AD No. 20713
ASTIA FILE COPY

THE GROUP THEORY METHOD IN QUANTUM MECHANICS

by

Dr. B. L. van der Waerden

"The translation reported in this document has been made possible through support and sponsorship extended by the Geophysical Research Directorate of the Air Force Cambridge Research Laboratories, A.F., under Contract No. AF 19(122)-450. It is published for technical information only and does not represent recommendations or conclusions of the sponsoring agency."

Copyright 1932, by Julius Springer, Berlin, vested in the Attorney General of the United States pursuant to law. Published under License No. A-1621 of the Attorney General.

Special Report

Translation by Jack Alster

Department of Chemistry
Boston University
Boston, Mass.

Contract No. AF 19(122)-450

Geophysical Research Directorate

Air Force Cambridge Research Laboratories
Cambridge, Massachusetts

PERSONNEL

Dr. Howard Reiss, Responsible Investigator

Mr. Herbert Glazer, Research Assistant

Mr. Jack Alster, Technical Assistant

We wish to express our sincere appreciation to Drs. Arthur Kohlenberg and Walter Hauser of the Boston University Physics Department for their many helpful suggestions.

TABLE OF CONTENTS

page

I. QUANTUM - MECHANICAL INTRODUCTION

1. The Schroedinger Differential Equation 1

2. Linear Operators. Orthogonal Systems . 5

3. The Wave Equation of an Atom and that of a Molecule 16

4. The Electron in a Spherically-Symmetrical Field 21

5. The Perturbation - Calculation 29

6. Moment of Momentum and Infinitesimal Rotations 34

II. GROUPS AND THEIR REPRESENTATIONS

7. Linear Transformations 42

8. Groups 50

9. Equivalence and Reducibility of Representations 60

10. Representation of Abelian Groups. Examples 68

11. The Uniqueness Theorems 75

12. The Kronecker Product Transformation 78

13. Commutation of Matrices with Representations 83

14. Representations of a Finite Group 89

15. The Characters 100

III. THE ROTATION GROUP AND THE LORENTZ GROUP

16. The Linear Group c_2 , the Unitary Group u_2 and its Relation to the Rotation Group $\mathfrak{G}_3$ 105

17. The Infinitesimal Transformations and the Representations of the Rotation Group 112

18. Examples and Applications 123

1) Reduction of the Product Representation $D_J \times D_{J'}$, of the Rotation Group 123

2) Applications of the Relation (18.1) 127

Table of Contents (Cont'd.)

3) The Reflection Character 131

19. Selection and Intensity Rules 132

20. The Representations of the Lorentz Group 139

1) The group C_2 and the Restricted Lorentz transformations 139

2) The Reflection S and the complete Lorentz Group 142

3) Spinor - Analysis 145

4) The Infinitesimal Transformations 148

IV. THE SPINNING ELECTRON

21. The Spin 154

22. The Wave Function of the Spinning Electron 157

23. The Lorentz - Invariant Wave Equation of Dirac 167

24. The Electron in the Central Field according to Dirac 174

25. The Multiple Electron Problem 178

1) Multiplet Structure. Zeeman Effect 178

2) The Anomalous Zeeman Effect 184

V. THE PERMUTATION GROUP AND THE PAULI EXCLUSION PRINCIPLE

26. Resonance of Equivalent Entities 189

27. The Pauli Exclusion Principle and the Periodic System of the Elements 196

28. The Eigenfunctions of the Atom according to the Pauli Exclusion Principle 204

29. The Approximate Calculation of the Energy 215

30. The Pure Spin Functions and their Transformation under Rotations and Permutations 225

VI. MOLECULAR SPECTRA

31. The Quantum Numbers of the Molecule 234

Table of Contents (Cont'd.)

32.	The Rotational Levels	242
33.	Inclusion of the Spin	250
34.	Homonuclear, Diatomic Molecules	255
35.	The Formation of the Molecule from Two Atoms	257
36.	Remarks about the Estimation of the Energy	265

PREFACE.

The quantum-mechanical treatment of atoms and molecules based on the Schroedinger wave-equation encounters difficulties which are due to the complexity of the problem. Nevertheless we can make general predictions about the eigenfunctions and eigenvalues which find their confirmation in observed spectroscopic regularities. We can make these predictions because of the symmetry properties of the wave equation, namely: the invariance towards rotations, reflections and permutations among the electrons (or nuclei). Group-Theory serves as a mathematical aid in the formulation of these rules. The "Darstellungs Theory" of finite and continuous groups is of particular help.

The aim of this small volume is to explain in as simple a manner as possible these mathematical concepts and their physical application. I have been careful to get along always with the simplest tools and, in the mathematical developments, I have tried not to go beyond the physically significant details. I have given special consideration to the most recent developments due to the work of Dirac, Slater and others, - thus, pushing to the background the complicated "Darstellungs-Theory" and the character-calculations of the symmetrical permutation group. Whoever wishes to delve deeper into the "Darstellung-Theory" of the symmetrical groups and into its connection with linear groups, should read the book by H. Weyl, "Gruppentheorie und Quantenmechanik", second edition, Leipzig 1931, - also, the original treatises by G.

Frobenius, I. Schur, and H. Weyl.

The core of this book, which also requires the greatest amount of attention on the part of the reader, consists of the Darstellungs-Theory of the rotational group in the third chapter - as well as of the spin theory in the fourth chapter, based on the former.

Although there appeared last year a book similar to mine by E. Wigner entitled "Gruppentheorie und ihre Anwendung auf die Quantenmechanik der Atome", Berlin 1931, I, nevertheless, feel the publication of this book to be justified on the basis of my treatment of the last chapter on molecules, the paragraphs on the Lorentzgroup and the relativistic wave equation - not to mention the separate treatment of many details.

The basic concepts of wave mechanics and spectroscopy are assumed as common knowledge. Group Theory and the theory of the spinning electron, however, are developed from the beginning.

Leipzig, January 1932.

B.L.v.d. Waerden

I. QUANTUM-MECHANICAL INTRODUCTION.

#1. The Schroedinger Differential Equation.

According to wave mechanics, all questions regarding the behavior of an electron, atom, or a system of electrons and atomic nuclei may be referred back for their solution to a study of the Schroedinger Differential Equation, which in its unrelativistic form is stated as follows:

$$H\Psi + \frac{h}{I}\frac{\partial \Psi}{\partial t} = 0 \qquad (1.1)$$

Here, Ψ, the wave function, is a time dependent complex function of the rectangular coordinates $q_0, q_1 \ldots q_f$ (or to be more exact: $x_0, y_0, z_0 \ldots x_f, y_f, z_f$) of the f+1 mass-points (electrons and nuclei) of the system. H is the energy-operator which arises out of the classical expression of the energy (the Hamiltonian Function), i.e.

$$T + U = \sum_{\lambda=0}^{f} \frac{1}{2\mu_\lambda}\left(p_{x\lambda}^2 + p_{y\lambda}^2 + p_{z\lambda}\right) + U(q) \qquad (1.2)$$

upon substituion of the momentum components P_x, P_y, P_z by $\frac{h}{i}\frac{\partial}{\partial x}, \frac{h}{I}\frac{\partial}{\partial y}, \frac{h}{I}\frac{\partial}{\partial z}$. That is:

$$H = \sum_{\lambda} - \frac{h^2}{2\mu_\lambda} \frac{\partial^2}{\partial x_\lambda^2} + \frac{\partial^2}{\partial y_\lambda^2} + \frac{\partial^2}{\partial z_\lambda^2} + U(q) = \sum_{\lambda} - \frac{h^2}{2\mu_\lambda} \Delta_\lambda + U(q) ,$$

where

μ_λ = mass of an electron of nucleus

$2\pi h$ = Planck's quantum of action

$U(q)$ = potential energy as a function of the co-ordinates q.*

The assumption is made that Ψ determines the state of the system at a definite time and that the probability for the system to be at time t in some region B of the q-space (configuration space) is proportional to the normalization

$$\int_{P_i} \bar{\Psi}\Psi dq.$$

If λ is a constant, then Ψ and $\lambda\Psi$ represent the same state.

The most important solutions of the differential equation (1.1) are the stationary waves or "Eigenschwingungen" (Eigen-oscillations):

$$\Psi = \psi(q)e^{i\omega t}$$

---

*Translator's Note:

The reader please note that the author's h is equivalent to the nowadays more commonly used $\hbar = \frac{h}{2\pi}$.

where ψ represents a time-independent function which plainly must satisfy the differential equation

$$H\psi = E\psi \quad (\text{where } E = h\omega) . \tag{1.3}$$

Equation (1.3) has the form of a linear eigenvalue problem, in which two unknowns appear: the eigenfunction ψ and the eigenvalue E. The eigenvalues E of the energy operator constitute the possible energy levels of the system; in spectroscopy they are also called "terms", since by virtue of the formula

$$E_1 - E_2 = h\nu$$

one can calculate the frequency[1] ν of the light emitted due to the quantum jump $E_1 \rightarrow E_2$, or of the light absorbed due to the quantum jump $E_2 \rightarrow E_1$[2].

---

1) By frequency we mean here the number of oscillations in 2π seconds.

(Translator's Note: Hence, $2\pi\nu$ is the number of vibrations per second. The reader will observe that this latter unit of frequency is more commonly employed nowadays. Now, since the author's h is equivalent to our $\frac{h}{2\pi}$, the equation above $E_1 - E_2 = h\nu$ ($\nu \equiv$ oscillations per 2π seconds) in terms of the more common units, therefore, becomes:

$$E_1 - E_2 = \frac{h}{2\pi} \cdot 2\pi\nu$$

($2\pi\nu \equiv$ oscillations per second).

2) The "terms" are usually measured in terms of the wave-numbers; i.e. the reciprocal wave-lengths. The term corresponding to the energy E is, thus: $\frac{1}{\lambda} = \frac{\omega}{2\pi c} = \frac{E}{2\pi hc}$.
The atomic energy is often measured, also, in terms of the volt, where $E = Ve$, e being the charge of the electron.

Only those eigenfunctions have physical significance which remain finite in the infinite region of the q-space. If one assumes, the potential energy U to go to zero at infinity, then two types of eigenfunctions become evident: firstly, those with $E > 0$, which can be represented as superpositions of smooth waves in the region $U=0$; these waves extend to infinity and their energy values form a continuous spectrum covering the entire positive E-axis. Secondly, there are the eigenfunctions with $E < 0$ which differ considerably from zero only in the "Potential-senken" (Potential wells) - to be more exact, in the region $U < E$; in the region $U > E$, however, they diminish rapidly at infinity (in an exponential manner); their eigenvalues form a discreet termspectrum, which can be arranged according to increasing energy-values: E_1, E_2 ..... . These matters are best clarified by simple examples having one degree of freedom or spherical symmetry - in which case, calculations can actually be carried out.

For the mathematical treatment of the problem, especially in the case of the formulation of the perturbation calculations, it is useful to modify somewhat the problem in the sense that the atom or molecule is considered to be contained in a reflecting (hollow) sphere having a very large radius R. The eigenfunctions ψ must then vanish at the walls of the sphere. By making this limitation the entire spectrum becomes discreet. In the region $E > 0$, the eigenvalues lie very close together[1] and at the limit

---

1) cf. Courant-Hilbert: Methoden der mathematischen Physik Vol.I 2nd Edition. (Vol.XII of this collection) 1931 Chapter VI.

for $R \to \infty$ they given rise to the continuous spectrum, whereas in the region $E < 0$, they lie much farther apart —— , so that for $R \to \infty$ they become the eigenvalues of the discreet term-spectrum.

The restrictions placed upon the configuration space just outlined, lead to the advantageous situation in which the eigen-functions are now normalizable and, in most cases, are also capable of forming a complete orthogonal system (cf §2). In the following sections we shall consider these restrictions to be always made where it is convenient to a mathematical treatment.

#2 Linear Operators, Orthogonal Systems.

By "functions" we shall always imply in this paragraph complex functions of the coordinates of a system of mass-points.

By the scalar product (φ, ψ) of two functions φ, ψ we shall mean the integral

$$(\varphi, \psi) = \int \bar{\varphi}\, \psi \, dV$$

extending over the entire phase space. Obviously (φ, ψ) is the complex-conjugate of (ψ, φ) and we have for constant α:

$$(\varphi, \alpha\psi) = \alpha \cdot (\varphi, \psi)$$
$$(\alpha\varphi, \psi) = \bar{\alpha} \cdot (\varphi, \psi)$$
$$(\varphi, \psi_1 + \psi_2) = (\varphi, \psi_1) + (\varphi, \psi_2)$$
$$(\varphi_1 + \varphi_2, \psi) = (\varphi_1 \psi) + (\varphi_2 \psi).$$

A special case of the scalar product is the "Quadratintegral" or the normalization:

$$N\psi = (\psi,\psi) = \int \bar{\psi}\psi dV = \int |\psi|^2 dV.$$

The Schwartz inequality is valid:

$$|(\varphi,\psi)|^2 \leqq N\varphi \cdot N\psi \ .$$

A function ψ is <u>normalized</u> if its normalization is equal to 1. Two functions are <u>orthogonal</u> if their scalar product is equal to zero, i.e. $(\varphi,\psi) = 0$.

The energy operator H of the previous chapter has the following properties: it is linear, i.e.

$$H(\varphi + \psi) = H(\varphi) + H(\psi)$$

$$H(\alpha\varphi) = \alpha H(\varphi)$$

it is, also, <u>symmetrical</u> or <u>self-adjunct</u>. In other words, for all functions φ,ψ vanishing at the boundary of a region (or fairly rapidly so, at infinity), the following relation obtains:

$$(\varphi, H\psi) = (H\varphi, \psi) \ . \qquad (2.1)$$

This can be easily shown by the method of partial integration.

In Quantum Mechanics we assign a linear operator not only to the energy but also to any other measurable quantity; e.g. to

the momentum-components $p_x = m\dot{x}$ (etc.) we assign the operators $\frac{h}{i}\frac{\partial}{\partial x}$ (etc.) and to the angular momentum components $yp_z - zp_y$ (etc.) we assign the operators

$$hL_x = \frac{h}{I}\left(y\frac{\partial}{\partial z} - z\frac{\partial}{\partial y}\right) \text{ etc.}$$

These operators are also self adjunct. If the wave function ψ is an eigenfunction of the operator Ω - that is, if it satisfies the boundary conditions, and if, also, the following relation holds:

$$\Omega\psi = \lambda\psi ,$$

then it may be stated that the physical quantity Ω has the discreet value λ in the state ψ. For every state ψ (also, for those ψ's which are not eigenfunctions of Ω), one may define the average value $\hat{\Omega}$ of the physical quantity Ω as follows:

$$\hat{\Omega} = (\psi, \Omega\psi) = \int\bar{\psi}\cdot\Omega\psi dV$$

where ψ is assumed to be normalized. The self-adjunct property of the operator Ω assures the reality of all the average values and of all eigenvalues, in particular.

Two eigenfunctions of a self-adjunct operator Ω, belonging to different eigenvalues, are always orthogonal.

Proof: From the equations

$$\Omega\psi_1 = \lambda_1\psi_1 \cdot \Omega\psi_2 = \lambda_2\psi_2, \text{ and } (\Omega\psi_1, \psi_2) = (\psi_1, \Omega\psi_2)$$

we get

$$(\lambda_1\Psi_1,\Psi_2) = (\Psi_1,\lambda_2\Psi_2)$$
$$(\lambda_1-\lambda_2)(\Psi_1\Psi_2) = 0$$
$$(\Psi_1\Psi_2) = 0.$$

Of special importance are those operators which can commute with the energy-operator. The following "Erhaltungs-satz" (conservation law) applies to them: (The conservation laws for energy, momentum, and angular momentum are but special cases here).

<u>If an operator Ω commutes with the energy operator H, then the eigenvalues as well as the average values of Ω remain constant throughout the time interval, in which the state Ψ varies in accordance with (1.1).</u>

Proof: a) Constancy of the eigenvalues:

At time t=0, $\Omega\Psi = \lambda\Psi$.

We consider the time derivative of the function

$$F = (\Omega - \lambda)\Psi.$$

That is:

$$\frac{h}{I}\frac{\partial F}{\partial t} = \frac{h}{i}(\Omega-\lambda)\frac{\partial\Psi}{\partial t} = -(\Omega-\lambda)H\Psi = -H(\Omega-\lambda)\Psi = -HF$$

By means of this differential equation and the initial value F=0 for t=0, the function F is fully determined. Hence, F=0 for all t - i.e. the function Ψ, at all times, remains the eigenfunction of Ω belonging to the eigenvalue λ.

b) Constancy of the average values:

$$\frac{h}{i}\frac{d\Omega}{dt} = \frac{h}{i}\frac{d}{dt}(\Psi, \Omega\Psi) = \frac{h}{i}\left(\frac{\partial\Psi}{\partial t}, \Omega\Psi\right) + \frac{h}{i}\left(\Psi, \Omega\frac{\partial\Psi}{\partial t}\right)$$

$$= \left(-\frac{h}{i}\frac{\partial\Psi}{\partial t}, \Omega\Psi\right) + \left(\Psi, \Omega\frac{h}{i}\frac{\partial\Psi}{\partial t}\right)$$

$$= (H\Psi, \Omega\Psi) - (\Psi, \Omega H\Psi)$$

$$\frac{h}{i}\frac{d}{dt}\hat{\Omega} = (\Psi, H\Omega\Psi) - (\Psi, H\Omega\Psi) = 0$$

Another important property of the operators Ω (which commute with H) is that they always transform the eigenfunctions ψ belonging to a certain energy-level into similar ones; from the equation $H\psi = E\psi$ we get $H\Omega\psi = E\Omega\psi$, as can be seen immediately.

Upon limiting ourselves to finite space, we find that for the quantum-mechanical energy-eigenvalue problem

$$H\psi = E\psi \qquad (2.2)$$

the following rules are valid: their exact proof has, according to my knowledge, not yet been demonstrated in all cases.

I. The eigenvalues constitute an infinite series increasing without limit: E_1, E_2, E_3.....

II. For every eigenvalue there exists only a finite number

of linearly independent eigenfunctions, from which all the other eigenfunctions (having complex coefficients) can be formed by linear combinations of these. If their number is $K>1$, then we speak of a K-fold degeneracy. One is then at liberty to always choose these K eigenfunctions orthogonal to one another. When this is done for all energy values and when all the resulting functions are arranged according to increasing eigenvalues, then we obtain a system which consists of an infinite amount of orthogonal eigenfunctions $\varphi_1, \varphi_2, \varphi_3 \ldots\ldots$.

III By continuously varying one of the parameters occurring in the operator H (e.g. the mass, or the field-strength of an external field), it is found that the eigenvalues depend on these parameters in a continuous and differentiable manner.

IV The eigenfunctions $\varphi_1, \varphi_2, \ldots$ constitute a <u>complete set of functions</u>. This means that the average value of every continuous function ψ can be approximated as accurately as desired by a properly chosen sum $\sum_1^n c_\nu \varphi_\nu$; i.e. it is possible to choose the c_ν and n in such a way that for every δ the mean square error ("mittlere Fehlerquadrat")

$$N(\psi - \sum_1^n c_\nu \varphi_\nu) \tag{2.3}$$

is always smaller than δ.

We consider the complete orthogonal system, $\varphi_1, \varphi_2, \varphi_3 \ldots$ to be normalized:

$$N(\varphi_\nu) = (\varphi_\nu, \varphi_\nu) = 1 \quad .$$

Now, in order that in the approximation of a function ψ by a sum $\sum_1^n c_\nu \varphi_\nu$ the mean square error (2.3) becomes as small as possible for a given n, one has to choose for the c_ν the "expansion coefficients":

$$c_\nu = (\varphi_\nu, \psi) \tag{2.4}$$

As a result

$$N(\psi - \sum_1^n c_\nu \varphi_\nu) = N(\psi) - \sum_1^n \bar{c}_\nu c_\nu \ .$$

The "Bessel inequality" immediately follows from this:

$$\sum_1^n \bar{c}_\nu c_\nu \leqq N(\psi)$$

The "completeness relation" also follows as a criterion for the completeness of the system of functions

$$N(\psi) = \sum_1^\infty \bar{c}_\nu c_\nu \tag{2.5}$$

which must be satisfied for all continuous functions ψ.

The expansion coefficients c_ν (2.4) completely determine the function ψ, for when ψ_1 and ψ_2 have the same expansior coefficients, then the expansion coefficients of their difference $\psi_1 - \psi_2$ are equal to zero, from which (according to 2.5) we get.

$$N(\psi_1 - \psi_2) = 0 \quad \text{and, of course,} \quad \psi_1 = \psi_2 \ .$$

In particular, a continuous functio n ψ, whose expansion

coefficients are all zero, (i.e. one which is orthogonal with respect to all the φ_ν) is identically equal to zero.

The series $\sum_1^\infty c_\nu \varphi_\nu$ is called the expansion of ψ in terms of the complete orthogonal system of the φ_ν . It need not necessarily converge, although most of the times it does so, whenever ψ is sufficiently often differentiable. Since, clearly, the function is determined by the series, one may write symbolically (even in the case of divergence):

$$\psi \sim \sum_1^\infty c_\nu \varphi_\nu \quad . \tag{2.6}$$

For purposes of forming the scalar products and the normalizations, one may replace the function ψ (according to the accuracy desired) by $\sum c_\nu \varphi_\nu$, since for every function φ , we have according to the Schwartz inequality:

$$\left|(\varphi,\psi) - (\varphi, \sum_1^n c_\nu \varphi_\nu)\right| = \left|(\varphi, \psi - \sum c_\nu \varphi_\nu)\right|$$

$$\leqq \sqrt{N(\varphi) \cdot N(\psi - \sum c_\nu \varphi_\nu)}$$

$$\leqq \sqrt{\varepsilon \cdot N(\varphi)}$$

It follows then, that

$$(\varphi,\psi) = \sum_1^\infty c_\nu (\varphi, \varphi_\nu) = \sum \overline{(\varphi_\nu, \varphi)} c_\nu = \sum \bar{b}_\nu c_\nu \quad , \tag{2.7}$$

where b_ν are the expansion coefficients of the function φ .

With the aid of the orthogonal set of the φ_ν , one may assign to every linear operator Ω an infinite matrix; this is done by expanding $\Omega \varphi_\nu$ in terms of the φ_μ .

$$\Omega \varphi_\nu \sim \sum \omega_{\mu\nu} \varphi_\mu \, .$$

The matrix elements $\omega_{\mu\nu}$ are defined as

$$\omega_{\mu\nu} = (\varphi_\mu, \Omega \varphi_\nu) \, .$$

If Ω is self-adjoint, then the matrix $(\omega_{\mu\nu})$ is "Hermetically symmetrical":

$$\bar{\omega}_{\mu\nu} = \omega_{\nu\mu} \, .$$

We now wish to calculate the expansion-coefficients of Ω, when those of

$$\psi \sim \sum c_\nu \varphi_\nu \tag{2.8}$$

are given.

If we set $\Omega \psi \sim \sum d_\nu \varphi_\nu$, then, according to the above-mentioned statement (2.7) for the case where Ω is self-adjoint, we have:

$$d_\mu = (\varphi_\mu, \Omega\psi) = (\Omega \varphi_\mu, \psi) = \sum_\nu \bar{\omega}_{\nu\mu} c_\nu$$
$$= \sum_\nu \omega_{\mu\nu} c_\nu \, .$$

This means. <u>The expansion coefficients of $\Omega\psi$ are the same as those obtained by the following procedure: One first makes a termwise application of the operator Ω on the left- and right-hand side of the series (2.8); one then expands everything in terms of φ and finally rearranges the resulting terms.</u>

If the operator Ω commutes with the energy-operator H, the eigenfunctions of which are the φ_λ, then all matrix-elements

$\omega_{\mu\nu}$ (whose indices μ, ν refer to the eigenfunctions φ_μ, φ_ν with different eigenvalues $E_\mu \neq E_\nu$) must be equal to zero. The φ_μ belonging to an energyvalue E are linearly transformed among one another by Ω, and can be so chosen (as we shall yet discuss in greater detail in §7) that they immediately become eigenfunctions of Ω. The commutative Ω and H have, therefore, a mutual, complete set of eigenfunctions.

The completeness requirement of the set of eigenfunctions is of great significance for the practical solution of eigenvalue problems. For example, we shall see in §5 that the perturbation-calculation is based essentially on this completeness requirement. A further application is exemplified by the method of the separation of variables, which in many cases, simplifies to a large extent the solution of eigenvalue problems. The method is based on the following: The variables $q_1, \ldots, q_s$ which occur in the functions ψ, are arbitrarily broken up into two subdivisions, $\{q_1 \ldots q_\eta\}$ and $\{q_{\eta+1} \ldots q_s\}$; the operator H, too, consists of two parts. $H = H_1 + H_2$, the first part referring only to the $q_1 \ldots q_\eta$, and the second part only to the $q_{\eta+1} \ldots q_s$. The eigenfunctions of H may then be formed as products $\varphi(q_1 \ldots q_\eta) \cdot \psi(q_{\eta+1} \ldots q_s)$, where φ is an eigenfunction of H_1 and ψ an eigenfunction of H_2. That the eigenfunctions of H are obtained in this manner is obvious, since

$$H\varphi\psi = (H_1+H_2)\varphi\psi = (H_1\varphi)\cdot\psi + \varphi\cdot(H_2\psi)$$

$$= E_1\varphi\psi + E_2\varphi\psi = (E_1+E_2)\varphi\psi .$$

The eigenvalue E is equal to E_1+E_2. The question, however, is

whether one can obtain in this manner <u>all</u> the eigenfunctions of the problem. The question can be answered in the affirmative as soon as one knows that at least φ (or the ψ) form a complete orthogonal system. For instance, χ can be expanded as a function of $q_1 \ldots q_n$ in terms of the eigenfunctions $\varphi_1, \varphi_2 \ldots$

$$\chi \sim \sum_{1 \ldots} \varphi_\nu(q_1, \ldots, q_n)\, c_\nu(q_{n+1}, \ldots, q_s) \qquad (2.9)$$

if χ is any arbitrary eigenfunction of the operator H.

According to the rule proven above, one can calculate $H_1\chi$ and $H_2\chi$ by a formal term-by-term operation on χ.

$$H\chi = H_1\chi + H_2\chi \sim \sum_1^\infty E_\nu \varphi_\nu c_\nu + \sum_1^\infty \varphi_\nu H_2 c_\nu .$$

On the other hand, χ is an eigenfunction of H:

$$H\chi = E\chi \sim \sum E \varphi_\nu c_\nu .$$

In both expansions of $H\chi$, all coefficients must agree with one another:

$$E_\nu c_\nu + H_2 c_\nu = E c_\nu ,$$

$$H_2 c_\nu = (E - E_\nu) c_\nu .$$

The c_ν, thus, are eigenfunctions of H_2 belonging to the eigenvalues $E'_\nu = E - E_\nu$. With increasing ν, E_ν increases beyond limit; hence, E'_ν will at one time be smaller than the smallest eigenvalue of H_2; — in that case, however, $c_\nu = 0$. Hence, we have the sum (2.9) only a finite number of terms, of which every single one is an eigenfunction of H, namely, a product $\varphi_\nu \psi_\mu$.

In reality, therefore, these products form a basis for all eigenfunctions of the operator H. This method, among others, can always be applied whenever the observed system consists of two sub-systems, whose interaction energy is zero, or else is negligible. The eigenfunctions, then, are the products of those eigenfunctions of the component parts and the energyvalues are the sums of the energies of the parts.

#3. <u>The Wave Equation of an Atom and that of a Molecule.</u>

We refer again to the Schroedinger equation of a system of f+1 mass-points with masses $\mu_0, \mu_1 \ldots \mu_f$.

$$(-\sum_{0}^{f} \frac{h^2}{2\mu_\lambda} \Delta_\lambda + U)\psi = E\psi \,. \qquad (3.1)$$

Specifically speaking, we are contemplating on an atom with f electrons or a diatomic molecule with f-1 electrons. In order to simplify the differential equation (3.1), we introduce in place of the co-ordinates $q_0 \ldots q_f$, the co-ordinates q_s of the center of mass and the coordinates q'_ν of the masses $\mu_1 \ldots \mu_n$ relative to the center of mass:

$$\left.\begin{aligned} Mq_s &= \mu_0 q_0 + \mu_1 q_1 + \ldots + \mu_f q_f; \quad M = \mu_0 + \mu_1 + \ldots \mu_f; \\ q'_\nu &= q_\nu - q_s \quad (\nu = 1,2,\ldots f) \end{aligned}\right\}$$

The equation then becomes:

$$\left\{ -\frac{h^2}{2M}\Delta_s - \sum_{1}^{f} \frac{h^2}{2\mu_\lambda}\Delta'_\lambda + \frac{h^2}{2M}\left(\sum_{1}^{f} \frac{\partial}{\partial q'_\lambda}\right)^2 + U \right\}\psi = E\psi \,, \qquad (3.2)$$

Here, $\frac{h}{I}\sum\frac{\partial}{\partial q_\lambda'}$ is the vector operator for the sum of the momenta of the masses $\mu_1,\ldots,\mu_f$ relative to the center of mass, or what amounts to the same,- the negative momentum of the mass-point μ_0. The square is to be considered a scalar vector-square.

In the absence of external forces, the potential energy U depends only on the relative co-ordinates q'. In other cases, U is mostly a sum of terms which depend only on the relative co-ordinates, and of terms, again, which depend only on q_s. In all these cases we can separate the variables and form the relation:

$$\psi = \psi_1(q_s)\cdot\psi_2(q_1',\ldots,q_f').$$

For the wave-function ψ_1 of the center of mass, we obtain a common Schroedinger equation with mass M. For ψ_2 one obtains, upon substituting again q for q' and E for E_2, the differential equation:

$$\left\{-\sum_1^f\frac{h^2}{2\mu_s}\Delta_s + \frac{h^2}{2M}\left(\sum_1^f\frac{\partial}{\partial q}\right)^2 + U\right\}\psi = E\psi\,, \qquad (3.3)$$

In the simplest case where we have one nucleus and an electron with masses μ_0 and μ_1 and with $M = \mu_0 + \mu_1$, we obtain;

$$-\frac{h^2}{2\mu_1}\frac{\mu_0}{\mu_0+\mu_1}\Delta + U)\psi = E\psi\,.$$

This is the same equation as that for an electron in the force-field of a <u>fixed</u> nucleus, except that it contains the correction

factor $\frac{\mu_o + \mu_1}{\mu_o}$ for the mass, thus making allowance also for the motion of the nucleus. This factor is very nearly equal to unity (it is known that for the H-atom, the value of $\mu_1 : \mu_0$ is 1/1850, and for the He-atom it is 1:4·1840); it may, therefore, be ignored where we are not interested in the exact position of the terms. This accounts, then, for the cancellation of the term containing the $\frac{h^2}{2M}$ in equation (3.3). This approximation is all the more permitted in the case of atoms, with more electrons, where the calculation of the position of the terms is, anyway, only an approximate one (due to the complexity of the differential equation), and where, moreover, the ratio $\mu : M$ is even smaller than in hydrogen. This consideration, therefore, gives rise to the theoretical justification for the customary procedure, of applying the Schroedinger equation to the atoms in such a manner that the nucleus can be considered as a fixed force-center.

The matter becomes more complicated, however, in the case of a di-atomic molecule. Here, μ_o and μ_1 are the masses of the nuclei, $\mu_2 = \ldots = \mu_f = \mu$ those of the electrons. In the differential equation (3.3), one might again wish to cancel out the small terms. In this case, however, one can not assume at the outset, that the terms with $\frac{1}{M}$ or $\frac{1}{\mu_1}$ are small compared to those with $\frac{1}{\mu}$, for, in the terms containing $\frac{1}{M}$, there occur also the factors $h^2 \frac{\partial^2}{\partial q^2}$, which represent the squares of the momentum components of the nuclei; these may be large compared to the squares of the momentum components of the electrons.

In order to separate completely the smaller terms from the larger ones, we transform the equation (3.3) by introducing a

"fictitious nucleus" with co-ordinates $q_* = q_1 - q_0$ which is arrived at by subtracting the vector combining the nuclei $q_1 - q_0$ from the center of mass. The position of the fictitious nucleus determines (for given positions of the center of mass and electrons) the positions of both nuclei. Upon converting (3.3) in terms of the new co-ordinates $q_*, q_2 \ldots q_f$, we obtain:

$$\left\{ -\frac{h^2}{2M_*}\Delta_* - \frac{h^2}{2\mu}\sum_2^f \Delta_\nu + \frac{h^2}{M}\left(\sum_2^f \frac{\partial}{\partial q_\lambda}\right)^2 + U \right\} \psi = E\psi \qquad (3.4)$$

where $M_* = \dfrac{\mu_0 \mu_1}{\mu_0 + \mu_1}$.

The third term in the parenthesis is negligibly small compared to the second[1] term and it may therefore be ignored. The potential energy U, associated with the electrical forces, may to an approximation be calculated in such a manner that the nuclei q_0 and q_1 can be considered to occupy positions $-\dfrac{\mu_1}{\mu_0+\mu_1} q_*$ and $\dfrac{\mu_0}{\mu_0+\mu_1} q_*$ on both sides of the center of mass.

Later in Chapter 30 we shall further elaborate on equation (3.4) and we shall show how to an approximation we can reduce the eigenvalue problem into the simpler problem in which we have 1. a system of electrons in the field of two fixed nuclei and 2. an equation of (harmonic) vibrations having only one degree of freedom. This latter equation, by the way, governs the splitting up of the vibrational and rotational levels of

---

1) An operator is "small" in comparison to another operator when its matrix elements (which appear in the perturbation calculation) are relatively small.

the electron terms.

THE TRANSITION-PROBABILITIES. We shall now calculate the probability for the occurrence of a quantum-jump of the system from state ψ_n having energy E to a state $\psi_{n'}$ with energy $E' < E$, this jump being accompanied by the simultaneous emission of a light-quantum (polarized parallel to the x- , y- or z- axis) with $h\nu = E - E'$[1]. The product $X\psi_n$ (or $Y\psi_n$ or $Z\psi_n$) is expanded, (here, $X = \sum e_\nu x_\nu$ is the electric dipole moment of the system along the x-direction) the expansion being in terms of the orthogonal system of the eigenfunctions; the coefficient $X_{n'n}$ of $\psi_{n'}$ is, then, determined in this expansion:

$$X_{n'n} = (\psi_{n'}, X\psi_n).$$

As a result,

$$|X_{n'n}|^2 \frac{4\nu^3}{3hc^3}$$

is the desired probability, to which, of course, the intensity of the emitted light is proportional. In like manner the intensity of the light absorbed during the jump $E' \rightarrow E$ is proportional to $|X_{n'n}|^2$. In case of a degeneracy, $|X_{n'n}|^2$ is to be replaced by the sum of the squares $\sum |X_{n'n}|^2$ extending over all the n and n' for which $E_n = E$ and $E_{n'} = E'$.

---

1) For the formulation of this method by means of the Theory of Light-quanta, see P.A.M.Dirac. Principles of Quantum Mechanics. Cambridge 1930. (in German. von W. Bloch, Berlin 1930).

The selection rule proceeds from the intensity rule, as follows. If $X_{n'n} = Y_{n'n} = Z_{n'n} = 0$ for all $E_n = E$ and $E_{n'} = E'$, then no quantum-jump $E \to E'$ occurs, – consequently, the corresponding spectral line is absent.

The electric moment X for an atom can be set up with reference to its center of mass. The contribution of the nucleus to the electric moment can be ignored in comparison with the contributions of the electrons, since its distance from the center of mass is of an order of magnitude of $\frac{\mu}{M}$ as small. We may therefore set up the relation

$$X = -e \sum_1^f x_\nu \ ,$$

#4. The Electron in a Spherically-Symmetrical Field.

If in the Schroedinger Differential Equation of an electron

$$-\frac{h^2}{2\mu}\Delta\psi - eV\psi = E\psi \qquad (4.1)$$

the potential V is a function of the radius r alone, (hydrogen-atom, helium-ion) then, clearly, the variables may be separated by introducing the polar co-ordinates r, ϑ, φ.

$$\Delta = \frac{\partial^2}{\partial r^2} + \frac{2}{r}\frac{\partial}{\partial r} + \frac{1}{r^2}\Lambda,$$
$$\Lambda = \frac{1}{\sin\vartheta}\frac{\partial}{\partial\vartheta}\sin\vartheta\frac{\partial}{\partial\vartheta} + \frac{1}{\sin^2\vartheta}\frac{\partial^2}{\partial\varphi^2} \ . \qquad (4.2)$$

The eigenfunctions are then obtained in the form of products:

$$\psi = f(r)Y_\ell(\vartheta,\varphi) \tag{4.3}$$

$$\Lambda Y_\ell = \lambda Y_\ell\,, \tag{4.4}$$

$$-\frac{h^2}{2\mu}\left(\frac{\partial^2}{\partial r^2}+\frac{2}{r}\frac{\partial}{\partial r}+\frac{\lambda}{r^2}\right)f - eVf = Ef\,. \tag{4.5}$$

The Y_ℓ are spherical functions of the ℓ-th order, which may be defined simply as the potential forms of ℓ-th degree in U_ℓ (i.e. homogeneous polynomials of degree ℓ in x y z, which satisfy the potential equation $\Delta U_\ell = 0$) divided by r^ℓ.

$$U_\ell = r^\ell Y_\ell\,, \qquad \Delta U_\ell = 0\,.$$

Since $\Delta U_\ell = 0$, we obtain for the eigenvalue λ in (4.4) the value

$$\lambda = -\ell(\ell+1). \tag{4.6}$$

The spherical functions Y_ℓ are linear combinations of $2\ell+1$ linearly independent functions of the form:

$$Y_\ell^{(m)} = e^{im\varphi}\,y_\ell^{(m)}(\vartheta) \qquad (-\ell \leqq m \leqq \ell)^*\,.$$

---

*Proof. If the potential form U_ℓ is put into the form

$$U_\ell = \sum\sum c_{pq}(x+iy)^p\,(x-iy)^q\,z^{\ell-p-q}$$

then the potential equation $\Delta U_\ell = 0$ yields the recursion formula:

$$2(p+1)(q+1)c_{p+1,q+1} + (\ell-p-q)(\ell-p-q-1)c_{pq} = 0$$

for the coefficients.

Since, according to (4.4), these are eigenfunctions of the self-adjoint operator Λ, then, according to Chapter 2, any two spherical functions of different order are orthogonal to one another:

$$\int \bar{Y}_{\ell'} Y_{\ell} \, dv = 0 \quad \text{for} \quad \ell' \neq \ell .$$

In like manner, any two spherical functions $Y_{\ell}^{(m)}$ corresponding to different values of m, are orthogonal to each other:

$$\int \bar{Y}_{\ell}^{(m)} Y_{\ell}^{(m')} dv = 0 \qquad \text{for } m' \neq m,$$

because, the factor $e^{-im\varphi} e^{im'\varphi} = e^{i(m'-m)\varphi}$ becomes zero when integrated with respect to φ from 0 to 2π.

The spherical functions $Y_{\ell}^{(m)}$ form a complete orthogonal set on the sphere: Every continuous function on the sphere can be equally well approximated (to any degree of accuracy) by a sum of spherical functions. The proof: Every continuous function on the sphere can be easily and continuously extended into the interior. Also, as is known, every continuous function in a closed region of space can be approximated to any desirable degree of accuracy by means of a polynomial in x,y,z. Hence, all we need to show is that every polynomial on the sphere is equal to a sum of spherical functions. We now assert that

---

*Proof, cont'd.
If by $U_{\ell}^{(m)}$ we mean that part of the expression of U_{ℓ} whose terms have a constant difference p-q=m, then the coefficients of $U_{\ell}^{(m)}$ are determined by the recursion formula, except for a mutual factor. Hence, $U_{\ell} = \sum_m U_{\ell}^{(m)}$ and $U_{\ell}^{(m)} = r^{\ell} e^{im\varphi} Y_{\ell}^{(m)}(\theta)$.

every form of the n'th degree F can be expressed by means of the potential forms U_ℓ , in the following way:

$$F = U_n + r^2U_{n-2} + r^4U_{n-4} + \ldots + r^{2h}U_{n-2h} \ , \qquad (4.7)$$

The statement is clear insofar as the forms of the zero'th and first degree are concerned, since they themselves are always potential forms. If the statement is assumed to be valid for all polynomials of degree $< n$, then one proceeds for the polynomial F of degree n in the following manner: First we represent the polynomial of the (n-2)-degree ΔF in the manner indicated:

$$\Delta F = U^*_{n-2} + r^2U^*_{n-4} + r^4U_{n-6} + \ldots \ , \qquad (4.8)$$

We then equate:

$$U^*_\ell = (n-\ell)(n+\ell+1)U_\ell \qquad (\ell = n-2,\ n-4, \ldots) \qquad (4.9)$$

and we determine U_n from (4.7). All that remains to be shown is that this U_n satisfies the potential equation. Upon applying the Δ-operator on both sides of (4.8), the simple calculation yields:

$$\Delta F = \Delta U_n + \sum_{\ell=n-2k} (n-\ell)(n+\ell+1)r^{2k-2}U_\ell$$

$$= \Delta U_n + U^*_{n-2} + r^2U^*_{n-4} + \ldots \qquad (4.10)$$

A comparison of (4.8) with (4.10) yields $\Delta U_n = 0$, as was to be proven.

From the completeness requirement of the spherical functions it turns out according to the end of Chapter 2, that, indeed, the products (4.3) all represent eigenfunctions of (4.1).

Now, for every value of the azimuthal quantum number ℓ the eigenvalues E of (4.5) are indexed in the order of increasing size by means of a principal-quantum number $n = \ell+1, \ell+2, \ell+3 \ldots$. Every energy level $E(\ell, n)$ has a $(2\ell+1)$-fold degeneracy, since it is possible to choose for Y_ℓ in (4.3) all the $Y_\ell^{(m)}$ with $-\ell \leqq m \leqq \ell$. The number m is called the magnetic quantum number. For a further discussion of the eigenvalue problem (4.5) we refer the reader to the well-known literature[1].

In case of a coulombic field of attraction (hydrogen, He^+) with the potential $V = \frac{Ze}{r}$, the terms with the same n and $\ell = 0, 1, \ldots, n-1$ coincide. We then have.

$$E(n,\ell) = E_n = -\frac{B}{n^2}, \quad B = \frac{Z^2 \mu e^4}{2\hbar^2}$$

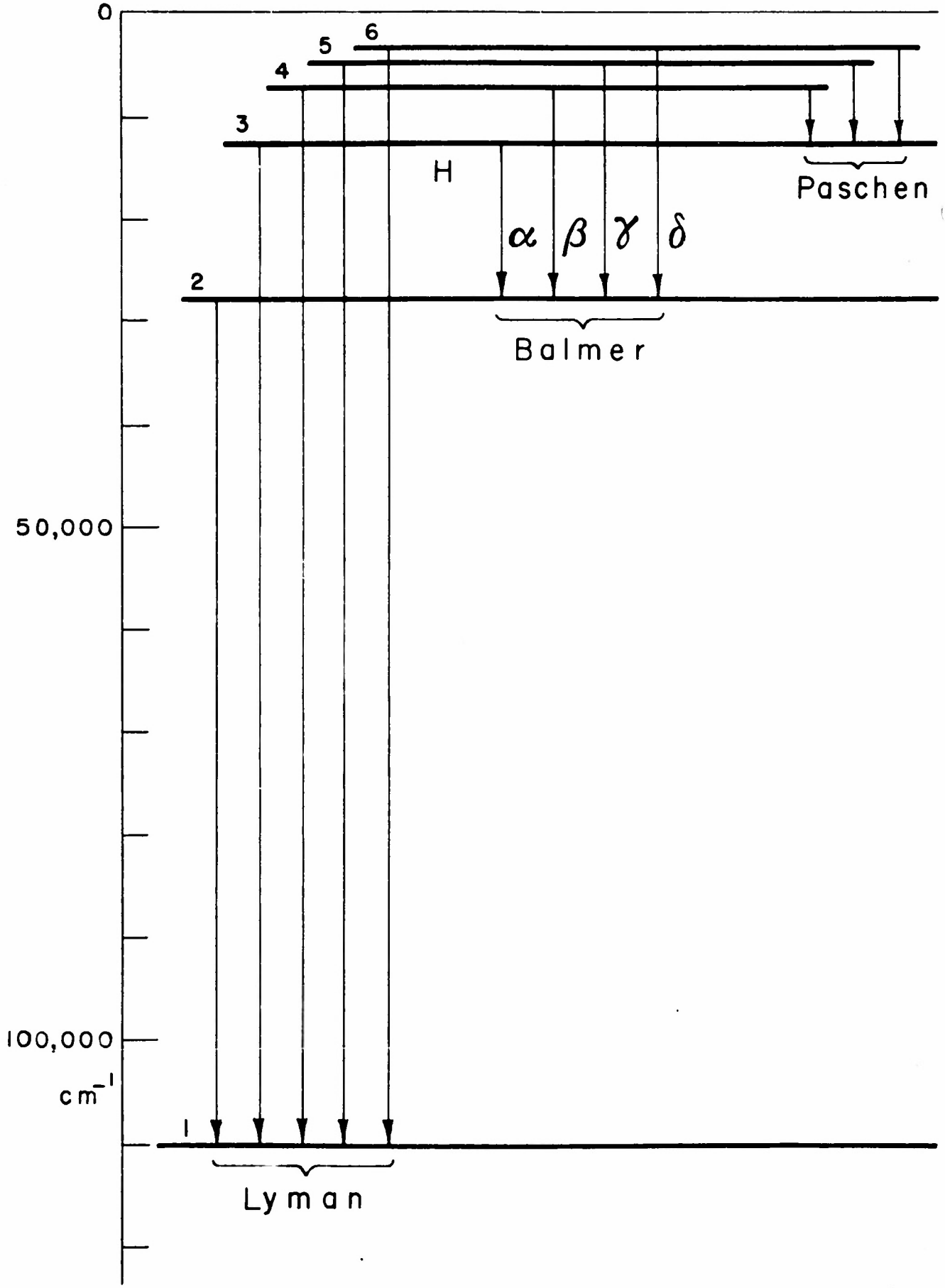
This yields for the terms, expressed in "wave-numbers" (reciprocal wave lengths), the following results.

$$\frac{1}{\lambda} = \frac{2\pi\nu}{c} = \frac{E}{2\pi\hbar c} = \frac{R}{n^2}, \quad R = \frac{Z^2 \mu e^4}{4\pi\hbar^3 c} = 109{,}722 \text{ cm}^{-1}.$$

This is, also, in harmony with the experimental observations. The transitions $E_n \to E_2$, $E_n \to E_3$, $E_n \to E_1$ give rise to the Balmer-, Paschen-, and Lyman series of the Hydrogen spectrum

---

1) Schroedinger, E. Abhandlungen zur Wellenmechanik, Leipzig, 1927.
Courant u. Hilbert. Methoden der mathematischen Physik, 2. Aufl. Kap. V, #12, 1931.



THE TERM-SPECTRUM OF THE "H" ATOM

Figure 1

(see. Fig. 1).

For a non-coulombic field, the terms with different l will, in general, also be different.

In accordance with section 3, the transition probabilities are obtained by expanding $X\psi_{nl} = -ex\psi_{nl}, Y\psi_{nl}, Z\psi_{nl}$ in terms of the eigenfunctions $\psi_{n'l'}$. In order to obtain the expansion of a quantity like

$$x\psi_{nl} = xf_{nl}(r)Y_l(\vartheta;\varphi) \tag{4.11}$$

in terms of $\psi_{n'l'}$, it is necessary to first expand xY_l on the sphere in terms of the sphercial functions. We apply the above expansion (4.7) to the polynomial $f = xU_l$. It is easily determined that

$$\Delta f = \Delta(xU_l) = \frac{\partial}{\partial x}U_l = U^*_{l-1}.$$

Since $\frac{\partial}{\partial x}U_l$ is, in fact, a potential form, only the first term appears in the expansion (4.8); consequently;

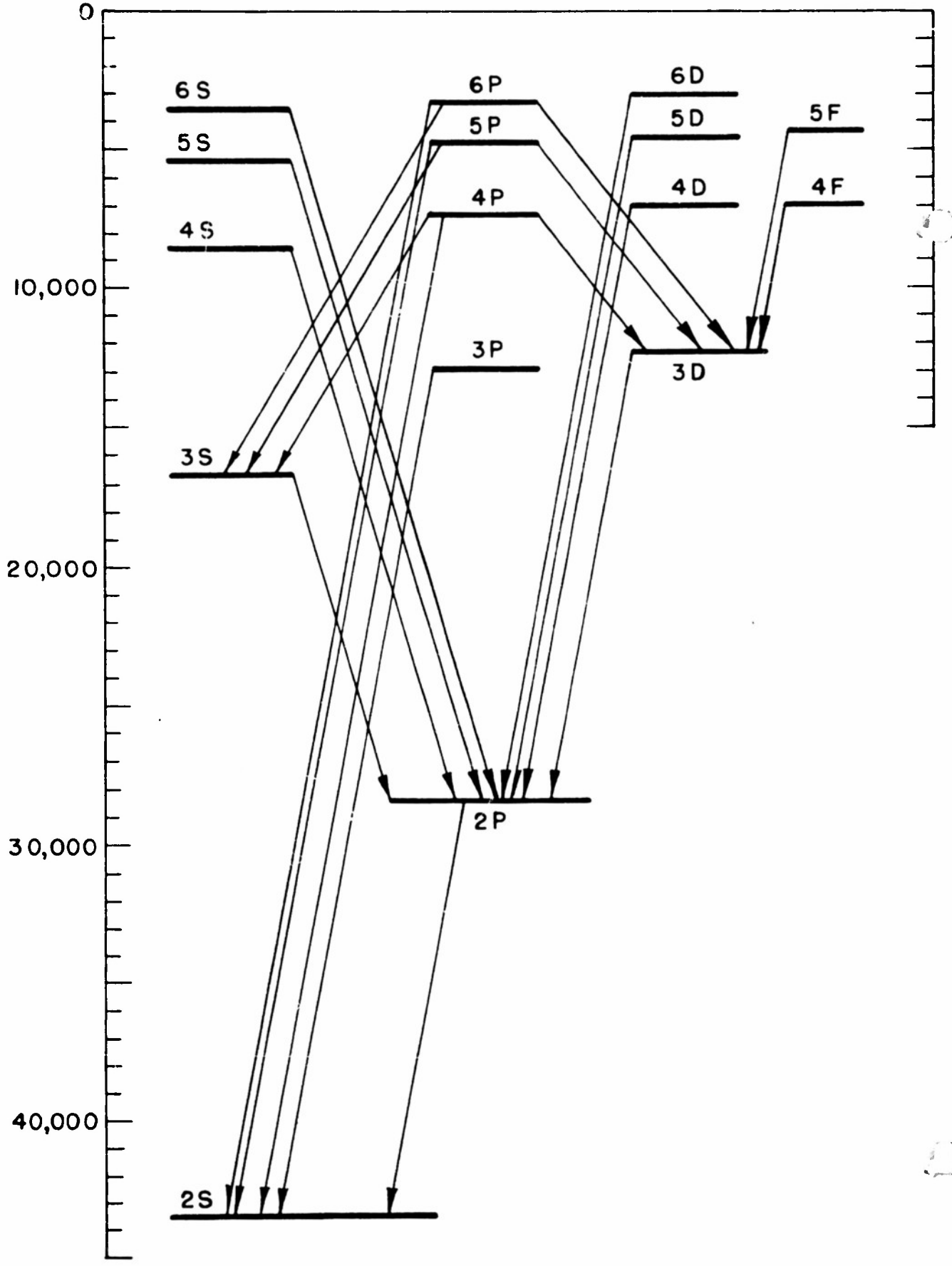
$$xU_l = U_{l+1} + r^2U_{l-1}.$$

If we again equate $U_l = r^lY_l$, then

$$xY_l = rY_{l+1} + r\,Y_{l-1} \tag{4.12}$$

In the expansion of $x\psi_{nl}$ (4.11) in terms of the eigenfunctions, only those terms appear $\psi_{n'l'} = f_{n'l'}(r)Y_{l'}$, for which $l' = l \pm 1$. The same is true of $y\psi_{nl}$ and $z\psi_{nl}$ hence, we have for the azimuthal quantum number the selection rule

$$l \rightarrow l \pm 1.$$



THE Li-ARC SPECTRUM

Figure 2

In the case of hydrogen, this rule can not be proven experimentally since the terms with different l coincide. The alkali metals Li, Na, K, Cs, however, have a "hydrogen-like" spectrum, as is borne out by experience. Their observed terms can be arranged in series. the s-,p-,d-,f-, and finally also the g- series. The terms of a single series are denoted by a principal quantum-number n. We thus write.

1s, 2s, 3s,.... for the terms of the s-series

2p, 3p,.... for the terms of the p-series, etc.

As for the term-values ns, np, nd, nf, ng, the following formula is valid:

$$-\frac{R}{(n-\varkappa)^2} \quad , \qquad (4.13)$$

where $\varkappa$ is a correction which depends on n only to a small extent and most of the time lies between $\frac{1}{2}$ and 0; for the higher series (d-, f-, etc) it is very nearly equal to zero. The spectral lines arise due to the transitions from a term of one series to a term of a neighboring series. That is, the s-terms combine only with the p-terms, the p-terms only with the s- and d-terms, etc. (see Fig. 2). In order that this be in agreement with the above selection rule for l, we must assign to the series s,p,d,f,.... the values $l=0,1,2,3,....$.

We can explain the hydrogen-like character of the spectrum by assuming that e.g. the Li-atom consists of 1. a spherically symmetrical core Li^{++} (nucleus and two electrons) which, as a rule, remains in the ground state, and 2. a

valence electron which makes the quantum-jumps. The field of a core is about the same as the field of a nucleus whose charge is screened off by the inner electrons[1]. The screening-off is complete as far as distantly removed points are concerned, it is incomplete in the vicinity of the nucleus. The screened-off field, therefore, is no pure coulombic field. Only at great distances away may we consider it roughly as a coulombic field as in the case of the H-atom; at smaller distances, though, a "perturbation" in the form of a stronger attraction takes place, bringing about a decrease in the potential energy and alongside it also a decrease in the eigenvalues. According to the "perturbation-theory" (see Chapter 5) the effect of the perturbation is small as long as the eigenfunctions in the unperturbed problem are small in the vicinity of the nucleus; the effect is large, however, when the eigenfunctions assume substantial values in the neighborhood of the nucleus. It so happens that for relatively large values of ℓ the eigenfunctions in the hydrogen-problem are small for small r, this being so, due to the fact that their power series expansion begins with r^{ℓ}. We, therefore, have an explanation for the steadily increasing degree of similarity to hydrogen on the part of the series with increasing ℓ.

---

1) To a second approximation, of course, we must take into account the "polarization" of the core by the valence electron.

#5. The Perturbation-Calculation.

The problem of the perturbation calculation is outlined as follows: Let the energy-operator H consist of two parts.

$$H = H^0 + \epsilon W$$

where the latter part, the "perturbation term" is provided with a small factor ϵ. The eigenvalue problem to be solved is, thus:

$$(H^0 + \epsilon W)\psi = E\psi \qquad (5.1)$$

Let us consider the unperturbed problem $H\psi^0 = E^0\psi$ to be already solved; We assume that the eigenfunctions $\varphi_1, \varphi_2, \ldots$ form a normalized, complete orthogonal system and that the corresponding eigenvalues $E_1^0, E_2^0, \ldots$ are arranged according to size. The problem is to determine the eigenfunctions and, above all, the eigenvalues of the perturbed problem to a first approximation; that is, we ignore any error of the order of magnitude of ϵ^2.

We assume that the eigenvalues depend on ϵ in a continuous and differentiable manner. Thus, the n'th eigenvalue E_n lies in the vicinity of the n'th eigenvalue E_n^0 of the unperturbed problem; it is expanded in powers of ϵ (Taylor Formula with the Remainder), and is represented by

$$E_n = E_n^0 + \zeta\epsilon + \ldots . \qquad (5.2)$$

We next omit the index n, and expand both sides of (5.1) in terms of the orthogonal system of φ_ν.

Let, $$\psi \sim \sum_1^\infty c_\lambda \varphi_\lambda ,$$
$$W\varphi_\mu \sim \sum_1^\infty w_{\lambda\mu}\varphi_\lambda ,$$
$$H^0\varphi_\lambda = E^0_\lambda \varphi_\lambda .$$

Upon equating the expansion-coefficients in (5.1) we obtain:

$$c_\lambda E^0_\lambda + \epsilon\sum_1^\infty w_{\lambda\mu} c_\mu = c_\lambda E. \qquad (5.3)$$

This equation is still rigorously valid. E lies in the vicinity of an E^0_n. If, now, $E^0 \neq E^0_n$, then for small values of ϵ, $|E^0_\lambda - E|$ lies above a fixed positive number; hence, we can solve (5.3) for c_λ.

$$c_\lambda = \frac{\epsilon\sum_1^\infty w_{\lambda\mu} c_\mu}{E^0_\lambda - E} . \qquad (5.4)$$

The question now is: Which $E^0_\lambda \neq E^0_n$?

Now, there may arise k consecutive, equal eigenvalues of the unperturbed problem which we denote by $E^0_n = E^0_{n+1} \ldots = E^0_{n+k-1}$ (k-fold degeneracy). We may then use the solution (5.4) for $\lambda \neq n, n+1, \ldots . n+k-1$. Here, we see that the c_λ with $\lambda \neq n, n+1, \ldots, n+k-1$ are small and of the order of magnitude ϵ.

But $c_n, c_{n+1}, \ldots . c_{n+k-1}$ are not small, at least not all of them. Their limiting values for $\epsilon=0$ are $c^0_n, c^0_{n+1}, \ldots . c^0_{n+k-1}$, hence

$$c_\lambda = c^0_\lambda + \ldots . , \qquad (5.5)$$

Upon substituting (5.2) and (5.5) into (5.3), and comparing

the δ-terms (keeping also in mind the order of magnitude of the c_λ with $\lambda \neq n, n+1, \ldots n+k-1$) we obtain,

$$\sum_{n}^{n+k-1} w_{\lambda\mu} c_\mu^\circ = \zeta c_\lambda^\circ \qquad (\lambda = n, n+1, n+k-1) \qquad (5.6)$$

Upon elimination of c_μ°, there follows the "secular equation":

$$\begin{vmatrix} w_{nn} - \zeta & w_{n,n+1} & \cdots & w_{n,n+k-1} \\ w_{n+1,n} & w_{n+1,n+1} - \zeta & \cdots & w_{n+1,n+k-1} \\ \cdots\cdots & \cdots\cdots\cdots\cdots & & \\ w_{n+k-1,n} & w_{n+k-1,n+1} & \cdots & w_{n+k-1,n+k-1} - \zeta \end{vmatrix} = 0 \qquad (5.7)$$

whose roots ζ_n, $\zeta_{n+1}, \ldots \zeta_{n+k-1}$ determine to a first approximation the energy levels of the perturbed problem with the aid of (5.2). The k-fold degenerate term E_n° is, therefore, "split up" into k (not necessarily different) terms. With the aid of (5.6) we find for each root ζ_λ certain values for $c_n^\circ, \ldots, c_{n+k-1}^\circ$, - the initial terms of the series for $c_n, \ldots, c_{n+k-1}$. The initial terms of the series for the remaining c_λ (i.e. the terms containing δ) are calculated from (5.4). The expansion coefficients of the solution thus found for the perturbed problem differ from those of a linear-combination $c_n^\circ \varphi_n + \ldots + c_{n+k-1}^\circ \varphi_{n+k-1}$ only by terms of the first order and higher. The "principal-axis" problem (5.6) will be discussed further in Chapter 8; it will be shown that i-linearly independent solutions of (5.6) correspond to an i-

fold root of (5.7), as our eigenvalue problem demands. As can be seen, the solution of the principal-axis problem (5.6) yields simultaneously the "first approximation" for the eigenvalue, and the "zero'th approximation" for the eigenfunctions. It is of course possible, also, to extend the approximation to higher powers of ϵ; we shall, however, not go into that.

The first approximation of the eigenvalues becomes inexact with increasing ϵ, as soon as in (5.4) the denominator becomes of the same order of magnitude as the numerator, so that the perturbed value of E approaches another term $E_\lambda^0 \neq E_n^0$. It is then said that the terms E_λ^0 and E_n^0 perturb one another. The mutual perturbation ceases, however, when the $w_{\mu\lambda}$ in the numerator vanish. That is, two terms E', E" do not perturb one another (no matter how close together they are) as long as $w_{\lambda\mu} = 0$ for $E_\lambda^0 = E'$, $E_\mu^0 = E''$.

The method of the perturbation calculation undergoes a modification in the case where the "unperturbed" eigenfunctions $\varphi_1, \varphi_2 \ldots$ are not eigenfunctions of a single operator H_0, but are, rather, the solutions of various approximate eigenvalue problems; accordingly, they do not form an orthogonal system. We assume, nevertheless, that the functions φ_ν do form a complete system of functions or, at least, that the eigenfunctions ψ of the proper operator H may be replaced (to a satisfactory approximation) by the sums $\sum_1^N c_\mu \varphi_\mu$, where the sum contains a large, though finite, number of terms. The eigenvalue problem $H\psi = E\psi$ now becomes

$$\sum c_\mu H\varphi_\mu = E\sum c_\mu \varphi_\mu . \qquad (5.8)$$

For the "zero'th approximation" we have $H\varphi_\mu = E_\mu \varphi_\mu$, and upon setting the coefficient on the left side ((of the equation)) equal to that on the right, we arrive at the result that in the zero'th approximation all c_μ are equal to zero, with the exception of those c_μ whose eigenvalues E_μ coincide with E in the zero'th approximation. These c_μ are again denoted by $c_n, c_{n+1}, \ldots c_{n+k-1}$. Instead of expanding, as before, both sides of (5.3) in terms of φ_μ , which would involve certain difficulties due to the non-orthogonality of the φ_λ , we form rather on both sides of (5.8) the scalar-product with φ_λ for $\lambda = n, n+1, \ldots n+k-1$:

$$\sum c_\mu (\varphi_\lambda , H\varphi_\mu) = E\sum c_\mu (\varphi_\lambda , \varphi_\mu).$$

The $(\varphi_\lambda, \varphi_\mu)$ with $\lambda \neq \mu$ are very small, since the φ_λ, being almost the eigenfunctions of the operator H, form almost an orthogonal system. Similarly, $(\varphi_\lambda, H\varphi_\mu)$ for $\lambda \neq \mu$ are very small, since the relation $H\varphi_\mu = E_\mu \varphi_\mu$ is <u>almost</u> true. Let us now omit these small terms completely,- inasmuch as they are also multiplied by small coefficients c_μ (i.e. for $\mu \neq n, n+1, \ldots n+k-1$), quite analogous to our previous omission of the terms with ϵ^2. We are left then with the finite system of linear equations, analogous to (5.6):

$$\sum_{n}^{n+k-1} c_\mu (\varphi_\lambda , H\varphi_\mu) = E\sum_{n}^{n+k-1} c_\mu (\varphi_\lambda , \varphi_\mu).$$

For the sake of brevity, we equate $(\varphi_\lambda, \varphi_\mu) = g_{\lambda\mu}$ and $(\varphi_\lambda, H\varphi_\mu) = h_{\lambda\mu}$, so that our system of equations now becomes:

$$\sum_{n}^{n+K-1} (h_{\lambda\mu} - E g_{\lambda\mu}) c_\mu = 0 ,$$

Elimination of c_μ again yields a type of secular equation

$$| h_{\lambda\mu} - E g_{\lambda\mu} | = 0 ,$$

from which the energy-values are determined to a first approximation.

If, in the above calculation, the small terms with $\mu \neq n, n+1, \ldots . n+k-1$ are not ignored, but rather the entire sum $\sum_{1}^{N} c_\mu \varphi_\mu$ is taken into account, and if, in addition, N is allowed to grow larger and larger, then a secular equation of ever increasing degree is obtained, from which the eigenvalues can be determined more and more accurately. This is the Ritz Method for the approximate solution of eigenvalue problems.

The perturbation calculation can, also, be used in the investigation of the behavior of a given wave function subject to a time-dependent perturbation. I refer the reader to the textbooks on Quantum-Mechanics.

#6. Moment of Momentum and Infinitesimal Rotations.

In accordance with Chapter 2, the following operators correspond to the angular momentum components of a system of f electrons:

$$hLi_x = \frac{h}{i}\sum\left(y\frac{\partial}{\partial z} - z\frac{\partial}{\partial y}\right)$$
$$hLi_y = \frac{h}{i}\sum\left(z\frac{\partial}{\partial x} - x\frac{\partial}{\partial z}\right) \qquad (6.1)$$
$$hLi_z = \frac{h}{i}\sum\left(x\frac{\partial}{\partial y} - y\frac{\partial}{\partial x}\right)$$

where the summation extends over all the electrons. Similarly, for the square of the total angular momentum we have the operator

$$h^2\vec{L}^2 = h^2(L_x^2 + L_y^2 + L_z^2)$$

The operators L_x, L_y, L_z bear a simple relation to the rotations in space. For a rotation of all the points (x,y,z) of the configuration space of a single electron about the z-axis, let us say - by an "infinitesimally small" angle $\delta\alpha$, the co-ordinates x,y,z undergo the transformations:

$$\delta x = -y\delta\alpha, \quad y = x\delta\alpha, \quad \delta z = 0$$

(except for magnitudes (of the coordinates) small as compared to $\delta\alpha$). A rotation D in space transforms a function $\psi(q) = \psi(x,y,z)$ into the function $D\psi = \psi'$, which is defined[1] by:

$$\psi'(Dq) = \psi(q) \quad \text{or} \quad \psi'(q) = \psi(D^{-1}q)$$

where D^{-1} is the reverse or inverse rotation.

---

1) This means that the graphic representation of the function (the system of its levels in space) undergoes a rotation D.

Hence,

$$\psi'(x,y,z) = \psi(x-\delta x, y-\delta y, z-\delta z)$$

$$\delta\psi = \psi' - \psi = -\frac{\partial\psi}{\partial x}\delta x - \frac{\partial\psi}{\partial y}\delta y - \frac{\partial\psi}{\partial z}\delta z$$

$$= \left(y\frac{\partial\psi}{\partial x} - x\frac{\partial\psi}{\partial y}\right)\delta\alpha .$$

Accordingly, the operation $-(x\frac{\partial}{\partial y} - y\frac{\partial}{\partial x})$ determines the increase of the function ψ with the rotation $\delta\alpha$ (except for magnitudes of higher order); it is known as an infinitesimal rotation about the z-axis. The increase of a function of several series of variables $\psi(q_1,\ldots,q_f)$ is determined by the operator

$$I_z = -\sum_1^f (x\frac{\partial}{\partial y} - y\frac{\partial}{\partial x}) = -iL_z \quad ;$$

here, the operation consists of a simultaneous rotation of all the points $q_1,\ldots,q_f$ about the same angle.

The operators I_x, I_y, I_z satisfy the transformation equations:

$$\begin{aligned} I_xI_y - I_yI_x &= I_z \\ I_yI_z - I_zI_y &= I_x \\ I_zI_x - I_xI_z &= I_y \end{aligned} \qquad (6.3)$$

as can be readily calculated.

In the case of a spherically symmetrical field, any rotation (hence, also, any infinitesimal rotation I_x, I_y, or I_z) sends the eigenfunctions of a certain energy level again into those of the same energy level; the linear family of eigenfunctions, thus, always undergoes a linear trans-

formation. In the case of a single electron, the eigenfunctions are products of functions f(r) and spherical functions of the ℓ'th degree U_ℓ. These undergo certain linear transformations due to the operators I_x, I_y and I_z. The factor f(r) remains constant and therefore, makes no contributions. For example, (since $Y_\ell^{(m)}$ depends on the angle φ merely due to the factor $e^{im\varphi}$):

$$I_z Y_\ell^{(m)} = - im\, Y_\ell^{(m)} ,$$

hence,

$$L_z Y_\ell^{(m)} = m Y_\ell^{(m)} .$$

Thus, the function $Y_\ell^{(m)}$ belongs to the eigenvalue m of L_z, or, in the state $f(r)Y_\ell^{(m)}$, L_z has the discreet value m.

We now proceed to calculate the operator $\vec{L}^2$. In the case of a single electron we have:

$$-\vec{L}^2 = \left(y\frac{\partial}{\partial z} - z\frac{\partial}{\partial y}\right)^2 + \left(z\frac{\partial}{\partial x} - x\frac{\partial}{\partial z}\right)^2 + \left(x\frac{\partial}{\partial y} - y\frac{\partial}{\partial x}\right)^2$$

$$= \left(y^2\frac{\partial^2}{\partial z^2} + z^2\frac{\partial^2}{\partial y^2} - 2yz\,\frac{\partial^2}{\partial y\,\partial z} - y\frac{\partial}{\partial y} - z\frac{\partial}{\partial z}\right) + \ldots$$

$$= r^2\Delta - \left(x\frac{\partial}{\partial x} + y\frac{\partial}{\partial y} + z\frac{\partial}{\partial z}\right)^2 - \left(x\frac{\partial}{\partial x} + y\frac{\partial}{\partial y} + z\frac{\partial}{\partial z}\right)$$

If we convert this into polar co-ordinates with the aid of (4.2) and the relation

$$x\frac{\partial}{\partial x} + y\frac{\partial}{\partial y} + z\frac{\partial}{\partial z} = r\frac{\partial}{\partial r} ,$$

we get:

$$-\vec{L}^2 = \Lambda . \qquad (6.4)$$

It follows from (4.4) and (4.6) that:

$$\vec{L}^2 Y_\ell = \ell(\ell+1) Y_\ell \quad ,$$

similarly, since the operator Λ acts only on ϑ and φ and not on r, we have:

$$\vec{L}^2 f(r) Y_\ell = \ell(\ell+1) f(r) Y_\ell \quad .$$

This means that <u>in the state $f(r)Y_\ell$, $\vec{L}^2$ has the discreet value $\ell(\ell+1)$</u>.

Thus, the moment of momentum $h\vec{L}$ is a vector, whose square has the value $h^2\ell(\ell+1)$ in the states $f(r)Y_\ell^{(m)}$ $(m=\ell, \ell-1, \ldots, -\ell)$ and whose z-component has the values hm. One can visualize this thing by drawing a vector of length $h\ell$, the direction of which is so chosen that its z-component assumes one of the only-allowed values $h\ell, h(\ell-1), \ldots, -h\ell$ If the z-component has a discreet value, then the component along another direction can <u>not</u> have any discreet value; the operators L_x and L_z are not commutative, and hence have no system of eigenfunctions in common. The spherical functions $Y_\ell^{(m)}$ which we deliberately assumed as a basis for all spherical functions, are exactly the eigenfunctions of L_z in the linear family of all spherical functions.

<u>ZEEMANEFFECT</u>. The linear magnetic perturbation term in the Schroedinger equation of an electron is as follows:

$$W = \frac{e}{\mu c} (\vec{A} \cdot \vec{p})$$

where,

μ = mass of the electron

e = charge of the electron

$\vec{A}$ = vector potential

$$(\text{rot } \vec{A} = H)$$

$\vec{p}$ = vector with components

$$p_x = \frac{h}{I} \frac{\partial}{\partial x}, \text{ etc.}$$

In the case of a constant homogeneous magnetic field of strength H_z along the z-direction ($A_x = -\frac{1}{t} yH_z$; $A_y = \frac{1}{t} xH_z$; $A_z = 0$), the above equation is equivalent to:

$$W = \varkappa H_z L_z; \qquad \varkappa = \frac{eh}{2\mu c} = \text{Bohr magneton.}$$

If H_0 is the unperturbed (spherically symmetrical) energy operator, then, according to the above, one can so choose the eigenfunctions of H_0 belonging to a certain eigenvalue E_0, that they simultaneously belong to a definite eigenvalue m of L_z. Thus, they are eigenfunctions of the sum $H = H_0 + W = H_0 + \varkappa H_z L_z$ belonging to the eigenvalue

$$E = E_0 + \varkappa H_z m . \qquad (6.5)$$

It is seen, therefore, that the Zeeman effect gives rise to a splitting of the term by the amount $\varkappa H_z m$. The same is true literally of a system of many electrons of equal charge and mass, provided, however, that the eigenfunctions belonging to each energy level are so chosen, that they are at the same time eigenfunctions of L_z. The eigenvalue m of L_z is called the "magnetic

quantum-number", since, according to the above, the atom behaves like a magnet whose magnetic moment in the z-direction amounts to m Bohr magnetons. The frequency ν of the emitted spectral lines is determined by:

$$h = E - E' = (E_0 - E_0') + \alpha H_z (m - m') . \qquad (6.6)$$

THE SELECTION RULE FOR m: The transition probabilities, to which the intensities of the emitted lines in the Zeeman-effect are proportional, are obtained (in accordance with Section 3) by expanding the products $X\psi_n$, $Y\psi_n$, $Z\psi_n$ in terms of the eigenfunctions $\psi_{n'}$. If, again, one so selects the ψ_n and $\psi_{n'}$, that they take on the factors $e^{-im\alpha}$ or $e^{-im'\alpha}$ (upon a rotation D_α about the z-axis by angle α) and if, also:

$$\begin{aligned} (X + iY)\psi_n &\sim \sum (X_{n'n} + iY_{n'n})\psi_{n'} , \\ (X - iY)\psi_n &\sim \sum (X_{n'n} - iY_{n'n})\psi_{n'} , \\ Z\psi_n &\sim \sum Z_{n'n}\psi_{n'} , \end{aligned}$$

then the left-hand sides of these series take on the factors $e^{-i(m+1)\alpha}$, $e^{-i(m-1)\alpha}$, $e^{-im\alpha}$, when undergoing a rotation D_α . On the right-hand sides, the rotation D_α can be carried out in two different ways: Either by carrying it out on all terms, thus yielding the factor $e^{-im'\alpha}$ for the term $\psi_{n'}$, or by multiplying the entire series by $e^{-i(m+1)\alpha}$, $e^{-i(m-1)\alpha}$ or $e^{-im\alpha}$. Both operations must yield the same results. It follows then that in the first series only terms with $m' = m+1$ can occur — in the second series

only those with $m' = m-1$; and in the third only those terms with $m' = m$. The following selection rule is then valid:

$$m' = m+1,\ m,\ m-1. \qquad (6.7)$$

For $m'=m$ only light linerarly polarized parallel to the z-axis is emitted; for the case $m' = m \pm 1$, an observer in the xy-plane would observe light linearly polarized parallel to this plane, whereas an observer looking along the z-direction would detect circularly polarized light[1].

The above deliberations were kept as general as possible, so that they would be valid for any axial-symmetrical force field and for any number of electrons, provided that the eigenfunctions may indeed be so chosen that they are multiplied by $e^{-im\alpha}$ corresponding to rotations D_α. In the case of the Zeeman-effect, equation (6.6) indicates that the wave numbers of the various splitting components of the spectral lines depend only upon the differences $m-m'$, which, according to (6.7), can only have the values 0 or ± 1. This means that <u>each line splits into three equidistant components (which correspond to the jumps $m \rightarrow m+1$, $m \rightarrow m$, $m \rightarrow m-1$) for which the above polarization rules are valid</u>. This is the normal Zeeman Effect. As for the Anomalous Zeeman Effect, no explanation can yet be offered at this stage; we shall come back to it in the fourth chapter when we speak about the "spinning electron".

1) This fact is a simple consequence of the classical laws of electro-dynamics. To show this, we associate with a change of the angular momentum by ± 1 a corresponding "circular orbital motion" of the electrons.

II. GROUPS AND THEIR REPRESENTATIONS.

#7. Linear Transformations.

An n-dimensional vector-space $(e_1,.....e_n)$ consists of the linear combinations $c_1e_1+.....+c_ne_n$ of n linearly independent base vectors with complex coefficients $c_1,.....,c_n$. Each group of quantities of any sort, which consists of the linear combinations of n-linearly independent quantities, are, likewise, termed as a vector-space. For example, the eigenfunctions of each individual energy-level, like those which have been treated in Chapter I, form a vector space.

A linear operator or a linear transformation A of the vector-space assigns to each base vector e_μ essentially a new vector e'_μ :

$$Ae_\mu = e'_\mu = \sum e_\lambda a_{\lambda\mu} \quad ; \tag{7.1}$$

in like manner it also assigns to each linear combination $v = \sum e_\lambda c_\lambda$ the same linear combination of the e'_λ :

$$Av = \sum\sum e_\lambda a_{\lambda\mu} c_\mu \quad .$$

The coefficients of Av are:

$$c'_\lambda = \sum a_{\lambda\mu} c_\mu \quad , \tag{7.2}$$

Accordingly, for a given choice of the base vectors $e_1,....e_n$, the transformation A is represented by the matrix:

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \cdot & & & \cdot \\ \cdot & & & \cdot \\ \cdot & & & \cdot \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix}$$

The matrix, too, is frequently denoted by A.[1)]

If two linear transformations A and B are carried out consecutively, first B then A, a product transformation AB is obtained which due to :

$$ABe_\mu = A(Be_\mu) = A \sum e_\lambda b_{\lambda\mu} = \sum \sum e_\varkappa a_{\varkappa\lambda} b_{\lambda\mu}$$

is represented by the product matrix $(a_{\varkappa\lambda}) \cdot (b_{\lambda\mu})$ with elements

$$\sum a_{\varkappa\lambda} b_{\lambda\mu} .$$

Equation (7.2) can be written as a matrix equation in the following manner:

$$\begin{pmatrix} c'_1 \\ \cdot \\ \cdot \\ \cdot \\ c'_n \end{pmatrix} = (a_{\lambda\mu}) \cdot \begin{pmatrix} c_1 \\ \cdot \\ \cdot \\ \cdot \\ c_n \end{pmatrix}$$

If the determinant $|A|$ of the matrix A is other than zero, then (7.2) can be solved for c_μ ; a transformation A^{-1} is then obtained which is the inverse of transformation A:

1) In the application of these formulæ, it is to be noted, that each individual equation (7.1) corresponds to a column of the matrix A, and each equation (7.2) to a row of A. Upon completely writing out the equations (7.1), one sees that A, the matrix of the coefficients, appears on the right-hand side in its transposed form:

$$Ae_1 = e_1 a_{11} + e_2 a_{21} + \cdots + e_n a_{n1}$$

$$Ae_2 = e_1 a_{12} + e_2 a_{22} + \cdots + e_n a_{n2}$$

. . . . . .

$A^{-1} Av = v$ for each vector v.

The product $A^{-1}A$, as well as AA^{-1}, is the "identity transformation" 1, represented by the "unit-matrix":

$$E = \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & & 0 \\ \cdot & & \cdot & \cdot \\ \cdot & & \cdot & \cdot \\ \cdot & & & \cdot \\ 0 & 0 & \dots & 1 \end{pmatrix}.$$

If, however, $|A| = 0$, then A has no inverse and A is known as a singular matrix.

If new base vectors $d_1 \dots d_n$ are introduced by means of a transformation of bases by the non-singular matrix $P = (p_{\lambda\mu})$:

$$d_\nu = Pe_\nu, \quad e_\nu = P^{-1}d_\nu, \quad P^{-1} = (q_{\varkappa\lambda})$$

then,

$$Ad_\nu = A \sum_\varkappa e_\varkappa p_{\mu\nu} = \sum_\lambda \sum_\mu e_\lambda a_{\lambda\mu} p_{\mu\nu}$$

$$= \sum_\varkappa \sum_\lambda \sum_\mu d_\varkappa q_{\varkappa\lambda} a_{\lambda\mu} p_{\mu\nu};$$

hence, the same transformation A, (relative to the new base now) can be represented by the matrix $P^{-1}AP$. If, on the other hand, the entire space is transformed by a transformation Q, then the transformation A becomes a new transformation QAQ^{-1}. For, if the old transformation A converts the vector v into w, then the new one must convert the vector Qv into Qw; this, in fact, is accomplished by the transformation QAQ^{-1}.

We speak of a unitary vector-space when we define a

Hermetian Form as follows:

$$(v,v) = \sum \sum g_{\lambda\mu}\, \bar{c}_\lambda c_\mu \; ; \qquad g_{\lambda\mu} = \bar{g}_{\mu\lambda}$$

the value it takes on for each vector $v = (c_1 \ldots, c_n)$ (except $v=0$) is always positive. (It is known as the positive-definite Hermitian Form.) The scalar-product of two vectors is defined by

$$(v,w) = \sum \sum g_{\lambda\mu}\, \bar{c}_\lambda d_\mu \; . \qquad (7.3)$$

The vectors are called orthogonal if the product is equal to zero.

The base vectors $e_1 \ldots e_n$ can always be chosen orthogonal, so that $(e_\lambda, e_\mu) = 0$ for $\lambda \neq \mu$ [1]; consequently, $g_{\lambda\mu} = 0$ for $\lambda \neq \mu$. They can next be normalized so that $(e_\lambda, e_\lambda) = 1$ or $g_{\lambda\lambda} = 1$. We then obtain:

$$(v,v) = \sum \bar{c}_\lambda c_\lambda \quad ; \quad (v,w) = \sum \bar{c}_\lambda d_\lambda \; ,$$

and the matrix $(g_{\lambda\mu})$ in (7.3) becomes the unitary matrix for the new, orthogonal base-vectors.

An m-dimensional linear sub-space $(v_1, \ldots . v_m)$ of a vector space $(e_1 \ldots . e_n)$ consists of all linear combinations of m-linearly independent vectors $v_1 \ldots . v_m$. Even when the v_λ are not linearly independent, their linear combinations still form a sub-space

---

1) PROOF: If e_1 to e_n are not orthogonal, then we replace every e_λ, for $\lambda = 2,3,\ldots,n$, by $e'_\lambda = e_\lambda + \beta_\lambda e_1$, where β is so chosen that $(e_1, e'_\lambda) = (e_1, e_\lambda) + \beta \cdot (e_\lambda, e_\lambda) = 0$. In the same manner we replace every e'_λ for values of $\lambda = 3,..,n$ by $e''_\lambda = e'_\lambda + \beta'_\lambda e'_1$ so that $(e'_2, e''_\lambda) = 0$, etc.

$(v_1, \ldots, v_m)$; the only difference is that the dimension of this sub-space is smaller than m.

For every m-dimensional sub-space $r = (v_1 \ldots v_m)$ of a unitary vector space R_n there exists a sub-space r' perpendicular to it consisting of all vectors w which are orthogonal to the range v_1 to v_m. Upon adding to the range v_1 to v_m additional base vectors v_{m+1} to v_n (orthogonal to the former) in order to form a base for the entire space, one can then easily see that the perpendicular space r' consists of all linear combinations of v_{m+1} to v_m. Hence, r' has the dimension n-m. r and r' together extend over the entire space R_n, i.e. each vector of R_n is the sum of a vector from r and one from r'.

A linear transformation A is <u>unitary</u> if it leaves all scalar products unaltered, i.e. when the following is always valid:

$$\sum g_{\lambda\mu}\, \bar{c}'_\lambda\, d'_\mu = \sum g_{\lambda\mu}\, \bar{c}_\lambda\, d_\mu \qquad \text{for} \quad c'_\lambda = \sum a_{\lambda\mu}\, c_\mu\,, \quad d'_\lambda = \sum a_{\lambda\mu}\, d_\mu\,.$$

For the case where $\tilde{A}$ is the complex conjugate transposed matrix to A and $G = (g_{\lambda\mu})$, we obtain the condition:

$$\tilde{A}\, G\, A = G$$

or, for the special case when G is the unit matrix, we have:

$$\tilde{A}\, A = E \quad \text{or} \quad \sum \bar{a}_{\lambda\mu}\, a_{\lambda\nu} = \delta_{\mu\lambda}\,. \tag{7.4}$$

Hence, $\tilde{A}$ is the inverse matrix of A:

$$A^{-1} = \tilde{A}. \tag{7.5}$$

A linear transformation A is Hermetian or self-adjoint (cf #1) whenever $(Av,w) = (v,Aw)$. If the matrix A relates to an ortho-normal system of base vectors, then the "self-adjoint" property is expressed by

$$a_{\lambda\mu} = \bar{a}_{\mu\lambda} . \tag{7.6}$$

Rule: If a self-adjoint or unitary transformation A transforms a linear sub-space, then it also transforms the sub-space which is completely perpendicular to it.

The rule is clear as far as the unitary transformations are concerned. As for the self-adjoint ones, it follows from the following considerations: If $v_1, \ldots v_m$ are the base-vectors of the first sub-space and if w is any arbitrary vector of the sub-space perpendicular to it, then

$$(Aw, v_\lambda) = (w, Av_\lambda) = 0$$

hence, Aw is orthogonal to all v_λ .

It follows easily from this rule that every self-adjoint transformation as well as every unitary one consists of a complete orthogonal-system of eigenvectors v, (i.e. one consisting of n vectors).

$$Av = \lambda v.$$

To wit:. With the aid of the secular equation (cf. # 5).

$$\begin{vmatrix} a_{11}-\lambda & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22}-\lambda & \cdots & a_{2n} \\ \cdots & \cdots\cdots & \cdots\cdots & \end{vmatrix} = 0 \quad , \qquad (7.7)$$

one immediately obtains an eigenvector v_1, whose components $c_1,\ldots,c_n$ form a solution of the system of equations:

$$\begin{aligned} (a_{11}-\lambda)c_1 + a_{12}c_2 + \ldots &= 0 \\ a_{21}c_1 + (a_{22}-\lambda)c_2 + \ldots &= 0 \\ \cdot \qquad\qquad\qquad\qquad &\cdot \\ \cdot \qquad\qquad\qquad\qquad &\cdot \\ a_{n1}c_1 + a_{n2}c_2 + \ldots &= 0 \end{aligned} \quad .$$

The space R_{n-1} which is completely perpendicular to it, is transformed by A in accordance with the above rule, thus, by the same method we find again in R_{n-1} an eigenvector v_2, etc.

If the $v_1,\ldots.v_n$ obtained are used as base vectors of R_n, then the transformation A can be represented by the diagonal-matrix:

$$\begin{pmatrix} \lambda_1 & 0 & \ldots. \\ 0 & \lambda_2 & \\ \cdot & & \cdot \\ \cdot & & \cdot \\ \cdot & & \cdot \\ \cdot & & \cdot \end{pmatrix} .$$

The transformation A is said to be transformed unto the principal axes. One arrives at the simple conclusion that every eigenvalue

λ of A occurs among the λ_ν and all the corresponding eigenvectors form linear combinations of those v_ν, for which $\lambda_\nu = \lambda$.

Rule: Every system of commutative unitary or self-adjoint matrices can be transformed unto the principal axes.

Proof: If all the matrices of a system are multiples of the unit matrix, then the case is trivial. A is then to be considered a matrix which is not a multiple of the unit matrix. Let a complete orthogonal system of eigenvectors of A be given by:

$$v_1, v_2, \ldots, v_k; \quad w_1, \ldots w_h; \quad \ldots\ldots;$$

and let the corresponding eigenvalues be:

$$\lambda_1, \lambda_1, \ldots, \lambda_1; \quad \lambda_2, \ldots \lambda_2; \quad \ldots \quad (\lambda_1 \neq \lambda_2, \text{ etc.}).$$

Now, to the eigenvalue λ_1 there corresponds a linear subspace $R_k = (v_1, \ldots v_k)$ which consists of all eigenvectors belonging to this eigenvalue; so, also, there corresponds to λ_2 a space R_h, etc. In order that another matrix B commute with A, B must transform the spaces R_k, R_h individually (for if v is a vector in R_k, i.e. $Av = \lambda_1 v$, then it follows that $ABv = BAv = B\lambda_1 v = \lambda_1 Bv$; in other words, Bv again belongs to R_k.) If our rule is now assumed as proven for spaces of an order of dimension smaller than that given (for dimension 1, the case is trivial, of course), then all transformations A, B, ... in the spaces R_k, $R_h \ldots R_m$ can be transformed simultaneously unto the principal axes. From these our rule follows quite obviously.

The left side of the secular equation (7.7) may be considered as the determinant $|A - \lambda E|$ of the matrix $A - \lambda E$. Among its

coefficients the one of special note is the coefficient of $(-\lambda)^{n-1}$) the trace of A.

$$S(A) = a_{11} + a_{22} + \ldots + a_{nn}.$$

Of importance, also, is the term which is independent of λ - the determinant $|A|$. Since the left side $|A - \lambda E|$ of the secular equation remains invariant under a transformation of the matrix A unto another base:

$$|P^{-1}AP - \lambda E| = |P^{-1}(A - \lambda E)P| = |P|^{-1}|A - \lambda E| \cdot |P| = |A - \lambda E|,$$

then all the coefficients on the left side are likewise invariant. In particular,

$$S(P^{-1}AP) = S(A).$$

8. GROUPS

A quantity g of elements a,b,.... of any sort (e.g. numbers or transformations) constitutes a group, if the following four conditions are satisfied.

(8.1) To every pair of elements a,b there corresponds a "product" ab (or a·b) which also belongs to g.

(8.2) The Associative Law must hold:

$$ab \cdot c = a \cdot bc$$

(8.3) There exists an "identity element" e or 1, having the property

$$ae = ea = a .$$

(8.4) For every a of g there exists an inverse a^{-1} in $\mathfrak{z}$, so that $a \cdot a^{-1} = a^{-1} \cdot a = 1$.

The group is Abelian, whenever $ab = ba$.

If the elements of the group are transformations (linear, to be sure) or permutations (the interchanging of a number of objects) and if the product $a \cdot b$ represents that transformation or permutation which is due to the carrying out first of b and then of a (as in #7), then the Associative Law is automatically satisfied; the unit e is the "identity transformation" which leaves all objects unchanged, and a^{-1} is the inverse transformation, which reverses the transformation a. Thus, a number of (non-singular) transformations constitutes a group, if for each pair of transformations it also contains their product as an element and if for each transformation it also contains the inverse as an element. The same holds for groups of permutations.

For instance, the real rotations of the 3-dimensional space about a fixed point constitute a non-Abelian group. the rotational group $\mathfrak{d}_3$. A more general group is obtained upon including also the reflections. The rotations about a fixed axis constitute an Abelian group. The Lorentz transformations, i.e. the real non-singular transformations of the four-dimensional vector-space, which leave the quadratic form $x^2+y^2+z^2-c^2t^2$ invariant and do not reverse the sign of the time interval, form the Lorentz-group. The linear transformations of the determinant One in the n-dimensional vector-space form the special linear group $\mathfrak{c}_n$; the unitary transformations of the determinant One form the special unitary group $\mathfrak{u}_n$. The permutations of n objects give rise to the

symmetrical group S_n.

The permutations of the symmetrical group are denoted in the following manner. For P, a permutation of the numbers 1,2,3,4,5, which sends 1 into 5, 5 into 4, 4 into 1, and furthermore, 2 into 3, and 3 into 2, we write: P = (154)(23). The manner of notation indicates that P is the product of the "cyclic interchanges" (154) and (23). Similarly, (1234) is the cyclic permutation which sends 1 into 2, 2 into 3, 3 into 4, and 4 into 1.

The "cyclic notation" just mentioned, is especially convenient when we wish to calculate the permutations QPQ^{-1} which are "conjugate" to a permutation P. For this purpose, one need only express P as the product of cyclic interchanges and carry out the permutation Q upon the symbols occurring in these cycles. For example, if

$$P = (12345), \quad Q = (23),$$

then

$$QPQ^{-1} = (13245).$$

The elements tst^{-1} conjugate to a group element s in any group g whatever, form a class of mutually conjugate group elements.

Every permutation can be written as a product of transpositions of type (ik) where only two numbers are interchanged. For example,

$$(12345) = (12)(23)(34)(45)$$

The even permutations, i.e. the products of an even number of transpositions, constitute in themselves a group called the alternating group A_n.

In some Abelian groups, we shall denote the element formed from a and b by a+b rather than by a·b as done so far. In this

case, we write 0 instead of 1 (since $a+0 = a$) and $-a$ instead of a^{-1} (since $-a + a = 0$). To these <u>additive groups</u> belong, for instance, all vector-spaces which satisfy, of course, the conditions (8.2) to (8.4) [in the additive notation, e.g. $(a+b)+c = a+(b+c)$].

The vector space has yet another peculiarity: it is possible not only to combine its elements (vectors) by means of addition, but also to combine them with the (complex) numbers by means of multiplication. The number times a vector yields a vector, thus

$$\theta(u+v) = \theta u + \theta v \qquad (8.5)$$

One speaks, in general, of a <u>group with operators</u> when certain "multipliers" or "operators" θ having the property (8.5) are applied to an additive group. For example, one may think of each system of linear transformations of a vector space along with all the numbers θ in terms of an operator domain for this vector space.

A subset of a group which forms a new group by virtue of the same rule of combination $a \cdot b$ is called a sub-group. The sub-group must contain in addition to a and b also $a \cdot b$, and in addition to a also a^{-1}. To wit: the alternating group A_n is a sub-group of the symmetrical group S_n. As for the additive groups, the sub-group must always contain $a+b$, and $-a$. And as far as a group with operators is concerned, the additional demand is made that the subset include a and θa as well (the admissible sub-group). <u>Examples</u>: the sub-spaces of a vector-space are admissible sub-groups, the integral multiples of a vector constitute a non-admissible sub-group.

A general example of a sub-group of a non-Abelian group G is

the central of G which consists of those groupelements z, which are permutable with all the elements of G.

The cosets or residue classes ag of a sub-group g in a group G are obtained by multiplying all elements of g on the left by any arbitrary element a of G. Two elements a, b belong to the same residue class, if $b^{-1}a$ belongs to g. Two different residue-classes have no element in common, and all residue-classes together make up the entire group G. Example: The residue classes of A_n in S_n are the classes of even and odd permutations.

Upon multiplying the elements of a residue-class ag by the elements of another residue-class bg, one does not always get a new residue-class of g. This is the case, though, when the group g is identical with all its "conjugate" subgroups aga^{-1}. Such sub-groups are known as normal divisors. For example, A_n is a normal divisor in S_n; similarly, the central of any group whatever is a normal divisor; S_n, on the other hand, is not a normal divisor of S_{n+1}. If the cosets ag and bg of a normal divisor g are multiplied by one another in the manner indicated and the resulting coset abg is described as the product of the two initial residue classes, then it is readily seen that these cosets, (to be thought of as elements) constitute a new group - namely, the factor group or coset group G/g. In the case of an Abelian group, (.e.g. a vector space) each subgroup is a normal divisor, it is, therefore, always possible to set up the factor group. In order to form the factor group of a vector space $R = (e_1, \ldots, e_n)$ corresponding to the sub-space $r = (v_1, \ldots, v_n)$, we combine into one class all the vectors which result from one vector upon addition of all

the vectors of the subspace; these residue classes are then to be thought of as elements of a new additive group. Upon multiplication by a number θ, the sub-space is mapped unto itself (i.e. it is an admissible sub-group) and, consequently, each residue class is carried unto a new residue class. Thus, in the case of the residue classes, too, we can explain the multiplication by a number θ. The coset-group is again a group with operators. If the base $(v_1, \ldots, v_m)$ of the sub-space is extended to a base for the entire space (by addition of more linearly independent $v_{m+1}, \ldots, v_n$), then we obtain a residue class involving all the vectors $c_1 v_1 + \ldots$ $\ldots + c_{m+1} v_{m+1} + \ldots + c_n v_n$ with fixed c_{m+1} to c_n. The cosets are thus uniquely represented by the vectors $c_{m+1} v_{m+1} + \ldots + c_n v_n$. We conclude that <u>the coset group R/r is an (n-m) - dimensional vector-space</u>.

Generally speaking, in the case of additive groups involving operators θ one may also multiply the residue classes of a sub-group g by the operators θ (provided the sub-group is an admissible one); hence, the factor-group G/g is a group having the same operator domain.

If to every element a of a group g there corresponds an element $\bar{a}$ of a group $\bar{g}$, in such a manner, that the product $\bar{a}\bar{b}$ corresponds to the product ab (and, likewise, the same relations hold for the identity elements and the inverses) and if, in addition, every element $\bar{a}$ of $\bar{g}$ appears as an image at least once, then we speak of a homomorphic mapping of g unto $\bar{g}$. (Example: Let g be the symmetrical group S_n. We assign to every even permutation the number +1, to every odd permutation the number -1. Then $\bar{g}$ is the group whose elements are the numbers +1 and -1.) If with

the different a, b there are always associated different images $\bar{a}, \bar{b}$ (which in our example above is true only for $n=2$), then the mapping is called an iso-morphism and the groups iso-morphic; they differ from another only insofar as the notation of the elements is concerned; the structure of the groups is the same, however. We then write

$$g \cong \bar{g}.$$

In the case of additive groups with operators, the occurrence of a homo-or iso-morphism is predicated on the basis that the sum $\bar{a} + \bar{b}$ is associated with the sum $a + b$ and also that θa is carried into $\theta \bar{a}$ (both operators must, therefore, have the same operator domain). We speak, thus, of an operator homomorphism or operator isomorphism. For instance, two vector-spaces are always operator-isomorphic if they have the same dimensionality, for then we can associate with the base vectors of one of the vector spaces those of the other vector space; likewise, we can associate with every linear combination of the base vectors of the first space the same linear combination of base vectors of the other space. Every linear transformation of a vector-space is an operator-homomorphism, as long as we consider only numbers as operators.

If a group g is mapped unto $\bar{g}$ homomorphically and not iso-morphically, it is easy to perceive that the elements of g, to which corresponds the identity element in $\bar{g}$, form a normal divisor h in g; also, the elements of g, with which we associate any arbitrarily fixed element of $\bar{g}$, form a residue-class of this normal divisor. Thus, there corresponds to every residue-class of h in a bi-unique manner an element of $\bar{g}$; it is this correspondence which is known

as an isomorphism. Thus we obtain the Homomorphy Rule.

If g is mapped homomorphically unto $\bar{g}$, then $\bar{g}$ is isomorphic with the factor-group g/h, where h consists of those elements of g to which corresponds the identity element in $\bar{g}$. In a reverse manner, g is always mapped homomorphically unto each factor-group g/h, whenever we associate with every element the particular residue class in which it lies. (In the above example where $\bar{g}$ consisted of the numbers +1, -1, h happens to be the alternating group.)

The homomorphy rule is valid, too, for groups with operators, provided one always substitutes for the expressions homo- and isomorphy, those of operator homo- and isomorphy, respectively.

An important special case of the homomorphism concept arises when the corresponding group $\bar{g}$ consists of linear transformations of a vector-space R. A non-singular linear transformation A of the space R is associated with every element of the group g, in such a manner that the product AB always corresponds to the product ab. We speak, thus, of a representation of the group g by means of linear transformations (or by means of matrices). The dimensionality n of the representation space is known as the order of the representation. If the correspondence is bi-unique or, in other words, isomorphic, then we speak of a faithful representation. If the representation is not a faithful one, then, in accordance with the homomorphy rule, it is nevertheless a faithful representation of a factor group g/h.

For a detailed presentation of the fundamentals of group theory the reader should consult A. Speiser, "Theorie der Gruppen von endlicher Ordnung" or B.L. van der Waerden, "Modern Algebra I";

both copies appear in this series.

The application of the concept of representations of groups to quantum mechanics is based on the following:

The Schroedinger Differential Equation of a system is carried into itself by certain transformations of the variables occurring in the wave function, as for example:

a) the interchanging of the co-ordinates of the various electrons (or eventually nuclei), which are indistinguishable in the differential equation.

b) the translations, rotations, and reflections of the space which leave the force-field in question unaltered. In an atom where the nucleus is considered as a fixed center of force, we consider the rotations of the space about this point as well as the reflection at this point. For an atom in a homogeneous magnetic or electric field we replace the entire rotational group by the subgroup of the rotations about a fixed axis. In the case of the diatomic molecule, both atomic nuclei are assumed to be fixed centers of force (to a first approximation); here, then, we must consider the rotations about the line joining the nuclei and also the reflections on planes passing through this line. As for the case of two nuclei of equal charge, one must also take into account the reflection on the plane passing perpendicularly through the center of the line joining the nuclei, etc.

The transformations of the Schroedinger Eigenvalue problem constitute a <u>group</u> in the following instances: a) the symmetrical permutation group S_f when f electrons are involved, b) a group due to rotations and reflections. The transformations of this

group give rise to transformations of the wave functions ψ, whenever it is determined that a transformation T in space (rotation or reflection), which carries the system of points $q_1, q_2 \ldots, q_f$ into $q_1', \ldots, q_f'$, also carries the wave function ψ into ψ', where:

$$\psi'(q_1', \ldots q_f') = \psi(q_1, \ldots q_f)$$

or, what amounts to the same thing:

$$\psi'(q_1, \ldots q_f) = \psi(T^{-1}q_1, \ldots, T^{-1}q_f).$$

In this manner the functions ψ are <u>linearly</u> transformed. Now, when S and T are both transformations, then

$$(ST)\psi = S(T\psi).$$

Since the Schroedinger differential equation is not altered by these transformations, its eigen-functions must then pass over into eigenfunctions belonging to the same eigenvalue. Thus, <u>the eigenfunctions corresponding to each energy level are linearly transformed, and it is these transformations which constitute a representation of the group in question</u>.

When one succeeds in setting up the various possible representations of the groups under consideration and in classifying them, one arrives then at a classification of the eigenfunctions and eigenvalues of the atoms and molecules. The "group-theoretical arrangement of the system of terms" is based on that.

§9. EQUIVALENCE AND REDUCIBILITY OF REPRESENTATIONS.

It is often convenient to think of the vector space R of a representation as an additive group with operators; the operators are the group elements a, and the multiplication av of a by a vector v means that the linear transformation A associated with a is to be carried out on the vector v. This manner of notation suggests itself as a result of the quantum-mechanical applications — namely, the result of a rotation D or permutation P applied to an eigenfunction ψ is quite naturally denoted by $D\psi$ and $P\psi$ respectively. By use of this notation one avoids introducing the corresponding matrices A; this is very convenient inasmuch as we often consider various representation spaces, subspaces, etc. simultaneously. In these cases then, we would have to introduce always new letters A, A', etc. to represent the various transformations. Finally, there is the additional advantage in that all group-theoretical concepts and rules can be applied directly to representation-spaces as soon as we think of the latter as additive groups with operators. Thus, the concept of operator-isomorphism as applied to two representation-spaces R, R' of the same group g, leads immediately to the concept of the equivalence of two representations. <u>The two representations of g in R and R' are equivalent, if one can choose the base vectors in R and R' in such a manner that each group element in both spaces can be represented by the same matrix</u>. That means, a representation generated by matrices A is equivalent to the representation $P^{-1}AP$, where P is any fixed matrix.

The concept of an admissible sub-group (i.e. one invariant with respect to the operators θ) leads to the concept of an <u>invariant subspace</u>, i.e. a linear subspace which is transformed into itself by the transformations of the representation. If there exists such an invariant subspace r which consists not merely of the null vector nor even comprises the total space R, then we call the representation as well as the vector space reducible (with respect to the group g)[1].

What is the form of the matrices of a reducible representation? We choose a system of base-vectors $u_1, \ldots u_h$ in the invariant subspace r, and extend it to a base $(u_1 \ldots u_n)$ for the entire space R. Then we have

$$\left.\begin{aligned} a u_\mu &= \sum_1^h u_\lambda p_{\lambda\mu} \qquad (\mu = 1, \ldots h) \\ a u_\nu &= \sum_1^h u_\lambda q_{\lambda\nu} + \sum_{h+1}^n u_\lambda s_{\lambda\nu} \qquad (\nu = h+1, \ldots, n) \end{aligned}\right\} \tag{9.1}$$

Hence, the matrix A is of the form:

$$A = \begin{pmatrix} P & Q \\ 0 & S \end{pmatrix},$$

where P, Q and S again are matrices, and 0 the null-matrix. The matrices P belong to the representation of g in the subspace r. What are the matrices S?

---

1) As far as the concept of reducibility is concerned, it is immaterial whether the operators (a or θ) form a group or not; any system of operators which give rise to linear transformations of R will suffice.

The factor space R/r admits again the operators a of g. If in equations (9.1) we pass from the vectors u to their corresponding residue classes $\bar{u}$ of R/r (cf. the second part of the homomorphy rule), then $\bar{u}_1, \ldots \bar{u}_h$ become equal to zero, since, of course, the $u_1, \ldots u_h$ all lie in the sub-space r, i.e. the coset of zero; $\bar{u}_{h+1} \ldots \bar{u}_n$, on the other hand, form a linearly independent base for the factor-space R/r. Hence,

$$a\bar{u}_\nu = \sum_{h+1}^{n} \bar{u}_\lambda s_{\lambda\nu} \qquad (\nu = h+1, \ldots n)$$

Therefore, the matrices S belong exactly to the representation which is generated by the factor-space R/r.

The choice of the complementing base vectors u_{h+1} to u_n involves, of course, a certain degree of arbitrariness.

If by a special choice of these base-vectors it is possible to make the matrix elements q all go to zero, then the $(u_{h+1} \ldots, u_n)$ again span an invariant sub-space s and it is said that the space R <u>breaks apart</u> into the subspaces r and s:

$$R = r + s \,. \tag{9.2}$$

It is also said of the representation D generated by R that it breaks apart into the representations D_1 and D_2, generated by r and s. We thus write.

$$D = D_1 + D_2$$

The matrices S belong, in this case, to the representation in the space s. This leads immediately to the isomorphy rule. From (9.2) there follows. $s \cong R/r$ (and likewise, $r \cong R/s$).*

In subsequent paragraphs we shall deal mainly with representations of groups by unitary transformations. Consequently, we have for every invariant subspace r an invariant subspace s completely perpendicular to it - hence, the reducibility, in this case, leads immediately to the breaking apart of the representation.

The union group which consists of all sums $a + b$ (a taken from a, b taken from b) is denoted by (a,b) where a and b are subgroups of an additive Abelian group. If each element of s =(a,b) can be represented even uniquely by the sum $a+b$, then s is the direct sum of a and b and we write s = a+b, as in equation (9.2). A criterion for the directness of the sum is that the subgroups a and b have only the zero element in common[1]. We define the direct sum of more than two subgroups in an analogous manner; the notation

$$s = a_1 + \ldots + a_n$$

signifies that every element of s can be uniquely represented by the sum $a_1 + \ldots + a_n$. The criterion for this is that every a_λ has only the zero element in common with the sum of all the preceding a_λ.

An additive group G with operators (in particular, a vector-

---

* This rule is a special case of the general group-theoretical isomorphy rule. $V/A \cong B/D$, where V and D represent the union and intersection respectively of the subgroups A,B; The rule is valid as soon as U becomes a normal divisor in V. See B.L. van der Waerden. Moderne Algebra I 40.

1) In the case of non-Abelian groups, it is required, moreover, that a and b be normal divisors in s.

space along with a representation) is <u>irreducible</u> or <u>minimal</u> if it contains no invariant subgroup (not counting itself and the zero element). It is <u>completely reducible</u> if it is a direct sum of irreducible (admissible) subgroups.

$$G = g_1 + \ldots + g_h .$$

In like manner, a representation is completely reducible if the vector-space belonging to it is completely reducible; the representation then appears in the "reduced form" :

$$D = D_1 + \ldots + D_h .^{*}$$

Among the D_ν there may, of course, occur equivalent pairs.

An example of a completely reducible representation: If one takes three base vectors and subjects them to all permutations of the symmetric group S_3:

(12) $e_1 = e_2$ (123) $e_1 = e_2$

(12) $e_2 = e_1$ (123) $e_2 = e_3$

(12) $e_3 = e_3$ (123) $e_3 = e_1$, etc. ,

then one obtains a representation of the permutation group by linear (unitary) transformations. The representation is reducible because the vector $s = e_1 + e_2 + e_3$ is invariant under all permutations; consequently, the same is true also of the sub-space or

---

* The matrices of the representation are of the form:

$$A = \begin{pmatrix} A_1 & & & 0 \\ & A_2 & & \\ & & \ddots & \\ 0 & & & A_h \end{pmatrix}$$

where the matrices A_ν belong to the irreducible representations D_ν .

"ray" r_1, which consists of all multiples $s\alpha$. The subspace r_2 perpendicular to this ray, which is generated by the differences e_1-e_2, e_2-e_3, is likewise invariant; as is obvious, it no longer contains any invariant subspace. Hence, the three-dimensional vector space G is completely reducible:

$$G = r_1 + r_2,$$

and the representation breaks apart into two irreducible representations of orders1 and 2. The matrices of these representations can be easily set up in an explicit form. The first order representation is the "identical representation". It associates with each permutation the unit matrix. The second order representation consists of the matrices.

$$(12) \sim \begin{pmatrix} -1 & 1 \\ 0 & 1 \end{pmatrix}; \quad (13) \sim \begin{pmatrix} 0 & -1 \\ -1 & 0 \end{pmatrix}; \quad (23) \sim \begin{pmatrix} 1 & 0 \\ 1 & -1 \end{pmatrix};$$

$$(123) \sim \begin{pmatrix} 0 & -1 \\ 1 & -1 \end{pmatrix} \quad (132) \sim \begin{pmatrix} -1 & 1 \\ -1 & 0 \end{pmatrix}; \quad 1 \sim \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

Unitary representations (representations by unitary transformations) and systems of unitary transformations in general, are always (either irreducible or) completely reducible, for if the vector-space R is reducible and r is an invariant subspace, then R breaks up into r_1 and an invariant subspace r_2 completely perpendicular to it. If one of these two spaces is again reducible, then it breaks apart in the same fashion, etc.

The representations occurring in quantum-mechanics are always unitary, since for each rotation or permutation the integral

$$N\psi = \int \bar{\psi}\psi dV$$

extending over the entire phase space remains invariant.

The quantum-mechanical significance of the reducibility of a representation is based on the following:

The energy levels of a system (such as a multiple electron system) are often ascertained upon ignoring certain terms in the energy-expression, these being again introduced into the problem later on as perturbation terms ϵW (where ϵ is allowwed to grow continuously from 0 and up); now, one often knows from the nature of the perturbation terms that they, as well as the unperturbed operator H_0, are invariant under a group g. The unperturbed as well as the perturbed eigenfunctions undergo under the operations of the group g a system of linear transformations which may be reducible or irreducible. At the limit $\epsilon \to 0$, the transformation of the perturbed eigenfunction must become that of the unperturbed one. Now it is clear, however, that at the limit a reducible transformation group can never go over into an irreducible one, for at the limit, the matrices

$$\begin{pmatrix} P & Q \\ 0 & S \end{pmatrix} \text{ or } \begin{pmatrix} P & 0 \\ 0 & S \end{pmatrix}$$

always give rise to matrices of the same form. Nor does the order of the representation ever change at the limit; the most that can happen is that the different energy levels may move closer together, so that consequently two or more spaces of dimensions n_1, n_2 are molten into one space of dimension n_1+n_2 in which we now have a reducible representation.

We conclude from here: If there exists at the limit for

$\epsilon=0$ an n-th order irreducible representation, then for small ϵ, too, there exists an irreducible representation of the same order. And since it is possible to make ϵ as large as desired by its continuous growth, therefore the same result is also valid for large perturbation terms. If prior to the perturbation we have an irreducible representation of the group g of order n and if the perturbation is invariant under the group g, then this perturbation, no matter how large it may be, can never cause a term-splitting; instead, it always remains an irreducible representation of order n. In like manner one observes: If for $\epsilon=0$ there exists an n-th order completely reducible representation which breaks up into the irreducible $\sum_1 + \sum_2 + \ldots + \sum_r$, then, as a result of the perturbation, the n-fold term will at most split up into r terms the eigenfunctions of which then undergo at the limit for $\epsilon \to 0$ the irreducible representations $\sum_1, \ldots \sum_r$.

We shall later set up all the possible irreducible representations for all the groups under consideration here. In all future cases, they will be differentiated from one another by means of numbers (quantum-numbers). A continuous variation of the quantity ϵ will not cause these numbers to become discontinuous, thus (except for equivalence) the representation must always remain the same for changing ϵ. Thus, in the above case, it always remains as one of the representations $\sum_1, \ldots \sum_r$ (or eventually as a sum of some of these).

10. REPRESENTATION OF ABELIAN GROUPS. - EXAMPLES

In the case of a unitary representation of an Abelian group the matrices of the representation all commute with one another, hence, according to the end of # 7, they can all together be transformed to the principal axes. If $v_1, \ldots, v_n$ are the principal axes or eigenvectors, then the one-dimensional sub-spaces $(v_1), \ldots, (v_n)$ are invariant under all the transformations of the group, hence the representation breaks apart into a number of first order representations which, of course, are irreducible.

Example 1: Let the group to be represented be cyclic of order n - i.e. it is to consist of the powers $1, a, a^2, \ldots, a^{n-1}$ of an element a, where $a^n = 1$. The element a in a first order representation is represented by the matrix (α). Consequently, a^2 is represented by (α^2), etc., and finally, $a^n = 1$ by (α^n). Hence, $\alpha^n = 1$ and α is an n-th root of one. There are n different n-th roots of one:

$$\alpha = e^{\frac{2\,im}{n}} \qquad (m = 0,1,\ldots,n-1) \quad ,$$

hence, a cyclic group of the n-th order has exactly n different representations of the first order. Any representation whatever breaks apart into similar representations of the first order where, of course, a certain representation of the first order may occur more than once.

A cyclic group of the 2nd order, for instance, is the group of permutations of two objects (electrons). The representations are given by

$$a \rightarrow (+1) \quad , \qquad a \rightarrow (-1)$$

where a represents the interchanging of the two objects. The vectors v_+ belonging to the representation (+1) remain invariant under the operation a; the v_- belonging to (-1) change their sign:

$$a v_+ = v_+ \qquad a v_- = -v_-.$$

The vectors v_+ are called "symmetric", the v_- "antisymmetric". Thus, the spectrum of the Helium atom, for example, breaks apart into two completely separated term-systems, of which one (the singlet system) corresponds to the symmetric eigenfunctions and the other (the triplet-system) corresponds to the anti-symmetric eigenfunctions.

Exactly the same is valid for the group consisting of the reflection at the origin of the co-ordinate system (in three-dimensional space):

$$x' = -x, \quad y' = -y, \quad z' = -z$$

and the identity. Here too, there exist two kinds of base vectors v_+, v_-. The vectors v_+ are said to belong to the "reflection-character +1" and the vectors v_- to the "reflection-character -1". This differentiation demands a breaking apart of the term-system of any atom into two component-systems, which differ from one another by the reflection-character $w = \pm 1$.

Example 2. <u>The Rotational Group about a fixed axis.</u>

Each rotation D_φ is denoted by an angle of rotation φ. If the rotation D_φ is represented by the first order matrix $\chi(\varphi)$, then, if product is to correspond to product:

$$\chi(\varphi_1 + \varphi_2) = \chi(\varphi_1) \cdot \chi(\varphi_2).$$

The continuous solutions of this function equation are:

$$\chi(\varphi) = e^{c\varphi}$$

Since $\chi(2\pi) = \chi(0)$, of necessity, (at least for a single-valued representation), it follows that

$$e^{2\pi c} = 1,$$

hence, $ic = m$, m having integral values. The representation is, therefore,

$$\chi(\varphi) = e^{-im\varphi}$$

Consequently, there exist infinitely many representations of the first order, which correspond to the values

$$m = 0, \pm 1, \pm 2, \ldots\ldots .$$

Any single-valued, continuous representation can be constructed from these.

This consideration applies mainly to the theory of molecular spectra. If a di-atomic molecule is considered (to a first approximation) as a system of two fixed nuclei about which the electrons revolve, then a rotation about the line joining the nuclei will send the eigenfunctions of each energy level into themselves. We thus obtain for every energy level a representation of the group of rotations about this axis, which we consider to be in reduced form. The eigenfunctions, as well as their

corresponding terms, therefore can be differentiated from one another according to the different representations of the first order with $m = 0, \pm 1, \pm 2, \ldots$, to which they belong. It is customary to denote the absolute value $|m|$ by the symbol Λ. We denote the terms with $\Lambda = 0$ $(m=0)$ as Σ- terms, those with $\Lambda = 1$ $(m = \pm 1)$ as Π-terms, those with $\Lambda = 2$ $(m = \pm 2)$ Δ-terms, etc. The reason why the terms with opposed values of m (e.g. with $m = +1$ and $m = -1$) are not distinguished from one another in the notation, will soon become evident.

Example 3. The Axial Rotation-Reflection Group.

The energy operator of the diatomic molecule with two fixed centers, discussed above, not only permits rotations about the fixed molecular axis α, but reflections on the planes passing through this axis as well. These rotations and reflections together form a non-Abelian group G, the axial rotation-reflection group. A fixed reflection may be selected and denoted by s_y; now, by combining s_y with any arbitrary rotation we can form all other reflections.

The following relations hold:

$$D_\varphi \cdot D_\psi = D_{\varphi+\psi}$$

$$D_\varphi s_y = s_y D_{-\varphi}$$

In each representation-space of the group G it is next possible to reduce out the sub-group of the rotations. As a result we obtain certain base vectors v_m (where m is a whole integer) which upon a rotation D_φ take on the factor $e^{-im\varphi}$.

Next, let $m = \Lambda > 0$. The reflection s_y sends v_λ into the vector $v_{-\lambda}$, since

$D_\varphi (s_y v_\Lambda) = s_y D_\varphi v_\Lambda = s_y e^{+i\Lambda\varphi} v_\Lambda = e^{i\Lambda\varphi} (s_y v_\Lambda)$.

v_Λ and $v_{-\Lambda}$ together span a two-dimensional subspace $r = (v_\Lambda, v_{-\Lambda})$, which obviously is invariant under the group G and no longer contains any smaller invariant sub-space. For if there did exist a one-dimensional invariant sub-space r' in r_Λ, it would at the same time be an irreducible representation-space of the rotational subgroup. Hence, a rotation by φ would confer upon the base vector belonging to this subspace the factor $e^{im\varphi}$, where m could only be $\pm\Lambda$, since these two representations of the rotational group are the only ones to occur as components in the representation-space r_Λ, a reflection however, would send this base-vector into another one belonging to the rotation character $-\Lambda$. Consequently, the sub-space r' would have to be, at least, two-dimensional. The representation-spaces $(v_\Lambda, v_{-\Lambda})$, therefore, are irreducible. It follows from here that in the case of a molecule with two fixed nuclei, a pair of eigenfunctions ψ_Λ, $\psi_{-\Lambda}$ always belong to the same energy value. The representationmatrices for the rotation D_φ and the reflection s_y are:

$$\begin{pmatrix} e^{-i\Lambda\varphi} & 0 \\ 0 & e^{i\Lambda\varphi} \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix},$$

This representation is denoted by A_Λ.

For $\Lambda = 0$, in which case the vectors v_0 are invariant under all rotations, $v_{-\Lambda}$ need not be different from v_Λ. If one thinks of the space of the vectors v_0 as a representation-space of the cyclic group consisting of the identity and a reflection, and one reduces this representation of the cyclic

group, then two types of first order representations result. They are characterized by the "reflection-characters" +1 and -1 and denoted by A^+ and A^-, respectively. The base-vector of an irreducible representation space $A_0^{\pm}$ takes on the factor ± 1 for every reflection and is invariant under all rotations. The irreducible representations of the rotation-reflection-group G are, thus: A_0^+, A_0^-, A_1, A_2, A_3, ...

As the above example shows, there exist also first order representations of a non-Abelian group g. These representations, however, are of necessity unfaithful since the representation matrices commute with one another whereas the group elements do not. Since we associate with the group-elements ab and ba the same matrix, we also associate with the "commutator".

$$ab(ba)^{-1} = aba^{-1}b^{-1}$$

the unit-matrix. All these commutators and their products form a subgroup, the "commutator-group" of g, the elements of which are all represented by the unit matrix. The representation, therefore, is a true representation of an (Abelian) factor-group g/h, where the normal-divisor h includes, at least, the commutator-group.

Example 4. The symmetrical group $S_n (n > 2)$ is non-Abelian. Among the commutators we find the permutations:

$$(ij)(ijk)(ij)^{-1}(ijk)^{-1} = (ijk),$$

that is, all the "three-cycles". Clearly these permutations and their products form the "alternating group" A_n $(\neq 8)$. Hence, every first order representation of S_n is, at the same time, a

representation of S_n/A_n. Since this factor-group is cyclic and of the order 2 it has only two representations of the first order. One is the <u>identical</u> or <u>symmetrical</u> representation which associates with all permutations the unit matrix (1), the other one is the <u>anti-symmetrical</u> representation which associates with the even permutations the matrix (1) and with the odd ones the matrix (-1). All remaining representations of S_n are of a higher order than one.

Example 5. <u>The Symmetrical group S_3</u>.

It is possible to determine the irreducible representations of this group exactly by the same method employed in the determination of the representations of the axial rotation-reflection-group G in example 3. One merely starts out with any arbitrary representation of the S_3 and then reduces from it the representation of the alternating group A_3 contained therein. A_3 is cyclic and consists of the three permutations 1(123), (132). According to example 1, there exist only three first-order representations of the cyclic three-group where a third unit root

$$1 \text{ or } \rho = e^{\frac{2\pi i}{3}} \qquad \text{or} \qquad \rho' = \rho^{-1} = e^{-\frac{2\pi i}{3}}$$

is associated with the resulting element 1.

A vector v_ρ, belonging to the unit-root ρ, yields, upon the carrying out of the permutation (12) a vector $v_{\rho'}$ that belongs to ρ^{-1}, for we have:

$$(123)v_{\rho'} = (123)(12)v_\rho = (12)(123)^{-1}v_\rho$$
$$= (12)\rho^{-1}v_\rho = \rho^{-1}v_{\rho'} .$$

The vectors v_ρ , $v_{\rho'}$ span the space of an irreducible representation of the second order. In the space of the vectors v_1, which are invariant under the permutations 1, (123), (132), we find yet (upon reduction of the cyclic groups 1, (12)) another two representations of the first order, namely the identical - and the antisymmetrical one. Hence, altogether there are two representations of the first order and an irreducible representation of the second order. From this proof it follows that every representation breaks up completely into irreducible ones of the three types mentioned above. The second order representation obtained in the example of #9, must be equivalent to the second order representation (v_ρ , $v_{\rho'}$) described here; that can be easily verified by calculation.

11. THE UNIQUENESS THEOREMS.

Theorem 1. If $G = g_1 + \ldots + g_k$ is a completely reducible additive group and H is any arbitrary (admissible) sub-group, then $G = H + g_{\nu_1} + \ldots + g_{\nu_\kappa}$ (for an appropriate choice of the g_{ν_i} from the groups g_1 to g_k).

Proof: We form:

$$H_1 = (H, g_1),$$
$$H_2 = (H_1, g_2),$$
$$\ldots\ldots\ldots\ldots\ldots$$
$$H_k = (H_{k-1}, g_k) = G.$$

The intersection of H and g_1 is an invariant sub-group of g_1,

or, since g_1 was assumed to be irreducible, it is either g_1 or zero (i.e. it consists only of the zero-element). If the intersection is g_1, then g_1 is contained in H. Hence, $H_1 = H$. If the intersection is zero, then the sum (H, g_1) is direct, hence, $H_1 = H + g_1$.

In a manner analogous to H_1 we proceed with all other groups H_ν ; we then succeed in either converting all parentheses $(H_{\nu-1}, g_\nu)$ into direct sums or reducing them down to one term $H_{\nu-1}$. From all these equations together we, therefore, find that $G = H_\ell$ is a direct sum of H and some g_ν, as was to be proven.

Theorem 2. If $G = g_1 + g_2 + \ldots + g_\ell$, and also

$$G = g_1' + g_2 + \ldots + g_\ell$$

then $g_1 \cong g_1'$.

Proof: As a consequence of the isomorphism rule of #9, g_1 and g_1' are both isomorphic with the factor-group $G/g_2 + \ldots + g_\ell$.

Theorem 3. If $G = g_1 + g_2 + \ldots + g_\nu$ and $G = h_1 + h_2 + \ldots + h_s$ are two decompositions of a completely reducible additive group into irreducible ones, then $r = s$, and the g_ν, in whatever sequence they occur, are isomorphic with the h_μ.

Proof. Upon applying theorem 1 with $H = h_2 + \ldots + h_s$, it follows that:

$$G = (h_2 + \ldots + h_s) + (\sum g_\nu),$$

where $\sum g_\nu$ represents the sum of a number of g_ν. From theorem 2, we have therefore $\sum g_\nu \cong h_1$. Since h_1 is irreducible $\sum g_\nu$

must also be irreducible; hence, it can consist of only one term, namely (when the g_ν are number appropriately) the term g_1. Hence, we have $g_1 \cong h_1$ and

$$G = h_2 + \ldots + h_s + g_1.$$

By applying theorem 1 again, this time with $H = h_3 + \ldots + h_s + g_1$, we obtain:

$$G = (h_3 + \ldots + h_s + g_1) + \sum{}' g_\nu$$

(the sum $\sum'$ does not contain g_1); hence, by comparison with the previous equation, according to rule 2. $\sum' g_\nu \cong h_2$. Thus, again, the sum consists of one term only, namely g_2. We therefore have $g_2 \cong h_2$ and

$$G = h_3 + \ldots + h_s + g_1 + g_2 .$$

By continuing thus, we obtain $g_\nu \cong h_\nu$ ($\nu = 1,2,\ldots s-1$), and

$$G = h_s + g_1 + \ldots g_{s-1}.$$

As a consequence of theorem 2, again, we have: $h_s \cong g_s + \ldots + g_r$. Therefore, since, as before, the last of the sums may contain one term only, $r = s$ and $g_s = h_s$; the proof now, is complete.

In particular, this rule leads to the fact that the irreducible components, into which a representation breaks apart, depend only on this representation itself and do not depend upon the manner in which the vector-space is broken apart into irreducible subspaces. It, therefore, makes sense to say that a given representation D contains the irreducible representation D_1 three times and another one D_2 once.

One more note-worthy conclusion follows from theorem 1

$$G/H \cong g_{\mu} + \ldots + g_{\nu\kappa} .$$

Since each homomorphic image of G is isomorphic with a factor-group G/H, there follows:

Theorem 4. Every homomorphic image of a completely reducible additive group is isomorphic with the sum of a number of components g_ϱ of g.

12. THE KRONECKER PRODUCT TRANSFORMATION.

Let there be given a transformation A of the n-dimensional vector space R_n and a transformation B of R_m. We may take for R_n the totality of linear combinations $c_1u_1 + \ldots + c_nu_n$ in n variables $u_1, \ldots;$ likewise, for R_m we may take the totality of the linear combinations $d_1v_1 + \ldots + d_mv_m$. The base vectors are the u_λ and the v_ϱ .

The n·m linearly independent products $u_\lambda v_\varrho$ may again be thought of as base elements of a vector space whose vectors are the combinations $\sum c_{\lambda\varrho} u_\lambda v_\varrho$. If we now transform the u according to A and the v according to B, then the products $u_\lambda v_\varrho$ also undergo a linear transformation

$$u'_\mu v'_\sigma = \sum u_\lambda v_\varrho \alpha_{\lambda\mu} \beta_{\varrho\sigma} ,$$

which is denoted as the product transformation A × B. Use is made of the product transformation, in particular, where A and B are representations of the same operator a of a group g. In that

case, $(A \times B)u_\lambda v_\rho$ is simply the result of the application of the operator a to the product $u_\lambda v_\rho$. If A and B form two representations D and D' of the group g, then it is clear that $A \times B$ again form a representation - the <u>product representation</u> $D \times D'$.

In the same manner we may also multiply more than two representations by one another, we then obtain representations of the type:

$$D \times D' \times D'' = (D \times D') \times D'' = D \times (D' \times D'').$$

<u>The product representation obtained from two unitary representations is unitary</u>. The proof of this rule is left to the reader.

Now, let D and $\tilde{D}$ be two irreducible representations of a group G. We ask: Under what condition does the identical representation appear as a component in the product representation? Or what amounts to the same thing: Under what condition do we find in the space of the representation $D \times \tilde{D}$ an invariant vector?

If $R = (u_1, \ldots, u_n)$ is the space of the representation D and $S = (v_1, \ldots, v_m)$ is the space of the representation $\tilde{D}$, then each vector of the product space is of the form:

$$w = \sum \sum c_{\lambda\rho} u_\lambda v_\rho = \sum_1^n u_\lambda v'_\lambda \qquad \left(v'_\lambda = \sum_1^m c_{\lambda\rho} v_\rho \right)$$

In order that w be invariant, we must have for each group element a

$$\alpha v = \sum_1^n (\alpha u_\lambda)(\alpha v_\lambda') = \sum_1^n u_\lambda v_\lambda'$$

If we set

$$(\quad)' = \alpha u_\lambda = \sum_1^n u_\mu \alpha_{\mu\lambda}$$

then we obtain

$$\sum \sum u_\mu \alpha_{\mu\lambda} \cdot \alpha v_\lambda' = \sum u_\mu v_\mu'$$

or

$$\sum_1^n \alpha_{\mu\lambda} \cdot \alpha v_\lambda' = v_\mu' . \qquad (12.1)$$

If we set $\alpha_{\mu\lambda} = \alpha'_{\lambda\mu}$ (the transposed matrix) and denote the inverse matrix to $(\alpha'_{\lambda\mu})$ by $(\beta_{\lambda\mu})$ then upon multiplying by $\beta_{\mu\nu}$ and summing over μ :

$$\alpha v_\nu' = \sum v_\mu' \beta_{\mu\nu} \qquad (12.2)$$

we can solve equations (12.1) for $\alpha v_\lambda'$. As a first consequence we find that $(v_1', \ldots, v_n')$ is an invariant subspace of S under the group g; in other words, since S is irreducible, $(v_1', \ldots, v_n')$ corresponds to S. Thus, m, the dimension of S, is at most equal to n: $m \leqq n$. Secondly, upon permuting the roles of R and S, it can be shown that $n \leqq m$; hence $m = n$. In other words, the vectors $v_1', \ldots, v_n'$ are linearly independent; they may be employed as base-vectors for S and as such are denoted by $v_1, \ldots, v_n$. Now, (12.2) implies that the group element α in the representation $\tilde{D}$ is represented by the matrix $(\beta_{\lambda\mu})$. Thus:

<u>In order that there be an invariant vector w in the space of the representation $D \times \tilde{D}$, the matrices of $\tilde{D}$ (with respect to an appropriate base $(v_1, \ldots, v_n)$ must be the inverses of the</u>

transposed matrices of the representation D; then the invariant vector is given by

$$w = \sum_1^h u_\lambda v_\lambda \quad .$$

The relation between two representations D and $\tilde{D}$ which demands that $\sum u_\lambda v_\lambda$ be invariant, is of course a reversible one; the matrices of D are also the inverses of the transposed matrices of the representation $\tilde{D}$. The representations D and $\tilde{D}$ connected by this condition are contragredient to another. . For every representation D there is a contragredient representation $\tilde{D}$. If D is reducible then $\tilde{D}$, too, is reducible and conversely so. In the case of unitary representations the inverse-transpose matrices $(\beta_{\lambda\mu})$ are complex conjugates of $(\alpha_{\lambda\mu})$; in this case, therefore, the contragredient representation is also the complex-conjugate.

In order that a completely reducible representation $D = D_1 + D_2 + \ldots + D_h$ contain as a component a certain irreducible representation $\tilde{F}$ that is contragredient to F, it is necessary and sufficient that the decomposition of the product transformation $D \times F = D_1 \times F + \ldots + D_h \times F$ yield but once the identical representation - for then $\tilde{F}$ is contragredient to a D_ν .

A further consequence is the following: In order that the product representation $D \times E$ contain the irreducible representation $\tilde{F}$ as a component, the product representation $D \times E \times F$ must contain the identical representation at least once. This relation is symmetric in $D_1 E$ and F.

In the case of Abelian groups and above all, in the case of first order representations, the product representation is rather

trivial. If $(\chi(a))$ and $(\chi'(a))$ are the representation matrices of the group elements a, then $(\chi(a)\,\chi'(a))$ is the representation matrix for a in the product representation. If $D: a \to \alpha$ is a representation of the first order and the other D' is any representation whatever $a \to A$, then $a \to \alpha A$ is the product representation. If D' is irreducible then $D \times D'$ is also irreducible because a reducible system αA, upon multiplication of all matrices by α^{-1}, would yield a reducible system A. In the case of representations of order higher than one, it is quite likely that a product representation resulting from irreducible representations is reducible.

Example. We wish to calculate as well as reduce the product representations of the representations A_0^+, A_0^-, A_1, A_2, ... of the axial rotation-reflection group (§10, example 3).

The base vectors of the representations A_λ and A_μ are (for $\lambda > 0$, $\mu > 0$) $u_{\pm\lambda}$ and $v_{\pm\mu}$. Their products are $u_\lambda v_\mu$, $u_{-\lambda} v_{-\mu}$, $u_\lambda v_{-\mu}$, and $u_{-\lambda} v_\mu$. The reflection s_y permutes the first two among one another as well as the last two. The first pair takes on the factors $e^{\pm i(\lambda+\mu)\varphi}$ for rotations D_φ; it is, therefore, transformed according to $A_{\lambda+\mu}$. The second pair takes on the factors $e^{\pm i(\lambda-\mu)\varphi}$ for D_φ. It is, therefore, in the case $\lambda \neq \mu$, transformed according to $A_{|\lambda-\mu|}$. For the case $\lambda = \mu$ the two vectors $u_\lambda v_{-\lambda}$ and $u_{-\lambda} v_\lambda$ are invariant under the D_φ; their sum $u_\lambda v_{-\lambda} + u_{-\lambda} v_\lambda$ takes on the factor +1 as it undergoes the reflection s_y whereas their difference $u_\lambda v_{-\lambda} - u_{-\lambda} v_\lambda$ takes on the factor -1. Hence

$$A_\lambda \times A_\mu = A_{\lambda+\mu} + A_{|\lambda-\mu|} \qquad \text{for } \lambda = \mu \text{ both} > 0.$$

$$A_\lambda \times A_\lambda = A_{2\lambda} + A_0^+ + A_0^- \qquad \text{for } \lambda = \mu > 0$$

If $\mu = 0^+$, then the products $u_\lambda v_0$ and $u_{-\lambda} v_0$ transform exactly like u_λ and $u_{-\lambda}$ — that is, according to A_1. Consequently,

$$A_\lambda \times A_0^+ = A_\lambda \quad \text{(This is true also for } \lambda = 0^-\text{)}.$$

For $\lambda = 0^-$ and $\lambda > 0$, $u_\lambda v_0$ and $u_{-\lambda} v_0$ transform exactly like u_λ and $u_{-\lambda}$ — that is, according to $A_{-\lambda}$. That yields

$$A_\lambda \times A_0^- = A_\lambda \qquad (\lambda > 0).$$

Finally, if $\lambda = \mu = 0^-$, then the product $u_0 v_0$ is invariant under the rotations D_φ and the reflection s_y; hence,

$$A_0^- \times A_0^- = A_0^+$$

13. COMMUTATION OF MATRICES WITH REPRESENTATIONS.

Let R and S be two vector spaces with a common operator domain G; the operators of G are such that they give rise to linear transformations in both spaces. Furthermore, let T be a linear transformation that maps the space R operator-homomorphically upon S or upon a sub-space of S. The expression "operator-homomorphically" means that if $Tv = w$, then the transformation T also carries αv into αw for every α of G, i.e.

$$T\alpha v = \alpha Tv$$

or a commutes with T. Now, the lemma due to Schur states.

If R is irreducible, then T is either an isomorphism or it maps every vector upon the null vector.

In the first case, R is equivalent to an irreducible sub-space of S. If S itself is irreducible then it follows that R is equivalent to S. In the special case where R = S and where we employ for R and S the same base vectors, we further assert: the matrix T is a multiple of the unit matrix.

Proof. According to the homomorphism rule, T is an isomorphic mapping of a factor space R/r. If R is irreducible, then, of necessity, r = (0) or R = r. The alternative in Schur's lemma follows immediately from here.

In order to prove the assertion $T = \tau E$ for the case R = S, we determine τ in such a manner that the determinant $|T - \tau E| = 0$. Since $T - \tau E$ commutes with all a just like T does, then, according to the already proven part of the rule, the matrix $T - \tau E$ either represents a single valued (that is, non-singular) transformation or it must be equal to zero. Thus, if $|T - \tau E| = 0$, then $T - \tau E = 0$ or $T = \tau E$.

The statement $T = \tau E$ is, of course, also valid for the case where $R \cong S$ (and not R = S) provided, however, the bases chosen for R and S bear an isomorphic correspondence to each other.

We wish to determine the linear transformations which commute with a completely reducible system G of linear transformations of a space R. Or, what amounts to the same (according to the remarks made above), we wish to determine the operator-homomorphisms of a completely reducible representation-space R having the operator

domain G.

$$\text{Let } R = r_1 + r_2 + \ldots + r_r. \tag{13.1}$$

When spaces such as $r_1, \ldots, r_r$ are transformed equivalently by G, one introduces into them suitable base vectors in such a manner that the transformations of these spaces are represented by the same matrices. Let therefore T be a linear transformation that maps R operator-homomorphically on to itself. In order to identify the transformation completely, one need merely know its effect on the vectors of $r_1, r_2, \ldots, r_r$. T maps r_1 on to the space Tr_1 that is isomorphic with r_1. The vectors $w = Tv$ of Tr_1 may again be broken up into the components:

$$Tv = w = w_1 + w_2 + w_3 + \ldots + w_r \ , \tag{13.2}$$

in accordance with (13.1). The association $w \to w_1$ or $w \to w_2$ is again an operator-homomorphism; hence, $v \to w \to w_1$ or $v \to w \to w_2$ etc. is an operator-homomorphism, too. Thus, according to the lemma of Schur, only those components can occur in the decomposition (13.2) which belong to the subspaces $r_1, \ldots, r_k$ that are equivalent to r_1; all remaining components must be zero. Moreover, according to the lemma of Schur (second part), the associations $v \to w_1$ etc. must be represented by mutliples of the unit matrix. When dealing with the mapping of r_1 on to r_λ we denote these multiples by $\tau_{\lambda 1} E$. Whatever is valid for r_1 is, of course, also valid for all the other r_μ. There are, for example, the mappings $\tau_{\lambda\mu} E$ of r_μ on to the equivalent r_λ.

If we now form the matrix of T with respect to a basis that is

constructed from the bases of $r_1,\ldots,r_k$, $r_{k+1},\ldots,r_r$, we then obtain

$$\left(\begin{array}{cc|c} \tau_{11}^E\ \tau_{12}^E\ldots\ \tau_{1k}^E & & \\ \vdots & & 0 \\ \tau_{k1}^E\ \tau_{k2}^E\ldots\ \tau_{kk}^E & & \\ \hline & \tau_{k+1k+1}^E\ldots & \\ 0 & \vdots & \\ & \ldots\ldots\ldots & \end{array}\right) \qquad (13.3)$$

The same result can also be formulated in another manner. First let us write down the base vectors of the equivalent spaces r_1 to r_k one beneath the other:

$v_{11}, v_{12},\ldots,v_{1n}$ (base of r_1)
$v_{21}, v_{22},\ldots,v_{2n}$ (base of r_2)
$v_{k1}, v_{k2},\ldots,v_{kn}$ (base of r_k) .

Now, the operations in G transform the <u>rows</u> of this rectangle linearly into themselves; in particular, all the rows are transformed in like manner. The operation T, however, (which commutes with G) transforms the <u>columns</u> of the same rectangle; and again all columns are transformed in the same, though arbitrary manner. Similar rectangles of base vectors are obtained for r_{k+1} to r_r.

In this manner we obtain all the matrices that commute with a completely reducible system. These matrices form a ring $\mathfrak{T}$, i.e. a system of quantities which contains for every pair of elements also their sum, difference and product. The ring $\mathfrak{T}$ is a "direct sum" of the rings $\mathfrak{T}_1, \ldots \mathfrak{T}_\varphi$ consisting of the matrices of a single block of (13.3) (with zeros in the other blocks); in other words, each matrix of the ring $\mathfrak{T}$ can be uniquely represented by a sum of matrices from the rings $\mathfrak{T}_1, \ldots, \mathfrak{T}_\varphi$. In addition, the products of matrices of two different rings $\mathfrak{T}_\nu$, $\mathfrak{T}_\mu$ is always zero. We write

$$\mathfrak{T} = \mathfrak{T}_1 + \ldots + \mathfrak{T}_\varphi .$$

The matrices of the ring $\mathfrak{T}_1$ add and multiply exactly like k-dimensional matrices

$$\begin{pmatrix} \tau_{11} & \tau_{12} \cdots & \tau_{1k} \\ \vdots & \vdots & \vdots \\ \tau_{k1} & \tau_{2k} \cdots & \tau_{kk} \end{pmatrix} \qquad (13.4)$$

having as elements the entirely arbitrary numbers $\tau_{\lambda\mu}$. The ring of all these matrices is called the <u>complete matrix-ring of order k</u>. Thus, the ring $\mathfrak{T}$ is a direct sum of complete matrix-rings.

By setting all the $\tau_{\lambda\mu}$ in (13.4) equal to zero, with the exception of a single $\tau_{\lambda\mu} = 1$, we obtain the basis quantities of a complete matrix-ring. The matrix thus obtained (13.4) is denoted by $C_{\lambda\mu}$. Consequently, each matrix can be uniquely represented by a sum $\sum C_{\lambda\mu}\tau_{\lambda\mu}$. The rules for calculating the $C_{\lambda\mu}$ are:

$$C_{\varkappa\lambda}\ C_{\lambda\mu} = C_{\varkappa\mu}\ ,$$

$$C_{\varkappa\lambda}\ C_{\lambda'\mu} = 0 \qquad (\lambda \neq \lambda')\ .$$

The laws set up above also permit one to answer the following question: Let there be given a vector space R operated upon by two commutative groups of linear transformations - or more generally speaking, by two commutative, completely reducible systems G, H of linear transformations. The question is: Is it possible to reduce the system in such a manner that the commutativeness in the reduced form meets the eye immediately?

Well, we first reduce out the system G; that leads to the above rectangular systems of base vectors

$$\begin{matrix} v_{11} v_{12} \cdots v_{1n} \\ \cdots\cdots\cdots \\ v_{k1} v_{k2} \cdots v_{kn} \end{matrix}\ ,$$

According to the result above, the columns of the individual rectangles are linearly transformed into themselves by the commutative system H; in fact, all columns are transformed in like manner. Thus, each column defines a subspace of R invariant to H; as such, it must again be completely reducible. One can, therefore, replace the base vectors in a column by suitable linear combinations so that the column breaks apart into individual segments which are transformed by H in an irreducible manner. In like manner we carry out this modification of the base vectors for all columns of a rectangle. As before, the rows of the rectangle are likewise transformed in an irreducible manner by the

system G. Consequently, by slicing the rectangle horizontally, it can be entirely broken apart into subsidiary rectangles, the columns of which are transformed in an irreducible manner by the system H.

Thus, if we are given two commutative systems G, H of linear transformations of a vector space R, both of which are completely reducible, then we may arrange the base vectors in the form of rectangles

$$\begin{matrix} v_{11} \cdots v_{1n} \\ \cdot \quad\quad \cdot \\ \cdot \quad\quad \cdot \\ \cdot \quad\quad \cdot \\ v_{s1} \cdots v_{sn} \end{matrix}$$

in such a manner that all the rows in each rectangle are irreducibly transformed by the system G in the same way, likewise, all the columns are irreducibly transformed by H in the same way.

14. REPRESENTATIONS OF A FINITE GROUP.[1)]

Let G be a finite group with h elements. In the space of any arbitrary representation let us choose a positive-definite

---

1) The contents of this section and the next is not absolutely necessary for the quantum-mechanical applications treated in this book. Nevertheless, for those who wish to delve deeper into the representation theory, this section is indispensable. The theory is due to G. Frobenius and the methods of proof employed here are due to E. Noether. For another simple proof see I. Schur. Sitzungsberichte, Berlin 1905, p.406.

Hermitian form, upon which we shall carry out all the transformations of the group, and add. That yields a positive-definite form which is invariant under the group. Thus, the matrices of the representation are unitary, and, the representation is always either irreducible or completely reducible.

A special representation is obtained when one takes the elements of the group as base vectors, i.e., we take as vectors all linear combinations

$$c = \sum_s \gamma_s s \tag{14.1}$$

with complex coefficients γ_s. These "group numbers" (14.1) form a ring R_g; not only can they be added (as well as multiplied) by numbers but they can also be multiplied by one another:

$$\sum_s \gamma_s s \cdot \sum_t \delta_t t = \sum_s \sum_t \gamma_s \delta_t s t.$$

This ring R_g is known as the group ring or the group algebra or the group domain. The identity element e of the group is at the same time also the identity element of the ring. Upon multiplying the elements of the ring by the base vectors s, we obtain each time a linear transformation of the ring into itself - hence, a representation of the group G. This representation is known as the regular representation of the group G; it is of the order h.

The invariant subspaces of the regular representation are such that for every group number a they also contain all the sa, where s is any arbitrary element of the group. The subspace then also contains all the $\sum_s \gamma_s s \cdot a$; that is, it

contains all ca, where c is a group number. Subspaces of this type are called left-ideals. According to what has been said above, the group ring is completely reducible - that is, it is a direct sum of irreducible left ideals.

We now present the proofs for the following rules:

Rule 1. Every irreducible representation of G is contained in the regular representation; (that is, it is equivalent to the representation generated by an irreducible left ideal).

Proof: First of all, we note that each representation of the group G can be extended to a representation of the ring. To do that, we let the matrix S correspond to the group element s and the matrix $\sum_s \gamma_s S$ to the ring element $\sum_s \gamma_s s$. The product and sum of the matrices then correspond respectively to the product and sum of two ring elements. Now, if v is any arbitrary vector of the representation space r, then $c \longrightarrow cv$ is a linear mapping of the group ring on to the representation space. This mapping constitutes an operator-homomorphism, (with respect to the operator domain G), for from (14.1) we have:

$$s \cdot c \longrightarrow sc \cdot v = s \cdot cv \ .$$

Hence, according to #11, Rule 4, the space r is isomorphic with a sum of irreducible subspaces of the regular representation — or, if r itself is irreducible, it is isomorphic with a single such irreducible subspace, Q.E.D.

Rule 2. The ring R_g is a direct sum of complete matrix rings.

Proof: We wish to determine the operator homomorphisms of

the ring R_g (or the transformations that commute with the regular representation). If T is the desired transformation which also carries over the identity element of the group G into an element t, then, since T commutes with all group elements s, we have:

$$T \sum_s c_s se = \sum_s c_s sTe = \sum_s c_s st.$$

We see, therefore, that the operation T amounts to a multiplication on the right of all ring elements by t. To every T there corresponds exactly one t and conversely. Corresponding to the product TU of two homomorphisms there is the reversed product ut, for we have (for any arbitrary c in R_g):

$$TU \cdot c = Tcu = cut.$$

Also, corresponding to the sum T + U we have the sum t + u. Thus, the ring R_g is "inversely isomorphic" with the ring $\mathfrak{J}$ of the operator homomorphisms - that is. it is isomorphic with an inversion of the products. According to #13, $\mathfrak{J}$ is a direct sum of complete matrix rings; Now, in order to find the ring that is inversely isomorphic with $\mathfrak{J}$ one need only transpose all the matrices. That again leads to a direct sum of complete matrix rings.

Now, which are the left-ideals in a ring

$$R_g = \mathfrak{J}_1 + \mathfrak{J}_2 + \cdots + \mathfrak{J}_g , \qquad (14.3)$$

where $\mathfrak{J}_1, \ldots, \mathfrak{J}_g$ are matrix rings?

We introduce as basis quantities in $\mathfrak{J}_1$ the n matrices $C_{\lambda\mu}$ (cf. #13). The elements $(C_{11}, C_{21}, \ldots, C_{n1})$ give rise to a left-ideal in $\mathfrak{J}_1$ and therefore also in R_g; the $(C_{12}, C_{22}, \ldots, C_{n2})$

likewise yield a left-ideal in $\mathfrak{J}_1$ etc. In this manner we obtain n left-ideals in $\mathfrak{J}_1$. Upon calculating which representation of R_g these ideals generate, we find that all quantities from $\mathfrak{J}_2, \ldots, \mathfrak{J}_f$ are represented by zero; also, the quantities $t = \sum \sum \alpha_{\lambda\mu} C_{\lambda\mu}$ from $\mathfrak{J}_1$ are represented exactly by the matrix $(\alpha_{\lambda\mu})$ in all the n representations. We have namely for each t_2 from $\mathfrak{J}_2$:

$$t_2 C_{\nu\varkappa} = 0$$

and for each $t_1 = \sum \sum \alpha_{\lambda\mu} C_{\lambda\mu}$ from $\mathfrak{J}_1$:

$$t_1 C_{\nu\varkappa} = \sum_{\lambda} \sum_{\mu} \alpha_{\lambda\mu} C_{\lambda\mu} C_{\nu\varkappa} = \sum_{\lambda} \alpha_{\lambda\nu} C_{\lambda\nu} C_{\nu\varkappa} = \sum_{\lambda} C_{\lambda\varkappa} \alpha_{\lambda\nu} .$$

From here it follows that the above left-ideals are equivalent and generate the same irreducible representation. The same is also true of the left-ideals of $\mathfrak{J}_2$. The latter, however, are not equivalent to the former; for in the representation generated by the left-ideals of $\mathfrak{J}_2$ the elements of $\mathfrak{J}_1$ are represented by zero. This is not the case, however, in the first-mentioned representation. Thus, we obtain just as many inequivalent representations as there are matrix rings in (14,3). If n_ν is the order of the matrices of $\mathfrak{J}_\nu$, then the representation Δ_ν formed by these matrices is likewise of order n_ν . Also, since $\mathfrak{J}_\nu$ breaks apart into n_ν equivalent left-ideals, the representation Δ_ν occurs in the regular representation n_ν -times.

Thus, the order of each irreducible representation indicates how often it occurs in the regular representation. The dimension of $\mathfrak{J}_\nu$ (i.e. the number of linearly independent base-vectors $C_{\lambda\mu}$) is n_ν^2 ; hence, the dimension h of R_g is given by:

$$h = \sum_{\nu=1}^{t} n_\nu^2 \quad . \tag{14.4}$$

Following is "Burnside's Theorem":

Each irreducible representation of order n_ν contains n_ν^2 linearly independent matrices.

That is so, for among the linear combinations of the matrices of the representation Δ_ν, there occur the representation matrices of all the elements of the ring R and, in particular, of the ring $\mathfrak{J}_\nu$; these representation matrices constitute all the possible matrices $(\alpha_{\lambda\mu})$ with arbitrary $\alpha_{\lambda\mu}$.

Finally, we wish to determine how many irreducible representations a given group can have. For this purpose we first determine the "central" of the ring R_g. The central constitutes the totality of those quantities $\sum \gamma_s s$ that commute with all other group numbers. In fact, they need but commute with all the group elements, that is:

$$\sum \gamma_s tst^{-1} = \sum \gamma_s s \quad .$$

In order that this be so, it is necessary and sufficient that in the sum $\sum \gamma_s s$ each s as well as each tst^{-1} (the group element conjugate to s) have the same coefficient. If we denote by k the sum of all the different group elements tst^{-1} conjugate to s, including s itself, then each element $\sum \gamma_s s$ of the

central must be of the form:

$$z = \sum \gamma_k k \quad . \tag{14.5}$$

Thus, the central of R_g is a vector space whose dimensionality q' is equal to the number of classes of conjugate group elements — for, according to (14.5), the sums k of the classes span the central.

We can also determine the central with the aid of the sum-representation (14.3). A quantity $t = t_1 + \ldots + t_q$ of R_g commutes with all quantities of R_g if t_1 commutes with all the matrices of $\mathfrak{J}_1$ — that is, if t_1 is a multiple $\lambda_1 e_1$ of the unit matrix e_1 of $\mathfrak{J}_1$ and if, likewise, $t_2 = \lambda_2 e_2, \ldots$ $\ldots, t_q = \lambda_q e_q$. Thus, the central of R_g is spanned by the linearly independent quantities $(e_1, \ldots, e_q)$; its dimensionality is therefore, q. Hence we have

$$q' = q$$

or. <u>the number of irreducible representations is equal to the number of classes of conjugate group elements</u>.

EXAMPLES

1. The Symmetrical Group S_3.

Number of elements: $3! = 6$. Classes of conjugate group elements: the class of (1), that of (12) and that of (123); hence, there are three representations. These are already known to us from #10, example 5. The orders of the representations are 1, 1, and 2, respectively. In fact, we have

$$6 = 1^2 + 1^2 + 2^2$$

The two first order representations are the symmetrical and anti-symmetrical ones. We define (in addition to two linearly independent base vectors e_1, e_2) a third vector e_3 by means of the relation $e_3 = -e_1 - e_2$ or $e_1 + e_2 + e_3 = 0$. We now obtain the second order representation simply by subjecting these three vectors to the permutations of S_3. Obviously, this representation is faithful.

2. The Symmetrical Group S_4.

Number of elements: $4! = 24$.

Classes: those of (1), (12), (123), (12)(34), and (1234); hence, there are five representations. The factor group S_4/V_4 is $\cong S_3$; according to example 1, therefore, it has two first order representations and one second order representation. So far then, we have three unfaithful representations of S_4 (of orders 1, 1, and 2). From

$$24 = 1^2 + 1^2 + 2^2 + n_4^2 + n_5^2$$

we obtain:

$$n_4^2 + n_5^2 = 18;$$

hence, $n_4 = n_5 = 3$. Of the four vectors e_1, e_2, e_3, e_4, the first three are linearly independent, whereas the fourth one is defined by $e_1 + e_2 + e_3 + e_4 = 0$. Now, by subjecting these four vectors to the permutations of S_4, we obtain one of the two third-order representations. And if in the representation just obtained, we reverse the signs of the matrices (which represent odd permutations) then we shall find the remaining third order repre-

sentation.

3. The Alternating Group A_4.

Number of elements: 12.

Classes: those of (1), (123), (132) and (12)(34); hence, there are four representations. The factor group A_4/V_4 is cyclic and of order three; hence, it has three first-order representations (with third-order roots of one). From

$$12 = 1^2 + 1^2 + 1^2 + n_4^2$$

we have

$$n_4 = 3.$$

The one remaining representation of order three is the same as the above two third-order representations of S_4 employed, of course, in connection with the permutations of A_4. The above second order representation of S_4 applied to A_4 breaks apart into two complex-conjugate representations of the first order.

GENERALIZATIONS.

The Burnside theorem is valid not only for the representations of finite groups but also for every irreducible system of matrices which for any two matrices contains also their product. Furthermore, the following generalization due to Frobenius and Schur is valid. A completely reducible system of matrices whose irreducible components (the equivalent ones being counted just once) are of the orders $n_1, n_2, \ldots, n_s$ contains exactly $n_1^2 + n_2^2 + \ldots + n_s^2$ linearly independent matrices provided the system contains for any two matrices also their product.

The group ring R_g is an example of a "hypercomplex number system". This is a vector space of finite dimensions which is made into a ring by defining in it a type of multiplication that need not be commutative, but otherwise must possess all the usual properties of a multiplication (including the law of association). We consider here only those hypercomplex number systems which have as their multiplicative domain the domain of the complex numbers.

Clearly, the theorems of this section which refer to the representations of the ring R_g have validity not only for the group ring but also for every hyper complex number system which is

1. completely reducible (i.e. it is a sum of irreducible left ideals) and

2. which contains an identity element.

Accordingly, every such system is a direct sum of complete matrix rings and consists of as many irreducible representations as there are matrix rings in its decomposition. Moreover, it can be shown that each reducible representation of a ring of the above type breaks apart completely into irreducible representations. In particular, a single complete matrix ring has only one irreducible representation; in fact, the latter is generated by the matrices of the ring itself.

Occasionally, one finds an application of this theorem to quantum mechanical problems. For example, since the days of Dirac one often poses the problem of determining a system of 4 matrices Γ_1, Γ_2, Γ_3, Γ_4 which satisfies the following

relations:

$$\Gamma_\lambda^2 = 1, \quad \Gamma_\lambda \Gamma_\mu = -\Gamma_\mu \Gamma_\lambda \qquad (\lambda \neq \mu). \tag{14.6}$$

If in addition to the Γ_λ we add yet their products in batches of two, three, etc., then with the aid of the relations (14.6) we can express all these through the 16 following entities:

$$1, \ \Gamma_\lambda, \ \Gamma_\lambda \Gamma_\mu, \ \Gamma_\lambda \Gamma_\mu \Gamma_\nu, \ \Gamma_1 \Gamma_2 \Gamma_3 \Gamma_4 \tag{14.7}$$
$$(\lambda, \mu, \nu = 1, 2, 3, 4; \ \lambda < \mu < \nu).$$

Now, if the matrices are not known,- in fact, if they are being sought,- then we next form a hypercomplex number system having 16 basis elements:

$$1, \ \gamma_\lambda, \ \gamma_{\lambda\mu}, \ \gamma_{\lambda\mu\nu}, \ \gamma_{1234} \qquad (\lambda, \mu, \nu = 1,2,3,4 \ \ \lambda<\mu<\nu). \tag{14.8}$$

We shall agree, also, that these basis elements are multiplied by one another exactly like the matrices (14.7) by means of the calculation rules (14.6). By means of this agreement the number system is uniquely determined. Thus, each system of matrices (14.7) having the properties (14.6) gives rise to a representation of the hypercomplex system and conversely. The problem above, therefore, has been effectively reduced to the problem of finding the representations of a hyper-complex system by matrices.

Well, according to Dirac[1] we know there exists a representation by 4-row matrices where the basis elements (14.8) are represented by 16 linearly independent matrices. Hence, the

---

1) Dirac, P.A.M. Proc. Roy. Soc. London, A vol. 117 (1928) p. 610.

hypercomplex system, in question, is isomorphic with the complete matrix ring of all the four-row matrices. According to the above theorems it follows from here that (barring equivalence) there exists but one irreducible representation and that of order 4. Also, every reducible representation breaks apart completely into irreducible ones, all of which are equivalent to the fourth-order representation. We conclude that (barring equivalence as well as entirely trivial repetitions) the only possible representation is the Dirac representation.

15. THE CHARACTERS

As is known, the traces $\sum_{\lambda} a_{\lambda\lambda}$ of the matrices $(a_{\lambda\mu})$ of a representation are invariants. The trace of the matrix that corresponds to the group element t in the representation D is denoted by $S(t)$ or $S_D(t)$. The traces of the irreducible representations are called characters.

Any two conjugate elements a and $b^{-1}ab$ of a group have the same traces, since

$$S(B^{-1}AB) = S(A).$$

Thus, the traces and characters are "class functions"; for every class of conjugate elements of a group they have a distinct value.

The traces and characters constitute a very useful tool often employed in breaking apart a given representation into irreducible ones. This decomposition is brought about with the aid of the "orthogonality relations" which we shall now derive.

Let

$$s \to A(s), \qquad s \to B(s)$$

be two irreducible representations of a finite group G. If C is any arbitrary matrix which maps the space of the second representation on to that of the first, then

$$P = \sum_t A(t)CB(t^{-1})$$

(summation over all elements of the group) is again a mapping of the second space on to the first, P commuting with all elements s of the group:

$$A(s)P = A(s)\sum_t A(t)CB(t^{-1}) = \sum_t A(st)CB(t^{-1}s^{-1})B(s)$$
$$= PB(s).$$

From Schur's lemma (# 13) it follows that:

P = 0 if the representations A(t), B(t) are non-equivalent

$P = \beta E$ if the representations are equal.

On writing that out in full, we get:

$$\sum_\lambda \sum_\mu \sum_t a_{\varkappa\lambda}(t)c_{\lambda\mu}b_{\mu\nu}(t^{-1}) = \begin{cases} 0 & \text{if A(s) and B(s) are non-equivalent} \\ \beta\delta_{\varkappa\nu} & \text{if A(s) = B(s),} \end{cases}$$

or since the $c_{\lambda\mu}$ are entirely arbitrary:

$$\sum_t a_{\varkappa\lambda}(t)\, b_{\mu\nu}(t^{-1}) = \begin{cases} 0 & \text{if A(s), B(s) are non-equivalent} \\ \beta_{\lambda\nu}\delta_{\varkappa\nu} & \text{for A(s) = B(s).} \end{cases} \qquad (15.1)$$

In order to determine the $\beta_{\lambda\mu}$ for the case A = B, we set $\varkappa = \nu$ and sum over ν. Since $B(s^{-1})A(s) = A(s^{-1})A(s) = 1$,

$a_{\varkappa\lambda}(t)\, b_{\mu\varkappa}(t^{-1})$ on the lefthand side of (15.1) becomes $\delta_{\lambda\mu}$; thus,

$$\sum_t \delta_{\lambda\mu} = \beta_{\lambda\mu} \sum_\nu \delta_{\nu\nu} \quad .$$

If h is the number of group elements and n the order of the representation, we obtain :

$$h\, \delta_{\lambda\mu} = n\, \beta_{\lambda\mu} \quad .$$

Hence (15.1) becomes:

$$a_{\varkappa\lambda}(t)\, b_{\mu\nu}(t^{-1}) = \begin{cases} 0 \ , & \text{if A,B are non-equivalent} \\ \frac{h}{n}\, \delta_{\lambda\mu}\, \delta_{\varkappa\mu} & \text{for A = B.} \end{cases}$$

For a unitary representation B(s),

$$B(t^{-1}) = \tilde{B}(t), \quad \text{consequently,}$$

$$b_{\mu\nu}(t^{-1}) = \bar{b}_{\nu\mu}(t), \quad \text{and}$$

$$\sum_t a_{\varkappa\lambda}(t)\, \bar{b}_{\nu\mu}(t) = \begin{cases} 0 & \text{for A, B non-equivalent} \\ \frac{h}{n}\, \delta_{\varkappa\nu}\, \delta_{\lambda\mu} & \text{for A = B.} \end{cases} \qquad (15.2)$$

These are the orthogonality relations for the matrix elements. Upon setting $\varkappa = \lambda$, $\nu = \mu$ and summing over λ and μ, we obtain the <u>orthogonality relations for the characters</u>:

$$\sum_t \chi^{(1)}(t)\, \overline{\chi^{(2)}(t)} = \begin{cases} 0, \\ h. \end{cases} \qquad (15.3)$$

The sum is equal to zero for the characters of non-equivalent representations and equal to h for the characters of equivalent representations.

If $\chi^{(1)}, \ldots, \chi^{(r)}$ are the characters of the various non-equivalent representations and

$$S(t) = \sum_t c_\lambda \chi^{(\lambda)}(t)$$

is the trace of any particular representation (which, in turn, contains the representation having the number λ exactly c_λ - times), then by (15.3) we have:

$$\sum_t S(t)\, \bar{\chi}^{(\lambda)}(t) = c_\lambda h. \qquad (15.4)$$

With the aid of this relation we can now calculate the numbers c_λ from the trace of the given representation and the characters of the irreducible representations. We notice also that, barring equivalence, the representation is uniquely determined by the traces S(t).

The relations (15.4) are especially suitable when one seeks to reduce out a product representation $D_\lambda \times D_\mu$. The trace of a product matrix $A \times B$ is

$$\sum_\lambda \sum_\mu a_{\lambda\lambda} b_{\mu\mu} = \Big(\sum_\lambda a_{\lambda\lambda}\Big)\Big(\sum_\mu b_{\mu\mu}\Big) = S(A)S(B);$$

hence, the trace of a product representation $D_\lambda \times D_\mu$ is the product of the traces of the factor representations.

If, for example, we denote the three representations of the S_3 by $\mathfrak{J}$ (identical rep.), A (antisymmetrical rep.), U(second order representation) then by this method we obtain:

$$\mathfrak{f} \times \mathfrak{f} = \mathfrak{f} \qquad A \times A = \mathfrak{f}$$

$$\mathfrak{f} \times A = A \qquad A \times U = U$$

$$\mathfrak{f} \times U = U$$

$$U \times U = \mathfrak{f} + A + U.$$

III THE ROTATION GROUP AND THE LORENTZ GROUP.

16. The Linear Group C_2, the Unitary Group U_2 and its Relation to the Rotation Group $\mathfrak{d}_3$.

We take for the vector space all the binary linear forms $c_1u_1 + c_2u_2$ in two variables u_1, u_2. The transformations A of the special linear group C_2 send the base vectors u_1, u_2 into

$$u_1' = u_1\alpha + u_2\gamma$$
$$u_2' = u_1\beta + u_2\delta.$$

The corresponding matrices are

$$A = \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix}; \quad \alpha\delta - \beta\gamma = 1$$

Clearly, the inverse transformation is

$$A^{-1} = \begin{pmatrix} \delta & -\beta \\ -\gamma & \alpha \end{pmatrix}.$$

Upon choosing the identity as the basic Hermetian form the unitary A have, according to #7, the following property:

$$A = A^{-1} \quad \text{or} \quad \begin{pmatrix} \bar{\alpha} & \bar{\gamma} \\ \bar{\beta} & \bar{\delta} \end{pmatrix} = \begin{pmatrix} \delta & -\beta \\ -\gamma & \alpha \end{pmatrix}.$$

Consequently, $\bar{\alpha} = \delta$, $\bar{\beta} = -\gamma$ and the special unitary group U_2 consists of the transformations:

$$\left.\begin{aligned} u_1' &= u_1\alpha - u_2\beta \\ u_2' &= u_1\beta + u_2\bar{\alpha} \end{aligned}\right\} \quad \text{or} \quad \left.\begin{aligned} c_1' &= \alpha c_1 + \beta c_2 \\ c_2' &= -\bar{\beta} c_1 + \bar{\alpha} c_2 \end{aligned}\right\} \tag{16.1}$$

where $\alpha\bar{\alpha} + \beta\bar{\beta} = 1$.

Let us consider a transformation with determinant 1 which transforms the vector coefficients $(d_1 d_2)$ co-variant to (c_1, c_2). We note that every such transformation leaves the expression $c_2 d_1 - c_1 d_2$ invariant. As a result, the coefficients $(c_2, -c_1)$ of this expression transform contravariant to (c_1, c_2). Moreover, in the case of the unitary group U_2 the expression $\bar{c}_1 c_1 + \bar{c}_2 c_2$, is invariant; consequently, $(\bar{c}_1, \bar{c}_2)$, too, transform contravariant to (c_1, c_2). Finally, in this latter case the $(\bar{c}_2, -\bar{c}_1)$ transform contravariant to $(c_2, -c_1)$ and co-variant to (c_1, c_2).

Let us choose as base vectors the products of degree v :

$$u_1^v,\ u_1^{v-1}u_2, \ldots, u_2^v\ , \tag{16.2}$$

or, what amounts to the same, let us erect the space of all forms

$$c_0 u_1^v + c_1 u_1^{v-1}u_2 + \ldots + c_v u_2^v.$$

Clearly, these products are linearly transformed by the transformations A, for A carries $u_1^r u_2^{v-r}$ over into

$$u_1'^r u_2'^{v-r} = (u_1\alpha + u_2\gamma)^r\ (u_1\beta + u_2\delta)^{v-r}.$$

The latter, of course, is a linear combination of the products (16.2); hence, we now have a representation of the groups C_2 and U_2.

The representation of C_2 and U_2 thus obtained, is denoted by D_J, where J is set equal to $\frac{1}{2}$ v. (The reason is based mainly upon spectroscopic applications). In particular, D_0 is the identical representation of the first order (where the only base vector is invariant under all transformations of the group). $D_{\frac{1}{2}}$ is the

representation of C_2 through itself. The representation D_J is of the order $v + 1 = 2J + 1$.

One of the proper ties of the representation D_1 in the space of the polynomials

$$c_0u_1^2 + c_1u_1u_2 + c_2u_2^2$$

is that the "discriminant"

$$c_1^2 - 4c_0c_2$$

is invariant under it. That is easily seen when one replaces the c_0, c_1, c_2 by new variables through the following substitutions:

$$\left.\begin{aligned} x &= -c_0 + c_2 \\ y &= -i(c_0 + c_2) \\ z &= c_1 \end{aligned}\right\} \qquad \left.\begin{aligned} x+iy &= 2c_2 \\ x-iy &= -2c_0 \\ z &= c_1 \end{aligned}\right\} \qquad (16.3)$$

From here we obtain:

$$x^2 + y^2 + z^2 = (x + iy)(x - iy) + z^2 = c_1^2 - 4c_0c_2.$$

Hence, the form $x^2 + y^2 + z^2$ is invariant under the transformations of D_1. The transformations are (complex) <u>rotations</u>.[1)]

Let us investigate the conditions for real behavior. We note that the coefficients c_0, c_1, c_2 of any quadratic form transform exactly like the coefficients a_1b_1, $a_1b_2 + a_2b_1$, a_2b_2 of the special form $(a_1u_1 + a_2u_2)(b_1u_1 + b_2u_2)$. In a unitary transformation (16.1) from U_2, however, b_1, b_2 transform just like $-\bar{a}_2, \bar{a}_1$; hence, the c transform like $-a_1\bar{a}_2$, $a_1\bar{a}_1 - a_2\bar{a}_2$, $a_2\bar{a}_1$.

---

1) They could not possibly be reflections inasmuch as they can go over into the identity ($\alpha = 1$, $\beta = 0$) continuously.

Thus, the x, y, z (of equation (16.3)) transform like

$$a_1\bar{a}_2 + a_2\bar{a}_1; \quad i(a_1\bar{a}_2 - a_2\bar{a}_1); \quad a_1\bar{a}_1 - a_2\bar{a}_2.$$

Now, these three numbers are consistently real and are always transformed into real numbers; hence, the transformation coefficients must also be real, i.e.

The vectors (x,y,z) always undergo real rotations under the representation D_1 of the group U_2.

The real group of rotations of the space is denoted by δ.

Clearly, the representation D_1 contains all real rotations of the space. This can be shown by calculating for the special unitary transformations

$$B(\beta) = \begin{pmatrix} \cos\beta & -\sin\beta \\ +\sin\beta & \cos\beta \end{pmatrix} \qquad C(\gamma) = \begin{pmatrix} e^{-i\gamma} & 0 \\ 0 & e^{+i\gamma} \end{pmatrix} \tag{16.4}$$

the associated rotations in the representations D_1. We then obtain

$$\left.\begin{aligned} B(\beta): &\begin{cases} x' = x\cos 2\beta + z\sin 2\beta \\ y' = y \\ z' = -x\sin 2\beta + z\cos 2\beta \end{cases} \\ C(\gamma): &\begin{cases} x' = x\cos 2\gamma - y\sin 2\gamma \\ y' = x\sin 2\gamma + y\cos 2\gamma \\ z' = z. \end{cases} \end{aligned}\right\} \tag{16.5}$$

These are rotations about the y- and z-axes by the angles of rotation 2β and 2γ respectively. From them we can form all other possible rotations. Now, a rotation by the Eulerian angles ϑ, φ, ψ is none other than the product $Z_\varphi Y_\vartheta Z_\psi$ of the rotations about the z-, y-, and z- axes by the angles of rotation φ, ϑ, ψ respectively.

According to the above calculation, the matrices $C(\frac{1}{2}\varphi)$, $B(\frac{1}{2}\vartheta)$, $C(\frac{1}{2}\psi)$ give rise to the rotations Z_φ, Y_ϑ, Z_ψ. Now, upon multiplication of these matrices by one another, we obtain the following explicit formula for the matrix of the transformation from U_2 (which gives rise to a rotation by given Eulerian angles):

$$C(\tfrac{1}{2}\varphi)\, B(\tfrac{1}{2}\vartheta)\, C(\tfrac{1}{2}\psi)$$

$$= \begin{pmatrix} e^{-\frac{1}{2}i\varphi} & 0 \\ 0 & e^{+\frac{1}{2}i\varphi} \end{pmatrix} \begin{pmatrix} \cos\frac{1}{2}\vartheta & -\sin\frac{1}{2}\vartheta \\ \sin\frac{1}{2}\vartheta & \cos\frac{1}{2}\vartheta \end{pmatrix} \begin{pmatrix} e^{-\frac{1}{2}i\psi} & 0 \\ 0 & e^{+\frac{1}{2}i\psi} \end{pmatrix}$$

$$= \begin{pmatrix} e^{-\frac{1}{2}i(\varphi+\psi)}\cos\frac{1}{2}\vartheta & -e^{-\frac{1}{2}i(\varphi-\psi)}\sin\frac{1}{2}\vartheta \\ e^{+\frac{1}{2}i(\varphi-\psi)}\sin\frac{1}{2}\vartheta & e^{+\frac{1}{2}i(\varphi+\psi)}\cos\frac{1}{2}\vartheta \end{pmatrix} = \begin{pmatrix} \alpha & -\bar{\beta} \\ \beta & \bar{\alpha} \end{pmatrix}. \qquad (16.6)$$

We now determine the faithfulness of the representation. According to the homomorphism rule (#8) one need only determine which transformations of U_2 give rise to the identity in the representation D_1. These transformations must leave the products u_1^2, u_1u_2 and u_2^2 invariant and this obviously is accomplished only by the two transformations:

$$E = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \qquad \text{and} \qquad -E = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}.$$

The above two transformations constitute the subgroup h referred to in the homomorphism rule. The cosets of h consist always of two transformations A and -A; hence, they give rise steadily to the same rotation. Consequently, the representation is not faithful. If in the group C_2 we limit ourselves, however, to that region about the unit matrix E which contains only one transformation from each coset,

then the representation is faithful within this region. Consequently, it is uniquely reversible, for to every rotation D by a sufficiently small angle of rotation there corresponds a single unitary transformation A in the vicinity of the identity. The matrix A varies continuously if the corresponding rotation D does the same (as can be seen, for instance, from the explicit equation (16.6)); but when the rotation D operating on the rotation group traverses a closed path (of 2π radians), the associated matrix A goes over into $-A$. Thus, the matrix A operating on the entire group $\mathfrak{d}$ is a double-valued continuous function of the rotation D. For this reason it is said that the group C_2 constitutes a <u>double-valued representation</u> of the real rotation group $\mathfrak{d}$.

Remarks: We can restore the single-valuedness of the representation by considering the two vectors $c_1u_1 + c_2u_2$ and $\lambda c_1u_1 + \lambda c_2u_2$ in the vector space (u_1,u_2) (differing by a factor $\lambda \neq 0$) as essentially not differing from one another. In the same manner we consider the linear transformations with matrices A and λA as essentially not differing from one another; in particular, the matrices A and $-A$ which represent the same rotation differ only in a trivial manner. The vectors λu (consisting of a $u \neq 0$ multiplied by an arbitrary λ) form a one-dimensional subspace, a <u>ray</u>. We speak of a <u>ray-representation</u> (rather than a vector representation) when dealing with a representation in which we are considering the transformation of the rays rather than that of the vectors and in which, moreover, we do not distinguish between A and λA. If the matrix A from a ray-representation corresponds to the group element α, the matrix B to β, then the product $AB\lambda$ (with arbitrary λ) rather than

AB must correspond to ab. Now, one can always pass from a ray representation to a vector representation that is at most finitely multi-valued. That can be accomplished by multiplying the matrices A by a factor λ such that their determinants equal 1. Except for a root of one the factor λ is uniquely determined. It is always so chosen that the unit matrix corresponds to the identity element. Where we have a continuous representation of a continuous group we are able to uniquely determine the factor λ in the vicinity of unity through a continuous process. As a result, the matrix product AB can be accurately associated with the product ab in this vicinity.

The representations D_J ($J = 0, \frac{1}{2}, 1, 1\frac{1}{2}, \ldots$) are representations of U_2, but U_2 is a double-valued representation of δ; hence, the D_J are, at the most, to be considered as double-valued representations of δ. D_0 is the identical representation. $D_{\frac{1}{2}}$ is the double-valued representation of δ through U_2. D_1 is the single-valued (unique) representation of δ through itself. It will be shown later that the representations D_J with integral values of J are single-valued representations of δ; on the other hand, those with "half-integral" values of J are double-valued. It will also be shown that under a rotation, the spherical functions of l-th degree (l an integer) transform exactly according to D_l. Finally, we shall later also establish the irreducibility of the representation D_J of U_2 or δ.

The representation D_J of U_2 possesses an invariant Hermetian form, namely:

$$\sum_0^{\nu} r!(\nu - r)!\, \bar{c}_r c_r \qquad (16.7)$$

Proof: The coefficients c_r transform just like the coefficients

$\binom{\nu}{r} a_1^{\nu-r} a_2^r$ of the special form $(a_1 u_1 + a_2 u_2)^\nu$. The $\bar{c}_r$ likewise transform like $\binom{\nu}{r} \bar{a}_1^{\nu-r} \bar{a}_1^r$. Now, since $\bar{a}_1 a_1 + \bar{a}_2 a_2$ is invariant, then also

$$\nu! (\bar{a}_1 a_1 + \bar{a}_2 a_2)^\nu = \nu! \sum_{r=0}^{\nu} \binom{\nu}{r} \bar{a}_2^{\nu-r} a_1^{\nu-r} \bar{a}_1^r a_1^r$$

$$= \sum_{r=0}^{\nu} r!\,(\nu-r)! \binom{\nu}{r}^2 a_1^{\nu-r} \bar{a}_1^r a_1^{\nu-r} a_1^r$$

is invariant; the latter expression transforms like (16.7).

Thus, we have shown that the representations D_J of U_2 or δ are all <u>unitary</u>. The vectors

$$\frac{u_1^{\nu-r} u_2^r}{\sqrt{r!\,(\nu-r)!}} \qquad (16.8)$$

form an ortho-normal system for the form (16.7). We further note for future reference that the vector (16.8) takes on the factor $e^{\frac{1}{2} i(\nu-2r)\gamma}$ when undergoing a rotation D_γ by the angle of rotation γ about the z-axis. (According to (16.4) also - due to D_γ, u_1 is multiplied by $e^{\frac{i\gamma}{2}}$ and u_2 by $e^{\frac{-i\gamma}{2}}$.) Furthermore, the rotation D_y by an angle of rotation π about the y-axis carries the product $u_1^{\nu-r} u_2^r$ over into $(-1)^{\nu-r} u_1^r u_2^{\nu-r}$. (According to (16.4), the same rotation also carries u_1 and u_2 over into $-u_2$ and u_1.

#17. The Infinitesimal Transformations and the Representations of the Rotation Group.

It is convenient to employ the following parametric representation of the rotations of space - especially so, in the vicinity of unity. We consider a rotation about a (directed) axis <u>a</u> by an

angle of rotation φ (measured along the proper direction of rotation). We represent this rotation by a vector in the direction $\underline{a}$ of length φ ; the orthogonal components $\alpha_1, \alpha_2, \alpha_3$ of this vector are employed as a parameter of the rotation group. Thus, the entire parameter space is a closed sphere of radius π ; points on the surface of this sphere lying diametrically opposite each other are imagined to be identifiable. The product of two rotations $d_\alpha(\alpha_1, \alpha_2, \alpha_3)$ and $d_\beta(\beta_1, \beta_2, \beta_3)$ is a rotation $d_\beta d_\alpha = d_\gamma$, where in the neighborhood of the zero point the $\gamma_\nu = \varphi_\nu(\alpha, \beta)$ are single-valued, even analytical functions of the α and β ; the latter can be solved uniquely for the β. For $\beta = 0$ we have $\gamma = \alpha$ and the two functional matrices

$$S_\lambda^\nu(\alpha) = \left(\frac{\partial \gamma_\nu}{\partial \beta_\lambda}\right)_{\beta=0} \qquad \text{and} \qquad T_\gamma^\nu(\alpha) = \left(\frac{\partial \beta_\nu}{\partial \gamma_\lambda}\right)_{\gamma=\alpha}$$

are inverses of each other: $ST = E$.

We seek all (single- or multi-valued) representations of the rotation group δ in which each rotation d_α (in the vicinity of unity) is represented by a linear transformation D_α ; also, the matrix of the latter must depend[1] in a continuous, differentiable manner on the parameters $\alpha_1, \alpha_2, \alpha_3$ and, besides, the representation condition

$$D_\beta D_\alpha = D_\gamma = D_{\varphi(\alpha,\beta)}$$

must, of course, be satisfied.

The method to be used is the Lie-Cartan Method of Infinitesimal

---

1) The requirement for continuity (with differentiability) would by itself have been sufficient; even then, it can be replaced by still weaker requirements. We shall, however, not be concerned with those at the present.

Transformations.[1)]

On carrying out the transformation D_β on a vector u of the representation space there arises a vector $v = D_\beta\, u$. When $\beta_\nu = 0$ we have $v = u$. For infinitely small β_ν we can expand $D_\beta u$ (upon ignoring all terms of order higher than one) as follows:

$$v = D_\beta\, u = u + \left(\frac{\partial v}{\partial \beta_1}\right)\beta_1 + \left(\frac{\partial v}{\partial \beta_2}\right)_0 \beta_2 + \left(\frac{\partial v}{\partial \beta_3}\right)\beta_3 + \cdots .$$

The quantities $\left(\frac{\partial v}{\partial \beta_\nu}\right)_0$ depend linearly on u; hence

$$\left(\frac{\partial v}{\partial \beta_\nu}\right) = I_\nu\, u \qquad (\nu = 1,2,3)$$

and we call the linear transformations I_ν ($\nu = 1,2,3$) the <u>infinitesimal transformations</u> of the representation. These, in turn, represent the <u>infinitesimal rotations</u> about the X -, Y -, and Z - axes, respectively. Instead of I_1, I_2, I_3 we write I_x, I_y, I_z.

We start out with a fixed vector u_0 of the representation space and we equate

$$u = D_\alpha\, u_0$$

and

$$v = D_\beta\, u = D_\beta\, D_\alpha\, u_0 = D_\gamma\, u_0$$

where, of course, $\gamma_\nu = \varphi(\alpha, \beta)$. Differentiating the last formula with respect to β_ν and then setting $\beta_\nu = 0$ yields:

$$I_\lambda u = \left(\frac{\partial v}{\partial \beta_\lambda}\right)_{\beta=0} = \sum_\nu \left(\frac{\partial v}{\partial \gamma_\nu}\right)_{\gamma=\alpha} \left(\frac{\partial \gamma_\nu}{\partial \beta_\lambda}\right)_{\beta=0}$$

$$= \sum_\nu \frac{\partial u}{\partial \alpha_\nu} S_\lambda^\nu(\alpha) .$$

---

1) A detailed analysis of this method as well as further references in the literature are to be found in. Math. Z. vol. 23 (1925) by H. Weyl; also vol. 24 (1926) ibid. Representations of a much more general class of groups are set up there - the rotation - and Lorentz group being only special cases.

On solving these linear equations by means of the inverse matrix T, we obtain

$$\frac{\partial u}{\partial \alpha_\nu} = \sum_\sigma I_\sigma u T^\sigma_\nu(\alpha). \qquad (17.1)$$

These are the Lie "characteristic differential equations" of the representations.

It is to be noted that the T^σ_ν depend only on the composition of the rotation group and not on the representation in question. Since the quantity $u = D_\alpha u_0$ is fully determined by the differential equation (17.1) and the initial condition $u = u_0$ for $\alpha = 0$, there follows the theorem:

<u>A representation of the rotation group is fully determined by its infinitesimal transformations</u> I_x, I_y, I_z.

The operations I_x, I_y, I_z, however, are not completely arbitrary, for they must satisfy the "integrability conditions". These are obtained from (17.1) by differentiating with respect to α_μ and equating the mixed partials $\frac{\partial^2}{\partial\alpha_\mu \partial\alpha_\nu}$ and $\frac{\partial^2}{\partial\alpha_\nu \partial\alpha_\mu}$. The calculation yields

$$\sum_\sigma I_\sigma u_0 \left(\frac{\partial T^\sigma_\nu}{\partial\alpha_\mu} - \frac{\partial T^\sigma_\mu}{\partial\alpha_\nu}\right) + \sum_\rho \sum_\sigma I_\rho I_\sigma u \left(T^\rho_\mu T^\sigma_\nu - T^\rho_\nu T^\sigma_\mu\right) = 0.$$

We shall later need this relation for $\alpha = 0$.* In this case T^ρ_μ becomes the unit matrix and we obtain

$$-\sum_\sigma I_\sigma u_0 c^\sigma_{\mu\nu} + I_\mu I_\nu u_0 - I_\nu I_\mu u_0 = 0,$$

---

* As for any arbitrary α, we obtain the same result, anyway; only here the calculation is somewhat longer.

where, for the sake of brevity, we put

$$- c^{\sigma}_{\mu\nu} = \left(\frac{\partial T^{\sigma}_{\nu}}{\partial \alpha_{\mu}} - \frac{\partial T^{\sigma}_{\mu}}{\partial \alpha_{\nu}} \right)_{\alpha = 0} . \qquad (17.2)$$

But the vector u_0 was chosen to be completely arbitrary; hence it follows that

$$I_{\mu} I_{\nu} - I_{\nu} I_{\mu} = \sum_{\sigma} I_{\sigma} c^{\sigma}_{\mu\nu} . \qquad (17.3)$$

From (17.2) we see that the <u>real</u> constants $c^{\sigma}_{\mu\nu}$ depend only on the group. They can be determined as soon as we select for our representation space the space of the linear functions $\psi(x,y,z) = a_1 x + a_2 y + a_3 z$; here, the operators I_{ν} are immediately specified (cf. #6). In general, the commutation rules (6.3) must be valid for each representation. They are

$$\left.\begin{aligned} I_x I_y - I_y I_x &= I_z \\ I_y I_z - I_z I_y &= I_x \\ I_z I_x - I_x I_z &= I_y . \end{aligned}\right\} \qquad (17.4)$$

These permutation rules are fundamental in that they form a basis for the determination of all the possible representations. In order to put them into a more suitable form we introduce the operators

$$L_x = iI_x, \quad L_y = iI_y, \quad L_z = iI_z ,$$

as was done in #6. Furthermore, we set:

$$\begin{aligned} L_x + iL_y &= L_p \\ L_x - iL_y &= L_q . \end{aligned}$$

Hence, the permutation rules (17.4) when put into the new form are:

$$\left.\begin{aligned} L_z L_p - L_p L_z &= L_p \\ L_z L_q - L_q L_z &= -L_q \\ L_p L_q - L_q L_p &= 2L_z \end{aligned}\right\} \quad 1) \qquad (17.5)$$

Let there be given an arbitrary representation of the rotation group in a vector space R (of finite dimensions). In the same vector space there occurs, of course, a representation of that Abelian subgroup which consists of the rotations $(0, 0, \gamma)$ about the z-axis. According to #10, example 2, the latter representation can be reduced. One then finds a series of base vectors v_M which when undergoing the rotation $(0, 0, \gamma)$ assume the factor $e^{iM\gamma}$. (In a single-valued representation M must be an integer; that need not be the case, however, if the representation is single-valued only in the vicinity of unity.) We have

$$L_z v_M = iI_z v_M = i\left(\frac{\partial}{\partial\gamma} D(0,0,\gamma) v_M\right)_{\gamma=0} = i\left(\frac{\partial}{\partial\gamma} e^{-iM\gamma} v_M\right)_{\gamma=0} = Mv_M;$$

hence, the vectors v_M are eigenvectors of the operator L_z belonging to the eigenvalue M. We could have obtained them also by means of a principal-axes-transformation of the operator L_z.

Lemma. If the vector v belongs to the eigenvalue M of L_z then $L_p v$ and $L_q v$ belong to the eigenvalues (M+1) and (M-1) of L_z, respect-

---

1) This form of the permutation rules is more suitable since in the first two relations there occur this time only two (instead of as before - all three) infinitesimal transformations. The expressions L_p and L_q suggest themselves quite naturally when one conceives of the formation $TL - L_z L - LL_z$ as a linear transformation T in the space of the linear combinations $L = \lambda L_x + \mu L_y + \nu L_z$, and when this transformation is transformed on to the principal axes. The eigenvectors are L_p, L_q, and L_z and the corresponding eigenvalues 1, -1, 0.

ively.

Proof: From $L_z v = Mv$ it follows that

$$L_z L_p v = (L_p L_z + L_p)v = L_p M v + L_p v = (M + 1)L_p v \ .$$

Analogously,

$$L_z L_q v = (M-1)L_q v.$$

Thus, the lemma is proven.

We seek in the space R a vector v_J belonging to the largest eigenvalue J of L_z (or when dealing with imaginary eigenvalues, we seek a v_J belonging to the eigenvalue of L_z with the largest real component). Then $L_p v_J$ belongs to the eigenvalue J + 1. Since, however, J is the largest eigenvalue occurring, $L_p v_J = 0$, of necessity. Furthermore,

$v_{J-1} = L_q v_J$ belongs to the eigenvalue J-1

$v_{J-2} = L_q v_{J-1}$ belongs to the eigenvalue J-2, etc.

The series is thus continued until a null-vector is reached. That must happen once, since only a finite number of eigenvalues can occur in the space R.

It can be easily proven that for M = J, J-1, J-2,

$$L_p v_M = \rho_M v_{M+1} \qquad (17.6)$$

where ρ_M is an integer. Well, in proving that, we first observe that the formula is valid for the largest value M = J with $\rho_J = 0$ for then we have $L_p v_J = 0$. We now show that (17.6) is valid for M = μ-1 as soon as it also holds for M = μ. We have namely:

$$L_p v_{\mu-1} = L_p L_q v_\mu = (L_q L_p + 2L_z)v$$

$$= L_q \rho_\mu v_{\mu+1} + 2\mu v = (\rho_\mu + 2\mu) v_\mu .$$

Hence, (17.6) is proven. We have for ρ_M the recursion formuli:

$$\rho_{\mu-1} = \rho_\mu + 2\mu \; ; \qquad \rho_J = 0 .$$

Their solution is

$$\rho_M = J(J+1) - M(M+1). \tag{17.7}$$

It must happen once that a $v_M = 0$ whereas the preceding $v_{M+1} \neq 0$. Then $\rho_M = 0$, of necessity. From here it follows, however, that $M = -(J+1)$, for the only roots of the equation

$$J(J+1) - x(x+1) = 0$$

are $x = J$ and $x = -(J+1)$. Moreover, the value $M=J$ is ignored since $v_J \neq 0$. Thus the first null-vector in the series of vectors v_J, v_{J-1}, v_{J-2},.... is the vector $v_{-(J+1)}$.
The number of terms in the series v_J, v_{J-1} ,....,v_{-J} is 2J+1; hence, 2J+1 is a natural number and J is a half-integer. The allowable values of J are

$$J = 0, \tfrac{1}{2}, 1, 1\tfrac{1}{2}, \ldots .$$

In order to make the formuli more symmetrical we can provide the v_M with numerical factors and we define

$$L_q v_M = \sqrt{J(J+1) - M(M-1)} \cdot v_{M-1} .$$

Consequently,

$$
\begin{aligned}
L_p v_M &= \sqrt{J(J+1) - M(M+1)}\cdot v_{M+1} \\
&= \sqrt{(J - M)(J + M + 1)}\cdot v_{M+1} \\
L_q v_M &= \sqrt{J(J+1) - M(M-1)}\cdot v_{M-1} \\
&= \sqrt{(J + M)(J - M + 1)}\cdot v_{M-1} \\
L_z v_M &= M v_M .
\end{aligned}
\tag{17.8}
$$

The subspace $(v_J v_{J-1}, \ldots, v_{-J})$ of our vector space R is transformed into itself by the operations L_p, L_q, L_z or, in other words, by the infinitesimal rotations I_x, I_y, I_z. Therefore, also, the same subspace is transformed into itself by all transformations of the representation of the rotation group. In other words,

the vectors v_J, $v_{J-1}, \ldots, v_{-J}$ determine an invariant subspace R_{2J+1}.

The transformations of this subspace form a representation of the rotation group which is completely determined by equations (17.8). In the space R_{2J+1} the operator L_z has the simple eigenvalues $M = J, J-1, \ldots, -J$ corresponding to the eigenvectors v_M. It is also noteworthy that all vectors of the space R_{2J+1} are eigenvectors of the operators

$$
\bar{L}^2 = L_x^2 + L_y^2 + L_z^2 = (L_p L_q + L_q L_p) + L_z^2,
$$

for from (17.8) we have by a simple calculation:

$$
\bar{L}^2 v_M = J(J+1) v_M \tag{17.9}
$$

The space R_{2J+1} is irreducible. That is so for if R' is an

invariant subspace of R_{2J+1} and v' an eigenvector of L_z in the space R', then v' must agree with one of the vectors $v_J, \ldots . v_{-J}$ to within a numerical factor. (It could not be otherwise, for there are no other eigenvectors of L_z in R_{2J+1}.) According to (17.8), one $v' = v_M$, gives rise to all other v_M ($M = J, J-1, \ldots, -J$) due to the transformations L_q and L_p. For this reason all v_M belong to R'. Therefore, R' constitutes the entire space R_{2J+1} as was to be proven.

The representation of order 2J+1 determined by (17.8) is equivalent to the representation denoted by D_J which was constructed in the previous section.

In fact, in the space $(\ldots, u_1^{v-r} u_2^r, \ldots)$ of the representation $D_J^!$, the base vectors $u_1^{v-r} u_2^r$ take on the factor $e^{iM\gamma} = e^{\frac{1}{2}i(v-2r)}$ as they undergo the rotation $(0,0,\gamma)$. Thus, the values $M = r - \frac{1}{2}v$ ($= \frac{1}{2}v, \frac{1}{2}v-1, \ldots, -\frac{1}{2}v$) each occur once. If now we form in this space the subspace R_{2J+1} according to the above construction, then R_{2J+1} coincides with the entire space (because both are of the same dimension).

The v_M of the space R_{2J+1} agree with the products $u_1^{J+M} u_2^{J-M}$ of the representation D_J to within numerical factors. On calculating the numerical factors we obtain

$$v_M = \frac{u_1^{J+M} u_2^{J-M}}{\sqrt{(J+M)!(J-M)!}} \quad . \tag{17.10}$$

In accordance with # 16, these v form simultaneously an orthonormal system. In exactly the same manner we show for $J = l$ (an integer) that the representation determined by (17.8) is in agreement with the representation generated by the spherical functions

$Y_\ell^{(m)}$ of order ℓ. For these, too, are included in the number $2\ell+1$ and the operator L_z has $m=\ell$ as the highest eigenvalue. Therefore, the spherical functions $Y_\ell^{(m)}$ transform according to the irreducible representation D_ℓ. In other words, the normalization factors connected with the spherical functions $Y_\ell^{(m)}$ can be so chosen that equations (17.8) are strictly valid for them. From here it follows also that D_ℓ is a single-valued representation[1]. For instance, the functions $rY_1^{(1)} = -(x+iy)$, $rY_1^{(0)} = \sqrt{2}z$, $rY_1^{(-1)} = x-iy$ transform according to the representation D_1.

Finally, we can prove that:

each irreducible representation is equivalent with one of the representations D_J determined by (17.8). That follows, since the space R_{2J+1} constructed in the representation space must necessarily coincide with the entire space when irreducibility prevails.

The above deliberations provide a method by means of which we are able to effect the decomposition of any arbitrary representation D into irreducible D_J. The method consists merely in setting up the eigenvectors of the operator L_z in the space in question and in specifying, too, the eigenvalues that occur together with their multiples. When J is the largest integral or half-integral eigenvalue occurring, then D contains a representation D_J, in the space of which the eigenvalues J, J-1,....,-J each occur once. Of the

---

1) For non-integral values J, the representation D_J is not single-valued. For the vector v_J when undergoing a rotation $(0,0,\gamma)$ is multiplied by a factor $e^{-i\pi j\gamma}$ and this yields the value -1 when $\gamma = 2\pi$ and j = integer + ½.

remaining eigenvalues we again select the largest J' and we cleave off a representation $D_{J'}$, etc., until all eigenvalues are used up.

18. Examples and Applications.

1. Reduction of the Product Representation $D_j \times D_{j'}$ of the Rotation Group.

Let $(u_j,\ldots,u_{-j})$ be the space of the representation D_j and $(v_{j'},\ldots,v_{-j'})$ the space of $D_{j'}$. Then the products $u_m v_{m'}$ are the base vectors in the space of the representation $D_j \times D_{j'}$. Upon a rotation $(0,0,\gamma)$ about the z-axis the product $u_m v_{m'}$ takes on the factor $e^{-i(m+m')\gamma}$ and it belongs to the eigenvalue $M = m + m'$ corresponding to the operator L_z. The permitted values of M are:

$$
\begin{array}{lll}
(m=j) & M = j + j',\ j + j' -1,\ldots,j-j', \\
m=j-1) & M = \quad j + j' -1, j+j'-2,\ldots,j-j'-1, \\
(m=j-2) & M = \quad\quad\quad\quad j+j'-2,\ldots\ldots j-j'-2, \\
\ldots\ldots & \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\
(m=-j) & M = \quad\quad\quad\quad\quad\quad -j+j',\ldots\ldots\ldots -j-j'.
\end{array}
$$

Let us assume tentatively that $j \geq j'$. It can then be seen that the largest value $M = j+j'$ occurs <u>once</u>, the next largest $M = j+j'-1$ occurs <u>twice</u>, etc., the frequency of the other values increasing always by one until the value $M = j-j'$ is reached which then occurs $(2j'+1)$ times. All smaller values of M (i.e. ranging between $M = j-j'$ and $M = -j+j'$) occur also $(2j'+1)$-times. As far as the negative values of M are concerned, we need not delay ourselves any further.

The above observation in conjunction with the rules of #17 lead us to the following conolusions:

The representation $D_{j+j'}$ which corresponds to the largest eigenvalue occurs but <u>once</u> in the representation $D_j \times D_{j'}$ in question. Also, the representation space of $D_j \times D_{j'}$ involves each of the eigenvalues $M = j + j', \ldots, -(j+j')$ only once. When these eigenvalues are cancelled the largest value that remains is $M = j+j'-1$ which now occurs only once; henoe, the representation $D_{j+j'-1}$ likewise occurs only once. (The corresponding eigenvalues are $j+j'-1$, $j+j'-2, \ldots, -(j+j-1)$). By continuing this line of reasoning one finally obtains the representation $D_{j-j'}$ which uses up all remaining eigenvalues. Hence, we have:

$$D_j \times D_{j'} = D_{j+j'} + D_{j+j'-1} + \ldots + D_{|j-j'|}. \qquad (18.1)$$

The addition of the 'absolute sign' to $j-j'$ makes the formula symmetrical in j and j'; it also makes the formula valid for $j' > j$. For example,

$$D_0 \times D_j = D_j,$$
$$D_1 \times D_1 = D_2 + D_1 + D_0$$
$$D_1 \times D_{\frac{1}{2}} = D_{1\frac{1}{2}} + D_{\frac{1}{2}}.$$

In order to carry out explicitly the reduction of the representation $D_j \times D_{j'}$ we must state the vectors w_M (in the space of the products $u_m v_{m'}$) which transform according to D_J ($J = j+j', j+j'-1, \ldots$). We shall write this time U_m, V_m, W_M^J instead of u_m, $v_{m'}$, w_M and, according to (17.10) we set:

$$U_m = \frac{u_1^{j+m} u_2^{j-m}}{\sqrt{(j+m)!(j-m)!}}; \qquad V_{m'} = \frac{v_1^{j'+m'} v_2^{j'-m'}}{\sqrt{(j'+m')!(j'-m')!}}$$

Presently, we form for $J = j+j'-\lambda$ $(\lambda=0,1,2,....)$ the expression:

$$A = (u_1v_2 - u_2v_1)^{\lambda}(u_1x_1+ u_2x_2)^{2j-\lambda}(v_1x_1+v_2x_2)^{2j'-\lambda}$$

and we state that the coefficients of

$$X_M^J = \frac{x_1^{J+M}x_2^{J-M}}{\sqrt{(J+M)!(J-M)!}}$$

in A for $M = J, J-1,....,-J$ represent the desired quantities W_M^J.

Proof: The expression A is invariant when the x_1, x_2 are transformed contragrediently to u_1, u_2 and v_1, v_2. The coefficients W_M^J, therefore, are transformed contragrediently to the X_M^J. Likewise, the expression $(u_1x_1+ u_2x_2)^{2J} = (2J)!\Sigma U_M^J X_J^J$ is invariant when the U_M^J are transformed like the $u_1^{J+M}u_2^{J-M}$: $\sqrt{(J+M)!(J-M)!}$. In other words, the U_M^J, too, are transformed contragrediently to the X_M^J. Hence, the W_M^J are transformed like the U_M^J i.e. according to D_J, Q.E.D.

The calculation of A proceeds as follows:

$$(u_1v_2 - u_2v_1)^{\lambda} = \sum_{\nu=0}^{\lambda}(-1)^{\nu}\binom{\lambda}{\nu}(u_1v_2)^{\lambda-\nu}(u_2v_1)^{\nu}$$

$$A = \sum\sum_{mm'} \sqrt{(j+m)!(j-m)!(j'+m')!(j'-m')!(J+M)!(J-M)!}$$
$$\cdot \sum_{\nu=0}^{\lambda}(-1)^{\nu}\binom{\lambda}{\nu}\binom{2j-\lambda}{j-m-\nu}\binom{2j'-\lambda}{j'+m'-\nu} U_mV_{m'}X^J_{m+m'}$$
$$= \lambda!(2j-\lambda)!(2j'-\lambda)!\sum\sum_{mm'} c^J_{mm'}U_mV_{m'}X^J_{m+m'}$$

$$c^J_{mm'} = \sum_{\nu}(-1)^{\nu}\frac{\sqrt{(j+m)!(j-m)!(j'+m')!(j'-m')!(J+M)!(J-M)!}}{(j-m-\nu)!(j+m-\lambda+\nu)!(j'+m'-\nu)!(j'-m'-\lambda+\nu)!\nu!(\lambda-\nu)!}$$

$$(M = m+m')^{*} \qquad (18.2)$$

*The fraction in (18.2) becomes zero as soon as one of the numbers $(j+m-\nu)$ etc. in the denominator becomes negative. Moreover, $0! = 1$.

We, therefore, have

$$W^J_M = \rho_J \sum'_{m+m'=M} c^J_{mm'} U_m V_{m'} \quad , \tag{18.3}$$

where ρ_J is a numerical factor that can be chosen quite arbitrarily. For instance, it can be chosen so that the W^J_M form[1] a normalized orthogonal system in the unitary vector space of the $U_m V_m$. If now we set $b^J_{mm'} = \rho_J c^J_{mm'}$ then for each fixed M the $b^J_{mm'}$ (with $m+m' = M$) form a unitary matrix B_M, where J is the column index and m or m' the row index. According to (7.5), the transposed matrix B_M^{-1} is simply the complex conjugate transposed matrix $\tilde{B}_M$. That is, equations (18.2) can be solved for the $U_m V_{m'}$ in the following manner:

$$U_m V_{m'} = \sum_J \rho_J c^J_{mm'} W^J_{m+m'} . \tag{18.4}$$

The values of the numbers ρ_J are unimportant. Equation (18.4) is known by the name: Clebsch-Gordan Series. The numbers $c^J_{mm'}$ can, in general, be calculated from (18.2). For the special case $J = j+j'(\lambda=0)$, equation (18.2) reduces to:

$$c^J_{mm'} = \sqrt{\frac{(J+M)!(J-M)!}{(j+m)!(j-m)!(j'+m')!(j'-m')!}} \quad ;$$

---

1) Since any two non-equivalent irreducible subspaces of the space of the unitary representation $D_j \times D_{j'}$ are always orthogonal to one another, it follows that any two vectors W_M belonging to different values of J are orthogonal to each other. That is so, for the vertical projection of one space upon the other gives rise to an operator-homomorphism that can be none other than the zero mapping. The orthogonality of any two W_M with equal J but different M can be explained in the same way - except that here we speak of the representations of the subgroup of the rotations about the Z-axis.

TABLE I

$j' = 1/2$

J	$m' = 1/2$	$m' = -1/2$
$j + 1/2$	$\sqrt{j + m + 1}$	$\sqrt{j - m + 1}$
$j - 1/2$	$-\sqrt{j - m}$	$+\sqrt{j + m}$

$j' = 1$

J	$m' = 1$	$m' = 0$	$m' = -1$
$j+1$	$\sqrt{\frac{(j+m+2)(j+m+1)}{2}}$	$\sqrt{(j+m+1)(j-m+1)}$	$\sqrt{\frac{(j-m+2)(j-m+1)}{2}}$
j	$-\sqrt{2(j+m+1)(j-m)}$	$+2m$	$+\sqrt{2(j+m)(j-m+1)}$
$j-1$	$\sqrt{\frac{(j-m)(j-m-1)}{2}}$	$-\sqrt{(j+m)(j-m)}$	$\sqrt{\frac{(j+m)(j+m-1)}{2}}$

likewise, for the special case $J = j-j'$ ($\lambda = 2j'$; $j \geq j'$) equation (18.2) reduces to:

$$c^{J}_{mm'} = (-1)^{j'+m'} \sqrt{\frac{(j+m)!(j-m)!}{(j'+m')!(j'-m')!(J+M)!(J-M)!}}$$

Table I displays the values of the $c^{J}_{mm'}$ for the simplest cases.

Applications of the Relation (18.1).

The state of a system of f electrons in a spherically symmetrical force field is described by a function $\psi(q_1, q_2, \ldots . q_f)$. When a linear family of eigenfunctions of a certain energy level undergoes a rotation it is transformed linearly into itself. Hence, the linear family breaks apart into subfamilies that transform in accordance with the representations D_J. In this case it is customary to write L instead of J. Also, since the representation is unique there can be only integral values of L.

In our case, the operator for an infinitesimal rotation of all electrons - let us say, about the z-axis - is:

$$I_z = -\sum_1^f \left(x \frac{\partial}{\partial y} - y \frac{\partial}{\partial x}\right).$$

Hence, $hiI_z = hL_z$ is the operator for the z-component of the angular momentum $h\vec{L}$. Let us consider for a moment an irreducible family of eigenfunctions that transforms according to D_L. Then according to #17 the operator $\vec{L}^2$ has the eigenvalue L(L+1) and L_z, the eigenvalues M = L,L-1,....,-L. In terms of the vector scheme, we imagine the atom to be represented by an angular momentum vector of length hL, the z-component of which takes on the values hM(M=L,...,-L) - just like in the one-electron problem (#6). We speak of S-, P-, D-, F- etc. terms of the atom for L = 0,1,2,3,..., respectively (analogous to the s-,p-,d-,f-,.... terms in the case of the electron). L is known as the azimuthal quantum number.

A continuous variation in the potential (brought about, for

example, by a diminution in the mutual repulsion of the electrons) does not affect the quantum number L, for the representation varies continuously and, therefore, the eigenvalue of $\bar{L}^2$ cannot possibly become discontinuous. Thus, we can tabulate the permitted values of L, excluding thereby the interaction forces between the electrons. Better still, the tabulation can be carried out upon replacing the interaction by an appropriately chosen screening of the field of the nucleus (cf #4); later on, this interaction can again be introduced gradually.

We shall consider two electrons. The state of each electron is described by a wave function $\varphi_{n\ell}^{(m)}$ to which, according to #4, there corresponds a definite term symbol ns or np, nd etc., depending on the value of ℓ taken. If the repulsion of the electrons is ignored, then the eigenfunctions of the entire system are the products $\varphi_{n\ell}^{(m)}(q_1)\varphi_{n'\ell'}^{(m')}(q_2)$; when undergoing rotations they transform according to $D_\ell \times D_{\ell'}$. According to (18.1) the reduction of this representation gives rise to a series of subfamilies which transform according to D_L ($L = \ell+\ell', \ell+\ell'-1, \ldots |\ell-\ell'|$). If now we introduce the interaction we observe that the atom terms belonging to the different L-values separate in accordance with #9; the (2L+1)-fold degeneracy of the individual term is, however, not removed and the representations D_L remain as they are.

In terms of the vector scheme, one must combine the two vectors (representing the angular momenta of the individual electrons) in such a manner that the length hL of the resultant is either equal to $h(\ell+\ell')$ or is less than that by an integral multiple of h. Then the smallest values of the resultant is $h|\ell-\ell'|$, as must be the case according to (18.1).

Example: Let $l = l' = 1$ - that is, we have two p-electrons. Also, let E_1 and E_2 be the energy levels of the electrons not including the mutual repulsion. Then, combining E_1 with E_2 we obtain the total energy $E_1 + E_2$. Now, the repulsion will cause the latter term to split into the three terms with $l = 0,1,2$; i.e., it splits into an S-, a P-, and a D-term, respectively. In exactly the same manner we proceed for all other cases.

When dealing with more than two electrons one simply applies the formula several times. For example, in the case of an s-, a p- and a d-electron the calculation proceeds as follows:

$$D_0 \times D_1 \times D_2 = (D_0 \times D_1) \times D_2 = D_1 \times D_2 = D_3 + D_2 + D_1 ;$$

hence, we can have an F-, a D- and a P-term. We denote each term by a symbol which consists of the symbols of the individual electrons and the symbol of the entire term. For instance, in a system of three electrons of which two are in the 1s-state and one is in the 2p-state and the resulting term is necessarily a P-term, we have: $1s^2 2p\,P$.

By definition, an S-state is always spherically symmetric, for the ψ-function is invariant under all rotations. Now, the addition of an s-electron does not affect the permitted values of L, for we have $D_l \times D_0 = D_l$.

In #4 we derived a number of rules (through approximate considerations) concerning the terms of an atom such as Li, Na, K - the atom consisting of one emitting electron and a spherically symmetric core. Now, considerations of the above type lead to the exact verification of just these rules. Before, we substituted for the inter-action between the emitting electron and the core

electrons a simple screening of the field of the nucleus - and for the possible values of the moment of momentum of the emitting electron we found $\ell = 0,1,2,....$ If now we assume the core to be spherically symmetric ($L'=0$) in the absence of the emitting electron, then we obtain for the entire system (where screening replaces interaction) the value $L = \ell$. Then when the perturbation (i.e. interaction minus screening) is introduced, no splitting occurs - in agreement with what has been said above. Instead, each term is still $(2\ell+1)$-fold degenerate and, as before, it transforms according to D_ℓ. In the next section we shall also see that the selection rule $\ell \to \ell \pm 1$ which explains the arrangement of the terms in series is strictly valid.

Now, hydrogen-like spectra (like those of Li, Na, K, etc.) are not the only type of line spectra that are characterized by a grouping of the lines in series. Let us consider any atom whatever in which the quantum numbers n, ℓ of all its electrons except one remain unchanged. The principal quantum number of the outermost electron, however, we allow to run through the series of all the allowable values $n = \ell+1, \ell+2, \ell+3, ...$. As a result, for each allowable value of the total azimuthal quantum number L there arises a <u>term series</u> of ever increasing energy values. The upper limit of this term series represents the energy of the state of the ion which is due to a complete removal of the outermost electron. An example of such a series in the He-atom is the 'principal series' 1s np P ($n=2,3,4,....$); its upper limit is the energy of the He-ion in the ground state (see fig.6, #26). Likewise, in the case of carbon C the following series among others are possible:

$Is^2 2s^2 2p\ ns\ P$, $Is^2 2s^2 2p\ np\ S$, $Is^2 2s^2 2p\ np\ P$, $Is^2 2s^2 2p\ np\ D$.

The common limit of these series is the ionic term $1s^2 2s^2 2p\ P$. Generally, such series can be represented by empirical formuli of the type

$$E_n = E_\infty - \frac{1}{(n-\varkappa)^2}$$

where E_∞ is the energy of the ion; as n increases, $\varkappa$ tends rapidly towards a constant limit. In the following section we shall derive the selection rules with the aid of which we shall then be able to determine as to which series combine with another to give series of spectral lines.

3. The Reflection Character

The field of a single nucleus is invariant not only under spacial rotations but under reflections, as well. By combining any rotation with the single "reflection at the point of origin"

$$x' = -x,\ y' = -y,\ z' = -z$$

(which commutes with all rotations) we can produce all the other reflections. This single reflection together with the identity form an Abelian group of order 2. Due to the above-mentioned commutability it is possible to reduce out this Abelian group simultaneously with the rotation group. In other words, it is possible to choose the base vectors of a representation (in particular, the eigenfunctions belonging to a certain energy level) in such a manner that they transform according to D_ℓ when undergoing rotations and simultaneously take on the factor $w = \pm 1$ when under-

going the reflection S. The factor w is known as the reflection character.

A specific example is provided by the one-electron problem in which the spherical functions belonging to the ℓ-th energy level correspond to the reflection character $(-1)^{\ell}$.

Let us consider f electrons having the azimuthal quantum numbers $\ell_1, \ell_2, \ldots\ldots, \ell_f$ and located in a spherically symmetrical field. On ignoring the mutual perturbation among the electrons and then bringing them together, there arise the products of the eigenfunctions

$$\psi = \psi_1(q_1)\psi_2(q_2)\ldots\ldots\psi_f(q_f)$$

which obviously correspond to the reflection character

$$w = (-1)^{\ell_1 + \ell_2 + \ldots\ldots + \ell_f} . \qquad (18.5)$$

Even after the perturbation is again taken into account - that is, when the eigenfunctions no longer are products $\psi_1\psi_2\ldots\psi_f$ the character still remains the same. The resulting terms are called even or odd depending on whether w = +1 or -1. For example, of the four carbon series mentioned before, the first two consist of odd terms, the other two of even terms. We shall soon understand the relations these terms bear to the spectrum.

19. Selection and Intensity Rules.

We begin by stating two lemmas in group theory.

Lemma 1: Let two representations D, D' of a group G in the spaces $R = (u_1,\ldots,u_n)$ and $R' = (v_1,\ldots,v_n)$ be given by exactly the same formulae

$$au_\mu = \sum_\lambda u_\lambda \alpha_{\lambda\mu}$$

$$av_\mu = \sum_\lambda v_\lambda \alpha_{\lambda\mu}$$

The only difference here is that the u_μ form a linearly independent base for R whereas the v_μ are linearly dependent. Let also the representation D be completely reducible, ie.

$$D = D_1 + ... + D_k. \qquad (19.1)$$

Consequently, D', too, is completely reducible and the decomposition of D' is brought about by cancellation of a number of representations in the sum on the right-hand side of (19.1).

Proof: On associating with each vector $u = \sum_\lambda c_\lambda u_\lambda$ the vector $v = \sum_\lambda c_\lambda v_\lambda$, there corresponds to the sum of two vectors $(u_1 + u_2)$ again the sum $(v_1 + v_2)$ and av corresponds to au; hence, the correspondence is an operator-homomorphism. Thus, according to # 11, Theorem 4

$$R' \cong r'_1 + ... + r'_h$$

where the r' are certain irreducible subspaces produced on decomposition of R.

Lemma 1 is applied mainly to product representations. For instance, one starts out with certain quantities (eigenfunctions or others) $U_j^{(m)}$ and $V_{j'}^{(m')}$ which transform according to D_j and $D_{j'}$, respectively. The question is: how do the products $U_j^{(m)} V_{j'}^{(m')}$ transform? Well, on replacing the U, V by as many independent variables, uv, the products $u_j^{(m)} v_{j'}^{(m')}$ transform according to $D_j \times D_{j'} = \sum_J D_J$;- $J = j + j', ..., |j-j'|$. Now, we replace the

u,v in these transformation formulae back again by U,V. The formulae then are still valid although the products may now be linearly dependent. The lemma then points out that the products transform according to a representation ΣD_J in which there occur some (and eventually all) of the permitted $J = j + j', \ldots, |j - j'|$.

Lemma 2: Let a complete orthogonal system

$$\phi_1^{(1)}, \ldots, \phi_1^{(h_1)}; \phi_2^{(1)}, \ldots, \phi_2^{(h_2)}; \ldots \qquad (19.2)$$

be determined in such a manner that under a given transformation group G (e.g. the rotations in space) the set of functions $\phi_\lambda^{(1)}, \ldots, \phi_\lambda^{(h_\lambda)}$ suffer one irreducible representation D_λ for each λ. Let us also expand a set of functions $\varphi^{(1)}, \ldots, \varphi^{(h)}$, which suffer under the same group a completely reducible representation D, in terms of the orthogonal system (19.2). As a result, only those φ_λ^σ can actually occur in the expansion, the representation D_λ of which is contained in the representation D as a constituent.

Proof: For φ a linear combination of $\varphi^{(1)}$ to $\varphi^{(h)}$, we have

$$\varphi \sim \sum_1^{h_1} a_{1\sigma} \phi_1^{(\sigma)} + \sum_1^{h_2} a_{2\sigma} \phi_2^{\sigma} + \ldots\ldots$$

$$= \omega_1 + \omega_2 + \ldots\ldots \qquad (19.3)$$

Since φ determines all the components $a_{\lambda\sigma}$ uniquely, it also determines the ω_1, ω_2, etc. uniquely. The association $\varphi \rightarrow \omega_\lambda$ constitutes a linear mapping of the set $(\varphi) = (\varphi^{(1)}, \ldots, \varphi^{(h)})$ on to the set $(\phi_\lambda^{(1)}, \ldots, \phi_\lambda^{(h_\lambda)})$, for sum corresponds to sum and the multiple $\alpha\omega_\lambda$ corresponds to the multiple $\alpha\varphi$. Furthermore, a transformation t of the group G can be carried out on (19.3) yielding:

$$t\varphi = t\omega_1 + t\omega_2 + \ldots\ldots,$$

where $t\omega_\lambda$ is again a linear combination of the same form as ω_λ. Thus, in our mapping $t\omega_\lambda$ corresponds to $t\varphi$. In other words, the mapping is an operator homomorphism of the set (φ) on to the set (ω_λ). This set (ω_λ) either consists only of the zero or it is identical with the entire irreducible set $(\phi_\lambda^{(1)}, \ldots\ldots, \phi_\lambda^{(h_\lambda)})$ and transforms according to D_λ. This D_λ must, according to the theorems of #11, occur as a constituent in the decomposition D. The proof is thus complete.

Corollary to Lemma 2. Let us replace the φ successively by the functions $\varphi^{(1)},\ldots,\varphi^{(h)}$ and, accordingly, let us provide the $a_{1\nu}$, $a_{2\nu},\ldots$ with a superscript μ $(= 1,2,3,\ldots,h)$. Then, as soon as we know the representations D and D_1, the coefficients $a_{1\nu}^{(\mu)}$ (and likewise $a_{2\nu}^{(\mu)}$ etc.) are uniquely determined (to within a common factor) in a purely group-theoretical manner, - provided, however, that no two equivalent irreducible constituents occur in the decomposition of the representation D.

Proof: According to the proof above, the $a_{1\nu}^{(\mu)}$ form the matrix of an operator homomorphic mapping of the set $(\varphi^{(1)},\ldots,\varphi^{(h)})$ on to the set $(\phi_1^{(1)}, \ldots, \phi_1^{(h_1)})$. We imagine the $\varphi^{(1)}$ to $\varphi^{(h)}$ to be so chosen that the non-equivalent irreducible subsets are given by $(\varphi^{(1)},\ldots,\varphi^{(a)})$, $(\varphi^{(a+1)},\ldots,\varphi^{(a+b)})$,.... If then the $a_{1\nu}^{(\mu)}$ are not all zero, then $(\phi_1^{(1)}, \ldots, \phi_1^{(h_1)})$ must be equivalent to one of the φ-subsets, let us say $(\varphi^{(1)},\ldots,\varphi^{(a)})$. By an appropriate choice of the $\varphi^{(1)},\ldots,\varphi^{(a)}$, the two equivalent sets are transformed not only equivalently but also exactly alike. The operator homomorphism of the set $(\varphi^{(1)},\ldots,\varphi^{(h)})$ also maps the subsets $(\varphi^{(1)},\ldots,\varphi^{(a)})$ etc. homomorphically. According to Schur's lemma, the subset

$(\varphi^{(1)}, \ldots, \varphi^{(a)})$ is mapped by virtue of a matrix λE whereas the remaining non-equivalent subspaces are mapped on to zero. Hence, the matrix of the mapping is uniquely determined to within a factor λ. The same is true even if one subsequently reverts to another base of the set $(\varphi^{(1)}, \ldots, \varphi^{(h)})$.

Lemma 2 is the basis for the selection rules. In # 4 and # 6, the selection rules for a single electron were derived. They are:

$$\left.\begin{array}{l} l \to l \pm 1 \text{ for a spherically symmetric field} \\ m \to m \text{ or } m \pm 1 \text{ for an axially symmetric field.} \end{array}\right\}$$

The rule for m (by virtue of its derivation) is also applicable to the case of many electrons. How about the rule for l, when l is replaced by L ?

According to # 3, the intensities of the emitted lines depend on the expansion coefficients a,b,c occurring in the expansions

$$\begin{aligned} X\varphi_L^{(m)} &= \Sigma \varphi_{L'}^{(m')} a_{L'L}^{(m'm)} \\ Y\varphi^{(m)} &= \Sigma \varphi_{L'}^{(m')} b_{L'L}^{(m'm)} \\ Z\varphi_L^{(m)} &= \Sigma \varphi_{L'}^{(m')} c_{L'L}^{(m'm)} \end{aligned} \qquad (19.4)$$

The quantities on the left or rather their linear combinations

$$-(X + iY)\varphi_L^{(m)}, \quad (X - iY)\varphi_L^{(m)}, \quad \sqrt{2}\, Z\, \varphi_L^{(m)} \qquad (19.5)$$

are, according to the denotation from before, the products $V_1^{(-1)} U_L^{(m)}$, $V_1^{(1)} U_L^{(m)}$, $V_1^{(0)} U_L^{(m)}$. These transform according to $\Sigma D_{L'}$, where L' can take on some of the values $L \pm 1$ and L. Therefore, also on the right-hand side of (19.4) only these $D_{L'}$ can occur.

That gives rise to the selection rule:

$$L \longrightarrow \begin{cases} L - 1 \\ L \\ L + 1 \end{cases} \qquad (0 \rightarrow 0 \text{ is forbidden}). \qquad (19.6)$$

In exactly the same manner (only much simpler here) we obtain the selection rule for the reflection character $w = (-1)^{\Sigma \ell_s}$:

$$w \longrightarrow -w \qquad (19.7)$$

or the Laporte Rule: $\Sigma \ell_s$ can jump only by an odd number. For if in (19.4), the $\varphi_\lambda^{(m)}$ take on the factor w under the reflection s then the left-hand sides of (19.4) take on the factor -w; hence, also the right-hand sides will contain only terms having the reflection character -w. Let us consider the case of an emitting (valence) electron and a spherically symmetric core. We imagine the valence electron to make a quantum jump and the core to remain unchanged (as a first approximation). The selection rule in this case then excludes the transition $L \longrightarrow L$ permitted in (19.6). In this case, therefore, the old selection rule $L \longrightarrow L \pm 1$ obtains, or what amounts to the same, we have $\ell \longrightarrow \ell \pm 1$.

Furthermore, according to the corollary of lemma 2 and by virtue of group-theoretic considerations alone we find that the coefficients $a_{L'L}^{(m'm)}$ in (19.4) are uniquely determined for each pair of values L, L' to within a factor $\varrho_{L'L}$ (independent of m and m'). Calculation of these coefficients yields information about the intensity conditions of the lines that arise due to the removal of the degeneracy of the terms by a non-spherically symmetric perturbation, (Zeeman - or Stark Effect). That is so, provided

the perturbation is so weak that the ψ-function of the unperturbed system may be employed in the calculation of the intensities. The calculation is actually very simple when we observe that (18.4) indeed constitutes an expansion for the products $U_m V_{m'}$ which are for $j = L$, $j' = 1$ of the same type as our $U_L^{(m)} = \psi_L^{(m)}$ and $V_1^{(1)} = -(X + iY)$, $V_1^{(-1)} = X - iY$, $Y_1^{(0)} = Z\sqrt{2}$. Therefore, according to lemma 2 (corollary), the expansion coefficients of the products $U_L^{(m)} \; V_1^{(m')}$ must agree for each L' with the coefficients of (18.4) to within a numerical factor. In order to bring the different notations into close agreement with each other we replace the symbols L, L', m' in (19.4) by j, J, M, respectively. Thus, we write:

$$\begin{aligned} -(X + iY)\psi_j^{(m)} &= -\Sigma \psi_r^M \, (a + ib)_{rj}^{(Mm)} \quad , \\ (X - iY)\psi_j^{(m)} &= \Sigma \psi_r^M \, (a + ib)_{rj}^{(Mm)} \quad , \\ \sqrt{2} Z \psi_j^{(m)} &= \Sigma \psi_r^M \, \sqrt{2} c_{rj}^{(Mm)} \quad . \end{aligned} \tag{19.8}$$

Well, the coefficients on the right-hand side of (19.8) must for each J be proportional to the expansion coefficients $c_{m,M-m}^{r}$ of (18.4). Thus, we obtain (cf. the other table on p. 126 b)

$$\begin{aligned} J = j + 1: \quad (a+ib)_{r,j}^{(m+1,m)} &= -\varrho \sqrt{(j+m+2)(j+m+1)}, \\ (a-ib)_{r,j}^{(m-1,m)} &= \varrho \sqrt{(j-m+2)(j-m+1)}, \\ c_{r,j}^{(m,m)} &= \varrho \sqrt{(j+m+1)(j-m+1)}; \\ J = j: \quad (a+ib)_{r,j}^{(m+1,m)} &= \sigma \sqrt{(j+m+1)(j-m)}, \\ (a-ib)_{r,j}^{(m-1,m)} &= \sigma \sqrt{(j+m)(j-m+1)}, \\ c_{r,j}^{(m,m)} &= \sigma \, m; \\ J = j - 1: \quad (a+ib)_{r,j}^{(m+1,m)} &= \tau \sqrt{(j-m)(j-m-1)}, \\ (a-ib)_{r,j}^{(m+1,m)} &= -\tau \sqrt{(j+m)(j+m-1)}, \\ c_{r,j}^{(m,m)} &= \tau \sqrt{(j+m)(j-m)}. \end{aligned} \tag{19.9}$$

Hereafter, the j,J in these formulae are to be replaced by L, L'. The a_{Jj}^{Mm} and b_{Jj}^{Mm} can be easily calculated from the $(a+ib)_{Jj}^{Mm}$ and $(a-ib)_{Jj}^{Mm}$ when desired. In accordance with # 3 the squares of these numbers yield the transition probabilities to which the intensities of the corresponding spectral lines are proportional. The direction of polarization of the emitted light was already specified in # 6.

20. The Representations of the Lorentz-Group.

In the same manner that we set up the representations of the rotational group $\mathfrak{D}$ in # 16 and # 17, we now wish to investigate the representations of the group of Lorentz-transformations (Lorentz Group).

1. The group c_2 and the restricted Lorentz-transformations.

We begin with the group c_2 of the uni-modular linear transformations in a two-dimensional vector-space. Since we shall later be concerned with co- and contra-gradient vectors, we adopt the notations of Ricci-Kalkuel; we denote the base-vectors by $\overset{1}{u}$, $\overset{2}{u}$ and their linear combinations by $a_1\overset{1}{u} + a_2\overset{2}{u}$. The transformation formulae are:

$$\begin{aligned} \overset{1}{u}' &= \overset{1}{u}\alpha + \overset{2}{u}\gamma \\ \overset{2}{u}' &= \overset{1}{u}\beta + \overset{2}{u}\delta \\ \alpha\delta - \beta\gamma &= 1 \end{aligned} \qquad (20.1)$$

We have recourse, also, to a second vector-space $(a_{\dot{1}}\overset{\dot{1}}{u} + a_{\dot{2}}\overset{\dot{2}}{u})$ which is to be transformed simultaneously with the first vector space - but always by the complex-conjugate matrix:

$$\dot{u}' = \dot{u}\bar{\alpha} + \dot{\dot{u}}\bar{\gamma}$$
$$\dot{\dot{u}}' = \dot{u}\bar{\beta} + \dot{\dot{u}}\bar{\delta} \qquad (20.2)$$

By convention, all quantities transformed according to (20.2) are denoted by punctuated indices ($\dot{1},\dot{2}$ or $1',2'$).

The $\overset{1}{u}$, $\overset{2}{u}$ may be interpreted as the numerical variables and the $\dot{u}$, $\dot{\dot{u}}$ as the complex-conjugate variables $\dot{u} = \bar{u}$, $\dot{\dot{u}} = \bar{\dot{u}}$. Occasionally we shall make use of this interpretation; at times, however, we shall also mean by $\overset{1}{u}$, $\overset{2}{u}$, $\dot{u}$, $\dot{\dot{u}}$ four entirely arbitrary base-vectors. If (a_1, a_2) and (b_1, b_2) are two vectors, both of which are transformed according to (20.1), then $(a_1 b_2 - a_2 b_1)$ is invariant; hence, the vector $(b_2, -b_1)$ is contra-gradient to (a_1, a_2). Thus, from each binary vector (b_1, b_2) we can form a contra-gradient vector (b^1, b^2) whose components are:

$$b^1 = b_2 \ , \quad b^2 = -b_1 \qquad (20.3)$$

In like manner we define for each vector $(b_{\dot{1}}, b_{\dot{2}})$ a contra-gradient vector $(b^{\dot{1}}, b^{\dot{2}}) = (b_{\dot{2}}, -b_{\dot{1}})$. The linear space of all bi-linear forms

$$c_{\dot{1}\dot{1}} \overset{1}{u}\dot{u} + c_{\dot{1}\dot{2}} \overset{1}{u}\dot{\dot{u}} + c_{\dot{2}\dot{1}} \overset{2}{u}\dot{u} + c_{\dot{2}\dot{2}} \overset{2}{u}\dot{\dot{u}} \qquad (20.4)$$

is linearly transformed into itself by the transformations (20.1) and (20.2) (i.e. by means of the substitution of the $\overset{\lambda}{u}$ by the $\overset{\lambda}{u}'$); in other words, the determinant

$$|C| = c_{\dot{1}\dot{1}} c_{\dot{2}\dot{2}} - c_{\dot{1}\dot{2}} c_{\dot{2}\dot{1}}$$

is invariant. In order that the form (20.4) assume only real values - according to the interpretation of the $\overset{\lambda}{u}, \dot{u}$ as pairs of complex-conjugate variables, the $c_{\dot{1}\dot{1}}$ and $c_{\dot{2}\dot{2}}$ must, therefore, be real and the $c_{\dot{1}\dot{2}}$, $c_{\dot{2}\dot{1}}$ complex-conjugate. We shall henceforth make this

assumption about real behavior. The above form is obviously invariant under the transformations (20.1).

We now introduce the real variables, x, y, z, t by means of the following relations:

$$\begin{aligned} o_{2\dot{1}} &= x + iy \\ o_{1\dot{2}} &= x - iy \\ o_{1\dot{1}} &= z + ct \\ o_{2\dot{2}} &= -z + ct \end{aligned} \qquad (20.5)$$

These new variables undergo linear transformations under the group (20.2), just like the $o_{\lambda\dot{\mu}}$. Since they are real and remain real even after the transformations, the transformation coefficients are also real. Moreover, the quadratic form

$$o_{1\dot{1}}\, o_{2\dot{2}} - o_{1\dot{2}}\, o_{2\dot{1}} = c^2 t^2 - z^2 - x^2 - y^2$$

is invariant; hence, we are dealing with <u>real Lorentz-transformations</u> of the variables x, y, z, t. The Lorentz-transformations thus obtained include, of course, all the spacial rotations, for if in (20.1) and (20.2) one chooses, in particular:

$$\delta = \bar{\alpha}, \quad \gamma = -\bar{\beta}$$

then $2ct = o_{1\dot{1}} + o_{2\dot{2}}$ is left invariant. Also, the transformations (20.1) become unitary and, hence, they transform $\dot{u}^{\dot{1}}$ and $\dot{u}^{\dot{2}}$ (or $\bar{u}^{\dot{1}}$ and $\bar{u}^{\dot{2}}$) just like u_2, $-u_1$. In addition, the form (20.4) is transformed like $-o_{1\dot{2}}\, \dot{u}^{\dot{1}}\dot{u}^{\dot{1}} + (o_{1\dot{1}} - o_{2\dot{2}})\dot{u}^{\dot{1}}\dot{u}^{\dot{2}} + o_{2\dot{1}}\dot{u}^{\dot{2}}\dot{u}^{\dot{2}} = 2(c_0 \dot{u}^{\dot{1}}\dot{u}^{\dot{1}} + c_1 \dot{u}^{\dot{1}}\dot{u}^{\dot{2}} + c_2 \dot{u}^{\dot{2}}\dot{u}^{\dot{2}})$, i.e., the variables x,y,z undergo exactly the transformations given in #16. But among the Lorentz-transformations the following also occurs:

$$\overset{1}{u}' = \alpha \overset{1}{u}$$
$$\overset{2}{u}' = \alpha^{-1} \overset{2}{u},$$

$$\begin{aligned} c'_{11} &= \alpha^2 c_{1\dot{1}} & x' &= x \\ c'_{2\dot{2}} &= \alpha^{-2} c_{2\dot{2}} & y' &= y \\ c'_{1\dot{2}} &= c_{1\dot{2}} & z' &= \tfrac{1}{2}(\alpha^2 + \alpha^{-2})z + \tfrac{1}{2}(\alpha^2 - \alpha^{-2})ct \\ c'_{2\dot{1}} &= c_{2\dot{1}} & ct' &= \tfrac{1}{2}(\alpha^2 - \alpha^{-2})z + \tfrac{1}{2}(\alpha^2 + \alpha^{-2})ct \end{aligned} \qquad (20.6)$$

Here, the new referenoe system (x',y',z',t') moves with an arbitrary velocity $v = c\frac{\alpha^4 - 1}{\alpha^4 + 1}$ in the z-direction. Now, all restricted Lorentz-transformation, (i.e. those which do not reverse the direction of rotation of the space) can be formed by combining this transformation with any arbitrary rotation. Hence, the transformations (20.1) yield all the restricted Lorentz-transformations. Or what amounts to the same, <u>the restricted Lorentz-group is a representation of the group</u> C_2 .

The only transformations (20.1) which yield the identity of the Lorentz-group are given by the matrices:

$$E = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \qquad \text{and} \qquad -E = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}$$

hence, conversely, the group C_2 is a double-valued representation of the restricted Lorentz-group.

<u>2. The Reflection s and the complete Lorentz-group</u>.

As mentioned above, the restricted Lorentz-group may be represented as a linear group in two variables which is double-valued. This representation, however, cannot be extended into a representation of the <u>complete Lorentz-group</u> without raising the number of variables. (The oomplete Lorentz-group is formed from the proper

Lorentz-group by introducing in addition the reflection s:

$$x' = -x,\ y' = -y,\ z' = -z,\ t' = t.$$

Since the reflection s commutes with all true rotations in space, the matrix S (representing the reflection) should also commute with the entire unitary group U_2 ; But U_2 is an irreducible system of transformations. Consequently, S should be a multiple of the unit matrix and commute with the entire group C_2 . But the reflection s transforms the Lorentz-transformation (20.6) with a velocity of the reference system different from and opposite to that due to S. Hence, the matrix S cannot commute with the matrix belonging to (20.6), and we are left with a contradiction.

A representation of the complete Lorentz-group can be attained, however, in four variables, when to the variables u_1, u_2 we add the punctuated quantities $\dot{u}_1, \dot{u}_2$ which are now accepted as new variables and not merely as complex-conjugates to u_1 and u_2 . As already mentioned, the $\dot{u}_1$ and $\dot{u}_2$ transform like the u_2 and $-u_1$, respectively, under the pure spatial rotations. If we set $\dot{u}_1 = v_2$ and $\dot{u}_2 = -v_1$, then the new base-vectors v_1 and v_2 are transformed under spatial rotations exactly like u_1 and u_2 , respectively. We now set up tentatively the following representation for the transformation s which commutes with all rotations:

$$su_\lambda = iv_{\dot{\lambda}}, \quad sv_{\dot{\lambda}} = iu_\lambda \quad (\lambda = 1,2). \tag{20.7}$$

The bi-linear form (20.4) or

$$c_{1\dot{1}}\, u_1 v_2 - c_{1\dot{2}}\, u_1 v_1 + c_{2\dot{1}}\, u_2 v_2 - c_{2\dot{2}}\, u_2 v_1$$

is transformed by the transformation s (20.7) into:

$$c_{\dot{2}\dot{2}}\,\overset{1}{u}\,v_{\dot{2}} + c_{\dot{1}\dot{2}}\,\overset{1}{u}\,v_{\dot{1}} - c_{\dot{2}\dot{1}}\,\overset{2}{u}\,v_{\dot{2}} - c_{\dot{1}\dot{1}}\,\overset{1}{u}\,v_{\dot{1}} \;;$$

whence we get

$$c'_{\dot{1}\dot{1}} = c_{\dot{2}\dot{2}}\;;\quad c'_{\dot{2}\dot{2}} = c_{\dot{1}\dot{1}}\;;\quad c'_{\dot{2}\dot{1}} = -c_{\dot{1}\dot{2}}\;;\quad c'_{\dot{1}\dot{2}} = -c_{\dot{2}\dot{1}}$$

and this on heeding (20.5) is exactly the reflection we are seeking;

$$x' = -x,\quad y' = -y,\quad z' = -z,\quad t' = +t.$$

Each unrestricted Lorentz-transformation can be obtained as the product of a restricted Lorentz-transformation a by the reflection s. On associating with each of these products(as)the product of the corresponding matrices, we obtain a double-valued representation of the complete Lorentz-group in four variables. Now, the representation (20.7) of the reflection s has the disadvantage of having its square equal to -E rather than to +E; that is due to the double-valuedness of the representation. We can do away with this disadvantage (without necessarily losing thereby the representation-properties) by multiplying the representation matrices of all the restricted Lorentz-transformations by -i; eg. we set instead of (20.7)

$$s\,\overset{\lambda}{u} = v_{\dot{\lambda}}\,,\qquad s\,v_{\dot{\lambda}} = \overset{\lambda}{u}.$$

Henceforth we shall do that.

A vector $a_{\dot{1}}\overset{\dot{1}}{u} + a_{\dot{2}}\overset{\dot{2}}{u}$ of the punctuated space, expressed in terms of the new base vectors $v_{\dot{1}} = -\overset{\dot{2}}{u}$, $v_{\dot{2}} = +\overset{\dot{1}}{u}$, has the components $a^{\dot{1}} = -a_{\dot{2}}$, $a^{\dot{2}} = a_{\dot{1}}$. The notation is in agreement with the convention (20.3). The components a_1, a_2, $a^{\dot{1}}$, $a^{\dot{2}}$ of any arbitrary vector $a_1\overset{1}{u} + a_2\overset{2}{u} + a^{\dot{1}}v_{\dot{1}} + a^{\dot{2}}v_{\dot{2}}$ are frequently denoted by a_1, a_2, a_3, a_4. This notation, however, does not permit one to recognize clearly the manner in which

these components have been transformed.

It is perhaps of use to point out in connection with this, that the components a_1, a_2, a_3, a_4 transform in an entirely different (in fact, non-equivalent) manner from the components of a space-time tensor (x,y,z,t). Apart from the double-valuedness of the transformation of the a_ν under a given Lorentz-transformation, the essential difference between the two lies in that the restricted Lorentz-transformations of the space-time tensors form an irreducible system; the transformations belonging to the a_ν, however, break up into two subsystems corresponding to the invariant sub-spaces $(\overset{1}{u}, \overset{2}{u})$ and $\underset{1}{v},\underset{2}{v})$.

3. Spinor - Analysis

The manner of writing the co- and contra-gradient vectors with upper and lower indices possesses, of course, the advantage in that the invariance of certain relations immediately catches the eye. For example, the system of equations

$$a_\lambda = \Sigma\, c_{\lambda\dot{\mu}} \cdot b^{\dot{\mu}}$$

is invariant under the group C_2, because the $c_{\lambda\dot{\mu}}$ (the coefficients of (20.4) transform exactly like the coefficients $c_\lambda c_{\dot{\mu}}$ of a reduced bi-linear form $(c_1 \overset{1}{u} + c_2 \overset{2}{u})\,(c_{\dot{1}} \overset{\dot{1}}{u} + c_{\dot{2}} \overset{\dot{2}}{u})$. That is, the matrix $C = (c_{\lambda\dot{\mu}})$ transforms a binary vector of the type $(b^{\dot{1}}, b^{\dot{2}})$ into one of the type (a_1, a_2). Conversely, the inverse matrix, multiplied by the invariant

$$|C| = c_{1\dot{1}}\, c_{2\dot{2}} - c_{1\dot{2}}\, c_{2\dot{1}}\ ,\ \text{that is:}$$

$$C' = |C|\, C^{-1} = \begin{pmatrix} +c_{22} & -c_{12} \\ -c_{21} & +c_{11} \end{pmatrix}$$

transforms a binary vector of type (a_1, a_2) into one of type (b^1, b^2).

According to (20.5), we have:

$$C = \begin{pmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{pmatrix} = \begin{pmatrix} z + ct & x - iy \\ x + iy & -z + ct \end{pmatrix}$$

whence we get

$$C' = \begin{pmatrix} -z + ct & -(x - iy) \\ -(x+iy) & (z + ct) \end{pmatrix}$$

On introducing the Pauli Matrices :

$$\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} = \sigma_x \quad ; \quad \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} = \sigma_y \quad ; \quad \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} = \sigma_z$$

and denoting the unit matrix by σ_0 , we obtain:

$$C = x\sigma_x + y\sigma_y + z\sigma_z + ct\sigma_0,$$

$$C' = -x\sigma_x - y\sigma_y - z\sigma_z + ct\sigma_0 \qquad (20.9)$$

We denote the variables x,y,z,ct by x^1, x^2, x^3, x^0 , and the matrices $\sigma_x, \sigma_y, \sigma_z, \sigma_0$ by $\sigma_1, \sigma_2, \sigma_3, \sigma_0$. - also the elements of the matrix σ_k by $\sigma_{k\lambda\dot{\mu}}$ ($k = 1,2,3,4$; $\lambda = 1,2$;

$\mu = 1,2$). We then obtain:

$$c_{\lambda\dot{\mu}} = \sum x^k \sigma_{k\lambda\dot{\mu}} \quad . \tag{20.10}$$

Thus, each equation of the form:

$$a_\lambda = \sum_k \sum_{\dot{\mu}} x^k \sigma_{k\lambda\dot{\mu}} b^{\dot{\mu}}$$

is left invariant under all restricted Lorentz-transformations - provided the a_λ and $b^{\dot{\mu}}$ are transformed according to the above-mentioned double-valued representation of the Lorentz-group and the x^k according to the Lorentz-group itself - whereas the pure numerical quantities $\sigma_{k\lambda\dot{\mu}}$ are not at all affected by the transformation.

Likewise, on setting $\sigma_1' = -\sigma_x$, $\sigma_2' = -\sigma_y$, $\sigma_3' = -\sigma_z$, $\sigma_0' = \sigma_0$ then each equation of the form:

$$a^{\dot{\mu}} = \sum_k \sum_\lambda x^k \sigma_k'^{\dot{\mu}\lambda} b_\lambda \tag{20.11}$$

also remains invariant under all restricted Lorentz-transformations.

Moreover, the pair of equations (20.10), (20.11) are left invariant under the reflection s which transforms a_λ into $a^{\dot{\lambda}}$, $b_{\dot{\mu}}$ into $+ b^{\dot{\mu}}$, and x^k into $-x^k$ ($k = 1,2,3$), for under this reflection (20.10) goes over into (20.11) and conversely.

The invariance properties just mentioned obtain, of course, also when we substitute the x^k by any other expressions which transform like the components of a space-time vector, as for example,

$$\frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z}, -\frac{1}{c}\frac{\partial}{\partial t}$$

4. The Infinitesimal Transformations.

We now seek to determine again by the method of the infinitesimal transformations all the differentiable representations of the Lorentz-group. Every restricted representation of the restricted Lorentz-group is simultaneously a representation of the group C_2 and conversely. Hence, we next seek to represent the C_2. We take for the matrix of a transformation of C_2:

$$A = \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix} = \begin{pmatrix} 1 + \alpha_1 + i\alpha_2 & \alpha_3 + i\alpha_4 \\ \alpha_5 + i\alpha_6 & \delta \end{pmatrix} \qquad (20.12)$$

where,

$$\delta = \frac{1 + \beta\gamma}{\alpha} = \frac{1 + (\alpha_3 + i\alpha_4)(\alpha_5 + i\alpha_6)}{1 + \alpha_1 + i\alpha_2}$$

and for the parameter in the vicinity of unity we take the real variables $\alpha_1, \ldots, \alpha_6$. We define the infinitesimal transformations $I_1, \ldots, I_6$, of any arbitrary representation like in #17 and we find, just as over there, that there must exist commutation rules of the form:

$$I_\mu I_\nu - I_\nu I_\mu = \sum_\sigma I_\sigma c^\sigma_{\mu\nu} .$$

The $c^\sigma_{\mu\nu}$ are real numbers which depend only on the composition of the group; they may, therefore, be determined from any representation, e.g., from the matrices of C_2 itself. As for the latter representation, we have according to (20.12):

$$I_1 = \frac{\partial A}{\partial \alpha_1} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \quad I_3 = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \quad I_5 = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}$$

$$I_2 = \frac{\partial A}{\partial \alpha_2} = \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix} \quad I_4 = \begin{pmatrix} 0 & i \\ 0 & 0 \end{pmatrix} \quad I_6 = \begin{pmatrix} 0 & 0 \\ i & 0 \end{pmatrix}$$

We then obtain the following commutation rules with real coefficients:[1]

$$I_1 I_3 - I_3 I_1 = 2I_3 \quad I_1 I_4 - I_4 I_1 = 2I_4 \quad I_2 I_3 - I_3 I_2 = 2I_4$$
$$I_1 I_5 - I_5 I_1 = -2I_5 \quad I_1 I_6 - I_6 I_1 = -2I_6 \quad I_2 I_5 - I_5 I_2 = -2I_6$$
$$I_3 I_5 - I_5 I_3 = I_1 \quad I_3 I_6 - I_6 I_3 = I_2 \quad I_4 I_5 - I_5 I_4 = I_2$$

$$I_2 I_4 - I_4 I_2 = -2I_3 \qquad I_1 I_2 - I_2 I_1 = 0$$
$$I_2 I_6 - I_6 I_2 = 2I_5 \qquad I_3 I_4 - I_4 I_3 = 0$$
$$I_4 I_6 - I_6 I_4 = -I_1 \qquad I_5 I_6 - I_6 I_5 = 0$$

These relations must also be valid for any arbitrary representation. They can be simplified by the introduction of new operators:

$$I_1 + iI_2 = 4A_3; \quad I_3 + iI_4 = 2A_p; \quad I_5 + iI_6 = 2A_q;$$
$$I_1 - iI_2 = 4B_3; \quad I_3 - iI_4 = 2B_p; \quad I_5 - iI_6 = 2B_q.$$

Calculation leads to the conclusion that the A's commute with the B's:

$$A_h B_k - B_k A_h = 0 \quad \text{for } h, k = z, p, q.$$

Moreover,

---

1) That the coefficients must be real, is evident from the general considerations of # 17. The coefficients are uniquely determined only as a result of the requirement of real behavior.

$$\left.\begin{aligned} A_3A_\rho - A_\rho A_3 &= A_\rho \\ A_3A_\sigma - A_\sigma A_3 &= -A_\sigma \\ A_\rho A_\sigma - A_\sigma A_\rho &= 2A_3 \end{aligned}\right\} \qquad \left.\begin{aligned} B_3B_\rho - B_\rho B_3 &= B_\rho \\ B_3B_\sigma - B_\sigma B_3 &= -B_\sigma \\ B_\rho B_\sigma - B_\sigma B_\rho &= 2B_3 \end{aligned}\right\}$$

These are for the A and the B the same commutation rules as (17.5). Hence, a number of conclusions deduced from (17.5) are valid here. Let v_J be a vector belonging to the highest eigenvalue J of A_3. It is then possible to find a whole series of eigenvectors $v_M (-J' \leq M' \leq J')$ which are transformed according to (17.8) by the operators A_k (with A_k instead of L_k). The v_J, belonging to the eigenvalue J, constitute a linear space which is invariant under B_k for they commute with A_3. In the same manner as above, we can find a series of vectors $v_{JM'}$ $(-J' \leq M' \leq J')$ which are transformed by B_k according to (17.8). Now, each of these $v_{JM'}$ yields an entire series of vectors $v_{MM'}$ $(-J \leq M \leq J)$, as the operator A_σ is repeatedly applied. Thus, we find $(2J + 1)(2J' + 1)$ vectors $v_{MM'}$, for which the following are valid:

$$\left.\begin{aligned} A_\rho v_{MM'} &= \sqrt{(J - M)(J + M + 1)}\, v_{M+1,M'} \\ A_\sigma v_{MM'} &= \sqrt{(J + M)(J - M + 1)}\, v_{M-1,M'} \\ A_3 v_{MM'} &= M v_{MM'} \\ B_\rho v_{MM'} &= \sqrt{(J' - M')(J' + M' + 1)}\, v_{M,M'+1} \\ B_\sigma v_{MM'} &= \sqrt{(J' + M')(J' - M' + 1)}\, v_{M,M'-1} \\ B_3 v_{MM'} &= M' v_{MM'} \end{aligned}\right\} \qquad (20.13)$$

The vectors $v_{MM'}$ also determine an irreducible representation of the group C_2. The irreducibility follows readily from the same line of reasoning employed in # 17. If the representation given originally is irreducible, then the $v_{MM'}$ necessarily span the entire space: hence, every irreducible representation is equivalent to one of the

representations $D_{JJ'}$ given by (20.13).

One readily finds a system of quantities which is transformed exactly according to $D_{JJ'}$; we need only set:

$$v_{MM'} = \frac{\overset{1}{u}^{J+M} \overset{2}{u}^{J+M}}{\sqrt{(J+M)!(J-M)!}} \cdot \frac{\overset{1}{u}^{J'+M'} \overset{2}{u}^{J'+M'}}{\sqrt{(J'+M')!(J'-M')!}}.$$

An arbitrary linear-combination of these $v_{MM'}$ is given by the expression

$$c_{\lambda\mu\ldots\nu\dot{\varrho}\dot{\sigma}\ldots\dot{\tau}} \; \overset{\lambda}{u}\overset{\mu}{u} \ldots \overset{\nu}{u}\dot{u}\dot{u} \ldots \dot{u}$$

where the tensor c is symmetric with respect to the 2J indices $\lambda,\ldots,\nu$ and, the 2J' indices $\dot{\varrho},\ldots,\dot{\tau}$. Thus, <u>the tensors c constitute a class of quantities which suffer under the group C_2 an irreducible group of transformations; these classes are also the only ones which (except for equivalenoe) have this property</u>.

I will not enter here into the proof (though not a very difficult one) showing that every representation is completely reducible and that, therefore, all possible quantities may be written as the sums of quantities of the types just enumerated.

According to the above considerations, all types of "quantities" which are in any way linearly transformed under the restricted Lorentz-group are determined. The simplest quantities are the invariants or scalars; next the binary vectors (a_1, a_2) and $(a_{\dot{1}}, a_{\dot{2}})$ (the contragradient vectors are transfromed equivalently to the cogradient ones - according to (20.3). Next in order of complexity are the tensors $c_{\lambda\dot{\mu}}$ which according to (20.4) are equivalent to the space-time tensors. Finally, the most complex of the quantities

are the symmetric tensors $c_{\lambda\mu}$ and $c_{\dot\lambda\dot\mu}$, each having three components, etc.

The overall name for all these quantities is "spinors", for they will later play an important role in the theory of the spinning electron.

The Lorentz-invariant equations can be formulated with the aid of all types of spinors. We sum here over all the superscripts and subscripts (which are either both punctuated or both unpunctuated) in the usual, patient manner. Moreover, there are certain numerical spinors which are of significance and whose components are each invariant. We have already become acquainted with a spinor of that type, namely, the quantity $\sigma_{k\lambda\mu}$ defined by (20.7). (To be sure, it is not a true spinor, for the index k does not run from 1 to 2 but rather from 0 to 3, and it is transformed in the manner of a space-time tensor.) This quantity occurs everywhere in the formulae (like in (20.10)) as a connecting link between spins and space-time vectors. With its aid, we may write (20.4) as follows:

$$c_{\dot\lambda\mu} = \sigma_{k\lambda\dot\mu} x^k \quad .$$

In like manner, we can, by virtue of the quantity σ, associate with every space-time vector or tensor a spinor, e.g.

$$f_{\varkappa\lambda\dot\mu\dot\nu} = \sigma_{k\varkappa\dot\mu}\sigma_{l\lambda\dot\nu} F^{kl} \qquad (20.13)$$

Another useful quantity is the true spinor $\varepsilon^{\lambda\mu}$ with components

$$\varepsilon^{12} = 1, \quad \varepsilon^{21} = -1, \quad \varepsilon^{11} = \varepsilon^{22} = 0.$$

In exactly the same way we define $\varepsilon_{\lambda\mu}$, $\varepsilon^{\dot\lambda\dot\mu}$ and $\varepsilon_{\dot\lambda\dot\mu}$. With the aid of these quantities we form invariants such as:

$$\varepsilon^{\lambda\mu} a_\lambda b_\mu = a_1 b_2 - a_2 b_1 ,$$

and we write (20.3) in the invariant form:

$$b^\lambda = \varepsilon^{\lambda\mu} b_\mu .$$

We shall not enter into the proof[1] showing that the symbols ε and σ are sufficient in the sense that we may consider every invariant system of equations among the spinors and space-time tensors as invariant. The following example should suffice. The scalar product $x_k y^k$ of two space-time vectors can be represented[2] (in terms of the ε-symbolism) by:

$$x_k y^k = -\tfrac{1}{2} \varepsilon^{\kappa\lambda} \varepsilon^{\dot\mu\dot\nu} \xi_{\kappa\dot\mu} \eta_{\lambda\dot\nu}$$

where

$$\xi_{\lambda\dot\mu} = \sigma_{k\lambda\dot\mu} x^k$$

and

$$\eta_{\lambda\dot\mu} = \sigma_{k\lambda\dot\mu} y^k$$

are the associated spinors.

---

1) In brief, the proof proceeds along the following lines. First, let us replace the space-time vectors and tensors by their equivalent spinors (as in (20.13). Then, the σ-symbols are no longer necessary. Now, by setting covariants equal to zero, we can obtain any and all invariant systems of equations (Gram's Theorem). Moreover, all covariants of binary tensors are composed of 'linear factors' $a_\mu x^\mu$ or $a_\dot\mu y^\dot\mu$ and "bracket factors" - $(ab) = a_1 b_2 - a_2 b_1 = \varepsilon^{\lambda\mu} a_\lambda b_\mu$ or $(\dot a \dot b) = \varepsilon^{\dot\lambda\dot\mu} a_\dot\lambda b_\dot\mu$ - in line with the symbolic breaking apart of the tensors. From here then the assertion follows. As for the theorems which we borrowed from the theory of invariants cf. R. Weitzenbroeck, Invarianten Theorie, Groningen 1923.

2) Uhlenbeck and Laporte render a more detailed treatment of spinor analysis involving many examples and applications, the Maxwell Field Equations being just one of many examples. See Phys.Rev. vol 37 (1931) p. 1380.

IV. THE SPINNING ELECTRON.

21. The Spin

We have seen in # 6 that the Schroedinger wave equation for the electron with the magnetic perturbation term in the energy operator

$$\frac{e}{\mu c}\vec{A}\cdot\vec{p} = \varkappa\,\vec{H}\cdot\vec{L}$$

(μ = mass, e = charge, $\varkappa = \frac{eh}{2\mu c}$)

is able to explain only the normal Zeeman-effect (as it occurs in connection with the singlet terms); it cannot explain, however, the anomalous Zeeman effect. It seems therefore indispensable to the explanation of the anomalous Zeeman effect to assume (besides the magnetic moment of the orbital motion which is always proportional to the mechanical moment) the presence of still another magnetic moment in the atom. According to the hypothesis of Uhlenbeck and Goudsmit[3] this moment is due to the so-called spin - the restricted angular momentum of the "spinning" electron.

We observe a direct mechanical effect of the electron spin in connection with the demagnetzation of ferromagnetic substances. The experiment shows that a variation in the mechanical angular momentum is to the variation in the magnetic momentum as 1 : $\frac{e}{\mu c}$ or as h : $2\varkappa$, instead of as h : $\varkappa$ as the case should be were the magnetization to depend upon the orbital motion of the electron alone.

---

3) Uhlenbeck and Goudsmit:
Naturwissenschaften vol.13 (1925) p.953
Nature vol.117(1926) p.264

In order to explain this anomaly by attributing the ferromagnetism to the spin alone, one must assume that the magnetic moment of the spinning electron is twice as large as the magnetic moment of an orbital motion having the same mechanical angular momentum.

That the spin is quantized (just like the mechanical angular momentum $\vec{L}$), i.e. that its components along any direction whatever are capable of assuming only discreet values, is borne out by the experiment of Stern and Gerlach. A ray of silver atoms in the ground state ($\ell = 0$) is passed along in the x-direction under the influence of a magnetic field, the magnitude of which varies sharply along the z-direction. This field exerts a force $\frac{\partial H_z}{\partial z}\mu_z$ on a magnet whose moment in the z-direction has the value μ_z. The ray is split into two sub-rays corresponding to the values $\mu_z = \pm \mathcal{H}$. On making the plausible assumption that only one electron is responsible for the magnetic moment (the spins of the other electrons mutually compensating for one another[1]) it turns out that the magnetic moment of the electron along any direction can assume only the values $\pm \mathcal{H}$ and the mechanical angular momentum (spin) only the values $\pm \frac{1}{2}$ h.

The quantization of the spin also enables one to explain the multiplet splitting of the spectral terms. In the simplest case (that of the alkali-metals) where only one electron plays the important role, the phenomenon looks somewhat as follows. The terms agree to a first approximation with the energy values calculated in #4 for the electron in the central field; all of them, however, (with the

---

1) The assumption is plausible for the reason that the silver ion Ag^+ (as well as the Na^+, K^+, etc) in its ground state, does not exhibit the Zeeman-effect.

exception of the s-terms, $\ell = 0$) consist of a fine doublet. Upon application of a non-spherically symmetric perturbation field, one of the terms in the doublet splits into $2\ell + 2$ terms and the other into 2ℓ terms, whereas in the "spin-less" theory, one would expect a splitting into $2\ell + 1$ terms. The two terms can be differentiated from one another by means of a quantum-number j; for the $(2\ell + 2)$-fold degenerate term j takes on the value $\ell + \frac{1}{2}$ and for the other term, the value $\ell - \frac{1}{2}$. For purposes of a (heuristic) illustration of the state of affairs, one imagines that the orbital angular momentum $h\ell$ and the spin angular momentum $\frac{1}{2}h$ combine to give a resultant hj with $j = \ell \pm \frac{1}{2}$. This total angular momentum, therefore, determines the degree of the degeneracy $2j + 1$, just as in the "spin-less" case the angular momentum $h\ell$ determines the degree of the degeneracy $2\ell + 1$.

One further imagines that the two terms $j = \ell + \frac{1}{2}$ and $j = \ell - \frac{1}{2}$ split apart as a result of the interaction of the spin $\frac{1}{2}h$ with the orbital angular momentum $h\ell$. The exact justification for this "vector-scheme" will be given later on. For the present, we shall be content with stating that vector-schemes of a similar type also lead (in a qualitative way) to the correct multiplet-splittings of complicated spectra.

In the case of a weak magnetic field Landé found empirically that the terms $j = \ell \pm \frac{1}{2}$ split into $2j + 1$ equidistant components; these components differ from the unperturbed terms by the displacements:

$$g\varkappa H_3 m \qquad \left(g = \frac{j+\frac{1}{2}}{\ell+\frac{1}{2}};\quad m = j,\ j-1,\ldots,-j\right) \qquad (21.1)$$

In the case $\ell = 0$ where the entire angular momentum is due to the

spin, $m = \pm \frac{1}{2}$ and $g = 2$. The values $m = \pm \frac{1}{2}$ multiplied by h are precisely the possible values of the z-components of the angular momentum and the factor $g = 2$ confirms anew that the magnetic moment $2\varkappa m$ corresponds to the angular momentum hm.

Thus, the following are the hypotheses we have employed:

1. The electron has a restricted mechanical angular momentum or spin $\frac{1}{2}h$, whose components along eaoh fixed space-direction can only take on the values $\pm \frac{1}{2}h$.

2. The energy effect of the spin is small as compared to that of charge and mass as long as there exists no external magnetic field.

3. To the spin $\frac{1}{2}h$ there corresponds a magnetic moment $\varkappa$.

22. The Wave Function of the Spinning Electron.

We now attempt to translate the hypotheses above into the language of wave-mechanics! Kinematically, the existence of spin implies that an electron has in addition to its three degrees of freedom x,y,z at least one more degree of freedom - that of the spin. It is represented by the z-component of the spin-angular momentum, measured in multiples of $\frac{1}{2}h$. This z-component is a variable σ_z which, according to hypothesis 1 of the previous section, can assume only the values +1 and -1. We now set up according to Pauli[1], a wave-function

$$\psi(x,y,z,\sigma_z) = \psi(q,\sigma_z)$$

where the space co-ordinates q extend over the entire space and σ_z assumes only the values +1 and -1. This function is equivalent to the pair of functions $\psi_{} = \psi(q,1)$; $\psi_{} = \psi(q,-1)$, or as we prefer to say, it is a wave function with two components

---

1) Pauli, W.: Z.f. Physik vol.43 (1927) p.601.

φ_1, φ_2 - which are the usual space-functions.

According to the statistical interpretation of wave-mechanics, $\int \varphi_1 \bar{\varphi}_1 dv$, (integrated over a portion of space) is proportional to the probability of finding the electron in that space with a spin that is directed parallel to the positive z-axis. Likewise, $\int \varphi_2 \bar{\varphi}_2 dv$ is proportional to the probability of finding it there with the opposite spin. Finally, the sum

$$\int (\varphi_1 \bar{\varphi}_1 + \varphi_2 \bar{\varphi}_2) dv$$

represents the probability of finding the electron at all in this space.

We may consider the two components φ_1, φ_2 of the ψ-function to be the components of a vector in a two-dimensional vector-space, called the "spin-space". The constant vectors of this vector-space are pairs of numbers; in other words, they are functions of the spin co-ordinates alone. If now we introduce into this vector-space any two constant base-vectors u_1, u_2, then all vectors may be expressed in terms of these:

$$\psi = w_1 u_1 + w_2 u_2 . \qquad (22.1)$$

The w_λ may be dependent on the space-co-ordinates q; they arise due to a linear transformation with constant coefficients of the previous φ_1, φ_2.

According to hypothesis 2 of # 21, the wave equation must, to a first approximation, read exactly as in the case of the Schroedinger ψ-function:

$$H w_\gamma = E w_\gamma ,$$

(i.e. for each ψ-component ψ_1, ψ_2 or w_1, w_2).

In the second approximation, we add to the H-operator smaller perturbation terms which operate on the spin co-ordinates σ . To a first approximation, therefore, we may choose for w_1 and w_2 in (22.1) two arbitrary eigenfunctions of the Schroedinger equation corresponding to the same energy-value. The number of linearly independent ψ corresponding to each energy level is now doubled. If $\varphi^{(1)}$,....,$\varphi^{(k)}$ are the spin-less eigenfunctions belonging to the eigenvalue E, then

$$\varphi^{(1)}\ u_1,\ \varphi^{(2)}\ u_1,\ldots,\ \varphi^{(k)}\ u_1,$$
$$\varphi^{(1)}\ u_2,\ \varphi^{(2)}\ u_2,\ldots,\ \varphi^{(k)}\ u_2, \qquad (22.2)$$

are the 2h linearly-independent eigenfunctions belonging to the same level. It is these which are capable of splitting apart when subjected to the spin-perturbation. This doubling of the degree of degeneracy is in harmony with the experimental facts already mentioned in connection with the doublet splitting of the alkali-terms. For the present, therefore, we are not compelled to assume any further degrees of freedom besides the σ_3 .

We must now investigate how the rotations of the set of axes transform[1] the function $\varphi(q,\sigma_3)$, which, up to now, was defined only in terms of a special choice of the z-axis. We carry out a rotation D^{-1} on the co-ordinate system - or what amounts to the same, we hold the co-ordinates fixed and carry out the rotation D

---

1) The derivation of these transformation formulae by means of group theory was first developed by J.v. Neuman and E. Wigner, Z.f.Physik vol. 47 (1927) p.203.

on the space of the spinning electron. Then the pure spin-functions u_1 and u_2 go over into the pure spin-functions Du_1 and Du_2 :

$$\left.\begin{aligned} Du_1 &= u_1\alpha_{11} + u_2\alpha_{21} \\ Du_2 &= u_1\alpha_{12} + u_2\alpha_{22} \end{aligned}\right\} \qquad (22.3)$$

We further assume that a product $w(q)u_\lambda$ of a pure space-function by a pure spin function is transformed in such a manner that the two factors are transformed independently of each other. Namely, the u_λ are transformed in accordance with (22.3) and the w according to the usual rules for the transformation of space functions:

$$\left.\begin{aligned} D(w(q)u_\lambda) &= w(D^{-1}q)Du_\lambda \\ Du_\lambda &= \Sigma\, u_\mu \alpha_{\mu\lambda} \end{aligned}\right\} \qquad (22.4)$$

Finally, we assume that the sum $w_1u_1 + w_2u_2$ is transformed into the sum $D(w_1u_1) + D(w_2u_2)$. On setting:

$$D(w_1u_1) + D(w_2u_2) = w_1'u_1 + w_2'u_2 ,$$

we obtain for the new components $w_1'w_1'$:

$$\begin{aligned} w_1'(q) &= \alpha_{11}w_1(D^{-1}q) + \alpha_{12}w_2(D^{-1}q), \\ w_2'(q) &= \alpha_{21}w_1(D^{-1}q) + \alpha_{22}w_2(D^{-1}q). \end{aligned}$$

The coefficients $\alpha_{\mu\lambda}$ depend only on the rotation D selected. Actually, they are determined to within a common factor λ, for the φ-function multiplied by λ represents the same state as the original one. Yet the coefficients can be so normalized that the determinant $\alpha_{11}\alpha_{22} - \alpha_{12}\alpha_{21}$ is always equal to 1 (cf. #16, small type). Thus they are determined to within a factor ± 1.

We allow the unit-matrix $\alpha_{\lambda\mu} = \delta_{\lambda\mu}$ to correspond to the identity D = 1. We further assume that, in the vicinity of the identity, the coefficients $\alpha_{\lambda\mu}$ depend in a continuous and differentiable manner on the parameters of the rotation D. To the product of two rotations there corresponds necessarily (to within an eventual factor λ which after the normalization can be only $\pm$ 1) the product of the corresponding transformations. Hence, we have before us in (22.3) a representation of the rotation group (which is, at most double-valued). It satisfies all the conditions given in # 17.

According to # 17, however, we obtain such a representation of the rotation group (barring equivalence) only through two-by-two matrices. For instance, we obtain the double-valued representation $D_{\frac{1}{2}}$ through the unitary matrices

$$\begin{pmatrix} \alpha & \beta \\ -\bar{\beta} & \bar{\alpha} \end{pmatrix}$$

of the determinant 1, which is given explicitly by (17.8). This means that for an appropriate choice of the base vectors u_1, u_2, our representation becomes identical with the representation $D_{\frac{1}{2}}$.

From this result we immediately obtain an exact qualitative explanation of the doublet-splitting of the alkali-terms. If we select for $\varphi^{(1)}, \ldots, \varphi^{(k)}$ in (22.2) the $2\ell + 1$ eigenfunctions $\varphi_\ell^{(m)}$ corresponding to a spin-less alkali-term (which also transform according to D_ℓ) - then, obviously, the $2\ell + 1$ products (22.2) transform according to the product-representation $D_{\frac{1}{2}} \times D_\ell$. However,

$$D_{\frac{1}{2}} \times D_\ell = D_{\ell+\frac{1}{2}} + D_{\ell-\frac{1}{2}} \quad (\text{or} = D_{\frac{1}{2}} \text{ for } \ell = 0).$$

Let us consider the spin-perturbation! (Its formula must, of course, be invariant under rotations.) It is then seen that only

the terms $D_{l+\frac{1}{2}}$ and $D_{l-\frac{1}{2}}$ are capable of splitting apart; a further splitting is impossible. The term $D_{l+\frac{1}{2}}$ is $(2l+2)$-fold degenerate; the other is $2l$-fold degenerate. This degeneracy can be removed only by non-spherically symmetric perturbation - a fact which is in agreement with the experiments. The number $l \pm \frac{1}{2}$, associated with the representation, is denoted by j and is known as the <u>inner quantum number</u> of the electron.

For the sake of completing the transformation formulae for the spin functions, we must yet state how these spin functions transform under the reflection s:

$$x' = -x, \quad y' = -y, \quad z' = -z.$$

We assume that the quantities u_1, u_2 are linearly transformed by s, like in (22.3), the matrix of transformation being S. Since the reflection s commutes with all rotations D, then the matrix S must also commute with the representation $D_{\frac{1}{2}}$. But $D_{\frac{1}{2}}$ is irreducible; hence, S is a multiple of the unit matrix

$$S = \lambda E.$$

The value of λ is entirely arbitrary since any φ-function multiplied by λ represents the same state as the original φ-function. We choose the simplest value of λ, namely, $\lambda = 1$. Consequently, <u>to the reflection s there corresponds the identity transformation of the quantities u_1, u_2.</u>

In the subsequent theory we shall need the operators for the components of the angular momentum. The operators

$$hL_3' = \frac{h}{i}\left(x\frac{\partial}{\partial y} - y\frac{\partial}{\partial x}\right), \text{ etc.} \qquad (22.5)$$

employed previously yield only the angular momentum of the orbital motion and not the spin. We seek, therefore, an additional term in the expression above, to give us the total angular momentum. We remember from #6 (in the case without spin) that the operators L_x, L_y, L_z are the i-fold infinitesimal rotation operators I_x, I_y, I_z. Now, let us consider the case of the spinning electron and let us, here too, form for the transformations (22.4) the infinitesimal rotations I_x, I_y, I_z or I_1, I_2, I_3 :

$$I_\varkappa = \left[\frac{\partial}{\partial\alpha_\varkappa} D(\alpha_1\alpha_2\alpha_3)\right]_{\alpha=0} . \quad (\varkappa = 1,2,3)$$

On applying these infinitesimal rotations to the products $w(q)u_\lambda$, we have by the rule of differentiation for products:

$$I_\varkappa (w(q)u_\lambda) = (I_\varkappa w(q))u_\lambda + w(q) I_\varkappa u_\lambda ,$$

or, written in a different notation:

$$I_\varkappa = I'_\varkappa + I''_\varkappa ,$$

where $I'_\varkappa$ is the operator of an infinitesimal rotation which is to be carried out only on the $w(q)$, i.e.

$$-I'_1 = y\frac{\partial}{\partial z} - z\frac{\partial}{\partial y} , \qquad -I'_2 = z\frac{\partial}{\partial x} - x\frac{\partial}{\partial z} ,$$

$$-I'_3 = x\frac{\partial}{\partial y} - y\frac{\partial}{\partial x} .$$

$I''_\varkappa$, however, is the operator of an infinitesimal rotation to be carried out only upon the u_λ. The expression for $I''_\varkappa$ can be taken from (17.8) with $J = \frac{1}{2}$:

$$\begin{aligned} &I''_x u_1 = -\tfrac{1}{2} iu_2, \quad I''_y u_1 = +\tfrac{1}{2} u_2, \quad I''_z u_1 = -\tfrac{1}{2} iu_1, \\ &I''_x u_2 = -\tfrac{1}{2} iu_1, \quad I''_y u_2 = -\tfrac{1}{2} u_1, \quad I''_z u_2 = \tfrac{1}{2} iu_2. \end{aligned} \qquad (22.6)$$

We now set up for the components hM_x, hM_y, hM_z of the angular momentum hM, the operators I_x multiplied by hi:

$$M_x = iI_x = L_x + S_x \; ; \quad L_x = iI'_x \; ; \quad S_x = iI''_x \; .$$

The first component $\vec{L}$ of the vector $\vec{M}$ is the angular momentum of the orbital motion; the second one - $\vec{S}$ is the spin. The addition of the latter term for the total angular momentum is justified by the fact that all three components M_x, M_y, M_z obviously commute with the spherically symmetric energy-operator - a fact which leads to the conservation principle for these components. Since this principle of conservation is the basis of all measurements of the angular momentum, it alone is sufficient for justifying the occurrence of the additional term in $\vec{M}$.

The components M_x, M_y, M_z of $\vec{M}$ are exactly the operators iI_x, iI_y, iI_z used in the derivation of the representations D_J and denoted there by L_x, L_y, L_z, respectively. It is true, therefore, that for a family of eigenfunctions which transform in accordance with D_J, the operator $\vec{M}^2 = M_x^2 + M_y^2 + M_z^2$ has the eigenvalue $J(J+1)$ and the operator M_z the eigenvalues M (= J, J-1,....,-J). Consequently, the inner quantum number j of the doublet term, multiplied by h, (which determines the manner of transformation D_j of the eigenfunctions) may be called the magnitude of the angular momentum. We have already done this in the vector scheme of the previous section.

The S_x are linear operators in spin space which according to (22.6) can be represented by the matrices:

$$S_x = \frac{1}{2}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \quad S_y = \frac{1}{2}\begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \quad S_z = \frac{1}{2}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \qquad (22.7)$$

We have already encountered these matrices in #20. They are the Pauli matrices $\sigma_x, \sigma_y, \sigma_z$ and are components of the double spin vector $2\vec{S}$.

The formulae thus far are valid for a certain choice of the base-vectors u_1, u_2; in fact, as can be seen from (22.7), u_1 and u_2 are eigenvectors of the operator S_z. In other words, u_1 and u_2 represent the states in which the moment of the momentum hS_z assumes the discreet values $\frac{1}{2}h$ and $-\frac{1}{2}h$, respectively. From here it follows when taking u_1 and u_2 as functions of the spin co-ordinate σ_z :

$$u_1(1) = \varrho_1 \neq 0 \qquad u_1(-1) = 0$$
$$u_2(1) = 0 \qquad u_2(-1) = \varrho_2 \neq 0.$$

Hence, the function $\psi = w_1 u_1 + w_2 u_2$ has the values:

$$\psi_1 = \psi(q,1) = w_1(q)\varrho_1$$
$$\psi_2 = \psi(q,-1) = w_2(q)\varrho_2$$

$\bar{w}_1 w_1 + \bar{w}_2 w_2$ remains invariant for all rotations. That holds also for $\bar{\psi}_1 \psi_1 + \bar{\psi}_2 \psi_2 = |\varrho_1|^2 \bar{w}_1 w_1 + |\varrho_2|^2 \bar{w}_2 w_2$ - in accordance with the physical significance of the ψ_ν. Hence, $|\varrho_1| = |\varrho_2|$[1]. Since the magnitude of the common factor of u_1 and u_2 is of little consequence, we are at liberty to choose $|\varrho_1| = |\varrho_2| = 1$. And since by definition ψ_1 and ψ_2 are determined to within the phase factors $e^{i\theta_1}$ and $e^{i\theta_2}$ (see beginning of this section), we may even choose $\varrho_1 = \varrho_2 = 1$. It follows then that $\psi_1 = w_1$, $\psi_2 = w_2$. In other words, the difference which was made in the beginning between the

---

1) Else, the invariance of the individual terms $\bar{w}_1 w_1$ and $\bar{w}_2 w_2$ would follow from the invariance of the two forms $\bar{w}_1 w_1 = \bar{w}_2 w_2$ and $\bar{w}_1 w_1 |\varrho_1|^2 + \bar{w}_2 w_2 |\varrho_2|^2$, which is contrary to fact.

φ_ν and the w_ν no longer exists. We shall therefore write henoeforth φ_ν instead of w_ν .

We are as yet unable to set up the wave-equation for the pair of functions (φ_1, φ_2) because we still do not know the addition terms which are responsible for the spin-perturbation (doublet splitting). We may, however, write down the magnetic addition term responsible for the anomalous Zeeman effect (on the basis of hypothesis 3 (#21) . The addition-term for a magnet the moment of which is $\frac{2\varkappa}{h}$ as large as the angular momentum $h\hat{S}$, is obviously:

$$2\varkappa(\hat{H}\hat{S}) \qquad (22.8)$$

or for a magnetic field in the z-direction:

$$\varkappa H_3 \sigma_3 \quad ; \quad \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} *$$

It is now only a matter of pure calculation to determine from here the extent of the Zeeman effect. We refer the reader to #25 where the calculation is applied to a multiple electron problem.

---

* There might be some confusion due to the fact that the matrix σ_3 is denoted by exactly the same symbol used for the spin variable σ_3 of the beginning of this section. On closer inspection, however, one easily verifies that operating with the matrix σ_3 is tantamount to multiplying the function $\varphi(q,\sigma_3)$ by σ_3 . Hence, no danger is involved in the interchanging of the two symbols σ_3 .

23. The Lorentz-Invariant Wave Equation of Dirac.

In #22 we purposely formulated the transformation properties of the wave function of the spinning electron in such a manner that the formulation was not limited in its application to any specific wave equation. Consequently, the results above assume a compelling character, and are also applicable to the many-electron problem. In the case of the one-electron problem, P.A.M.Dirac[1] has set up an equation which, like the relativistic Schroedinger Equation, is invariant under Lorentz-transformations. Above all, it leads in a rather natural manner to the correct magnetic spin effect (22.8) and electric spin perturbation. These effects are, of course, responsible for the doublet-splitting of the alkali and hydrogen terms.

As is known, the relativistic Schroedinger equation is:

$$(c^{-2} d_t^2 - d_x^2 - d_y^2 - d_z^2)\Psi = \mu^2 c^2 \Psi , \qquad (23.1)$$

where

$$\begin{aligned} d_x &= p_x + \frac{e}{c} A_x \qquad \left(p_x = \frac{h}{i} \frac{\partial}{\partial x} \quad \text{etc.}\right) \\ d_t &= -ih \frac{\partial}{\partial t} - e\varphi \end{aligned} \qquad (23.2)$$

and where φ is the (scalar) electrical potential and $\vec{A}$ the magnetic potential, i.e.

$$\begin{aligned} \vec{E} &= -\nabla\varphi - \frac{1}{c}\frac{\partial \vec{A}}{\partial t} \\ \vec{H} &= \text{rot } \vec{A} \\ \text{div } \vec{A} + \frac{1}{c}\frac{\partial \varphi}{\partial t} &= 0. \end{aligned}$$

---

1) Dirac, P.A.M.: Proc.Roy.Soc.(A) vol.117 (1928) p.610; vol.118 (1928) p.351. Darwin, C.G. ibid. vol.118 p.654.

If in (23.1) we set:

$$\Psi = e^{-ih^{-1}(\mu c^2 + E)t}\,\varphi(x,y,z)$$

and we divide by $2\mu e^{-ih^{-1}(\mu c^2+E)t}$ then we obtain:

$$-\frac{h^2}{2\mu}\Delta\varphi + \frac{e}{\mu c}(\vec{A}\vec{p})\varphi - (E + e\varphi)\varphi + \frac{1}{2\mu c^2}(e^2\vec{A}^2 - (E + e\varphi)^2)\varphi = 0. \qquad (23.3)$$

Except for the last term (the " relativity correction") equation (23.3) agrees with the Schroedinger equation without spin, which we have been using until now. Unfortunately, however, equation (23.3) does not have the exact form of an eigenvalue problem since E occurs as a square, which in turn is due to the fact that the original equation (23.1) is a differential equation of the second order with respect to time. It is due to Dirac that the equation was transformed and modified in such a manner that it becomes one of first order.

In order to arrive at the Dirac equation, we first replace the function Ψ in (23.1) by a pair of functions (Ψ_1, Ψ_2) in accordance with #22. Next, we seek to split the operator on the left hand side of the equation

$$c^{-2}\,d_t^2 - d_x^2 - d_y^2 - d_z^2$$

into two factors. We accomplish that with the aid of the two-by-two matrices σ_x, σ_y, σ_z (20), (not mentioning the occurrence of small addition terms to which we shall yet come back later on). Thus, by virtue of the easily verified relations

$$\sigma_x^2 = 1 \qquad \sigma_y\sigma_z = i\sigma_x \qquad \sigma_z\sigma_y = -i\sigma_x$$
$$\sigma_y^2 = 1 \qquad \sigma_z\sigma_x = i\sigma_y \qquad \sigma_x\sigma_z = -i\sigma_y$$
$$\sigma_z^2 = 1 \qquad \sigma_x\sigma_y = i\sigma_z \qquad \sigma_y\sigma_x = -i\sigma_z$$

we achieve the factorization in the following manner:

$$c^{-2}d_t^2 - d_x^2 - d_y^2 - d_z^2$$
$$= (c^{-1}d_t - d_x\sigma_x - d_y\sigma_y - d_z\sigma_z)(c^{-1}d_t + d_x\sigma_x + d_y\sigma_y + d_z\sigma_z).$$

The factorization is valid so long as we assume the operators d_t, d_x, d_y, d_z to commute with one another. That, in turn, holds only for the case of constant potentials $\vec{A}, \varphi$. Consequently, we obtain a wave equation different from (23.1), i.e.

$$(c^{-1}d_t - d_x\sigma_x - d_y\sigma_y - d_z\sigma_z)(c^{-1}d_t + d_x\sigma_x + d_y\sigma_y + d_z\sigma_z)\Psi = \mu^2c^2\Psi \qquad (23.4)$$

which agrees with (23.1) only for the case of constant potentials. We next assume tentatively that the equation thus factored is the proper one. In the case of variable potentials d_x, d_y, d_z do not commute with one another; we have instead:

$$\left.\begin{aligned}
d_yd_z - d_zd_y &= \frac{he}{ic}\left(\frac{\partial A_z}{\partial y} - \frac{\partial A_y}{\partial z}\right) = \frac{he}{ic}H_x\\
d_zd_x - d_xd_z &= \frac{he}{ic}\left(\frac{\partial A_x}{\partial z} - \frac{\partial A_z}{\partial x}\right) = \frac{he}{ic}H_y\\
d_xd_y - d_yd_x &= \frac{he}{ic}\left(\frac{\partial A_y}{\partial x} - \frac{\partial A_x}{\partial y}\right) = \frac{he}{ic}H_z\\
d_td_x - d_xd_t &= \frac{he}{i}\left(\frac{1}{c}\frac{\partial A_x}{\partial t} + \frac{\partial \varphi}{\partial x}\right) = \frac{he}{i}E_x\\
d_td_y - d_yd_t &= \frac{he}{i}\left(\frac{1}{c}\frac{\partial A_y}{\partial t} + \frac{\partial \varphi}{\partial y}\right) = \frac{he}{i}E_y\\
d_td_z - d_zd_t &= \frac{he}{i}\left(\frac{1}{c}\frac{\partial A_z}{\partial t} + \frac{\partial \varphi}{\partial z}\right) = \frac{he}{i}E_z
\end{aligned}\right\} \qquad (23.5)$$

When we take these commutation rules into account in (23.4), we obtain in the previous wave-equation (23.1) the following addition terms (on the left-hand side):

$$-\frac{he}{ic}(E_x\sigma_x + E_y\sigma_y + E_z\sigma_z)\Psi - \frac{he}{c}(H_x\sigma_x + H_y\sigma_y + H_z\sigma_z)\Psi.$$

If we wish to have the corresponding addition terms in (23.3) we must divide this expression by -2μ. And if we again introduce

the spin-vector $\vec{S}$ with components $\frac{1}{2}\sigma_x$, $\frac{1}{2}\sigma_y$, $\frac{1}{2}\sigma_z$ (cf. (22.7)), then the addition terms in (23.3) become:

$$\frac{he}{i\mu c}(\vec{E}\vec{S})\psi + \frac{he}{\mu c}(\vec{H}\vec{S})\psi. \tag{23.6}$$

The magnetic addition term is exactly that of (22.8), a fact which speaks in favor of the factorization (23.4). The electric addition term leads to the "spin-perturbation" of the terms in the absence of an external magnetic field.

Equation (23.4) is obviously equivalent to the following pair of equations:

$$\left.\begin{aligned}\left(\tfrac{1}{c}d_t + d_x\sigma_x + d_y\sigma_y + d_z\sigma_z\right)\Psi &= -\mu c\dot{\Psi}\\ \left(\tfrac{1}{c}d_t - d_x\sigma_x - d_y\sigma_y - d_z\sigma_z\right)\dot{\Psi} &= -\mu c\Psi\end{aligned}\right\} \tag{23.7}$$

where $\dot{\Psi}$ is some other function with the components $\Psi^{\dot{i}}$ and $\Psi^{\dot{i}}$[1].

The relativistic invariance of equation (23.7) becomes immediately evident, upon introducing the notations of # 20 and setting $\frac{1}{c}d_t = d_0 = -d^0$, $d_x = d_1 = d^1$, etc. The equations then become:

$$\left.\begin{aligned}d^k\sigma_k^{\dot{\nu}\lambda}\Psi_\lambda &= \mu c\,\Psi^{\dot{\nu}}\\ d^k\sigma_{k\lambda\dot{\nu}}\Psi^{\dot{\nu}} &= \mu c\,\Psi_\lambda\end{aligned}\right\} \tag{23.8}$$

The invariance of this pair of equations under restricted and unrestricted Lorentz transformations was demonstrated in # 20.

We already pointed out in # 20 that it is necessary to introduce a second pair of components $\Psi^{\dot{\nu}}$ (in addition to the Ψ_λ) if

---

1) The function Ψ then satisfies the differential equation of the second order which is obtained from (23.4) by interchanging the two factors on the left-hand side. A change of sign in the first term of (23.6) corresponds to this interchange.

one wishes to extend the representation C_2 of the restricted Lorentz-group to a representation of the complete Lorentz-group. Expressed in another manner that means that the introduction of the Ψ^* is necessary in order that the wave equation be not only invariant under restricted Lorentz-transformations but also under spacial reflections. Moreover, there is the additional advantage of having the wave-equation (23.8) linear in $\frac{\partial}{\partial t}$ and assume the form $(\frac{1}{i}\frac{\partial}{\partial t} + H)\Psi = 0$, where H is a linear self-adjoint operator. Accordingly, the stationary states are given by an equation in the form of a linear self-adjoint eigenvalue-problem $H\varphi = E\varphi$, whose eigenvalues then are definitely real. All these arguments point to the soundness of the Dirac wave equation. A further confirmation of the equation is due to the derivation of the Hydrogen-fine structure in the next section (# 24).

A problem thus far unsolved consists in that the Dirac equation (just like the relativistic Schroedinger equation) not only possesses positive eigenvalues but also negative values of the order of magnitude $-mc^2$. To the latter no physical significance can be ascribed. This difficulty is intimately connected with the one which consists in that the relativistic Ψ -function has four components instead of two - the implication being that the electron must have, in addition to its spin - degree of freedom, a still further degree of freedom[1], thus far undetected.

In many investigations it is convenient to introduce instead of the four components Ψ_λ, Ψ_μ , four other components Ψ^s_λ, Ψ^a_λ,

---

1) Compare to this, E. Schroedinger: Berl.Ber.1931, as well as V. Fock: Z.f. Physik vol.68 (1931) p.522-534.

by means of the relations:

$$\Psi^s_\lambda = \Psi_\lambda + \Psi^\lambda$$
$$\Psi^a_\lambda = \Psi_\lambda - \Psi^\lambda \qquad (\lambda = 1,2).$$

In exactly the same manner that the Ψ_λ, Ψ^λ correspond to the breaking apart of the four-dimensional vector space into irreducible sub-spaces under the restricted Lorentz-group, the Ψ^s_λ, Ψ^a_λ correspond to the breaking apart of the vector space into irreducible subspaces under the rotation-reflection group. Namely, under rotations, the Ψ^s_λ, Ψ^a_λ transform exactly like the Ψ_λ, and Ψ^λ, whereas under the reflection s, the following obtains:

$$s\,\Psi^s_\lambda = \Psi^s_\lambda \quad ; \quad s\,\Psi^a_\lambda = -\Psi^a_\lambda.$$

On multiplying the two equations (23.7) by c and subsequently adding and subtracting them we obtain:

$$\left.\begin{aligned}(d_t + \mu c^2)\Psi^a + c(d_x\sigma_x + d_y\sigma_y + d_z\sigma_z)\Psi^a = 0\\ (d_t - \mu c^2)\Psi^a + c(d_x\sigma_x + d_y\sigma_y + d_z\sigma_z)\Psi^s = 0.\end{aligned}\right\}$$

These are the equations originally set up by Dirac. The stationary states are determined by the additional demand:

$$\Psi^{s,a}_\lambda = e^{-ih^{-1}Et}\,\varphi^{s,a}_\lambda \quad ; \qquad (\lambda = 1,2)$$

their differential equations are:

$$\left.\begin{aligned}(E + e\varphi - \mu c^2)\psi^s &= c(d_x\sigma_x + d_y\sigma_y + d_z\sigma_z)\varphi^a\\ (E + e\varphi + \mu c^2)\psi^a &= c(d_x\sigma_x + d_y\sigma_y + d_z\sigma_z)\varphi^s\end{aligned}\right\} \qquad (23.9)$$

If one is interested in the states of positive energy where E lies

close to μc^2, then the factor $E + e\varphi + \mu c^2$ is very large compared to $E + e\varphi - \mu c^2$; hence, the φ^a must be very small compared to the φ^s. We can thus identify the φ^s with the components of the ψ-function in the non-relativistic (Pauli) theory, whereas the φ^a represent, to a certain degree, the relativistic perturbation. Setting $E = \mu c^2 + E'$, the differential equation of the second order for φ^s becomes:

$$\left.\begin{aligned} -\frac{\hbar^2}{2\mu}\Delta\varphi^s + \Big\{ -E' - e\varphi - \frac{1}{2\mu c^2}(E' + e\varphi)^2 + \frac{e}{\mu c}(\vec{A}\vec{p}) \\ + \frac{e^2}{2\mu c^2}\vec{A}^2 + 2\varkappa(\vec{H}\cdot\vec{S})\Big\}\varphi^s - 2\varkappa i(\vec{E}\cdot\vec{S})\varphi^a = 0 \end{aligned}\right\} \qquad (23.10)$$

where

$$\varphi^a = (E + e\varphi + \mu c^2)^{-1} c(d_x\sigma_x + d_y\sigma_y + d_z\sigma_z)\varphi^s .$$

The setting up of a satisfactory relativistic wave equation for more than one electron has not yet met with success. This is due to the fact that, in addition to the 3f space coordinates of the f-electrons, the wave-equation must (as a result of its Lorentz invariance) also contain f-different times. In other words, the equation cannot be of the desired form:

$$\frac{\hbar}{i}\frac{\partial}{\partial t}\Psi + H\Psi = 0.$$

The solution of this difficulty would probably be facilitated through the development of a logical quantum-mechanics of wave-fields[1].

---

1) See W. Heisenberg and W. Pauli: Z.f.Physik vol.56 (1929) p.1
vol.59 (1930) p.168

24. The Electron in the Central Field according to Dirac.

The differential equation for the Dirac wave function in an electrostatic force field with the potential $\varphi(r)$ is, according to (23.9):

$$\left.\begin{aligned}(E + e\varphi - \mu c^2)\varphi^s &= c(\vec{p}\sigma)\varphi^a \\ (E + e\varphi + \mu c^2)\varphi^a &= c(\vec{p}\sigma)\varphi^s\end{aligned}\right\} \qquad (24.1)$$

where

$$(\vec{p}\sigma) = p_x\sigma_x + p_y\sigma_y + p_z\sigma_z \; ; \quad p_x = \frac{\hbar}{i}\frac{\partial}{\partial x} \quad , \text{ etc.}$$

The problem is to find a family of solutions ϕ (each having four components φ_1^s, φ_2^s, φ_1^a, φ_2^a) which transform according to the irreducible representation D_j of the rotation group and which, moreover, correspond to a definite reflection character w. We introduce into spin space four base vectors u_1^s, u_2^s, u_1^a, u_2^a and expand the function

$$\varphi = \varphi_1^s\, u_1^s + \varphi_2^s\, u_2^s + \varphi_1^a\, u_1^a + \varphi_2^a\, u_2^a$$

in terms of spherical functions $Y_l^{(n)}(\vartheta, \varphi)$. That yields:

$$\varphi = \Sigma f_{l n \lambda}(r) Y_l^{(n)}\, u_\lambda^s + \Sigma g_{l n \lambda}(r) Y_l^{(n)}\, u_\lambda^a = \Sigma P_l + \Sigma Q_l. \qquad (24.2)$$

The individual aggregates P_l and Q_l transform under rotations just like ϕ - in other words, according to D_j. They are, however, linear combinations of functions $Y_l^{(n)} u_\lambda^s$ or $Y_l^{(n)} u_\lambda^a$ which transform in accordance with $D_l \times D_{\frac{1}{2}} = D_{l+\frac{1}{2}} + D_{l-\frac{1}{2}}$. Hence, $j = l \pm \frac{1}{2}$, i.e. j is a half-integer and for l the two values $j \pm \frac{1}{2}$ alone count. (One of these values, of course, is even and the other odd.) We wish to denote these two values by l' and l'' so that $(-1)^{l'} = w$ and $(-1)^{l''} = -w$; that obviously can always be achieved.

In (24.2) the terms P_l belong to the reflection character $(-1)^l$ and the terms Q_l belong to the reflection-character $(-1)^{l+1}$. Thus, if φ is to belong to the reflection character $w = (-1)^{l'}$, then of the two possible terms P_l in (24.2) only the $P_{l'}$ and, likewise, of the two Q_l only the $Q_{l''}$ can actually occur. Hence, $\varphi = P_{l'} + Q_{l''}$ or

$$\left.\begin{aligned} \varphi^s = P_{l'} &= \Sigma f_{l'n\lambda}(r) Y_{l'}^{(n)} u_\lambda^s \\ \varphi^a = Q_{l''} &= \Sigma f_{l''n\lambda}(r) Y_{l''}^{(n)} u_\lambda^a \end{aligned}\right\}$$

The linear combinations of the $Y_l^{(n)} u_\lambda$ which transform according to $D_{l+\frac{1}{2}}$ or $D_{l-\frac{1}{2}}$, are, in accordance with #18, the following:

$$\left.\begin{aligned} W_{l,l+\frac{1}{2}}^{(m)} &= \sqrt{l+m+\tfrac{1}{2}}\, Y_l^{(m-\frac{1}{2})} u_1 + \sqrt{l-m+\tfrac{1}{2}}\, Y_l^{(m+\frac{1}{2})} u_2, \\ W_{l,l-\frac{1}{2}}^{(m)} &= -\sqrt{l-m+\tfrac{1}{2}}\, Y_l^{(m-\frac{1}{2})} u_1 + \sqrt{l+m+\tfrac{1}{2}}\, Y_l^{(m+\frac{1}{2})} u_2. \end{aligned}\right\} \quad (24.3)^*$$

It follows from here that

$$\left.\begin{aligned} \varphi^s = P_{l'}^{(m)} &= f(r) W_{l',j}^{(m)} \\ \varphi^a = Q_{l''}^{(m)} &= g(r) W_{l'',j}^{(m)} \end{aligned}\right\} \quad (24.4)$$

As a result, the problem is reduced to the determination of two functions f(r) and g(r). On inserting (24.4) into (24.1) we obtain for these functions the differential equations:

$$\left.\begin{aligned} (E + e\varphi - \mu c^2) f(r) W_{l',j}^{(m)} &= c(\vec{p}\sigma) g(r) W_{l'',j}^{(m)} \\ (E + e\varphi + \mu c^2) g(r) W_{l'',j}^{(m)} &= c(\vec{p}\sigma) f(r) W_{l',j}^{(m)} . \end{aligned}\right\} \quad (24.5)$$

According to the rule of differentiation for products, we have:

---

* Without having to depend on #18 one can also easily demonstrate the soundness of this statement by carrying out the operations M_p, M_q, M_s of #22 on both sides of (24.3).

$$(\vec{p}\cdot\sigma)f(r)W^{(m)}_{l,j} = f(r)(\vec{p}\sigma)W^{(m)}_{l,j} + f'(r)\frac{\hbar}{i}\varepsilon W^{(m)}_{l,j}$$

where

$$\varepsilon = \frac{x}{r}\sigma_x + \frac{y}{r}\sigma_y + \frac{z}{r}\sigma_z$$

Since the expressions $(\vec{p}\sigma)W^{(m)}_{l',j}$ and $\varepsilon W^{(m)}_{l',j}$ transform under rotations according to D_j, (although as space functions they belong to the opposite reflection character $-w$) they can be only numerical multiples of $hr^{-1}W^{(m)}_{l'',j}$ and $W^{(m)}_{l'',j}$, respectively. The same is valid on interchanging the l' with l''. The calculation (e.g. from the explicit expressions (24.3) and the formulae for the spherical functions) is not difficult; it yields (for an appropriate choice of the proportionality factors in the $W_{l,j}$):

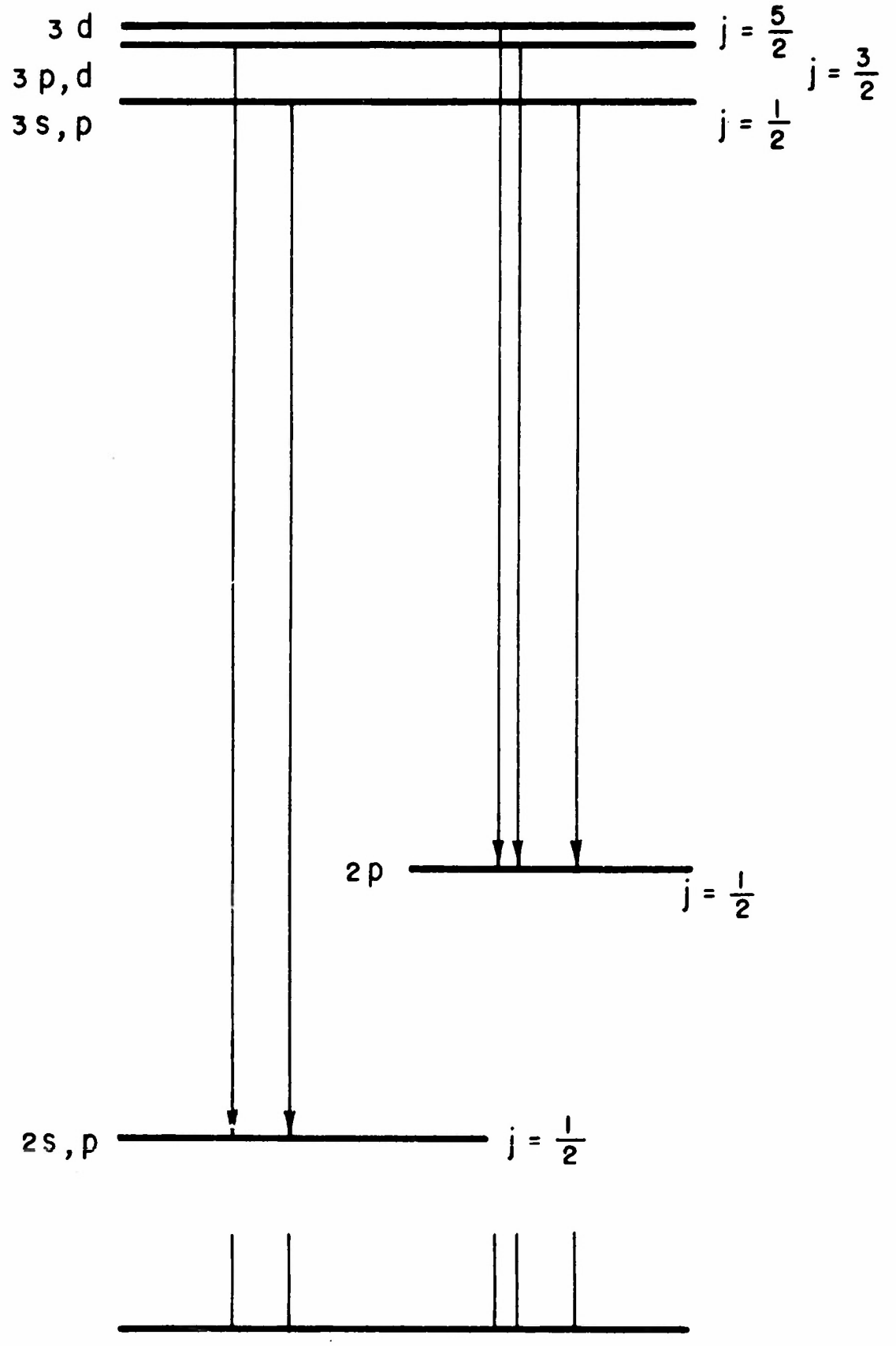
$$\begin{aligned}
\varepsilon W^{(m)}_{j\pm\frac{1}{2},j} &= W^{(m)}_{j\pm\frac{1}{2},j} \\
(\vec{p}\sigma)W_{j+\frac{1}{2},j} &= -(j+\tfrac{3}{2})\frac{\hbar i}{r}W_{j-\frac{1}{2},j} \\
(p\sigma)W_{j-\frac{1}{2},j} &= (j-\tfrac{1}{2})\frac{\hbar i}{r}W_{j+\frac{1}{2},j} \quad .
\end{aligned}$$

The two last formulae can be put into a more suitable form by substituting for the quantum numbers w,j a new integral number k, defined by:

$$\begin{aligned}
k &= j+\tfrac{1}{2} = l'+1 \qquad \text{for } l' = j-\tfrac{1}{2} \\
k &= -(j+\tfrac{1}{2}) = -l' \qquad \text{for } l' = j+\tfrac{1}{2}.
\end{aligned}$$

We then have:

$$\begin{aligned}
(\vec{p}\sigma)W_{l',j} &= (k-1)\frac{\hbar i}{r}W_{l'',j} = (1-k)\frac{\hbar}{ir}W_{l'',j} \\
(\vec{p}\sigma)W_{l'',j} &= -(k-1)\frac{\hbar i}{r}W_{l',j} = (1+k)\frac{\hbar}{ir}W_{l',j}
\end{aligned}$$



FINE-STRUCTURE OF THE Hα LINE

Figure 3

Upon inserting this into (24.5) there follows:

$$(E + e\varphi - \mu c^2)f = \frac{\hbar c}{i}\left(\frac{1-k}{r}g + g'\right).$$
$$(E + e\varphi + \mu c^2)g = \frac{\hbar c}{i}\left(\frac{1+k}{r}f + f'\right).$$

As for the determination of the pair of functions f,g and the eigenvalue E, I refer the reader to the text-literature. The calculation leads to the fine structure of the hydrogen and He^+ terms as well as to the doublet splitting of the light alkalis, in agreement with the experiments. Since $\varphi^s = P_{l'}$, l' is the usual azimuthal quantum number l ; its value is $k - 1$ for $k > 0$ and $-k$ for $k \leq 0$. In the case of a pure coulombic field (H, He^+) we find that for each principal quantum number n there are two terms with equal $j = |k| - \frac{1}{2}$ and different $l = j \pm \frac{1}{2}$ which coincide. (see Fig. 3).

The problem treated in this section can also be treated as a perturbation problem by starting out with the differential equation (23.10) and considering the terms containing $(E' + e\varphi)^2$ and $(\vec{E}\cdot\vec{S})$ as perturbation terms. In particular, this method can be used to advantage in the case of a non-Coulombic field. The term $(E' + e\varphi)^2$ leads merely to a term displacement independent of the spin-orientation; the other term:

$$-2\mathcal{H}i(\vec{E}.\vec{S})\varphi^a = -\mathcal{H}i(\vec{E}\cdot\sigma)(E' + e\varphi + 2\mu c^2)^{-1}c(\vec{p}\cdot\sigma)\varphi^s$$

is responsible for the doublet splitting. On ignoring $E' + e\varphi$ as against $2\mu c^2$, we obtain:

$$-\frac{\mathcal{H}i}{2\mu c}(\vec{E}\cdot\sigma)(\vec{p}\cdot\sigma)\varphi^s = -\frac{\mathcal{H}i}{2\mu c}\left\{(\vec{E}\cdot\vec{p}) + i([\vec{E}\vec{p}]\cdot\sigma)\right\}\varphi^s .$$

As for the splitting, only the second term in the brackets counts. Since

$$\vec{E} = -\frac{F}{r}\frac{\partial \varphi}{\partial r} \quad \text{and} \quad (\vec{r}\cdot\vec{p}) = \vec{L},$$

the splitting term is simply:

$$-\frac{\varkappa}{2\mu c r}\frac{\partial \varphi}{\partial r}(\vec{L}\cdot\sigma)\psi^s = -\frac{\varkappa}{\mu c r}\frac{\partial \varphi}{\partial r}(\vec{L}\cdot\vec{S})\psi^s.$$

Now,

$$(\vec{L}\cdot\vec{S}) = (\vec{L}+\vec{S})^2 - \vec{L}^2 - \vec{S}^2 = \vec{M}^2 - \vec{L}^2 - \vec{S}^2,$$

$$(\vec{L}\cdot\vec{S})\psi^s = \left\{j(j+1) - l(l+1) - \tfrac{1}{2}\left(\tfrac{1}{2}+1\right)\right\}\psi^s$$
$$= (k-1)\psi^s$$

where $k = l+1$ or $-l$ as above. From this the extent of the splitting can be calculated without any difficulty.

In the case of the many-electron problem, formula (24.6) also gives a useful addition term corresponding to the interaction between the orbital momentum of an electron and its own spin, provided the field in which the single electron moves does not deviate too much from a central field.

25. The Multiple Electron Problem. Multiplet Structure. Zeeman - effect.

We refer back to the non-relativistic theory.

The state of a system of f electrons is given by a function

$$\phi(q_1, q_2, \ldots, q_f, \sigma_1, \ldots, \sigma_f),$$

where q_h are the space coordinates and σ_h the spin coordinates (defined with respect to the z-axis) of the h-th electron. If as in #22 we introduce into the spin space of the first electron the base-vectors u_1, u_2 and do likewise for the second electron v_1, v_2

etc., then our function may be written in the form:

$$\varphi(q_1,\ldots,q_f,\ \sigma_1,\ldots,\sigma_f) = \sum_{\lambda,\ldots,\nu} \varphi_{\lambda\mu\ldots\nu}(q)u_\lambda v_\mu,\ldots w_\nu \ . \quad (25.1)$$

Thus, instead of taking as a basis the function $\varphi(q,\sigma)$, we choose rather a system of functions $\varphi_{\lambda\mu\ldots\nu}$ of q alone.

Rotations of the space transform the functions (25.1) in such a manner that each pair of base-vectors (like u_1, u_2) are transformed according to the representation $D_{\frac{1}{2}}$ of the rotational group; the $\varphi_{\lambda\mu\ldots\nu}$, however, are transformed like the usual space-functions. Consequently, the products $u_\lambda v_\mu \ldots . w_\nu$ transform according to the representation $D_{\frac{1}{2}} \times D_{\frac{1}{2}} \times \ldots . \times D_{\frac{1}{2}}$. The reflection s leaves the $u_\lambda, v_\mu \ldots .$, invariant.

If we have a system of eigenfunctions $\varphi^{(1)}(q),\ldots .\varphi^{(k)}(q)$ of the Schroedinger equation without spin and they belong to the eigenvalue E, then the $k\cdot 2^f$ products

$$\varphi^{(\alpha)} \quad u_\lambda v_\mu \ldots . \ w_\nu \qquad (25.2)$$

satisfy the Schroedinger equation - the spin perturbation terms being ignored. In order to find out how this $(k.2^f)$-fold term is split by the effect of the spin, we first investigate as to how it transforms under rotations. Let the $\varphi^{(\alpha)}$ transform according to D_L (S-,P-,D-terms, cf. #17). The products (25.2) then transform according to the representation:

$$D_L \times D_{\frac{1}{2}} \times D_{\frac{1}{2}} \times \ldots . \times D_{\frac{1}{2}} \ . \qquad (25.3)$$

By reducing this representation we obtain the irreducible subspaces which subsequently break apart due to the spin perturbation.

It is convenient to carry out the reduction of the represen-

tation (25.3) in such a manner that one first multiplies the factors $D_{\frac{1}{2}}$ by one another:

$$D_{\frac{1}{2}} \times D_{\frac{1}{2}} = D_0 + D_1$$
$$D_{\frac{1}{2}} \times D_{\frac{1}{2}} \times D_{\frac{1}{2}} = D_{\frac{1}{2}} + D_{\frac{1}{2}} + D_{1\frac{1}{2}}$$

and then multiplies each individual term D_S thus obtained by D_L as follows:

$$D_L \times D_S = \Sigma D_J \qquad (J = L+S, \ldots, |L-S|). \qquad (25.4)$$

In terms of the vector-scheme, we first combine the spins $\frac{1}{2}h$ of the individual electrons to form a resultant hS which is then added[1] vectorially to the total orbital momentum hL to form a resultant of length hJ; the components of hJ in the z-direction may take on the values hM (M = J,J-1,....,-J). L is known as the azimuthal quantum-number, S, the spin number, J the inner quantum number, and M the magnetic quantum number. S and J are integers for an even number of electrons; otherwise they are half-integers.

The different D_J terms which arise from each of the products (25.4) belong to the same multiplet[2)(p.181)]. The multiplet is normal if the terms with greatest J lie highest in the energy diagram; else the converse holds.

---

1) This approach is especially practical in the case where the multiplet-splitting is small as compared to the term-splitting (mentioned in #18.2) caused by the interactions among the electrons - that is, in the case of the Russell-Saunders coupling. If other coupling conditions prevail, e.g. the so-called (j,j)-coupling where the interaction between spin and the orbital motion of the individual electron outweighs all other interactions, then we first combine the orbital-momenta $h\ell$ of the individual electrons with their spins $\frac{1}{2}h$ according to the scheme:

$$D_\ell \times D_{\frac{1}{2}} = D_{\ell+\frac{1}{2}} + D_{\ell-\frac{1}{2}}$$

and then we mutliply the representations D_j of the individual electrons by one another. In the final analysis the J-values obtained are the same as those in the Russell-Saunders coupling - the only difference is that they are arranged differently.

If $L \geq S$, then the number of terms of a multiplet (the multiplicity) is, according to (25.4), equal to $2S + 1$. If, however, $L < S$, then the multiplicity "cannot fully develop itself"; only $2L + 1$ terms occur. For the case $L = 0$ (S-terms), in particular, only one term occurs (a singlet). In spite of that, $S = \frac{1}{2}$ always represents a doublet term and $S = 1$ always a triplet term, etc. Accordingly, we differentiate among the terms, as follows:

singlet terms: $^1S, ^1P, ^1D, \ldots$ $\quad (S = 0)$

doublet terms: $^2S, ^2P, ^2D, \ldots$ $\quad (S = \frac{1}{2})$

triplet terms: $^3S, ^3P, ^3D, \ldots$ $\quad (S = 1)$

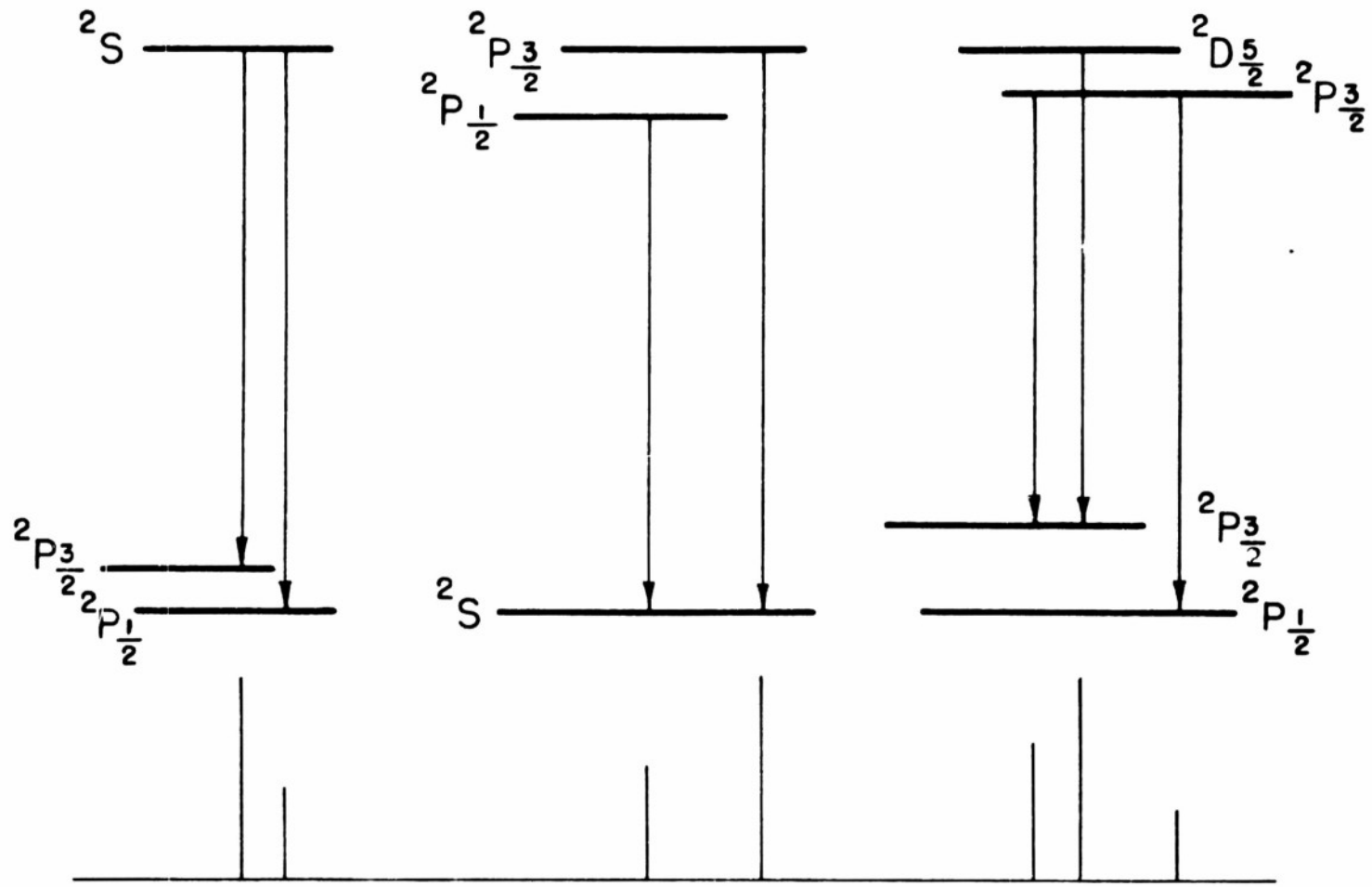
etc. The symbol ^2P is expressed as doublet P .

The reason for this terminology is based on the selection rule for S which we shall establish right away. The components of a multiplet are distinguished from one another by means of an index J attached to the letters S,P,D, etc. on their lower right, e.g. A ^3P-term consists of the components $^3P_0, ^3P_1, ^3P_2$.

The determination of the behavior of the eigenfunctions (25.2) under the reflection s at the zero point is quite simple inasmuch as the u_λ etc. are thereby left invariant. If the $\phi^{(k)}$ belong to the reflection-character

---

2) In the next chapter we shall see that of the various theoretically possible values of S which arise due to the multiplication $D_{\frac{1}{2}} \times D_{\frac{1}{2}} \times \ldots$, only one occurs actually in nature; this value of S, however, gives rise to a complete multiplet (25.4) in which all theoretically possible values of J occur. Hence, in the case of two electrons (e.g. Helium), we do not find a singlet term in the immediate neighborhood of each triplet, a fact contrary to theoretical expectations ($D_{\frac{1}{2}} \times D_{\frac{1}{2}} = D_0 + D_1$).



NORMAL DOUBLETS

Figure 4

$$w = (-1)^{l_1 + \cdots + l_f}$$

then the products (25.4), too, belong to the reflection-character and nothing is changed upon inclusion of the spin-perturbation.

The following exact selection rules are valid:

$$\left.\begin{aligned} J &\longrightarrow J-1,\ J,\ J+1 \qquad (\text{except for } 0 \longrightarrow 0) \\ M &\longrightarrow M-1,\ M,\ M+1 \\ w &\longrightarrow -w \end{aligned}\right\} \qquad (25.6)$$

The corollaries concerning the intensity and polarization of the emitted light that we have derived in #19, are also valid here. For proof of this one need only refer back to #19 and replace there all the L by J. As will be recalled, the proof was based exclusively upon the properties of the representation D_L.

The selection rule for J indicates which transitions are possible between the terms of any two multiplets. The intensities of the emitted spectral lines are approximately proportional to the products of the degrees of degeneracy $(2J + 1)(2J' + 1)$ of the initial and final level. In Fig. 4 we display for a few doublet terms the allowed inter-combinations as well as the positions and intensities of the lines.

The selection rule for w is precisely the Laport Rule (cf.#19). The selection rule for M prevails in the case of an axially symmetric perturbation which removes the $(2J + 1)$-fold rotational degeneracy (the Zeeman or Stark effect). The intensity relations of the split-up lines produced by a small perturbation of the above kind can be inferred from equations (19.9).

As long as the multiplet-splitting (spin-effect) is small -

i.e. especially in the case of the lighter elements, the following selection-rules are valid:

$$\left.\begin{aligned} &L \longrightarrow L-1,\ L,\ L+1 \quad (\text{except for } 0 \longrightarrow 0) \\ &S \longrightarrow S \end{aligned}\right\} \qquad (25.7)$$

That is so, for in multiplying the approximate eigenfunctions (25.2) by x, y, or z and expanding the products in terms of all eigenfunctions, then the products $u_\lambda v_\mu \ldots$ remain unchanged and the $x\varphi^{(\alpha)}$ etc. are in fact expanded in terms of the $\varphi^{(\beta)}$; thus, the same terms $\varphi^{(\beta)}$ occur which occurred even in the absence of the spin perturbation and must, therefore, satisfy the old selection rule for L. However, the spin-functions $u_\lambda v_\mu \ldots$ as well as their linear combinations belonging to the representation D_S remain unchanged in the expansion.

It happens sometimes that, due to the spin-perturbation, lines are emitted (especially in the case of the heavier elements) which are in contradiction with the restrictions (25.7). For example, we find among the heavier elements that there frequently occur combinations among triplet-and singlet-terms.

The rule $S \longrightarrow S$ signifies the breaking apart of the entire spectrum series of an element into various line systems, each of which belonging to one term system with one S-value. These term systems are known as singlet-, doublet-,.... systems - according to the scheme (25.5). As was already remarked, inter-combinations among the various systems are possible under certain conditions.

Example: In the case of the lighter atoms with two emitting electrons (He, Be, Mg), there occur singlet- and one triplet-system which do not combine with one another (see Fig. 6 on p. 192). The S-terms in both systems are singlets; nevertheless, we speak of

a ^{3}S-term (expressed as: triplet S) if the term belongs to the triplet-system. The extent of the multiplet-splitting can be calculated[1] on the basis of plausible assumptions regarding the interaction energy of spin with orbital motion.

The Anomalous Zeeman-Effect: According to #22, the linear perturbation term of the wave-equation due to the magnetic field strength is (for a homogeneous field H_3 along the z-direction):

$$\varkappa \vec{H}.(\vec{L} + 2\vec{S}) = \varkappa \vec{H}\cdot(\vec{M} + \vec{S}) = \varkappa H_3(M_3 + S_3).$$

If we assume the perturbation to be small in relation to the multiplet splitting (a weak magnetic field) then with the aid of the perturbation calculation we form for a family of eigenfunctions R_{2J+1} belonging to one line of the multiplet the expression $(M_3 + S_3)\varphi_J^{(M)}$ and expand it in terms of the $\varphi_{J'}^{(M')}$. We then seek out from the expansion the terms $\varphi_J^{(M')}$ which belong to the same linear family R_{2J+1}. Since $M_3\varphi_J^{(M)} = M\varphi_J^{(M)}$, we need only evaluate $S_3\varphi_J^{(M)}$. If we add to the above also $S_x\varphi_J^{(M)}$ and $S_y\varphi_J^{(M)}$, then we have to expand in terms of $\varphi_{J'}^{(M')}$, $3(2J + 1)$ functions which all transform according to $D_1 \times D_J$. According to #19, all coefficients in such an expansion (which relate to the space R_{2J+1}) are uniquely determined (to within a factor) by pure group-theoretical considerations. If, therefore, S_x', S_y', S_3' are the operators which are obtained from the S_x, S_y, S_3 by discarding from the series expansion all terms which do not relate to the spaoe R_{2J+1}, and if also M_x', M_y', M_3' are

---

1) See W. Heisenberg: Z.f.Phys. vol.39 (1926) p.499
as well as, S. Goudsmit: Phys.Rev. vol. 31 (1928) p.946.

formed in an analogous fashion, then the S'_x, S'_y, S'_z must, of necessity, agree with the M'_x, M'_y, M'_z - to within a factor β:

$$S'_x = \beta M'_x, \quad S'_y = \beta M'_y, \quad S'_z = \beta M'_z .$$

From here there follows:

$$(M'_z + S'_z)\varphi_J^{(M)} = (1+\beta)M'_z\,\varphi_J^{(M)} = (1+\beta)M\varphi_J^{(M)} ;$$

hence, the $\varphi_J^{(M)}$ are to a first approximation the eigenfunctions of the perturbation problem and $(1+\beta)M\varkappa H_z$ is the extent of the magnetic splitting.

In order to determine the splitting-factor $g = 1 + \beta$, we make use of the following artifice. We form the scalar-product:

$$(S'M') = (M'S') = \beta M' = \beta J(J + 1).$$

Then

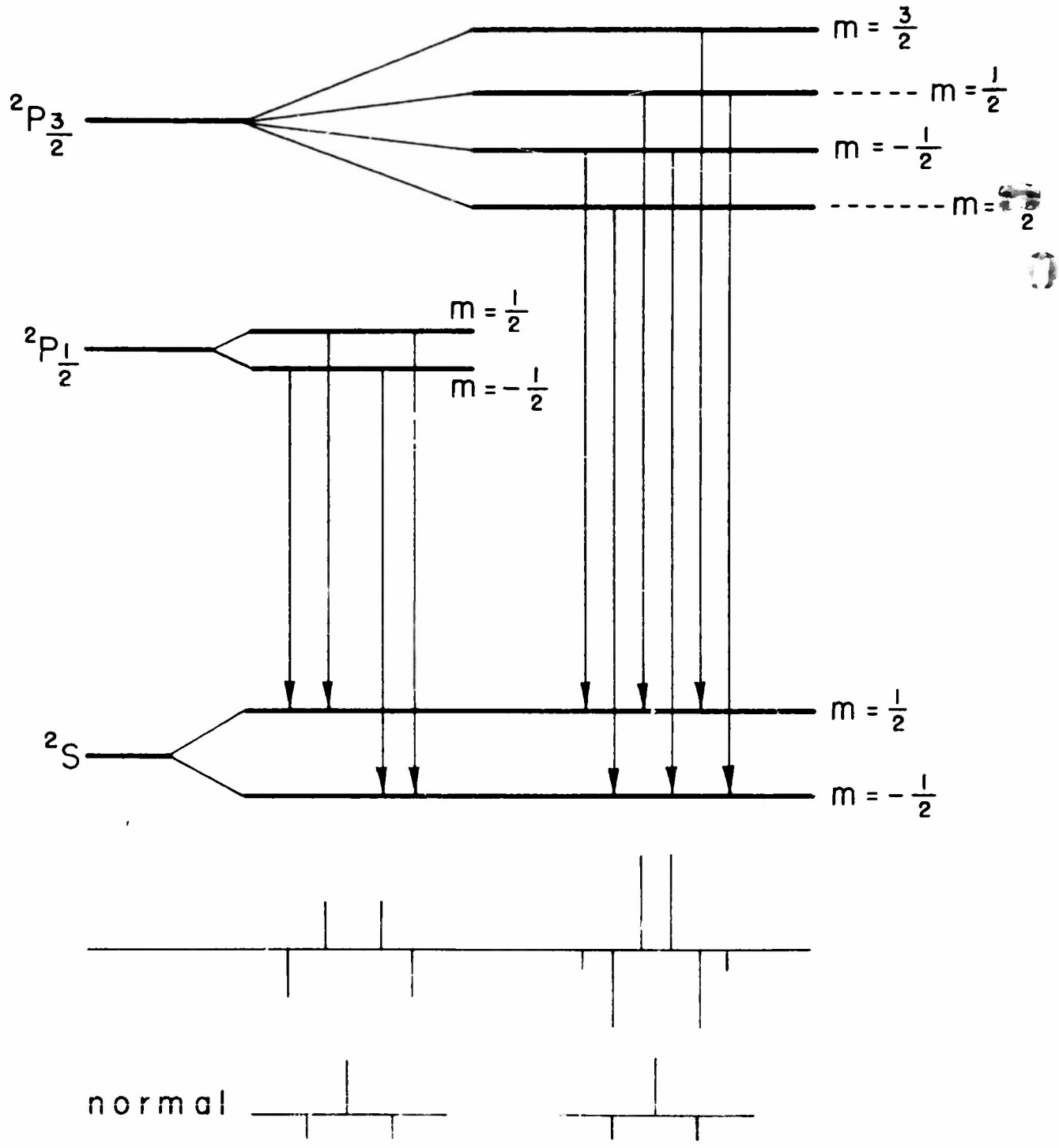
$$\vec{L}^2 = (\vec{M} - \vec{S})^2 = \vec{M}^2 - \vec{M}\vec{S} - \vec{S}\vec{M} + \vec{S}^2 .$$

If in the last equation (on the left-hand side and the right-hand side) one limits oneself to those operators which relate to the space R_{2J+1} and if one also observes that all the lines of a multiplet belong approximately to the eigenvalue $L(L + 1)$ of $\vec{L}^2$ and to the eigenvalue $S(S + 1)$ of $\vec{S}^2$, then we obtain (for a small multiplet splitting):

$$L(L + 1) = J(J + 1) - 2\beta J(J + 1) + S(S + 1).$$

From here we calculate β and $g = 1 + \beta = 1 + \dfrac{J(J+1) + S(S+1) - L(L+1)}{J(J + 1)}$

The above formula is in agreement with the experiments (cf. the Landé empirical equation (21.1) for $S = \frac{1}{2}$). It determines together with the selection rule $M \longrightarrow M + 1, M, M - 1$ and the intensity rules the typical Zeeman-splitting which recurs in each quantum-jump $L \longrightarrow L'$



ZEEMAN-TYPE SPECTRUM $^2P \leftarrow \rightarrow {}^2S$

Figure 5

$S \longrightarrow S'$, $J \longrightarrow J'$. Two examples of this are illustrated in Fig. 5.

The lines polarized parallel to the magnetic field are directed upwards in the diagram; the others are directed downwards. For purposes of comparison, the normal Zeeman-effect pattern was drawn in both cases on the same scale.

If the magnetic splitting is of the same magnitude as the multiplet-splitting (a stronger magnetic field) then both perturbations must be treated simultaneously.

Acting on the linear family of the $(2L + 1)(2S + 1)$ eigenfunctions $\phi_L^{(M)}\, w_S^{(M')}$ are the magnetic perturbation W

$$W\phi_L^{(M)}\, w_S^{(M')} = \mathcal{H}\, H_3 (L_3 + 2S_3)\phi_L^{(M)}\, w_S^{(M')}$$

$$= \mathcal{H}\, H_3 (m + 2m')\phi_L^{(M)}\, \varphi_S^{(M')}$$

and the spherically symmetric spin-perturbation V, the eigenfunctions of which are the linear combinations

$$\phi_J^{(M)} = \sum_{m} c^J_{mm'}\, \phi_L^{(m)}\, w_S^{(m')} \qquad (m + m' = M)$$

(cf. (183), and the eigenvalues of which can be determined in a rather empirical manner from the position of the multiplet terms:

$$V\phi_J^{(M)} = \varepsilon_J\, \phi_J^{(M)} .$$

As a result, the matrix for V (with respect to the base of the $\phi_J^{(M)}$) is known. In order to calculate the total perturbation V + W, one must first refer V to the old base $\phi_L^{(m)}\, w_S^{(m')}$. On denoting the matrix of the $c^J_{mm'} = c^{JM}_{mm'}$ (where J and M are the splitting indices and m and m' the numerical indices) by Q and the diagonal matrix of the ε_J by R, the matrix for V, (with respect to the old base) becomes:

$$Q\,R\,Q^{-1}.$$

Hence, if W is the diagonal matrix of the $\mathcal{H}\,H_z(m + 2m')$, the secular equation becomes:

$$|W + Q\,R\,Q^{-1} - \zeta\,E| = 0, \qquad (25.8)$$

or, on multiplying through by the determinant $|Q|$:

$$|W\,Q + Q\,R - \zeta Q| = 0 \qquad (25.9)$$

The solving of this equation is facilitated by the recognition that all the matrices in question break apart into component parts which correspond to the individual values of $M = m + m'$. To each M there belong certain allowable values of J in the form of column indices - and just as many pairs of values m, m' in the form of row indices. The sub-determinant of (25.9) (corresponding to one value of M) is:

$$|\mathcal{H}\,H_z(m + 2m')c^J_{mm'} + c^J_{mm'}\,(\varepsilon_J - \zeta)| = 0 \qquad (25.10)$$

The numbers $c^J_{mm'}$ can be calculated from (18.2). In the case of the doublet all equations (25.10) are either linear or quadratic; they can therefore be easily solved[1].

Let us likewise break up equation (25.8) into component equations each of which corresponds to one value of M. We then notice that the sum of the roots is equal to the trace of the matrix $W + Q\,R\,Q^{-1}$. Consequently, we have the <u>rule of sums</u>: <u>The sum of the splittings $(\zeta - \varepsilon_J)$ is for each value of M a</u>

---

1) Heisenberg, W. and P. Jordan: Anwendung der Quanten-mechanik auf das Problem der anomalen Zeeman-effecte - Z.f. Physik, vol. 37 (1926) p.263.
Darwin, K.: Proc.Roy.Soc.(A) vol 118 (1928) p.264.

linear function of the fieldstrength H_z, namely:

$$\mathcal{H}\, H_z \sum_{m+m'=M} (m + 2m') \ .$$

The coefficient of $\mathcal{H}H_z$ must, of course, agree with the value $M\Sigma g(J)$ obtained before for weak fields.

For the case of a very strong field, where the magnetic splitting is large compared to the multiplet splitting ε_z we may to a first approximation completely ignore the ε_J; so we take for the eigenfunctions $\phi_L^{(m)}\, \phi_S^{(m')}$ and for the eigenvalues $m + 2m'$. Here the selection rules

$$m \longrightarrow m + 1,\ m,\ m - 1$$

$$m' \longrightarrow m'$$

are valid; hence, only a normal Zeeman-effect is obtained. For very strong fields, therefore, the anomalous Zeeman-effect becomes the normal one (Paschen-Backet Effect). The terms are, of course, further split up by the spin perturbation.

V. THE PERMUTATION GROUP AND THE PAULI EXCLUSION PRINCIPLE.

26. Resonance of Equivalent Entities.[1)]

A stationary state of a system of two electrons (without spin) is described by an eigenfunction of the form

$$\varphi(q_1, q_2) = \phi_1(q_1)\phi_2(q_2) \qquad (26.1)$$

where ϕ_1 and ϕ_2 are eigenfunctions of the individual electrons and where, also, the interaction energy between the electrons is ignored. If E_1 and E_2 are the energy values of the individual electrons, then $E = E_1 + E_2$ is the energy value belonging to (26.1). Now, to this energy value there also belongs another eigenfunction - that which is obtained from (26.1) by an interchange (12) of the electrons:

$$\phi' = (1\ 2)\ \phi = \phi_1(q_2)\phi_2(q_1). \qquad (26.2)$$

We now consider the interaction as a perturbation. Since the interaction energy commutes with the permutation (1 2) both must be capable of being simultaneously transformed on to the principal axes. The principal axes transformation of the permutation (1 2) leads to the following linear combinations:

$$\phi_s = \phi + (1\ 2)\phi$$
$$\phi_a = \phi - (1\ 2)\phi.$$

These symmetrical and anti-symmetrical eigenfunctions belong to the two different first-order representations of the permutation group $\mathfrak{S}_2$. As a result of the electron interaction the terms corresponding to ϕ_s and ϕ_a separate; nevertheless, they still remain

---

1) Heisenberg, W. Z.f.Physik vol.38 (1926) p.411.

symmetrical or anti-symmetrical for each possible perturbation acts on the two electrons according to the same law and therefore converts symmetrical functions ϕ_s again into symmetrical ones and likewise, anti-symmetrical functions ϕ_a again into anti-symmetrical ones.

Let W be the interaction energy (i.e. $H = H_0 + W$ is the entire energy) and let the series expansion for $W\phi$ be as follows:

$$W\phi = w\phi + w'\phi' + \ldots . ; \quad w = (\phi, W\phi); \quad w' = (\phi', W\phi)$$

Consequently, upon carrying out the operation (1 2), we obtain:

$$W\phi' = w\phi' + w'\phi + \ldots .$$

If there exist no further eigenfunctions belonging to the same eigenvalue, then (according to the perturbation calculation) the term-splitting is found by means of the principal-axes transformation of the matrix:

$$\begin{pmatrix} w & w' \\ w' & w \end{pmatrix}.$$

As was noted already, the term splitting is brought about by the introduction of the linear combinations

$$\phi + \phi' = \phi_s \quad \text{and} \quad \phi - \phi' = \phi_a.$$

We have:

$$W(\phi + \phi') = (w + w')(\phi + \phi') + \ldots ,$$
$$W(\phi - \phi') = (w - w')(\phi - \phi') + \ldots .$$

Thus, the term-values are (to a first approximation):

for ϕ_s: $E_1 + E_2 + w + w'$

for ϕ_a: $E_1 + E_2 + w - w'$.

We see, therefore, the terms lie on both sides of the average value

$$E_1 + E_2 + w = (\phi, H_o\phi) + (\phi, W\phi)$$

$$= (\phi, H\phi),$$

which is equal to the average value of the energy in the state ϕ. The extent of the splitting is equal to double the <u>exchange integral</u>

$$w' = (\phi', W\phi).$$

Since $W = \frac{e^2}{r_{12}}$ is the electrostatic interaction energy, we have:

$$w' = e^2\int \frac{\bar{\phi}'\phi}{r_{12}}\, dq = e^2 \int \frac{\bar{\varphi}_1(q_2)\bar{\varphi}_2(q_1)\varphi_1(q_1)\varphi_2(q_2)}{r_{12}}\, dq$$

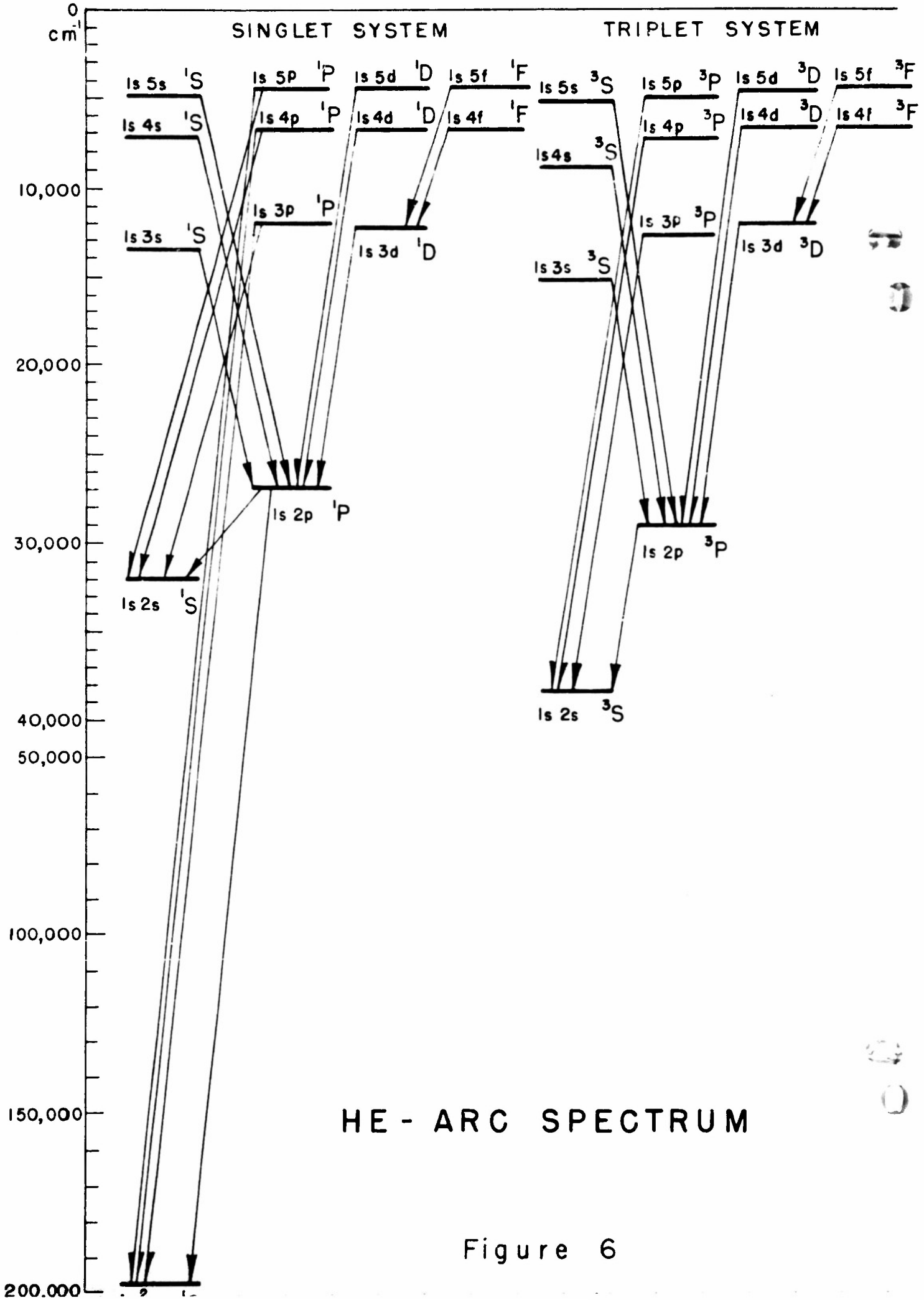
The factor $\frac{1}{r_{12}}$ is largest when q_2 is almost equal to q_1 ; but then the numerator is almost equal to the positive expression $\varphi_1(q_1)\bar{\varphi}_1(q_1)\varphi_2(q_1)\bar{\varphi}_2(q_1)$. Hence, it may be assumed that in the atom the exchange integral is generally positive. It follows then that as a rule the symmetrical term (ϕ_s) has the higher energy and the antisymmetrical (ϕ_a) the lower energy.

We can differentiate between the symmetrical and anti-symmetrical states by means of a "symmetry-character" $\chi = \pm 1$ which is defined by

$$(1\ 2)\phi = \chi\, \omega.$$

If the two electrons are in the same states: $\varphi_1 = \varphi_2$, then only the symmetrical state ϕ_s of the system is allowed since $\phi_a = 0$.

The selection rule for the symmetry character χ can be easily obtained. A symmetrical or anti-symmetrical eigenfunction when multiplied by Σx, Σy, or Σz yields always similar functions, the



HE - ARC SPECTRUM

Figure 6

series expansion of which also contains only symmetrical or anti-symmetrical terms respectively. Thus, the selection rule is:

$$\chi \longrightarrow \chi$$

The calculation for Helium (without spin) lead to magnitudes of the symmetrical and anti-symmetrical terms which were in agreement with the experiments[1]. According to the theory of #25, each term should further split up into one singlet and one triplet. In reality, the symmetrical terms display only singlets, the anti-symmetrical terms only triplets (see Fig. 6). The reason for this will be taken up later. The symmetrical terms combine only with one another; the anti-symmetrical ones act likewise. As for the groundstate of He, both electrons are in the lowest s-level; thus $\phi_1 = \phi_2$ and $\chi = 1$.

We have a similar state of affairs in the case of the hydrogen molecule H_2 (as well as in He_2, N_2, O_2 etc.). The eigenfunction ϕ involves the co-ordinates of two nuclei and two electrons. Upon interchanging the nuclei the symmetrical eigenfunctions assume the factor $\chi = 1$ and the anti-symmetrical ones the factor -1. Transitions are here possible only as a result of the nuclear moments, the magnetic effects of which, however, are extremely small. In a way, therefore, it is possible to distinguish between two types of hydrogen molecules: the symmetrical H_2-molecules and the anti-symmetrical ones. The two types also have different spectra (ortho- and para-hydrogen).

Now, the exchange-phenomenon for two or more electrons relates to the rotational degeneracy. The main principle involved is that the permutations not only commute with the energy but also with all rotations. Consequently (according to #13), it is possible to

simultaneously reduce out for each energy level the rotation group and the permutation group. In the case of two electrons, that simply means that the symmetrical and anti-symmetrical eigenfunctions each form for every energy level a rotationally invariant family of eigenfunctions; each of these can then be reduced out separately in terms of the rotation group.

A ϕ-function of a single electron without spin will henceforth be represented by the symbol $\phi(n\ell m|q)$, where n is the principal quantun number, ℓ the inner-, and m or m_ℓ the magnetic-quantum number. Let us now consider two electrons and tentatively disregard their interaction! We form from the two rotationally invariant linear families $\phi(n\ell|q_1)$ and $\phi(n'\ell'|q_2)$ of eigenfunctions of the individual electrons the $(2\ell+1)(2\ell'+1)$ products

$$\phi = \phi(n\ell m|q_1)\phi(n'\ell'm'|q_2); \qquad (26.3)$$

we form also the permuted products $(1\ 2)\phi$. As a result. we have for $n \neq n'$ or $\ell \neq \ell'$ a total of $2(2\ell+1)(2\ell'+1)$ (approximate) eigenfunctions belonging to the same energy value $E = E_1 + E_2$. The sums $\phi + (1\ 2)\phi$ and the differences $\phi - (1\ 2)\phi$ determine two linear subsets each consisting of $(2\ell+1)(2\ell'+1)$ linearly independent, symmetrical or anti-symmetrical functions. When subjected to rotations, the two sets transform just like the sets (ϕ) and $(1\ 2)(\phi)$ (from which they originated) according to the representation

$$\begin{aligned} D_\ell \times D_{\ell'} &= D_{\ell+\ell'} + D_{\ell+\ell'-1} + \cdots + D_{|\ell-\ell'|} \\ &= \Sigma\, D_L\ . \end{aligned} \qquad (26.4)$$

As a result of the interaction, the symmetrical and anti-symmetrical terms, as well as the terms corresponding to the various L-values

separate out; we then obtain for each of the L-values indicated in (26.3) one symmetrical term and one anti-symmetrical term. As a rule, the symmetrical terms lie higher in the energy diagram than the anti-symmetrical ones.

The matter becomes somewhat more complicated when both electrons are "in the same orbit", i.e. when $n = n', \ell = \ell'$. In that case, the products $(1\ 2)\phi$ can be obtained not only by permuting q_1 with q_2 in (26.3), but also by permuting m with m'. In other words, the $(1\ 2)\phi$ are already contained in the set (26.3) and there are only $(2\ell + 1)^2$ linearly independent eigenfunctions in the set. The symmetrical eigenfunctions are (when in (26.3) we omit the indices n, ℓ):

$$\phi + (1\ 2)\phi = \phi(m|q_1)\phi(m'|q_2) + \phi(m'|q_1)\phi(m|q_2),$$

and the anti-symmetrical ones:

$$\phi - (1\ 2)\phi = \phi(m|q_1)\phi(m'|q_2) - \phi(m'|q_1)\phi(m|q_1).$$

We assume that in the symmetrical case $m \geq m'$ and in the anti-symmetrical case $m > m'$. The eigenvalue of the operator L_z is for both functions $M = m + m'$. We consider the case $\ell = 2$. The allowable values of M are in the symmetric case:

$$\begin{aligned}
(m = \ \ 2)\ M &= 4,3,2,1,0 \\
(m = \ \ 1)\ M &= \quad 2,1,0,-1 \\
(m = \ \ 0)\ M &= \quad\quad 0,-1,-2 \\
(m = -1)\ M &= \quad\quad\quad -2,-3 \\
(m = -2)\ M &= \quad\quad\quad\quad -4;
\end{aligned}$$

In the anti-symmetric case:

(m = 2) M = 3,2,1,0.

(m = 1) M = 1,0,-1

(m = 0) M = -1,-2

(m = -1) M = -3.

It can be seen from here that the two largest values M = 4,3 (or M = 3,2 in the anti-symmetric case) occur each but once; the two next-smaller values M = 2,1 (or 1,0) occur each twice; the next one M = 0, three times. We need not concern ourselves with the negative values of M. If now we collect the values of M into the series L, L-1,....,-L beginning always with the largest M (as prescribed in #17) then we obtain for the symmetrical eigenfunctions the representations:

$$D_{2l} + D_{2l-2} + ... + D_o \quad \text{(in our case: } D_4 + D_2 + D_0 \text{)};$$

and for the anti-symmetrical eigenfunctions the representations

$$D_{2l-1} + D_{2l-3} + ... + D_1 \quad \text{(in our case: } D_3 + D_1\text{)}.$$

The matter becomes even more complicated in the many-electron problem. In addition to the symmetrical and the anti-symmetrical representations, we have still otherpossible representations of the permutation-group (cf. examples in #14). The only representations of the first order are the symmetrical and anti-symmetrical ones; all others are of a higher order. There is no sense, however, in continuing with the discussion prior to acquainting ourselves with a law which limits the number of possibilities quite severely; that is the Pauli-Exclusion Principle.

27. The Pauli - Exclusion Principle and the Periodic System of the Elements.[1)]

We recall the following facts. The allowable states for an emitting electron in the field of a screened-off nucleus are the following (in the order of increasing energy):

1s, 2s, 2p, 3s, 3p,..... (see Fig. 2 and 6).

In the Coulombic field (H, He^+), the principal quantum number alone is responsible for the position of the terms; the more the field deviates from a coulombic field, however, the further down the p-terms and especially the s-terms are displaced. Whenever the atomic core has the total charge of a hydrogen nucleus, the deviation from the coulombic field generally increases with increasing atomic number.

As to whether the term following 3p is the 3d - or the 4s - term depends completely upon the screening conditions. In the case of multi-valent ions, 3d is the lower term; in the case of atomic cores of charge 1, however, 4s is generally the lower term. The positions of the lowest terms 1s, 2s, 2p, 3s are so far removed from one another that in the case of unknown low terms one can generally predict the number of the term by merely considering the magnitude of the ionization energy alone.

In the case of unexcited atoms one would expect all electrons to be situated in the lowest level, that is in state 1s. That,

---

1) cf. F. Hund: Linienspektren und periodisches System der Elemente. Berlin, 1927.

however, is not the case. Instead, a completely different behavior is displayed which is intimately connected with the existence of the Periodic System.

The periodic system of the elements begins as follows:

0	I	II	III	IV	V	VI	VII	VIII
	1. H							
2.He	3.Li	4. Be	5. B	6. C	7. N	8. O	9. F	
10.Ne	11.Na	12.Mg	13.Al	14.Si	15. P	16. S	17.Cl	
18.Ar	19. K	20.Ca	(21.Sc	(22.Ti	(23. V	(24.Cr	(25.Mn	(26.Fe (27.Co (28.Ni

The ground state of H as well as of He^+ (ion) is the lowest s-orbital with the principal quantum number n = 1; in other words, the 1s orbital. The ground state of the He-atom consists in that the two electrons are both in the 1s-orbital. The transition to the next element Li in the periodic chart consists in the increase of the atomic number by 1 and the taking on of a new electron which acts as a valence electron.

As Fig. 2 (p. 15) indicates, this valence electron, however, is not in the 1s-orbital (when in its ground state) but rather in the 2s-orbital (else, the term 1s would have to lie lower than the ground-term of the H - or even of the He - atom!). In like manner, the ground state in which both of the valence electrons of Be are situated is the 2s orbital, as can be deduced from the ionization energies of these electrons. The emitting electrons of the elements

B, C, N, O, F are located neither in the lowest orbital 1s nor in the next lowest 2s, but rather in the 2-p orbital. In Neon, the number of electrons has grown to 2 + 2 + 6 and the loosest electrons are still in the 2p-orbital (as is deducible from the slowly increasing ionization potential).

The next element Na again has a hydrogen-like spectrum (like Li). The lowest term of the emitting electron is an s-term which has a higher energy than the ground-term 2s of Li; it must therefore be (at least) a 3s term. The next element Mg has two 3s electrons. We have in the second row of the periodic system an exact repetition of the first row. For K in the third row the ionization potential again decreases; the ground term of K is an s-term and lies above the ground term of Na; we therefore assign to the emitting electron the principal quantum number 4.

This shell-like structure of the atom is confirmed by x-ray spectroscopy. All experimental data point to the following Rule of Stoner. Each s-orbital (with a fixed principal quantum number) may be occupied by two electrons, each p-orbital by six electrons; as a rule, each orbital having the fixed quantum numbers n and ℓ may be occupied by the maximum number of $2(2\ell + 1)$ electrons[1].

Accordingly, in the case of He which has two electrons, the 1s orbital is completely occupied; in the case of Be having 2 + 2 electrons the 1s and 2s orbitals are fully occupied; in Neon having 2 + 2 + 6 electrons the 1s-, 2s- and 2-p orbitals are filled; in Mg also the 3s and in Ar the 3s- and 3p-orbitals are fully occupied, exactly in consonance with the periodic system. K and Ca are

---

1) Stoner, E.C.: Phil.Mag.vol. 48 (1924) p.71.

analogous to Na and Mg; here, the 3s-orbital becomes filled. Starting with Sc the behavior is different from that of the first two periods because now the 3d-electrons, too, begin to compete. First, ten 3d-electrons add on, then six 4p-electrons, until the 3d- and 4p-orbitals are filled with 16 electrons. The first "long period" of the system terminates with the noble gas Xe which has 36 electrons. Then the second long period begins; it proceeds in an entirely analogous fashion. With the advent of the f-electrons (the group of the Rare Earths), the present order breaks up completely as should be the case in view of existing chemical experimental data.

In order to explain the Stoner Rule, Pauli[1] set up the Exclusion Principle for equivalent orbitals: No two electrons of an atom may occupy the same quantum state, i.e. have the same quantum numbers (n, ℓ, j, m). Consequently, for fixed n, ℓ the values are $j = \ell \pm \frac{1}{2}$ and $m = j, j-1, \ldots, -j$. Hence, a total of $(2\ell + 2) + 2\ell = 2(2\ell + 1)$ combinations of these values are possible, as the Stoner Rule requires.

In the form given here the Pauli Exclusion Principle is not yet invariant under rotations. Let ustake, for example, the case of two s-electrons having the eigenfunctions

$$\varphi^{(1)} = \varphi(q)u_1 \qquad \text{and} \qquad \varphi^{(2)} = \varphi(q)u_2$$

(the spin being either parallel or anti-parallel to the z-axis). According to the Pauli Exclusion Principle, the eigenfunctions[2] $\varphi^{(a)}(1)\omega^{(a)}(2)$ and $\varphi^{(a)}(1)\varphi^{(a)}(2)$ for the pair of electrons are not allowed (we neglect here the interactions among the electrons). On

---

1) Pauli, W.: Z.f. Physik vol. 31, (1925) p.765.
2) 1 represents the argument q_1, σ_{1z}, likewise 2, the q_2, σ_{2z}, etc.

the other hand though, $\phi^{(1)}(1)\phi^{(2)}(2)$ is allowed. Now, the operation Lq (cf. #17), which is a combination of infinitesimal rotations ($L_q = iL_x + L_y$), transforms the above mentioned permissible product into $\phi^{(1)}(1)\phi^{(1)}(2)$ which is forbidden.

We do, however, obtain an exclusion principle which is invariant under rotations by demanding that the eigenfunctions of a system of electrons (as functions of the space and spin coordinates) must be anti-symmetric, i.e. they must change sign for every permutation of any two electrons.

In our case,

$$\phi^{(1)}(1)\phi^{(2)}(2) - \phi^{(2)}(1)\phi^{(1)}(2)$$

is the only allowable eigenfunction. In the general case, where an anti-symmetric eigenfunction is to be constructed from f-electron eigenfunctions $\phi_1, \ldots\ldots, \phi_f$, we set up the alternating sum:

$$\psi = \Sigma \delta_p P\phi_1(q_1, \sigma_1)\phi_2(q_2, \sigma_2)\ldots\ldots\phi_f(q_f, \sigma_f) \quad (27.1)$$

where P runs through all permutations and $\delta_p = \pm 1$, according as to whether P is an even or an odd permutation. One is easily convinced that the sum (27.1) is the only antisymmetric linear combination of its terms. The expression (27.1) goes to zero as soon as two ϕ become equal to one another (or, in general, if there is a linear dependence among the ϕ). Hence the exclusion of equivalent quantum orbitals is a consequence of the antisymmetry.

It follows easily from the time-dependent Schroedinger equation

$$\frac{h}{i}\frac{\partial \psi}{\partial t} + H\psi = 0$$

that a function ψ which is initially anti-symmetric remains as such forever, provided all electrons in the energy operator H are equi-

valent. Thus, the Pauli Exclusion Principle need only be satisfied for a specific time and no physical perturbation could possible ever annul it.

It is worthy of note that (upon neglecting the interactions among the electrons and the spin-perturbation) one may choose for the quantum numbers of a single electron (instead of n, ℓ, j, m) the following: $(n, \ell, m_\ell, m_s)(m_s = \pm\frac{1}{2};\ m_\ell = \ell, \ell-1, \ldots, -\ell)$. The simplest eigenfunctions of a single electron are namely the products

$$\phi(n \ell m_\ell) u_2 \qquad \begin{matrix} (m_s = \frac{1}{2} \quad \text{for } u_1); \\ (m_s = -\frac{1}{2} \quad \text{for } u_2) \ . \end{matrix} \tag{27.2}$$

The antisymmetric expression (27.1) can of course be formed from any arbitrary system of f linearly independent ϕ-functions of a single electron.

We now wish to consider the case in which the maximum number $f = 2(2\ell + 1)$ of electrons having the same quantum numbers n, ℓ unite to form an eigenfunction like (27.1). When a rotation D is carried out simultaneously on the co-ordinates $q_1, \ldots, q_f$ of the orbitals, then the ϕ_r become linearly transformed and the expression (27.1) is left unchanged (to within a factor). Thus, the entire rotationally invariant family to which the function ϕ belongs, consists of only one term. In other words, we are dealing here with an S-term ($L = 0$). The same result is obtained when one carries out the rotation D on the spin-coordinates; hence, the spin number is $S = 0$ and also $J = 0$. Consequently, <u>a shell filled with $2(2\ell + 1)$ electrons always has spherical symmetry and does not include the spin</u>. As was seen, the atoms He, Be, Ne, Mg, Ar, Ca in their ground state, and likewise, the atomic-cores of the alkalis Li, Na, K consist of such filled shells. Accordingly, the ground states of the first-mentioned

atoms and of the latter-mentioned ions consist always of the S-terms.

The completed shell is "closed off" (i.e. invulnerable to decomposition), especially when a considerably large number of electrons are held tight and close to the nucleus. We find this situation in the elements

He - (two 1s-electrons)

Ne - (two 1s, two 2s, six 2p-electrons)

Ar - (the same plus two 3s and six 3p-electrons).

These are, of course, the noble gases which are known not to enter into any chemically observable bond formations[1]. In the case of the alkali-earth element Be (two 1s and two 2s electrons) this closure effect does not yet commence. That is due to the fact that for the relatively small atomic number 4, the 2s electrons are not held tight enough by the nucleus as are, for example, the 1s electrons in He. In the noble gas Ne, however, where an additional six 2p electrons with the same principal quantum number 2 are involved, all the electrons are held tightly by the nucleus; there is no more room for any additional electrons.

In connection with the next noble gas Ar, one would expect to find enough room for ten 3d-electrons besides the two 3s and the six 3p-electrons. That is not the case, however, since for a d-electron the screening-off of the nucleus by the remaining electrons is rather complete; the d-electrons, therefore, experience a relatively small attraction by the nucleus[2](p.203). Accordingly, in the case of the elements

---

1) Spectroscopic methods, however, indicate the existance of the molecule He_2. It is not very stable, though, and exists only among excited He-atoms. It is certainly not formed by two He-atoms in their ground state.

K, Ca,.... follow Ar, the first shell to be filled is not the 3d-shell but rather the 4s-shell (for K and Ca). After the complete filling of the 4s, 3d, and 4p-shells, there again appears a noble gas, namely Xe.

In the Periodic Table every noble gas is succeeded by an alkali metal followed, in turn, by an alkali-earth metal. Here the one or two additional electrons are held somewhat loosely and, therefore, are more easily removable. That explains the easy formation of the monovalent ions Li^+, Na^+, K^+, Rb^+, Cs^+ and the divalent ions Be^{++}, Mg^{++}, Ca^{++}, Sr^{++}, Ba^{++}. The closed core of the Li, Na, K etc. atom behaves in a spherically symmetric manner ($L = 0$, $S = 0$, $J = 0$) with respect to its environment. It therefore causes no increase in the term-multiplicity of the spectrum emitted by the valence electron. For this reason, all the metals have a "hydrogen-like spectrum". The same is true of the "spark-spectra" of the ions Be^+, Mg^+, Ca^+, etc.

As can be seen, the Pauli-Exclusion Principle provides an explanation for the structure of the Periodic System and the typical chemical and spectroscopic properties of e.g. the noble gases, the alkalis and alkali-earth metals. These properties could otherwise not be explained. In the following section we shall discuss more fully how the term spectra of the various elements are formed according to the Pauli Exclusion Principle.

---

2) The higher the l, the less the "orbital penetrates into the atom", or put into exact quantum-mechanical terms: the farther away from the nucleus lie the largest values of the ϕ-function.

28. The Eigenfunctions of the Atom according to the Pauli Exclusion Principle.

The eigenfunctions of a many-electron system, need not be antisymmetric with respect to both the space coordinates and the spin coordinates. In the case of two electrons, for example, the ψ-function can be either symmetric or antisymmetric with respect to the space coordinates; but then if the former case obtains it must be antisymmetric with respect to the spin coordinates and if the latter is the case, then the ψ function must, of necessity, be symmetric with respect to the spin coordinates. That guarantees the change of sign of the ψ-function upon permutation of the space - and spin coordinates of the two electrons.

If we again start out from the base quantities u_1, u_2 and $v_1 v_2$ for the spin-functions of the two individual electrons, then the only antisymmetric spin function in question is $u_1 v_1 - u_2 v_2$. Here we have $S = 0$, since rotations will transform the spin function into itself only. The symmetric spinfunctions in question are

$$u_1 v_1, \quad u_1 v_2 + u_2 v_1, \quad u_2 v_2 .$$

These determine a linear family invariant under rotations. The eigenvalues of the operator S_3 for our three functions are

$$m = 1, \quad 0, \quad -1;$$

hence, the family of functions above necessarily transform according to D_1, i.e. $S = 1$. We note that as a consequence of the Pauli Exclusion Principle, there corresponds to each antisymmetric spin-function a symmetric space-function. The rule below then follows.

In the case of two electrons, there corresponds to each symmetric space-function a spin-function with S = 0 and hence a singlet term; to each antisymmetric space-function, however, there corresponds a spin-function with S = 1 and, hence, a triplet term.

That explains the fact (already mentioned in #26) that in Helium the symmetric eigenfunctions have only singlet terms and the antisymmetric eigenfunctions only triplet terms. The same is also true of all atoms which have beyond a close core two emitting electrons, like Be, Mg, Ca etc. The ground state of these atoms is always a ^1S-term, for when both electrons are in the lowest permissible s-orbital the eigenfunction is necessarily symmetric.

In the case of more than two electrons the same question arises: How do the eigenfunctions transform under permutations when they are functions of the space coordinates alone or when they are functions of the spin coordinates alone ? Only here the problem is considerably more complicated since we are now dealing with non-linear representations of the permutation group. Let us ignore the spin perturbation and consider the eigenfunctions as products

$$\phi_b\,(q_1,\ldots,q_f)u_\lambda\, v_\mu \ldots\, w_\varphi \qquad (28.1)$$

or linear combinations of such products. Here the ϕ_b are eigenfunctions which incorporate the electrostatic interaction but not the spin. Now, the expressions (28.1) undergo on each energy level four groups of linear transformations which all commute with one another. They are

the rotations of the q-space,

the rotations of spin space,

the permutations of the $q_1,\ldots,q_f$,

the permutations of the u,v,....,w.

The first method of investigating the behavior of the functions (28.1) under these groups and of meeting the requirements of the Pauli principle consists in that one first reduces out the two commutative transformation groups for the space functions only. We do that by arranging the eigenfunctions in the form of rectangles whose rows transform according to any irreducible representation D_L of the rotation group and whose columns transform according to an irreducible representation Δ of the permutation group (see #13). The columns are indicated by the numbers m_L (= L,L-1,....,-L). In general, to each energy level there corresponds a rectangle of the above type only. We would indeed have a case of a "fortuitous degeneracy" were this not so.

The same type of reduction is then carried out for the spin functions only. Here again we obtain rectangles whose column-indices are m_S = S, S-1,....,-S and whose representations are D_S and Δ'. Finally, from the space- and spin-functions thus obtained, we seek out those which satisfy the Pauli principle - i.e. those which are antisymmetric with respect to permutations. Obviously, the products of the space- and spin-functions (which under the permutations are individually transformed according to Δ and Δ') belong to the product representation $\Delta \times \Delta'$. The question only is whether as a result of the reduction this product representation $\Delta \times \Delta'$ has as a constituent the antisymmetric representation $\vec{A}$. Well, according to a theorem in #12 this question is equivalent to the one which inquires whether the product $\Delta \times \Delta' \times \vec{A}$ contains the identical representation; it is also equivalent to the one which inquires whether the product $\Delta \times \vec{A}$ contains the representation $\bar{\Delta}'$ which is contragradient to Δ'. Now, since the representation $\Delta \times \vec{A}$ is irre-

ducible, $\tilde{\Delta}'$ is a constituent of the former only if

$$\Delta \times \vec{A} = \tilde{\Delta}' \quad \text{or} \quad \Delta = \tilde{\Delta}' \times \vec{A}.$$

As long as this relation exists between the representations Δ and Δ' we have in the product space one antisymmetric eigenfunction, else none. In other words, as long as this relation is satisfied, we can form from each column of the rectangle (D_L, Δ) involving the orbital functions and from each column of the spin-rectangle one antisymmetric orbital spin-function $\phi^{(m_L, m_S)}$. Since that is also valid for all pairs of columns, m_L can assume all the values L, L-1, ..., -L and m_S all the values S, S-1, ..., -S. The crux of it all is that we obtain a multiplet term which transforms according to $D_L \times D_S$ and which satisfies the Pauli-principle.

This "first method" has always been pursued in the early treatises concerning the group-theoretical arrangement of the line-spectrum. In order to pursue this method to completion (up to the point of classifying the allowable terms and calculating to a first approximation their eigenvalues) it is necessary to actually set up the representations Δ and Δ', to calculate their characters and determine their relation to the representations D_L and D_S. The reader is referred to the book by H. Weyl[1] for this treatment.

There is, however, an older method which has recently been applied with great success by J. C. Slater[2]. It makes use of much simpler tools and, in particular, it is independent of the representation theory of the permutation group. In this method we do not discuss the permutations of the orbitals separately

---

1) Weyl, H.: Group theory and Quantum Mechanics, 2nd Edition, 1931.
2) Slater, J.C.: Phys.Rev. vol. 34 (2) 1929 p. 1293.

from those of the spins; instead, we permute the orbitals and spins simultaneously. In this manner we limit ourselves at the very outset to the antisymmetric eigenfunctions

$$\Sigma \delta_P P\phi_b(q_1,\ldots,q_f)u_\lambda v_\mu \ldots w_\nu ; \qquad \delta_P = \begin{cases} +1 \text{ for P even} \\ -1 \text{ for P odd.} \end{cases} \tag{28.3}$$

Rotations of the q-space as well as those of the spin space always again convert these antisymmetric eigenfunctions into exactly similar ones. It should, therefore, be possible to reduce out in the space of the functions (28.3) the two commutative groups of transformations which are induced by these rotations. We then obtain a rectangle whose rows are transformed according to D_L by spacial rotations and whose columns are transformed according to D_S by spin rotations. When undergoing simultaneous spin - and space-rotations, the entire rectangle suffers naturally the transformation $D_L \times D_S = \Sigma D_J$. The terms D_J split apart due to the spin perturbation, exactly as in #25. We see, therefore, that the Pauli-Exclusion Principle never rules out part of a multiplet; it either forbids the entire multiplet or permits it in its entirety.

With the aid of the second method we shall now endeavor to answer the question: which multiplets can be obtained from given electron terms?

If at first we ignore the interaction among the electrons then we may replace the function ϕ_b in (28.3) by the product

$$\phi_b = \phi(n_1|q_1)\ (n_2|q_2)\ \ldots\ (n_f|q_f).$$

Here, the single quantum number n_ν actually represents three numbers

(n, ℓ, m_ℓ). By indicating the three values (n, ℓ, m_ℓ) we therefore approximately determine the function ω. In order to also indicate the particular product $u_\lambda v_\mu \ldots w_\nu$ with which the function ϕ_b in (28.3) is to be multiplied, we introduce into each bracket the + or − symbol; the latter represent respectively $m_s = +\frac{1}{2}$ or $-\frac{1}{2}$. (That is, + represents a factor u_1; − represents a factor u_2). For example, the eigenfunction

$$\Sigma c_p P\phi(2\ 1\ 1 \mid q_1)\phi(2\ 1\ 0 \mid q_2)\phi(2\ 1\ 0 \mid q_3)u_1 v_1 w_2 \qquad (28.4)$$

is indicated by the symbol

$$(2\ 1\ 1\ +)\ (2\ 1\ 0\ +)\ (2\ 1\ 0\ -).$$

Inasmuch as the sign of the function (28.4) is of no consequence, the sequential order of the symbols (n, ℓ, m_ℓ, m_s), too, is of no consequence. No two electrons with the same symbols (n, ℓ, m_ℓ, m_s) may occur since then the sum (28.4) would be equal to zero.

If the n, ℓ of the various electrons are given and the m_ℓ, m_s are allowed to vary, we obtain an entire series of symbols. Corresponding to each symbol we can calculate the sums $M_L = \Sigma m_\ell$ and $M_S = \Sigma m_s$ and make an index of the values M_L, M_S in question. For example, in the case of three 2p-electrons the index looks like this[1]!

---

1) The negative values of M_S are omitted in the table since they are merely repetitions of the positive values with signs reversed.

	M_L	M_s
(2 1 1 +) (2 1 0 +) (2 1 -1 +)	0	3/2
(2 1 1 +) (2 1 0 +) (2 1 1 -)	2	1/2
(2 1 1 +) (2 1 0 +) (2 1 0 -)	1	1/2
(2 1 1 +) (2 1 0 +) (2 1 -1 -)	0	1/2
(2 1 1 +) (2 1 -1 +) (2 1 1 -)	1	1/2
(2 1 1 +) (2 1 -1 +) (2 1 0 -)	0	1/2
(2 1 1 +) (2 1 -1 +) (2 1 -1 -)	-1	1/2
(2 1 0 +) (2 1 -1 +) (2 1 1 -)	0	1/2
(2 1 0 +) (2 1 -1 +) (2 1 0 -)	-1	1/2
(2 1 0 +) (2 1 -1 +) (2 1 -1 -)	-2	1/2

We now collect the pairs of values (M_L, M_S) thus obtained into a double series (M_L = L, L-1, ... -L; M_S = S, S-1, ..., -S). We begin with the largest value S of M_S and seek the largest value L of M_L corresponding to it. In our case the largest values are L = 0, S = $\frac{3}{2}$. The corresponding double series includes the values M_L = 0 and $M_S = \pm\frac{3}{2}$. We next cancel these values in the table and we again seek out (among the remaining values) the largest M_S. That value is S = $\frac{1}{2}$ and the L belonging to it is L = 2. As a result, we obtain the series M_L = 2, 1, 0, -1, -2 and $M_S = \pm\frac{1}{2}$. Finally, we are left with a series with M_L = 1, 0, -1 and $M_S = \pm\frac{1}{2}$. Hence, three 2-p electrons give rise tc the allowable terms

$$^4S(L = 0,\ S = \tfrac{3}{2}),\quad ^2D\ (L = 2,\ S = \tfrac{1}{2}),\quad ^2P\ (L = 1,\ S = \tfrac{1}{2}).$$

The result, is of course, independent of the principal quantum number 2. We call the electrons belonging to the same quantum numbers n,ℓ <u>equivalent electrons</u>. The above calculation then leads

to the fact that three equivalent p-terms give rise to the terms ^4S, ^2D, ^2P.

In exactly the same manner we calculate all other cases. For the case of two electrons we obtain, of course, the following known results. Two non-equivalent electrons (n, ℓ) and (n', ℓ'), first of all, give rise to symmetric singlet terms with $L = \ell + \ell'$, $\ell + \ell' - 1, \ldots, |\ell - \ell'|$, and secondly, to antisymmetric triplet terms having the same L-values. On the other hand, two non-equivalent (n, ℓ) electrons give rise to symmetric singlet terms with $L = 2\ell, 2\ell-2, \ldots, 0$ and antisymmetric triplet terms with $L = 2\ell-1, 2\ell-3, \ldots, 1$. As a rule, the triplet terms lie below the singlet terms.

Before we set up the corresponding rules for more than two electrons we shall formulate a few general rules which follow quite automatically upon application of the above method.

Rule 1: A completed shell (n, ℓ) in which all pairs of values $m_\ell = \ell, \ell-1, \ldots, -\ell$, $m_s = s, \ldots, -s$ each occur once does not increase the term-multiplicity. The latter is left simply unaltered in all the rows of the index and it exerts no influence on the values of M_L and M_S since $\Sigma m_\ell = 0$ and $\Sigma m_s = 0$.

By virtue of this rule, we need consider only the relatively few electrons which are beyond the completed shell in the atom. These are known as emitting electrons.

Rule 2: If (after discarding the completed shell) one is left with two (or more) non-equivalent groups of equivalent electrons, one first calculates for both groups separately the values of Σm_ℓ and Σm_s. In this manner one obtains certain pairs of values M_L', M_S' for one group and certain other pairs M_L'', M_S''

for the other group. The next step consists in forming through all possible combinations $M_L = M'_L + M''_L$ and $M_S = M'_S + M''_S$ and then seeking out (by known method,) the double series ($M_L = L, L-1, \ldots, -L$, $M_S = S, S-1, \ldots, -S$). An alternate and rather obvious method proceeds by first finding for the two groups separately the double series (L', S') and (L'', S'') and then combining every L' with every L'' according to the rule $L = L' + L'', L' + L'' - 1, \ldots, |L' - L''|$ and likewise combining every S' with every S'' according to the formula $S = S' + S'', S' + S'' - 1, \ldots, |S' - S''|$. The result is the same in either case.

Example: In nitrogen (7 electrons), where the 1s- and 2s-orbitals are filled and where there are two 2p-electrons and one 3s-electron, what terms are formed? The two 2p- electrons yield the terms $^3P(L' = S' = 1)$, $^1D(L' = 2, S' = 0)$ and $^1S(L' = 0, S' = 0)$. On combining these with the 3s - electron ($\ell = 0$, $s = \frac{1}{2}$) one obtains the terms ^4P, ^2P, ^2D, ^2S. The term symbol for the ^4P-term is for example: $1s^2\ 2s^2\ 2p^2\ 3s\ ^4P$ - or in short, (if we omit the completed shells) $2p^2\ 3s\ ^4P$. The term is only a triplet although it belongs to the quartet system (cf. #25).

In general all permitted values are easily calculated by rule 2, provided we know the permitted values in the case of equivalent electrons. There are at most two equivalent s-electrons, six equivalent p-electrons, etc. The cases regarding one, two or three equivalent p-electrons we have already discussed. Let us now write down for each of 4 equivalent p-electrons the allowable series of 4 symbols (n, ℓ, m_s, m_ℓ). We note that our work can be considerably simplified when instead of writing down four such series of symbols we write down only two, corresponding to the two electrons which

are needed to complete the shell. For since Σm_l and Σm_s are always equal to zero for a completed shell, it follows that for the two missing electrons these sums must have a value opposite to that of the other four electrons. In cases of four electrons, therefore, we need only consider two of the electrons and then reverse the signs of M_L and M_S. The double series (M_L = L, L-1,....-L; M_S = S,S-1,....,-S) however are not affected by this change of sign. In other words, four equivalent p-electrons yield the same term multiplicity as two.

The same consideration is evidently also true of other cases. Hence, the following rule obtains:

Rule 3: Four equivalent p-electrons yield the same term multiplicity as do two of the same and five equivalent p-electrons the same as a single one. Likewise, six equivalent d-electrons yield the same term multiplicity as four, seven the same as three, eight the same as two, and nine the same as a single d-electron.

Below we tabulate the allowable terms for the most important cases of equivalent s-, p- and d-electrons:

s^2 : ^1S.

s^1 : ^2S.

p^6 : ^1S.

p^1 or p^5 : ^2P.

p^2 or p^4 : ^3P, ^1D, ^1S.

p^3 : ^4S. ^2D, ^2P.

d^0 : ^1S.

d^1 or d^9 : ^2D.

d^2 or d^8 : ^3F, ^3P, ^1G, ^1D, ^1S.

d^3 or d^7 : ^4F, ^4P, ^2H, ^2G, ^2F, ^2D, ^2D, ^2P.

d^4 or d^6 : ^5D; ^3H; ^3G; ^3F, ^3D; ^3P, etc.

d^5 : ^6S, ^4G, ^4F, ^4D, ^4P, etc.

Example: What is the ground state of the oxygen atom? When all eight electrons are placed one by one into the lowest orbitals one obtains the symbol $1s^2\ 2s^2\ 2p^4$. In addition to the completed shells we have four equivalent p-electrons; that is, we have the terms ^1S, ^3P, ^1D. According to a verified empirical rule the terms of highest multiplicity have the lowest energy; In our case, the triplet term has the lowest energy. Hence, the ground state is specified by the symbol: $1s^2\ 2s^2\ 2p^4\ ^3P$ or, in short, $2p^4\ ^3P$. The term is "even": $w = +1$ (cf. #18, 3).

The lowest terms to be combined with the ground state are obtained on raising one of the 2s- or 2p-electrons into a higher orbital. On raising a 2s-electron into the 2p-orbital we obtain the electron configuration $1s^2\ 2s\ 2p^5$ and the terms ^3P and ^1P; of these, only the triplet term combines with the ground state. If we raise a 2p-electron into the 3s-orbital, we obtain the configuration $2p^3\ 3s$. The three 2p-electrons by themselves would yield ^4S, ^2P, ^2D; when we include also the 3s-electron we obtain ^5S, ^3S; ^3P, ^1P; ^3D, ^1D. Only ^3S, ^3P, ^3D combine with the ground state. If, instead of raising the 2p electron into the 3s orbital we were to raise it to the 4s-, 5s- etc. orbitals, we would then obtain three term series ^3S, ^3P, ^3D whose limits are the ionic terms $2p^3\ ^4S$, $2p^3\ ^4P$, $2p^3\ ^2D$.

It is recommended to the reader that he likewise determine the ground states of all the elements in the first row of the periodic system. For a detailed discussion of the spectra of the individual elements, I refer the reader to the book by F. Hund[1].

---

1) Hund, F.: Linienspektren und periodisches System der Elemente. Berlin, 1927.

#29 Approximate Calculation of the Energy.

When the interaction of the electrons is completely ignored, one does not get a good approximation for the energy-values and eigen-functions of an atom. A much better approximation is gotten when, for each individual electron one replaces the effect of the other electrons by a screening of the nuclear field in an appropriate manner. The "method of the self-consistent field" due to Hartree[1] gives a very useful form for this screening. In this method one attempts to determine, for each single electron, the potential of the screened-off field such that it satisfies the following: One calculates the wave functions $\psi_\alpha(q)$ of the individual electrons in this field. One then averages the charge density of all electrons except the ν-th,

$$-e\sum_{\alpha\neq\nu}\bar{\psi}_\alpha\psi_\alpha$$

over each sphere r = const. and adds this potential field and the effect of the nucleus to give a total potential. For the 'self-consistent field' this potential will agree with that assumed orginally in calculating $\psi_\alpha(q)$. It turns out that these fields which are "compatible with one another", can be determined with sufficient accuracy by means of successive approximations. It is found that in all the cases calculated, the energy-values thus obtained (we denote them by E_o) agree quite well with the observed term-values. We may, therefore, assume that the product of the Hartree electron-eigenfunctions

---

1) Hartree, D.R.: Proc. Cambr. Phil. Soc. vol.24 (1928) p. 89.

$$\psi_b = \psi_1(q_1)\ \psi_2(q_2) \ \ldots\ \psi_f(q_f) \qquad (29.1)$$

gives us a useful approximation for the eigenfunction of the system. This assumption is confirmed by theoretical considerations regarding the order of magnitude of the off-diagonal terms of the energy-matrix corresponding to the functions (29.1).[1] Now, in order to calculate with greater accuracy the atom's terms and their splitting due to the interaction (without spin), one applies perturbation theory in which the functions (29.1) are used as a first approximation. It is, then, suitable to choose the screening-fields and, along with them, the ψ-functions in a slightly different manner from that of Hartree - namely, by choosing for all electrons and all states one and the same (average) screening-field. The gain consists in that the functions (29.1) and their permutations are all eigenfunctions of one and the same "unperturbed" operator

$$H_0 = -\frac{h^2}{2\mu}\sum_\alpha \Delta_\alpha + \sum_\alpha \left(-\frac{e^2 Z}{r_\alpha} + eU(r_\alpha)\right)$$

($U(r)$ = screening-potential). These form, therefore, an orthogonal system, a convenient thing for the perturbation calculation. We now write for the eigenfunctions ψ_b, more specifically,

$$\psi_b = \psi(n_1/q_1)\psi(n_2/q_2)\ \ldots\ldots\ \psi(n_f/q_f) \qquad (29.2)$$

where each n_ν denotes a system of three quantum-numbers (n, l, m_l). If in (29.2) we vary the quantumnumbers m_l and also permute the elec-

---

1) Gaunt, J.A.: Proc. Cambr. Phil. Soc., vol.24 (1928) p.328
Slater, J.C.: Phys. Rev., vol. 32 (1928) p. 339.

trons, then there arises a system of functions ψ_b which may be differentiated from one another by the numbers b and all of which belong to the same eigenvalue E_o of the operator H_o. If we assume that this term E_o is so far removed from the neighboring terms that no mutual perturbation sets in, then the splitting of this term is calculated, according to perturbation theory, by means of a transformation of the principal-axes of the perturbation-matrix (ω_{ab}). This (the splitting) is obtained by applying the perturbation-operator (interaction minus screening)

$$W = \sum_{\varkappa,\lambda} \frac{e^2}{r_{\varkappa\lambda}} - \sum_{\lambda} e\,U(r_\lambda)$$

on ψ_b and expanding in terms of the eigen-functions of H_o:

$$W\psi_b \sim \sum \omega_{ab}\psi_a + \quad \ldots\ldots \qquad (29.3)$$

where, on the right-hand side, only those terms enter which belong to the system of the ψ_b .

Let us, now, consider any term of the operator W, e g. the term $\frac{e^2}{r_{12}}$. If we multiply (29.2) by this expression the factors $(n_3|q_3) \ldots (n_f|q_f)$ remain unaltered; hence, one need but expand the product $\frac{e^2}{r_{12}}\psi(n_1|q_1)\psi(n_2|q_2)$ in terms of the products $\psi(n_1'|q_1)\psi(n_2'|q_2)$ where only those terms enter which, as a result of variation of the quantum-numbers m_λ or exchange of q_1 and q_2, arise from $\psi(n_1|q_1)\psi(n_2|q_2)$. The expansion coefficients are:

$$A(n_1'n_2'|n_1n_2) = \iint \bar{\psi}(n_1'|q_1)\bar{\psi}(n_2'|q_2)\frac{e^2}{r_{12}}\psi(n_1|q_1)\psi(n_2|q_2)dq_1dq_2 \qquad (29.4)$$

The expressions $A(n_\lambda' n_\mu' \mid n_\lambda n_\mu)$ are formed in an analogous fashion. It is even easier to evaluate the terms of (29.3) due to the screening potentials $U(r_\lambda)$. One first expands the expression $-eU(r_\lambda)\psi(n_\lambda \mid q_\nu)$ in terms of the $\psi(n_\lambda \mid q_\lambda)$, where the factor $Y_l^{(m)}$, (spherical harmonics of the order l) which is independent of r, is unaltered by $\psi(n_\lambda \mid q_\lambda)$, and where only the term having the same principal quantum-number $n' = n$ enters. The only expansion coefficient different from zero is, therefore:

$$B(n_\lambda) = B(n_\lambda \mid n_\lambda) = -\int \overline{\psi}(n_\lambda \mid q_\lambda) eU(r_\lambda) \psi(n_\lambda \mid q_\lambda) dq_\lambda \qquad (29.5)$$

in which the integral may even be replaced by an integral on r_λ only, so that it is independent of the quantum-number m.

The addition of all these expressions A and B gives the elements ω_{ab} of the perturbation-matrix, whose eigenvalues ζ_ν determine the corrected energy-values

$$E_\nu = E_0 + \zeta_\nu$$

In order to carry out the principal-axes transformation of this matrix (ω_{ab}), one has to introduce instead of the ω_b new linear combinations which can be determined by reduction of the rotation and the permutational groups. Of these linear combinations only those are required which satisfy the Pauli-principle. Both methods of the previous section may, therefore, be applied. One can proceed either by first reducing out the spin-less functions (29.2) and then introducing the Pauli principle or, in accordance with Slater, by immediately and simultaneously introducing the spin and forming the anti-symmetric linear combinations. In both methods one may

be spared the explicit calculation of the correct linear combinations by virtue of certain considerations regarding the traces. In the first method, one needs the characters of the symmetric group; in the second method that is unnecessary. We shall follow the second, simpler method of Slater.

We, thus, immediately introduce the spin-coordinates σ_z along with the space-coordinates q. In place of the pure space-functions $\psi(n|q)$ (which are specified by three quantum-numbers (n, ℓ, m_ℓ) there now are the space-spin-functions

$$\psi(n, m_s | q, \sigma_z) = \psi(n|q) \cdot u_\lambda$$

which are specified by four quantum-numbers (n, ℓ, m_ℓ, m_s). We write for these four quantum numbers (n, ℓ, m_ℓ, m_s) one single symbol ϱ; and for the space- and spin-coordinates q, σ_z we write a single symbol χ. Instead of (29.2) we thus have:

$$\psi_\beta = \psi(\varrho_1|\chi_1)\psi(\varrho_2|\chi_2)\cdots\psi(\varrho_f|\chi_f) \qquad (29.6)$$

The number β represents a row of symbols $\varrho_1, \varrho_2, \cdots, \varrho_f$. Upon multiplication of (29.3) by the spin-functions $u_\lambda v_\mu \ldots$, there follows:

$$W\psi_\beta = \sum \omega_{\alpha\beta}\psi_\alpha + \cdots \qquad (29.7)$$

where the $\omega_{\alpha\beta}$ are sums of expressions of the following type:

$$A(\varrho'_\lambda \varrho'_\mu | \varrho_\lambda \varrho_\mu) = \begin{cases} A & (n'_\lambda n'_\mu \mid n_\lambda n_\mu) \text{ for} \begin{cases} m'_{s\lambda} = m_{s\lambda} \\ m'_{s\mu} = m_{s\mu} \end{cases} \\ 0 & \text{otherwise} \end{cases} \qquad (29.8)$$

and

$$B(\varrho_\lambda | \varrho_\lambda) \quad B(n_\lambda | n_\lambda).$$

If, now, we carry out the permutation P upon the arguments $x_1, \cdots x_f$ in (29.6), or, what amounts to the same thing, if we carry out the permutation P^{-1} upon the quantum number symbols $e_1 \cdots e_f$, then there arises from ψ_β a function $P\psi_\beta$ which, again, belongs to the system of the ψ_β and may, thus, be denoted by $\psi_{P\beta}$.

We now form the anti-symmetric linear combinations:

$$\Psi_\beta = \sum_P \delta_P P\psi_\beta$$

Upon carrying out the operations $\Sigma \delta_P P$, which commute with W, we get from (29.7):

$$W\Psi_\beta = \Sigma \omega_{P\alpha,\beta} \Psi_\alpha + \cdots$$

The same term Ψ_α may occur several times on the right-hand side, since there occur also along with the Ψ_α, the $\Psi_{P\alpha} = \delta_P \Psi_\alpha$ having the coefficients $\omega_{P\alpha,\beta}$. Collecting all these terms we have:

$$W\Psi_\beta = \Sigma \Psi_\alpha \Omega_{\alpha\beta} \; ; \; \Omega_{\alpha,\beta} = \sum_P \delta_P \omega_{P\alpha,\beta} \tag{29.9}$$

It is convenient for future reference to replace the perturbation operator by the complete energy operator $H = H_o + W$. Its matrix $(\theta_{\alpha\beta})$ is the same as $(\Omega_{\alpha\beta})$ except for the diagonal terms:

$$\theta_{\beta\beta} = \Omega_{\beta\beta} + E_o \tag{29.10}$$

We denote this matrix $(\theta_{\alpha\beta})$ simply by H. Its eigen-values are

the energy-values $E_\nu = E_0 + \zeta_\nu$.

The principal-axis transformation of the matrix H is considerably simplified by the following observation: If the function ψ_β in (29.7) belongs to definite eigen-values $M_L = \Sigma m_l$ and $M_S = \Sigma m_s$ of the operators L_3 and S_3, then all the terms on the right-hand side, likewise, belong to the same values. Consequently, the matrix $(\Omega_{\alpha\beta})$ and thus also the energy matrix H break up into as many sub-matrices $H(M_L, M_S)$ as there are pairs of values (M_L, M_S).

We now form, for each pair of values (M_L, M_S), the trace of the matrix $H(M_L, M_S)$:

$$Sp(M_L, M_S) = \sum_{\substack{\Sigma m_l = M_L \\ \Sigma m_s = M_S}} \theta_{\beta\beta} \qquad (29.11)$$

This trace must have the same value in the matrix $H(M_L, M_S)$ transformed to diagonal form. Now, there correspond to each energy-term E_ν certain quantum numbers L,S and also $(2L+1)(2S+1)$ eigen-functions

$$\Psi_{L,S}^{(M_L, M_S)} \qquad (-L \leq M_L \leq L, \quad -S \leq M_S \leq S).$$

To each one of these there corresponds a diagonal term E_ν of the diagonalized matrices $H(M_L, M_S)$. Consequently, the trace $Sp(M_L, M_S)$ of the matrix $H(M_L, M_S)$ is equal to the sum of all those terms E_ν for which $L \geq |M_L|$ and $S \geq |M_S|$:

$$Sp(M_L, M_S) = \sum_{\substack{L \geq |M_L| \\ S \geq |M_S|}} E_\nu (L,S) . \qquad (29.12)$$

As soon as the traces on the left-hand side are known, we have in (29.12) a linear system of equations for the determination of the terms E_γ. If, in particular, there corresponds to each pair of values (L,S) only <u>one</u> term $E_\gamma(L,S)$, as is often the case, then the system of equations (29.12) is sufficient for the determination of all the E_γ .

The diagonal terms $\theta_{\beta\beta}$, of which $\Sigma(M_L, M_S)$ is composed, are to be gotten from (29.9) and (29.10):

$$\theta_{\beta\beta} = \Omega_{\beta\beta} + E_0 = \sum_P \delta_P \omega_{P\beta,\beta} + E_0 \qquad (29.13)$$

In order to calculate them one has to seek in (29.7) on the right-hand side the terms with $\alpha = P\beta$ or $\psi_\alpha = P\psi_\beta$. Now in reality there occur in (29.7) on the right-hand side only those terms for which at most two quantum-number symbols e_λ , e_μ of ψ_β are changed into e'_λ , e'_μ ; hence, the permutation P can only be the identity ($e'_\lambda = e_\lambda$; $e'_\mu = e_\mu$) or a transposition ($\lambda\mu$)($e'_\lambda = e_\mu$; $e'_\mu = e_\lambda$). Accordingly, there actually occur in (29.13) only the terms

$$\left.\begin{aligned} \omega_{\beta\beta} &= \sum_{\lambda,\mu} A(e_\lambda e_\mu | e_\lambda e_\mu) + \sum_\lambda B(e_\lambda | e_\lambda) \\ \omega_{(\lambda,\mu)\beta,\beta} &= A(e_\lambda e_\mu | e_\mu e_\lambda) \end{aligned}\right\} \qquad (29.14)$$

The last term, "the exchange-integral" is, according to (29.8) different from zero only when the spins of the λ -th and the μ-th electron are in the same direction. The first term, $\omega_{\beta\beta}$ however, is independent of the spins. It merely represents the average

of the perturbation energy W in the state ψ_β . The terms $B(\varrho_\lambda | \varrho_\lambda)$ are even independent of the quantum-numbers m_ℓ . We wish to collect them along with the term E_0 of (29.13) into a single expression:

$$I = E_0 + \sum_\lambda B(\varrho_\lambda, \varrho_\lambda)$$

If we set:

$$J(\varrho, \varrho') = A(\varrho\varrho' | \varrho\varrho'); \quad K(\varrho, \varrho') = A(\varrho\varrho' | \varrho'\varrho),$$

then according to (29.13):

$$\theta_{\beta\beta} = \theta(\varrho_1, \varrho_2, \cdots, \varrho_f) = I + \sum_{\lambda,\mu} J(\varrho_\lambda, \varrho_\mu) - \sum_{\lambda,\mu} K(\varrho_\lambda, \varrho_\mu) \quad (29.15)$$

The terms $I + \sum J$ in this expression give the average energy of the state ψ_β . As was already noted in #26, the exchange-integral K is chiefly positive. Since, for the sum $\sum K$, only those electron-pairs having their spins in the same direction enter our consideration, then $\sum K$ is largest when there are as many spins as possible in the same direction. That is, for the largest values of M_S and S. Thus we see an explanation for the empirical rule which states that the terms of greatest multiplicity $2S + 1$ are usually lowest. Equations (29.12) indicate, for each individual case, the positions of the terms.

Example: We shall consider two electrons, one of which is in an s-orbital. That is, ns, n'p. (For the cases ns, n's, or ns, n'd etc. the calculation is exactly the same.) The possible terms are: ^1P, ^3P; the corresponding term-values are denoted in like manner. In equations (29.12) it is sufficient to choose $M_L = 0$ and $M_S = 0$. From (29.12) we obtain:

$$Sp(0,0) = {}^1P + {}^3P,$$
$$Sp(0,1) = {}^3P$$

Upon writing out the symbols $\varrho_\lambda = (n\, l\, m_l\, m_s)$, we get according to (29.11) and (29.15):

$$\begin{aligned} Sp(0,1) &= \theta\,(n00+,\ n'10+) \\ &= I + J(n00,n'10) - K(n00,n'10), \\ Sp(0,0) &= \theta\,(n00+,n'10-) + \theta\,(n00-,n'10+) \\ &= 2\theta\,(n00-,n'10+). \\ &= 2I + 2J(n00,\ n'10) \end{aligned}$$

There follows, thus:

$${}^3P = I + J - K$$

$${}^1P = I + J + K$$

The difference between the two terms, therefore, is equal to twice the exchange-integral, as in #26.

The discussion of more complicated examples is facilitated by means of the following useful considerations: If we consider a completed shell and an additional electron χ_f (with quantum-numbers n, l, m_l, m_s) then each eigen-function $\psi(\varrho_f, \chi_f)$ of the individual electron gives rise to only one eigen-function ψ_β of the system, whose energy E_γ is independent of the quantum-numbers $M_S = m_s$ and $M_L = m_l$. Each of the sub-matrices $H(M_L, M_S)$ has only one single element $\theta_{\beta\beta} = I + \sum J - \sum K = E_\gamma$. Hence $\sum J - \sum K$ must be independent of m_l and m_s. Those summations in $\sum J$ and $\sum K$ which describe the interaction of the electron-pairs inside of the closed shell are, of course, independent of

m_l and m_s since they have no effect at all upon the quantum numbers of the outer electron. Hence the sum of the terms in $\sum J - \sum K$ which describe the interaction of the outer electron α_f with the electrons of the closed shell is independent of the quantum-numbers m_l, m_s of this outer electron.

Naturally, this rule is still valid if in addition to the one electron and the closed shell there are yet other electrons, for the values of the integrals J and K (which always refer to two electrons) are unaffected by the presence of the other electrons. The terms in $\sum J$ and $\sum K$ which refer to those electron-pairs, of which one or both electrons are in the closed shell, make a fixed contribution to all matrix-elements $\theta_{\beta\beta}$ and thus to all energy terms E_γ. This means that the presence of the closed shell does affect the position of the term-system but not the splitting. As for the calculation of the splitting, one may limit oneself to those electron-pairs which are situated outside of the closed shell.

For the evaluation of the integrals J and K, as well as for the discussion of further examples, we refer to the original treatise by J. C. Slater.[1)]

#30. The Pure Spin-functions and their Transformation due to Rotations and Permutations.

In making use of the "second method" in #28 and #29, two questions remained unanswered, namely: to which representation of the Permutation-Group do the eigen-functions we set up belong if

---

1) Slater, J.C.: Phys Rev., vol. 34 (2) (1929) p.1293.

we ignore the spin completely? And with which spin-functions must they be multiplied in order to obtain the correct (anti-symmetric) wave functions?

In order to answer these questions, we must refer back to the "first method"; that is, we must separate the space-functions from the spin-functions and we must separately reduce out, for both of them, the rotation and permutation groups. According to #25 we know that two representations Δ and Δ' of the permutation group, belonging together, always enter consideration insofar as the space- and spin-functions are concerned. Between these there exists the relation (28.2). Consequently, it is sufficient to determine Δ'. We may limit ourselves, then, to the pure spin-functions.

All spin-functions of f electrons are linearly composed of 2^f products

$$u_\lambda v_\mu \ldots t_\nu \qquad (\lambda, \mu, \ldots = 1,2). \tag{30.1}$$

They thus form a 2^f-dimensional vector-space R which is linearly transformed, firstly, by permutations of the electrons and, secondly, by rotations of space; or, what amounts to the same thing, it is linearly transformed by a simultaneous unitary transformation of the pairs of variables $u_\lambda, v_\mu, \ldots\ldots, t_\nu$. We thus find in R a representation π of the permutation-group S_f and a representation δ of the unitary group u_2. According to #13, the reduction of these two representations can occur simultaneously since the operations of the two groups commute with one another. Even more can be said, namely:

<u>All matrices T which commute with the matrices of the system</u>

Π are linear combinations of the matrices of the system $\mathfrak{S}$.

Proof: The transformation T is given by:

$$T\, u_\lambda v_\mu \cdots t_\nu = \sum c_{\lambda'\lambda,\mu'\mu,\ldots,\nu'\nu}\, u_{\lambda'} v_{\mu'} \cdots t_{\nu'} \qquad (30.2)$$

If, now, T commutes with the transformations brought about by the permutations of the letters, u, v, ...w, then the coefficients $c_{\lambda'\lambda,\mu'\mu,\ldots,\nu'\nu}$ must transform into themselves upon permutations of the index-pairs. If we write the index ℓ instead of the pair of indices λ, λ', and, likewise, m instead of μ,μ', etc., then $c_{\ell,m,\ldots n}$ must be symmetric in all the indices.

The system $\mathfrak{S}$ consists of all transformations which are brought about by:

$$u'_\lambda = \sum c_{\lambda'\lambda} u_{\lambda'} \quad ; \quad v'_\mu = \sum c_{\mu'\mu} u_{\mu'} \quad ; \quad \cdots$$

with

$$\begin{pmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{pmatrix} = \begin{pmatrix} \alpha & \beta \\ -\bar{\beta} & \bar{\alpha} \end{pmatrix}$$

and,

$$\alpha\bar{\alpha} + \beta\bar{\beta} = 1$$

This yields,

$$u'_\lambda u'_\mu \cdots w'_\nu = \sum c_{\lambda'\lambda} c_{\mu'\mu} \cdots c_{\mu'\mu}\, u_{\lambda'} v_{\mu'} \cdots w_{\nu'}$$

that is, it yields a transformation (30.2), with coefficients:

$$c_{\lambda'\lambda,\mu'\mu,\ldots,\nu'\nu} = c_{\lambda'\lambda}\, c_{\mu'\mu} \cdots c_{\nu'\nu}$$

In brief:

$$c_{lm\ldots n} = c_l c_m \cdots c_n . \qquad (30.3).$$

It is yet to be proved that all symmetrical $c_{lm\ldots n}$ are linear combinations of the specific $c_{lm\ldots n}$ of (30.3), or that all linear equations

$$\sum \mathcal{H}_{lm\ldots n}\, c_{lm\ldots n} = 0 \qquad (30.4)$$

which are valid for the specific $c_{lm\ldots n}$ (30.3), are also valid for each symmetrical $c_{lm\ldots n}$.

We set:

$$\left.\begin{aligned} c_{11} = \alpha &= a_1 + ia_2 , \quad & c_{12} = \beta &= a_3 + ia_4 \\ c_{12} = \bar{\alpha} &= a_1 - ia_2 , \quad & c_{21} = \bar{\beta} &= -a_3 \quad ia_4 \end{aligned}\right\} \qquad (30.5)$$

If, now, the following equations holds:

$$\sum \mathcal{H}_{lm\ldots n}\, c_l c_m \cdots c_n = 0 \qquad (30.6)$$

Then upon inserting (30.5) we obtain an equation which is valid for all real a_1, a_2, a_3, a_4 for which $a_1^2 + a_2^2 + a_3^2 + a_4^2 = 1$. Due to the homogeneity, this equation still holds when all the a_k are multiplied by a common real factor τ , that is, the equation holds identically in a_1, a_2, a_3, a_4 and hence all its coefficients vanish. Now, in a reverse manner, we can calculate from these coefficients the coefficients of (30.6) since we can find a unique solution for the a_k in (30.5). Hence the coefficients of (30.6) also vanish. In other words,

$$\sum_P P\, \mathcal{H}_{lm\ldots n} = 0$$

where P extends over all the permutations of the indices. From

this, however, (30.4) follows for arbitrary $c_{lm\ldots u}$ which are symmetrical in the indices, which we set out to prove.

Thus, the system of matrices which commute with Π, known as σ, consists of all the linear combinations of the matrices of δ. If a sub-space of the vector-space is invariant under σ, then it is also invariant under δ. Conversely, if it is irreducible with respect to σ, then it also is irreducible with respect to δ. If two sub-spaces are equivalent with respect to σ, then they are also equivalent with respect to δ, and conversely.

According to #13, the system σ is easily set up. If we reduce completely the representation Π and arrange the base-vectors in rectangles

$$\begin{matrix} V_{11},\ldots,V_{1n} & V'_{11},\ldots, & V_{1n'} \\ \vdots \quad\quad \vdots & \vdots \quad\quad \vdots & \\ V_{k1},\ldots,V_{kn} & V'_{k'1},\ldots, & V'_{k'n'} \end{matrix} \quad ;\ldots; \qquad (30.7)$$

the rows of which are equally and irreducibly transformed by the group Π, then the columns of this rectangle are equally and, otherwise, arbitrarily transformed by the transformations of σ. From this it follows that for the representation δ also the columns of this rectangle determine equivalent irreducible representation-spaces, whereas the columns of different rectangles undergo nonequivalent transformations. To each irreducible representation of the permutation-group contained in Π there belongs by virtue of the rectangle (30.7) a definite irreducible representation of the unitary group u_2 or the rotational group d. .

Since the various irreducible representations of u_2 can be differentiated from one another by means of their spin-number s, this spin-number may simultaenously serve as a means of differentiating between the irreducible components of π as follows: For a given electron-number f there corresponds to each s a very definite irreducible representation Δ' of the permutation-group contained in π; and to different s there belong different representations. We denote this representation Δ' by Δ'_s. The only values of s that enter our consideration are: $s = \frac{f}{2} - g$ (g an integer, $\leq f/2$).

In order to write down the rectangle (30.7) explicitly we next seek to determine, for each $s = (f/2) - g$, such quantities in the vector-space R which are transformed according to the representation D_s for the group u_2. We may proceed as in #13. We introduce a contra-gradient pair of variables x_1, x_2 and we form from g "commutators" of the type $u_1v_2 - u_2v_1$ and f - 2g "linear-factors" of the type $u_1x_x + u_2x_2$ the invariant expression:

(30.8)

$$B = (u_1v_2 - u_2v_1)\ldots(p_1q_2 - p_2q_1)(r_1x_1 + r_2x_2)\ldots(t_1x_1 + t_2x_2).$$

The coefficients X^S_M of the monomial

$$X^S_M = \frac{x_1^{S+M} \quad x_2^{S-M}}{\sqrt{(S+M)!\ (S-M)!}}$$

are then transformed in R according to the representation D_S. As a result, we have found for each of the rectangles (30.7) a column. The other columns are found by permuting the letters u up to t. Among all the expressions resulting from permutation we retain only the system of linearly-independent ones. Each row of the rect-

angle thus obtained is transformed by permutations according to the representation Δ'_S .

Example: For f = 3 we have in the $\mathfrak{f}^3$ -dimensional space of the products $u_\lambda v_\mu w_\nu$, the following rectangles:

$S=\frac{3}{2}$:

$$\begin{array}{c} \sqrt{3}\ u_1v_1w_1 \\ u_1v_1w_2 + u_1v_2w_1 + u_2v_1w_1 \\ u_1v_2w_2 + u_2v_1w_2 + u_2v_2w_1 \\ \sqrt{3}\ u_2v_2w_2 \end{array}$$

$S=\frac{1}{2}$:

$$\begin{array}{cc} (u_1v_2 - u_2v_1)w_1 & (u_1w_2 - u_2w_1)v_1 \\ (u_1v_2 - u_2v_1)w_2 & (u_1w_2 - u_2w_1)v_1 \end{array}$$

It is to be noted that the spin-functions η^S_M , the coefficients of (30.2), can be characterized by the fact that they are anti-symmetric in the first g electron-pairs and symmetric in the other f - 2g electrons. In terms of the vector-scheme, one should imagine the spins of the g pairs to be oriented in opposite directions and those of the remaining f - 2g electrons to be in the same direction. Accordingly, the resultant spin is

$$S = \frac{f-2g}{2} = \frac{1}{2} f - g.$$

It follows from our construction of the representations Δ'_S that all the matrix-elements of these representations are rational numbers. Hence, the representation Δ'_S is equivalent to its conjugate-complex or contra-gradient representation $\overline{\Delta}_S$ (cf.#12). As a result, the relation (28.2) reduces to:

$$\Delta_S = \Delta'_S \times \vec{A}$$

The calculation of the characters of the representations Δ'_S of S_f, is not difficult when we keep in mind what has been said before. It is sufficient to determine in two ways the trace of a transformation AP in the space R, where A is a special unitary transformation of the form

$$u_1 = \zeta u_1 \quad ; \quad u_2 = \zeta^{-1} u_2$$

and where P is a permutation. Two ways: the first, on the basis of the "rectangle-basis" (30.7) and, the second, on the basis of the base $u_\kappa v_\mu \ldots . w_\nu$. The following is the result of the calculation: If the permutation P of the letters uv.... breaks up into cyclic permutations each of $\alpha_1, \alpha_2, \ldots \alpha_k$ letters, and if χ'_S (p) denotes the character of P in the representation Δ'_S, then the following formula holds:

$$\sum_S \chi'_S(P)(\zeta^{2S} + \zeta^{2S-2} + \ldots + \zeta^{-2S})$$
$$= (\zeta^{\alpha_1} + \zeta^{-\alpha_1})(\zeta^{\alpha_2} + \zeta^{-\alpha_2}) \ldots (\zeta^{\alpha_k} + \zeta^{-\alpha_k})$$

where the summation extends over all $S = \frac{1}{2} f - g$. If we multiply both sides by $\zeta^f(1 - \zeta^2)$ and we set $\zeta^2 = z$, then there follows:

χ'_S (P) is the coefficient of z^g in the polynomial

$$(1 + z^{\alpha_1})(1 + z^{\alpha_2}) \ldots . (1 + z^{\alpha_k})(1 - z).$$

In order to obtain from this the characters of the repre-

sentation Δ_s, we need only multiply the characters of the odd permutations by -1.

We note that for cases dealing with series of n (instead of 2) variables ($\lambda, \mu \ldots . \nu = 1,2,\ldots .n$), we can also apply (with a few modifications) the method used in this section for the investigation of transformations of the power-products $u_\lambda v_\mu \ldots . t_\nu$ brought about by permutations and linear transformations of the series of variables uv..... If we choose $n \geq f$ then each irreducible representation of S_f occurs at least once as a component in the representation π of the permutation-group S_f. This fact can be made use of in the calculation of the characters of the symmetrical group. For the further treatment of these matters, I refer the reader to the original treatises of I. Schur and H. Weyl[1] as well as to H. Weyl.[2]

---

1) Schur, I.: Dissertation, Berlin 1901.
Weyl, H.: Math. Z. vol. 23 (1925) p. 271.
Schur, I.: "Sitzungsber." Berlin 1927, p.58.

2) Weyl, H.: Grouptheory and Quantum Mechanics. 2. Edition, Chapter V.

VI. MOLECULAR - SPECTRA.[1)]

#31. The Quantum-numbers of the Molecule.

In order to obtain a rough view of the possible energy levels of a molecule, and in order to treat questions regarding stability, one conceives of the molecule as a system of two fixed nuclei K,K' and f electrons q_1 to q_f in the field of these two force-centers. If we consider the nuclei to lie on the z-axis at distances $\beta\varrho$ and $\beta'\varrho$ from the center of mass, where

$$\beta = \frac{M'}{M^\circ + M'}, \quad \beta' = \frac{M^\circ}{M^\circ + M'};$$

M°, M' = nuclear masses, ϱ = internuclear distance

then this two-center-problem is invariant under the group of rotations and reflections about the z-axis, the representations of which we have already determined in #10 (example 3). The result was the following: The eigen-functions $\varphi_{\pm\Lambda}$ each have an axial quantum-number Λ whose meaning is that for a rotation $(0, 0, \mathcal{P})$ the functions $\varphi_{\pm\Lambda}$ take on the factor $e^{\mp i\Lambda\mathcal{P}}$. In the case $\Lambda = 0$, there are two types of eigen-functions, φ_o^+ and φ_o^-, which under reflection s_y $(y' = -y)$ assume the factors +1 and -1. We write for these cases $\Lambda = 0^+$ and $\Lambda = 0^-$; the corresponding representations (of the first degree) of the rotation-reflection-group are A_0^+ and A_0^-. For $\Lambda > 0$, on the other hand, there are two eigen-functions φ_Λ and $\varphi_{-\Lambda}$ belonging to the same

1) For a thorough presentation of the Theory of Molecular-Spectra cf. R. DE L. KRONIG: Band Spectra and Molecular Structure, Cambr. 1930. This book by Kronig and the previous chapter complement each other in that the group-theoretical considerations presented here are lacking in Kronig.

energy-value. These transform into one another under the reflection s_y, and together undergo the irreducible representation of the second degree A_Λ.

The terms with $\Lambda = 0^+, 0^-, 1, 2, 3, \ldots$ are denoted by the Greek letters $\Sigma^+, \Sigma^-, \Pi, \Delta, \Phi, \ldots$, which correspond to the Latin letters S, P, D, F, ... for atom-terms. When considering the spin, it is found that there is a further splitting of these terms. To these we shall come back later.

For the infinitesimal rotation I_z we have, $I_z \varphi_\Lambda = -i\Lambda\varphi_\Lambda$ hence $L_z\varphi_\Lambda = iI_z\varphi_\Lambda = \Lambda\varphi_\Lambda$; i.e., the component hL_z of the angular momentum has the discreet value $h\Lambda$ for the state φ_Λ. The remaining components hL_x, hL_y are, of course, not constants. In terms of the vector-scheme, one imagines that the instantaneous angular-momentum-vector carries out a precession about the inter-nuclear bonding line. Its Z-component remains constant $= h\Lambda$.

In reality, however, a molecule is not a system of two fixed nuclei, but rather a system of two moving nuclei k, k' and f moving electrons. If we keep the center of mass fixed at the origin then an imaginary nucleus (cf. #3) with coordinates q_o and f electrons $q_1, \ldots . q_f$ remain in motion. The entire problem is invariant to rotations and under rotations the eigen-functions undergo a representation D_K having the reflection-character $w = \pm 1$. The questions which we must answer are: What relations exist among the eigen-functions $\varphi_{\pm\Lambda}$ of the two-center-problem and the eigen-functions $\psi_K^{(m)}$ of the freely rotating molecule? What relations exist between the quantum-number Λ and the quantum-numbers K, m, w? What relations exist between the energy- value $E(\rho)$ of the two-

center problem with inter-nuclear distance ϱ, and the actual distance is not fixed?

For thetime being we shall disregard the spin. We shall expand the 2K+1 functions

$$\psi^{(m)}(q_0, q_1, \ldots, q_f)$$

into a family of eigen-functions of the free molecule which are transformed by rotations according to D_K. Here q_0 signify the coordinates of an imaginary nucleus (like q_* in #3) - this being located in the direction kk' at the distance ϱ from the center of mass.

The point $(0,0,\varrho)$ on the Z-axis into which the point q_0 moves as a result of an appropriate rotation D is denoted by Q. If, as a result of a rotation D, the function $\psi^{(m)}$ goes over into $'\psi^{(m)}$, then we have:

$$\psi^{(m)}(q_0, \ldots, q_f) = {}'\psi^{(m)}(Dq_0, \ldots, Dq_f) = \sum a_{gm}(D)\, \psi^{(g)}(Dq_0, \ldots, Dq_f),$$

where $a_{gm}(D)$ signify the elements of the representative matrix of D in the representation D_K. If, now, we choose the rotation D in such a manner that $Dq_0 = Q$, then there follows:

$$\psi^{(m)}(q_0, \ldots, q_f) = \sum_g a_{gm}(D)\, \psi^{(g)}(Q, Dq_1, \ldots, Dq_f). \qquad (31.1)$$

This basic formula takes our ψ-function back to the 2K+1 functions

$$\psi_Q^{(g)} = \psi^{(g)}(Q, q_1, \ldots, q_f) \qquad (g = K, K-1, \ldots, -K),$$

which contain two degrees of freedom less.

Naturally, the rotation D is not yet fully specified by stating the points q_o and Q, for it can be replaced by $D_{\gamma'} D$, where $D_{\gamma'}$ is a rotation $(0,0,\gamma')$ which leaves the point Q invariant. The rotation $D_{\gamma'}$ is indicated in the representation D_K by a diagonal-matrix having the elements $e^{-im\gamma'}$. If we replace D in (31.1) by $D_{\gamma'} D$, we get:

$$\psi^{(m)}(q_o,\ldots,q_f) = \sum_g e^{-ig\gamma'} a_{gm}(D)\psi^{(g)}\ (Q,D_{\gamma'} Dq_1,\ldots,D_{\gamma'} Dq_f).$$

In order that this expression agree with (31.1), the following must hold:

$$e^{-ig\gamma'}\psi^{(g)}(Q,D_{\gamma'} q_1,\ldots,D_{\gamma'} q_f) = \psi^{(g)}\ (Q,q_1,\ldots,q_f),$$

or

$$D_{\gamma'}\psi_Q^{(g)} = e^{-ig\gamma'}\psi_Q^{(g)}. \qquad (31.2)$$

Actually, we knew this from before; the present consideration shows, however, that the property (31.2) of the functions $\psi_Q^{(g)}$ is also <u>sufficient</u>, in the sense that the functions (31.1) depend only upon the co-ordinates q_o to q_f, and not upon the choice of D.

It is easily verified that, for an arbitrary choice of the functions $\psi_Q^{(g)}$ according to the condition (31.2), the functions (31.1) do actually determine a linear group which undergoes the representation D_K under rotations.

For a definite position of the nuclei on the Z-axis and for a fixed Q, the functions $\psi_Q^{(g)}$ are the eigen-functions of the free molecule. <u>We now assume that these functions agree approximately (except for a factor depending on Q) with the eigen-functions</u>

φ_Λ of the two-center problem discussed in the beginning of this section. We shall see more clearly in the next section that this assumption is justified for the case where the nuclear mass is large compared to the electron-mass. Tentatively, it is sufficient to note that for purposes of studying the electron-motion, we may as an approximation consider the much heavier nuclei to be at rest. Since the function $\psi_Q^{(g)}$ assumes the factor $e^{-ig\varphi}$ for the rotation D, then, of necessity: $g = \pm\Lambda$. In general, functions φ_Λ with different Λ belong to entirely different energy-terms; it would, thus, be pointless to combine into a single expression (31.1) such $\varphi_{\pm\Lambda}$ having different Λ , for the purpose of obtaining approximations for $\psi^{(g)}$. We therefore assume that all $\psi_Q^{(g)}$ on the right-hand side of equation (31.1) are approximately zero, with the exception of at most two amongst them which correspond to the values $g = \pm\Lambda$ and which are approximately given by $f_+(\rho)\varphi_\Lambda$ and $f_-(\rho)\varphi_{-\Lambda}$.

The relation between the functions $f_+(\rho)$ and $f_-(\rho)$ is easily determined from the behavior of the family (31.1) as it undergoes the reflection $s(x' = -x, y' = -y, z' = -z)$. If s_y is the reflection $y' = -y$ and D_y the rotation about the y-axis ($x' = -x$, $z' = -z$), then $D_y s = s_y$. Now, in order to calculate $\psi^{(m)}(sq_0,\ldots,sq_f)$ according to (31.1), we need to know the rotation which transfers the point sq_0 into Q. This is $D_y D$, for D transfers sq_0 into $sQ = (0,0,-\rho)$ and D_y transfers sQ into Q. Hence,

$$\psi^{(m)}(sq_0,\ldots,sq_f) = \sum_g a_{gm}(D_y D)\,\psi^{(g)}(Q, D_y D sq_1,\ldots,D_y D sq_f),$$

or since $Ds = sD$ and $D_y s = s_y$:

$$\psi^{(m)}(sq_0,\ldots,sq_f) = \sum_h \sum_g a_{gh}(D_y)a_{hm}(D)\,\psi^{(g)}(\varrho, s_y Dq_1,\ldots,s_y Dq_f). \tag{31.3}$$

The matrix $(a_{gh}(D_y))$ indicates how the rotation D_y transforms the base-vectors v_g of the representation D_K. Now, the v_g are transformed like $u_1^{K+g}\,u_2^{K-g}$: $\sqrt{(K+g)(K-g)}$ [cf.(17.10)] and the rotation D_y takes $u_1^{K+g}u_2^{K-g}$ into $(-1)^{K-g}\,u_1^{K-g}u_2^{K-g}$ (cf. #16); hence,

$$a_{gh}(D_y) = \begin{cases} (-1)^{K-g} & \text{for } h = -g \\ 0 & \text{otherwise.} \end{cases}$$

We insert this into (31.3):

$$\psi^{(m)}(sq_0,\ldots,sq_f) = \sum_h (-1)^{K+h} a_{hm}(D)\,\psi^{(-h)}(\varrho, s_y Dq_1,\ldots,s_y Dq_f).$$

In order that the functions $\psi^{(m)}$ correspond to the reflection-character w, this expression must agree with $w\cdot\psi^{(m)}$, i.e. the following must obtain:

$$(-1)^{K+g}\psi^{(-g)}(\varrho, s_y q_1,\ldots,s_y q_f) = w\cdot\psi^{(g)}(\varrho,q_1,\ldots,q_f),$$

or, the reflection s_y sends the function $\psi_Q^{(g)}$ into $(-1)^{K+g}\,w\,\psi_Q^{(-g)}$. Now, we have assumed above that all the $\psi_Q^{(g)}$ are approximately equal to zero, with the exception of one or two of them where $g = \pm\Lambda$. These are given by $f_\pm(\rho)\varphi_{\pm\Lambda}$. For the case $\Lambda = 0$ it turns out that as $\psi_Q^{(0)}$ undergoes the reflection s_y it assumes the factor $(-1)^K w$, i.e.

$$\left.\begin{aligned} (-1)^K &= w \quad \text{for} \quad \Lambda = 0^+ \\ (-1)^K &= -w \quad \text{for} \quad \Lambda = 0^- \end{aligned}\right\} . \tag{31.3}$$

For the case $\Lambda > 0$ we find that as $f_+(\varrho)\varphi_\Lambda$ undergoes the reflection s_y, it must go over into $(-1)^{K+\Lambda} w f_-(\varrho)\varphi_{-\Lambda}$. Since, though, φ_Λ is sent into $\varphi_{-\Lambda}$ by the same reflection, we have

$$f_-(\varrho) = (-1)^{K+\Lambda} w f_+(\varrho). \qquad (31.4)$$

Thus, for these cases, both values of w are always possible. To each $w = \pm 1$ there corresponds a family of eigen-functions (31.1). For $\Lambda = 0$, on the other hand, w is determined by (31.3). We write, henceforth, $f(\varrho)$ instead of $f_+(\varrho)$.

The next section will show that the function $f(\varrho)$ is an eigen-function of a vibration-problem and is dependent upon a vibration-quantum number $v = 0,1,2,\ldots$. The quantum-number K which determines the total angular momentum $h\bar{K}$ of the molecule is known as the rotational quantum-number. It may assume the values $K = \Lambda, \Lambda+1, \Lambda+2,\ldots$. The energy of the entire molecule depends: firstly, upon the electron-state, secondly, upon the vibrational quantum-number v and finally, to a smaller extent, upon the rotational quantum number K. Thus, for each electron-state, there exists a system of vibration-terms, each of which appears to be further split up by the rotation.

As is evident from the meaning of the quantum-number Λ for the case of the fixed nuclei, $h\Lambda$ represents the magnitude of the component of the total angular momentum along the direction of the inter-nuclear-bonding line. As was already mentioned, the terms with $\Lambda = 0,1,2,3,4$ are denoted by the letters $\Sigma, \Pi, \Delta, \Phi, \Gamma$.

Due to reasons of a general, group-theoretical type, the following selection rules obtain:

$$\left.\begin{array}{l} K \longleftrightarrow K-1,\ K,\ K+1 \quad (\text{except } 0 \longleftrightarrow 0) \\ \\ w \longleftrightarrow -w. \end{array}\right\} \tag{31.5}$$

In order to derive the selection rule for Λ we shall consider the matrix-elements of the electric moment of electrons and nuclei together. One is easily convinced that due to the slow motion of the nuclei their contributions to the rather large frequencies encountered are entirely negligible. Practically speaking, therefore, we are concerned only with the electrons. We must multiply the operators $X = \Sigma ex_\gamma$, Y, Z by the eigenfunctions (31.1) and then expand the result in terms of the eigenfunctions (31.1). The expansion holds identically in q_0 and in particular for $q_0 = Q$, $D = 1$, $\psi^{(m)} = \psi_Q^{(m)}$. Now we have the approximate equation:

$$\psi_Q^{(m)} = f_\pm(\varrho)\varphi_{\pm\Lambda} \qquad \text{for } m = \pm\Lambda$$

$$\psi_Q^{(m)} = 0 \qquad \text{otherwise.}$$

In the expansion of $(X + iY)\varphi_\Lambda$, $(X - iY)\varphi_\Lambda$, and $Z\varphi_\Lambda$ in terms of the $\varphi_{\pm\Lambda'}$, actually only the values $\Lambda' = \Lambda + 1, \Lambda, \Lambda - 1$ occur since the rotations D_γ confer upon the products $(X + iY)\varphi_\Lambda$ etc., the factors $e^{-i(\Lambda+1)\gamma}$. Likewise, in the expansion of $(X \pm iY)\varphi_0^+$ and $Z\varphi_0^+$ only the values $\Lambda = 1$ or 0^+ occur and in the expansion of $(X \pm iY)\varphi_0^-$ and $Z\varphi_0^-$ only the values $\Lambda = 1$ or 0^- occur. This is so since $Z\varphi_0^+$ and $Z\varphi_0^-$ behave, under the influence of the rotations D_γ and the reflection s_y, just like φ_0^+ and φ_0^- respectively. Hence the selection-rule for Λ is:

$$\Lambda \leftrightarrow \Lambda+1,\ \Lambda\ ,\ \Lambda-1, \quad \underline{\text{but not}}\ 0^+ \longleftrightarrow 0^-. \tag{31.6}$$

It is to be noted that for the transitions $0^+ \longleftrightarrow 0^-$ and $0^- \longleftrightarrow 0^-$ K must, of necessity, skip by 1 since otherwise, by (31.3), the selection-rule for w would be contradicted.

#32. The Rotational Levels.

We refer back to the wave-equation of the molecule, set up in #3, in order to justify accurately the rule arrived at in the preceding section and in order to calculate the rotational-splitting. Upon holding the center-of-mass fixed and ignoring the smallest terms the wave-equation becomes:

$$\frac{\hbar^2}{2M}\Delta_0\psi - \frac{\hbar^2}{2\mu}\sum_1^f \Delta_\kappa\psi + U\psi = E\psi \tag{32.1}$$

where μ is the electron-mass and $M = \frac{M^0 M'}{M^0 + M'}$ an imaginary nuclear-mass. In the equation we replace ψ by the function (31.1). In order to evaluate $\Delta_0\psi$ we write the Δ -operator in terms of polar-coordinates:

$$\Delta_0 = \frac{\partial^2}{\partial\rho^2} + \frac{2}{\rho}\frac{\partial}{\partial\rho} + \frac{1}{\rho^2}\Lambda_0$$

where Λ_0 is the known operator, depending only on θ_0 and φ_0, for which we can write, according to (6.4)

$$\Lambda_0 = -\vec{L}_0^2 \qquad -(L_{0x}^2 + L_{0y}^2 + L_{0z}^2).$$

The direct evaluation of $\Lambda_0\psi$ according to (31.1) is not an

easy matter since D and $a_{gm}(D)$ depend on θ and φ in a rather complicated manner We therefore introduce instead the total angular momentum $\vec{L}$ having the components:

$$L_x = L_{0x} + L_{1x} + \cdots + L_{fx}, \text{ etc.}$$

and the electronic spin momentum $\vec{L}'$ having the components

$$L'_x = L_{1x} + \cdots + L_{fx}, \text{ etc.}$$

L_x commutes with L'_x, L_y with L'_y and L_z with L'_z. Moreover,

$$\vec{L}_0 = \vec{L} - \vec{L}' \quad , \text{ hence:}$$

$$-\Lambda_0 = \vec{L}_0^2 = (\vec{L} - \vec{L}')^2 = \vec{L}^2 - 2\vec{L}'\cdot\vec{L} + \vec{L}'^2 \tag{32.2}$$

Now, $\vec{L}'^2\psi = K(K+1)\psi$ since ψ belongs to the representation D_K and to the angular-momentum $\hbar K$. The operator $\vec{L}'^2$ refers only to the electrons and is of the same type and magnitude as are the operators Δ_α. It is multiplied (in 32.1) however, by the thousandfold smaller mass-factor $\frac{\hbar^2}{2M}$. We therefore combine it with the operator $\Sigma\Delta_\alpha$, as a small perturbation term. We thus obtain:

$$\frac{\hbar^2}{2M}\left(-\frac{\partial^2}{\partial\rho^2} - \frac{2}{\rho}\frac{\partial}{\partial\rho} + \frac{K(K+1)}{\rho^2}\right)\psi - \frac{\hbar^2}{M\rho^2}\vec{L}'\cdot\vec{L}\psi \tag{32.3}$$

$$+\left(-\frac{\hbar^2}{2\mu}\Sigma\Delta_\alpha + \frac{\hbar^2}{2M\rho^2}\vec{L}'^2\right)\psi + U\psi = E\psi$$

The most difficult term of this sum is that containing $\vec{L}'\cdot\vec{L}$. According to (17.8) we have:

$$L_x \psi^{(m)} = \tfrac{1}{2}\sqrt{(K+m)(K-m+1)}\,\psi^{(m-1)} + \tfrac{1}{2}\sqrt{(K-m)(K+m+1)}\,\psi^{(m+1)}$$

$$L_y \psi^{(m)} = -\tfrac{1}{2i}\sqrt{(K+m)(K-m+1)}\,\psi^{(m-1)} + \tfrac{1}{2i}\sqrt{(K-m)(K+m+1)}\,\psi^{(m+1)} \qquad (32.4)$$

$$L_z \psi^{(m)} = m\,\psi^{(m)}$$

Upon introducing this into (32.3) we find that all the differential-operators acting on θ_1, φ_1 vanish. Since the family (31.1) and the differential equation (32.3) are invariant under rotations, the differential equation (32.3) is satisfied identically in D when it is satisfied in the special case D = 1, $q_0 = Q$. In this case we have:

$$\psi = \psi_Q^{(m)} = \psi(m|\rho, q_1, \ldots, q_f) \qquad (m = K, K-1, \ldots, -K).$$

The differential-eqaution (32.3) connects, as a result of (32.4), the functions $\psi_Q^{(m)}$ corresponding to the different values of m. Thus a solution of (32.3) is given by a system of 2K+1 functions $\psi_Q^{(m)}$.

As follows from (32.4): for large K the term containing $\vec{L}'\cdot\vec{L}$ in (32.3) lies, in order of magnitude, between the not-very-large term containing K(K+1) and the very small term with $\vec{L}'^2$. If we next disregard the two small terms containing $\vec{L}'\cdot\vec{L}$ and $\vec{L}'^2$, then we are left with a differential-equation in which only <u>one</u> of the 2K+1 functions $\psi_Q^{(m)}$ occurs and which is independent of the index m:

$$\frac{\hbar^2}{2M}\left(-\frac{\partial^2}{\partial\rho^2}-\frac{2}{\rho}\frac{\partial}{\partial\rho}+\frac{K(K+1)}{\rho^2}\right)\psi-\frac{\hbar^2}{2\mu}\Sigma\Delta_\alpha\psi \qquad (32.5)$$
$$+U\psi=E\psi$$

We may, therefore, choose for m = K, K-1,....,-K any <u>arbitrary</u> solutions of (32.5) which belong to the same eigen-value E and which satisfy the condition $D_\gamma\,\psi_Q^{(m)}=e^{-im\gamma}\psi_Q^{(m)}$. In particular, it is permissible to set all the $\psi_Q^{(m)}$ equal to zero except one ψ_Q^0 or two $\psi_Q^{(\pm\Lambda)}$ as we did in #31. The system of solutions of (32.5) thus found we denote by $\psi_0^{(m)}$ because we wish to take it as a first approximation to the exact solution of (32.3).

If we now add to (32.5) the very small term $\frac{\hbar^2}{2M}\vec{L}'^2$, then the eigen-functions do not change to any considerably extent, only a tiny term-displacement occurs. Of the three terms of the expression $\vec{L}'^2 = L_x'^2 + L_y'^2 + L_z'^2$, the third one can be evaluated most easily. It is: $L_z'\psi_Q^{(m)} = m\,\psi_Q^{(m)}$, hence $L_z'^2\psi_Q^{(m)} = m^2\,\psi_Q^{(m)}$. We now wish to keep this third term and ignore the other two, in spite of the fact that all three terms, of course, are of the same order of magnitude. We consider the term with $\vec{L}'\cdot\vec{L}$ in (32.3) as a second perturbation term. We apply to the approximate eigen-functions $\psi_0^{(m)}$ the perturbation operator

$$\frac{\hbar^2}{2M}(L_z'^2 - 2\vec{L}'\cdot\vec{L}) = \frac{\hbar^2}{2M}\left\{L_z'^2 - 2(L_x'L_x + L_y'L_y \quad L_z'L_z)\right\}$$

and we expand in terms of these. The application of the operators L_x or L_y upon a system $\psi_0^{(K)}$, $\psi_0^{(K-1)}$,...., $\psi_0^{(-K)}$, of which only the $\psi_0^{(\pm\Lambda)}$ are different from zero yields, according to (32.4), a system $\psi^{(K)}$, $\psi^{(K-1)}$, . . . , $\psi^{(-K)}$ in which only the $\psi^{(\pm\Lambda\pm1)}$ are different from zero. If, subsequently, we carry out the operations L_x' and L_y', which do not alter the upper

index m, then all $\psi^{(m)}$ except the $\psi^{(\pm\Lambda\pm1)}$ remain zero. If we, then, expand the functions of m, ϱ , q obtained, in terms of the functions $\psi_0^{(m)}$, only those functions actually occur which, for $m = \pm\Lambda \pm 1$, are different from zero. These belong, therefore, to the value $\Lambda' = \Lambda \pm 1$, and consequently, they belong, in general, also to another energy-value from the $\psi_0^{(\pm\Lambda)}$. The terms containing $L'_x L_x$ and $L'_y L_y$ in the perturbation-operator, therefore, make no contribution to our perturbation-calculation. There remains:

$$\frac{\hbar^2}{2M\varrho^2}(L_z'^2 - 2L'_z L_z)\psi_0^{(m)} = \frac{\hbar^2}{2M\varrho^2}(m^2 - 2mm)\psi_0^{(m)}$$

$$= -\frac{\hbar^2}{2M\varrho^2}m^2\psi_0^{(m)} = -\frac{\hbar^2}{2M\varrho^2}\Lambda^2\psi_0^{(m)} .$$

We absorb this term also into the differential equation (32.5) and replace the equation by

$$\frac{\hbar^2}{2M}\left(-\frac{\partial^2}{\partial\varrho^2} - \frac{2}{\varrho}\frac{\partial}{\partial\varrho} + \frac{K(K+1)-\Lambda^2}{\varrho^2}\right)\psi - \frac{\hbar^2}{2\mu}\sum_1^f \Delta_\alpha\psi + U\psi = E\psi \qquad (32.6)$$

The solution $\psi^{(\pm\Lambda)}$ of this differential-equation represents therefore, the first approximation of the perturbation-calculation, As is seen, there is still in this approximation no difference between the terms of different reflection-characters w. The splitting into two terms with $w = \pm 1$ (the so called σ-type-doublet) first appears in the next approximation. We shall not go into any greater detail here but just remark that the σ-type-splitting does not become noticeable for small values of K since the perturbation-term $\vec{L}'\cdot\vec{L}$ is then very small in accordance with (32.4).

We now solve (32.6) by means of the rule:

$$\psi = f(\varrho)\varphi(\varrho, q_2, \ldots, q_f) \tag{32.7}$$

where $f(\varrho)$ may be a function which changes rapidly with ϱ, but φ, on the other hand, changes so slowly with ϱ that as a result we may ignore that part of the differentation $\frac{\partial}{\partial \varrho}$ referring to φ along with the small mass-factor $\frac{\hbar^2}{2M}$ in (32.6).

The function φ is determined from the differential equation of the two-center problem:

$$-\frac{\hbar^2}{2\mu}\sum_1^f \Delta_\alpha \varphi + U\varphi = E(\varrho)\varphi \tag{32.8}$$

The function $\varrho f = F$, on the other hand, is to satisfy:

$$\left(-\frac{\hbar^2}{2M_0}\frac{\partial^2}{\partial \varrho^2} + E(\varrho) + \frac{\hbar^2}{2M_0}\frac{K(K+1)-\Lambda^2}{\varrho^2}\right)F = EF \tag{32.9}$$

One easily calculates that the function (32.7) thus determined does actually satisfy the differential equation (32.6) (ignoring the terms as prescribed) and with the aid of the assumption of the completeness of the system of the φ, it turns out as in #2, that all solutions of (32.6) may be obtained in this manner. It seems quite justifiable to have assumed that the solution φ of the two-center problem - appropriately normalized - does not depend too strongly on ϱ, considering this dependency in the perturbation - calculation would yield only a term-displacement but no splitting anyway.

The equation (32.9) has the same form as the equation for the

vibration of a material point (oscillator) in one dimension having the potential energy:

$$E(\varrho) + \frac{\hbar^2}{2M_0} \frac{K(K+1) - \Lambda^2}{\varrho} \qquad (32.10)$$

A stable molecule is, of course, only possible when this expression possesses a minimum somewhere. According to (32.8) the decisive term in (32.10), namely $E(\varrho)$, is just the energy of a fictitious molecule with fixed nuclei at a distance ϱ, whereas the additional term represents the energy of the "centrifugal - force". As $\varrho \to 0$, $E(\varrho)$ tends towards ∞, and for $\varrho \to \infty$, $E(\varrho)$ tends towards the energy $E(\infty)$ of a system of two separated atoms or ions (see Fig. 7).

When the expression (32.10) is given and has a minimum, then the differential equation (32.9) possesses a finite or infinite number of eigenvalues $E < E(\infty)$ or <u>vibration - terms</u> which are distinguished from one another by means of a <u>vibrational quantum-number</u> $v = 0,1,2,3,\ldots$. Generally speaking, a vibration-term of this type does not change very much as K runs through the values Λ, $\Lambda + 1$, $\Lambda + 2,\ldots$ for the changes in $\frac{\hbar^2}{2M_0} \frac{K(K+1)}{\varrho^2}$ are small compared to the distances between the vibration terms. For this reason, there correspond to each vibration-term $E_{v\Lambda}$ a series of densely populated <u>rotational-levels</u> which belong to the different values of K. The approximate position of the rotational levels is easily given by the perturbation-calculation. That is, if F_k is any normalized solution of (32.9) for any average value k of K, then the perturbation-term $\frac{\hbar^2}{2M_0} \frac{K(K+1) - k(k+1)}{\varrho^2}$ brings about an increase in the eigen-value above the average-value of this

expression, as follows:

$$\frac{\hbar^2}{2M}\left\{K(K+1) - k(k+1)\right\}\widehat{\varrho^{-2}}$$

where

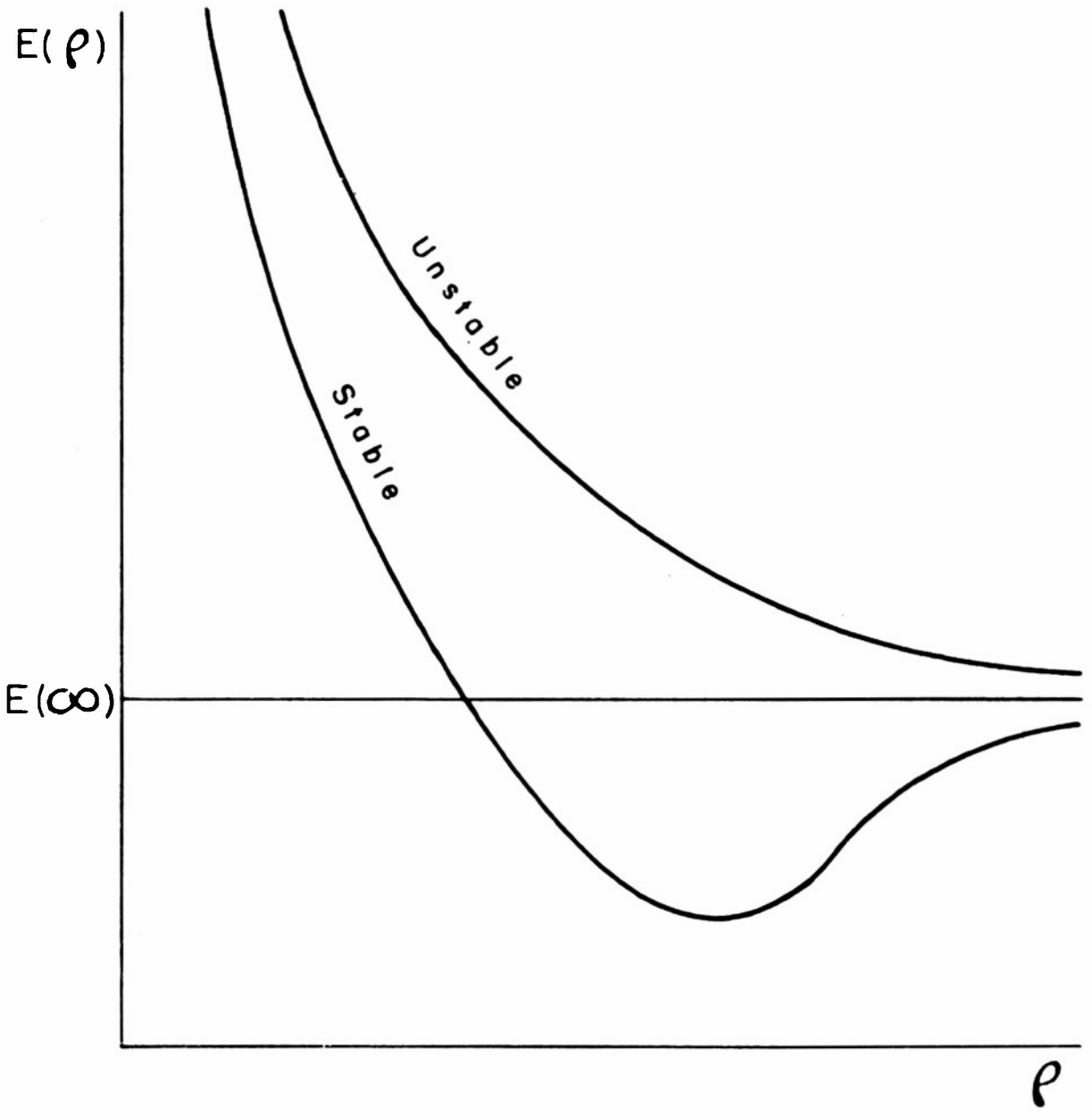
$$\widehat{\varrho^{-2}} = \int_0^\infty F_k\, \varrho^{-2}\, F_k\, d\varrho$$

A definite quantum-jump of the electron configuration (i.e. of the eigen-function $\varphi_{\pm\Lambda}$ of the two-center problem), which is connected with a jump $v \longrightarrow v'$ of the vibration-quantum-number, gives rise to a "band" in the spectrum, a group of very many, usually densely spaced lines which correspond to the various possible values of K and K'. The band breaks up into two or three "branches", the P-branch with $K \longrightarrow K + 1$, the Q-branch with $K \longrightarrow K$ and the R-branch with $K \longrightarrow K - 1$ (the arrows represent emission; for the case of absorption, they point in the opposite direction).[1] If the initial state and the final state are both Σ -states then K must jump by 1 (cf. #31, end) and the Q-branch drops out. If neither the initial- nor the final state is a Σ -state, then all three branches become doubled, since the two jumps $+1 \longrightarrow -1$ and $-1 \longrightarrow +1$ are possible as far as the reflection-character w is concerned (σ -type-doublet). The doubling appears first for the larger values of K.

All these considerations are strictly valid for singlet-terms ($S = 0$). If, however, a resultant spin-moment is present, then a further splitting occurs with which we shall now be concerned.

---

1) To be exact, one would denote the branches by P(K), Q(K), and R(K), in accordance with the newer conventions.



THE FUNCTION E(ρ)

Figure 7

#33. Inclusion of the Spin.

We have to distinguish between two cases:

a) when the multiplet-splitting (spin-effect) is large in comparison to the rotational-splitting,

b) when the multiplet-splitting is small compared to the rotational splitting.

Case a) occurs in molecules composed of heavy atoms (I_2, Hg_2, etc.) and case b) occurs in the lighter molecules (H_2, He_2, etc.) as well as always in the case of Σ -terms. The reason for this will become evident later. The transitional region where the rotational and multiplet-splitting are of the same order of magnitude is fortunately not large, since the rotational splitting decreases with increasing atomic weight and the multiplet-splitting increases at the same time. In the trasition-region the terms having a large rotational-quantum-number K tend more to case b), those with small K tend more to a).

In case b) one simply applies the theory of #32 first and then considers the spin. From each term of rotational quantum-number K and spin-quantum number S there arises, in accordance with our familiar scheme, a multiplet with $J = K + S, K + S - 1, \ldots, |K - S|$, and the same selection-rules that are valid for an atom (cf. #24) are valid here.

In the case a) of the two-center problem we must introduce the spin-coordinates before discussing the rotational-splitting. Each spin-less eigen-function $\varphi_\Lambda (q_1, \ldots, q_f)$ belongs to a definite irreducible representation of the permutation group, to which, in turn, there corresponds, according to the Pauli principle, a definite

spin-number S. The projection of the spin-vector upon the Z-axis (inter-nuclear bonding line) has the eigen-values $h\Sigma$ (Σ = S, S-1, ..., -S). To each value of Σ belongs a definite function $u_\Sigma(\sigma_1, \ldots, \sigma_f)$ of the spin-coordinates. The products $\varphi_\Lambda u_\Sigma$, or rather their anti-symmetric linear combinations

$$\Phi_{\Lambda,\Sigma} = \Sigma \delta_P P \varphi_\Lambda u_\Sigma$$

are, to a first approximation, the eigen-functions of the entire system. For the rotations D_γ $(0,0,\gamma)$ they assume the factors $e^{-i\gamma(\Lambda+\Sigma)}$. We therefore set $\Omega = \Lambda + \Sigma$.

The reflection s_y sends φ_Λ into $\varphi_{-\Lambda}$. In order to obtain the transformation of the u_Σ we note that the reflection s_y is composed of the reflection s at the initial-point and the rotation D_y about the y-axis. Under the influence of s the spin-functions u_Σ remain invariant whereas under the influence of D_y the u_Σ go into $(-1)^{S+\Sigma} u_{-\Sigma}$. Hence $\Phi_{\Lambda,\Sigma}$ becomes $\Phi_{-\Lambda,-\Sigma}$ and the two functions $\Phi_{\Lambda,\Sigma}$, $\Phi_{-\Lambda,-\Sigma}$ together undergo an irreducible representation of the rotational-reflection group for the case $\Omega > 0$. For the case $\Omega = 0$ we should, by right, have to form the sum and the difference

$$\Phi_o^+ = \Phi_{\Lambda,\Sigma} + \Phi_{-\Lambda,-\Sigma}$$

and

$$\Phi_o^- = \Phi_{\Lambda,\Sigma} - \Phi_{-\Lambda,-\Sigma}$$

which undergo the representations A_o^+ and A_o^-. However, we ignore this differentiation since it corresponds to no significant spin-splitting. The splitting in A_o^+ and A_o^- may later be considered simultaneously with the σ-type-splitting since it is of the same

order of magnitude.

Due to the spin-perturbation the $2S+1$ terms with different Σ separate from one another. Considerations similar to those which were applied to the spin-orbit interaction in atoms lead us to set the interaction-energy between the vectors $\vec{L}$ and $\vec{S}$ proportional to the scalar product

$$\vec{L}\cdot\vec{S} = L_x S_x + L_y S_y + L_z S_z .$$ [1)]

In the first approximation of the perturbation-calculation there remains, of this product $\vec{L}\cdot\vec{S}$, only the term $L_z S_z = \Lambda\Sigma$, and this gives rise to a splitting which is proportional to Σ, as is borne out by experiment. The multiplet is called normal when the energy increases with Σ; it is called reversed when it decreases with increasing Σ. For $\Lambda = 0$ the coupling-energy is zero. In this approximation no spin-splitting occurs in the case of the Σ-terms. This is the reason why, in the case of the Σ-terms, case b) always obtains. We therefore now assume $\Lambda > 0$.

Now we again go from the two-center problem to the freely rotating molecule. Let

$$\psi^{(m)}(q_0,\ldots,q_f;\sigma_1,\ldots,\sigma_f) = \sum_\nu \psi_\nu^{(m)}(q_0,\ldots,q_f) w_\nu$$

represent a family of eigen-functions of the molecule, with spin, belonging to the representation D_J of the rotational group, where the w_ν are any linearly independent functions of the spin-coordinates only. If on both sides, we carry out a rotation D^{-1},

---

1) For a more exact proof see W. Kramers: Z.f.Physik, 53 (1929) p. 429.

which sends the point $\zeta = (0,0,\varrho)$ into q_0, then we obtain:

$$\sum a_{gm}(D^{-1})\,\psi^{(g)}(q_0,\ldots,q_f;\sigma) = \sum_{\nu} \psi_{\nu}^{(m)}(Dq_0,\ldots,Dq_f)\,D^{-1}w_{\nu} \qquad (33.1)$$

where $(a_{gm}(D^{-1}))$ are the matrices for D^{-1} in the representation D_J and where $D^{-1}w_{\nu} = \sum_{\mu} b_{\mu\nu}(D^{-1})w_{\mu}$ is the transformation-formula of the spin-functions for the rotation D^{-1}. Since $dq_0 = \zeta$, the solution of (33.1) in terms of the $\psi^{(g)}$ on the left-hand side yields:

$$\psi^{(m)}(q_0,\ldots,q_f;\sigma) = \sum_{g} a_{gm}(D) \sum_{\nu} \psi_{\nu}^{(g)}(Q,Dq_1,\ldots,Dq_f)\,D^{-1}w_{\nu} \qquad (33.2)$$

This equation is the analogue of equation (31.1). The sum $\sum_{\nu}$ on the right-hand side arises from the function

$$\psi_Q^{(m)} = \sum_{\nu} \psi_{\nu}^{(m)}(Q,q_1,\ldots,q_f)\,w_{\nu} = \psi^{(m)}(\varrho,q_1,\ldots,q_f;\sigma_1,\ldots,\sigma_f)$$

due to the rotation D^{-1} of all the electrons and the spins.

We again assume that these functions $\psi_Q^{(g)}$ are approximately equal to zero with the exception of one or two amongst them: $\psi_Q^{(\Omega)}$ and $\psi_Q^{(-\Omega)}$ which are set equal to $f_+(\varrho)\Phi_{\Lambda,\Sigma}$ and $f_-(\varrho)\Phi_{-\Lambda,-\Sigma}$, respectively. The justification of this approximation can be carried out as in #32 as soon as the spin-terms in the wave-equation are known or are taken in a reasonable form. We do not here carry out this rather awkward calculation. The relation between the functions f_+ and f_- may be determined, as in #31, from the reflection-character of the eigenfunctions (33.2). We find:

$$f_-(\varrho) \quad = \quad (-1)^{J+\Omega} \quad w f_+(\varrho).$$

The function $f(\varrho) = f_+(\varrho)$ is determined as an eigen-function of an oscillator-equation which is constructed analogously to (32.9), the only difference being that J occurs instead of K.

Thus, each term of the two -body problem including spin, having definite quantum-numbers $\Lambda > 0, S, \Sigma$ and $\Omega = \Lambda + \Sigma$, gives rise to a series of vibration-terms with v = 0,1,2,....

Each of these is again split up into rotational levels which are distinguished from one another by means of the rotational-quantum-number J and the relfection-character w = ± 1. The following exact selection rules hold:

$$\left.\begin{aligned} &J \longrightarrow J-1,\ J,\ J+1 \quad (P\text{-}, Q, \text{and } R\text{-branch}) \\ &w \longrightarrow -w \end{aligned}\right\} \qquad (33.3)$$

In order to derive the selection-rules for Λ and Σ we proceed as in #31. In the series-expansions for $X\psi$, $Y\psi$, $Z\psi$ we simply equate $q_o = \varrho$ and D = 1. In this manner we come back to the selection-rules for the two-body problem including spin. As is easily evident, the selection rules for the case of a not-too-large spin-effect are:

$$\left.\begin{aligned} &\Lambda \longrightarrow \Lambda+1,\ \Lambda,\ \Lambda-1 \\ &S \longrightarrow S \\ &\Sigma \longrightarrow \Sigma \end{aligned}\right\} \qquad (33.4)$$

The first of these three rules is very well satisfied in practice. If, in the case of the heavier elements, the two other

rules are violated at times, we still have for the sum $\Lambda + \Sigma = \Omega$ the rule:

$$\Omega \longrightarrow \Omega-1, \Omega, \Omega+1$$

For transitions from case a) to case b), or the reverse, (Σ-Π transitions in heavier elements), as well as in the transition-region where the spin-perturbation is to be treated simultaneously with the rotational-splitting, the only selection-rules that are valid are those for J, w, Λ, and + but not those for K and Σ.

In case a) as well as in case b) we add to the term-symbols Σ, Π, Δ etc., as an index, the "multiplicity" 2S + 1, as we do in the case of the atom-terms. Thus, $^3\Sigma^+$ (say: triplet sigma plus) signifies a term of the two-body-problem having $\Lambda = 0^+$, S = 1.

#34. Homonuclear, Diatomic Molecules.

If the two nuclei of the two-body problem have equal charges the, aside from the axial rotation-reflection group, the problem also permits the reflection s at the center of mass (which commutes with all the elements of the rotation-reflection group). Under the influence of this reflection the spin-less eigen-functions assume the factors $\epsilon = \pm 1$. The same is true, also, upon addition of the spin since the pure spin-functions are invariant under the reflection s. Each of two eigen-functions $\varphi_{\pm\Lambda}$ belonging together always has the same quantum number ϵ. The terms are denoted as follows:

$\varepsilon = +1$: $\Sigma_g, \Pi_g, \Delta_g$,....: "even terms",

$\varepsilon = -1$: $\Sigma_u, \Pi_u, \Delta_u$,....: "odd terms".

If we combine the reflection s with the rotation D_ϑ about the z-axis, for which the eigen-functions $\varphi_{\pm\Lambda}$ take on the factor $(-1)^\Lambda$, we then obtain the reflection in the middle-plane $s_\vartheta = s \cdot D_\vartheta$ and we see that the eigen-functions $\varphi_{\pm\Lambda}$ then take on the factor $(-1)^\Lambda \varepsilon$. We shall <u>not</u> use this reflection since the reflection s leads us to much simpler rules.

If the nuclei of equal charge also have equal masses then an exchange of the two nuclei would also send the differential equation of the freely-rotating molecule back into itself (the exchange amounting to a replacement of the fictitious nucleus q_o by $-q_o$). The eigen-function ψ may either be symmetrical or anti-symmetrical with respect to the nuclei, i.e., under the influence of the transformation $q_o \longrightarrow -q_o$, it takes on the factor $\chi = \pm 1$. We wish to derive the relation between this symmetry-character χ and ε.

If we first carry out the exchange of the nuclei $q_o \longrightarrow -q_o$ and then the reflection s of the entire system ($q_o \longrightarrow -q_o$, $q_1 \longrightarrow -q_1$, etc.) we obtain the transformation $q_1 \longrightarrow -q_1, \ldots, q_f \longrightarrow -q_f$, where the reflection s is carried out only upon the electrons and where the eigen-functions ψ assume the factor $w \cdot \chi$. This is especially true when the nucleus q_1 is held fixed on the Z-axis, that is, for $q_1 = \zeta$, according to the notations of #32. For $q_1 = \zeta$, however, the function ψ approximates an eigen-function $\varphi_{\pm\Lambda}$ of the two-body problem, and this

assumes the factor ε upon the reflection s of all the electrons. Hence

$$\varepsilon = w.\chi \tag{34.1}$$

This result is independent of all assumptions regarding the magnitude of the spin-effect; it is valid for the two cases a) and b) of #33 as well as for the transition-cases. For the symmetry-character χ the selection-rule $\chi \rightarrow \chi$ is valid, for if ψ is symmetrical or anti-symmetrical with respect to the nuclei, then $X\psi$, $Y\psi$, $\Delta\psi$ are also symmetrical or anti-symmetrical.

In its series-expansion, too, there occur only similar functions. Since, moreover, $w \longrightarrow -w$, there follows for ε the strict selection-rule:

$$\varepsilon \longrightarrow -\varepsilon \tag{34.2}$$

That is: <u>Even terms combine only with odd ones and conversely</u>.

<u>#35. The Formation of the Molecule from Two Atoms</u>.

When the nuclei of the two-body problem are adiabatically drawn apart the molecule splits into two atoms (or ions) and the molecular terms $u(\varrho)$ transform continuously into terms of the pair of atoms. This process can even be followed spectroscopically. Classically speaking, by raising the vibrational quantum number the nuclear distance reaches an ever larger maximum-value and the energy of the molecule approaches the energy of the sparated pair of atoms. I.e., for increasing v, the vibrational-terms converge to a sum of two atom-terms. We now wish to investigate what relation-

ships exist in this arrangement between the symmetry properties of the molecule and those of the separated atoms.

We begin with two separated atoms. One is in a state $\varphi = \varphi_L^{(m)}$ with reflection-character w. The other one is in the state $\varphi' = \varphi_{L'}^{'(m')}$ having the reflection-character w'. We shall completely disregard the spin at first. Thus φ is a function of the space-coordinates q_1 to q_f of the electrons of the first atom, and φ' , of the space-coordinates q_{f+1} to $q_{f+f'}$ of the second atom The nuclei lie on the Z-axis at fixed, widely-separated points in such positions that their center of mass is at the origin.

The product $\varphi\varphi'$ is an eigen-function of the pair of atoms. Due to the interaction between the electrons and nuclei of the two atoms the corresponding term $E + E'$ splits into several terms, the eigenfunctions of which to the "zero-th approximation," are found by reducing of the <u>axial</u> rotation-reflection group in the space of the products $\varphi^{(m)}\varphi'^{(m')}$. This reduction proceeds as follows: We first collect the functions $\varphi_L^{(m)}$ in pairs $\varphi_L^{(\pm\lambda)}$ ($\lambda = 0,1,....L$), where each pair is transformed by the rotation-reflection group according to the representation A_λ. For $\lambda=0$ we deal only with a function $\varphi_L^{(0)}$ and the representation A_0^+ or A_0^- , depending on whether $(-1)^L w = +1$ or -1.[1] The most frequent case is $(-1)^L w = +1$. In like manner we obtain for the second atom the pairs of eigen-functions $\varphi_{L'}^{'(\pm\lambda')}$ and the representations $A_{\lambda'}$.

---

1) As to whether we deal with A_0^+ or A_0^- depends, namely on the behavior of the function as it undergoes the reflection s_y . This reflection s_y is composed of the rotation D_y about the Y-axis and the reflection in the nucleus of the first atom, in which $\varphi_L^{(m)}$ takes on the factors $(-1)^L$ and w.

The products $\varphi(\pm\lambda,\pm\lambda') = \varphi_L^{(\pm\lambda)}\varphi_{L'}^{(\pm\lambda')}$ are transformed in accordance with the product representation $A_\lambda \times A_{\lambda'}$, which, according to #12, breaks up as follows:

$$\left.\begin{aligned}
&A_\lambda \times A_{\lambda'} = A_{\lambda+\lambda'} + A_{|\lambda-\lambda'|} && \text{for } \lambda_1 \neq \lambda_2, \text{ both} > 0\\
&A_\lambda \times A_\lambda = A_{2\lambda} + A_0^+ + A_0^- && \text{for } \lambda > 0,\\
&A_\lambda \times A_0^\pm = A_\lambda && \text{for } \lambda > 0\\
&A_0^+ \times A_0^+ = A_0^- \times A_0^- = A_0^+\\
&A_0^+ \times A_0^- = A_0^- \ .
\end{aligned}\right\} \quad (35.1)$$

The corresponding eigenfunctions are also found without effort. In the first case it is the pairs:

$$\varphi(\lambda,\lambda'), \varphi(-\lambda,-\lambda') \text{ and } \varphi(\lambda,-\lambda'), \varphi(-\lambda,\lambda')$$

in the second case it is the pair $\varphi(\lambda,\lambda)$, $\varphi(-\lambda,-\lambda)$ and the single functions $\varphi(\lambda,-\lambda)+\varphi(-\lambda,\lambda)$ and $\varphi(\lambda,-\lambda)-\varphi(\lambda,-\lambda)$; the other cases are trivial. From equations (35.1) we read off the possible <u>Λ-values</u> for the molecule. The eigenfunctions just set up may be denoted by $\Phi(\pm\Lambda)$. For the lowest cases, the results are compiled in the following table:

Atomic terms	Representations A		Molecular terms
S,S(L = L' = 0)	$\lambda=\lambda'=0$	hence $\Lambda=0$	$\Sigma^{\pm}$ *
P,S(L = 1,L' = 0)	$\lambda=1, \lambda'=0$	hence $\Lambda=1$	Π
	$\lambda=0, \lambda'=0$	hence $\Lambda=0$	$\Sigma^{\pm}$ *

* See p. 260

Atomic terms	Representations Λ		Molecular terms
D,S($L=2, L'=0$)	$\lambda=2, \lambda'=0$	hence $\Lambda=2$	Δ
	$\lambda=1, \lambda'=0$	hence $\Lambda=1$	Π
	$\lambda=0, \lambda'=0$	hence $\Lambda=0$	$\Sigma^{\pm}$ *
P,P($L=L'=1$)	$\lambda=1, \lambda'=1$	hence $\Lambda=2, 0^+, 0^-$	Δ, Σ^+ and Σ^-
	$\lambda=1, \lambda'=0$	hence $\Lambda=1$	Π
	$\lambda=0, \lambda'=1$	hence $\Lambda=1$	Π
	$\lambda=0, \lambda'=0$	hence $\Lambda=0$	$\Sigma^{\pm}$ *

In the case of <u>equivalent nuclei</u> we add to each eigen-function $\Phi(\pm\Lambda)$ another one $s\Phi(\pm\Lambda)$ which arises from the first due to the reflection s in the center of mass. We then have to form the sums and differences, $(1+s)\Phi(\pm\Lambda)$ and $(1-s)\Phi(\pm\Lambda)$, for which the reflection-quantum number η has the values +1 and -1. Each term of the above scheme thus splits into an even term and an odd one. As we shall see, according to the Pauli exclusion principle, under certain conditions some of the terms may drop out, provided the atoms are in the same states.

We now include the spin and the exclusion principle. The spin-effect is small in comparison to the electrostatic interaction and will, at first, be neglected. Instead of the previous m we now write m_L and instead of $\varphi_L^{(m)}(q)$ we write $\varphi(m_L|q)$ where q represents q_1 to q_f. Let the first atom have the spin-number s and the eigenfunctions[1] $\psi(m_L, m_s|q,\sigma) = \varphi(m_L|q)\cdot u(m_s|\sigma)$. Likewise, let the second atom have the spin-number s' and the eigenfunctions[1]

---

* The sign (+ or -) is determined by $(-1)^L w\cdot(-1)^{L'} w'$.

1) From these eigenfunctions, we may also first form for each individual atom the anti-symmetric linear-combinations $\Sigma\delta_Q Q\psi$; the result (35.2) is the same.

$$\psi'(m'_L, m'_s \mid q,\sigma) = \varphi'(m'_L \mid q)\cdot u'(m'_s \mid \sigma) \quad .$$

Upon multiplying the two we get for the molecule the approximate eigenfunctions $\psi\psi'$ from which we must form, according to the Pauli-principle, the anti-symmetric linear combinations

$$\psi_a = \sum \delta_P P\psi\psi' = \sum \delta_P P\varphi\varphi' uu' \qquad (35.2)$$

We can write this expression in another manner. If in the permutation group $S_{f+f'}$ we denote by S all the permutations which permute only the first f electrons amongst one another and leave the remaining ones unchanged, and if we denote by S' all those which permute only the last f' electrons, then the products QQ' form a sub-group g of $S_{f+f'}$ whose cosets we may denote by Rg. Here R is any arbitrarily chosen element of the coset. Consequently (35.2) is equivalent to

$$\psi_a = \sum_R \delta_R R\left(\sum_Q \delta_Q Q\psi\right)\left(\sum_{Q'} \delta_{Q'} Q'\psi\right) \qquad (35.3)$$

The individual terms of this sum correspond to states in which certain electrons are found at one nucleus and the remaining ones at the other nucleus. It is clear that the sum (35.3) does not vanish as long as the individual factors $\sum \delta_Q Q\psi$ (the anti-symmetric eigenfunctions of the individual atoms) do not vanish. The number of linearly-independent functions (35.3) is equal to the product of the number of linearly-independent anti-symmetric eigenfunctions of the individual atoms. That is, it is equal to the number of possible numerical combinations m_L, m_S, m'_L, m'_S.

The rotations and reflections of the electrons and of their spins commute with all permutations and may, therefore, be carried

out termwise on (35.2). We now reduce the products $\varphi\varphi'$ in (35.2) in accordance with the axial rotation-reflection group, and the products uu' according to the spacial-rotation group. The former is achieved by the above equations (35.1) and gives rise to the various values of Λ . The latter occurs according to the known equation for the combination of spin-vectors:

$$S = s + s',\ s + s' - 1,\ldots,\ |s - s'| \ . \qquad (35.4)$$

In this manner we obtain the new linear combinations of the functions (35.2):

$$\psi_\alpha' = \sum_P \delta_P P\varphi(\pm\Lambda|q)r(S, M_s|\sigma) \qquad (35.5)$$

The terms belonging to the different values of and s split apart due to the interaction of the electrons and the nuclei. The total term-multiplicity is obtained upon combining each value of Λ (as often as it occurs in equation (35.1)) with each value of S from (35.4).

In the case of <u>equivalent nuclei</u> the reflection s at the center of mass yields, for each eigen-function ψ_a', a new one, $s\psi_a'$, having the same quantum-numbers Λ, S, M_s and the same energy. In case the factors φ, φ' in (35.2) belong to different terms of the two atoms, then these new functions $s\psi_a'$ are linearly independent of the ψ_a' already mentioned. In this case we may form $(1 + s)\psi_a'$ and $(1 - s)\psi_a'$, as above. <u>All counted term-types thus occur twice: once with $\varepsilon = +1$ and once with $\varepsilon = -1$.</u>

If, however, the factors φ, φ' belong to the same terms of the two atoms, i.e., if the atoms are in equivalent states, then the $s\psi_a'$ are already contained in the linear family of the ψ_a. All the terms then occur only once with $\varepsilon = +1$ or -1. We shall now derive the rules applying to this case.

We may replace the reflection s by a reflection at the first nucleus k (in which the function ψ in (35.2) takes on the factor w) followed by a parallel displacement by ϱ in the direction kk', (in which the function ψ becomes the function ψ' having the same quantum-numbers). Thus

$$s\psi(m_L, m_s | q_1, \ldots, q_f ; \sigma_1, \ldots, \sigma_f) = w \cdot \psi'(m_L, m_s | q_1, \ldots, q_f ; \sigma_1, \ldots, \sigma_f).$$

Likewise,

$$s\psi'(m_L' \; m_s' \mid q_{f+1}, \ldots, q_{2f} ; \sigma_{f+1}, \ldots, \sigma_{2f})$$

$$= w \cdot \psi(m_L', m_s' \mid q_{f+1}, \ldots, q_{2f}, \sigma_{f+1}, \ldots, \sigma_{2f})$$

hence, (since $w^2 = 1$),

$$s\psi(m_L, m_s | q_1, \ldots; \ldots, \sigma_f)\psi'(m'_L, m'_s | q_{f+1}, \ldots; \ldots, \sigma_{2f})$$

$$= \psi(m'_L, m'_s | q_{f+1}, \ldots; \ldots, \sigma_{2f})\psi'(m_L, m_s\ q_1, \ldots; \ldots \sigma_f)$$

In other words, the effect of the reflection s on a product $\psi\psi'$ consists of interchanging m_L with m'_L, m_s with m'_s, and the electrons with numbers 1 to f with the electrons f + 1 to 2f. The latter electron-interchange P*, being a product of f transpositions, is an even permutation or an odd one, depending on whether f is even or odd. If we now apply the operation $\sum \delta_P P$ (which commutes with s) to both sides of the expression just obtained, then upon adding the factor $\delta_{P^*} = (-1)^f$ we can reverse the electron-interchange P* on the right-hand side. This means, therefore, that the reflection s applied to the functions ψ_α (35.2) amounts to an interchange of m_L with m'_L and of m_s with m'_s plus the addition of a factor $(-1)^f$.

When we say that we go from the functions ψ_α (35.2) to their linear combinations ψ'_α (35.5), we mean that we introduce instead of the spin-functions un' their linear combinations $v(S, M_S)$, and instead of the products $\varphi\varphi'$ their linear combinations Φ $(\pm\Lambda)$. It was proven in #26 that upon interchange of the spins m_L, m'_L the $v(S, M_S)$ are symmetrical when S = 2s, 2s-2, ..., 0 and anti-symmetrical when S = 2s-1, 2s-3, ..., 1. I.e., the $v(S, M_S)$ assume the factor $(-1)^{2s-S}$ in this inter-change. Here then, 2s is even or odd according to whether f is even or odd. Therefore, $(-1)^f (-1)^{2s-S} = (-1)^S$. The functions $\Phi(\pm\Lambda)$ have the following form:

for $\lambda \neq \lambda'$, $\Lambda = \lambda + \lambda'$: $\Phi(\Lambda) = \varphi(\lambda)\varphi'(\lambda')$; $\Phi(-\Lambda) = \varphi(-\lambda)\varphi'(-\lambda')$

for $\lambda > \lambda' > 0$, $\Lambda = \lambda - \lambda'$: $\Phi(\Lambda) = \varphi(\lambda)\varphi'(-\lambda')$; $\Phi(-\Lambda) = \varphi(-\lambda)\varphi'(\lambda')$

for $\lambda = \lambda' > 0$, $\Lambda = 2\lambda$: $\Phi(\Lambda) = \varphi(\lambda)\varphi'(\lambda)$; $\Phi(-\Lambda) = \varphi(-\lambda)\varphi'(-\lambda)$

for $\lambda = \lambda' > 0$, $\Lambda = 0^{\pm}$: $\Phi(0^{\pm}) = \varphi(\lambda)\varphi'(-\lambda) \pm \varphi(-\lambda)\varphi'(\lambda)$

for $\lambda = \lambda' = 0^{\pm}$, $\Lambda = 0^{+}$: $\Phi(0^{+}) = \varphi(0)\varphi'(0)$

In the first two cases ($\lambda \neq \lambda'$) we have, in addition to the Φ-function above, another one with λ and λ' interchanged and belonging to the same energy-value. Instead of considering the Φ-functions above we may consider their sums Φ_+ and differences Φ_-. These are transformed exactly according to the same representation A_Λ :

for $\lambda \neq \lambda'$, $\Lambda = \lambda + \lambda'$: $\Phi_{\pm}(\Lambda) = \varphi(\lambda)\varphi'(\lambda') \pm \varphi(\lambda')\varphi'(\lambda)$,

$$\Phi_{\pm}(-\Lambda) = \varphi(-\lambda)\varphi'(-\lambda') \pm \varphi(-\lambda')\varphi'(-\lambda)$$

for $\lambda > \lambda' > 0$, $\Lambda = \lambda - \lambda'$: $\Phi_{\pm}(\Lambda) = \varphi(\lambda)\varphi'(-\lambda') \pm \varphi(\lambda')\varphi'(-\lambda)$

$$\Phi_{\pm}(-\Delta) = \varphi(-\lambda)\varphi'(\lambda') - \varphi(-\lambda')\varphi'(\lambda)$$

If we consider the sums and differences we can, in all cases, read off directly from the form of the functions as to which factors they take on upon an interchange of the quantum-numbers $m_L = \pm\lambda$ and $m_L' = \pm\lambda'$. Upon adding the factor $(-1)^S$ we obtain the desired values of ε .

For $\lambda \neq \lambda'$ both values $\varepsilon = \pm 1$ occur, but each pair of values λ, λ' is to be considered just once regardless of the order of sequence. For $\lambda = \lambda' > 0$, both terms $\lambda = 2\lambda$ and $\Lambda = 0^+$ have $\varepsilon = (-1)^S$. The term $\Lambda = 0^-$, though, has $\varepsilon = (-1)^{S+1}$. For $\lambda = \lambda' = 0$, $\Lambda = 0^+$, we have again $\varepsilon = (-1)^S$.

Collecting the rules together, we have:

A. Different Atoms

$\lambda = 0^{\pm}, 1, \ldots, L$ (0^+ for $w = (-1)^L$, else 0^-),

$\lambda' = 0^{\pm}, 1, \ldots, L'$ (0^+ for $w = (-1)^L$, else 0^-),

$\Lambda = |\lambda \pm \lambda'|$ or/and $\Lambda = 2\lambda, 0^+, 0^-$ etc. according to (35.1),

$S = s + s', s + s' - 1, \ldots, |s - s'|$ independent of Λ.

B. Same Atoms in Different States:

The terms are as above, all having $\varepsilon = 1$ (even) and $\varepsilon = -1$ (odd).

C. Same Atoms in Same States:

a) $\lambda \neq \lambda'$: Terms as above with $\varepsilon = \pm 1$:
each pair λ, λ', however, occurs only once regardless of order of sequence.

b) $\lambda = \lambda' > 0$: $\Lambda = 2\lambda, 0^+$ with $\varepsilon = (-1)^S$

$\Lambda = 0^-$ with $\varepsilon = (-1)^{S+1}$

c) $\lambda = \lambda' = 0$: $\Lambda = 0^+$ $\varepsilon = (-1)^S$

#36. Remarks about the Estimation of the Energy.

The estimation of the split-up molecular terms and of the stability of the molecule present a difficult problem. There are three methods available, each one having its own disadvantages and

thus, of necessity, complementing one another.

The first method is the perturbation-calculation applied to the non-orthogonal approximate eigen-functions (35.2) or (35.5). The method essentially offers nothing new as compared to the calculation for the atom in #29. One can also perform the calculation either with the aid of the group-characters (Heitler) or without their aid (Slater). The method was used successfully by Heitler and London[1] and for the ground-state of the H_2-molecule. It was then used[2] by the same authors for the explanation of chemical binding. It was simplified[3] by Born and Weyl and extended by Heitler and Rumer[4] to poly-atomic molecules. Only when all the component atoms (or all but one) are in S-states and there is no accidental degeneracy does the method lead to clear formulae. The result is that the state having the lowest energy is the one for which the spin-number S has the smallest value $|s - s'|$, assuming that a certain exchange integral is positive and that it overweighs the remaining perturbation terms. For the ground state of H_2 this exchange-integral is large and positive, and experience indicates that the same holds in most of the other cases also. The result can be explained in terms of a "saturation" of the spin-

---

1) Heitler, W., and F. London: Z.f.Physik, vol.44 (1927) p.455 to 472.
2) Z.f. Physik, vol. 46, p.455; vol. 47, p.835; vol. 50 p. 24. Summarizing presentation by W. Heitler: Phys.Z. vol. 31 (1930) p.185.
3) Born,M.: Z.f.Physik, vol. 64, (1927). p.729 to 740. Weyl,H.: Gött.Nachr. (1930) p.285; (1931) p.33. Summarizing presentation by M. Born: "Ergebnisse der exakten Naturwissenschaften," vol.10 (1931) p.387 to 444.
4) Heitler, W. and G. Rumer: Z.f. Physik, vol. 68, (1931), p.12 - 41. See also the report by M. Born cited above.

vectors corresponding to the saturation of the chemical valences. The "valence with respect to hydrogen" of an atom in its ground state is then to be set equal to doubled(the) spin-number, 2s.

The perturbation calculation just sketched is valid only for large inter-nuclear distances ϱ , of course, it does not hold as $\varrho \longrightarrow \infty$ since as $\varrho \longrightarrow \infty$ the second approximation of the perturbation calculation which considers the polarization and leads to the van der Waals forces, outweighs[1] the first approximation. At infinity, namely, the polarization forces behave like ϱ^{-7} , whereas the exchange-integral approaches zero at the rate of $e^{-\alpha\varrho}$. One hopes that the exchange integral will behave in many cases as in hydrogen, that is, that for moderately large ϱ the first approximation will yield the result. This hope is quite well justified as long as the atom-terms under consideration do not possess any very close neighboring terms. If these are present, however, one has to include them in the perturbation calculation, (Ritz Method). For example, despite the triplet-ground-state ($2s^2\,2p^2\,{}^3P$) of the C-atom, its quadri-valence can first be understood through the presence of a close, excited term $2s2p^3\,{}^5S$. For small values of ϱ, such as we have in the actual molecule, the method given here fails. As a result of the very frequent high degeneracy associated with excited molecular states the actual calculations of this method become unusually complicated, if not actually impossible. Also, the method is essentially not sensitive to certain fine points like the directed valences. A modified method by Slater[2] which starts from the eigen-functions

---

1) London, F. and R. Eisenschitz: Z.f.Physik, vol. 60 (1930), p. 491 - 572.
2) Slater, J.C.: Physic. Rev. vol. 38 (1931), p. 1109

of the individual electrons (instead of from the actual atoms) is able to consider these fine points.

The second method which is applicable to small distances consists in considering the limiting case $\rho=0$, in which the nuclei coincide with one another and the molecule becomes an atom. We next investigate again how the symmetry quantum numbers of the molecule behave in this limiting case. We prefer to proceed by starting with an atom whose nucleus then splits up into two force-centers along the Z-direction Due to the removal of the central symmetry of the field, each group of eigen-functions $\psi_L^{(m)}$ splits into sub-groups $\psi_L^{(\pm\Lambda)}$ with $\Lambda=0,1,2,\ldots,L$. Whether to set $\Lambda = 0^+$ or 0^- depends upon whether $(-1)^L$ $w = +1$ or -1. The spin-number S is maintained during the pulling-apart process. In the case of equivalent nuclei, $\varepsilon=w$, because ε as well as w indicates the behavior of the eigen-functions of the electron-configuration under the influence of reflection at the center of mass. Hence we have a complete picture of the term-types which can arise from one atom term during the nuclear-splitting.

An approximate picture of the positions of the terms for small ρ shows the following rule to hold: Of those terms, which arise from one atomic term, the one that lies lowest is the one for which the absolute values of the electron-eigenfunctions (or the density of the electron cloud) grow the fastest, from the zero-point towards the positive and negative Z-direction (along which the two halves of the nucleus are pulled). Disregarding the repulsion of the nuclei and considering only very small ρ, the terms agree with those of the atom. If the repulsion is included all the energy-

values, naturally, are raised by a fixed amount for each ϱ .

In order to bridge the gap between the large and the small ϱ -values and in order to find out more about the behavior and the arrangement of the molecular terms corresponding to the actual intermediate ϱ -values, we make use of a third method which was developed by Mulliken and Hund.[1] The method is as follows: We study the behavior of the individual electrons under the influence of the two nuclei, excluding, however, the interaction among the electrons or replacing it by means of a screening of the nuclear fields. The method corresponds to the Hartree method for atomic-spectra and leads qualitatively to very good results. Each individual electron has a quantum-number $\lambda = 0,1,2,\ldots$ and is, therefore, denoted as a σ , π , or δ electron. For the σ electrons we always have $\lambda=0^+$, and never $\lambda=0^-$.* The combination of the λ values for the individual electrons follows the known rules (35.1). Due to the Pauli-principle, however, not all the calculated terms occur in reality. According to the exclusion-principle there is room for only two electrons (having opposite spins) in the same σ orbital. Likewise, there is room for only four electrons in the same π or δ orbital corresponding to the values $m_\Lambda = \pm\Lambda$, $m_s = \pm\frac{1}{2}$ of the Z-components of the orbital- and spin-angular momenta. For two equivalent σ -electrons we have the symbol σ^2. Likewise, for two, three, or four equivalent

---

1) Hund, F.: For the Interpretation of Molecular Spectra V. Z.f. Physik vol. 63 (1930), p.719. Application to the question of chemical binding. Z.f Physik vol. 73 (1931) p. 1. Compare also G. Herzberg: Z.f. Physik, vol. 57 (1928), p.601.

* For if the eigenfunction ψ (written in terms of cylindrical coordinates r,z,φ) does not depend on φ , as must be the case for $\lambda=0$ then ψ remains invariant under the reflection s_y.

π -electrons we have the symbols π^2, π^3, π^4, etc. A completed shell σ^2, π^4, or δ^4 does not raise the term-multiplicity of the remaining electrons. It itself is in a $^1\Sigma^+$ -state, since all the spin- and orbital angular momenta are removed. The non-filled shells of equivalent electrons give rise to the following terms:

a σ-electron: $^2\Sigma$

π or π^3 : $^2\Pi$

π^2 : $^3\Sigma^-$, $^1\Sigma^+$, $^1\Delta$

δ or δ^3 : $^2\Delta$

δ^2 : $^3\Sigma^-$, $^1\Sigma^+$, $^1\Gamma$

For inequivalent electrons or electron-groups, the λ-values and spin-numbers are to be combined simply according to (35.1) and (35.4). No combination is excluded, just as in #28. For example, in the case of $\sigma\sigma$ (two inequivalent σ-electrons) we have the terms $^3\Sigma^+$ and $^1\Sigma^+$; likewise, in the case $\sigma\pi$ or $\sigma\pi^3$ we have the terms $^3\Pi$, $^1\Pi$; in the case $\sigma\pi^2$ we obtain, upon combining $^2\Sigma$ with $^3\Sigma^-$, $^1\Sigma^+$, $^1\Delta$, the terms $^4\Sigma^-$, $^2\Sigma^-$, $^2\Sigma^+$, $^2\Delta$; etc.

In the case of equivalent nuclei, each electron has still another quantum-number $\varepsilon = \pm 1$ (or an index g or u associated with the electron-symbol: σ_g , σ_u etc) and we have: $\varepsilon = \varepsilon_1 \varepsilon_2 \cdots \varepsilon_{2f}$.

As for the relative positions of the individual energy-levels, one proceeds either through the two boundary transitions $\rho \to \infty$ and $\rho \to 0$, which, of course, have meaning also for an individual electron, or by means of the direct calculation of the eigenfunctions of the two-center-problem for an electron, with the aid of elliptical coordinates or perturbation methods. For a detailed treatment

we refer the reader to the work of F. Hund cited on page 269, footnote 1.

In the arranging of the terms for small, intermediate, and large ϱ, it is to be noted that only terms of the same "race", i.e. having the same symmetry-quantum-numbers (in our case, Λ, S, and ε), can pass over into one another. Terms of different "races" may cross one another without incurring any mutual perturbation On the other hand, terms of the same 'race' do not as a rule cross one another and, therefore, can easily be arranged in order (starting from the bottom up).