QS^{-1}P}. \tag{29}$$

These equations will be satisfied if they are satisfied for all c' which assume the value 1 in one subclass, 0 in all others and if this is true for all subclasses. Since (see (g) of §3) the P^{-1} form a subclass if the P do, (29) can be written as

$$\sum_{P \subset s} z'_{PR} = \sum_{P \subset s} z'_{RP}, \tag{30}$$

QS^{-1} has been replaced by R, the summation over P is to be extended over one subclass but (30) must be valid no matter over which subclass it is extended; z' is itself a subclass function.

Because of this circumstance, (30) is equivalent to

$$\sum_p \sum_{P \subset s} z'_{pPRp^{-1}} = \sum_p \sum_{P \subset s} z'_{pRPp^{-1}}, \tag{31}$$

the summation over p to be extended over the subgroup. Next, the index $pPRp^{-1}$ on the left of (31) can be written as $pPp^{-1}pRp^{-1}$ and pPp^{-1} can be replaced by Q. The summation over Q is then to be extended over $p^{-1}sp$, but this is identical with s. This gives for the left side of (31)

$$\sum_p \sum_{Q \subset s} z'_{QpRp^{-1}}. \tag{31a}$$

A similar transformation can be carried out on the right side of (31). Finally, the summation over p lets pPp^{-1} run over the subclass of R—as many times as the normalizer of R in the subgroup indicates. This is, however, the same number on the left and right sides so that (31) finally reduces to

$$\sum_{Q \subset s} \sum_{R \subset s'} z'_{QR} = \sum_{Q \subset s} \sum_{R \subset s'} z'_{RQ}. \tag{32}$$

In order to analyse further the implication of (32) for the z', we recall the observation (h) of §3, that the product of two subclasses consists of entire subclasses. We denote, by $(s, s'; s'')$, the number of times the subclass s'' is contained in the product ss'. Hence, (32) can be written in the form

$$\sum_{s''} (s, s'; s'') z'_{s''} = \sum_{s''} (s', s; s'') z'_{s''}, \tag{33}$$

where

$$z'_{s''} = \sum_{P \subset s''} z'_P \tag{33a}$$

is z'_P multiplied with the number of elements in its subclass. Thus (33) is the condition the z' must satisfy for each pair of subclasses s and s', if the corresponding

Z' is to be in the centre of the set of the C'. The question which has yet to be answered concerns the number of equations (33) that are independent.

If the subclasses s and s' commute, (33) becomes trivial. If they do not, it gives a relation between the z'. However, they also furnish then an element

$$\sum_{s''}[(s,s';s'')-(s',s;s'')]\,s'' = ss'-s's \tag{34}$$

for the set of commutators of the algebra of the subclasses. If an equation of the form (33) can be obtained as a linear combination of similar equations, with other s, s', the corresponding commutator (34) also can be obtained as a linear combination of other commutators. The case that the only solution of the equations (33) is $z'_s = 0$ need not be considered since at least the z' of the unit element can surely be arbitrary. Hence, as one looks successively at the equations (33), at each step one either finds that a new condition for the z' has been added—in which case one also finds a commutator of two subclasses which is linearly independent from the commutators found before—or one finds that the equation in question is a linear combination of equations obtained before, in which case the commutator of the two subclasses is also a linear combination of commutators found before. It follows that the number of equations (33) which are independent is equal to the dimension of the set of commutators of the algebra of the subclasses. Hence, the dimension of the centre Z' of the matrices C' is the difference between the number of subclasses and the number of linearly independent commutators of the algebra of subclasses

Comparison with (25) therefore gives

$$\sum_{J,j}(J,j)^e = \text{number of subclasses, minus the number of linearly independent commutators of these.} \tag{III}$$

Combining this with (II), we find

$$\sum_{J,j}(J,j)^2-(J,j)^e = \text{number of linearly independent commutators of the algebra of subclasses.} \tag{IIIa}$$

The summations over J in (III) and (IIIa) have to be extended over all irreducible representations D^J, each to be taken once. We recall that $(J,j)^e = 1$, if d^j is contained in the restriction of D^J to the subgroup; it is 0 otherwise.

If all (J,j) are either 0 or 1, the left side of (IIIa) vanishes—and it vanishes only in this case—so that the algebra of subclasses must be commutative. This is the theorem (Wigner 1968) which is generalized by (II) and (III).

If the subgroup is the full group, the (J,j) are all 1 or 0 and (IIIa) only gives the well known result that the classes commute. However, if the subgroup consists only of the unit element, one obtains the

Corollary. The order of the group, minus the number of classes (the left side of (IIIa)) is equal to the number of linearly independent commutators of the group.

This can be seen also directly: the commutator algebra contains all the differences between any two elements of any class. The elements of the algebra which are not commutators are the sums of the elements of the classes.

References

Wigner, E. P. 1968 Article in *Spectroscopic and group theoretical methods in physics* (Racah Memorial Volume). Amsterdam: North Holland Publishing Co. (The readers familiar with the literature will note the similarity of the basic idea of this article with that of the next reference.)

Schur, I. 1905, 1908 *Ber. Preuss. Akad. Wiss.* 1905, p. 406 and, 1908, p. 664. (In one form or another, the methods and ideas of these articles and those of G. Frobenius, also published by the Preuss. Akademie, remain integral parts of most treatments of representation theory, even the modern ones. However, see also W. Burnside 1911 *Theory of groups* (2nd ed.) Cambridge University Press.

Condition That All Irreducible Representations of a Compact Lie Group, If Restricted to a Subgroup, Contain No Representation More Than Once

F. E. Goldrich and E. P. Wigner

Canadian Journal of Mathematics 24, 432–438 (1972)

1. One of the results of the theory of the irreducible representations of the unitary group in n dimensions U_n is that these representations, if restricted to the subgroup U_{n-1} leaving a vector (let us say the unit vector e_1 along the first coordinate axis) invariant, do not contain any irreducible representation of this U_{n-1} more than once (see [**1**, Chapter X and Equation (10.21)]; the irreducible representations of the unitary group were first determined by I. Schur in his doctoral dissertation (Berlin, 1901)). Some time ago, a criterion for this situation was derived for finite groups [**3**] and the purpose of the present article is to prove the aforementioned result for compact Lie groups, and to apply it to the theory of the representations of U_n.

The physicist's interest in the question whether a group and a corresponding subgroup satisfy the criterion of the title of this article is motivated by the phenomenon of "symmetry breaking". Let us consider a quantum mechanical operator—usually the Hamiltonian—which is invariant under a group of operators. When the operators of the group are applied to the characteristic vectors of a characteristic value of the quantum mechanical operator, these vectors transform, except for the rare case of "accidental degeneracy", according to an irreducible representation of the invariance group. The physicist says that the characteristic vectors of each characteristic value "belong" to the different rows of the irreducible representation in question. If the original quantum mechanical operator is now modified by a small perturbation which is not invariant under the whole group but only under a subgroup, the original characteristic value splits up in as many characteristic values as are irreducible representations of the subgroup contained in the irreducible representation of the whole group to which this characteristic value belongs. If each irreducible representation of the subgroup is contained only once in the irreducible representation of the full group, only one new characteristic value belonging to any irreducible representation of the subgroup will arise from any characteristic value of the original quantum mechanical operator. The corresponding characteristic vectors belong to the irreducible representations of the subgroup and are completely determined by this criterion. This is not the case if several characteristic values of the new problem arise from a characteristic value of

Received April 19, 1971 and in revised form, November 17, 1971.

the original problem, i.e., if the irreducible representation of the original, full group contains any irreducible representation of the subgroup more than once. The knowledge of the characteristic vectors of the new problem, containing the small perturbation, reduces the calculation of the characteristic values of the new problem to the calculation of a matrix element or scalar product. The calculation of the splitting of the atomic energy levels by a magnetic field (Zeeman effect) is an example herefor; the original group is R_3 (or U_2) in this case, the subgroup is R_2 (or U_1). There are, however, also other questions which can be answered with relative ease if the characteristic vectors of the new problem are known.

In order to formulate the criterion for the situation given in the title of this article, one introduces the concept of subclasses. These consist of the elements of the group (in our case U_n) which can be transformed into each other by an element of the subgroup (in our case U_{n-1}). In the case of finite groups, it was shown that the criterion for no irreducible representation of the whole group to contain any representation of the subgroup more than once is that the subclasses commute. In the case of continuous groups, the subclasses are, as a rule, manifolds of lower dimensionality than the whole group. As a result, the commuting nature of the subclasses is less easily defined than in the case of finite groups. In the case of the latter, the statement that two subclasses C_1 and C_2 commute means that every element R of the group appears as often in the complex C_1C_2 as in the complex C_2C_1. Both these complexes are, naturally, finite if the whole group is finite. If the group is continuous, the sets C_1C_2 and C_2C_1 are, as a rule, continuous and infinite manifolds and the frequency of the occurrence of an element R in them is not defined. Hence, the criterion for finite groups must be reformulated and this is the principal subject of the present article, leading to Equation (11). The f in this equation is an arbirtary numerical function on the group. Naturally, this equation is also an expression for the identity of all complexes C_1C_2 and C_2C_1 in the case of finite groups, C_1 and C_2 being subclasses, if the invariant integrations over the subgroup—indicated by the $\int du$ and $\int ds$ symbols—are replaced by summations over the elements of the subgroup.

In the third part of the article, (11) will be shown to hold for all pairs of elements R and T of U_n if the integrations over u and s are the invariant integrations over the subgroup U_{n-1} which leaves a vector e_1 invariant. Thus, a proof for the result of the theory of the representations of U_n, mentioned in the first paragraph, will be produced.

Lastly, a sufficient but not necessary property of subgroups is given which assures that the irreducible representations of the whole group, if restricted to the subgroup in question, do not contain any representation of the latter more than once.

2. Let G be a compact Lie group, and let H be a subgroup of G. Consider now those functions $a(R)$ defined and continuous on the group space, for which

$\int a(R)D(R)dR$ commutes with $D(s)$ for all s in H, and for all irreducible representations D of G. The dR denotes the invariant integration over the group. Now suppose that the matrix $C^{(0)}$ commutes with all $D^{(0)}(s)$ for a single $D^{(0)}$, all s in H. Then the matrix $C^{(0)}$ can be written as $\int f(R)D^{(0)}(R)dR$, where

$$f(R) = \frac{l}{v}\sum_{i,k} C_{ik}{}^{(0)}D_{ki}{}^{(0)}(R^{-1}), \tag{1}$$

where l is the dimension of $D^{(0)}$ and $v = \int dR$ is the group volume, since

$$\begin{aligned} &\frac{l}{v}\int\sum_{i,k} C_{ik}{}^{(0)}D_{ki}{}^{(0)}(R^{-1})D_{jl}{}^{(0)}(R)dR \\ &= \frac{l}{v}\sum_{i,k} C_{ik}{}^{(0)}\frac{v}{l}\delta_{kl}\delta_{ji} \\ &= C_{jl}{}^{(0)}. \end{aligned} \tag{2}$$

It is clear that $f(R)$ is continuous on the group, since it is a linear functional of the matrix elements of a group representation.

Consider now the matrix $C^{(m)}$, defined by

$$\begin{aligned} C_{ij}{}^{(m)} &= \int f(R)D_{ij}{}^{(m)}(R)dR \\ &= \int \frac{l}{v}\sum_{n,k} C_{nk}{}^{(0)}D_{kn}{}^{(0)}(R^{-1})D_{ij}{}^{(m)}(R)dR \\ &= \begin{cases} 0, & m \neq 0 \\ C_{ij}{}^{(0)}, & m = 0. \end{cases} \end{aligned} \tag{3}$$

The matrix $C^{(m)}$ is equal to either $C^{(0)}$ or zero, so $\int f(R)D_{ij}{}^{(m)}(R)dR$ certainly commutes with $D^{(m)}(s)$, and so $f(R)$ is contained in the set of $a(R)$ defined above. Hence, any matrix which commutes with the $D^{(j)}(s)$, for every s but any single j, is contained in the set $\int a(R)D(R)dR$.

The requirement that $\int a(R)D(R)dR$ commute with $D(s)$ for s in the subgroup H imposes certain conditions on the functions $a(R)$. In particular,

$$\int a(R)D(R)D(s)dR = \int a(R)D(s)D(R)dR; \tag{4}$$

multiplication of (4) by $D(s^{-1})$ on the right yields

$$\int a(R)D(R)dR = \int a(R)\, D(sRs^{-1})dR \tag{4a}$$

and replacing the variable R on the right by $s^{-1}Rs$ then gives

$$\int a(R)D(R)dR = \int a(s^{-1}Rs)D(R)dR. \tag{5}$$

Since this holds for all irreducible representations D, the completeness of the matrix elements of the irreducible representations requires that

$$a(R) = a(s^{-1}Rs). \tag{5a}$$

We will call the set of elements of the group which can be written in the form $s^{-1}Rs$ the *subclass* containing the element R. Evidently, the product of

two subclasses contains full subclasses. A subclass either is contained in the subgroup, in which case it consists of exactly one class of the subgroup, or contains no elements therein.

According to (5a), then, the functions in our set a are simply the continuous, subclass-valued functions on G. Any such function can be written in the form

$$a(R) = \int \tilde{a}(s^{-1}Rs)ds, \tag{6}$$

where $\tilde{a}$ is any continuous function on the group.

Consider now the $\int a(R)D(R)dR$. These matrices commute with all the $D(s)$. If the representation $D(R)$ is considered as a representation of the subgroup H, then all the matrices which commute with the $D(s)$ will commute with each other, if and only if $D(s)$ contains no irreducible representation of H more than once. Thus, the condition that $D(R)$ restricted to H contain no irreducible representation of H more than once is equivalent to the condition that the $\int a(R)D(R)dR$ commute with each other.

Writing $a(R)$ as in (6), we have, instead of

$$\left(\int a(R)D(R)dR\right)\left(\int a'(T)D(T)dT\right) = \left(\int a'(T)D(T)dT\right)\left(\int a(R)D(R)dR\right) \tag{7}$$

the equation

$$\begin{aligned} &\iiiint \tilde{a}(s^{-1}Rs)\tilde{a}'(u^{-1}Tu)D(RT)dsdudRdT \\ &= \iiiint \tilde{a}(s^{-1}Rs)\tilde{a}'(u^{-1}Tu)D(TR)dsdudRdT, \end{aligned} \tag{8}$$

or

$$\begin{aligned} &\iiiint \tilde{a}(R)\tilde{a}'(T)D(sRs^{-1}uTu^{-1})dsdudRdT \\ &= \iiiint \tilde{a}(R)\tilde{a}'(T)D(uTu^{-1}sRs^{-1})dsdudRdT. \end{aligned} \tag{9}$$

Since $\tilde{a}$ and $\tilde{a}'$ are arbitrary functions, this requires that for any D

$$\iint D(sRs^{-1}uTu^{-1})dsdu = \iint D(uTu^{-1}sRs^{-1})dsdu. \tag{10}$$

But since the matrix elements of the irreducible representations form a complete set of functions on group space, (10) says that any function integrated over the set $sRs^{-1}\ uTu^{-1}$ is equal to the same function integrated over the set $uTu^{-1}\ sRs^{-1}$:

THEOREM 1. *The necessary and sufficient condition for no irreducible representations of a compact Lie group to contain any representation of a subgroup H more than once is the validity of*

$$\iint f(sRs^{-1}uTu^{-1})dsdu = \iint f(uTu^{-1}sRs^{-1})dsdu \tag{11}$$

for all elements R, T of the full group and any continuous function f on this group.

The ds and du in (11) denote the invariant integration over the subgroup H.

Theorem 1 corresponds to the criterion for finite groups that the subclasses commute.

3. We want to show next that Equation (11) holds when the group G is U_n, and the subgroup H is U_{n-1}.

In order to do this, it will be demonstrated, first, that if a subclass contains a transformation R, it also contains its transpose R^T. In order to demonstrate this, we transform R to the principal axes

$$R = V^\dagger \omega V, \tag{12}$$

where ω is a diagonal matrix with diagonal elements of modulus 1, V is unitary. The cross denotes hermitean adjoint. The transpose of (12) reads $R^T = V^T \omega \bar{V}$, but we may write as well (the bar denotes complex conjugation)

$$R^T = V^T \eta \omega \bar{\eta} \bar{V}, \tag{13}$$

where η is an arbitrary diagonal matrix the diagonal elements of which are also of modulus 1 so that $\eta\bar{\eta} = 1$ and also $\eta\omega\bar{\eta} = \omega$. One can express ω by means of both (12) and (13); equating the two expressions gives

$$VRV^\dagger = \bar{\eta}\bar{V}R^T V^T \eta$$

so that

$$V^T \eta V R (V^T \eta V)^{-1} = R^T; \tag{14}$$

i.e., the unitary matrix $V^T \eta V$ transforms R into its transpose R^T and this is true for any η as defined above. It may be observed, parenthetically, that these matrices are symmetrical. This observation represents a special case of a more general theorem due to O. Taussky and H. Zassenhaus [**2**]. Equation (14) now leads to Lemma 1.

LEMMA 1. *Any subclass of U_n with respect to U_{n-1} which contains R also contains R^T.*

In order to prove this, it is necessary only to prove that η in (14) can be so chosen that $V^T \eta V$ leaves e_1 invariant. This will be the case if

$$(V^T \eta V)_{i1} = \delta_{i1} \qquad (i = 1, 2, \ldots, n), \tag{15}$$

or

$$\sum_l V_{li} \eta_l V_{l1} = \delta_{i1}, \tag{15a}$$

and this will be satisfied if one chooses

$$\eta_l = \bar{V}_{l1}/V_{l1} \tag{16}$$

if $V_{l1} \neq 0$ and arbitrarily, though of modulus 1, if $V_{l1} = 0$. With these η_l one has

$$(V^T \eta V)_{i1} = \sum_l V_{li} \bar{V}_{l1} = \delta_{i1}. \tag{17}$$

This then proves Lemma 1.

A further lemma concerns the group integral.

LEMMA 2. *If H is a compact Lie group, such that for all u in H there exists an element w of H such that $u^T = wuw^{-1}$, then, for any function f,*

$$\int f(u^T)du = \int f(u)du. \tag{18}$$

In order to prove Lemma 2, consider

$$\begin{aligned}\iint f(v^{-1}uv)dudv &= \int[\int f(v^{-1}uv)du]dv \\ &= \int[\int f(u)du]dv \\ &= V_H \int f(u)du\end{aligned} \tag{19}$$

where V_H is the group volume. But we have also

$$\iint f(v^{-1}uv)dudv = \int[f(v^{-1}uv)dv]dv. \tag{19a}$$

Choose now a w such that $u^T = wuw^{-1}$. Then

$$\begin{aligned}\iint f(v^{-1}uv)dudv &= \int[\int f(v^{-1}w^{-1}u^Twv)dv]du \\ &= \int[\int f(v^{-1}u^Tv)dv]du \\ &= \int[\int f(v^{-1}u^Tv)du]dv = V_H \int f(u^T)du,\end{aligned} \tag{20}$$

which proves the lemma.

We wish to show now that (11) holds for U_n with respect to the subgroup U_{n-1} which leaves a vector e_1 invariant. Replacing u by sq one obtains

$$\iint f(sRs^{-1}uTu^{-1})dsdu = \iint f(sRqTq^{-1}s^{-1})dsdq. \tag{21}$$

According to Lemma 1, every element of U_n and its transpose are in the same subclass. In particular, there must be an element z of U_{n-1} such that $RqTq^{-1} = z(RqTQ^{-1})^Tz^{-1}$. The z will be a function of q but not of s. Hence, s can be replaced by sz^{-1} so that

$$\begin{aligned}\iint f(sRs^{-1}uTu^{-1})dsdu &= \iint f(sz(RqTq^{-1})^Tz^{-1}s^{-1}dsdq \\ &= \iint f(s(RqTq^{-1})^Ts^{-1})dsdq \\ &= \iint f(s(q^T)^{-1}T^Tq^TR^Ts^{-1})dsdq.\end{aligned} \tag{22}$$

Because of Lemma 2, q^T can be replaced herein by q. There is, furthermore, some x in U_{n-1} such that $T^T = xTx^{-1}$ and some y such that $R^T = yRy^{-1}$. Hence (22) goes over into

$$\iint f(sRs^{-1}uTu^{-1})dsdu = \iint f(sq^{-1}xTx^{-1}qyRy^{-1}s^{-1})dsdq. \tag{23}$$

However, the integral over q remains invariant if one replaces q by xqs. After that, one can replace s by sy^{-1} and, finally, q by u^{-1}. Hence

$$\begin{aligned}\iint f(sRs^{-1}uTu^{-1})dsdu &= \iint f(q^{-1}TqsyRy^{-1}s^{-1})dsdq \\ &= \iint f(q^{-1}TqsRs^{-1})dsdq \\ &= \iint f(uTu^{-1}sRs^{-1})dsdu.\end{aligned} \tag{24}$$

This verifies (11) for our case leading to the result anticipated:

THEOREM 2. *The irreducible representations of U_n do not contain any representation of U_{n-1} more than once.*

One easily verifies that *the same is true for the subgroup $U_1 \times SU_{n-1}$*. However, it is not true for SU_{n-1} alone, no matter whether the original group were U_n or SU_n. This is particularly obvious for $n = 2$: in this case SU_1 contains only the identity element, whereas SU_2 is not commutative and has many dimensional irreducible representations.

4. The preceding demonstration of Theorem 2 may appear complicated. However, the only essential point is that an anti-isomorphism $R \to R'$ (resulting in $R'T' = (TR)'$) can be generated by the elements s of the subgroup ($R' = sRs^{-1}$) on all elements R of the full group. In the case considered, the anti-isomorphism was $R' = R^T$. This observation renders it possible to generalize Theorem 2.

THEOREM 3. *If an anti-isomorphism of the full group $R \to R' = sRs^{-1}$ can be induced on all elements R of the full group by the elements s of the subgroup (s depending on R), then no irreducible representation of the full group contains any representation of the subgroup more than once.*

Theorem 3, in contrast to (11), gives a condition for its conclusion which is sufficient but not necessary.

REFERENCES

1. M. Hamermesh, *Group theory and its application to physical problems* (Addison Wesley, London, 1962).
2. O. Taussky and H. Zassenhaus, *On the similarity transformation between a matrix and its transpose*, Pacific J. Math. *9* (1959), 893–896.
3. E. P. Wigner, *Condition that the irreducible representations of a finite group, considered as representations of a subgroup, do not contain any representation more than once*, Spectroscopic and Group Theoretical Methods in Physics, F. Loebl, Editor (North Holland, Amsterdam, 1968).

Princeton University,
Princeton, New Jersey;
University of Massachusetts,
Amherst, Massachusetts

Reduction of Direct Products and Restriction of Representations to Subgroups: The Everydays Tasks of the Quantum Theorists*

E. P. Wigner †

SIAM Journal on Applied Mathematics 25, 169–185 (1973)

Abstract. The direct product of two representations plays a fundamental role if the same principles of symmetry apply to two particles, or more generally to two physical systems, and if the two particles, or the two physical systems, are united and considered to be a single system. If one disregards the interaction of the two systems, the representations of their union is the direct product of the representations of the individual systems but if the interaction is taken into account, this direct product splits up into its irreducible components. Physicists have determined, therefore, the reduction of the direct products of the representations of almost all the groups which play an important role in physical theories, such as the groups of rotation in two-, three- and four-dimensional spaces, the various inhomogeneous groups in these dimensions, including the basic invariance group, the Poincaré group.

If the symmetry of a system is perturbed by an external field or otherwise, and hence the symmetry group reduced to a subgroup, the representations of the various states must be restricted to the subgroup. The resulting problem of the reduction of the restricted representation into its irreducible parts has also been considered, and largely solved, for all the groups mentioned before, the subgroups being almost all possible subgroups.

Some general results for both problems are presented.

The writer of this article wishes to apologize first, for the terminology he uses. It is the terminology used by his colleagues, physicists, and will appear antiquated and crude to the mathematician readers. Physicists are less addicted to an abstract language than are modern mathematicians and this applies particularly to the writer of the present article.[1] Second, he wishes to apologize for the incompleteness of the references to be furnished.[2] He is well aware that he has not kept up with the literature as well as did those for whom representation theory constitutes the primary interest. In fact, he has probably forgotten some of the articles he did read and the conclusions of which are reproduced below.

The difference between the roles which symmetry plays in classical and quantum mechanics. It may be good to recall the reason for the different role and the greater effectiveness of symmetry considerations in quantum than in classical mechanics. The reason is, briefly, that the set of quantum mechanical states forms a linear manifold, i.e., if ψ_1 and ψ_2 represent two states there are states $\alpha_1\psi_1 + \alpha_2\psi_2$ for any two complex numbers α_1 and α_2. In addition, it can be proved that the

* Received by the editors November 27, 1972. Presented by invitation at the Drexel University–Society for Industrial and Applied Mathematics Conference on Lie Algebras: Applications and Computational Methods, sponsored in part by a grant from the United States Air Force Office of Scientific Research, held at Drexel University, Philadelphia, Pennsylvania, June 15 and 16, 1972.

† Department of Physics, Princeton University, Princeton, New Jersey 08540.

[1] See, however, L. C. Biedenharn's article in *Group Theoretical Concepts and Methods in Elementary Particle Physics*, F. Gürsey, ed., Gordon and Breach, New York, 1964.

[2] For a general review of representation theory, see F. D. Murnaghan's book, *The Theory of Group Representations*, Dover, New York, 1963. Collections of articles dealing with our subject include (a) that mentioned in footnote 1; (b) *Symmetry Groups in Nuclear and Particle Physics*, F. J. Dyson, ed., Benjamin, New York, 1966; (c) *Spectroscopic and Group Theoretical Methods in Physics*, G. Racah, Memorial Volume, North Holland, Amsterdam, 1968; (d) *Group Theory and its Applications*, E. M. Loebl, ed., Academic Press, New York, 1968.

symmetry operations can be represented by unitary, or at worst antilinear, antiunitary operators. This means that if $\psi_1, \psi_2, \cdots$ are the state vectors of certain states, the state vectors of the states obtained from these by a symmetry transformation can be either written as $U\psi_1, U\psi_2, \cdots$, where U is a unitary linear operator, or as $U\psi_1^*$, $U\psi_2^*$, where the star denotes transition to the conjugate complex. If a symmetry operation can be written as the square of another symmetry operation, its representative will be the square of an operator which is either unitary or antiunitary and which is, hence, unitary. This applies to all rotations, displacements, proper Lorentz transformations and we shall restrict ourselves, hence, to the case of unitary symmetry operations. It can be shown then—and I will not repeat the proof—that a new basis ψ_k^j of the state vectors can be formed—linear combinations of any originally chosen set—which transform by irreducible representations of the symmetry operations R:

$$O_R\psi_\kappa^j = \sum_\lambda j(R)_{\lambda\kappa}\psi_\lambda^j; \tag{1}$$

O_R is the (unitary) operator which corresponds to the symmetry operation R, the $j(R)_{\lambda\kappa}$ is the $\lambda\kappa$ element of the matrix of the unitary irreducible representation j which corresponds to the operation R. It is the possibility of the formation of states ψ_κ^j, with transformation properties (1), which constitutes the greater effectiveness of symmetry considerations in quantum than in classical theory. It has no counterpart in the latter. Thus, if the symmetry operations consist of all rotations in ordinary space, one can superpose the states ψ rotated in all possible directions, that is, the states $O_R\psi$ with all R, and obtain a spherically symmetric state (unless the superposition gives the null vector). This corresponds to a state ψ_0^0 in (1), the $j(R)$ of which is 1 for all R. No corresponding construction is possible in classical theory: the states can be subjected to symmetry transformations, and thus new states obtained, but they cannot be superposed to obtain states with particularly simple properties, such as the states ψ_κ^j; each operation gives only a new state with properties similar to the original one.

The question naturally comes up concerning which augmentations of the symmetry group lead to new properties of the members of the basis sets ψ_κ^j and which will simply produce, from each such element, another element, i.e., have the role which all symmetry transformations have in classical theory. The symmetry operations of the former set will be called property providing transformations, those of the latter set relating transformations. Naturally, the character of an added transformation, whether property providing or relating, will in general depend on the set $\psi_1^j, \psi_2^j, \cdots$ and their irreducible representation $j(R)$. If we start from the group of space and time displacement transformations, and augment this group by the rotations in space and the Lorentz transformations, i.e., enlarge it to the Poincaré group, we have the group of property providing transformations which was called[3] "little group." Its elements had the property to transform the ψ_κ^j with the same j among themselves. Actually, the irreducible representations of the group of space-time displacements are one-dimensional

[3] E. P. Wigner, Ann. of Math., 40 (1939), p. 149. Also reprinted in footnote 2(b).

and can be characterized by a four-dimensional momentum vector p such that the operation of the space-time displacement T_a by a vector a has the effect

$$T_a\psi^p = e^{ip\cdot a}\psi^p. \tag{2}$$

The members of the property providing transformations (i.e., the elements of the little group) were then those Lorentz transformations L which left p unchanged: $Lp = p$. For a vector parallel to the time axis they were the ordinary rotations. The cosets of the group of the little group had effects similar to the effects of all symmetry operations of classical theory: they transformed the states ψ^p into states $\psi^{p'}$, with $p' \neq p$, and thus established relations between different states but did not furnish information on the properties of the states ψ^p. The elements of the little group, of the property providing transformations, did that.

I went into the question of property providing and relating transformations in some detail for three reasons. First, the question of the extension of the symmetry groups, to create greater though only approximate symmetry groups, forms and will continue to form an important endeavor in our theories. The question regarding which of the added transformations is in one or the other category gives clues concerning the significance of these transformations. Second, the augmentation of the symmetry group is the operation opposite to that of our last subject; of the restriction of the group to a subgroup. It is not the exact opposite because the augmentations considered are all such that the original group remains an invariant subgroup of the augmented group. But third, I wanted to call attention to a rather interesting mathematical problem. Let us consider a group g and embed it, as an invariant subgroup, into a larger group. Let us then consider an irreducible representation $j(R)$ of the original group g, and ask for the subgroup of the larger group into which that irreducible representation $j(R)$ can be extended without increasing its dimension. This will be, then, the property providing transformation in that case. It is easy to show that the elements L of the property providing transformation will be such that the representation

$$j_L(R) = j(LRL^{-1}) \tag{3}$$

is equivalent to the original $j(R)$. The right side of (3) is defined because LRL^{-1} is an element of the original, smaller, group. Naturally, the product of two L for which (3) is equivalent to the original $j(R)$ has also this property and the elements of the form LR, where L has the property just given, $R \subset g$, form a group which is a subgroup of the whole augmented group. However, the determination of the properties of this group, i.e., the set of the L for which there is a matrix $d(L)$ depending on j and L but independent of R such that for all $R \subset g$,

$$j(LRL^{-1}) = d(L)j(R)d(L)^{-1}, \tag{3a}$$

has not been carried out to my knowledge, not even in the case of finite groups. Only if the augmented group consists only of the elements L^qR, $0 \leqq q \leqq p$, where $R \subset g$ and L^p is also a member of g, can this be done easily. In this case the number of irreducible representations for which (3a) has a solution is equal to the number of classes C of g which are left invariant by L so that $LCL^{-1} = C$.

One can see that C and LCL^{-1} are either identical or have no element in common: if $LXL^{-1} = ZXZ^{-1}$ with $X, Z \subset g$, then if also $Y \subset g$, one has also for YXY^{-1},

$$\begin{aligned} L(YXY^{-1})L^{-1} &= (LYL^{-1})(LXL^{-1})(LYL^{-1})^{-1} \\ &= (LYL^{-1})ZXZ^{-1}(LYL^{-1})^{-1}, \end{aligned} \tag{4}$$

but LYL^{-1} is in g since this is an invariant subgroup. The case that the subgroup is augmented by a single element and its powers is far from the most general case but it is an example and can be handled quite easily.

Because of (3a) and the irreducible nature of the $j(R)$ representation, the representative of L must be a constant times the $d(L)$ of (3a), $\omega d(L)$. Let p be the lowest power of L such that $L^p = s \subset g$. Then the representative of s, that is, $j(s) = \omega^p d(L)^p$ and we choose ω in such a way that this equation becomes valid; in fact we assume that the indeterminate constant was so chosen in $d(L)$ that $d(L)^p = j(s)$. The representative of $L^k r$ then becomes $d(L)^k j(r)$ and we should verify that this is a representation. In fact,

$$\begin{aligned} d(L)^k j(r) d(L)^m j(r') &= d(L)^{k+m} d(L)^{-m} j(r) d(L)^m j(r') \\ &= d(L)^{k+m} j(L^{-m} r L^m) j(r') = d(L)^{k+m} j(L^{-m} r L^m r'). \end{aligned} \tag{4a}$$

The third member is obtained by a repeated application of (3a). If $k + m < p$, the verification is complete; if $k + m \geqq p$, one writes instead of the last member

$$d(L)^{k+m-p} d(L)^p j(L^{-m} r L^m r') = d(L)^{k+m-p} j(s L^{-m} r L^m r'),$$

and this completes the verification.

The number of representations $j(R)$ for which $j'(r) = j(LrL^{-1})$ is equivalent to $j(r)$ is given by

$$\frac{1}{h} \sum_j \sum_r \chi^j(LrL^{-1}) \chi^j(r)^*, \tag{5}$$

where h is the order of g, and $\chi^j(r)$ is the character of the element r in the representation $j(r)$. In fact, the sum over r is h if the representations $j(r)$ and $j'(r) = j(LrL^{-1})$ are equivalent, 0 otherwise. However, because of the orthogonality relations of the characters of the irreducible representations, the sum over j, carried out first, gives $(h/n_r)\Delta(LrL^{-1}, r)$ where n_r is the member of elements in the conjugacy class of r, and $\Delta(r', r) = 1$ if r' and r are in the same conjugacy class, 0 otherwise. Hence (5), the number of irreducible representations which can be augmented by themselves to representations of the large group $L^k g$, $0 \leqq k < p$, becomes

$$\sum_r \Delta(LrL^{-1}, r)/n_r. \tag{5a}$$

As we saw at (4), if $r \subset C$ and $LrL^{-1} \subset C$, this applies for all elements of C. These then contribute 1 to the sum (5a), the r in classes which are not invariant under conjugation by L contribute 0 and the theorem follows.

Direct products of irreducible representations. If the union of two physical systems, for example, of two atoms, is considered as a single physical system, the Hilbert space of the union is the direct product of the Hilbert spaces of the constituents. If the invariance group G of both is the same, and if the union is only conceptual, the invariance group of the union will consist of all the transformations

$R \times R'$, both R and R' being members of G and R acting in the Hilbert space of the first, R' in that of the second system. However, if the two systems interact, the invariance group of the union will be reduced to the elements $R \times R$, what Mackey calls the diagonal elements of the direct product: the invariance applies only if both systems of the union are subjected to the same transformation. If we think of the states of the union in which the first system is in one of the states ψ_κ^j, transforming by $j(R)$, the second system in one of the states $\varphi_{\kappa'}^{j'}$ the union will transform by the Kronecker product $j(R) \times j'(R)$ of the two representations j and j'. The problem then is to find the linear combinations of the $\psi_\kappa^j \times \varphi_{\kappa'}^{j'}$ which transform by an irreducible representation of the symmetry group, i.e., the reduction of the representation $j \times j'$. This will provide important information on the properties of the states of the interacting union.

The situation is particularly simple in the case of the Poincaré group. One can start with the invariant subgroup of space-time displacements; the representations of this with respect to the two subsystems are characterized, according to (2), by four-vectors p and p'. The direct product of these is still one-dimensional and is characterized by the four-vector $p + p'$. If p and p' are parallel, their property providing transformations are the same: the two groups of transformations which leave two parallel axes invariant are identical. The property providing transformations of the union are also the same, for the same reason, and the direct product of the property providing transformations of the constituents gives the property providing transformations of the union. Hence, in the case considered, as far as these transformations are concerned, the problem reduces to the reduction of the Kronecker product of two representations of their group which is, if p (and hence p') is time-like, the three-dimensional rotation group R_3 (or SU_2). The relating transformations of the union's representation $p + p'$ of the space-time displacement group are also the same as those of the constituents and do not contribute additional information on the total representations of the union.

In the more general case in which p and p' are not parallel the situation is somewhat more complicated[4] and will not be discussed here in detail. Essentially, one has to deal in that case with the reduction of the Kronecker product $j(R) \times j'(R) \times j''(R)$ of three representations of R_3 (or SU_2).

The significance of the problem of the reduction of the Kronecker product $j(R) \times j'(R)$ of representations of R_3 and SU_2 was recognized quite early and there are many articles and several books on the subject.[5] It will not be taken

[4] L. Michel, article in the Report of the 1953 IUPAP Meeting (p. 272); A. S. Wightman, article in *Relations de dispersion et particules élémentaires*, Cecile de Witt, ed., Hermann, Paris, 1960, and several other articles; W. Rühl, *The Lorentz Group and Harmonic Analysis*, Benjamin, New York, 1970.

[5] See, for example, E. Feenberg and G. E. Pake, *Notes on the Quantum Theory of Angular Momentum*, Addison-Wesley, Reading, Mass., 1953; A. R. Edmonds, *Angular Momentum in Quantum Mechanics*, Princeton University Press, Princeton, N.J., 1957; M. E. Rose, *Elementary Theory of Angular Momentum*, John Wiley, New York, 1957; W. T. Sharp, *The Quantum Theory of Angular Momentum*, Rep. CRL-43, Chalk River, Ontario, 1957. Among more recent publications, see *Quantum Theory of Angular Momentum*, L. C. Biedenharn and H. VanDam, eds., Academic Press, New York, 1965, E. El-Baz, *Traitement graphique de l'algebre des moments angulaires*, Masson et Cie, Paris, 1969. The most advanced book on the question, incorporating some results to be discussed later, remains, in this writer's opinion, A. P. Yutsis, I. B. Levinson and V. V. Vanagas, *The Theory of Angular Momentum*, Vilnius, USSR, 1960 (translated by A. and A. R. Sen, Jerusalem, 1962).

up here again. Instead, some less well-known results will be presented, relating to the general problem of reducing Kronecker products of irreducible representations of arbitrary groups.

The irreducible representations of the underlying group will be denoted by $j, k, l, \cdots$ with or without primes. The sign $\langle f(R)\rangle$ will denote the group average of the expression $f(R)$ in the braces; this is the sum over all group elements divided by the order of the group in the case of finite groups, the group integral divided by the group integral of 1 in the case of compact continuous groups and we shall restrict the rest of this section to such groups. All representations will be assumed to be in unitary form. Thus the orthogonality relations read in this notation,

$$\langle j(R)_{\iota\iota'} k(R)^*_{\kappa\kappa'}\rangle = \delta_{jk}\delta_{\iota\kappa}\delta_{\iota'\kappa'}/[j], \tag{6}$$

where $[j]$ is the dimension of the representation j, the star denotes complex conjugation. The character of $j(R)$ will be denoted by $\chi^{(j)}(R) = \sum_\iota j(R)_{\iota\iota}$. We further denote the number of times the conjugate complex of $l(R)$, that is $l(R)^*$, is continued in the direct product of the representations j and k by

$$(j, k, l) = \langle \chi^{(j)}(R)\chi^{(k)}(R)\chi^{(l)}(R)\rangle. \tag{7}$$

This symbol also gives the number of times $j(R)^*$ is contained in $k(R) \times l(R)$, or $k(R)^*$ in $j(R) \times l(R)$: the symbol (j, k, l) is symmetric in j, k and l.

It is useful next to define the matrix[6]

$$[j, k, l]_{\iota\kappa\lambda;\iota'\kappa'\lambda'} = \langle j(R)_{\iota\iota'} k(R)_{\kappa\kappa'} l(R)_{\lambda\lambda'}\rangle. \tag{8}$$

Since $\langle f(R)\rangle = \langle f(RS)\rangle_R = \langle f(SR)\rangle_R$ (the index on the brace indicates the variable over which the average is taken), one obtains

$$[j, k, l]_{\iota''\kappa''\lambda'';\iota\kappa\lambda} = \sum_{\iota'\kappa'\lambda'} [j, k, l]_{\iota''\kappa''\lambda'';\iota'\kappa'\lambda'} j(S)_{\iota'\iota} k(S)_{\kappa'\kappa} l(S)_{\lambda'\lambda}. \tag{8a}$$

Averaging this over S, one finds with (8),

$$[j, k, l] = [j, k, l]^2. \tag{8b}$$

Since we deal only with finite or compact groups, we also have $\langle f(R)\rangle = \langle f(R^{-1})\rangle$ and this gives, because of the assumed unitary form of the representations,

$$[j, k, l] = [j, k, l]^\dagger, \tag{8c}$$

the dagger denoting the Hermitian adjoint: the $[j, k, l]$ are idempotent self-adjoint matrices, that is, projection operators. Their characteristic values are all 0 to 1 and the number of the latter is, because of (7),

$$\operatorname{tr}[j, k, l] = (j, k, l). \tag{8d}$$

The $[j, k, l]$ are (j, k, l)-dimensional projection operators in their $[j][k][l]$-dimensional space. They can be given the form

$$[j, k, l]_{\iota'\kappa'\lambda';\iota\kappa\lambda} = \sum_\rho \begin{pmatrix} j & k & l \\ \iota' & \kappa' & \lambda' \end{pmatrix}^*_\rho \begin{pmatrix} j & k & l \\ \iota & \kappa & \lambda \end{pmatrix}_\rho, \tag{9}$$

[6] This matrix was first defined in G. W. Mackey's work (cf. below) and then independently, by Sharp and Derome (cf. below).

the summation over ρ running over (j, k, l) values. The "three-j-symbols" on the right[7] are the conjugate complexes of the characteristic vectors of $[j, k, l]$ which belong to the characteristic value 1. They form an orthonormal system, which is, as a rule, incomplete.

$$\sum_{\iota\kappa\lambda} \begin{pmatrix} j & k & l \\ \iota & \kappa & \lambda \end{pmatrix}_\rho^* \begin{pmatrix} j & k & l \\ \iota & \kappa & \lambda \end{pmatrix}_\sigma = \delta_{\rho\sigma}. \tag{10}$$

The three-j-symbols are determined, by their definition (9), only up to a unitary transformation: the original symbol can be replaced by

$$\begin{pmatrix} j & k & l \\ \iota & \kappa & \lambda \end{pmatrix}_\rho' = \sum_s u_{\rho\sigma} \begin{pmatrix} j & k & l \\ \iota & \kappa & \lambda \end{pmatrix}_\sigma, \tag{10a}$$

u being any (j, k, l)-dimensional unitary matrix. Most of the mathematical theorems which will be reviewed below relate to the numbers (j, k, l), not to the matrices $[j, k, l]$. Even though this will be well known to most readers, it may be useful, therefore, to give an additional and, in fact, the principal reason for the physicist's interest in the matrices $[j, k, l]$ and their characteristic vectors.

The physicist often considers two sets of states φ_ι, $\iota = 1, 2, \cdots, [j]$, and ψ_λ, $\lambda = 1, 2, \cdots, [l]$ (this is not the usual notation) which transform, under the unitary operations O_R of a group by irreducible representations

$$O_R\varphi_\iota = \sum_{\iota'} j(R)_{\iota'\iota}\varphi_{\iota'}, \tag{11a}$$

$$O_R\psi_\lambda = \sum_{\lambda'} l(R)_{\lambda'\lambda}\psi_{\lambda'}. \tag{11b}$$

He is then interested in the matrix elements of a set of self-adjoint operators q_κ which he calls the components of an irreducible tensor operator. They transform as

$$O_R q_\kappa O_R^{-1} = \sum_{k'} k(R)_{\kappa'\kappa} q_{\kappa'}. \tag{11c}$$

The physicist is then interested in the scalar products ("matrix elements")

$$Q^\kappa_{\iota\lambda} = (\varphi_\iota, q_\kappa\psi_\lambda). \tag{12}$$

The group may be the group of rotations in ordinary space, the φ_ι the state vectors of a given energy in a spherically symmetric field, the ψ_λ the state vectors of another energy. The operators q_κ may represent the components of the electric dipole moment, i.e., the sum of the coordinate of the particles, multiplied with the electric charges of these. The squares of the $Q^\kappa_{\iota\lambda}$ are then proportional to the optical transition probabilities from state φ_ι to ψ_λ caused by an oscillating electric field in the κ direction.

Because of the unitary nature of the O_R one has

$$\begin{aligned} Q^\kappa_{\iota\lambda} &= (\varphi_\iota, q_\kappa\psi_\lambda) = (O_R\varphi_\iota, O_R q_\kappa\psi_\lambda) = (O_R\varphi_\iota, O_R q_\kappa O_R^{-1} O_R\psi_\lambda) \\ &= \sum_{\iota'} \sum_{\kappa'\lambda'} j(R)^*_{\iota'\iota} k(R)_{\kappa'\kappa} l(R)_{\lambda'\lambda} Q^{\kappa'}_{\iota'\lambda'}. \end{aligned} \tag{13}$$

[7] Three-j-symbols were first defined by the present writer in the article reprinted in *Quantum Theory of Angular Momentum* (see footnote 5). See also *Group Theory and its Application to the Quantum Theory of Atomic Spectra*, Chap. 4, Academic Press, New York, 1959.

Averaged over all group elements, this gives

$$Q^{\kappa}_{\iota\lambda} = \sum_{\iota'\kappa'\lambda'} [j^*, k, l]_{\iota'\kappa'\lambda';\iota\kappa\lambda} Q^{\kappa'}_{\iota'\lambda'} \tag{14}$$

or, because of (8c),

$$Q^{\kappa}_{\iota\lambda} = \sum_{\iota'\kappa'\lambda'} [j^*, k, l]^*_{\iota\kappa\lambda;\iota'\kappa'\lambda'} Q^{\kappa'}_{\iota'\lambda'}. \tag{15}$$

It follows that the physicist's "matrix elements" $Q^{\kappa}_{\iota\lambda}$ are the components of a characteristic vector of the matrix $[j^*, k, l]^*$ belonging to the characteristic value 1—or the conjugate complexes of the components of such a characteristic value of $[j^*, k, l]$: they are linear combinations of the three-j-symbols

$$Q^{\kappa}_{\iota\lambda} = \sum_{\rho} c_{\rho} \begin{pmatrix} j^* & k & l \\ \iota & \kappa & \lambda \end{pmatrix}_{\rho}. \tag{16}$$

The c_ρ depend on the set φ, the set ψ, the q_κ, but not on the ι, κ, λ. The formula for the matrix elements is most telling if the sum contains only one term, i.e., if $(j, k, l) = 1$. A group for which all (j, k, l) are 0 or 1 is called multiplicity free[8]; the group of rotations in ordinary space, R_3, does have this property. Naturally, if $(j, k, l) = 0$, all $Q^{\kappa}_{\iota\lambda} = 0$.

The matrix $[j, k, l]$ and the three-j-symbols have several important properties, the principal ones among them deriving from their relation to the matrix U^{jk} which brings the Kronecker product $j(R) \times k(R)$ into the reduced form. The reciprocal of this U^{jk} brings, conversely, the reduced form into the form

$$j(R)_{\iota\iota'} k(R)_{\kappa\kappa'} = \sum_{\substack{l'\rho \\ \lambda\lambda'}} U^{jk}_{\iota\kappa;l'\lambda\rho} l'(R)^*_{\lambda\lambda'} (U^{jk}_{\iota'\kappa';l'\lambda'\rho})^*. \tag{17}$$

The irreducible parts of $j \times k$ are denoted by $l'(R)^*$; the index ρ, running from 1 to the number of $l'(R)^*$ in the reduced form, that is, to (j, k, l'), takes care of the multiplicity of $l'(R)^*$ in the reduced form. The U^{jk} can be transformed in a way similar to the three-j-symbols, that is, as (16) shows, without altering the right side of (17). Introducing (17) into the definition (8) of $[j, k, l]$, one obtains by means of the orthogonality relation

$$[j, k, l]_{\iota\kappa\lambda;\iota'\kappa'\lambda'} = \sum_{\rho} U^{jk}_{\iota\kappa;l\lambda\rho} (U^{jk}_{\iota'\kappa';l\lambda'\rho})^* / [l]. \tag{18}$$

Comparison of this with (9), taking into account the arbitrariness in the definition of the three-j-symbols (or of the U^{jk}), permits one to write

$$\begin{pmatrix} j & k & l \\ \iota & \kappa & \lambda \end{pmatrix}_{\rho} = [l]^{-1/2} U^{jk}_{\iota\kappa;l\lambda\rho}. \tag{19}$$

[8] Multiplicity free groups were defined and studied by G. W. Mackey, Amer. J. Math., 75 (1953), p. 387, Pacific J. Math., 8 (1958), p. 503. A special class of such groups and their representations were studied by the present writer, Amer. J. Math., 63 (1941), p. 57. See also footnote 2(b). (The R_3 and SU_2 groups belong to this class.) The general case was studied by W. T. Sharp and J. R. Derome, J. Mathematical Phys., 6 (1965), p. 1584, and J. R. Derome, Ibid., 7 (1966), p. 612.

Hence, because of the unitary nature of the U^{jk} one has

$$\sum_{\iota\kappa} \begin{pmatrix} j & k & l \\ \iota & \kappa & \lambda \end{pmatrix}_\rho \begin{pmatrix} j & k & l' \\ \iota & \kappa & \lambda' \end{pmatrix}^*_\sigma = \delta_{ll'}\delta_{\lambda\lambda'}\delta_{\rho\sigma}/[l] \tag{20}$$

and similar equations apply for the other pairs j, l and k, l. Thus, for example,

$$\sum_{\iota\lambda} \begin{pmatrix} j & k & l \\ \iota & \kappa & \lambda \end{pmatrix}_\rho \begin{pmatrix} j & k' & l \\ \iota & \kappa' & \lambda \end{pmatrix}_\sigma = \delta_{kk'}\delta_{\kappa\kappa'}\delta_{\rho\sigma}/[k]. \tag{20a}$$

Actually, (10) is a special case of (20), or of (20a) obtained by setting $l' = l$, $\lambda' = \lambda$ and summing over λ. The other orthogonality relations are

$$\sum_\rho \sum_{l\lambda} [l] \begin{pmatrix} j & k & l \\ \iota & \kappa & \lambda \end{pmatrix}_\rho \begin{pmatrix} j & k & l \\ \iota' & \kappa' & \lambda \end{pmatrix}^*_\rho = \delta_{\iota\iota'}\delta_{\kappa\kappa'}. \tag{21}$$

These, if expressed in terms of the U^{jk} (or U^{jl} in the case of (20a)), are obvious; the purpose of the derivation by means of the projection operator $[j, k, l]$ was only to bring out the connection between U^{jk} and U^{jl}, U^{kl}. Thus, it appears that it is possible to set

$$[l]^{-1/2} U^{jk}_{\iota\kappa;l\lambda\rho} = [k]^{-1/2} U^{jl}_{\iota\lambda;k\kappa\rho} = [j]^{-1/2} U^{kl}_{\kappa\lambda;j\iota\rho}. \tag{22}$$

However, the U will not be used below and we shall use, instead of (17),

$$\begin{aligned} j(R)_{\iota\iota'} k(R)_{\kappa\kappa'} &= \sum_{l\rho} \sum_{\lambda\lambda'} [l] \begin{pmatrix} j & k & l \\ \iota & \kappa & \lambda \end{pmatrix}_\rho l(R)^*_{\lambda\lambda'} \begin{pmatrix} j & k & l \\ \iota' & \kappa' & \lambda' \end{pmatrix}^*_\rho \\ &= \sum_l \sum_{\lambda\lambda'} [l][j, k, l]_{\iota\kappa\lambda;\iota'\kappa'\lambda'} l(R)^*_{\lambda\lambda'}. \end{aligned} \tag{23}$$

Similar equations hold, naturally for the Kronecker product[9]

$$j(R)_{\iota\iota'} \times l(R)_{\lambda\lambda'} = \sum_k \sum_{\kappa\kappa'} [k][j, k, l]_{\iota\kappa\lambda;\iota'\kappa'\lambda'} k(R)^*_{\kappa\kappa'} \tag{23a}$$

and also for $k(R) \times l(R)$. Actually, these equations are direct consequences of (8). The unitary nature of U does result, however, in significant relations for the $[j, k, l]$:

$$\sum_{l\lambda} [l][j, k, l]_{\iota\kappa\lambda;\iota'\kappa'\lambda} = \delta_{\iota\iota'}\delta_{\kappa\kappa'}, \tag{24}$$

$$\sum_{\iota\kappa} [j, k, l]_{\iota\kappa\lambda;\iota\kappa\lambda'} = (j, k, l)\delta_{\lambda\lambda'}/[l]. \tag{24a}$$

Again, because of the invariance of the $[j, k, l]$ with respect to the interchanges of the triples (j, ι, ι'), (k, κ, κ'), (l, λ, λ'), equations similar to (24) apply for the sum over k and κ, or j and ι, and the analogue of (24a) applies if the sum is extended over the $\iota\lambda$ or $\kappa\lambda$ pairs. Equation (8d) is, incidentally, a direct consequence of (24a).

[9] The considerations which follow can be regarded to constitute a reformulation of the elementary parts of the Racah algebra for not multiplicity free groups. For reviews of multiplicity free groups' Racah algebra, cf. the references in footnote 5. The SU_n groups' Racah algebra was studied by J. R. Derome and G. Jakimow, Canad. J. Phys., 48 (1970), p. 2169.

Another set of relations can be derived either from the consideration of expressions similar to (8) but with the product of four representations averaged over the group, or by considering the operation $T^{j\iota\iota'}$ of multiplication by $j(R)_{\iota\iota'}$. The matrix elements of this are given by

$$j(R)_{\iota\iota'}k(R)_{\kappa\kappa'} = \sum_{l}\sum_{\lambda\lambda'} T^{j\iota\iota'}_{k\kappa\kappa';l\lambda\lambda'}l(R)_{\lambda\lambda'}. \tag{25}$$

Multiplication by $m(R)^*_{\mu\mu'}$ and averaging over the group then gives

$$[j,k,m^*]_{\iota\kappa\mu;\iota'\kappa'\mu'} = [m]^{-1}T^{j\iota\iota'}_{k\kappa\kappa';m\mu\mu'} \tag{25a}$$

Since one obtains the same result whether one multiplies with $j(R)_{\iota\iota'}$ first and with $n(R)_{\nu\nu'}$ afterwards, or the other way around, the corresponding T, and hence the projection matrices, also commute. With a slight change in the symbols considered one obtains[10]

$$\begin{aligned}&\sum_{m}\sum_{\mu\mu'}[j,k,m^*]_{\iota\kappa\mu;\iota'\kappa'\mu'}[m][n,l,m]_{\nu\lambda\mu;\nu'\lambda'\mu'}\\ &= \sum_{m}\sum_{\mu\mu'}[n,k,m^*]_{\nu\kappa\mu;\nu'\kappa'\mu'}[m][j,l,m]_{\iota\lambda\mu;\iota'\lambda'\mu'}.\end{aligned} \tag{26}$$

The representation m^* is assumed to have the form of the conjugate complex of m; the right side of (26) differs from the left side by having j, ι, ι' and n, ν, ν' interchanged. A relation similar to (26) can be derived also, and in the same way, for the (j,k,l) symbols defined in (7). We have, in fact, in this case,

$$\sum_{m}(j,k,m^*)(n,l,m) = \sum_{m}(n,k,m^*)(j,l,m) = \sum_{m}(k,l,m^*)(j,n,m). \tag{26a}$$

Naturally, the role of m and m^* can be interchanged. There is one more formula which seems worth mentioning: setting $\nu = \nu'$, $\kappa = \kappa'$ in (26) and summing it over these indices, one obtains because of (24a),

$$\begin{aligned}&\sum_{m}\sum_{\mu\mu'}\sum_{\kappa}[k,j,m^*]_{\kappa\iota\mu;\kappa\iota'\mu'}[m]\sum_{\nu}[n,l,m]_{\nu\lambda\mu;\nu\lambda'\mu'}\\ &= \sum_{m\mu}(n,k,m^*)[j,l,m]_{\iota\lambda\mu;\iota'\lambda'\mu}.\end{aligned} \tag{26b}$$

Actually, setting $\iota = \iota'$, $\lambda = \lambda'$, and summing over these indices one recovers (26a). If one were to define self-adjoint matrices K, N, etc., with elements

$$K_{j\iota\iota';m\mu\mu'} = [j]^{1/2}\sum_{\kappa}[k,j,m^*]_{\kappa\iota\mu;\kappa\iota'\mu'}[m]^{1/2} \tag{27}$$

and similarly for N and M (their dimension is of the order of the group), (26b) would assume the form

$$KN = \sum_{m}(k,n,m^*)M = NK, \tag{27a}$$

the summation extending over all irreducible representations.

As the preceding discussion indicates, the number of identities obeyed by the $[j,k,l]$, etc., projection operators is very large; they lead one in fact to suspect

[10] Equation (26) is, for multiplicity free groups, a Racah identity.

that there is a simpler way to deal with them. They do give directly expressions for all averages (sums or integrals) of the form

$$\langle j(R^\alpha)_{\iota\iota'} k(R^\beta)_{\kappa\kappa'} \cdots \rangle \tag{28}$$

with an arbitrary number of factors in the braces, the $\alpha, \beta, \cdots$ being integers. These expressions are simplest, of course, in the case of multiplicity free groups where the expression (9) for them is just a product of two factors. Actually, few of the preceding formulae are *needed* to obtain an expression for (28) in terms of the $[j, k, l]$; (23a) suffices in principle. The other formulae can, however, simplify the resulting expressions; the purpose of (26) is, principally, to transform into each other the different expressions resulting from applying (23a) to (28) in different orders.

The $[j, k, l]$ are trivial for commutative groups; they are well known for the multiplicity free R_3 and SU_2 (the most important applications in physics), for SU_3 and some more complex groups.[11] These will not be reviewed here. Instead, some theorems will be presented—few of them new—for the formulation of which it is not necessary to assume a definite form of the representations. They will deal principally with the characters of representations and the quantities (j, k, l) defined in (7).

Some identities connecting multiplicities (j, k, l) with other group averages. Before entering into a detailed discussion of the identities in question we note that if two—or all three—of the representations j, k are the same, there are, in addition to the

$$(j, k, l) = \sum_{\iota\kappa\lambda} \langle j(R)_{\iota\iota} k(R)_{\kappa\kappa} l(R)_{\lambda\lambda} \rangle \tag{29}$$

defined in (7), other quantities which remain unchanged if the representations are subjected to a similarity transformation. Such a quantity is, for instance,

$$\sum_{\iota\kappa\lambda} \langle j(R)_{\iota\kappa} j(R)_{\kappa\iota} l(R)_{\lambda\lambda} \rangle = \langle \chi^j(R^2) \chi^l(R) \rangle . \tag{30}$$

One can find an expression for (30) by noting that the matrix $[j, j, l]_{\iota\kappa\lambda;\iota'\kappa'\lambda'}$ remains invariant if the ι and κ are interchanged. Hence, a characteristic vector thereof,

$$\begin{pmatrix} j & j & l \\ \iota & \kappa & \lambda \end{pmatrix}_\rho ,$$

remains a characteristic vector if its ι and κ are interchanged. It is possible, therefore, to choose these characteristic vectors in such a way that they belong to an irreducible representation of the group containing the interchanges of these two indices, i.e., to choose some of them to be symmetric, some antisymmetric with respect to this interchange. We denote the number of the former by $(j, j, l)_s$, of the latter by $(j, j, l)_a$. Naturally,

$$(j, j, l)_s + (j, j, l)_a = (j, j, l) = \langle \chi^j(R)^2 \chi^l(R) \rangle \tag{31}$$

[11] See footnote 9 and also J. R. Derome, J. Mathematical Phys., 8 (1967), p. 714.

but we have also, for the expression in (30), because of (9),

$$\langle \chi^j(R^2)\chi^l(R)\rangle = \sum_{\iota\kappa\lambda} [j,j,l]_{\iota\kappa\lambda;\kappa\iota\lambda}$$

(31a)

$$= \sum_{\iota\kappa\lambda}\sum_{\rho} \begin{pmatrix} j & j & l \\ \iota & \kappa & \lambda \end{pmatrix}_{\rho}^{*} \begin{pmatrix} j & j & l \\ \kappa & \iota & \lambda \end{pmatrix}_{\rho} = (j,j,l)_s - (j,j,l)_a.$$

The last member follows because the interchange of κ and ι in the second three-j-symbol will leave it unchanged for $(j, j, l)_s$ of the ρ, and change its sign for $(j, j, l)_a$ of them. After the change, the summation can be carried out by (10).

If all three representations are identical, the three-j-symbols' indices ι, κ, λ can be arbitrarily interchanged. Hence, one can choose them in such a way that they belong, with respect to the interchange of their indices ι, κ, λ, to an irreducible representation of the symmetric group S_3. We denote the number of those which are totally symmetric by $(j)_s$, the number of those which are totally antisymmetric by $(j)_a$, and the number of *pairs* which belong to the two-dimensional, degenerate representation by $(j)_d$. Again,

(32) $$(j)_s + (j)_a + 2(j)_d = (j,j,j) = \langle \chi^j(R)^3\rangle$$

and one can verify also that

(32a) $$\langle \chi^j(R^2)\chi^j(R)\rangle = (j)_s - (j)_a$$

and

(32b) $$\langle \chi^j(R^3)\rangle = (j)_s + (j)_a - (j)_d.$$

These formulae are contained, essentially, already in the articles of Sharp and Derome (see footnote 8). Incidentally, (31a) contains, in a sense, the identity of Frobenius and Schur.[12] Setting for l the trivial representation by the one-dimensional unit matrix to be denoted by 0 one obtains

(31b) $$\langle \chi^j(R^2)\rangle = (j,j,0)_s - (j,j,0)_a.$$

Now $j \times 0 = j$ and if this is not equivalent to j^*, the $(j, j, 0)$ vanishes. If it is equivalent, (31b) shows that if the left side is 1, the transformation can be effected by a unitary symmetric matrix (cf. (23)) which can be diagonalized by a real orthogonal matrix. One can then transform j by the product of the square root of the diagonal part and this real matrix and bring it thereby into real form. If the left side of (31b) is -1, the matrix transforming j into j^* is necessarily a skew symmetric matrix and j cannot be made real by a similarity transformation—if it were real, it could be transformed into its conjugate complex by the unit matrix. However, this derivation of the Frobenius–Schur relation

(33) $$\langle \chi^j(R^2)\rangle = C_j$$

(with $C_j = 0$ if some χ^j is complex, 1 if all $j(R)$ can be made real, $C_j = -1$ if $\chi^j(R)$ is real but the $j(R)$ cannot be made real by a similarity transformation) is not fundamentally different from the original one.

[12] G. Frobenius and I. Schur, Berl. Ber., 1906, p. 186.

Let me quote, without proof, three theorems for which I am responsible:[13]

$$\sum_{ljk}(l,j,k)^2 = \langle (h/n_R)^2 \rangle = \sum_R h/n_R^2, \tag{34}$$

where h is the order of the group, n_R the number of elements in R's conjugacy class;

$$\sum_j C_j \chi^j(S) = \zeta(S), \tag{34a}$$

$\zeta(S)$ being the number of the square roots of S;

$$\langle \zeta(R)^3 \rangle = \sum_{ljk} C_l C_j C_k (l,j,k). \tag{34b}$$

Since every group element is the square root of a group element,

$$\langle \zeta(R) \rangle = 1.$$

The very simple equation

$$\langle \xi(R)^2 \rangle = \text{number of real classes}, \tag{34c}$$

"real" meaning that it contains with an element also its reciprocal, must have been known before.[14] Let me also mention an equation derived originally by Frame:[15]

$$\langle \chi^k(R)/n_R \rangle = \sum_j (j,j^*,k)/h. \tag{35}$$

Without modification, these equations apply only to finite groups; a method to formulate them for compact groups was, however, proposed lately.[16]

Let me present, finally, three identities which I have not seen before in the literature. They do not appear to be of great importance but they do illustrate the great variety of identities. The first of them is an analogue of (34a) but gives an expression for the number of cube roots of an element, to be denoted by $\eta(S)$, rather than for the number of square roots. The second one is an analogue of (32b) or (33), but refers to the fourth rather than the third or second powers. The last one is the analogue of (34c) but refers again to the cube rather than square roots.

Clearly,

$$\eta(R) = \sum_S \delta_{R,S^3} = \sum_S \Delta_{R,S^3}/n_R, \tag{36}$$

where $\delta_{R,T}$ is 1 or 0 depending on whether $R = T$ does or does not hold; $\Delta_{R,T}$ is 1 if R and T are in the same class, 0 otherwise. It then follows from the orthogonality relation

$$\sum_j \chi^j(R)\chi^j(T)^* = h\Delta_{R,T}/n_R$$

[13] E. P. Wigner, Amer. J. Math., 63 (1941), p. 57.

[14] As was pointed out already in the article in footnote 13, the relation of equation (34c) was probably known already to the German founders of representation theory (see footnote 12).

[15] J. S. Frame, Bull. Amer. Math. Soc., 53 (1947), p. 584. Cf. also Ibid., 54 (1948), p. 740.

[16] F. E. Goldrich and E. P. Wigner, Canad. J. Math., 24 (1972), p. 432.

that

$$\eta(R) = \sum_S \sum_j \chi^j(R)\chi^j(S^3)^*/h = \sum_j \langle \chi^j(R)\chi^j(S^3)^* \rangle_S$$
$$= \sum_j \chi^j(R)\{(j)_s + (j)_a - (j)_d\}. \tag{36a}$$

This is the first identity. The second one, the analogue of (32b), refers to

$$\langle \chi^j(R^4) \rangle = \sum_{\iota\kappa} \langle j(R^2)_{\iota\kappa} j(R^2)_{\kappa\iota} \rangle$$
$$= \sum_{\iota\kappa} \sum_{l\rho} \sum_{\lambda\lambda'} [l] \begin{pmatrix} j & j & l \\ \iota & \kappa & \lambda \end{pmatrix}_\rho \langle l(R^2)^*_{\lambda\lambda'} \rangle \begin{pmatrix} j & j & l \\ \kappa & \iota & \lambda' \end{pmatrix}^*_\rho . \tag{37}$$

Now $\langle l(R^2) \rangle$ contains equally many elements from any class. It is, therefore, a multiple of the unit matrix (the same applies to any $\langle l(R^n) \rangle$), and it is given by the Frobenius–Schur formula (33):

$$\langle l(R^2)_{\lambda\lambda'} \rangle = C_l \delta_{\lambda\lambda'}/[l]. \tag{33a}$$

Inserting this into (37), one obtains

$$\langle \chi^j(R^4) \rangle = \sum_\rho \sum_{\iota\kappa} \sum_{\lambda l} \begin{pmatrix} j & j & l \\ \iota & \kappa & \lambda \end{pmatrix}_\rho \begin{pmatrix} j & j & l \\ \kappa & \iota & \lambda \end{pmatrix}^*_\rho C_l$$
$$= \sum_l [(j,j,l)_s - (j,j,l)_a] C_l .$$

For $(j,j,l)_s$ values of ρ the three-j-symbol does not change when ι and κ are interchanged; for $(j,j,l)_a$ values thereof there is a change of sign. After the interchange, the $\iota\kappa\lambda$ sum is given by (10).

The last identity is the analogue of (34c). It refers to $\langle \eta(R)^2 \rangle$. We have

$$\sum_R \eta(R)^2 = \sum_R \sum_{S,T} \delta_{R,S^3}\delta_{R,T^3} = \sum_{ST} \delta_{S^3,T^3} = \sum_{S,T} \Delta_{S^3,T^3}/n_{S^3}$$
$$= \sum_{S,T} \sum_j \chi^j(S^3)\chi^j(T^3)^*/h = \sum_j h \langle \chi^j(S^3) \rangle \langle \chi^j(T^3)^* \rangle . \tag{38}$$

Hence it follows from (32b) that

$$\langle \eta(R)^2 \rangle = \sum_j [(j)_s + (j)_a - (j)_d]^2 . \tag{38a}$$

This is the last identity to be derived here. Evidently, however, further similar identities exist.[17]

Restriction of irreducible representations to subgroups. There are at least two reasons for the physicist's interest in this question. The first one is a more formal one: he likes to specify the representations completely, i.e., not only up to a similarity transformation. This makes it possible to have definite expressions

[17] For multiplicity free groups, and even some generalizations of these, similar and even further reaching results were communicated to me soon after our meeting by Professor H. J. Groenewold of the University of Groningen. They are contained in the doctoral theses of A. J. Van Lanten and E. deVries.

for the matrix elements. The specification which is most common and which has been used most generally starts with a series of subgroups $g \supset g_1 \supset g_2 \cdots$. These are so chosen that no irreducible representation of g_n contains an irreducible representation of g_{n+1} more than once, the last g being commutative. Even though there are groups for which no such sequence of subgroups exists,[18] in the groups of interest sequences of subgroups satisfying this condition could always easily be obtained. The representation of g is then assumed to have the form in which its restriction to g_1 is in the reduced form. It is then specified that these representations have the form that their restriction to g_2 be also in the reduced form, and so on. The rows and columns of the representation of the original g can then be specified by a sequence $j_1^\alpha, j_2^\beta, j_3^\gamma, \cdots$ giving the representations $\alpha, \beta, \gamma, \cdots$ of the subgroups $g_1, g_2, g_3, \cdots$ in which the row is situated. It is easy to see then that the only similarity transformation $j(R) \to Uj(R)U^{-1}$ which preserves this situation is a similarity transformation by a diagonal matrix. If the diagonal matrix U is assumed to be unitary, leaving $j(R)$ unitary, one tries to impose reality conditions to specify the form of $j(R)$ except for some signs $\pm$.

Actually, the procedure is, in most practical cases, quite simple. Thus, in the case of R_3, one chooses a subgroup R_2 for g_1—almost invariably rotations about the Z-axis. This already is commutative so that the series of subgroups contains only one member. In order to specify the representation more fully, one demands either that the representations of rotations about the X-axis or those about the Y-axis be real. This leaves a transformation with a diagonal matrix with diagonal elements ± 1 free—and the signs are indeed not quite easy to specify. In the case of U_n the series of subgroups would have $n-1$ members $U_{n-1} \supset U_{n-2} \supset \cdots \supset U_1$. The reality conditions are again rather easily specified, the signs are not.

The second, more intrinsic reason for the physicist's interest in the restriction of irreducible representations to subgroups is motivated by the phenomenon[19] of "symmetry breaking". Let us consider, for an example, a spherically symmetric system—the symmetry group is then R_3 or SU_2. If one then introduces a homogeneous electric (or magnetic) field, let us say in the Z-direction, the original symmetry is "broken": only an R_2 symmetry remains. However, if the field is not too strong, the state vectors of the original situation will be altered only slightly, except that they will belong to representations of R_2. Hence, the character of the states originating from a representation $j(S)$ of R_3 will belong to those representations $j^1(S)$ of R_2 which arise from the restriction of $j(S)$ to the R_2 which remains as a symmetry group. If the restriction of the representation $j(S)$ of R_3 contains the representation $j^1(S)$ of R_2 only once, the state vector of that representation is, in the approximation considered (that is, for not too strong electric or magnetic fields), defined except for a factor. Actually, a factor in the state vector is physically

[18] Personal communication of G. W. Mackey.

[19] The literature on this subject is too extended to be quoted here in detail. The first kind of symmetry breaking is discussed in most books on the application of group theory to spectroscopy, etc. (cf. footnotes 5 and 7). The literature on the more intricate problem of the second kind of symmetry breaking is even more extended. As examples one may quote L. Michel's and L. A. Radicati's articles, Ann. of Phys., 66 (1971), p. 758 and *Evolution of Particle Physics*, M. Conversi, ed., Academic Press, New York, 1970.

irrelevant and the practically realizable electric or magnetic fields are "weak" in the sense here considered.

A second example of symmetry breaking is the intrinsic one. It refers to the situation that the original symmetry is only approximate, such as the SU_3 symmetry for particles[20] (or the SU_4 symmetry for nuclei). There is then, usually, a subgroup—SU_2 in the case of SU_3—which is much more accurate and one wishes to know which states of SU_2 symmetry, and how many, are contained in the set of states transformed by a representation of SU_3. The solution is again provided by the restriction of the representation of SU_3 to SU_2.

As these examples, and the first reason in particular, illustrate, the symmetry relations are most effective if the restriction of the irreducible representation $j(R)$ of the full group contains each irreducible representation of the subgroup either only once, or not at all. In order to specify such a situation, and also to treat the problem more generally, I found it useful to introduce the concept of the subclass.[21]

A subclass, with respect to a subgroup, contains those elements which can be transformed into each other by an element of the subgroup. If s is an element of the subgroup, the subclass which contains the element R of the group also contains sRs^{-1}, and this is true for all elements s of the subgroup. The subclasses are defined with respect to a subgroup. The subclasses have several properties analogous to the properties of the ordinary conjugacy classes but they are, as a rule, not commutative. In fact, the condition which originated the interest in the concept of subclasses is the fact that if they do commute no irreducible representation of the full group contains, if restricted to the subgroup defining the subclass, any representation of the subgroup more than once.

Let me give, in conclusion, a few formulae for the numbers (J,j) which specify the number of times the irreducible representation j of the subgroup is contained in the irreducible representation J of the full group. The first and rather obvious one is

$$\sum_J [J](J,j) = (G/g)j, \tag{39}$$

where G and g are orders of the full group and of the subgroup considered, respectively. Next, we have

$$\sum_{J,j} (J,j)^2 = \text{number of subclasses}, \tag{40}$$

the summation over J and j to be extended over all representations of the full group and the subgroup, respectively. Finally,

$$\sum_{J,j} (J,j)^2 - (J,j)^c = \text{number of linearly independent commutators of the algebra of subclasses}, \tag{41}$$

[20] For the SU_3 model for particles, see M. Gell-Mann and Y. Ne'eman, *The Eight-Fold Way*, Benjamin, New York, 1964. The SU_4's role in nuclear physics was initiated by the present writer's article, Phys. Rev., 15 (1937), p. 106. It became dormant but was revived by P. Franzini and L. A. Radicati, Phys. Rev. Letters, 13 (1964), p. 173. See also Radicati's closing lecture in the Proc. Intern. School of Physics Ettore Majorana, Academic Press, New York. 1966, p. 275.

[21] E. P. Wigner, footnote 2(c) and Proc. Roy. Soc., A322 (1971), p. 181. The latter article contains also equations (39), (40), (41).

$(J, j)^c$ being 1 if the representation j of the subgroup is contained in J's restriction, 0 otherwise. Evidently, the sum on the left side of (41) can be 0 only if all (J, j) are either 1 or 0, i.e., if the subclasses commute. As was mentioned before, these formulae, valid literally only for finite groups, can be modified to apply also to compact Lie groups, but the modification will not be discussed here.

Bibliography of Mathematical Papers

1929

(With J. von Neumann) Über merkwürdige diskrete Eigenwerte. Physikalische Zeitschrift *30*, 465–467 (1929)

(With J. von Neumann) Über das Verhalten von Eigenwerten bei adiabatischen Prozessen. Physikalische Zeitschrift *30*, 467–470 (1929)

1934

(With P. Jordan and J. von Neumann) On an Algebraic Generalization of the Quantum Mechanical Formalism. Annals of Mathematics *35*, 29–64 (1934)

1939

On Unitary Representations of the Inhomogeneous Lorentz Group. Annals of Mathematics *40*, 149–204 (1939)

Translated into Hungarian: Az inhomogén Lorentz-csoport unitér ábrázolásairól. Magyar Fizikai Folyóirat *21*, 21–72 (1973)

1940

(With J. von Neumann) Minimally Almost Periodic Groups. Annals of Mathematics *41*, 746–750 (1940)

1941

On Representations of Certain Finite Groups. American Journal of Mathematics *63*, 57–63 (1941)

Translated into Hungarian: Bizonyos véges csoportok ábrázolásairól. Magyar Fizikai Folyóirat *19*, 159–165 (1971)

1949

(With T. D. Newton) Localized States for Elementary Systems. Reviews of Modern Physics *21*, 400–406 (1949)

Translated into Hungarian: Elemi rendszerek lokalizált állapotai. Magyar Fizikai Folyóirat *22*, 87–98 (1974)

1951

On a Class of Analytic Functions from the Quantum Theory of Collisions. Annals of Mathematics *53*, 36–67 (1951)

1952

On the Connection Between the Distribution of Poles and Residues for an R Function and Its Invariant Derivative. Annals of Mathematics *55*, 7–18 (1952)

Derivative Matrix and Scattering Matrix. Revista Mexicana de Fisica *1*, 91–101 (1952)

(With E. Inönü) Representations of the Galilei Group. Il Nuovo Cimento *9*, 705–718 (1952)
Translated into Hungarian: A Galilei csoport ábrázolásai. Magyar Fizikai Folyóirat *22*, 183–193 (1974)

Simplified Derivation of the Properties of the Elementary Transcendentals. American Mathematical Monthly *59*, 669–683 (1952)

1953

(With E. Inönü) On the Contraction of Groups and Their Representations. Proceedings of the National Academy of Sciences (USA) *39*, 510–524 (1953)
Translated into Hungarian: Csoportok és ábrázolásaik kontrakciójáról. Magyar Fizikai Folyóirat *22*, 265–278 (1974)

1954

(With E. Inönü) On a Particular Type of Convergence to a Singular Matrix. Proceedings of the National Academy of Sciences (USA) *40*, 119–121 (1954)

Application of the Rayleigh-Schrödinger Perturbation Theory to the Hydrogen Atom. The Physical Review *94*, 77–78 (1954)

(With J. von Neumann) Significance of Loewner's Theorem in the Quantum Theory of Collisions. Annals of Mathematics *59*, 418–433 (1954)

1955

Characteristic Vectors of Bordered Matrices with Infinite Dimensions I. Annals of Mathematics *62*, 548–564 (1955)

1957

Characteristic Vectors of Bordered Matrices with Infinite Dimensions II. Annals of Mathematics *65*, 203–207 (1957)

(With Edwin Hewitt) On a Theorem of Magnus. Proceedings of the American Mathematical Society *8*, 740–744 (1957)

1960

Normal Form of Antiunitary Operators. Journal of Mathematical Physics *1*, 409–413 (1960)

Phenomenological Distinction Between Unitary and Antiunitary Symmetry Operators. Journal of Mathematical Physics *1*, 414–416 (1960)

1963

On Weakly Positive Matrices. Canadian Journal of Mathematics *15*, 313–317 (1963)

1964

Unitary Representations of the Inhomogeneous Lorentz Group Including Reflections. In: Group Theoretical Concepts in Elementary Particle Physics (Feza Gursey, ed.). Gordon and Breach, New York 1964, pp. 37–80

1965

On the Matrices Which Reduce the Kronecker Products of Representations of S.R. Groups. In: Quantum Theory of Angular Momentum (L. C. Biedenharn and H. van Dam, eds.). Academic Press, New York 1965, pp. 87–133

Translated into Hungarian: A Mátrixokról, amelyak egyszerüen reducibilis csoportok ábrázolásainak Kronecker-szorzatát kiredukálják. Magyar Fizikai Folyóirat *19*, 165–207 (1971)

1967

(With A. W. Sáenz) Unimodular Matrices Homomorphic to Lorentz Transformations in $n \geq 2$ Spacelike Dimensions. Zeitschrift für Naturforschung *22a*, 1293–1299 (1967)

1968

On a Generalization of Euler's Angles. In: Group Theory and its Applications. Academic Press, New York 1968, pp. 119–129

Condition That the Irreducible Representations of a Group, Considered as Representations of a Subgroup, Do Not Contain Any Representation of the Subgroup More Than Once. In: Spectroscopic and Group Theoretic Methods in Physics (F. Bloch et al., eds.). North-Holland, Amsterdam 1968, pp. 131–136

1970

(With H. C. Schweinler) Orthogonalization Methods. Journal of Mathematical Physics *11*, 1693–1694 (1970)

1971

Restriction of Irreducible Representations of Groups to a Subgroup. Proceedings of the Royal Society of London A *322*, 181–189 (1971)

1972

(With F. E. Goldrich) Condition That All Irreducible Representations of a Compact Lie Group, If Restricted to a Subgroup, Contain No Representation More Than Once. Canadian Journal of Mathematics *24*, 432–438 (1972)

1973

Reduction of Direct Products and Restriction of Representations to Subgroups: The Everyday Tasks of the Quantum Theorists. SIAM Journal on Applied Mathematics *25*, 169–185 (1973)

Related but not reprinted in the Collected Works

Special Functions – A Group Theoretic Approach by James Talman based on 1955 Lectures by E. P. Wigner. Benjamin, New York 1968

Springer-Verlag and the Environment

We at Springer-Verlag firmly believe that an international science publisher has a special obligation to the environment, and our corporate policies consistently reflect this conviction.

We also expect our business partners – paper mills, printers, packaging manufacturers, etc. – to commit themselves to using environmentally friendly materials and production processes.

The paper in this book is made from low- or no-chlorine pulp and is acid free, in conformance with international standards for paper permanency.